



Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Trabajo de Fin de Grado

Grado en Finanzas Banca y Seguros Oportunidades de inversión para el pequeño inversor español en albor de la inteligencia artificial

Presentado por:

Pablo Lozano Álvarez

Tutelado por

Gabriel de la Fuente Herrero

Valladolid, 29 de Junio de 2023

RESUMEN

La Inteligencia Artificial (IA) puede ayudar a pequeños inversores en la tarea de canalizar el ahorro y optimizar la relación riesgo-rentabilidad. En este trabajo se analizan los principios básicos de la Inteligencia Artificial a lo largo de su historia y su aplicación a las finanzas. También se apuntan posibles aplicaciones futuras sugeridas por la literatura más actual. Para finalizar, se analiza la situación actual de la inversión financiera de las familias españolas y los usos actuales de la inteligencia artificial a través los robo-advisor y los fondos de inversión gestionados por esta tecnología.

Clasificación JEL: G11, G24, C45

Palabras clave: Fondos de inversión, rentabilidad, inteligencia artificial, robo-advisor.

ABSTRACT

Artificial Intelligence (AI) can help small investors in the task of channelling savings and optimising the risk-return trade-off. This paper analyses the basic principles of Artificial Intelligence throughout its history and its application to finance. It also points out possible future applications suggested by the most current literature. Finally, we analyse the current situation of financial investment by Spanish households and the current uses of artificial intelligence through robo-advisors and investment funds managed by this technology.

Classification: G11, G24, C45

Key words: Investment fund, profitability, artificial intelligence, robo-advisor.

ÍNDICE DE CONTENIDOS

1. INTRODUCCIÓN	4
2. INTELIGENCIA ARTIFICIAL	5
2.1 INTRODUCCIÓN E HISTORIA	5
2.2 MODELOS.....	7
2.2.1 <i>Redes Neuronales Artificiales</i>	<i>7</i>
2.2.2 <i>Árbol de decisiones</i>	<i>8</i>
2.2.3 <i>Algoritmos genéticos</i>	<i>9</i>
2.2.4 <i>Procesamiento del lenguaje natural.....</i>	<i>10</i>
3. APLICACIONES DE LA INTELIGENCIA ARTIFICIAL EN FINANZAS....	12
3.1 OPTIMIZACIÓN DE CARTERAS	12
3.2 ROBO-ADVISOR	17
3.3 ANÁLISIS DE DATOS.....	21
3.4 PREDICCIÓN BURSÁTIL DE QUIEBRAS Y RIESGO DE CRÉDITO	24
4. INVERSIÓN FINANCIERA DE LAS ECONOMÍAS DOMÉSTICAS E INTELIGENCIA ARTIFICIAL.....	28
4.1 DATOS GENERALES DE LAS FAMILIAS	28
4.2 ROBO-ADVISORS.....	29
4.3 SITUACIÓN ACTUAL DE FONDOS DE INVERSIÓN GESTIONADOS CON IA EN ESPAÑA	29
5. CONCLUSIONES	31
6. REFERENCIAS BIBLIOGRÁFICAS.....	33
7. ANEXOS	41

ÍNDICE DE GRÁFICOS

GRÁFICO 1 RIQUEZA MEDIA NETA DE LAS FAMILIAS A LO LARGO DE UN PERIODO DE AÑOS.....	42
GRÁFICO 2 VALOR MEDIO DE DESCOMPOSICIÓN EN TANTO POR CIENTO DE LAS INVERSIONES DE LAS FAMILIAS 2002-2020.	43
GRÁFICO 3 VALOR MEDIO DE ACTIVOS FINANCIEROS DE LAS FAMILIAS 2002-2020....	43

ÍNDICE DE FIGURAS

FIGURA 1 ARQUITECTURA DE SISTEMA DE UN ALGORITMO OPTIMIZADOR DE CARTERA	44
FIGURA 2 TOPOLOGÍA GENERAL DE UNA RED NEURONAL ARTIFICIAL	44
FIGURA 3 UN ÁRBOL DE DECISIONES QUE ILUSTRA EL DESARROLLO GENÉRICO	45
FIGURA 4 EJEMPLO DEL PROCESO DE UN ALGORITMO GENÉTICO.....	46
FIGURA 5 COMPONENTES DEL PROCESAMIENTO DE LENGUAJE NATURAL	47

ÍNDICE DE TABLAS

TABLA 1 DESCOMPOSICIÓN DEL VALOR MEDIO DE ACTIVOS FINANCIEROS DE LAS FAMILIAS 2002 -2020.	48
TABLA 2 DATOS BÁSICOS DE FONDOS CON IA EN ESPAÑA.....	49
TABLA 3 DATOS RENTABILIDAD ANUALIZADA DE FONDOS IA EN ESPAÑA.....	50
TABLA 4 RATIOS DE RENTABILIDAD-RIESGO DE FONDOS IA EN ESPAÑA	51

1. INTRODUCCIÓN

El objetivo del presente Trabajo Fin de Grado es analizar cómo puede influir la Inteligencia Artificial en la inversión financiera de las familias. Las familias buscan maneras de invertir en los mercados financieros diversificando sus inversiones y tratando de conseguir la mayor rentabilidad posible con un riesgo determinado. Un pequeño inversor no se puede permitir tener un gestor propio que administre su patrimonio. Tampoco, todos los inversores tienen los suficientes conocimientos financieros para realizar sus inversiones de forma directa. En consecuencia, el uso de los robo-advisor o los fondos de inversión gestionados por inteligencia artificial puede ser una buena opción para ellos.

Parece que la inteligencia artificial (en adelante, IA) puede transformar muchos sectores de formas que incluso ahora no podemos imaginar. Uno de esos afectados puede ser el sector financiero. Las grandes entidades empiezan a emplear esta tecnología, por ejemplo, en “chats digitales vitaminados” con lenguajes de aprendizaje autónomo. Las razones de estos movimientos pueden ser ahorrar costes laborales, entrar en la moda de la IA de cara a accionistas y clientes o anticiparse al futuro del sector. La pregunta es si subirán la apuesta y dejarán que la IA gobierne a sus decisiones de inversión.

En este trabajo se pretende analizar cómo puede ser el futuro de las inversiones del pequeño ahorrador con la ayuda de la IA, cuál es su presente en España, cómo pueden mejorar a la productos financieros convencionales y qué problemas presentan.

El trabajo se ha documentado considerando distintas publicaciones que se enumeran en la bibliografía. También se ha utilizado páginas web de referencia como la Comisión Nacional del Mercado de Valores (CNMV), Banco de España, Morningstar, de donde se extrae una extensa muestra de datos de fondos de inversión, al igual que Yahoo! Finance, e Inverco, que es la asociación española de Instituciones de Inversión Colectiva y Fondos de Pensiones. Estas páginas listadas previamente se han usado para extraer datos reales acerca de datos macroeconómicos y fondos de inversión.

El resto del trabajo se organiza de la siguiente forma. En primer lugar, la siguiente sección detalla conceptos generales de la IA, su historia relativa a las finanzas y principales modelos que conectan la IA con las finanzas. La sección tercera continua con las principales aplicaciones de la IA a las finanzas. La sección cuarta presenta la situación de inversión de las familias y más concretamente de la IA. Para finalizar, la última sección resume las conclusiones más importantes del trabajo.

2. INTELIGENCIA ARTIFICIAL

2.1 Introducción e historia

Los modelos de IA son sistemas basados en máquinas con diversos niveles de autonomía que pueden, para un conjunto dado de objetivos definidos por humanos, hacer predicciones, recomendaciones o tomar decisiones utilizando cantidades masivas de datos de fuentes alternativas y análisis de datos denominados *big data* (OECD, 2021).

A partir de 1980 se populariza el aprendizaje automático o *machine learning* (ML) que entra dentro del paraguas de la IA. Se entiende como un subconjunto de técnicas que tiene como finalidad el aprendizaje autónomo de ordenadores sin instrucciones directas. Para ello es necesario un conjunto de datos y modelos matemáticos. Estos nuevos modelos, al igual que los de econometría tradicional, tienen por objeto principal extraer información de los datos y realizar predicciones (Carbó y Manuel, 2022).

Los primeros esfuerzos en IA se concentraron en reglas que incluían algoritmos basados en la lógica. Turing detalla en 1950 una prueba operativa (la Prueba de Turing) para el comportamiento inteligente. Turing proporcionó los componentes principales para el trabajo posterior de la IA con el lenguaje, el razonamiento, el conocimiento, el aprendizaje y la comprensión. A través de su prueba sentó la base del *machine learning*, los algoritmos genéticos y el aprendizaje por refuerzo (Buchanan, 2019).

El sector financiero es uno de los primeros ámbitos en suscitar interés por el uso de la IA, incluso antes de que se dispusiera de ordenadores. En los años 60, muchas investigaciones se centraron en la estadística bayesiana, un método primigenio en el aprendizaje autónomo de la IA. La estadística bayesiana se basa en la idea de expresar la incertidumbre sobre el estado (desconocido) de la naturaleza en términos de probabilidad. Durante las décadas de 1970 y 1980 se produjo “el invierno de la IA” sin mucho avance en este campo (Durkin, 2002; Raftery y Gill, 2002). La estadística bayesiana se sigue utilizando hoy en día, a menudo por parte de los analistas cuantitativos. Su función sigue siendo similar a la sugerida en los años sesenta. Es una herramienta para hacer frente a la incertidumbre, pero no debe ser la única que se utilice al hacer una evaluación. Algunos de sus usos son la predicción bursátil y labores en el sector de la auditoría (Grosse, 2013).

Durante la década de 1980, se utilizaron múltiples técnicas en la construcción de IA para las finanzas, como White (1988) con las Redes Neuronales o los Sistemas Difusos o *Fuzzy Systems* (J.J. Buckley, 1987). Sin embargo, gran parte del interés se centró en el uso de sistemas expertos o *expert systems*, que fueron usados para predecir tendencias del mercado y proveer planes financieros personalizados.

El sector bancario y las instituciones financieras empezaron a utilizar cada vez más los sistemas expertos para reducir los riesgos de los errores humanos. Ayudó en el análisis financiero, el análisis de mercado, el desarrollo empresarial, los negocios internacionales, el cambio de divisas y la gestión bancaria. Su uso directo se redujo en los 2000 pero han sido fundamentales para poner en marcha la IA e introducir su aplicación en las finanzas en el mundo moderno (Smigel, 2022). Otro sus usos fue la aplicación de la IA para la detección de fraude en los años 1990 (Senator y otros, 1995).

El *trading* algorítmico tiene origen en 1970 pero no fue hasta 2017 cuándo se creó el primer fondo gestionado enteramente por IA: AIEQ (Buchanan, 2019; Rui Chen, 2022). Aun así hay que tener en cuenta que existe un gran avance actualmente en todo lo relativo con la IA en todos los campos. Solo en 2021 la

inversión global en IA superó los 200.000 millones de dólares en total (OECD, 2021).

Fuera del sector tecnológico, la industria de servicios financieros es la que más gasta en servicios de IA y está experimentando un crecimiento muy rápido (Citi, 2018). Esta influencia se está viendo en todas las partes del ámbito financiero como bancos, *hedge funds*, aseguradoras, empresas *fintech* entre otros.

En el sector de los servicios financieros, las aplicaciones de la IA incluyen la negociación algorítmica, la composición y optimización de carteras, la validación de modelos, las pruebas retrospectivas, el robo-asesoramiento, los asistentes virtuales al cliente, el análisis del impacto en el mercado, el tratamiento de datos y un largo etcétera (Buchanan, 2019). A continuación se describen los modelos más importantes utilizados por la IA en finanzas.

2.2 Modelos

A lo largo de este trabajo, he repasado una serie enfoques que exploraban cómo usar la IA para aplicarlo a las finanzas, la valoración de activos y fondos. Todos ellos utilizaban diversos modelos para aplicar de forma más eficiente la funcionalidad seleccionada. En este punto explicaré varios modelos que creo que son los más usados o los más influyentes cuando van a basarse en ellos.

2.2.1 Redes Neuronales Artificiales

La red neuronal artificial (*artificial neural network*) sigue el modelo de una red neuronal biológica. Al igual que la red neuronal biológica, la red artificial es una interconexión de nodos, análogos a las neuronas. Cada neurona de la red es capaz de recibir señales de entrada, procesarlas y enviar una señal de salida. Cada red neuronal tiene tres componentes críticos: el nodo, la topología de la red y las reglas de aprendizaje. El nodo determina cómo procesa las señales, por ejemplo, el número de entradas y salidas asociadas al nodo y la función de activación. La topología de la red determina cómo se organizan y conectan los nodos (Livingstone, 2009).

Los nodos se organizan en matrices lineales, denominadas capas. Diseñar la topología de la red implica determinar el número de nodos de cada capa, el número de capas de la red y la ruta de las conexiones entre los nodos. Normalmente, esos factores se fijan inicialmente por intuición y se optimizan mediante múltiples ciclos de experimentos. También pueden utilizarse algunos métodos racionales para diseñar una red neuronal.

A la hora de entrenar a las redes, se realiza un proceso de aprendizaje. Los valores de entrada en la primera capa se conectan a las siguientes capas, llamadas capas ocultas, con varios porcentajes conocidos como pesos. Esto se puede ver en la Figura 2. Siendo un nodo de una capa oculta un 80% de un nodo de entrada y un 20% de otro, por ejemplo. El aprendizaje puede clasificarse en dos grandes categorías: aprendizaje supervisado y no supervisado (Hao y otros, 2016).

El aprendizaje supervisado consiste en entrenar una red utilizando las muestras de entrenamiento existentes, incluidos los datos y su correspondiente salida. Este se utiliza normalmente para clasificar y predecir datos. Los pesos se ajustan para minimizar el error entre la salida de la red y la salida correcta (Hao y otros, 2016).

En cambio, el aprendizaje no supervisado no utiliza valores de salida objetivo de un conjunto de entrenamiento. La red intenta descubrir el patrón o la tendencia subyacente sólo en los datos de entrada. Los distintos tipos de redes requieren procesos de aprendizaje diferentes (LeCunn y otros, 2015).

2.2.2 Árbol de decisión

Los árboles de decisión son modelos secuenciales que combinan de forma lógica una secuencia de pruebas simples. Cada prueba compara un atributo numérico con un valor umbral o un atributo nominal con un conjunto de valores posibles. Este modelo tiene una ventaja con el modelo de redes neuronales explicado antes, en términos de comprensibilidad. Las reglas lógicas que sigue un árbol de

decisión son mucho más fáciles de interpretar que los pesos numéricos de las conexiones entre los nodos de una red neuronal.

El algoritmo intenta generalizar o encontrar patrones en los datos introducidos. Para ello, determina qué pruebas (preguntas) dividen mejor las instancias en clases separadas, formando un árbol. Este procedimiento puede concebirse como una búsqueda codiciosa a través del espacio de todos los árboles de decisión posibles, eligiendo la única división, “la rama” que proporcione la mayor ganancia para el objetivo a lograr. La Figura 3 muestra un esquema del desarrollo de este procedimiento. Como ocurre con otros modelos de clasificación de patrones, los más complejos (árboles de decisión más grandes) tienden a producir peores resultados generados. Por tanto, se suele optar a producir árboles de decisión de menor tamaño (Kotsiantis, 2011).

El proceso de ejecución del modelo tiene dos fases principales: la fase de crecimiento y la fase de poda. La fase de crecimiento consiste en una separación repetitiva de los datos de entrenamiento que da lugar a un árbol de forma que se elige entre dos posibilidades, “las hojas”, de forma continua hasta llegar al objetivo. El objetivo de la fase de poda es eliminar las partes generadas en la fase de crecimiento que no aporten para evitar un número excesivo de datos. La fase de poda tiene como objetivo evitar divisiones que no cumplan ciertos umbrales específicos, se requiere por ejemplo un número máximo de hojas (Bengio y otros 2010).

Los árboles de decisión son la base de los bosques aleatorios “*random forest*” que tienen gran aplicación práctica debido a su mejor rendimiento y son principalmente una suma de árboles de decisión (Breiman, 2001).

2.2.3 Algoritmos genéticos

Los algoritmos genéticos se inspiran en la teoría darwiniana de la evolución, en la que se simula la supervivencia de las criaturas más aptas y sus genes. Es un algoritmo basado en las poblaciones en donde cada solución corresponde a un cromosoma y cada parámetro representa un gen. El algoritmo evalúa la aptitud

de cada individuo de la población mediante una función de aptitud (objetivo). Para mejorar las soluciones deficientes, las mejores soluciones se eligen aleatoriamente con un mecanismo de selección. Este operador tiene más probabilidades de elegir las mejores soluciones, ya que la probabilidad es proporcional a la aptitud (valor del objetivo).

El algoritmo comienza con una población aleatoria que puede generarse a partir de una distribución aleatoria gaussiana para aumentar la diversidad. Esta población incluye múltiples soluciones potenciales donde el objetivo principal de la fase inicial es repartir las soluciones por el espacio de búsqueda de la forma más uniforme posible. Lo que puede ayudar a aumentar la diversidad de la población y tener más posibilidades de encontrar regiones prometedoras (Mirjalili, 2019).

A continuación, se produce un cruce entre genes dando lugar a “hijos”. Los genes con soluciones más adecuadas, medido por una función de aptitud, son más probables de ser seleccionados. Después de este proceso llegamos a la mutación, que se basa en cambios aleatorios de la solución. Teniendo en cuenta estos dos procesos, cada vez existen unos genes más capaces con una mejor solución. Los individuos más aptos se seleccionan y el proceso de muestreo se orienta hacia las regiones más aptas en donde la solución se ha probado más concluyente (Wang, 1991).

La condición imprescindible para dar por concluido el algoritmo es que finalice el bucle evolutivo principal. A menudo, se ejecuta durante un número predefinido de generaciones. El tiempo y el coste de las evaluaciones de las funciones de aptitud pueden limitar la duración del proceso de optimización. La última condición de fin es la convergencia del proceso de optimización. Todo este proceso queda detallado en la Figura 4. Se ha demostrado ser especialmente útil en el problema de optimización de cartera (Kramer, 2017).

2.2.4 Procesamiento del lenguaje natural

El procesamiento del lenguaje natural (*Natural Language Processing*) es un conjunto de técnicas computacionales para el análisis y la representación

automáticos de las lenguas humanas. Es esencial en el análisis de datos como información financiera. Casi todos los sistemas de *NLP* tienen una gramática y un analizador sintáctico asociado. La gramática es una especificación finita de un número potencialmente infinito de oraciones. Un analizador sintáctico, es un algoritmo que analiza una oración y asigna una o más descripciones estructurales de acuerdo con la gramática. Las oraciones, tras ser caracterizadas con descripciones estructurales, deben someterse a un procesamiento posterior (Chowdhary, 2020; Joshi, 1991).

El análisis de oraciones se divide a su vez en análisis sintáctico y análisis semántico. El objetivo general de este análisis es determinar qué "significa" una frase. En la práctica, esto implica traducir la entrada de lenguaje natural a un lenguaje con semántica simple, o a un lenguaje de comandos de base de datos. En la mayoría de los sistemas, la primera etapa es el análisis sintáctico. Este proceso se muestra en la Figura 5 (Chowdhary, 2020).

Las técnicas del procesamiento del lenguaje natural incluyen:

- Identificación de frases: son utilizadas al extraer información y pueden arrojar resultados dispares dependiendo en gran medida de la tecnología de reconocimiento específica, el tipo de frase y la estrategia de concordancia.
- Extracción de entidades con nombre: una entidad con nombre es una frase especial que identifica un concepto o entidad, como nombres propios, de empresas, organizaciones, etc. Obviamente, las entidades con nombre expresan mejor el concepto de entidad. Además, las entidades con nombre expresan información más precisa que las frases generales.
- Extracción de conceptos: es una frase especial más genérica que la entidad con nombre. La entidad con nombre identifica un concepto, por lo que podemos considerar que pertenece al mismo. El concepto también incluye otras frases que no pertenecen a la entidad nominal.
- Anáfora y correferencias: consiste en encontrar las cosas reales para el pronombre o la frase desconocida que aparece en el documento. Esta técnica parece haber contribuido a la recuperación de información, ya que

es capaz de eliminar las expresiones que resulten poco claras en el documento (Xi, 2013).

3. Aplicaciones de la Inteligencia Artificial en finanzas

3.1 Optimización de carteras

El proceso de toma de decisiones financieras suele basarse en la elección de activos prometedores y la asignación de recursos entre ellos. La optimización de carteras requiere encontrar una frontera eficiente que se caracteriza por ofrecer la mayor rentabilidad esperada para cada nivel de riesgo. El problema tiene varios objetivos y un amplio espacio de decisión. El problema de optimización cuadrática para maximizar la rentabilidad esperada y minimizar el riesgo se formula en la teoría moderna de carteras (Kabašinskas, 2022). La Teoría Moderna de Carteras (TMC) fue una contribución pionera propuesta por Markowitz en 1952 para optimizar la gestión de carteras de inversión. Este modelo introdujo dos métricas fundamentales: la rentabilidad esperada y la varianza (o desviación típica) de la rentabilidad.

La rentabilidad esperada refleja la idea de que un activo con un buen rendimiento en el pasado tiende a mantener ese rendimiento en el futuro. Sin embargo, la verdadera innovación de la TMC fue la introducción de la varianza en la medición del riesgo de la cartera. De esta forma, el riesgo se identifica con la variabilidad de los rendimientos históricos de una cartera en relación con su valor medio.

Con el tiempo, se han realizado numerosos avances para mejorar esta teoría, lo que ha llevado a la creación de más medidas de riesgo y la inclusión de restricciones prácticas en la negociación bursátil. Además, se han propuesto diversos métodos de optimización exactos e híbridos para resolver los modelos de optimización de carteras, los cuales se han vuelto cada vez más complejos (Ferreira y Gandomi, 2021).

Sin embargo, el uso de la varianza de la cartera como medida de riesgo ha sido objeto de críticas, debido a que la varianza considera tanto las desviaciones

negativas como las positivas. A pesar de estas críticas, la TMP sigue siendo una referencia importante en la gestión de carteras y ha sentado las bases para el desarrollo de otras teorías y enfoques más sofisticados en el campo de las inversiones. El uso de la IA es una solución propuesta para estas críticas, sin embargo, no existe un único modelo, por lo que nombraré las más importantes.

Rather y otros (2017) proponen como los modelos más utilizados los algoritmos genéticos (*genetic algorithms*), redes neuronales artificiales (*artificial neural networks*) y optimización por enjambre de partículas (*particle swarm optimization*). Mientras que Santos y otros (2022) enumeran los modelos difusos (*fuzzy*) de refuerzo (*reinforcement*), algoritmos evolutivos o genéticos (*evolutionary algorithms*) y aprendizaje profundo (*deep learning*). En lo que ambos trabajos están de acuerdo es que suele ser bastante frecuente los sistemas híbridos que engloban varios de estos modelos y que pueden llegar a ser los más prometedores.

Una de las diferencias entre modelos es el objetivo que sigue la IA programada cuando optimiza la cartera. Se divide en dos grupos, aquellos que siguen un objetivo único y los que tratan de conseguir objetivos múltiples. Dentro de los objetivos únicos están aquellos que buscan la minimización del riesgo y aquellos que intentan de maximizar el beneficio. La mayoría de trabajos difieren en la duración del periodo que estudia el algoritmo y con qué frecuencia varía los activos de la cartera.

Zhang y Liu (2013) desarrollan un modelo multiobjetivo que junta la programación difusa y los algoritmos genéticos. Lo novedoso de este modelo es que tiene en cuenta la probabilidad de bancarrota y lo analiza con la programación difusa para tras conseguir una correcta optimización de cartera con los algoritmos genéticos.

Solin y otros (2019) utilizan las redes neuronales artificiales para la predicción bursátil de acciones, para después realizar una cartera nuevamente con los algoritmos genéticos. Para ajustar la rentabilidad al riesgo en un modelo

multiobjetivo usan varios ratios como son los de Sharpe, Treynor y Jensen¹. Entrenaron las redes neuronales con precios máximos, mínimos y cierres históricos, además del volumen de transacciones en búsqueda de patrones para la predicción del precio al día siguiente. Skolpadungket y otros (2016) presentan un trabajo similar.

Silva y otros (2014) optimizan un modelo de inversión utilizando indicadores fundamentales y técnicos como se puede ver en la Figura 1. El objetivo del modelo, que usa algoritmos evolutivos multiobjetivo, es elegir las acciones que deben incluirse en una cartera de valores. Tiene dos objetivos: la rentabilidad y la varianza de las rentabilidades. El algoritmo elige el peso que debe darse a cada uno de los indicadores fundamentales y elige los valores de los indicadores técnicos. Los indicadores se basan en la información de los estados financieros de cada empresa. De esta forma, se pueden obtener diferentes estrategias con mayor y menor rentabilidad y varianza.

La preselección de indicadores fundamentales busca descubrir empresas estables, con buen crecimiento, un precio razonable y que tengan potencial para obtener mayores rentabilidades. No obstante, es el algoritmo el encargado de decir qué indicador debe tener un mayor o menor peso en la decisión final. Algunos de los indicadores fundamentales usados son:

- El rendimiento de los fondos propios (ROE) que mide el capacidad de la empresa a la hora de remunerar al inversor en función del capital invertido. Es la relación entre el beneficio neto y la cifra de fondos propios.
- El margen de beneficio que representa la parte de los ingresos por ventas de una empresa que se queda como beneficio, una vez restados todos sus costes.
- El ratio precio/beneficio (PER) que es la relación entre el precio actual de las acciones de una empresa y sus beneficios por acción.
- El ratio de reparto que mide el porcentaje de los ingresos netos que se distribuye como dividendos.

¹ Ratios explicados en el Anexo I.

Además también utiliza indicadores técnicos como:

- Seleccionador de mercados que invertir: indicador que considera el presupuesto disponible, los posibles costes en las transacciones, el conocimiento que tiene sobre el negocio de las empresas a invertir, el acceso a la información, y la experiencia a la hora de elegir un mercado.
- Momento de entrada al mercado: precio y condiciones de mercado que deben materializarse para entrar en una posición bursátil ya elegida. El objetivo es definir un conjunto de reglas que den una señal de entrada para mejorar el *timing* de compra y aumentar la fiabilidad del sistema.
- *Stop loss*: orden predeterminada y automática que reduce la exposición de una cartera y hace saltar de una posición perdedora. De esta forma, protege el capital restante disponible para continuar la actividad inversora.

Freitas y otros (2009) usan las redes neuronales para predicciones que pueden captar oportunidades de inversión a corto plazo. En concreto, consiste en obtener una frontera eficiente a partir de los datos disponibles en el momento t , seleccionar la cartera eficiente deseada, invertir toda la riqueza disponible en las acciones según las ponderaciones de la cartera y vender toda la cartera en el momento $t+1$. Finalmente, reinvertir toda la riqueza obtenida en una nueva cartera obtenida del mismo modo. No obstante, los autores reconocen que el modelo de optimización de carteras, basado en predicciones con oportunidades a corto plazo, utiliza una estrategia muy agresiva.

Jang y Seong (2023) comentan que el número de intentos de previsión y optimización de carteras en el mercado financiero ha aumentado debido a la necesidad de profundizar en el análisis de datos y al incremento de la cantidad de datos en el mercado financiero. Sin embargo, los estudios recientes de optimización de carteras se basan en análisis técnicos y aprendizaje por refuerzo (*deep learning*). El aprendizaje por refuerzo define el entorno (en este caso, el mercado de valores), el agente (es decir, el operador) y la recompensa (al ser multiobjetivo el riesgo y la rentabilidad) y optimiza la red de políticas para maximizar la recompensa optimizada de la cartera. Utiliza una red de políticas basada en una red neuronal convolucional con análisis técnico y un gradiente de políticas determinístico.

Steinbacher (2016), alternativamente, desarrolla un procedimiento de selección con aprendizaje por refuerzo en el que los agentes tienden a adoptar carteras que les han proporcionado altos rendimientos en el pasado. Un extra a este concepto es introducir elecciones imperfectas de los agentes y denominarlas “desconfianza”. Se dice que un agente es suspicaz si existe una probabilidad estrictamente positiva de que no adopte una cartera con mayor rentabilidad. En este modelo se hace una clara separación entre mercados alcistas y bajistas en donde los agentes suspicaces y no suspicaces utilizan estrategias distintas. Se observa que los agentes son mucho más susceptibles al riesgo en una tendencia bajista y sincronizan extremadamente sus elecciones en las carteras menos arriesgadas. En un mercado alcista se percibe una ligera desviación de las carteras de la frontera eficiente, mientras que el nivel de sincronización es más débil que en el mercado bajista. Además, se utiliza el método de Monte Carlo y el coeficiente de variación para examinar la consistencia en las selecciones de los agentes y descubrir que los agentes se comportan de forma más consistente en las elecciones más deseadas o en las menos deseadas. Vercher y Bermúdez (2015) desarrollan un modelo computacional del enfoque evolutivo multiobjetivo y el rendimiento del modelo de credibilidad se entena usando el mercado español.

Golia y otros (2019) aplican un híbrido de IA mejorada y optimización robusta para calcular el riesgo de una cartera de productos. Mediante la aplicación de una red neuronal mejorada con algoritmo de raíz-corredor (*runner-root algorithm*), se predice la demanda futura de cada producto y se calcula el índice de riesgo de cada producto en función de su demanda futura predicha. Se propone un modelo matemático de dos objetivos (minimizar el riesgo y maximizar la rentabilidad) en el que se tiene en cuenta el efecto de las inversiones, la fiabilidad y la pérdida de ventas permitida en la cartera de productos diseñada. Debido a la incertidumbre de la rentabilidad, se proponen dos modelos robustos de contrapartida.

Meghwani y Thakur (2018) proponen un modelo de tres objetivos. Además de tener en cuenta la rentabilidad y el riesgo, también tiene en cuenta los costes de transacción.

Otra opción es estimar coeficientes de regresión utilizando datos de expectativas de beneficios, variables de impulso de precios y datos financieros declarados (Guerard y otros, 2015).

El aprendizaje automático para la gestión de carteras también emplea la minería de reglas de asociación y la lógica difusa (*fuzzy logic*) con el fin de construir una lista de activos recomendados (Cavalcante y otros, 2016; Preeti Paranjape-Voditel, 2011). Además, la lógica difusa puede combinarse con indicadores técnicos para apoyar las decisiones de construcción de carteras (Yunusoğlu, 2012). También es posible el uso de técnicas de minería de datos que se combinan con reglas de asociación y análisis de agrupación para investigar el movimiento conjunto de los mercados financieros y el modelo resultante se utiliza para la recomendación de carteras (Liao y Chou, 2013).

Una de las mayores limitaciones de los modelos basados en IA es su lenta convergencia y que las soluciones óptimas no siempre están garantizadas. Por ejemplo, la optimización por enjambre de partículas y los algoritmos genéticos son técnicas de IA basadas en poblaciones en las que la solución óptima puede obtenerse o no en la primera pasada. Una práctica común consiste en reiniciar el algoritmo genético o la optimización por enjambre de partículas con un nuevo conjunto de poblaciones y esperar el resultado final. Muchos estudiosos han informado sobre la lenta convergencia de las redes neuronales artificiales. Por ejemplo, se trata de un método de ensayo y error para seleccionar el número óptimo de neuronas en las capas ocultas (Rather y otros, 2017).

3.2 Robo-Advisor

Los robo asesores ofrecen asesoramiento de inversión en línea basado en las respuestas del usuario a un cuestionario cumplimentado en línea. Para mantener los costes bajos y simplificar el proceso, utilizan algoritmos automatizados para hacer recomendaciones sobre cómo asignar recursos entre distintos tipos de

activos. En la mayoría de los casos, estos algoritmos se basan también en la teoría moderna de carteras. Construyen carteras con mayor riesgo aumentando la proporción de renta variable respecto a la renta fija y, dentro de cada tipo de instrumento, invirtiendo en activos de mayor riesgo (Abraham y otros, 2019).

El cuestionario está diseñado para obtener información que permita establecer los parámetros básicos de riesgo y las preferencias de inversión del usuario, pero no necesariamente obtiene información completa sobre su situación financiera. Basándose en las respuestas del usuario, el robo-advisor formula un programa de asignación de activos para el usuario y hace recomendaciones de inversiones específicas. Los clientes con objetivos de inversión similares suelen recibir el mismo asesoramiento de inversión y pueden tener las mismas o inversiones muy similares en sus cuentas (Fein, 2015).

Una característica clave de un robo-advisor es la ausencia de contacto humano entre el asesor y los inversores. Están diseñados para evitar la necesidad de una relación personal de asesoramiento con el cliente. Los principales robo-advisors ofrecen servicios de inversión discrecional con capacidad para efectuar transacciones de valores y llevar a cabo un programa de inversión para el usuario. Acto seguido se debe transferir dinero al robo-advisor o a sus filiales para que lo inviertan de acuerdo con el programa de inversión recomendado. La cuenta puede estructurarse para reasignar o reequilibrar automáticamente las inversiones en función de la información recibida. Las transacciones de valores suelen efectuarse a través del corredor de bolsa y/o el depositario afiliado al robo-asesor. Los robo-advisors menos completos sólo ofrecen recomendaciones de asignación de activos y reequilibrio, que el inversor debe aplicar en otro lugar. Algunos recomiendan inversiones en acciones individuales pero suele ser muy común que recomienden incluir fondos de inversión y fondos cotizados (ETFs) (Fein, 2015).

El saldo mínimo requerido para abrir una cuenta de robo-advisor suele ser pequeño, entre 1.000 y 10.000 euros. Algunos aceptan cuentas sin saldo mínimo. Por ejemplo, un robo-asesor totalmente automatizado puede cobrar una comisión tan baja como el 0,25% de los activos gestionados, mientras que las

comisiones de los asesores humanos tradicionales no suelen bajar del 0,75% (López y otros, 2015). Pueden obtener una remuneración adicional a través de intermediarios afiliados y no afiliados que prestan servicios de inversión al usuario del robo-advisor, por los que el usuario paga una comisión y/o acuerdos de reparto de ingresos con fondos de inversión. Los usuarios de robo-advisors también pagan comisiones directamente a los fondos de inversión y ETF en los que invierten a través del programa.

Una de las principal críticas de los robo-advisor es como obtienen el perfil de riesgo del cliente. El procedimiento actual de los robo-advisors es la aplicación de cuestionarios individuales basados en la web. Los roboasesores exigen que se introduzcan los datos en forma de una estructura predefinida. La mayoría de los robo-asesores ofrecen entre 1.000 y 5.000 combinaciones de respuestas diferentes para sus cuestionarios, pero algunos utilizan formularios dinámicos, formulando preguntas diferentes o adicionales en función de las respuestas anteriores; esto añade más complejidad a la evaluación y puede dar lugar a más de 1 millón de permutaciones. Algunos robo-asesores comprueban la coherencia de las respuestas de sus formularios con otras respuestas y supuestos razonables. Como resultado, algunas opciones de respuesta (por ejemplo, un horizonte de inversión de un año o un objetivo de inversión de liquidez segura) ponen fin al proceso de asesoramiento. También es posible que no pregunten al cliente el resto de las inversiones que tienen para realizar la cartera (Tertilt y Scholz, 2018).

Otra queja recurrente es la ausencia de contacto personal con un asesor humano que pueda evaluar con más detenimiento las necesidades y circunstancias de inversión del cliente. También es posible que el número de preguntas sea apabullante y que no haya ningún tipo de incentivo en complementar la lista de preguntas, en las que además, un 60% de ellas son intrascendentes para el perfil de riesgo final. Un asesor humano puede ofrecer orientación personalizada y animar a los inversores a ahorrar más, diversificar y realizar operaciones menos especulativas. Los asesores tradicionales pueden estar a disposición del inversor en momentos cruciales, como durante la volatilidad del mercado,

cuando es más probable que los inversores entren en pánico y cometan errores de inversión (Fein, 2015; Tertilt y Scholz, 2018).

Si se tiene en cuenta las aportaciones positivas de los robo-asesores, una adopción de estos asesores podrá disminuir la prevalencia de sesgos del comportamiento del inversor, como por ejemplo:

- El efecto de disposición (falacia del jugador) que es la tendencia a obtener ganancias con más frecuencia que pérdidas. Aquí los inversores venden acciones que ganan porque esperan que pierdan en el futuro. Al mismo tiempo, los inversores no quieren vender acciones que pierden, porque esperan que repunten y ganen más en el futuro.
- La persecución de tendencias, que es la propensión de los inversores a comprar valores tras una serie de subidas de precio posteriores. Este comportamiento sugiere que los inversores creen que el precio de una acción tiene más probabilidades de subir que de bajar tras una serie de subidas.

Siendo los resultados positivos estos sesgos no desaparecen por completo después de la intervención del robo-asesoramiento. Además se observó que los inversores sin el consejo de un asesor aumentaron la diversificación de su cartera tanto en términos del número de valores que poseían como de la volatilidad de su cartera ajustada al mercado y mostraron un mayor rendimiento en términos de rentabilidad de las operaciones ajustada por su riesgo sistemático (D'Acunto y otros, 2017).

Back y otros (2023) coinciden que los robo asesores pueden facilitar la toma de decisiones de inversiones difíciles, sobre todo en términos de reducción de los sesgos de comportamiento que pueden ocasionar costes económicos sustanciales.

Tao y otros (2021) argumentan que entre 2016 y 2019, en periodo alcista y en la bolsa estadounidense, los robo advisors han superado a otros fondos de inversión en los ratios riesgo/rentabilidad y el alfa de rendimiento y también a algunos de los principales índices de renta variable. Es decir, los robo-advisors

pueden llegar a proporcionar a los inversores rendimientos superiores ajustados al riesgo.

Caballero-Fernández y otros (2022) son menos entusiastas con los resultados de los robo-advisors en comparación de resultados con un grupo de analistas desde 2015 a 2019 de mercados latinoamericanos. Concluyen que ambas estrategias lograron superar al índice de referencia, sin embargo la cartera de inversión de analistas aceleró su crecimiento a partir de 2018, aumentando su diferencia positiva frente a la cartera de su competencia. Además, esta cartera tiene una mayor posibilidad de obtener rentabilidades anormales o inesperadas que la cartera de los analistas, dado el riesgo sistemático que conlleva.

Torno y Schildmann (2020) opinan que los inversores interesados no deberían confiar ciegamente en un robo-advisor. Especialmente cuando se consideran horizontes de inversión más cortos, los inversores deben ser conscientes de que los robo asesores, suelen utilizar grandes porcentajes de renta variable en sus carteras recomendadas, por lo que presentan alta volatilidad. Esto puede deberse a su necesidad de ofrecer a los clientes una alta rentabilidad, pero también tiene el inconveniente de que, si se produce una crisis y se supera el horizonte de inversión, puede incurrir en pérdidas.

3.3 Análisis de datos

En los últimos años, con la aparición de Internet y de los cada vez más estrictos reguladores, se empuja a las empresas hacia una mayor publicación de informes. El número de datos disponibles a la hora de invertir pueden llegar a ser inabarcables. Es por ello por lo que, desde hace poco tiempo, empiezan a aparecer las empresas que establecen equipos de ciencia de datos para analizar todos esos reportes e informes publicados. La iniciativa la tomaron ciertos *hedge funds* y ya se empezó a extender por un amplio número de empresas que adoptaron nuevos métodos analíticos en múltiples puntos de la cadena de valor de la gestión de activos (Doshi y otros, 2019). En esta tendencia la IA puede desempeñar un papel vital.

Doshi y otros (2019) explican que las gestoras que se dedican al análisis de datos se centran sobre todo en utilizar ciertas fuentes de datos para mejorar los modelos predictivos existentes para la toma de decisiones de inversión. Algunos de estos enfoques son el análisis de sentimientos con los datos de redes sociales para tomar decisiones óptimas de inversión y negociación, aprovechando el procesamiento del lenguaje natural o el ya mencionado análisis de conjuntos de datos estructurados, como pueden ser la información financieras de empresas.

Con relación al análisis de sentimientos, García-Méndez (2023) toma los datos aportados por usuarios expertos en redes españolas de *microblogging* para cribar en tiempo real información financiera. Estos datos son útiles para generar indicadores sobre la fijación de precios de los activos (Houlihan y Creamer, 2021). Se basa en técnicas de Procesamiento del Lenguaje Natural (*Natural Language Processing*) clasificando los textos en negativos neutros y positivos, denominados como oportunidades y precauciones, para conocer en tiempo real las opiniones que influyen en el mercado. Además se basan en características del lenguaje para clasificar las emociones unidas a la información financiera para dar una valoración más precisa.

Gupta y otros (2020) clasifican el reconocimiento de emociones en dos enfoques principalmente: el basado en el léxico (basado en diccionarios) y el basado en el aprendizaje automático (*machine learning*). Este último se clasifica a su vez en enfoques de aprendizaje supervisado y no supervisado.

Los pasos a seguir por los modelos a la hora analizar los textos financieros son:

- La selección de información usada para clasificar algunos tipos de datos predefinidos en un documento de texto. Los sistemas de extracción buscan y eligen la información pertinente de los fragmentos, para después colocar todas las piezas extraídas en un contexto prefijado.
- Aplicación de modelos de procesamiento del lenguaje natural (*natural language processing*), que como su nombre indica, intenta ayudar a transformar mensajes imprecisos y ambiguos en mensajes legibles y listos para clasificar.

- Por último la clasificación de textos es un procedimiento que comprende la extracción de características del texto ya extraído, la reducción de la dimensionalidad debido al gran volumen de texto procesado, la selección del clasificador prefijados en un principio y al final la evaluación (Gupta y otros, 2020).

Para la predicción de precios de las acciones se utiliza principalmente las redes neuronales convolucionales (*convolutional neural networks*), las redes neuronales recurrentes (*recurrent neural networks*) y la memoria a largo plazo (*long short-term memory*). Lin y otros (2022) citan que a la hora de construir modelos de procesamiento de texto se basan en “algunos métodos de incrustación de texto, como la incrustación de palabras, frases y eventos, cuyo objetivo es traducir un vector de alta dimensión a un espacio de baja dimensión. Es decir, una incrustación captura la semántica específica de la entrada colocando entradas semánticamente similares cerca unas de otras en el espacio de incrustación. En otras palabras, la incrustación de palabras se refiere a un grupo de algoritmos de aprendizaje automático que aprenden representaciones vectoriales de palabras densas, contextuales y de alta dimensión. Estos modelos tratan de identificar qué tipo información corresponde para cada activo. Uno de los grandes problemas presentes en este tipo de trabajos es cómo verificar que el modelo entiende las etiquetas dadas en el texto analizado (Maia y otros, 2021).

Parece ser que la información negativa es más sensible ante la reacción de los mercados cuando se publican los textos financieros, por lo menos, es lo que argumenta Mushtaq y otros (2022). Por ello, los ejecutivos suelen incluir una mayor complejidad a los reportes a la hora de transmitir información negativa de la empresa. La consecuencia es una necesidad de aplicar modelos cada vez más precisos (Souza y otros, 2019). Creen además que puede ser una tecnología clave a la hora de mejorar la eficiencia del mercado.

Hernández-Nieves y otros (2021) da una primera aproximación al análisis de datos y al uso combinado de algoritmos y técnicas de *machine learning*, con el análisis tradicional de mercado. Sus resultados indican que cuando los algoritmos de *machine learning* se entrenan con una cantidad de datos

suficientemente grande, es posible predecir con cierto éxito el valor de cierre basándose en el valor de apertura actual del mercado. Así, tras identificar señales de compra y venta, se ha podido crear un sistema que recomienda al usuario comprar, mantener o vender una acción en un momento determinado del día, según la predicción obtenida por los algoritmos de regresión.

3.4 Predicción bursátil de quiebras y riesgo de crédito

Uno de los usos de la IA que se han utilizado durante más tiempo en las finanzas es la predicción del riesgo. Los bancos utilizan la IA para calcular la probabilidad de impago de cada individuo. A esto se le denomina *credit scoring* y fue una de las primeras aplicaciones reales en el sector. Actualmente, y viendo los beneficios que pueden traer la aplicación de esta tecnología, se puede empezar a aplicar para la valoración de activos financieros y la predicción de sucesos que afectan a su valoración como las quiebras.

La investigación sobre la predicción de la quiebra puede centrarse en gran medida en la identificación de los "síntomas" que conducen a la quiebra. Los modelos recientes llevan a cabo procedimientos internos de selección de variables. Esto significa que pueden decidir entre una amplia gama de variables lo que considera importante sin intervención humana. Los modelos estadísticos tradicionales se han abandonado en gran medida en favor de modelos de alta dimensionalidad como la Importancia de Gini (*Gini Importance*), la Ganancia de información (*Information Gain*) y la Frecuencia de división (*Split Frequency*), así como sus equivalentes de medida relativa como la importancia relativa de la variable. Todos ellos son enfoques centrados en los datos que analizan la capacidad predictiva de un parámetro en función de las variables seleccionadas y clasificadas en los nodos de los árboles en lugar de las pruebas de significación (Snow, 2020).

Existen muchas arquitecturas utilizadas para la detección del riesgo. Una de ellas son las reglas difusas (*Fuzzy rules*) se utilizan para capturar la incertidumbre directamente en el proceso de previsión y a menudo se combinan con otros métodos. Otras son las redes neuronales recurrentes (*Recurrent neural networks*) que se utilizan principalmente para modelizar datos de series

temporales. Relacionado con la anterior, las redes con memoria a corto plazo (*long short term memory*) aplican una arquitectura específica al marco de redes neuronales recurrentes para ampliar y perfeccionar el uso de la memoria en la previsión. Otras arquitecturas serían vecinos más próximos (*Nearest Neighbors*), árboles (*trees*) y agrupación (*clustering*) (Breedem, 2020).

Aly y otros (2022) distinguen las metodologías tradicionales para predecir la quiebra, como el Análisis Discriminante Múltiple y la Regresión Logística. Y aquellas que usan la IA que tienen un rendimiento superior, como la máquina de vectores de soporte, el árbol de decisión, el bosque aleatorio, la base ingenua, la red neuronal artificial y el algoritmo genético.

La utilización de estimadores y optimizadores es clave a la hora de calcular el riesgo. los estimadores se basan en un principio estadístico para estimar los valores de los parámetros del modelo, normalmente con sus correspondientes intervalos de confianza y pruebas estadísticas en el marco de la estadística tradicional. Los optimizadores suelen especificar un criterio de adecuación que debe optimizarse. Es posible que a veces sea difícil trazar líneas claras entre estas categorías, ya que los estimadores y los optimizadores pueden adoptar propiedades unos de otros. La estimación por máxima verosimilitud es el estimador estadístico dominante mientras que el descenso de gradiente (*gradient descent*) sería el optimizador más representativo (Breedem, 2020).

La modelización conjunta (*ensemble modeling*) es una técnica general que puede combinar predicciones de distintos tipos de modelos y parece especialmente adecuada para el riesgo de crédito, debido a los conjuntos de datos disponibles son normalmente limitados. El principio básico que subyace al modelado de conjuntos es que diferentes modelos pueden capturar diferentes aspectos de los datos. Esto puede proporcionar solidez frente a valores atípicos y anomalías, así como frente a los factores que se incluyen en el modelado. Uno de los mayores problemas de la predicción de quiebras es que la mayoría de los conjuntos de datos no están disponibles públicamente o sólo se refieren a escenarios económicos específicos, como empresas privadas de distintos países. En el caso de las empresas privadas, en general se dispone de poca

información, lo que dificulta la explotación de otras fuentes de información que pueden mejorar el rendimiento de la predicción de quiebra (por ejemplo, divulgaciones textuales, informes anuales, datos bursátiles) y que pueden ser utilizadas por modelos más complejos; la predicción de quiebra en realidad implica diferentes tareas: la predicción de impago en tareas para el año siguiente, utilizando datos pasados, y la tarea de predicción de la probabilidad de supervivencia que pretende predecir la probabilidad de que una empresa se enfrente a dificultades financieras años venideros.

La mayoría de los conjuntos de datos no permiten realizar ambas tareas, lo que supone una clara limitación para el desarrollo de modelos inteligentes que pretendan generalizar; La métrica más común para la predicción de quiebras es el Área Bajo la Curva (*Area Under the Curve*) (Lombardo, y otros, 2022). Estos mismos autores, definen que las redes neuronales deberían ser preferibles entre todos los modelos para las tareas de predicción de quiebras, ya que presentan la mejor capacidad de generalizar en casos nuevos y no vistos con cada métrica adoptada.

Snow (2020) propone un modelo para predecir la bancarrota de una empresa basado en una máquina de refuerzo de gradiente (*gradient boosting machine*). Una máquina de refuerzo de gradiente construye secuencialmente múltiples modelos de árboles de decisión a partir de los cuales se predice el resultado final. Los resultados de la declaración incluyen, entre otros, cuánto durará el proceso de quiebra, si la empresa saldrá con éxito, y si implicará una venta de activos.

Tiene en cuenta tanto valores contables como ratios a la hora de predecir bancarrotas. Los valores de activo y pasivo, los ratios de solvencia y los valores de ingresos son las dimensiones más importantes, pero también importan los valores de rentabilidad, valoración y liquidez. En concreto, las variables más fuertes identificadas son los ingresos antes de impuestos y los ingresos antes de gastos extraordinarios. Los dos ratios más influyentes son el BPA (beneficio por acción) y el ratio Precio/Ventas.

El trabajo concluye que las empresas que tienen grandes inversiones de I+D, es menos probable que quiebren, en igualdad de condiciones. En el entrenamiento del modelo se tienen en cuenta observaciones pasadas en las que se conocen los datos. El modelo, por tanto, aprende las asociaciones de clase a partir de los patrones pasados de las variables explicativas comúnmente denominadas características y mapea estos datos de entrada en un resultado de datos de acuerdo con funciones recién aprendidas, ponderadas y aproximadas.

Este estudio también hace un seguimiento de la duración del proceso de quiebra y enseña que las empresas con más efectivo e inversiones a corto plazo, pasarán, en igualdad de condiciones, más de un año en el procedimiento de quiebra. La razón de ello es posiblemente que las empresas con beneficios netos elevados son más complejas de desenmarañar. En cuanto a la supervivencia del proceso de quiebra, las empresas con un bajo coeficiente entre el pasivo corriente y el pasivo total, una elevada inversión y anticipos, así como un buen coeficiente de rotación de activos tienden a sobrevivir al proceso de quiebra. Si el pasivo corriente es proporcionalmente bajo, significa que hay menos demandas urgentes a corto plazo, lo que daría a la empresa tiempo suficiente para recuperarse y restablecerse como empresa en funcionamiento.

Muñoz-Izquierdo y otros (2019) relacionan los informes de auditoría para analizar la probabilidad de declararse en quiebra y anticipar condiciones financieras precarias con gran precisión. Modelizan un algoritmo que utiliza tres metodologías no paramétricas con sólo variables extraídas del informe de auditoría (la opinión de auditoría, el tipo de párrafos y el número de comentarios en el informe) para pronosticar una situación de quiebra. Usan el algoritmo *PART* que es un clasificador de inducción de reglas incorporando una forma modificada del árbol de decisión y eliminando algunos de los caminos hasta llegar a la forma de árbol parcial. Una vez que el algoritmo encuentra el árbol parcial, la construcción del árbol se detiene y se genera una regla con la hoja que representa el mayor número de situaciones. Pese a ser simple, promete un rendimiento similar al de otros algoritmos de aprendizaje automático.

4. Inversión financiera de las economías domésticas e Inteligencia Artificial

4.1 Datos generales de las familias

La riqueza neta media de las familias presenta una tendencia positiva los últimos años. No obstante, según el Gráfico 1, no han logrado alcanzar el máximo de 318.000 euros de 2008. Actualmente la riqueza neta media de las familias es de 272.970 euros. Una gran parte del patrimonio de las familias son activos reales (Gráfico 2) y esa es una de las razones por las que la riqueza no ha llegado a niveles de 2008. El precio de la vivienda en España llegó a máximos en junio de 2007 con 2.115 euros el metro cuadrado, mientras, en diciembre de 2020 el valor de las viviendas era de 1.754 euros el metro cuadrado (Idealista, 2023).

El porcentaje de riqueza bruta de los activos reales siempre ha presentado valores mayores al 70% en los años analizados en el Gráfico 2. Por contraste, las familias españolas no invertían en activos financieros más de un 12% de su riqueza bruta en la primera década de los 2000, pero eso está cambiando. En el Gráfico 3 se puede ver la evolución de la inversión en activos financieros, en donde se duplica la cantidad en poco más de un decenio. Con datos de 2020, los activos financieros como porcentaje de activos totales brutos representaron el 20,1% (Banco de España, 2020).

Descomponiéndolo por categorías, muchos de los activos financieros han ido aumentando con los años. Si atendemos a los datos de la Tabla 1, puede verse que el activo que más ha crecido son las cuentas y depósitos para pagos. El crecimiento del total de activos financieros puede llevar a engaño. Aun así, es reseñable el aumento de acciones y fondos frente a valores de renta fija en retroceso en el último decenio. Se puede explicar este hecho debido a los bajos tipos de interés presentes durante esos años.

Los fondos de inversión han mostrado un crecimiento constante desde principios de siglo. Actualmente representan el 11,5% de los activos financieros de los hogares. En valores absolutos, los fondos de inversión mueven 620.018 millones

a fecha de 2021, mientras que en 2015 era de 373.189 millones de euros (Inverco, 2022).

4.2 Robo-advisors

Según Statista (2023), a finales de 2022 2,29 billones de euros eran manejados a nivel global por parte de robo-advisors entre los más de 200 millones de usuarios que utilizaban algún tipo de robo asesor. De acuerdo con esta misma página se espere que se duplique esta cantidad para 2027. Si tenemos en cuenta solo España, Statista informa que cuentan con casi 45.000 millones de euros como activos gestionados y que se espere que lleguen a 81.000 millones de euros para 2027.

Los usuarios de robo asesores tienden a ser inversores más jóvenes con valores globales de cartera más bajos, que poseen activos, como fondos cotizados en bolsa (ETF), en mayor grado. Según Statista el 74,6% son menores de 44 años a nivel nacional, lo que contrasta con el perfil de edad en los fondos en España. Estos inversores son casi 16 puntos porcentuales menos propensos a buscar ayuda de asesores financieros humanos (Brenner y Meyll, 2020; Nourallah, 2022).

Según una encuesta elaborada por Statista (2023), el 14% de los encuestados preven usar en el futuro robo-advisor. Otro 30% se imaginaría consultando en algún momento un robo asesor. La cifra más alta entre los encuestados (40%) son los que no consideran consultar en un futuro los robo asesores. Para finalizar el 15% no sabe qué responder a esta encuesta. La edad de los encuestados varía entre los 18 y los 64 años (Statista, 2023).

4.3 Situación actual de fondos de inversión gestionados con IA en España

La aplicación directa de IA a fondos de inversión no tiene un largo recorrido en España, donde otros países con mayor innovación financiera como los anglosajones destacan. Los pocos fondos que se han atrevido a intentar ser dirigidos por un algoritmo han encontrado resultados mixtos, pero pueden ser indicativo de los puntos positivos y negativos de este tipo de fondos.

La Tabla 2 detalla las características básicas de los fondos que han usado IA. En su mayoría son de renta variable exceptuando uno que es de renta mixta, es decir, que tiene un máximo de 30% de activos de renta variable. En cuanto a su patrimonio, omitiendo el segundo de Santander, no superan los 30 millones de euros que puede considerarse como fondos poco cuantiosos. En cuanto a los gastos, con un promedio de 1,2% de los fondos analizados no son especialmente altas. Hay que tener en cuenta que los fondos gestionados por bancos repercuten de media un 1,7% de gastos TER. Salen perdiendo en este punto con los robo-advisor teniendo en cuenta los datos aportados con anterioridad. (Toledo y Marco, 2010).

En términos generales, la rentabilidad de los fondos españoles presenta un rendimiento bajo, sobre todo si se compara con el de otros activos. El promedio de los fondos de inversión actuales presenta una rentabilidad anualizada de 1,91% en los últimos 15 años. En comparación, para este mismo periodo, la bolsa nacional obtiene una rentabilidad anualizada media del 1,35%, muy inferior a otros índices extranjeros como pueden ser Eurostoxx 50 con un 4,2% y la del S&P500 con 10,7% (Fernández y otros, 2022)

La Tabla 3 muestra las rentabilidades de los fondos gestionados en España con estas características. Como se puede ver, existe mucha disparidad entre los años analizados, incluso de un año al siguiente del mismo fondo. Hay que comentar que los tres últimos fondos analizados fueron liquidados en el año 2022 puesto que la gran parte de su patrimonio era propiedad de un único inversor. Su rentabilidad anualizada es deficiente al compararla con los benchmark aportados por Morningstar y Yahoo! Finance. Únicamente Gestión Boutique VIII Global Gradient y Anattea Kutema FI mantienen un saldo positivo con respecto a sus benchmark.

Con respecto a la volatilidad, son resultados que no aporta mucha información por sí misma, por lo que se puede comparar con el índice estándar de renta variable que utiliza Morningstar y es válido para todos los fondos de este tipo. En este caso corresponde a “Morningstar EU Mod Gbl Tgt Alloc NR EUR” que

presenta una volatilidad a 3 años de 8,63%. Los resultados son muy dispares existiendo un fondo que casi triplica el índice estándar y otros que se quedan algo por detrás. Mientras, el único de renta mixta mantiene una volatilidad inferior que la de su índice ajustado “Morningstar EU Cau Gbl Tgt Alloc NR EUR”. Este índice obtiene una desviación de 5,85% comparada con 1,88% de nuestro fondo analizado.

Los resultados de los ratios rentabilidad-riesgo, expuestos en la Tabla 4, no pueden ser considerados como óptimos en la mayoría de los fondos, en especial el ratio de Sharpe. Sin embargo, según Ferruz y Sarto (2003), cuando se utiliza el ratio de Sharpe para medir el rendimiento de las carteras de inversión, puede haber situaciones en las que las clasificaciones resultantes sean inconsistentes. Esto ocurre especialmente cuando los rendimientos promedio de las carteras son más bajos que los rendimientos promedio del activo sin riesgo que se utiliza como referencia en el caso español. Lo anterior dicho puede suceder en algunos casos de los mostrados.

En definitiva, el análisis de los fondos que han tenido alguna experiencia reconocida con la inteligencia ha tenido resultados con margen de mejora en la mayoría de los casos hasta el momento. Seguramente habrá que esperar un mayor desarrollo de la tecnología para que más gestoras intenten abrir fondos con IA. Aun así, es posible que fuera de los principales entidades bancarias alguna startup apueste por este concepto con la revolución presente de la IA.

5. CONCLUSIONES

Este trabajo pretende hacer un examen de la IA y su relación con la inversión financiera, además de qué ruta puede presentar en años venideros.

La IA puede haber llegado para quedarse. Es la conclusión que puede extraerse en el último año en los reportes de prensa y declaración de ejecutivos. Afectando a tantos sectores y con las autoridades preocupados por regular la nueva tecnología que podría afectar a millones de empleos a nivel mundial. Concretamente, según Goldman Sachs (2023), el 35% de operaciones

comerciales y financieras podrían ser automatizadas por la IA en EE. UU. y Europa.

La realidad en el presente es que la IA está lejos de sustituir totalmente a los humanos y presenta serias dudas si lo hará en un futuro. Tanto los robo-advisor como algoritmos de optimización de carteras han presentado resultados que no destacan entre otros fondos comparativos, sobre todo en nuestro país. Los fondos de inversión que declaran utilizar esta tecnología no despuntan entre los fondos convencionales.

Parece que el principal problema es el conjunto de datos que usa el algoritmo para entrenarse. Con tantas variables a tener en cuenta y un entorno macroeconómico turbulento, utilizar modelos basados en tendencias de datos históricos puede presentar predicciones imprecisas a largo plazo.

Lo que sí parece que brindará una herramienta conveniente para el gestor en tiempos futuros. El estudio de la información relevante a la hora de valorar activos es el punto fuerte de esta tecnología. En definitiva, el análisis de datos es la aplicación más confiable para su uso en el presente.

Este trabajo ha tenido un enfoque genérico, puesto que trata una tecnología que le falta mucho desarrollo de su potencial aplicación. La elaboración de este proyecto ha tenido en cuenta algunos de las investigaciones con resultados que tienen más trascendencia a nivel teórico. A partir de esta visión más general del presente de la IA, surgen líneas de investigación más concretas en alguna aplicación específica, además de realizar un seguimiento del uso de esta tecnología en España.

Para finalizar, y aún presentando ciertas ventajas, los productos financieros ofertados con IA no sobresalen entre sus homónimos teniendo en cuenta sus resultados por el momento. O por lo menos no presentan soluciones claras a los problemas existentes de las inversiones financieras de este país. Una unión entre la IA y los gestores podrían aumentar la eficiencia de algunas de sus labores. Ciertamente nos podemos encontrar ante la revolución silenciosa de la

IA, en cuanto que cada vez más decisiones financieras recibirán el soporte de la IA.

6. REFERENCIAS BIBLIOGRÁFICAS

- Abraham, F., Schmukler, S. L., y Tessada, J. (2019). "Robo-Advisors: Investing through Machines.". *World Bank Research and Policy Briefs*, 134881.
- Aly, S., Alfonse, M., y Salem, A.-B. M. (2022). "Bankruptcy Prediction Using Artificial Intelligence Techniques: A Survey". En: Magdi, D.A., Helmy, Y.K., Mamdouh, M., Joshi, A. (Eds.) *Digital Transformation Technology. Lecture Notes in Networks and Systems*, 224. Springer.
- Back, C., Morana, S., y Spann, M. (2023). "When do robo-advisors make us better investors? The impact of social design elements on investor behavior". *Journal of Behavioral and Experimental Economics*, 103.
- Banco de España. (2020). *Encuesta financiera de los hogares (EFF)*. <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/ArticulosAnaliticos/22/T3/Fich/be2203-art21.pdf>.
- Bengio, Y. Delalleau, Olivier y Simard, Clarence. (2010). "Decision trees do not generalize to new variations". *Computational Intelligence*, 26, pp. 449-467.
- Breeden, J. L. (30 de mayo, 2020). "A Survey of Machine Learning in Credit Risk." SSRN <https://ssrn.com/abstract=3616342>.
- Breiman, L. (2001). "Random Forests". *Machine Learning* 45, pp. 5–32.
- Brenner, L., y Meyll, T. (2020). "Robo-advisors: A substitute for human financial advice?" *Journal of Behavioral and Experimental Finance*, 25.
- Brown, C.E., N., N.L., y Phillips, M. (1990). Expert Systems for Personal Financial Planning. *Journal of Financial Planning*, 3, pp. 137-143.
- Brown, K. C. Reilly, F. K. (2012) "Analysis of Investments and Management of Portfolios". *Cengage Learning*, 941.
- Buchanan, B. G. (2019). "Artificial intelligence in finance". *The Alan Turing Institute*.
- Caballero-Fernández, R., Cortez, K., y Ceballos-Hornero, D. (2022). *Robo Advisors vs. Value Investing Strategies: A Fuzzy Jensen's Alpha*

- Assessment*. En: Rodríguez García, M.d.P., Cortez Alejandro, K.A., Merigó, J.M., Terceño-Gómez, A., Sorrosal Forradellas, M.T., Kacprzyk, J. (Eds.) *Digital Era and Fuzzy Applications in Management and Economy*. XX SIGEF 2021. Lecture Notes in Networks and Systems, 384. Springer.
- Carbó, A. A.-R., y Manuel, J. (2022). "Inteligencia Artificial y finanzas: Una alianza estratégica". *Banco de España*. <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSerias/DocumentosOcasiones/22/Fich/do2222.pdf>.
- Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., y Oliveira, A. L. (2016). "Computational Intelligence and Financial Markets: A Survey and Future Directions." *Expert Systems With Applications* 55, pp. 194–211
- Chowdhary, K.R. (2020). "Natural Language Processing". *Fundamentals of Artificial Intelligence* (603-649) Springer.
- Citi. (2018). "Bank of the Future: the ABCs of Digital Disruption in Finance." *CitiReport March 2018*. <https://ir.citi.com/CiDxU7p7pAittTmqzfMCS9%2F91IS21vljJXbn3wjppSEYiTXJ8FvEPRWx8WmmrKNgBSzDi8E2mGOI%3D>.
- Comisión Nacional del Mercado de Valores (2022). "Los fondos de inversión y la inversión colectiva" [https://www.cnmv.es/DocPortal/Publicaciones/Guias/G06_Fondos_II C.pdf](https://www.cnmv.es/DocPortal/Publicaciones/Guias/G06_Fondos_II_C.pdf).
- D'Acunto, F., Prabhala, N., y Rossi, A. G. (2017). "The Promises and Pitfalls of Robo-Advising". [Comunicación en congreso] 8th Miami Behavioral Finance Conference 2017, <https://ssrn.com/abstract=3122577>.
- De Souza, J.A.S., Rissatti, J.C., Rover, S., Borba, J.A., (2019). "The linguistic complexities of narrative accounting disclosure on financial statements: An analysis based on readability characteristics". *Research in International Business and Finance*, 48, pp. 59–74.
- De Ville, B. (2013). "Decision trees." *WIREs Comput Stat*, 5, pp. 448-455
- Doshi, S., Kwek, J. y Lai, J. (2019). "Advanced analytics in asset management: Beyond the buzz". *McKinsey & Company*.

<https://www.mckinsey.com/~media/McKinsey/Industries/Financial%20Services/Our%20Insights/Advanced%20analytics%20in%20asset%20management%20Beyond%20the%20buzz/Advanced-analytics-in-asset-management-vf.pdf>.

Durkin, J. (2002). "History and applications". En Leondres, C. T. (Ed.) *Expert Systems: The Technology of Knowledge Management and Decision Making for the 21st Century (1-22)* Academic Press.

Lopez, J. C., Babacic, S., y De la Ossa, A. (2015). "Advice Goes Virtual: How New Digital Investment Services Are Changing" *Journal of Financial Perspective*, 3.

Fein, M. L. (2015). "Robo-advisors: A closer look". SSRN <https://ssrn.com/abstract=2658701>.

Fernandez, P., García, T. S., Acín, J. F. y Acín, P. F. (2022) "Rentabilidad de los Fondos de Inversión en España. 2006-2021" <https://ssrn.com/abstract=4022725>.

Ferreira, F. G., y Gandomi, A. H. (2021). "Artificial Intelligence Applied to Stock Market Trading: A Review". *IEEE Access*, 9, pp. 30898-30917.

Ferruz, L. y Sarto, J. L. (2004) "An analysis of Spanish investment fund performance: some considerations concerning Sharpe's ratio" *Omega*, 32, pp. 273-284.

Freitas, F. D., Souza, A. F., y Almeida, A. R. (2009). "Prediction-based portfolio optimization model using neural networks." *Neurocomputing*, 72, pp. 2155-2170.

Goldman Sachs (2023) "The Potentially Large Effects of Artificial Intelligence on Economic Growth" https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf

Golia, A., Zare, H. K., Tavakkoli-Moghaddam, R., y Sadeghieh, A. (2019). "Hybrid artificial intelligence and robust optimization for a multi-objective product portfolio problem Case study: The dairy products industry." *Computers & Industrial Engineering*, 137.

- Grosse, R. (2013). "Bayesian machine learning" *Metacademy* https://metacademy.org/roadmaps/rgrosse/bayesian_machine_learning.
- Guerard Jr, J. B., Markowitz, H., y Xu, G. (2015). "Earnings forecasting in a global stock selection model and efficient portfolio construction and management." *International Journal of Forecasting*, 31, pp. 550-560.
- Gupta, A., Dengre, V., Kheruwala, H.A. (2020) "Comprehensive review of text-mining applications in finance." *Financial Innovation*, 6, pp. 39.
- Hao X., Zhang, G., y Ma, S. (2016). "Deep Learning". *International Journal of Semantic Computing*. 10, pp. 417-439.
- Hernández-Nieves, E., Canto, Á. B., Chamoso-Santos, P., Prieta-Pintado, F. d., y Corchado-Rodríguez, J. M. (2021). "A Machine Learning Platform for Stock Investment Recommendation Systems". En Dong, Y., Herrera-Viedma, E., Matsui, K., Omatsu, S., González Briones, A., Rodríguez González, S. (Eds.) *Distributed Computing and Artificial Intelligence, 17th International Conference. DCAI 2020. Advances in Intelligent Systems and Computing*, 1237. Springer.
- Holland, J. H. (1992) "Genetic Algorithms", *Scientific American*, 267, pp. 66-73.
- Houlihan, P., Creamer, G. G. (2021). "Leveraging social media to predict continuation and reversal in asset prices." *Computational Economics*, 57(2), pp. 433–453.
- Idealista (Mayo, 2023) "Evolución del precio de la vivienda en venta en España" *Idealista/data* <https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/>.
- Inverco (2023) "Qué son los Fondos y Sociedades De Inversión (IIC)" *Inverco* <https://www.inverco.es/20/0/133>.
- Inverco. (2022). "Las Instituciones de Inversión Colectiva y Fondos de Pensión." <https://www.inverco.es/archivosdb/ahorro-financiero-de-las-familias-iics-y-fp-2021.pdf>.
- J.J. Buckley. (1987). "The fuzzy mathematics of finance." *Fuzzy Sets and Systems*, 21, pp. 257-273.

- Jang, J., y Seong, N. (2023). "Deep reinforcement learning for stock portfolio optimization by connecting with modern portfolio theory." *Expert Systems with Applications*, 218.
- Jensen, M. C (1969). "Risk, the pricing of capital assets, and the evaluation of investment portfolios". *The Journal of Business*, 42(2), 167-247.
- Joshi, A. K. (1991). "Natural Language Processing". *Science*, 253, pp. 1242–1249.
- Kabašinskas, Č. J. (2022). "Review of Multi-Criteria Decision-Making Methods in Finance Using Explainable Artificial Intelligence." *Frontiers in Artificial Intelligence*, 5.
- Kotsiantis, S.B. (2013). "Decision trees: a recent overview." *Artificial Intelligence Review*. 39, pp. 261–283.
- Kramer, O. (2017). "Genetic Algorithm Essentials". *Springer Publishing Company* 679.
- LeCun, Y., Bengio, Y. y Hinton, . (2015). "G. Deep learning". *Nature*, 521, pp. 436–444.
- Liao, S.-H., y Chou, S.-Y. (2013). "Data mining investigation of co-movements on the Taiwan and China stock markets for future investment portfolio." *Expert Systems with Applications*, 40(5), pp. 1542-1554.
- Lin, W., Tsai, C., Chen, H. (2022). "Factors affecting text mining-based stock prediction: Text feature representations, machine learning models, and news platforms" *Applied Soft Computing*, 130.
- Livingstone, D.J. (2009). "Artificial Neural Networks: Methods and Applications". *Methods in Molecular Biology*, 458.
- Lombardo, G., Pellegrin, M., Adosoglou, G., Cagnoni, S., Pardalos, P. M., y Poggi, A. (2022). "Machine Learning for Bankruptcy Prediction in the American Stock Market: Dataset and Benchmarks." *Future Internet*, 14(8), pp. 244.
- M. Maia, J. E. Sales, A. Freitas, S. Handschuh y M. Endres, (2021). "A Comparative Study of Deep Neural Network Models on Multi-Label Text Classification in Finance," [Comunicación en congreso] IEEE 15th International Conference on Semantic Computing (ICSC),

Laguna Hills, CA, USA,

<https://ieeexplore.ieee.org/document/9364465>.

- Markowitz, H. (1952). "Portfolio selection". *Journal of Finance*, 7(1).
- Meghwani, S. S., y Thakur, M. (2018). "Multi-objective heuristic algorithms for practical portfolio optimization and rebalancing with transaction cost." *Applied Soft Computing*, 67, pp. 865-894.
- Mirjalili, S. (2018). "Evolutionary Algorithms and Neural Networks: Theory and Applications". Springer Publishing Company, Incorporated.
- Muñoz-Izquierdo, N., Camacho-Miñano, M.-d.-M., Segovia-Vargas, M.-J., y Pascual-Ezama, D. (2019). *Is the External Audit Report Useful for Bankruptcy Prediction? Evidence Using Artificial Intelligence* .
- Mushtaq, R., Gull, A. A., Shahab, Y. Derouiche, I., (2022). "Do financial performance indicators predict 10-K text sentiments? An application of artificial intelligence". *Research in International Business and Finance*, 61.
- Nourallah, M. (2022). "One size does not fit all: Young retail investors' initial trust in financial robo-advisors." *Journal of Business Research*, 156.
- OECD. (2021). "Artificial Intelligence, Machine Learning and Big Data in Finance." <https://www.oecd.org/finance/financial-markets/Artificial-intelligence-machine-learning-big-data-in-finance.pdf>.
- Preeti Paranjape-Voditel, U. D. (2011). "An Association Rule Mining Based Stock Market Recommender System." [Comunicación en congreso] Second International Conference on Emerging Applications of Information Technology, Kolkata, India, pp. 21-24, <https://ieeexplore.ieee.org/document/5734909>.
- Raftery, A., y Gill, H. (2002) "Bayesian Statistics" [Presentación de PowerPoint] *American Sociological Association* <https://sites.stat.washington.edu/raftery/Research/PDF/bayescourse.pdf>.
- Rather, A. M., Sastry, V. N., y Agarwal, A. (2017). "Stock market prediction and Portfolio selection models: a survey." *OPSEARCH*, 54, pp. 558-579
- Rui Chen, J. R. (2022). "Do AI-powered mutual funds perform better?" *Finance Research Letters*, 47(A).

- Santos, G. C., Veiga, A. C., y Gomes, K. (2022). "Portfolio optimization using artificial intelligence: a systematic literature review." *Exacta*.
- Senator, T. E., Goldberg, H. G., Wooton, J., Cottini, M. A., y Khan, A. F. (1995). "FinCEN Artificial Intelligence System: Identifying Potential Money Laundering From Reports of Large Cash Transactions." *Counterdrug Law Enforcement: Applied Technology for Improved Operational Effectiveness International Technology Symposium*.
<https://www.ojp.gov/ncjrs/virtual-library/abstracts/fincen-artificial-intelligence-system-identifying-potential-money>.
- Sharpe, William F (1994). "The Sharpe Ratio". *The Journal of Portfolio Management*, 21(1), pp. 49–58.
- Silva, A., Neves, R., y Horta, N. (2014). "A hybrid approach to portfolio composition based on fundamental and technical indicators". *Expert Systems with Application*, 42(4), pp. 2036-2048.
- Skolpadungket, P., Dahal, K., y Harnpornchai, N. (2016). "Handling Model Risk in Portfolio Selection Using Multi-Objective Genetic Algorithm." En Dunis, C., Middleton, P., Karathanasopolous, A., Theofilatos, K. (Eds.) *Artificial Intelligence in Financial Markets*, 1, pp. 285-310.
- Smigel, L. (27 de Abril, 2022). "History of AI in Finance." *Analyzing Alpha*.
<https://analyzingalpha.com/history-of-ai-in-finance>.
- Snow, D. (6 de Noviembre, 2019). "Financial Event Prediction using Machine Learning." SSRN <https://ssrn.com/abstract=3481555>.
- Solin, M. M., Alamsyah, A., Rikumahu, B., y Saputra, M. A. (2019). "Forecasting Portfolio Optimization using Artificial Neural Network and Genetic Algorithm". [Comunicación en congreso] *7th International Conference on Information and Communication Technology (ICoICT)*, Kuala Lumpur, Malaysia, 2019, pp. 1-7
<https://ieeexplore.ieee.org/document/8835344>.
- Statista. (18 de Febrero, 2023). "Robo-Advisor: Spain." *Statista Market Insights*
<https://es.statista.com/outlook/dmo/fintech/digital-investment/robo-advisors/spain>.
- Statista. (8 de Junio, 2023). "RoboAdvisor usage (prospective) in Spain"
Statista Market Insights

<https://www.statista.com/forecasts/1001481/roboadvisor-usage-prospective-in-spain>.

- Steinbacher, M. (2016). "Portfolio Selection as a Multi-period Choice Problem Under Uncertainty: An Interaction-Based Approach" . En Dunis, C., Middleton, P., Karathanasopolous, A., Theofilatos, K. (Eds.) *Artificial Intelligence in Financial Markets*. (245–284) New Developments in Quantitative Trading and Investment.
- Tao, R., Su, C.-W., Xiao, Y., Dai, K., y Khalid, F. (2021). "Robo advisors, algorithmic trading and investment management: Wonders of fourth industrial revolution in financial markets". *Technological Forecasting and Social Change*, 163.
- Tertilt, M., & 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), pp. 70-84.
- Toledo, I. y Marco, R. (2010). "Costs associated with mutual funds in Spain". *International Journal Of Finance And Economics*, 15, pp. 165-179.
- Torno, A., y Schildmann, S. (2020). "What Do Robo-Advisors Recommend? - An Analysis of Portfolio Structure, Performance and Risk." En: Clapham, B., Koch, JA. (Eds.) *Enterprise Applications, Markets and Services in the Finance Industry*, (92–108). Springer.
- Vercher, E., y Bermúdez, J. D. (2015). "Portfolio optimization using a credibility mean-absolute semi-deviation model." *Expert Systems with Applications*, 42(20), pp. 7121-7131.
- Wang, L., y Bayer, S. E. (1991). "Genetic algorithms". En: Interdisciplinary Computing in Java Programming (Ed.), *The Springer International Series in Engineering and Computer Science*, (101-116) Springer.
- White. (1988). "Economic prediction using neural networks: the case of IBM daily stock returns". [Comunicación en congreso] IEEE 1988 International Conference on Neural Networks, San Diego, CA, USA <https://ieeexplore.ieee.org/document/23959>.
- Xi, S. M. (2013). "Application of Natural Language Processing for Information Retrieval". *Applied Mechanics and Materials*, 380, pp. 2614–2618.

- Yunusoğlu, M. G. (2012). "A fuzzy rule based expert system for stock evaluation and portfolio construction: an application to Istanbul stock exchange." *Expert Systems with Applications*, 40(3), pp. 908-920.
- Zhang, W.-G., y Liu, Y.-J. (2013). "Credibilitic mean-variance model for multi-period portfolio selection problem with risk control." *OR Spectrum*, 36, pp. 113–132.
- Zou, J., Han, Y., So, SS. (2008). "Overview of Artificial Neural Networks." En: Livingstone, D.J. (Ed.) *Artificial Neural Networks. Methods in Molecular Biology*, (14-22). Humana Press.

7. ANEXOS

ANEXO I

Ratio de Sharpe: expresión matemática ideada por el economista William F. Sharpe para medir el resultado de los fondos de inversión. Un ratio de Sharpe negativo significa que el tipo de interés sin riesgo de referencia es superior a la rentabilidad histórica o prevista de la cartera. (Sharpe, 1994)

$$S_h = \frac{R_{Ft} - R_{Bt}}{\sigma_D}$$

R_{Ft} = La rentabilidad del fondo en el periodo t.

R_{Bt} = El rendimiento de la cartera o valor de referencia en el periodo t.

σ_D = La desviación típica durante el período.

Ratio de Treynor: está basado en el ratio de Sharpe. Usa, sin embargo, la beta en vez de la desviación típica de la cartera. Asume que el riesgo relevante es el sistemático y mide el exceso de rentabilidad por unidad de riesgo sistemático. (Brown y Reilly, 2012)

$$T_i = \frac{\bar{R}_i - \overline{RFR}}{\beta_i}$$

\bar{R}_i = La tasa media de rentabilidad de la cartera durante un período de tiempo.

\overline{RFR} = La tasa media de rendimiento de una inversión sin riesgo.

β_i = Beta de la cartera.

Ratio alfa de Jensen: Es la diferencia entre la rentabilidad real de un fondo y la rentabilidad esperada que podría haber obtenido la cartera de referencia con el mismo nivel de riesgo de mercado. Jensen asegura que “si el gestor de la cartera tiene capacidad para prever los precios de los valores, la ecuación será positiva”. (Jensen, 1968)

$$\alpha_j = (R_{jt} - R_{Ft}) - \beta_j[R_{Mt} - R_{Ft}]$$

R_{jt} = Rendimiento de una cartera durante un período de tiempo t.

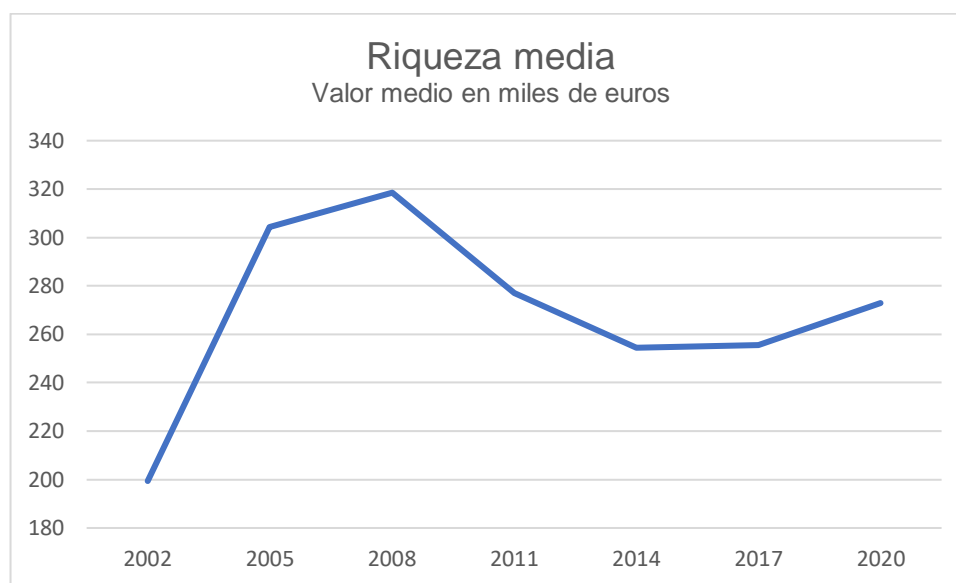
R_{Ft} = Rendimiento del tipo libre de riesgo durante un tiempo t.

β_j = Beta de la cartera.

R_{Mt} = Rendimiento del índice de mercado apropiado durante un tiempo t.

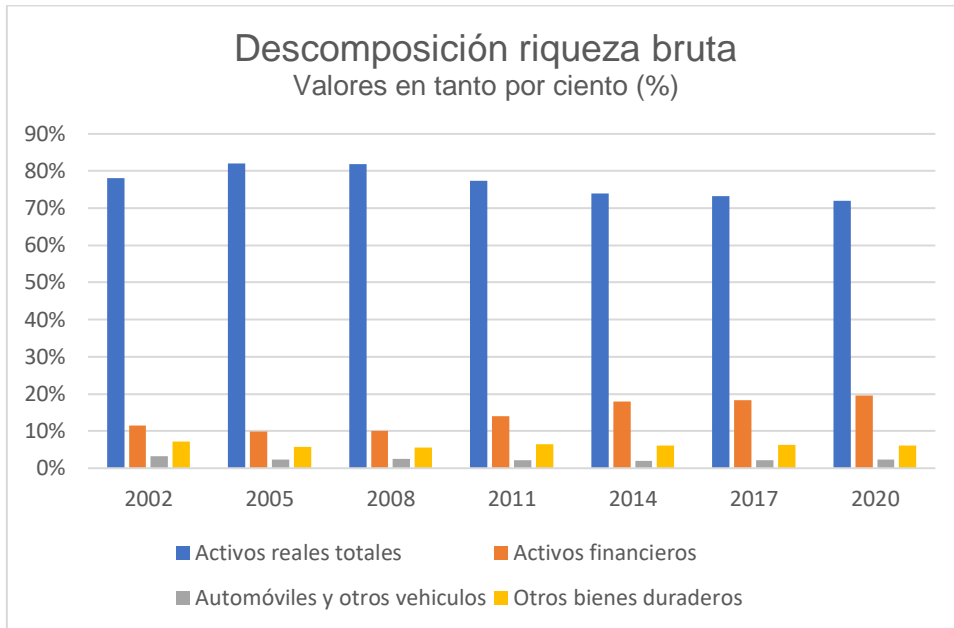
ANEXO II

Gráfico 1 Riqueza media neta de las familias a lo largo de un periodo de años.



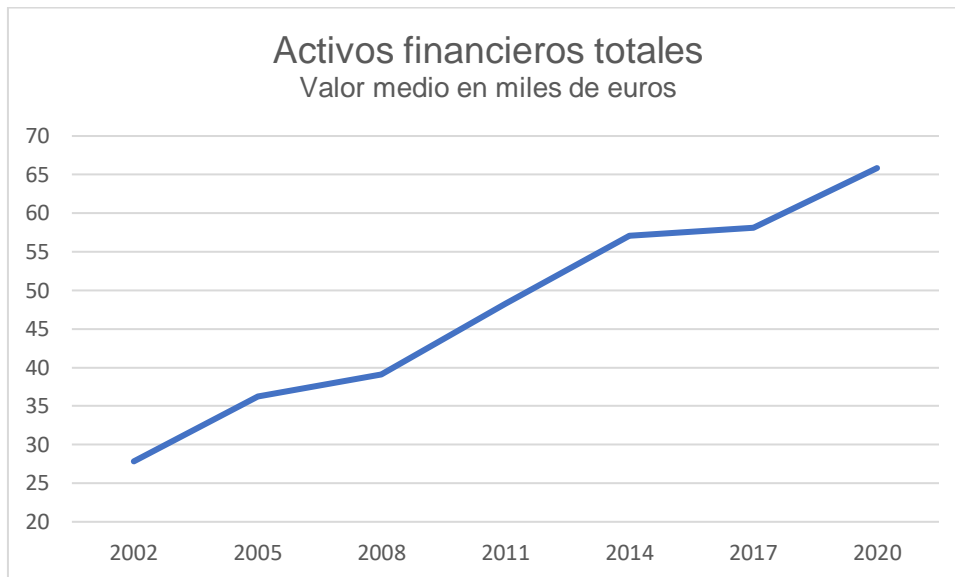
Fuente: Banco de España, Encuesta Financiera de las Familias. Euros (€), año base 2020.

Gráfico 2 Valor medio de descomposición en tanto por ciento de las inversiones de las familias 2002-2020.



Fuente: Banco de España, Encuesta Financiera de las Familias. Tanto por ciento (%) año base 2020.

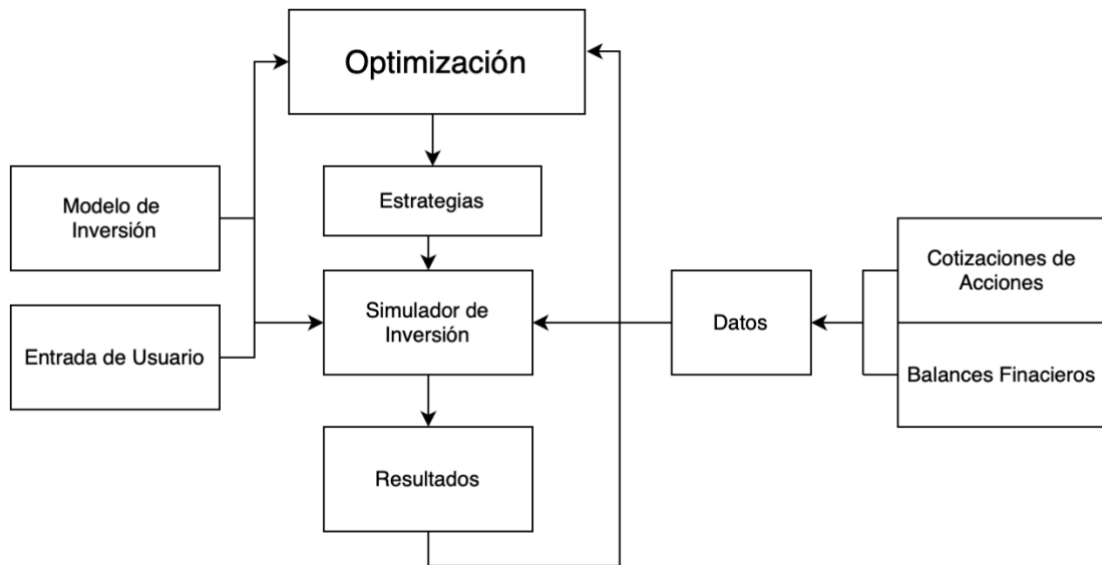
Gráfico 3 Valor medio de activos financieros de las familias 2002-2020.



Fuente: Banco de España, Encuesta Financiera de las Familias. Euros (€), año base 2020.

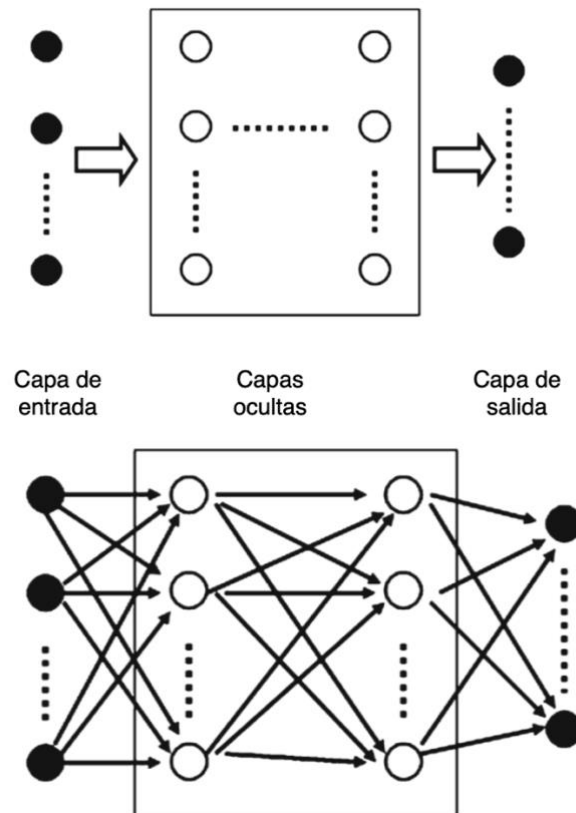
ANEXO III

Figura 1 Arquitectura de sistema de un algoritmo optimizador de cartera.



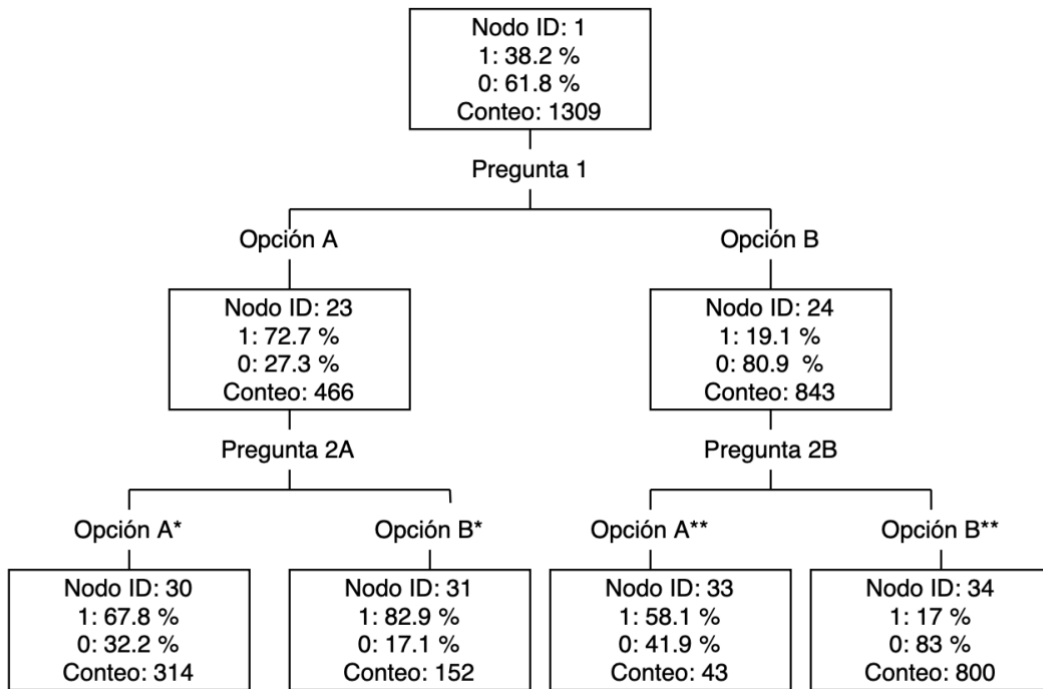
Fuente: Silva y otros (2014).

Figura 2 Topología general de una red neuronal artificial.



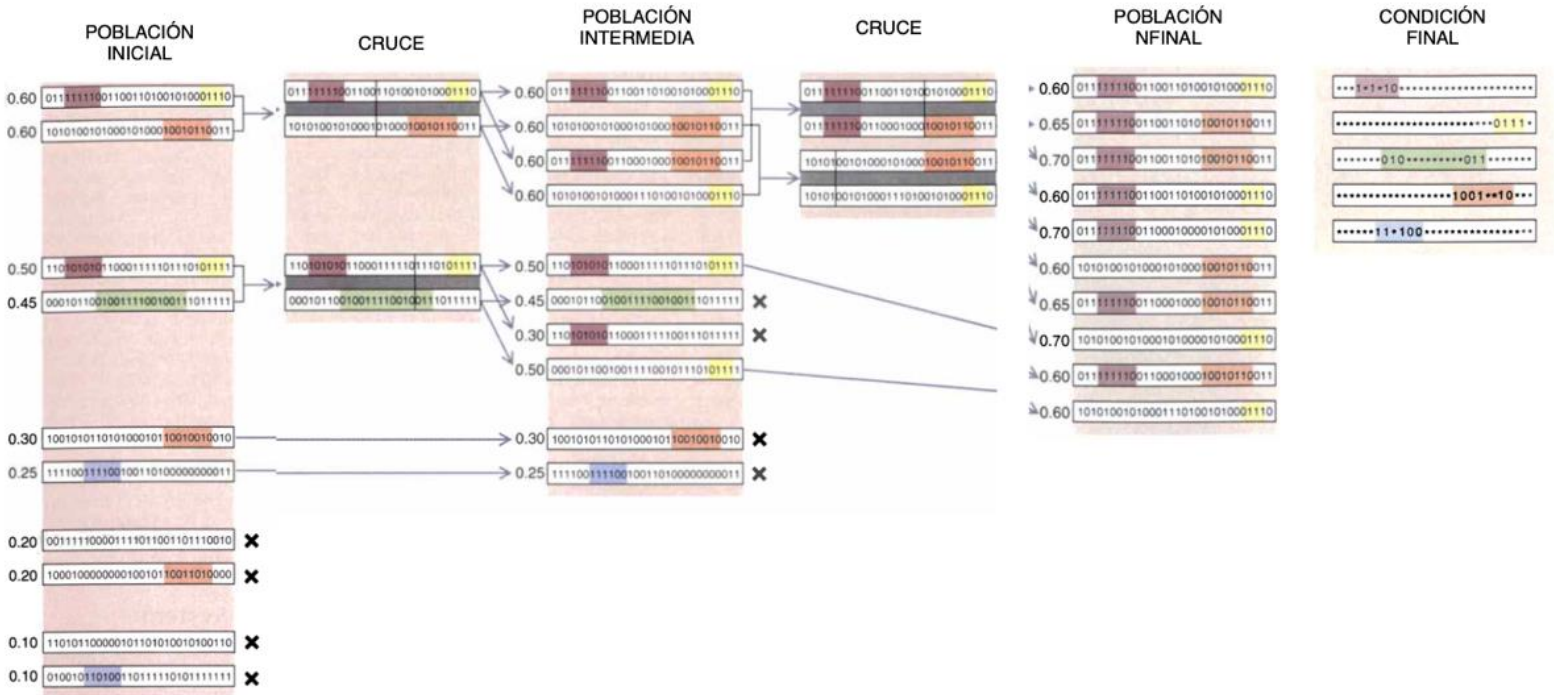
Fuente: Zou y otros (2008).

Figura 3 Un árbol de decisiones que ilustra el desarrollo genérico



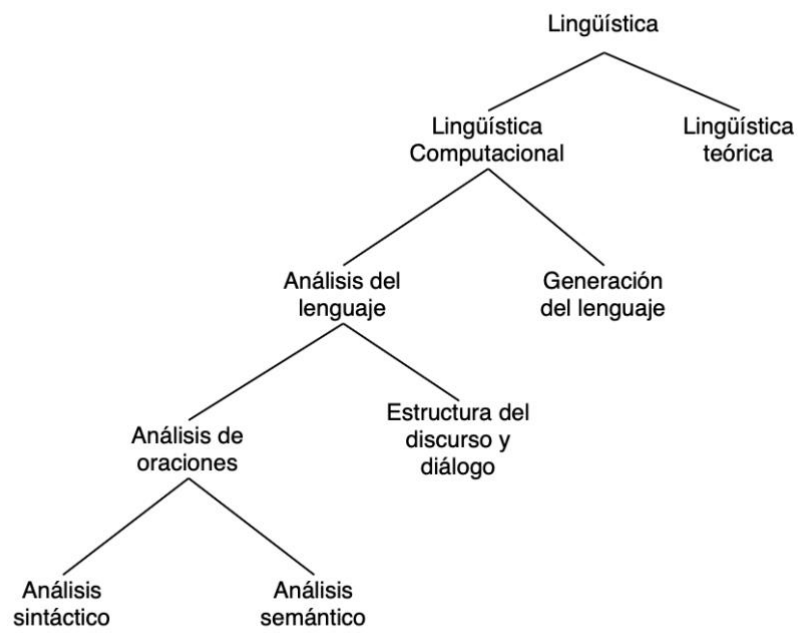
Fuente: de Ville (2013).

Figura 4 Ejemplo del proceso de un algoritmo genético



Fuente: Holland (1992).

Figura 5 Componentes del procesamiento de lenguaje natural



Fuente: Chowdhary (2020).

ANEXO IV

Tabla 1 Descomposición del valor medio de activos financieros de las familias 2002 -2020.

	2002	2005	2008	2011	2014	2017	2020
Cuentas y depositos para pagos	5.992,23	9.849,76	9.195,01	8.695,39	9.764,14	14.196,81	22.124,47
Acciones no cotizadas	3.351,73	3.244,40	3.155,59	7.743,87	9.428,10	7.978,98	9.419,37
Planes de pensiones y seguros vida	4.726,69	7.030,92	6.959,02	9.025,35	8.640,44	9.488,23	8.802,64
Fondos de inversión	2.614,60	4.945,61	2.990,33	2.708,77	5.277,98	5.587,26	7.572,62
Acciones cotizadas en bolsa	3.958,73	3.964,17	3.441,39	4.456,29	7.317,97	7.727,98	6.473,79
Cuentas no utilizables pagos	5.214,00	5.379,62	10.029,54	11.038,96	10.354,26	5.425,04	4.348,81
Valores de renta fija	609,31	667,02	720,66	862,74	336,86	159,78	184,22

Fuente: Banco de España, Encuesta Financiera de las Familias. Euros (€), año base 2020.

Tabla 2 Datos básicos de fondos con IA en España

	Categoría	Patrimonio	Gastos TER ²	Volatilidad ³
Santander PB Strategic Allocation F	Renta variable	16.176.000 €	1,62%	6,53%
Santander International Fund SICAV	Renta mixta	289.182.039 €	0,51%	1,88%
Gestión Boutique VIII Global Gradient	Renta variable	10.662.000 €	1,37%	22,28%
Anattea Kunie Quant FI	Renta variable	22.750.000 €	1,35%	6,95%
Anattea Kutema FI	Renta variable	3.200.000 €	1,00%	5,47%
Anattea Mayak FI	Renta variable	588.000 €	1,35%	10,88%

Fuente: Morningstar y Yahoo Finance!

² El ratio de gastos totales (TER) es una medida de los costes totales asociados a la gestión y el funcionamiento de un fondo de inversión. Estos costes consisten en comisiones de gestión y otros gastos adicionales. El coste total aplicado por el fondo se divide por el patrimonio total para obtener una cantidad porcentual, que representa el TER (CNMV, 2022).

³ La volatilidad es el movimiento observado del precio de un activo. En este caso Morningstar y Yahoo Finance! utilizan la desviación típica. La desviación típica es la medida estadística de la dispersión en torno a una media, que representa la amplitud de la variación de los rendimientos de una acción o cartera durante un determinado período de tiempo histórico. (Morningstar, 2023).

Tabla 3 Datos rentabilidad anualizada de fondos IA en España

	3 años anualizada	Variación con Benchmark ⁴	YTD	2022	2021	2020	2019
Santander PB Strategic Allocation F	-0,71%	-5,67%	1,70%	-15,32%	2,61%	-0,60%	4,83%
Santander International Fund SICAV	-0,11%	-5,21%	1,26%	-3,68%	-0,61%	-	-
Gestión Boutique VIII Global Gradient	10,16%	5,20%	-1,07%	-25,97%	38,88%	-	-
Anattea Kunie Quant FI	-0,91%	-7,51%	0%	-5,18%	5,05%	3,88%	0,38%
Anattea Kutema FI	3,37%	0,67%	0%	-0,78%	8,70%	5,32%	6,39%
Anattea Mayak FI	2,38%	-5,52%	0%	-4,23%	8,99%	5,68%	-0,75%

Fuente: Morningstar y Yahoo! Finance.

⁴ Datos de Benchmark, o categorías, suministrados por Morningstar y Yahoo! Finance. Diferencia entre la rentabilidad del fondo de los últimos 3 años anualizada menos la rentabilidad del benchmark correspondiente según la fuente.

Tabla 4 Ratios de rentabilidad-riesgo de fondos IA en España

	Alfa 3a ⁵	Beta ⁶	Ratio Sharpe ⁷
Santander PB Strategic Allocation F	-5,11	0,7	-0,29
Santander International Fund SICAV	-1,24	0,18	-0,18
Gestión Boutique VIII Global Gradient	7	1,22	0,56
Anattea Kunie Quant FI	0,82	0,47	-0,24
Anattea Kutema FI	3,93	0,11	0,14
Anattea Mayak FI	1,35	0,51	0,46

Fuente: Morningstar y Yahoo Financ

⁵ Alfa a 3 años

⁶ Calculado para un periodo de 3 años

⁷ Ratio de Sharpe calculado para los últimos 3 años de datos del fondo

