

Ricardo Gimeno\*  
José Manuel Marqués\*\*

# TRADICIÓN E INTELIGENCIA ARTIFICIAL: OPORTUNIDADES Y RETOS DEL *MACHINE* *LEARNING* PARA LOS SERVICIOS FINANCIEROS

*Los métodos basados en la inteligencia artificial están transformando multitud de sectores al permitir automatizar tareas rutinarias e importantes mejoras en el análisis de la información. El sector financiero no es ajeno a esta tendencia y está tratando de aprovechar las oportunidades de estas técnicas al tiempo que debe ser consciente y actuar ante los riesgos y limitaciones que implican.*

## Tradition and artificial intelligence: opportunities and challenges of machine learning for financial services

*The methods based on artificial intelligence are changing numerous sectors, allowing for the automation of repetitive tasks and significant improvements in information analysis. The financial industry is also affected by these changes and is trying to take advantage of the opportunities that arise from them. Nevertheless, the industry also needs to be aware of their risks and limitations.*

**Palabras clave:** inteligencia artificial, aprendizaje automático.

**Keywords:** artificial intelligence, machine learning.

**JEL:** C18, O33.

### 1. Introducción

La inteligencia artificial y los modelos de *machine learning* no son algo nuevo, sus propuestas datan de más de medio siglo y, de hecho, el término «inteligencia artificial» se acuñó por primera vez durante la

Conferencia de Dartmouth en 1956, si bien, en gran parte estas propuestas permanecían en un ámbito más bien teórico y limitado por la escasa capacidad de computación existente. Sin embargo, en los últimos años, en un entorno de aumento en la capacidad de almacenamiento y procesamiento de la información, sus aplicaciones han dejado de estar vinculadas a casos anecdóticos —como los juegos de ajedrez o de preguntas para empezar a realizar tareas que los humanos desarrollamos de modo habitual, como el reconocimiento de voz y de imágenes<sup>1</sup> o la identificación de patrones

---

\* Jefe de la División de Inteligencia de Mercados. Banco de España.  
Contacto: [ricardo.gimeno@bde.es](mailto:ricardo.gimeno@bde.es)

\*\* Jefe de la División de Innovación Financiera. Banco de España.  
Contacto: [manuel-m@bde.es](mailto:manuel-m@bde.es)

Este artículo refleja la opinión de los autores y no se corresponde necesariamente con la opinión del Banco de España o del Eurosistema. Los autores agradecen la colaboración y comentarios de Andrés Alonso y José Manuel Carbó.

Versión de mayo de 2022.

<https://doi.org/10.32796/ice.2022.926.7403>

---

<sup>1</sup> La posibilidad de reconocer la voz no deja de ser una predicción sobre un objeto que se vincula con unos fonemas como variables explicativas o una imagen como asimilación de una combinación de píxeles con un determinado concepto.

de comportamiento—. Estos avances han hecho que el uso de la inteligencia artificial se extienda progresivamente a más sectores económicos y en más entornos. El caso de la industria financiera y la profesión económica no es una excepción y cada vez asistimos a numerosos ejemplos de aplicaciones innovadoras de estas técnicas en distintos casos de uso (Organisation for Economic Co-operation and Development [OECD], 2021). A lo largo de este artículo se pretende, en primer término, revisar brevemente los motivos que subyacen al uso de esta tecnología en el sector financiero. A continuación, se tratará de comparar las características de este tipo de modelización con otros enfoques econométricos usados tradicionalmente en la predicción de eventos financieros y, por último, se discutirán algunos de los retos que conlleva el uso de estas herramientas y la forma en la que pueden ser abordados.

## 2. El uso de la inteligencia artificial en el sector financiero: un nuevo paradigma

En finanzas, el principal uso de las técnicas de inteligencia artificial se enfoca básicamente en predecir una variable (generalmente binaria, como, por ejemplo, sería el caso de estimar un posible impago o la satisfacción de un cliente bancario) manejando la información disponible de un modo muy flexible y sin suponer ninguna relación subyacente entre las variables, es decir, desde un punto de partida completamente agnóstico sobre el proceso de generación de datos (Agrawal *et al.*, 2018). Esta flexibilidad, unida a la mejora en la precisión que puede alcanzarse, hace que las aplicaciones sean mayores que con las técnicas econométricas tradicionales, en las que para predecir necesitamos una relación teórica subyacente entre las variables. Así, podemos encontrarnos que el uso de estos algoritmos abarca no solo aquellos entornos donde venía siendo habitual emplear análisis de regresiones lineales, sino que también incluye el propio tratamiento de datos —agrupando las observaciones en distintos *clusters* o reduciendo la dimensión de

nuestros datos—, la toma de decisiones o incluso la inferencia causal, renunciando, en este caso (al menos en parte), a la flexibilidad sobre el proceso generador de los datos que se ha comentado anteriormente.

Adicionalmente, un elemento importante de estas técnicas es la posibilidad de ir «aprendiendo», es decir, la posibilidad de corregir los errores en la predicción e ir incorporando esa información de modo que el algoritmo cada vez es más preciso (es lo que se conoce como aprendizaje automático o *machine learning*, ML). Este proceso puede realizarse inclusive de un modo continuo, conforme se incorpora nueva información acerca de los errores cometidos, en lo que se conoce como aprendizaje por refuerzo. Por otro lado, estos algoritmos pueden realizar su trabajo predictivo de un modo supervisado, es decir, disponiendo el analista de un ejemplo previo con la verdadera relación entre las variables explicativas y el objeto a predecir, o de forma no supervisada, de modo que el analista desconoce el nombre o «etiqueta» de las variables y debe inferir el patrón de comportamiento en los datos sin dicha información previa.

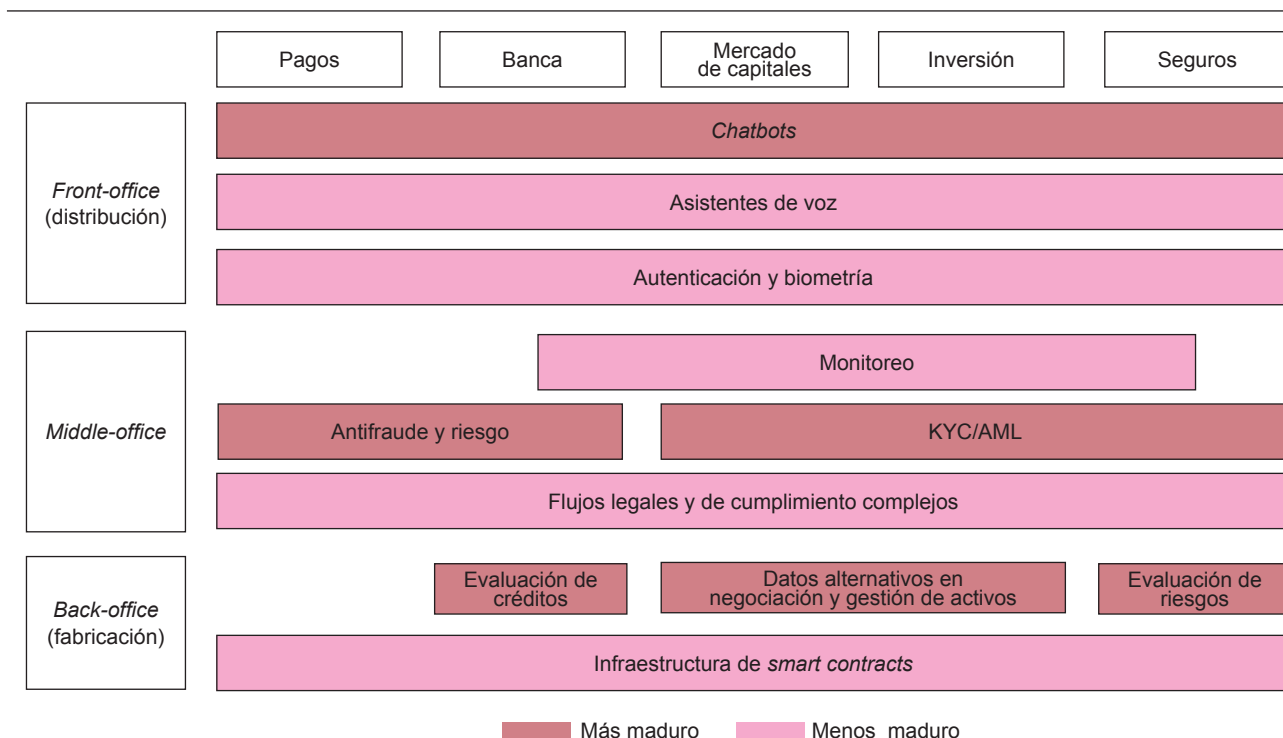
Las aplicaciones prácticas de estas técnicas son múltiples y conllevan, en muchos casos, ganancias importantes de eficiencia, al lograr realizar con menor coste actividades intensivas en trabajo humano, ya sea por su naturaleza repetitiva o su complejidad.

Dentro del sector financiero, las principales ganancias se centran en las posibilidades de automatizar procesos manuales de modo que no solo se reduce el coste, sino que, en muchos casos, mejora la satisfacción del cliente (al reducir el tiempo de respuesta) o se amplía la cobertura del servicio. Adicionalmente, también resulta relevante el aumento de la capacidad analítica al permitir procesar y sacar conclusiones de modo más rápido y preciso e incorporar fuentes de información que en el pasado no era posible analizar.

Tal y como se muestra en la Figura 1, las aplicaciones incluyen distintos segmentos del sector financiero y, además, comprenden toda la cadena de valor, encontrándose aplicaciones que van desde las tareas

FIGURA 1

## EJEMPLOS DE USO DE INTELIGENCIA ARTIFICIAL



NOTA: KYC y AML son los acrónimos de Know Your Customer (conoce a tus clientes) y Anti-Money Laundering (prevención de blanqueo de capitales).

FUENTE: Autonomous NEXT.

de procesamiento (*back-office*) hasta la interacción con el cliente y el mercado (*front-office*), pasando por las relacionadas con el control y monitoreo de la propia operativa financiera. En algunos casos, como en la parte de *back-office*, la integración con otras tecnologías, como la posibilidad de desplegar contratos inteligentes o *smart contracts*, podría ampliar significativamente el potencial de estas técnicas.

De un modo más concreto, y sin ánimo de ser exhaustivo, a continuación, se enumeran algunos ejemplos en los que ya se están desplegando aplicaciones que conllevan el uso de algoritmos basado en la inteligencia artificial (Fernández Bedoya, 2019):

— *Chatbot y asistentes virtuales*, que permiten no solo canalizar las peticiones de los clientes, sino

incluso incorporar labores de asesoramiento. La automatización permite incrementar la disponibilidad de atención sin restricciones horarias.

— *Personalización de productos y marketing*, de forma que la oferta se ajusta de modo más preciso al cliente al que va dirigido, integrando en ocasiones productos adicionales al financiero e incorporando información sobre cuestiones como la geolocalización del cliente en el proceso.

— *Procesos de onboarding* o captación de clientes usando, por ejemplo, datos biométricos de forma que se simplifica la identificación del cliente, mejorando los estándares de seguridad que pueden tener otros medios.

— *Detección de fraude y prevención de operaciones de blanqueo de capitales*, de modo que se

incorporan nuevas fuentes de información, se reducen los falsos positivos y se mejoran los sistemas de alerta temprana.

— *Evaluación de créditos o credit scoring*, mejorando tanto la propia concesión de crédito como el cálculo de provisiones o asignación de capital en función del nivel de riesgo del acreditado.

— *Cumplimiento regulatorio o Regtech*, de modo que se facilita tanto el cumplimiento con los requerimientos regulatorios como el seguimiento de los cambios regulatorios.

— *Otros ámbitos*, como el propio uso entre los supervisores (*Suptech*) o la valoración de las reclamaciones en el sector de los seguros, la detección de errores o prácticas fraudulentas en las operaciones de *trading*, la gestión de riesgos, los análisis de sentimiento económico o cálculo de indicadores macroeconómicos en tiempo real son campos en los que también se encuentran aplicaciones de este tipo de herramientas (Bank of England [BoE], 2019; Financial Stability Board [FSB], 2017; Institute of International Finance [IIF], 2019; Königstorfer y Thalmann, 2020; Goodell *et al.*, 2021).

### 3. La inteligencia artificial frente a la tradición estadística en el sector financiero

Evidentemente, la mayoría de estas aplicaciones no han nacido con el ML pero, frente a los modelos tradicionales econométricos que podrían usarse, por ejemplo, para la toma de decisiones de si dar un crédito o no o para realizar la predicción de una variable macroeconómica, las técnicas de ML suponen un importante complemento que, en muchos casos, pueden mejorar el uso que se hacía con esas herramientas estadísticas (Mullainathan y Spiess, 2017; Athey, 2018; Athey e Imbens, 2019). De hecho, en ciertos campos, como el reconocimiento facial, o los *chatbots*, no existen alternativas viables al ML dentro de la econometría tradicional. Esto es así porque la flexibilidad en la modelización y el poder predictivo que aporta el ML es muy superior a lo

que pueden proporcionar los modelos necesariamente más simples a los que se ve constreñida la econometría tradicional.

El ML está especialmente preparado para tratar fenómenos o problemas complejos en los que se pueden producir situaciones tales como datos no estándar, alta dimensionalidad de nuestra base de datos, múltiples interacciones (incluso no lineales) entre estas variables, o correlaciones en las colas de las distribuciones. La econometría tradicional ha desarrollado soluciones puntuales para muchos de estos problemas: modelos factoriales para la alta dimensionalidad, modelos de diferencias en diferencias o de regresiones en discontinuidad que están diseñados específicamente para interacciones, o modelos de regresión cuantílica para las colas de las distribuciones, entre otras muchas. Sin embargo, estas soluciones son particulares para situaciones concretas, y difícilmente combinables si se dan dos o tres de estos problemas a la vez. Sin embargo, la versatilidad del ML permite explorar cualquier combinación de estos fenómenos a la que nos podamos enfrentar, lo que lleva a una identificación de las especificaciones óptimas de forma mucho más agnóstica, sin los supuestos *a priori* que pueden sesgar la ciencia económica tradicional (Breiman, 2001; Maurya *et al.*, 2020). Por decirlo de otra forma, con el ML se prima la aproximación lo más fidedigna posible a una realidad compleja y la mayor capacidad de predicción posible, frente a la aproximación común que parte de un modelo (por definición y etimológicamente, una simplificación de la realidad) predeterminado. Un caso paradigmático de esto son los populares (en economía) *Dynamic Stochastic General Equilibrium Models*, que tienen un fuerte componente paramétrico/estructural teórico en el que muchas veces se acaba más adaptando la realidad al modelo, que al revés (véase, entre otros, Blanchard, 2017; Romer, 2016).

Adicionalmente, técnicas de ML como los algoritmos evolutivos permiten encontrar soluciones dinámicas, adaptándose a una realidad cambiante (muchas veces de modo automático a través de lo que se viene

a denominar como técnicas de aprendizaje) sin la restricción que puede implicar querer «ajustarse o cumplir» con un modelo predeterminado que, en muchas ocasiones, acumula errores de predicción importantes (si bien, podría argumentarse que aproximaciones bayesianas también permitirían adaptaciones cambiantes de modelos tradicionales a ciertos cambios de la realidad).

En este ámbito de predicción/clasificación/decisión, las técnicas de ML deberían considerarse como la técnica de referencia a batir, ya que suponen la aproximación más amplia o carente de restricciones. Frente a esta solución, debería analizarse la capacidad predictiva de cualquier otra alternativa de modelización u optimización más tradicional (que a nivel estadístico, generalmente, puede reducirse a ser considerado como un caso particular, y restringido, del resultado del ML) que aportase beneficios adicionales a su capacidad de predicción/clasificación/explicación y que trataremos más adelante en este apartado. Solo de esa manera puede valorarse el coste en términos de la pérdida de precisión estadística en la que incurriríamos por renunciar al ML frente a técnicas econométricas tradicionales. En esencia y en términos matemáticos, es equivalente a la comparación entre un óptimo restringido (econometría tradicional) frente a un óptimo libre (ML). El primero solo debería ser preferible si la simplicidad que aporta no va acompañada de una pérdida significativa de la conexión con la realidad (Hoepner *et al.*, 2021).

No obstante, el ML, tal y como se ha comentado, no está exento de desafíos. La aplicación del ML, especialmente en casos como el *credit scoring* (Alonso y Carbó, 2022), o en otros casos de uso financiero (European Banking Authority [EBA], 2020; Dupont *et al.*, 2020; Deutsch BundesBank Eurosystem [DBE] y Bundesanstalt für Finanzdienstleistungsaufsicht [BaFin], 2021), debe tenerlos en cuenta para realizar un uso responsable tras un análisis coste/riesgo y beneficio. Esto también debería realizarse si se opta por la econometría tradicional, en la que raramente se ve un análisis de los costes derivados de la (sobre)

simplificación inherente al modelo especificado (con la posible excepción de los modelos de ecuaciones estructurales en los que sí es habitual contrastar el modelo especificado frente a la alternativa del sistema completamente determinado, aún más restrictivo que hacerlo frente a una alternativa ML).

Precisamente, la gran potencia que tienen las técnicas de ML para encontrar la solución a un problema lleva a que corramos el riesgo de que esa solución concreta no sea generalizable (*overfitting*). Así, podríamos acabar con una herramienta perfecta para tomar la decisión de si deberíamos haber dado el crédito a Juan (que formó parte del conjunto de datos usado para entrenar/enseñar al algoritmo de ML), pero cuyas conclusiones no me valgan para decidir si dárselo a Roberto (que no formó parte de ese conjunto de datos de entrenamiento). Esto conlleva la necesidad de diseñar correctamente los procesos de aprendizaje, dividiendo el conjunto de datos en una parte usada para el entrenamiento, y otra para la evaluación de la calidad del ML, y la elección del modelo de ML que mejor ajuste este segundo. No obstante, este proceso, conocido como *cross-validation*, tampoco está exento de problemas, ya que la cantidad prácticamente infinita de modelos ML que se pueden llegar a obtener, puede llevar a un abuso del número de algoritmos entrenados, de tal manera que acabemos sobreajustando la submuestra usada para la evaluación, en vez de sobreajustar la submuestra de entrenamiento —Breiman (2001) y su explicación sobre la «multiplicidad de buenos modelos»—. De nuevo pueden existir soluciones como dividir de nuevo la submuestra de evaluación en otras dos muestras, usando la primera para elegir el mejor ML entrenado y la segunda para hacer la evaluación de la calidad del ML, y, de este modo, continuar en este proceso iterativo entre número de ML entrenados y número de grupos de (evaluación/selección). Una alternativa a esta estrategia de «carrera armamentística» para la selección del mejor modelo consiste en buscar indicadores que combinen la capacidad de ajuste del algoritmo de ML, pero penalizando la complejidad

del mismo, de forma análoga al papel jugado en la econometría tradicional por el coeficiente de determinación corregido, el criterio de información de Akaike o el bayesiano.

El segundo problema clásico del ML es el conocido como el de la «caja negra», ya que, si además de predecir queremos extraer conclusiones del ejercicio de ML, este debería ser interpretable. Esto es especialmente relevante en finanzas, por ejemplo, si vamos a decidir si dar un crédito a una persona en función de un algoritmo de ML, es necesario que esta persona sepa qué debería hacer para mejorar su evaluación de riesgo de impago (ver, más adelante, el apartado 4). De igual forma, si no solo queremos predecir una variable económica, sino que, además, nos interesa conocer las variables que influyen en la misma para desarrollar distinto tipo de políticas económicas, las técnicas de ML de modo aislado pueden no darnos esa solución. La enorme flexibilidad del ML hace que también sea más complicado evaluar los efectos de una variable explicativa sobre el objeto a predecir. Así, la presencia de interacciones entre variables lleva a que los efectos de una determinada variable puedan llegar a ser diametralmente opuestos dependiendo de qué valores se observen en las otras variables. De igual forma, la posibilidad de existir correlaciones en las colas puede llevar a distintos efectos de una misma variable, simplemente dependiendo del valor de partida de esa misma variable. No obstante, recientemente se han propuesto algunas vías para solucionarlo dentro del campo conocido como «inteligencia artificial explicable» (Alonso y Carbó, 2022; Ribeiro *et al.*, 2016; Lundberg y Lee, 2017; Molnar, 2020).

El tercer problema está ligado al anterior, y es que precisamente el carácter de caja negra puede llevar a que se magnifiquen sesgos no deseados (debido a una mayor dificultad de controlar plenamente los patrones implícitos en los datos), que acabe perpetuando problemas previos. Si, por ejemplo, a pesar de tratar de restringir el entrenamiento de nuestro algoritmo no mostrando variables de carácter personal como sexo o raza,

usamos una gran cantidad de datos por parte del ML, es probable que la máquina acabe aprendiendo a través de otras variables esas clases que queríamos proteger como el género o la raza, por ejemplo, a través del código postal o los estudios. En última instancia, una discriminación estadística (Phelps, 1972), como la que a veces se da entre géneros en la obtención de créditos (De Andrés *et al.*, 2021), derivará en una mayor discriminación entre individuos en mi resultado (Fuster *et al.*, 2022). Como ejemplo de este fenómeno, la existencia del *gender gap* en sueldos de partida puede llevar a que el algoritmo de ML pueda erróneamente identificar el género como la variable relevante para el *credit scoring*, en vez del sueldo. El carácter de caja negra del ML hace muy difícil detectar este tipo de discriminaciones que la econometría tradicional detecta mucho más fácilmente. Lógicamente, la importancia del sesgo depende del contexto, no es lo mismo que el sesgo se produzca a la hora de conceder un crédito, con efectos económicos graves, que en la complejidad sintáctica con la que nos responde un asistente virtual (Bono *et al.*, 2021; Dobbie *et al.*, 2021; Bartlett *et al.*, 2022).

Un desafío adicional de los modelos de ML vinculado también con el carácter de caja negra es el relacionado con la posibilidad de ser *hackeado* (Goldwasser *et al.*, 2022). Estos autores señalan cómo se podrían provocar comportamientos erróneos influyendo en los procesos de entrenamiento. Para ello, solo es necesario añadir en el conjunto de datos de entrenamiento un conjunto de datos ficticios para que, por ejemplo, el ML decidiera dar un crédito siempre que la solicitud fuera por una cuantía que acabara en 23. La entidad de crédito que habría solicitado el ML para hacer, por ejemplo, el *credit scoring* no tendría forma de detectarlo, y los autores podrían obtener un crédito de ella, siempre que pidan uno de la cantidad «adecuada».

Por todo lo anterior, es fundamental que los profesionales que se dediquen al ML conozcan también las peculiaridades de la muestra de datos que manejan, la forma en la que se ha obtenido (para saber si hay sesgos en su obtención), y sean conscientes de los



riesgos que existen en el caso de uso en el que se está empleando estas técnicas, diseñando, en consecuencia, un sistema de gobernanza de los modelos y los datos que los mitigue. Como señala IIF (2020), no hay una regla que satisfaga todas las condiciones de uso, que pueden tener muy distintos objetivos, e incluso condiciones distintas en cada país, dependiendo de diferencias culturales o legislaciones vigentes.

En este sentido, a la hora de decidir el mejor método analítico es importante desarrollar una agenda completa que realice un completo análisis coste/beneficio de los métodos empleados, no dando por supuesto que las herramientas usadas en el pasado son las más adecuadas por estar más familiarizadas con ellas, pero tampoco abrazando entusiásticamente las propuestas de alternativas más complejas sin analizar previamente los riesgos nuevos a los que se enfrenta la modelización del ML. La capacitación de los profesionales que usan ML debe ser una combinación de habilidades para el uso de herramientas complejas de ML, pero también de conocimientos sólidos de la fundamentación teórica económica-financiera del problema que se va a tratar (IIF, 2020).

Adicionalmente, hay que tener en cuenta que, de modo paralelo al aumento en el uso de estas técnicas, también ha comenzado a aparecer un intenso trabajo académico y aplicado que trata de mejorar aspectos como la interpretabilidad o la causalidad o integrar las ventajas de los métodos econométricos tradicionales con los del ML, y que requiere una actualización constante de cualquier investigador de datos que quiera dedicarse al ML.

#### 4. Retos para las autoridades en el uso de las técnicas de *machine learning/artificial intelligence*

A lo largo del apartado anterior, se han mencionado algunas de las cuestiones a las que el analista debe hacer frente a la hora de utilizar las técnicas de ML. Adicionalmente, estas técnicas plantean una serie de cuestiones a las autoridades financieras relacionadas tanto con la forma en la que pueden mitigar riesgos

conocidos como con la aparición de nuevas cuestiones que se plantean en relación con valores fundamentales de la sociedad que deben ser preservados. En este sentido, el analista debe ser consciente de la normativa que afecta a estos procedimientos, mucha de la cual, como se verá en este apartado, está en proceso de elaboración.

El uso de la inteligencia artificial, desde un punto de vista general, y no centrado exclusivamente en el sector financiero, plantea un importante dilema para las autoridades que, por un lado, tratan de fomentar su uso y trasladar a los usuarios las ganancias en eficiencia derivadas de la innovación tecnológica y, por otro lado, deben preservar el cumplimiento de valores fundamentales como la no manipulación del usuario, la no discriminación entre individuos por causas no deseables o la protección de la privacidad recogidas en una normativa desarrollada con carácter previo al uso de estos algoritmos. En el caso de Europa, este dilema se está abordando, además de con iniciativas generales que afectan al uso de la inteligencia como la estrategia de computación en la nube o la de datos, con dos iniciativas de carácter más específico: un plan coordinado para potenciar el uso de la inteligencia artificial, y una propuesta de Reglamento para regular y mitigar los riesgos derivados del uso de la inteligencia artificial (European Commission [EC], 2021).

En relación con la primera línea de actuación, se trata de desarrollar un plan ambicioso que promueva el uso de la inteligencia artificial desde distintos frentes:

- Estableciendo las condiciones iniciales en coordinación de políticas entre los países; recopilación y accesibilidad a los datos; y desarrollo de la infraestructura computacional.

- Medidas horizontales que permitan movilizar el talento investigador, la experimentación, la financiación y el crecimiento de las aplicaciones de inteligencia artificial exitosas.

- Fomentar los conocimientos y las políticas para garantizar la confianza en los sistemas de inteligencia artificial.

- Construir y fomentar el uso de la inteligencia artificial en sectores estratégicos de alto impacto.

En cuanto a la segunda línea de actuación, el Consejo y el Parlamento Europeo propusieron un Reglamento que regula el uso de la inteligencia artificial en lo que se ha denominado como *Artificial Intelligence Act*. Esta propuesta, que aún se encuentra en fase de aprobación, constituye uno de los primeros intentos a nivel internacional en tratar de definir un marco ordenado para permitir el uso generalizado de estas técnicas, preservando los valores fundamentales de la sociedad que, en el caso europeo, se recogen en la Carta de los Derechos Fundamentales de la Unión Europea. Para ello, se parte de una definición relativamente amplia de inteligencia artificial, que incluye tanto los métodos de aprendizaje automático o ML, estrategias basadas en la lógica y el conocimiento, o herramientas de naturaleza más estadística. Posteriormente, se identifican aquellas categorías de uso de la inteligencia artificial en las que, dado el nivel de riesgo, se considera que requieren especial atención:

— Casos en los que se considera prohibido el uso de la inteligencia artificial: aquellos que buscan manipular la conducta, opinión o decisión de los ciudadanos, explotar la información de una persona para abusar de sus debilidades, o sistemas de clasificación que conlleven a un trato perjudicial o desfavorable a colectivos o personas de un modo desproporcionado o no relacionado con el contexto donde se han originado los datos. Asimismo, se prohíbe el uso de las técnicas de identificación biométrica en tiempo real en espacios de acceso público, excepto en unos casos concretos previstos en la ley.

— Casos de alto riesgo en el uso de la inteligencia artificial: se definen una serie de ámbitos en los que, por su relevancia e impacto, se les considera de alto riesgo. Entre estos ámbitos se incluyen distintas categorías como la identificación biométrica para captar nuevos clientes, las aplicaciones en la selección o contratación de personas físicas, los asuntos relacionados con la aplicación de la ley, o el acceso a servicios públicos o privados esenciales entre los que se encuentra, por ejemplo, la concesión de crédito.

En estos casos de uso de alto riesgo, se establecen una serie de requisitos relativos a los datos usados y a la

gobernanza de los mismos, la transparencia que hay que proporcionar a los usuarios, la necesidad de que exista supervisión humana (incluida la posibilidad de poder interpretar los resultados) o el mandato de desarrollar mecanismos de gestión del riesgo que permitan identificar de un modo continuo los riesgos que puedan existir y la forma de mitigarlos.

Adicionalmente, en función del contexto en el que se desarrollan los algoritmos, es preciso cumplir con la normativa de carácter más general que resulte de aplicación. Así, por ejemplo, en el caso de emplear datos de carácter personal, es necesario respetar las normas del Reglamento General de Protección de Datos (GDPR, por sus siglas en inglés) o, en el caso de técnicas aplicadas a la identificación en servicios financieros, cumplir con los requisitos de las normas de prevención de blanqueo de capitales.

Un caso de especial interés lo constituyen la aplicación de las técnicas de inteligencia artificial de cara a determinar la solvencia de un acreditado y poderle conceder un crédito. En esta situación, además de los requerimientos como sistema de alto riesgo que podrían derivarse del reglamento de inteligencia artificial una vez que se apruebe, deben de cumplirse las exigencias derivadas de los supervisores financieros en relación con aspectos como la interpretabilidad, la intervención humana en el proceso, las implicaciones derivadas de la externalización de proveedores, la gestión e identificación de riesgos de modelo, los procesos de validación de sistemas de *rating*, o las vulnerabilidades cibernéticas y de resiliencia (Alonso y Carbó, 2020, 2022; BaFin, 2021; Dupont *et al.*, 2020; EBA, 2021).

En definitiva, tal y como se ha esbozado en este apartado, el uso de estos algoritmos conlleva una serie de requisitos normativos, muchos de los cuales se encuentran todavía en proceso de discusión<sup>2</sup>, para poder conjugar los avances que ofrecen estas técnicas sin perder

<sup>2</sup> Véase, por ejemplo, el caso de la consulta de la EBA a la industria sobre el uso de ML en modelos internos de determinación de capital regulatorio, finalizada en febrero de 2022.



de vista los valores fundamentales que como sociedad hay que preservar. Esto conlleva un escenario complejo, ya que, al tiempo que se avanza en las mejoras predictivas y aplicaciones que se derivan de estos modelos, las autoridades deben, no solo desarrollar y adaptar la normativa, sino, también, dotarla del nivel de coordinación y flexibilidad necesarios para que no dificulten innecesariamente las ganancias que pueden alcanzarse gracias a este proceso de innovación tecnológica y financiera.

## 5. Conclusiones

Las nuevas posibilidades tecnológicas, unidas a la capacidad para recopilar y procesar fuentes de información, hacen que las técnicas de inteligencia artificial y aprendizaje automático (*machine learning*) puedan aplicarse en múltiples procesos tanto básicos como analíticos en prácticamente todos los sectores de la actividad económica. El sector financiero no es ajeno a este fenómeno, sin duda alguna disruptivo y que puede ayudarle a obtener mejoras significativas de eficiencia y precisión en procesos como la concesión de crédito y ahorros importantes en tareas repetitivas. No obstante, aunque estas técnicas mejoran en términos predictivos las capacidades de los modelos econométricos tradicionales, también tienen importantes desafíos como la dificultad para explicar los resultados, la estabilidad de los mismos o la posibilidad de perpetuar sesgos discriminatorios. Estos desafíos deben ser tenidos en cuenta por parte de los profesionales y empresas que usen estos modelos, si bien, algunos de ellos también afectan a otros modelos tradicionales y deben formar parte de una estrategia integral de las empresas a la hora de gestionar y explotar la información. Por otro lado, la importancia de muchos de estos desafíos depende, en gran medida, del contexto en el que se estén empleando los procesos de AI&ML, en este sentido, las autoridades públicas están planteando marcos normativos que realicen una serie de requerimientos en función del nivel de riesgo que implica la automatización y el análisis avanzado de los datos para los derechos fundamentales de los ciudadanos. Las entidades financieras, al igual que

otros sectores, no pueden dejar pasar las oportunidades que ofrecen estas tecnologías, pero deben anticiparse a las implicaciones que pueden tener para esos valores fundamentales y desarrollar las herramientas de gobernanza adecuadas, en especial en aquellas actividades como la concesión de crédito que son consideradas como esenciales para la actividad económica.

## Referencias bibliográficas

- Agrawal, A., Gans, J., & GoldFarb, A. (2018). *Prediction Machines: The Simple Economist of Artificial Intelligence*. Harvard Business Review Press.
- Alonso, A., & Carbó, J. M. (2020). *Machine learning in credit risk: measuring the dilemma between prediction and supervisory cost* (Documento de trabajo del Banco de España n.º 2032).
- Alonso, A., & Carbó, J. M. (2022). Accuracy and stability of explanations of machine learning models for credit default prediction. *Journal of Financial Innovation*. Forthcoming.
- Athey, S. (2018). The impact of machine learning on economics. *The economics of artificial intelligence: An agenda*, 507-547.
- Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685-725.
- BaFin, Bundesanstalt für Finanzdienstleistungsaufsicht. (2021). *Big data and AI: Principles for the use of algorithms in decision-making processes*.
- Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, 143(1), 30-56.
- Blanchard, O. (2017). Do DSGE models have a future? In R. S. Gürkaynak, & C. Tille (Eds.), *DSGE Models in the Conduct of Policy: Use as intended* (pp. 93-100). CEPR Press.
- BoE, Bank of England. (2019). *Machine learning in UK financial services*.
- Bono, T., Croxson, K., & Giles, A. (2021). Algorithmic fairness in credit scoring. *Oxford Review of Economic Policy*, 37(3), 585-617.
- Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, 16(3), 199-215.
- DBE/BaFin, Deutsch BundesBank Eurosystem/Bundesanstalt für Finanzdienstleistungsaufsicht. (2021). *Machine learning in risk models – Characteristics and supervisory priorities*.
- De Andrés, P., Gimeno, R., & Mateos de Cabo, R. (2021). The gender gap in bank credit access. *Journal of Corporate Finance*, 71, 101782.
- Dobbie, W., Liberman, A., Paravisini, D., & Pathania, V. (2021). Measuring bias in consumer lending. *The Review of Economic Studies*, 88(6), 2799-2832.

- Dupont, L., Fliche, O., & Yang, S. (2020). *Governance of Artificial Intelligence in Finance*. Banque De France.
- EBA, European Banking Authority. (2020). *Report on Big Data and Advanced Analytics*.
- EBA, European Banking Authority. (2021). *Discussion Paper on Machine Learning for IRB Models*.
- EC, European Commission. (2021). *Fostering a European approach to Artificial Intelligence*. COM(2021) 205 final.
- Fernández Bedoya, A. (2019). Inteligencia Artificial en los servicios financieros. *Boletín Económico del Banco de España* n.º 2/2019. Artículos analíticos.
- FSB, Financial Stability Board. (2017). *Artificial intelligence and machine learning in financial services. Market developments and financial stability implications*.
- Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance*, 77(1), 5-47.
- Goldwasser, S., Kim, M. P., Vaikuntanathan, V., & Zamir, O. (2022). Planting Undetectable Backdoors in Machine Learning Models. *arXiv preprint arXiv:2204.06974*.
- Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance*, 32, 100577.
- Hoepner, A. G. F., McMillan, D., Vivian, A., & Simen, C. W. (2021). Significance, relevance and explainability in the machine learning age: an econometrics and financial data science perspective. *The European Journal of Finance*, 27(1-2), 1-7.
- IIF, Institute of International Finance. (2019). *Machine Learning in Credit Risk* (2<sup>nd</sup> edition).
- IIF, Institute of International Finance. (2020). *Machine Learning governance*.
- Königstorfer, F., & Thalmann, S. (2020). Applications of Artificial Intelligence in commercial banks – A research agenda for behavioral finance. *Journal of Behavioral and Experimental Finance*, 27, 100352.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In S. Thrun, L. K. Saul, & B. Schölkopf (Eds.), *Advances in neural information processing systems* (pp. 4765-4774).
- Maurya, N., Kumar, N., & Maurya, V. K. (2020). A Review on Machine Learning (Feature Selection, Classification and Clustering) Approaches of Big Data Mining in Different Area of Research. *Journal of Critical Reviews*, 7(19), 2610-2626.
- Molnar, C. (2020). *Interpretable machine learning. A guide for making black box models interpretable*. Lulu.com.
- Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87-106.
- OECD, Organisation for Economic Co-operation and Development. (2021). *Artificial Intelligence, Machine Learning and Big Data in Finance. Opportunities, Challenges and Implications for Policy Makers*.
- Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *American Economic Review*, 62(4), 659-661.
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22<sup>nd</sup> ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144).
- Romer, P. (2016). The Trouble with Macroeconomics. *Commons Memorial Lecture of the Omicron Delta Epsilon Society*.