



INSTITUTO ESPAÑOL  
DE ANALISTAS DESDE 1965  
FUNDACIÓN

# El impacto de la inteligencia artificial en la industria financiera: promesas y amenazas

CARLOS CONTRERAS

ISBN: 978-84-09-60967-3  
Edita: Instituto Español de Analistas  
Imprime: Reimpventa

# Índice

Lista de términos .....	7
Lista de tablas .....	10
Lista de gráficos .....	12
Presentación .....	13
1. Introducción .....	16
<b>PARTE I</b>	
2. Situación actual de la inteligencia artificial .....	23
3. Principales técnicas de AI .....	35
4. Implantando una estrategia de AI en instituciones financieras .....	41
<b>PARTE II</b>	
5. Aplicaciones comunes de AI en la industria .....	56
5.1. Procesos de administración y gestión de información ..	56
5.2. Contabilidad y auditoría interna .....	58

5.3. Determinación del capital regulatorio.....	63
5.4. Gestión de riesgos y detección y prevención del fraude.....	65
5.5. Cumplimiento normativo.....	72
5.6. Gestión del cambio (CM) y gestión del conocimiento (KM).....	75
5.7. Educación financiera.....	79
<b>6. Banca comercial.....</b>	<b>90</b>
6.1. Actividad de marketing y ventas.....	91
6.2. Chatbots y asistentes virtuales inteligentes (IVAs).....	94
6.3. Gestión de Pagos.....	98
6.4. Creditscoring de préstamos hipotecarios y financiación al consumo.....	100
<b>7. Banca Privada y Banca Personal.....</b>	<b>109</b>
7.1. Aplicaciones de AI en la industria de gestión patrimonial.....	110
7.2. Robo-advisors.....	112
<b>8. Banca de inversión y corporativa.....</b>	<b>130</b>
8.1. Actividad de trading.....	132
8.2. Gestión de carteras.....	137
8.3. Evaluación de solvencia y rating de bonos.....	142
8.4. Mercados primarios de capitales.....	145
8.5. Capital privado y capital riesgo.....	147
<b>9. Industria aseguradora.....</b>	<b>165</b>
9.1. Gestión contractual y documental.....	166
9.2. Detección de fraudes.....	167
9.3. Marketing y Distribución.....	168

9.4. Evaluación de riesgos .....	169
9.5. Análisis de comportamiento de los asegurados.....	171
9.6. Desarrollo de productos de Seguro.....	172
9.7. Adopción de la tecnología en el sector de seguros .....	174
9.8. Principios de gobernanza de la AI.....	176
<b>10. Bancos centrales, supervisores del mercado y autoridades de competencia .....</b>	<b>182</b>
10.1. Bancos Centrales.....	183
10.2. Supervisores de los mercados de valores .....	187
10.3. Autoridades de defensa de la competencia .....	191
<b>PARTE III</b>	
<b>11. Riesgos, amenazas y desafíos de la AI en la industria.....</b>	<b>201</b>
11.1. Inversiones improductivas .....	202
11.2. Riesgo de penetración de nuevos competidores .....	204
11.3. Impacto sobre los recursos humanos.....	207
11.4. Cambios en las estructuras organizativas.....	214
11.5. Comportamientos de alucinación y opacidad de los resultados.....	215
11.6. Exceso de confianza en la AI y dependencia .....	221
11.7. Ciberdelincuencia vinculada a AI .....	223
11.8. Problemas éticos: privacidad, prejuicios y discriminación.....	226
<b>12. Conclusiones .....</b>	<b>238</b>
<b>Anexo A. Detalle de algunas técnicas de inteligencia artificial...</b>	<b>244</b>
<b>Referencias .....</b>	<b>311</b>



<b>Papeles de la Fundación</b> .....	321
<b>Entidades Patrono de la Fundación del Instituto Español de Analistas</b> .....	325



## Principales términos

Actor-Critic	Aprendizaje supervisado
Advantage Actor-Critic	Aprendizaje no supervisado
Advanced Reinforcement Learning	Árboles de clasificación y regresión (CART)
Agrupación jerárquica	Árboles de decisión
Algoritmos de Aprendizaje Profundo Q-Networks	Artificial neuronal networks (ANNs)
Algoritmos Genéticos (AG)	Asistente virtual inteligente (IVA)
Análisis de Componentes Independientes (ICA)	Asynchronous Advantage Actor-Critic (A3C)
Análisis de Componentes principales (ACP)	Auto machine learning (AML)
Análisis de Factores (AF)	Bidirectional Encoder Representations from Transformers
Aprendizaje automatizado en el borde	Combined Algorithm Selection and Hyper- parameter tuning (CASH)
Aprendizaje por refuerzo	Convolutional Neural Networks (CNN)
Aprendizaje por refuerzo avanzado	



Decision support system (DSS)	Grid Search
Deep Learning	Hierarchical Reinforcement Learning
Deep Deterministic Policy Gradient (DDPG)	Hybrid intelligent system (HIS)
Deep Reinforcement Learning from Human Feedback (RLHF)	Inteligencia artificial explicable (XAI)
Density-Based Spatial Clustering of Applications with Noise (DBSCAN)	Isolation Forest
Deterministic Policy Gradient (DPG)	K-means
Edge Learning	K-medoids
Elastic Net	K-Nearest Neighbors o KNN
Evolutionary methods	Kernel
Extreme gradient boosting (XGBoost)	Least Angle Regression (LARS)
Federated learning	Leaky ReLU
Fuzzy Logic	Light gradient boosting machine (LightGBM)
Generative Adversarial Imitation Learning (GAIL)	Local Interpretable Model-agnostic Explanations
Generative Adversarial Network (GAN)	Mapa de características autoorganizadas (SOFM)
Generative Pre-trained Transformer	Minería de datos
Gradiente Boosting (GB)	Métodos Evolutivos
	Método Ridge



Model-based Reinforcement Learning

Modelos generativos probabilísticos (PGM).

Naïve Bayes Classifier (NBC)

Neural Architecture Search (NAS)

One-Class SVM

Optimización de la Política Próxima (PPO)

Parametric ReLU (PReLU)

Procesamiento de lenguaje natural (NLP)

Proximal policy optimization (PPO)

Q-Learning

Random forest

Rectified Linear Unit

Recurrent Neural Networks (RNN)

Red Generativa Adversarial (GAN)

Red Generativa Adversarial Condicional (cGAN)

Red Generativa Adversarial de Ciclo Consistente (CycleGAN)

Red neuronal *autoencoder*

Red Neuronal Convolutacional

Red Neuronal Explicable Mejorada (ExNN)

Red Neuronal Profunda

Red Neuronal Recurrente

Regresión LASSO

REINFORCE

ReLU (Rectified Linear Unit)

Retropropagación con descenso de gradiente

Robo-advisor (RA)

Robotic process automation o RPA

Self-Organizing Map (SOM)

SHapley Additive exPlanations

Sistemas basados en conocimientos (KBS)

Sistemas expertos (ES)

Support vector machine (SVM)

T-distributed Stochastic Neighbor Embedding (t-SNE)



## Lista de tablas

<b>Tabla A.1.</b> Ventajas e inconvenientes de los sistemas expertos ...	247
<b>Tabla A.2.</b> Ventajas e inconvenientes del algoritmo de vecino más cercano.....	251
<b>Tabla A.3.</b> Ventajas e inconvenientes del Naïve Bayes Classifier..	252
<b>Tabla A.4.</b> Ventajas e inconvenientes de las máquinas de soporte vectorial.....	254
<b>Tabla A.5.</b> Ventajas e inconvenientes de los modelos Random Forest.....	262
<b>Tabla A.6.</b> Ventajas e inconvenientes de los modelos Gradient Boosting.....	264
<b>Tabla A.7.</b> Ventajas e inconvenientes de los modelos basados en CART.....	267
<b>Tabla A.8.</b> Ventajas e inconvenientes de Neural Architecture Search.....	277
<b>Tabla A.9.</b> Ventajas e inconvenientes del Análisis de Componentes Principales (ACP) .....	282
<b>Tabla A.10.</b> Ventajas e inconvenientes del modelo REINFORCE ..	291



<b>Tabla A.11.</b> Ventajas e inconvenientes del modelo Deterministic Policy Gradient.....	296
<b>Tabla A.12.</b> Ventajas e inconvenientes del modelo Deep Deterministic Policy Gradient.....	297
<b>Tabla A.13.</b> Ventajas e inconvenientes del Model-based Reinforcement Learning.....	299
<b>Tabla A.14.</b> Ventajas e inconvenientes del enfoque Hierarchical Reinforcement Learning.....	300
<b>Tabla A.15.</b> Ventajas e inconvenientes del enfoque de Aprendizaje Federado.....	308

## Lista de gráficos

<b>Gráfico 2.1.</b> Evolución de capacidades de la inteligencia artificial.....	26
<b>Gráfico 2.2.</b> Fuentes de mejora del valor de producción inducida por la aplicación de herramientas de AI.....	30
<b>Gráfico 2.3.</b> Distribución de los ahorros de costes por áreas en la industria bancaria .....	30
<b>Gráfico 2.4.</b> Distribución por funciones de la creación de valor en la industria bancaria .....	31
<b>Gráfico 2.5.</b> Grado de importancia estratégica de la AI en el plazo de 2 años para compañías del sector financiero .....	32
<b>Gráfico 3.1.</b> Técnicas de aprendizaje supervisado.....	36
<b>Gráfico 3.2.</b> Técnicas de aprendizaje no supervisado .....	37
<b>Gráfico 3.3.</b> Técnicas de aprendizaje por refuerzo .....	38
<b>Gráfico 3.4.</b> Principios y técnicas clave en el diseño y entrenamiento de redes neuronales.....	39
<b>Gráfico 4.1.</b> Impacto de los proyectos de AI.....	48



INSTITUTO  
ESPAÑOL  
DE ANALISTAS  
DESDE 1965  
FUNDACIÓN

El impacto de la inteligencia artificial en la  
industria financiera: promesas y amenazas

1

# Presentación



## Presentación

La inteligencia artificial ha emergido como una fuerza transformadora en diversas industrias, y el sector financiero no es una excepción. Desde la automatización de tareas hasta el análisis predictivo, la inteligencia artificial está remodelando profundamente la forma en que operan las instituciones financieras. Por este motivo, el Instituto Español de Analistas decidió elaborar un estudio de carácter divulgativo sobre los diferentes efectos que la aplicación de herramientas de inteligencia artificial está teniendo en la actualidad y pueden tener en el futuro en la industria financiera.

El estudio se titula *“El impacto de la inteligencia artificial en la industria financiera: promesas y amenazas”* y ha sido elaborado por Carlos Contreras al que quiero felicitar por el excelente trabajo realizado.

En la primera parte del informe se comenta la situación actual del fenómeno de la inteligencia artificial y se describen brevemente las principales técnicas y los problemas habituales que

enfrentan las entidades financieras en la implementación de su estrategia en esta materia. En la segunda parte se comentan las herramientas de inteligencia artificial en los diferentes ámbitos de la industria: banca comercial, banca de inversión, banca privada, gestión de activos y compañías de seguros. También se describe cómo los bancos centrales y los supervisores del mercado y las autoridades de defensa de la competencia están utilizando este tipo de aplicaciones. La

tercera parte del informe se dedica a analizar los principales riesgos y amenazas que se derivan de la generalización del uso de tecnología de inteligencia artificial y se destacan algunas conclusiones.

En las conclusiones se señala que estamos tan sólo al “principio del principio” de la era de la inteligencia artificial, y que la industria financiera ni siquiera ha arañado la superficie de cómo puede transformar su operativa y la relación con sus clientes. No cabe duda de que estas aplicaciones tecnológicas pueden contribuir a que los servicios financieros sean más baratos, rápidos, accesibles y más eficientes en muchos sentidos.

Por lo tanto, la inteligencia artificial está revolucionando la industria financiera, ofreciendo una amplia gama de beneficios, que incluyen la mejora de la productividad, la posibilidad de reducir el fraude e impulsar la innovación, así como mejorar la experiencia del cliente. Sin embargo, también plantea desafíos significativos, como el riesgo de que las millonarias inversiones realizadas sean improductivas, que su implantación tenga un impacto negativo en las políticas de

recursos humanos, y que exija cambios profundos en sus estructuras organizativas.

Para aprovechar al máximo el potencial de la inteligencia artificial y mitigar sus riesgos, las instituciones financieras deben adoptar un enfoque equilibrado que priorice la equidad, la transparencia y la responsabilidad ética en el desarrollo y la implementación de tecnologías de inteligencia artificial.

En el Instituto Español de Analistas confiamos que este informe sea de utilidad para sus lectores y que contribuya a un debate constructivo sobre los retos y desafíos que la inteligencia artificial plantea a la industria de servicios financieros.

**Lola Solana**

*Presidenta del Instituto  
Español de Analistas*



INSTITUTO  
ESPAÑOL  
DE ANALISTAS  
DESDE 1965  
FUNDACIÓN

El impacto de la inteligencia artificial en la  
industria financiera: promesas y amenazas

1

# Introducción

## Introducción

La actual oleada de avances de la inteligencia artificial aún no ha cumplido una década, y su impulso parece estar, en todo caso, acelerándose. La AI está afectando a todos los ámbitos de la sociedad y los negocios, transformando naciones, creando nuevas oportunidades y cambiando la forma en la que consumimos, aprendemos y trabajamos.

*Evident Insights (2023)*

La inteligencia artificial (AI) está demostrando ser una poderosa herramienta en los servicios financieros ya que la capacidad de los ordenadores para imitar las competencias de la mente humana tiene múltiples aplicaciones en este ámbito. Estamos justo “al principio del principio” de la era de la inteligencia artificial<sup>1</sup> y cabe afirmar que la industria financiera ni siquiera ha arañado

la superficie de cómo la AI puede transformar su operativa y su relación con el cliente. Uno de los mayores potenciales de la AI es ayudar a los bancos a convertir su *tesoro oculto* de datos en información útil para los clientes y, como resultado, rentable para las entidades<sup>2</sup>. La inteligencia artificial tiene el potencial de cambiar la naturaleza y las fronteras de actividad de la industria financiera,

<sup>1</sup> Véase Lin (2019).

<sup>2</sup> Véase Yurcan (2022).



no simplemente cómo ésta lleva a cabo sus actividades. Mediante un despliegue masivo de herramientas de AI las entidades pueden aspirar a convertirse en el centro de la vida financiera de sus clientes y en parte de un ecosistema más grande<sup>3</sup>. La AI puede ayudar a lograr una transformación digital real, no sólo a conseguir cambios incrementales y reducción lineales de costes.

La AI puede contribuir a que los servicios financieros sean más baratos, más rápidos, más accesibles y más eficiente en muchos sentidos<sup>4</sup>. Estas capacidades incluyen la posibilidad de aprender de experiencias previas para comprender y responder, mediante modelos de lenguaje denominados *large language models* (LLM) en su terminología en inglés, así como para adoptar decisiones y para ofrecer soluciones a una amplia variedad de problemas. La AI mejora el reconocimiento de patrones adaptativos mediante el uso de métodos estadísticos modernos y grandes volúmenes de datos para proporcionar como respuesta

la «mejor conjetura» a cualquier conjunto de problemas específicos y estrechamente definidos.

En la industria financiera, la disponibilidad de datos legibles por máquinas que se generan en los mercados financieros, los indicadores macroeconómicos, y los datos corporativos de empresas cotizadas es enorme. Además, los profesionales de la industria financiera están cada vez más interesados en datos alternativos que incluyen publicaciones en redes sociales, grabaciones de voz, videos, artículos de noticias e incluso imágenes satelitales<sup>5</sup>. El uso combinado de estas fuentes de datos en herramientas de AI puede resultar útil para la adopción de decisiones de trading, de inversión, de concesión de créditos y otros aspectos comerciales, así como para la gestión de riesgos. La utilidad de aplicar software de AI en un contexto de persistente aumento de la capacidad de computación y de almacenamiento de datos resulta innegable.

Por otra parte, aunque en muchos proyectos en los que se ha aplicado la AI en la industria

<sup>3</sup> Véase Yurcan (2021).

<sup>4</sup> Véase Lin (2019).

<sup>5</sup> Véase In et al. (2019).

financiera las promesas han sido exageradas y los resultados a menudo decepcionantes, mirando al futuro las expectativas parecen altas y el potencial, enorme. La AI puede convertirse en una parte esencial de la tecnología en el sector de la banca, los servicios financieros y los seguros, y puede cambiar la forma de ofrecer productos y servicios. Analizando la naturaleza de las actividades que integran el mundo de las finanzas, es más que probable que esta industria se vea tan impactada por los desarrollos de inteligencia artificial como por el propio proceso de digitalización que ya ha tenido lugar. El impacto aislado de la AI previsiblemente sea incluso superior en esta industria que el que han tenido otras tecnologías emergentes como Internet de las Cosas (IoT), el Metaverso, e incluso la *distributed ledger technology* (DLT). Pero es que, además, los impactos de todas estas nuevas tecnologías se refuerzan entre sí.

Como ha ocurrido anteriormente con Fintech y con DeFI, la inteligencia artificial puede debilitar los lazos que han mantenido unidos los componentes de las instituciones financieras tradicionales y abrir la puerta a más innovaciones y nuevos modelos operativos. Esto

representa una gran oportunidad, pero también una seria amenaza para los incumbentes actuales. En este sentido, parece razonable pensar que la AI tiene un alto potencial de impacto positivo cuando se implementa con prudencia y diligencia. Y la aplicación de herramientas de AI puede traducirse en cambios significativos en el nivel de competencia, dependiendo de la capacidad de los agentes financieros a la hora de implementar adecuadamente este tipo de proyectos. La mayoría de las instituciones financieras tienen proyectos (al menos de tipo piloto) en este terreno, cuyos resultados se materializarán en cambios en las operaciones del día a día en un futuro próximo. Por ello, indudablemente, la AI va transformar el sector financiero en los próximos años.

Este Informe ofrece una visión amplia, pero en un nivel divulgativo, de los diferentes efectos que la aplicación de herramientas de inteligencia artificial está teniendo ya y puede tener en un futuro en la industria financiera. El trabajo no ofrece una visión de las experiencias particulares de las diferentes entidades financieras en este terreno. En ocasiones estas experiencias y proyectos

en inteligencia artificial son reveladas al mercado por motivos de comunicación corporativa, pero en muchas ocasiones son confidenciales, para preservar las ventajas competitivas esperadas de las inversiones realizadas.

En la primera parte del informe se comenta la situación actual del fenómeno de la AI y se describen brevemente las principales técnicas. También se revisan los problemas habituales que enfrentan las entidades financieras en la implementación de su estrategia de inteligencia artificial. La segunda parte del informe está dedicada a comentar las aplicaciones de tecnología de inteligencia artificial en los diferentes ámbitos de la industria: banca comercial, banca de inversión, banca privada, gestión de activos y compañías de seguros. También se describe cómo los bancos centrales y los

supervisores del mercado y las autoridades de defensa de la competencia están aplicando este tipo de herramientas. La tercera parte del informe se dedica a analizar los principales riesgos y amenazas que se derivan de la generalización del uso de tecnología de inteligencia artificial y a destacar algunas conclusiones.

El ámbito de investigación sobre el uso y sobre los efectos de AI en el mundo actual es sencillamente inabarcable, como también lo es el análisis de las aplicaciones de herramientas de inteligencia artificial en el ámbito más reducido de la industria financiera. Por ello, este informe únicamente ofrece una visión superficial y resumida de un fenómeno que está afectando ya al espacio de las finanzas de forma muy intensa.



## REFERENCIAS

Evident Insights (2023) *The Evident AI index Key Findings report*. Enero.

In, S. Y., Rook D., and Monk A. (2019) "Integrating alternative data (also known as ESG data) in investment decision making", *Global Economic Review*, 48 (3): 237-260.

Lin, T. C. (2019) "Artificial intelligence, finance, and the law", *Fordham Law Review*, 88: 531-551.



Yurcan, B. (2021) "Three Practical Applications for Artificial Intelligence in Payments", *The Financial Brand*. 20 de abril.

Yurcan, B. (2022) "AI in Banking: Top Priorities for 2022 (And Beyond)", *The Financial Brand*. 2 de mayo.



INSTITUTO  
ESPAÑOL  
DE ANALISTAS  
DESDE 1965  
FUNDACIÓN

El impacto de la inteligencia artificial en la  
industria financiera: promesas y amenazas

# Parte I





## 2. Situación actual de la inteligencia artificial

“Posiblemente la AI es la tecnología más avanzada en la que la humanidad está trabajando hoy en día y afectará a todas las industrias y a todos los aspectos de nuestra vida”.

*Sundar Pichai (2023) CEO de Google.*

“La AI ya está aquí. Ahora viene la parte dura: aprender cómo gestionarla y gobernarla”

*Mustafa Suleyman y Eric Schmidt (2023).*

La inteligencia artificial es la ciencia orientada a crear máquinas capaces de hacer cosas que requerirían inteligencia si lo hicieran los hombres. También se define como la tecnología aplicada a la simulación mediante software de algunas de las funciones cognitivas de los cerebros biológicos. O como el uso de ordenadores y algoritmos para simular y complementar la inteligencia humana. Así,

la inteligencia artificial puede entenderse como la capacidad de una máquina para realizar las funciones cognitivas que generalmente asociamos con la mente humana, tales como percibir, razonar, aprender, interactuar con un entorno, resolver problemas e incluso ejercitar la creatividad. La AI tiene sus fundamentos en la informática, las matemáticas y la estadística, pero también en

la lingüística, la psicología y la filosofía. Las herramientas de AI permiten el reconocimiento adaptativo de patrones utilizando grandes volúmenes de datos y métodos estadísticos para dar la respuesta de “mejor suposición” a cualquier conjunto de problemas definidos<sup>1</sup>. Se afirma también que en la actualidad la inteligencia artificial es, por encima de todo, la técnica enfocada a desarrollar máquinas de optimización.

En el mundo de la tecnología los eslóganes, las modas y las exageraciones son una constante en el tiempo, y se afirma que estamos hoy asistiendo a una auténtica hipérbole (*hype*) en la consideración del impacto de la inteligencia artificial. Se compara el impacto potencial de AI con el que tuvo en su día la electricidad en la economía y en la vida cotidiana. Se afirma que la inteligencia artificial es la revolución más grande de nuestro tiempo y el principal catalizador para la Industria 4.0. Incluso Marc Andreessen ha comparado la AI con la “piedra filosofal”. Lo cierto es que la inteligencia artificial se encuentra ya por doquier de forma visible e invisible para los consumidores y trabajadores.

Lo que hoy conocemos como inteligencia artificial tuvo su origen en los desarrollos matemáticos e informáticos de los años 50. El término se acuñó en 1956 en una conferencia en el Dartmouth College en Estados Unidos. Ahora estamos viviendo el tercer boom en AI. La primera y segunda etapa tuvieron lugar en los años 70 y en los 80. No mucho después de la exuberancia en ambos periodos se asistió a sendas desilusiones sobre el alcance real de las aplicaciones basadas en herramientas AI. Esto condujo a una caída en la inversión y en la financiación de la investigación sobre esta rama de la informática, lo que vino a llamarse el “invierno de la AI”. Son numerosos los expertos que predicen que esta vez es definitivamente diferente, porque se ha producido un cambio relevante en la capacidad de computación, en las técnicas de optimización y en la disponibilidad de datos.<sup>2</sup> Estos desarrollos han permitido avances espectaculares tanto en inteligencia artificial específica como en inteligencia artificial general y generativa, de forma que nos encontramos en un estadio AI 2.0. La mayor parte de

<sup>1</sup> Véase Chan *et al.* (2023)

<sup>2</sup> Véase Chow *et al.* (2019).

los profesionales considera hoy que la inteligencia artificial tendrá en nuestras vidas una repercusión al menos tan fuerte como la que implicó la penetración del ordenador personal, Internet y los teléfonos inteligentes. Y el mundo de las finanzas no va a ser una excepción en cuanto a sus efectos.

Por una vez, la exageración quizá está justificada ya que la AI presenta una velocidad de adopción sin precedentes, puede derribar barreras de entrada en diversas industrias y, de alguna forma, alterar las relaciones económicas y sociales. La AI es una tecnología de propósito general, polivalente, que puede transformar el funcionamiento económico, alterar las capacidades competitivas de empresas, las relaciones sociales, la política, la educación y la ciencia y también, por tanto, alterar el poder relativo de grupos y países. La AI puede incluso cambiar la forma en la que pensamos y creamos y, en definitiva, cómo entendemos nuestro lugar en el mundo<sup>3</sup>. En opinión de algunos analistas, la inteligencia artificial puede, por otra parte, crear la mayor concentración conocida de poder económico en un puñado de empresas de

tecnología e incluso anular nuestra comprensión de la verdad.

El avance en los desarrollos en inteligencia artificial ha sido espectacular en AI específica tanto predictiva como generativa. La AI ha dejado de ser materia de ciencia ficción para convertirse en una realidad. En todo caso, a diferencia de la *inteligencia artificial débil*, que sólo es capaz de ejecutar una tarea, la *inteligencia de propósito general* (AGI) o *God-like AI* igualaría o excedería la *inteligencia humana media* en varios aspectos. Esto implicaría alcanzar un sistema con capacidades de rivalizar con las de un humano, por lo que respecta a capacidades cognitivas, incluyendo el razonamiento, la resolución de problemas, la percepción, el aprendizaje y la comprensión del lenguaje, así como en capacidades emocionales y sociales (como la empatía) y en otras capacidades como la percepción visual, auditiva y motora, en orientación y en creatividad. En este estadio, las capacidades de la IA serían indistinguibles de las de un ser humano y se habría superado lo que se conoce como prueba de Turing. Muchos investigadores creen que todavía

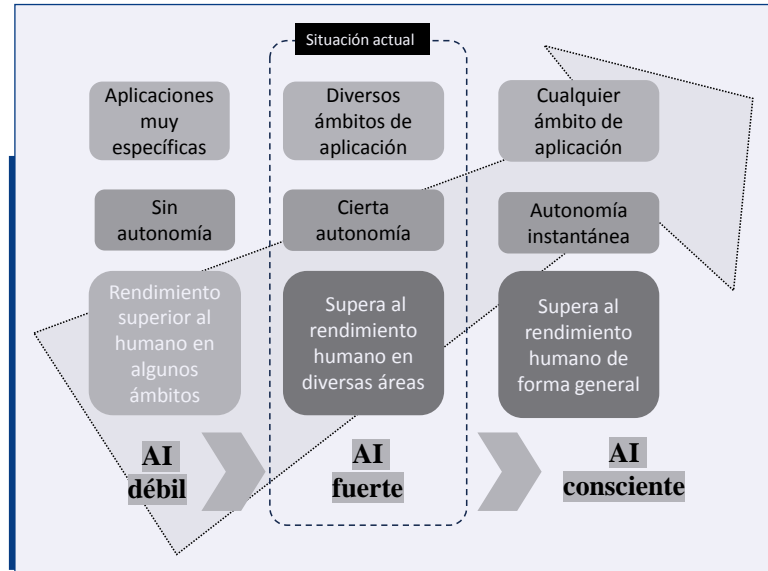
<sup>3</sup> Véase Wolf (2023).



estamos a décadas, si no siglos, de alcanzar el AGI<sup>4</sup>. En este estadio, *una superinteligencia artificial* sería incluso capaz de entender el mundo de forma relativamente global, tener sentido común, gestionar varios tipos de aprendizaje por sí

misma, comprender sistemas simbólicos y usar conocimiento metacognitivo. Una máquina con superinteligencia podría educarse a sí misma con una velocidad inaudita y multiplicar sus capacidades en poco tiempo.

Gráfico 2.1.  
Evolución de  
capacidades de la  
inteligencia artificial.



Fuente: elaboración propia.

El avance de las funcionalidades de la AI mejora con la potencia informática instalada, la disponibilidad de datos y el número de usuarios. Se trata de una tecnología cuyo desempleo en muchas ocasiones se comporta

mejor de forma inesperada a la hora de filtrar datos, reconocer imágenes y voz, identificar sentimientos en los textos, generar texto, imágenes, e incluso código. Es lo que se conoce como *propiedades emergentes*. Algunos

<sup>4</sup> McKinsey and Company (2024).

sistemas de AI están aprendiendo por sí mismos habilidades que no se esperaba que tuvieran y, además, no se comprende bien cómo sucede esto. Las iteraciones y la retroalimentación de las herramientas generarán en el futuro nuevas sorpresas, lo que para algunos resulta preocupante.

El aprendizaje de las herramientas de AI, se verá reforzado por el aumento en el número de usuarios. Y, en este contexto, las barreras de entrada para la adopción de este tipo de tecnología por parte de los usuarios son decrecientes. Casi cualquiera de los cinco mil millones de personas con acceso a Internet puede, ahora mismo, operar un *chatbot* basado en un modelo LLM de forma gratuita y con poca o ninguna capacitación. El ritmo de desarrollo de la AI es asombroso y la atención científica, empresarial y de los medios de comunicación en torno a estas herramientas, que había venido aumentando en intensidad desde 2017, estalló a mediados de 2022, y particularmente a principios de 2023. A ello contribuyó el lanzamiento de proyectos, como [ChatGPT](#) y Dall-E y API por

parte de OpenAI (asociada con Microsoft), Cloud AI Platform, Bard, Magi y Gemini por parte de Google, Midjourney por Midjourney Lab, Gato por parte de DeepMind, Tongyi Qianwen por parte Alibaba, x.AI por parte del entorno corporativo de Elon Musk, Claude por parte de Anthropic, o RAG por parte de Cohore, entre otros. Para dar una idea del ritmo de penetración cabe señalar que ChatGPT solo tardó cinco días en alcanzar un millón de usuarios<sup>5</sup>.

La AI generativa está dando lugar a la creación de todo un ecosistema, desde proveedores de hardware hasta creadores de aplicaciones, que ayudará a que su impacto en los negocios y en la vida personal sea determinante. La AI «tradicional» se caracteriza por que el hecho de que los algoritmos se entrenan para realizar una sola tarea, el entrenamiento se realiza con supervisión humana y utilizando datos específicos para esa tarea. Como resultado, se pueden realizar ajustes para alcanzar una alta precisión, pero el entrenamiento debe repetirse para cada nuevo caso de uso. En contraste, en la *Gen-AI* los

<sup>5</sup> Recordemos que Apple había tardado más de dos meses en alcanzar el mismo nivel de adopción para su iPhone; Facebook (ahora Meta) tuvo que esperar diez meses y Netflix más de tres años para construir la misma base de usuarios.

modelos básicos (*foundations models*) se entrenan previamente con conjuntos de datos masivos y los algoritmos se pueden adaptar a una amplia variedad de tareas posteriores, incluida la generación de contenido.

En esta situación, se ha despertado un debate sobre las oportunidades que ofrece la *Gen-AI* y los riesgos que un desarrollo sin control puede significar para la humanidad. Algunos esperan que su uso masivo contribuya a que la humanidad dé un importante salto adelante, al permitir un intenso y continuado avance en la productividad, al tiempo que ayuda a la especie humana en el desarrollo de la ciencia. Además, la *AI* generativa está llevando la tecnología a un ámbito que se creía exclusivo de la mente humana: la creatividad. El uso de estas aplicaciones podría traducirse en un mayor nivel de creatividad e innovación en el mundo si los humanos son estimulados con nuevas ideas iniciales. Otros, por el contrario, predicen un escenario de mayor concentración de poder económico y político, más discriminación y un mayor riesgo de desinformación generalizada. Las máquinas han hecho posible la vida tal como la conocemos, pero los humanos han temido siempre

la posibilidad de que éstas adquiriesen algún día inteligencia humana y se “lanzasen por su cuenta”. Se ha afirmado que el desarrollo de los sistemas de inteligencia artificial va demasiado rápido y que los desarrolladores de *AI* están compitiendo en una carrera fuera de control para desplegar mentes digitales cada vez más poderosas que nadie, ni siquiera sus creadores, pueden entender, predecir o controlar de manera confiable. El panorama actual se ha descrito como un escenario en el que un conjunto de empresas y gobiernos está corriendo hacia una línea de meta sin comprender suficientemente lo que hay al otro lado. El *Future of Life Institute* ha solicitado un freno a la vorágine del desarrollo de herramientas *AI* para evitar riesgos profundos para la sociedad y la humanidad en sí misma. Esta pausa daría tiempo para introducir protocolos de seguridad en *AI* y asegurar que su avance sólo se produce cuando existan garantías de que sus efectos serán positivos y sus riesgos manejables. Los analistas menos recelosos en este ámbito recuerdan que aún no se ha llegado a construir un ordenador superinteligente que aprenda y se desarrolle de forma autónoma, que comprenda su entorno sin necesidad de supervisión y que pueda



transformar el mundo que le rodea. Pero como señala Hogarth, la naturaleza de la tecnología hace que sea excepcionalmente difícil predecir con exactitud cuándo podría llegarse a ese punto, ya que la capacidad de computación para entrenar modelos de aprendizaje en AI se ha multiplicado por un factor de cien millones en los últimos 10 años. Resulta complicado pensar que las empresas y los gobiernos vayan a realizar una pausa en el desarrollo de AI, porque también en esta tecnología existe la ventaja del primer movimiento (*first-mover-advantage*). Tanto en el terreno de la competencia corporativa como en geopolítica el desarrollo de AI se ha convertido en una cuestión crítica.

Según Autonomous, una firma de investigación financiera, el número de dispositivos informáticos inteligentes (que se estima en 22 mil millones) supera a la población actual en proporción de tres a uno. Y la cantidad de dispositivos inteligentes seguirá aumentando, con empresas de capital de riesgo invirtiendo decenas de miles de millones de dólares anuales en empresas que impulsa dispositivos de AI conectados digitalmente.

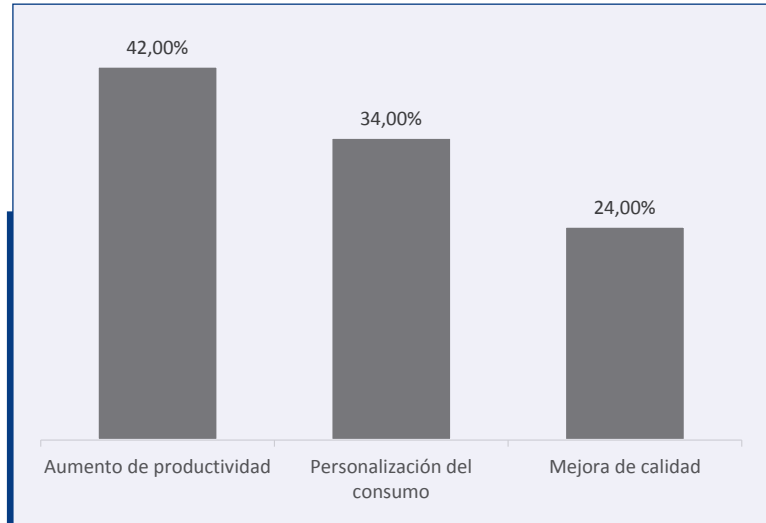
Según un trabajo de IDC, el gasto mundial de gobiernos y empresas en estas tecnologías habría superado los 500.000 millones de dólares en 2023. Sólo las empresas privadas especializadas en AI y GenAI han recibido recientemente más de 21.000 millones de dólares en inversiones, de los cuales 11.000 millones se han concentraron en el primer trimestre de 2023. La firma de investigación de mercado IDC ha proyectado que el mercado global de AI alcanzará un tamaño de más de 1,5 billones de dólares en 2030.

Respecto de los efectos económicos del despliegue de la inteligencia artificial, un estudio de PWC estima que la AI podría aportar en torno a 15,7 billones de dólares a la economía mundial en 2030, más que la producción actual de China e India juntas. De este importe, 6,6 billones de dólares procederían de aumentos en la productividad<sup>6</sup>. Véase en el Gráfico 2.2 el desglose de las fuentes de aumento del PIB.

<sup>6</sup> Véase Rao et al. (2017).



Gráfico 2.2. Fuentes de la mejora del valor de producción inducida por la aplicación de herramientas de AI.



Fuente: Basado en Rao et al. (2017).

En el caso del sector financiero, según las estimaciones de Insider, las aplicaciones de AI podrían estar ya ahorrando en una cifra cercana a 440.000

millones de dólares para las entidades de esta industria, con el desglose por áreas que se muestra en el Gráfico 2.3.

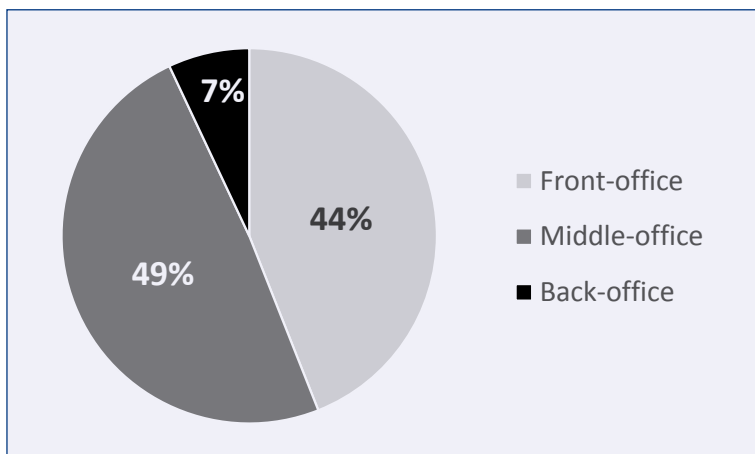


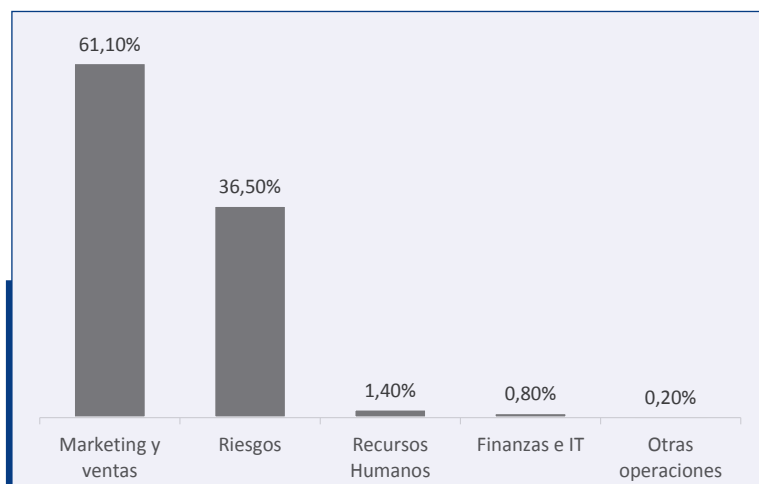
Gráfico 2.3. Distribución de los ahorros de los costes por áreas en la industria bancaria.

Fuente: Basado en Digalaki (2022).

Se ha calculado también, por parte de McKinsey and Co, que el uso de la inteligencia artificial en la industria financiera puede potencialmente desbloquear

en torno a un billón de dólares anuales de valor incremental, con la distribución por funciones que se muestra en el Gráfico 2.4.

Gráfico 2.4.  
Distribución por funciones de la creación de valor en la industria bancaria derivada de la implementación de herramientas de AI.



Fuente: Basado en Biswas et al. (2020).

De acuerdo con el McKinsey Global Institute (MGI) en el sector bancario global, la AI generativa podría agregar un valor anual de entre \$200 mil millones y \$340 mil millones de dólares, lo que implicaría entre el 2,8 y el 4,7 por ciento de los ingresos totales de la industria. Este efecto de derivaría en gran medida de mejoras en la productividad<sup>7</sup>.

Según Autonomous, en 2030

las instituciones financieras tradicionales podrán reducir los costes en un 22 %. Pero, además de mejorar la eficiencia, mediante el uso de herramientas de AI contribuirá al diseño de nuevos productos y servicios. Según el informe de *Insider Intelligence* en torno a un 80% de los bancos son muy conscientes de los beneficios potenciales que puede presentar la adopción de tecnologías AI. En la misma línea según la encuesta

<sup>7</sup> McKinsey Global Institute (2023).

de NVIDIA, el 83% de los directivos (de los cuales el 81% pertenecen a la alta dirección) coincidían en que la AI era importante para el éxito futuro de su organización. En la misma encuesta, el 34% de los profesionales de servicios financieros consideraba que la AI aumentaría los ingresos anuales de su empresa en al menos un 20%<sup>8</sup>.

Por otra parte, según una encuesta de Deloitte, el 86% de los que adoptan la inteligencia artificial en los servicios financieros afirman que la AI será muy importante o crítica para el éxito estratégico de su empresa en los próximos dos años<sup>9</sup>. Véase gráfico 2.5.

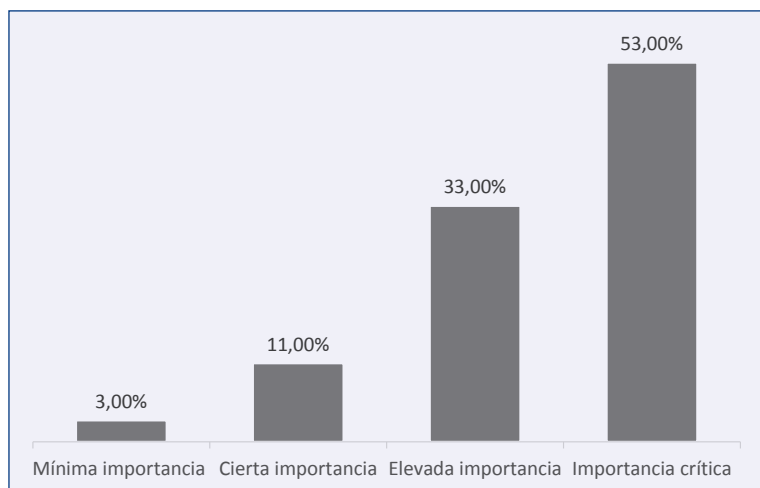


Gráfico 2.5. Grado de importancia estratégica de la AI en el plazo de 2 años para compañías del sector financiero.

Fuente: Basado en Deloitte (2022).

Según el informe del IBM Institute for Business Value (2024), el 78% de los bancos de todo el mundo ya están probando o implementando aplicaciones tácticas de IA generativa que van desde la gestión de riesgos hasta los servicios al cliente.

En un análisis de Accenture se indicaba que, entre 2018 y 2022, los bancos que invirtieran en AI y en instrumentos de colaboración hombre-máquina al mismo ritmo que las empresas con mejores resultados podrían aumentar sus ingresos un 34%

<sup>8</sup> Véase NVIDIA (2021).

<sup>9</sup> Véase Deloitte (2022).

de media y su empleo un 14%. Otra investigación de Accenture muestra que el 90% de todas las horas de trabajo en el sector bancario pueden verse afectadas de algún modo por los modelos LLM y que el 54% del tiempo de trabajo del sector financiero tiene un mayor potencial de automatización mediante AI. El informe prevé que para 2028, el sector registrará un aumento del 30% en la productividad del conjunto de empleados dedicados a operaciones desde atención al

público y a taras administrativas<sup>10</sup>. Según la investigación de Accenture, utilizando herramientas de inteligencia artificial los bancos podrían lograr un aumento de entre 2 y 5 veces en el volumen de interacciones o transacciones con la misma plantilla, lo que implica que podrían alcanzar ratios de eficiencia (coste/ingresos) por debajo del 30%<sup>11</sup>.



## REFERENCIAS

- Abbot, M., Murray, J., and Smith, K. (2023) *Breaking barriers: Exploring how banks scale generative AI for growth*. 23 de mayo. Accenture.
- Adora, P. (2021) *Where smart banking meets AI*. Accenture.
- Biswas, S., Carson, B., Chung, V., Singh, S. and Thomas, R. (2020) "AI-bank of the future: Can banks meet the AI challenge?", *McKinsey and Co. Global Banking & Securities*.
- Chan, C, Chow, C., Wong, J., Dimakis, N., Nayler, D., Bermudes, J., Raman, J., Lam, R., and Baker, M. (2023) *Artificial Intelligence Applications in Financial Services. Asset Management, Banking, and Insurance*. Marsh and McLennan Companies.
- Chow, C., Frame, K., Likhman, S., Spooner, N. y Wong, J. (2019) *Investor's expectations on responsible artificial intelligence and data governance*. Hermes Investment Management.
- Deloitte (2022) *Artificial intelligence: Transforming the future of banking*.
- Digalaki, E. (2022) "The impact of artificial intelligence in the banking sector & how AI is being used in 2022", *Business Insider*.
- IBM Institute for Business Value (2024) "Global Outlook for Banking and Financial Markets: Regenerate banking with AI". IBM.

<sup>10</sup> Véase Abbot et al. (2023).

<sup>11</sup> Véase Adora (2021).



- McKinsey Global Institute (2023) "The economic potential of generative AI: the next productivity frontier". 14 junio de 2023.
- McKinsey and Company (2024) "What is artificial general intelligence (AGI)?", 21 de marzo.
- NVIDIA (2021) *The State of AI in Financial Services: at a Glance*.
- Pichai, S. (2023) "Building AI responsibly is the only race that really matters", *Financial Times*, 23 de mayo.
- Rao, A. S. and Verweij, G. (2017) *Sizing the prize What's the real value of AI for your business and how can you capitalise?* Price Waterhouse Coopers.
- Suleyman, M. and Schmidt, E. (2023) "We need an AI equivalent of the IPCC", *Financial Times*. 19 de octubre.
- Wolf, M. (2023) "The threat and the promise of AI", *Financial Times*, 9 de mayo.



### 3. Principales técnicas de inteligencia artificial<sup>12</sup>

“Desde que se inventaron los ordenadores nos hemos preguntado si se podría hacer que aprendieran. Si pudiéramos entender cómo programarlos para que aprendiesen, y mejorasen automáticamente con la experiencia, el impacto sería espectacular”.

*Tom Mitchell (1997).*

La inteligencia artificial abarca una amplia variedad de enfoques, cada uno con sus propias limitaciones y ventajas. En este capítulo se ofrece una descripción divulgativa básica de algunas de las técnicas que dan soporte a los desarrollos de inteligencia artificial. Se comenta para cada una de ellas, su fundamento, en algunos

casos los pasos que requieren, así como sus principales ventajas e inconvenientes.

Principales técnicas:

1. **Sistemas expertos.** Los sistemas expertos son programas de computadora que emulan la toma de decisiones de un experto humano en un

<sup>12</sup> En la elaboración de este capítulo el autor se ha beneficiado del uso de ChatGPT.4 de OpenAi.



dominio específico.

## 2. Aprendizaje automático

(*machine learning*). El

aprendizaje automático es un enfoque clave en la AI, que implica el uso de algoritmos y modelos estadísticos para permitir que las máquinas

aprendan a partir de datos y mejoren su rendimiento con la experiencia.

Algunas subcategorías del aprendizaje automático incluyen diversas técnicas de *aprendizaje supervisado*, como las que se muestran en el Gráfico 3.1.

Regresión lineal	Técnica utilizada para predecir un valor numérico (una variable continua) basada en una o más variables de entrada. El objetivo es encontrar la mejor línea recta que se ajuste a los datos de entrenamiento.
Regresión logística	Se utiliza para tareas de clasificación binaria. Estima la probabilidad de que un ejemplo pertenezca a una de las dos clases posibles.
Árboles de decisión	Modelos de predicción creados a partir de datos de entrenamiento. Los árboles se dividen en nodos, y las decisiones se toman a medida que se desciende por el árbol. Se pueden utilizar tanto para tareas de clasificación como de regresión.
Random Forest	Conjunto de árboles de decisión, que se combinan para mejorar la precisión y reducir el sobreajuste. Se utiliza comúnmente para clasificación y regresión.
SMV	Las máquinas de soporte vectorial se utilizan en problemas de clasificación. Buscan encontrar el hiperplano que mejor separa las clases en un espacio de alta dimensión. Pueden manejar problemas de clasificación lineal y no lineal.
K-NN	K-Nearest Neighbors es una técnica de clasificación y regresión. Clasifica los datos basados en la mayoría de los "k" ejemplos más cercanos en el espacio de características.
Naive Bayes	Algoritmo para tareas de clasificación de texto, como el filtrado de spam o la clasificación de documentos. Se basa en el teorema de Bayes.
Gradient Boosting	Un enfoque de conjunto que combina múltiples modelos para crear uno más poderoso. Ejemplos: Gradient Boosting Machines (GBM) y XGBoost.
Deep Learning	El aprendizaje profundo se enfoca en el uso de redes neuronales profundas para el procesamiento de imágenes, de lenguaje natural, etc. Ejemplos: redes neuronales convolucionales (CNN) y redes neuronales recurrentes (RNN).

Gráfico 3.1. Algunas técnicas de aprendizaje supervisado.

Fuente: elaboración propia.

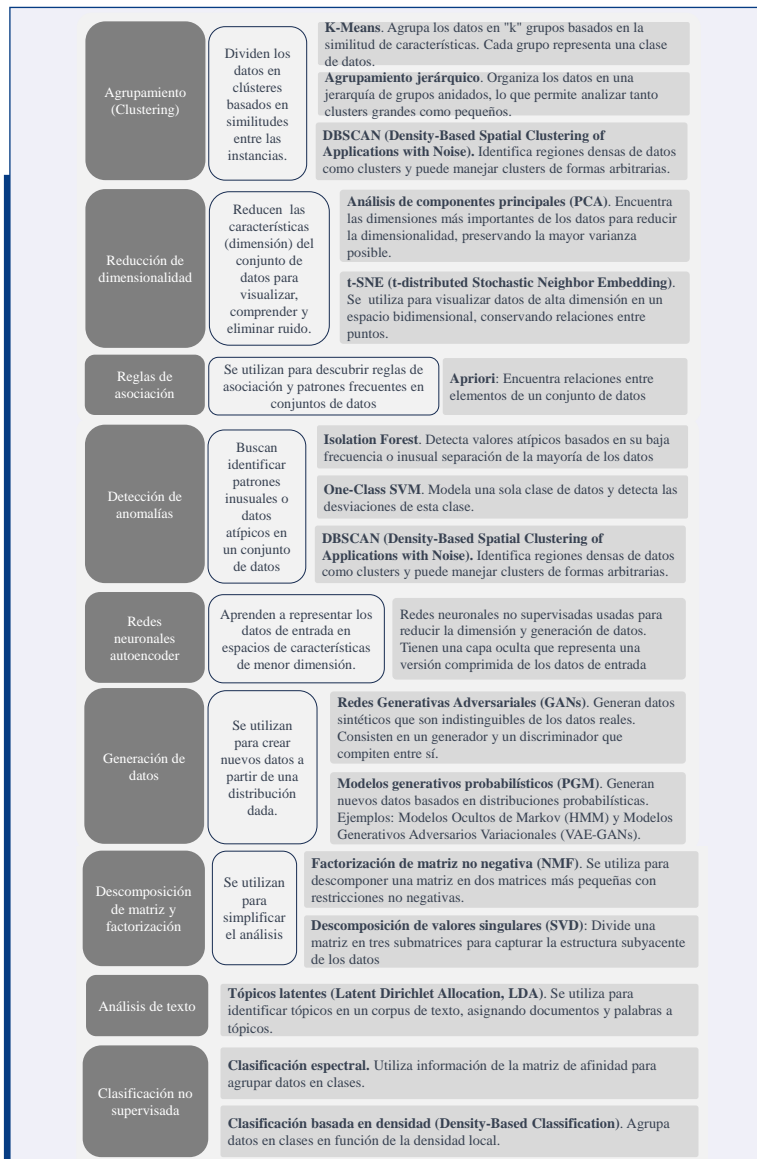
Por otra parte, se han desarrollado diversas técnicas de *aprendizaje no supervisado*, en la que el modelo se entrena utilizando datos no etiquetados, es decir, datos que no tienen etiquetas de

clase o categorías predefinidas. El objetivo principal del aprendizaje no supervisado es descubrir patrones, estructuras o relaciones en los datos sin una orientación específica hacia una tarea de clasificación

o regresión. Algunas de estas técnicas y sus subcategorías

se muestran en el Gráfico 3.2.

Gráfico 3.2. Algunas técnicas de aprendizaje supervisado.



Fuente: elaboración propia.



También se han desarrollado técnicas denominadas *aprendizaje por refuerzo*. Se trata de un enfoque de aprendizaje automático en el que un agente interactúa con un entorno y toma decisiones para maximizar una recompensa acumulada a

lo largo del tiempo. El agente aprende a través de la prueba y error, ajustando sus acciones en función de la retroalimentación que recibe del entorno. Un resumen de diversas técnicas de aprendizaje por refuerzo se ofrece en el Gráfico 3.3.

Gráfico 3.3.  
Algunas técnicas  
de aprendizaje por  
refuerzo.

RL basado en valor	<b>Q-Learning.</b> Un algoritmo de aprendizaje por refuerzo que aprende una función Q que estima la recompensa acumulativa esperada para cada par estado-acción. Se utiliza para problemas de aprendizaje por refuerzo discretos.
RL basado en políticas	<b>Métodos de Monte Carlo.</b> Estiman directamente la política óptima y el valor esperado de una política utilizando muestreo. Son útiles para problemas en los que se desconoce el modelo del entorno. <b>Aprendizaje profundo de políticas (DRL).</b> Utiliza redes neuronales profundas para aprender políticas óptimas en entornos complejos y de alta dimensión
Exploración y explotación	<b><math>\epsilon</math>-Greedy:</b> Un enfoque simple en el que el agente elige la acción óptima con probabilidad $1-\epsilon$ y una acción aleatoria con probabilidad $\epsilon$ . <b>Upper Confidence Bound (UCB).</b> Un algoritmo que selecciona acciones basadas en una combinación de estimaciones de valor y una medida de incertidumbre. <b>Thompson Sampling.</b> Utiliza distribuciones Bayesianas para modelar la incertidumbre y tomar decisiones basadas en muestras aleatorias de éstas.
Aprendizaje profundo en RL	<b>Deep Q-Network (DQN):</b> Combina el aprendizaje por refuerzo con redes neuronales profundas para manejar problemas con espacios de estado y acción continuos. <b>Policy Gradients.</b> Optimiza directamente la política del agente utilizando gradientes y se utiliza para problemas de control continuo.
RL basado en modelos	<b>Model-Based RL.</b> El agente aprende un modelo del entorno y luego utiliza ese modelo para planificar y tomar decisiones. <b>Imaginative Planning (Imagination).</b> Este enfoque utiliza simulaciones del entorno para planificar y mejorar la política del agente.
Aprendizaje profundo de adversarios (Adversarial RL):	<b>Generative Adversarial Imitation Learning (GAIL).</b> Utiliza una red generativa adversaria para aprender políticas de comportamiento humano a partir de datos de expertos. <b>Proximal Policy Optimization (PPO).</b> Este algoritmo entrena una política basada en gradiente con penalización de la desviación KL para mejorar la estabilidad y el rendimiento.
Aprendizaje por refuerzo con restricciones	<b>Constraint Learning:</b> Introduce restricciones en la política del agente para garantizar un comportamiento deseado o seguro..
Aprendizaje por refuerzo profundo en tiempo real	<b>Deep Reinforcement Learning from Human Feedback (RLHF).</b> Los agentes de RL pueden aprender de la retroalimentación de los humanos, lo que es útil en entornos donde la recompensa es difícil de definir.

Fuente: *Elaboración propia.*

3. **Redes neuronales artificiales:**  
 Estas son estructuras de aprendizaje profundo inspiradas en el funcionamiento del cerebro humano. Las redes neuronales se utilizan en una

amplia gama de aplicaciones, como el procesamiento de imágenes, el procesamiento de lenguaje natural y la visión por computadora. Véase Gráfico 3.4.

<b>Arquitectura de la red neuronal</b>	<b>Perceptrón:</b> La unidad básica de una red neuronal es el perceptrón. Varias capas de perceptrones se organizan en una red para formar la arquitectura de la red.	
<b>Funciones de activación</b>	<b>Sigmoide:</b> Utilizado en redes neuronales tradicionales para clasificación binaria	<b>ReLU (Rectified Linear Unit)</b> comúnmente usado en redes profundas. Ayuda a evitar el problema de desaparición del gradiente
<b>Feedforward</b>	<b>Alimentación directa:</b> los datos fluyen a través de la red desde la capa de entrada hasta la capa de salida sin bucles ni retroalimentación	
<b>Backpropagation</b>	<b>Retropropagación:</b> algoritmo para ajustar las ponderaciones de las conexiones en función de la diferencia entre la salida deseada y la salida real. Ayuda a minimizar el error de la red	
<b>Capas ocultas</b>	Capas entre la entrada y la salida que extraen características y representaciones intermedias de los datos	
<b>Regularización</b>	<b>Dropout:</b> Técnica que apaga aleatoriamente un porcentaje de las neuronas durante el entrenamiento para evitar el sobreajuste	<b>L1 y L2 Regularización:</b> añaden términos de penalización a la función de coste para evitar ponderaciones excesivas
<b>Inicialización de ponderaciones</b>	Asegura que los pesos iniciales de la red no se acercan a cero para acelerar el proceso de entrenamiento	

Gráfico 3.4. Principios y técnicas clave en el diseño y entrenamiento de redes neuronales.

4. **Procesamiento de lenguaje natural (NLP).** El NLP se centra en la interacción entre las computadoras y el lenguaje humano. Se utiliza para tareas como traducción automática, *chatbots*, análisis de sentimientos, resumen de texto y generación de texto.
5. **Visión por computadora.** La visión por computadora permite que las máquinas interpreten y comprendan

imágenes y videos. Se aplica en reconocimiento de objetos, detección de rostros, análisis de imágenes médicas y vehículos autónomos.

6. **Minería de datos.** La minería de datos implica el análisis de grandes conjuntos de datos para descubrir patrones, tendencias y conocimientos ocultos. Es una técnica clave para la toma de decisiones basada en datos.



7. **Lógica difusa** (*Fuzzy Logic*). La lógica difusa se utiliza en situaciones donde la lógica binaria (verdadero/falso) no es suficiente. Permite representar grados de verdad en lugar de valores absolutos, lo que es útil en sistemas de control y toma de decisiones vagas.

8. **Algoritmos genéticos**. Inspirados en la teoría de la evolución, los algoritmos genéticos se utilizan para la optimización y búsqueda heurística. Se aplican en problemas de optimización combinatoria y diseño de sistemas.

9. **Aprendizaje profundo** (*Deep Learning*). El aprendizaje profundo es una rama del aprendizaje automático que utiliza redes neuronales profundas para modelizar datos complejos y extraer características de alto

nivel. Es fundamental en aplicaciones de visión por computadora y procesamiento de lenguaje natural.

10. **Robótica**. La robótica combina la AI con la ingeniería para crear máquinas autónomas capaces de realizar tareas físicas.

La inteligencia artificial continúa evolucionando, y se desarrollan constantemente nuevas técnicas y enfoques para abordar una amplia gama de desafíos. En el Anexo A se ofrece una descripción más detallada de algunas de las principales técnicas de inteligencia artificial, así como de sus procesos de construcción y entrenamiento. También se comentan las ventajas y limitaciones a las que se enfrentan estos modelos.



## REFERENCIAS

Mitchell, T. (1997) *Machine Learning*. Springer.



## 4. Implantando una estrategia de AI en instituciones financieras

Nuestra mejor oportunidad de reinventarnos es responder a una pregunta aparentemente sencilla: ¿Qué es posible ahora en la era de las máquinas inteligentes que antes ni siquiera era concebible?

*Mike Walsh (2020)*

El diseño de una estrategia de implantación de la inteligencia artificial en las entidades financieras implica un dilema en cuanto a su alcance y, básicamente, hay dos enfoques polares que pueden ser adoptados:

- Por una parte, un *incrementalismo digital*, orientado a la automatización con el fin de reducir costes y mejorar los procesos mediante la incorporación gradual de proyectos de inteligencia artificial.
- Por otra parte, es posible visualizar una *transformación digital completa basada en la AI* para convertir a la entidad en el centro de la vida financiera de los clientes y formar parte de un ecosistema más amplio. Esto implica automatizar los flujos de trabajo manual, pero también llevar a cabo una revisión completa de los modelos de negocio<sup>13</sup>, la

<sup>13</sup> Véase Ryll et al. (2020).

implantación de una cultura de innovación respaldada por la AI, la reinención y mejora de productos, servicios y procesos, así como la implantación de un sistema hiperpersonalizado de comunicación y recomendaciones a los clientes.

La elección de uno u otro enfoque por parte de las entidades depende de las expectativas de alcanzar un grado suficiente de madurez en el ámbito de la inteligencia artificial. A tal efecto, cabe reconocer que la implantación de la tecnología AI en las empresas y en las instituciones financieras enfrenta un conjunto de retos relativos a aspectos como:

- **Talento.** La inteligencia artificial es un campo altamente especializado y en constante evolución. Resulta determinante encontrar profesionales con habilidades técnicas en este ámbito y conocimientos sólidos para diseñar, implementar y mantener este tipo de proyectos. Dado que la implantación de proyectos de AI es un fenómeno relativamente reciente, la falta de experiencia y conocimientos técnicos es un desafío común en la implementación de este tipo de proyectos. La demanda

creciente de profesionales en este ámbito ha generado una escasez de talento para el desarrollo de este tipo de proyectos, con el consiguiente incremento en el coste laboral. Conforme resulta más evidente el alcance y la tipología de proyectos de AI que desean desarrollar los diferentes bancos están creando puestos ad-hoc en diferentes niveles incluyendo posiciones "Director of Generative AI", "Generative AI Solution Architect", "Senior Data Scientist with Generative AI focus", etc.

Una estrategia que algunas entidades están desarrollando para mitigar el problema de la escasez de talento es la colaboración con instituciones académicas, centros de investigación y consultores especializados en AI. A medio y largo plazo la solución pasa por la combinación de la capacitación interna de los empleados y el fomento de la participación en comunidades y redes profesionales relacionadas con la AI.

- **Infraestructuras técnicas.** La recopilación, el almacenamiento y el procesamiento de grandes conjuntos de datos puede



plantear desafíos en términos de infraestructura y capacidad de almacenamiento. Los proyectos de AI a menudo requieren recursos computacionales potentes, como servidores de alto rendimiento (unidades centrales de procesamiento y unidades de procesamiento gráfico) así como alta capacidad de almacenamiento. La infraestructura y los recursos adecuados son costosos y requieren inversiones significativas. En cuanto a la potencia de cálculo, los proyectos de inteligencia artificial exigen habitualmente de un importante poder de procesamiento para entrenar y ejecutar sistemas complejos. En cuanto al almacenamiento, la elevada capacidad viene requerida por el hecho de que los proyectos de AI suelen trabajar con grandes volúmenes de datos. Además, la velocidad de acceso a los datos es también importante para garantizar un rendimiento eficiente durante el entrenamiento y la inferencia de los modelos.

Uno de los principales desafíos a los que se enfrentan los gestores de proyectos de inteligencia artificial es su escalabilidad. Esto

es, la capacidad que agregar recursos cuando es necesario, sin afectar significativamente el rendimiento y la disponibilidad del sistema. Sucede que este tipo de proyectos necesitan escalar en términos de recursos computacionales a medida que crecen y manejan mayores volúmenes de datos. Las soluciones en la nube pueden ofrecer ventajas significativas para la implementación de este tipo de proyectos. En este sentido, es preciso adoptar decisiones basadas en criterios de coste, seguridad y rapidez para desplegar infraestructuras propias y/o contratar proveedores de servicios en la nube, que ofrecen recursos computacionales y de almacenamiento escalables bajo demanda. Del mismo modo, es necesario tener disponibilidad de *frameworks* y bibliotecas de software para el procesamiento y entrenamiento de modelos. Se trata de herramientas con interfaces fáciles de usar, diseñadas para optimizar el consumo de recursos computacionales, el rendimiento y la eficiencia. Finalmente, la infraestructura debe incorporar herramientas de gestión y monitorización para supervisar el uso de CPU y GPU, la memoria, la velocidad de acceso



a los datos y otros indicadores clave para identificar cuellos de botella y realizar ajustes cuando sea necesario.

- **Gestión de datos.** Los proyectos de AI requieren grandes cantidades de datos de calidad para entrenar y mejorar los modelos. La disponibilidad de datos en cantidad y calidad es un desafío crítico en la implementación de proyectos de inteligencia artificial. En algunos casos, la dificultad de acceder a datos necesarios se debe a restricciones legales, a políticas de privacidad o a limitaciones de acceso a sistemas externos. El uso de datos en proyectos de AI plantea preocupaciones éticas y de privacidad. Es esencial asegurarse de que los datos utilizados sean legítimos y estén protegidos adecuadamente para garantizar la privacidad de los individuos, y para cumplir con las regulaciones de protección de datos. La calidad de los datos, por otra parte, es esencial para que los resultados sean precisos y fiables. Esto obliga en muchas ocasiones a realizar una limpieza y un preprocesamiento exhaustivo de los datos para eliminar valores duplicados, errores y

valores atípicos, así como para manejar el problema de los “datos que faltan” y garantizar la coherencia y consistencia del conjunto de datos. Cuando se trata de proyectos de aprendizaje automático supervisado, los datos necesitan ser etiquetados para su incorporación en los modelos, pero esta tarea puede ser lenta y costosa.

---

*Uno de los principales desafíos a los que se enfrentan los gestores de proyectos de inteligencia artificial es su escalabilidad.*

---

Aunque en el caso de las técnicas de aprendizaje semi-supervisado y no supervisado se reduce el esfuerzo de etiquetado, continúa siendo necesario garantizar la calidad de las etiquetas. Para mitigar los problemas de resultados poco equitativos o imparciales en las predicciones y en las decisiones adoptadas por los modelos de



AI es necesario, además, asegurar que los datos están exentos de sesgos sociales y culturales, lo que exige un trabajo de supervisión previa. Finalmente, la implantación de proyectos de AI exige diseñar actualizaciones regulares de los datos, en función de los cambios en el entorno y las necesidades del proyecto, para que los modelos mantengan un rendimiento óptimo. Es importante establecer mecanismos para recopilar y actualizar los datos.

- **Interpretación y explicabilidad de los resultados.** Los resultados derivados de los modelos de AI son a menudo complejos y difíciles de interpretar para los usuarios no técnicos. Los esfuerzos por generar confianza pasan por reducir el problema de *caja negra*. Desarrollar protocolos para interpretar los resultados puede contribuir a mitigar el problema de desconfianza por parte de los usuarios.

Algunas recomendaciones para garantizar la adecuada implementación de proyectos de inteligencia artificial son las siguientes<sup>14</sup>:

- Una implementación exitosa de proyectos de AI requiere de una intensa involucración de la alta dirección y del consejo de administración. La implementación de este tipo de proyectos puede generar resistencias y preocupaciones entre los empleados. Es necesario gestionar adecuadamente la comunicación y la capacitación para garantizar que los empleados comprendan la relevancia del impacto.
- Claridad de los objetivos. Los empleados deben tener claro cómo se utilizará la nueva tecnología y por qué se está trabajando en ella. Resulta importante construir casos de uso que expliquen por qué deben implantarse los proyectos. Los bancos necesitan una hoja de ruta clara para la innovación y la adopción de la AI. Esto orienta las prioridades de inversión y ayuda a evaluar el progreso.
- Transparencia respecto de los efectos esperados. Deben comunicarse los beneficios de la implantación de los proyectos de AI. La disposición de los empleados para adoptar

<sup>14</sup> Véase Evident AI Innovation Report (2022).

nuevos enfoques y tecnologías depende en gran medida de la calidad de la comunicación interna. La implantación de la tecnología de inteligencia artificial requiere un ciclo completo de despliegue. En cada fase los expertos deben comunicar claramente cómo se utilizará la nueva tecnología y por qué. Puede resultar útil desglosar el plan de alto nivel para el despliegue completo en fases más pequeñas. Comenzar con mini-proyectos de menor complejidad puede reducir la resistencia al cambio.

- Diseñar un enfoque multidisciplinario. Además de los conocimientos técnicos, es necesario contar con experiencia en el dominio específico de la aplicación de la AI, comprensión del negocio y habilidades de gestión de proyectos. Enfrentar estos desafíos requiere una planificación cuidadosa, una sólida comprensión de los requisitos y objetivos del negocio, así como una gestión efectiva del cambio y una colaboración estrecha entre equipos técnicos y no técnicos.
- La implantación de proyectos

de AI pueden implicar cambios organizativos y nuevas formas de trabajar que implican una mayor colaboración interdisciplinaria y una toma de decisiones basada esencialmente en datos.<sup>15</sup>

- Proteger la propiedad intelectual de los proyectos mediante patentes puede resultar clave para obtener y consolidar ventajas competitivas derivadas de la implantación de los proyectos de AI.
- Crear o participar en ecosistemas de inteligencia artificial puede permitir aprovechar la innovación compartida a través de colaboraciones diversas con comunidades de código abierto, universidades, aceleradores y proveedores de soluciones de terceros y *startups* de AI.
- Realismo respecto del retorno de la inversión y el calendario. La implementación de proyectos de AI suele implicar costes significativos, tanto en términos de recursos humanos como de inversión en infraestructuras. Es importante evaluar el retorno esperado y garantizar que estos proyectos estén

<sup>15</sup> Véase Fountaine *et al.* (2021).

alineados con los objetivos y las estrategias de la empresa. También es fundamental tener expectativas realistas respecto del período temporal en el que los proyectos pueden llegar a ofrecer resultados tangibles. Un problema que surge a menudo es que, como la AI puede ser difícil de implementar correctamente en un principio (algo ocurre con muchas tecnologías emergentes), es posible que los beneficios no se perciban inmediata o directamente. En consecuencia, el ROI debe hacerse para periodos de medio y largo plazo<sup>16</sup>. Tampoco parece aconsejable realizar una estimación del retorno proyecto por proyecto, ya que en la implementación de la tecnología AI existe una importante reutilización y los resultados de unos proyectos se apoyan en otros<sup>17</sup>.

- Gestión continuada de los proyectos. Una vez implantados los proyectos de inteligencia artificial es necesario una monitorización permanente de los sistemas para identificar problemas de rendimiento, errores o cambios inesperados

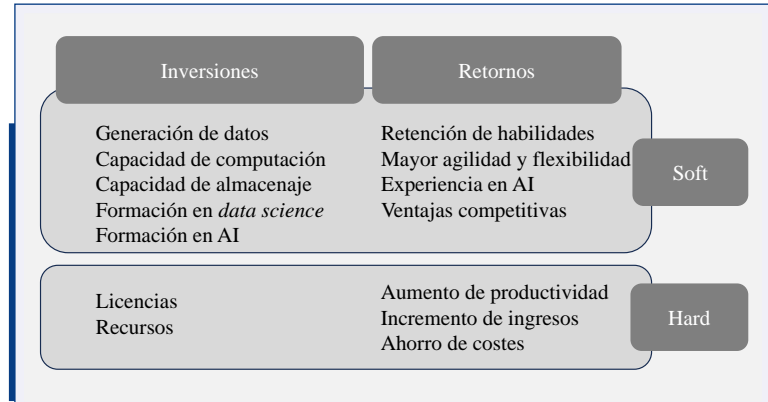
en los resultados. A tal fin, es preciso diseñar e implementar sistemas orientados a obtener métricas de rendimiento, a realizar un seguimiento de los factores clave, y a identificar alertas. También deben diseñarse sistemas de *reporting* y resulta necesario establecer procedimientos orientados a la resolución de los problemas que surgen tras el despliegue de los sistema de AI.

- Establecimiento de equipos centralizados de investigación de AI. Esta estrategia es un sello distintivo de las organizaciones financieras más avanzadas en el ámbito de la inteligencia artificial. La creación de equipos internos de investigación en AI implica invertir en el desarrollo de capacidades internas, lo que fomenta una cultura de innovación y prepara a los bancos para implementar nuevos avances. La investigación tanto pura como aplicada sobre soluciones de AI suele acelerar la madurez de las entidades financieras en este terreno y puede establecer un liderazgo intelectual en este ámbito.

<sup>16</sup> Véase Kejriwal (2022).

<sup>17</sup> Véase Kejriwal (2022).

Gráfico 4.1. Impacto de los proyectos de AI.



Fuente: Basado en Kejrival (2022).

En la misma línea, la investigación de Gartner<sup>18</sup> muestra que las entidades financieras líderes en la implementación de proyectos de inteligencia artificial, en el sentido de alcanzar o superar rápidamente el impacto esperado de sus proyectos de AI y ofrecer resultados financieros y comerciales críticos, comparten cuatro comportamientos comunes:

- La contratación de talento externo especializado en AI suele resultar más costosa pero puede traducirse en una implementación más rápida.
- La adquisición de software con funcionalidad en AI integrada permite a las

empresas experimentar rápidamente con proyectos.

- La experimentación debe ser temprana y la gama de proyectos piloto implementados debe ser amplia.
- Resulta conveniente asignar un líder de implementación de AI con sólida experiencia analítica y de datos.
- La AI debe verse como una prioridad estratégica para toda la organización.

Una de las cuestiones críticas para el éxito del impulso de la IA generativa en una entidad de servicios financieros es la elección de un modelo operativo eficiente,

<sup>18</sup> Véase Turner (2022).



en el sentido de permitir su escalabilidad y el alineamiento con la estructura organizacional y con la cultura de la empresa. Un modelo operativo adecuado debe permitir a la institución financiera realizar eficientemente tres tipos de actividades<sup>19</sup>:

- Dirección estratégica que implica identificar aquellos casos de uso de IA que se alineen con los objetivos estratégicos de la empresa; ordenarlos por prioridades (considerando la maximización del valor y la gestión del riesgo); y monitorizar la creación de valor para garantizar una asignación eficiente de recursos.
- Configuración, que implica definir estándares comunes (en cuanto a arquitectura tecnológica, prácticas de datos y marcos y controles de riesgo) para aumentar la eficiencia y la reutilización de conocimientos adquiridos.
- Ejecución que implica el diseño y prueba de las soluciones técnicas de los casos de uso, la puesta en producción según criterios determinados en cuanto a rendimiento y

seguridad y garantizar que el impacto puede ser medido.

Básicamente hay cuatro tipos de arquetipos de modelos operativos: i) modelo altamente centralizado que optimiza la velocidad y la homogeneidad, pero que puede generar un aislamiento respecto del resto de la organización; ii) modelo basado en una unidad de dirección centralmente que se integra mejor en la organización pero cuyo desarrollo depende de la colaboración de las unidades de negocio; iii) modelo basado en el liderazgo por cada unidad de negocio pero con un apoyo centralizado que permite integración aunque puede implicar falta de homogeneidad; y iv) modelo altamente descentralizado que facilita el apoyo pero cuya implementación puede implicar un menor nivel de calidad y la dificultad para la reutilización de recursos<sup>20</sup>.

Del mismo modo, se han identificado algunos errores que resulta muy aconsejable evitar y que están relacionados con los siguientes aspectos:

- Rechazo de empleados a los proyectos de AI por una

<sup>19</sup> Véase Buehler *et al.* (2024).

<sup>20</sup> Véase Behler *et al.* (2024)

insuficiente involucración inicial y una comunicación deficiente.

- Los empleados pueden percibir una excesiva exigencia diaria en su contribución en la implantación de proyectos de AI y tener dificultades para compatibilizar su dedicación a estos proyectos con su trabajo habitual.
- Existencia de contradicciones en los mensajes transmitidos por la alta dirección<sup>21</sup>.
- Dificultad para mejorar los datos, si éstos no han sido previamente utilizados para orientar las decisiones.

El despliegue de proyectos de inteligencia artificial en la industria financiera no para de crecer y su aceleración se justifica por factores técnicos, sociales y económicos.

Entre los factores técnicos que explican esta evolución destacan:

- La mejora en las herramientas de AI. El avance reciente en términos de capacidades de las propias herramientas de AI en funciones de análisis y predicción, así como en funciones de relación

con clientes y con otros *stakeholders* explican la adopción más intensa de estas tecnologías por parte de las entidades financieras. Pueden mencionarse avances notables en reconocimiento y generación de voz y el mayor realismo de los avatares que constituyen los instrumentos básicos para *chatbots*, asistentes virtuales y *robo-advisors*. Estas tecnologías resultan ahora más sencillas de implementar tras la digitalización previa de las arquitecturas de las entidades del sector financiero.

- La disponibilidad de datos de mayor calidad. También es relevante la acumulación de datos útiles, limpios y de calidad que han realizado las entidades en los últimos años
- El avance en la capacidad de computación instalada como consecuencia de la reducción en el precio relativo de la red física de ordenadores.
- La disponibilidad de herramientas AI. La aparición en el mercado de herramientas de inteligencia artificial que las entidades financieras pueden contratar en régimen

<sup>21</sup> Véase Jochim *et al.* (2023).

de “software as a service”, es uno de los factores más críticos del aumento de los despliegues, particularmente en el caso de proyectos piloto.

- La digitalización y adopción de soluciones en la nube. Parece existir una correlación positiva entre la intensidad con que una entidad implementa proyectos de AI, y su liderazgo en materia de digitalización y de adopción temprana de la tecnología de iCloud.

Además de los factores técnicos pueden destacarse algunos factores sociales que explican el avance de las iniciativas de AI en las instituciones financieras (así como en otras industrias):

- El impacto de la mayor aceptación por parte de los usuarios a la hora de interactuar con *interfaces* de aplicaciones de AI. Esta circunstancia se debe en parte a la migración ya realizada de los canales bancarios y de inversión al mundo online y móvil por los clientes, así como por el uso previo de videojuegos *apps* y redes sociales por parte de éstos. Particularmente relevante es la plena incorporación como usuarios de servicios financieros y de inversión de la Generación

Y o *millennials*, así como la progresiva incorporación de la Generación Z, que son ya *nativos digitales*, como consumidores de este tipo de servicios.

- Interés de los nuevos usuarios a la hora de lograr independencia en la gestión de su propia *salud financiera*. Una tendencia que se ha visto gradualmente reforzada por la emergencia de proyectos Fintech y DeFi.

Finalmente, hay determinados factores económicos que contribuyen a explicar el despliegue más veloz de los proyectos de AI en las entidades financieras:

- La necesidad por parte de las entidades del sector de ahorrar costes en los procesos de *back-office* y *middle-office*, mediante la automatización de tareas repetitivas. Los costes de ventas, generales y administrativos (SG&A por sus siglas en inglés) están sometidos a un escrutinio constante por parte de los analistas. Las funciones de soporte están siendo objeto no sólo de automatización de procesos mediante AI analítica, pero también de aprovechamiento de la tecnología para crear nuevos contenidos y generar nuevos



conocimientos (AI generativa).

- También en el *front-office* hay un campo de mejora significativo. En este sentido, los *chatbots* basados en AI pueden gestionar las consultas rutinarias de los clientes, permitiendo que los empleados se centren en tareas más complejas y de mayor valor añadido para los clientes.
- Resulta evidente la urgencia de mejorar la eficiencia en tareas de *compliance*, de *reporting* financiero y no financiero, así como de control de riesgos, que se derivan de nuevas normativas sobre ESG y otras, en un entorno como el financiero, particularmente regulado.
- Reducción en el coste y el tiempo de desarrollo de proyectos de AI, ante la aparición en el mercado de aplicaciones y servicios de inteligencia artificial y la posibilidad de acceder a consultoría externa a precios más competitivos.
- Reducción del problema de escasez de personal. La implementación de proyectos de AI se contempla por parte de algunas entidades financieras como una forma de intentar cubrir los gaps de recursos

*La implantación de la tecnología de inteligencia artificial requiere un ciclo completo de despliegue. En cada fase los expertos deben comunicar claramente cómo se utilizará la nueva tecnología y por qué.*

humanos, que se ponen de manifiesto como consecuencia de mercados laborales, ya bastante tensionados en muchos países. Una tendencia que será más evidente en un futuro, como consecuencia del envejecimiento de la población.

- Mejora de la calidad de los servicios y atención al cliente. Las entidades financieras necesitarán aprovechar las ventajas que ofrecen las herramientas de AI a la hora de crear nuevos productos y servicios, así como de mejorar la atención de los clientes. En concreto, la adopción de estas



herramientas permite conocer mejor el comportamiento de los clientes y anticipar con mayor precisión sus necesidades para ofrecer servicios más personalizados. Pero también el uso de aplicaciones AI permite mejorar la experiencia de uso por parte de los clientes en los servicios ofrecidos, así como aumentar la intensidad de contacto con éstos al permitir interacciones 24 horas al día, 7 días a la semana. La clave es lograr experiencias “personalizadas” que generan confianza en los clientes.

- Los sistemas de inteligencia artificial pueden también ayudar a crear nuevos productos y servicios, contribuyendo a acelerar la innovación financiera<sup>22</sup>.
- Y, en otros ámbitos, la AI permite mejorar la capacidad predictiva y la gestión de riesgos de las entidades.

En definitiva, el uso de sistemas de AI será decisiva en

la posibilidad de mantener o mejorar la posición competitiva de los incumbentes, y se mostrará como el principal instrumento de los nuevos jugadores para intentar alcanzar cuotas de mercado en esta industria. Esta estrategia puede desarrollarse de forma interna o en colaboración con compañías Fintech o grandes proveedores de tecnología utilizando APIs y servicios en almacenamiento en la nube. La inteligencia artificial desplegada a suficiente escala tendrá el potencial de hacer sentir al consumidor que la oferta financiera se construye en torno a sus especificaciones y preferencias exactas. El consumidor financiero confiará en que su banco sepa, sin ninguna solicitud de su parte, cuándo está dispuesto a actuar sobre una oferta específica ultrapersonalizada, y que se la presente de manera proactiva en el momento adecuado, para poder decidir por vía digital a cualquier hora del día y de la noche<sup>23</sup>.



<sup>22</sup> Véase Chui et al. (2022).

<sup>23</sup> Véase Ogan (2023)

## REFERENCIAS

- Buehler, K., Weintraub, B. and Lerner, L. (2024) "Scaling gen AI in banking: Choosing the best operating model". McKinsey and Company. 22 de marzo.
- Chui, M., Roberts, R. and Yee, L. (2022) "Generative AI is here: How tools like ChatGPT could change your business", *Quantum Black McKinsey and Company*.
- Evident AI Innovation Report (2022) *Mapping the race for AI innovation in banking*. 07.
- Fountaine, T., McCarthy, B. and Saleh, T. (2021) "Reimagining your business for AI" *McKinsey Analytics*.
- Jochim, M., Rosendahl, M and Thomas, S. (2023) "Reset and reimagine: The role of generative AI in SG&A success", 23 de Agosto. *McKinsey and Co*.
- Kejriwal, M. (2022) "AI in Practice and Implementation: Issues and Costs" en *Artificial Intelligence for Industries of the Future: Beyond Facebook, Amazon, Microsoft and Google*: 25-45. Cham: Springer International Publishing.
- Ogan, L. (2023) "Failing to Experiment with AI Tools Like ChatGPT Is an Inexcusable Mistake", *The Financial Brand*, 10 de julio.
- Ryll, L., Barton, M. E., Zhang, B. Z., McWaters, R. J., Schizas, E., Hao, R., ... and Yerolemou, N. (2020) "Transforming paradigms: A global AI in financial services survey" Disponible en SSRN.
- Turner, J. (2022) "CFOs: Here Are 4 Actions to Ensure You Implement AI – The Right Way", 26 de Julio.
- Walsh, M. (2020) "AI Should Change What You Do – Not Just How You Do It", *Harvard Business School. Technology And Analytics*. 21 de Septiembre.



INSTITUTO  
ESPAÑOL  
DE ANALISTAS

DESDE 1965

FUNDACIÓN

El impacto de la inteligencia artificial en la  
industria financiera: promesas y amenazas



# Parte II. Aplicaciones



## 5. Aplicaciones comunes de AI en la industria

“Así como la electricidad transformó casi todo hace cien años, hoy en día me cuesta pensar en una sola industria que la AI no vaya a transformar en los próximos años”.

*Andrew Ng.*

Las entidades que operan en los diferentes subsectores de la industria financiera están utilizando, de forma más o menos común, herramientas de inteligencia artificial en ámbitos tan diversos como los procesos administrativos, documentales y de gestión de información, contabilidad y auditoría,

determinación y asignación del capital regulatorio, gestión de riesgos, prevención del fraude, cumplimiento normativo, gestión del cambio, gestión del conocimiento, así como en los proyectos de educación financiera. En este capítulo se revisan algunos de estos procesos.

### 5.1. PROCESOS DE ADMINISTRACIÓN Y GESTIÓN DE INFORMACIÓN

Aunque en las entidades financieras todavía no se ha

adaptado completamente el flujo de trabajo a las nuevas tecnologías

disponibles, y hay algunos procesos que siguen haciéndose de forma cuasi-manual, lograr la automatización completa de ciertos procesos de *back* y *middle office* se ha convertido en un desafío muy relevante en la industria. Actualmente se están utilizando herramientas de inteligencia artificial en la automatización de tareas como la incorporación de la información de clientes en los sistemas, los procesos de tramitación de préstamos, las solicitudes de diversos servicios, el ingreso de cheques o la cumplimentación de formularios contractuales. La AI se está aplicando para eliminar la mayor parte de las tareas (lentas y propensas a errores) asociadas a la importación de datos de clientes procedentes de contratos, formularios y otras fuentes<sup>1</sup>. La mejora en el reconocimiento de la escritura manual, mediante el uso de reconocimiento óptico de caracteres (OCR, por sus siglas en inglés) hace posible la digitalización de los documentos físicos, lo que permite la reducción de costes operativos implicados en procesos dedicados al papeleo realizados por humanos.

Los bancos están utilizando software de automatización robótica de procesos -*robotic process automation* (RPA) en su terminología en inglés-. Estas aplicaciones informáticas imitan las tareas realizadas por humanos mediante reglas digitales<sup>2</sup>. Las herramientas de inteligencia artificial permiten descifrar cadenas insondables de caracteres, que contienen información sobre transacciones y clientes y las convierten en texto legible de información estructurada fácilmente comprensible por empleados y clientes<sup>3</sup>. Una mejor comprensión de los datos por parte de los clientes reduce, además, la carga que supone para los bancos millones de llamadas de consultas innecesarias, al tiempo que fomenta la confianza de los clientes en los servicios bancarios<sup>4</sup>.

El sector bancario realiza millones de transacciones al día, lo que exige la recopilación y el registro de un enorme volumen de información, que por primera vez deja de convertirse en una tarea abrumadora para sus de empleados

<sup>1</sup> Véase Cavus *et al.* (2021)

<sup>2</sup> Véase Kurode (2018), Indriasari *et al.* (2019) y Kochhar *et al.* (2019).

<sup>3</sup> Véase Abusalma (2021).

<sup>4</sup> Véase Kaur *et al.* (2020) y Husain *et al.* (2022).

gracias al uso de aplicaciones basadas en AI<sup>5</sup>. Estas técnicas junto con el procesamiento del lenguaje natural y otras herramientas inteligentes de automatización de procesos están cambiando el panorama del flujo de trabajo en las oficinas bancarias y en los centros de administración en la industria<sup>6</sup>.

Además, al sustituir ciertos

procesos humanos por la automatización basada en AI los bancos pueden imponer controles normativos y reglamentarios que antes eran imposibles, al tiempo que se reduce el tiempo y los recursos dedicados a los procesos, con la consiguiente reducción de los costes operativos y la mejora de las ratios de eficiencia<sup>7</sup>.

## 5.2. CONTABILIDAD Y AUDITORÍA INTERNA

Los defensores de la revolución que representa la inteligencia artificial consideran que la aplicación de este tipo de herramientas puede guiar un cambio positivo en el papel que desempeña la profesión contable en las organizaciones<sup>8</sup>. Entre los beneficios que se asignan al uso de la inteligencia artificial destacan las mejoras de eficacia y eficiencia, una mayor coherencia en la estructuración de las tareas de auditoría, la mejora en la toma de decisiones, una mayor

calidad en la comunicación, la mejor capacitación del personal y la reducción de la probabilidad de fraude<sup>9</sup>. Además, el uso de herramientas de AI puede dar un importante impulso a la denominada *auditoría continua*, entendida como la recopilación metódica y exhaustiva de evidencia de auditoría electrónica como una base razonable para emitir una opinión sobre la presentación fiel de los estados financieros realizados en un sistema de contabilidad

<sup>5</sup> Véase Yu y Song (2020).

<sup>6</sup> Véase Makhija y Chacko.

<sup>7</sup> Véase Kaur *et al.* (2020).

<sup>8</sup> Véase Stancheva-Todorova (2018).

<sup>9</sup> Véase Omoteso (2012), Chukwuani y Egiji (2020) o Mohammad *et al.* (2020).

en tiempo real que permite producir un *informe permanente o informes bajo demanda*<sup>10</sup>.

Las capacidades de las aplicaciones de inteligencia artificial en contabilidad y auditoría se centraron inicialmente en la automatización de tareas intensivas en trabajo, estructuradas y repetitivas que se realizan a lo largo de los procesos<sup>11</sup>. Tareas como revisión de documentos, confirmaciones, recuentos de inventario, análisis predictivo de riesgos, listas de solicitudes de verificación a clientes finales. En la actualidad, el uso de herramientas de AI en la contabilidad y la auditoría es especialmente intenso en el ámbito de la gestión de datos (extracción, comparación y validación), ya que esta tecnología puede localizar información relevante, extraerla de los documentos y hacerla utilizable para el contable y auditor humano, que puede dedicar más tiempo a las áreas que requieren un juicio de alto nivel<sup>12</sup>. Son muy diversas las tareas en las que las herramientas de AI pueden ayudar

a la automatización, incluyendo la verificación, el recálculo, la comprobación de la exactitud de los cálculos, así como de los asientos del libro mayor y los extractos bancarios. Con el tiempo, los sistemas de AI aprenden de la interacción humana y mejoran sus capacidades para extraer información importante y relevante<sup>13</sup>. Las herramientas de AI son capaces de escanear palabras clave y patrones en documentos electrónicos complejos para identificar y extraer información contable relevante de diversas fuentes, como facturas de ventas, de compras y otros contratos<sup>14</sup>. Las herramientas de AI pueden ayudar a la auditoría ya que permiten la evaluación de ficheros de texto, imágenes y voz, de forma que pueden introducir en el análisis documentación, conferencias telefónicas, correos electrónicos, comunicados de prensa y noticias de fuentes internas y externas<sup>15</sup>.

El foco esencial de la auditoría interna (y externa) consiste en proporcionar una opinión experta e independiente

<sup>10</sup> Véase Zhao et al. (2004)

<sup>11</sup> Véase Rapoport (2016).

<sup>12</sup> Véase Brennan et al. (2017).

<sup>13</sup> Véase Whitehouse (2015).

<sup>14</sup> Véase Agnew (2016) o Hasan (2021).

<sup>15</sup> Véase Bizarro y Dorian (2017).

sobre la veracidad y equidad de la información financiera y su adecuación con las normas contables aplicables y la legislación pertinente. La auditoría comprende un conjunto de actividades intensivas en información que implican la recopilación, organización, procesamiento, evaluación y presentación de datos con miras a generar una opinión confiable, que suele ser una amalgama de juicios basados en evidencia relevante, apropiada, adecuada y convincente sobre diferentes aspectos de los estados financieros. Hay dos tecnologías que van a alterar profundamente el trabajo de contables y auditores: *blockchain* e inteligencia artificial<sup>16</sup>. El área de la auditoría es ideal para la automatización parcial debido a su intensidad de mano de obra y la variedad de estructuras de decisión implicadas<sup>17</sup>.

A medida que las auditorías se automaticen se hará menos hincapié en la comprobación de los datos y se prestará mayor atención a la comprensión de

la imagen general que ofrecen los datos, y a la evaluación de tendencias, patrones y valores atípicos<sup>18</sup>. Pero también para tareas que requieren una mayor capacidad analítica se vienen utilizando tecnologías de AI. Por ejemplo, se ha extendido el uso de sistemas expertos en evaluaciones de materialidad, evaluaciones de control interno y juicios sobre empresas en funcionamiento<sup>19</sup>. Del mismo modo, las redes neuronales y los algoritmos genéticos se están utilizando en revisiones analíticas y en tareas de clasificación de transacciones<sup>20</sup> o, por ejemplo, para discriminar entre deudas cobrables o incobrables. Existe consenso en que se pueden utilizar aplicaciones de AI basadas *lógica difusa* y redes neuronales para resolver algunos problemas contables de forma más completa<sup>21</sup>. Por ejemplo, las herramientas de AI pueden detectar si una empresa registra cifras de ventas inusualmente altas justo antes del final de un periodo de información, o desembolsa pagos inusualmente elevados justo después de finalizar

<sup>16</sup> Véase Han *et al.* (2023).

<sup>17</sup> Véase Issa *et al.* (2016).

<sup>18</sup> Véase Accounting Today (2016).

<sup>19</sup> Véase O'Leary (1995) y Issa *et al.* (2016).

<sup>20</sup> Véase Zhang *et al.* (2022).

<sup>21</sup> Véase Baldwin *et al.* (2006).

el periodo de referencia<sup>22</sup>. Las herramientas de AI también pueden detectar anomalías en los datos, como un aumento inesperado de pedidos en una región concreta, partidas de gastos inusualmente elevadas contabilizadas por una persona o condiciones de arrendamiento de equipos excepcionalmente favorables para un proveedor<sup>23</sup>. Los contables y auditores utilizan cada vez más la inteligencia artificial para el análisis y la comunicación de resultados<sup>24</sup>, así como para realizar análisis predictivos sobre eventos financieros y delitos<sup>25</sup>. Del mismo modo se están utilizando herramientas de inteligencia artificial basada en procesamiento del lenguaje natural para la creación de informes de actividades sospechosas para los procesos de lucha contra el blanqueo de capitales en los servicios financieros<sup>26</sup> y algoritmos genéticos para detección de quiebras<sup>27</sup>.

En definitiva, la auditoría

interna de las entidades financieras puede beneficiarse de la introducción de la AI en tareas con diferente tipo de complejidad. La automatización de los procesos puede lograr mayor rapidez y eficiencia y, como resultado, aumentar el grado de nivel añadido de las tareas que pueden realizar los auditores internos humanos. Esto mejora las habilidades de los auditores internos en campos como la detección anticipada de fraude y el seguimiento de anomalías en tiempo real<sup>28</sup>.

El uso de herramientas de inteligencia artificial en el ámbito de la contabilidad y la auditoría interna no está exenta de problemas. Se han destacado, entre otros, algunos posibles problemas entre los que destacan los siguientes:

- Construir, actualizar y mantener sistemas complejos de AI implica un coste potencialmente elevado<sup>29</sup>.

<sup>22</sup> Véase Rapoport (2016).

<sup>23</sup> Véase Brennan *et al.* (2017).

<sup>24</sup> Véase Schneider *et al.* (2015).

<sup>25</sup> Véase Tschakert *et al.* (2016).

<sup>26</sup> Véase Chukwuani y Egji (2020).

<sup>27</sup> Véase Zemánková (2019).

<sup>28</sup> Véase Baldwin *et al.* (2006) y Couceiro *et al.* (2020).

<sup>29</sup> Véase Pieptea y Anderson (1987).

- Los procesos de decisión de auditoría puedan prolongarse en mayor medida como resultado de la exploración de un mayor número de alternativas<sup>30</sup>.
- El uso de aplicaciones de AI puede generar inhibición para ejercer el juicio profesional por parte de los empleados que se ocupan de estas tareas<sup>31</sup>.
- Existe la posibilidad de que se acuse ante los tribunales a la empresa y/o a los auditores de haber confiado excesivamente en las pruebas de las ayudas a la toma de decisiones si se producen errores<sup>32</sup>.

En otro orden de cosas, se ha puesto énfasis en la importancia de que se examine la parcialidad en los resultados basados en AI y la capacidad de los humanos que utilizan estas aplicaciones a la hora de emitir juicios y tomar decisiones apropiadas. Cuando se usan máquinas inteligentes, los resultados tienden a reflejar los prejuicios de los seres humanos que las crean o interactúan con ellas.

Estos sesgos incluyen aquellos que son debidos a asimetrías en los datos subyacentes, el *sesgo por interacción* (las máquinas aprenden los sesgos de las personas que las entrenan) y el *sesgo de confirmación*, que se produce cuando las máquinas protegen a los humanos de puntos de vista contradictorios y les proporcionan información que confirman sus preferencias o creencias (*burbuja de personalización*)<sup>33</sup>.

Además, resulta importante dotar las decisiones de contabilidad y auditoría basadas en inteligencia artificial de suficiente transparencia. Mientras que las primeras versiones de herramientas de AI (como, por ejemplo, los sistemas expertos basados en reglas) permitían a los observadores humanos comprender con relativa facilidad las relaciones entre entradas, transformaciones y salidas de los modelos, las técnicas de aprendizaje automático y las redes neuronales de aprendizaje profundo son, a menudo, *cajas negras* cuyos resultados son difíciles de entender e interpretar.

<sup>30</sup> Véase Mackay et al. (1992).

<sup>31</sup> Véase Yuthas y Dillard (1996).

<sup>32</sup> Véase Bizarro y Dorian (2017)

<sup>33</sup> Véase Kokina y Davenport (2017).

A lo largo de los años, el foco de la investigación y aplicación de modelos de AI en contabilidad y auditoría ha consistido en mejorar su rendimiento predictivo<sup>34</sup> pero, a medida que éste aumenta la *explicabilidad* de sus resultados generalmente disminuye, ya que se requieren muchas variables y cálculos de múltiples capas para hacer que un modelo sea más efectivo, lo que hace que el modelo sea más opaco<sup>35</sup>. Si bien uno de los principales desafíos de la adopción de la inteligencia artificial en la auditoría es la falta de *explicabilidad* de sus resultados, a medida que esta tecnología

madura, también lo hacen las herramientas que pueden mejorar la interpretabilidad de los resultados<sup>36</sup>.

Finalmente, con el creciente uso de la AI en las actividades de auditoría, también están surgiendo tensiones entre las diferentes partes interesadas, incluyendo contables, auditores internos, firmas de auditoría externa, clientes e inversores. Esta tensión surge del conflicto entre los derechos de las partes interesadas, las brechas de responsabilidad, las brechas de expectativas y también el conflicto ocasional de principios éticos<sup>37</sup>.

### 5.3. DETERMINACIÓN DEL CAPITAL REGULATORIO

Tras los acuerdos de Basilea II (2004), los bancos tienen la posibilidad de utilizar modelos internos (enfoque avanzado) para estimar sus requisitos de fondos propios, siempre que cuenten con la autorización previa de su supervisor. Sin esta autorización, los bancos

estiman la cantidad mínima de capital que deben mantener utilizando ponderaciones de riesgo basadas en calificaciones externas atribuidas a sus contrapartes (método estándar). Los bancos que utilizan un enfoque avanzado llevan años usando técnicas estadísticas

<sup>34</sup> Véase Perols (2011), Perols et al. (2016) o Bao et al. (2020).

<sup>35</sup> Véase Virág y Nyitrai (2014) o Baryannis et al. (2019).

<sup>36</sup> Véase Zhang et al. (2022).

<sup>37</sup> Véase Munoko et al. (2020).

tradicionales, típicamente de regresión logística. Recientemente están comenzando a utilizar herramientas de inteligencia artificial. Algunos trabajos sobre el tema indican que el uso de redes neuronales y técnicas de XGBoost en la estimación del riesgo crediticio puede traducirse en una reducción del nivel requerido de capital. La necesidad de *explicabilidad* e interpretabilidad de los resultados es uno de los posibles obstáculos a los que se enfrenta la dirección de los bancos para adoptar esta decisión basándose exclusivamente en modelos de *machine learning*<sup>38</sup>.

Además de la interpretabilidad, la estabilidad de las predicciones y la gobernanza de los modelos se encuentran entre las preocupaciones que surgen en los supervisores a la hora de evaluar modelos de ML en la estimación de riesgos financieros. El 11 de noviembre de 2021, la Autoridad Bancaria Europea (EBA) publicó un informe y consultó a la industria sobre el uso del aprendizaje automático en modelos

elaborados bajo el enfoque *Internal ratings-based (IRB)*<sup>39</sup> con el objetivo de involucrar a la industria y a la comunidad supervisora en la investigación sobre esta cuestión, y construir un entendimiento común sobre los retos relacionados con el cumplimiento de los requisitos reglamentarios<sup>40</sup>. De la respuesta de las entidades se concluía que en las principales fases de modelización de los modelos internos las técnicas de *machine learning* se utilizaban principalmente para la *diferenciación del riesgo*, mientras que para la *cuantificación del riesgo* se ponía énfasis en la importancia de los posibles problemas de disponibilidad de datos. El motivo es que esta tarea exige periodos de observación histórica prolongados, que podrían no estar aún disponibles como fuentes de datos utilizadas en las técnicas de aprendizaje automático<sup>41</sup>. Por otra parte, la EBA ha analizado la interacción entre los requisitos prudenciales sobre modelos IRB y la normativa sobre AI de la Unión Europea, y aunque apoya esta última, opina que el

<sup>38</sup> Véase Fraisse y Laporte (2022) y Alonso-Robisco y Carbó (2022a).

<sup>39</sup> El enfoque del riesgo crediticio basado en calificaciones internas permite a los bancos modelizar la probabilidad de incumplimiento (PoF) y, en el caso del enfoque IRB avanzado, también los niveles de pérdida en caso de incumplimiento (LGD) y la exposición en caso de incumplimiento (EAD).

<sup>40</sup> Véase EBA (2021).

<sup>41</sup> Véase EBA (2021)

texto se beneficiaría de nuevas aclaraciones para garantizar

que no queden incertidumbres jurídicas tras su adopción<sup>42</sup>.

## 5.4. GESTIÓN DE RIESGOS Y DETECCIÓN Y PREVENCIÓN DEL FRAUDE

Por lo que respecta a la gestión del riesgo en el sector bancario, la inteligencia artificial se puede aplicar en tres niveles: protección, detección y respuesta<sup>43</sup>. Las pérdidas sufridas por los bancos tras los ciberataques se pueden clasificar en directas e indirectas. Las pérdidas directas corresponden a robos de dinero, pagos de chantajes ante violaciones de datos y bloqueos y costes de restablecimiento de los sistemas. Las pérdidas indirectas están asociadas a pérdida de reputación y daños a la marca, problemas de imagen y repercusiones económicas derivadas de frustración e insatisfacción de los clientes.

La AI puede ofrecer a los bancos herramientas poderosas para prevenir y detectar ciberdelitos y gestionar sus amenazas. En concreto, se han utilizado con éxito métodos como las redes

neuronales artificiales, la lógica difusa y los algoritmos genéticos. Las RNAs son robustas al ruido estadístico y suficientemente flexibles para manejar fenómenos dinámicos complejos. En términos más generales, las empresas han utilizado herramientas de inteligencia artificial para detectar amenazas y detener ataques, analizar puntos finales móviles y mejorar el análisis humano en este terreno.

La aplicación de herramientas de AI puede ayudar a las instituciones financieras en la identificación de la exposición al riesgo operativo, su medición, estimación y evaluación de sus efectos, así como proponer estrategias de mitigación del riesgo. La tecnología de AI es particularmente útil en la evaluación de riesgos debido a su capacidad para detectar, analizar y responder rápidamente a las

<sup>42</sup> Véase EBA (2023).

<sup>43</sup> Véase Goosen *et al.* (2018) o Ling *et al.* (2019).

amenazas mediante el análisis de comportamiento de eventos y anomalías. Su uso reduce la cantidad de falsos positivos generados por las herramientas tradicionales de detección de vulnerabilidades. A diferencia de los sistemas tradicionales de clasificación de riesgos, la AI puede medir las exposiciones y las contramedidas de forma independiente. Además, mediante técnicas basadas en AI pueden evaluarse datos no estructurados y los patrones relacionados con incidentes pasados pueden identificarse y convertirse en predictores de riesgos con los que construir escenarios plausibles prospectivos<sup>44</sup>.

Las herramientas de inteligencia artificial hacen posible incorporar mediciones automatizadas para mejorar la precisión de la predicción de los resultados esperados y verificar instantáneamente que los valores reales coincidan con las predicciones. Este enfoque crea una forma innovadora de verificación del control que es de naturaleza proactiva<sup>45</sup>. Los algoritmos de aprendizaje

automático tienen el potencial de analizar millones de puntos de datos para detectar transacciones fraudulentas que probablemente pasarían desapercibidas para los humanos. Las técnicas de AI mejoran la precisión de las aprobaciones en tiempo real y reducen los resultados falsos positivos. Las combinaciones de datos comerciales y registros de comunicaciones electrónicas y de voz permiten a los bancos observar las pautas de comportamiento que surgen para predecir riesgos latentes. También permite a las entidades generar y priorizar alertas en función de los tipos de actividad sospechosa y del nivel de riesgo<sup>46</sup>.

El reconocimiento facial es una de las numerosas formas en que los bancos pueden aumentar la seguridad al tiempo que mejoran la accesibilidad a sus sistemas. El uso del reconocimiento facial para iniciar sesión en la operativa bancaria online y en los sistemas de pago seguros y robustos utilizando reconocimiento facial es una de las aplicaciones basadas en inteligencia artificial<sup>47</sup>.

<sup>44</sup> Véase Chan (2023).

<sup>45</sup> Véase Chan (2023).

<sup>46</sup> Véase Aziz y Dowling (2019).

<sup>47</sup> Véase Adesuyi *et al.* (2013), Manju y Madhumathi (2019), Jain *et al.* (2021) y Mohammed



La actividad bancaria se enfrenta a muy diversos intentos de fraude. Habitualmente se habla de fraude bancario para referirse al uso no autorizado de información confidencial de un individuo para suplantar la identidad, realizar compras o retirar fondos de la cuenta del usuario. El fraude con tarjeta de crédito implica utilizar una tarjeta de crédito física o virtual para obtener dinero o realizar pagos sin el permiso del propietario. La tecnología permite ocultar la identidad y ubicación de los usuarios que realizan transacciones a través de Internet, lo que contribuye a aumentar el fraude en la red. Este tipo de fraude es un problema de rápido crecimiento como consecuencia del aumento del comercio electrónico y por el hecho de que el canal móvil facilita casi cualquier tipo de pago. Las transacciones fraudulentas se mezclan con transacciones genuinas y las técnicas simples de coincidencia de patrones a menudo no son suficientes para detectar con precisión este tipo de fraude. Por lo tanto, la implementación de sistemas eficientes de detección de fraude se ha vuelto imperativa para que todos los bancos emisores de tarjetas de

crédito minimicen sus pérdidas. El fraude se puede detectar analizando el comportamiento de gasto de los clientes a partir de datos de transacciones anteriores. Si se observa alguna desviación en el comportamiento de gasto con respecto a los patrones disponibles, se puede etiquetar como una transacción potencialmente fraudulenta. Uno de los problemas asociados con la investigación sobre el fraude es la falta de datos reales, debido a la sensibilidad de los datos y a cuestiones de privacidad. El segundo problema es lidiar con el desequilibrio de los datos o su distribución sesgada, ya que la cantidad de transacciones fraudulentas es mucho menor en comparación con las transacciones legítimas. En el intento de superar este problema, se utilizan métodos de *sobre-muestreo minoritario sintético* para aumentar la cantidad de datos de baja incidencia en el conjunto de datos que generan transacciones fraudulentas sintéticas relacionadas con el conjunto de datos original<sup>48</sup>. Los sistemas de detección tradicionales dependen de criterios definidos o registros aprendidos que dificultan la detección de nuevos patrones

(2021).

<sup>48</sup> Véase Modi y Dayma (2017).



de ataque. Para descubrir nuevos patrones y lograr una mayor precisión de detección, se están utilizando actualmente herramientas de inteligencia artificial<sup>49</sup> como métodos de aprendizaje automático basados en aprendizaje supervisado<sup>50</sup> y no supervisado, así como aprendizaje profundo, redes neuronales artificiales<sup>51</sup>, máquina de vectores de soporte, regresión logística<sup>52</sup>, lógica difusa, programación genética, árboles de decisión<sup>53</sup>, K- vecino más cercano<sup>54</sup>, etc. Por ejemplo, el uso comparado de técnicas de redes neuronales y de regresión logística ofrece una ventaja sustancial respecto de los métodos tradicionales. También se han comparado los resultados de técnicas obtenidas con el algoritmo Naïve Bayes y el árbol de decisión y varios métodos SVM diferentes (con funciones de núcleo polinomiales, sigmoideas, lineales y RBF)<sup>55</sup>. En la detección del fraude, algunos estudios

han utilizado un sistema híbrido basado en la técnica de red neuronal artificial y el algoritmo Harmony Search (HAA). El sistema utiliza HSA para optimizar los parámetros de ANN, mientras que la propia ANN se utiliza para la detección<sup>56</sup>. También se han utilizado enfoques basados en redes neuronales convolucionales (CNNs) - que constan de más de una capa oculta - para encontrar transacciones fraudulentas<sup>57</sup>.

Una parte significativa de los fraudes actuales se realizan mediante el uso de *ingeniería social*. En este caso, los estafadores utilizan sus habilidades para obtener información detallada sobre el sistema. En lugar de ataques técnicos a los sistemas, los piratas sociales apuntan a humanos con acceso a la información, manipulándolos para que divulguen información confidencial o incluso para que

<sup>49</sup> Véase Perols (2011), Chaudhary *et al.* (2012), Choi y Lee (2018) o Minastireanu y Mesnita (2019).

<sup>50</sup> Véase Patel y Gond (2014), Banerjee *et al.* (2018).

<sup>51</sup> Véase Patidar y Sharma (2011), Raj y Portia (2011), Mishra y Dash (2014), Rajamani y Rathika (2015), Soltaniziba y Balafar (2015), Mishra *et al.* (2017), Monirzadeh *et al.* (2018), Thennakoon *et al.* (2019), Nkomo y Breetzke (2020).

<sup>52</sup> Véase Zareapoor y Shamsolmoali (2015).

<sup>53</sup> Véase Mahmoudi y Duman (2015), Bahnsen *et al.* (2015) y Jain *et al.* (2016).

<sup>54</sup> Véase Sudha y Raj (2017).

<sup>55</sup> Véase Khan *et al.* (2014) o Mubarek y Adali (2017).

<sup>56</sup> Véase Daliri (2020).

<sup>57</sup> Véase Fu *et al.* (2016).

lleven a cabo ataques maliciosos mediante la influencia y la persuasión. Los fraudes están relacionados con el robo o la puesta en peligro de datos bancarios mediante diversas técnicas. El *phishing* es el ataque de ingeniería social más popular en el sistema bancario en línea. Normalmente, un pirata informático envía un correo electrónico utilizando la marca registrada de la organización legítima para llamar la atención de su objetivo. El correo electrónico falso parece provenir de un banco confiable que solicita que el cliente actualice la información de su cuenta utilizando el enlace proporcionado (que es un enlace falso). En el *spear phishing* se utilizan los datos personales de una víctima potencial para adaptar el contenido del correo electrónico, con una mayor probabilidad de éxito. El *vishing* es una forma de ataque de ingeniería social en la que un atacante utiliza una llamada telefónica para engañar a una víctima y revelar información confidencial como el número de tarjeta de crédito, el código PIN o la dirección particular detallada. El *caballo de troya* utiliza dispositivos digitales como unidades USB o RAM para llamar la atención de la víctima y

perpetrar un ataque. Esta técnica de ataque se basa en la curiosidad humana y puede lograr que se propague el *malware* instalado en su dispositivo. También se utilizan sitios web fraudulentos. En este tipo de ataque, el pirata informático explota la confianza de la víctima y la lleva a acceder a su sitio web falso, que descarga automáticamente archivos maliciosos en la computadora de la víctima. Al igual que con el ataque del caballo de troya, el archivo descargado le da al atacante acceso a información confidencial desde el navegador local de la víctima. El *pretexto* utiliza un escenario programado para engañar a la víctima para que revele información confidencial o realice otras actividades maliciosas sin saberlo. La *ingeniería social inversa* es el mejor ejemplo de pretexto, en el que un atacante crea una escena o situación y una víctima inocente cree que el *hacker* puede proporcionar una solución<sup>58</sup>. La facilidad de las operaciones bancarias conlleva varias amenazas y vulnerabilidades que afectan a esta industria. El proceso de ataque de *ingeniería social* implica habitualmente un ciclo de ataque en cuatro fases: i) investigación (recopilación) de información sobre el objetivo;

<sup>58</sup> Véase Airehrour et al. (2018).

ii) desarrollo de simpatía para generar confianza; iii) explotación de la confianza (manipulan del comportamiento humano); y iv) uso de la información (perpetración del ataque). La explotación puede ser un ataque de una o varias etapas (un solo ataque o un ataque secuencial). Las tecnologías de inteligencia artificial, como el aprendizaje profundo, pueden introducirse en la seguridad cibernética para construir modelos inteligentes que implementen la clasificación de *malware*. En particular, la hibridación de sistemas de K-NN y SVM se han utilizado con cierto éxito para la detección de este tipo de fraude<sup>59</sup>. También se han propuesto sistemas de clasificación de *malware* basada en secuencias genéticas estáticas, utilizando CNNs.

Una acción importante en la seguridad de la red y de lucha contra el fraude es la detección de ataques de intrusión. Cualquier dispositivo expuesto a Internet corre el riesgo de ser atacado y comprometido por lo que detectar intentos de ataque, ya sean exitosos o no, es importante para proteger las redes (servidores,

hosts finales y otros dispositivos), así como para el análisis forense<sup>60</sup>. Con el tráfico de red en constante crecimiento, el sistema de detección de intrusiones se ve afectado por datos de intrusión de alta dimensión y tipos de ataques complejos, y el sistema tradicional de detección resulta insuficiente para enfrentar los desafíos en este ámbito. El sistema de detección de intrusos (IDS por sus siglas en inglés) es la “alarma antirrobo” en el campo de la seguridad de las redes informáticas. Su objetivo es proteger el sistema combinando alarmas emitidas cuando la seguridad de la red se ve amenazada y emprender las acciones correspondientes. IDS puede detectar, asegurar e identificar el uso, replicación, cambio y destrucción no autorizados de sistemas de información. Actualmente existen tres tipos principales de detección de intrusiones: basada en mal uso, en anomalías y mixta. En este contexto, herramientas de inteligencia artificial que incluyen minería de datos, árboles de decisión, máquinas de vectores de soporte, algoritmos genéticos, redes neuronales artificiales (ANN), etc. están siendo muy

<sup>59</sup> Véase Dada (2017) y Chowdhury *et al.* (2017)

<sup>60</sup> Véase Divakaran *et al.* (2017).

utilizadas<sup>61</sup>. En particular, las ANNs se han utilizado ampliamente en sistemas de detección de intrusiones debido a su estructura de modelo única, capacidad de simulación no lineal, fuerte aprendizaje y capacidad de adaptación<sup>62</sup>. De igual modo se están aplicando algoritmos del tipo *long-short term memory (LSTM)* con objeto de reducir la tasa de falsos positivos<sup>63</sup>. Asimismo, se han propuesto métodos híbridos de aprendizaje automático combinando el método de selección de características (aprendizaje supervisado) y el método de reducción de datos (aprendizaje no supervisado)<sup>64</sup>. La aparición del aprendizaje profundo proporciona una nueva dirección para la detección de intrusiones. El aprendizaje profundo descubre estructuras intrincadas en grandes conjuntos de datos mediante el uso del algoritmo de retro propagación y parece adecuado para detectar una variedad de comportamientos de invasión de alta dimensión<sup>65</sup>. También se

están comparando (en términos de las medidas habituales de rendimiento como son precisión, especificidad, sensibilidad y tasa de falsos positivos) los resultados de detección de sistemas basados en anomalías utilizando máquinas impulsada por gradiente (GBM), bosque aleatorio, red neuronal profunda, máquina de vectores de soporte y árbol de clasificación y regresión<sup>66</sup>. Por su parte, el *aprendizaje con pocos disparos –Few-Shot Learning (FSL)* por su terminología en inglés– es un método novedoso de aprendizaje profundo que se ha vuelto popular en los últimos dos años. Su objetivo es aprender de una pequeña cantidad de datos etiquetados. FSL es un método eficaz para resolver el problema de la pequeña cantidad de datos de detección de intrusiones en la red, pero también necesita un conjunto de datos equilibrado para el entrenamiento<sup>67</sup>. En otra línea, se han propuestos métodos de detección de intrusiones basados en la combinación de aprendizaje

<sup>61</sup> Véase Kumar y Kumar (2012), Shah *et al.* (2018), Nguyen *et al.* (2021), Singh y Khare (2022).

<sup>62</sup> Véase Shah y Trivedi (2012), Sarvari *et al.* (2020).

<sup>63</sup> Véase Staudemeyer (2015) y Al-Jallad *et al.* (2020).

<sup>64</sup> Véase Megantara y Ahmad (2021).

<sup>65</sup> Véase Zhao *et al.* (2017).

<sup>66</sup> Véase Tama y Rhee (2019).

<sup>67</sup> Véase Yu y Bian (2020).



profundo y aprendizaje por transferencia, que transforma el problema de detección de intrusiones en un problema de reconocimiento de imágenes.

La industria de la ciberseguridad también está desarrollando sistemas de AI de defensa que permiten una protección de forma más rápida y sistemática. En lugar de operar estáticamente en un solo entorno y depender de la lectura humana de los registros (para verificar si se

ha producido alguna actividad sospechosa en la red), los nuevos sistemas inteligentes de detección de intrusos permiten que la máquina gestione el ciclo inicial de detección.

Estas soluciones consisten en colocar recolectores de datos en los puntos finales de la red y monitorizar el comportamiento típico de cada usuario. El componente de AI realiza una detección automática, genera alertas, bloquea los dispositivos y solicita comprobaciones.

## 5.5. CUMPLIMIENTO NORMATIVO

Las entidades financieras necesitan mitigar tanto los riesgos financieros como los legales, así como mantener intacta la reputación de su marca. Debido al aumento de las transacciones y al panorama regulatorio, en constante cambio, mantener estándares de calidad en materia de *compliance* se ha vuelto mucho más complejo y costoso. De hecho, en la industria financiera se está produciendo un crecimiento continuo en el gasto destinado a tareas relacionadas con cumplimiento normativo. Aspectos como el reconocimiento de clientes

–*know-your-customer* (KYC) en su terminología en inglés–, la lucha contra el blanqueo de capitales –*anti-money laundering* en su terminología en inglés–, la lucha contra la financiación de terrorismo y, más recientemente, el cumplimiento y la elaboración de información relativa a cuestiones medioambientales, sociales y de gobernanza –*environmental, social and governance* (ESG) por su terminología en inglés–, son una fuente relevante de consumo de recursos para las empresas.

El lavado de dinero y la financiación del terrorismo

han sido una de las principales amenazas a la integridad del sistema financiero internacional en las últimas décadas y ha habido numerosos casos en los que las instituciones financieras han sido penalizadas por no poder mantener controles efectivos. Para combatir los delitos financieros y cumplir con los requisitos contra el lavado de capitales, los bancos tradicionalmente han estado utilizando sistemas basados en reglas que ayudan a identificar transacciones en efectivo más allá de los valores umbral, así como transferencias internacionales de fondos y otras transacciones que pueden resultar sospechosas. Estos sistemas detectan las transacciones sospechosas según reglas predefinidas y generan alertas, lo que permite realizar investigaciones internas de los casos. De resultar positivas, se preparan informes de materia sospechosa –*suspicious matter report* (SMR) por su terminología en inglés– que son enviados al regulador para garantizar la presentación de informes de cumplimiento. Se estima que la tasa de falsos positivos de estas alertas es muy elevada. En el caso de KYC, no es fácil obtener datos basados en la información básica de identidad y dirección de los

clientes. La calidad de los datos obtenidos también puede ser baja debido a su gran volumen y la velocidad a la que se generan. Por lo tanto, para preprocesar los datos desde múltiples formatos no estándar es necesaria su limpieza y transformación. En segundo lugar, los patrones de comportamiento en materia de lavado de dinero pueden diferir entre países y empresas, lo que hace que los mecanismos de seguimiento actuales, simplistas y estáticos sean propensos a errores. En tercer lugar, puede resultar difícil diferenciar las transacciones normales de las anormales, lo que permite a los blanqueadores de dinero evitar la detección<sup>68</sup>.

Los modelos con problemas de *caja negra*, como las redes neuronales y los modelos de aumento de gradiente pueden predecir resultados muy precisos, sin embargo, a menudo tienen retrasos en la interpretabilidad, por lo que es difícil explicar el fundamento de las decisiones tomadas. Por otro lado, los modelos de *caja blanca*, como el árbol de decisión y la regresión lineal, son altamente interpretables y más fáciles de explicar las decisiones tomadas por los modelos, pero proporcionan

<sup>68</sup> Véase Chen (2020).

resultados menos precisos. Por lo general, los modelos de *caja negra* se eligen considerando la alta precisión esperada en la detección de transacciones sospechosas. Sin embargo, la falta de transparencia sobre cómo estos modelos toman decisiones para concluir que una transacción es sospechosa; y una solicitud obligatoria por parte de los supervisores para proporcionar una justificación de la transacción sospechosa, está dificultando la adopción de esta tecnología en las unidades de delitos financieros en las instituciones financieras<sup>69</sup>. La capacidad de los métodos de aprendizaje automático para analizar cantidades muy grandes de datos, al tiempo que ofrecen una gran granularidad y profundidad de análisis predictivo, puede mejorar significativamente las capacidades analíticas en las áreas de cumplimiento normativo<sup>70</sup> sin embargo, como contrapartida, el modelo es más complejo y falta conocimiento explicativo. Este es un problema especialmente relevante cuando los análisis se aplican en un contexto normativo, ya que un supervisor o un equipo de cumplimiento querrán auditar y comprender el modelo aplicado. Como consecuencia, se han

puesto de manifiesto algunas limitaciones frente al entusiasmo desenfrenado y la aceptación incondicional de la aplicación de AI en este terreno<sup>71</sup>. Las técnicas de aprendizaje no supervisados pueden ser una posible solución para abordar el problema que plantea la insuficiencia de datos bancarios reales. Es habitual tener que realizar un reentrenamiento completo basado en todos los datos, lo que provoca repetidos retrasos tras la introducción de nuevos registros. De ahí que el tiempo del proceso de aprendizaje sea otro reto importante en el contexto de la detección del blanqueo de capitales.

Por otra parte, el cumplimiento de criterios de ESG, que incluye también aspectos relativos a equidad, inclusión y diversidad, se está convirtiendo en una tarea muy relevante de todas las empresas, incluyendo las entidades financieras. En el caso de la industria financiera, que es el núcleo de la implementación de las finanzas sostenibles, las tareas de control de ESG resultan particularmente relevantes. En este contexto parece necesario implementar soluciones efectivas

<sup>69</sup> Véase Kute et al. (2021).

<sup>70</sup> Véase Van Liebergen (2017) y Dey (2017), Chen (2020) y Aggarwal et al. (2020).

<sup>71</sup> Véase Van Liebergen (2017) y Butler y O'Brien (2019).

para mitigar el avance del gasto y, al mismo tiempo, mejorar el flujo de trabajo en este ámbito. Dado que la garantía del cumplimiento implica monitorizar un enorme volumen de datos, a menudo en tiempo real, procedentes de fuentes dispares tanto estructuradas como no estructuradas, las herramientas de inteligencia artificial están ganando terreno en la gestión del cumplimiento tanto por parte de las propias instituciones financieras, como también de los supervisores<sup>72</sup>.

En otro orden de cosas, no debe subestimarse la cuestión del coste de las herramientas de inteligencia artificial, ya que es posible que los pequeños

intermediarios no tengan suficiente presupuesto para comprar soluciones sofisticadas, mientras que las grandes organizaciones que operan en múltiples jurisdicciones necesitan soluciones estandarizadas que pocos proveedores pueden ofrecer<sup>73</sup>. Además, la adopción de nuevas tecnologías de AI puede requerir inversiones adicionales en otras partes de la organización, para actualizar los sistemas de bases de datos heredados y hacerlos compatibles con las nuevas herramientas<sup>74</sup>. Y hay que considerar también costes indirectos, como la contratación de personal con experiencia y conocimientos técnicos necesarios<sup>75</sup>.

## 5.6. GESTIÓN DEL CAMBIO (CM) Y GESTIÓN DEL CONOCIMIENTO (KM)

Las capacidades emergentes de la inteligencia artificial van a impregnar casi todos los contornos y actividades organizacionales, incluida la gestión del cambio

y la gestión del conocimiento. Gestionar el cambio incide de forma decisiva en la capacidad competitiva de una empresa, ya que proporciona diferenciación

<sup>72</sup> Véase Al-Shabandar *et al.* (2019) y Bedi *et al.* (2020).

<sup>73</sup> Véase Grint *et al.* (2017).

<sup>74</sup> Véase Canhoto (2021).

<sup>75</sup> Véase Merendino *et al.* (2018).

y permite una adaptación eficaz ante la transformación del entorno. La cultura y la estructura organizacional juegan un papel importante a la hora de permitir u obstaculizar la adopción de soluciones ante el cambio. La calidad de los datos, el talento y el desarrollo de habilidades son también muy relevantes en la gestión del cambio. En los tres aspectos, las herramientas de inteligencia artificial se están mostrando como ayudas potencialmente muy interesantes<sup>76</sup>. Los trabajadores del conocimiento deben estar conectados con personas o recursos adecuados para obtener conocimientos en los momentos adecuados para permitir una mejor toma de decisiones. Es posible que, debido al surgimiento de las capacidades de la AI, sea necesario implementar nuevas divisiones del trabajo entre humanos y máquinas inteligentes, diferentes de las utilizadas anteriormente en las organizaciones<sup>77</sup>. Los algoritmos de aprendizaje automático transformarán la experiencia en las organizaciones, remodelarán los límites laborales y ocupacionales y ofrecerán formas novedosas

de coordinación y control<sup>78</sup>.

Las finanzas son esencialmente una industria basada en el conocimiento. La gestión del conocimiento –*knowledge management* (KM) por su terminología en inglés– es el proceso de creación de valor a partir de los activos intangibles de una organización. Se trata de la mejor manera de aprovechar el conocimiento internamente en la organización y externamente en clientes y partes interesadas. También se afirma que KM es el proceso orientado a generar, difundir, utilizar y gestionar la información y el conocimiento de una organización. Las entidades financieras necesitan KM porque fomenta el aprendizaje, el crecimiento, la innovación y el éxito institucional, y porque las empresas basadas en el conocimiento han demostrado ser más resilientes y competitivas. La inteligencia artificial se ha convertido en una piedra angular de la gestión del conocimiento en el siglo XXI. Dado que tanto la AI como la KM están inexorablemente ligadas a la naturaleza del conocimiento y el aprendizaje, los avances

<sup>76</sup> Véase Banwo (2018) y Butler y O'brien (2019), Mercier (2021) y Guo y Polak (2023).

<sup>77</sup> Véase Taherdoost y Madanchian (2023).

<sup>78</sup> Véase Farajá *et al.* (2018).

recientes en inteligencia artificial pueden proporcionar nuevas bases para transformar la KM en las organizaciones<sup>79</sup>.

Una actividad relevante en el ámbito de la gestión del conocimiento es la creación de conocimiento inteligente –*intelligent knowledge* (IK) por su terminología en inglés–. El IK es una forma de conocimiento que comparte algunas características de la inteligencia como la memoria, el reconocimiento, la actualización y raciocinio y que se deriva de la intersección del conocimiento humano y conocimiento útil basado en *data-mining*. Mediante la construcción de un sistema de IK una organización puede alcanzar un determinado nivel de sabiduría de negocio (*business wise*)<sup>80</sup>. La diferencia básica entre información y conocimiento es la capacidad de acción del segundo concepto. Las personas que adoptan las decisiones siempre combinan diferentes tipos de datos y conocimientos disponibles en diversas formas en la organización. Uno de los pilares clave para avanzar

en el campo de la gestión del conocimiento y, en consecuencia, apoyar la toma de decisiones es la inteligencia artificial<sup>81</sup>. Compartir conocimientos, transformar conocimiento individual en conocimiento organizacional colectivo y reencarnar las organizaciones en “organizaciones de conocimiento” son ámbitos en los que las herramientas de inteligencia artificial pueden ayudar<sup>82</sup>. La captura y representación del conocimiento son básicas en la ingeniería del conocimiento. La inteligencia artificial puede resultar particularmente útil a la hora de representar el conocimiento adquirido, así como en su codificación y en los métodos de búsqueda y recuperación de éste en los sistemas de gestión del conocimiento. Las tecnologías avanzadas de AI, como las redes neuronales y los algoritmos genéticos proporcionan herramientas, por ejemplo, para el análisis semántico de textos, la minería de textos, la elaboración de perfiles de usuario, o la comparación de patrones. Como estas funciones están presentes en las tareas de gestión de

<sup>79</sup> Véase Sanzogni *et al.* (2017).

<sup>80</sup> Véase Shi y Li (2007).

<sup>81</sup> Véase Metaxiotis (2003).

<sup>82</sup> Véase Liebowitz (2001).

conocimientos las tecnologías basadas en inteligencia artificial pueden apoyar a toda la organización en este terreno<sup>83</sup>.

El potencial de la AI de aprendizaje profundo para la creación de conocimiento radica en su poder predictivo en situaciones como la previsión de probabilidades. A través de capacidades analíticas de autoaprendizaje y características para el reconocimiento de patrones, las organizaciones pueden aprovechar *big data* de manera sin precedentes<sup>84</sup>. Otra función clave de la KM es crear y mantener una memoria organizacional que rastree los recursos de conocimiento generados y adquiridos. En este ámbito, los algoritmos de autoaprendizaje basados en datos abren nuevas posibilidades para recolectar, clasificar, organizar, almacenar y recuperar *big data* que se generan en las organizaciones, incluidos datos que anteriormente se consideraban difíciles de manejar y analizar<sup>85</sup>. Además, si distribuir el conocimiento en toda la

organización es un requisito previo para aplicarlo eficazmente en la resolución de problemas y la toma de decisiones, existen barreras temporales, espaciales y funcionales que obstaculizan el intercambio. En este contexto, la inteligencia artificial puede ayudar a evitar que el conocimiento sea local y esté fragmentado en las organizaciones. Y, finalmente, por lo que respecta a la aplicación del conocimiento, herramientas impulsadas por AI, como los asistentes inteligentes, pueden mejorar las capacidades de recuperación y representación del conocimiento y acercar el conocimiento adecuado a los trabajadores de primera línea en tiempo real<sup>86</sup>. En este contexto, se espera una simbiosis entre humanos y máquinas, logrando así una inteligencia colaborativa, en la que la inteligencia artificial y los trabajadores mejorarán las fortalezas complementarias de cada uno<sup>87</sup>.

<sup>83</sup> Véase Birzniece (2011).

<sup>84</sup> Véase Farajá *et al.* (2018).

<sup>85</sup> Véase Paschen *et al.* (2020).

<sup>86</sup> Véase Maedche *et al.* (2019).

<sup>87</sup> Véase Wilson y Daugherty (2018) y Paschen *et al.* (2020).

## 5.7. EDUCACIÓN FINANCIERA

Desde la perspectiva de los clientes de servicios financieros, la inclusión financiera implica actualmente la posibilidad de acceso, alfabetización digital, conciencia sobre aspectos básicos de ciberseguridad y cierto nivel de educación financiera. La importancia de la alfabetización financiera ha cobrado mayor impulso desde la crisis de 2008 porque los ahorradores y prestatarios más vulnerables y menos educados financieramente parecen estar más expuestos a fraudes, muestran menor capacidad para hacer frente a gastos de emergencia, menor propensión a retirar depósitos de bancos en dificultades, y son más lentos a la hora de abandonar los mercados de valores antes situaciones de crisis<sup>88</sup>. Un conocimiento financiero básico ayuda a planificar la riqueza, evaluar los riesgos y adoptar decisiones de ahorro más sensatas<sup>89</sup>. La OCDE ha desarrollado un conjunto de recomendaciones para ayudar a los gobiernos y otras autoridades

a diseñar, implementar y evaluar políticas para apoyar la resiliencia financiera y el bienestar, así como para abordar las necesidades de los grupos vulnerables<sup>90</sup>.

Recientemente, la integración de tecnologías de inteligencia artificial como el aprendizaje automático y el procesamiento del lenguaje natural está proporcionado a los particulares y a los gestores de pequeñas empresas un acceso sin precedentes a información y conocimientos financieros, lo que les permite adoptar decisiones mejor informadas sobre sus préstamos e inversiones. La reducción de los costes de adquisición de educación o información financiera mejora acceso a las oportunidades de inversión para las personas que no son profesionales financieros o inversores cualificados y tecnologías como ChatGPT son capaces de proporcionar a los ahorradores ciertos conocimientos financieros, mitigando la asimetría de información en el proceso de

<sup>88</sup> Véase Lusardi y Mitchell (2014), Guiso y Viviano (2015), Brown *et al.* (2016) y Hasler *et al.* (2018).

<sup>89</sup> Véase Ameriks *et al.* (2003), Grohmann *et al.* (2018) y Bianchi (2018).

<sup>90</sup> Véase OCDE (2020).

inversión. Los recientes avances tecnológicos en IA explicable (XAI) y modelos lingüísticos como ChatGPT tienen el potencial de revolucionar la forma en que profesionales no financieros acceden a conocimientos sobre finanzas<sup>91</sup>. ChatGPT, es un *chatbot* asistente lanzado por OpenAI, que utiliza IA y procesamiento del lenguaje natural para responder consultas de entrada de los usuarios y generar respuestas similares a las humanas, y ha recibido una importante atención pública desde su introducción en noviembre de 2022. Las investigaciones sobre el posible éxito del uso de aplicaciones del tipo de ChatGPT en la formación financiera de los usuarios arrojan diversos resultados. Por una parte, la abundancia de datos y bibliografía académica sobre la que se ha formado ChatGPT permite que explique con eficacia el impacto de un conjunto de factores en los mercados financieros y, además, muestra capacidad para explicar conceptos financieros complejos de forma fácilmente comprensible para un profano. En segundo término, la especificación del público destinatario resulta determinante para la eficacia de la transmisión de conocimientos por parte

de ChatGPT. Así, cuando no se especifica el nivel, tiende a proporcionar respuestas de naturaleza más técnica y puede incluir términos y conceptos que no son fácilmente comprensibles para los no profesionales de las finanzas. Por el contrario, si se define claramente el público destinatario puede ajustar el lenguaje y la terminología de sus respuestas, para que sean más accesibles y fáciles de entender por parte de personas con conocimientos financieros limitados, o bien ofrecer explicaciones precisas y técnicas a un público más avanzado. En la misma línea, los resultados indican que la falta de precisión o claridad en la redacción de la pregunta puede desviar las respuestas de ChatGPT hacia terrenos menos útiles. Incluso una pregunta redactada con precisión puede ser problemática si ChatGPT la interpreta de manera diferente a la intención del usuario. Realizar las preguntas de forma adecuada es clave para el éxito del resultado. Pero interpretar correctamente la respuesta es también necesario. La mayoría de las respuestas proporcionadas por ChatGPT en materia financiera son correctas y elocuentes, pero algunas son engañosas y otras equivocadas,

<sup>91</sup> Véase Yue *et al.* (2023) y Ali y Aysan (2023).

y no es seguro que usuarios con escasa alfabetización financiera puedan discriminar. Además, para discriminar correctamente necesitan comprender el razonamiento de la inteligencia artificial. Por lo tanto, aunque la AI desempeñará un papel importante en el tratamiento de la educación financiera, todavía existen algunas limitaciones<sup>92</sup>. Las *alucinaciones* que a menudo se reportan en los resultados de inteligencia artificial generativa pueden hacer que su uso tenga un beneficio limitado cuando la precisión es esencial.

En otro orden de cosas, a pesar de que los *chatbots* tienen la capacidad de mejorar

la educación financiera y, en principio, podrían reducir la asimetría de conocimientos financieros entre tipos de usuarios, también se han puesto de manifiesto los riesgos que pueden conllevar en el contexto de *miseria digital* y *brecha digital*. Los *chatbots* y otros recursos tecnológicos para el aprendizaje pueden no estar disponibles para usuarios sin acceso a conexiones estables de Internet o que no disponen de materiales necesarios para participar en reuniones virtuales<sup>93</sup>.



## REFERENCIAS

- Abdolmohammadi, M. J. (1999) "A comprehensive taxonomy of audit task structure, professional rank and decision aids for behavioral research", *Behavioral Research in Accounting*, 11: 51-92.
- Abusalma, A. (2021) "The effect of implementing artificial intelligence on job performance in commercial banks of Jordan", *Management Science Letters*, 11 (7): 2061-2070.
- Accounting Today (2016). *The audit of the future*. Accounting Today.
- Adesuyi, F. A., Oluwafemi, O., Oludare, A. I., and Rick, A. V. (2013) "Secure authentication for mobile banking using facial recognition.
- Aggarwal, N., Wareham, S., and Lehmann, R. (2020) "Applications of machine learning in the identification, measurement and mitigation of money laundering", *Journal of Financial Compliance*, 4 (2): 140-166.

<sup>92</sup> Véase Robert (2023).

<sup>93</sup> Véase Gill et al. (2024).

- Agnew, H. (2016). Auditing: Pitch battle. Financial Times. 9 de mayo.
- Airehrour, D., Vasudevan Nair, N., and Madanian, S. (2018) "Social engineering attacks and countermeasures in the New Zealand banking system: Advancing a user-reflective mitigation model", *Information*, 9 (5): 110.
- Ali, H. and Aysan, A. F. (2023) "What will ChatGPT Revolutionize in Financial Industry?". Disponible en SSRN 4403372.
- Al-Jallad, K., Aljnidi, M., and Desouki, M. S. (2020) "Anomaly detection optimization using big data and deep learning to reduce false-positive", *Journal of Big Data*, 7 (1): 1-12.
- Alonso-Robisco, A. and Carbó, J. M. (2022a) "Can machine learning models save capital for banks? Evidence from a Spanish credit portfolio", *International Review of Financial Analysis*, 84: 102372.
- Alonso-Robisco, A. and Carbó, J.M. (2022b) "Inteligencia artificial y finanzas: una alianza estratégica", *Documentos Ocasionales*, 2222. Banco de España.
- Al-Shabandar, R., Lightbody, G., Browne, F., Liu, J., Wang, H., and Zheng, H. (2019) "The application of artificial intelligence in financial compliance management" en *Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing*: 1-6.
- Ameriks, J., Caplin, A., and Leahy, J. (2003) "Wealth accumulation and the propensity to plan", *The Quarterly Journal of Economics*, 118(3): 1007-1047.
- Aziz, S. y Dowling, M. (2019a) "Machine Learning and AI for Risk Management" en *Disrupting finance: FinTech and strategy in the 21st century*: 33-50. Lynn, T., Mooney, J. G., Rosati, P., y Cummins, M. (Editores). Springer Nature. Palgrave. MacMillan.
- Bahnsen, A. C., Aouada, D., Stojanovic, A., and Ottersten, B. (2016) "Feature engineering strategies for credit card fraud detection", *Expert Systems with Applications*, 51: 134-142.
- Baldwin, A. A., Brown, C. E., and Trinkle, B. S. (2006) "Opportunities for artificial intelligence development in the accounting domain: the case for auditing", *Intelligent Systems in Accounting, Finance & Management: International Journal*, 14(3): 77-86.
- Banerjee, R., Bourla, G., Chen, S., Kashyap, M., and Purohit, S. (2018) "Comparative analysis of machine learning algorithms through credit card fraud detection" en *IEEE MIT Undergraduate Research Technology Conference (URTC)*: 1-4. IEEE.
- Bank of England/Financial Conduct Authority (2019) *Machine learning in UK financial services*.
- Banwo, A. (2018) "Artificial intelligence and financial services: Regulatory tracking and change management", *Journal of Securities Operations & Custody*, 10 (4): 354-365.
- Banwo, A. (2018) "Artificial intelligence and financial services: Regulatory tracking and change management", *Journal of Securities Operations & Custody*, 10 (4): 354-365.
- Bao, Y., Ke, B., Li, B., Yu, Y. J., and Zhang, J. (2020) "Detecting accounting fraud in publicly traded US firms using a machine learning approach", *Journal of Accounting Research*, 58(1): 199-235.
- Baryannis, G., Samir, D., and Grigoris, A. (2019) "Predicting supply chain risks using machine learning: The trade-off between performance and interpretability", *Future Generation Computer Systems*, 101: 993-1004.
- Bedi, P., Goyal, S. B., and Kumar, J. (2020) "Basic structure on artificial intelligence: a revolution in risk management and compliance" en *3rd International Conference on Intelligent Sustainable Systems (ICISS)*: 570-576. IEEE.

- Bianchi, M. (2018) "Financial literacy and portfolio dynamics", *The Journal of Finance*, 73: 831-59.
- Birzniece, I. (2011) "Artificial intelligence in knowledge management: Overview and trends", *Computer Science* 1407-7493, 46: 5-11.
- Bizarro, P. A. and Dorian, M. (2017) "Artificial Intelligence: The Future of Auditing", *Internal Auditing*, 5: 21-26.
- Brennan, B., M. Baccala, and M. Flynn (2017) "Artificial intelligence comes to financial statement audits". *CFO.com*. 2 de febrero.
- Brown, M., Guin, B., and Morkoetter, S. (2016) "Deposit withdrawals from distressed commercial banks: The importance of switching costs" en *University of St. Gallen, School of Finance Research Paper: 2013-2019*.
- Butler, T. and O'Brien, L. (2019) "Artificial intelligence for regulatory compliance: Are we there yet?", *Journal of Financial Compliance*, 3 (1): 44-59.
- Canhoto, A. I. (2021) "Leveraging machine learning in the global fight against money laundering and terrorism financing: An affordances perspective", *Journal of Business Research*, 131: 441-452.
- Cavus, N., Mohammed, Y. B., and Yakubu, M. N. (2021) "An artificial intelligence-based model for prediction of parameters affecting sustainable growth of mobile banking apps", *Sustainability*, 13(11): 6206.
- Chan, A. (2023) "Can AI Be Used for Risk Assessments?", *CISM*. 28 de abril.
- Chaudhary, K., Yadav, J., and Mallick, B. (2012) "A review of fraud detection techniques: Credit card", *International Journal of Computer Applications*, 45 (1): 39-44.
- Chen, T. H. (2020) "Do you know your customer? Bank risk assessment based on machine learning", *Applied Soft Computing*, 86: 105779.
- Choi, D. and Lee, K. (2018) "An artificial intelligence approach to financial fraud detection under IoT environment: A survey and implementation", *Security and Communication Networks*.
- Chowdhury, M. M. U., Hammond, F., Konowicz, G., Xin, C., Wu, H., and Li, J. (2017) "A few-shot deep learning approach for improved intrusion detection" en *IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON): 456-462*. IEEE.
- Chukwuani, V. N. and Egiji, M. A. (2020) "Automation of Accounting Processes: Impact of Artificial Intelligence", *International Journal of Research and Innovation in Social Science (IJRISS)*, 4: 444-449.
- Couceiro, B., Pedrosa, I., and Marini, A. (2020) "State of the art of artificial intelligence in internal audit context" en *15th Iberian Conference on Information Systems and Technologies (CISTI): 1-7*. IEEE. Junio.
- Dada, E.G. (2017) "A hybridized approach to intrusion detection system", *Faculty Seminar Series: 1-8*.
- Dalal, S., Seth, B., Radulescu, M., Secara, C., and Tolea, C. (2022) "Predicting fraud in financial payment services through optimized hyper-parameter-tuned XGBoost model", *Mathematics*, 10(24): 4679.
- Daliri, S. (2020) "Using harmony search algorithm in neural networks to improve fraud detection in banking system", *Computational Intelligence and Neuroscience*.
- Deveau, R., Griffin, S.J. y Reis, S. (2023) "AI-powered marketing and sales reach new heights with generative AI", *Growth, Marketing & Sales Practice*. McKinsey & Company.

- Dey, D. (2017) "Growing importance of machine learning in compliance and regulatory reporting", *European Journal of Multidisciplinary Studies*, 2 (7): 255-258.
- Divakaran, D. M., Fok, K. W., Nevat, I., and Thing, V. L. (2017) "Evidence gathering for network security and forensics", *Digital Investigation*, 20, S56-S65.
- European Banking Authority (2021) *Discussion paper on machine learning for IRB models. Public hearing*. Diciembre.
- European Banking Authority (2023) *Follow-Up Report from The Consultation On The Discussion Paper On Machine Learning For IRB Models*. Agosto.
- Farajá, S., Pachidi, S., and Sayegh, K. (2018) "Working and organizing in the age of the learning algorithm", *Information and Organization*, 28 (1): 62-70.
- Fraisse, H. and Laporte, M. (2022) "Return on investment on artificial intelligence: The case of bank capital requirement", *Journal of Banking & Finance*, 138: 106401.
- Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., ... and Buyya, R. (2024) "Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots", *Internet of Things and Cyber-Physical Systems*, 4: 19-23.
- Goosen, R., Rontojannis, A., Deutscher, S., Rogg, J., Bohmayr, W., and Mkrтчian, D. (2018) "Artificial Intelligence is a Threat to Cybersecurity. It's also a Solution", *Boston Consulting Group (BCG), Tech. Rep.*
- Grint, R., O'driscoll, C., and Paton, S. (2017) *New technologies and anti-money laundering compliance*. London: Financial Conduct Authority. Search.
- Grohmann, A., Klühs, T., and Menkhoff, L. (2018) "Does financial literacy improve financial inclusion? Cross country evidence", *World Development*, 111: 84-96
- Guiso, L. and Viviano, E. (2015) "How much can financial literacy help?", *Review of Finance*, 19: 1347-82.
- Guo, H. and Polak, P. (2023) "Intelligent finance and change management implications", *Humanities and Social Sciences Communications*, 10 (1): 1-8.
- Han, H., Shiwakoti, R. K., Jarvis, R., Mordi, C., and Botchie, D. (2023) "Accounting and auditing with blockchain technology and artificial Intelligence: A literature review", *International Journal of Accounting Information Systems*, 48: 100598.
- Hasan, A. R. (2021) "Artificial Intelligence (AI) in accounting & auditing: A Literature review", *Open Journal of Business and Management*, 10(1): 440-465.
- Hasler, A., Lusardi, A., and Oggero, N. (2018) "Financial fragility in the US: Evidence and implications", en *GFLEC Working Paper 18/1*. Washington, DC: Global Financial Literacy Excellence Center, The George Washington University School of Business.
- Husain, A. R. A. M., Hamdan, A., and Fadhul, S. M. (2022) "The Impact of Artificial Intelligence on the Banking Industry Performance", *Future of Organizations and Work After the 4th Industrial Revolution: The Role of Artificial Intelligence, Big Data, Automation, and Robotics*: 145-156.
- Indriasari, E., Gaol, F. L., and Matsuo, T. (2019) "Digital banking transformation: Application of artificial intelligence and big data analytics for leveraging customer experience in the Indonesia banking sector", *Paper presented at the 8th International Congress on Advanced Applied Informatics*.
- Issa, H., Sun, T., and Vasarhelyi, M. A. (2016) "Research ideas for artificial intelligence in auditing: The



- formalization of audit and workforce supplementation”, *Journal of Emerging Technologies in Accounting*, 13(2): 1-20.
- Jain, A., Arora, D., Bali, R., and Sinha, D. (2021) “Secure authentication for banking using face recognition”, *Journal of Informatics Electrical and Electronics Engineering (JIEEE)*, 2(2): 1-8.
- Jain, R., Gour, B., and Dubey, S. (2016) “A hybrid approach for credit card fraud detection using rough set and decision tree technique”, *International Journal of Computer Applications*, 139(10): 1-6.
- Jarrahi, M. H., Askay, D., Eshraghi, A., and Smith, P. (2023) “Artificial intelligence and knowledge management: A partnership between human and AI”, *Business Horizons*, 66(1): 87-99.
- Kaur, N., Sahdev, S. L., Sharma, M., and Siddiqui, L. (2020) “The influence of artificial intelligence on the banking industry and how AI is changing the face of modern day banks”, *International Journal of Management*, 11(6): 577-585.
- Khan, A. U. S., Akhtar, N., and Qureshi, M. N. (2014) “Real-time credit-card fraud detection using artificial neural network tuned by simulated annealing algorithm”, en *Proceedings of International Conference on Recent Trends in Information, Telecommunication and Computing, ITC*: 113-121.
- Kim, G., Yi, H., Lee, J., Paek, Y., and Yoon, S. (2016) “LSTM-based system-call language modeling and robust ensemble method for designing host-based intrusion detection systems”, *arXiv*.
- Kochhar, K., Purohit, H., and Chutani, R. (2019) “The rise of artificial intelligence in banking sector” en *The 5th International Conference on Educational Research and Practice (ICERP)*: 127-135.
- Kokina, J. and Davenport, T. H. (2017) “The emergence of artificial intelligence: How automation is changing auditing”, *Journal of Emerging Technologies in Accounting*, 14(1): 115-122.
- Kumar, G. and Kumar, K. (2012) “The use of multi-objective genetic algorithm-based approach to create ensemble of ANN for intrusion detection”, *International Journal of Intelligence Science*, 2 (4A): 1-13.
- Kurode, T. (2018) “Review of applicability of Artificial Intelligence in various financial services in India”, *Journal of Advance Management Research*, 6.
- Kute, D. V., Pradhan, B., Shukla, N., and Alamri, A. (2021) “Deep learning and explainable artificial intelligence techniques applied for detecting money laundering—a critical review”. *IEEE Access*, 9: 82300-82317.
- Liebowitz, J. (2001) “Knowledge management and its link to artificial intelligence”, *Expert Systems with Applications*, 20 (1): 1-6.
- Ling, L., Gao, Z., Silas, M. A., Lee, I., and Le Doeuff, E. A. (2019) “An AI-based, multi-stage detection system of banking botnets”, *arXiv:1907.08276*.
- Lusardi, A. and Mitchell, O.S. (2014) “The Economic Importance of Financial Literacy: Theory and Evidence”, *Journal of Economic Literature* 52 (1): 34.
- Mackay, J. M.; Barr, S.H. and Kletke, M.G. (1992) “An Empirical Investigation of the Effects of Decision Aids on Problem Solving Processes” *Decision Sciences*, 23 (3): 648-672.
- Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., ... and Söllner, M. (2019) “AI-based digital assistants: Opportunities, threats, and research perspectives”, *Business & Information Systems Engineering*, 61: 535-544.
- Mahmoudi, N. and Duman, E. (2015) “Detecting credit card fraud by modified Fisher discriminant analysis”, *Expert Systems with Applications*, 42 (5): 2510-2516.



- Makhija, P. and Chacko, E. (2021) "Efficiency and advancement of artificial intelligence in service sector with special reference to banking industry" *Fourth Industrial Revolution and Business Dynamics: Issues and Implications*: 21-35.
- Manju, V. and Madhumathi, S. (2019) "Improving net banking security with face recognition-based biometric verification", *International Journal of Scientific Research in Computer Science, Engineering, and Information Technology*, 5(3): 82-91.
- Megantara, A. A. and Ahmad, T. (2021) "A hybrid machine learning method for increasing the performance of network intrusion detection systems", *Journal of Big Data*, 8(1): 1-19.
- Mercier, F. (2021) "Beyond the hype: How can the financial industry benefit from artificial intelligence?", *Journal of Securities Operations & Custody*, 13 (2): 163-169.
- Merendino, A., Dibb, S., Meadows, M., Quinn, L., Wilson, D., Simkin, L., and Canhoto, A. (2018) "Big data, big decisions: The impact of big data on board level decision-making", *Journal of Business Research*, 93: 67-78.
- Met, İ., Kabukçu, D., Uzunoğulları, G., Soyalp, Ü., and Dakdevir, T. (2020) "Transformation of business model in finance sector with artificial intelligence and robotic process automation", *Digital Business Strategies in Blockchain Ecosystems: Transformational Design and Future of Global Business*: 3-29.
- Metaxiotis, K., Ergazakis, K., Samouilidis, E., and Psarras, J. (2003) "Decision support through knowledge management: the role of the artificial intelligence", *Information Management & Computer Security*, 11 (5): 216-221.
- Minastireanu, E. A. and Mesnita, G. (2019) "An Analysis of the Most Used Machine Learning Algorithms for Online Fraud Detection", *Informatica Economica*, 23 (1).
- Mishra, M. K., and Dash, R. (2014) "A comparative study of Chebyshev functional link artificial neural network, multi-layer perceptron and decision tree for credit card fraud detection" en *International Conference on Information Technology*: 228-233. IEEE.
- Mishra, C., Gupta, D. L., and Singh, R. (2017) "Credit card fraud identification using artificial neural networks", *International Journal of Computer Systems*, 4 (07): 151-159.
- Modi, K. and Dayma, R. (2017) "Review on fraud detection methods in credit card transactions" en *International Conference on Intelligent Computing and Control (I2C2)*:1-5. IEEE.
- Mohammad, S. J., Hamad, A. K., Borgi, H., Thu, P. A., Sial, M. S., and Alhadidi, A. A. (2020) "How artificial intelligence changes the future of accounting industry", *International Journal of Economics and Business Administration*, 8(3): 478-488.
- Mohammed, I. A. (2021) "The interaction between artificial intelligence and identity and access management: an empirical study", *International Journal of Creative Research Thoughts (IJCRT)*: 668-671.
- Monirzadeh, Z., Habibzadeh, M., and Farajian, N. (2018) "Detection of violations in credit cards of banks and financial institutions based on artificial neural network and metaheuristic optimization algorithm", *International Journal of Advanced Computer Science and Applications*, 9 (1).
- Mubarek, A. M., and Adali, E. (2017) "Multilayer perceptron neural network technique for fraud detection" en *International Conference on Computer Science and Engineering (UBMK)*: 383-387. IEEE.



- Munoko, I., Brown-Liburd, H. L., and Vasarhelyi, M. (2020) "The ethical implications of using artificial intelligence in auditing", *Journal of Business Ethics*, 167: 209-234.
- Ng, A. (2017) "Why AI could be as transformational as electricity", *Artificial Intelligence, Computation & Data, Stanford Engineering*. 14 de marzo.
- Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., y Sun, X. (2011) "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature", *Decision Support Systems*, 50(3): 559-569.
- Nguyen, P. C., Nguyen, Q. T., and Le, K. H. (2021) "An ensemble feature selection algorithm for machine learning based intrusion detection system" en *8th NAFOSTED Conference on Information and Computer Science (NICS)*: 50-54. IEEE.
- Nkomo, B. K., and Breetzke, T. (2020) "A conceptual model for the use of artificial intelligence for credit card fraud detection in banks" en *Conference on Information Communications Technology and Society (ICTAS)*:1-6. IEEE.
- OECD (2020) *Recommendation of the Council on Financial Literacy*. OECD/LEGAL/0461. Paris.
- O'Leary, D. E. (1995) "AI in Accounting, Finance and Management. Intelligent Systems in Accounting", *Finance and Management*, 4: 149-153.
- Omoteso, K. (2012) "The Application of Artificial Intelligence in Auditing: Looking Back to the Future", *Expert Systems with Applications*, 39: 8490-8495.
- Paschen, J., Wilson, M., and Ferreira, J. J. (2020) "Collaborative intelligence: How human and artificial intelligence create value along the B2B sales funnel", *Business Horizons*, 63(3): 403-414.
- Patel, S. and Gond, S. (2014) "Supervised Machine (SVM) learning for credit card fraud detection", *International Journal of Engineering Trends and Technology*, 8(3), 137-139.
- Patidar, R., and Sharma, L. (2011) "Credit card fraud detection using neural network", *International Journal of Soft Computing and Engineering (IJSCE)*, 1: 32-38.
- Perols, J. (2011) "Financial statement fraud detection: An analysis of statistical and machine learning algorithms", *Auditing A Journal of Practice & Theory*, 30 (2): 19-50.
- Perols, J. L., Bowen, R. M., Zimmermann, C., and Samba, B. (2017) "Finding needles in a haystack: Using data analytics to improve fraud prediction", *The Accounting Review*, 92(2), 221-245.
- Pieptea, D. R. and Anderson, E. (1987) "Price and value of decision support systems", *MIS Quarterly*: 515-528.
- Raj, S. B. E. and Portia, A. A. (2011) "Analysis on credit card fraud detection methods" en *International Conference on Computer, Communication and Electrical Technology (ICCCET)*: 152-156. IEEE.
- Rajamani, R. and Rathika, M. (2015) "Credit Card Fraud Detection using Hidden Markov Model and Neural Networks" en *Proceedings of the UGC Sponsored National Conference on Advanced Networking and Applications*.
- Rani, T.O., Saravanan, M., Sahu, AK, Sagayam, KM and Elngar, AA (2022) "Predicting Online Fraudulent Transactions Using Machine Learning", *Research Square*.
- Rapoport, M. (2016) "Auditors count on tech for backup", *Wall Street Journal*, 8 de marzo.
- Robert, M. (2023) "Does Generative AI Solve the Financial Literacy Problem?" *Knowledge at Wharton*, 28 de junio.

- Rodríguez, J. F., Papale, M., Carminati, M., and Zanero, S. (2022) "A natural language processing approach for financial fraud detection" en *Proceedings of the Italian Conference on Cybersecurity ITASEC 2022*, Rome, Italy,, 3260: 135-149. CEUR-WS. org. 20 al 23 de Junio.
- Sanzogni, L., Guzman, G., and Busch, P. (2017) "Artificial intelligence and knowledge management: questioning the tacit dimension", *Prometheus*, 35 (1): 37-56.
- Sarvari, S., Sani, N. F. M., Hanapi, Z. M., and Abdullah, M. T. (2020) "An efficient anomaly intrusion detection method with feature selection and evolutionary neural network", *IEEE Access*, 8: 70651-70663.
- Schneider, G. P., Dai, J., Janvrin, D. J., Ajayi, K. and Raschke, R. L. (2015) "Infer, predict, and assure: Accounting opportunities in data analytics", *Accounting Horizons*, 29 (3): 719-742.
- Shah, S. A. R., Issac, B. and Jacob, S. M. (2018) "Intelligent intrusion detection system through combined and optimized machine learning", *International Journal of Computational Intelligence and Applications*, 17(02): 1850007.
- Shah, B. and Trivedi, C.B. H (2012) "Artificial Neural Network based Intrusion Detection System; A Survey", *International Journal of Computer Applications*, vol. 39 (6): 13-18.
- Shi, Y. and Li, X. (2007) "Knowledge management platforms and intelligent knowledge beyond data mining" en *Advances in Multiple Criteria Decision Making and Human Systems Management: Knowledge and Wisdom*, eds. Y. Shi, DL Olsen and A. Stam (IOS Press, Amsterdam, 2007): 272-288.
- Singh, G. and Khare, N. (2022) "A survey of intrusion detection from the perspective of intrusion datasets and machine learning techniques", *International Journal of Computers and Applications*, 44 (7): 659-669.
- Soltaniziba, S., and Balafar, M. A. (2015) "The study of fraud detection in financial and credit institutions with real data", *Computer Science and Engineering*, 5(3): 30-36.
- Stancheva-Todorova, E. P. (2018) "How artificial intelligence is challenging accounting profession", *Journal of International Scientific Publications. Economy & Business*, 12: 126-141.
- Staudemeyer, R. C. (2015) "Applying long short-term memory recurrent neural networks to intrusion detection", *South African Computer Journal*, 56 (1): 136-154.
- Sudha, C. and Raj, T. N. (2017) "Credit card fraud detection in internet using k-nearest neighbor algorithm", *International Journal of Computing Science*, 5: 22-30.
- Taherdoost, H. and Madanchian, M. (2023) "Artificial Intelligence and Knowledge Management: Impacts, Benefits, and Implementation", *Computers*, 12(4), 72.
- Tama, B. A. and Rhee, K. H. (2019) "An in-depth experimental study of anomaly detection using gradient boosted machine", *Neural Computing and Applications*, 31: 955-965.
- Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S., and Kuruwitaarachchi, N. (2019) "Real-time credit card fraud detection using machine learning" en *9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*: 488-493. IEEE
- Tschakert, N., Kokina, J., Kozłowski, S. and Vasarhelyi, M. (2016) "The next frontier in data analytics", *Journal of Accountancy*. Agosto.
- Van Liebergen, B. (2017) "Machine learning: A revolution in risk management and compliance?" *Journal of Financial Transformation, Capco Institute*, 45: 60-67.

- Virág, M. and Nyitrai, T. (2014) "Is there a trade-off between the predictive power and the interpretability of bankruptcy models? The case of the first Hungarian bankruptcy prediction model", *Acta Oeconomica*, 64 (4): 419-440.
- Whitehouse, T. (2015) "The technology transforming your annual audit", *Compliance Week*. 1 de diciembre.
- Wilson, H. J. and Daugherty, P. R. (2018) "Collaborative intelligence: Humans and AI are joining forces", *Harvard Business Review*, 96(4): 114-123.
- Yu, Y. and Bian, N. (2020) "An intrusion detection method using few-shot learning", *IEEE Access*, 8: 49730-49740.
- Yu, T. R. and Song, X. (2020) "Big data and artificial intelligence in the banking industry. Chapter 117 en *Handbook of Financial Econometrics, Statistics and Machine Learning*: 4025-4041.
- Yue, T., Au, D., Au, C. C., and Lu, K. Y. (2023) "Democratizing financial knowledge with ChatGPT by OpenAI: Unleashing the Power of Technology". Disponible en SSRN 4346152.
- Zareapoor, M. and Shamsolmoali, P. (2015) "Application of credit card fraud detection: Based on bagging ensemble classifier", *Procedia Computer Science*, 48: 679-685.
- Zemánková, A. (2019) "Artificial Intelligence and Blockchain in Audit and Accounting: Literature Review", *WSEAS Transactions on Business and Economics*, 16: 568-581.
- Zhang, C. A., Cho, S., and Vasarhelyi, M. (2022) "Explainable artificial intelligence (xai) in auditing", *International Journal of Accounting Information Systems*, 46: 100572.
- Zhang, Y., Xiong, F., Xie, Y., Fan, X., and Gu, H. (2020a) *The Impact of Artificial Intelligence and Blockchain on the Accounting Profession*. IEEE
- Zhao, N., Yen, D. C., and Chang, I. (2004) "Auditing in the e-Commerce Era", *Information Management & Computer Security*, 12: 389-400.
- Zhao, G., Zhang, C., and Zheng, L. (2017) "Intrusion detection using deep belief network and probabilistic neural network" en *IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, 1, (1): 639-642. IEEE.



## 6. Banca comercial

“La capacidad de los ordenadores para recopilar, digerir y actuar ha transformado fundamentalmente las finanzas, que han pasado de ser un negocio dominado por humanos a ser un negocio cohabitado por humanos y ordenadores en sinergia, en una especie de *cyborg finance*”

*Tom C.W. Lin (2012).*

Los bancos comerciales son los principales operadores del sistema de pagos, así como los principales tomadores de depósitos y proveedores de crédito a empresas y hogares. Si durante mucho tiempo los bancos comerciales operaron en nichos regionales y en mercados protegidos, el panorama de competencia cambió radicalmente con los procesos de desregulación,

y posteriormente con la digitalización, la disponibilidad de servicios electrónicos a través de ordenadores y teléfonos móviles y, más recientemente, con el advenimiento del fenómeno *Fintech* y de las *finanzas descentralizadas (DeFi)*<sup>94</sup>. Actualmente cada vez más organizaciones no bancarias dirigen sus actividades a las partes más rentables de la banca<sup>95</sup>, de forma que nuevos competidores

<sup>94</sup> Véase Hu (2005).

<sup>95</sup> Véase McKinsey (2016).

están ofreciendo servicios especializados antes reservados a los bancos. En palabras de Bill Gates, “la banca es necesaria, pero los bancos no lo son”.

En este entorno, la inteligencia artificial puede ser un remedio para muchos males de la banca minorista ya que puede aumentar la productividad y mejorar la experiencia del cliente. Además, la AI puede facilitar una próxima generación de productos y servicios impulsados digitalmente y, por otra parte, al arrojar luz sobre los factores clave que determinan el margen financiero puede contribuir a mejorar la gestión financiera.

Del mismo modo, los algoritmos de inteligencia artificial pueden contribuir a perfeccionar el análisis de riesgos<sup>96</sup>. La percepción sobre el impacto de la AI en esta industria es diferente desde una perspectiva operativa o estratégica<sup>97</sup>. No obstante, en la medida que la inteligencia artificial representa el triunfo de las capacidades *soft* frente a las capacidades *hard*, la generalización de los algoritmos de inteligencia artificial y su uso por parte de nuevos entrantes puede también representar un formidable desafío para los actuales jugadores de la banca comercial.

## 6.1. ACTIVIDAD DE MARKETING Y VENTAS

Existe consenso sobre la creciente importancia de la inteligencia artificial en relación con el papel de los gerentes de marketing<sup>98</sup>. En particular, la AI generativa tiene un enorme potencial para transformar la forma en que las instituciones financieras abordan esta actividad

ya que puede ayudar en:

- La automatización de tareas rutinarias comerciales repetitivas y de bajo valor, como la entrada de datos y la clasificación de respuestas, etc. Esta contribución permite liberar a los especialistas humanos

<sup>96</sup> Véase Boukadakis (2023).

<sup>97</sup> Véase Åberg y Khati (2018) y Ourdedine (2019).

<sup>98</sup> Véase Mogaji y Nguyen (2022).



que pueden concentrarse en tareas más complejas y de valor añadido con clientes actuales y potenciales<sup>99</sup>.

- El análisis de mercado. Los algoritmos de inteligencia artificial generativa pueden analizar los datos y comentarios de los clientes para proporcionar información y recomendaciones en la investigación de mercado y en el desarrollo de estrategias. Sus capacidades para combinar y analizar grandes cantidades de datos, como información demográfica, datos de clientes existentes y tendencias del mercado, permite que las herramientas de AI identifiquen nuevas audiencias en la industria financiera.
- El apoyo en el diseño de las campañas. Los algoritmos de AI pueden brindar ideas y sugerencias para campañas de marketing, crear descripciones de productos e incluso escribir publicaciones y artículos en blogs<sup>100</sup>. La adopción de AI puede mejorar significativamente la competitividad y agilidad de una institución, permitiéndole anticipar y responder a las

necesidades de los clientes de manera rápida y eficiente.

- La segmentación de clientes. Las herramientas de inteligencia artificial basadas en “modelos similares” permiten identificar atributos comunes en sus bases de clientes y otros usuarios de productos y servicios similares, así como modelos de propensión para determinar qué características hacen que, ante una determinada oferta, un consumidor tenga más probabilidades de realizar una conversión hacia la compra. Las iniciativas de venta cruzada o prospección han sido una fuente importante de crecimiento de los ingresos para los bancos. La mayoría de las veces se dirigen a clientes existentes, respecto de los que los bancos ya tienen información sobre atributos tales como situación financiera, comportamiento histórico de gastos y situación y movimiento de las carteras de valores. La oportunidad de la AI radica en su capacidad para anticipar con precisión y prontitud las necesidades no atendidas de estos clientes. Muchos clientes califican para múltiples productos, lo que dificulta la

<sup>99</sup> Deveau et al. (2023).

<sup>100</sup> Véase Marous (2022a).



tarea de los gerentes de ventas de productos financieros. Para mejorar el interés de los clientes a través de campañas digitales, los equipos de marketing digital pueden, en primer término, realizar exclusiones a nivel de producto y luego utilizar las listas proporcionadas por los sistemas de relaciones con el cliente (CRM) para encontrar el/los mejor/es producto/s para cada tipo cliente. Esto implica agrupar a los clientes y las audiencias en función de las pruebas para conocer sus necesidades y/o deseos.

- La creación de contenido. La AI generativa puede colaborar en el diseño y la redacción de borradores de textos de anuncios y del contenido de los correos electrónicos con ofertas comerciales.
- La personalización de campañas. Utilizando AI resulta posible crear ofertas *hiperpersonalizadas* basadas en el comportamiento, la personalidad y el historial de financiación y de ahorro/inversión del cliente individual. Para ejecutar estas estrategias los algoritmos de AI pueden

además generar contenido específico personalizado. Además, la incorporación de conocimientos de *finanzas conductuales* en los sistemas de AI pueden mejorar la eficacia de la venta<sup>101</sup>.

- La evaluación de campañas. Aprovechando el análisis predictivo y las recomendaciones basadas en datos los algoritmos de inteligencia artificial pueden mejorar el retorno de la inversión<sup>102</sup>.
- El soporte crítico al equipo de ventas. Las herramientas de AI también pueden ayudar a los expertos de ventas durante el proceso completo desde la propuesta hasta el cierre de las transacciones. Los gestores encargados de la venta pueden recibir soporte de un asistente virtual interno en un espacio temporal 24x7. Este asistente virtual puede también proporcionar orientación para la negociación en tiempo real con información predictiva basada en un análisis exhaustivo de los datos históricos de las transacciones, el comportamiento de otros

<sup>101</sup> Véase Königstorfer y Thalmann (2020)

<sup>102</sup> Véase Königstorfer y Thalmann (2020)

clientes y sobre los precios de la competencia<sup>103</sup>.

- La atención post-cierre. Incluso después del cierre de la transacción, el uso de AI puede ser utilizado para generar una “bienvenida” al servicio para el cliente con contenido, al tiempo que un

*chatbot* puede proporcionar respuestas inmediatas a las preguntas iniciales de los clientes. Adicionalmente, funcionan ya sistemas basados en AI para detectar síntomas de alerta respecto de la decisión de abandono del servicio por parte de los clientes.

## 6.2. CHATBOTS Y ASISTENTES VIRTUALES INTELIGENTES (IVAS)

A menudo los bancos de *retail* tienen dificultades para lidiar con un volumen creciente de consultas de los centros de llamadas y de correos electrónicos de los clientes, ya que su modelo tradicional de servicio al cliente tiene economías de escala limitadas. Desde hace años, los centros de atención de los bancos comerciales se enfrentan a un drástico crecimiento en el volumen de llamadas, de forma que mantener personal de apoyo capacitado para manejar las llamadas entrantes es cada vez más difícil. Los síntomas de este desafío se manifiestan en largos tiempos de espera de consumidores desesperados por

recibir ayuda y una experiencia frustrante para el cliente, que erosiona su lealtad. Para reducir el número de contactos telefónicos, las empresas han introducido masivamente sistemas de atención a través de sus *websites*. La comunicación con los clientes a través de mensajes digitales (inicialmente de texto) surgieron hace ya varias décadas. El modelo tradicional más simple consistía en una conversación *on line* entre un cliente que preguntaba y un experto humano al otro extremo de la línea, escribiendo las respuestas. Posteriormente surgieron los *chatbots* o programas de software diseñados para simular

<sup>103</sup> Véase Deveau *et al.* (2023).

conversaciones con seres humanos. El modelo básico de *chatbot* se basaba en respuestas programadas. Inicialmente se trataba de herramientas digitales, sin funcionalidad de lenguaje natural. El usuario introducía una pregunta o un fragmento de texto y, si éste coincidía con algo parecido a una *pregunta frecuente* en la base de datos del *bot*, el usuario recibía una respuesta en la que el texto apenas variaba. Estos *bots* seguían un conjunto predefinido de reglas y patrones de conversación, y respondían a comandos específicos, de forma que no podían comprender ni generar respuestas más allá de su programación inicial. Bajo este enfoque el sistema aprendía mediante la intervención humana a posteriori, ya que se revisaban las respuestas inexactas y se introducían nuevas respuestas enlatadas más precisas y nuevas reglas de relación. El resultado era en muchos casos decepcionante, ya que los *chatbots* de segunda generación no eran capaces de determinar el contexto de las preguntas (una tarea relativamente fácil para un ser humano, pero para la que la mayoría de las soluciones automatizadas mostraban dificultades insalvables en una conversación). Comprender el contexto es necesario para proporcionar una

respuesta precisa que evite que un consumidor necesite buscar ayuda adicional en sus consultas. En consecuencia, el sistema evolucionó hacia *chatbots* con mensajería bidireccional. Bajo este enfoque, el *bot* se basaba también en respuestas enlatadas ante preguntas frecuentes, pero se involucraba a un humano cuando el usuario indicaba que la respuesta no era adecuada. En este contexto, se requería una solución tecnológica capaz de interpretar la compleja mecánica del lenguaje natural para atender adecuadamente la mayoría de las consultas de los usuarios y reducir de forma efectiva el tráfico hacia los *call-centres*.

La emergencia de herramientas de inteligencia artificial con capacidad para comprender y generar lenguaje natural ha permitido el desarrollo de *chatbots* más avanzados, utilizando LLM que permite comprender el lenguaje mediante modelos de *deep learning* entrenados con millones de parámetros y gran cantidad de datos recogidos de Internet. Estos modelos procesan y analizan la información a través de redes neuronales para comprender y responder preguntas, capturando la semántica y la sintaxis del lenguaje humano y memorizando a su vez

datos de la fase de entrenamiento. Pueden aprender de las conversaciones pasadas y mejorar sus respuestas con el tiempo, sin necesidad de intervención humana. Una vez se ha entrenado el modelo éste realiza dos tareas principales: predicción y generación de texto. El entrenamiento requiere convertir el texto en una representación numérica que sea capaz de interpretar el modelo, además de los parámetros y una función de pérdida para medir los resultados. El objetivo del modelo será optimizar los parámetros para minimizar la pérdida, repitiendo el proceso hasta que el nivel de precisión sea aceptable. Los *chatbots de procesamiento de lenguaje natural han continuado* mejorando y pueden comprender y generar respuestas más humanas, identificando el contexto de una conversación, lo que les permite responder de manera más natural.

Cuando los *chatbot* incorporan funcionalidades de procesamiento de lenguaje natural generativo se habla de *asistentes virtuales inteligentes* –IVAs por sus siglas en inglés–, que pueden además integrar avatares con mensajería

de voz. Los asistentes virtuales avanzados, que incorporan inteligencia artificial, análisis predictivo y procesamiento de lenguaje natural, son el tipo que muchas entidades financieras están ya introduciendo. Los asistentes virtuales pueden conversar y responder preguntas con una interacción similar a la humana durante toda la franja horaria cada día, lo que brinda la oportunidad de mejorar la satisfacción del cliente y reducir la carga de los equipos de soporte. En este contexto, son numerosos los estudios orientados a determinar el grado de aceptación de esta nueva tecnología desde la perspectiva del cliente, y a determinar si la introducción de un IVA logra reducir los costes, manteniendo al mismo tiempo la satisfacción del cliente<sup>104</sup>. En un sentido similar las entidades desean conocer si los consumidores prefieren un uso exclusivo de los IVA o un modelo híbrido (complementado parcialmente con asistencia humana. Y, del mismo, se está analizando en qué medida el conocimiento y la experiencia tecnológica previa tienen influencia en las actitudes de los clientes<sup>105</sup>.

<sup>104</sup> Véase Ates (2017).

<sup>105</sup> Véase Bergström et al. (2018).

Algunas entidades están además “humanizando” sus asistentes virtuales inteligentes (IVAs) dándoles nombres y dotándoles de avatares, como un paso que enlazará además con el desarrollo de sus servicios financieros en el Metaverso. Esto ha dado lugar a determinadas investigaciones sobre cómo los IVAs y el antropomorfismo percibido influyen en las respuestas de los clientes en el entorno de servicios financieros. El aumento en la adopción de *chatbots* e IVAs por parte de las entidades financieras se ha ido producido a medida que los consumidores se sienten más cómodos usando dispositivos móviles para comunicarse sobre sus finanzas. Una de las conclusiones de los estudios es que los consumidores tienden a desconocer la información disponible y las aplicaciones instaladas por las entidades, así como las capacidades que proporcionan los algoritmos de inteligencia artificial integrados en sus sistemas<sup>106</sup>. Además, se ha concluido que únicamente existe una aceptación elevada de esta tecnología cuando el cliente percibe una conversación continua, lo que se conoce como

«experiencia sincrónica».

De cara al futuro, se espera que los *chatbots* y los IVAs puedan incluso reconocer cientos de «intenciones» de los clientes, lo que mejorará su capacidad. Y se anticipa que los bancos entrenarán modelos individuales de inteligencia artificial para cada cliente utilizando los datos financieros personales y el historial de transacciones de éste. Con el tiempo, estos “robots/banqueros personales” podrían obtener conocimientos profundos sobre las circunstancias únicas de cada cliente para ofrecer mediante un interfaz conversacional recomendaciones personalizadas sobre su situación financiera global, el impacto de la adquisición de una vivienda, la planificación de su jubilación y otras cuestiones financieras<sup>107</sup>. Además, con los algoritmos de aprendizaje automático, la interacción continuada con el cliente permitirá a los asistentes virtuales “comprender” mejor las preferencias de éstos y adaptar sus recomendaciones.

Los modelos de inteligencia artificial general tienen la capacidad de usar la razón,

<sup>106</sup> Véase El-Gohary *et al.* (2021).

<sup>107</sup> Véase Marous (2023b)

representar conocimiento, comunicarse y aprender, pero no tienen un conocimiento real del mundo, no poseen conciencia como un ser humano ni son capaces de sentir ni empatizar. Actualmente los robots simplemente aplican patrones estadísticos y reglas aprendidas para generar un texto coherente. Por ello, previsiblemente los *chatbots* e IVAs nunca reemplazarán por completo a los humanos en el

servicio de atención, aunque se implementarán para brindar una especie de experiencia «cyborg», que integra la tecnología de AI con asesores humanos en sucursales y centros de llamadas. Humanos y máquinas trabajarán en conjunto para garantizar que no surjan problemas en las transiciones de los clientes entre los distintos canales digitales y físicos en una «mezcla sinérgica» de personas y tecnología.

### 6.3. GESTIÓN DE PAGOS

El uso de la inteligencia artificial en la gestión de pagos puede ayudar a que una entidad financiera forme parte de un ecosistema financiero, donde los usuarios pueden acceder a servicios con otras empresas y plataformas con las que hacen negocios. Esto convierte a las instituciones financieras en un centro esencial para sus clientes cuando participan en el comercio, en lugar de simplemente un *back-end* como procesador de pagos<sup>108</sup>. Además, el uso de nuevas aplicaciones de pago impulsadas por AI puede ayudar a los

consumidores a gastar de manera más inteligente e incorporar ciertas protecciones. Dado el poder de las herramientas de inteligencia artificial para analizar datos de flujos de pago diferentes es posible construir modelos altamente precisos de predicción de gasto. Pero, además, estos modelos tienen capacidad para estimar el riesgo de morosidad de los clientes. Como resultado, las aplicaciones de pago basadas en inteligencia artificial ofrecen a las entidades la oportunidad de interactuar con un cliente logrando mayor involucración en su

<sup>108</sup> Véase Yurcan (2021).

actividad financiera. Las entidades financieras pueden ahora ser proactivos y trabajar con los clientes para brindarles una mejor experiencia posible y ayudarles a gastar de manera más inteligente y, en última instancia, administrar mejor su *bienestar financiero*. Lo que puede traducirse en clientes más leales y rentables<sup>109</sup>.

Por otra parte, las transacciones fraudulentas en línea están causando pérdidas a personas y empresas de forma creciente como consecuencia de la generalización del uso de pagos on-line y el comercio electrónico. El diseño de algoritmos eficientes de detección de fraude es una oportunidad para reducir estas pérdidas. El aprendizaje automático y las técnicas estadísticas desempeñan un papel vital en la detección de transacciones fraudulentas. Las estafas en procesos del tipo *authorised push payments* (APP), en su terminología en inglés, se producen como consecuencia de que se engaña a los consumidores para que envíen pagos con pretextos falsos a cuentas bancarias controladas por estafadores.

Como los pagos realizados mediante sistemas de pago en tiempo real son irrevocables, las víctimas no pueden revertir las transacciones una vez que se dan cuenta de que han sido estafadas. Resulta posible aprovechar las aplicaciones de AI para detectar y gestionar las debilidades humanas<sup>110</sup>. A tal fin resulta posible generar detectores de abuso que puntúan periódicamente todas las transacciones e identifican casos de alto nivel de riesgo de abuso en millones de registros<sup>111</sup>. Sin embargo, la implementación de los modelos de detección de fraude mediante inteligencia artificial no está exenta de dificultades debido a la falta de datos, la naturaleza confidencial de los datos y el hecho de que el muestreo se realiza en presencia de sesgo o desequilibrio de clases (dado que las transacciones no fraudulentas representan un porcentaje elevadísimo de las transacciones totales).

<sup>109</sup> Véase Yurcan (2021).

<sup>110</sup> Véase Ma et al. (2018)

<sup>111</sup> Véase Leontjeva et al. (2023).

## 6.4. CREDIT SCORING DE PRÉSTAMOS HIPOTECARIOS Y FINANCIACIÓN AL CONSUMO

El éxito de la banca depende directamente de su capacidad de controlar y gestionar los riesgos asumidos y uno de los más relevantes es el riesgo de crédito. El *credit scoring* o calificación crediticia se convirtió hace tiempo en uno de los principales métodos analíticos de las entidades financieras para evaluar el riesgo crediticio. Las técnicas de calificación crediticia han sido vitales para permitir el enorme crecimiento del crédito minorista en las últimas décadas. Sin este tipo de herramientas de evaluación de riesgos precisa y operada automáticamente les hubiera resultado imposible a las entidades prestamistas ampliar sus carteras de préstamos tal como lo han hecho. El proceso de calificación crediticia como evaluación independiente tiene como objetivo evaluar hasta qué punto un potencial prestatario es capaz y está dispuesto a cumplir con sus obligaciones de pago, basándose en un complejo análisis de factores de riesgo del agente evaluado. Estos factores se refieren a aspectos personales del prestatario (edad, sexo, estado

civil, formación, profesión, sector ocupacional, vida laboral, lugar de residencia, etc.), su reputación de pagos, sus ingresos, su situación financiera y patrimonial, las garantías aportadas en la transacción, las condiciones financieras del préstamo y las características económicas de su entorno. Desarrollar sistemas fiables de calificación crediticia permite no sólo asegurar en mayor medida el cobro de los créditos y disminuir el riesgo, pero también reducir el coste del análisis crediticio, y adoptar decisiones de forma más rápida. Contar con un sistema eficaz de evaluación del riesgo crediticio se ha convertido en un factor crucial para obtener ventajas competitivas, ya que incluso una ligera mejora en la precisión de la calificación puede traducirse en una importante reducción en las ratios de morosidad y, consecuentemente en las pérdidas asumidas.

En el desarrollo de modelos de *credit scoring* los modelos tradicionales más comúnmente utilizados son el *análisis discriminante lineal* (LDA)<sup>112</sup>, la

<sup>112</sup> La LDA es una técnica de minería de datos para manejar problemas de clasificación que

*regresión logística* (LR)<sup>113</sup>, y el *análisis probit* (PA)<sup>114</sup>. El punto débil del LDA es el supuesto de relación lineal entre las variables (que en la práctica suele ser no lineal) y una excesiva sensibilidad de los resultados a la desviación del supuesto de normalidad multivariante. Por su parte, la LR predice resultados dicotómicos y una relación lineal entre variables en el exponente de la función logística, pero no requiere el supuesto de normalidad multivariante<sup>115</sup>. No obstante, presenta como restricción el hecho de que la variable de salida debe ser binaria y que no exista multicolinealidad entre las variables predictoras. En los tres casos, los modelos parecen útiles, pero adolecen de precisión<sup>116</sup>. En este contexto, a diferencia de los métodos estadísticos, los métodos

basados en AI no asumen distribuciones específicas de los datos. Estos métodos extraen automáticamente el conocimiento de las muestras de entrenamiento.

Las herramientas de inteligencia artificial aplicadas al *crédit scoring* no dejan de aumentar e incluyen árboles de decisión (DT), máquinas de vectores soporte (SVM), algoritmos genéticos (GA), redes neuronales artificiales ANN y KNN, incluyendo redes de retroalimentación multicapa y redes de retropropagación, así como *multi-classifer system* (MCS), bosques aleatorios (RF), Naïve Bayes, Elastic Net, algoritmos neurodifusos, Splines de regresión adaptativa multivariante (MARS), árbol de clasificación y regresión (CART), razonamiento basado en casos (CBR) y otros<sup>117</sup>. Los

---

reduce las variables observadas a un número menor de dimensiones, lo disminuye el número de características que deben ser consideradas por los clasificadores. En lugar de eliminar directamente variables irrelevantes o redundantes del espacio de características original, LDA transforma las variables originales mediante combinación lineal en un nuevo subconjunto de variables.

<sup>113</sup> La LR es una técnica de modelización estadística en la que la probabilidad de un resultado dicotómico (cero o uno) se relaciona con un conjunto de posibles variables predictivas. El objetivo de un modelo de regresión logística en el *scoring* crediticio es determinar la probabilidad condicional de que una observación específica pertenezca a una clase, dados los valores de las variables independientes de ese solicitante de crédito

<sup>114</sup> En el análisis probit, una combinación lineal de variables independientes se transforma en su valor de probabilidad acumulada a partir de una distribución normal. El método requiere encontrar el valor de los coeficientes en esta combinación lineal, de modo que esta probabilidad acumulada sea igual a la probabilidad real de que el resultado binario sea uno.

<sup>115</sup> Véase Ghodselahi y Amirmadhi (2011).

<sup>116</sup> Véase Sustersic *et al.* (2009).

<sup>117</sup> Véase West (2000), Malhotra y Malhotra (2003), Baesens *et al.* (2003), Van Gestel *et al.* (2003), Lee y Chen (2005), Bensus *et al.* (2005), Angelini *et al.* (2008), Abdou *et al.* (2008), Wang (2008),

resultados indican que las herramientas con inteligencia artificial pueden mejorar significativamente la capacidad predictiva de los impagos.

Por ejemplo, el modelo de árbol de decisión aplicado en el análisis de *credit scoring* es un método no paramétrico, en el sentido de que no intenta aprender parámetros sobre los cuales calificar los atributos de los solicitantes del crédito. El sistema memoriza ciertas características clave de los datos (las divisiones binarias y los valores de corte correspondientes para cada característica) y considera todas las divisiones posibles para encontrar la mejor, y el subárbol que mejor se ajusta se selecciona en función de su tasa de error general (clasificación errónea). La clasificación de un nuevo solicitante de crédito se determina en función de la clasificación del *nodo hoja* resultante después de atravesar el modelo de árbol utilizando los atributos del solicitante como entrada<sup>118</sup>.

Otro modelo gráfico aplicado al *scoring* crediticio

es la red neuronal, una técnica estadística avanzada que calcula las ponderaciones en forma de puntos para las variables independientes de casos anteriores de solicitantes solventes y no solventes. Las redes neuronales son una técnica que aprende a través de un proceso de entrenamiento de prueba y error multicapa (una capa de entrada, una capa oculta y una capa de salida). Cada nodo en la capa de entrada denota una variable predictiva y aporta el valor a la red. Los nodos en las capas ocultas combinan y transforman estos valores para que coincidan con los valores objetivo en la capa de salida. Cada nodo en la capa de salida representa una variable dependiente y, en el caso de la calificación crediticia, generalmente hay solo un nodo en esta capa. Las redes neurales se han aplicado a la calificación crediticia debido a su capacidad para entrenar funciones lineales y no lineales para describir datos<sup>119</sup>.

Por lo que respecta a la técnica de máquina de vectores de soporte también empleada para resolver problemas de clasificación

Tsai y Wu (2008), Yu *et al.* (2008 y 2009), Zhou *et al.* (2009), Luo *et al.* (2009), Yap *et al.* (2011), Cao *et al.* (2013) Ala'Raj y Abbod (2016a y 2016b), Bacham y Zhao (2017), Wójcicka (2017), Zhang (2017), o Guegan *et al.* (2018).

<sup>118</sup> Véase Harris (2015).

<sup>119</sup> Véase Harris (2015).

de prestatarios, el objetivo de este algoritmo es encontrar el hiperplano de separación óptimo entre las clases maximizando el margen entre ellas. Aquí, los puntos que se encuentran en los límites se denominan vectores de soporte, mientras que el centro del margen se denomina hiperplano de separación óptimo. Al igual que la red neural artificial, esta técnica a menudo es criticada por comportarse como una *caja negra*.<sup>120</sup>

A pesar de estos avances, los analistas de crédito de las instituciones financieras se ven presionados para buscar continuamente mejoras en el desempeño de los clasificadores en un intento de mitigar el riesgo crediticio que enfrentan sus instituciones, por lo que la investigación continúa avanzando. Recientemente, los investigadores han propuesto *enfoques híbridos* para mejorar las evaluaciones ya que, en general, la aplicación de un único modelo con un grupo fijo de muestras de entrenamiento puede generar algún tipo de sesgo inductivo. Una forma eficaz de reducir el sesgo es el modelo de conjunto, que aprovechan la

diversidad para reducir el error de varianza y generar mejores resultados. La idea principal del método de conjunto es combinar un conjunto de modelos, cada uno de los cuales resuelve el mismo problema original, para obtener un modelo compuesto con estimaciones o decisiones más precisas y confiables. En general, el aprendizaje conjunto es un enfoque que combina los resultados de dos o más clasificadores entrenados individualmente para realizar predicciones más precisas. Aunque los clasificadores se entrenan individualmente para producir sus propias decisiones, luego se combinan mediante votación para producir una decisión de conjunto<sup>121</sup>.

Adicionalmente, los análisis de crédito han comenzado a incorporar el hecho de que los costes asociados con la clasificación incorrecta de un cliente poco solvente son mayores que los asociados con la clasificación errónea de un cliente solvente. Es decir, el hecho de que diferentes clases de errores afectan de forma cuantitativamente

<sup>120</sup> Véase Harris (2015).

<sup>121</sup> Véase Huang *et al.* (2007), Chen *et al.* (2009), Chen y Li (2010), Zhou *et al.* (2010), Ghodselahi (2011), Koutanaei *et al.* (2015), Nalić *et al.* (2020).

diferente, lo que implica ajustes en el modelo<sup>122</sup>.

El uso de herramientas de inteligencia artificial en el *credit scoring*, que tienden a mejorar los resultados de la clasificación de potenciales prestatarios, está sujeta al problema del elevado consumo de recursos computacionales y de tiempo cuando el entrenamiento utiliza muestras de datos del mundo real de un tamaño razonable. La capacidad de computación utilizada en muchos de estos algoritmos crece exponencialmente a medida que aumenta el tamaño del conjunto de datos de entrenamiento. Desde esta perspectiva, la aplicación de algoritmos computacionalmente

más eficientes en el ámbito de la calificación crediticia se está volviendo cada vez más importante a medida que crece el tamaño de los conjuntos de datos históricos. En este contexto, una parte de la investigación se está dirigiendo al análisis del balance o *trade-off* entre tiempo de cálculo y precisión para métodos alternativos con el objetivo de optimizar el resultado. Un método compatible con este enfoque consiste en elegir la propuesta de muestra estratificada que optimiza la reducción del tiempo de cómputo y la desviación/precisión del resultado<sup>123</sup>.



## REFERENCIAS

- Abdou, H., Pointon, J., and El-Masry, A. (2008) "Neural nets versus conventional techniques in credit scoring in Egyptian banking", *Expert Systems with Applications*, 35(3): 1275-1292.
- Åberg, E. and Khati, Y. (2018) "Artificial Intelligence in Customer Service: A Study on Customers' Perceptions regarding IVR Services in the Banking Industry". Disponible en Digitala Vetenskapliga Arkivet.
- Ala'Raj, M. and Abbod, M. F. (2016a) "A New Hybrid Ensemble Credit Scoring Model Based on Classifiers Consensus System Approach", *Expert Systems with Applications*, 64: 36-55.

<sup>122</sup> Véase Yu *et al.* (2009a).

<sup>123</sup> Véase Hens y Tiwari (2012)

- Ala'Raj, M. and Abbod, M. F. (2016b) "Classifiers Consensus System Approach for Credit Scoring", *Knowledge-Based Systems* 104: 89-105.
- Angelini, E., Tollo, G. D., and Roil, A. (2008) "A neural network approach for credit risk evaluation", *The Quarterly Review of Economics and Finance*, 48: 733-755.
- Ates, M. (2017) "Artificial intelligence in banking: A Case Study of the Introduction of a Virtual Assistant into Customer Service".
- Bacham, D, and Zhao, J. (2017) "Machine Learning: Challenges and Opportunities in Credit Risk Modeling", *Moddy's Analytics*, Julio.
- Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., and Vanthienen, J. (2003) "Benchmarking state-of-art classification algorithm for credit scoring", *Journal of Operational Research Society*, 54: 627-635.
- Bensic, M, Sarlija, N. and Zekic-Susac, M. (2005) "Modeling Small-Business Credit Scoring by Using Logistic Regression, Neural Networks and Decision Trees", *Journal of Intelligent Systems in Accounting, Finance and Management*, 13 (3): 133-150.
- Bergström S.I., Svenningsson, P., and Thoresson, A. (2018) "The Era of Artificial Intelligence in Swedish Banking: Exploring Customer Attitudes Towards AI as a Substitute to Brick and Mortar Offices", disponible en *Digitala Vetenskapliga Arkivet*.
- Boukadakis, M. (2023) "Fintech Vendors Must Stop Giving Small Institutions Short Shrift on AI Tech", *The Financial Brand*. 12 de mayo.
- Cao, J., Lu, H., Wang, W. and Wang, J. (2013) "A Loan Default Discrimination Model Using Cost-Sensitive Support Vector Machine Improved by PSO", *Information Technology and Management*, 14: 193-204.
- Chen, F. L. and Li, F. C. (2010) "Combination of feature selection approaches with SVM in credit scoring", *Expert Systems with Applications*, 37(7): 4902-4909.
- Chen, W., Ma, C., and Ma, L. (2009) "Mining the customer credit using hybrid support vector machine technique", *Expert Systems with Applications*, 36(4): 7611-7616.
- Cocheo, S. (2020) "AI's Real Impact on Banking: The Critical Importance of Human Skills", *The Financial Brand*, 3 de noviembre.
- Deveau, R., Griffin, S.J. and Reis. S. (2023) "AI-powered marketing and sales reach new heights with generative AI", *Growth, Marketing & Sales Practice*. McKinsey & Company.
- El-Gohary, H., Thayaseelan, A., Babatunde, S., and El-Gohary, S. (2021) "An exploratory study on the effect of artificial intelligence-enabled technology on customer experiences in the banking sector", *Journal of Technological Advancements (JTA)*, 1(1):1-17.
- Figini, S., Bonelli, F., y Giovannini, E. (2017) "Solvency prediction for small and medium enterprises in banking", *Decision Support Systems*, 102: 91-97.
- Ghodselahe, A. (2011) "A hybrid support vector machine ensemble model for credit scoring", *International Journal of Computer Applications*, 17(5): 1-5.
- Ghodselahe, A. and Amirmadhi, A. (2011) "Application of artificial intelligence techniques for credit risk evaluation", *International Journal of Modeling and Optimization*, 1(3): 243.
- Guegan, D. Addo, P. and Hassani, B. (2018) "Credit risk analysis using machine and deep learning models", *Risks* 6: 38.



- Harris, T. (2015) "Credit scoring using the clustered support vector machine", *Expert Systems with Applications*, 42(2): 741-750.
- Hens, A. B. and Tiwari, M. K. (2012) "Computational time reduction for credit scoring: An integrated approach based on support vector machine and stratified sampling method", *Expert Systems with Applications*, 39(8): 6774-6781.
- Hu, X. (2005) "A data mining approach for retailing bank customer attrition analysis", *Applied Intelligence*, 22: 47-60.
- Huang, C. L., Chen, M. C., and Wang, C. J. (2007) "Credit scoring with a data mining approach based on support vector machines", *Expert Systems with Applications*, 33(4): 847-856.
- Joyce, L. (2018) "AI and The Banking Industry's \$1 Trillion Opportunity", *The Financial Brand*. 29 de Mayo.
- Jung, C., Mueller, H., Pedemonte, S., Plances, S., y Thew, O. (2019). *Machine learning in UK financial services*. Bank of England and Financial Conduct Authority.
- Khandani, A. E., Kim, A. J., y Lo, A. W. (2010) "Consumer credit-risk models via machine-learning algorithms", *Journal of Banking & Finance*, 34(11): 2767-2787.
- Königstorfer, F. and Thalmann, S. (2020) "Applications of Artificial Intelligence in commercial banks—A research agenda for behavioral finance", *Journal of Behavioral and Experimental Finance*, 27, 100352.
- Koutanaei, F. N., Sajedi, H., and Khanbabaee, M. (2015) "A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit scoring", *Journal of Retailing and Consumer Services*, 27: 11-23.
- Lee, T. S. and Chen, I. F. (2005) "A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines", *Expert Systems with Application*, 28: 743-752.
- Leontjeva, A., Richards, G., Sriskandaraja, K., Perchman, J., and Pizzato, L. (2023). "Detection of Abuse in Financial Transaction Descriptions Using Machine Learning", *arXiv*: 2303.08016.
- Li, S. T., Shiue, W. and Huang, M. H. (2006) "The evaluation of consumer loans using support vector machines", *Expert Systems with Applications*, 30(4): 772-782.
- Lin, T. C. (2012) "The new investor", *UCLA Law Review*, 60: 678-735.
- Luo, S. T., Cheng, B. W., and Hsieh, C. H. (2009) "Prediction model building with clustering-launched classification and support vector machines in credit scoring", *Expert Systems with Applications*, 36(4): 7562-7566.
- Ma, X., Sha, J., Wang, D., Yu, Y., Yang, Q., and Niu, X. (2018) "Study on a prediction of P2P network loan default based on the machine learning LightGBM and XGboost algorithms according to different high dimensional data cleaning", *Electronic Commerce Research and Applications*, 31: 24-39.
- Malhotra, R. and Malhotra, D. K. (2003) "Evaluating Consumer Loans Using Neural Networks", *Omega* 31: 83-96.
- Marous, J. (2022a) "AI Maturity in Banking Lags All Other Industries" *The Financial Brand*, 10 de Agosto.
- Marous, J. (2020) "Artificial Intelligence in Banking: More Hype Than Reality", *The Financial Brand*, 24 de Febrero.



- Marous, J. (2022b) "Our ChatGPT Interview Shows AI Future in Banking Is Scary-Good" *The Financial Brand*, 17 de diciembre de 2022.
- Marous, J. (2023a) "ChatGPT in Banking: Balancing Its Promise and Its Risks", *The Financial Brand*, 5 de Septiembre.
- Marous, J. (2023b) "The Importance of Continuous AI Innovation in Banking", *The Financial Brand*, 21 de Agosto.
- McKinsey (2016) *Cutting through the noise around financial technology*.
- Mogaji, E., and Nguyen, N. P. (2022) "Managers' understanding of artificial intelligence in relation to marketing financial services: insights from a cross-country study", *International Journal of Bank Marketing*, 40(6): 1272-1298.
- Mori, M. (2021) "AI-powered virtual assistants in the realms of banking and financial services" en *Virtual Assistant*. Edited by Ali Soofastaei. IntechOpen.
- Nalić, J., Martinović, G., and Žagar, D. (2020) "New hybrid data mining model for credit scoring based on feature selection algorithm and ensemble classifiers", *Advanced Engineering Informatics*, 45: 101130.
- Ourdedine, K. (2019) "General perception of Artificial Intelligence and impacts on the financial sector in Finland", *Haaga-Helia University of Applied Science*.
- Pun, J. and Lawryshyn, Y. (2012) "Improving credit card fraud detection using a meta-classification strategy", *International Journal of Computer Applications* 56: 41-46.
- Sartwell, L, White Booker, S and Jeyasekar, P. (2022) "The Hidden Risks of Artificial Intelligence in Bank Marketing", *The Financial Brand*. 30 de mayo.
- Sustersic, M., Mramor, D., and Zupan, J. (2009) "Consumer credit scoring models with limited data", *Expert System with Application*, 36: 4736-4744.
- Tsai, C.F and Wu, J.W. (2008) "Using neural network ensembles for bankruptcy prediction and credit scoring", *Expert Systems with Applications*, 34: 2639-2649.
- Van Gestel, T., Baesens, B., Garcia, I. J. and Van Dijcke, P. (2003) "A support vector machine approach to credit scoring" en *Forum Financier-Revue Bancaire Et Financiera Bank En Financiewezen*. Bruxelles. Larcier: 73-82.
- Wang, Y. Q. (2008) "Building credit scoring systems based on support-based support vector machine" *Proceedings of Fourth international Conference on Natural computation*, (ICNC): 323-327.
- West, D. (2000) "Neural network credit scoring models", *Computers and Operations Research*, 27: 1131-1152.
- Wójcicka, A. (2017) "Neural Networks vs. Discriminant Analysis in the Assessment of Default", *Electronic Economy*, 339-49.
- Yap, B. W., Ong, S. H., and Husain, N. H. M. (2011) "Using data mining to improve assessment of credit worthiness via credit scoring models", *Expert Systems with Applications*, 38(10): 13274-13283.
- Yu, L., Wang, S.Y., and Cao, J. (2009a) "A modified least squares support vector machine classifier with application to credit risk analysis", *International Journal of Information Technology and Decision Making*, 8: 697-710.
- Yu, L., Wang, S., and Lai, K. K. (2008) "Credit risk assessment with a multistage neural network ensemble learning approach", *Expert Systems with Applications*, 34: 1434-1444.



- Yu, L., Wang, S., and Lai, K. K. (2009b) "An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: the case of credit scoring", *European Journal of Operational Research*, 195: 942-959.
- Yurcan, B. (2021) "Three Practical Applications for Artificial Intelligence in Payments", *The Financial Brand*. 20 de abril.
- Zhang, Z. (2017) "Machine Learning Approaches to Predicting Company Bankruptcy", *Journal of Financial Risk Management*, 6: 364-74.
- Zhou, L., Lai, K. K., and Yu, L. (2009) "Credit scoring using support vector machines with direct search for parameters selection", *Soft Computing*, 13: 149-155.
- Zhou, L., Lai, K. K., and Yu, L. (2010) "Least squares support vector machines ensemble models for credit scoring", *Expert Systems with Applications*, 37(1): 127-133.

## 7. Banca privada y banca personal

“El asesoramiento automatizado basado en inteligencia artificial representa una aplicación FinTech que ya está sustituyendo a los asesores minoristas de inversión. En la banca privada los clientes también esperan cada vez más la prestación de servicios a través de diferentes canales digitales, pero con un mayor grado de personalización”.

*Christian Dietzmann, Timon Jaeggi y Rainer Alt (2023)*

En la primera parte del capítulo se comenta cómo el uso progresivo de algoritmos de inteligencia artificial en los diferentes eslabones de la cadena de producción de la banca privada está transformando esta industria, incluso en su modelo más tradicional. La segunda parte se dedica al fenómeno de las plataformas de gestión automatizada. En particular, se

abordan cuestiones relativas a su definición, las ventajas que justifican su crecimiento, las principales limitaciones del modelo y los obstáculos a los que se enfrenta este segmento de mercado. También se discute en qué medida el futuro pasará por un modelo híbrido de atención conjunta de humanos y máquinas en la gestión de patrimonios.

## 7.1. APLICACIONES DE AI EN LA INDUSTRIA DE GESTIÓN PATRIMONIAL

La inteligencia artificial está ya contribuyendo a mejorar la eficiencia operativa y la calidad del servicio ofrecido por las entidades de banca privada y banca personal a sus clientes. Son tan diversas las tareas encomendadas a las técnicas de inteligencia artificial en este terreno que podría afirmarse que, la AI está impulsando una auténtica transformación silenciosa en esta industria.

Por un lado, las técnicas de inteligencia artificial pueden identificar segmentos de clientes basados en comportamientos financieros y preferencias, lo que permite a las entidades de banca privada ofrecer servicios y productos más específicos y personalizados. Esto puede traducirse en un desarrollo empresarial más rentable en este ámbito, al permitir que las instituciones dirijan sus productos y servicios a clases más homogéneas de clientes. La tecnología AI está ayudando significativamente a las entidades desde hace tiempo mediante el uso de algoritmos como

Naïve Bayes, K-NN, Random Forest, SVM, o AdaBoost <sup>124</sup>.

En segundo lugar, el uso de la inteligencia artificial permite una mayor personalización en la gestión de carteras, al analizar el perfil de riesgo y las metas financieras de los clientes para recomendar de forma mucho más eficiente carteras de inversión a medida. Además, gracias a las herramientas de AI los asesores de *private banking* pueden ofrecer recomendaciones más precisas y adaptadas a las necesidades de cada cliente. Los algoritmos de AI permiten analizar grandes cantidades de datos financieros y no financieros en tiempo real para identificar patrones, tendencias y oportunidades de inversión, lo que ayuda a los asesores/gestores a tomar decisiones informadas y a ajustar las estrategias de inversión de manera proactiva.

Además, mediante técnicas de AI las entidades de banca privada pueden mejorar la identificación de riesgos potenciales en las carteras de los clientes y generar alertas cuando es necesario. Es

<sup>124</sup> Véase Smeureanu *et al.* (2013), Fitriani y Febrianto (2021).

posible diseñar algoritmos de inteligencia artificial orientados a analizar el sentimiento del mercado y las noticias financieras en tiempo real para evaluar cómo pueden afectar al precio de los activos contenidos en las carteras de los clientes y proporcionar recomendaciones basadas en ese análisis. Estas alertas se refieren tanto a cambios repentinos en las condiciones del mercado como al anticipo de eventos económicos relevantes. Pero también es posible mediante algoritmos de AI prever cambios que podrían sobrevenir en el perfil de riesgo del propio inversor, y que estarían asociados a situaciones personales o laborales de los clientes.

En otro orden de cosas, las herramientas de AI pueden ayudar a las entidades de la industria de banca privada y personal a automatizar tareas administrativas y rutinarias, como la recopilación y el análisis de datos financieros, la generación de informes y la gestión de documentación. Esto libera tiempo para que los asesores se centren en actividades de mayor valor, como la interacción con los clientes y la toma de decisiones de inversión estratégica.

En general, el uso de la tecnología digital puede mejorar

sustancialmente la calidad de los servicios de banca privada y de banca personal. La digitalización permite una mejor accesibilidad, ya que a través de aplicaciones móviles los inversores tienen acceso a sus cuentas y el seguimiento de su rendimiento en cualquier momento y lugar. Además, mediante herramientas de inteligencia artificial en las que se basan los *chatbots* y asistentes virtuales, las entidades pueden brindar respuestas rápidas a las preguntas de los clientes y ayudar en la gestión de consultas y solicitudes de manera eficiente, mejorando la experiencia del cliente, sin que ello implique una pérdida en la percepción de trato personalizado ni en la calidad del servicio.

En otro orden de cosas, las herramientas de AI pueden ayudar a las instituciones de *private banking* a ajustar el impacto de los frecuentes cambios en las regulaciones financieras y a garantizar el cumplimiento normativo. Pueden diseñarse algoritmos de inteligencia artificial para automatizar procesos de *compliance* y ayudar en la generación de informes requeridos por las autoridades reguladoras.

## 7.2. ROBO-ADVISORS

La implantación de *robo-advisors* (RAs) está ampliamente reconocida como una de las tendencias disruptivas más importantes en la industria de gestión patrimonial en la actualidad. Los RAs son uno de los mejores ejemplos de sistemas autónomos de apoyo a la decisión -*decision support system* (DSS) en su terminología en inglés<sup>125</sup> y prometen un asesoramiento en materia de inversiones eficiente, racional y transparente. La definición del término “robo” aplicada al ámbito de los servicios financieros fue explorada por primera vez en una edición de 2002 de la revista *Financial Planning Magazine*<sup>126</sup>, para referirse a una “automatización sin interacción humana”. En 2008 se convirtió en un modelo de negocio real con el lanzamiento de la primera plataforma digital; y en 2013, las empresas de robo-asesoramiento dieron un gran salto y se convirtieron en una de las principales disrupciones en el mercado financiero al

integrar inteligencia artificial y tecnología financiera (Fintech) en el proceso de gestión de carteras.

Un *robo-advisor* es un software basado en algoritmos que ofrece servicios de planificación financiera automatizada<sup>127</sup>. Las plataformas de RA están diseñadas para ofrecer una cobertura integral del servicio, incluyendo funciones de registro, análisis de necesidades y objetivos de los clientes, capacidad de asesoramiento, de asignación de activos, y de recomendaciones de reequilibrio de cartera ajustadas al perfil de riesgo de los clientes<sup>128</sup>.

De acuerdo con el alcance de los servicios ofrecidos, se han clasificado las plataformas de *robo-advisors* en cuatro modalidades: i) *auto-pilot*, ii) *direct plan-based*; iii) *goal-based advisory* y iv) *full-service*<sup>129</sup>. Los RAs recopilan información sobre los objetivos financieros y la tolerancia al riesgo de un determinado cliente y pueden

<sup>125</sup> Véase Watson (2017).

<sup>126</sup> Por parte de Richard J. Koreto.

<sup>127</sup> Véase Maedche *et al.* (2016).

<sup>128</sup> Véase Bourgeron *et al.* (2019).

<sup>129</sup> Véase Singh y Kaur (2017).

diseñar de forma automática carteras de inversión diversificadas y equilibradas adaptadas a dicho perfil. Además, estos algoritmos pueden ajustar automáticamente la composición recomendada de la cartera de inversión en función de cambios en los objetivos del cliente y de las condiciones del mercado. El flujo de trabajo típico de un *robo-advisor* incluye siete componentes principales:

- Selección del universo de activos.
- Identificación del perfil del inversor (esencialmente en cuanto a su aversión al riesgo, su conocimiento de los mercados y sus necesidades potenciales de liquidez).
- Asignación de activos/ optimización de la cartera.
- Seguimiento de la cartera y los mercados involucrados.
- Propuesta de reequilibrio de cartera.
- Revisión de desempeño.
- Presentación de informes.

El diseño de los RAs puede incluir o no el concurso de un asesor humano.

Los robo-asesores son especialmente adecuados para clientes con presupuestos de gasto menores que el requerido para mantener un servicio de asesoramiento tradicional de gestión patrimonial, que es costoso y suele ser más exclusivo. La gestión automatizada de patrimonios, en forma de robo-asesores, pretende cambiar esta situación y acercar la gestión de patrimonios a un espectro mucho más amplio de inversores y a un precio asequible<sup>130</sup>. Los modelos de negocio en robo-asesoramiento son diversos, algunos se dirigen directamente a los consumidores (*business-to-consumers*), mientras que otros ofrecen su plataforma tecnológica a asesores financieros incumbentes en forma de marca blanca (*business-to-business*) y otros intentan combinar ambos enfoques<sup>131</sup>. Las entidades de banca personal pueden extender los servicios de asesoramiento y gestión patrimonial a una base de clientes con menor patrimonio a un coste satisfactorio mediante *robo-advisors* de producción

<sup>130</sup> Véase Rättyä (2016).

<sup>131</sup> Véase Puhle (2016).

propia o implantados en régimen de marca blanca<sup>132</sup>. Desde la perspectiva de la oferta, durante la última década los *robo-advisors* han pasado de ser un puñado de empresas emergentes a un segmento establecido y de rápido crecimiento de la industria de gestión patrimonial<sup>133</sup>. Desde el lado de la oferta, la gran ventaja de los “neo-advisors” es que su diseño está concebido de forma digital, lo que implica una mejora significativa en términos de eficiencia. Sin embargo, el establecimiento de plataformas de *robo-advisory* implica fuertes inversiones en tecnología. Cuando la introducción de algoritmos de asesoramiento automatizado no se realiza en el ámbito de una entidad incumbente ni mediante acuerdos de B2B, sino mediante el lanzamiento de nuevas plataformas, se ha puesto de manifiesto que una limitación para el avance de la oferta de esta modalidad de servicios es el riesgo asociado a no alcanzar una masa crítica y a que la falta de lealtad de los usuarios se traduzca en abandonos rápidos. En definitiva, la incertidumbre respecto de

la posibilidad de amortizar adecuadamente las inversiones en tecnología y los costes de adquisición de clientes (gastos de marketing) es elevada<sup>134</sup>.

Desde la perspectiva de la situación competitiva, se ha afirmado que las entidades de banca privada, que ofrecen servicios más o menos *premium*, podrían ser algo más inmunes a la generalización de los RAs. Y respecto de los gestores de los clientes en el segmento de banca personal, se argumenta que el asesoramiento automatizado puede ser tanto una ayuda muy valiosa, pero también significar una auténtica amenaza, ya que pueden sustituir los servicios especializados de las entidades incumbentes. La aparición de plataformas digitales independientes diseñadas para ayudar a las personas a invertir su dinero de manera eficiente y diversificada, sin necesidad de la intervención de un asesor financiero humano o con mínima intervención representa una amenaza de disrupción en el estatus de la competencia en esta industria<sup>135</sup>. El valor añadido

<sup>132</sup> Véase Deloitte (2016a y 2016b) y Eule (2017).

<sup>133</sup> Véase Grealish y Kolm (2021).

<sup>134</sup> Véase Puhle (2016) y Luo y Ye (2019).

<sup>135</sup> Véase Phoon y Koh (2017).

de los gestores humanos debe aumentar para permitir una diferenciación en el servicio.

Se ha afirmado también que puede producirse un crecimiento en la industria global de gestión patrimonios, conforme se extienda la penetración de estos servicios (con diferente alcance de valor añadido) desde segmentos de clientes con patrimonio neto ultra-elevado hacia segmentos de clientes con patrimonio elevado, personas ricas (segmento *affluent*), clientes acomodados (segmento *mass-affluent*) y hasta el mercado masivo. En un contexto de crecimiento de este tipo, el uso de gestores automatizados de cartera podría permitir cerrar la brecha que, de otro modo, se produciría entre una demanda creciente y una oferta inicialmente inelástica de estos servicios. Los RAs pueden no sólo convertirse en la solución de inversión preferida para los clientes minoristas o aquellos muy sensible al coste de los servicios, sino que también ganarán importancia en los segmentos individuales ricos y de alto patrimonio. A través de la cooperación con empresas de tecnología financiera o la adquisición de tecnología

innovadora, el robo-asesoramiento podría convertirse en un elemento cada vez más importante para los bancos tradicionales y los administradores de patrimonio<sup>136</sup>.

Continuando el análisis desde el lado de la oferta, algunas de las limitaciones de las plataformas de RAs que se han señalado incluyen las siguientes<sup>137</sup>:

- La dificultad de programar los algoritmos para proporcionar una orientación personal. Los sistemas no pueden pensar más allá de su capacidad programada y no son flexibles para adaptarse a una situación única del mercado para la que no han sido preparados, no muestran empatía ni ofrecen un “consejo amistoso” que muchos inversores buscan en momentos de mercados volátiles.
- Dificultad para integrar en su programación asesoramiento más allá de los asuntos financieros que los inversores suelen necesitar (aspectos relacionados con cuestiones fiscales, del sector inmobiliario, de derecho de familia, normas sobre de movilidad entre jurisdicciones etc.). La respuesta

<sup>136</sup> Véase Uhl y Rohner (2018).

<sup>137</sup> Véase Singh y Kaur (2017).

a estas demandas suele requerir del concurso de asesores humanos experimentados.

- La necesidad de reprogramación ante cambios en las normas o en las políticas de inversión implica costes y tiempo, mientras que el asesoramiento humano suele ser más flexible e inmediato.
- Difícil equilibrio entre sencillez y rigor. No resulta sencillo reconciliar una incorporación de los clientes a un servicio de robo-asesoramiento que sea por una parte sencilla y rápida y, por otra parte, que garantice un riguroso perfil de aversión al riesgo del inversor, siendo el perfilador del riesgo un input clave en la correcta asignación de activos. En el ámbito de los *robo-advisors*, el asesoramiento de *persona a persona* se transforma en un proceso digital de *persona a computadora*<sup>138</sup>, de forma que la elaboración de perfiles de inversores tradicionales realizada durante entrevistas en persona y la interacción bilateral se reemplaza por cuestionarios en línea y procesos de autoinforme. En este sentido, algunas

investigaciones han mostrado que determinados *robo-advisors* suelen utilizar cuestionarios excesivamente simplificados y, para evitar procesos tediosos no asistidos por humanos, el número de preguntas para evaluar el perfil de riesgo de sus usuarios resulta insuficiente. Sin embargo, en la práctica, el perfil de riesgo parece condicionado por un amplio conjunto de circunstancias, que incluyen aspectos relativos a la situación financiera, los objetivos de inversión, la edad del inversor, la experiencia de inversión previa y otra información relevante del cliente. Además, los cuestionarios tradicionales de tolerancia al riesgo no suelen tener en cuenta los hallazgos de la investigación en finanzas conductuales. En algunas investigaciones sobre el tema un resultado que sorprende es que en determinadas plataformas algunas de las preguntas no parecen tener ningún impacto en la categorización del riesgo<sup>139</sup>. Es cierto, sin embargo, que otras herramientas más modernas de asesoramiento automatizado ofrecen perfiladores muy sofisticados, en los que la evaluación es

<sup>138</sup> Véase Jung *et al.* (2018).

<sup>139</sup> Véase Faloon y Scherer (2017), Tertilt y Scholz (2018), Ostern *et al.* (2020).

exhaustiva y, además, no se limita a la elaboración de perfiles de riesgo, sino que también incluyen preferencias éticas específicas (por ejemplo, relativas a objetivos de propósito corporativo, sostenibilidad medioambiental, aspectos sociales y de gobernanza, así como de inclusión, equidad y diversidad)<sup>140</sup>. Finalmente, se ha recomendado que adicionalmente a formular preguntas teóricas sobre el nivel de riesgo con el que se sienten cómodos los clientes, los *robo-advisors* podrían solicitar el historial de transacciones inversión para obtener, mediante inteligencia artificial conclusiones sobre el estilo de inversión de sus clientes<sup>141</sup>. En definitiva, parece claro que la calidad de los perfiladores de riesgo resulta un rasgo decisivo que determina la bondad del servicio.

- Crear sistemas adecuados para el cumplimiento de los procesos de *know-your-customer* (orientados a satisfacer

los requisitos de las normas de *anti-money laundering*) implica costes elevados para el oferente de una nueva plataforma independiente bajo modelo de negocio de B2C. Una estructura de la que sí disponen las entidades incumbentes. El propio avance tecnológico está permitiendo, sin embargo, integrar soluciones digitales que están reduciendo gradualmente dichos costes<sup>142</sup>.

Existe consenso en que las autoridades encargadas de formular las políticas deben lidiar con el impacto de los *robo-advisors* en el sistema financiero en general, así como reevaluar sus prácticas regulatorias y de supervisión<sup>143</sup>. Resulta crítico asegurar que los oferentes de servicios de robo-asesoramiento cumplen las obligaciones fiduciarias legalmente establecidas para los asesores de inversión, incluyendo los criterios de protección del consumidor MIFID II y las salvaguardias de competencia justa (*level playing field*)<sup>144</sup>.

<sup>140</sup> Véase Hakala (2019) y Au *et al.* (2021).

<sup>141</sup> Véase Puhle (2016).

<sup>142</sup> Véase Puhle (2016).

<sup>143</sup> Véase Clarke (2020) y Iannarone (2021).

<sup>144</sup> Véase Ji (2017), Baker y Dellaert (2017), Strzelczyk (2017), Edwards (2018), Carney (2018), Sanz-Bayón y Vega (2018), Rifkin (2018), Seidt *et al.* (2018), Maume (2019a y 2019b) o Mezzanotte

Desde el lado de la demanda, las ventajas que justifican el crecimiento del segmento de robo-asesoramiento son diversas e incluyen, entre otras, una buena relación calidad-coste para el usuario, comodidad y facilidad de uso, atractiva experiencia de usuario, completa accesibilidad, transparencia y ausencia de conflictos de interés. Los *robo-advisors* amplían el acceso a los servicios de gestión patrimonial al hacer que sea más fácil y menos costoso abrir cuentas de inversión y recibir asesoramiento financiero, así como planificar y automatizar decisiones de inversión. Estas plataformas suelen cobrar tarifas más bajas en comparación con los asesores financieros tradicionales, gracias a la automatización casi total de los procesos y por la forma en la que conciben el proceso de inversión<sup>145</sup>. Aunque las plataformas de *robo-advisors* pueden basar la inversión recomendada en una amplia variedad de clases de activos, como acciones, bonos y fondos especializados y temáticos, así como otros instrumentos financieros, en la práctica suelen utilizar predominantemente productos que requieren

una gestión de cartera nula o menos activa, como los fondos cotizados en bolsa (ETF). Los ETF replican índices y, por lo tanto, no requieren una toma de decisiones activa por parte de los administradores de cartera con respecto a la selección y asignación de valores específicos, sino por tipos de activos.

Se ha señalado como posible ventaja significativa de los *robo-advisors* frente a con los gestores de patrimonio que desarrollan productos propios, la ausencia de conflictos de interés y la mitigación de los sesgos en las recomendaciones<sup>146</sup>. También se ha argumentado que como la filosofía de inversión de los *robo-advisors* es habitualmente bastante pasiva, las posiciones en cartera de inversiones se mantienen a medio y largo plazo. De esta forma, los inversores no se ven compelidos a realizar operaciones frecuentes que generan comisiones de *brokerage*. Transacciones que no siempre redundan en mejoras de diversificación o en un aprovechamiento de estrategias de *market-timing*. Todo ello significaría una reducción de costes para los

(2020).

<sup>145</sup> Véase Ringe y Ruof (2018).

<sup>146</sup> Véase Edwards (2018), Rourke (2019).

clientes. Adicionalmente, los servicios de robo-asesoramiento suelen ofrecer a los clientes una accesibilidad en toda la franja horaria a través de ordenadores y dispositivos móviles.

El avance del patrimonio gestionado bajo plataformas de RAs se ha justificado también por el hecho de que, desde la perspectiva de la demanda, se ha incorporado una nueva generación de potenciales clientes con patrimonio suficiente que tienen mejor alfabetización tecnológica y que son más receptivos a las tecnologías digitales. Además, respecto del servicio ofrecido por los asesores financieros tradicionales, perciben una mejora de la calidad del servicio proporcionado por los RAs, y valoran más la reducción del coste que este enfoque representa. Otra explicación del crecimiento del patrimonio gestionado bajo servicios de robo-asesoramiento es la existencia de procesos financieros a gran escala, como la concentración de la riqueza global y la adopción del asesoramiento automatizado en Asia<sup>147</sup>.

De cara al futuro, la transferencia de riqueza hacia las generaciones

de *millennials* y la generación Z, implicará previsiblemente una creciente aceptación del concepto de robo-asesoramiento entre los inversores individuales. Dada la caracterización actual de los usuarios de servicios de asesoramiento automatizado (clientes más jóvenes y actualmente con menor nivel de patrimonio<sup>148</sup>) este proceso de transferencia puede incrementar significativamente el importe de inversión bajo gestión de modelos puros e híbridos de *robo-advisory*.

Es elevada la atención que la investigación está prestando a los factores que determinan la aceptación o el rechazo de la adopción de este modelo de servicio bajo plataformas de robo-asesoramiento en la industria de gestión patrimonial. Los principales resultados son los siguientes:

El factor clave principal que cabría esperar determinase la demanda comparada de servicios de robo-asesoramiento frente al servicio tradicional es la diferencia en las expectativas de desempeño. El desempeño puede aproximarse por la relación entre los ratios de rentabilidad-neta-de-gastos-

<sup>147</sup> Véase Béketov *et al.* (2018).

<sup>148</sup> Véase Fulk *et al.* (2018).

totales/riesgo-asumido de la inversión realizada que ofrecen ambos sistemas<sup>149</sup>. En principio, en contextos en los que la toma de decisiones requiere el análisis de grandes cantidades de datos, las decisiones algorítmicas a menudo superan la toma de decisiones humana para alcanzar resultados de decisión objetivamente óptimos<sup>150</sup>. Sin embargo, en este ámbito particular, los estudios no ofrecen hasta la fecha conclusiones definitivas, lo que se justifica porque el tamaño de la muestra es reducido y la calidad de los datos baja, ya que las plataformas operativas no suelen ofrecer información pública al respecto<sup>151</sup>.

Se ha señalado, como un obstáculo relevante para la adopción del asesoramiento automatizado, la resistencia a perder el control de las decisiones de inversión<sup>152</sup>. La creciente complejidad de los sistemas y la consiguiente percepción de incertidumbre, reforzada por una cierta asimetría de información

respecto de la calidad del sistema, despiertan un deseo de control por parte del usuario<sup>153</sup>. La controlabilidad del sistema puede ser percibida por el usuario como un mitigante del riesgo y, en consecuencia, un factor determinante de la aceptación del sistema<sup>154</sup>. Alguno de los estudios sobre el tema encuentra que los usuarios son más propensos a delegar en robo-asesores que les dan mayor control. Por ejemplo, introduciendo un requisito de aprobación previa a la ejecución, frente a aquellos que realizan la inversión de forma autónoma sin intervención del cliente. En todo caso, no existe una explicación concluyente de por qué determinados inversores superan la aversión que potencialmente implica el uso de servicios basados en algoritmos<sup>155</sup>.

Por otra parte, de forma contraintuitiva respecto del argumento anterior, algunos estudios reflejan que la intención de delegar en los sistemas automatizados por

<sup>149</sup> Véase Venkatesh *et al.* (2012).

<sup>150</sup> Véase Bigman y Gray (2018) y Diab *et al.* (2011).

<sup>151</sup> Véase Kordela (2018).

<sup>152</sup> Véase Rühr *et al.* (2019) y Rühr (2020).

<sup>153</sup> Véase Nussbaumer *et al.* (2012) y Stout *et al.* (2017).

<sup>154</sup> Véase Rhee *et al.* (2005) y Elie-Dit-Cosaque y Straub (2011).

<sup>155</sup> Véase Dietvorst *et al.* (2018).

parte de los clientes aumenta a medida que se incrementa la *percepción de responsabilidad* del sistema. Los usuarios pueden llegar a percibir la delegación como un medio para culpar al “agente” de posibles resultados negativos de las decisiones. En este contexto, la traslación de la responsabilidad aliviaría la percepción de pérdida de *utilidad*. Este resultado sería compatible con algunos hallazgos en ámbito de las finanzas conductuales<sup>156</sup>.

Relacionado con la cuestión de la retención del control está el tema de la transparencia del proceso. La transparencia del algoritmo se definiría como el grado de monitorización que el usuario puede realizar del sistema y el nivel que puede alcanzar de comprensión sobre las diferentes etapas del proceso, incluyendo el uso de la información y la forma en la que razona el algoritmo<sup>157</sup>. Los resultados de algunos estudios sugieren que los inversores están más inclinados a delegar la toma de decisiones financieras cuando reciben información detallada sobre los principios de inversión subyacentes en el algoritmo de

*robo-advisory*. Mejorar el nivel de transparencia, familiarizando al usuario con el sistema, puede crear una *ilusión de control*, más allá de que éste pueda ser ejercido de forma efectiva. Esta *ilusión de control* actuaría con un *signal* en un entorno con asimetría de información<sup>158</sup>. De esta forma, el nivel de transparencia y la aceptabilidad del sistema se relacionarían de forma positiva.

En otro orden de cosas, el nivel de educación financiera, la afinidad por la interacción tecnológica del consumidor y el nivel comparado de “inteligencia percibida” parecen ser factores que condicionan positivamente la demanda relativa de los servicios de gestión patrimonial bajo esquemas de RAs<sup>159</sup>. El auge de los *robo-advisors* requiere que los consumidores comprendan las limitaciones de estos servicios y obtengan una educación financiera adecuada. En este sentido, los estudios indican que la demanda de asesoramiento automatizado se beneficia del hecho de que los RAs pueden brindar educación financiera a posibles inversores

<sup>156</sup> Véase Stout et al. (2017).

<sup>157</sup> Véase Wang y Benbasat (2007) y Garfinkel et al. (2017), Diakopoulos (2020).

<sup>158</sup> Véase Rühr (2020).

<sup>159</sup> Véase Menon (2021) y Aw et al. (2023).

de forma muy eficiente<sup>160</sup>.

Los resultados de la investigación sobre la demanda comparada de los servicios tradicionales de gestión patrimonial y del segmento de gestión automatizada indican que numerosos inversores perciben los *robo-advisors* esencialmente como una alternativa para el análisis cuantitativo, pero creen necesaria la intervención humana para medir las emociones de los inversores. De esta forma, los algoritmos de RAs se contemplarían preferentemente como un servicio complementario y no como un sustituto de los asesores financieros humanos<sup>161</sup>. En ningún caso se espera un escenario de “destrucción del asesoramiento tradicional”<sup>162</sup>. Los modelos híbridos de asesoramiento automatizado con intervención parcial humana podrían imponerse y prevalecer en el futuro, en la medida en que puedan combinar lo mejor de ambos mundos<sup>163</sup>.

Otro de los ámbitos de

investigación respecto de la relación entre los *robo-advisors* y la banca tradicional especializada en gestión patrimonial se ha centrado en el papel del valor de los activos intangibles de los bancos. Esencialmente, en qué medida la confianza en la institución bancaria y en el asesor financiero determina el efecto de estas relaciones<sup>164</sup>. Vinculado con esta cuestión está el tema de la confianza de los clientes a la hora de ofrecer información y ceder fondos bajo gestión a un ente no humano como factor determinante de la adopción de sistemas de *robo-advisory*<sup>165</sup>. Los resultados de la investigación en este ámbito indican que la confianza es un juicio subjetivo multifacético que abarca aspectos asociados con la confiabilidad del sistema, su solidez, la familiaridad, la comprensibilidad y, entre otros factores, la previsibilidad de los resultados<sup>166</sup>. De igual forma, algunos trabajos sugieren que la confianza en el asesor automatizado está relacionada positivamente

<sup>160</sup> Véase Litterscheidt y Streich (2020).

<sup>161</sup> Véase Jung *et al.* (2019), Fisch *et al.* (2019), Bhatia *et al.* (2021) y Coggan (2023).

<sup>162</sup> Véase Raaij (2017).

<sup>163</sup> Véase Salampasis *et al.* (2017), Rourke (2019) y Hakala (2019) y Scholz y Tertilt (2021).

<sup>164</sup> Véase Chou *et al.* (2023).

<sup>165</sup> Véase Reuba, K. (2017), Lee *et al.* (2018) y Belanche *et al.* (2019).

<sup>166</sup> Véase Bartlett y McCarley (2019).

con la confiabilidad del asesor (reputación de la marca), la previsibilidad del entorno (grado de incertidumbre) y la consistencia de las recomendaciones con las señales ambientales (normalmente recibidas a través de los *mass-media* y las redes sociales)<sup>167</sup>.

En otro orden de cosas, se ha destacado como un obstáculo para la adopción de estas herramientas la percepción por parte de los clientes de un bajo nivel de personalización del servicio de robo-asesoramiento, en ausencia de intervención humana. En este contexto, se han propuesto soluciones que podrían ayudar a solucionar este problema. La incorporación de herramientas adicionales de inteligencia artificial como es el caso del procesamiento del lenguaje natural, permite desarrollar servicios de asesoramiento automatizado basados en voz (*voice-based robotic advisor* por su terminología en inglés)<sup>168</sup>. En este sentido, los sistemas de *robo-asesoramiento conversacionales* comienzan a ofrecer resultados progresivamente más potentes<sup>169</sup>.

Adicionalmente, se viene

analizando en qué medida el uso de *robo-advisors* mejora o empeoran la situación respecto del impacto de sesgos emocionales y cognitivos que a menudo se ponen de manifiesto en la toma de decisiones relativas a la gestión de cartera por parte de los inversores. El aprendizaje automático promete mejorar la calidad de las decisiones, debido a la supuesta ausencia de sesgos humanos. Las decisiones adoptadas por gestores humanos podrían, por ejemplo, incorporar una propensión a dar más peso a las alternativas que son consistentes con sus experiencias personales. Se trata de una forma de sesgo conocida como *anclaje*, una de muchas que pueden afectar las decisiones empresariales. El *sesgo de disponibilidad* es otro posible problema. Se trata de un atajo mental (heurístico) mediante el cual las personas hacen suposiciones familiares cuando se enfrentan a decisiones. Los supuestos habrán servido adecuadamente en el pasado, pero podrían resultar inmerecidos en situaciones nuevas. El *sesgo de confirmación* es la tendencia a seleccionar evidencia que respalde creencias preconcebidas,

<sup>167</sup> Véase Sutherland *et al.* (2016) y Hertz y Wiese (2019).

<sup>168</sup> Véase Boschke (2018) y Ostern *et al.* (2020).

<sup>169</sup> Véase Day *et al.* (2018).

mientras que el *sesgo de aversión a las pérdidas* impone un conservadurismo indebido en los procesos de toma de decisiones<sup>170</sup>. Otros aspectos conductuales en este ámbito incluyen exceso de *confianza*, *efecto rebaño*, así como capacidades cognitivas limitadas. Se ha afirmado que los asesores automatizados pueden proporcionar procesos de toma de decisión más simples y menos emocionales, así como ayudar a los inversores individuales a mitigar sus sesgos

de comportamiento en la toma de decisiones de inversión<sup>171</sup>. Sin embargo, algunas investigaciones han concluido también que parte del deseo de retención de control de las decisiones de inversión en procesos con gestores automatizados procede de la existencia de sesgos conductuales como la exagerada apreciación sobre la propia opinión y el exceso de confianza<sup>172</sup>.



## REFERENCIAS

- Abraham, F., Schmukler, S. L. and Tessada, J. (2019) "Robo-advisors: Investing through machines", *World Bank Research and Policy Briefs*, 134881.
- Au, C. D., Klingenberger, L., Svoboda, M., and Frère, E. (2021) "Business model of sustainable robo-advisors: Empirical insights for practical implementation", *Sustainability*, 13(23): 13009.
- Aw, E. C. X., Zha, T., and Chuah, S. H. W. (2023) "My new financial companion! non-linear understanding of Robo-advisory service acceptance", *The Service Industries Journal*, 43(3-4): 185-212.
- Baba, B. and Sevil, G. (2020) "Predicting IPO initial returns using random forest", *Borsa Istanbul Review*, 20 (1): 13-23.
- Baer, T. and Kamalnath, V. (2017) "Controlling machine-learning algorithms and their biases", McKinsey and Co.
- Baker, T., and Dellaert, B. (2017) "Regulating robo-advice across the financial services industry", *Iowa Law Review*, 103: 713-749.
- Bartlett, M. L. and McCarley, J. S. (2019) "Human interaction with automated aids: Implications for robo-advisors", *Financial Planning Review*, 2(3-4).

<sup>170</sup> Véase Baer y Kamalnath (2017).

<sup>171</sup> Véase D'Acunto *et al.* (2019), Shanmuganathan (2020) o Bhatia *et al.* (2020 y 2022).

<sup>172</sup> Véase Yaniv y Kleinberger (2000), Tan *et al.* (2012).

- Béketov, M., Lehmann, K., and Wittke, M. (2018) "Robo Advisors: quantitative methods inside the robots", *Journal of Asset Management*, 19(6): 363-370.
- Belanche, D., Casalo, L., and Flavian, C. (2019) "Artificial intelligence in fintech: understanding robo-advisors adoption among customers", *Industrial Management & Data Systems*, 119: 1411-1430.
- Bhatia, A., Chandani, A., Atiq, R., Mehta, M., and Divekar, R. (2021) "Artificial intelligence in financial services: qualitative research to discover robo-advisory services", *Qualitative Research in Financial Markets*, 13(5): 632-654.
- Bhatia, A., Chandani, A., and Chhateja, J. (2020) "Robo advisory and its potential in addressing the behavioral biases of investors—A qualitative study in Indian context", *Journal of Behavioral and Experimental Finance*, 25: 100281.
- Bhatia, A., Chandani, A., Divekar, R., Mehta, M., and Vijay, N. (2022) "Digital innovation in wealth management landscape: the moderating role of robo advisors in behavioural biases and investment decision-making", *International Journal of Innovation Science*, 14(3/4): 693-712.
- Bigman, Y. E. and Gray, K. (2018) "People are Averse to Machines Making Moral Decisions," *Cognition*, 181: 21-34.
- Boreiko, D. and Massarotti, F. (2020) "How risk profiles of investors affect robo-advised portfolios", *Frontiers in Artificial Intelligence*, 3: 60.
- Boschke, H. (2018) "Using Artificial Intelligence in Wealth Management" en *The WealthTech Book: The FinTech Handbook for Investors, Entrepreneurs and Finance Visionaries*, S. Chishti, T. Puschmann (eds.), Hoboken, NJ: Wiley.
- Bourgeron, T., Lezmi, E., and Roncalli, T. (2019) "Robust asset allocation for robo-advisors", *arXiv:1902.07449*.
- Carney, C. (2018) "Robo-advisers and the suitability requirement: how they fit in the regulatory framework", *Columbia Business Law Review*, 2018(2): 586-616.
- Chou, S. Y., Lin, C. W., Chen, Y. C., and Chiou, J. S. (2023) "The complementary effects of bank intangible value binding in customer robo-advisory adoption", *International Journal of Bank Marketing*, 41(4): 971-988.
- Clarke, D. (2020) "Robo-advisors-market impact and fiduciary duty of care to retail investors" Disponible en SSRN: 3539122.
- Coggan, P. (2023) "Advance of the AI Adviser. Wealth managers are embracing machine learning for research and client relations", *FT Wealth*, 79. Septiembre.
- D'Acunto, F., Prabhala, N., and Rossi, A. G. (2019) "The promises and pitfalls of robo-advising", *The Review of Financial Studies*, 32(5): 1983-2020.
- Day, M. Y., Lin, J. T., and Chen, Y. C. (2018) "Artificial intelligence for conversational robo-advisor" en *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM):1057-1064*. IEEE. Agosto.
- Deloitte (2016a) *Cost-Income Ratios and Robo-Advisory. Why Wealth Managers Need to Engage with Robo-Advisors*.
- Deloitte (2016b) *The expansion of Robo-Advisory in Wealth Management*.
- Diab, D.L., Pui, S.Y., Yankelevich, M., and Highhouse, S. (2011) "Lay Perceptions of Selection Decision Aids

- in US and Non-US Samples" *International Journal of Selection and Assessment*, 19 (2): 209-216.
- Diakopoulos, N. (2020) "Accountability, Transparency, and Algorithms" en *The Oxford Handbook of Ethics of AI*, 17(4): 197.
- Dbouk, W. and Jamali, I. (2018) "Predicting daily oil prices: Linear and non-linear models", *Research in International Business and Finance*, 46: 149-165.
- Dietvorst, B. J., Simmons, J. P., and Massey, C. (2018) "Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them", *Management Science*, 64(3): 1155-1170.
- Edwards, B. P. (2018) "The rise of automated investment advice: can robo-advisers rescue the retail market", *Chicago Kent Law Review*, 93: 97-111.
- Elie-Dit-Cosaque, C.M., and Straub, D.W. (2011) "Opening the Black Box of System Usage: User Adaptation to Disruptive IT", *European Journal of Information Systems*, 20 (5): 589-607.
- Dietzmann, C., Jaeggi, T., and Alt, R. (2023) "Implications of AI-based robo-advisory for private banking investment advisory", *Journal of Electronic Business & Digital Economics*.
- Eule, A. (2017) "Rating the Robo-Advisors", *Barron's*, 29 de Julio.
- Faloon, M., and Scherer, B. (2017) "Individualization of robo-advice", *The Journal of Wealth Management*, 20(1): 30-36.
- Fisch, J. E., Laboure, M., and Turner, J. A. (2019) The Emergence of the Robo-advisor. *The Disruptive Impact of FinTech on Retirement Systems*, 13.
- Fitriani, M. A. and Febrianto, D. C. (2021) "Data mining for potential customer segmentation in the marketing bank dataset". *JUITA: Jurnal Informatika*, 9(1): 25-32.
- Fulk, M., Watkins, K., and Kruger, M. (2018) "Who uses robo-advisory services, and who does not?", *Financial Services Review*, 27(2): 173-188.
- Garfinkel, S., Matthews, J., Shapiro, S. S., and Smith, J. M. (2017) "Toward algorithmic transparency and accountability", *Communications of the ACM*, 60 (9): 5-5.
- Godarzi, A. A., Amiri, R. M., Talaei, A., and Jamasb, T. (2014) "Predicting oil price movements: A dynamic Artificial Neural Network approach", *Energy Policy*, 68: 371-382.
- Grealish, A. and Kolm, P. N. (2021) "Robo-Advisors Today and Tomorrow: Investment Advice Is Just an App Away", *The Journal of Wealth Management*, 26 (2).
- Hakala, K. (2019) *Robo-advisors as a form of artificial intelligence in private customers' investment advisory services*. Aalto University Business School.
- Hertz, N. and Wiese, E. (2019) "Good advice is beyond all price, but what if it comes from a machine?", *Journal of Experimental Psychology: Applied*, 25(3): 386.
- Hu, J. W. S., Hu, Y. C., and Lin, R. R. W. (2012) "Applying neural networks to prices prediction of crude oil futures", *Mathematical Problems in Engineering*.
- Iannarone, N. G. (2021) "Rethinking automated investment adviser disclosure" en *The Routledge Handbook of FinTech*: 347-357. Routledge.
- Ji, M. (2017) "Are robots good fiduciaries: regulating robo-advisors under the investment advisers act of 1940", *Columbia Law Review*, 117: 1543-1584.

- Jung, D., Dorner, V., Glaser, F. and Morana, S. (2018) "Robo-advisory: digitalization and automation of financial advisory", *Business & Information Systems Engineering*, 60: 81-86.
- Jung, D., Glaser, F., and Köpplin, W. (2019) "Robo-advisory: opportunities and risks for the future of financial advisory", *Advances in Consulting Research: Recent Findings and Practical Cases*: 405-427.
- Kordela, D. (2018) "Robo-advisors in asset management-the experience from Germany", *European Journal of Service Management*, 28(4/1): 151-157.
- Lee, S., Choi, J., Ngo-Ye, T. L., and Cummings, M. (2018) "Modeling trust in the adoption decision process of robo-advisors: an agent-based simulation approach", *WAIS 2018 Proceedings*, 24.
- Litterscheidt, R., and Streich, D. J. (2020) "Financial education and digital asset management: What's in the black box?", *Journal of Behavioural and Experimental Economics*, 87:101573.
- Luo, Y. and Ye, Q. (2019) "Understanding consumers' loyalty to an online outshopping platform: the role of social capital and perceived value", *Sustainability* 11:5371.
- Maedche, A., Morana, S., Schacht, S., Werth, D., and Krumeich, J. (2016) "Advanced user assistance systems", *Business & Information Systems Engineering*, 58: 367-370.
- Maume, P. (2019a) "Regulating robo-advisory", *Texas International Law Journal*, 55: 49-87.
- Maume, P. (2019b) "Reducing legal uncertainty and regulatory arbitrage for robo-advice", *European Company and Financial Law Review*, 16(5): 622-651.
- Menon, M. M. M. (2021) "An empirical analysis of the factors influencing the adoption of robo-advisory in wealth management", *Information Technology in Industry*, 9(3): 560-566.
- Mezzanotte, F. E. (2020) "An examination into the investor protection properties of robo-advisory services in Switzerland", *Capital Markets Law Journal*, 15(4): 489-508.
- Munshi, M., Patel, M., Alqahtani, F., Tolba, A., Gupta, R., Jadav, N. K., ... and Dragomir, A. (2022) "Artificial Intelligence and Exploratory-Data-Analysis-Based Initial Public Offering Gain Prediction for Public Investors", *Sustainability*, 14(20): 13406.
- Nussbaumer, P., Matter, I., and Schwabe, G. (2012) "Enforced" vs. "Casual" Transparency - Findings from IT-Supported Financial Advisory Encounters," *ACM Transactions on Management Information Systems* (3:2): 1-19.
- Ostern, N. K., Schöler, J., and Moormann, J. (2020) "Toward Voice-Enabled Robotic Advisory for Personalized Wealth Management" en AMCIS.
- Phoon, K., and Koh, F. (2017) "Robo-advisors and wealth management", *The Journal of Alternative Investments*, 20(3): 79-94.
- Puhle, M. (2016) "The FinTech Revolution: A Closer Look at Robo-Advisors", *Economy and Finance* 4 (3): 256-271.
- Raaij, W. F. (2017) "Explaining Customer Experience of Digital Financial Advice", *Economics World*, 5 (1): 69-84.
- Rättyä, J. (2016) "March of the Robo-advisors: The potential for global expansion of digital asset management platforms", *Metropolia, University of Applied Science, Thesis*. Helsinki.
- Reuba, K. (2017) "Robo-Advisors: Early Disruptors in Private Wealth Management", *Allianz Investment Themes and Strategy*.

- Rhee, H.-S., Ryu, Y., and Kim, C. T. (2005) "I Am Fine but You Are Not: Optimistic Bias and Illusion of Control on Information Security", *Twenty-Sixth International Conference on Information Systems*, Las Vegas, NV, USA.
- Rifkin, J. G. (2018) "Robo-Advisers Jumping on the Bandwagon: Yet Another Cry for a Uniform Standard". *North Carolina Law Review*, 97(3), 673-709.
- Ringe, W. G. and Ruof, C. (2018) "A regulatory sandbox for robo-advice", *European Banking Institute Working Paper Series*, 26.
- Rourke, C. (2019) *Leveraging the Competition: How Wealth Managers Can Use Robo-Advisors to their Advantage*. Doctoral dissertation, University of Iowa.
- Rühr, A. (2020) "Robo-advisor configuration: An investigation of user preferences and the performance-control dilemma", *Ludwig-Maximilians-Universität Research Paper*. Munich. Germany.
- Rühr, A., Berger, B., and Hess, T. (2019) "Can I control my robo-advisor? Trade-offs in automation and user control in (digital) investment management".
- Salampasis, D., Mention, A. L., and Kaiser, A. O. (2017) "Wealth management in times of robo: Towards hybrid human-machine interactions", Disponible en SSRN: 3111996.
- Sanz-Bayón, P. and Vega, L. G. (2018) "Automated Investment Advice: Legal Challenges and Regulatory Questions", *Banking & Financial Services Policy Report*, 37(3): 1-11.
- Scholz, P. and Tertilt, M. (2021) *Robo-advisory: The rise of the investment machines. Robo-Advisory, Investing in the Digital Age*: 3-19. Palgrave. MacMillan.
- Seidt, A. L., Zaharis, N., and Jarrett, C. (2018) "Paying Attention to that Man behind the Curtain: State Securities Regulators' Early Conversations with Robo-Advisors", *University of Toledo Law Review*, 50, 501-523.
- Shanmuganathan, M. (2020) "Behavioural finance in an era of artificial intelligence: Longitudinal case study of robo-advisors in investment decisions", *Journal of Behavioral and Experimental Finance*, 27: 100297.
- Singh, I. and Kaur, N. (2017) "Wealth management through robo advisory", *International Journal of Research-Granthaalayah*, 5(6): 33-43.
- Smeureanu, I., Ruxanda, G., and Badea, L. M. (2013) "Customer segmentation in private banking sector using machine learning techniques", *Journal of Business Economics and Management*, 14(5): 923-939.
- Stout, N., Dennis, A., and Wells, T. (2017) "The Buck Stops There: The Impact of Perceived Accountability and Control on the Intention to Delegate to Software Agents", *AIS Transactions on Human-Computer Interaction* 6 (1):1-15.
- Strzelczyk, B. E. (2017) "Rise of the machines: the legal implications for investor protection with the rise of Robo-Advisors", *Depaul Business & Commercial Law Journal*, 16 (1): 54-85.
- Sutherland, S. C., Hartevelde, C., and Young, M. E. (2016) "Effects of the advisor and environment on requesting and complying with automated advice", *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 6 (4): 1-36.
- Tan, W.K., Tan, C.H., and Teo, H.H. (2012) "Consumer-Based Decision Aid that Explains Which to Buy: Decision Confirmation or Overconfidence Bias?", *Decision Support Systems*, 53 (1): 127- 141.



- Tertilt, M. and Scholz, P. (2018) "To advise, or not to advise—how robo-advisors evaluate the risk preferences of private investors", *The Journal of Wealth Management*, 21(2): 70-84.
- Uhl, M. W. and Rohner, P. (2018) "Robo-advisors versus traditional investment advisors: An unequal game", *The Journal of Wealth Management*, 21(1): 44-50.
- Venkatesh, V., Thong, J.Y.L., and Xu, X. (2012) "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology", *MIS Quarterly*, 36 (1): 157-178.
- Wang, W. and Benbasat, I. (2007) "Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs", *Journal of Management Information Systems* 23 (4): 217-246.
- Watson, H. J. (2017) "Preparing for the cognitive generation of decision support", *MIS Quarterly Executive*, 16(3): 153-169.
- Yaniv, I., and Kleinberger, E. (2000) "Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation", *Organizational Behavior and Human Decision Processes*, 83 (2): 260- 281.



## 8. Banca de inversión y corporativa

En el dinámico mercado de capitales actual, la inteligencia artificial demuestra una capacidad de aprendizaje automático, rapidez y precisión para manejar datos complejos. Gradualmente ha adquirido el potencial de convertirse en un “cerebro financiero”.

*Xiao-lin Zheng, Meng-ying Zhu, Qi-bing Li,  
Chao-chao Chen, and Yan-chao Tan (2019).*

Buena parte de las actividades de la banca de inversión y corporativa se basan en la capacidad de los gestores a la hora de convertir conocimientos y análisis en soluciones financieras orientadas a actividades como trading a corto plazo, inversiones a medio y largo plazo mediante valores y préstamos; determinación del precios de los bonos y las acciones en los mercados primarios, cálculo del precio de instrumentos derivados; y asesoramiento a clientes en relación con la

asignación a lo largo del tiempo de capital y deuda en proyectos alternativos, así como elaboración de recomendaciones en relación con la coberturas de riesgos, etc.

En consecuencia, los bancos inversión desarrollan modelos de *trading* especulativo orientados a predecir el precio y la volatilidad de diferentes clases de activos (acciones, bonos, instrumentos derivados, instrumentos de tipo de cambio y *commodities*). También se incluyen en este ámbito actividades de gestión

de cartera para el propio banco y para vehículos institucionales (como fondos de inversión, de pensiones, de capital privado, fondos de infraestructuras y *hedge funds*). Esta gestión se materializa en la selección óptima de activos, la búsqueda de *alpha*, y el diseño de estrategias eficientes de cobertura. Del mismo modo, una actividad que desarrolla la banca de inversión es la simulación financiera, la planificación y el asesoramiento a clientes en transacciones de fusiones y adquisiciones y en el recurso a los mercados de capitales (acciones, instrumentos *equity-related* y deuda). Finalmente, en la banca corporativa resulta necesaria la evaluación crediticia (que implica la calificación de riesgo de crédito, la estimación de probabilidades de quiebra y la predicción de tasas de recuperación), ya que también desarrollan actividad de financiación de proyectos (*project finance*), de adquisiciones y de activos.

En el contexto de la banca de inversión y corporativa, los algoritmos de inteligencia artificial orientados al análisis financiero están mostrando algunas ventajas respecto de las herramientas estadísticas tradicionales, ya que estas últimas exigen requisitos más estrictos

para convertir la información en datos de entrada, lo que origina pérdida de información ante intervalos de datos inadecuados y conversiones diferentes. Por el contrario, determinadas técnicas de inteligencia artificial permiten introducir datos numéricos directamente como entrada para su procesamiento. Además, en las técnicas estadísticas como la regresión o el análisis discriminante se requieren supuestos de distribución relativos a los datos de entrada, lo que no es necesario en determinados algoritmos de inteligencia artificial. Por ejemplo, las redes neuronales permiten agregar directamente nuevos datos a una red entrenada para actualizar el resultado del entrenamiento anterior. Así, en aplicaciones financieras dinámicas, los algoritmos de inteligencia artificial pueden acomodar información nueva sin reprocesar información antigua. Adicionalmente, algunos algoritmos de AI son estimadores sin modelo, lo que permite capturar el efecto de interacción entre variables sin que los usuarios tenga que especificar modelos econométricos. Finalmente, esta tecnología permite abordar análisis progresivamente más complejos, aumentando las capas ocultas de la red. Es habitual combinar varias herramientas

de inteligencia artificial para un determinado análisis<sup>173</sup>.

## 8.1. ACTIVIDAD DE *TRADING*

Los *hedge fund* y las mesas de *proprietary trading* de los bancos y fondos de inversión han venido utilizando durante las últimas décadas técnicas de negociación cuantitativa para lograr una ejecución más rápida y a mejores precios, conseguir una mayor precisión en las predicciones y evitar sesgos por causas psicológicas o emocionales. La operativa automatizada de alta frecuencia –*high frequency trading* (HFT) en su terminología en inglés– constituye desde hace tiempo el núcleo cuantitativamente más relevante de los mercados financieros. En este ámbito, la inteligencia artificial está ayudando no sólo a determinar la estrategia de negociación, pero también a reducir los errores de ejecución mediante el incremento en la capacidad de verificar automáticamente y de forma simultánea múltiples condiciones de mercado. Además del análisis técnico, en la actividad de *trading*

se habían venido aplicado metodologías estadísticas de series temporales basadas en datos históricos, para pronosticar los precios y rendimientos de las acciones incluyendo modelos de heterocedasticidad condicional autorregresiva (ARCH), de media móvil autorregresiva (ARMA), de media móvil autorregresiva integrada (ARIMA), filtrado de Kalman y suavizado exponencial. Pero el mercado de valores es caótico, complejo, no lineal, dinámico y volátil, de forma que la predicción es una de las tareas más desafiantes en el pronóstico de series temporales.

Las técnicas de inteligencia artificial están ayudando en la actividad de *trading* permitiendo la incorporación de información cada vez más diversa procedente de las redes sociales y otras fuentes. Al mismo tiempo, a diferencia de los métodos tradicionales de series de tiempo, las herramientas de inteligencia

<sup>173</sup> Véase Bahrammirzaee (2010).

artificial permiten manejar datos no lineales, caóticos, ruidosos y complejos del mercado de valores, lo que lleva a predicciones más precisas. La adopción de métodos de negociación asistidos por computadora basados en aprendizaje automático viene ganando popularidad en los últimos años debido a su capacidad para procesar de manera eficiente información pasada. En concreto, el uso de redes de aprendizaje profundo para el análisis y predicción de valores de mercado ha mejorado la extracción de características de un gran conjunto de datos sin procesar y sin depender del conocimiento previo de los predictores. Estos rasgos hacen que el aprendizaje profundo sea atractivo en la predicción utilizada en la negociación de valores de alta frecuencia en los mercados de valores. Tanto en el *trading* direccional como en la negociación de valor relativo y en la contratación de derivados resulta necesario realizar predicciones sobre el precio futuro de los activos y su volatilidad. Las aplicaciones de algoritmos de inteligencia artificial por parte de la industria y en el ámbito académico se han

multiplicado en los últimos años.

El *trading* de instrumentos de tipos de cambio (*forex*) es el ámbito en el que antes comenzaron a aplicarse algoritmos de inteligencia artificial. El principal problema al modelizar series temporales de precios de activos financieros es que su dinámica y propiedades cambian con el tiempo. Esto es particularmente cierto con los tipos de cambio, el mayor mercado financiero del mundo, debido a la cantidad de “flujo de información” inconstante y a su imprevisibilidad lineal. Los estudios empíricos muestran que la distribución de los rendimientos diarios es aproximadamente simétrica y leptocúrtica<sup>174</sup>. Esta circunstancia se ha justificado por el hecho de que las muestras se distribuyen como una distribución cuya media y varianza cambian con el tiempo, y también porque los rendimientos observados no provienen de una distribución normal sino de una combinación de distribuciones normales, por lo que resulta complicado que un solo modelo capture la dinámica de toda la serie temporal. Además, el rendimiento promedio de los datos sobre tipos de cambio de

<sup>174</sup> Es decir, que las colas de la distribución son más pesadas, lo que da como resultado una mayor probabilidad de eventos extremos positivos o negativos.

alta frecuencia puede resultar insignificante en comparación con su volatilidad. En este contexto, diversos estudios han demostrado que las redes neuronales adaptativas superan significativamente a los modelos lineales como ARMA y el modelo nativo de paseo aleatorio. En todo caso, las técnicas de inteligencia artificial aplicadas en el *trading* de instrumentos Forex son diversas incluyendo no sólo RNN y algoritmos genéticos, pero también SVM, regresiones de vectores de soporte (SVR), percepciones multicapa, funciones de base radial, redes neuronales dinámicas, sistemas neuro-difusos, etc.<sup>175</sup>.

Cuando se trata de la predicción del precio de las acciones, los inversores están utilizando un amplio conjunto de técnicas de inteligencia artificial incluyendo redes neuronales artificiales (ANN)

y recurrentes (RNN), *machine learning*, Support Vector Machine (SVM), bosques aleatorios, Modelo bayesiano Naive, modelos K-vecino más cercano (KNN), modelos de árboles de decisión (DT), Random Forest, *adaptive boosting* (Adaboost), *gradient boosting* y *eXtreme gradient boosting* (XGBoost), para predecir los rendimientos esperados de las acciones<sup>176</sup>.

Una de las novedades que permite el uso de algoritmos de inteligencia artificial es la incorporación de información no financiera procedente de las redes sociales en el análisis de su impacto en la variación del precio de los valores bursátiles. En este sentido, se están empleando algoritmos de caracterización de texto, técnicas de aprendizaje profundo como Long Short Term Memory (LSTM), Logistic Regression (LR) y Multilayer Perceptron (MLP). Esta línea de

<sup>175</sup> Véase Refenes *et al.* (1993), Tenti (1996), Hann y Steurer (1996), Franses y Van Griensven (1998), Franses y Van Homelen (1998), Zhang y Hu (1998), El Shazly y El Shazly (1999), Hu *et al.* (1999), Yao y Tan (2000), Leung *et al.* (2000), Dempster *et al.* (2001), Zhang y Berardi (2001), Andreou *et al.* (2002), Nag y Mitra (2002), Kodogiannis y Lolis (2002), Kamruzzaman y Sarker (2004), Chen *et al.* (2004), Dempster y Leemans (2006), Yu *et al.* (2007), Eng *et al.* (2008), Ni y Yin (2009), Pacelli, *et al.* (2011), Galeschuk (2016), Chaudhuri y Ghosh (2016), Huang *et al.* (2004b) o Rundo (2019).

<sup>176</sup> Véase entre otros muchos, Jangmin *et al.* (2006), Soni (2011), Shen *et al.* (2012), Hegazy *et al.* (2014), Dash y Dash (2016), Chen y Hao (2017), Chong *et al.* (2017), Singh y Srivastava (2017), Gudelek, *et al.* (2017), Hiransha *et al.* (2018), Dang (2019), Wu *et al.* (2020), Shen y Shafiq (2020), Li *et al.* (2020b), Yang *et al.* (2020), Liu *et al.* (2020), Nabipour *et al.* (2020), Carta *et al.* (2021), Lin *et al.* (2021), Mascio *et al.* (2021), Ahmed *et al.* (2022), Kumbure *et al.* (2022) o Leippold *et al.* (2022).

investigación está recibiendo una enorme atención por parte de la industria y la academia<sup>177</sup>.

En los ámbitos de la renta fija y los tipos de interés existe una notable dificultad en la predicción de los precios futuros de los instrumentos. Los precios de los swaps de tipos de interés, de los bonos gubernamentales y los bonos corporativos son un reflejo de políticas e interacciones de mercado extremadamente complejas y se ven afectados por muy diversas variables. En este contexto, algoritmos de inteligencia artificial, y en particular los métodos de *machine learning*, están siendo muy utilizados para intentar mejorar las estrategias de trading en la curva de tipos y en el análisis de los diferenciales de crédito de los bonos<sup>178</sup>. Debido a que los métodos de aprendizaje automático pueden modelizar relaciones entre variables explicativas y dependientes basadas en estructuras complejas,

no lineales y/o no paramétricas, no es sorprendente que este enfoque esté mostrando resultados de pronóstico prometedores y superen significativamente a los métodos de regresión tradicionales<sup>179</sup>.

Relacionado con el *trading* de bonos, la negociación de derivados de crédito, y específicamente de *credit default swaps* (CDS), está también siendo objeto de revisión a la luz de las herramientas de inteligencia artificial. Así, algoritmos de ANN, *Deep learning*, modelos de aprendizaje automático LSTM, unidad recurrente cerrada (GRU), red de perceptrones multicapa (MLP), redes generativas adversariales (GAN) y SVMs están siendo utilizados en estimación de las primas de estos instrumentos<sup>180</sup>.

Por otra parte, hasta hace algún tiempo, en la valoración de opciones se habían venido

<sup>177</sup> Véase Mittal y Goel (2012), Nguyen y Shirai (2015), Nguyen *et al.* (2015), Kordonis *et al.* (2016), Chiong *et al.* (2018), Yoo *et al.* (2018), Mohan *et al.* (2019), Guo y Li (2019), Mehtab y Sen (2019), Gupta y Chen (2020), Maqsood *et al.* (2020), Li *et al.* (2020a), Huang y Liu (2020), Mehta *et al.* (2021), Gite *et al.* (2021), Jing *et al.* (2021), Koukaras *et al.* (2022), Gupta *et al.* (2022b), Swathi *et al.* (2022).

<sup>178</sup> Véase Matsui *et al.* (2009), Ganguli y Dunnmon (2017), Pendharkar y Cusatis (2018), Wright *et al.* (2018), Zhang *et al.* (2020), Jabeur *et al.* (2020), Kim (2021), Kim *et al.* (2021), Fedenia *et al.* (2021), Nazemi *et al.* (2022).

<sup>179</sup> Véase Götze *et al.* (2020 y 2023).

<sup>180</sup> Véase Shaban *et al.* (2010), Feser y Broby (2020), Beytollahi y Zeinali (2020), Lin *et al.* (2022), Koy y Çolak (2023), Mao *et al.* (2023).

utilizando tradicionalmente métodos como Black-Scholes, modelos binomiales, el método de diferencias finitas (FD) y el método de simulación de Monte Carlo (MC). Sin embargo, recientemente se están aplicando también algoritmos de *machine learning* (en concreto métodos de árbol de decisión, Random Forest, XGBoost y LightGMB, regresiones de vectores de soporte (SVR) y particularmente redes neuronales (tanto redes híbridas que incorporan información de los modelos paramétricos como redes neuronales modulares (MNN) no paramétricas, para determinar el precio de las primas de las opciones y valorar las estrategias de cobertura.

Un componente básico de la creación de mercado de opciones y de la gestión del libro de volatilidad en una entidad es la estimación de esta última variable. Los algoritmos de inteligencia artificial están diseñados para captar patrones no lineales de series temporales. Además, bajo este tipo de herramientas de inteligencia artificial, no se

requieren suposiciones respecto de la distribución de los datos y, adicionalmente, el modelo aprende las relaciones entre los datos de entrada financieros y el precio de la opción a partir de los datos históricos. Las comparaciones revelan que el error cuadrático medio de la red neuronal puede ser menor que el que se deriva del uso del modelo Black-Scholes y de los métodos binomiales. Además, los resultados para estimaciones fuera de la muestra parecen mejores.<sup>181</sup>

La actividad de *trading* de instrumentos de futuros sobre materias primas está también experimentando cambios significativos en la forma en la que se determinan las estrategias de negociación, con la creciente aplicación de herramientas de inteligencia artificial. Los algoritmos utilizados son variados, incluyendo particularmente redes neuronales de varias modalidades, como *convolutional neural networks* y *time-delayed neural network*, y también algoritmos de *deep learning*, LSTM, algoritmos de *neuro-fuzzy*, SVR,

<sup>181</sup> Véase Malliaris y Salchenberger (1993a y 1993b), Hutchinson et al. (1994), Niranjan et al. (1996), Lajbcygier y Connor (1997), Gonzalez-Miranda y Burgess (1997), Anders et al. (1998), Yao et al. (2000), Ghaziri et al. (2000), Gençay y Qi (2001), Refenes y Holt (2001), Amilon (2003), Lajbcygier (2004), Hamid e Iqbal (2004), Andreou et al. (2006), Blynski y Faseruk (2006), Gradojevic et al. (2009), Liang et al. (2009), Lin y Yeh (2009), Mitra (2012), Gaspar et al. (2020), Ivaşcu (2021), Qian et al. (2022), Cohen et al. (2022) o Gierjatowicz et al. (2022).

análisis DTL, etc. Los mercados en los que se está aplicando AI incluyen aquellos en los que se negocian instrumentos sobre

petróleo, gas, electricidad y otros instrumentos de energía<sup>182</sup>, oro<sup>183</sup>, y *commodities* agrícolas y otros productos no agrícolas<sup>184</sup>.

## 8.2. GESTIÓN DE CARTERAS

La teoría moderna de carteras asume que el inversor enfoca su decisión esencialmente en obtener resultados en términos de media y varianza de su riqueza final<sup>185</sup>. Para que sea compatible con la teoría normativa de la maximización de la utilidad esperada, hay que suponer que los rendimientos de los activos se distribuyen normalmente de forma conjunta o que las preferencias del inversor pueden modelizarse mediante una función de utilidad cuadrática. Los datos de entrada para el problema de optimización de la cartera son los rendimientos esperados de los activos y su matriz de covarianza. Además, se necesita una medida de la tolerancia al riesgo del

inversor, lo que implica determinar una volatilidad aceptable de la riqueza terminal. Es sabido que el marco de la teoría moderna de carteras presenta algunos problemas prácticos relacionados con el hecho de exhibir una sensibilidad elevada a los inputs de entrada utilizados en los modelos, maximización del error de estimación y excesiva concentración de la cartera, por lo que se han introducido ciertas mejoras multidimensionales. Las herramientas de inteligencia artificial están incorporando también, y con mayor facilidad, estas modificaciones en sus modelos de gestión<sup>186</sup>. La optimización de la cartera implica

<sup>182</sup> Véase Hu *et al.* (2012), Godarzi *et al.* (2014), Hamdi y Aloui (2015), Dbouk y Jamali (2018), Luo *et al.* (2019), Butler *et al.* (2021), Assaad y Fayek (2021), Wang y Wang (2021), Xu y Zhang (2022a, 2022b, y 2023), Jiang *et al.* (2022), Bouteska *et al.* (2023) o Pei *et al.* (2023).

<sup>183</sup> Véase Hussein *et al.* (2011), Yazdani-Chamzini *et al.* (2012), KangaraniFarahani y Mehralian (2013), Malliaris y Malliaris (2013 y 2015), Li (2014), Lin (2015), Dubey (2016), Hafezi y Akhavan (2018), Bin-Khamis y Yee (2018), Zhu y Zhang (2018), Manjula y Karthikeyan (2019).

<sup>184</sup> Véase Jha y Sinha (2013), Manogna y Mishra (2021), Rasheed *et al.* (2021), Chen *et al.* (2021) o Jaiswal *et al.* (2022).

<sup>185</sup> Véanse modelos de Markowitz y de Black-Litterman.

<sup>186</sup> Véase Béketov *et al.* (2018).

que se cumplen las restricciones del inversor respecto de la asignación de activos y que se puede controlar el riesgo dentro de los parámetros establecidos, permitiendo alcanzar ratios de rentabilidad-riesgo aceptables. Con las herramientas financieras disponibles, el número de combinaciones posibles de carteras que pueden sintetizarse es astronómica, pero buscar asignaciones que coincidan con los objetivos y limitaciones de un gestor es un proceso complejo y que requiere tiempo<sup>187</sup>.

Por otra parte, en los modelos econométricos tradicionales se especifica un modelo teórico y se asume que los datos generados proceden de un proceso estocástico. Además, los modelos de riesgo convencionales asumen que los mercados se comportan de acuerdo con relaciones lineales. En el enfoque tradicional de gestión del riesgo se desarrollan modelos econométricos, se prueba la multicolinealidad y se ajustan los factores explicativos para minimizar las correlaciones entre ellos y eliminar la autocorrelación. Posteriormente se asignan

ponderaciones a los factores que contribuyen significativamente a explicar la variable estudiada y se utilizan éstas como si fueran estables. Este proceso se traduce en un riesgo de modelo<sup>188</sup>. Es sabido que la estimación del Valor en Riesgo (VaR) y del Valor en Riesgo Condicional (CVaR), como medidas ampliamente utilizadas del riesgo de pérdida en las inversiones, están seriamente expuestas a riesgos de modelo y, por lo tanto, su aplicación resulta de utilidad sólo en condiciones normales de mercado. La inteligencia artificial está siendo una ayuda importante para los analistas de mercado en la detección de relaciones no lineales entre variables con impacto en los mercados financieros. En un modelo de aprendizaje automático no lineal la ponderación cambia con el tiempo, dependiendo de una multitud de factores<sup>189</sup>. La AI se está convirtiendo en una ayuda fundamental para superar las deficiencias de los modelos econométricos clásicos en la detección de valores atípicos, la extracción de características y la clasificación y regresión de datos complejos. La aplicación de algoritmos de inteligencia artificial

<sup>187</sup> Véase Bahrammirzaee (2010).

<sup>188</sup> Véase Kilburn (2019).

<sup>189</sup> Véase Kilburn (2023).

se está generalizando para predecir los mejores resultados futuros para indicadores del binomio rentabilidad-riesgo (esto es, para establecer fronteras óptimas de eficiencia). En concreto, se están utilizando técnicas como redes residuales (ResNet), memoria a largo plazo (LSTM), unidad recurrente cerrada (GRU), autoatención (SA), atención aditiva (AA), redes neuronales profundas (DNN), entre otras<sup>190</sup>. En el análisis agregado de riesgos Value at Risk se están aplicando también algoritmos de inteligencia artificial del tipo SOM, Gaussian Mixtures, Cluster Analysis, NN o GELM<sup>191</sup>. Del mismo modo están siendo aplicadas también diferentes modalidades de algoritmos de inteligencia artificial en la estimación de la volatilidad<sup>192</sup>.

En definitiva, la gestión de una cartera de activos implica la aceptación de riesgo de mercado relacionados con la evolución adversa de los precios

tanto en estrategias de apuestas direccionales como de valor relativo (*relative value*) y de riesgo de iliquidez y, en este ámbito, se está dedicando un notable esfuerzo para determinar si los modelos basados en tecnología de inteligencia artificial, como el aprendizaje automático, generan resultados mejores que aquellos obtenidos mediante métodos estadísticos tradicionales. En las comparaciones resultan ganadores unos y otros según los estudios. En aquellos casos en los que los algoritmos de aprendizaje automático parecen superiores las ventajas se basan en su capacidad para identificar relaciones no lineales entre los factores explicativos y las variables objetivo, su adaptación a distribuciones-no-normales de los factores explicativos, la búsqueda de efectos de interacción entre los factores explicativos, y su capacidad para lograr previsiones más sólidas mediante enfoques de conjunto<sup>193</sup>. Además, se ha demostrado que,

<sup>190</sup> Véase Cao *et al.* (2020), Gašperov *et al.* (2020), Wang y Zhou (2020), Clark *et al.* (2020), Tsang y Wong (2020), Warin (2021), Zhang *et al.* (2021c), Pretorius y Van Zyl (2022), Huang *et al.* (2022), Van Staden (2023) o Papanicolaou *et al.* (2023).

<sup>191</sup> Véase, por ejemplo, Kanevski y Timonin (2010), Kristjanpoller *et al.* (2014), Monfared y Enke, D. (2014), Zhang *et al.* (2017) o Mahdavi-Damghani y Roberts (2017).

<sup>192</sup> Véase Chen *et al.* (2008), Bildirici y Ersin (2009), Chen *et al.* (2010), Matías *et al.* (2010), Arnerić *et al.* (2014), Kristjanpoller y Minutolo (2015), Lu *et al.* (2016) Kristjanpoller y Hernández (2017) y Hajizadeh *et al.* (2019).

<sup>193</sup> Véase Breeden (2020).

si bien es cierto que un equipo de analistas con talento y tiempo suficiente puede conseguir el mismo resultado, los algoritmos de inteligencia artificial parecen conseguir estos resultados en periodos de tiempo más cortos.

La asignación de activos implica inherentemente el riesgo de que el modelo que se utiliza para fijar los criterios de composición óptima de cartera sea incompleto u obsoleto a la hora de establecer relaciones estables entre precios y variables en cada momento. El análisis de las implicaciones de utilizar modelos que han dejado de ser válidos se conoce como gestión del riesgo de modelo. Las técnicas de AI parecen ser especialmente adecuadas para profundizar en este análisis<sup>194</sup>. Por ejemplo, se están utilizando técnicas basadas en la combinación de redes neuronales y árboles de decisión, así como algoritmos de aprendizaje automático del tipo SVM, para generar advertencias (incluso en tiempo real) a los gestores respecto de cambios en los patrones de variación de precios subyacentes en los mercados.

En un entorno de *big data*, los gestores de carteras de valores se

enfrentan al reto de la sobrecarga de datos y noticias falsas. Los proveedores de análisis de datos están integrando información procedente de las redes sociales de determinadas empresas y personas influyentes. Pero muchos profesionales se sienten abrumados por la cantidad de información disponible a la hora organizar y analizar de forma sistemática grandes cantidades de datos. En este contexto, los algoritmos de inteligencia artificial están ayudando a extraer correlaciones entre los datos masivos y ofrecen apoyo a la toma de decisiones. En la búsqueda de alfa (*alpha seeking*) mediante el desarrollo de estrategias orientadas a generar rendimientos superiores al mercado o a un índice de referencia, la inteligencia artificial puede desempeñar un papel importante ya que ayuda a identificar patrones, señales y oportunidades que los gestores de inversiones pueden aprovechar. Estas funciones incluyen:

- El análisis de información (procesando gran cantidad de datos financieros y no financieros de manera eficiente para identificar patrones y tendencias no evidentes para los humanos).

<sup>194</sup> Véase Wu y Olson (2015) y Woodall (2017).

- La generación de modelos predictivos de movimientos futuros en los mercados financieros.
- La identificación de *sentimiento del mercado* a partir de fuentes de información en tiempo real, como noticias, redes sociales y blogs, que ayuda a los inversores a comprender mejor cómo se sienten los participantes del mercado y cómo podría afectar este sentimiento al comportamiento de los precios.
- La detección de anomalías (por ejemplo, patrones inusuales en el volumen de operaciones o en los precios de determinados activos).

Adicionalmente, en las actividades de gestión de carteras, los algoritmos de inteligencia artificial están ayudando en la gestión del riesgo de liquidez. Este riesgo está relacionado con los costes de entrar y salir con grandes posiciones cuando el mercado pierde profundidad y los diferenciales entre oferta y demanda se amplían rápidamente. Las entidades con

fuertes posiciones abiertas en los mercados (como los *hedge funds*, los bancos de inversión y ciertos fondos de inversión) están aplicando sistemas de inteligencia artificial para comprender mejor el impacto de su negociación en los precios y en la profundidad de mercado y existe ya cierta investigación sobre el tema<sup>195</sup>. Las técnicas de aprendizaje profundo ayudan a identificar conexiones entre activos que no son fácilmente observables y, así, permiten entrar en posiciones de riesgo deseadas a través de una serie de activos correlacionados, en lugar de tomar una gran posición en un solo activo<sup>196</sup>. En el mismo sentido, se están utilizando técnicas de aprendizaje por refuerzo, para dotar a los algoritmos de capacidad para aprender cuál es la reacción del mercado, de forma que en operaciones futuras pueda tenerse en cuenta cómo repercutirán las órdenes en los precios del mercado<sup>197</sup>.

Además del análisis de riesgo de mercado, de liquidez y crediticio, los analistas y gestores están integrando en sus decisiones de inversión

<sup>195</sup> Véase Petchamé (2011), Day (2017) y Tavana *et al.* (2018).

<sup>196</sup> Véase Cavalcante *et al.* (2016) y Heaton *et al.* (2017).

<sup>197</sup> Véase Hendricks y Wilcox (2014).

criterios de ESG. Los inversores institucionales encuentran dificultades serias para dotar de coherencia y comparabilidad entre empresas grandes cantidades de información disponible no estructurada para realizar evaluaciones de ESG. Están surgiendo proveedores externos de servicios de evaluación del

compromiso y las capacidades de las empresas para combatir el cambio climático, aplicar criterios sociales y de implantación de modelos adecuados de gobernanza. Estos proveedores están desarrollando indicadores directos e indirectos basados en algoritmos de inteligencia artificial.

### 8.3. EVALUACIÓN DE SOLVENCIA Y RATING DE BONOS

Al igual que en el caso del análisis de riesgo de financiación hipotecaria y de crédito al consumo, la inteligencia artificial está desempeñando hoy un papel crucial en el análisis de riesgo crediticio de las empresas (riesgo corporativo) y de los gobiernos (riesgo soberano) que llevan a cabo las agencias de calificación crediticia y los bancos de inversión. Estas técnicas hacen posible un análisis de datos más completo ya que permiten analizar grandes volúmenes de información financiera y no financiera, historiales de crédito, registros de transacciones, y otros datos de manera rápida y precisa. Además, estas técnicas ayudan en el desarrollo de modelos de crédito más sofisticados que tienen en cuenta una gama más

amplia de variables y factores, y permiten la detección de patrones y tendencias en los datos que podrían pasar desapercibidos en los modelos tradicionales. Estas técnicas también pueden ser utilizadas para identificar señales de riesgo tempranas y comportamientos de pago irregulares. Al proporcionar a los profesionales del crédito información más precisa y detallada, la inteligencia artificial puede ayudar a las instituciones financieras a tomar decisiones de crédito más informadas y a reducir el riesgo de incumplimiento.

El análisis del riesgo crediticio se basa en el cálculo de la pérdida esperada que, como es sabido, es el producto de tres medidas: la exposición

(EAD), la probabilidad de incumplimiento (PD) y la pérdida en caso de incumplimiento (LGD). Tradicionalmente, en la evaluación de la solvencia las entidades financieras han empleado modelos de clasificación simples, como análisis discriminante lineal (LDA) y modelos *logit* y *probit* clásicos. Ya desde 1994<sup>198</sup>, el uso de nuevas técnicas está mostrando que el análisis de solvencia puede mejorarse significativamente mediante el aprovechamiento de técnicas de AI, dada su capacidad de comprensión semántica de datos no estructurados<sup>199</sup>. Actualmente, en el análisis de riesgo corporativo basado en la estimación de parámetros de probabilidad de impago se están utilizando algoritmos de redes neurales, SVM, SVR, *boosting*, *bagging*, *random forest*, etc.<sup>200</sup> Y, del mismo modo, técnicas de inteligencia artificial como LASSO *regression*, Cluster análisis o MARS son aplicadas en los análisis de *stress testing*<sup>201</sup>. Por otra parte, el análisis de la pérdida

en caso de incumplimiento está teniendo mayor protagonismo como consecuencia del acuerdo de Basilea II, ya que se ha convertido en una medida mucho más crítica para los bancos y otras instituciones financieras<sup>202</sup>. Se ha observado que la distribución de la tasa de recuperación de los bonos y préstamos corporativos en mora es bimodal o multimodal. Además, se han identificado tanto factores transversales invariantes en el tiempo como factores con una dependencia temporal que afectan a la tasa de recuperación. Por ejemplo, las condiciones macroeconómicas en el momento del incumplimiento parecen fuentes potencialmente explicativas en la determinación del nivel de recuperación. Los estudios tradicionales han venido utilizando modelos paramétricos estándar y su poder predictivo para las tasas de recuperación ha resultado bajo. Puesto que, en los últimos años, la cantidad y calidad de los datos financieros y macroeconómicos disponibles han aumentado debido a mejoras

<sup>198</sup> Véase Altman *et al.* (1994).

<sup>199</sup> Véase Aziz y Dowling (2019).

<sup>200</sup> Véase Čížek *et al.* (2005), Hua *et al.* (2007), Chakraborty y Sharma (2007), Ahn y Kim (2009), Härdle *et al.* (2009), Yang *et al.* (2011), Härdle *et al.* (2012), Bastos (2014), Yao *et al.* (2015 y 2017), Raei *et al.* (2016), Mittal *et al.* (2016), Attigeri *et al.* (2017), Barboza *et al.* (2017), Aziz y Dowling (2019) y Aphale y Shinde (2020).

<sup>201</sup> Véase Blom (2015) o Chan-Lau (2017).

<sup>202</sup> Véase Loterman *et al.* (2012).

en la potencia computacional y de almacenamiento, el nuevo desafío, tanto teórico como computacional, es lidiar con los grandes conjuntos de datos disponibles para aplicaciones empíricas, en constante aumento<sup>203</sup>. En este contexto, se están utilizando métodos de aprendizaje automático del tipo de regresión gaussiana inversa, bosque aleatorio, propagación de expectativas de potencia escasa y regresión de vectores de soporte (SVR), así como operador LASSO<sup>204</sup>.

Una de las ventajas de los métodos de aprendizaje automático es que permiten ampliar los factores de análisis para crear una medida integral del riesgo crediticio basada en información cualitativa divulgada por las compañías que puede ser obtenida de informes anuales, conferencias telefónicas utilizadas para la presentación de resultados, folletos de emisión de bonos o de IPOs y de otras fuentes. Así, se están utilizando técnicas como *support vector regressions* (SVR), *supervised LDA* (sLDA), o

*random forest regression trees* (RF) con este propósito<sup>205</sup>.

Por lo que respecta a las calificaciones crediticias (*rating*), hay básicamente dos resultados la “calificación crediticia de incumplimiento” o *rating* del emisor y a la “calificación específica de la emisión de deuda”. La primera ofrece una opinión sobre la capacidad general y la predisposición de un emisor para pagar sus obligaciones financieras, de acuerdo con su solvencia fundamental; mientras la segunda informa sobre la probabilidad de que el inversor reciba los pagos prometidos de principal e intereses asociados con una particular emisión de bonos. Para el análisis de ambos tipos de calificaciones se están utilizando desde hace tiempo herramientas de inteligencia artificial del tipo de sistemas expertos, redes neuronales de retropropagación, redes neuronales artificiales competitivas, Support Vector Machines (SVM), Naïve Bayes Classification, árboles de clasificación y regresión CART, MPL, etc.<sup>206</sup>

<sup>203</sup> Véase Nazemi y Fabozzi (2018).

<sup>204</sup> Véase Yao *et al.* (2015), Nazemi *et al.* (2018 y 2022).

<sup>205</sup> Véase Frankel *et al.* (2027 y 2018), Donovan *et al.* (2021).

<sup>206</sup> Véase Surkan y Singleton (1990), Kim *et al.* (1993), Maher y Sen (1997), Chaveesuk *et al.* (1999), Shin y Han (1999), Kim y Han (2001), Huang *et al.* (2004a), Kumar y Bhattacharya (2006), Chen y Shih (2006), Lee (2007), Zhong *et al.* (2014), Wu *et al.* (2014), Daniel *et al.* (2019), Ozturk

## 8.4. MERCADOS PRIMARIOS DE CAPITALES

El éxito de los bancos de inversión en su actividad en los mercados primarios depende en gran medida de su capacidad para establecer correctamente el precio de los activos de deuda, capital e instrumentos híbridos (como bonos convertibles y canjeables) al que los inversores institucionales (fondos de inversión, *hedge funds* y compañías de seguros) están dispuestos a poner órdenes en los procesos de *bookbuilding*. En mercados en los que el aseguramiento (*underwriting*) no es habitual, la credibilidad sobre la fijación del precio adecuado de colocación es crucial en el posicionamiento competitivo de las entidades financieras en esta actividad.

Por lo que respecta al ámbito de las salidas a bolsa (IPOs), el mercado está utilizando diversas soluciones basadas en inteligencia artificial para predecir

el rendimiento de las acciones implicadas tras la colocación. Se han aplicado modelos de memoria a largo plazo (LSTM), redes neuronales artificiales<sup>207</sup> y redes neuronales recurrentes (RNN), así como redes neuronales difusas (FNN)<sup>208</sup> y algoritmos genéticos<sup>209</sup>. En general, las investigaciones han demostrado que los algoritmos K-Vecino más cercano, Random Forest (RF)<sup>210</sup>, XGBoost Regressor<sup>211</sup> y SVM<sup>212</sup>, son particularmente eficientes y aptos para abordar problemas que los modelos de regresión lineal no pueden resolver.

En relación con el mercado primario de deuda, parece evidente que para mantener una relación de confianza entre inversores y bancos colocadores (*bookrunners*) resulta determinante establecer con precisión el resultado de los proceso de *bookbuilding* en las emisiones públicas de deuda corporativa.

*et al.* (2016), Saha y Waheed, S. (2017), Golbayani *et al.* (2020), Jabeur *et al.* (2020), Takawira y Mwamba (2020 y 2022) y Overes y Van der Wel (2023).

<sup>207</sup> Véase Robertson *et al.* (1998), Reber *et al.* (2005), Mitsdorffer y Diederich (2008).

<sup>208</sup> Véase Wang *et al.* (2018).

<sup>209</sup> Véase Luque *et al.* (2012), Huang *et al.* (2012).

<sup>210</sup> Véase Basti *et al.* (2015), Quintana *et al.* (2017) y Baba y Sevil (2020).

<sup>211</sup> Véase Munshi *et al.* (2022),

<sup>212</sup> Véase Basti *et al.* (2015).

También resulta importante anticipar el impacto que, tras la colocación, pueden tener en el *mercado gris* y posteriormente en el mercado secundario determinadas informaciones corporativas. Las herramientas de inteligencia artificial también están ayudando en este ámbito específicamente en tres tareas:

- En la determinación del *fair value* del diferencial de crédito respecto de los pares (*peers*) más relevantes del emisor en virtud de su *rating*, sector, fortaleza y vulnerabilidades ante el entorno económico y financiero.
- En la predicción de la prima de nueva emisión (*new issue premium*) en función de las condiciones de mercado (apetito por el riesgo por parte de los inversores institucionales, acumulación de *pipeline*, nivel de *tráfico* en el mercado primario, etc.)
- En la identificación del nivel óptimo del precio indicativo inicial (*initial price thoughts*). El nivel que se muestra en la apertura del libro de órdenes influye en la creación de

*momento* en la demanda de los inversores durante la fase inicial de construcción del libro.

Por otra parte, recientemente, el mercado está prestando una importancia creciente a los aspectos corporativos de ESG, concediendo incluso una cierta prima de precio o (*greenium*) a aquellas empresas con buenas historias en este terreno. Los analistas, además de contar con *rating* de sostenibilidad independientes y con proveedores externos de segunda opinión (SPOs) para el análisis de los instrumentos de finanzas sostenibles<sup>213</sup>, están recabando por sí mismos información sobre el comportamiento real de las empresas al respecto. Su objetivo consiste en anticipar oportunidades y evitar potenciales problemas vinculados a comportamientos de *greenwashing*. Sin embargo, los datos, directos o indirectos, rara vez están estructurados de forma que sean legibles por las máquinas, de manera que las algunas compañías generadoras de software de AI están ofreciendo herramientas entrenadas para procesamiento del lenguaje natural y el reconocimiento de

<sup>213</sup> Como bonos y préstamos verdes e instrumentos vinculados a objetivos de sostenibilidad (SLBs y SLLs en su siglas en inglés).

imágenes que permite utilizar aplicaciones que interpretan

documentos de texto no estructurados e imágenes<sup>214</sup>.

## 8.5. CAPITAL PRIVADO Y CAPITAL RIESGO

Las implicaciones del desarrollo de la inteligencia artificial sobre la actividad de capital privado (PC) y capital riesgo (VC) son muy diversas, y muestran efectos positivos y negativos sobre la industria.

Por lo que respecta a la base de inversión, cabe destacar dos aspectos positivos y una potencial amenaza. Por una parte, se está produciendo una intensa generación de *startups* en las que la inteligencia artificial es sustancial a su modelo de negocio. De hecho, la AI es una de las áreas más dinámicas y la emergencia de empresas especializadas en este ámbito es enorme. Esta circunstancia significa una ampliación de la base de empresas objetivo para la industria del capital riesgo (*venture capital*)<sup>215</sup>. Investigaciones

recientes sugieren que las nuevas empresas de AI emplean modelos de negocio novedosos o diferentes y que sus capacidades únicas conducirán a nuevos productos, servicios y modelos de negocio<sup>216</sup>. A pesar de todo, se están observando desafíos importantes para la creación exitosa de valor a partir de la AI<sup>217</sup>. El ámbito de la inteligencia artificial continúa siendo un dominio altamente técnico y existe una brecha de habilidades y escasez de mano de obra en los roles laborales necesarios para implementar sistemas complejos de inteligencia artificial<sup>218</sup>. Además, resulta complicado para los emprendedores incipientes formar y pagar equipos con las habilidades necesarias para aprovechar con éxito las oportunidades.

<sup>214</sup> Véase DigFin (2019).

<sup>215</sup> Véase OECD (2018) y Obschonka y Audretsch (2020).

<sup>216</sup> Véase Brynjolfsson y McAfee (2017).

<sup>217</sup> Véase Jöhnk *et al.* (2021) y Benbya *et al.* (2020).

<sup>218</sup> Véase Marr (2018) y Chalmers *et al.* (2018).

Por otra parte, se está investigando en qué medida la AI tendrá un impacto sobre el emprendimiento. Si el impacto general sobre el mercado de trabajo derivado de la generalización de aplicaciones de AI es dramático en términos de reducción neta en el número total de horas trabajadas<sup>219</sup>, muchas personas podrían centrar sus esfuerzos en la creación de nuevas empresas. Una nueva generación de emprendedores podría proporcionar nuevas oportunidades de inversión, en sectores muy diversos, para los fondos de *venture capital* y *private equity*<sup>220</sup>. Desde la perspectiva negativa, sin embargo, existe el riesgo de las plataformas tecnológicas impulsadas por la AI, en poder de “grandes tecnologías”, se vuelvan más dominantes en la vida económica lo que llevaría a una creciente consolidación y una disminución a largo plazo del dinamismo empresarial independiente<sup>221</sup>.

En segundo término, la aplicación de herramientas de

inteligencia artificial puede afectar de forma decisiva a la función de gestión de inversiones, que es nuclear en la industria del capital privado, en sus tres etapas relevantes. Así, resulta posible entrenar algoritmos de AI para la detección de oportunidades de inversión (empresas en crecimiento) mediante el uso de “grandes datos”. Por otra parte, pueden utilizarse algoritmos de inteligencia artificial para ayudar a las empresas ya en cartera de los fondos de capital privado a evaluar opciones estratégicas de crecimiento<sup>222</sup>. Por ejemplo, estas herramientas pueden ayudar en la simulación de escenarios sobre procesos de aceptación de productos en nuevos mercados, sobre el éxito de la internacionalización y sobre los impactos de procesos de *building-up*<sup>223</sup>. Finalmente, los sistemas de AI pueden ayudar en el proceso de desinversión, detectando demanda de inversores industriales y financieros, así como evaluando la demanda esperada, cuando la salida de las empresas en

<sup>219</sup> Incluyendo el incremento de las horas de trabajo dedicadas al desarrollo e implantación de aplicaciones de inteligencia artificial en el conjunto de todas las industrias y sectores económicos.

<sup>220</sup> Véase D’Mello (2019), Levine (2019) y Choi y Kang (2019).

<sup>221</sup> Véase Decker *et al.* (2016 y 2017), Zhu y Liu (2018), Shapiro (2019) y Hartmann y Henkel (2020).

<sup>222</sup> Véase Chui *et al.* (2018), Tarafdar *et al.* (2019) y Soni *et al.* (2020).

<sup>223</sup> Véase Agrawal *et al.* (2018 y 2019).

cartera se produce mediante procesos de salidas a bolsa.

diario automatizando procesos de reporting, relación con inversores.

En otro orden de cosas, la AI puede ayudar a las gestoras de PE y VC en su funcionamiento



## REFERENCIAS

- Agrawal A., Gans J. S., and Goldfarb A. (2018) "Prediction machines: The simple economics of artificial intelligence", *Harvard Business Review Press*.
- Agrawal, A., Gans, J. S. and Goldfarb, A. (2019) "Artificial intelligence: the ambiguous labor market impact of automating prediction", *Journal of Economic Perspectives*, 33 (2): 31-50.
- Ahmed, S., Alshater, M. M., El Ammari, A., and Hammami, H. (2022) "Artificial intelligence and machine learning in finance: A bibliometric review", *Research in International Business and Finance*, 61:101646.
- Ahn, H., and Kim, K. J. (2009) "Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach", *Applied Soft Computing*, 9(2): 599-607.
- Altman, E. I., Marco, G., and Varetto, F. (1994) "Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience)", *Journal of Banking & Finance*, 18(3), 505-529.
- Amilon, H. (2003) "A neural network versus Black-Scholes: a comparison of pricing and hedging performances", *Journal of Forecasting*, 22(4): 317-335.
- Anders, U., Korn, O. and Schmitt, C. (1998) "Improving the pricing of options: A neural network approach", *Journal of Forecasting*, 17(5): 369-388.
- Andreou, P. C., Charalambous, C., and Martzoukos, S. H. (2006) "Robust artificial neural networks for pricing of European options", *Computational Economics*, 27: 329-351.
- Andreou, P. C., Charalambous, C., and Martzoukos, S. H. (2008) "Pricing and trading European options by combining artificial neural networks and parametric models with implied parameters", *European Journal of Operational Research*, 185(3): 1415-1433.
- Andreou, A. S., Georgopoulos, E. F., and Likothanassis, S. D. (2002) "Exchange-rates forecasting: A hybrid algorithm based on genetically optimized adaptive neural networks", *Computational Economics*, 20: 191-210.
- Aphale, A. S., and Shinde, S. R. (2020) "Predict loan approval in banking system machine learning approach for cooperative banks loan approval", *International Journal of Engineering Trends and Applications (IJETA)*, 9(8).
- Aragonés, J. R., Blanco, C., and Estévez, P. G. (2007) "Neural network volatility forecasts", *Intelligent Systems in Accounting, Finance & Management: International Journal*, 15(3-4), 107-121.

- Arner, D. W., Barberis, J. N., and Buckley, R. P. (2016) "The emergence of RegTech 2.0: From know your customer to know your data", *Journal of Financial Transformation*, 79 (UNSW Law Research Paper: 17-63).
- Assaad, R. H. and Fayek, S. (2021) "Predicting the price of crude oil and its fluctuations using computational econometrics: deep learning, LSTM, and convolutional neural networks", *Econometric Research in Finance*, 6(2): 119-137.
- Attigeri, G. V., Pai, M. M., and Pai, R. M. (2017) "Credit risk assessment using machine learning algorithms", *Advanced Science Letters*, 23(4): 3649-3653.
- Bahrammirzaee, A. (2010) "A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems", *Neural Computing and Applications*, 19(8), 1165-1195.
- Barboza, F., Kimura, H. and Altman, E. (2017) "Machine learning models and bankruptcy prediction", *Expert Systems with Applications* 83: 405-17.
- Basti, E., Kuzey, C., and Denle, D. (2015) "Analyzing initial public offerings short-term performance using decision trees and SVMs", *Decision Support Systems*, 73:15-27.
- Bastos, J. A. (2014) "Ensemble Predictions of Recovery Rates", *Journal of Financial Services Research* 46: 177-93.
- Béketo, M., Lehmann, K., and Wittke, M. (2018) "Robo Advisors: quantitative methods inside the robots", *Journal of Asset Management*, 19(6): 363-370.
- Benbya, H., Davenport, T. H., y Pachidi, S. (2020) "Artificial intelligence in organizations: current state and future opportunities", *MIS Quarterly Executive*, 19 (4): ix-xxi.
- Beytollahi, A. and Zeinali, H. (2020) "Comparing prediction power of artificial neural networks compound models in predicting credit default swap prices through Black-Scholes-Merton model", *Iranian Journal of Management Studies*, 13(1): 69-93.
- Bildirici, M., and Ersin, Ö. (2009) "Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul Stock Exchange", *Expert Systems with Applications*, 36(4): 7355-7362.
- Bin Khamis, A. and Yee, P. H. (2018) "A Hybrid Model of Artificial Neural Network and Genetic Algorithm in Forecasting Gold Price", *European Journal of Engineering and Technology Research*, 3(6): 10-14.
- Blom, T. (2015) "Top-down Stress Testing: An Application of Adaptive Lasso to Forecasting Credit Loss Rates", *Master's Thesis, Faculty of Science, Hong Kong, China*.
- Blynski, L. and Faseruk, A. (2006) "Comparison of the effectiveness of option price forecasting: Black-Scholes vs. simple and hybrid neural networks", *Journal of Financial Management & Analysis*, 19(2): 46.
- Bouteska, A., Hajek, P., Fisher, B., and Abedin, M. Z. (2023) "Nonlinearity in forecasting energy commodity prices: Evidence from a focused time-delayed neural network", *Research in International Business and Finance*, 64, 101863.
- Bozanic, Z., and Kraft, P. (2015) "Qualitative corporate disclosure and credit analysts' soft rating adjustments", *European Accounting Review*.
- Breeden, J. L. (2020) "A Survey of Machine Learning in Credit Risk", Mimeo.

- Brynjolfsson, E. and McAfee, A. (2017) "The business of artificial intelligence", *Harvard Business Review*.
- Butler, S., Kokoszka, P., Miao, H., and Shang, H. L. (2021) "Neural network prediction of crude oil futures using B-splines", *Energy Economics*, 94: 105080.
- Cao, H. K., Cao, H. K., and Nguyen, B. T. (2020) "Delafo: An efficient portfolio optimization using deep neural networks" en *Advances in Knowledge Discovery and Data Mining: 24th Pacific-Asia Conference, PAKDD: 11-14, Singapore, Proceedings, Part I 24: 623-635*. Springer International Publishing.
- Carta, S., Corrigan, A., Ferreira, A., Podda, A. S., and Recupero, D. R. (2021) "A multi-layer and multi-ensemble stock trader using deep learning and deep reinforcement learning", *Applied Intelligence*, 51: 889-905.
- Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., and Oliveira, A. L. (2016) "Computational intelligence and financial markets: A survey and future directions", *Expert Systems with Applications*, 55: 194-211.
- Chakraborty, S. and Sharma, S. K. (2007) "Prediction of corporate financial health by artificial neural network", *International Journal of Electronic Finance*, 1(4): 442-459.
- Chalmers, D., MacKenzie, N. G., and Carter, S. (2021) "Artificial intelligence and entrepreneurship: Implications for venture creation in the fourth industrial revolution", *Entrepreneurship Theory and Practice*, 45(5): 1028-1053.
- Chandrinou, S. K., Sakkas, G., and Lagaros, N. D. (2018). "AIRMS: A risk management tool using machine learning", *Expert Systems with Applications*, 105: 34-48.
- Chan-Lau, J. (2017) "Lasso Regressions and Forecasting Models in Applied Stress Testing", *IMF Working Papers*, 17 (1).
- Chaudhuri, T. D. and Ghosh, I. (2016) "Artificial neural network and time series modeling based approach to forecasting the exchange rate in a multivariate framework", *arXiv:1607.02093*.
- Chaveesuk, R., Srivaree-Ratana, C., and Smith, A. E. (1999) "Alternative neural network approaches to corporate bond rating", *Journal of Engineering Valuation and Cost Analysis*, 2(2): 117-131.
- Chen, Z., Goh, H. S., Sin, K. L., Lim, K., Chung, N. K. H., and Liew, X. Y. (2021). "Automated agriculture commodity price prediction system with machine learning techniques", *arXiv:2106.12747*.
- Chen, T. and Guestrin, C. (2016) "Xgboost: A scalable tree boosting system", *Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining: 785-794*.
- Chen Y., Hao Y. (2017) "A feature weighted support vector machine and K-nearest neighbour algorithm for stock market indices prediction", *Expert Systems With Applications*, 80: 340-355.
- Chen, S., Härdle, W. K., and Jeong, K. (2010) "Forecasting volatility with support vector machine-based GARCH model", *Journal of Forecasting*, 29(4): 406-433.
- Chen, S., Jeong, K., and Härdle, W. K. (2008) "Support vector regression based GARCH model with application to forecasting volatility of financial returns", *SFB 649 Discussion Paper*, 14/08. Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin.
- Chen, A. S. and Leung, M. T. (2004) "Regression neural network for error correction in foreign exchange forecasting and trading", *Computers & Operations Research*, 31(7): 1049-1068.
- Chen, W. H. and Shih, J. Y. (2006) "A study of Taiwan's issuer credit rating systems using support vector machines", *Expert Systems with Applications*, 30(3): 427-435.

- Chiong, R., Fan, Z., Hu, Z., Adam, M. T., Lutz, B., and Neumann, D. (2018) "A sentiment analysis-based machine learning approach for financial market prediction via news disclosures" en *Proceedings of the genetic and evolutionary computation conference companion*: 278-279. Julio.
- Choi, T. M., Chan, H. K., y Yue, X. (2017) "Recent development in big data analytics for business operations and risk management", *IEEE Transactions on Cybernetics*, 47(1): 81-92.
- Choi D. Y., Kang J. H. (2019) "Net job creation in an increasingly autonomous economy: The challenge of a generation", *Journal of Management Inquiry*, 28(3): 300-305.
- Chong, E., Han, C., and Park, F. C. (2017) "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies", *Expert Systems with Applications*, 83: 187-205.
- Chui, M., Manyika, J., and Miremadi, M. (2018) "What AI can and can't do (yet) for your business", *McKinsey Quarterly*, 1: 97-108.
- Čížek, P., Weron, R., Härdle, W., Härdle, W., Moro, R., and Schäfer, D. (2005) "Predicting bankruptcy with support vector machines": 225-248. Springer Berlin Heidelberg.
- Clark, B., Feinstein, Z., and Simaan, M. (2020) "A machine learning efficient frontier", *Operations Research Letters*, 48(5): 630-634.
- Cohen, S. N., Reisinger, C., and Wang, S. (2022) "Estimating risks of European option books using neural stochastic differential equation market models", *Journal of Computational Finance*, 26(3).
- Dang, Q. V. (2019) "Reinforcement learning in stock trading" en *International conference on computer science, applied mathematics and applications*: 311-322). Cham: Springer International Publishing.
- Daniel, C., Hančlová, J., and El Woujoud Bousselmi, H. (2019) "Corporate rating forecasting using Artificial Intelligence statistical techniques", *Investment Management & Financial Innovations*: 16(2), 295.
- Dash, R. and Dash, P. K. (2016) "A hybrid stock trading framework integrating technical analysis with machine learning techniques", *The Journal of Finance and Data Science*, 2 (1), 42-57.
- Day, S. (2017) "Quants turn to machine learning to model market impact". *Risk.net*.
- Dbouk, W. and Jamali, I. (2018) "Predicting daily oil prices: Linear and non-linear models", *Research in International Business and Finance*, 46: 149-165.
- Decker R. A., Haltiwanger J., Jarmin R. S. and Miranda J. (2016) "Declining business dynamism: What we know and the way forward", *American Economic Review*, 106(5): 203-207.
- Decker R. A., Haltiwanger J., Jarmin R. S. and Miranda J. (2017) "Declining dynamism, allocative efficiency, and the productivity slowdown", *American Economic Review*, 107(5): 322-326.
- Dempster, M. A. and Leemans, V. (2006) "An automated FX trading system using adaptive reinforcement learning", *Expert Systems with Applications*, 30(3): 543-552.
- Dempster, M. A., Payne, T. W., Romahi, Y., and Thompson, G. W. (2001) "Computational learning techniques for intraday FX trading using popular technical indicators", *IEEE Transactions on Neural Networks*, 12(4): 744-754.
- DigFin (2019) "ESG now a third of MioTech's A.I. business", 9 de septiembre.
- D'Mello, J. F. (2019) "Universal basic income and entrepreneurial pursuit in an autonomous society", *Journal of Management Inquiry*, 28(3): 306-310.

- Donovan, J., Jennings, J., Koharki, K., and Lee, J. (2021) "Measuring credit risk using qualitative disclosure", *Review of Accounting Studies*, 26: 815-863.
- Dubey, A. D. (2016) "Gold price prediction using support vector regression and ANFIS models" en *International Conference on Computer Communication and Informatics (ICCCI)*. Enero: 1-6. IEEE.
- El-Shazly, M. R. and El Shazly, H. E. (1999) "Forecasting currency prices using a genetically evolved neural network architecture", *International Review of Financial Analysis*, 8(1): 67-82.
- Eng, M. H., Li, Y., Wang, Q. G., and Lee, T. H. (2008) "Forecast forex with ANN using fundamental data" en *International Conference on Information Management, Innovation Management and Industrial Engineering*, 1: 279-282. IEEE.
- Fedenia, M., Nam, S., and Ronen, T. (2021) "Machine Learning in the Corporate Bond Market and Beyond: A New Classifier", disponible en SSRN 3848068.
- Feser, J. A. and Broby, D. (2020) "The Determinants of Credit Default Swap Premia and the Use of Machine Learning Techniques for their Estimation", *Centre for Financial Regulation and Innovation. Financial Technology Paper*: 1-126.
- Frankel, R. M., Jennings, J. N., and Lee, J. A. (2017) "Using natural language processing to assess text usefulness to readers: The case of conference calls and earnings prediction". Disponible en SSRN 3095754.
- Frankel, R., Jennings, J., and Lee, J. (2018) "How well do machines mimic readers? comparing analyst revisions to machine-based revisions generated from conference-call narratives". *Working paper*.
- Franses, P. H., and Van Griensven, K. (1998) "Forecasting exchange rates using neural networks for technical trading rules", *Studies in Nonlinear Dynamics & Econometrics*, 2(4).
- Franses, P. H. and Van Homelen, P. (1998) "On forecasting exchange rates using neural networks", *Applied Financial Economics*, 8(6): 589-596.
- Galeshchuk, S. (2016) "Neural networks performance in exchange rate prediction", *Neurocomputing*, 172: 446-452.
- Ganguli, S. and Dunnmon, J. (2017) "Machine learning for better models for predicting bond prices" *arXiv:1705.01142*.
- Gaspar, R. M., Lopes, S. D., and Sequeira, B. (2020) "Neural network pricing of American put options", *Risks*, 8(3): 73.
- Gašperov, B., Šarić, F., Begušić, S., and Kostanjčar, Z. (2020) "Adaptive rolling window selection for minimum variance portfolio estimation based on reinforcement learning" en *43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*: 1098-1102. IEEE. Septiembre.
- Gençay, R. and Qi, M. (2001) "Pricing and hedging derivative securities with neural networks: Bayesian regularization, early stopping, and bagging", *IEEE transactions on neural networks*, 12(4): 726-734.
- Ghaziri, H., Elfakhani, S., and Assi, J. (2000) "Neural, networks approach to pricing, options, *Neural Network World*, 1(2/00): 271-277.
- Gierjatowicz, P., Sabate-Vidales, M., Siska, D., Szpruch, L., and Zuric, Z. (2022). "Robust pricing and hedging via neural stochastic differential equations", *Journal of Computational Finance*, 26(3).



- Gite, S., Khatavkar, H., Kotecha, K., Srivastava, S., Maheshwari, P., and Pandey, N. (2021) "Explainable stock prices prediction from financial news articles using sentiment analysis", *PeerJ Computer Science*, 7: e340.
- Godarzi, A. A., Amiri, R. M., Talaei, A., and Jamasb, T. (2014) "Predicting oil price movements: A dynamic Artificial Neural Network approach", *Energy Policy*, 68: 371-382.
- Golbayani, P., Florescu, I., and Chatterjee, R. (2020) "A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees", *The North American Journal of Economics and Finance*, 54: 101251.
- Gonzalez Miranda, F. and Burgess, N. (1997) "Modelling market volatilities: the neural network perspective", *The European Journal of Finance*, 3(2): 137-157.
- Gotoh, J. Y., Takeda, A., and Yamamoto, R. (2014) "Interaction between financial risk measures and machine learning methods", *Computational Management Science*, 11(4): 365-402.
- Götze, T., Gürtler, M., and Witowski, E. (2020) "Improving CAT bond pricing models via machine learning", *Journal of Asset Management*, 21(5): 428-446.
- Götze, T., Gürtler, M., and Witowski, E. (2023) "Forecasting accuracy of machine learning and linear regression: evidence from the secondary CAT bond market", *Journal of Business Economics*: 1-32.
- Gradojevic, N., Gençay, R., and Kukolj, D. (2009) "Option pricing with modular neural networks", *IEEE Transactions on Neural Networks*, 20(4): 626-637.
- Gudelek, M. U., Boluk, S. A., and Ozbayoglu, A. M. (2017) "A deep learning based stock trading model with 2-D CNN trend detection" en *IEEE Symposium Series on Computational Intelligence (SSCI)*:1-8. IEEE.
- Guo, X. and Li, J. (2019) "A novel twitter sentiment analysis model with baseline correlation for financial market prediction with improved efficiency" en *Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*: 472-477. IEEE. Octubre.
- Gupta, R. and Chen, M. (2020) "Sentiment analysis for stock price prediction" en *2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*: 213-218. IEEE. Agosto.
- Gupta, I., Madan, T. K., Singh, S., and Singh, A. K. (2022b) "HiSA-SMFM: historical and sentiment analysis-based stock market forecasting model", *arXiv* :2203.08143.
- Hafezi, R. and Akhavan, A. (2018) "Forecasting gold price changes: Application of an equipped artificial neural network". *AUT Journal of Modeling and Simulation*, 50(1): 71-82.
- Hajizadeh, E., Mahootchi, M., Esfahanipour, A., and Massahi Kh, M. (2019) "A new NN-PSO hybrid model for forecasting Euro/Dollar exchange rate volatility", *Neural Computing and Applications*, 31: 2063-2071.
- Hamdi, M. and Aloui, C. (2015) "Forecasting crude oil price using artificial neural networks: a literature survey", *Economics Bulletin*, 35(2): 1339-1359.
- Hamid, S. A and Iqbal, Z. (2004) "Using neural networks for forecasting volatility of S&P 500 Index futures prices", *Journal of Business Research*, 57(10): 1116-1125.
- Hann, T. H. and Steurer, E. (1996) "Much ado about nothing? Exchange rate forecasting: Neural networks vs. linear models using monthly and weekly data", *Neurocomputing*, 10(4): 323-339.
- Härdle, W., Lee, Y. J., Schäfer, D., and Yeh, Y. R. (2009) "Variable selection and oversampling in the use

- of smooth support vector machines for predicting the default risk of companies”, *Journal of Forecasting*, 28(6): 512-534.
- Härdle, W. K., Prastyo, D., and Hafner, C. (2012) “Support vector machines with evolutionary feature selection for default prediction”. *SFB 649 Discussion Paper 2012-030*, Disponible en SSRN:2894201.
- Hartmann, P., and Henkel, J. (2020) “The rise of corporate science in AI: Data as a strategic resource”, *Academy of Management Discoveries*, 6(3): 359-381.
- Heaton, J. B., Polson, N. G., y Witte, J. H. (2017). Deep learning for finance: deep portfolios”, *Applied Stochastic Models in Business and Industry*, 33(1): 3-12.
- Hegazy, O., Soliman, O. S., and Salam, M. A. (2014) “A machine learning model for stock market prediction”, *arXiv:1402.7351*.
- Hendricks, D. y Wilcox, D. (2014) “A reinforcement learning extension to the Almgren-Chriss framework for optimal trade execution. en *IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFER)*: 457-464.
- Hiransha, M., Gopalakrishnan, E. A., Menon, V. K., and Soman, K. P. (2018) “NSE stock market prediction using deep-learning models”, *Procedia Computer Science*, 132: 1351-1362.
- Hu, J. W. S., Hu, Y. C., and Lin, R. R. W. (2012) “Applying neural networks to prices prediction of crude oil futures”, *Mathematical Problems in Engineering*.
- Hu, Y., Zhang, X., Feng, B., Xie, K., y Liu, M. (2015) “iTrade: A mobile data driven stock trading system with concept drift adaptation”, *International Journal of Data Warehousing and Mining*, 11(1): 66-83.
- Hu, M. Y., Zhang, G., Jiang, C. X., and Patuwo, B. E. (1999) “A cross-validation analysis of neural network out-of-sample performance in exchange rate forecasting”, *Decision Sciences*, 30(1): 197-216.
- Hua, Z., Wang, Y., Xu, X., Zhang, B. and Liang, L. (2007) “Predicting corporate financial distress based on integration of support vector machine and logistic regression”, *Expert Systems with Applications*, 33(2): 434-440.
- Huang, C. Chang, C., Kuo, L., Lin, B., Hsieh, T., and Chang, B. (2012) “A genetic-search model for first-day returns using IPO fundamentals”, en *International Conference on Machine Learning and Cybernetics, IEEE, Xian*: 1662-1667.
- Huang, Z., Chen, H., Hsu, C. J., Chen, W. H., and Wu, S. (2004a) “Credit rating analysis with support vector machines and neural networks: a market comparative study”, *Decision Support Systems*, 37(4): 543-558.
- Huang, W., Lai, K. K., Nakamori, Y., and Wang, S. (2004b) “Forecasting foreign exchange rates with artificial neural networks: A review”, *International Journal of Information Technology & Decision Making*, 3(01): 145-165.
- Huang, J. Y. and Liu, J. H. (2020) “Using social media mining technology to improve stock price forecast accuracy”, *Journal of Forecasting*, 39(1): 104-116.
- Huang, Y., Jia, Y., and Zhou, X. (2022) “Achieving mean-variance efficiency by continuous-time reinforcement learning” en *Proceedings of the Third ACM International Conference on AI in Finance*: 377-385. Noviembre.
- Hussein, S. F. M., Shah, M. B. N., Abd Jalal, M. R., and Abdullah, S. S. (2011) “Gold price prediction using

- radial basis function neural network”, *Fourth International Conference on Modeling, Simulation and Applied Optimization*: 1-11. Abril. IEEE.
- Hutchinson, J. M., Lo, A. W., and Poggio, T. (1994) “A nonparametric approach to pricing and hedging derivative securities via learning networks”, *The Journal of Finance*, 49(3): 851-889.
- Ivaşcu, C. F. (2021) “Option pricing using machine learning”, *Expert Systems with Applications*, 163: 113799.
- Jabeur, S. B., Sadaoui, A., Sghaier, A., and Aloui, R. (2020) “Machine learning models and cost-sensitive decision trees for bond rating prediction”, *Journal of the Operational Research Society*, 71(8): 1161-1179.
- Jaiswal, R., Jha, G. K., Kumar, R. R., and Choudhary, K. (2022) “Deep long short-term memory-based model for agricultural price forecasting”, *Neural Computing and Applications*, 34(6): 4661-4676.
- Jangmin, O., Lee, J., Lee, J. W., and Zhang, B. T. (2006) “Adaptive stock trading with dynamic asset allocation using reinforcement learning”, *Information Sciences*, 176(15): 2121-2147.
- Jha, G. K. and Sinha, K. (2013) “Agricultural price forecasting using neural network model: An innovative information delivery system”, *Agricultural Economics Research Review*, 26(347-2016-17087): 229-239.
- Jiang, H., Hu, W., Xiao, L., and Dong, Y. (2022) “A decomposition ensemble based deep learning approach for crude oil price forecasting”, *Resources Policy*, 78, 102855.
- Jing, N., Wu, Z., and Wang, H. (2021) “A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction”, *Expert Systems with Applications*, 178: 115019.
- Jöhnk, J., Weißert, M., and Wyrski, K. (2021) “Ready or not, AI comes -an interview study of organizational AI readiness factors”, *Business & Information Systems Engineering*, 63(1): 5-20.
- Kamruzzaman, J. and Sarker, R. A. (2004) “ANN-based forecasting of foreign currency exchange rates”, *Neural Information Processing-Letters and Reviews*, 3(2): 49-58.
- KangaraniFarahani, M. and Mehralian, S. (2013) “Comparison between artificial neural network and neuro-fuzzy for gold price prediction” en *13th Iranian Conference on Fuzzy Systems (IFSC)* Agosto: 1-5. IEEE.
- Kilburn, F. (2019) “Can robots learn to manage risk?”, *RiskNet*, 11 de julio.
- Kilburn, F. (2023) “The haves and the ‘have bots: can AI give vol forecasters an edge?” *RiskNet*, 1 de marzo.
- Kim, M. (2021) “Adaptive trading system integrating machine learning and back-testing: Korean bond market case”, *Expert Systems with Applications*, 176, 114767.
- Kim, J. M., Kim, D. H., and Jung, H. (2021) “Applications of machine learning for corporate bond yield spread forecasting”, *The North American Journal of Economics and Finance*, 58: 101540.
- Kanevski, M. F. and Timonin, V. (2010) “Machine learning analysis and modeling of interest rate curves”, *Paper presented at the 18th European Symposium on Artificial Neural Networks ESANN*, Bruges, Belgium, 28-30 de Abril.
- Kim, K. S. and Han, I. (2001) “The cluster-indexing method for case-based reasoning using self-organizing maps and learning vector quantization for bond rating cases”, *Expert Systems with Applications*, 21(3): 147-156.

- Kim, J. W., Weistroffer, H. R., and Redmond, R. T. (1993) "Expert systems for bond rating: a comparative analysis of statistical, rule-based and neural network systems", *Expert Systems*, 10(3): 167-172.
- Kodogiannis, V. and Lolis, A. (2002) "Forecasting financial time series using neural network and fuzzy system-based techniques", *Neural Computing & Applications*, 11: 90-102.
- Kordonis, J., Symeonidis, S., and Arampatzis, A. (2016) "Stock price forecasting via sentiment analysis on Twitter" en *Proceedings of The 20th Pan-Hellenic Conference on Informatics*: 1-6. Noviembre.
- Koukaras, P., Nousi, C., and Tjortjis, C. (2022) "Stock market prediction using microblogging sentiment analysis and machine learning" en *Telecom, MDPI*, 3 (2):358-378.
- Koy, A. and Çolak, A. B. (2023) "Predicting Stock Market Index and Credit Default Swap Spreads Using Artificial Intelligence and Determining Non-Linear Relations", *Archives of Advanced Engineering Science*, 1-18.
- Kristjanpoller, W. and Hernández, E. (2017) "Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors", *Expert Systems with Applications*, 84, 290-300.
- Kristjanpoller, W., Fadic, A. and Minutolo, M. C. (2014) "Volatility forecast using hybrid neural network models", *Expert Systems with Applications*, 41(5): 2437-2442.
- Kristjanpoller, W., and Minutolo, M. C. (2015) "Gold price volatility: A forecasting approach using the Artificial Neural Network-GARCH model", *Expert systems with applications*, 42(20): 7245-7251.
- Kumar, P. P. (2018) *Machine learning for model development in market risk*. GARP Institute.
- Kumar, K., and Bhattacharya, S. (2006) "Artificial neural network vs linear discriminant analysis in credit ratings forecast: A comparative study of prediction performances", *Review of Accounting and Finance*, 5(3), 216-227.
- Kumbure, M. M., Lohrmann, C., Luukka, P., and Porras, J. (2022) "Machine learning techniques and data for stock market forecasting: A literature review", *Expert Systems with Applications*, 197, 116659.
- Lajbcygier, P. (2004) "Improving option pricing with the product constrained hybrid neural network", *IEEE Transactions on Neural Networks*, 15(2): 465-476.
- Lajbcygier, P. R. and Connor, J. T. (1997) "Improved option pricing using artificial neural networks and bootstrap methods", *International Journal of Neural Systems*, 8(04): 457-471.
- Lee, Y. C. (2007) "Application of support vector machines to corporate credit rating prediction", *Expert Systems with Applications*, 33(1): 67-74.
- Leippold, M., Wang, Q., and Zhou, W. (2022) "Machine learning in the Chinese stock market", *Journal of Financial Economics*, 145(2): 64-82.
- Leung, M. T., Chen, A. S., and Daouk, H. (2000) "Forecasting exchange rates using general regression neural networks", *Computers & Operations Research*, 27(11-12): 1093-1110.
- Levine, D. I. (2019) "Automation as part of the solution", *Journal of Management Inquiry*, 28(3): 316-318.
- Li, B. (2014) "Research on WNN modeling for gold price forecasting based on improved artificial bee colony algorithm", *Computational intelligence and neuroscience*, 14: (2)-2.
- Li, Y., Ni, P., and Chang, V. (2020a) "Application of deep reinforcement learning in stock trading strategies and stock forecasting", *Computing*, 102(6): 1305-1322.
- Li, X., Wu, P., and Wang, W. (2020b) "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong", *Information Processing & Management*, 57(5): 102212.

- Lin, C. T., and Yeh, H. Y. (2009) "Empirical of the Taiwan stock index option price forecasting model-applied artificial neural network", *Applied Economics*, 41(15): 1965-1972.
- Lin, C. (2015) "Build prediction models for gold prices based on back-propagation neural network" en *International Conference on Modeling, Simulation and Applied Mathematics*: 155-158. Atlantis Press.
- Lin, S. Y., Liu, D. R., and Huang, H. P. (2022) "Credit default swap prediction based on generative adversarial networks", *Data Technologies and Applications*, 56 (5): 720-740.
- Lin, Y., Liu, S., Yang, H., Wu, H., and Jiang, B. (2021) "Improving stock trading decisions based on pattern recognition using machine learning technology". *PloS one*, 16 (8): e0255558.
- Liu, X. Y., Yang, H., Chen, Q., Zhang, R., Yang, L., Xiao, B., and Wang, C. D. (2020). FinRL: A deep reinforcement learning library for automated stock trading in quantitative finance. *arXiv:2011.09607*.
- Loterman, G., Brown, I., Martens, D., Mues, C. and Baesens, B. (2012), "Benchmarking regression algorithms for loss given default modelling", *International Journal of Forecasting*, 28 (1):161-170.
- Lu, X., Que, D., and Cao, G. (2016) "Volatility forecast based on the hybrid artificial neural network and GARCH-type models", *Procedia Computer Science*, 91: 1044-1049.
- Luo, Z., Cai, X., Tanaka, K., Takiguchi, T., Kinkyo, T., and Hamori, S. (2019) "Can we forecast daily oil futures prices? Experimental evidence from convolutional neural networks", *Journal of Risk and Financial Management*, 12(1): 9.
- Luo, C., Wu, D., and Wu, D. (2017) "A deep learning approach for credit scoring using credit default swaps", *Engineering Applications of Artificial Intelligence*, 65, 465-470.
- Luque, C., Quintana, D. and Isasi, P. (2012) "Predicting IPO underpricing with genetic algorithms", *International Journal of Artificial Intelligence*, 8: 133-146.
- Lynn, T., Mooney, J. G., Rosati, P., y Cummins, M. (2019). *Disrupting finance: FinTech and strategy in the 21st century*. Springer Nature. Palgrave. MacMillan.
- Mahdavi-Damghani, B. and Roberts, S. (2017) "A Proposed Risk Modeling Shift from the Approach of Stochastic Differential Equation towards Machine Learning Clustering: Illustration with the Concepts of Anticipative and Responsible VaR", *SSRN Electronic Journal*.
- Maher, J. J., and Sen, T. K. (1997) "Predicting bond ratings using neural networks: a comparison with logistic regression", *Intelligent Systems in Accounting, Finance & Management*, 6(1): 59-72.
- Malliaris, A. G. and Malliaris, M. (2013) "Are oil, gold and the euro inter-related? Time series and neural network analysis", *Review of Quantitative Finance and Accounting*, 40: 1-14.
- Malliaris, A. G. and Malliaris, M. (2015) "What drives gold returns? A decision tree analysis", *Finance Research Letters*, 13: 45-53.
- Malliaris, M. and Salchenberger, L. (1993a) "A neural network model for estimating option prices", *Applied Intelligence*, 3: 193-206.
- Malliaris, M. and Salchenberger, L. (1993b) "Beating the best: A neural network challenges the Black-Scholes formula" en *Proceedings of 9th IEEE Conference on Artificial Intelligence for Applications*: 445-449. Marzo. IEEE.
- Manjula, K. A. and Karthikeyan, P. (2019) "Gold price prediction using ensemble-based machine

- learning techniques”, en *3rd International Conference on Trends in Electronics and Informatics (ICOEI)*: 1360-1364. Abril. IEEE.
- Manogna R. L. and Mishra, A. K. (2021) “Forecasting spot prices of agricultural commodities in India: Application of deep-learning models”, *Intelligent Systems in Accounting, Finance and Management*, 28(1): 72-83.
- Mao, W., Zhu, H., Wu, H., Lu, Y., and Wang, H. (2023) “Forecasting and trading credit default swap indices using a deep learning model integrating Merton and LSTMs”, *Expert Systems with Applications*, 213: 119012.
- Maqsood, H., Mehmood, I., Maqsood, M., Yasir, M., Afzal, S., Aadil, F., ... and Muhammad, K. (2020) “A local and global event sentiment based efficient stock exchange forecasting using deep learning”, *International Journal of Information Management*, 50: 432-451.
- Marr, B. (2018) “The AI skills crisis and how to close the gap”, *Forbes*. 25 de Junio.
- Mascio, D. A., Fabozzi, F. J., and Zumwalt, J. K. (2021) “Market timing using combined forecasts and machine learning”, *Journal of Forecasting*, 40(1): 1-16.
- Matías, J. M., Febrero-Bande, M., González-Manteiga, W., and Reboredo, J. C. (2010). “Boosting GARCH and neural networks for the prediction of heteroskedastic time series”, *Mathematical and Computer Modelling*, 51(3-4): 256-271.
- Matsui, T., Goto, T., and Izumi, K. (2009) “Acquiring a government bond trading strategy using reinforcement learning”, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 13(6): 691-696.
- Mehta, P., Pandya, S., and Kotecha, K. (2021) “Harvesting social media sentiment analysis to enhance stock market prediction using deep learning”. *PeerJ Computer Science*, 7: e476.
- Mehtab, S. and Sen, J. (2019) “A robust predictive model for stock price prediction using deep learning and natural language processing”, *arXiv:1912.07700*.
- Mitra, S. K. (2012) “An option pricing model that combines neural network approach and Black Scholes formula”, *Global Journal of Computer Science and Technology*, 12(4): 6-16.
- Mitsdorffer, R. and Diederich, J. (2008) “Prediction of first-day returns of initial public offering in the US stock market using extraction from support vector machines” en J. Diederich (Ed.), *Rule extraction from support vector machines*, 80, Springer, Berlin: 185-203.
- Mittal, A. and Goel, A. (2012) “Stock prediction using twitter sentiment analysis”, *Stanford University WP CS229*:15, 2352.
- Mittal, L., Gupta, T., and Sangaiah, A. K. (2016) “Prediction of credit risk evaluation using naive bayes, artificial neural network and support vector machine”, *The IIOAB Journal*, 7(2): 33-42.
- Mohan, S., Mullapudi, S., Sammeta, S., Vijayvergia, P., and Anastasiu, D. C. (2019) “Stock price prediction using news sentiment analysis” en *IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)*: 205-208. IEEE. Abril.
- Monfared, S. A. and Enke, D. (2014) “Volatility Forecasting Using a Hybrid GJR-GARCH Neural Network Model”, *Procedia Computer Science* 36: 246-53.
- Nag, A. K. and Mitra, A. (2002) “Forecasting daily foreign exchange rates using genetically optimized neural networks”, *Journal of Forecasting*, 21(7): 501-511.

- Nazemi, A., Baumann, F., and Fabozzi, F. J. (2022) "Intertemporal defaulted bond recoveries prediction via machine learning", *European Journal of Operational Research*, 297(3): 1162-1177.
- Nazemi, A. and Fabozzi, F. J. (2018) "Macroeconomic variable selection for creditor recovery rates", *Journal of Banking & Finance*, 89: 14-25.
- Nazemi, A., Heidenreich, K., and Fabozzi, F. J. (2018) "Improving corporate bond recovery rate prediction using multi-factor support vector regressions", *European Journal of Operational Research*, 271(2): 664-675.
- Nguyen, T. H., and Shirai, K. (2015) "Topic modeling based sentiment analysis on social media for stock market prediction" en *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing 1: Long Papers*:1354-1364).
- Nguyen, T. H., Shirai, K., and Velcin, J. (2015) "Sentiment analysis on social media for stock movement prediction", *Expert Systems with Applications*, 42(24): 9603-9611.
- Ni, H. and Yin, H. (2009) "Exchange rate prediction using hybrid neural networks and trading indicators", *Neurocomputing*, 72(13-15): 2815-2823.
- Niranjan, M. (1996) "Sequential tracking in pricing financial options using model based and neural network approaches", *Advances in Neural Information Processing Systems*, 9.
- Obschonka, M. and Audretsch, D. B. (2020) "Artificial intelligence and big data in entrepreneurship: a new era has begun", *Small Business Economics*, 55: 529-539.
- OECD. (2018) *Private equity investment in artificial intelligence*. OECD.
- Overes, B. H., and Van der Wel, M. (2023) "Modelling sovereign credit ratings: Evaluating the accuracy and driving factors using machine learning techniques", *Computational Economics*, 61(3): 1273-1303.
- Ozturk, H., Namli, E., and Erdal, H. I. (2016) "Reducing overreliance on sovereign credit ratings: Which model serves better?", *Computational Economics*, 48: 59-81.
- Pacelli, V., Bevilacqua, V., and Azzollini, M. (2011) "An artificial neural network model to forecast exchange rates", *Journal of Intelligent Learning Systems and Applications*, 3(02): 57-69.
- Papanicolaou, A., Fu, H., Krishnamurthy, P., and Khorrami, F. (2023) "A Deep Neural Network Algorithm for Linear-Quadratic Portfolio Optimization with MGARCH and Small Transaction Costs", *IEEE Access*, 11: 16774-16792.
- Pei, Y., Huang, C. J., Shen, Y., and Wang, M. (2023) "A Novel Model for Spot Price Forecast of Natural Gas Based on Temporal Convolutional Network", *Energies*, 16(5): 2321.
- Pendharker, P. C., and Cusatis, P. (2018) "Trading financial indices with reinforcement learning agents", *Expert Systems with Applications*, 103, 1-13.
- Petchamé-Sala, J. (2011) "Liquidity risk modeling using artificial neural network", *Master's thesis, Universitat Politècnica de Catalunya*.
- Posthaus, D. A. (2019) "Yield curve fitting with artificial intelligence: a comparison of standard fitting methods with artificial intelligence algorithms", *Journal of Computational Finance*.
- Pretorius, R. and Van Zyl, T. (2022) "Deep reinforcement learning and convex mean-variance optimisation for portfolio management", *arXiv*: 2203.11318.



- Qian, L., Zhao, J., and Ma, Y. (2022) "Option Pricing Based on GA-BP neural network", *Procedia Computer Science*, 199: 1340-1354.
- Quintana, D., Saez, Y. and Isasi, P. (2017) "Random forest prediction of IPO underpricing", *Applied Sciences*, 6 (7).
- Raei, R., Saeidi Kousha, M., Fallahpour, S. and Fadaeinejad, M. (2016) "A Hybrid Model for Estimating the Probability of Default of Corporate Customers", *Iranian Journal of Management Studies* 9: 651-73.
- Rasheed, A., Younis, M. S., Ahmad, F., Qadir, J., and Kashif, M. (2021) "District wise price forecasting of wheat in Pakistan using deep learning", *arXiv:2103.04781*.
- Reber, B., Berry, B., Toms, S. (2005) "Predicting mispricing of initial public offerings", *Intelligent Systems in Accounting, Finance and Management*, 13 (1): 41-59.
- Refenes, A. N., Azema-Barac, M., Chen, L., and Karoussos, S. A. (1993) "Currency exchange rate prediction and neural network design strategies", *Neural Computing & Applications*, 1: 46-58.
- Refenes, A. P. and Holt, W. T. (2001) "Forecasting volatility with neural regression: A contribution to model adequacy", *IEEE Transactions on Neural Networks*, 12(4): 850-864.
- Robertson, S. J., Golden, B.L., Runger, G.C., and Wasil E.E. (1998) "Neural network models for initial public offerings", *Neurocomputing*, 18 (3):165-182.
- Rundo, F. (2019) "Deep LSTM with reinforcement learning layer for financial trend prediction in FX high frequency trading systems", *Applied Sciences*, 9(20): 4460.
- Saha, S., and Waheed, S. (2017) "Credit risk of bank customers can be predicted from customer's attribute using neural network", *International Journal of Computer Applications*, 161(3): 39-43.
- Sanford, A. and Moosa, I. (2015) "Operational risk modelling and organizational learning in structured finance operations: A Bayesian network approach", *Journal of the Operational Research Society*, 66(1): 86-115.
- SaShaban, K., Younes, A., Lam, R., Allison, M., and Kathirgamanathan, S. (2010) "Credit default swap pricing using artificial neural networks" en *The 2010 International Joint Conference on Neural Networks (IJCNN)*: 1-8. IEEE.
- Shapiro C. (2019) "Protecting competition in the American economy: Merger control, tech titans, labor markets", *Journal of Economic Perspectives*, 33(3): 69-93.
- Shen, S., Jiang, H., and Zhang, T. (2012) "Stock market forecasting using machine learning algorithms", *Department of Electrical Engineering, Stanford University, Stanford, CA*: 1-5.
- Shen, J. and Shafiq, M. O. (2020) "Short-term stock market price trend prediction using a comprehensive deep learning system", *Journal of big Data*, 7(1): 1-33.
- Shin, K. S. and Han, I. (1999) "Case-based reasoning supported by genetic algorithms for corporate bond rating", *Expert Systems with Applications*, 16(2): 85-95.
- Singh, R. and Srivastava, S. (2017) "Stock prediction using deep learning", *Multimedia Tools and Applications*, 76: 18569-18584.
- Soni, S. (2011) "Applications of ANNs in stock market prediction: a survey", *International Journal of Computer Science & Engineering Technology*, 2(3): 71-83.
- Soni, N., Sharma, E. K., Singh, N., and Kapoor, A. (2020) "Artificial intelligence in business: from research and innovation to market deployment", *Procedia Computer Science*, 167: 2200-2210.

- Surkan, A. J., and Singleton, J. C. (1990) "Neural networks for bond rating improved by multiple hidden layers" en *IJCNN international joint conference on neural networks*: 157-162). IEEE. Junio.
- Swathi, T., Kasiviswanath, N., and Rao, A. A. (2022) "An optimal deep learning-based LSTM for stock price prediction using twitter sentiment analysis", *Applied Intelligence*, 52(12): 13675-13688.
- Takawira, O., and Mwamba, W. M. (2020) "Determinants of sovereign credit ratings: An application of the Naïve Bayes classifier", *Eurasian Journal of Economics and Finance*, 8(4): 279-299.
- Takawira, O. and Mwamba, J. W. M. (2022) "An Analysis of Sovereign Credit Ratings Using Random Forest", *International Journal of Economics and Finance Studies*, 14(1): 29-87.
- Tarafdar, M., Beath, C. M., and Ross, J. W. (2019) "Using AI to enhance business operations", *MIT Sloan Management Review*, 60 (4).
- Tavana, M., Abtahi, A. R., Di Caprio, D., and Poortarigh, M. (2018) "An Artificial Neural Network and Bayesian Network model for liquidity risk assessment in banking", *Neurocomputing*, 275: 2525-2554.
- Tenti, P. (1996) "Forecasting foreign exchange rates using recurrent neural networks", *Applied Artificial Intelligence*, 10(6): 567-582.
- Tsang, K. H., and Wong, H. Y. (2020) "Deep-learning solution to portfolio selection with serially dependent returns", *SIAM Journal on Financial Mathematics*, 11(2): 593-619.
- Van Staden, P. M., Forsyth, P. A., and Li, Y. (2023) "A parsimonious neural network approach to solve portfolio optimization problems without using dynamic programming", *arXiv:2303.08968*.
- Wang, D., Qian, X., Quek, C. Tan, A., Miao, C, Zhang, X. et al. (2018) "An interpretable neural fuzzy inference for predictions of underpricing in initial public offerings", *Neurocomputing*, 319: 102-117.
- Wang, B. and Wang, J. (2021) "Energy futures price prediction and evaluation model with deep bidirectional gated recurrent unit neural network and RIF-based algorithm", *Energy*, 216: 119299.
- Wang, H. and Zhou, X. Y. (2020) "Continuous-time mean-variance portfolio selection: A reinforcement learning framework", *Mathematical Finance*, 30(4): 1273-1308.
- Warin, X. (2021) "Deep learning for efficient frontier calculation in finance" *arXiv:2101.02044*.
- Williams, J. (2018) "Can machine learning tools improve portfolio risk management...?" *Hedgeweak*, 28 de noviembre.
- Woodall, L. (2017) "Model risk managers eye benefits of machine learning". *Risk.net*.
- Wright, D., Capriotti, L., and Lee, J. (2018) "Machine learning and corporate bond trading", *Algorithmic Finance*, 7(3-4): 105-110.
- Wu, X., Chen, H., Wang, J., Troiano, L., Loia, V., and Fujita, H. (2020) "Adaptive stock trading strategies with deep reinforcement learning methods", *Information Sciences*, 538: 142-158.
- Wu, H. C., Hu, Y. H., and Huang, Y. H. (2014) "Two-stage credit rating prediction using machine learning techniques", *Kybernetes*, 43(7): 1098-1113.
- Wu, D. D. and Olson, D. L. (2015) *Enterprise Risk Management in Finance*. Springer, New York.
- Xu, X., and Zhang, Y. (2022a) "Thermal coal price forecasting via the neural network", *Intelligent Systems with Applications*, 14: 200084.

- Xu, X. and Zhang, Y. (2022b) "Canola and soybean oil price forecasts via neural networks", *Advances in Computational Intelligence*, 2(5):32.
- Xu, X. and Zhang, Y. (2023) "Coking coal futures price index forecasting with the neural network", *Mineral Economics*, 36(2): 349-359.
- Yang, H., Liu, X. Y., Zhong, S., & Walid, A. (2020) "Deep reinforcement learning for automated stock trading: An ensemble strategy" en *Proceedings of the first ACM international conference on AI in finance* (Ocutbre): 1-8.
- Yang, Z., You, W., and Ji, G. (2011) "Using partial least squares and support vector machines for bankruptcy prediction", *Expert Systems with Applications* 38: 8336-42.
- Yao, X., Crook, J., and Andreeva, G. (2015) "Support vector regression for loss given default modelling", *European Journal of Operational Research*, 240(2): 528-538.
- Yao, X., Crook, J. and Andreeva, G. (2017) "Enhancing two-stage modelling methodology for loss given default with support vector machines", *European Journal of Operational Research* 263: 679-89.
- Yao, J., Li, Y., and Tan, C. L. (2000) "Option price forecasting using neural networks", *Omega*, 28(4): 455-466.
- Yao, J. and Tan, C. L. (2000) "A case study on using neural networks to perform technical forecasting of forex", *Neurocomputing*, 34(1-4): 79-98.
- Yao, Q., Wang, M., Chen, Y., Dai, W., Li, Y. F., Tu, W. W., ... and Yu, Y. (2018) "Taking human out of learning applications: A survey on automated machine learning", *arXiv:1810.13306*.
- Yazdani-Chamzini, A., Yakhchali, S. H., Volungevičienė, D., and Zavadskas, E. K. (2012) "Forecasting gold price changes by using adaptive network fuzzy inference system", *Journal of Business Economics and Management*, 13(5): 994-1010.
- Yoo, S., Song, J., and Jeong, O. (2018) "Social media contents-based sentiment analysis and prediction system", *Expert Systems with Applications*, 105: 102-111.
- Yu, L., Wang, S., and Lai, K. K. (2007) *Foreign-exchange-rate forecasting with artificial neural networks* (Vol. 107). Springer Science & Business Media.
- Yu, L., Yao, X., Wang, S., and Lai, K. K. (2011) "Credit risk evaluation using a weighted least squares SVM classifier with design of experiment for parameter selection", *Expert Systems with Applications*, 38(12): 15392-15399.
- Zhang, G. P. and Berardi, V. L. (2001) "Time series forecasting with neural network ensembles: an application for exchange rate prediction", *Journal of the operational research society*, 52: 652-664.
- Zhang, G. and Hu, M. Y. (1998) "Neural network forecasting of the British pound/US dollar exchange rate", *Omega*, 26(4): 495-506.
- Zhang, H. G., Su, C. W., Song, Y, Qiu, S., Xiao, R., and Su, F. (2017) "Calculating Value-at-Risk for High-Dimensional Time Series Using a Nonlinear Random Mapping Model", *Economic Modelling* 67: 355-67.
- Zhang, C., Zhang, Z., Cucuringu, M., and Zohren, S. (2021c) "A universal end-to-end approach to portfolio optimization via deep learning", *arXiv: 2111.09170*.
- Zhang, Z., Zohren, S., and Stephen, R. (2020b) "Deep reinforcement learning for trading", *The Journal of Financial Data Science*, 5(3):1-29629616.



- Zheng, X. L., Zhu, M. Y., Li, Q. B., Chen, C. C., and Tan, Y. C. (2019) "FinBrain: when finance meets AI 2.0", *Frontiers of Information Technology & Electronic Engineering*, 20 (7): 914-924.
- Zhong, H., Miao, C., Shen, Z., and Feng, Y. (2014) "Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings", *Neurocomputing*, 128: 285-295.
- Zhu F., Liu Q. (2018). "Competing with complementors: An empirical look at Amazon.com.", *Strategic Management Journal*, 39 (10): 2618-2642.
- Zhu, Y. and Zhang, C. (2018) "Gold price prediction based on PCA-GA-BP neural network", *Journal of Computer and Communications*, 6(7): 22-33.

## 9. Industria aseguradora

La inteligencia artificial se puede utilizar en la prospección de clientes para una variedad de pólizas de seguros, incluidos, entre otros, los de vida, salud, rentas vitalicias, incapacidad, automóvil, propietarios de viviendas, inquilinos, inundaciones, terremotos, incendios forestales, granizo, agricultura y reaseguros.

*Sushant K Singh y Muralidhar Chivukula (2020)*

El uso de la inteligencia artificial está transformando la gestión de riesgos y, en consecuencia, la industria del seguro está fundamentalmente afectada por este fenómeno<sup>224</sup>. El sector de seguros, por naturaleza, ha sido una industria históricamente basada en el uso intensivo de datos. Las compañías de seguros gestionan grandes cantidades de datos de sus clientes y su negocio se basa en el análisis de dichos datos para comprender y evaluar eficazmente el riesgo. No es extraño, por lo

tanto, que la inteligencia artificial tenga un amplio impacto a lo largo de la cadena de valor de la industria aseguradora, desde la suscripción y la gestión de reclamaciones hasta la distribución y el servicio al cliente, y también en la gestión de activos. En este capítulo se revisan los efectos que la aplicación de herramientas de AI está teniendo en tareas asociadas con la gestión contractual y documental, la detección de fraudes, la distribución y el marketing, así como en aspectos como la

<sup>224</sup> Véase Aziz y Dowling (2018).

evaluación del riesgo asegurado,  
el análisis del comportamiento

de los asegurados, o el  
desarrollo de productos.

## 9.1. GESTIÓN CONTRACTUAL Y DOCUMENTAL

En el ámbito de gestión de contratos y la atención administrativa a clientes, las herramientas de AI facilitan las tareas relacionadas con cambios de datos en los contratos, organización de correos electrónicos recibidos de los clientes, así como en las respuestas a las solicitudes de los asegurados sobre aspectos contractuales, reclamaciones y otras cuestiones. Además, tareas administrativas simples y de bajo valor, como la reorganización de datos, pueden automatizarse y llevarse a cabo mediante herramientas de inteligencia artificial. Una arquitectura de datos mejorada para poder aplicar herramientas de AI en un contexto de creciente cantidad de datos disponibles a través de dispositivos de Internet de las Cosas (IoT), redes sociales y otras fuentes, así como una mejor comprensión de las leyes, las regulaciones y la normas de diferentes países y regiones,

gracias al uso de AI, permite garantizar el cumplimiento de los requisitos legales.

Adicionalmente, la gestión automatizada de reclamaciones mediante algoritmos de AI reduce el ciclo de vida de liquidación y aumenta la precisión de los pagos al tiempo que, al mejorar la detección automatizada de reclamaciones fraudulentas, reduce las posiciones con pérdidas<sup>225</sup>. El reconocimiento de imágenes mediante AI permite a las compañías aseguradoras agilizar determinadas tareas. Por ejemplo, los peritos de seguros que antes evaluaban fotografías manualmente para determinar el nivel de daño o pérdida se apoyan hoy en herramientas AI para agilizar el proceso y proporcionar una estimación inmediata del coste de las reparaciones, lo que acelera el procedimiento de reclamaciones.

<sup>225</sup> Véase Eling *et al.* (2021).

## 9.2. DETECCIÓN DE FRAUDES

El fraude de seguros es un fenómeno complejo cuyos resultados son casi imposibles de verificar con absoluta certeza, aunque se lleven a cabo investigaciones exhaustivas que, además, requieren tiempo y esfuerzo considerable. Las herramientas de inteligencia artificial están siendo aplicadas de forma creciente en esta función. Estas herramientas incluyen algoritmos de aprendizaje supervisados<sup>226</sup>, no supervisados<sup>227</sup> y modelos semi-supervisados. Se ha argumentado que el aprendizaje supervisado no puede aprender variables latentes tan complejas como las que subyacen en el caso del fraude en el seguro, lo que dificulta el uso de estos modelos en este ámbito, por ello las técnicas de inteligencia artificial basadas en aprendizaje profundo no supervisado han aumentado significativamente<sup>228</sup>. El aprendizaje semi-supervisado permite aprovechar las grandes cantidades de datos disponibles en muchos casos no etiquetados

en combinación con conjuntos de datos etiquetados normalmente de menor dimensión. En el ámbito de la predicción del fraude se están utilizando algoritmos de *logistic regression*, *support vector machine* y *Naïve Bayes* en el análisis de fraudes, por ejemplo, para el caso del seguro de automóviles<sup>229</sup>. En otras ocasiones, el origen del fraude se encuentra en un ámbito distinto del asegurado. Por ejemplo, en las pólizas de seguro de salud que brindan un reembolso completo de los gastos médicos o tratamientos gratuitos para el beneficiario, el fraude (en forma de facturación por servicios no prestados, cargos excesivos por servicios prestados y prestación de servicios médicamente innecesarios) puede proceder de los proveedores de servicios médicos. Estos comportamientos provocan un aumento innecesario en los costes de los servicios y una disminución en calidad de los servicios sanitarios. Las compañías aseguradoras

<sup>226</sup> Véase Khatri et al. (2020).

<sup>227</sup> Véase Srivastava y Salakhutdinov (2014).

<sup>228</sup> Véase Ekin et al. (2018 y 2019), Zafari y Ekin (2019) o Gomes et al. (2021).

<sup>229</sup> Véase Aslam et al. (2022).

están utilizando algoritmos de aprendizaje supervisado<sup>230</sup> y no supervisado<sup>231</sup>, así como redes

neuronales<sup>232</sup> para la detección de este tipo de fraudes.

### 9.3. MARKETING Y DISTRIBUCIÓN

Por lo que respecta a la distribución de productos de seguro, la combinación de técnicas de la inteligencia artificial orientadas a la generación de lenguaje natural, al análisis predictivo y al uso de motores de recomendación, está apoyando a los agentes de seguros ofreciendo ya conocimientos avanzados útiles en el proceso de venta. Por ejemplo, los *asistentes de ventas virtuales inteligentes* incorporan ahora estrategias de venta cruzada. Basándose en el análisis de datos, la AI puede mejorar la consulta al cliente y las recomendaciones de productos y, además, la oferta de coberturas puede hiper-personalizarse. Además, estos algoritmos permiten una gestión proactiva de las relaciones con los clientes y mejores servicios postventa

aumentando la transparencia en la relación con el asegurado<sup>233</sup>.

En el contexto del marketing, las herramientas de AI se están aplicando en la investigación de mercados y clientes para el desarrollo de productos, análisis de grupos objetivo, diseño de publicidad y elaboración de planes de marketing y comunicación. A tal fin, estos algoritmos están ayudando en el análisis predictivo y detección de patrones, así como en la predicción del valor de vida del cliente y en la segmentación de clientes para proporcionar estrategias de comunicación personalizadas. La AI permite el desarrollo de estrategias de marketing sofisticadas, por ejemplo, mediante el diseño de marketing de eventos en vivo

<sup>230</sup> Véase Dua y Bais (2014) y Gupta *et al* (2021).

<sup>231</sup> Véase Settipalli y Gangadharan (2023).

<sup>232</sup> Véase Obodoekwe y Van der Haar (2019), Shamitha y Ilango (2020) y Matschak *et al.* (2022).

<sup>233</sup> Véase Eling *et al.* (2021) y Zakirxodjaeva (2023).



para mejorar la experiencia del cliente y las tasas de respuesta<sup>234</sup>. El análisis de las preferencias en el comportamiento de compra del asegurado mediante AI se está utilizando para la identificación de acciones de venta cruzada potencial y para proponer productos específicos<sup>235</sup>.

El papel de los agentes y corredores de seguros también se verá significativamente alterado por la penetración masiva de las herramientas de inteligencia artificial. Cabe esperar una reducción en el número de agentes de seguros

operando en la industria, así como una dependencia mayor en la tecnología de aquellos que permanezcan. Y, en general, un aumento en la productividad. La misión fundamental de los agentes y corredores de seguros no cambiará, pero su valor añadido dependerá de su capacidad para ayudar a los potenciales asegurados en la combinación de una más amplia gama de pólizas de cobertura de riesgos relativos a salud, vida, movilidad, propiedad personal y residencial mediante el uso de asistentes personales inteligentes<sup>236</sup>.

## 9.4. EVALUACIÓN DE RIESGOS

Por lo que respecta a la evaluación de riesgos y la fijación de precios, las herramientas de AI pueden automatizar los procesos de gestión de solicitudes, suscripción y la evaluación de riesgos. Las decisiones sobre los precios adecuados para la cobertura de los riesgos,

sobre el nivel de reservas y el nivel de capital se adoptan en condiciones de incertidumbre, en el que el enfoque bayesiano es recurrentemente utilizado, pudiendo aplicarse diversas técnicas de inteligencia computacional<sup>237</sup>. Entre las técnicas de AI más utilizadas

<sup>234</sup> Un ejemplo aplicación del algoritmo de clasificación del árbol de decisión bajo la guía de big data en la planificación del marketing de seguros puede verse en Long (2022).

<sup>235</sup> Véase Ayuso *et al.* (2019), Eling *et al.* (2021).

<sup>236</sup> Véase Balasubramanian *et al.* (2021).

<sup>237</sup> Véase Parodi (2012), Mamoudan *et al.* (2021) y Andry *et al.* (2022).

en la industria de seguros para este tipo de análisis destaca la máquina de vectores de soporte, el regresor de bosque aleatorio, el XGBoost y las redes neuronales profundas (DNNs), las redes recurrentes (RNNs) y las redes convolucionales (CNNs)<sup>238</sup>. El uso de mayor cantidad de datos permite la formación de grupos de riesgo de menor tamaño y más homogéneos, con lo que se consigue mejorar la precisión en la estimación de probabilidades de siniestro y de exposición a riesgos. Las técnicas de máquina de vectores de soporte, redes neuronales artificiales, regresores de *random forest*, regresión logística y otras se están utilizando en la predicción de ocurrencia de siniestros. En particular, se vienen utilizando herramientas de AI tanto en el caso de modelos de mortalidad para la valoración de seguros de vida<sup>239</sup>, como en el caso de estimación de riesgo enfermedad para los seguros sanitarios<sup>240</sup>, de la probabilidad de accidentes para el caso

seguro de automóviles<sup>241</sup>, de la probabilidad de accidentes y catástrofe para los seguros de propiedad<sup>242</sup>, la probabilidad de pérdida de trabajo para los seguros de desempleo<sup>243</sup>, etc.

En general, una mejor evaluación de los riesgos puede reducir las pérdidas que se derivan de situaciones de información asimétrica. En particular, la evaluación de más y mejores datos disponibles antes de la celebración del contrato puede reducir la *selección adversa* en seguros sanitarios y en seguros de robo, que se produce porque el *agente* (el asegurado) tiene mejor información sobre sus propios rasgos de riesgo y los de su entorno que el *principal* (la compañía aseguradora), que es quien determina el precio de la póliza. El comportamiento adverso de los asegurados es típico y aumenta el riesgo de inestabilidad en el mercado de seguros. Los asegurados de alto riesgo proporcionan

<sup>238</sup> Véase Shinde y Raut (2020).

<sup>239</sup> Véase Deprez *et al.* (2017), Riley (2020), Blier-Wong *et al.* (2020), Wüthrich (2020), Balona y Richman (2020), Gabrielli *et al.* (2020), Richman (2021 y 2022), Perla *et al.* (2021), Kiermayer (2022), Taha *et al.* (2022) y Lozano-Murcia *et al.* (2023).

<sup>240</sup> Véase Ul-Hassan *et al.* (2021), Pfitzenreuter y Lima (2021), Vandrangí (2022), Stern *et al.* (2022), Ramya y Deepa (2022) o Patidar y Dudi (2023).

<sup>241</sup> Véase Gao y Wüthrich (2018), Lorentzen y Mayer (2020) y Gao *et al.* (2022).

<sup>242</sup> Véase Kuo y Lupton (2020).

<sup>243</sup> Véase Senousy *et al.* (2020).

deliberadamente información falsa a la aseguradora para escapar de primas más altas o para evitar ser excluidos de la elegibilidad. Además, en productos de seguros basados en el uso, el riesgo de manipulación (*moral hazard*) puede también reducirse gracias a una mejor alineación de intereses en términos de comportamiento *ex post* relevante para la asunción del riesgo y los esfuerzos de prevención. Como consecuencia, la inteligencia artificial puede facilitar a las

compañías de seguros cotizar de forma más precisa y competitiva pólizas y hacerlo en cuestión de minutos. También es posible desarrollar de forma automatizada recomendaciones personalizadas para que los clientes mejoren la prevención de pérdidas<sup>244</sup>. Mediante herramientas de AI es posible identificar de antemano los riesgos potenciales, informar al asegurado y ayudarlo para que adopte las precauciones necesarias<sup>245</sup>.

## 9.5. ANÁLISIS DE COMPORTAMIENTO DE LOS ASEGURADOS

Las compañías de seguros están particularmente interesadas en la predicción del comportamiento de los asegurados antes y después de la contratación de las pólizas. Antes de la contratación, el ámbito de preocupación es el problema de *selección adversa* que se deriva de la asimetría de la información. La predicción de este tipo de comportamientos está siendo objeto de análisis creciente bajo herramientas de inteligencia artificial<sup>246</sup>. Y tras la contratación

de las pólizas, el comportamiento de los asegurados también es relevante. Por ejemplo, en el ámbito de los seguros de vida las compañías del sector se interesan por los motivos que generan la cancelación de pólizas y el ejercicio de determinadas opciones de salida del contrato. Este comportamiento está asociado con aspectos relativos al envejecimiento, la situación económica o las características del contrato. Las opciones a favor

<sup>244</sup> Véase Kelley *et al.* (2018) y Eling *et al.* (2021).

<sup>245</sup> Véase Erem-Ceylan (2022).

<sup>246</sup> Véase Islam *et al.* (2021).

del asegurado se refieren a las de rescate, renta garantizada, de cambio de producto, de pago, etc. Por varias razones, la opción de rescate (que garantiza la equidad del contrato al permitir al tomador recuperar su reserva, una vez deducida una penalización) tiene un impacto significativo en las carteras de las compañías de seguros. Dado que la opción de rescate puede ejercerse en cualquier momento antes del vencimiento (es como una opción americana) la aseguradora está expuesta

de forma continua al riesgo de rescate. El ejercicio de estas opciones afecta a la tarificación, las reservas, la rentabilidad, la liquidez, la gestión del riesgo y, por ende, a la propia solvencia de la aseguradora. En este ámbito se están aplicando métodos de inteligencia artificial, como Random Forest, Extreme Gradient Boosting (XGBoost), máquina de vectores soporte y regresión logística (LR) que parecen aportar resultados superiores en la predicción respecto del ejercicio de las opciones<sup>247</sup>.

## 9.6. DESARROLLO DE PRODUCTOS DE SEGURO

El análisis predictivo y la detección de patrones y anomalías basadas en herramientas AI permite diseñar seguros más innovadores, gracias a la posibilidad de obtener información a partir *big data* recopilados en tiempo real mediante dispositivos varios, incluidos aquellos vinculados con IoT. La penetración y conexión entre dispositivos ya hoy instalados (como GPS y monitores en automóviles,

monitores de actividad física, asistentes domésticos, teléfonos inteligentes, relojes inteligentes, en electrodomésticos, etc.) y aquellos que se instalarán mañana (en ropa, gafas, etc.) generarán una enorme cantidad de nuevos datos. La disposición de parte de estos datos por parte de las aseguradoras permitirá una mejor comprensión de sus clientes y la posibilidad de ofrecer nuevas categorías de productos, precios más

<sup>247</sup> Véase Guillen *et al.* (2011), Hu *et al.* (2021), Loisel *et al.* (2021), Azzone *et al.* (2022) y Aleandri *et al.* (2022).

personalizados y una prestación de servicios casi en tiempo real<sup>248</sup>.

Además, con la implementación de la inteligencia artificial nuevos riesgos se vuelven asegurable aumentando el alcance de la industria. La suscripción automatizada y continua reduce los costes de transacción y permitirá la extensión del seguro *on-demand* para diversos activos y durante espacios temporales más cortos. Los productos de seguros basados en uso *-usage-based insurance* (UBI) en su expresión en inglés- son un reflejo de cómo las ofertas de seguros pueden adaptarse constantemente a los patrones de comportamiento de los individuos y permiten la existencia de microcoberturas antes inexistentes. Gracias a las herramientas de AI estos nuevos productos pueden ser diseñados, y se pueden evaluar los riesgos implicados para asignarles un precio<sup>249</sup>.

Del mismo modo, estos algoritmos permiten el establecimiento de servicios complementarios y nuevas fuentes de ingresos. En este sentido, las compañías de

seguros pueden plantearse la entrada en nuevos mercados mediante la colaboración con instituciones de otros sectores (conducción autónoma, consultas de salud en tiempo real, prestación de servicios de cuidado de personas mayores con *nanobots*, gestión de catástrofes naturales, creación ecosistemas de hogares inteligentes).<sup>250</sup>

Por otra parte, el uso masivo de inteligencia artificial en otros sectores, puede ser un nuevo origen de siniestros y las aplicaciones de inteligencia artificial tienen el potencial de transformar la naturaleza de los eventos cuyos riesgos pueden ser cubiertos. Con la AI el sector asegurador se enfrenta a un nuevo horizonte en el que trabajar en la cobertura de riesgos. Si aumenta la dependencia de muchas empresas respecto de sus sistemas de inteligencia artificial algunos riesgos podrían transformarse de alta frecuencia/baja gravedad a baja frecuencia/alta gravedad, ya que la mayoría de las pérdidas podrían ocurrir simultáneamente. Por otra parte, en la industria del seguro de automóviles y de salud el análisis de los riesgos

<sup>248</sup> Véase Balasubramanian *et al.* (2021).

<sup>249</sup> Véase Balasubramanian *et al.* (2021).

<sup>250</sup> Véase Lorenz *et al.* (2020), Eling *et al.* (2021).

puede complicarse. La irrupción del coche autónomo o de los *nanobots* sanitarios, dirigidos por algoritmos de AI, implica que el número de total siniestros puede reducirse, pero ahora los riesgos pueden proceder de actos de piratería, ya que la AI amplía la superficie de riesgo. Además, los siniestros pueden no proceder de la conducta humana, sino también de anomalías de software

de AI y puede comprometer la responsabilidad de las empresas que lo desarrollan, que también necesitarán cobertura de responsabilidades civiles<sup>251</sup>. Finalmente, la emergencia en el uso de herramientas de AI conlleva nuevos riesgos derivados de conductas de cibercrimen y, en consecuencia, la existencia de nuevos mercados de riesgos asegurables<sup>252</sup>.

## 9.7. ADOPCIÓN DE LA TECNOLOGÍA AI EN EL SECTOR DE SEGUROS

Como en otras industrias, en el sector asegurador la adopción de las herramientas de inteligencia artificial no está exenta de problemas. Diversos trabajos están analizando cuales son los factores que influyen en la intención de los empleados de la industria de seguros a la hora de adoptar aplicaciones de inteligencia artificial. Factores de dimensión tecnológica, organizativa y ambiental son objeto de análisis en este terreno. Algunos resultados indican que las variables tecnológicas

(ventaja relativa y complejidad) y ambientales (dinámica del mercado, apoyo regulatorio y presión competitiva) predicen significativamente la intención de comportamiento de los empleados en cuanto a la adopción de herramientas de inteligencia artificial. Los resultados indican también que el apoyo de la alta dirección es determinante, mientras que las competencias técnicas no parecen tener un impacto tan significativo en la intención de comportamiento<sup>253</sup>.

<sup>251</sup> Véase Biener *et al.* (2015).

<sup>252</sup> Véase Biener *et al.* (2015) y KPMG (2018), Talesh y Cunningham (2021) o Lior (2022).

<sup>253</sup> Véase Gupta *et al.* (2022a).

En definitiva, se espera que la inteligencia artificial contribuya a ampliar el alcance de la cobertura de riesgos, en nuevos clientes a los que se les puede facilitar el acceso a productos personalizados, así como ampliar la gama de riesgos para los que se dispone de cobertura de seguro, gracias a una mejor comprensión de los riesgos. Al tiempo, parece factible que el uso de algoritmos de inteligencia artificial consiga reducir el coste de la cobertura de riesgos, pudiendo además ofrecerse seguros más rentables mediante la automatización de tareas específicas, mejores evaluaciones de riesgos y la reducción del riesgo de selección adversa y de manipulación (*moral hazard*). De igual modo, la inteligencia artificial se muestra como un conjunto de técnicas con cierta capacidad para mitigar y prevenir riesgos, aportando nuevos conocimientos y habilitando sistemas de alerta temprana que permitan reducir las pérdidas<sup>254</sup>.

Sin embargo, la industria del seguro es un ejemplo de por qué existirá resistencia a la adopción de herramientas de AI por parte de algunos empleados y

también de por qué resulta poco probable que los mecanismos tradicionales de control estatales, como la legislación y las normas vinculantes, sirvan como una solución completa e inmediata para gestionar los posibles efectos perversos de la AI en la sociedad<sup>255</sup>. En este contexto, se ha destacado que además de promulgar normas legales, que incentiven a las aseguradoras a que adopten un enfoque centrado en las personas, en el diseño y uso de análisis de *big data* e inteligencia artificial en un marco de gobernanza de datos claro y eficaz, también resulta necesario implementar un proceso claro y responsable para explicar qué información se puede utilizar y cómo, así como empoderar a las personas cuyos datos va a ser utilizados mediante su participación en la determinación de la gestión de los datos personales. Además, las aseguradoras y los órganos de gobernanza, incluidos los reguladores y los formuladores de políticas, deben trabajar juntos para garantizar que los análisis de *big data* basados en inteligencia artificial que se desarrollen sean transparentes y precisos<sup>256</sup>.

<sup>254</sup> Véase Keller (2020).

<sup>255</sup> Véase Tournas y Bowman (2021).

<sup>256</sup> Véase Riyanti (2023).

## 9.8. PRINCIPIOS DE GOBERNANZA DE LA AI

Después de la publicación del Libro Blanco sobre Inteligencia Artificial (2020) y la Propuesta de un nuevo reglamento sobre AI de abril de 2021 de la Comisión Europea, la Autoridad Europea de Seguros y Pensiones de Jubilación (EIOPA), con la participación del Grupo Consultivo de Expertos de EIOPA sobre Ética Digital en los seguros, publicó en junio de 2021, un informe orientado a lograr una inteligencia artificial ética y fiable en el sector de los seguros europeo, que intenta ayudar a las compañías de seguros cuando implementan aplicaciones/sistemas de AI.

Los principios de gobernanza se articulan en torno a seis claves para una gobernanza eficaz:

- **Proporcionalidad.** Las medidas de gobernanza y control deben ser proporcionales al impacto de estos sistemas en los clientes y los negocios.
- **Equidad y no discriminación.** Las empresas deben diseñar sus sistemas de AI para lograr equidad procesal y distributiva que permita equilibrio entre las partes involucradas. Debe mitigarse la subjetividad y vigilar que la introducción de los datos en el sistema está libre de sesgos.
- **Transparencia y explicabilidad.** Las aseguradoras deben explicar a sus clientes cómo se obtienen los resultados derivados del uso de herramientas de AI. Los usuarios deben tener conocimiento de los datos de entrada, el significado de los resultados y las instrucciones para utilizar los sistemas de AI y sus limitaciones, especialmente en los casos de fijación de precios de seguros, detección de fraude y gestión de reclamaciones.
- **Supervisión humana.** Los sistemas diseñados por las aseguradoras deben permitir la supervisión humana y el personal responsable de supervisar los sistemas de AI debe estar adecuadamente capacitado.
- **Gestión de datos y mantenimiento de registros.** Debe evitarse la existencia de errores y sesgos en los datos de entrada utilizados en los sistemas de inteligencia artificial con objeto de mitigar el riesgo



de resultados defectuosos en los modelos de fijación de precios de las pólizas, en el diseño de productos y en la prevención de pérdidas. La gestión del mantenimiento de los registros de datos debe permitir su trazabilidad y auditabilidad.

- **Robustez y rendimiento.** Uso de sistemas de inteligencia artificial adecuados para su propósito y diseñados específicamente para las

tareas en cada caso relevantes. Garantía de la calibración, validación y reproducibilidad de los sistemas y realización de pruebas exhaustivas para lograr solidez y precisión, así como para confirmar la resistencia de las infraestructuras de TI ante potenciales ataques cibernéticos.



## REFERENCIAS

- Aleandri, M. (2017) "Modeling dynamic policyholder behaviour through machine learning techniques".
- Andry, J. F., Hartono, H., and Honni, A. C. (2022) "Data Set Analysis Using Rapid Miner to Predict Cost Insurance Forecast with Data Mining Methods", *Journal of Hunan University Natural Sciences*, 49(6).
- Aslam, F., Hunjra, A. I., Ftiti, Z., Louhichi, W., and Shams, T. (2022) "Insurance fraud detection: Evidence from artificial intelligence and machine learning", *Research in International Business and Finance*, 62: 101744.
- Ayuso, M., Guillen, M. and Perch Nielsen J. (2019) "Improving automobile insurance ratemaking using telematics: Incorporating mileage and driver behaviour data", *Transportation*, 46: 735-752.
- Aziz, S. y Dowling, M. (2019a) "Machine Learning and AI for Risk Management" en *Disrupting finance: FinTech and strategy in the 21st century*: 33-50. Lynn, T., Mooney, J. G., Rosati, P., y Cummins, M. (Editores). Springer Nature. Palgrave. MacMillan.
- Azzone, M., Barucci, E., Moncayo, G. G., and Marazzina, D. (2022) "A machine learning model for lapse prediction in life insurance contracts", *Expert Systems with Applications*, 191: 116261.
- Balasubramanian, R, Libarikian, A. and McElhane, D. (2021) "Insurance 2030—The impact of AI on the future of insurance", *McKinsey & Company*. 12 de marzo.
- Balona, C. and Richman, R. (2020) "The Actuary and IBNR Techniques: A Machine Learning Approach", Disponible en SSRN 3697256.
- Biener, C., Eling, M., and Wirfs, J. H. (2015) "Insurability of cyber risk: An empirical analysis", *The Geneva Papers on Risk and Insurance—Issues and Practice*, 40 (1): 131-158.

- Blier-Wong, C., Cossette, H., Lamontagne, L., and Marceau, E. (2020) "Machine learning in P&C insurance: A review for pricing and reserving", *Risks*, 9 (1): 4.
- Cankett, M. and Liddy, B. (2021) "Artificial Intelligence and the Insurance Industry Building trustworthy AI systems in the Insurance Sector from the EIOPA AI guidelines". *Deloitte. Blog: Audit and Assurance*. 18 de octubre.
- Deprez, P., Shevchenko, P. V., and Wüthrich, M. V. (2017) "Machine learning techniques for mortality modelling", *European Actuarial Journal*, 7: 337-352.
- Dua, P. and Bais, S. (2014) "Supervised learning methods for fraud detection in healthcare insurance", *Machine Learning in Healthcare Informatics*: 261-285.
- Ekin, T., Ieva, F., Ruggeri, F., and Soyer, R. (2018) "Statistical medical fraud assessment: Exposition to an emerging field", *International Statistical Review*, 86 (3): 379-402.
- Ekin, T., Lakomski, G., and Musal, R. M. (2019) "An unsupervised Bayesian hierarchical method for medical fraud assessment", *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 12(2): 116-124.
- Eling, M., Nuessle, D., and Staubli, J. (2021) "The impact of artificial intelligence along the insurance value chain and on the insurability of risks", *The Geneva Papers on Risk and Insurance-Issues and Practice*: 1-37.
- Erem-Ceylan, I. (2022) "The Effects of Artificial Intelligence on the Insurance Sector: Emergence, Applications, Challenges, and Opportunities", *The Impact of Artificial Intelligence on Governance, Economics and Finance*, 2: 225-241.
- Gabrielli, A., Richman, R., and Wüthrich, M. V. (2020) "Neural network embedding of the over-dispersed Poisson reserving model", *Scandinavian Actuarial Journal*, 1: 1-29.
- Gao, G., Meng, S. and Wüthrich, M. V. (2022) "What can we learn from telematics car driving data: A survey", *Insurance: Mathematics and Economics*, 104: 185-199.
- Gao, G. and Wüthrich, M. V. (2018) "Feature extraction from telematics car driving heatmaps", *European Actuarial Journal*, 8: 383-406.
- Gomes, C., Jin, Z., and Yang, H. (2021) "Insurance fraud detection with unsupervised deep learning", *Journal of Risk and Insurance*, 88(3): 591-624.
- Guillen, M., Perez, A. M., and Alcañiz, M. (2011) "A logistic regression approach to estimating customer profit loss due to lapses in insurance", *Document de Treball, XREAP*, 13.
- Gupta, S., Ghardallou, W., Pandey, D. K., and Sahu, G. P. (2022a) "Artificial intelligence adoption in the insurance industry: Evidence using the technology-organization-environment framework", *Research in International Business and Finance*, 63: 101757.
- Gupta, R. Y., Mudigonda, S. S., Baruah, P. K., and Kandala, P. K. (2021) "Markov model with machine learning integration for fraud detection in health insurance". *arXiv*: 2102.10978.
- Ho, C.W.L., Ali, J. and Caals, K. (2020) "Ensuring trustworthy use of artificial intelligence and big data analytics in health insurance", *Bulletin of the World Health Organization*, 98 (4): 263.
- Hu, S., O'Hagan, A., Sweeney, J., and Ghahramani, M. (2021) "A spatial machine learning model for analysing customers' lapse behaviour in life insurance", *Annals of Actuarial Science*, 15(2): 367-393.



- Islam, M. R., Liu, S., Biddle, R., Razzak, I., Wang, X., Tilocca, P., and Xu, G. (2021) "Discovering dynamic adverse behavior of policyholders in the life insurance industry", *Technological Forecasting and Social Change*, 163: 120486.
- Keller, B. (2020) "Promoting responsible artificial intelligence in insurance", *Geneva Association-International Association for the Study of Insurance Economics*.
- Kelley, K.H., Fontanetta, L.M., Heintzman M., and Pereira, N. (2018) "Artificial intelligence: Implications for social inflation and insurance", *Risk Management and Insurance Review*, 21 (3): 373-387.
- Khatri, S., Arora, A., and Agrawal, A. (2020) "Supervised machine learning algorithms for credit card fraud detection: A comparison", en *10th International Conference on Cloud Computing, Data Science & Engineering*: 680-683.
- Kiermayer, M. T. (2022). *Machine and deep learning in present actuarial challenges*. Doctoral dissertation, Universität Ulm.
- KPMG (2018) *Neues Denken, Neues Handeln. Insurance Thinking Ahead: Versicherungen im Zeitalter von Digitalisierung und Cyber, Studienteil B: Cyber*.
- Kuo, K. and Lupton, D. (2020) "Towards explainability of machine learning models in insurance pricing", *arXiv*: 2003.10674.
- Lior, A. (2022) "Insuring AI: The role of insurance in artificial intelligence regulation", *Harvard Journal of Law and Technology*, 35 (2): 467-530.
- Loisel, S., Piette, P., and Tsai, C. H. J. (2021) "Applying economic measures to lapse risk management with machine learning approaches", *ASTIN Bulletin: The Journal of the IAA*, 51 (3): 839-871.
- Long, J. (2022) "Analysis of Insurance Marketing Planning Based on BD-Guided Decision Tree Classification Algorithm", *Security and Communication Networks*.
- Lorentzen, C. and Mayer, M. (2020) "Peeking into the black box: An actuarial case study for interpretable machine learning", disponible en SSRN 3595944.
- Lorenz, J. T., Deetjen, U. and Van Ouwerkerk, J. (2020). "Ecosystems in insurance: The next frontier for enhancing productivity", *McKinsey & Company*.
- Lozano-Murcia, C., Romero, F. P., Serrano-Guerrero, J., and Olivas, J. A. (2023) "A Comparison between Explainable Machine Learning Methods for Classification and Regression Problems in the Actuarial Context", *Mathematics*, 11(14): 3088.
- Mamoudan, M. M., Forouzanfar, D., Mohammadnazari, Z., Aghsami, A., and Jolai, F. (2021) "Factor identification for insurance pricing mechanism using data mining and multi criteria decision making", *Journal of Ambient Intelligence and Humanized Computing*: 1-20.
- Matschak, T., Prinz, C., Rampold, F., and Trang, S. (2022) "Show Me Your Claims and I'll Tell You Your Offenses: Machine Learning-Based Decision Support for Fraud Detection on Medical Claim Data", *Scholar Space*.
- Obodoekwe, N., and Van der Haar, D. T. (2019) "A comparison of machine learning methods applicable to healthcare claims fraud detection" en *Information Technology and Systems: Proceedings of ICITS*: 548-557. Springer International Publishing.
- Parodi, P. (2012) "Computational intelligence with applications to general insurance: a review: II. Dealing with uncertain knowledge", *Annals of Actuarial Science*, 6 (2): 344-380.



- Patidar, S., and Dudi, S. (2023) "Estimating Medical Insurance Cost using Linear Regression with HyperParameterization, Decision Tree and Random Forest Models" en *13th International Conference on Cloud Computing, Data Science & Engineering*: 504-508). Enero. IEEE.
- Perla, F., Richman, R., Scognamiglio, S., and Wüthrich, M. V. (2021) "Time-series forecasting of mortality rates using deep learning", *Scandinavian Actuarial Journal*, 7: 572-598.
- Pfutzenreuter, T. C. and Lima, E. P. (2021) "Machine learning in healthcare management for medical insurance cost prediction" en *XLI Encontro Nacional de Engenharia de Produção*.
- Prince, R.A. (2015) "How Artificial Intelligence Will Eliminate the Need For The Vast Majority Of Life Insurance Agents", *Forbes*. 12 de abril.
- Ramya, D., and Deepa, J. (2022) "Health Insurance Cost Prediction using Machine Learning Algorithms" en *International Conference on Edge Computing and Applications (ICECAA)*: 1381-1384). Octubre. IEEE.
- Richman, R. (2021) "AI in actuarial science—a review of recent advances—part 2", *Annals of Actuarial Science*, 15(2): 230-258.
- Richman, R. (2022) "Mind the gap—safely incorporating deep learning models into the actuarial toolkit", *British Actuarial Journal*, 27, e21.
- Riley, J. (2020) "AI and Machine Learning Usage in Actuarial Science", *Williams Honors College. Honors Research Projects*.
- Riyanti, R. (2023) "Legal status of artificial intelligence-based health insurance services: Challenges, opportunities for customer protection", *International Journal of Health Sciences*, 6 (2): 1023-1034.
- Senousy, Y., Shehab, A., Hanna, W. K., Riad, A. M., El-bakry, H. A., and Elkhamisy, N. (2020) "A Smart social insurance big data analytics framework based on machine learning algorithms", *Cybernetics and Information Technologies*, 20 (1): 95-111.
- Settipalli, L., and Gangadharan, G. R. (2023) "WMTDBC: An unsupervised multivariate analysis model for fraud detection in health insurance claims", *Expert Systems with Applications*, 215: 119259.
- Shamitha, S. K. and Ilango, V. (2020) "A time-efficient model for detecting fraudulent health insurance claims using artificial neural networks" en *International Conference on System, Computation, Automation and Networking (ICSCAN)*: 1-6. IEEE. Julio.
- Shinde, A., and Raut, P. (2020) "Comparative study of regression models and deep learning models for insurance cost prediction" en *Intelligent Systems Design and Applications: 18th International Conference on Intelligent Systems Design and Applications (ISDA)*, 1:1102-1111 Vellore (India). 6-8 de diciembre. Springer International Publishing.
- Singh, S. K. and Chivukula, M. (2020) "A commentary on the application of artificial intelligence in the insurance industry", *Trends in Artificial Intelligence*, 4 (1), 75-79.
- Srivastava, N. and Salakhutdinov, R. (2014) "Multimodal learning with deep Boltzmann machines", *Journal of Machine Learning Research*, 15(1): 2949-2980.
- Stern, A. D., Goldfarb, A., Minssen, T., and Price II, W. N. (2022) "AI insurance: how liability insurance can drive the responsible adoption of artificial intelligence in health care", *NEJM Catalyst Innovations in Care Delivery*, 3(4): CAT-21.
- Taha, A., Cosgrave, B., and McKeever, S. (2022) "Using feature selection with machine learning for generation of insurance insights", *Applied Sciences*, 12(6): 3209.



- Talesh, S. A. and Cunningham, B. (2021). "The Technologization of Insurance: An Empirical Analysis of Big Data an Artificial Intelligence's Impact on Cybersecurity and Privacy", *Utah Law Review*, 5: 967-1027.
- Tournas, L. N. and Bowman, D. M. (2021) "AI Insurance: Risk Management 2.0", *IEEE Technology and Society Magazine*, 40 (4): 52-56.
- Ul-Hassan, C. A., Iqbal, J., Hussain, S., AlSalman, H., Mosleh, M. A., and Sajid Ullah, S. (2021) "A computational intelligence approach for predicting medical insurance cost", *Mathematical Problems in Engineering*, 2021: 1-13.
- Vandrangi, S. K. (2022) "Predicting the Insurance Claim By Each User Using Machine Learning Algorithms", *Journal of Emerging Strategies in New Economics*, 1(1): 1-11.
- Weber, M., Limmer, N., and Weking, J. (2022) "Where to start with AI? Identifying and prioritizing use cases for health insurance", *Proceedings of the 55th Hawaii International Conference on System Sciences*: 3962-3971.
- Wüthrich, M. V. (2020) "Bias regularization in neural network models for general insurance pricing", *European Actuarial Journal*, 10 (1): 179-202.
- Zafari, B. and Ekin, T. (2019) "Topic modelling for medical prescription fraud and abuse detection", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 68(3): 751-769.
- Zakirxodjaeva, S. A. (2023) "Methods of Using Artificial Intelligence in Insurance Companies and its Promising Directions", *International Journal on Economics, Finance and Sustainable Development*, 5(6): 9-15.



## 10. Bancos centrales, supervisores del mercado y autoridades de competencia

“Los macrodatos se utilizan por parte de los bancos centrales con aplicaciones de aprendizaje automático en diversos ámbitos, como la investigación macroeconómica, la política monetaria y el análisis de la estabilidad financiera. También afirman utilizar big data para la supervisión y la regulación (aplicaciones suptech y regtech)”

*Sebastian Doerr, Leonardo Gambacorta y  
José María Serena Garralda (2021).*

No sólo el sector privado tiene interés en adoptar tecnologías de inteligencia artificial, también las entidades públicas y otros organismos institucionales en el ámbito financiero están utilizando este tipo de herramientas. Por ejemplo, los Bancos Centrales pueden utilizar software AI para mejorar su comprensión de la evolución macroeconómica y financiera, para hacer más eficiente su intervención de política monetaria en el mercado,

o para mejorar su evaluación sobre riesgo sistémico de la industria financiera. Del mismo modo, las autoridades financieras pueden utilizar sistemas AI para mejorar la supervisión micro prudencial y macro prudencial y la realización de pruebas de resistencia (*stress test*), así como para mejorar su evaluación de la gobernanza y la cultura del riesgo en las entidades financieras supervisadas.

Por otra parte, los supervisores



de los mercados de valores pueden encontrar en los sistemas de AI un aliado para mejorar la detección de comportamientos de manipulación de precios y prácticas de *insider trading*.

Finalmente, las autoridades que velan por la competencia en el mercado comienzan a utilizar herramientas de

inteligencia artificial para detectar comportamientos colusorios. Si bien, se pone también de manifiesto que los algoritmos de inteligencia artificial pueden servir para generar y enmascarar comportamientos anticompetitivos.

En este capítulo se comentan algunas de estas cuestiones.

## 10.1. BANCOS CENTRALES

Actualmente se está produciendo una transformación de la forma en la operan los bancos centrales como consecuencia del hecho de que las capacidades de procesamiento y almacenamiento han aumentado espectacularmente y están desarrollándose aplicaciones basadas en el aprendizaje automático, que combinan elementos de estadística computacional, optimización matemática, reconocimiento de patrones, y análisis predictivo. Además, hay una cantidad cada vez mayor de datos granulares (*big data*) que permiten alimentar estos nuevos modelos<sup>257</sup>.

La previsión de variables macroeconómicas desempeña un papel importante en el diseño de las políticas públicas. Para determinar la política monetaria óptima los bancos centrales necesitan predecir variables como la tasa de inflación y la tasa de crecimiento de la renta y de la producción, interpretar esta información en relación con los objetivos encomendados (estabilidad de precios y estabilidad financiera), y diseñar las medidas adecuadas (estableciendo los tipos de interés oficiales y la cantidad de dinero en el sistema). Los modelos macroeconómicos utilizados se han enfrentado tradicionalmente

<sup>257</sup> Véase Einav y Levin (2013) y Bholat (2015).

a una disyuntiva entre un enfoque basado en la teoría económica y otro enfoque basado en los datos<sup>258</sup>. En la actualidad, la creciente popularidad de las herramientas de aprendizaje automático en la predicción económica proviene de su capacidad para descubrir patrones complejos en datos que no han sido preespecificados *a priori*. Esta flexibilidad contrasta fuertemente con los métodos de previsión utilizados tradicionalmente, cuya estructura lineal preestablecida era un punto débil. Las herramientas de inteligencia artificial utilizadas en los modelos macroeconómicos son diversas, incluyendo redes neuronales artificiales, modelos basados en árboles, máquinas de vectores de soporte, etc.<sup>259</sup>. Una opción habitual en la modelización y estimación de parámetros macroeconómicos es la aplicación de redes neuronales, debido a su capacidad de aproximar cualquier relación lineal y no lineal con un grado razonable de precisión<sup>260</sup>. Estos algoritmos de redes neuronales no imponen ninguna forma funcional de relación entre las variables explicativas

y el objetivo de previsión. En su lugar, la forma funcional es el resultado del algoritmo de la red, aunque las variables incluidas en el modelo continúan estando determinadas por el investigador. Los algoritmos de aprendizaje automático imponen relativamente poca estructura *a priori* y se añaden relaciones para mejorar la capacidad explicativa. Por ejemplo, el número de nodos de un árbol de decisión determina su complejidad y su capacidad para explicar los datos a los que se enfrenta. En los algoritmos de *machine learning* se aprovecha un conjunto de datos de muestra para entrenar el modelo. Si el modelo se entrena durante demasiado tiempo con datos de muestra o si el modelo es demasiado complejo, puede comenzar a aprender del "ruido" (información irrelevante dentro del conjunto de datos). Cuando el modelo memoriza el ruido y se ajusta demasiado al conjunto de entrenamiento, el modelo se "sobreajusta" y no puede generalizar bien los nuevos datos, lo que le impide realizar correctamente las tareas de clasificación o predicción para

<sup>258</sup> Básicamente entre los modelos del tipo VAR (Vector Autoregression) sin restricciones, que se basan en la abundancia de datos, y los modelos de equilibrio general dinámico estocástico (DSGE), que suelen calibrarse con estimaciones de parámetros estructurales obtenidas de otros estudios. Véase Genberg y Karagedikli (2021).

<sup>259</sup> Véase Doerr *et al.* (2021).

<sup>260</sup> Véase Kasai y Naraidoo (2011) o Hinterlang (2020).

las que fue diseñado. El riesgo de sobreajuste puede reducirse limitando *ex ante* la complejidad del algoritmo utilizado en el análisis<sup>261</sup>. Para realizar el análisis se divide la muestra de datos en dos partes, una de entrenamiento (para generar un modelo explicativo eficaz) y otra de validación (para comprobar en qué medida el modelo explica datos que se supone que están impulsados por el mismo proceso de generación de datos, pero no por el mismo proceso). A continuación, se ajustan los parámetros fundamentales del algoritmo y se repite el proceso de entrenamiento y validación de forma iterativa<sup>262</sup>.

El entorno de *big data* y los métodos de aprendizaje automático han atraído recientemente mucha atención en el análisis económico porque tienen capacidad de aproximar funciones no lineales y, además, los parámetros se actualizan de forma adaptativa<sup>263</sup>. La mayoría de los bancos centrales han debatido formalmente el uso de *big data* dentro de su institución. Los resultados de

las investigaciones indican que, en el ámbito de la predicción macroeconómica, las herramientas de *machine learning* funcionan particularmente bien para procesar datos no tradicionales y no estructurados, capturar una fuerte no linealidad y mejorar la precisión de la predicción. Los modelos de aprendizaje profundo son adecuados para datos no tradicionales, mientras que los modelos de aprendizaje conjunto son preferidos para conjuntos de datos tradicionales<sup>264</sup>.

Además de llevar a cabo el diseño y la ejecución de la política monetaria, otra de las responsabilidades de los bancos centrales es preservar la estabilidad financiera del sistema, lo que implica realizar evaluación de riesgo crediticio del sistema bancario y estimación del riesgo sistémico. Debido a la fuerte interconexión entre las entidades bancarias y su papel esencial en la financiación de la economía, la quiebra de un banco puede provocar un efecto en cadena, que puede extenderse rápidamente y tener un impacto negativo en otros bancos (efecto

<sup>261</sup> Véase Chakraborty y Joseph (2017) y Genberg y Karagedikli (2021).

<sup>262</sup> Véase Athey e Imbens (2019).

<sup>263</sup> Véase Maehashi y Shintani (2020).

<sup>264</sup> Véase Desai (2023).

sistémico). En relación con la supervisión del sistema bancario, se ha comparado el rendimiento estadístico de un modelo simple como la regresión logística (Logit), tradicionalmente utilizada, con otros modelos la regresión logística penalizada por LASSO, el árbol de clasificación y regresión (CART), el bosque aleatorio, XGBoost y las redes neuronales profundas. Siguiendo el proceso implementado para la validación supervisora de los sistemas basados en calificaciones internas (IRB), los resultados indican beneficios del uso del *machine learning* en términos de poder predictivo, tanto en clasificación como en calibración. Esta ventaja de los algoritmos de inteligencia artificial podría traducirse en ahorros en capital regulatorio.<sup>265</sup> En otros entornos, los resultados de las investigaciones han mostrado que los métodos de red neuronal artificial y K-vecino más cercano son más precisos que dos enfoques estadísticos tradicionales (análisis discriminante y regresión logística) en la predicción de quiebras bancarias<sup>266</sup>. Asimismo, algunos otros resultados empíricos

muestran que la herramienta de red neuronal es un método competitivo entre los existentes para evaluar la probabilidad de quiebras bancarias, especialmente para reducir la tasa de clasificación errónea de tipo I (falso positivo).<sup>267</sup> Y en otros contextos, la investigación indica que aunque las arquitecturas de redes neuronales MLP y LVQ parecen ser los modelos de predicción más exitosos en los experimentos, el rendimiento de predicción de las SVM es comparablemente satisfactorio<sup>268</sup>. Adicionalmente, algunos estudios proporcionan evidencia en el sentido que XGBoost supera en poder predictivo de quiebra bancaria a otros modelos de la misma familia, incluido el impulso de gradiente y AdaBoost<sup>269</sup>.

Por otra parte, para la evaluación del riesgo de estabilidad del sistema, los bancos centrales pueden ahora, además, acceder y recopilar gran cantidad de datos de numerosas fuentes de datos, lo que sitúa también en este ámbito en el terreno de *big data*. Los datos útiles disponibles no sólo

<sup>265</sup> Véase Alonso y Carbo (2020 y 2021).

<sup>266</sup> Véase Le y Viviani (2018).

<sup>267</sup> Véase Bell *et al.* (1990) y Tam y Kiang (1990).

<sup>268</sup> Véase Boyacioglu *et al.* (2009) y Ecer (2013).

<sup>269</sup> Véase Pham y Ho (2021).



son mayores en volumen, sino que también contienen información más granular, vienen en una gama más amplia de formatos y se actualizan con mayor frecuencia. Algunos ejemplos clave son el acceso a los datos sobre mercados *over-the-counter* de derivados, garantías depositadas previamente por las entidades financieras (*colateralización*) en las transacciones ejecutadas fuera de las bolsas de valores o bien a través de cámaras de compensación, información detallada de sus balances, etc. Del mismo modo los bancos centrales, tienen ahora a su disposición un conjunto cada vez mayor de datos microeconómicos, desde la concesión de hipotecas y uso de tarjetas de crédito, sentimientos de mercado incluidos en las noticias, percepción de riesgos en las redes sociales, etc.

Las herramientas de inteligencia artificial también están siendo aplicadas en la realización de las pruebas de esfuerzo o *stress test*. En particular, los modelos entrenados por SVM muestran una capacidad significativa para pronosticar correctamente los resultados en simulaciones de tensión de solvencia. Este enfoque ofrece una interpretación geométrica simple de los resultados, brindando la capacidad de medir la distancia hasta el incumplimiento para instituciones de las que se predijo que serían solventes. Así, las instituciones bancarias no sólo reciben una clasificación binaria como "solventes" o "insolventes", sino que también pueden clasificarse según su distancia o margen de seguridad frente a la insolvencia<sup>270</sup>.

## 10.2. SUPERVISORES DE LOS MERCADOS DE VALORES

Los supervisores de los mercados de valores pueden encontrar en el uso de las herramientas de inteligencia artificial un aliado útil en la detección de conductas de

manipulación de precios y de uso de información confidencial (*insider trading*) aprovechando su capacidad para reconocer patrones y detectar anomalías. Pero, al mismo tiempo, los

<sup>270</sup> Véase Gogas et al. (2018).

supervisores se enfrentan a nuevos desafíos que son consecuencia de los propios avances en la inteligencia artificial.

Los conceptos jurídicos establecidos para hacer frente a las conductas indebidas, asignando responsabilidad a comportamiento de abuso de mercado, están bajo tensión con la aplicación de la inteligencia artificial. Un ejemplo de lo que está sucediendo se encuentra en el ámbito de la negociación por cuenta propia de las entidades. Desde hace tiempo los operadores humanos no son el epicentro de los mercados financieros. Los ordenadores que ejecutan programas algorítmicos se han convertido en los operadores principales, mientras que los humanos ejecutan un porcentaje cada vez menor de las operaciones. Con el advenimiento de las herramientas de inteligencia artificial, los algoritmos tradicionales basados en reglas preestablecidos han quedado obsoletos. A los algoritmos basados en AI, se les asigna la tarea objetivo de maximizar la rentabilidad y se les deja que descubran la mejor estrategia para lograrlo. Estos algoritmos aprenden de decisiones anteriores, evalúan dinámicamente la nueva información y optimizan

sus soluciones para reflejar los nuevos datos. Posteriormente ejecutan complejas estrategias de negociación que, además, responden casi instantáneamente ante cambios en las condiciones del mercado. Una herramienta de AI puede ser entrenada para ser rentable, y para ello los desarrolladores instruyen la máquina para reconocer patrones de mercado. La máquina a medida que aprende asigna mayor ponderación a los factores que se correlacionan con la rentabilidad, pudiendo reconocer el volumen de negociación como uno de ellos. En algún momento, esto podría llevar al algoritmo a “reconocer” que las propias transacciones ejecutadas por el algoritmo aumentan el volumen de negociación y, como resultado, a crear estrategias de compras y ventas simultáneas para afectar a los precios. En definitiva, el algoritmo puede aprender a ejecutar una *wash sale*, que es un ejemplo clásico de conducta de manipulación. Debido a la capacidad de la inteligencia artificial de “aprender” sin instrucciones explícitas del desarrollador, puede ser difícil para los humanos concluir cómo estas herramientas adoptan las decisiones, incluso cuando éstas son potencialmente ilegales. En el caso del trading basado

en decisiones autónomas (que es diferente del *trading automatizado* basado en reglas), la cuestión es en qué medida el desarrollador del programa es culpable (responsable) de manipulación de mercado. Es una cuestión problemática ya que la figura legal de manipulación implica una conducta diseñada para engañar o defraudar a los inversores mediante el control o el impacto artificial en el precio de los valores. La manipulación exige intención en la conducta y, en el caso de un modelo de AI, podría no haber intervención humana en la toma de decisiones, siendo ésta atribuible exclusivamente al algoritmo<sup>271</sup>.

Por otra parte, teniendo en cuenta la capacidad superior de las herramientas de AI para reconocer patrones, no sólo es concebible que aprendan formas de manipulación formas clásicas, sino que también inventen otras formas de manipulación no evidentes ni reconocibles para los humanos<sup>272</sup>. En este contexto, la dificultad de evitar conductas de manipulación relacionada con el uso de algoritmos de AI plantea un desafío para la

eficacia del marco jurídico, la reputación de los reguladores y la viabilidad del mercado<sup>273</sup>.

Por lo que respecta a las conductas de *insider trading*, las normas del mercado definen el uso ilegal de información privilegiada como la compra o venta de valores por parte de personas en posesión de información material no pública o no disponible para el mercado. Este tipo de información puede referirse al informe financiero trimestral de una empresa, el lanzamiento de un nuevo producto, la consecución de un mandato o proyecto, un nuevo plan de adquisición, de fusión o de enajenación de activos, una pérdida inesperada, etc. Las empresas que cotizan en bolsas de valores están obligadas a anunciar sus declaraciones a través sistemas electrónico de divulgación de información. Además, los sistemas de vigilancia del mercado se centran en monitorizar y revisar las actividades comerciales en torno a eventos corporativos clave, como fusiones y adquisiciones, ofertas públicas de adquisición, informes de ganancias, regulaciones, etc. Cuando en los mercados de

<sup>271</sup> Véase Bathaee (2017).

<sup>272</sup> Véase Slemmer (2019).

<sup>273</sup> Véase Fletcher (2021).



valores se producen movimientos anormales de precios o volúmenes de negociación extraordinarios, el sistema de vigilancia ayuda a los funcionarios a comprender la actividad de negociación en el contexto del mercado y proporciona evidencia preliminar sobre prácticas comerciales ilegales, de forma que pueda iniciarse una investigación en profundidad sobre el caso. El uso ilegal de información privilegiada es un delito económico difícil de detectar en tiempo real y la identificación de operaciones sospechosas se basa en la clasificación de las transacciones según características como su tamaño, el precio de ejecución, la fecha de la transacción, etc. Detectar el uso ilegal de información privilegiada es difícil debido a la naturaleza compleja, no lineal y no estacionaria del mercado de valores. Las autoridades investigan las operaciones sospechosas obteniendo pruebas de diversas fuentes de datos, incluidas transacciones internas, pistas de inversores, testigos, informes de los medios, informantes y consultas informales. No es extraño que el uso de herramientas de inteligencia artificial en este ámbito no pare de crecer. La *minería de datos* de series temporales y

de datos procedentes de redes sociales es una metodología eficaz progresivamente utilizada para la detección de fraude financiero, incluyendo el uso de información privilegiada. Se han desarrollado también sistemas expertos basados en reglas para detectar actividades financieras ilícitas, con capacidad de reconstruir los patrones de transacciones financieras sospechosas de acuerdo con su esquema temporal, la existencia de los rendimientos anormales o la existencia de tergiversaciones de presentaciones y eventos. Se han utilizado, asimismo, árboles de decisión, redes neuronales y técnicas de agrupamiento de K-medias para encontrar actividades de uso de información privilegiada en los mercados de opciones y acciones y detectar en tiempo real transacciones sospechosas, utilizando el papel de los iniciados, el número de acciones negociadas y su registro temporal. También se han desarrollado herramientas que detectan y predicen el uso de información privilegiada ilegal de manera proactiva a partir de grandes fuentes heterogéneas de datos estructurados y no estructurados, utilizando un enfoque basado en el aprendizaje profundo combinado con procesamiento de señales

discretas en los datos de series temporales. Se han adoptado técnicas clásicas de aprendizaje supervisado como Random Forest, Naive Bayes (NB), Support Vector Machine (SVM) y Artificial Neural Networks (ANN) para detectar actividades transaccionales sospechosas. Del mismo modo, algunos estudios han aplicado un método de integración de Análisis de Componentes Principales

(PCA) y de Random Forest (RF) para identificar operaciones con información privilegiada en el mercado de valores. Y, en otros trabajos, se propone un enfoque combinado de GBDT (árbol de decisión de impulso de gradiente) y DE (evolución diferencial) para identificar actividades de uso de información privilegiada mediante el uso de datos de indicadores relevantes<sup>274</sup>.

### 10.3. AUTORIDADES DE DEFENSA DE LA COMPETENCIA

Los avances en la tecnología digital y en la inteligencia artificial han creado nuevas oportunidades para que las empresas privadas adopten formas sutiles y evasivas de comportamiento anticompetitivo. Pero estos mismos avances tecnológicos pueden también ayudar a los reguladores antimonopolio a mejorar su desempeño en la detección y respuesta a conductas privadas ilegales. Las autoridades antimonopolio están recurriendo cada vez más de algoritmos de aprendizaje

automático para supervisar el comportamiento del mercado<sup>275</sup>. La detección y la adquisición de fundamentos probatorios para iniciar investigaciones sobre leyes de competencia podrían depender en un futuro del uso de inteligencia artificial. En octubre de 2017, la Comisión Europea publicó una licitación de consultoría recabando opiniones sobre cómo la inteligencia artificial podría mejorar sus procesos de gestión de evidencia, redacción legal y su recopilación de inteligencia de mercado. Y posteriormente, en su

<sup>274</sup> Véase Donoho (2004), Diaz *et al.* (2011), Tamersoy *et al.* (2014), Fan *et al.* (2014), Golmohammadi *et al.* (2014), Kulkarni *et al.* (2017), Islam *et al.* (2018), Esen *et al.* (2019), Deng *et al.* (2019 y 2021), Seth y Chaudhary (2020) o Nti *et al.* (2020).

<sup>275</sup> Véase Coglianese y Lai (2022).

Libro Blanco sobre AI de febrero de 2020, volvió a expresar su intención de comprender cómo la AI puede dotar a las autoridades encargadas de hacer cumplir la ley de competencia con herramientas adecuadas<sup>276</sup>. Las autoridades *antitrust* encargadas de defender la competencia en los mercados (financieros y de bienes y servicios) están utilizando de forma creciente herramientas de inteligencia artificial para detectar prácticas anticompetitivas. Son varios los motivos por los que la AI puede resultar útil en este terreno<sup>277</sup>. En primer término, por eficiencia. Lo que ha venido a denominarse *antimonopolio computacional* puede reducir significativamente el tiempo y el esfuerzo necesarios para analizar un caso. Y, lo que es más importante, los tribunales y los organismos podrían tener más posibilidades de aplicar los principios jurídicos de forma coherente, incluso cuando los hechos son idiosincrásicos<sup>278</sup>. Aprovechando la capacidad de los algoritmos de AI, para reconocer anomalías y patrones de comportamiento, las autoridades pueden estimar con mayor precisión el punto de referencia

competitivo y así detectar cuánto se desvía el precio de mercado de este nivel. Además, dado que las autoridades deben desarrollar nuevas herramientas para poder proteger mejor a los consumidores y a la competencia de los comportamientos anticompetitivos, especialmente en el mundo digital, el uso de herramientas de inteligencia artificial por parte de las agencias encargadas de la supervisión podría ayudar a éstas a lograr un mejor entendimiento de los algoritmos que utilizan las empresas. En general, las agencias gubernamentales y los reguladores se encuentran en una enorme desventaja informativa al respecto, particularmente en relación con las empresas tecnológicas<sup>279</sup>.

Pero el creciente uso de las herramientas de inteligencia artificial en el ámbito de defensa de la competencia implica diversos desafíos:

- Se presenta un posible dilema, ya que la recopilación y el procesamiento de grandes cantidades de datos por

<sup>276</sup> Véase Von Bonin y Malhi (2020).

<sup>277</sup> Véase Lorenzoni (2022).

<sup>278</sup> Véase Lim (2021).

<sup>279</sup> Véase Treasury (2019).

parte de los algoritmos inteligentes pueden implicar un compromiso entre la protección de la competencia y la privacidad de los consumidores<sup>280</sup>.

- En una transición hacia la implementación exitosa de las leyes antimonopolio mediante algoritmos para la monitorización del mercado, las autoridades *antitrust* enfrentan desafíos institucionales como la necesidad de desarrollar su capacidad organizacional, evitar obstáculos legales y lograr confianza pública en esta estrategia<sup>281</sup>.
- Es preciso determinar el papel que debe darse a las herramientas informáticas en los procesos de toma de decisiones. Por ejemplo, estableciendo la fuerza probatoria procedente de estas herramientas y hasta qué punto es posible justificar decisiones basadas exclusivamente en los resultados obtenidos mediante algoritmos en las investigaciones anticompetitivas<sup>282</sup>.

En relación con la protección de la competencia en los mercados existe una cara menos positiva del uso de herramientas de inteligencia artificial. Se ha afirmado que tras el advenimiento de la inteligencia artificial veremos el fin de la competencia tal como la conocemos hasta ahora<sup>283</sup>. La propia implementación de algoritmos de AI por parte de las empresas puede facilitar comportamientos anticompetitivos. El fenómeno conocido como *colusión algorítmica*, es objeto de una enorme atención actualmente. Si las empresas del mercado bajo observación utilizan algoritmos de inteligencia artificial, el problema de la detección de conductas anticompetitivas se ve alterado. Una respuesta más rápida de las empresas modificando los precios ante cambios del mercado dificulta distinguir entre el comportamiento competitivo y los patrones de precios interdependientes. La aplicación de herramientas de redes neuronales artificiales por parte de las empresas y entidades financieras permite una respuesta en las políticas de

<sup>280</sup> Véase Šmejkal (2021).

<sup>281</sup> Véase Coglianesi y Lai (2022).

<sup>282</sup> Véase Schrepel (2021).

<sup>283</sup> Véase Ezrahi (2016) y Gurkaynak *et al.* (2018).

precios más rápida ante cambios en los costes y en la demanda<sup>284</sup>.

En este contexto, se están poniendo de manifiesto los límites que muestran las reglas actuales para responder a los tipos de comportamientos anticompetitivos que plausiblemente pueden surgir derivados del uso comerciales de la inteligencia artificial en el ámbito de la política comercial. Las grandes empresas, particularmente las grandes tecnológicas, disponen del poder y los recursos para recopilar, consolidar y controlar el recurso más valioso en el mundo digital actual: la información<sup>285</sup>. Una colusión tiene lugar cuando las empresas utilizan estrategias en la que se recompensa a una empresa por respetar el resultado infra-competitivo y la castiga por apartarse de él. Las tecnologías de AI pueden tener la capacidad de analizar y supervisar el mercado a una escala sin precedentes, examinando el comportamiento de los competidores y realizando cambios rápidos de precios para adaptarse al entorno. Los algoritmos de fijación de precios pueden aprender cómo implementar estrategias

colusorias basadas en esquemas de recompensa y castigo, que se traduzcan en precios infra-competitivos. Además, estos algoritmos podrían sostener resultados colusorios sin intervención humana. A diferencia de los cárteles tradicionales, no existe un acuerdo de colusión explícito. Los algoritmos en sí no están programados con la intención de coludir ni están sesgados para favorecer la formación de cárteles. Sin embargo, adoptan una estrategia colusoria mediante una toma de decisiones autónoma. Esto es lo que se conoce como *colusión tácita algorítmica*<sup>286</sup>.

Los estudiosos del derecho de la competencia están cada vez más preocupados por la capacidad de los *agentes autónomos basados en AI* a la hora de crear nuevas formas de acuerdos colusorios<sup>287</sup>. Se han destacado cuatro tipos de estrategias anticompetitivas en los que las herramientas de inteligencia artificial pueden tener un papel:

- La primera categoría –*Messenger*– se refiere al uso de ordenadores para ejecutar la

<sup>284</sup> Véase Beneke y Mackenrodt (2019).

<sup>285</sup> Véase Quaid (2021).

<sup>286</sup> Véase Gautier *et al.* (2020).

<sup>287</sup> Véase Mehra (2015), Ezrachi y Stucke (2017), Petit (2017) y Azzutti *et al.* (2021).

voluntad de los humanos en su intento de coludirse y restringir la competencia. En este escenario básico, los humanos están de acuerdo con el cártel y utilizan su ordenador para poner en práctica, supervisar y vigilar el cártel. Desde el punto de vista de la aplicación de la ley, en este caso el concepto jurídico de acuerdo puede aplicarse directamente, y los fiscales, con pruebas suficientes, no tendrán dificultad en condenar el uso de máquinas para facilitar el cártel.

- La segunda categoría –*Hub and Spoke*– se refiere al uso orquestado e intencionado de un algoritmo único para determinar el precio de mercado cobrado por un conjunto de empresas. Son las máquinas las que determinan los precios, pero una programación común garantiza un proceso de convergencia.
- En la tercera categoría, el *agente previsible*, presenta un escenario más complejo en el que los humanos diseñan unilateralmente la máquina para que ofrezca resultados predecibles y reaccione de una manera determinada ante las condiciones cambiantes del

mercado. En esta categoría cada operador desarrolla su máquina unilateralmente, pero con conocimiento del desarrollo las otras máquinas utilizadas por los competidores. La adopción de algoritmos similares por parte de toda la industria puede tener efectos anticompetitivos al crear una acción interdependiente que se traduce en un aumento de precios como resultado de una colusión tácita. En este escenario, la posibilidad de obtener pruebas sobre la intención de cambiar la dinámica del mercado resulta fundamental.

- La cuarta estrategia –*Ojo digital*– es la más compleja, ya que los competidores crean y utilizan unilateralmente algoritmos informáticos para maximizar beneficios. Las máquinas, mediante autoaprendizaje y experimentación, determinan de forma independiente los medios para alcanzar este objetivo, y ejecutan la estrategia óptima. En este caso no es fácil aplicar directamente conceptos jurídicos como la intención o el acuerdo, y la cuestión de la responsabilidad plantea un reto jurídico relevante<sup>288</sup>.

<sup>288</sup> Véase Azzutti et al. (2021).

En este contexto, se ha argumentado que la capacidad de las redes neuronales artificiales puede traducirse en una ampliación del entorno en el que es posible fijar un precio de maximización conjunta por parte de las empresas sin la necesidad de recurrir a comunicaciones abiertas. Y, por otra parte, el uso de redes neuronales artificiales puede relajar el tensionamiento de precios competitivos en el contexto de un oligopolio. La explicación es que una fuente común de inestabilidad del equilibrio en este tipo de

entorno son los cambios en la demanda que, en presencia de incertidumbre, pueden incentivar a las empresas a ajustar sus precios para no perder cuota de mercado. Ahora bien, los algoritmos pueden aprender a hacer predicciones precisas y evitar guerras de precios innecesarias<sup>289</sup>. En este sentido, los algoritmos pueden hacer que la colusión tácita sea más probable, estable y duradera al facilitar la detección y las represalias<sup>290</sup>.



## REFERENCIAS

- Alonso-Robisco, A. and Carbo, J. M. (2020) "Machine learning in credit risk: Measuring the dilemma between prediction and supervisory cost", *Documentos de Trabajo del Banco de España*, 2032.
- Alonso-Robisco, A. and Carbo, J. M. (2021) "Understanding the performance of machine learning models to predict credit default: a novel approach for supervisory evaluation", *Documentos de Trabajo del Banco de España*, 2105.
- Athey, S. and Imbens, G. (2019) "Machine Learning Methods Economists Should Know About". *arXiv.org Papers*:1903.10075.
- Azzutti, A., Ringe, W., and Stiehl, H. (2021) "Machine learning, market manipulation, and collusion on capital markets: why the "black box" matters", *University of Pennsylvania Journal of International Law*, 43(1), 79-136.
- Bathae, Y. (2017) "The artificial intelligence black box and the failure of intent and causation", *Harvard Journal of Law & Technology*, 31: 889.
- Bell, T. B., Ribar, G. S., and Verichio, J. (1990) "Neural nets versus logistic regression: A comparison of each model's ability to predict commercial bank failures", *Proceedings of the University of Kansas Symposium on Auditing Problems*.

<sup>289</sup> Véase Beneke y Mackenrodt (2019).

<sup>290</sup> Véase Rab (2019).

- Beneke, F. and Mackenrodt, M. O. (2019) "Artificial intelligence and collusion", *IIC-International Review of Intellectual Property and Competition Law*, 50: 109-134.
- Bholat, D. (2015) "Big data and central banks", *Bank of England Quarterly Bulletin*, 55(1):86-93.
- Boyacioglu, M. A., Kara, Y., and Baykan, Ö. K. (2009) "Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey", *Expert Systems with Applications*, 36(2): 3355-3366.
- Chakraborty, C. and Joseph, A. (2017) "Machine learning at central banks", *Staff Working Paper Bank of England*, 674.
- Coglianesi, C. and Lai, A. (2022) "Antitrust by Algorithm. Stanford Computational Antitrust", *All Faculty Scholarship*. 2755.
- Deng, S., Wang, C., Fu, Z., and Wang, M. (2021) "An intelligent system for insider trading identification in Chinese security market", *Computational Economics*, 57, 593-616.
- Deng, S., Wang, C., Wang, M., and Sun, Z. (2019) "A gradient boosting decision tree approach for insider trading identification: An empirical model evaluation of China stock market", *Applied Soft Computing*, 83: 105652.
- Desai, A. (2023) "Machine Learning for Economics Research: When What and How?", *arXiv*: 2304.00086.
- Diaz, D., Theodoulidis, B., and Sampaio, P. (2011) "Analysis of stock market manipulations using knowledge discovery techniques applied to intraday trade prices", *Expert Systems with Applications*, 38(10): 12757-12771.
- Doerr, S., Gambacorta, L., and Garralda, J. M. S. (2021) "Big data and machine learning in central banking", *BIS Working Papers*, 930.
- Donoho, S. (2004) "Early detection of insider trading in option markets" en Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: 420-429.
- Ecer, F. (2013) "Comparing the bank failure prediction performance of neural networks and support vector machines: The Turkish case", *Economic research-Ekonomska istraživanja*, 26(3): 81-98.
- Einav, L. and Levin, J. D. (2013) "The data revolution and economic analysis", *National Bureau of Economic Research Working Paper*, 19035.
- Esen, M. F., Bilgic, E., and Basdas, U. (2019) "How to detect illegal corporate insider trading? A data mining approach for detecting suspicious insider transactions", *Intelligent Systems in Accounting, Finance and Management*, 26(2): 60-70.
- Ezrachi, A. (2016) *Virtual competition: The promise and perils of the algorithm-driven economy*. Harvard University Press.
- Ezrachi, A. and Stucke, M. E. (2017) "Artificial intelligence & collusion: When computers inhibit competition", *University of Illinois Law Review*, 5: 1775-1810.
- Fan, W., Yu, P. S., and Wang, H. (2004) "Mining extremely skewed trading anomalies" en *International Conference on Extending Database Technology*: 801-810. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Fletcher, G. G. S. (2021) "Deterring Algorithmic Manipulation", *Vanderbilt Law Review*, 74 (2): 259-326.
- Gautier, A., Ittoo, A., and Van Cleynenbreugel, P. (2020) "AI algorithms, price discrimination and



- collusion: a technological, economic and legal perspective”, *European Journal of Law and Economics*, 50(3): 405-435.
- Genberg, H. and Karagedikli, Ö. (2021) “Machine Learning and Central Banks: Ready for Prime Time?”, *South East Asian Central Banks (SEACEN) Research and Training Centre*.
- Gogas, P., Papadimitriou, T., and Agrapetidou, A. (2018) “Forecasting bank failures and stress testing: A machine learning approach”, *International Journal of Forecasting*, 34(3): 440-455.
- Golmohammadi, K., Zaiane, O. R., and Díaz, D. (2014) “Detecting stock market manipulation using supervised learning algorithms” en *International Conference on Data Science and Advanced Analytics (DSAA)*: 435-441). IEEE. Octubre.
- Gurkaynak, G., Altinsoy, I. N., and Rona, U. (2018) “Competition Law Consequences of Artificial Intelligence”, *The Academic Gift Book of Elig. Attorneys-At-Law in Honor of the 20th Anniversary of Competition Law Practice in Turkey*: 289-312.
- Hinterlang, N. (2020) “Predicting monetary policy using artificial neural networks”, *Discussion Paper Deutsche Bundesbank*, 44/20.
- Islam, S. R., Ghafoor, S. K., and Eberle, W. (2018) “Mining illegal insider trading of stocks: A proactive approach” en *IEEE International Conference on Big Data*: 1397-1406. IEEE. Diciembre.
- Kasai, N. and Naraidoo, R. (2011) “Evaluating the forecasting performance of linear and nonlinear monetary policy rules for South Africa”, *Munich Personal RePEc Archive*.
- Kulkarni, A., Mani, P., and Domeniconi, C. (2017) “Network-based anomaly detection for insider trading”, *arXiv:1702.05809*.
- Le, H. H. and Viviani, J. L. (2018) “Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios”, *Research in International Business and Finance*, 44: 16-25.
- Lim, D. (2021) “Can Computational Antitrust Succeed?”, *Stanford Computational Antitrust*, 1.
- Lorenzoni, I. (2022) *Why do Competition Authorities need Artificial Intelligence? Yearbook of Antitrust and Regulatory Studies*, 33.
- Maehashi, K. and Shintani, M. (2020) “Macroeconomic forecasting using factor models and machine learning: an application to Japan”, *Journal of the Japanese and International Economies*, 58: 101104.
- Mehra, S. K. (2015) “Antitrust and the robo-seller: Competition in the time of algorithms”, *Minnesota Law Review*, 100 (4): 1323- 1376.
- Nti, I. K., Adekoya, A. F., and Weyori, B. A. (2020) “A comprehensive evaluation of ensemble learning for stock-market prediction”, *Journal of Big Data*, 7(1): 1-40.
- Öğüt, H., Aktaş, R., Alp, A., and Doğanay, M. M. (2009). Prediction of financial information manipulation by using support vector machine and probabilistic neural network”, *Expert Systems with Applications*, 36(3): 5419-5423.
- Önder, E., Bayır, F., and Hepsen, A. (2013) “Forecasting macroeconomic variables using artificial neural network and traditional smoothing techniques”, *Journal of Applied Finance and Banking*, 3(4): 73-104.
- Petit, N. (2017) “Antitrust and artificial intelligence: a research agenda”, *Journal of European Competition Law & Practice*, 8(6): 361-362.



- Pham, X. T. and Ho, T. H. (2021) "Using boosting algorithms to predict bank failure: An untold story", *International Review of Economics & Finance*, 76: 40-54.
- Quaid, J. (2021) "AI and Competition Law" en Florian Martin-Bariteau & Teresa Scassa, eds., *Artificial Intelligence and the Law in Canada*. Lexis Nexi. Toronto.
- Rab, S. (2019) "Artificial intelligence, algorithms and antitrust", *Competition Law Journal*, 18(4): 141-150.
- Schrepel, T. (2021) "Computational antitrust: An introduction and research agenda", *Stanford Computational Antitrust*, 1: 1-15.
- Seth, T. and Chaudhary, V. (2020) "A predictive analytics framework for insider trading events" en *IEEE International Conference on Big Data*: 218-225). IEEE. Diciembre.
- Slemmer, D. W. (2019) "Artificial Intelligence and Artificial Prices: Safeguarding Securities Markets from Manipulation by Non-Human Actors", *Brooklyn Journal of Corporate, Financial & Commercial Law*, 14: 149-180.
- Šmejkal, V. (2021) "Three challenges of artificial intelligence for antitrust policy and law", *Journal for the International and European Law, Economics and Market Integrations*, 8 (2): 97-118.
- Son, Y., Byun, H., and Lee, J. (2016) "Nonparametric machine learning models for predicting the credit default swaps: An empirical study", *Expert Systems with Applications*, 58, 210-220.
- Tam, K. Y. and Kiang, M. (1990) "Predicting bank failures: A neural network approach", *Applied Artificial Intelligence an International Journal*, 4 (4): 265-282.
- Tamersoy, A., Khalil, E., Xie, B., Lenkey, S. L., Routledge, B. R., Chau, D. H., and Navathe, S. B. (2014) "Large-scale insider trading analysis: patterns and discoveries", *Social Network Analysis and Mining*, 4: 1-17.
- Treasury, H. M. (2019) *Unlocking digital competition, Report of the Digital Competition Expert Panel*.
- Von Bonin, A. and Malhi, S. (2020) "The use of artificial intelligence in the future of competition law enforcement", *Journal of European Competition Law & Practice*, 11(8): 468-471.



INSTITUTO  
ESPAÑOL  
DE ANALISTAS  
DESDE 1965  
FUNDACIÓN

El impacto de la inteligencia artificial en la  
industria financiera: promesas y amenazas



## Parte III



## 11. Riesgos, amenazas y desafíos de la AI en la industria

“A medida que aumenta la aplicación de herramientas de inteligencia artificial en muchas áreas relacionadas con la prestación de servicios, aumentan las cuestiones controvertidas sobre las consideraciones éticas de su uso”

*Aleksandrina Aleksandrova, Valentina  
Ninova, y Zhelyo Zhele (2023).*

La inteligencia artificial desempeña un papel central en el proceso de cambio tecnológico actual en la industria de los servicios financieros. Su lugar destacado en las agendas de innovación indica la percepción sobre los importantes beneficios que estas tecnologías pueden aportar a las entidades financieras, a los consumidores y a los mercados. La adopción de tecnologías de inteligencia artificial promete acelerar la innovación, mejorar el marketing mediante la generación de

campañas de publicidad más enfocadas, utilizar asistentes virtuales inteligentes que atienden a los clientes en cualquier momento, reducir los costes de cumplimiento corporativo y de *back-office* mediante la automatización de tareas, así como aumentar la productividad (al poder destinar los recursos humanos a tareas de mayor valor añadido). Pero, al mismo tiempo, los sistemas de AI tienen el potencial de causar daños significativos. La implantación de la inteligencia artificial puede

convertirse en una espada de doble filo si la gestión de los datos no es prudente, si se aplica con objetivos de manipulación y de fraude, si se traduce en una mayor invasión de la privacidad, si conduce a una mayor polarización social o a mayor discriminación.<sup>1</sup> En vista de ello, en los últimos años se ha reconocido cada vez más la importancia de que se investiguen los desafíos que representa su implantación y que la adopción de la AI se guíe por consideraciones de innovación responsable.

En este capítulo se comentan algunas de las amenazas y desafíos que implica el uso de la inteligencia artificial en la industria financiera, incluyendo la posibilidad de que las inversiones realizadas sean improductivas; que su desarrollo facilite la penetración de nuevos competidores en el sector, generando turbulencias

en la industria (lo que puede ser positivo para los consumidores pero también alterar la solvencia de los incumbentes actuales); que tenga un impacto dramático en las políticas de recursos humanos de los intermediarios financieros; y que implique cambios drásticos en las estructuras organizativas. Estos desafíos también se refieren a la necesidad de lidiar con problemas de opacidad de los resultados obtenidos, con la posibilidad de exceso de confianza en los sistemas de AI implantados, con el riesgo de enfrentarse a una excesiva dependencia de algunos proveedores; y, desde luego, se refieren también a la necesaria acción proactiva para lograr que la implantación de proyectos de AI sea compatible con criterios éticos. Muchos de los retos comentados no sólo afectan al sector financiero, sino que son comunes a otras industrias.

## 11.1. INVERSIONES IMPRODUCTIVAS

Hay esencialmente tres riesgos implicados en la adopción de sistemas de inteligencia artificial: una baja calidad de las aplicaciones implementadas,

el posible rechazo de los consumidores a la hora de interactuar con interfaces de AI y la posible resistencia por parte de los empleados a la hora de incorporar

<sup>1</sup> Véase Chow *et al.* (2019).



estas herramientas en su trabajo. La insuficiente utilización de los sistemas de inteligencia artificial por parte de estos últimos puede estar vinculada a la falta de confianza sobre la integridad del sistema o a la desconfianza en los resultados, por ejemplo, derivada de la baja calidad de la estructura de los datos (incluso aunque éstos sean masivos). Ambas circunstancias pueden traducirse en que las entidades incurran en fuertes inversiones que no ofrezcan un retorno suficiente. Se ha acuñado incluso un nombre para el indicador del retorno sobre las inversiones en inteligencia artificial, el ROAI.

Más allá de la rentabilidad de las inversiones, el hecho de que los resultados ofrecidos por los sistemas implantados no sean suficientemente acertados o no sean robustos puede acarrear otros problemas económicos en las compañías. La calidad del sistema AI puede ser baja si los modelos de diseño y entrenamiento no han sido correctos. Los resultados obtenidos pueden, además, ser poco robustos. La robustez en el sentido estadístico significa que un modelo continúa proporcionando estimaciones de los parámetros fiables y coherentes si las desviaciones en los supuestos

del modelo no son materiales. Dado que los algoritmos de aprendizaje automático trabajan con pocas hipótesis respecto de la distribución estadística, suelen ser robustos, pero también deben serlo ante cambios en la distribución de la población, (*population drift*). La sensibilidad a la deriva de la población puede comprobarse mediante un estudio de simulación (modificando artificialmente la distribución) o mediante pruebas retrospectivas sobre un largo periodo muestral. Las estructuras de aprendizaje complejas pueden ser más vulnerables a los cambios en la población a lo largo del tiempo que los modelos lineales más sencillos, por lo que en el aprendizaje automático se ha despertado un interés reciente por los *métodos de adaptación de dominio*. Estos métodos se orientan a la construcción de una solución de aprendizaje automático en un ámbito que pueda ser aplicarla con éxito en otro distinto. En el aprendizaje automático, si los datos de entrenamiento son una muestra insesgada de una distribución subyacente, la función de clasificación aprendida hará predicciones precisas para las nuevas muestras. En caso contrario, habrá diferencias entre la distribución de los datos de

entrenamiento y la de los datos de prueba. Como los clasificadores estándar no pueden hacer frente a los cambios en la distribución de los datos entre las fases de entrenamiento y de prueba, se utilizan subcampos del aprendizaje que se ocupan de tener en cuenta este tipo de cambios como los métodos de adaptación al dominio (también conocidos como de *aprendizaje por transferencia*)<sup>2</sup>. Cuando no se corrigen los problemas de calidad de los resultados, el riesgo de confiar en un sistema de AI para adoptar decisiones estratégicas, tácticas u operativas puede ser muy elevado.

Otra de las causas señaladas por las que inversiones millonarias en proyectos de AI en los bancos incumbentes pueden no ser suficientemente rentables tiene su origen en aspectos organizativos. La mayoría de los bancos tradicionales están organizados en torno a distintas líneas de negocio con equipos centralizados de tecnología y estructurados como centros de costes, mientras que los gestores de los negocios definen objetivos unilateralmente, y la alineación con la estrategia tecnológica y analítica de la empresa puede ser débil o inadecuada<sup>3</sup>.

## 11.2. RIESGO DE PENETRACIÓN DE NUEVOS COMPETIDORES

Desde la década de los años 60 y hasta los años 90 en el siglo XX, los bancos experimentaron una transformación radical como resultado de la automatización de las tareas bancarias y la informatización de las sucursales. Aún en un entorno de desregulación, las entidades bancarias no estuvieron

amenazadas por la entrada en la industria de competidores de naturaleza diferente. Posteriormente, con la emergencia de Internet y el proceso de digitalización se produjo el fenómeno Fintech, que comenzó a desdibujar las fronteras de la industria bancaria. Se lanzaron, por parte de empresas de perfil no

<sup>2</sup> Véase Kouw y Loog (2018 y 2019) o Redko et al. (2020) o Stojanov et al. (2021).

<sup>3</sup> Véase Biswas et al. (2020).

sólo financiero sino tecnológico, plataformas de financiación hipotecaria y financiación al consumo, así como de *préstamos entre pares* y plataformas de inversión. Estas plataformas combinan la evaluación automatizada del comportamiento de los clientes minoristas con la oferta de interfaces más fáciles de usar, para brindar una asistencia más práctica y con un menor coste para los usuarios. Las compañías de Fintech pueden asesorar en decisiones de financiación e inversión con menos requisitos de colaboración y asociación humana<sup>4</sup>. Aunque los bancos tradicionales están gastando cifras millonarias en desarrollos tecnológicos la penetración de nuevos competidores apoyada en la reducción de las barreras de entrada ha sido una realidad. Posteriormente, las innovaciones basadas en *blockchain* han permitido generar confianza descentralizada y acelerar el tiempo de liquidación de las transacciones, ya que posibilita a individuos y empresas ejecutar transacciones directamente y ver el mismo registro de intercambios actualizado y protegido mediante

criptografía<sup>5</sup>. Estos desarrollos DeFi también están siendo una fuente de competencia de nuevos entrantes en la industria financiera. Del mismo modo, el desarrollo continuo de tecnologías de *big data*, computación en la nube e inteligencia artificial están barriendo las barreras de entrada en la industria. Como ha ocurrido con Fintech y con DeFi anteriormente, el uso de aplicaciones de inteligencia artificial puede debilitar los lazos que han mantenido unidos los componentes de las instituciones financieras tradicionales, abriendo la puerta a más innovaciones y nuevos modelos operativos<sup>6</sup>. Las empresas especializadas en tecnología de inteligencia artificial, entre las que destacan gigantes tecnológicos incluidos en el grupo GAFAMN<sup>7</sup>, tienen la capacidad de agregar habilidades de servicios financieros a sus asistentes domésticos inteligentes y aprovechar esta interfaz de datos para penetrar en la industria financiera. Las ventajas competitivas de las GAFAMN en el seguimiento de la participación de los usuarios en las redes sociales y la observación de sus

<sup>4</sup> Véase Mor y Gupta (2021).

<sup>5</sup> Véase Hassani *et al.* (2018).

<sup>6</sup> Véase Kumar *et al.* (2022 y 2023).

<sup>7</sup> Google, Apple, Meta, Amazon, Microsoft y Nvidia.

patrones de gasto *on-line* cambian por completo las oportunidades de crecimiento orgánico, con la aplicación de estrategias que van desde la venta cruzada hasta la prestación de asesoramiento financiero proactivo. El uso inteligente de los datos (aún con el riesgo de convertirse en una amenaza para la privacidad del consumidor) puede proporcionar una propuesta de valor que el cliente aprecia, e incluso por la que puede estar dispuesto a pagar. Esta circunstancia da una oportunidad a las empresas de tecnología para que emerjan como jugadores relevantes del ecosistema de servicios financieros en un futuro.

La combinación DeFi y uso de la inteligencia artificial podría cambiar el equilibrio entre los bancos tradicionales y las FinTech, con la emergencia de lo que ha venido a denominarse *AI-first entities*. Ser una institución financiera basada en datos ya no es opcional (si alguna vez lo fue). Las entidades ganadoras serán aquellas con mejores capacidades para usar los datos y la AI en beneficio del consumidor.

Se ha argumentado, por

otra parte, que la inteligencia artificial también contribuye a la estandarización, la modelización y la inteligencia de negocio, al tiempo que ayuda a la toma de decisiones crediticias, permite una alerta temprana del fraude y contribuye a la supervisión de riesgos. En este contexto, el uso de AI puede favorecer una mayor fidelidad de los clientes<sup>8</sup>. La satisfacción del cliente decide la supervivencia y el éxito de la organización en el actual entorno competitivo, y el uso de tecnologías de inteligencia artificial puede contribuir a mejorar la experiencia de los clientes en este ámbito<sup>9</sup>. Dado que los desarrollos de AI exigen fuertes inversiones en tecnología, pero también gran experiencia y conocimiento sobre el negocio, los bancos incumbentes tienen ciertas ventajas competitivas y esta tecnología puede actuar como una nueva barrera de entrada. En este sentido, las entidades financieras tradicionales parecen estar en disposición de bloquear ataques disruptivos de nuevos entrantes aprovechando la fortaleza de su capital, su experiencia en el sector y su posición de dominio. Un freno a esta reacción, sin embargo, puede ser una actitud

<sup>8</sup> Véase Hu y Wang (2020).

<sup>9</sup> Véase Satheesh y Nagaraj (2021).

de reticencia a la hora de aplicar las nuevas herramientas de AI. Se ha destacado como un posible error en la orientación estratégica de la implantación de AI, por parte de las entidades tradicionales, el hecho de concentrar los recursos exclusivamente en proyectos enfocados en la reducción de costes, la mejora de la productividad, la gestión de riesgos y la protección contra pérdidas por fraude. Si bien estos casos de uso ciertamente pueden ayudar a las instituciones financieras a cumplir con sus objetivos financieros a corto plazo, pueden perder la ocasión de que el consumidor perciba los beneficios que personalmente les puede proporcionar una transición hacia el uso de la inteligencia artificial. En definitiva, podrían renunciar a la capacidad que representa la AI para mejorar la experiencia de uso de los clientes, reforzar su fidelidad y permitirles

participar en un nuevo ecosistema financiero más avanzado.

Un escenario que se contempla como factible es aquel en el que los gigantes tecnológicos actuales establecen relaciones con proveedores bancarios tradicionales o con *neobancos* para participar en el negocio financiero. No es descartable que las empresas de tecnología y de finanzas se fusionen o al menos colaboren para crear perfiles psicográficos<sup>10</sup> completos de los consumidores a través de datos sociales, comerciales, personales y financieros.

Además, la colaboración entre entidades financieras tradicionales y empresas tecnológicas podría además traducirse en una transferencia de talento y conocimientos enriquecedora para ambos tipos de empresas.

### 11.3. IMPACTO SOBRE LOS RECURSOS HUMANOS

El análisis del impacto de las tecnologías de inteligencia artificial en los recursos

humanos en los diferentes sectores, incluyendo la industria financiera, se concentra en cinco

<sup>10</sup> La segmentación psicográfica toma en cuenta los rasgos psicológicos de los consumidores, su estilo de vida, sentimientos, intereses, deseos, etc.

ámbitos: i) las estrategias de sustitución de empleos; ii) la colaboración entre humanos y algoritmos en la toma de decisiones; iii) las oportunidades de aprendizaje, formación y capacitación; iv) el reclutamiento; y v) el desempeño laboral.

Las innovaciones tecnológicas pueden afectar al nivel del empleo esencialmente por dos vías: desplazando directamente a los trabajadores de tareas que antes desempeñaban (efecto de desplazamiento o fenómeno de *desempleo tecnológico*) y aumentando la demanda de empleo para nuevas tareas vinculadas con el desarrollo y aplicación de las nuevas tecnologías (*empleo emergente*). ¿Cuál de los dos efectos prevalecerá en la industria financiera? En general, las revoluciones industriales pasadas sugieren que, a corto plazo, el efecto de desplazamiento puede dominar. A largo plazo, cuando los mercados y la sociedad están plenamente adaptados a los grandes shocks de automatización, el efecto del empleo emergente supera con creces, de forma que se produce

un efecto positivo neto en el empleo. Además, la experiencia informa sobre *efectos externos* entre sectores relevantes, ya que los nuevos empleos no se generan necesariamente en los mismos sectores en los que se destruyen<sup>11</sup>. Las nuevas tecnologías pueden reducir la cantidad total de tiempo de trabajo necesario para producir el consumo actual, pero los mayores ingresos resultantes (junto con la característica de no saturación de la utilidad en el consumo) pueden provocar un aumento de la demanda de productos existentes, así como de nuevos productos y servicios (algunos de los cuales serán creados como consecuencia de la implementación de las nuevas tecnologías). El aumento resultante del empleo puede intensificarse si el gasto adicional se orienta a servicios con elasticidad renta superior a la unidad que sean intensivos en mano de obra<sup>12</sup>. Históricamente, la informatización se había limitado en gran medida a tareas rutinarias manuales y cognitivas que implicaban actividades explícitas basadas en reglas<sup>13</sup>. Con la inteligencia artificial, sin

<sup>11</sup> Véase Acemoglu y Restrepo (2016).

<sup>12</sup> Véase Borland y Coelli (2017).

<sup>13</sup> Véase Autor y Dorn (2013) y Goos *et al.* (2009).

embargo, la informatización se está extendiendo a ámbitos comúnmente definidos como no rutinarios. Son muy numerosas las tareas que hace sólo una década se definían como no rutinarias y que hoy se han vuelto claramente “computarizables”<sup>14</sup>. Se argumenta que quizá se está produciendo una sobreestimación de la “automatizabilidad” del trabajo, ya que las ocupaciones etiquetadas como de alto riesgo a menudo todavía contienen una proporción sustancial de tareas que son difíciles de automatizar<sup>15</sup>. Sin embargo, los avances en campos relacionados con el aprendizaje automático, la minería de datos, la visión artificial, la estadística computacional y otros subcampos de la inteligencia artificial están poniendo de manifiesto que el desarrollo de algoritmos permite automatizar tareas de naturaleza cada vez más cognitiva<sup>16</sup>.

La disponibilidad de *big data* pone en valor una de las principales ventajas comparativas de los ordenadores en relación con el trabajo humano: la escalabilidad. Se necesita poca

evidencia para demostrar que, al realizar tareas de cálculo laborioso, las redes de máquinas escalan mejor que el trabajo humano<sup>17</sup>. Los avances en las interfaces de usuario también permiten que las computadoras respondan directamente a una gama más amplia de solicitudes humanas, pudiendo sustituir parcialmente trabajo de mano de obra altamente calificada, al tiempo que permiten que algunos tipos de trabajos se automaticen por completo. Ocupaciones que requieren un juicio sutil también son cada vez más susceptibles a la informatización.

La inteligencia artificial tiene ya hoy la capacidad de realizar tareas de alta cualificación y trabajos no rutinarios, lo que afectará de forma notable al empleo en la industria financiera. Una parte de los empleados actuales aumentarán su productividad, se producirán redundancias y la industria necesitará de nuevos perfiles para desarrollar y mantener nuevos sistemas basados en AI. Algunos de los nuevos trabajos se crearán dentro de la industria financiera, pero otros serán creados en

<sup>14</sup> Véase Brynjolfsson y McAfee (2011).

<sup>15</sup> Véase Arntz *et al.* (2016).

<sup>16</sup> Véase Frey y Osborne (2017).

<sup>17</sup> Véase Campbell-Kelly (2009).

proveedores externos de servicios de las entidades financieras, que actuarán como clientes.

Los empleos necesarios para la transición hacia un modelo híbrido de colaboración entre humanos y máquinas inteligencias son aquellos destinados a realizar tres tipos de roles esenciales<sup>18</sup>:

- Los *entrenadores* que se ocupan de hacer posible la implementación de los sistemas de AI.
- Los *explicadores* que ayudan a proporcionar claridad a medida que aumenta la opacidad de los sistemas de inteligencia artificial.
- Los *encargados de mantenimiento* que ayudan a garantizar que los sistemas de AI estén funcionando según lo diseñado y que las consecuencias no deseadas se aborden con la urgencia adecuada.

Se afirma que la repercusión a medio plazo en el mundo laboral de la AI en la industria financiera puede ser tremenda. Si la innovación tecnológica afectó

principalmente a la mano de obra menos calificada, la AI reemplazará parcialmente a trabajadores del conocimiento y empoderará más a aquellos con capacidades de creatividad e intuición. Existe el riesgo de que a corto plazo la promesa de más empleos y tiempo libre no compense un impacto rápido negativo en la actividad laboral de la industria financiera, lo que puede generar tensiones dentro del sector.

Respecto del impacto de la generación del uso de inteligencia artificial en la estructura del empleo de la industria financiera, existe todavía debate sobre si se producirá un fenómeno de *polarización laboral*, lo que se ha asociado al mayor uso de tecnología de la información y los ordenadores de forma transversal por sectores y países. Este fenómeno es un patrón de cambios en la composición del empleo en el que hay un aumento en la proporción del rango de alta calificación, una disminución en el caso de calificación media y un aumento en la proporción de empleos de baja calificación<sup>19</sup>. La polarización implica que, en términos netos, los cambios

<sup>18</sup> Véase Wilson *et al.* (2017).

<sup>19</sup> Véase, Acemoglu y Autor (2011), Autor y Dorn (2013), Goos y Manning (2007) y Coelli y Borland (2016).

en el empleo muestran una forma de U en cuanto al nivel de habilidades, con caídas relativas del empleo en el medio de la distribución y ganancias relativas en los extremos.

Y, del mismo modo, no está claro que en el caso de la generalización del uso de la tecnología AI vaya a repetirse la evidencia previa de lo que ocurrió con las TICs. En este último caso se produjo una sustitución de trabajo realizado en tareas rutinarias, pero en el trabajo dedicado a la realización de tareas cognitivas no rutinarias la tecnología actuó como un complemento<sup>20</sup>. Las tecnologías informáticas y de comunicación se consideran complementarias y, por tanto, tienen una mayor demanda de trabajadores que puedan realizar tareas cognitivas no rutinarias, lo que requiere trabajadores con altos niveles de educación<sup>21</sup>. Se argumenta que un aumento en la proporción de la fuerza laboral con altos niveles de educación favorece la posibilidad de aprovechar las oportunidades de creación de empleo que brindan las tecnologías informáticas. Esta argumentación parece, en principio, también

aplicable en el caso de la AI.

Por otra parte, cabe destacar que los avances recientes en inteligencia artificial están impulsados principalmente por el aprendizaje automático, una tecnología de predicción. Siendo la predicción como insumo crucial en la toma de decisiones, el análisis del impacto en los puestos de trabajo, exige en todo caso, distinguir entre roles relativos de tareas de predicción y decisión. La inteligencia artificial afecta de forma diferente cuando la automatización de la predicción conduce a la automatización de decisiones que cuando simplemente contribuye a la mejora de la toma de decisiones por parte de los humanos<sup>22</sup>. Dado que se espera que humanos y robots trabajen juntos de manera antes no imaginable, se producirán cambios en las habilidades que determinan la empleabilidad. La perturbación causada por la inteligencia artificial está sucediendo 10 veces más rápido y a 300 veces la escala de la revolución industrial de finales del siglo XVIII y principios del XIX, lo que implica un impacto aproximadamente 3.000 veces

<sup>20</sup> Véase Gaggl y Wright (2017).

<sup>21</sup> Véase Goldin y Katz (2009).

<sup>22</sup> Véase Agrawal *et al.* (2019).

superior<sup>23</sup>. Debido a la magnitud de estos impactos en el empleo, las organizaciones tendrán que repensar su función de recursos humanos para realinear su experiencia con la realidad de la AI<sup>24</sup>. La *inteligencia aumentada* es un modelo de asociación entre personas y algoritmos de AI, que trabajan juntas para mejorar el rendimiento cognitivo. La *inteligencia aumentada* se centra en el papel de asistencia de la AI en el avance de las capacidades humanas. En el ámbito de la industria financiera se traduce en una interacción del algoritmo con las personas que mejora lo que ya saben, contribuye a reducir los errores y el trabajo rutinario, y puede mejorar las relaciones con los clientes. El objetivo de la *inteligencia aumentada* es hacer más eficiente las organizaciones, al mismo tiempo que se complementa con un toque humano y sentido común para gestionar los riesgos de la automatización de decisiones.

Otro de los retos a los que se enfrentan actualmente las entidades financieras que desean desarrollar proyectos ambiciosos en el ámbito de la AI es la escasez

de recursos humanos con conocimientos especializados en esta área. Las entidades inmersas en un despliegue importante de aplicaciones de AI pueden enfrentarse, transitoriamente al menos, a déficit de competencias y a la indisponibilidad de recursos humanos con las capacidades necesarias. La implantación de la AI en el sector financiero requiere competencias y conocimientos especializados, que pueden ser difíciles de encontrar. Las instituciones financieras deben invertir en formación y contratar a expertos en inteligencia artificial para garantizar el éxito de sus iniciativas en este terreno.

La formación de los empleados a través de asistentes virtuales basados en inteligencia es otro de los aspectos de cómo el despliegue de herramientas de AI influye en la gestión de los recursos humanos. El uso de personajes animados inteligentes con fines de capacitación permite retroalimentación y apoyo, como ocurre en el caso de los entrenadores humanos<sup>25</sup>. En el mismo contexto, se ha analizado en qué medida agentes basados en AI son herramientas útiles para

<sup>23</sup> Véase Dobbs *et al.* (2015).

<sup>24</sup> Véase Poba-Nzaou *et al.* (2021).

<sup>25</sup> Véase Behrend y Thompson (2011).

mejorar las habilidades de los empleados cuando interactúan en entornos estratégicos y de negociación, ahorrando esfuerzo considerable y ofreciendo un mejor rendimiento<sup>26</sup>.

Por lo que respecta al clima laboral, se ha resaltado que el uso de herramientas de inteligencia artificial en la toma de decisiones que puede aumentar la precisión y reducir el tiempo de respuesta con la mejora en el rendimiento de los empleados<sup>27</sup>, aprovechando su capacidad para procesar grandes cantidades de datos a altas velocidades<sup>28</sup>. Pero desde una perspectiva negativa, también se ha puesto de manifiesto que su uso extensivo en la toma de decisiones puede percibirse por parte de los empleados como una amenaza a su autonomía, estatus y seguridad laboral<sup>29</sup>. Además, la aplicación de herramientas de AI, al aumentar el número de opciones en la decisión y la complejidad percibida del proceso, puede generar insatisfacción en los trabajadores.

Finalmente, las herramientas

de inteligencia artificial pueden ayudar en los procesos de selección y contratación. Por ejemplo, permite a los empleados trabajando en los departamentos de recursos humanos agilizar las verificaciones de antecedentes de solicitudes de empleo y extrapolar posibles comportamientos en términos de adecuación y desempeño laboral. Además, el uso de entrevistas por video asincrónicas (AVI por sus siglas en inglés) y de agentes de decisión basados en inteligencia artificial se está incorporando en el proceso de selección de personal y permite que las entrevistas previas a la contratación se realicen en cualquier momento y en cualquier parte del mundo. Los empleadores pueden utilizar estos sistemas para entrevistar a un número enorme de candidatos para un determinado puesto de trabajo y registrar sus respuestas automáticamente, realizando un filtrado rápido<sup>30</sup>. Se argumenta que además de la eficiencia, un proceso de selección basado en algoritmos de AI puede tener como ventaja el hecho de resultar más sistemático y

<sup>26</sup> Véase Lin *et al.* (2014).

<sup>27</sup> Véase Lawler y Elliot (1996) y Hooper *et al.* (1998).

<sup>28</sup> Véase Lindebaum *et al.* (2020).

<sup>29</sup> Véase Lawler y Elliot (1996).

<sup>30</sup> Véase Naim *et al.* (2015), Brenner *et al.* (2016), Rasipuram *et al.* (2016) o Torres y Gregory (2018).

podría contener menores sesgos personales de los reclutadores<sup>31</sup>.

## 11.4 CAMBIOS EN LAS ESTRUCTURAS ORGANIZATIVAS

El desarrollo y avance de las tecnologías de la información y la comunicación cambiaron significativamente la forma de hacer negocios en la industria financiera. La transformación digital no fue una opción para las organizaciones, sin importar el tamaño de su balance, su madurez, el número de empleados, o el tipo de clientes, etc. Para sobrevivir y seguir siendo competitivos en la era digital los bancos tradicionales reconocieron los beneficios potenciales del uso de las nuevas tecnologías digitales. Al igual que ha ocurrido en el pasado reciente en el que las organizaciones financieras han tenido que crear una *cultura organizacional digital*, la emergencia de la inteligencia artificial y la adopción del concepto de *big data* imponen desafíos equivalentes. La cultura es el elemento más importante para la continuación de los valores corporativos

fundamentales, permite una mayor participación de los empleados en las transformaciones e implica una menor resistencia para el cambio y la adopción de nuevas tecnologías<sup>32</sup>. El uso generalizado de herramientas de inteligencia artificial puede obligar a las entidades financieras a realizar ajustes en la estructura organizativa del negocio y en el modelo operativo, para garantizar que existan niveles apropiados de supervisión de riesgos y que las evaluaciones de desempeño se alineen con las nuevas formas de trabajar<sup>33</sup>.

La inteligencia artificial puede conducir a una transformación en el modelo de servicio al cliente y los procesos operativos internos en el sector financiero. Las instituciones deberían preparar sus modelos de negocio y a sus empleados para este futuro, de forma que pueden

<sup>31</sup> Véase Van Esch *et al.* (2019) y Sajjadi *et al.* (2019).

<sup>32</sup> Véase Çetin-Gürkan y Çiftci (2020).

<sup>33</sup> Véase Kruse *et al.* (2019).

convertir estos desarrollos en una oportunidad. No hay duda de que estos procesos tendrán

repercusiones significativas en las estructuras organizativas de las entidades bancarias.

## 11.5 COMPORTAMIENTOS DE ALUCINACIÓN Y OPACIDAD DE LOS RESULTADOS

Los retos generales que plantea la inteligencia artificial para la innovación responsable, combinados con los daños concretos que puede causar su uso en los servicios financieros, hacen necesario garantizar y demostrar que los sistemas de AI son fiables y se utilizan de forma responsable. Por otra parte, la transparencia de la AI, que en este caso se refiere a la disponibilidad de información relevante sobre estos sistemas para las partes interesadas pertinentes, es también crucial. Sin embargo, la falta de fiabilidad y falta de transparencia son dos riesgos asociados al uso de herramientas de inteligencia artificial que a menudo se ponen de relieve. Una consecuencia imprevista en cuanto a los resultados de diagnóstico o a las soluciones proporcionadas por los sistemas de AI es que

estos sistemas pueden ser propensos a la *alucinación*.

En general, las alucinaciones de modelos de lenguaje LLM consisten en la generación de contenidos que no son representativos o no tienen sentido para la fuente proporcionada, por ejemplo, debido a errores en la codificación y decodificación entre el texto y las representaciones<sup>34</sup>. Los modelos previamente entrenados almacenan conocimiento en sus parámetros y pueden generar respuestas informativas cuando se utilizan en sistemas conversacionales. Sin embargo, los modelos de lenguaje pueden generar declaraciones aparentemente plausibles que, sin embargo, son irrelevantes u objetivamente incorrectas<sup>35</sup>. También se han detectado alucinaciones en

<sup>34</sup> Véase Beutel *et al.* (2023).

<sup>35</sup> Véase Sun *et al.* (2023).

tareas de reconocimiento facial y de imágenes mediante herramientas de inteligencia artificial. Inspirándose en las capacidades lingüísticas superiores de los modelos de lenguaje de gran tamaño, se han explorado también modelos de lenguaje-visión de gran tamaño (LVLM) y también estos modelos sufren el problema de las alucinaciones, es decir, tienden a generar objetos que son inconsistentes con las imágenes objetivo en las descripciones<sup>36</sup>.

El problema de la *alucinación* en AI, que se caracteriza por el hecho de que con falsa confianza se generan respuestas incoherentes y, a veces, incluso disparatadas, dificulta confiar en estos sistemas para funciones críticas sin una adecuada supervisión humana. Además, las alucinaciones generadas por sistemas de AI pueden dar lugar a lugar a problemas éticos y legales en la industria financiera. Se ha destacado que mejorar los insumos de capacitación incluyendo información diversa, precisa y conjuntos de datos contextualmente relevantes junto con actualizaciones frecuentes de los modelos de

entrenamiento podrían ayudar en el tratamiento de este problema<sup>37</sup>.

Un segundo problema es el de la opacidad de los resultados. La necesidad de transparencia incluye la posibilidad de acceso a información relativa a la lógica del modelo de inteligencia artificial (*transparencia del sistema*), así como a información sobre los procesos que rodean el diseño, el desarrollo y el despliegue del sistema (*transparencia del proceso*). Para ambas categorías, las partes interesadas con acceso conveniente a la información incluyen: participantes que se ocupan de diferentes funciones dentro de la empresa y accionistas (*transparencia interna*), así como inversores, proveedores, clientes y reguladores (*transparencia externa*).

Los problemas de opacidad tienen que ver con el propio modelo y también con los datos que lo alimentan. Las aplicaciones de AI son complejas y los modelos que utilizan suelen ser muy grandes, no lineales y no paramétricos, lo que dificulta la comprensión global de sus propiedades y su validación. Además, utilizan una gama más

<sup>36</sup> Véase Li et al. (2023).

<sup>37</sup> Véase Athaluri et al. (2023).



amplia de datos, datos complejos, y en ocasiones no estructurados (por ejemplo, procedentes de fuentes de noticias, imágenes de satélite o redes sociales). De forma creciente, la utilización de sistemas de AI se basa en la mezcla de datos propios y datos proporcionados por terceras fuentes, de forma que resulta complicado de verificar la calidad y el posible sesgo de los datos.

Por otra parte, los nuevos sistemas de IA combinan varios modelos y constan de forma creciente de multitud de componentes que interactúan entre sí. Esto dificulta validar si lo hacen como se pretende. Los métodos avanzados de aprendizaje algorítmico (como el aprendizaje profundo) son intrínsecamente poco transparentes o explicables. Los modelos de AI a menudo se consideran *cajas negras*, no necesariamente porque sean complicados, sino porque no se pueden interpretar o visualizar fácilmente, ya sea porque son muy recursivos (como en el caso del aprendizaje profundo), porque incluyen gran cantidad de dimensiones (como las máquinas de vectores de soporte) o porque son multicapa y algunas están ocultas (el caso de las redes neuronales). Otra razón es que

pueden no conocerse las señales de entrada. Como además, para mejorar los resultados, en vez de utilizar un único modelo independiente suele emplearse un conjunto de modelos, el rastreo de cómo se ponderan y generan los resultados no es sencillo.

En la industria financiera, el problema de la *caja negra* tiene implicaciones obvias para su uso en gestión de carteras y en gestión de riesgos, donde la opacidad en los resultados puede ser una fuente de riesgo relevante para las entidades. Las técnicas de aprendizaje automático extraen automáticamente conocimiento de un conjunto de datos y construyen diferentes representaciones de modelos para explicar el conjunto de datos. La principal diferencia entre los métodos estadísticos tradicionales y los métodos de aprendizaje automático es que los métodos estadísticos generalmente necesitan que los investigadores impongan estructuras, como la linealidad en el análisis de regresión múltiple, y construyan el modelo estimando parámetros para ajustarse a los datos o a la observación. Sin embargo, las técnicas de aprendizaje automático pueden aprender la estructura particular del modelo a partir de los datos y sin restricciones previas. Como

resultado, las estructuras impuestas por los analistas en los modelos estadísticos son relativamente simples y fáciles de interpretar, mientras que los modelos obtenidos con métodos de aprendizaje automático suelen ser muy complicados y difíciles de explicar. Si no es fácil trazar la generación de los resultados el sistema puede despertar desconfianza.

El peligro de que la aplicación de herramientas de AI opere en contextos de *caja negra* no se circunscribe al ámbito interno de las entidades, sino que se extiende a la relación con los clientes. Puede no comprenderse completamente cómo funciona el sistema (qué datos utiliza y cómo adopta las decisiones), cómo decide y, por tanto, cuándo y por qué se equivoca en su caso respecto a decisiones que afectan a clientes. Esta opacidad puede generar conflictos, por ejemplo, con los solicitantes de créditos y préstamos. Esta circunstancia obliga a establecer *sistemas de respuesta* a las solicitudes de información, así como *políticas de rectificación y compensación*. En el mundo de las finanzas la *explicabilidad* de los resultados y la mitigación del problema de

*caja negra* no es sólo un requisito interno razonable, y una parte del proceso de validación del modelo, puede ser también una condición legal o reglamentaria.

El antídoto contra situaciones de *caja negra* es crear un sistema de inteligencia artificial que pueda explicarse a sí mismo, donde las decisiones y predicciones se hagan de forma transparente. Sin embargo, la realidad demuestra que desafortunadamente existe tensión entre precisión y *explicabilidad*. Los modelos más precisos suelen ser los menos transparentes y explicables, mientras que los que permiten mayor grado de claridad son menos precisos. Para promover la rendición de cuentas y facilitar la auditoría de las decisiones, recientemente se ha producido una tendencia hacia la AI explicable, o XAI. El campo de la *inteligencia artificial explicable* ha sido testigo del surgimiento de numerosos métodos y técnicas destinados a comprender el intrincado funcionamiento de los modelos de aprendizaje profundo<sup>38</sup>. Así, la *inteligencia artificial explicable* es un campo de la informática de la AI que promueve un conjunto de herramientas, técnicas y

<sup>38</sup> Véase Yang et al. (2023).



algoritmos que pueden generar explicaciones de alta calidad interpretables, intuitivas y comprensibles para los humanos en relación con los resultados y decisiones obtenidas mediante AI<sup>39</sup>. La XAI se considera parte de la *tercera ola de AI*, cuyo objetivo es generar algoritmos que puedan explicarse por sí mismos<sup>40</sup>. En el ámbito de las finanzas, la XAI se está utilizando, por ejemplo, para documentar, comprender y validar modelos de aprendizaje automático de *caja negra* utilizados para las decisiones de aprobación de préstamos<sup>41</sup>.

En el proceso de explicación de cómo se producen los resultados en un sistema de AI se han destacado tres enfoques principalmente<sup>42</sup>:

- Una interpretación local de los algoritmos de aprendizaje automático, mediante un modelo lineal dentro de una región local del espacio de características (*Local Interpretable Model-agnostic Explanations*).

- Una medición y visualización de la contribución marginal de cada característica del aprendizaje automático (*SHapley Additive exPlanations*).
- El uso de enfoques más sencillos basados, por ejemplo, en reglas o árboles de decisión, para explicar el comportamiento general de un algoritmo de aprendizaje automático.

En la elección del enfoque adecuado entre los anteriores es conveniente determinar con precisión si lo que realmente se necesita es la comprensión completa de cómo funciona el algoritmo predictivo, o bien de cómo el algoritmo realiza una predicción o toma una decisión en un caso específico. El primer problema parece más difícil de solucionar que el segundo ya que puede requerir la validación general del modelo, pero es posible que el banco sólo necesite en cada caso explicar un caso específico (por ejemplo, una decisión de préstamo a un cliente concreto)<sup>43</sup>. Cada uno de los enfoques tiene sus puntos

<sup>39</sup> Véase Das y Rad (2020).

<sup>40</sup> Véase Zhang *et al.* (2022).

<sup>41</sup> Véase Hall y Gill (2019) o Bussmann *et al.* (2020).

<sup>42</sup> Véase Quell *et al.* (2021).

<sup>43</sup> Véase Sudjianto *et al.* (2020).

débiles. Por ejemplo, los dos primeros enfoques señalados pueden no ser apropiados en el caso de aplicaciones críticas en finanzas, donde se requieren explicaciones precisas.<sup>44</sup>

Si bien algunos algoritmos de aprendizaje son intrínsecamente explicables (por ejemplo, los árboles de decisión del tipo ID3 o CART), las redes Deep ReLU, pueden interpretarse a su vez mediante conjuntos equivalentes de modelos lineales locales (LLU). Con el conjunto de herramientas adecuado, el analista puede determinar cómo se comporta localmente la red con tanta precisión como sea necesario<sup>45</sup>. Por otra parte, dado que las redes neuronales poseen un buen rendimiento de predicción, pero adolecen de una falta de interpretabilidad del modelo, se ha propuesto mejorar su *explicabilidad* mediante restricciones en la arquitectura. En este sentido, el objetivo es lograr una red neuronal explicable mejorada (ExNN, por sus siglas en inglés) con un equilibrio superior entre el rendimiento de la predicción y la interpretabilidad del modelo<sup>46</sup>.

Un proceso XAI es consistente con el concepto de «humano en el circuito», ya que su objetivo es permitir una mejor comprensión de los sistemas de AI opacos para que los humanos puedan utilizar mejor dichas herramientas en su trabajo. Esto nos lleva a la cuestión de la necesidad de supervisión de los sistemas de AI ante la posibilidad de propensión a sobreajustar los datos de entrenamiento (*overfitting*), como consecuencia de la complejidad de los algoritmos de aprendizaje automático. Este problema consiste en que se ajustan aspectos aleatorios espurios de los datos de entrenamiento, así como la estructura verdadera para toda la población. Una técnica para manejar el sobreajuste es la *regularización*, que consiste en comparar el rendimiento de los datos de entrenamiento con un conjunto de datos independientes. El objetivo es decidir si, no siendo posible aumentar el tamaño de la muestra, se debe reducir la complejidad del modelo y el número de parámetros utilizados. Muchos métodos de aprendizaje automático utilizan el rendimiento en un conjunto de datos fuera de la muestra para determinar cuándo

<sup>44</sup> Véase Quell et al. (2021).

<sup>45</sup> Véase Quell et al. (2021).

<sup>46</sup> Véase Yang et al. (2020).

dejar de entrenar el modelo, pero la significación del resultado debe ajustarse en función del número de pruebas realizadas, ya que un buen resultado con

menos pruebas fuera de la muestra es mejor que un resultado ligeramente mejor después de muchas más pruebas<sup>47</sup>.

## 11.6. EXCESO DE CONFIANZA EN LA AI Y RIESGO DE DEPENDENCIA

La confianza es un estado psicológico que comprende la intención de aceptar la vulnerabilidad basada en expectativas positivas respecto del comportamiento de otro ente. En el caso que nos ocupa la confianza respecto de los resultados aportados por un sistema de inteligencia artificial. La confianza sólo es relevante en condiciones de riesgo e incertidumbre, en las que una confianza mal depositada conlleva potencialmente pérdida o daño<sup>48</sup>. La confianza en los algoritmos se apoya en tres niveles<sup>49</sup>: confianza disposicional (tendencia a confiar en la automatización en general), confianza situacional (en la que influyen componentes del entorno contextual inmediato) y confianza

interaccional (que depende de experiencias inmediatas con la automatización en cuestión).

Un reto clave para tener confianza en los sistemas de AI es alcanzar un nivel adecuado de precisión de los resultados. Cabe esperar que resultados inexactos generen desconfianza y un rechazo al uso de los algoritmos. Sin embargo, algunas investigaciones sugieren que no siempre la percepción de un comportamiento inexacto de los sistemas de AI se traduce en un menor seguimiento de las instrucciones que ofrecen estos algoritmos<sup>50</sup>. Este comportamiento inesperado es una cuestión relacionada con aspectos conductuales.

<sup>47</sup> Véase Quell *et al.* (2021).

<sup>48</sup> Véase Lockey *et al.* (2021).

<sup>49</sup> Véase Hoff y Bashir (2015)

<sup>50</sup> Véase Salem *et al.* (2015).

La familiaridad y la percepción de calidad de la gobernanza son otros componentes básicos de la confianza en los sistemas de inteligencia artificial. Los estudios empíricos indican que la familiaridad con los sistemas de AI facilita la confianza. Un mayor conocimiento sobre el funcionamiento de la AI afecta la evaluación de los beneficios percibidos y, por tanto, existe una correlación positiva entre familiaridad y confianza<sup>51</sup>. Los estudios indican, además, que la gobernanza en forma de controles adecuados para garantizar un adecuado desarrollo y el despliegue de un sistema AI, facilita también la confianza en el sistema.

La relación óptima de confianza en los algoritmos AI exige que el usuario tenga una comprensión adecuada y bien calibrada de las capacidades del sistema, de modo que no desconfíe demasiado sobre la calidad de su desempeño (lo que podría dar lugar a un desuso), pero que tampoco confíe en exceso (lo que podría dar lugar a un uso indebido)<sup>52</sup>. De hecho, existe el riesgo de un exceso de confianza en la AI por parte de las entidades financieras que adoptan

estos sistemas. Las instituciones financieras deben asegurarse de que no utilizan ciegamente los resultados obtenidos de sistema de inteligencia artificial y que existen controles humanos para supervisar y anular decisiones adoptadas en caso necesario.

Por otra parte, un universo competitivo en el ámbito de la inteligencia artificial caracterizado por en una excesiva concentración de proveedores externos de software de AI, puede conducir a que las entidades financieras tengan una elevada dependencia de un número limitado de oferentes de servicios y aplicaciones en este terreno. En este contexto, adquirir capacidades propias suficientes para supervisar adecuadamente los sistemas AI y, en su caso, para diseñar estrategias de migración es una condición necesaria para la correcta gestión de este riesgo en las entidades financieras. Además, una recomendación estratégica para mitigar el exceso de dependencia y los riesgos derivados de la implementación de la AI en las organizaciones consiste en realizar un inventario exhaustivo del entorno AI. Este ejercicio implica la identificación

<sup>51</sup> Véase Said *et al.* (2023).

<sup>52</sup> Véase Sanders *et al.* (2019).

de todas las áreas en las que la institución depende de la inteligencia artificial, incluyendo tanto los modelos creados

internamente, como todas las funciones proporcionadas por terceros en cualquier momento del ciclo de vida del producto.

## 11.7 CIBERDELINCUENCIA VINCULADA A AI

Como ya hemos visto, el uso de la AI para combatir la ciberdelincuencia se está generalizando en el sector financiero, pero al tiempo, los ciberataques impulsados por sistemas basados en AI se han vuelto más comunes últimamente. Las herramientas de inteligencia artificial están siendo utilizadas tanto para mejorar como para alterar la ciberseguridad<sup>53</sup>. El desarrollo y la generalización del uso de la inteligencia artificial puede tener tres efectos en términos de seguridad cibernética: i) la ampliación del conjunto de actores que pueden llevar a cabo ataques; ii) un posible aumento en la velocidad a la que se pueden llevar a cabo estos ataques y iii) una ampliación en el alcance de los objetivos potenciales<sup>54</sup>.

Se ha afirmado que confiar en herramientas de aprendizaje automático y en redes neuronales para realizar tareas de ciberseguridad es un arma de doble filo: puede mejorar sustancialmente las prácticas de ciberseguridad, pero también puede ser una invitación a formas de ataques a las propias aplicaciones de AI, lo que puede plantear graves amenazas a la seguridad<sup>55</sup>.

Por otra parte se argumenta que los bancos tradicionales se encuentran en una desventaja inherente en comparación con las empresas tecnológicas cuando se trata de utilizar nuevas tecnologías<sup>56</sup>. En un contexto de implementación ambiciosa de una estrategia AI, esta circunstancia puede

<sup>53</sup> Véase AL-Dosari *et al.* (2022).

<sup>54</sup> Véase Meinert (2018).

<sup>55</sup> Véase Taddeo *et al.* (2021).

<sup>56</sup> Véase Shmuratko y Sheludko (2019).

generar riesgos de seguridad adicionales, debido a que los sistemas de software financieros y de seguridad heredados pueden ser incompatibles con las soluciones innovadoras.

En el contexto de la inteligencia artificial cabe distinguir dos aspectos clave de la ciberseguridad: la protección de todos los aspectos de la transformación y resiliencia digital y de los datos<sup>57</sup>; y la disponibilidad de sistemas, almacenes de datos y redes esenciales para garantizar la continuidad del servicio. En relación con la primera cuestión, cabe plantearse cómo afecta el uso de la AI en el sector bancario a la privacidad de los datos. Dado que los marcos normativos relacionados con esta cuestión aún están en fase de desarrollo, existe un área gris respecto de la cuestión de en qué medida los bancos pueden confiar en proveedores de servicios externos para garantizar la privacidad de los datos de sus clientes<sup>58</sup>. Se ha puesto de manifiesto que el uso de herramientas basadas en AI, como el procesamiento del lenguaje natural para analizar las comunicaciones de empleados

y clientes, o de *chatbots* para comunicarse con los clientes, puede en algunas circunstancias conducir a la vulneración de la privacidad de los clientes<sup>59</sup>.

En relación con la cuestión relativa a la garantía de continuidad en la prestación de los servicios financieros, los ciberdelitos basados en inteligencia artificial que pueden poner en riesgo la operación suelen agruparse en cinco categorías:

- La AI puede utilizarse como herramienta para perpetrar delitos relacionados con la suplantación de identidad mediante audio y vídeo, *phishing* personalizado, predicción del comportamiento de personas o instituciones para descubrir y explotar vulnerabilidades, generación de contenidos falsos (*fake news*) para mancillar la reputación corporativa, interrupción de los sistemas, corrupción de datos y chantaje a gran escala. En el nuevo mundo de la inteligencia artificial no hay texto, ni grabación de voz, ni videoclip en cuyo contenido se pueda

<sup>57</sup> Esto incluye la confidencialidad de la información sensible y la integridad del procesamiento.

<sup>58</sup> Véase Truby *et al.* (2020).

<sup>59</sup> Véase Lai *et al.* (2018) y Caldwell *et al.* (2020).

confiar sin contrastar que se trata de una fuente confiable. Cualquier voz o imagen puede ser una creación manipulada. Los riesgos de desinformación mediante sobrecarga de datos y noticias falsas aumentan en este nuevo entorno.

- El uso de inteligencia artificial hace posibles ataques más complejos y debilita la eficiencia de los métodos de protección tradicionales. Las máquinas destacan por velocidad, escala y alcance. Los algoritmos pueden lanzar ataques en milisegundos e infectar millones de ordenadores. Además, pueden escanear código informático a gran escala en busca de determinados tipos de vulnerabilidades<sup>60</sup>.
- Los piratas informáticos pueden emplear algoritmos de inteligencia artificial para descubrir vulnerabilidades en los sistemas de seguridad de los bancos<sup>61</sup>. El *aprendizaje automático adversario* implica que los sistemas impulsados por AI tienen vulnerabilidades

inherentes que deben abordarse al utilizar dichos sistemas en ciberseguridad<sup>62</sup>.

- Un sistema de AI puede ser en sí mismo el objetivo de un delito. Los piratas informáticos pueden atacar el sistema de AI para que la seguridad de los bancos se comporte de manera errática para causar daños<sup>63</sup>.
- La AI permite reducir las barreras de entrada a nuevos cibercriminales, que utilizan esta tecnología para ciber-delinquir. A diferencia de muchos delitos tradicionales los que se producen en el ámbito digital suelen ser muy reproducibles: una vez desarrolladas, las técnicas pueden compartirse, repetirse e incluso venderse, lo que permite la posible comercialización de técnicas delictivas. Los delincuentes pueden subcontratar los aspectos más difíciles de sus delitos basados en la AI<sup>64</sup>. Se está produciendo, de hecho, una automatización en la creación de modelos de fraude y una comercialización como servicio

<sup>60</sup> Véase Schneier (2018).

<sup>61</sup> Véase Kaloudi y Li (2020).

<sup>62</sup> Véase Geluvaraj *et al.* (2019).

<sup>63</sup> Véase Wolff (2020).

<sup>64</sup> Véase Caldwell *et al.* (2020).

(*malware-as-a-service*), lo que permite a los delincuentes con menos habilidades monetizar

el cibercrimen también en el ámbito financiero<sup>65</sup>.

## 11.8. PROBLEMAS ÉTICOS: PRIVACIDAD, PREJUICIOS Y DISCRIMINACIÓN

Las directrices éticas publicadas por la Comisión Europea, que ilustran tanto la importancia de las cuestiones éticas como los límites con otros principios, incluyen: agencia humana y supervisión, robustez técnica y seguridad, privacidad y gobernanza de datos, transparencia, diversidad, no discriminación y equidad, bienestar social y ambiental, y responsabilidad. Estas directrices subrayan el amplio espectro de desafíos relacionados con la ética y la justicia en el ámbito de la inteligencia artificial. El uso de herramientas de AI puede originar problemas de carácter ético relacionados con: i) el uso de datos personales; ii) la inclusión de prejuicios en determinadas decisiones y recomendaciones y iii) la emergencia de situaciones de discriminación.

En el caso de la industria financiera cabe destacar al menos cuatro tipos de riesgos: falta de

privacidad de los datos, sesgos en la concesión de créditos y préstamos a clientes, sesgos en las recomendaciones de inversión a clientes y discriminación en la contratación de trabajadores de las propias entidades.

Se ha destacado como uno de los riesgos asociados al uso de herramientas de AI en la industria financiera la mayor probabilidad de disolución de la privacidad individual. La tecnología de inteligencia artificial aumenta la capacidad de recopilar, analizar y sintetizar cantidades increíblemente grandes de datos, de una gran variedad de fuentes y a una escala antes imposible. Además, este tipo de tecnología está ahora ampliamente disponible para personas y organizaciones en todo el mundo. Si bien la tecnología computacional anteriormente ya impactaba negativamente la privacidad, las nuevas capacidades

<sup>65</sup> Véase Chan (2023).

de la AI contribuyen a aumentar el potencial de daños<sup>66</sup>. Si no se controlan, las capacidades de los sistemas de inteligencia artificial podrían representar una amenaza fundamental a la privacidad tanto durante su funcionamiento como por la posibilidad de que estos sistemas no sean inmunes a la filtración de información en condiciones adversas. En este ámbito, se han distinguido tres tipos básicos de problemas:

Por un lado, la *privacidad en la adquisición de datos* en el contexto de *big data*, en el que millones de datos son recopilados, estructurados y vendidos. Existe el riesgo de que información privada sea utilizada ilegalmente con fines comerciales. Además, con la introducción de la *computación en la nube*, muchas empresas están migrando sus datos a estos entornos porque es una estrategia barata, fácil de usar y conveniente para obtener acceso a la red bajo demanda. Los mayores requisitos de computación de la inteligencia artificial se han traducido en que la *computación en la nube* se ha configurado como la infraestructura principal de muchas aplicaciones de AI. Pero ello implica que la privacidad

de los datos de los clientes debe estar garantizada también en dicho entorno, y los clientes de estos servicios difícilmente pueden monitorizar este riesgo.

Por lo que respecta a la *privacidad en la extracción de conocimientos*, a medida que las herramientas utilizadas a este fin se vuelven cada vez más poderosas, se pueden integrar numerosos fragmentos de datos aparentemente no relacionados para identificar características de comportamiento individuales, exponiendo así la privacidad personal. Se puede delinear el mapa de comportamiento de una persona y analizar sus preferencias personales y hábitos de comportamiento, prediciendo así en mayor medida las necesidades potenciales de los clientes. La personalización se ha convertido en una característica destacada de las aplicaciones inteligentes actuales, pero estos procesos pueden conllevar una invasión de la privacidad<sup>67</sup>.

El análisis de privacidad incluye la tabulación de todas las características funcionales y técnicas del sistema. Compartir datos involucrando identidades

<sup>66</sup> Véase Curzon et al. (2021).

<sup>67</sup> Véase Li y Zhang (2017).

de individuos, implica un cambio en las protecciones existentes. Por su parte, en cuanto a las características técnicas, el análisis debe distinguir si implica observar o rastrear personas (por ejemplo, cámaras de video, teléfonos celulares o GPS) o una comunicación a través de medios que no son de confianza<sup>68</sup>. La evaluación de impacto sobre la privacidad demuestra que la efectividad exige identificar los flujos de información, categorizar la información, incluidos todos los tipos de privacidad (personal, de ubicación, de comportamiento, de comunicación, etc.), documentar cómo se procesa la información y quién lo hace, determinar cómo se almacena y protege la información, y enumerar quién tiene acceso a la misma<sup>69</sup>. En este contexto, la *privacidad diferencial* es una herramienta de preservación de la privacidad predominante. Al ocultar a un individuo en la información agregada, se puede preservar la privacidad de los participantes en un conjunto de datos. La *privacidad diferencial* garantiza que, incluso si un agente malicioso conoce todos los demás registros de un conjunto de datos

excepto los de un individuo desconocido, no pueda inferir la información de dicho registro<sup>70</sup>.

El segundo riesgo de carácter ético tiene que ver con la existencia de sesgos. En el ámbito de la inteligencia artificial, junto con una validación insuficiente de los modelos, la inmadurez de los planes de contingencia y otras amenazas, la existencia de sesgos involuntarios también expone a la empresa financiera a riesgos operativos y de reputación<sup>71</sup>. Existe la posibilidad de que existan sesgos en el aprendizaje automático introducidos durante proceso de desarrollo (*development bias*). Esta circunstancia es el resultado de la repetición iterativa del ciclo de desarrollo, que se repite hasta que se consigue un determinado resultado. En la modelización estadística tradicional este sesgo puede ser consecuencia de un ajuste continuado del modelo mediante cambios en la muestra, cambios en las variables explicativas, o transformación de las variables de entrada. Un comportamiento conocido como *p-value hacking*.

<sup>68</sup> Véase Curzon et al. (2021).

<sup>69</sup> Véase Curzon et al. (2021).

<sup>70</sup> Véase Zhu et al. (2020).

<sup>71</sup> Véase Zhang y Zhou (2019).

En el caso del aprendizaje automático, la complejidad es mucho mayor y resulta más difícil detectar la manipulación, ya que los desarrolladores suelen tener que intervenir para lograr una mejor solución de aprendizaje automático. Por ejemplo, en las redes neuronales son posibles organizaciones alternativas de neuronas, organizadas en las diferentes capas. Y, en algoritmos del tipo SVM, los conjuntos hiperparámetros deben ser ajustados<sup>72</sup>.

Por otra parte, se ha demostrado que, aunque los algoritmos en sí mismos no contengan sesgos, el uso de datos históricos sesgados puede generar sesgos en los resultados automatizados, lo que puede derivar en un trato injusto de los clientes, de los empleados y/o de los potenciales empleados. Además, la experiencia ha mostrado que un algoritmo de aprendizaje automático suficientemente sofisticado para mejorar sus predicciones es capaz de re-incorporar en el análisis de forma autónoma algunas variables nuevas y datos no proporcionados como inputs. Y ello incluye datos que habían sido

intencionadamente eliminados por el analista para evitar sesgos<sup>73</sup>. Por ejemplo, esto puede ocurrir en la evaluación de calificaciones crediticias personales dentro de los procesos de concesión de préstamos. Las instituciones financieras deben asegurarse de que sus sistemas de AI sean justos y transparentes, y no discriminen de forma arbitraria a determinados colectivos de personas o de empresas.

La inteligencia artificial es tan propensa a sesgos como la humanidad, pero la buena noticia es que los sesgos de los algoritmos también pueden diagnosticarse y tratarse<sup>74</sup>. Para combatir el problema de los sesgos, que pueden traducirse en resultados discriminatorios y problemas reputacionales para las entidades, debe llevarse a cabo una investigación activa que comprende esquemáticamente las siguientes etapas:

- Definición de qué constituye un sesgo problemático (ya sea de clasificación, de predicción o estadístico) presente en los datos de entrada.

<sup>72</sup> Véase Véase Quell et al. (2021).

<sup>73</sup> Véase Véase Quell et al. (2021).

<sup>74</sup> Véase Baer y Kamalnath (2017) y Holzfeind (2023).



- Establecer las métricas para caracterizar y cuantificar dichos sesgos.
- Determinar en qué medida los sesgos presentes en los datos se reflejan en los algoritmos de AI.
- Mitigar esos sesgos cuando sea posible, ya sea a nivel de datos o del algoritmo.

Se ha demostrado que el uso de características sensibles es insuficiente para eliminar sesgos, debido a la influencia indirecta de información sensible. Y se ha propuesto un enfoque de regularización que es aplicable a cualquier algoritmo de predicción con modelos discriminativos probabilísticos<sup>75</sup>. En todo caso, resulta complicado eliminar completamente la generación de resultados que reflejan falta de equidad residual en presencia de datos con sesgos, incluso en el caso de aplicar algoritmos “justos” de aprendizaje automático<sup>76</sup>.

Existe una creciente atención al análisis de las causas que subyacen en la toma de decisiones injustas cuando se utilizan algoritmos de aprendizaje

automático. Es decir, decisiones algorítmicas que discriminan a unos grupos sobre otros, especialmente con grupos que se definen sobre atributos protegidos, como género, raza, nacionalidad, etc. En este ámbito, se está investigando un problema diferente de la existencia de sesgos en los datos y que se refiere a la “falta de datos” dentro de las muestras, como causa posible de resultados discriminatorios. Los *valores faltantes* son una manifestación frecuente de este problema y se debe a que determinados grupos protegidos pueden ser más reacios a dar información que podría usarse en su contra. Además, la información sensible para algunos grupos puede ser borrada por operadores humanos, o la adquisición de datos puede ser simplemente menos completa y sistemática para algunos grupos<sup>77</sup>. Dos vías alternativas para mitigar este problema es la aplicación de nuevos métodos para imputar *valores faltantes* (basados en técnicas de aprendizaje automático como las redes generativas adversarias, los conjuntos difusos, los modelos generativos profundos y los

<sup>75</sup> Véase Kamishima *et al.* (2012).

<sup>76</sup> Véase Kallus y Zhou (2018).

<sup>77</sup> Véase Martínez-Plumed *et al.* (2021).

codificadores automáticos). Y, por otra parte, diseñar métodos de generación artificial de datos faltantes sintéticos<sup>78</sup>.

Finalmente, en relación con la selección de personal, los bancos y las compañías de seguros utilizan cada vez más algoritmos de inteligencia artificial para evaluar a los solicitantes de empleo, así como *chatbots* para realizar entrevistas en línea. A diferencia de las entrevistas por video sincrónicas (SVI) “bidireccionales” las AVI se conocen como entrevistas “unidireccionales”. En éstas últimas, las respuestas de los solicitantes de empleo se registran en un momento dado y pueden evaluarse en otro momento. Para aumentar la disponibilidad de recursos de personal, esta tecnología ha comenzado a desplazar las tradicionales entrevistas telefónicas. Algunas empresas desarrollan sus propias herramientas de AI para estas tareas y otros subcontratan los sistemas a proveedores externos. Estas herramientas clasifican a los empleados potenciales en segmentos para optimizar los efectos del entrenamiento para los sistemas de AI a través de programas de desarrollo

personalizados. Los algoritmos son capaces de evaluar el desempeño de los candidatos en función de cientos de rasgos diferentes, incluidos las respuestas, pero también sus movimientos oculares, el lenguaje corporal, los patrones del habla y el parpadeo. Los defensores de la AI en el reclutamiento afirman que el uso de herramientas de inteligencia artificial reduce el coste y minimiza los efectos potenciales de prejuicios y la intuición propios del comportamiento humano. Sin embargo, una posible consecuencia no deseada es la discriminación contra grupos de personas, si existe un sesgo inherente en los datos de entrada. Además, el uso de diferentes modalidades de entrevista por video (sincrónica versus asincrónica) y diferentes agentes de decisión (humanos versus AI) puede provocar diferentes respuestas tanto de los entrevistadores como de los solicitantes de empleo. Esta circunstancia puede provocar situaciones de discriminación o impactos sociales no deseados<sup>79</sup>.



<sup>78</sup> Véase Martínez-Plumed et al. (2021).

<sup>79</sup> Véase Straus et al. (2001) Torres y Mejía (2017) y Ver Haan et al. (2017).

## REFERENCIAS

- Acemoglu, D. and Autor, D. (2011) "Skills, tasks and technologies: Implications for employment and earnings" en *Handbook of Labor Economics*, 4B, eds O. Ashenfelter and D. Card, Elsevier, Amsterdam.
- Acemoglu, D. and Restrepo, P. (2016) "The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment", *NBER Working Paper*, 22252.
- Agrawal, A., Gans, J. S. and Goldfarb, A. (2019) "Artificial intelligence: the ambiguous labor market impact of automating prediction", *Journal of Economic Perspectives*, 33 (2): 31-50.
- AL-Dosari, K., Fetais, N., and Kucukvar, M. (2022) "Artificial Intelligence and Cyber Defense System for Banking Industry: A Qualitative Study of AI Applications and Challenges", *Cybernetics and Systems*, 1-29.
- Aleksandrova, A., Ninova, V., and Zhelev, Z. (2023) "A Survey on AI Implementation in Finance, (Cyber) Insurance and Financial Controlling", *Risks*, 11(5): 91.
- Arntz, M., Gregory, T. and Zierahn, U. (2016) "The risk of automation for jobs in OECD countries: A comparative analysis", *OECD Social, Employment and Migration Working Papers*.
- Athaluri, S. A., Manthena, S. V., Kesapragada, V. K. M., Yarlagadda, V., Dave, T., and Duddumpudi, R. T. S. (2023) "Exploring the boundaries of reality: investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references", *Cureus*, 15(4).
- Autor, D. H. and Dorn, D. (2013) "The growth of low-skill service jobs and the polarization of the US labor market", *American Economic Review*, 103 (5): 1553-1597.
- Baer, T. and Kamalnath, V. (2017) "Controlling machine-learning algorithms and their biases", McKinsey and Co.
- Behrend, T. S. and Thompson, L. F. (2011) "Similarity effects in online training: Effects with computerized trainer agents", *Computers in Human Behavior*, 27(3): 1201-1206.
- Beutel, G., Geerits, E., and Kielstein, J. T. (2023) "Artificial hallucination: GPT on LSD?" *Critical Care*, 27(1): 148.
- Borland, J. and Coelli, M. (2017) "Are robots taking our jobs?", *Australian Economic Review*, 50 (4): 377-397.
- Brenner, F. S., Ortner, T. M., and Fay, D. (2016) "Asynchronous video interviewing as a new technology in personnel selection: The applicant's point of view", *Frontiers in Psychology*, 7: 863.
- Brynjolfsson, E. and McAfee, A. (2011) *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*.
- Bussmann, N., Giudici, P., Marinelli, D., and Papenbrock, J. (2020) "Explainable AI in fintech risk management", *Frontiers in Artificial Intelligence*, 3, 26.
- Caldwell, M., Andrews, J. T., Tanay, T., and Griffin, L. D. (2020) "AI-enabled future crime", *Crime Science*, 9 (1), 1-13.

- Campbell-Kelly, M. (2009) "Origin of computing", *Scientific American*, 301(3): 62-69.
- Çetin-Gürkan, G., and Çiftci, G. (2020) "Developing a supportive culture in digital transformation". *Digital Business Strategies in Blockchain Ecosystems: Transformational Design and Future of Global Business*: 83-102.
- Chan, A. (2023) "Can AI Be Used for Risk Assessments?", *CISM*. 28 de abril.
- Chow, C., Frame, K., Likhtman, S., Spooner, N. y Wong, J. (2019) *Investor's expectations on responsible artificial intelligence and data governance*. Hermes Investment Management.
- Coelli, M. and Borland, J. (2017) "Estimating the automation risk of Australian jobs", *Department of Economics Working Paper, University of Melbourne*.
- Curzon, J., Kosa, T. A., Akalu, R., and El-Khatib, K. (2021) "Privacy and artificial intelligence", *IEEE Transactions on Artificial Intelligence*, 2 (2): 96-108.
- Das, A. and Rad, P. (2020) "Opportunities and challenges in explainable artificial intelligence (XAI): A survey", *arXiv:2006.11371*.
- Dobbs, R., J. Manyika and J. Woetzel (2015) "The Four Global Forces Breaking all the Trends", *McKinsey Global Institute*. London, San Francisco, Shanghai.
- Frey, C. B. and Osborne, M. A. (2017) "The future of employment: How susceptible are jobs to computerisation?" *Technological forecasting and social change*, 114: 254-280.
- Gaggl, P. and Wright, G. (2017) "A short-run view of what computers do: Evidence from a UK tax incentive", *American Economic Journal: Applied Economics*, 9 (3): 262-94.
- Gelubaraj, B., Satwik, P. M. and Ashok-Kumar, T. A. (2019) "The future of cybersecurity: Major role of artificial intelligence, machine learning, and deep learning in cyberspace" en *International Conference on Computer Networks and Communication Technologies: 739-747*. Springer, Singapore.
- Goldin, C. and Katz, L. F. (2009) *The race between education and technology*. Harvard University Press.
- Goos, M. and Manning, A. (2007) "Lousy and lovely jobs: The rising polarisation of work in Britain", *Review of Economics and Statistics*, 89: 118-33.
- Goos, M., Manning, A., and Salomons, A. (2009) "Job polarization in Europe", *American Economic Review*, 99 (2): 58-63.
- Haan, M., Ongena, Y. P., Vannieuwenhuyze, J. T., and De Glopper, K. (2017) "Response behavior in a video-web survey: A mode comparison study", *Journal of Survey Statistics and Methodology*, 5(1): 48-69.
- Hall, P. and Gill, N. (2019). *An introduction to machine learning interpretability*. O'Reilly Media, Incorporated.
- Hassani, H., Huang, X., and Silva, E. (2018) "Banking with blockchain-ed big data", *Journal of Management Analytics*, 5(4): 256-275.
- Hoff K. A. and Bashir M. (2015) "Trust in automation: Integrating empirical evidence on factors that influence trust", *Human Factors*, 57: 407-434.
- Holzfeind, D. (2023) "Overcoming bias in machine learning", *McKinsey Classic*, McKinsey and Co. Octubre.

- Hooper, R. S., Galvin, T. P., Kilmer, R. A., and Liebowitz, J. (1998) "Use of an expert system in a personnel selection process", *Expert Systems with Applications*, 14(4), 425-432.
- Hu, X. and Wang, K. (2020) "Bank financial innovation and computer information security management based on artificial intelligence" en *2nd international conference on machine learning, Big Data and Business Intelligence (MLBDBI)*: 572-575). IEEE. Octubre.
- Kallus, N. and Zhou, A. (2018) "Residual unfairness in fair machine learning from prejudiced data" en *International Conference on Machine Learning*: 2439-2448). PMLR. Julio.
- Kaloudi, N and Li, J. (2020) "The AI-based cyber threat landscape: A survey", *ACM Computing Surveys*, 53 (1): 1-34.
- Kamishima, T., Akaho, S., Asoh, H., and Sakuma, J. (2012) "Fairness-aware classifier with prejudice remover regularizer" en *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2012, Bristol, UK, September 24-28, 2012. Proceedings, Part II* 23: 35-50. Springer Berlin Heidelberg.
- Kouw, W. M. and Loog, M. (2018) "An introduction to domain adaptation and transfer learning", *arXiv:1812.11806*.
- Kouw, W. M. and Loog, M. (2019) "A review of domain adaptation without target labels", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(3): 766-785.
- Kruse, L., Wunderlich, N., and Beck, R. (2019) "Artificial intelligence for the financial services industry: What challenges organizations to succeed", *ScholarSpace*.
- Kumar, V., Saheb, S. S., Ghayas, A., Kumari, S., Chandel, J. K., Pandey, S. K., and Kumar, S. (2023) "AI-Based Hybrid Models for Predicting Loan Risk in the Banking Sector", *Big Data Mining and Analytics*, 6(4): 478-490.
- Kumar, A., Srivastava, A., and Gupta, P. K. (2022) "Banking 4.0: The era of artificial intelligence-based fintech", *Strategic Change*, 31 (6): 591-601.
- Lai, S. T., Leu, F. Y., and Lin, J. W. (2018) "A banking chatbot security control procedure for protecting user data security and privacy" en *International conference on broadband and wireless computing, communication and applications*:561-71. Springer, Cham. Octubre.
- Lawler, J. J., and Elliot, R. (1996) "Artificial intelligence in HRM: An experimental study of an expert system", *Journal of Management*, 22(1): 85-111.
- Li, Y., Du, Y., Zhou, K., Wang, J., Zhao, W. X., and Wen, J. R. (2023) "Evaluating object hallucination in large vision-language models". *arXiv*: 2305.10355.
- Li, X. and Zhang, T. (2017) "An exploration on artificial intelligence application: From security, privacy and ethic perspective" en *IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*: 416-420). IEEE. Abril
- Lin, R., Gal, Y., Kraus, S., and Mazliah, Y. (2014) "Training with automated agents improves people's behavior in negotiation and coordination tasks", *Decision Support Systems*, 60: 1-9.
- Lindebaum, D., Vesa, M., and den Hond, F. (2020) "Insights from The Machine Stops to better understand rational assumptions in algorithmic decision-making and its implications for organizations", *Academy of Management Review*, 45(1), 247-263.
- Lockey, S., Gillespie, N., Holm, D., and Someh, I. A. (2021) "A review of trust in artificial intelligence:

- Challenges, vulnerabilities and future directions”, *Proceedings of the 54th Hawaii International Conference on System Sciences*.
- Martínez-Plumed, F., Ferri, C., David, N., Hernández-Orallo J. (2021) “Missing the missing values: The ugly duckling of fairness in machine learning”, *International Journal of Intelligent Systems*, 36 (7): 3217-3258.
- Meinert, M. C. (2018). Artificial Intelligence: The Next Frontier of Cyber Warfare?. *American Bankers Association. ABA Banking Journal*, 110 (3): 43-43.
- Mor, S. and Gupta, G. (2021) “Artificial intelligence and technical efficiency: The case of Indian commercial banks”, *Strategic Change*, 30(3): 235-245.
- Naim, I., Tanveer, M. I., Gildea, D., and Hoque, M. E. (2015) “Automated prediction and analysis of job interview performance: The role of what you say and how you say it” en *11th IEEE international conference and workshops on automatic face and gesture recognition (FG)*: 1-6). IEEE. Mayo.
- Petropoulos, G. (2018) “The impact of artificial intelligence on employment”, en *Work in The Digital Age*: 119-132. Editors M. Neufeind, J. O’Reilly and F. Ranft. *Foundation for European Progressive Studies (FEPS)*.
- Poba-Nzaou, P., Galani, M., Uwizeyemungu, S., and Ceric, A. (2021) “The impacts of artificial intelligence (AI) on jobs: an industry perspective”, *Strategic HR Review*, 20(2): 60-65.
- Quell, P., Bellotti, A. G., Breeden, J. L., and Martin, J. C. (2021) “Machine learning and model risk management”. *MRMIA Best Practices*, 1.
- Rasipuram, S., Rao, S. P., and Jayagopi, D. B. (2016) “Automatic prediction of fluency in interface-based interviews” en *IEEE Annual India Conference (INDICON)*: 1-6. IEEE. Diciembre.
- Redko, I., Morvant, E., Habrard, A., Sebban, M., and Bennani, Y. (2020) “A survey on domain adaptation theory: learning bounds and theoretical guarantees”, arXiv:2004.11829.
- Said, N., Potinteu, A. E., Brich, I., Buder, J., Schumm, H., and Huff, M. (2023) “An artificial intelligence perspective: How knowledge and confidence shape risk and benefit perception”, *Computers in Human Behavior*: 107855.
- Sajjadiani, S., Sojourner, A. J., Kammeyer-Mueller, J. D., and Mykerezzi, E. (2019) “Using machine learning to translate applicant work history into predictors of performance and turnover”, *Journal of Applied Psychology*, 104 (10): 1207-1225.
- Salem, M., Lakatos, G., Amirabdollahian, F., and Dautenhahn, K. (2015) “Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust” en *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*:141-148. Marzo.
- Sanders, T., Kaplan, A., Koch, R., Schwartz, M., and Hancock, P. A. (2019) “The relationship between trust and use choice in human-robot interaction”, *Human Factors*, 61(4): 614-626.
- Satheesh, M. K., and Nagaraj, S. (2021) “Applications of artificial intelligence on customer experience and service quality of the banking sector”, *International Management Review*, 17(1): 9-17.
- Schneier, B. (2018) “Artificial intelligence and the attack/defense balance”, *IEEE Security & Privacy*, 16(02): 96-96.
- Shmuratko, Y. A. and Sheludko, S. A. (2019) “Financial Technologies’ Impact on the Development of Banking”, *Financial and credit activity problems of theory and practice*, 4 (31): 61-69.



- Stojanov, P., Li, Z., Gong, M., Cai, R., Carbonell, J., y Zhang, K. (2021) "Domain adaptation with invariant representation learning: What transformations to learn?", *Advances in Neural Information Processing Systems*, 34: 24791-24803.
- Straus, S. G., Miles, J. A., and Levesque, L. L. (2001) "The effects of videoconference, telephone, and face-to-face media on interviewer and applicant judgments in employment interviews", *Journal of Management*, 27(3): 363-381.
- Sudjianto, A., Knauth, W., Singh, R., Yang, Z., and Zhang, A. (2020) "Unwrapping the black box of deep RELU networks: Interpretability, diagnostics, and simplification", *arXiv:2011.04041*.
- Suen, H. Y., Chen, M. Y. C., and Lu, S. H. (2019) "Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes?" *Computers in Human Behavior*, 98: 93-101.
- Sun, W., Shi, Z., Gao, S., Ren, P., de Rijke, M., and Ren, Z. (2023) "Contrastive learning reduces hallucination in conversations" en *Proceedings of the AAAI Conference on Artificial Intelligence*, 37 (11): 13618-13626. Junio.
- Taddeo, M., McCutcheon, T., and Floridi, L. (2021) "Trusting Artificial Intelligence in Cybersecurity Is a Double-Edged Sword", *Ethics, Governance, and Policies in Artificial Intelligence*: 289-297.
- Torres, E. N., and Gregory, A. (2018) "Hiring manager's evaluations of asynchronous video interviews: The role of candidate competencies, aesthetics, and resume placement", *International Journal of Hospitality Management*, 75: 86-93.
- Torres, E. N. and Mejia, C. (2017) "Asynchronous video interviews in the hospitality industry: Considerations for virtual employee selection", *International Journal of Hospitality Management*, 61: 4-13.
- Truby, J., Brown, R. and Dahdal, A. (2020) "Banking on AI: Mandating a proactive approach to AI regulation in the financial sector", *Law and Financial Markets Review*, 14 (2): 110-20.
- Van Esch, P., Black, J. S., and Ferolie, J. (2019) "Marketing AI recruitment: The next phase in job application and selection", *Computers in Human Behavior*, 90: 215-222.
- Vrontis, D., Christofi, M., Pereira, V., Tarba, S., Makrides, A., and Trichina, E. (2022). "Artificial intelligence, robotics, advanced technologies and human resource management: a systematic review", *The International Journal of Human Resource Management*, 33(6), 1237-1266.
- Wilson, H. J., Daugherty, P., and Bianzino, N. (2017) "The jobs that artificial intelligence will create", *MIT Sloan Management Review*, 58(4): 14.
- Wolff, J. (2020) "How to improve cybersecurity for artificial intelligence", *Brookings Institution. United States of America*.
- Yang, W., Wei, Y., Wei, H., Chen, Y., Huang, G., Li, X., ... and Kang, B. (2023) "Survey on Explainable AI: From Approaches, Limitations and Applications Aspects", *Human-Centric Intelligent Systems*: 1-28.
- Yang, Z., Zhang, A., and Sudjianto, A. (2020) "Enhancing explainability of neural networks through architecture constraints", *IEEE Transactions on Neural Networks and Learning Systems*, 32(6): 2610-2621.
- Zhang, C. A., Cho, S., and Vasarhelyi, M. (2022) "Explainable artificial intelligence (xai) in auditing", *International Journal of Accounting Information Systems*, 46: 100572.



- Zhang, Y. and Zhou, L. (2019) "Fairness assessment for artificial intelligence in financial industry", *arXiv:1912.07211*.
- Zhu, T., Ye, D., Wang, W., Zhou, W., and Philip, S. Y. (2020) "More than privacy: Applying differential privacy in key areas of artificial intelligence", *IEEE Transactions on Knowledge and Data Engineering*, 34(6): 2824-2843.



## 12. Conclusiones

La inteligencia artificial es la ciencia informática aplicada a la simulación de algunas de las funciones cognitivas de los cerebros biológicos. La AI es una tecnología de propósito general polivalente que tendrá en nuestras vidas una repercusión al menos tan fuerte como la que implicó la penetración del ordenador personal, Internet y los *smart phones*.

Estamos tan sólo al “principio del principio” de la era de la inteligencia artificial, pero la capacidad de computación para entrenar modelos de aprendizaje automático se ha multiplicado ya por un factor de cien millones en los últimos 10 años. En consecuencia, no se descarta la posibilidad de que en algún momento pueda llegar a existir una *superinteligencia artificial*.

La industria financiera ni siquiera ha arañado la superficie de cómo la AI puede transformar su operativa y la relación con sus clientes, pero parece claro que esta tecnología tiene

el potencial de cambiar la naturaleza y sus fronteras.

La AI puede contribuir a que los servicios financieros sean más baratos, más rápidos, más accesibles y más eficiente en muchos sentidos. Y las entidades financieras pueden aspirar a convertirse en el centro de la vida financiera de sus clientes y parte de un ecosistema superior.

El impacto aislado de la AI previsiblemente sea incluso superior en la industria financiera que el que han tenido otras tecnologías emergentes como Internet de las Cosas,

el Metaverso, e incluso la *distributed ledger technology* (DLT). Pero es que, además, los impactos de todas estas nuevas tecnologías se refuerzan entre sí.

Como ha ocurrido con Fintech y con DeFI anteriormente, la inteligencia artificial puede debilitar los lazos que han mantenido unidos los componentes de las instituciones financieras tradicionales, y abrir la puerta a más innovaciones y nuevos modelos operativos. Esto representa una gran oportunidad, pero también una seria amenaza para los incumbentes actuales de la industria.

Las entidades financieras que operan en los diferentes subsectores de la industria están utilizando, de forma más o menos común, herramientas de inteligencia artificial en ámbitos tan diversos como los procesos administrativos, documentales y de gestión de información, contabilidad y auditoría, gestión de riesgos, prevención del fraude, cumplimiento normativo, gestión del cambio, gestión del conocimiento, e incluso en proyectos de educación financiera.

En el caso de la banca comercial, además de automatizar con técnicas de AI buena parte

de los procesos de *back&middle office*, la inteligencia artificial está utilizándose en tareas de marketing para lograr una hiperpersonalización de las campañas. En la atención a los clientes se están usando *chatbots* y *asistentes virtuales inteligentes* de forma creciente para mejorar la experiencia del cliente. También es está usando la AI en el *credit scoring* de préstamos hipotecarios y de financiación al consumo, así como en la gestión de pagos. Del mismo modo, algoritmos basados en inteligencia artificial se aplican en los modelos internos de riesgo, para una determinación más precisa de los requerimientos de capital regulatorio.

En la banca privada y la banca personal se está produciendo un uso progresivo de algoritmos de inteligencia artificial en los diferentes eslabones de la cadena de producción. En este ámbito se espera incluso que se imponga un modelo híbrido (humanos y máquinas) de gestión y atención comercial. Desde la perspectiva de la situación competitiva, se ha afirmado que las entidades de banca privada, que ofrecen servicios más o menos *premium*, podrían ser algo más inmunes a la generalización de los modelos basados en robo-advisors. Respecto del segmento de banca

personal, se argumenta que el asesoramiento automatizado puede ser tanto una ayuda muy valiosa, pero también significar una auténtica amenaza.

En el ámbito de la banca de inversión la inteligencia artificial está siendo gradualmente aplicada, mediante algoritmos de negociación automatizada y, en algunos casos, trading autónomo. También se utiliza en la predicción de precios y volatilidades de diferentes tipos activos financieros. Del mismo modo, se están usando herramientas de AI en la gestión del riesgo de los modelos aplicados en el trading que los bancos realizan empleando sus propios balances. Y los bancos están apoyándose en algoritmos de inteligencia artificial para tareas de análisis en procesos de fusiones y adquisiciones y de asesoramiento en mercados primarios de capitales.

En la banca corporativa, se está utilizando la inteligencia artificial en la evaluación crediticia (calificación de riesgo de crédito, estimación de probabilidades de quiebra y predicción de tasas de recuperación).

En la actividad de *asset management*, los gestores de fondos de inversión, de

pensiones, de capital privado y los *hedge funds* están empleando herramientas de AI en la búsqueda de *alfa*, y en el diseño de estrategias eficientes de cobertura.

Los gestores de capital privado y *venture capital* están utilizando algoritmos de inteligencia artificial y *big data* para detectar oportunidades de inversión, encontrando compañías objetivo no cotizadas que encajan en sus estrategias de inversión. También utilizan herramientas de AI para ayudar a las empresas ya en las carteras de sus fondos a la hora de evaluar opciones estratégicas de crecimiento. Y, finalmente, utilizan algoritmos inteligentes en proceso de desinversión para identificar potenciales compradores.

En la industria del seguro, un sector históricamente basado en el uso intensivo de datos, el uso de la inteligencia artificial está transformando toda la operativa a lo largo de su cadena de valor: desde la comercialización, el diseño de productos, la evaluación del riesgo y la fijación del precio de las pólizas, hasta los procesos de suscripción, el servicio de atención al cliente y la gestión de reclamaciones. Mediante herramientas de AI y el uso *big data* es posible identificar grupos de riesgo



más pequeños y homogéneos y mejorar la precisión de la predicción sobre exposición y pérdida. La inteligencia artificial permite mitigar los problemas de *selección adversa y moral hazard* y, *por ende, reducir el fraude*. Además, la AI amplía el ámbito de aseguramiento, al surgir nuevos mercados de microcoberturas y nuevos riesgos derivados de la generalización del uso de AI.

Pero no sólo el sector privado tiene interés en adoptar tecnologías de inteligencia artificial, también las entidades públicas y otros organismos institucionales en el ámbito financiero están incorporando estas herramientas en su gestión.

Los bancos centrales utilizan software AI para mejorar su comprensión de la evolución económica y financiera, hacer más eficiente el diseño de sus políticas monetarias, y mejorar su evaluación del riesgo sistémico. Las autoridades monetarias encargadas de la estabilidad del sistema financiero utilizan también la inteligencia artificial para mejorar la supervisión micro y macro prudencial, y en la realización de pruebas de resistencia de los bancos (stress test), así como para evaluar la gobernanza y la cultura

del riesgo en las entidades financieras supervisadas.

Los supervisores de los mercados de valores están encontrando en los sistemas de AI un aliado eficiente para mejorar la detección de patrones de manipulación de precios, de conductas de abuso de mercado, así como de prácticas de *insider trading*. Sin embargo, se están también produciendo casos de manipulación de precios derivados de la aplicación de estrategias por parte de sistemas inteligentes de *trading* autónomo, en los que resulta difícil imputar intención a los desarrolladores. Esto plantea para los supervisores del mercado un desafío notable en la asignación de responsabilidades.

Las agencias gubernamentales que velan por la competencia en el mercado comienzan a emplear herramientas de inteligencia artificial para detectar comportamientos colusorios. Pero se enfrentan a un dilema, ya que la recopilación y el procesamiento de grandes cantidades de datos por parte de sus algoritmos inteligentes pueden implicar un *trade-off* entre protección de la competencia y privacidad de los consumidores. Además, se está poniendo de manifiesto que los algoritmos de inteligencia artificial

pueden utilizarse para generar y enmascarar comportamientos anticompetitivos (*colusión algorítmica*) que resultan ser más complejos y difíciles de detectar.

La inteligencia artificial desempeña un papel central en el proceso de cambio tecnológico actual en la industria de los servicios financieros. Su lugar destacado en las agendas de innovación indica la percepción sobre los importantes beneficios que estas tecnologías pueden aportar a las entidades financieras, a los consumidores y a los mercados. La adopción de tecnologías de inteligencia artificial promete acelerar la innovación, mejorar el marketing mediante la generación de campañas de publicidad más enfocadas, reducir los costes de cumplimiento corporativo y de *back-office* mediante la automatización de tareas, así como aumentar la productividad (al poder destinar los recursos humanos a tareas de mayor valor añadido), mientras que asistentes virtuales contribuyen en las tareas de atención al cliente. Pero, al mismo tiempo, los sistemas de AI tienen el potencial de causar daños significativos. La implantación de la inteligencia artificial puede convertirse en una espada de

doble filo si la gestión de los datos no es prudente, si se aplica con objetivos de manipulación, si se traduce en una mayor invasión de la privacidad, o si conduce a una mayor polarización social y a mayor discriminación.

Los desafíos que representa la implantación de la AI en la industria financiera son numerosos. Para las entidades financieras, existe el riesgo de que las millonarias inversiones realizadas sean improductivas, que su implantación tenga un impacto dramático en las políticas de recursos humanos, y que exija cambios profundos en sus estructuras organizativas. La *inteligencia aumentada* que surge de la colaboración entre humanos y máquinas hará más eficientes las organizaciones, pero el clima laboral puede resentirse si los empleados perciben que su autonomía en la toma de decisiones queda debilitada y su seguridad laboral amenazada.

Estos desafíos también se refieren a la necesidad de lidiar con problemas como la opacidad de los resultados obtenidos mediante herramientas AI (el problema de *caja negra*); el riesgo de incurrir en excesos de confianza en los sistemas de inteligencia artificial implantados;



y el riesgo de enfrentarse a una excesiva dependencia de algunos proveedores.

Por otra parte, se pone de manifiesto la necesidad de llevar a cabo una acción proactiva para lograr que la implantación de proyectos de AI sea compatible con criterios éticos. Por ejemplo, en la concesión de créditos y en la selección de empleados.

Finalmente, el desarrollo de la AI puede facilitar la penetración de nuevos competidores en el sector, generando turbulencias en la industria.





INSTITUTO  
ESPAÑOL  
DE ANALISTAS  
DESDE 1965  
FUNDACIÓN

El impacto de la inteligencia artificial en la  
industria financiera: promesas y amenazas

# Anexos





## Anexo A. Detalle de algunas técnicas de inteligencia artificial

“El estado actual de la inteligencia artificial nos pone al borde de algo maravilloso, algo terrible o ambas cosas”.

*Chris Geczy (2024).*

### A.1. SISTEMAS EXPERTOS (ES) Y SISTEMAS BASADOS EN CONOCIMIENTOS (KBS)

Los *sistemas expertos* y los *sistemas basados en conocimientos* (*knowledge-based systems* en su terminología en inglés) son técnicas iniciales de inteligencia artificial<sup>1</sup>. Se trata de sistemas informáticos que contienen un cuerpo de conocimiento bien organizado, que imitan a los expertos y tienen capacidad para resolver problemas en un dominio

limitado de experiencia<sup>2</sup>. Estos *algoritmos basados en reglas* son programas de ordenador capaces de usar información en un ámbito específico de conocimiento, y de utilizar un conjunto de procedimientos de inferencia, para resolver problemas suficientemente complejos (como para que su solución requiriese de experiencia humana significativa en ausencia

<sup>1</sup> Véase Schoech *et al.* (1985).

<sup>2</sup> Véase Jackson (1986), Patterson (1990) y Hunt (2012).



de tales algoritmos). El conjunto de procedimientos de inferencia tiene que ser proporcionado por un experto humano en el área particular de interés, mientras que la base de conocimiento es una acumulación de información relevante que incluye datos, hechos, juicios y resultados.<sup>3</sup> Los sistemas expertos constan de tres componentes principales:

- Base de conocimiento. Es la parte del sistema que almacena el conocimiento adquirido de expertos humanos en forma de reglas, que representan el razonamiento y las estrategias utilizadas por los expertos para los resolver los problemas. El conocimiento se puede representar usando una variedad de técnicas entre las cuales la más utilizada es la regla de producción "si-entonces". La base de conocimientos se puede aumentar continuamente según sea necesario con la experiencia acumulada
- Mecanismo de adquisición de conocimiento. Es el proceso mediante el cual el sistema adquiere el conocimiento de los expertos humanos. Esto puede implicar consultas, análisis de documentos o cualquier otra técnica utilizada para extraer y formalizar el conocimiento experto.
- El motor de inferencia es el componente que realiza el razonamiento y la deducción a partir de la base de conocimiento. Se emplea durante una sesión de consulta, examina el estado de la base de conocimientos, maneja el contenido de la base de conocimientos y determina el orden en que se realizan las inferencias. Utiliza las reglas almacenadas para procesar la información proporcionada y llegar a conclusiones o tomar decisiones.
- La interfaz de usuario habilita comunicación entre el sistema y el usuario. Principalmente incluye visualizaciones en pantalla, un diálogo de consulta y un componente de explicación.
- Los sistemas expertos no están limitados por esquemas matemáticos o analógicos rígidos y pueden manejar conocimiento fáctico o heurístico. Además, pueden

<sup>3</sup> Algunos estudios de interés sobre el tema incluyen Harmon y King (1985), Horvitz *et al.* (1988), Rolston (1988), Sol *et al.* (2013) o Gupta y Nagpal (2020)



dotarse de capacidades para manejar cambios cualitativos en la información y para hacer frente a situaciones inciertas, poco fiables o incluso la pérdida

de datos. Las principales ventajas e inconvenientes de los sistemas expertos se describen en la Tabla A.1.

Tabla A.1. Ventajas e inconvenientes de los sistemas expertos.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o Capacidad para capturar y retener el conocimiento experto.</li> <li>o Capacidad para compartir y distribuir ese conocimiento.</li> <li>o Capacidad de realizar razonamientos consistentes y explicables.</li> </ul>	<ul style="list-style-type: none"> <li>o Dificultad para adquirir y actualizar el conocimiento experto.</li> <li>o Dificultad para lidiar con problemas fuera del alcance de la base de conocimiento obtenido.</li> </ul>

Fuente: elaboración propia.

## A.2. APRENDIZAJE AUTOMÁTICO SUPERVISADO

En las décadas de 1930 y 1940, los pioneros de la computación, incluido el matemático teórico Alan Turing, comenzaron a trabajar en las técnicas básicas para el aprendizaje automático. El término *machine learning* (ML) fue introducido por Arthur Samuel mientras trabajaba para IBM en 1959, para describir las tareas de reconocimiento de patrones que aportaban el componente de "aprendizaje" en los sistemas pioneros de la inteligencia artificial. En un mundo cada vez

más digitalizado se generan datos de forma masiva y un análisis adecuado de los mismos es una poderosa herramienta para mejorar los modelos de negocio, la toma de decisiones e incluso los productos y servicios en todos los sectores. Sin embargo, el volumen y la complejidad de los datos es excesivo para que los humanos puedan tenerlos en cuenta razonablemente sin la ayuda de la computación. En este contexto, la construcción y el entrenamiento de modelos



de aprendizaje se ha convertido en una necesidad y, además, la propia disponibilidad de datos masivos aumenta el potencial del aprendizaje automático. Este proceso conlleva pasos complejos como el preprocesamiento de características, la selección de algoritmos, o el ajuste de hiperparámetros<sup>4</sup>. En la práctica, el proceso de construcción de un modelo de aprendizaje automático de alta calidad es un proceso iterativo, complejo y lento que implica probar diferentes algoritmos y técnicas. Además, exige sólidos conocimientos y experiencia. De acuerdo con la teoría de *no free-lunch*<sup>5</sup>, parece imposible lograr un buen rendimiento para todas estas tareas sin la intervención de expertos humanos<sup>6</sup>. En este contexto, *auto machine learning* o AML por sus siglas en inglés se ocupa de desarrollar métodos que construyan modelos de aprendizaje automático adecuados sin intervención humana (o con la mínima posible)<sup>7</sup>.

En los últimos años, se han

introducido varias técnicas y marcos para abordar el desafío de automatizar el proceso de selección combinada de algoritmos y ajuste de hiperparámetros -*Combined Algorithm Selection and Hyperparameter Tuning* (CASH) en terminología inglesa- en el dominio del aprendizaje automático con el objetivo de reducir el papel del ser humano en este proceso. Se ha generado una creciente demanda de sistemas de aprendizaje automático y la aparición de numerosas ofertas comerciales orientadas a satisfacer esta demanda<sup>8</sup>.

El aprendizaje automático, permite que un ordenador ajuste modelos o reconozca patrones a partir del análisis de los datos proporcionados para obtener respuestas y predicciones a cuestiones diversas. Las técnicas de *machine learning* son utilizadas para entrenar sistemas informáticos a aprender de los datos y a realizar tareas específicas, sin ser explícitamente programados,

<sup>4</sup> Véase Elshawi *et al.* (2019).

<sup>5</sup> Véase Wolpert y Macready (1997).

<sup>6</sup> Véase Yao *et al.* (2018).

<sup>7</sup> Véase Tuggener *et al.* (2019) y Chauhan *et al.* (2020).

<sup>8</sup> Por ejemplo, BigML.com, Wise.io, SkyTree.com, RapidMiner.com, Dato.com, Prediction.io, DataRobot.com, Microsoft's Azure Machine Learning, Google's Prediction API y Amazon Machine Learning. Véase Feurer *et al.* (2015).

es decir sin proporcionar instrucciones directas. AML se basa en procedimientos que van mejorando automáticamente conforme se incorpora más evidencia al algoritmo. A diferencia de los “sistemas basados en reglas” en los que el programador humano decide explícitamente qué decisiones se toman y en qué estados del mundo, ML permite que el algoritmo aprenda con intervención humana limitada o nula.

Como resultado del crecimiento exponencial de la potencia de cálculo y la cantidad de datos disponibles en la última década, los modelos de ML han podido avanzar y se han sofisticado, pudiendo mejorar las predicciones de los modelos tradicionales y encontrar patrones en grandes cantidades de datos procedentes de fuentes cada vez más diversas. Bajo esquemas de AML puede lograrse la automatización de la recopilación de datos, su limpieza y la imputación de *datos faltantes*; también puede automatizarse la selección y transformación de características; la adecuación de los modelos, así como la asignación de recursos informáticos, los hiper-

parámetros, la inferencia, la supervisión de modelos y la detección de anomalías<sup>9</sup>.

El aprendizaje automático incluye técnicas de aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo. El aprendizaje supervisado usa datos clasificados y etiquetados. Cuando el resultado es correcto se produce una retroalimentación. Existen varias técnicas utilizadas para entrenar y construir modelos predictivos a partir de datos de entrenamiento etiquetados. La *regresión lineal*, es una técnica utilizada para predecir un valor numérico continuo. El objetivo es encontrar la línea que mejor se ajusta a los datos de entrenamiento. Reproduce en parte el enfoque estadístico tradicional, en el que hay un rango de variables independientes y un conjunto de datos de entrada y se intenta determinar la relación con la variable dependiente. Pero difiere de éste en que la regresión tradicional utiliza técnicas orientadas a utilizar un gran número de variables como variables independientes y luego descartarlas automáticamente si carecen de poder explicativo. A diferencia del método utilizado en

<sup>9</sup> Véase Hutter *et al.* (2019).

econometría tradicional el modelo teórico necesario para determinar las variables independientes adecuadas no es necesario en este caso<sup>10</sup>. La *regresión logística*,

a diferencia de la regresión lineal, se utiliza para problemas de clasificación binaria. El objetivo es predecir la probabilidad de pertenencia a una clase específica.

### A.3. ALGORITMO DE VECINO MÁS CERCANO

El algoritmo de vecino más cercano –*K-Nearest Neighbors* (KNN) por su terminología en inglés–, es un mecanismo de aprendizaje supervisado utilizado para la clasificación y regresión en el que la predicción se basa en la similitud con los ejemplos de entrenamiento más cercanos en el espacio de características.<sup>11</sup>

El proceso básico del *algoritmo de vecino más cercano* se resume en los siguientes pasos:

- Preparación de los datos de entrenamiento o ejemplos etiquetados, donde cada uno tiene un conjunto de características y una etiqueta de clase o valor objetivo.
- Definición del parámetro K o número entero positivo que

se utiliza para determinar el número de vecinos más cercanos a considerar para realizar una predicción. El valor de K puede afectar la precisión del algoritmo.

- Cálculo de la distancia euclidiana entre el ejemplo de prueba y todos los ejemplos de entrenamiento.
- Selección de los K vecinos más cercanos en función de la distancia calculada.
- Votación (clasificación) o promedio (regresión). En el caso de clasificación, se realiza una votación mayoritaria entre los vecinos más cercanos para determinar la clase de predicción. En el caso de regresión, se realiza un

<sup>10</sup> Véase Aziz y Dowling (2019).

<sup>11</sup> Véase Patrick y Fischer (1970), Jiang et al. (2007), Peterson (2009), Sun y Huang (2010), Kramer y Kramer (2013) o Larose y Larose (2014).



promedio de los valores de los vecinos más cercanos para obtener la predicción.

- Predicción: Se asigna la clase de predicción al ejemplo de prueba o se devuelve

el valor de predicción en el caso de regresión.

El algoritmo de vecino más cercano tiene algunas ventajas y limitaciones que se describen en la Tabla A.2.

Tabla A.2. Ventajas e inconvenientes del algoritmo de vecino más cercano.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o Es relativamente simple de implementar.</li> </ul>	<ul style="list-style-type: none"> <li>o Es computacionalmente costoso en grandes conjuntos de datos (al no realizar un proceso de entrenamiento en sí, sino que almacena todo el conjunto de datos de entrenamiento y realiza comparaciones cada vez que se hace una predicción).</li> </ul>

Fuente: elaboración propia.

## A.4. CLASIFICADOR BAYES INGENUO (NBC)

El clasificador Bayes ingenuo -*Naïve Bayes Classifier* (NBC) por su terminología en inglés es un algoritmo de aprendizaje supervisado basado en el teorema de Bayes y en la suposición de independencia condicional entre las características<sup>12</sup>. A pesar de su simplicidad, es ampliamente utilizado y funciona bien en una variedad de problemas de

clasificación. Dada una instancia a clasificar, se calcula la probabilidad de que pertenezca a cada clase posible y se asigna a la clase con la probabilidad más alta. Aunque la suposición de la independencia de los atributos es generalmente una hipótesis pobre, y se viola a menudo para los conjuntos de datos verdaderos, a menudo proporciona una mejor precisión

<sup>12</sup> Véase Zhang y Su (2004), Dai et al. (2007), Leung (2007), Huang y Li (2011), Ting et al. (2011).



de clasificación en conjuntos de datos en tiempo real que cualquier otro clasificador. También requiere una pequeña cantidad de datos de entrenamiento. El

clasificador Naïve-Bayes aprende de los datos de entrenamiento y luego predice la clase de la instancia de prueba con la mayor probabilidad posterior.<sup>13</sup>

Tabla A.3. Ventajas e inconvenientes del Naïve Bayes Classifier.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o Eficiencia computacional en el cálculo de las probabilidades condicionales.</li> </ul>	<ul style="list-style-type: none"> <li>o La precisión del clasificador está afectada por la verosimilitud de la suposición de independencia condicional.</li> <li>o El requerimiento de gran cantidad de muestras implica importantes recursos de almacenamiento y computación.</li> </ul>

Fuente: elaboración propia.

El proceso básico del clasificador Bayes ingenuo se puede resumir en los siguientes pasos:

- Preparación de datos de entrenamiento: El conjunto de entrenamiento consiste en ejemplos etiquetados, donde cada ejemplo tiene un conjunto de características y una etiqueta de clase.
- Cálculo de las probabilidades a priori de que una instancia aleatoria pertenezca a cada clase sin tener en cuenta las características.
- Cálculo de las probabilidades condicionales para cada clase. Para ello, se asume que las características son independientes entre sí, lo que se conoce como la suposición de independencia condicional. Esta hipótesis es lo que le da el nombre de "naive" o ingenua a la técnica.
- Clasificación: Dada una instancia a clasificar, se calculan las probabilidades condicionales para cada clase utilizando las características de la instancia. Luego, se utiliza el teorema de Bayes para calcular la

<sup>13</sup> Véase Mosquera *et al.* (2018).



probabilidad posterior de cada clase. Finalmente, se asigna la clase con la probabilidad más alta como la clase de predicción.

El clasificador Bayes ingenuo puede manejar tanto características categóricas o

discretas como continuas. En el caso de características continuas, se asume una distribución de probabilidad para modelizar la probabilidad condicional. Algunas de sus ventajas e inconvenientes de este clasificador se muestran en la Tabla A.3.

## A.5. MÁQUINAS DE SOPORTE VECTORIAL

Las máquinas de soporte vectorial –*support vector machine* (SVM) por su terminología en inglés– son modelos de aprendizaje supervisado utilizados para la clasificación y regresión de datos. Las SVM ofrecen un rendimiento predictivo equilibrado de manera distintiva, incluso en estudios donde los tamaños de muestra pueden ser relativamente limitados<sup>14</sup>. Un algoritmo de entrenamiento SVM considera dos categorías y asigna los nuevos ejemplos a una u otra, lo que lo convierte en un clasificador binario no probabilístico.<sup>15</sup> El modelo SVM es útil en el desarrollo de aplicaciones multidominio en un entorno de *big data*<sup>16</sup>. Su

objetivo principal es encontrar un hiperplano que mejor separe los datos en diferentes clases o que mejor se ajuste a un conjunto de puntos. Este hiperplano se define por un subconjunto de los datos de entrenamiento llamados vectores de soporte (*support vectors*), que son los puntos más cercanos al hiperplano de separación. El SVM puede manejar tanto problemas de clasificación lineal como no lineal. En el caso de problemas linealmente separables, el SVM busca un hiperplano lineal que separe las clases, mientras que, en el caso de problemas no lineales, el SVM utiliza técnicas como el *kernel trick*, que permiten transformar el espacio de

<sup>14</sup> Pisner y Schnyer (2020).

<sup>15</sup> Véase Hearst *et al.* (1998), Meyer y Wien (2001), Chen *et al.* (2005), Noble (2006), Mammone *et al.* (2009).

<sup>16</sup> Véase Suthaharan y Suthaharan (2016).



características original en uno de mayor dimensión donde los datos sean linealmente separables.

Las funciones de *kernel*, son utilizadas para transformar datos de baja dimensionalidad a un espacio de mayor dimensionalidad, donde

los patrones y estructuras subyacentes pueden ser más fácilmente separables. Estas funciones son esenciales para resolver problemas de clasificación no lineales, donde los límites de decisión no pueden ser delineados linealmente en el espacio de características original.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>○ Es eficaz en espacios de alta dimensión y puede manejar conjuntos de datos con un número significativo de características.</li> <li>○ Busca el hiperplano que maximiza la distancia entre las clases, lo que puede ayudar a reducir el riesgo de sobreajuste.</li> <li>○ Los vectores de soporte (o puntos críticos para definir el hiperplano de separación) permiten la construcción del modelo sin dependencia de la totalidad del conjunto de datos.</li> </ul>	<ul style="list-style-type: none"> <li>○ Suele ser computacionalmente costoso (requiere una cantidad considerable de tiempo y recursos para conjuntos de datos muy grandes).</li> <li>○ La selección de un buen kernel y la optimización de sus parámetros puede ser un desafío (ya que un mal ajuste de kernel afecta el rendimiento del modelo).</li> </ul>

Tabla A.4. Ventajas e inconvenientes de las máquinas de soporte vectorial.

Fuente: elaboración propia.

En términos generales, una función de *kernel* toma un par de puntos de datos en el espacio original y calcula su similitud o producto interno en el espacio de características transformado. Este cálculo de similitud se utiliza para determinar la proximidad entre los puntos y, en última instancia, para

construir un hiperplano óptimo de separación entre las clases en el espacio de características transformado. Las funciones de *kernel* más comunes incluyen:

- *Kernel* lineal: Se utiliza para problemas linealmente separables y simplemente



calcula el producto interno entre dos puntos.

- *Kernel* polinómico: Realiza una transformación polinómica de los datos para aumentar la dimensionalidad.
- *Kernel* de base radial (RBF): Es uno de los *kernels* más utilizados y aplica una transformación no lineal basada en la distancia euclidiana entre los puntos.
- *Kernel* sigmoide: Realiza una transformación no lineal utilizando funciones sigmoides. Las funciones sigmoides son un tipo de función matemática que se caracterizan por tener una forma de "S". La función sigmoide más comúnmente utilizada es la función de activación sigmoide, también conocida como función logística<sup>17</sup>.

En la Tabla A.4 se muestran las ventajas e inconvenientes de este tipo de modelo.

Cada tipo de *kernel* tiene propiedades matemáticas y características específicas que las hacen más adecuadas para

diferentes tipos de problemas y conjuntos de datos. La elección adecuada del *kernel* es crucial para obtener un rendimiento óptimo en los modelos de SVM. Pero las funciones sigmoides pueden tener problemas de desvanecimiento o saturación de gradientes en modelos de redes neuronales más profundos, por lo que han surgido otras funciones de activación, como la ReLU (*Rectified Linear Unit*). Se trata de una función de activación que establece el valor de salida en cero para todos los valores de entrada negativos y mantiene el valor de entrada para los valores de entrada positivos. Esto significa que la función es lineal y no saturada en el rango de valores positivos, lo que facilita el cálculo del gradiente y evita el problema de los gradientes que desvanecen. Tiene como ventajas su eficiencia computacional, ya que solo implica una comparación y no requiere cálculos exponenciales o logarítmicos y, además, puede generar salidas dispersas, lo que permite modelos más eficientes y fáciles de interpretar. Sin embargo, la ReLU tiene el problema de las "neuronas muertas", que consiste en que, si una neurona

<sup>17</sup> Las funciones sigmoides son monótonas, diferenciables, y no lineales. Se utilizan también como funciones de activación en las capas ocultas de las redes neuronales, donde ayudan a introducir no linealidad y a regular la salida de cada neurona.

*ReLU* se activa en cero, es posible que nunca se recupere durante el entrenamiento, lo que significa que no aprenderá y no contribuirá al modelo. Esto puede abordarse mediante algunas modificaciones. Por ejemplo, la *Leaky ReLU* permite que cuando el input es positivo se comporta como *ReLU*, manteniendo el valor de entrada, pero cuando es negativo, multiplica el valor de entrada por un valor, lo que evita que la neurona se vuelva completamente inactiva.

Por otra parte, la *Parametric ReLU* (PReLU), introduce parámetros que pueden aprenderse en lugar de un valor constante para controlar la inclinación de la función en la región negativa. A diferencia de *Leaky ReLU*, donde se fija un valor pequeño para los valores negativos, PReLU permite que los parámetros se ajusten durante el entrenamiento, lo que le otorga una mayor flexibilidad y capacidad de adaptación. La función de activación PReLU

tiene como ventaja principal su mayor flexibilidad. A diferencia de *Leaky ReLU*, donde la inclinación se fija de antemano, PReLU permite que la inclinación se adapte durante el entrenamiento. Esto puede ayudar al modelo a aprender la mejor inclinación para mejorar el rendimiento. Permite la superación del problema de “neuronas muertas” al proporcionar un gradiente no nulo para los valores negativos. Tiene mayor capacidad de representación al introducir parámetros que se pueden aprender, lo que puede ser especialmente útil en casos donde la relación entre las variables de entrada y la salida es altamente no lineal. Pero también presenta algunos inconvenientes. Requiere más parámetros que las funciones *ReLU* o *Leaky ReLU*, lo que puede aumentar la complejidad del modelo y requerir más datos de entrenamiento. En el ámbito de las finanzas la técnica de *Support Vector Machine* está resultando de una notable utilidad.<sup>18</sup>

<sup>18</sup> Véase Trafalis y Ince (2000), Fan y Palaniswami (2000), Van Gestel et al. (2001), Cao y Tay (2001 y 2003), Tay y Cao (2001a, 2001b y 2002), Yang et al. (2002), Pérez-Cruz et al. (2003), Huang et al. (2004), Čížek et al. (2005), Shin et al. (2005), Huang (2005), Kumar y Thenmozhi (2006), Ince y Trafalis (2007), Boyacioglu et al. (2009), Bellotti y Crook (2009), Lee y To (2010), Chen et al. (2010), Chaudhuri y De (2011), Trustorff et al. (2011), Sun y Li (2012), Das y Padhy (2012), Erdogan (2013), Okasha (2014), Gui, et al. (2014), Li et al. (2014), Danenas y Garsva (2015), Madge y Bhatt (2015), Kumar et al. (2016), Gong et al. (2016), Kewat et al. (2017), Altan y Karasu (2019), Horak et al. (2020), Sivaram et al. (2020), Zahariev, et al. (2020), Zhang et al. (2021b), Sun et al. (2021) y Ali et al. (2021).

## A.6. REGRESIÓN LASSO, RIDGE Y ELASTIC NET

La regresión de LASSO (*least absolute shrinkage and selection operator* en su terminología en inglés) se encuadra dentro del ámbito del aprendizaje automático supervisado. Se trata de una técnica utilizada para la selección de características y la regularización de modelos de regresión.<sup>19</sup> Su objetivo principal es realizar una selección de características automática, es decir, identificar aquellas variables o características más relevantes o que más contribuyen a la predicción del valor objetivo. En el ámbito de las finanzas LASSO tiene algunas aplicaciones.<sup>20</sup>

Las técnicas de LASSO y Ridge son métodos utilizados en la regresión lineal penalizada, también conocida como regresión regularizada. Estos métodos son variantes de la regresión lineal tradicional y se utilizan para abordar problemas de multicolinealidad, para reducir la complejidad del modelo y para evitar problemas de sobreajuste.

La regresión penalizada introduce un término de penalización en la función objetivo utilizada para ajustar el modelo de regresión. Este término penaliza los coeficientes de regresión grandes, lo que permite controlar y reducir su magnitud. La regresión LASSO permite identificar y destacar las variables más importantes y, al mismo tiempo, controlar la complejidad del modelo. Esta técnica se basa en la minimización de una función que combina el error de ajuste (ajuste a los datos de entrenamiento) con una penalización de regularización L1 (norma L1) de los coeficientes del modelo. Esta penalización L1 impone una restricción sobre los coeficientes, favoreciendo la reducción de algunos de ellos a cero, lo que lleva a la selección de características. En particular, el método LASSO agrega un término de penalización basado en la norma L1 (suma de los valores absolutos de los coeficientes) a la función objetivo de la regresión. La introducción de este término

<sup>19</sup> Véase Tibshirani (1996), Wang *et al.* (2007), Owen (2007), Huang *et al.* (2008), Reid *et al.* (2016) y Ranstam y Cook (2018).

<sup>20</sup> Véase Véase Salehi *et al.* (2015), Pereira *et al.* (2016), Härdle, *et al.* (2017), Xu *et al.* (2019), Januaviani y Bon (2019), Shrivastava *et al.* (2020), Nie y Deng (2020), Sant'Anna *et al.* (2020), Wang y Liu (2020), Geraldo-Campos *et al.* (2022), Gao y Yang (2022) y Yang, *et al.* (2023).

lleva a que algunos coeficientes de regresión se vuelvan exactamente cero, lo que permite la eliminación automática de algunas características y la construcción de modelos más simples. La regresión LASSO tiene la capacidad de realizar selección de características y es útil cuando se sospecha que solo un subconjunto de las características es relevante.

A diferencia de la regresión LASSO, el *método Ridge* utiliza una norma L2 (suma de los cuadrados de los coeficientes) como término de penalización. Esto conduce a que los coeficientes de regresión se reduzcan, pero no lleguen a ser exactamente cero. La técnica *Ridge* es útil para reducir la influencia de características irrelevantes y mejorar la estabilidad del modelo. También es útil cuando se tienen múltiples características altamente correlacionadas. Este sistema se utiliza para resolver el problema de la multicolinealidad que suele plantearse en la regresión lineal múltiple<sup>21</sup> y se utiliza en diversos análisis financieros.<sup>22</sup>

Además de LASSO y *Ridge*, existen otras técnicas de regresión penalizada, como la *Elastic Net*, que combina los términos de penalización L1 y L2. La técnica *Elastic Net* es un enfoque de regularización utilizado en el aprendizaje automático y la estadística para abordar el problema de la selección de características y la reducción de la dimensionalidad en conjuntos de datos de alta dimensionalidad. El procedimiento de Red Elástica es una forma de optimización regularizada para la regresión lineal que sirve de puente entre las regresiones LASSO y *Ridge*.<sup>23</sup> La técnica *Elastic Net* combina estas dos formas de regularización mediante una combinación lineal de ambas penalizaciones. Introduce un parámetro de regularización adicional, llamado "alfa" ( $\alpha$ ), que controla la mezcla entre las penalizaciones L1 y L2. Cuando alfa es igual a cero, *Elastic Net* se reduce a la regresión de *Ridge*, y cuando alfa es igual a uno, se reduce a la regresión de Lasso. La ventaja de *Elastic Net* es que puede superar algunas

<sup>21</sup> Véase Hoerl y Kennard (1970), Osborne *et al.* (2000), Owen (2007), McDonald (2009) o Dorugade (2014).

<sup>22</sup> Véase Kalatzis *et al.* (2011), Annaert *et al.* (2013), Pereira *et al.* (2016), Fičura (2017), Geraldo-Campos *et al.* (2022), Soria Quijaite *et al.* (2022) o Chen (2023).

<sup>23</sup> Véase Zou y Hastie (2005), Hans (2011), Wu y Yang (2014), Zhang *et al.* (2017), Uchiyama *et al.* (2019), Shu *et al.* (2020) o He *et al.* (2022).

limitaciones de la regresión de *Ridge* y de la regresión de LASSO. La combinación de las penalizaciones L1 y L2 permite que *Elastic Net* seleccione un conjunto de características más pequeño que la regresión de *Ridge* y sea más estable que la regresión de LASSO cuando hay características altamente correlacionadas. Esto es especialmente útil cuando se trabaja con conjuntos de datos de alta dimensionalidad y multicolinealidad. La técnica *Elastic Net*, como herramienta de regularización que combina las penalizaciones L1 y L2 permite un equilibrio entre la estabilidad y la capacidad de selección automática de características. En el entorno de regresión de alta dimensión, la

Red Elástica produce un modelo parsimonioso al reducir todos los coeficientes hacia el origen. Sin embargo, en ciertos entornos, este comportamiento podría no ser deseable dado que, si algunas características están altamente correlacionadas entre sí y asociadas con la respuesta, podría ser deseable reducir menos los coeficientes correspondientes a ese subconjunto de características. A tal fin se ha propuesto una alternativa denominada *Red Elástica Cluster*, que reduce selectivamente los coeficientes de dichas variables entre sí, en lugar de hacia el origen.<sup>24</sup> Las regresiones del tipo *Elastic Net* han tenido diversas aplicaciones en el ámbito financiero.<sup>25</sup>

## A.7. REGRESION LARS

La regresión LARS (*Least Angle Regression*) es otro algoritmo utilizado en la regresión lineal que tiene como objetivo la selección de características y la regularización de los coeficientes. LARS ayuda a solucionar el problema de construir un modelo

de predicción lineal cuando la cantidad de predictores candidatos es grande y los datos posiblemente contienen anomalías que son difíciles de visualizar y limpiar.<sup>26</sup> La misión de la regresión LARS consiste en encontrar la dirección en la que

<sup>24</sup> Véase Witten *et al.* (2014).

<sup>25</sup> Véase Cui *et al.* (2021).

<sup>26</sup> Véase Khan *et al.* (2007).



los coeficientes se mueven en el espacio de características de manera más rápida. Proporciona una solución de regresión en función de una combinación lineal de las características seleccionadas.<sup>27</sup> A diferencia de LASSO y *Ridge*, LARS no utiliza una penalización específica. En cada paso, se agrega una variable que tiene la mayor correlación absoluta con el residuo parcial actual, y luego se actualizan los coeficientes en la dirección de esa variable. En cuanto a los resultados, si bien LASSO produce una solución única, donde algunos coeficientes se establecen en cero y solo un subconjunto de características se selecciona en el modelo

final, LARS proporciona una ruta de solución completa, lo que significa que puede obtener todas las soluciones intermedias entre el modelo nulo (sin características) y el modelo completo (con todas las características). Esto permite un análisis más detallado de cómo se agregan las características a medida que se avanza en el proceso. En definitiva, LARS se enfoca en encontrar la dirección de mayor correlación con el residuo parcial, mientras que LASSO se enfoca en lograr la selección automática de características y la reducción de la dimensionalidad a través de la penalización L1. Este tipo de regresión está siendo utilizada en el mundo de las finanzas.<sup>28</sup>

## A.8. ÁRBOLES DE DECISIÓN

Los árboles de decisión son estructuras utilizadas para realizar decisiones basadas en condiciones en los datos de entrada. Cada nodo del árbol representa una característica y cada rama representa una regla de decisión. La decisión se basa en una estructura jerárquica que

divide el espacio de características en regiones rectangulares. Cada nodo interno del árbol representa una prueba sobre una característica, mientras que las hojas representan las clases o las salidas de la tarea de aprendizaje. Los árboles de decisión son populares

<sup>27</sup> Véase Efron *et al.* (2004), Madigan y Ridgeway (2004) o Hu *et al.* (2018).

<sup>28</sup> Véase Wang y Tan (2009), Salehi *et al.* (2015) o Chopani *et al.* (2017).

debido a su interpretabilidad y facilidad para capturar relaciones

no lineales en los datos.

## A.9. BOSQUE ALEATORIO

Una modalidad de aprendizaje automático supervisado es el *bosque aleatorio (random forest)*, un algoritmo desarrollado por Breiman y Cutler en 2001. Es una técnica muy utilizada en la inteligencia artificial debido a su capacidad para manejar problemas de clasificación y regresión. El bosque aleatorio puede gestionar miles de variables explicativas. Esta técnica se puede utilizar para clasificar la importancia de las variables y para mostrar el efecto no lineal de las mismas. Su uso suele resultar robusto para valores atípicos. Mediante esta técnica se puede modelizar interacciones complejas entre variables.<sup>29</sup> Se ejecuta mediante la construcción de múltiples árboles de decisión mientras entrena y genera la clase que es el modo de salida de las clases por árboles individuales. Se basa en la construcción de múltiples árboles de decisión y la combinación de sus predicciones

para obtener resultados más precisos y estables.<sup>30</sup> Cada árbol se entrena en una muestra aleatoria de los datos y las predicciones se promedian o ponderan para obtener la predicción final. El uso de múltiples árboles en paralelo ayuda a reducir el sobreajuste y mejorar la generalización del modelo.

La construcción de un modelo de inteligencia artificial basado en *random forest* implica los siguientes pasos:

- Muestreo: Se toma una muestra aleatoria (*bootstrap*) con reemplazo del conjunto original de datos. Los ejemplos se seleccionan aleatoriamente y algunos pueden ser seleccionados más de una vez.
- Se construye un árbol de decisión para cada muestra *bootstrap*. Cada árbol se construye seleccionando un

<sup>29</sup> Véase Li (2013).

<sup>30</sup> Véase Grömping (2009).



subconjunto aleatorio de características (atributos) del conjunto original de datos en cada división del árbol. Esto ayuda a garantizar la diversidad de los árboles.

- Votación o promedio de predicciones. Para realizar una predicción, se aplica cada árbol a una nueva instancia de datos y se obtiene una

predicción individual. En el caso de la clasificación, se utiliza una votación mayoritaria para determinar la clase final. En la regresión, se promedian las predicciones para obtener el valor final.

Las principales ventajas e inconvenientes de los modelos de bosque aleatorio se muestran en la tabla A.5.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o Robustez frente a valores atípicos y frente a ruido en los datos por la combinación de múltiples árboles.</li> <li>o Potencia: puede manejar conjuntos de datos con un gran número de características al seleccionar aleatoriamente subconjuntos de características en cada división.</li> <li>o Capacidad informacional: proporciona una medida de la importancia de cada característica en el proceso de predicción.</li> <li>o Capacidad de generalización: captura patrones y realiza predicciones precisas fuera de la muestra.</li> </ul>	<ul style="list-style-type: none"> <li>o Suele ser computacionalmente costoso.</li> <li>o Puede requerir ajustes de hiper-parámetros para obtener el mejor rendimiento en un problema específico.</li> </ul>

Tabla A.5. Ventajas e inconvenientes de los modelos *random forest*.

Fuente: elaboración propia.

Los algoritmos de bosque aleatorio han sido utilizados

en el ámbito de las finanzas en los últimos años.<sup>31</sup>

<sup>31</sup> Véase Kumar y Thenmozhi (2006), Mori y Umezawa (2007) Liu *et al.* (2015), Zou *et al.* (2015), Khaidem *et al.* (2016), Thakur y Kumar (2018), Patel *et al.* (2019), Zhang *et al.* (2021a), Ghosh *et al.* (2022) o Wang (2022).

## A.10. GRADIENT BOOSTING

Otra modalidad basada en árboles de decisión es el *gradient boosting*. Esta técnica se origina a raíz de un conjunto de trabajos sobre la clasificación iterativa ponderada.<sup>32</sup> Esta

visión ha llevado al desarrollo de “algoritmos de potenciación” en muchas áreas del aprendizaje automático más allá de la regresión y la clasificación.<sup>33</sup> Se trata de un algoritmo de aprendizaje automático utilizado para construir modelos predictivos de alta precisión. En lugar de entrenar los árboles de forma independiente, el sistema *gradient boosting* lo hace en secuencia, de forma que cada árbol se ajusta para corregir los errores de los árboles anteriores. Esto permite mejorar el rendimiento del modelo a medida que se construyen iterativamente árboles adicionales. En este sentido, es una técnica de ensamblado que combina múltiples modelos de aprendizaje débiles para crear un modelo fuerte.<sup>34</sup> A diferencia de

otros métodos de ensamblado, como el *random forest*, el *gradient boosting* se enfoca en mejorar iterativamente el rendimiento del modelo mediante la optimización de funciones de pérdida. El objetivo de este tipo de algoritmos es “potenciar” la pequeña ventaja que una hipótesis producida por un “aprendiz débil” puede lograr sobre una conjetura aleatoria, utilizando el procedimiento de aprendizaje en una secuencia de distribuciones cuidadosamente construidas.<sup>35</sup>

El proceso habitual para la construcción de un modelo basado en *Gradient Boosting* se puede resumir en los siguientes pasos:

- Se construye un modelo inicial simple de árbol de decisión para realizar predicciones iniciales.
- Se calculan los residuos (diferencia entre las predicciones iniciales y los valores reales del conjunto de entrenamiento). Estos errores

<sup>32</sup> Véase Freund y Schapire (1997), Breiman (1996), Schapire y Singer (1998) o Mason et al. (1999).

<sup>33</sup> Véase Friedman et al. (2000), Friedman (2001), Sipper y Moore (2022), Blanchard et al. (2003), Lugosi y Vayatis (2004), Bickel et al. (2006) o Bühlmann y Hothorn (2007).

<sup>34</sup> Véase Natekin y Knoll (2013).

<sup>35</sup> Véase Zemel y Pitassi (2000).



representan la cantidad de información no capturada por el modelo inicial.

- Se reconstruye el modelo para capturar la información residual del modelo anterior.
- Se actualizan las predicciones combinando las predicciones del modelo inicial con las del nuevo modelo. En cada iteración, las predicciones se ajustan para acercarse cada vez más a los valores reales.
- Se repite el proceso para

mejorar gradualmente las predicciones. Cada nuevo modelo se ajusta a los residuos del modelo anterior.

- Finalización del proceso: El proceso se detiene cuando se alcanza un número predefinido de modelos o cuando se alcanza un criterio de parada, como un límite de rendimiento o un número máximo de iteraciones.

Las principales ventajas e inconvenientes de este tipo de herramientas se ofrecen en la tabla A.6.

Tabla A.6. Ventajas e inconvenientes de los modelos *gradient boosting*.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o Alto rendimiento. Puede generar modelos de alta precisión, utilizando múltiples modelos débiles y optimizando la función de pérdida adecuadamente.</li> <li>o Flexibilidad. Puede manejar tanto problemas de clasificación como de regresión.</li> <li>o Regularización. Permite aplicar técnicas para evitar el sobreajuste, como la limitación de la profundidad de los árboles o la introducción de parámetros de regularización.</li> <li>o Capacidad informacional. Puede proporcionar una medida de la importancia de las características utilizadas en el proceso de predicción.</li> </ul>	<ul style="list-style-type: none"> <li>o Sensibilidad a valores atípicos en los datos de entrenamiento, ya que intenta ajustarse a ellos, por lo que es relevante para tratar adecuadamente los valores atípicos antes de aplicar el algoritmo.</li> <li>o Mayor complejidad computacional. Puede ser computacionalmente costoso, especialmente cuando se utilizan conjuntos de datos grandes o se requiere un número significativo de iteraciones.</li> </ul>

Fuente: elaboración propia.

El uso de algoritmos de *Gradient Boosting* en finanzas es extenso.<sup>36</sup>

## A.11. AJUSTES AL SISTEMA DE *GRADIENT BOOSTING*

Existen algunas implementaciones mejoradas de *gradient boosting* que utilizan técnicas adicionales para mejorar su eficiencia y el rendimiento. Estas implementaciones incluyen características como la regularización, el muestreo de columnas y la poda de árboles para obtener modelos más precisos y eficientes. Por ejemplo, *Extreme gradient boosting* (XGBoost) y *LightGBM* son dos técnicas de aprendizaje automático basadas en árboles de decisión que se han vuelto muy populares debido a su rendimiento y eficiencia en problemas de clasificación y regresión.

XGBoost es una implementación optimizada y escalable de *gradient boosting* que destaca por su capacidad para manejar problemas

complejos y grandes volúmenes de datos, así como por su rendimiento y velocidad. Utiliza una combinación de algoritmos de *boosting*, regularización y técnicas avanzadas de paralelización con objeto de mejorar la precisión y eficiencia del modelo. Tiene la capacidad de manejar características perdidas y gestionar datos desequilibrados. La diferencia entre un algoritmo *gradient boosting* simple y XGBoost es que, a diferencia del primero, en este último el proceso de adición de los “aprendices débiles” no ocurre de forma sucesiva, sino que toma un enfoque de subprocesos múltiples mediante una adecuada utilización del CPU de la máquina. Esto conduce a una mayor velocidad y rendimiento.<sup>37</sup> Son diversas las aplicaciones de XGBoost en el ámbito de las finanzas.<sup>38</sup>

<sup>36</sup> Véase, por ejemplo, Salehi *et al.* (2015), Agapitos *et al.* (2017), Carmona *et al.* (2019), Climent *et al.* (2019), Liu *et al.* (2019), Lao *et al.* (2021), Liu *et al.* (2022) o Torky *et al.* (2023).

<sup>37</sup> Véase Ramraj *et al.* (2016), Mitchell y Frank (2017), Davis *et al.* (2020).

<sup>38</sup> Véase Chang *et al.* (2018), Huang y Yen (2019), Nobre y Neves (2019), Li *et al.* (2022), Nti y Somanathan (2022) o Qin (2022).



Por su parte, *light gradient boosting machine (LightGBM)* es otra técnica basada en árboles de decisión que también se enfoca en la eficiencia y el rendimiento, y está diseñado para manejar conjuntos de datos de gran tamaño. Utiliza una estrategia de crecimiento basada en hojas (*leaf-wise*) en lugar de la

estrategia tradicional basada en niveles (*level-wise*), lo que permite construir árboles más profundos y capturar patrones más complejos en los datos. En consecuencia, pueden lograrse modelos más precisos y una mayor capacidad para aprender a partir de datos desequilibrados y con ausencia de características.

## A.12 ÁRBOLES DE CLASIFICACIÓN Y REGRESIÓN (CART).

La técnica de Árboles de Clasificación y Regresión (CART por sus siglas en inglés) es un algoritmo de aprendizaje automático utilizado tanto para problemas de clasificación como de regresión. Utiliza un enfoque recursivo para dividir los datos en diferentes regiones de características, minimizando una función objetivo específica. CART es utilizado en algoritmos de árboles de decisión, *random forest* y *gradient boosting*. Los árboles CART construyen modelos de árbol de decisión que dividen el conjunto de datos en subconjuntos más pequeños basados en características o atributos, con el objetivo de predecir la variable objetivo. En el caso de la clasificación, los árboles CART dividen el conjunto

de datos en subconjuntos de manera recursiva, donde cada división se basa en un atributo y un umbral particular. Estas divisiones se realizan de tal manera que maximizan la pureza de las clases en cada subconjunto resultante. El proceso continúa hasta que se alcanza un criterio de parada, como un número máximo de niveles en el árbol o una pureza mínima. Una vez construido el árbol de clasificación, se puede utilizar para predecir la clase de nuevas instancias de datos, siguiendo el camino en el árbol que corresponde a las características de la instancia y tomando la decisión basada en las etiquetas de las hojas.

En el caso de las tareas de regresión, los árboles CART



también dividen el conjunto de datos en subconjuntos, pero en lugar de clasificar, predicen valores continuos. Las divisiones se realizan en función de la reducción de la varianza en cada subconjunto resultante. De manera similar a la clasificación, se sigue el camino en el árbol hasta llegar

a una hoja que proporciona la predicción de regresión.

En la tabla A.7 se muestran algunas ventajas e inconvenientes de los modelos basados en árboles de clasificación y regresión.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o Fácil interpretabilidad: representan reglas de decisión basadas en atributos.</li> <li>o Capacidad para gestionar datos con valores faltantes al tratarlos como una categoría adicional durante el proceso de división.</li> <li>o Robustez ante valores atípicos y ruido en los datos.</li> <li>o Flexibilidad y no linealidad: pueden capturar relaciones no lineales entre variables y no hacen suposiciones específicas sobre la forma funcional de los datos.</li> </ul>	<ul style="list-style-type: none"> <li>o Tendencia al sobreajuste especialmente cuando crecen demasiado y se ajustan demasiado a los datos de entrenamiento.</li> <li>o Inestabilidad (pequeños cambios en los datos de entrenamiento pueden llevar a cambios significativos en la estructura del árbol).</li> </ul>

Tabla A.7. Ventajas e inconvenientes de los modelos basados en CART.

Fuente: elaboración propia.

## A.13. REDES NEURONALES ARTIFICIALES

Las redes neuronales artificiales *-artificial neuronal networks (ANNs)* por su terminología en inglés- son un tipo de algoritmo utilizado en el campo del aprendizaje automático inspirado en el funcionamiento

del cerebro humano y su red de neuronas. Las ANNs consisten en un conjunto interconectado de unidades de procesamiento llamadas neuronas artificiales o nodos, que funcionan en



paralelo para procesar y transmitir información. Estas neuronas están organizadas en capas, generalmente divididas en una capa de entrada, una o varias capas ocultas y una capa de salida. Cada neurona artificial recibe entradas de datos ponderadas, las procesa mediante una función de activación y produce una salida que se transmite a otras neuronas en la red. Las conexiones entre las neuronas están asociadas con ponderaciones que se ajustan durante el proceso de entrenamiento de la red. El entrenamiento de una ANN implica la presentación de un conjunto de datos de entrada (datos de entrenamiento), junto con las salidas esperadas. Durante el entrenamiento, la red ajusta las ponderaciones de las conexiones internas para minimizar la diferencia entre las salidas producidas por la red y las salidas esperadas. Este proceso de ajuste se realiza utilizando algoritmos de optimización, como el *descenso del gradiente*, para encontrar los valores óptimos de los pesos. Las ANNs son conocidas por su capacidad para modelar relaciones complejas y realizar tareas de clasificación, reconocimiento de patrones, regresión y otras aplicaciones de aprendizaje automático. Dado su enfoque en la conexión y procesamiento

paralelo de información, las redes neuronales pueden detectar y aprender características no lineales en los datos, lo que las hace efectivas en muchos ámbitos. Especialmente, como resultado de su adaptabilidad, las ANNs representan potentes soluciones para el procesamiento subjetivo de la información, toma de decisiones y previsión. Las redes neuronales conectan entradas con salidas, pero con capas intermedias de neuronas ocultas. Cada unidad en cada capa está conectada a otras unidades y solo realiza un cálculo. Cada una de estas conexiones tiene ponderaciones que pueden ser fijas o pueden evolucionar, y algunas redes incluso permiten que las conexiones se eliminen y se generen nuevas.

Dependiendo de cómo se utilicen y se configuren las redes neuronales artificiales pueden utilizarse tanto en técnicas de aprendizaje supervisado como en técnicas de aprendizaje no supervisado. En el aprendizaje supervisado, las redes neuronales artificiales se utilizan para entrenar modelos que puedan hacer predicciones o clasificaciones basadas en ejemplos de entrenamiento etiquetados. En este caso, se proporcionan pares de entrada-salida (datos de entrenamiento etiquetados)

a la red neuronal, y el objetivo es ajustar las ponderaciones y las conexiones de la red para minimizar el error entre las salidas predichas y las salidas reales conocidas. En cambio, en el aprendizaje no supervisado, las redes neuronales artificiales se utilizan para encontrar patrones o estructuras ocultas en los datos sin la necesidad de etiquetas. Se utilizan para realizar tareas como agrupamiento (*clustering*), reducción de dimensionalidad y detección de anomalías. En este caso, la red neuronal se entrena para aprender representaciones internas de los datos sin una respuesta esperada predefinida. Además, existen también técnicas de aprendizaje semi-supervisado, que combinan características de ambos enfoques y utilizan redes neuronales artificiales para aprovechar tanto datos etiquetados como no etiquetados. En resumen, las redes neuronales

artificiales son un tipo de modelo que se puede aplicar en diferentes contextos de aprendizaje, ya sea supervisado, no supervisado o semi-supervisado, dependiendo del objetivo y de la configuración de entrenamiento.

Las ANNs están teniendo una amplia aplicación en el ámbito del análisis financiero.<sup>39</sup>

Dentro de esta técnica se utilizan dos algoritmos principales. El primero es la corrección de errores entrenada mediante aprendizaje no supervisado, como *retropropagación con descenso de gradiente*, con el objetivo de reducir gradualmente el error al mínimo respecto a una función de pérdida (o coste). El segundo es el mapa autoorganizado (SOM) y el mapa de características autoorganizadas (SOFM) que aplica el aprendizaje competitivo.

## A.14. ALGORITMO DE RETROPROPAGACIÓN

El algoritmo de retropropagación (*backpropagation*) se utiliza para entrenar redes neuronales

artificiales con múltiples capas, también conocidas como *redes neuronales profundas*. Se basa en la regla de la cadena de cálculo

<sup>39</sup> Véase Boyacioglu et al. (2009), Horak et al. (2020) o Ali et al. (2021).



diferencial, que permite calcular el gradiente del error con respecto a cada peso y sesgo de la red a partir del gradiente del error en la capa de salida. El uso del algoritmo *backpropagation* ha permitido avances significativos en el campo del aprendizaje profundo ya que, al permitir que la red neuronal ajuste sus ponderaciones y conexiones, se produce una mejora progresiva su capacidad de hacer predicciones precisas conforme aumenta el entrenamiento.

El proceso de entrenamiento de una red neuronal que utiliza retropropagación se puede resumir en los siguientes pasos:

- Inicialmente se asignan de forma aleatoria las ponderaciones y los sesgos de las conexiones entre las neuronas de la red neuronal.
- Se realiza una propagación hacia adelante (*forward propagation*) mediante la introducción de una muestra de entrenamiento en la red neuronal y se calcula el resultado para dicha muestra, propagando la información a través de las capas desde la entrada hasta la salida.
- Se compara el resultado de la red neuronal con el resultado esperado (etiqueta o valor objetivo) de la muestra de entrenamiento. Se calcula el valor del error que mide la discrepancia entre la salida obtenida y la esperada.
- El error se propaga hacia atrás (*back propagation*) a través de la red neuronal, desde la capa de salida hasta la capa de entrada, para calcular el *gradiente del error* con respecto a los pesos y sesgos de la red.
- Utilizando el *gradiente del error* calculado, se ajustan las ponderaciones y sesgos de las conexiones de la red mediante un algoritmo de optimización, como el descenso del gradiente, para minimizar el error de manera iterativa.
- Estos últimos procesos se repiten con diferentes muestras de entrenamiento hasta que el error de la red neuronal sea suficientemente bajo o hasta que se alcance un número máximo de iteraciones.



## A.15. MAPA AUTOORGANIZADO

El Mapa Autoorganizado, también conocido como SOM (*Self-Organizing Map* por su terminología en inglés), es una técnica de aprendizaje no supervisado propuesto por el investigador finlandés Teuvo Kohonen en la década de 1980. Es un tipo de red neuronal artificial que tiene como objetivo organizar y visualizar datos complejos de alta dimensionalidad en una estructura bidimensional o tridimensional. Su objetivo principal consiste en encontrar relaciones y patrones subyacentes en los datos sin la necesidad de etiquetas o valores objetivo.

El proceso básico de construcción de un modelo basado en un mapa autoorganizado (SOM) se puede resumir en los siguientes pasos:

- Se asignan las ponderaciones de las neuronas de manera aleatoria. Cada neurona representa un nodo en la estructura del mapa.
- Competencia: Se selecciona una muestra de entrada y se calcula la distancia entre esa muestra y las ponderaciones de todas las neuronas. La neurona con

el peso más cercano se activa y se considera la ganadora.

- Cooperación: Las neuronas vecinas a la ganadora también se actualizan para que se asemejen más a la muestra de entrada. Este proceso fomenta la cooperación y la organización topológica en el mapa.
- Se ajustan las ponderaciones de las neuronas ganadora y vecinas utilizando un proceso de aprendizaje iterativo basado en una función de vecindad y una tasa de aprendizaje. Las neuronas más cercanas a la ganadora se actualizan más que las neuronas lejanas.
- Iteración: Los pasos de competencia, cooperación y actualización de ponderaciones se repiten para varias muestras de entrada hasta que se alcance un estado de convergencia o hasta que se cumpla un número máximo de iteraciones.

El resultado final de un SOM es un mapa bidimensional o tridimensional donde las neuronas están organizadas de tal manera que las muestras de entrada similares están cercanas entre sí



en el mapa. Esto permite visualizar las relaciones y agrupamientos en los datos y proporciona una

herramienta para la exploración y análisis de patrones en los datos.

## A.16. MAPA DE CARACTERÍSTICAS AUTOORGANIZADAS (SOFM)

El mapa de características autoorganizadas –*Self-Organizing Feature Map* (SOFM) según sus siglas en inglés– es una variante del mapa autoorganizado que se utiliza para realizar una reducción de la dimensionalidad y visualización de datos. A diferencia del SOM, que organiza los datos en una estructura bidimensional o tridimensional, el SOFM busca organizar los datos en una red de neuronas unidimensional. El objetivo principal del SOFM es encontrar patrones y representaciones significativas en los datos de entrada. El proceso básico del mapa de características autoorganizadas es muy similar al de SOM ya que se asignan inicialmente ponderaciones de las

neuronas de manera aleatoria, se produce una competencia entre neuronas, activando la ganadora y las vecinas. Se produce un ajuste de las ponderaciones y se repite el proceso hasta que se alcanza un estado de convergencia o hasta que se cumpla un número máximo de iteraciones. El resultado final del mapa de características autoorganizadas es una representación en la cual las neuronas en el mapa unidimensional organizan las características de las muestras de entrada. Esto permite visualizar las relaciones y agrupamientos en los datos, así como identificar características significativas y representaciones latentes en los datos de alta dimensionalidad.

## A.17. REDES GENERATIVAS ADVERSARIALES

Las Redes Generativas Adversariales –*Generative*

*Adversarial Network* (GAN) por su terminología en inglés– son un



tipo de modelo de aprendizaje automático que consiste en dos redes neuronales que se enfrentan en un juego de competencia. La primera red es conocida como *generadora* y es responsable de crear muestras sintéticas, generando nuevos datos a partir de un espacio de ruido aleatorio. Su objetivo es producir datos que sean lo más similares posible a los datos reales que se utilizan para entrenar el modelo. La segunda red, conocida como *discriminadora*, intenta distinguir entre las muestras generadas y las reales. Su objetivo es aprender a distinguir correctamente entre ambos tipos de datos. Juntas, estas redes aprenden a generar datos realistas que son indistinguibles de los datos reales y se utilizan para generar imágenes, textos y otros tipos de datos sintéticos. El proceso de entrenamiento de una GAN implica que las redes generadora y discriminadora se entrenen de manera simultánea, pero en direcciones opuestas. El *generador* trata de producir datos que engañen al *discriminador* como si fueran reales, mientras que el *discriminador* intenta mejorar su capacidad para distinguir entre datos reales y generados. A medida que ambos se entrenan y compiten, se busca alcanzar un equilibrio

en el cual el *generador* produce datos cada vez más realistas y el *discriminador* se vuelve más hábil en la detección de datos falsos.

La clave del éxito de las GANs radica en su capacidad para aprender de forma no supervisada, lo que significa que no requieren etiquetas o datos de entrenamiento previamente clasificados, sino que se basan en la competencia entre ambas redes para aprender a capturar las características estadísticas y la estructura de los datos reales. Las GANs han sido fundamentales en el avance de la investigación en inteligencia artificial y el desarrollo de técnicas más avanzadas, como las Redes Generativas Adversariales Condicionales (cGAN, por sus siglas en inglés) y las GAN de Ciclo Consistente (CycleGAN por sus siglas en inglés), que permiten un control más específico y la transferencia de estilo entre diferentes dominios de datos. Las cGANs son una extensión de las GANs estándar que permiten controlar y guiar la generación de datos mediante el uso de información adicional, conocida como condición o etiqueta de clase. En una cGAN, tanto el *generador* como el *discriminador* reciben una entrada adicional, que puede ser cualquier tipo de

información auxiliar relevante para el problema en cuestión. Esta información condicional se proporciona como una etiqueta o un vector de características que especifica las características deseadas del dato generado. El *generador* utiliza tanto el ruido aleatorio de entrada como la información condicional para generar datos sintéticos. La información condicional permite al *generador* producir muestras más específicas y controladas. El *discriminador*, por su parte, también recibe la información condicional junto con los datos reales y generados. Su tarea es clasificar correctamente los datos reales y generados, teniendo en cuenta tanto las características visuales como la información condicional. El proceso de entrenamiento de una cGAN es similar al de una GAN estándar, aunque en una cGAN la información condicional también se utiliza para guiar el entrenamiento. Esto permite un control más preciso sobre la generación de datos y la capacidad de generar muestras específicas según las condiciones dadas. Al agregar información condicional a las GANs, las cGANs brindan mayor flexibilidad y control sobre la generación de datos, permitiendo la creación de modelos más especializados

y adaptados a las necesidades específicas de cada problema.

Por su parte, las *CycleGANs* son un tipo de arquitectura de Redes Generativas Adversariales utilizado para realizar la transferencia de estilo y la traducción entre diferentes dominios de datos sin necesidad de pares de datos de entrenamiento. A diferencia de las GAN tradicionales, donde se requiere una coincidencia uno a uno de datos en los conjuntos de entrenamiento, las *CycleGANs* permiten aprender relaciones de mapeo entre dos dominios diferentes sin necesidad de tener pares de datos correspondientes. Esto es particularmente útil cuando se desea realizar la traducción de un dominio a otro sin disponer de ejemplos emparejados directos. La idea central de las *CycleGANs* es utilizar dos *generadores* y dos *discriminadores* en un enfoque de ciclo cerrado, con el objetivo principal de lograr la consistencia del ciclo entre los dos dominios. Para lograrlo, se utiliza una pérdida de ciclo que penaliza la diferencia entre los datos originales y los datos reconstruidos, que asegura que la información se conserve a través del ciclo de traducción y ayuda a mantener la coherencia entre los dominios.

## A.18. LAS REDES NEURONALES CONVOLUCIONALES (CNN) Y RECURRENTES (RNN)

Las primeras redes neuronales simples de alimentación directa fueron propuestas en 1958. En este enfoque la información se mueve en una sola dirección: hacia adelante desde la capa de entrada del modelo a su capa de salida, sin viajar hacia atrás para ser reanalizada por el modelo. Eso significa que se pueden introducir datos en el modelo y luego entrenar éste para predecir algo sobre diferentes conjuntos de datos. Posteriormente se han desarrollado las denominadas Redes Neuronales Convolucionales –*Convolutional Neural Networks* (CNN) en terminología en inglés–. Se trata de un tipo de arquitectura de red neuronal artificial diseñada específicamente para procesar datos con estructura de cuadrícula, como imágenes, señales de audio y datos secuenciales. Las CNNs se han vuelto muy populares en el campo del procesamiento de imágenes y visión por computadora debido a su capacidad para extraer características relevantes y realizar clasificación, detección de objetos, segmentación y otras tareas relacionadas. Las CNNs

se inspiran en la organización y funcionamiento del sistema visual humano, donde las neuronas se especializan en la detección de características locales en regiones específicas del campo visual. Al igual que las redes neuronales tradicionales, las CNNs están compuestas por capas de neuronas, pero su arquitectura se adapta a la estructura de cuadrícula de los datos de entrada. Los principales componentes de las CNN son las siguientes:

- Capas convolucionales que aplican operaciones de convolución a los datos de entrada utilizando filtros o *kernels*. La convolución es un proceso que permite extraer características locales y aprender patrones visuales, como bordes, texturas y formas, en diferentes niveles de abstracción.
- Capas de *pooling* que reducen la dimensionalidad de las características obtenidas de las capas convolucionales, lo que ayuda a disminuir la cantidad de parámetros y la complejidad computacional.



El *pooling* se realiza mediante operaciones como el *max pooling*, donde se selecciona el valor máximo dentro de una ventana de píxeles.

- Capas totalmente conectadas que reciben las características extraídas de las capas anteriores y las utilizan para realizar la clasificación o realizar otras tareas específicas. Estas capas están formadas por neuronas que se conectan a todas las neuronas de la capa anterior.
- Se utilizan funciones de activación no lineales, como ReLU (*Rectified Linear Unit*), para introducir no linealidad en la red y permitir el aprendizaje de representaciones más complejas.
- Las CNNs se entrenan utilizando el algoritmo de retropropagación del error, que ajusta las ponderaciones y los sesgos de las neuronas para minimizar la diferencia entre las salidas predichas y las salidas deseadas.

Las CNNs tienen capacidad para aprender características relevantes de forma automática sin la necesidad de una extracción manual de características, y muestran invarianza a la traslación

y a pequeñas deformaciones, lo que las hace efectivas en tareas de clasificación y detección de objetos; y permiten manejar conjuntos de datos grandes y complejos debido a su estructura de capas y al uso de operaciones compartidas.

Por su parte, en las redes neuronales recurrentes -*Recurrent Neural Networks* (RNN) en terminología en inglés- las conexiones incluyen bucles, lo que significa que el modelo mueve los datos hacia adelante y los hace retroceder para volver a ejecutarse a través de las capas anteriores. Cada entrada individual se introduce en el modelo por sí misma, así como en combinación con la entrada anterior. Por ejemplo, las RNNs pueden ayudar a detectar transacciones financieras fraudulentas de manera más compleja que lo hacen las CNNs, ya que pueden “aprender” del comportamiento financiero de un individuo, como una secuencia de transacciones como el historial de una tarjeta de crédito, y medir cada transacción contra el registro de la persona en su conjunto.



## A.19. BÚSQUEDA DE ARQUITECTURA NEURONAL

La búsqueda de arquitectura neuronal –*Neural Architecture Search* (NAS) por su terminología en inglés– es un enfoque automatizado para encontrar la arquitectura óptima de una red neuronal. En lugar de diseñar manualmente la estructura de la red, NAS utiliza algoritmos de búsqueda para explorar un espacio de posibles arquitecturas y encontrar automáticamente la mejor arquitectura para una tarea específica. La idea detrás

de una NAS es permitir que los ordenadores encuentren de manera eficiente estructuras de redes neuronales más efectivas, superando las limitaciones de los enfoques de diseño manual tradicionales. La búsqueda de la arquitectura se realiza mediante técnicas de optimización y aprendizaje automático, como algoritmos genéticos, búsqueda en rejilla, búsqueda aleatoria o métodos basados en aprendizaje por refuerzo.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o Efectividad en el diseño de redes neuronales altamente eficientes y precisas en una amplia gama de aplicaciones.</li> <li>o Esta técnica ha permitido avances significativos en la capacidad de las redes neuronales para adaptarse a tareas específicas y superar el rendimiento de las arquitecturas diseñadas manualmente.</li> </ul>	<ul style="list-style-type: none"> <li>o Elevado coste computacional debido a la gran cantidad de arquitecturas a evaluar (aunque se han desarrollado métodos para acelerar la búsqueda, como el uso de técnicas de aprendizaje automático para modelizar el rendimiento de las arquitecturas y la transferencia de conocimiento de arquitecturas previamente encontradas).</li> </ul>

Tabla A.8. Ventajas e inconvenientes de *Neural Architecture Search*.

Fuente: elaboración propia.

El proceso de búsqueda de arquitectura implica evaluar y comparar múltiples arquitecturas candidatas en un conjunto

de datos de validación o mediante técnicas de validación cruzada. Se utilizan métricas de rendimiento, como precisión,



tasa de error o pérdida, para medir el desempeño de cada arquitectura y guiar la búsqueda hacia las arquitecturas más

prometedoras. Las principales ventajas e inconvenientes de las NAS se muestran en la tabla A.8.

## A.20. DEEP LEARNING

Las redes neuronales profundas, también conocidas como *deep learning*, son un tipo de ANNs con múltiples capas ocultas. Estas redes profundas han demostrado tener un gran éxito en áreas como el procesamiento de imágenes, reconocimiento de voz, traducción automática y muchas otras aplicaciones de inteligencia artificial. La intuición detrás del aprendizaje profundo es modelizar con mayor precisión relaciones complejas entre variables y, en última instancia, imitar mejor la toma de decisiones humana. El aprendizaje profundo utiliza una red neuronal de varias capas para imitar el aprendizaje humano mediante el concepto de *descenso gradiente* y se utiliza a menudo en lenguaje natural. Con este método, un algoritmo puede ir aprendiendo las características de una imagen paso a paso, o capa a capa. Puede realizar una propagación hacia adelante y luego retropropagar para encontrar el mejor ajuste

de una imagen basándose en las características aprendidas. Cuando los algoritmos informáticos que aprenden con este método se basan en datos de entrada masivos los resultados pueden ser sorprendentes. Además, los resultados obtenidos con estos algoritmos a menudo se convierten de nuevo en datos de entrada para el sistema, en un bucle de retroalimentación como es el caso del aprendizaje por refuerzo.

Una característica clave del aprendizaje profundo es la adición de "capas ocultas" que permite determinar influencias múltiples y combinadas entre las variables de entrada mediante una modelización. A medida que los datos de entrada avanzan a través de las capas ocultas, las variables se combinan y recombinan en nuevos factores re-ponderados según la influencia de la capa anterior. Esta adición de capas ocultas entre la entrada



y la salida es, sin embargo, la fuente del problema conocido como *caja negra*, ya que no

siempre está claro cómo se han re combinado las entradas para crear un determinado resultado.

## A.21. BÚSQUEDA EN REJILLA

La búsqueda en rejilla (*grid search* en su terminología en inglés) es una técnica de optimización utilizada en el aprendizaje automático para encontrar la combinación óptima de hiperparámetros de un modelo. Los hiperparámetros son parámetros configurables externamente al modelo que afectan a su rendimiento y al comportamiento de variable como la tasa de aprendizaje, el número de capas ocultas en una red neuronal o la profundidad de un árbol de decisiones. La búsqueda en rejilla se utiliza para encontrar los valores de estos hiperparámetros que maximicen el rendimiento del modelo en función de una métrica de evaluación, como la precisión o minimizan su medida inversa, el error.

La búsqueda en rejilla es una técnica exhaustiva que explora sistemáticamente todas las combinaciones posibles de valores de hiperparámetros dentro de los rangos especificados. Aunque

puede ser computacionalmente costosa, es útil para encontrar una buena configuración de hiperparámetros en problemas donde el espacio de búsqueda no es demasiado grande. Sin embargo, a medida que aumenta el número de hiperparámetros y sus rangos posibles, la búsqueda en rejilla puede volverse impracticable debido a la explosión combinatoria. En tales casos, se utilizan técnicas más avanzadas, como la búsqueda aleatoria, la búsqueda en cuadrícula aleatorizada o la optimización bayesiana, para reducir la complejidad de la búsqueda.

El proceso básico de construcción de un modelo basado en *Grid Search* se puede resumir en los siguientes pasos:

- Se define una lista de hiperparámetros que hay que ajustar y se especifican los rangos de valores posibles para cada uno.



- Se genera un conjunto de combinaciones posibles de valores de hiperparámetros a través de una rejilla o malla, donde cada combinación representa un punto en la rejilla.
- Para cada combinación de hiperparámetros, se entrena y evalúa el modelo utilizando validación cruzada u otra técnica de evaluación. Se registra la métrica de evaluación, como la precisión o el error, para cada combinación.
- Se selecciona la combinación de hiperparámetros que proporciona el mejor rendimiento en función de la métrica de evaluación.
- Una vez seleccionada la mejor combinación de hiperparámetros, se entrena el modelo final utilizando todos los datos de entrenamiento disponibles.

## A.22. APRENDIZAJE NO SUPERVISADO

Es una rama del *machine learning* en la que se busca descubrir patrones y estructuras ocultas en los datos, sin la necesidad de etiquetas o información de salida previamente definida. Además, puede aprender por sí mismo. Los algoritmos de aprendizaje no supervisado exploran los datos para encontrar relaciones y agruparlos en categorías o realizar otras tareas de descubrimiento. Entre las técnicas de aprendizaje no supervisado, están las de agrupamiento que tienen similitud con el sistema de *support vector machine* (SVM) ya que involucran características

de mapeo en un plano, pero se diferencian de aquel en que no tratan de predecir resultados sino de crear grupos similares para identificar patrones y estructuras en conjuntos de datos. Hay tres modalidades de aprendizaje no supervisado esencialmente:

- *K-means*: Es uno de los algoritmos de agrupación más populares. Trata de dividir los datos en un número predefinido de clústeres (K) de manera que las instancias dentro de cada clúster sean similares entre sí y distintas de las instancias en otros clústeres. Comienza seleccionándose K centroides



iniciales y luego se asigna cada instancia al clúster cuyo centroide esté más cercano. Posteriormente se actualizan los centroides calculando la media de las instancias asignadas a cada clúster. Este proceso se repite iterativamente hasta que la asignación de instancias y los centroides converjan.

- *K-medoids*: Es similar al algoritmo *K-means*, pero en lugar de utilizar la media de las instancias para representar el centroide de cada clúster, utiliza una instancia real del conjunto de datos, llamada *medoide*. El *medoide* es la instancia más representativa del clúster en términos de distancia a las demás instancias del mismo clúster. A diferencia de *K-means*, *K-medoids* es más robusto ante valores atípicos o datos no numéricos, pero también puede ser más costoso computacionalmente.
- Agrupación jerárquica: Es un

enfoque de agrupación que crea una estructura jerárquica de clústeres. Comienza considerando cada instancia como un clúster individual y luego fusiona iterativamente los clústeres más cercanos hasta obtener un solo clúster que englobe todos los datos. Existen dos enfoques principales para la agrupación jerárquica: *aglomerativo (bottom-up)* y *divisivo (top-down)*. En el enfoque *aglomerativo*, se inicia con clústeres individuales y se fusionan gradualmente en clústeres más grandes. En el enfoque *divisivo*, se parte de un único clúster grande y se divide en clústeres más pequeños.

Tanto *K-means* como *K-medoids* son técnicas de agrupación basadas en partición, donde se asigna cada instancia a un único clúster. En contraste, la agrupación jerárquica permite tener una representación más detallada de las relaciones entre clústeres a través de la estructura jerárquica.

## A.23. ANÁLISIS DE COMPONENTES PRINCIPALES (ACP).

El análisis de componentes principales es una técnica de aprendizaje no supervisado

utilizada para reducir la dimensionalidad de un conjunto de datos. A través del



ACP, se busca encontrar una representación más compacta del conjunto de datos mediante la proyección de las variables originales en un nuevo espacio de menor dimensión, llamados componentes principales. Los componentes principales se eligen de manera que capturen la mayor cantidad posible de variabilidad

presente en los datos originales. ACP reduce la dimensionalidad de un conjunto de datos, pero preservando al mismo tiempo la mayor cantidad posible de información contenida en los datos originales y busca transformar un conjunto de variables correlacionadas en un conjunto de variables no correlacionadas.

Tabla A.9. Ventajas e inconvenientes del Análisis de Componentes Principales (ACP).

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o La reducción de la dimensionalidad permite disminuir la cantidad de variables en un conjunto de datos, lo que facilita la visualización y el análisis.</li> <li>o La transformación de las variables originales en componentes principales no correlacionados ayuda a eliminar la multicolinealidad (correlación entre las variables explicativas).</li> <li>o Puede revelar patrones y estructuras ocultas en los datos al destacar las variables que contribuyen más a la variabilidad total.</li> <li>o La compresión de datos que supone la reducción de la dimensionalidad facilita el almacenamiento y la transmisión eficiente de datos.</li> </ul>	<ul style="list-style-type: none"> <li>o Se basa en la covarianza o correlación de las variables y asume una estructura lineal en los datos.</li> <li>o Puede no ser adecuado para todos los conjuntos de datos, especialmente cuando éstos no siguen una distribución normal o cuando hay relaciones no lineales importantes.</li> </ul>

Fuente: elaboración propia.

El proceso de construcción de un modelo de Análisis de Componentes Principales implica los siguientes pasos:

- Se realiza una estandarización de los datos para asegurarse de que todas las variables tengan la misma escala y que ninguna

variable tenga una influencia desproporcionada en el análisis debido a su rango de valores.

- Se calcula la matriz de covarianza o correlación entre las variables del conjunto de datos.
- A partir de la matriz de covarianza o correlación, se calculan los componentes principales. Estos componentes principales son combinaciones lineales de las variables originales que capturan la mayor varianza posible en los datos.
- Se seleccionan los componentes principales más importantes en función de su contribución a la varianza total. Se puede establecer un umbral

para determinar cuántos componentes principales se deben retener.

- Se proyectan los datos originales en el nuevo espacio de componentes principales. Esto implica multiplicar los datos por los vectores de carga de los componentes principales.

Las principales ventajas e inconvenientes de este tipo de modelos se muestran en la Tabla A.9.

El análisis de componentes principales y las regresiones por mínimos cuadrados parciales son bastante similares, ya que ambos pretenden reducir el número de variables combinándolas y extrayendo factores comunes.

## A.24. EL ANÁLISIS DE COMPONENTES INDEPENDIENTES (ICA)

El Análisis de Componentes Independientes –*Independent Component Analysis* (ICA) por su terminología en inglés– es una técnica de aprendizaje automático no supervisado que se utiliza para descomponer una mezcla de señales en sus componentes originales,

asumiendo que éstos son estadísticamente independientes entre sí. A diferencia del Análisis de Componentes Principales, que se centra en encontrar las componentes principales con mayor varianza, el ICA busca identificar las fuentes o señales subyacentes que dieron lugar



a la mezcla observada. Por lo tanto, es especialmente útil en escenarios donde se desea separar y recuperar señales originales que se mezclaron, como en el procesamiento de señales o imágenes. El ICA asume ciertas condiciones, como la no *gaussianidad* de las fuentes y la linealidad de la mezcla, y puede requerir suposiciones adicionales según el contexto específico de aplicación.

Los pasos del proceso básico para construir un modelo basado en el Análisis de Componentes Independientes son los siguientes:

- Preparación de los datos. Se obtiene un conjunto de señales mezcladas u observaciones. Estas señales pueden ser datos de sensores, señales de audio, imágenes o cualquier otro tipo de datos donde se sospeche que existen fuentes subyacentes independientes.
- Estimación de las componentes independientes. Utilizando técnicas de optimización y estadística, se busca estimar las componentes independientes a partir de las señales mezcladas. El objetivo es encontrar una transformación lineal que permita recuperar las fuentes originales.
- Separación de las señales. Aplicando la transformación estimada a las señales mezcladas, se obtienen las componentes independientes separadas, que representan las fuentes originales deseadas.
- Interpretación y análisis de los componentes. Los componentes independientes pueden ser analizadas e interpretadas para comprender mejor las fuentes subyacentes y las estructuras presentes en los datos originales.

## A.25. EL ANÁLISIS DE FACTORES (AF)

El Análisis de Factores (AF) una técnica estadística utilizada para explorar la estructura subyacente de un conjunto de variables observadas y reducir la dimensionalidad de los datos.

Aunque el AF no se considera una técnica específica de inteligencia artificial, puede ser utilizada como parte de un enfoque más amplio de aprendizaje automático. Este tipo de modelo



busca identificar las relaciones subyacentes entre las variables observadas y agruparlas en factores latentes. Estos factores latentes representan constructos no observables o conceptos subyacentes que influyen en las variables observadas. La idea es que las variables observadas sean influidas por un número menor de factores, lo que permite una representación más compacta y comprensible de los datos.

En el contexto de la inteligencia artificial, el AF puede utilizarse como una técnica previa al entrenamiento de modelos de aprendizaje automático para reducir la dimensionalidad de los datos y extraer características significativas. Esto puede ayudar a mejorar la eficiencia y la precisión de los modelos al trabajar con conjuntos de datos de alta dimensionalidad.

El proceso básico del Análisis de Factores se puede resumir en los siguientes pasos:

- Se prepara un conjunto de

variables observadas que se espera estén relacionadas o sean influidas por factores latentes comunes.

- Utilizando técnicas de análisis matemático, como el método de máxima verosimilitud o el método de los componentes principales, se estima el número de factores y se calculan los valores de carga de los factores para cada variable. Las cargas de los factores indican la fuerza y dirección de la relación entre las variables observadas y los factores latentes.
- Los factores obtenidos pueden ser interpretados y etiquetados según su contenido y relación con las variables observadas.
- Se analizan las cargas de los factores, los valores propios y otros indicadores para evaluar la calidad y la validez del modelo de factores. Esto incluye examinar la varianza explicada por los factores y determinar si el modelo de factores se ajusta bien a los datos.

## A.26. APRENDIZAJE POR REFUERZO

El aprendizaje por refuerzo es una técnica en la que un agente

aprende a través de la experiencia al tomar acciones en un entorno



y a recibir retroalimentación en forma de recompensas o penalizaciones. Su objetivo es aprender una política óptima o estrategia que indica qué acción tomar en cada estado para maximizar la recompensa acumulada a largo plazo. Se trata de un proceso de optimización para encontrar la mejor solución a una situación y tiende a orientar el algoritmo hacia los criterios definidos por el agente que diseña del modelo<sup>40</sup>. Permite que la estructura aprenda con tiempo real. En el aprendizaje por refuerzo un agente interactúa con un entorno y aprende a través de ensayo y error. El agente toma decisiones secuenciales y recibe retroalimentación positiva o negativa en función de las acciones realizadas. Los métodos de aprendizaje por refuerzo se ocupan de cómo los agentes virtuales eligen sus acciones para maximizar una función de recompensa definida por un

ser humano. Estos métodos no requieren pares de entrada/salida etiquetados y no es necesario corregir explícitamente las acciones subóptimas. El objetivo es encontrar un equilibrio entre la exploración del territorio desconocido y la explotación del conocimiento actual.

El Aprendizaje por Refuerzo Avanzado (*Advanced Reinforcement Learning* por su terminología en inglés) es una rama del aprendizaje automático que se enfoca en la toma de decisiones secuenciales y la interacción de un agente con un entorno para maximizar una recompensa acumulada. Se basa en el concepto de aprendizaje por refuerzo clásico, pero con el uso de técnicas más avanzadas como: los algoritmos de aprendizaje profundo *Q-networks* y los métodos de aprendizaje por refuerzo basados en políticas.

## A.27. ALGORITMOS DE APRENDIZAJE PROFUNDO Q-NETWORKS

Los Algoritmos de Aprendizaje Profundo –*Deep Q-Networks*

(DQN) por su terminología en inglés– son una clase de

<sup>40</sup> Chow et al. (2019).



mecanismo de aprendizaje por refuerzo profundo utilizados para entrenar modelos de redes neuronales profundas en la toma de decisiones, que permite manejar espacios de estados y acciones de alta dimensionalidad. Su objetivo principal es aprender una función de valor  $Q$ , que estima el valor esperado de tomar una acción en un estado dado. La función de valor  $Q$  representa la calidad esperada de la acción y se utiliza para guiar al agente de aprendizaje por refuerzo a tomar decisiones óptimas en diferentes entornos. Los DQN se basan en la combinación de dos conceptos clave: las redes neuronales profundas y el algoritmo  $Q$ -Learning. En lugar de usar tablas para almacenar los valores  $Q$  como en el  $Q$ -Learning tradicional, los DQNs utilizan redes neuronales profundas para aproximar la función de valor  $Q$ .

Durante el entrenamiento, la red neuronal se actualiza utilizando el algoritmo de propagación hacia atrás (*backpropagation*) para minimizar la diferencia entre los valores  $Q$  predichos y los valores  $Q$  objetivo. El valor  $Q$  objetivo se calcula utilizando la ecuación de Bellman, que combina la recompensa inmediata y el valor  $Q$  máximo estimado del estado siguiente. Este proceso

se repite iterativamente para mejorar la aproximación de la función  $Q$  a medida que el agente interactúa con el entorno.

El proceso de entrenamiento de *Deep Q-Networks* implica iterar a través de diferentes etapas:

- El agente interactúa con el entorno, tomando acciones y recibiendo recompensas. Durante estas interacciones, se recopila una experiencia que consiste en pares de estados, acciones, recompensas y estados siguientes.
- La experiencia recopilada se almacena en una memoria de repetición (*replay memory*), que es un *buffer* que conserva una muestra de las experiencias pasadas del agente.
- El DQN selecciona aleatoriamente una muestra de experiencias de la memoria de repetición y utiliza estos datos para entrenar la red neuronal profunda. La red neuronal toma un estado como entrada y genera una estimación del valor  $Q$  para cada acción posible en ese estado.
- Para estabilizar el entrenamiento y mejorar la convergencia, los DQNs utilizan una segunda red



neuronal, conocida como red objetivo (*target network*). Esta red se utiliza para calcular los objetivos de entrenamiento y se actualiza periódicamente con los parámetros de la red neuronal principal.

- A medida que el entrenamiento progresa, el DQN utiliza una política de exploración-

explotación para seleccionar las acciones a tomar. Al principio, el agente tiende a explorar el entorno de manera más aleatoria para descubrir nuevas estrategias. Con el tiempo, la política se desplaza hacia la explotación de las acciones que se han demostrado más efectivas.

## A.28. OPTIMIZACIÓN DE LA POLÍTICA PRÓXIMA (PPO)

Los métodos de aprendizaje por refuerzo basados en políticas (*policy-based*) pertenecen al enfoque del aprendizaje por refuerzo que se centran en aprender directamente una política, es decir, una estrategia que indica qué acciones tomar en cada estado para maximizar la recompensa acumulada a lo largo del tiempo. A diferencia de otros enfoques de aprendizaje por refuerzo, como los métodos basados en valores (*value-based methods*), que buscan estimar y aprender funciones de valor para determinar la calidad de cada estado-acción, los métodos basados en políticas se centran en aprender directamente una política. Existen diferentes técnicas y algoritmos utilizados en los métodos de aprendizaje por

refuerzo basados en políticas. Uno de ellos es el método *gradient policy*, que utiliza el cálculo de gradientes de la función objetivo con respecto a los parámetros de la política. La idea es actualizar los parámetros de la política en la dirección del gradiente para maximizar la recompensa esperada. En este ámbito destaca la técnica conocida como Optimización de la Política Próxima –*Proximal Policy Optimization* (PPO) por su terminología en inglés–. Se trata de un algoritmo de aprendizaje por refuerzo que se utiliza para entrenar agentes de inteligencia artificial en entornos de toma de decisiones basado en la optimización de políticas para mejorar gradualmente el rendimiento del agente a través de iteraciones de entrenamiento.



En un modelo PPO, se define una política que especifica la distribución de probabilidades sobre las acciones que el agente puede tomar en un estado dado. El objetivo es aprender una política óptima que maximice la recompensa acumulada a lo largo del tiempo. El enfoque central de PPO es la optimización de políticas mediante la actualización de los parámetros de la política a través de iteraciones de entrenamiento. A diferencia de otros algoritmos de aprendizaje por refuerzo que utilizan métodos basados en gradiente, PPO utiliza una estrategia de optimización basada en restricciones y una medida de similitud llamada *clip*. Esto ayuda a garantizar que las actualizaciones de la política no sean excesivas y evita cambios drásticos que puedan perjudicar el rendimiento.

En un modelo del tipo *proximal policy optimization* el proceso de recolección de datos, el cálculo de ventajas y la actualización de las políticas se repite en iteraciones de entrenamiento para mejorar gradualmente el rendimiento del agente. Así, el proceso de entrenamiento de este tipo de modelos implica las siguientes etapas:

- El agente interactúa con el entorno y recopila muestras de

experiencia, incluyendo estados, acciones y recompensas.

- Se calculan las ventajas de las acciones tomadas en comparación con las acciones alternativas posibles, lo que ayuda a estimar el valor de cada acción en relación con las otras posibles acciones.
- Se actualizan los parámetros de la política utilizando métodos de optimización que buscan maximizar las recompensas acumuladas. PPO utiliza el algoritmo de optimización de gradiente conjugado (CG) para actualizar los parámetros.
- Para evitar variaciones drásticas en la política, se aplica una restricción a las actualizaciones de políticas utilizando la medida de similitud *clip*. Esto ayuda a garantizar que las actualizaciones no se desvíen demasiado de la política actual.

Por lo tanto, el enfoque de optimización de PPO está basado en restricciones y la medida de similitud *clip* ayudan a estabilizar el proceso de entrenamiento y mejorar el rendimiento del agente. Esta medida permite limitar las actualizaciones de la política durante el proceso de entrenamiento y estabiliza el

proceso de optimización, ya que realizar actualizaciones demasiado grandes que impliquen cambios bruscos en la política puede conducir a rendimientos inestables y subóptimos. Mediante la medida *clip* se restringe el rango en el que se permiten las actualizaciones de los parámetros. En el sistema PPO se utiliza una función objetivo llamada *surrogate objective* que mide la mejora esperada

en el rendimiento de la política. En lugar de utilizar la función objetivo original directamente, se utiliza el ya citado factor *clip* que considera el mínimo entre dos términos en la función objetivo: la probabilidad de la acción seleccionada por la nueva política dividida por la probabilidad de la acción seleccionada por la política anterior y una versión modificada de la probabilidad.

## A.29. REINFORCE

Una técnica englobada dentro del enfoque de *gradient policy* es REINFORCE (también conocida como método de gradiente de Montecarlo)<sup>41</sup>. Es un algoritmo de aprendizaje por refuerzo utilizado para entrenar agentes en el aprendizaje de políticas estocásticas, es decir, políticas que toman decisiones basadas en distribuciones de probabilidad. El objetivo del algoritmo es también ajustar los parámetros de la política para maximizar la recompensa acumulada a lo largo del tiempo y se basa en la estimación del gradiente de la función de valor esperado de la recompensa acumulada. Esto

implica calcular la diferencia entre la recompensa obtenida y una línea base (recompensa promedio) y multiplicarla por el gradiente de log-verosimilitud de la acción tomada en función de los parámetros de la política.

El proceso de entrenamiento con REINFORCE incluye los siguientes pasos:

- Inicializar los parámetros de la política aleatoriamente.
- Generar una trayectoria en el entorno utilizando la política actual y registrar las acciones tomadas y las

<sup>41</sup> Fue propuesto por Ronald J. Williams en 1992



recompensas obtenidas.

- Calcular la recompensa acumulada a lo largo de la trayectoria.
- Calcular el gradiente de log-verosimilitud para cada acción tomada a lo largo de la trayectoria.
- Actualizar los parámetros de la política utilizando el gradiente calculado y una tasa de aprendizaje.
- Los últimos pasos del

proceso se repiten hasta que la política converge.

El algoritmo de REINFORCE utiliza el método de Montecarlo para estimar la recompensa acumulada, y se espera hasta el final del episodio para calcularla. En consecuencia, el agente debe interactuar con el entorno y completar un episodio antes de realizar actualizaciones en los parámetros de la política. Las principales ventajas e inconvenientes del modelo REINFORCE se describen en la Tabla A.10.

Tabla A.10. Ventajas e inconvenientes del modelo REINFORCE.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>○ La reducción de la dimensionalidad permite disminuir la cantidad de variables en un conjunto de datos, lo que facilita la visualización y el análisis.</li> <li>○ Es un enfoque simple pero efectivo para el aprendizaje de políticas en entornos de toma de decisiones.</li> </ul>	<ul style="list-style-type: none"> <li>○ Se basa en la covarianza o correlación de las variables y asume una estructura lineal en los datos.</li> <li>○ Alta varianza de las estimaciones del gradiente.</li> <li>○ Escasa eficiencia en entornos con grandes espacios de acción.</li> </ul>

Fuente: elaboración propia.

### A.30. ACTOR-CRITIC Y VARIANTES

La técnica de gradiente de política *Actor-Critic* es un enfoque avanzado dentro del aprendizaje

por refuerzo que combina elementos de dos enfoques: el método de gradiente de política



(*Actor*) y el método de evaluación de valor (*Critic*). Esta combinación permite un aprendizaje más eficiente y estable al utilizar una red neuronal (el *Actor*) para aprender la política y otra red neuronal (el *Critic*) para estimar la función de valor. El *Actor* es responsable de aprender y mejorar la política del agente. Generalmente, se modeliza como una red neuronal que toma el estado del entorno como entrada y produce una distribución de probabilidad sobre las posibles acciones a tomar. Por su parte, el *Critic* es responsable de estimar la función de valor, que indica la calidad de un estado o una acción en términos de la recompensa esperada. También se modeliza como una red neuronal que toma el estado del entorno como entrada y produce una estimación del valor. El agente interactúa con el entorno, seleccionando acciones según la política aprendida por el *Actor*. A medida que el agente toma acciones, el entorno proporciona recompensas y transiciones de estado. El *Critic* utiliza la función de valor estimada para evaluar la calidad de las acciones tomadas por el *Actor*. Esto proporciona una señal de retroalimentación para actualizar la política. La política del *Actor* se actualiza utilizando el gradiente de la función de valor estimada por

el *Critic*. El objetivo es maximizar la recompensa esperada al ajustar los parámetros del *Actor* en la dirección sugerida por el gradiente. La función de valor del *Critic* se actualiza utilizando el error entre la estimación del valor actual y la estimación objetivo. Esto ayuda a mejorar la precisión de la función de valor a lo largo del tiempo. La técnica *Actor-Critic* tiene como la ventaja de combinar las fortalezas de los métodos de gradiente de política y evaluación de valor. El *Critic* proporciona una señal de retroalimentación más precisa y estable para actualizar la política, mientras que el *Actor* permite una mayor flexibilidad y capacidad para explorar diferentes acciones.

Existen varias variantes y mejoras de la técnica *Actor-Critic*. Por ejemplo, *Advantage Actor-Critic* (A2C) es una variante que combina elementos de los métodos de gradiente de política (*Actor*) y evaluación de valor (*Critic*) para mejorar el rendimiento y la estabilidad del entrenamiento mediante actualizaciones de la política y del valor en paralelo utilizando múltiples agentes o hilos de ejecución. Esto permite un entrenamiento más eficiente y acelerado al aprovechar la capacidad de cómputo de



los sistemas modernos con múltiples núcleos o unidades de procesamiento.

El modelo A2C es reconocido por su eficiencia y estabilidad en comparación con otros enfoques de aprendizaje por refuerzo. Al utilizar múltiples agentes o hilos de ejecución, permite una exploración más diversa del espacio de acciones y una mejor estimación de los valores y ventajas. Además, A2C puede aprovechar el paralelismo disponible en los sistemas modernos para acelerar el entrenamiento y mejorar la eficiencia computacional.

El proceso básico de los modelos *Advantage Actor-Critic* (A2C) se puede describir de la siguiente manera:

- Se inicializan las redes neuronales que representan el *Actor* y el *Critic* con sus respectivos parámetros.
- Se ejecutan múltiples agentes o hilos de ejecución en paralelo para interactuar con el entorno. Cada agente selecciona acciones basadas en la política del *Actor* y registra las transiciones de estado, las acciones tomadas y las recompensas obtenidas.
- Se calcula la función de ventaja para cada transición, que mide la ventaja de tomar una acción en un estado en comparación con el valor estimado del *Critic*. La función de ventaja ayuda a reducir la varianza y proporciona una señal más precisa para actualizar la política.
- El *Critic* se actualiza utilizando el algoritmo de gradiente descendente para minimizar el error entre los valores estimados y los valores objetivo, utilizando la función de ventaja como objetivo.
- El *Actor* se actualiza utilizando el gradiente ascendente para maximizar la función objetivo que combina el logaritmo de la probabilidad de las acciones seleccionadas y la función de ventaja. Esto guía al *Actor* a mejorar la política en dirección a las acciones más ventajosas.
- Los parámetros tanto del *Actor* como del *Critic* se actualizan utilizando técnicas de optimización, como el descenso de gradiente estocástico.
- Los últimos pasos del proceso se repiten hasta que las redes converjan.



Por otra parte, el A3C (*Asynchronous Advantage Actor-Critic*) es una mejora del algoritmo A2C en el aprendizaje por refuerzo. Al igual que A2C, A3C combina elementos de los métodos de gradiente de política (*Actor*) y evaluación de valor (*Critic*) para entrenar agentes en problemas de control continuo. La principal característica de A3C es que utiliza un enfoque asíncrono y distribuido para acelerar el proceso de entrenamiento y mejorar la eficiencia computacional. En lugar de tener un solo hilo de ejecución que interactúa con el entorno y actualiza los parámetros de la red, A3C utiliza múltiples hilos de ejecución que operan de forma asíncrona.

El enfoque asíncrono y distribuido de A3C permite una mayor exploración del espacio de acciones y una mejor utilización de la capacidad de cómputo de los sistemas modernos con múltiples núcleos o unidades de procesamiento. Además, A3C aprovecha la experiencia y la diversidad de los hilos de ejecución para mejorar la calidad de las actualizaciones de las redes y acelerar el entrenamiento.

El proceso operativo del modelo *Asynchronous Advantage Actor-*

*Critic* (A3C) puede describirse de la siguiente forma:

- Se inicializan las redes neuronales que representan el *Actor* y el *Critic* con sus respectivos parámetros.
- Se crean múltiples hilos de ejecución, cada uno con una copia propia de las redes *Actor* y *Critic*.
- Cada hilo de ejecución interactúa con el entorno, seleccionando acciones basadas en la política del *Actor* y registrando las transiciones de estado, las acciones tomadas y las recompensas obtenidas.
- Se calcula la función de ventaja para cada transición, que mide la ventaja de tomar una acción en un estado en comparación con el valor estimado del *Critic*.
- Cada hilo de ejecución actualiza de forma asíncrona las redes *Actor* y *Critic* utilizando los datos recopilados en su propia interacción con el entorno. Esto incluye actualizar los parámetros de la red *Critic* mediante el algoritmo de gradiente descendente y actualizar los parámetros de la red *Actor* mediante el gradiente ascendente.



- De manera periódica, los hilos de ejecución sincronizan sus parámetros con una copia global de las redes *Actor* y *Critic* para evitar que los hilos se desvíen demasiado en sus actualizaciones.
- Los últimos pasos del proceso se repiten hasta que las redes converjan.

### A.31. DETERMINISTIC POLICY GRADIENT (DPG)

Como alternativa a los métodos que consisten en aprender una política estocástica que elige acciones de forma probabilística, los métodos de *determinist policy* aprenden una política determinista que mapea directamente los estados a las acciones. Estos métodos buscan encontrar la mejor acción posible en cada estado. En particular, el *Deterministic Policy Gradient* (DPG) es una técnica utilizada en el aprendizaje por refuerzo para entrenar agentes con políticas que seleccionan una acción específica en lugar de una distribución de probabilidad sobre las acciones. El DPG se enfoca en aprender políticas determinísticas para problemas de control continuo. Su objetivo es aprender una función de valor acción-estado óptima y una política que maximice esa función, utilizando una arquitectura de red neuronal para representar tanto la función de valor como la política determinística. La

función de valor se utiliza para evaluar la calidad de una acción en un estado determinado, mientras que la política determinística se utiliza para seleccionar acciones directamente basadas en los estados.

Una descripción de los principales pasos de un modelo DPG es la siguiente:

- Se inicializan los parámetros de las redes neuronales que representan la función de valor y la política determinística.
- El agente interactúa con el entorno, seleccionando acciones determinísticas basadas en la política actual. Se registran las transiciones de estado y las recompensas obtenidas.
- La función de valor se actualiza utilizando el gradiente descendente para minimizar el error cuadrático medio entre los

valores estimados y los valores objetivo-calculados utilizando la ecuación de Bellman.

- La política determinística se actualiza utilizando el gradiente ascendente para maximizar la función de valor aprendida. Este gradiente se estima utilizando la propiedad de diferenciación de

la función de valor respecto a las acciones tomadas por la política.

- Los últimos pasos se repiten hasta que la función de valor y la política converjan.

Las principales ventajas e inconvenientes de este modelo se describen en la Tabla A.11.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o Elevada capacidad de aprender políticas determinísticas en problemas de control continuo.</li> <li>o Estabilidad del entrenamiento.</li> <li>o Notable capacidad de lidiar con espacios de acción continuos.</li> </ul>	<ul style="list-style-type: none"> <li>o Puede sufrir de problemas de ajuste de hiperparámetros.</li> <li>o Exige de una modelización precisa de la función de valor y la política.</li> </ul>

Tabla A.11. Ventajas e inconvenientes del modelo *Deterministic Policy Gradient*.

Fuente: elaboración propia.

## A.32. DEEP DETERMINISTIC POLICY GRADIENT (DDPG).

La técnica *Deep Deterministic Policy Gradient* (DDPG) es una extensión del algoritmo DPG que combina la idea de aprendizaje de políticas determinísticas con el poder de las redes neuronales profundas (*deep learning*). DDPG se utiliza en problemas de control continuo en el aprendizaje por refuerzo. Su principal característica

es que utiliza las redes neuronales profundas *Actor* y *Critic*. El Actor se encarga de aprender la política determinística, mientras que el *Critic* se encarga de aprender la función de valor acción-estado.

El proceso básico de construcción de un modelo del tipo *Deep Deterministic*



*Policy Gradient* es el que se describe a continuación:

- Se inicializan tanto las redes *Actor* como el *Critic* con sus respectivos parámetros.
- El agente interactúa con el entorno, seleccionando acciones determinísticas basadas en la política del *Actor*. Se registran las transiciones de estado, las acciones tomadas y las recompensas obtenidas.
- La red *Critic* se actualiza utilizando el algoritmo de gradiente de error cuadrático medio (MSE) para minimizar la diferencia entre los valores estimados por la red y los valores objetivo-calculados utilizando la ecuación de Bellman.
- El *Actor* se actualiza utilizando la regla del gradiente ascendente para maximizar la función de valor aprendida por el *Critic*. El gradiente se estima utilizando la propiedad de diferenciación de la función de valor respecto a las acciones tomadas por la política del *Actor*.
- Los parámetros de las redes *Actor* y *Critic* se actualizan utilizando técnicas de optimización, como el descenso de gradiente estocástico.
- Los últimos pasos se repiten hasta que la función de valor y la política converjan.

En la Tabla A.12 se muestran algunas ventajas en inconvenientes de este tipo de modelo.

Tabla A.12. Ventajas e inconvenientes del modelo *Deep Deterministic Policy Gradient*.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>○ Permite abordar problemas con espacios de acción continuos, ya que las redes pueden aproximar funciones no lineales de manera más efectiva.</li> <li>○ Ofrece un aprendizaje efectivo y estable por utilizar memoria de repetición (replay buffer) y almacenar y reutilizar transiciones pasadas.</li> </ul>	<ul style="list-style-type: none"> <li>○ Exige un ajuste adecuado de los hiperparámetros.</li> <li>○ Exige considerar el equilibrio entre la exploración y la explotación para un desempeño óptimo.</li> </ul>

Fuente: elaboración propia.



### A.33. APRENDIZAJE POR REFUERZO BASADO EN MODELOS

El Aprendizaje por Refuerzo Basado en Modelos –*Model-based Reinforcement Learning* por su denominación en inglés– es un enfoque en el campo del aprendizaje por refuerzo que se basa en la construcción y utilización de un modelo del entorno para tomar decisiones y mejorar el rendimiento del agente de aprendizaje. Mientras que, en el aprendizaje por refuerzo tradicional, el agente interactúa directamente con el entorno, toma acciones y recibe retroalimentación en forma de recompensas, en el aprendizaje por refuerzo basado en modelos, se construye un modelo del entorno, que es una representación aproximada o una estimación del comportamiento del entorno. Este modelo puede ser aprendido a partir de datos recopilados de las interacciones con el entorno o puede ser una representación simplificada del entorno basada en conocimiento experto o simulaciones. Una vez que se tiene un modelo del entorno, el agente puede utilizarlo para simular y predecir el comportamiento del entorno antes de tomar acciones reales.

El proceso de construcción del *Model-based Reinforcement Learning* se basa en los siguientes pasos:

- El agente interactúa con el entorno y recopila datos de entrenamiento, como pares de estados, acciones y recompensas.
- Utilizando los datos recopilados, se construye un modelo que pueda simular el comportamiento del entorno. Esto puede implicar el uso de técnicas de aprendizaje automático, como redes neuronales o métodos de regresión.
- Una vez se dispone de un modelo del entorno, el agente puede realizar simulaciones utilizándolo para evaluar el resultado esperado de diferentes acciones en diferentes estados. Esto permite al agente planificar y tomar decisiones óptimas basadas en las predicciones del modelo.
- A medida que el agente interactúa con el entorno real y recibe retroalimentación,



puede actualizar y mejorar el modelo utilizando técnicas de aprendizaje automático, como el ajuste fino o la actualización de los parámetros del modelo.

En la tabla A.13 se ofrecen ventajas e inconvenientes del modelo de aprendizaje por refuerzo basado en modelos.

Ventajas/Fortalezas	Inconvenientes/Limitaciones	Tabla A.13. Ventajas e inconvenientes del <i>Model-based Reinforcement Learning</i> .
<ul style="list-style-type: none"> <li>○ Permite una exploración eficiente y segura del entorno (el agente puede simular y predecir el resultado de las acciones antes de realizarlas).</li> <li>○ Reduce el riesgo de acción en el caso de entornos complejos, costosos o peligrosos.</li> </ul>	<ul style="list-style-type: none"> <li>○ Exige la construcción y el mantenimiento de modelos precisos y actualizados.</li> <li>○ El rendimiento del agente está condicionado por la posible existencia de discrepancias entre el modelo y el entorno real.</li> </ul>	

Fuente: elaboración propia.

### A.34. APRENDIZAJE POR REFUERZO CON HORIZONTES TEMPORALES EXTENDIDOS (*HIERARCHICAL RL*)

El Aprendizaje por Refuerzo con Horizontes Temporales Extendidos -*Hierarchical Reinforcement Learning* (HRL) por su terminología en inglés- es un enfoque de aprendizaje por refuerzo orientado a analizar problemas de largo plazo y de gran escala mediante la descomposición jerárquica de las tareas en múltiples niveles de abstracción. En lugar de tratar de aprender directamente una política óptima en un nivel de granularidad fina, el enfoque

jerárquico divide la tarea en subtareas más manejables y aprende políticas en el nivel de cada subtask. Estas políticas se organizan en una estructura jerárquica en la que las acciones en un nivel superior representan las decisiones de más alto nivel y las acciones en niveles inferiores representan las acciones detalladas dentro de una subtask. El HRL puede tener múltiples niveles de jerarquía, lo que permite una descomposición de



la tarea en diferentes niveles de abstracción. Por ejemplo, el nivel más alto de la jerarquía podría referirse a decisiones estratégicas de alto nivel, mientras que los niveles inferiores pueden centrarse

en movimientos y acciones más específicas. En la Tabla A.14 se muestra las principales ventajas e inconvenientes de este tipo de enfoque.

Tabla A.14. Ventajas e inconvenientes del enfoque *hierarchical reinforcement learning*.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>o Reducción de la complejidad al dividir una tarea compleja en subtareas más pequeñas.</li> <li>o La reutilización del conocimiento (las políticas aprendidas en niveles superiores pueden ser reutilizadas en tareas similares o relacionadas) lo que acelera el aprendizaje en nuevos escenarios.</li> <li>o Al tener niveles de abstracción más altos en la jerarquía, se puede realizar una planificación a largo plazo para optimizar las recompensas acumuladas en el horizonte temporal extendido.</li> <li>o Exploración más eficiente ya que permite enfocar en políticas a nivel de subtareas.</li> </ul>	<ul style="list-style-type: none"> <li>o Existen desafíos en términos de estabilidad del entrenamiento.</li> <li>o El uso eficiente de los recursos no está garantizado.</li> <li>o La generalización a entornos no analizados puede resultar poco eficiente.</li> </ul>

Fuente: elaboración propia.

## A.35. MÉTODOS EVOLUTIVOS

Las técnicas de Métodos Evolutivos (*Evolutionary Methods* por su denominación en inglés) se basan en un enfoque de inteligencia artificial inspirado en la

evolución biológica para resolver problemas de optimización y aprendizaje automático. El proceso evolutivo se apoya en la idea de que las soluciones más



aptas tienen más probabilidades de ser seleccionadas y transmitir sus características beneficiosas a las generaciones futuras de algoritmos. Estos métodos se basan en algoritmos genéticos y operadores de selección, reproducción y mutación para evolucionar soluciones a través de generaciones. A diferencia de los enfoques tradicionales de AI que se centran en el ajuste de parámetros mediante técnicas de optimización basadas en gradiente, los métodos evolutivos exploran el espacio de soluciones de manera más amplia y no requieren información detallada del gradiente. Estas técnicas se utilizan en problemas donde la formulación analítica de una función objetivo es difícil o desconocida, y se busca encontrar soluciones de calidad sin depender de un modelo matemático preciso.

Los pasos principales en la construcción de los Métodos Evolutivos son los siguientes:

- Se crea una población inicial de individuos, donde cada individuo representa una solución potencial al problema.
- Cada individuo en la población se evalúa utilizando una función de aptitud que mide su rendimiento en el problema. La función de aptitud define el criterio de éxito y determina cómo se seleccionan los individuos para la reproducción.
- Los individuos con una mayor aptitud tienen más probabilidades de ser seleccionados para la reproducción y generar descendencia para la próxima generación. Este proceso está influenciado por operadores de selección como la ruleta, torneo o clasificación.
- Los individuos seleccionados se combinan entre sí mediante operadores de cruce o recombinación, que mezclan características de los padres para producir descendientes.
- Los descendientes pueden sufrir mutaciones aleatorias en sus características, lo que introduce diversidad en la población y permite la exploración de nuevas soluciones en el espacio de búsqueda.
- Los descendientes reemplazan a una parte de la población anterior, lo que conduce a una nueva generación de individuos.
- Los últimos pasos se repiten durante varias generaciones de



algoritmos hasta que se cumpla un criterio de terminación, como alcanzar un número máximo de generaciones o una aptitud deseada.

En particular, la técnica de optimización y búsqueda Algoritmos Genéticos (AG) está basada en principios inspirados en la teoría de la evolución biológica y la genética. Estos algoritmos se utilizan para encontrar soluciones aproximadas a problemas complejos que implican la optimización de una función objetivo. Los componentes clave de un algoritmo genético son:

- Las soluciones candidatas, también conocidas como cromosomas, se representan mediante una codificación adecuada, como cadenas de bits o estructuras de datos más complejas. Esta codificación define el espacio de búsqueda en el que el algoritmo opera.
- Se define una función objetivo que evalúa la calidad de cada solución candidata en términos de aptitud. Esta función el grado de ajuste de una solución a los requisitos del problema y

se utiliza para asignar un valor de aptitud a cada individuo.

- Los operadores genéticos, como la selección, la reproducción, la recombinación y la mutación, simulan los procesos de selección natural y reproducción. La selección favorece a los individuos más aptos para sobrevivir y reproducirse. La reproducción implica la combinación de características de los individuos seleccionados mediante operaciones como la recombinación y la mutación.
- El algoritmo itera a través de generaciones sucesivas, aplicando los operadores genéticos en cada iteración. Con el tiempo, se espera que la población evolucione y mejore, generando soluciones óptimas.
- Se establece un criterio de parada para detener el algoritmo, como alcanzar un número máximo de generaciones, lograr un nivel de aptitud deseado o agotar los recursos computacionales disponibles.

## A.36. TÉCNICAS DE LÓGICA DIFUSA (*FUZZY LOGIC*)

La lógica difusa (*fuzzy logic*) se basa en la observación de que las personas toman decisiones basadas en información imprecisa y no numérica. Los modelos borrosos se utilizan para representar datos incompletos e imprecisos. Este tipo de modelos pueden reconocer, representar, manipular, interpretar y utilizar datos e información que son vagos e inciertos. La teoría de conjuntos borrosos se asemeja al razonamiento humano en el uso de información aproximada y con incertidumbre para generar decisiones. La principal contribución de la teoría del control difuso es su capacidad para manejar muchos problemas prácticos que los métodos de control convencionales no pueden manejar adecuadamente.

Las técnicas de lógica difusa son un enfoque de modelización y razonamiento que permite manejar la incertidumbre y la imprecisión en los datos. A diferencia de la lógica clásica, que utiliza valores binarios (verdadero/falso), la lógica difusa permite la asignación de grados de pertenencia a diferentes categorías. Se basa en la idea

de que los objetos o conceptos del mundo real pueden tener grados de pertenencia a una determinada categoría en lugar de ser estrictamente verdaderos o falsos. Esto es especialmente útil cuando los datos son ambiguos o subjetivos, y no se pueden describir de manera precisa mediante valores discretos. El principal componente de la lógica difusa es la *función de membresía*, que asigna un grado de pertenencia a un elemento en relación con una categoría específica. Esta función puede tener diferentes formas, como triangular, trapezoidal o gaussiana, dependiendo del contexto y la naturaleza de los datos. La lógica difusa también utiliza operadores lógicos difusos, como "y" difuso, "o" difuso y "no" difuso, que permiten combinar los grados de pertenencia y realizar inferencias difusas. Estas inferencias difusas se basan en reglas "si-entonces" que relacionan las entradas difusas con las salidas difusas.

Esta técnica que permite modelizar y razonar con información incierta y subjetiva, se está mostrando especialmente útil en situaciones donde no



se dispone de datos precisos o donde la precisión absoluta no es necesaria. Además, esta técnica ha demostrado ser efectiva para capturar la

naturaleza imprecisa del mundo real y mejorar la capacidad de los sistemas de inteligencia artificial para tomar decisiones en entornos complejos.

### A.37. PROCESAMIENTO DEL LENGUAJE NATURAL (NLP)

El Procesamiento del Lenguaje Natural -*Natural Language Processing* (NLP) por su terminología en inglés- es una rama de la inteligencia artificial que se ocupa de la interacción entre las computadoras y el lenguaje humano en forma escrita o hablada. Su objetivo es permitir a las máquinas comprender, interpretar, analizar y generar texto de manera similar a como lo hacen los seres humanos. El NLP abarca una amplia gama de tareas y técnicas, incluyendo:

- Reconocimiento del habla. El NLP se ocupa de la conversión de voz en texto, lo que permite a las máquinas transcribir y procesar la entrada de voz.
- Análisis morfológico. Consiste en descomponer palabras en unidades más pequeñas, como raíces y afijos, para comprender la estructura morfológica de las palabras.
- Análisis sintáctico. Se refiere a analizar la estructura gramatical de las oraciones y determinar la relación entre las palabras. Esto incluye tareas como el etiquetado gramatical, el análisis de dependencia y el análisis de constituyentes.
- Análisis semántico. Se enfoca en comprender el significado de las oraciones y las relaciones entre las palabras. Involucra tareas como el análisis de sentimientos, la desambiguación léxica y la identificación de entidades nombradas.
- Traducción automática. El NLP se utiliza para desarrollar sistemas de traducción automática que permiten la traducción de texto de un idioma a otro.
- Generación de lenguaje natural. Consiste en generar texto o habla en lenguaje humano de manera

comprensible y coherente. Esto incluye tareas como la generación de resúmenes, la escritura automática y la generación de diálogos.

- Búsqueda y recuperación de información. El NLP se aplica en sistemas de búsqueda de información que permiten a los usuarios encontrar documentos o información relevante en

grandes conjuntos de datos.

El NLP utiliza una combinación de enfoques basados en reglas, aprendizaje automático y técnicas estadísticas para desarrollar modelos y algoritmos que permiten a las máquinas comprender y procesar el lenguaje humano de manera efectiva. Hay varios sistemas de NLP destacables.

### **A.38. GPT (GENERATIVE PRE-TRAINED TRANSFORMER)**

Es un modelo de lenguaje basado en la arquitectura de *Transformers* y entrenado en un amplio conjunto de datos de texto no etiquetado. Fue desarrollado por OpenAI y ha sido ampliamente reconocido por su capacidad para generar texto coherente y de alta calidad. El GPT utiliza un enfoque de aprendizaje automático no supervisado o preentrenamiento, donde se expone al modelo a grandes cantidades de texto para que aprenda a capturar patrones y estructuras del lenguaje. Este proceso de preentrenamiento permite al modelo adquirir conocimiento general sobre el lenguaje y desarrollar una representación interna de las relaciones entre palabras y frases.

La arquitectura del GPT se basa en la red neuronal transformadora (*Transformer*), que es conocida por su capacidad para modelizar dependencias a largo plazo en secuencias de datos. Esta arquitectura permite que el GPT capture relaciones contextuales y genere texto coherente. Una vez que el modelo GPT ha sido preentrenado, se puede ajustar o afinar para tareas específicas, como generación de texto, traducción automática, resumen de textos, respuesta a preguntas y muchas otras aplicaciones relacionadas con el procesamiento del lenguaje natural. El GPT ha sido entrenado en grandes conjuntos de datos de texto, como libros, artículos de noticias

y páginas web, lo que le permite tener un conocimiento profundo del lenguaje y la capacidad de generar texto que se asemeje a un habla humana coherente y convincente. Una de las características distintivas del GPT

es su capacidad para generar texto de forma autónoma. GPT ha demostrado ser una herramienta valiosa en muchas aplicaciones y ha impulsado avances significativos en el procesamiento del lenguaje natural.

### **A.39. BERT (BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS)**

Es un modelo de lenguaje basado en la arquitectura de *Transformers* y es uno de los avances más destacados en el procesamiento del lenguaje natural. Fue presentado por Google en 2018 y ha logrado un gran impacto en una amplia gama de tareas relacionadas con el lenguaje. El enfoque principal de BERT es capturar el contexto bidireccional de las palabras en una oración, lo que significa que considera tanto el contexto anterior como el posterior al procesar cada palabra. Esto se logra utilizando una técnica llamada "atención bidireccional" en la arquitectura de *Transformers*. A diferencia de los modelos de lenguaje tradicionales que procesan las palabras en una oración de manera secuencial, BERT procesa todas las palabras simultáneamente y captura las

dependencias contextuales en ambas direcciones. Esto permite que BERT tenga una comprensión más profunda del contexto y mejore el rendimiento en tareas que requieren un entendimiento más sofisticado del lenguaje. BERT se entrena en grandes cantidades de texto no etiquetado utilizando el aprendizaje no supervisado. Durante el entrenamiento, el modelo se expone a múltiples tareas de preentrenamiento, como la predicción de palabras ocultas en una oración o la clasificación de oraciones como pares coherentes o incoherentes. Una vez que BERT ha sido preentrenado en estas tareas, se puede ajustar o afinar para tareas específicas utilizando datos etiquetados. Este ajuste fino permite que BERT sea utilizado para una amplia gama de tareas de procesamiento del lenguaje



natural, como la clasificación de textos, la extracción de información, el análisis de sentimientos y muchas otras. Una de las ventajas clave de BERT es su capacidad para capturar las relaciones contextuales en el lenguaje, lo que permite un rendimiento excepcional en tareas que requieren un entendimiento profundo del

contexto. Además, BERT ha demostrado ser altamente efectivo en la transferencia de conocimiento entre diferentes tareas de procesamiento del lenguaje natural, lo que lo convierte en una herramienta valiosa para el desarrollo de aplicaciones y sistemas de AI basados en el lenguaje.

## A.40. APRENDIZAJE FEDERADO

El Aprendizaje Federado (*Federated Learning* por su terminología en inglés) es una técnica que permite entrenar modelos de AI utilizando datos distribuidos en diferentes dispositivos o ubicaciones sin necesidad de transferir los datos a un servidor centralizado. El aprendizaje federado garantiza la privacidad y seguridad de los datos al mantenerlos en los dispositivos locales y sólo compartir los conocimientos agregados del modelo. En lugar de enviar los datos a un servidor central, el proceso de aprendizaje federado se lleva a cabo en los dispositivos locales, como teléfonos móviles, dispositivos de Internet de las cosas o computadoras personales. Estos

dispositivos mantienen sus datos de manera segura y únicamente comparten los conocimientos agregados o actualizaciones del modelo con un servidor central.

El proceso de aprendizaje federado consta de los siguientes pasos:

- En la configuración inicial se establece una arquitectura de modelo y una estrategia de entrenamiento.
- Se distribuye una versión inicial del modelo a los dispositivos locales.
- Cada dispositivo local entrena el modelo utilizando sus propios datos locales. Este proceso se



realiza de manera autónoma y segura en cada dispositivo.

- Los dispositivos locales envían únicamente las actualizaciones del modelo, en lugar de los datos brutos, al servidor central.
- El servidor central recibe las actualizaciones de los dispositivos locales y las utiliza para actualizar el modelo global.

- Se realizan iteraciones sucesivas lo que permite al modelo mejorar a medida que se recopila más información de los dispositivos locales.

En la tabla A.15 se ofrecen las principales ventajas e inconvenientes del enfoque de Aprendizaje Federado.

Tabla A.15. Ventajas e inconvenientes del enfoque de Aprendizaje Federado.

Ventajas/Fortalezas	Inconvenientes/Limitaciones
<ul style="list-style-type: none"> <li>○ Mejora la privacidad y la seguridad de los datos, ya que los datos permanecen en los dispositivos locales.</li> <li>○ Se reduce la necesidad de transferir grandes volúmenes de datos, lo que ahorra ancho de banda y reduce la carga en los servidores centrales</li> <li>○ Permite el entrenamiento de modelos en entornos distribuidos, como dispositivos móviles con recursos limitados, y en entornos con conectividad intermitente.</li> </ul>	<ul style="list-style-type: none"> <li>○ Exige de un plan detallado para la selección de datos que deben ser transferidos.</li> <li>○ Requiere de una organización robusta para ejecutar las transferencias de datos.</li> </ul>

Fuente: elaboración propia.

## A.41. EL APRENDIZAJE AUTOMATIZADO EN EL BORDE

El *aprendizaje automático en el borde* se refiere a la aplicación de técnicas de aprendizaje automático directamente en dispositivos o sistemas ubicados

en la periferia de una red, en lugar de depender de servidores o centros de datos remotos para realizar el procesamiento y la toma de decisiones. Este enfoque se

relaciona estrechamente con el concepto de “*edge computing*” y se utiliza para aplicaciones que requieren una respuesta rápida y un procesamiento en tiempo real, como en el Internet de las cosas.

Al realizar el procesamiento de datos y la inferencia de modelos de aprendizaje automático en el borde, se pueden tomar decisiones en tiempo real sin depender de una conexión de red constante, por lo que la latencia es reducida. Esto es esencial para aplicaciones como vehículos autónomos, sistemas de seguridad y equipos médicos. Por otra parte, al mantener ciertos datos y procesos de aprendizaje automático en el borde, y limitar la transmisión de

datos sensibles a través de la red puede mejorarse la privacidad y seguridad del sistema, al tiempo que se produce una mayor eficiencia de ancho de banda y se reduce la congestión en la red.

Dado que el aprendizaje automático en el borde permite la distribución de modelos de aprendizaje automático a través de múltiples dispositivos, se mejora la escalabilidad. Los dispositivos en la periferia pueden tomar decisiones localmente sin necesidad de comunicarse con un servidor central, lo que puede ser crítico en situaciones en las que la comunicación con un servidor central podría fallar o ser ineficiente.

## A.42. SISTEMAS DE INTELIGENCIA HÍBRIDA

Aunque queda mucho por descubrir respecto de cómo entrena y se autoorganiza un cerebro biológico para procesar y extraer información diversa y compleja, algunos recientes avances en “neurobiología” permiten ya destacar ciertos mecanismos clave. Aprovechando esta base, el sistema de inteligencia híbrida –*Hybrid*

*Intelligent System* (HIS), por su terminología en inglés– es un enfoque que combina múltiples técnicas y metodologías de inteligencia artificial para resolver problemas complejos. Estos sistemas buscan aprovechar las fortalezas de diferentes enfoques de AI y combinarlos de manera sinérgica para obtener un rendimiento mejorado y



una solución más completa.

La idea es que, puesto que ningún enfoque único de inteligencia artificial puede resolver todos los problemas de manera óptima, la combinación de diferentes técnicas permite la superación de las limitaciones individuales lo que hace posible abordar problemas más complejos y obtener resultados más precisos. Por ejemplo, un sistema de inteligencia híbrida puede combinar el aprendizaje supervisado con el aprendizaje por refuerzo. El primero se utilizaría para aprovechar datos etiquetados y aprender a clasificar o predecir, mientras que el segundo se puede utilizar

para que el sistema aprenda a través de la interacción con un entorno y la retroalimentación de recompensas. Al combinar estos enfoques, se puede lograr un sistema que aprenda de manera eficiente a través de la experiencia y también utilice conocimiento previo para realizar predicciones precisas. Otros ejemplos de técnicas que se pueden combinar en sistemas de inteligencia híbrida incluyen redes neuronales profundas, lógica difusa, algoritmos genéticos, sistemas expertos, procesamiento del lenguaje natural, entre otros.



## REFERENCIAS

- Agapitos, A., Brabazon, A. and O'Neill, M. (2017) "Regularised gradient boosting for financial time-series modelling", *Computational Management Science*, 14: 367-391.
- Ali, M., Khan, D. M., Aamir, M., Ali, A., and Ahmad, Z. (2021) "Predicting the direction movement of financial time series using artificial neural network and support vector machine", *Complexity*: 1-13.
- Altan, A. and Karasu, S. (2019) "The effect of kernel values in support vector machine to forecasting performance of financial time series", *The Journal of Cognitive Systems*, 4(1): 17-21.
- Annaert, J., Claes, A. G., De Ceuster, M. J., and Zhang, H. (2013) "Estimating the spot rate curve using the Nelson-Siegel model: A ridge regression approach", *International Review of Economics & Finance*, 27: 482-496.
- Aziz, S. y Dowling, M. (2019) "Machine Learning and AI for Risk Management" en *Disrupting finance: FinTech and strategy in the 21st century*: 33-50. Lynn, T., Mooney, J. G., Rosati, P., y Cummins, M. (Editores). Springer Nature. Palgrave. MacMillan.
- Bellotti, T. and Crook, J. (2009) "Support vector machines for credit scoring and discovery of significant features", *Expert Systems with Applications*, 36(2): 3302-3308.
- Bickel, P. J., Ritov, Y. A., Zakai, A., and Yu, B. (2006) "Some Theory for Generalized Boosting Algorithms", *Journal of Machine Learning Research*, 7(5).
- Blanchard, G., Lugosi, G., and Vayatis, N. (2003) "On the rate of convergence of regularized boosting classifiers", *Journal of Machine Learning Research*, 4: 861-894.
- Boyacioglu, M. A., Kara, Y., and Baykan, Ö. K. (2009) "Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey", *Expert Systems with Applications*, 36(2): 3355-3366.
- Breiman, L. (1996) "Bagging predictors", *Machine Learning*, 24: 123-140.
- Bühlmann, P. and Hothorn, T. (2007) "Boosting algorithms: Regularization, prediction and model fitting", *Statistical Science*, 22 (4): 477-505. *Institute of Mathematical Statistics*, 200.
- Cao, L. and Tay, F. E. (2001) "Financial forecasting using support vector machines", *Neural Computing & Applications*, 10: 184-192.
- Cao, L. J. and Tay, F. E. H. (2003) "Support vector machine with adaptive parameters in financial time series forecasting". *IEEE Transactions on Neural Networks*, 14(6): 1506-1518.
- Carmona, P., Climent, F., and Momparler, A. (2019). Predicting failure in the US banking sector: An extreme Gradient Boosting Approach", *International Review of Economics & Finance*, 61: 304-323.
- Chang, Y. C., Chang, K. H., and Wu, G. J. (2018) "Application of eXtreme gradient boosting trees in the construction of credit risk assessment models for financial institutions", *Applied Soft Computing*, 73: 914-920.
- Chaudhuri, A. and De, K. (2011) "Fuzzy support vector machine for bankruptcy prediction", *Applied Soft Computing*, 11(2): 2472-2486.

- Chauhan, K., Jani, S., Thakkar, D., Dave, R., Bhatia, J., Tanwar, S., and Obaidat, M. S. (2020) "Automated machine learning: The new wave of machine learning" en *2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 205-212). IEEE.
- Chen, Q. (2023) "Pricing multi-dimensional American options using kernel ridge regression".
- Chen, P. H., Lin, C. J. and Schölkopf, B. (2005) "A tutorial on  $\nu$ -support vector machines. Applied Stochastic Models", *Business and Industry*, 21(2): 111-136.
- Chopani, M. R., Nassirzadeh, F., and Saehi, M. (2017) "Capability of models Support Vector Regression, Least Angle Regression and Adaptive Neuro Fuzzy Inference System for Earning Per Share forecasting", *Monetary & Financial Economics*, 23(12), 161-188.
- Chow, C., Frame, K., Likhtman, S., Spooner, N. y Wong, J. (2019) *Investor's expectations on responsible artificial intelligence and data governance*. Hermes Investment Management.
- Čížek, P., Weron, R., Härdle, W., Härdle, W., Moro, R., and Schäfer, D. (2005) "Predicting bankruptcy with support vector machines": 225-248. Springer Berlin Heidelberg.
- Climent, F., Momparler, A., and Carmona, P. (2019) "Anticipating bank distress in the Eurozone: An extreme gradient boosting approach", *Journal of Business Research*, 101: 885-896.
- Cui, L., Bai, L., Wang, Y., Jin, X., and Hancock, E. R. (2021) "Internet financing credit risk evaluation using multiple structural interacting elastic net feature selection", *Pattern recognition*, 114: 107835.
- Dai, W., Xue, G. R., Yang, Q., and Yu, Y. (2007) "Transferring naive bayes classifiers for text classification" en AAAI, (7): 540-545.
- Danenas, P. and Garsva, G. (2015) "Selection of support vector machines based classifiers for credit risk domain", *Expert Systems with Applications*, 42(6): 3194-3204.
- Das, S. P. and Padhy, S. (2012) "Support vector machines for prediction of futures prices in Indian stock market", *International Journal of Computer Applications*, 41(3).
- Davis, J., Devos, L., Reyners, S., and Schoutens, W. (2020) "Gradient boosting for quantitative finance", *Journal of Computational Finance*, 24 (4).
- Dorugade, A. V. (2014) "New ridge parameters for ridge regression", *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 15: 94-99.
- Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004) "Least angle regression", *The Annals of Statistics*, 32 (2): 407-499.
- Elshawi, R., Maher, M., and Sakr, S. (2019) "Automated machine learning: State-of-the-art and open challenges", *arXiv:1906.02287*.
- Erdogan, B. E. (2013) "Prediction of bankruptcy using support vector machines: an application to bank bankruptcy", *Journal of Statistical Computation and Simulation*, 83(8): 1543-1555.
- Fan, A. and Palaniswami, M. (2000) "Selecting bankruptcy predictors using a support vector machine approach", *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, 6: 354-359. IEEE.
- Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., and Hutter, F. (2015). "Efficient and robust automated machine learning", *Advances in Neural Information Processing Systems*, 28.
- Fičura, M. (2017) "Forecasting Foreign Exchange Rate Movements with k-Nearest-Neighbour, Ridge



- Regression and Feed-Forward Neural Networks”, *Ridge Regression and Feed-Forward Neural Networks* (22 enero).
- Freund, Y. and Schapire, R. (1997) “A decision-theoretic generalization of on-line learning and an application to boosting”, *Journal of Computer and System Science*, 55, 119-139.
- Friedman, J. H. (2001) “Greedy function approximation: A gradient boosting machine”, *The Annals of Statistics*, 29 (5): 1189- 1232.
- Friedman, J., Hastie, T., and Tibshirani, R. (2000) “Additive logistic regression: a statistical view of boosting”, *The Annals of Statistics*, 28(2): 337-407.
- Gao, W. and Yang, H. (2022) “Time-varying group lasso Granger causality graph for high dimensional dynamic system”, *Pattern Recognition*, 130: 108789.
- Geczy, C. (2024) “Capitalizing on the Potential of Artificial Intelligence”. Wharton’s Future of Finance Forum (Mesa Redonda).
- Geraldo-Campos, L. A., Soria, J. J., and Pando-Ezcurra, T. (2022) “Machine Learning for credit risk in the Reactive Peru Program: a comparison of the Lasso and Ridge regression models”, *Economies*, 10(8):188.
- Ghosh, P., Neufeld, A., and Sahoo, J. K. (2022) “Forecasting directional movements of stock prices for intraday trading using LSTM and random forests”, *Finance Research Letters*, 46: 102280.
- Gong, X., Si, Y. W., Fong, S., and Biuk-Aghai, R. P. (2016) “Financial time series pattern matching with extended UCR suite and support vector machine”, *Expert Systems with Applications*, 55: 284-296.
- Goodfellow, I., Bengio, Y., Courville, A., and Bengio, Y. (2016). *Deep learning* (Vol. 1). Cambridge: MIT Press.
- Grömping, U. (2009) “Variable importance assessment in regression: linear regression versus random forest”, *The American Statistician*, 63(4): 308-319.
- Gui, B., Wei, X., Shen, Q., Qi, J., and Guo, L. (2014) “Financial time series forecasting using support vector machine”, en *Tenth International Conference on Computational Intelligence and Security*: 39-43. IEEE.
- Gupta, I., and Nagpal, G. (2020) *Artificial intelligence and expert systems*. Mercury Learning and Information.
- Hans, C. (2011) “Elastic net regression modeling with the orthant normal prior”, *Journal of the American Statistical Association*, 106(496): 1383-1393.
- Härdle, W. K., Wang, W., and Zboňáková, L. (2017) “Time varying quantile Lasso”, *Applied Quantitative Finance*: 331-353. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Harmon, P. and King, D. (1985) *Expert systems: Artificial intelligence in business*. John Wiley & Sons, Inc.
- He, L., Chen, Y., Zhong, C. and Wu, K. (2022) “Granular elastic network regression with stochastic gradient descent”, *Mathematics*, 10(15): 2628.
- Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., and Scholkopf, B. (1998) “Support vector machines”, *IEEE Intelligent Systems and their applications*, 13(4): 18-28.
- Hoerl, A. E., and Kennard, R. W. (1970) “Ridge regression: applications to nonorthogonal problems”, *Technometrics*, 12(1): 69-82.



- Horak, J., Vrbka, J., and Suler, P. (2020) "Support vector machine methods and artificial neural networks used for the development of bankruptcy prediction models and their comparison", *Journal of Risk and Financial Management*, 13(3): 60.
- Horvitz, E. J., Breese, J. S., and Henrion, M. (1988) "Decision theory in expert systems and artificial intelligence," *International Journal of Approximate Reasoning*, 2(3), 247-302.
- Hu, R., Monebhurrin, V., Himeno, R., Yokota, H., and Costen, F. (2018) "An adaptive least angle regression method for uncertainty quantification in FDTD computation", *IEEE Transactions on Antennas and Propagation*, 66(12): 7188-7197.
- Huang, Y. and Li, L. (2011) "Naive Bayes classification algorithm based on small sample set" en *IEEE International Conference on Cloud Computing and Intelligence Systems*: 34-39. IEEE.
- Huang, J., Ma, S., and Zhang, C. H. (2008) "Adaptive Lasso for sparse high-dimensional regression models", *Statistica Sinica*:1603-1618.
- Huang, W., Nakamori, Y., and Wang, S. Y. (2005) "Forecasting stock market movement direction with support vector machine", *Computers & Operations Research*, 32(10): 2513-2522.
- Huang, Y. P. and Yen, M. F. (2019) "A new perspective of performance comparison among machine learning algorithms for financial distress prediction", *Applied Soft Computing*, 83: 105663.
- Hunt, V. D. (2012) *Artificial intelligence & expert systems sourcebook*. Springer Science & Business Media.
- Hutter, F., Kotthoff, L., and Vanschoren, J. (2019) *Automated machine learning: methods, systems, challenges*. Springer Nature.
- Ince, H., and Trafalis, T. B. (2007) "Kernel principal component analysis and support vector machines for stock price prediction", *IIE Transactions*, 39(6): 629-637.
- Jackson, P. (1986) *Introduction to expert systems*. United States.
- Januaviani, T. M. A., and Bon, A. T. (2019) "The LASSO (Least absolute shrinkage and selection operator) method to predict Indonesian foreign exchange deposit data", *Proceedings of the International Conference on Industrial Engineering and Operations Management*: 5-7.
- Jiang, L., Cai, Z., Wang, D., and Jiang, S. (2007) "Survey of improving k-nearest-neighbor for classification" en *Fourth international conference on fuzzy systems and knowledge discovery (FSKD)*, (1): 679-683. IEEE.
- Kalatzis, A. E., Bassetto, C. F., and Azzoni, C. R. (2011) "Multicollinearity and financial constraint in investment decisions: a Bayesian generalized ridge regression", *Journal of Applied Statistics*, 38(2): 287-299.
- Kewat, P., Sharma, R., Singh, U., and Itare, R. (2017) "Support vector machines through financial time series forecasting" en *International conference of Electronics, Communication and Aerospace Technology (ICECA)*, 2: 471-477. IEEE.
- Khaidem, L., Saha, S., and Dey, S. R. (2016) "Predicting the direction of stock market prices using random forest". *arXiv:1605.00003*.
- Khan, J. A., Van Aelst, S., and Zamar, R. H. (2007) "Robust linear model selection based on least angle regression", *Journal of the American Statistical Association*, 102(480): 1289-1299.
- Kramer, O. and Kramer, O. (2013) "K-nearest neighbors. Dimensionality reduction with unsupervised



- nearest neighbors," *Intelligent Systems Reference Library*, vol 51. Springer, Berlin, Heidelberg: 13-23.
- Kumar, D., Meghwani, S. S., and Thakur, M. (2016) "Proximal support vector machine based hybrid prediction models for trend forecasting in financial markets", *Journal of Computational Science*, 17: 1-13.
- Kumar, M., and Thenmozhi, M. (2006) "Forecasting stock index movement: A comparison of support vector machines and random forest" en *Indian institute of Capital Markets 9th Capital Markets Conference Paper*.
- Lao, Y., Qi, F., Zhou, J., and Fang, X. (2021) "A prediction method based on extreme gradient boosting tree model and its application", *Journal of Physics: Conference Series*, 1995 (1): 012017.
- Larose, D. T. and Larose, C. D. (2014). *K-Nearest Neighbor algorithm*.
- Lee, M. C. and To, C. (2010) "Comparison of support vector machine and back propagation neural network in evaluating the enterprise financial distress". *arXiv:1007.5133*.
- Leung, K. M. (2007) "Naive bayesian classifier", *Polytechnic University Department of Computer Science/ Finance and Risk Engineering WP*: 123-156.
- Li, X. (2013) "Using random forest for classification and regression", *Chinese Journal of Applied Entomology*, 50(4): 1190-1197.
- Li, H., Li, C. J., Wu, X. J., and Sun, J. (2014) "Statistics-based wrapper for feature selection: An implementation on financial distress identification with support vector machine", *Applied Soft Computing*, 19: 57-67.
- Li, Y., Stasinakis, C., and Yeo, W. M. (2022) "A hybrid XGBoost-MLP model for credit risk assessment on digital supply chain finance", *Forecasting*, 4(1): 184-207.
- Liang, X., Zhang, H., Xiao, J., and Chen, Y. (2009) "Improving option price forecasts with neural networks and support vector regressions", *Neurocomputing*, 72(13-15): 3055-3065.
- Liebowitz, J. (1988) *Introduction to expert systems*. Mitchell Publishing, Inc.
- Liu, C., Chan, Y., Alam Kazmi, S. H., and Fu, H. (2015) "Financial fraud detection model: Based on random forest", *International Journal of Economics and Finance*, 7(7): 178-188.
- Liu, J., Li, C., Ouyang, P., Liu, J., and Wu, C. (2022) "Interpreting the prediction results of the tree-based gradient boosting models for financial distress prediction with an explainable machine learning approach", *Journal of Forecasting*, 42 (5): 1112-1137.
- Liu, J., Wu, C., and Li, Y. (2019) "Improving financial distress prediction using financial network-based information and GA-based gradient boosting method", *Computational Economics*, 53: 851-872.
- Lugosi, G. and Vayatis, N. (2004) "On the Bayes-risk consistency of regularized boosting methods", *The Annals of Statistics*, 32(1): 30-55.
- Madge, S. and Bhatt, S. (2015) "Predicting stock price direction using support vector machines". *Independent work report spring*, 45.
- Madigan, D., and Ridgeway, G. (2004) "Discussion of Least angle regression by Efron et al.", *arXiv/0406469*.
- Mammone, A., Turchi, M., and Cristianini, N. (2009) "Support vector machines", *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(3): 283-289.



- Mason, L., Baxter, J., Bartlett, P., and Frean, M. (1999) "Boosting algorithms as gradient descent in function space, *NIPS*, 11.
- McDonald, G. C. (2009) "Ridge regression", *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1): 93-100.
- Meyer, D. and Wien, F. T. (2001) "Support vector machines", *R News*: 1(3): 23-26.
- Michie, D. (1985) "Current developments in artificial intelligence and expert systems", *Zygon*<sup>®</sup>, 20(4), 375-389.
- Mitchell, R. and Frank, E. (2017) "Accelerating the XGBoost algorithm using GPU computing", *Peer J Computer Science*, 3, e127.
- Mori, H. and Umezawa, Y. (2007) "Credit risk evaluation in power market with random forest" en *IEEE International Conference on Systems, Man and Cybernetics*: 3737-3742. IEEE.
- Mosquera, R., Castrillón, O. D., and Parra, L. (2018) "Máquinas de soporte vectorial, clasificador Naïve Bayes y algoritmos genéticos para la predicción de riesgos psicosociales en docentes de colegios públicos colombianos", *Información Tecnológica*, 29(6): 153-162.
- Natekin, A., and Knoll, A. (2013) "Gradient boosting machines, a tutorial", *Frontiers in neurobotics*, 7: 21.
- Nie, X. and Deng, G. (2020) "Enterprise financial early warning based on lasso regression screening variables", *Journal of Financial Risk Management*, 9(04): 454.
- Noble, W. S. (2006) "What is a support vector machine?," *Nature Biotechnology*, 24 (12): 1565-1567.
- Nobre, J., and Neves, R. F. (2019) "Combining principal component analysis, discrete wavelet transform and XGBoost to trade in the financial markets", *Expert Systems with Applications*, 125: 181-194.
- Nti, I. K. and Somanathan, A. R. (2022) "A Scalable RF-XGBoost Framework for Financial Fraud Mitigation", *IEEE Transactions on Computational Social Systems*.
- Okasha, M. K. (2014) "Using support vector machines in financial time series forecasting", *International Journal of Statistics and Applications*, 4(1): 28-39.
- Osborne, M. R., Presnell, B., and Turlach, B. A. (2000) "On the LASSO and its dual", *Journal of Computational and Graphical Statistics*, 9(2): 319-337.
- Owen, A. B. (2007) "A robust hybrid of lasso and ridge regression", *Contemporary Mathematics*, 443(7): 59-72.
- Patel, H., Parikh, S., Patel, A., and Parikh, A. (2019) "An application of ensemble random forest classifier for detecting financial statement manipulation of Indian listed companies", *Recent Developments in Machine Learning and Data Analytics*: IC3,18: 349-360. Springer Singapore.
- Patrick, E. A. and Fischer III, F. P. (1970) "A generalized k-nearest neighbor rule," *Information and Control*, 16(2): 128-152.
- Patterson, D. (1990). *Introduction to artificial intelligence and expert systems*. Prentice-Hall, Inc.
- Pereira, J. M., Basto, M., and Da Silva, A. F. (2016) "The logistic lasso and ridge regression in predicting corporate failure", *Procedia Economics and Finance*, 39: 634-641.
- Pérez-Cruz, F., Afonso-Rodríguez, J. A., and Giner, J. (2003) "Estimating GARCH models using support vector machines", *Quantitative Finance*, 3(3):163.

- Peterson, L. E. (2009) "K-nearest neighbor", *Scholarpedia*, 4(2), 1883.
- Pisner, D. A. and Schnyer, D. M. (2020). "Support vector machine" en *Machine Learning*: 101-121. Academic Press.
- Qin, R. (2022) "The Construction of Corporate Financial Management Risk Model Based on XGBoost Algorithm", *Journal of Mathematics*. ID 2043369.
- Ramraj, S., Uzir, N., Sunil, R., and Banerjee, S. (2016) "Experimenting XGBoost algorithm for prediction and classification of different datasets", *International Journal of Control Theory and Applications*, 9(40): 651-662.
- Ranstam, J. and Cook, J. A. (2018) "LASSO regression", *Journal of British Surgery*, 105(10): 1348-1348.
- Reid, S., Tibshirani, R., and Friedman, J. (2016) "A study of error variance estimation in lasso regression", *Statistica Sinica*: 35-67.
- Rolston, D. W. (1988) *Principles of artificial intelligence and expert systems development*. McGraw-Hill, Inc.
- Salehi, M., Moradi, M., and Molaei, S. (2015) "Forecasting systematic risk by Least Angel Regression, AdaBoost and Kernel Ridge Regression", *Modern Applied Science*, 9(11): 135.
- Sant'Anna, L. R., Caldeira, J. F., and Filomena, T. P. (2020) "Lasso-based index tracking and statistical arbitrage long-short strategies", *The North American Journal of Economics and Finance*, 51: 101055.
- Schapire, R. E. and Singer, Y. (1998) "Improved boosting algorithms using confidence-rated predictions", *Proceedings of COLT*, 11.
- Schoech, D., Jennings, H., Schkade, L. L., and Hooper-Russell, C. (1985) "Expert systems: Artificial intelligence for professional decisions", *Computers in Human Services*, 1(1), 81-115.
- Segal, M. R. (2004) *Machine learning benchmarks and random forest regression*. Kluwer Academic Publishers.
- Shin, K. S., Lee, T. S., and Kim, H. J. (2005) "An application of support vector machines in bankruptcy prediction model", *Expert Systems with Applications*, 28(1), 127-135.
- Shrivastava, S., Jeyanthi, P. M., and Singh, S. (2020) "Failure prediction of Indian Banks using SMOTE, Lasso regression, bagging and boosting", *Cogent Economics & Finance*, 8(1): 1729569.
- Shu, L., Shi, F., and Tian, G. (2020) "High-dimensional index tracking based on the adaptive elastic net", *Quantitative Finance*, 20(9): 1513-1530.
- Sipper, M., and Moore, J. H. (2022) "AddGBoost: A gradient boosting-style algorithm based on strong learners", *Machine Learning with Applications*, 7, 100243.
- Sivaram, M., Lydia, E. L., Pustokhina, I. V., Pustokhin, D. A., Elhoseny, M., Joshi, G. P., and Shankar, K. (2020) "An optimal least square support vector machine based earnings prediction of blockchain financial products", *IEEE Access*, 8: 120321-120330.
- Sol, H. G., Cees, A. and de Vries Robbé, P. F. eds. (2013) *Expert systems and artificial intelligence in decision support systems: proceedings of the Second Mini Euroconference*, Lunteren, The Netherlands, 17-20 Noviembre 1985. Springer Science & Business Media, 2013.
- Soria-Quijaite, J. J., Geraldo-Campos, L. A., and Pando-Ezcurra, T. (2022) "Machine learning for credit risk in the Reactive Peru Program: A comparison of the Lasso and Ridge regression models", *Economies*, 10: 188.



- Sun, J., Fujita, H., Zheng, Y., Ai, W. (2021) "Multi-class financial distress prediction based on support vector machines integrated with the decomposition and fusion methods", *Information Sciences*, 559: 153-170.
- Sun, S. and Huang, R. (2010) "An adaptive k-nearest neighbor algorithm" en *Seventh International Conference on Fuzzy Systems and Knowledge Discovery* (1): 91-94. IEEE.
- Sun, J. and Li, H. (2012) "Financial distress prediction using support vector machines: Ensemble vs. individual", *Applied Soft Computing*, 12(8): 2254-2265.
- Suthaharan, S. and Suthaharan, S. (2016) "Support vector machine. Machine learning models and algorithms for big data classification: thinking with examples for effective learning": 207-235.
- Tay, F. E. and Cao, L. J. (2001a) "Application of support vector machines in financial time series forecasting", *Omega*, 29(4): 309-317.
- Tay, F. E. and Cao, L. J. (2001b) "Improved financial time series forecasting by combining support vector machines with self-organizing feature map", *Intelligent Data Analysis*, 5(4): 339-354.
- Tay, F. E. and Cao, L. J. (2002) "Modified support vector machines in financial time series forecasting", *Neurocomputing*, 48(1-4): 847-861.
- Thakur, M. and Kumar, D. (2018) "A hybrid financial trading support system using multi-category classifiers and random forest", *Applied Soft Computing*, 67: 337-349.
- Tibshirani, R. (1996) "Regression shrinkage and selection via the lasso", *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267-288.
- Ting, S. L., Ip, W. H., and Tsang, A. H. (2011) "Is Naive Bayes a good classifier for document classification?", *International Journal of Software Engineering and Its Applications*, 5(3): 37-46.
- Torky, M., Gad, I., and Hassanien, A. E. (2023) "Explainable AI Model for Recognizing Financial Crisis Roots Based on Pigeon Optimization and Gradient Boosting Model", *International Journal of Computational Intelligence Systems*, 16(1): 50.
- Trafalis, T. B., and Ince, H. (2000) "Support vector machine for regression and applications to financial forecasting", *Proceedings of the IEEE-INNS-ENNS*.
- Trustorff, J. H., Konrad, P. M., and Leker, J. (2011). Credit risk prediction using support vector machines, *Review of Quantitative Finance and Accounting*, 36: 565-581.
- Tuggener, L., Amirian, M., Rombach, K., Lörrwald, S., Varlet, A., Westermann, C., and Stadelmann, T. (2019) "Automated machine learning in practice: state of the art and recent results" en *6th Swiss Conference on Data Science (SDS)*: 31-36. IEEE.
- Uchiyama, Y., Kadoya, T., and Nakagawa, K. (2019) "Verification of lead-lag effect in financial markets by the adaptive elastic net regression", en *8th International Congress on Advanced Applied Informatics (IIAI-AAI)*: 693-696. IEEE.
- Van Gestel, T., Suykens, J. A., Baestaens, D. E., Lambrechts, A., Lanckriet, G., Vandaele, B., ... and Vandewalle, J. (2001) "Financial time series prediction using least squares support vector machines within the evidence framework", *IEEE Transactions on neural networks*, 12(4): 809-821.
- Wang, Z. (2022) "A Study on Early Warning of Financial Indicators of Listed Companies Based on Random Forest", *Discrete Dynamics in Nature and Society*.
- Wang, H., Li, G., and Tsai, C. L. (2007) "Regression coefficient and autoregressive order shrinkage and



- selection via the LASSO”, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69(1), 63-78.
- Wang, W. and Liu, F. (2020) “Research on Identification of Fraud in Financial Reporting Based on Adaptive Group Lasso-Logistic Regression” en *IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, 9: 1274-1279. IEEE.
- Wang, Z. and Tan, S. (2009) “Identifying idiosyncratic stock return indicators from large financial factor set via least angle regression”, *Expert Systems with Applications*, 36(4), 8350-8355.
- Waterman, D. A. (1985) *A guide to expert systems*. Addison-Wesley Longman Publishing Co., Inc.
- Witten, D. M., Shojaie, A., and Zhang, F. (2014) “The cluster elastic net for high-dimensional regression with unknown variable grouping”, *Technometrics*, 56 (1): 112-122.
- Wolpert, D. H. and Macready, W. G. (1997) “No free lunch theorems for optimization,” *IEEE Transactions on Evolutionary Computation*, 1 (1).
- Wu, L. and Yang, Y. (2014) “Nonnegative elastic net and application in index tracking”, *Applied Mathematics and Computation*, 227: 541-552.
- Yang, H., Chan, L., and King, I. (2002) “Support vector machine regression for volatile stock market prediction”, en *International Conference on Intelligent Data Engineering and Automated Learning*: 391-396. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Xu, H., Caramanis, C., and Mannor, S. (2008) “Robust regression and lasso”, *Advances in Neural Information Processing Systems*, 21.
- Xu, Q., Li, M., Jiang, C., and He, Y. (2019) “Interconnectedness and systemic risk network of Chinese financial institutions: A LASSO-CoVaR approach”, *Physica A: Statistical Mechanics and its Applications*, 534: 122173.
- Yang, Y., Li, Y., and Chen, Y. (2023) “How green finance affects automobile innovation? The perspective of lasso-based multi-mediation effect”, *Thermal Science*, (00), 34-34.
- Yao, Q., Wang, M., Chen, Y., Dai, W., Li, Y. F., Tu, W. W., ... and Yu, Y. (2018) “Taking human out of learning applications: A survey on automated machine learning”, *arXiv:1810.13306*.
- Zahariev, A., Zveryakov, M., Prodanov, S., Zaharieva, G., Angelov, P, Zarkova, S., and Petrova; M. (2020) “Debt management evaluation through support vector machines: on the example of Italy and Greece”, *Entrepreneurship and Sustainability Issues*, 7.3: 2382.
- Zemel, R. and Pitassi, T. (2000) “A gradient-based boosting algorithm for regression problems”, *Advances in neural information processing systems*, 13.
- Zhang, Z. (2016). Introduction to machine learning: k-nearest neighbours. *Annals of translational medicine*, 4(11).
- Zhang, Z., Lai, Z., Xu, Y., Shao, L., Wu, J., & Xie, G. S. (2017) “Discriminative elastic-net regularized linear regression”, *IEEE Transactions on Image Processing*, 26(3): 1466-1481.
- Zhang, H., Shi, Y., and Tong, J. (2021a) “Online supply chain financial risk assessment based on improved random forest”, *Journal of Data, Information and Management*, 3, 41-48.
- Zhang, H., Shi, Y., Yang, X., and Zhou, R. (2021b) “A firefly algorithm modified support vector machine for the credit risk assessment of supply chain finance”, *Research in International Business and Finance*, 58: 101482.



- Zhang, H. and Su, J. (2004) "Naive Bayesian classifiers for ranking" en *European conference on machine learning*: 501-512. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Zou, H. and Hastie, T. (2005) "Regularization and variable selection via the elastic net", *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2): 301-320.
- Zou, Z. B., Peng, H., and Luo, L. K. (2015) "The application of random forest in finance", *Applied Mechanics and Materials*, 740: 947-951.

## Papeles de la Fundación

### FUNDACIÓN INSTITUTO ESPAÑOL DE ANALISTAS

- N.º 1 Estudio para la mejora del Gobierno Corporativo, la Transparencia Informativa y los Conflictos de Interés.
- N.º 2 Guía de Principios de Buen Gobierno Corporativo, Transparencia Informativa y los Conflictos de Interés en las Sociedades Cotizadas.
- N.º 3 Estudio sobre los efectos de la aplicación de las Normas Contables del IASB a los sectores cotizados de la Bolsa Española.
- N.º 4 Buen Gobierno, Transparencia y Ética en el sector público.
- N.º 5 Estudio sobre la Reforma de los Mercados Financieros Europeos (2 vols.):
- Resumen del estudio en Español.
  - Resumen del estudio en Inglés.
- N.º 6 Guía de Buenas Prácticas para el desarrollo de la Junta General de accionistas de Sociedades Cotizadas.
- N.º 7 Observatorio de Gobierno Corporativo de las Grandes Sociedades Cotizadas en el Mercado de Valores español (Ibex-35), 2003.
- N.º 8 Índice de Excelencia Informativa de las Sociedades del Ibex-35 de la Bolsa Española IEI-FEF.
- N.º 9 Pensiones y Ahorro a largo plazo: un viejo problema en busca de nuevas soluciones.
- N.º 10 El ciclo del Capital Riesgo en Europa: su gestión y aportación de valor.

- N.º 11 España y las nuevas Perspectivas Financieras de la Unión Europea 2007-2013: Nuevos condicionantes, nuevos objetivos, nuevas estrategias.
- N.º 12 Diversidad de Género en los Consejos de Administración de las Sociedades Cotizadas y Cajas de Ahorros españolas.
- N.º 13 España y la nueva arquitectura económica y financiera internacional. El desafío del G-8.
- N.º 14 Observatorio de Gobierno Corporativo de las Grandes Sociedades Cotizadas en el Mercado de Valores español (Ibex-35), 2004.
- N.º 15 Índice de Excelencia Informativa de las Sociedades de la Bolsa Española IEI-FEF.
- N.º 16 La Responsabilidad Corporativa: una propuesta para un entorno empresarial más eficiente y socialmente comprometido.
- N.º 17 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2006).
- N.º 18 Las Cajas de Ahorros: Modelo de Negocio, Estructura de la Propiedad y su Gobierno Corporativo.
- N.º 19 Observatorio de Gobierno Corporativo y Transparencia Informativa de las Sociedades Cotizadas en el Mercado Continuo Español, 2006.
- N.º 20 Observatorio sobre el Gobierno de la Economía Internacional (2007).
- N.º 21 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2007).
- N.º 22 Observatorio de Gobierno Corporativo y Transparencia Informativa de las Sociedades Cotizadas en el Mercado Continuo Español, 2007.
- N.º 23 El Papel de las Instituciones Públicas en las Operaciones de Control Corporativo Nacionales y Transnacionales.
- N.º 24 Instrumentos Financieros para la Jubilación.
- N.º 25 Observatorio sobre el Gobierno de la Economía Internacional (2008).
- N.º 26 La Filantropía: tendencias y perspectivas. Homenaje a Rodrigo Uría Meruéndano.
- N.º 27 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2008).

- N.º 28 El sistema bancario español ante el nuevo entorno financiero.
- N.º 29 Observatorio de Gobierno Corporativo y Transparencia Informativa de las Sociedades Cotizadas en el Mercado Continuo Español, 2008.
- N.º 30 Situación actual y perspectivas de las Instituciones de Inversión Colectiva.
- N.º 31 Observatorio de la Economía Internacional (2009).
- N.º 32 Observatorio de Gobierno Corporativo (2009).
- N.º 33 La Ley Concursal y su aplicación.
- N.º 34 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2009).
- N.º 35 Estudio sobre el Sector Asegurador en España.
- N.º 36 Nuevos Enfoques para Viejos Riesgos Financieros.
- N.º 37 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2010).
- N.º 38 Estudio sobre el Sector Asegurador en España, 2010. Los pilares cualitativos de Solvencia II.
- N.º 39 El Ahorro Familiar en España.
- N.º 40 Observatorio de Gobierno Corporativo (2010).
- N.º 41 Los Nuevos Mercados Energéticos.
- N.º 42 Mecanismos de prevención y Gestión de futuras crisis bancarias.
- N.º 43 Las Economías emergentes y el reequilibrio global: retos y oportunidades para España.
- N.º 44 La crisis en Europa: ¿un problema de deuda soberana o una crisis del Euro?
- N.º 45 Pequeña y mediana empresa: impacto y retos de la crisis en su financiación.
- N.º 46 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2012).
- N.º 47 Desapalancamiento y crecimiento en España.
- N.º 48 El Sector Asegurador ante las transformaciones del Estado de bienestar.

- N.º 49 La arquitectura Institucional de la refundación del euro.
- N.º 50 España: crecer en la nueva economía global.
- N.º 51 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2013).
- N.º 52 Nuevos desafíos del sector financiero: recuperando la confianza y mejorando la cultura financiera.
- N.º 53 La regulación Financiera: ¿Solución o problema?
- N.º 54 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2015).
- N.º 55 Acabar con el paro, ¿queremos?, ¿podemos? Una propuesta analítica sobre la ocupación y el empleo en España.
- N.º 56 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2016).
- N.º 57 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2017).
- N.º 58 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2018).
- N.º 59 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2020).
- N.º 60 El rol de las finanzas en una economía sostenible.
- N.º 61 Financiación de intangibles.
- N.º 62 DeFi: ilusiones, realidades y desafíos.
- N.º 63 Observatorio sobre la Reforma de los Mercados Financieros Europeos (2022).
- N.º 64 El impacto de la inteligencia artificial en la industria financiera: promesas y amenazas.

## ENTIDADES PATRONO DE LA FUNDACIÓN DEL INSTITUTO ESPAÑOL DE ANALISTAS

BANCO SANTANDER, S.A.

BANCO SABADELL

FIDELITY WORLDWIDE INVESTMENT

LA CAIXA

BOLSAS Y MERCADOS ESPAÑOLES

INDRA

URIA & MENENDEZ

ACS

EY

FUNDACIÓN MUTUA MADRILEÑA

KPMG

JB CAPITAL

SOLUNION

DELOITTE

INSTITUTO CAJASOL

J&A GARRIGUES, S.L.

CECA

FINALBION

FINREG360

FLUIDRA

GRUPO INSUR

GVC GAESCO

ORGANON

WORKDAY



**INSTITUTO ESPAÑOL  
DE ANALISTAS** DESDE 1965  
FUNDACIÓN