

Recuadro 2.3

INTELIGENCIA ARTIFICIAL: UN CAMPO DE OPORTUNIDADES CRECIENTES Y NUEVOS RETOS SUPERVISORES

En los últimos años, la popularidad de la inteligencia artificial (IA) ha experimentado un auge vertiginoso. Esto se debe tanto a los continuos avances en las capacidades de procesamiento informático como a la irrupción de los grandes modelos fundacionales¹ de última generación, capaces de procesar ingentes cantidades de información. Los nuevos modelos de IA permiten abordar tareas normalmente asociadas a la inteligencia humana, tales como traducir idiomas, contestar preguntas o incluso efectuar razonamientos.

Estas tecnologías son utilizadas en las entidades financieras, por ejemplo, para la automatización y optimización de tareas administrativas, para el análisis de grandes volúmenes de datos o en aplicaciones utilizadas en la relación con el cliente. De este modo, vemos asistentes conversacionales (*chatbots*) impulsados por IA que pueden gestionar consultas de clientes las 24 horas del día, el uso de algoritmos de clasificación avanzados para detectar fraudes financieros o la aplicación de IA para evaluar el riesgo crediticio de los clientes, entre otras aplicaciones.

Desde un punto de vista general, dichas tecnologías de propósito general tienen el potencial de aumentar la productividad en el conjunto de la economía. Ahora bien, su uso también entraña riesgos, algunos de ellos especialmente relevantes en el ámbito de la supervisión de conducta. Aquí, los modelos, a partir de los sesgos presentes en los datos utilizados para su desarrollo, pueden causar discriminaciones en la concesión de crédito que vulneren los derechos fundamentales de la clientela. Conviene recordar la dificultad que entraña la explicabilidad de estos modelos, pues, al usar miles de millones de parámetros, se hace complejo entender los factores que determinan los resultados obtenidos, y de ahí que se conozcan coloquialmente como «cajas negras».

Para hacer frente a esta situación, el Reglamento 2024/1689, de 13 de junio de 2024, sobre IA, regula en la Unión Europea (UE) la utilización de sistemas de IA con un enfoque basado en los niveles de riesgo y en la protección de los derechos fundamentales. Por ello, identifica como sistemas de IA de alto riesgo aquellos que se utilicen en la concesión de crédito a personas físicas. En el caso de estos sistemas de alto riesgo, los desarrolladores y usuarios deberán realizar evaluaciones exhaustivas de los riesgos asociados a sus modelos, antes de su puesta en producción. En particular, el reglamento europeo indica

que deben cumplir con requisitos de gestión de riesgos, gobierno de los datos, poseer la adecuada documentación técnica y realizar un análisis de sus resultados, así como tener niveles suficientes de precisión, robustez y ciberseguridad. Se enfatiza especialmente que los sistemas de IA deberán garantizar que son transparentes y explicables. Esto último implica proporcionar información clara sobre cómo funcionan los modelos y cómo se toman las decisiones basadas en ellos. De igual forma, deberán aplicar mecanismos de supervisión que garanticen la revisión (y, si es necesario, la corrección por parte de personas) de las decisiones automatizadas.

Una vez que estos sistemas de IA de alto riesgo entran en producción, el reglamento identifica que deberán ser supervisados por las autoridades de vigilancia del mercado. Estas autoridades se designan a escala nacional y, en el caso de los modelos de crédito, será la autoridad supervisora bancaria la que deberá reportar además los incidentes al Banco Central Europeo (BCE).

Ante la previsible designación del Banco de España como autoridad de vigilancia de mercado en el uso de la IA en la concesión de crédito a personas físicas, el Banco está ya trabajando para poder asumir dichas competencias. Esta labor de vigilancia conllevará, entre otras, las funciones de supervisión y control del cumplimiento del reglamento por las entidades supervisadas; la evaluación de la conformidad con la normativa de estos modelos; la colaboración internacional con otras autoridades de vigilancia de mercado para el intercambio de información y la aplicación coherente del reglamento en toda la UE; el seguimiento de incidentes, y la promoción de la innovación.

En el desarrollo de estas funciones, el Banco de España deberá coordinarse, entre otros organismos, con la Agencia Española de Supervisión de la Inteligencia Artificial. Así, por ejemplo, la Autoridad Bancaria Europea y el BCE, de forma conjunta con las autoridades nacionales competentes del Mecanismo Único de Supervisión, han estado trabajando en el uso de IA en los modelos de cálculo de capital regulatorio, fijando expectativas regulatorias y técnicas de supervisión. Estas expectativas aparecerán reflejadas en la actualización de la *Guía de Modelos Internos* del BCE, cuya publicación está prevista en 2025.

En definitiva, el auge de la IA y las oportunidades y retos que conlleva han llevado al Banco de España a crear un

¹ Los modelos fundacionales son modelos entrenados sobre grandes cantidades de datos y que pueden ser posteriormente especializados para realizar tareas concretas a partir del reentrenamiento con datos más específicos.

Recuadro 2.3

INTELIGENCIA ARTIFICIAL: UN CAMPO DE OPORTUNIDADES CRECIENTES Y NUEVOS RETOS SUPERVISORES (cont.)

grupo de trabajo transversal en el que, junto con el despliegue de la función de autoridad de vigilancia de mercado, se están incluyendo diversas líneas de trabajo para asegurar los recursos humanos y computacionales

necesarios, el despliegue de casos de uso en las diferentes funciones que realiza el Banco o el desarrollo de un modelo de gobierno interno que garantice un uso ético y responsable de la IA.