



**INTEGRACIÓN DE INTELIGENCIA ARTIFICIAL Y TEORÍA DE PORTAFOLIO EN  
LA EVALUACIÓN Y OPTIMIZACIÓN DE PRONÓSTICOS PARA ETFs**



**Juan David Murcia**

**Pontificia Universidad Javeriana Cali**  
**Facultad de ciencias económicas y administrativas**  
**Maestría en finanzas**  
**Santiago de Cali, 12 de febrero del 2024**

**PROYECTO DE INVESTIGACIÓN  
INTEGRACIÓN DE INTELIGENCIA ARTIFICIAL Y TEORÍA DE PORTAFOLIO EN  
LA EVALUACIÓN Y OPTIMIZACIÓN DE PRONÓSTICOS PARA ETFs**

**JUAN DAVID MURCIA PIEDRAHITA**

**Trabajo de grado presentado como requisito parcial para optar por el título  
de Magíster en Finanzas.**

**Director del trabajo de grado: Orlando Joaqui Barandica  
PhD(C) en ingeniería industrial.**

**PONTIFICIA UNIVERSIDAD JAVERIANA  
FACULTAD DE CIENCIAS ECONOMICAS Y ADMINISTRATIVAS  
MAESTRÍA EN FINANZAS  
SANTIAGO DE CALI**

**2024**

Santiago de Cali, 19 de julio de 2024

Doctor (a)

**Fabián Fernando Osorio Tinoco**

Decano

Facultad De Ciencias Económicas y Administrativas

Pontificia Universidad Javeriana

Cali

Por medio de la presente estoy entregando a usted el Trabajo de Grado cuyo título es “INTEGRACIÓN DE INTELIGENCIA ARTIFICIAL Y TEORÍA DE PORTAFOLIO EN LA EVALUACIÓN Y OPTIMIZACIÓN DE PRONÓSTICOS PARA ETFs”.

Espero que este Trabajo cumpla con los requisitos académicos exigidos y que alcance el propósito para el cual fue elaborado.

Atentamente,



---

Juan David Murcia Piedrahita

**1032499146**

Doctor  
Fabian Fernando Osorio Tinoco  
Decano  
Facultad de Ciencias Económicas y Administrativas  
Pontificia Universidad Javeriana

La Ciudad

Por medio de la presente me permito comunicarle, que en mi calidad de director de trabajo de grado he leído detenidamente el informe final del estudio titulado **“INTEGRACIÓN DE INTELIGENCIA ARTIFICIAL Y TEORÍA DE PORTAFOLIO EN LA EVALUACIÓN Y OPTIMIZACIÓN DE PRONÓSTICOS PARA ETFs”**, realizado por el estudiante de Maestría en Finanzas de la Facultad de Ciencias Económicas y Administrativas de la Universidad Javeriana nombres: **Juan David Murcia C.C 1032499146**, y considero que cumple con todos los requisitos requeridos para ser presentada a evaluación.

Atentamente

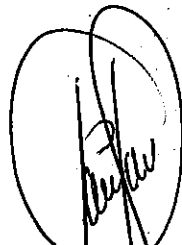
A handwritten signature in black ink, appearing to read 'O. Barandica', written over a horizontal line.

**ORLANDO JOAQUÍN BARANDICA**  
Director del Trabajo de Grado

ARTÍCULO 23 de la resolución N° 13 de julio 6 de 1946

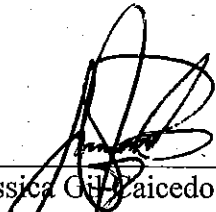
“La Universidad no se hace responsable por los conceptos emitidos por sus alumnos en sus trabajos de Tesis. Sólo velará porque no se publique nada contrario al dogma y a la moral católica y porque la Tesis no contenga ataques o polémicas puramente personales; antes bien, se vea en ellas al anhelo de buscar la Verdad y la Justicia”.

**“INTEGRACIÓN DE INTELIGENCIA ARTIFICIAL Y TEORÍA DE PORTAFOLIO EN LA EVALUACIÓN Y OPTIMIZACIÓN DE PRONÓSTICOS PARA ETFs**, Aprobado por el Comité de Trabajos de Grado en cumplimiento de los requisitos exigidos por la Pontificia Universidad Javeriana para optar por el título de Magíster en Finanzas.



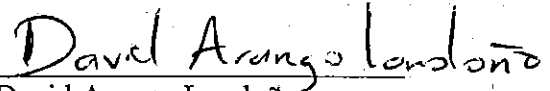
---

Fabián Fernando Osorio Tinoco  
Decano  
Facultad de Ciencias Económicas y Administrativas



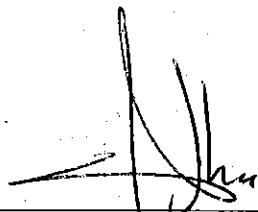
---

Jessica Gil Paicedo  
Directora de Maestría en Finanzas.



---

David Arango Londoño  
Jurado



---

Orlando Joaquín Barandica  
Director del Trabajo de Grado

Santiago de Cali, 10 de julio del 2024

## Resumen

Ante un mercado financiero global en constante evolución, los Exchange Traded Funds (ETFs) han surgido como herramientas versátiles de inversión, accesibles para diversificar portafolios a bajo costo y con alta flexibilidad. A continuación, esta investigación aborda la integración de la inteligencia artificial, especialmente el aprendizaje automático, para perfeccionar la evaluación y predicción de precios en 5 ETFs: SPDR S&P 500 ETF Trust (SPY), iShares Russell 2000 ETF (IWM), Invesco QQQ Trust (QQQ), Grayscale Bitcoin Trust (GBTC), Vanguard Total Stock Market ETF (VTI). A causa de la volatilidad y diversidad del mercado, existe una necesidad crítica de desarrollar estrategias de inversión que no solo optimicen los retornos, sino que también minimicen los riesgos asociados.

Este estudio utiliza técnicas de aprendizaje automático, como el método KNN para la clasificación de tendencias de precios, combinadas con la teoría de portafolio de Markowitz, para formular enfoques de inversión que maximicen la rentabilidad ajustada al riesgo. Así pues, el análisis descriptivo inicial reveló patrones que influyen en el rendimiento de los ETFs, estableciendo una base sólida para la predicción precisa de movimientos de precios y la construcción de portafolios optimizados.

Los resultados muestran que la integración de técnicas avanzadas de aprendizaje automático con principios fundamentales de inversión mejora notablemente la gestión de portafolios de ETFs, evidenciando una capacidad de predicción superior al 50% para tendencias de precios. Siendo así, esta capacidad para equilibrar el rendimiento y el riesgo es crucial en el entorno financiero actual, caracterizado por su rápida evolución y complejidad. Además, los hallazgos sugieren que adoptar enfoques tecnológicos avanzados puede superar las limitaciones de las estrategias de inversión tradicionales, ofreciendo a los inversores herramientas más robustas para navegar mercados financieros volátiles.

En conclusión, este estudio subraya la importancia de combinar la inteligencia artificial con la teoría de portafolios para optimizar la gestión de inversiones en ETFs, proponiendo futuras investigaciones que extiendan estos métodos a otros tipos de activos financieros, lo que podría ampliar aún más las estrategias de inversión efectivas y seguras.

# 1. Introducción

La revista Forbes define a los Exchange trust funds (ETF's) como uno de los instrumentos mas rápidos y adoptados en la historia de los mercados financieros, estos han surgido como alternativas de inversión esenciales, facilitando el acceso a diferentes activos y mercados para inversores de todo tipo. De acuerdo con Lettau y Madhavan (2018), los ETFs representan una de las innovaciones financieras más importantes de las últimas décadas, destacándose por su capacidad para replicar el rendimiento de índices específicos, al igual que los fondos mutuos indexados. La diversidad y significancia en el mercado de los ETFs, incluidos los avances en ETFs de bonos y fondos "beta inteligente", han crecido sustancialmente, con el potencial de transformar radicalmente el panorama de inversión global (pp. 135-154). Dicho esto, el crecimiento exponencial que ha tenido el uso de estos instrumentos financieros trae consigo diferentes oportunidades, especialmente en lo que respecta a la evaluación y optimización de pronósticos que permitan a los inversores mejorar sus rentabilidades.

Paralelamente, la inteligencia artificial (IA) y el aprendizaje automático han revolucionado diferentes sectores, proveyendo herramientas avanzadas capaces de analizar grandes volúmenes de datos, identificar patrones complejos y predecir tendencias futuras con precisión. Burrell y Folarin (1997) destacan en una investigación realizada en 1997 que las redes neuronales, una rama de la IA, representa una innovación significativa en el ámbito financiero, ofreciendo diferentes aplicaciones desde la evaluación de créditos hasta la gestión de grandes carteras de inversiones y demostrando, ya en ese entonces, un aumento sustancial en la productividad, remodelando el paisaje de la inversión financiera. De la misma forma, Kamalov (2020) demuestra cómo los modelos de redes neuronales, incluidos MLP, CNN y LSTM, pueden aplicarse efectivamente para prever cambios significativos en los precios de las acciones, ofreciendo un nuevo enfoque para la predicción de precios en el mercado financiero y destacando la capacidad superior de estos modelos en comparación con los métodos tradicionales.

Un estudio hecho por J.P. MORGAN asset management (2019) demostró que hasta el 85% de inversores institucionales en estados unidos tienen en su portafolio inversiones en ETFs y en Europa la cifra aumenta hasta el 88%, y la mayoría de estos inversionistas los utilizan por sus estrategias pasivas para lograr un crecimiento de capital. Entonces, la idea de crear nuevas estrategias que permitan inversiones más seguras y rentables que las metodologías tradicionales suena bastante atractiva, pero existe una dificultad al momento de minimizar el riesgo en la gestión de las inversiones en ETF. Entonces, el desafío reside en la naturaleza volátil y diversificada del mercado, donde las fluctuaciones son frecuentes y a menudo impredecibles. Esto hace que la tarea de minimizar el riesgo sea compleja, especialmente cuando se emplean estrategias pasivas que tradicionalmente dependen de la consistencia del mercado a largo plazo.

En este entorno, la capacidad para desarrollar metodologías avanzadas que no solo rastreen el mercado, sino que también anticipen cambios, se convierte en un aspecto crucial para el éxito de la inversión en ETFs. Por ende, la integración de la inteligencia artificial y la teoría de portafolio promete abordar esta problemática, ofreciendo un enfoque sistemático y basado en datos para la optimización de las inversiones en ETFs.

Dicho esto, en este documento se describen los resultados de la construcción un modelo de predicción mediante métodos de aprendizaje supervisado (método KNN), para identificar con precisión las tendencias direccionales en los precios de los ETFs. La precisión en la clasificación de las tendencias "Up" o "Down" no solo se traduce en una mejora de las capacidades predictivas de los precios históricos de cinco ETFs prominentes: iShares Russell 2000 ETF (IWM), Invesco QQQ Trust (QQQ), SPDR S&P 500 ETF Trust (SPY), Grayscale Bitcoin Trust (GBTC) y Vanguard Total Stock Marke Index Fund ETF Shares (VTI). Estos ETFs representan diferentes segmentos del mercado y son de gran interés para inversores y analistas debido a su impacto en carteras de inversión diversificadas. Además, se implementará un modelo de teoría de portafolio de Markowitz para construir y analizar un portafolio óptimo que maximice los rendimientos y minimice el riesgo a través de la combinación de cinco ETF (IWM, QQQ, SPY, GBTC y VTI).

En la primera sección de este documento se realizará una revisión de la literatura relacionada con las teorías tradicionales de mercado, los ETF, inteligencia artificial e inteligencia artificial aplicada a diferentes activos financieros. Segundo a esto, se desarrollará el marco conceptual donde se exponen diferentes términos claves para el desarrollo de esta investigación. Posteriormente, se presenta la estrategia metodológica empleada para este proyecto de investigación. Finalmente se presentan los resultados y las conclusiones.

## **2. Justificación y pertinencia**

La elección de la implementación de métodos de inteligencia artificial (IA), como el método KNN, y la aplicación de la Teoría de Portafolio de Markowitz en este trabajo se justifica en su capacidad para adaptarse y mejorar la gestión de inversiones en un entorno financiero en constante evolución. Aunque el CAPM y la EMH han sido fundamentos teóricos sólidos en la práctica financiera, su aplicación directa puede no ser suficiente para abordar las complejidades y dinámicas actuales del mercado.

En primer lugar, si bien el CAPM ofrece un marco para calcular el rendimiento esperado de los activos, basado en una relación lineal con su riesgo sistemático, la realidad del mercado moderno presenta una serie de desafíos que van más allá de esta concepción simplificada. La IA, con su capacidad para detectar patrones y correlaciones ocultas en los datos del mercado, proporciona una herramienta invaluable para

comprender y aprovechar estas complejidades. Esto desafía la noción de que los precios reflejan completamente toda la información disponible, como sugiere la EMH de Fama.

Además, la IA puede mejorar la eficiencia de los mercados al incorporar rápidamente información de tendencias en la toma de decisiones de inversión. Esto sugiere que existen oportunidades para que los modelos predictivos superen a aquellas estrategias tradicionales, proporcionando así una ventaja competitiva en el mercado.

Finalmente, estudios recientes que integran técnicas de IA desafían la idea de que no se pueden obtener ganancias sistemáticas al predecir los precios de las acciones, como postula la EMH. La precisión predictiva superior obtenida a través de estos enfoques no solo tiene implicaciones teóricas significativas, sino que también ofrece un valor práctico considerable para los gestores de portafolios que buscan maximizar los rendimientos y minimizar el riesgo en un entorno financiero cada vez más competitivo y dinámico.

La adopción de métodos de IA y la Teoría de Portafolio de Markowitz en este trabajo se justifica en su capacidad para mejorar la comprensión del mercado, optimizar la gestión de inversiones y ofrecer ventajas competitivas en un entorno financiero complejo y en constante cambio.

### **3. Marco teórico**

En este trabajo, el marco teórico se construye sobre tres pilares que juntos proporcionan una comprensión de la problemática estudiada. Primeramente, se examina la aplicación de la teoría de portafolio moderna de Markowitz. En segundo lugar, se profundiza en la naturaleza y características de los Exchange-Traded Funds (ETFs), instrumentos financieros que han ganado relevancia en el panorama de inversión global. Finalmente, se investiga cómo la inteligencia artificial se aplica en el ámbito financiero, enfocándose en la capacidad predictiva y las herramientas de análisis que la IA pone a disposición de los inversores para mejorar la gestión y pronóstico de sus inversiones en ETFs.

#### **3.1. Modelo de valoración de activos de capital (CAPM) y Teoría del portafolio moderna**

En 1964, William F. Sharpe introduce el Modelo de Valoración de Activos de Capital (CAPM) el cual explica la relación entre el riesgo sistemático y el rendimiento esperado de los activos. El CAPM postula que el rendimiento esperado de un activo es igual a la tasa de rendimiento libre de riesgo más una prima de riesgo, que es el producto del coeficiente beta del activo y la prima de riesgo del mercado. Sharpe (1964) determina que:

los precios de los activos de capital en el mercado se ajustan para reflejar tanto el precio del tiempo (la tasa de interés pura) como el precio del riesgo (el retorno esperado adicional por unidad de riesgo soportado). Este equilibrio asegura que, mediante la diversificación, los inversores puedan alcanzar cualquier punto deseado a lo largo de una línea de mercado de capitales, obteniendo retornos esperados más altos solo al asumir riesgos adicionales. (pp. 425-442)

Así mismo, este modelo es esencial para la construcción de un portafolio eficiente, ya que permite a los inversores estimar el retorno esperado de un activo en función de su nivel de riesgo no diversificable. En la evaluación y optimización de pronósticos para ETFs, el CAPM sirve como una herramienta teórica para fundamentar estrategias de inversión que buscan equilibrar la expectativa de ganancias con la exposición al riesgo de mercado.

Por lo tanto, la conexión entre la teoría del CAPM, la diversificación y los ETFs radica en el principio de que los ETFs, como vehículos de inversión, ofrecen una forma eficiente de lograr la diversificación recomendada por Sharpe. Al seguir índices o sectores específicos, los ETFs permiten a los inversores exponerse a una amplia variedad de activos, reduciendo así el riesgo no sistemático de sus carteras. Esta diversificación alinea los portafolios de inversión con la idea del CAPM de que los inversores pueden mejorar su ajuste riesgo-rendimiento mediante la adquisición de activos que, en conjunto, ofrecen un equilibrio óptimo entre el riesgo sistemático y los retornos esperados.

Por otro lado, la teoría de selección de portafolio de Harry Markowitz (1952) introdujo el concepto de diversificación como una estrategia para minimizar el riesgo sin sacrificar el rendimiento esperado. Su teoría propone que los inversores pueden construir una cartera óptima de activos, equilibrando el retorno esperado contra el riesgo, medido como la varianza de los retornos de la cartera. Este enfoque revolucionario enfatiza la importancia de considerar tanto el rendimiento como el riesgo al tomar decisiones de inversión, y subraya cómo la diversificación efectiva puede reducir el riesgo total de una cartera.

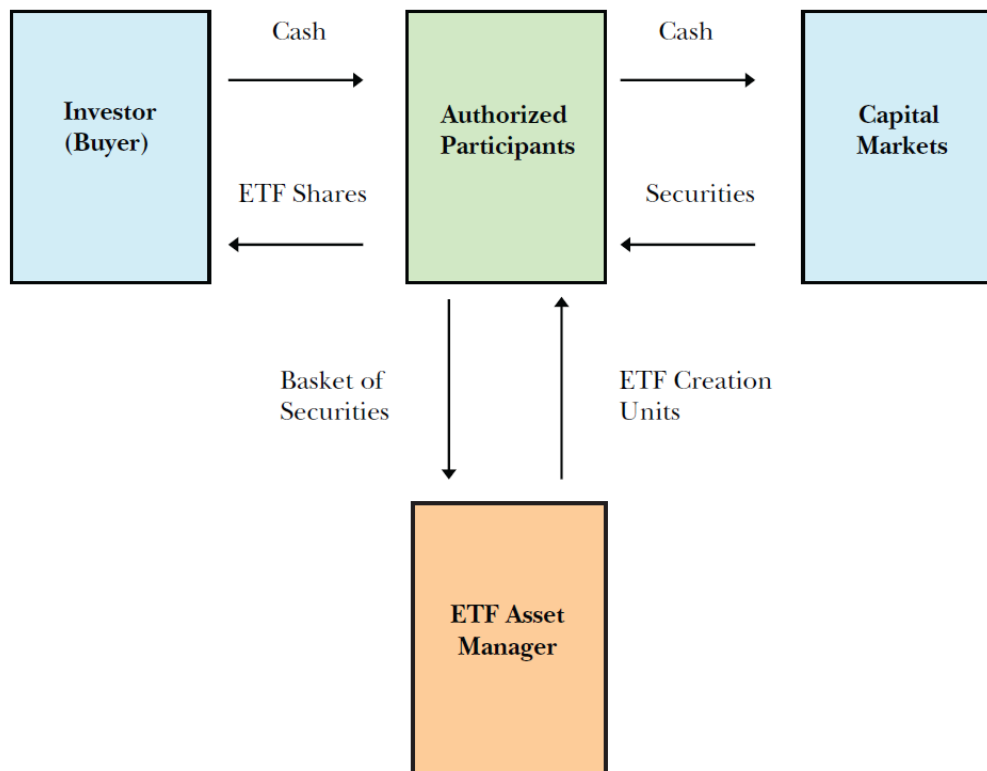
Estas dos teorías son fundamentales para este proyecto por la capacidad para proporcionar un marco robusto para la gestión de inversiones en ETFs. La aplicación del Modelo de Valoración de Activos de Capital (CAPM) de Sharpe y la Teoría de Selección de Carteras de Markowitz ofrece una base teórica sólida para entender cómo la diversificación y el equilibrio entre riesgo y retorno pueden optimizar las carteras de inversión. Mientras que el CAPM ayuda a identificar la relación entre el riesgo de mercado y el retorno esperado de los activos, la teoría de Markowitz presenta una metodología para construir una cartera óptima que minimice el riesgo para un nivel de retorno esperado dado. Juntas, estas teorías justifican la incorporación de ETFs como instrumentos clave en la construcción de carteras diversificadas.

### **3.2. Fondos cotizados en bolsa o Exchange-Traded Funds (ETF's)**

Los Exchange-Traded Funds (ETFs) constituyen una clase de instrumentos financieros que han revolucionado el panorama de la inversión desde su introducción. Estos fondos combinan características de los fondos mutuos y las acciones “ofreciendo una variedad de clases de activos (por ejemplo acciones, renta fija, materias primas y monedas) estrategias (por ejemplo, índices pasivos, basados en modelos y activos) y regiones” (Lettau & Madhavan, 2018), permitiendo a los inversores comprar y vender participaciones en una cartera diversificada de activos a través de transacciones en bolsas de valores, similar a las acciones individuales. La naturaleza única de los ETFs les permite ofrecer una liquidez significativa, transparencia en sus activos subyacentes y eficiencia en costos, lo que los hace atractivos tanto para inversores individuales como institucionales.

Estos instrumentos tienen una arquitectura que combina elementos de los fondos mutuos y las acciones, ofreciendo una estructura única que les permite ser negociados en bolsas de valores como las acciones individuales. Así mismo, rastrean índices, sectores, commodities o cestas de activos, replicando su rendimiento a través de la posesión directa de los activos subyacentes o mediante el uso de derivados. Esta construcción permite a los inversores diversificar sus carteras con una sola transacción, proporcionando liquidez, transparencia en la valoración de activos y eficiencia en costos. Además, la creación y redención de acciones de ETFs se realizan a través de entidades autorizadas, lo que asegura que el precio del ETF se mantenga cercano al valor neto de los activos (NAV) que representa.

## The ETF Architecture



*Figura 1. Lettau, M., & Madhavan, A. (2018). Exchange-Traded Funds 101 for Economists. Journal of Economic Perspectives, 32(1), 136.*

Por otro lado, estos fondos representan un instrumento financiero de inversión que está en constante crecimiento, "Los fondos han crecido sustancialmente en diversidad y tamaño en los últimos años junto con el auge de la inversión en índices pasivos. Las acciones aún representan más del 78 por ciento de los activos bajo gestión en ETFs a partir de 2017 (pero hay un crecimiento rápido en todas las clases de activos y en particular en renta fija con activos ahora en exceso de \$740 mil millones o el 17 por ciento de los activos" (Lettau & Madhavan, 2018, pp. 396-399). Por ende, encontrar nuevas formas de optimizarlos es un tema relevante en la actualidad de las finanzas, lo que nos lleva a la integración de nuevas tecnologías como la IA para evaluar diferentes estrategias que permitan crear valor a los inversionistas.

### 3.3. Inteligencia artificial

La inteligencia artificial (IA) es una rama de la informática que se enfoca en la creación de máquinas capaces de realizar tareas que, tradicionalmente, requerirían inteligencia humana. Esto incluye el aprendizaje, el razonamiento, la percepción, la comprensión del lenguaje y la creatividad. En las últimas

décadas, la IA ha experimentado avances significativos, impulsados por mejoras en el poder computacional, la disponibilidad de grandes conjuntos de datos y el desarrollo de algoritmos sofisticados. Su aplicación en el ámbito financiero, conocida como fintech, ha transformado radicalmente el sector, ofreciendo desde la automatización de procesos hasta la predicción de tendencias de mercado y la gestión personalizada de inversiones. La integración de la IA en las finanzas permite analizar vastas cantidades de información rápidamente, identificar patrones ocultos, mejorar la toma de decisiones y optimizar las estrategias de inversión, lo que abre un amplio espectro de posibilidades para aumentar la eficiencia y la rentabilidad en el mercado financiero, incluida la gestión de ETFs.

A modo de ejemplo, Uysal, A. S., Li, X., & Mulvey, J. M. (2021) destaca la integración de tareas de predicción y optimización en un enfoque de aprendizaje de extremo a extremo mediante redes neuronales, proporcionando decisiones de cartera directamente desde las características de entrada, lo que representa una metodología innovadora frente a los modelos tradicionales de dos pasos que pueden maximizar errores de predicción y llevar a decisiones subóptimas de asignación de activos. Esto refuerza la propuesta del uso de la inteligencia artificial para mejorar el rendimiento de los portafolios de inversión, encontrando oportunidades nuevas como manifiestan Oliveira, A. S., Ceretta, P. S., & Albrecht, P.:

“Una de las alternativas a los portafolios tradicionales, MF-DFA, fue confirmada para detectar autocorrelaciones a largo plazo, multifractalidad, asimetría, persistencia y dependencia de memoria. La efectividad de la técnica MF-DFA nos motivó a comparar su rendimiento con el enfoque de las redes neuronales artificiales (ANNs), que tienen la ventaja de poder manejar grandes cantidades de datos y relaciones complejas entre variables con alta precisión" (2023).

En este orden de ideas, existen métodos como el de KNN que es un algoritmo de aprendizaje supervisado simple y efectivo que se utiliza para clasificación y regresión. Funciona identificando los 'k' vecinos más cercanos a un punto de datos no clasificado, basándose en una medida de distancia como la euclidiana, y asignando la etiqueta o valor más común entre estos vecinos como la predicción para el punto. En clasificación, esto significa asignar la clase predominante entre los vecinos; en regresión, se asigna el promedio de los valores. El método K-Nearest Neighbors (KNN) puede aplicarse en la predicción de precios de activos financieros, como los ETFs, utilizando características históricas del mercado para pronosticar futuras fluctuaciones de precios.

Entonces, surge la cuestión de si realmente estos métodos son más precisos que los métodos tradicionales, por lo que diferentes autores como Kamalov, F., Smail, L., & Gurrib, I. (2021) cuya investigación consistía en predecir el valor del próximo día del índice S&P 500 basándose en información de los 14 días de negociación anteriores. Compararon ocho diferentes modelos de redes neuronales basados en capas

conectadas, convulsiónales y recurrentes, revelando que una red neuronal recurrente de una sola capa con el optimizador RMSprop produce resultados óptimos con un Error Absoluto Medio (MAE) de validación y prueba de 0.0150 y 0.0148, respectivamente. Confirmando la hipótesis de que este nuevo tipo de estrategias predictivas a pesar de ser tan complejas son más precisas.

## **4. Revisión de la literatura**

### **4.1 Revisión**

La literatura existente revela un creciente interés en la intersección de la Inteligencia Artificial (IA) y la teoría de portafolios tradicional. Dicho esto, las redes neuronales, las máquinas de vectores de soporte y los algoritmos de aprendizaje profundo son solo algunas de las aplicaciones de IA que han sido exploradas para mejorar la selección de portafolios y la predicción de precios de activos (Zhang, et al., 2019; Huang, et al., 2020).

En primera instancia, durante la revisión de la literatura, se encontró que la IA tiene una tendencia en volver más eficientes los procesos. Un estudio comparativo entre un enfoque lineal y uno de red neuronal demostró la superioridad de este último, con una tasa de clasificación correcta del 67% frente al 53% del modelo lineal, destacando su capacidad para capturar comportamientos no lineales en los datos financieros (Daniels & Kamp, 1999). Esta eficiencia en la identificación de patrones complejos en los datos de mercado es crucial para la predicción de precios y la selección de portafolios.

Por otro lado, se encontró que las técnicas de IA han evolucionado significativamente la teoría tradicional de portafolios, proporcionando herramientas para capturar y procesar grandes cantidades de datos no estructurados, lo que a menudo queda fuera del alcance de los métodos convencionales. Por ejemplo, se ha demostrado que las redes neuronales y las técnicas de aprendizaje profundo son capaces de identificar patrones complejos y dinámicos en los datos de mercado que pueden predecir con mayor precisión los movimientos de precios (Deng, et al., 2017).

Sin embargo, una premisa emergente en la literatura que se utilizó es la eficacia superior de los portafolios desarrollados con IA en la predicción de precios en activos financieros. Para ilustrar, en una investigación comparativa de técnicas multifractales frente a redes neuronales artificiales, se encontró que las redes neuronales proporcionaban señales de inversión más precisas, desafiando la Hipótesis del Mercado Eficiente (EMH) (Kumar & Thenmozhi, 2006). Además, estudios como el de Dixon, et al., (2020) han enfatizado la capacidad de las técnicas de IA para superar a las carteras tradicionales basadas en el modelo de Markowitz en términos de rendimiento ajustado al riesgo. En una línea similar, una investigación

reciente sobre la optimización de portafolios utilizando redes neuronales reportó un rendimiento superior en las pruebas fuera de muestra entre 2017 y 2021, logrando un ratio de Sharpe de 1.16, superando tanto a la paridad de riesgo nominal (0.79) como a la mezcla de igual peso (0.83). Este estudio también destacó la implementación de un mecanismo de selección de activos en la red neuronal, que mejoró aún más el rendimiento, elevando el ratio de Sharpe a 1.24 (Uysal et al., 2021).

Además, estudios como el que comparó el rendimiento de los portafolios tradicionales del S&P500, Markowitz y Sharpe con portafolios basados en técnicas multifractales (MF-DFA) y redes neuronales con clasificación de probabilidad asimétrica de Student (ANN-t), han revelado resultados interesantes. Este estudio utilizó datos diarios de acciones del S&P500 entre enero de 2018 y julio de 2022 para probar métricas de rendimiento y riesgo, como la volatilidad anual, el Valor en Riesgo, el Ratio de Sharpe, el Ratio de Sortino, Beta y el Alfa de Jensen. Los resultados confirmaron que la técnica ANN-t podría indicar mejores entradas de inversión, lo que contradice la Hipótesis del Mercado Eficiente. (Oliveira, Ceretta & Albrecht, 2023)

Además, es importante considerar los costos de transacción en la optimización de carteras. Un estudio que evaluó diferentes modelos de Deep Learning y Machine Learning en el contexto de inversión diaria en operaciones bursátiles reveló insights significativos. El estudio no solo comparó el rendimiento de los modelos sin considerar los costos de transacción, sino que también evaluó su eficacia cuando se incluyeron estos costos, proporcionando una visión más realista de su aplicación en entornos de mercado prácticos. (Ma, Han & Wang, 2021)

Tras considerar los costos de transacción en la optimización de carteras, es esencial destacar también los avances en la precisión de la predicción de precios de acciones mediante el uso de IA. Un estudio que analizó una variedad de arquitecturas de aprendizaje profundo para la predicción de series temporales mostró resultados notables. En este estudio, se predijo el valor del próximo día del índice S&P 500 basándose en información de los 14 días comerciales anteriores, comparando ocho modelos diferentes de redes neuronales. Los experimentos numéricos revelaron que una red neuronal recurrente de una sola capa con optimizador RMSprop produjo los mejores resultados, con un MAE de validación y prueba de 0.0150 y 0.0148 respectivamente. (Kamalov, Smail & Gurrib, 2021)

En conjunto, la literatura analizada sugiere que, aunque el CAPM y la EMH ofrecen una comprensión fundamental de la valoración de activos y la formación de precios y la IA representan vanguardias prometedoras que pueden complementar y mejorar estos modelos tradicionales. La capacidad de la IA para procesar y aprender de grandes volúmenes de datos puede ofrecer una ventaja considerable en la predicción

de los precios del mercado, como se demostró en un estudio que investigó el rendimiento de modelos de redes neuronales (MLP, CNN y LSTM) en la predicción de cambios significativos en los precios de las acciones usando datos diarios de 10 años de cuatro grandes empresas estadounidenses. Los resultados indicaron que estos modelos son capaces de predecir cambios significativos en el precio de los activos con un alto grado de precisión, especialmente para cambios más notables (Kamalov, 2020). Sin embargo, es importante reconocer que incluso las redes neuronales simples pueden enfrentar desafíos al intentar desafiar la Hipótesis del Mercado Eficiente, como se vio en un estudio sobre los rendimientos diarios de las acciones de IBM, donde se observó que obtener evidencia en contra de mercados eficientes con una red simple no es tarea fácil. (Burrell & Folarin, 1997).

## **4.2 Marco conceptual**

En el contexto financiero los “*Exchange trust funds - ETF*” o fondos cotizados en bolsa son “Un vehículo de inversión, que típicamente busca seguir el rendimiento de un índice específico, como lo hace un fondo mutuo indexado, estos difieren en que los ETF no interactúan directamente con los mercados de capital” Lettau, M., & Madhavan, A. (2018)

La inteligencia artificial permite “crear máquinas de propósito especial que realizan tareas cognitivas. El aprendizaje automático facilita que los algoritmos aprendan de los datos y desarrollen soluciones a problemas, utilizados para analizar datos, predecir tendencias, automatizar tareas y actuar como el "cerebro”.” Technology & National security (2018)

El aprendizaje automático es “Un enfoque poderoso para generar comportamiento inteligente, permitiendo a las máquinas ajustar su comportamiento para optimizar el rendimiento. Utiliza datos etiquetados en el aprendizaje supervisado, para aprender a distinguir entre categorías.” Technology & National security (2018)

En cuanto al aspecto más importante del trabajo de Markowitz fue demostrar que “el riesgo de una inversión no radica en el riesgo propio de cada valor, sino en la contribución que dicho valor aporta a la varianza del portafolio completo. Esto se debe principalmente a su covarianza con los demás valores en el portafolio.” Rubinstein (2002)

El termino Fintech procede de las palabras en ingles *Finance and technology* y “hace referencia a todas aquellas actividades que impliquen el empleo de la innovación y los desarrollos tecnológicos para el diseño, oferta y prestación de productos y servicios financieros.” CNMV (2021)

Por otro lado, el aprendizaje supervisado es un tipo de aprendizaje automático en el que se enseña a los modelos a hacer predicciones o decisiones basándose en ejemplos de entrada-salida previamente

etiquetados. El objetivo es que, tras entrenarse en un conjunto de datos donde cada ejemplo está asociado con una etiqueta o un resultado correcto, el modelo pueda identificar patrones y relaciones en los datos, permitiéndole hacer predicciones precisas sobre datos nuevos o no vistos. Este enfoque se utiliza ampliamente para tareas de clasificación y regresión, donde las entradas deben ser mapeadas a categorías discretas o valores continuos, respectivamente.

La optimización de portafolios se refiere al proceso de seleccionar la mejor combinación de activos para maximizar el retorno esperado para un nivel de riesgo dado, o minimizar el riesgo para un nivel de retorno esperado. Este concepto se basa en la teoría moderna de portafolios y es fundamental para el diseño y gestión de portafolios de inversión, especialmente en el contexto del uso de inteligencia artificial para mejorar las decisiones de inversión.

## **5. Objetivos**

### **5.1 Pregunta de investigación:**

¿Cómo pueden los métodos de aprendizaje supervisado y la teoría de portafolio de Markowitz ser integrados eficazmente para minimizar el riesgo y maximizar los rendimientos en la gestión de inversiones en ETFs, específicamente en los ETFs SPY, QQQ, IWM, GBTC y VTI?

### **5.2 Objetivo general**

Minimizar el Riesgo en la Gestión de Inversiones en ETFs mediante la implementación de métodos de Aprendizaje Supervisado y la Teoría de Portafolio de Markowitz.

### **5.3 Objetivos específicos**

1. Realizar un Análisis Estadístico Descriptivo de las Tendencias Históricas en IWM, QQQ, SPY, GBTC y VTI, para identificar patrones y comportamientos.
2. Aplicar el Método KNN para Clasificar con Precisión las Tendencias "Up" o "Down" en los ETFs, evaluando el porcentaje de exactitud.
3. Implementar la Teoría de Portafolio de Markowitz para Construir y Analizar un Portafolio Óptimo que maximice los rendimientos y minimice el riesgo a través de la combinación de IWM, QQQ, SPY, GBTC y VTI.

## 6. Metodología

Para minimizar el riesgo y maximizar los rendimientos en las inversiones Markowitz sugiere “Para reducir significativamente el riesgo, es necesario combinar títulos que no estén perfectamente correlacionados” (Markowitz, 1952)”. Esta investigación seguirá un enfoque metodológico que combina el análisis cuantitativo y la modelización, se utilizará R para implementar una metodología enfocada en la teoría de Markowitz y el aprendizaje supervisado para optimizar portafolios de ETFs. Inicialmente, se realizará una revisión de la literatura para establecer un marco teórico sólido que vincule el aprendizaje automático, la teoría financiera y la gestión de inversiones en ETFs.

### 6.1 Método KNN

Se utilizará el método KNN (K-Nearest Neighbors) para predecir el comportamiento de los ETFs. El método KNN es un algoritmo de aprendizaje supervisado que clasifica nuevas observaciones basándose en la similitud de las características con ejemplos etiquetados en el conjunto de datos de entrenamiento. Este método opera bajo el principio de que la similitud entre los ejemplos puede ser calculada a través de una distancia métrica, como la distancia euclidiana, Manhattan o Minkowski. La elección de la métrica depende de la naturaleza de los datos y del contexto específico de la aplicación. En esta investigación, la distancia euclidiana será usada por su simplicidad y eficacia.

De esta misma forma, Para clasificar una nueva observación, KNN identifica los K ejemplos más cercanos (vecinos) en el conjunto de datos de entrenamiento. La etiqueta de la nueva observación se determina por mayoría de votos entre sus vecinos: la clase que más se repite entre los K vecinos es asignada a la nueva observación. En el caso de regresión, el valor de la nueva observación se calcula como el promedio o la mediana de los valores de los vecinos más cercanos.

Este método tiene dos parámetros claves: Numero de vecinos (afectando la precisión y el rendimiento del modelo) y Métrica de distancia (Como se miden las similitudes entre observaciones)

Para comprender mejor el método K-Nearest Neighbors (KNN), es esencial destacar que este algoritmo utiliza cálculos de distancias para identificar los K vecinos más cercanos a una nueva muestra y, basándose en estos, realiza predicciones. La similitud entre los puntos en un espacio de características se calcula empleando la distancia euclidiana, que se define matemáticamente como:

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^n (\mathbf{p}_i - \mathbf{q}_i)^2} \quad (1)$$

Donde  $p, q$  representan dos puntos en un espacio de  $n$  dimensiones y  $p_i$  y  $q_i$  son los valores de la  $i$ -ésima característica de  $p$  y  $q$ , respectivamente. Esta fórmula (1) permite medir cuán cerca están dos puntos, facilitando así la tarea de clasificación o regresión dentro del conjunto de datos.

Posteriormente se incorporará el paquete CARET de R para optimizar la selección y gestión de portafolios financieros. CARET facilita una serie de procedimientos de aprendizaje automático como la clasificación y la regresión, proveyendo un marco cohesivo para la validación cruzada y la selección de modelos. Se utilizará CARET por su capacidad para simplificar la construcción y evaluación de modelos predictivos complejos. Esto es particularmente relevante en el análisis financiero donde la precisión y la capacidad de manejar múltiples variables son cruciales. La metodología se centrará en el uso práctico de CARET para analizar tendencias de mercado y predecir el rendimiento de ETFs específicos, evaluando la precisión de estos modelos a través de la matriz de clasificación. También, se construirá y se entrenará un modelo predictivo mediante el aprendizaje supervisado que permitirá la clasificación de cinco series temporales distintas constituidas por los datos históricos de los precios de los 5 ETFs desde el 1 de enero del 2020 hasta el 29 de diciembre del 2023.

Consecutivamente se entrena el modelo en R mediante la función “train” de la biblioteca “caret”, especificando que todas las columnas, excepto “direction”, se usen como predictores. Esta columna “direction” sirve como la variable objetivo a predecir. Un aspecto importante de este entrenamiento es el ajuste del parámetro  $K$ , para lo cual se establece “tuneLength = 50”. Esto instruye a “caret” para evaluar 50 valores distintos de  $K$ , optimizando así el modelo mediante la validación cruzada, donde se divide el conjunto de entrenamiento en subconjuntos para evaluar la eficacia del modelo en datos no vistos, asegurando una generalización.

Finalmente, se grafican los resultados de la validación para analizar cómo varía el rendimiento del modelo con diferentes configuraciones de  $K$ . Este gráfico sirve para identificar el número óptimo de vecinos y para comprender cómo afectan los distintos valores de  $K$  a la precisión del modelo, ayudando a evitar tanto el sobreajuste como el subajuste.

## **6.2 Proceso de optimización de Hiperparámetros**

El proceso de optimización de hiperparámetros es fundamental para mejorar el rendimiento de los modelos de aprendizaje automático implementados en esta investigación. Así, los hiperparámetros son parámetros que se establecen antes de que el proceso de aprendizaje comience y que no se ajustan durante el entrenamiento del modelo. Estos parámetros influyen significativamente en el comportamiento y la

precisión del modelo, y su correcta configuración puede marcar la diferencia entre un modelo subóptimo y uno altamente efectivo.

En el caso del método K-Nearest Neighbors (KNN), los hiperparámetros clave que se deben optimizar son el número de vecinos ( $k$ ) y la métrica de distancia utilizada. En ese orden de ideas, el número de vecinos ( $k$ ) define cuántos puntos de datos cercanos se considerarán al hacer una predicción para un nuevo punto, mientras que la métrica de distancia determina cómo se calcula la proximidad entre puntos de datos en el espacio de características.

Para optimizar dichos hiperparámetros en esta investigación, se utilizó la técnica de búsqueda en cuadrícula (Grid Search). Esta técnica implica probar todas las combinaciones posibles de un conjunto predefinido de valores de hiperparámetros para encontrar la combinación que produce el mejor rendimiento del modelo. Posteriormente el rendimiento del modelo se evaluó utilizando validación cruzada, específicamente una validación cruzada de 10-fold, que divide los datos en 10 subconjuntos, entrenando el modelo en 9 de ellos y validándolo en el restante, repitiendo este proceso 10 veces.

Durante la búsqueda en cuadrícula, se evaluaron múltiples valores para el número de vecinos ( $k$ ), que van desde 1 hasta 100, y se probaron diferentes métricas de distancia como la distancia euclidiana y la distancia Manhattan. Así mismo, el rendimiento del modelo para cada combinación de hiperparámetros se midió en términos de precisión en la clasificación de tendencias de precios de los ETFs. Este proceso se llevó a cabo de manera independiente para cada uno de los ETFs estudiados: SPY, QQQ, IWM, GBTC y VTI.

Los resultados de la optimización de hiperparámetros mostraron variabilidad en los valores óptimos de  $k$  para los diferentes ETFs, lo que refleja las diferencias en las dinámicas de mercado subyacentes de cada uno. Por ejemplo, se encontró que el valor óptimo de  $k$  para el ETF IWM era 5, mientras que para el ETF VTI era 83. Esta variabilidad destaca la importancia de ajustar cuidadosamente los hiperparámetros específicos de cada modelo para maximizar su rendimiento predictivo.

Una vez determinados los valores óptimos de los hiperparámetros, estos se utilizaron para entrenar los modelos finales. Los valores de los hiperparámetros óptimos se documentaron detalladamente y se incluyen en los anexos de este documento para referencia futura y reproducibilidad del estudio.

### **6.3 Elección del método y limitaciones**

La elección del método  $k$ -NN ( $k$ -Nearest Neighbors) para explicar la categorización del valor de las acciones (es decir, si suben o bajan) se fundamenta en su simplicidad y efectividad en tareas de clasificación

exploratoria. El enfoque principal de esta investigación se orienta hacia la conformación del portafolio, y el uso de k-NN sirve como una herramienta exploratoria para clasificar las acciones, proporcionando una base intuitiva para entender los movimientos del mercado. Así mismo, en esta tesis la comparación con otros enfoques de aprendizaje automático y estadístico no se realizó por varias razones. En primer lugar, el objetivo principal de esta investigación se centra en la conformación del portafolio, no en la evaluación exhaustiva de métodos de clasificación, dicha comparación con otros métodos desviaría el enfoque principal de esta investigación. En segundo lugar, el uso de k-NN se plantea de manera exploratoria, para proporcionar una visión inicial de la clasificación de las acciones. La intención no es demostrar que k-NN es superior, sino utilizarlo como una herramienta inicial para esta tarea específica.

Sin embargo, es importante mencionar que la literatura existente ya ha explorado diversos enfoques de aprendizaje automático y estadístico en el contexto de la predicción de precios de acciones y la gestión de portafolios. Por ejemplo, investigaciones como la de Kamalov (2020) han demostrado cómo los modelos de redes neuronales, incluidos MLP, CNN y LSTM, pueden aplicarse efectivamente para prever cambios significativos en los precios de las acciones. Estos estudios destacan la capacidad superior de estos modelos en comparación con los métodos tradicionales. A pesar de estos avances, la implementación de k-NN en esta investigación se justifica por su simplicidad y aplicabilidad inmediata en el contexto exploratorio inicial de la categorización de acciones.

Por otro lado, se decidió trabajar con un modelo de clasificación en lugar de uno de regresión porque predecir directamente el valor de las acciones es considerablemente más complejo debido a la alta volatilidad y la influencia de múltiples factores externos en el mercado financiero. La clasificación de las acciones para determinar si su valor sube o baja es una tarea más manejable y robusta. Este enfoque simplifica el análisis y proporciona información práctica y directa para la toma de decisiones en la conformación del portafolio, ya que identificar tendencias generales del mercado (alza o baja) es más útil y aplicable en la práctica de inversión que predecir valores exactos, los cuales pueden estar sujetos a variaciones y fluctuaciones impredecibles. La volatilidad del mercado financiero y la cantidad de variables que influyen en los precios de las acciones hacen que los modelos de regresión sean más susceptibles a errores significativos. En cambio, un modelo de clasificación puede proporcionar una visión más clara y estable de las tendencias del mercado, lo cual es fundamental para la construcción y gestión de portafolios. Además, el uso de un enfoque de clasificación permite a los inversores centrarse en las decisiones estratégicas sobre la dirección del mercado, que es una información más accionable en comparación con las predicciones exactas de los precios.

Finalmente, aunque la comparación con otros métodos podría enriquecer el análisis, se decidió mantener el enfoque en el uso del método k-NN para no desviar la atención del objetivo principal de la investigación, que es la conformación del portafolio. Los estudios citados demuestran que otros autores ya han explorado enfoques alternativos y, en muchos casos, no se encontraron resultados consistentemente superiores a los obtenidos con k-NN para tareas exploratorias específicas.

## 6.4 Modelo de Markowitz

Terminada la aplicación y ejecución del punto anterior, se continua con la conformación de un portafolio siguiendo la teoría de Markowitz. Esta teoría publicada en 1952 por Harry Markowitz introduce un enfoque cuantitativo para la inversión en portafolios que busca maximizar el retorno esperado para un determinado nivel de riesgo. Esta teoría es conocida como la teoría moderna del portafolio y se basa en el concepto de diversificación eficiente de activos.

La diversificación consiste en la practica de repartir las inversiones en diferentes activos financieros, y tiene como objetivo la reducción del riesgo global de un portafolio, o como lo define Markowitz:

"Al diversificar, el inversor puede reducir su exposición al riesgo individual de cualquier inversión. Este proceso de diversificación puede reducir la varianza del portafolio de inversiones, ya que no todos los activos se verán afectados de la misma manera por eventos económicos o del mercado" (Markowitz, 1952)"

Con lo que en un portafolio optimo, se espera "La menor varianza de retorno para un nivel dado de retorno esperado o el mayor retorno esperado para un nivel dado de riesgo" (1952) cuya relación se representa con una curva que se denomina la frontera eficiente de un portafolio, donde cada punto en esta curva representa un portafolio que ofrece el máximo retorno esperado para un nivel específico de riesgo.

Por otro lado, esta el riesgo y el retorno esperado. El riesgo se cuantifica típicamente como la varianza o la desviación estándar del portafolio. El retorno esperado se calcula como un promedio ponderado de los retornos esperados de los activos individuales, donde las ponderaciones reflejan la proporción del total del portafolio invertido en cada activo. La formula para el calculo del retorno esperado es la siguiente:

$$R_e = \bar{R}_i = \sum_{i=1}^n \omega_i R_i \quad (2)$$

Donde  $R_i$  es el retorno esperado del activo  $i$ , y  $\omega_i$  es la proporción del total del portafolio invertido en el activo  $i$ .

De esta forma, el retorno esperado del portafolio sería:

$$R_{Portafolio} = \omega_x \bar{R}_x + \omega_y \bar{R}_y \dots \omega_i \bar{R}_i \quad (3)$$

Siendo  $\omega_i$  el peso del activo  $i$  en el portafolio y  $\bar{R}_i$  la rentabilidad esperada del activo  $i$

Por otro lado, el riesgo del portafolio, expresado como la varianza, se calcula utilizando la formula:

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (R_t - \bar{R})^2 \quad (4)$$

Y la desviación estándar del portafolio como:

$$\sigma = \sqrt{(\sigma^2)} \quad (5)$$

El software utilizado para la optimización y entrenamiento de los modelos fue RStudio. Las librerías principales fueron caret, utilizada para la implementación de la búsqueda en cuadrícula y la validación cruzada; PortfolioAnalytics, empleada para la optimización de portafolios según la teoría de Markowitz; y ggplot2 y dplyr, destinadas a la visualización y manipulación de datos, respectivamente. La combinación de estas herramientas permitió realizar un análisis exhaustivo y riguroso, asegurando la robustez y validez de los resultados obtenidos.

En cuanto a la optimización de portafolios, se utilizó el paquete PortfolioAnalytics en R. En primera instancia, se calculó la frontera eficiente para identificar las combinaciones de portafolios que ofrecen el mejor retorno por nivel de riesgo. Luego, se determinó el portafolio con la mínima varianza global, destacando la opción de menor riesgo. Se trazó una línea tangente desde el punto libre de riesgo hasta la frontera eficiente para identificar el portafolio de mercado. Además, se evaluó el riesgo y retorno de cada activo individualmente y se construyó un portafolio donde cada activo tenía el mismo peso, para comparar su desempeño contra portafolios más diversificados.

Asimismo, se analizaron las fronteras de dos activos para comprender las implicaciones de combinar diferentes pares de ETFs en términos de riesgo y retorno. Se implementaron simulaciones de Monte Carlo para explorar posibles combinaciones de portafolios, ayudando a visualizar el espacio de posibilidades y a identificar portafolios óptimos bajo diferentes criterios. Finalmente, se utilizó el ratio de Sharpe para evaluar la eficiencia de cada portafolio, midiendo el exceso de retorno por unidad de riesgo, facilitando así la selección del portafolio que ofrece el mejor equilibrio entre riesgo y retorno. Esta metodología permitió una comprensión profunda de cómo los métodos de aprendizaje supervisado y la teoría de Markowitz pueden integrarse para la gestión de inversiones en ETFs

## **7. Resultados**

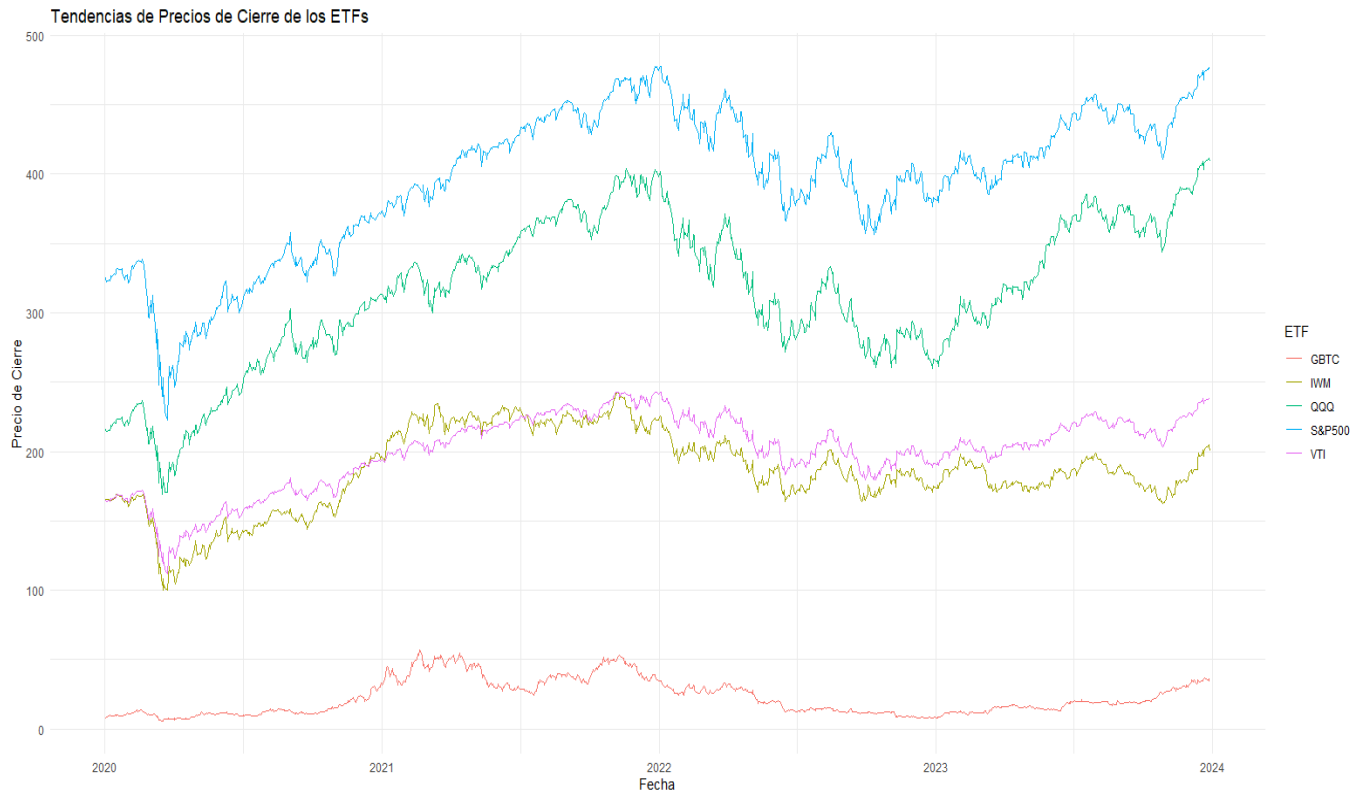
En la sección de esta tesis, abordaremos cómo los objetivos específicos han sido alcanzados mediante métodos estadísticos y de aprendizaje automático. Primero, se desglosará el análisis estadístico descriptivo realizado sobre las tendencias históricas de los ETFs IWM, QQQ, SPY, GBTC, y VTI, identificando patrones y comportamientos clave. Seguidamente, se discutirá la construcción del modelo predictivo, cuyo inicio se da con la selección y preparación de los datos, luego la aplicación y evaluación de la eficacia del método KNN para clasificar las tendencias como "Up" o "Down", enfocándose en el porcentaje de exactitud alcanzado para determinar la efectividad del modelo predictivo construido. Finalmente, detallaremos cómo se implementó la teoría de portafolio de Markowitz para construir y evaluar un portafolio óptimo, destacando cómo esta estrategia ha permitido maximizar rendimientos minimizando el riesgo.

### **7.1 Análisis estadístico descriptivo de las tendencias historias de los ETFs**

Para este estudio, se empleó RStudio, una herramienta mencionada en la metodología, que facilita la visualización de series temporales para identificar tendencias en los datos. Esta plataforma permitió una exploración detallada de las dinámicas de precios de los ETFs a lo largo del tiempo.

Por otro lado, el análisis de los precios de los ETFs reveló un patrón común de fluctuación durante el periodo estudiado. Específicamente, en 2020, marcado por la pandemia del COVID-19, se registró una notable disminución en el valor de estos activos, atribuida a la incertidumbre global en los mercados financieros. Sin embargo, con la gradual recuperación económica y la adaptación a nuevos contextos, se evidenció un incremento en los precios, demostrando la resiliencia y capacidad de adaptación del mercado financiero.

**Figura 2.** Precio de cierre de todos los ETF seleccionados desde el 01/01/2020 hasta el 31/12/2024



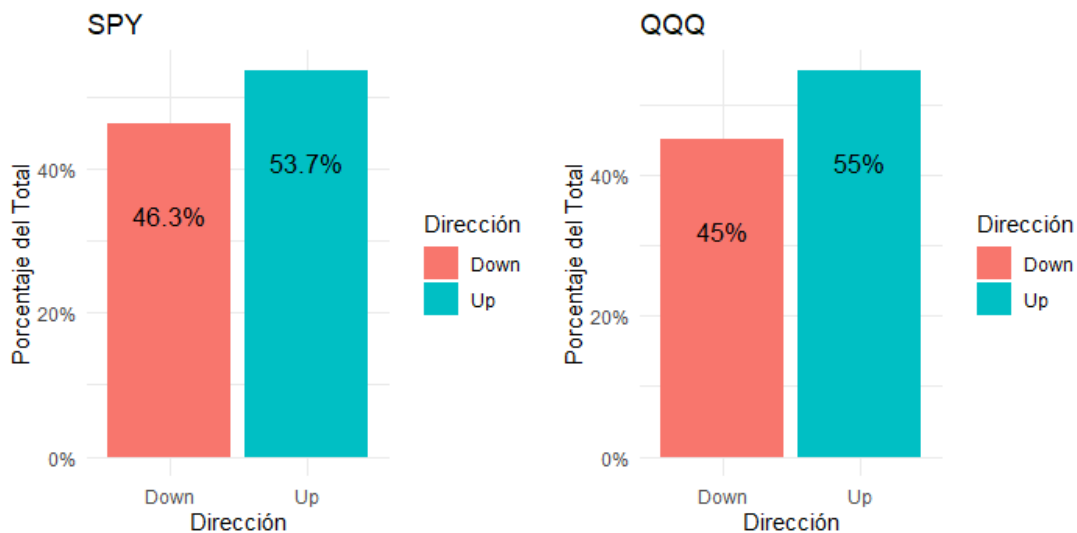
Se observó que los ETFs vinculados a acciones experimentaron caídas significativas a principios de 2020, reflejando las tendencias de los mercados financieros globales durante la pandemia. Contrariamente, el ETF GBTC, que sigue a Bitcoin, no mostró la misma tendencia descendente, manteniendo una estabilidad relativa durante ese periodo. Esto se debe a la percepción de Bitcoin como un posible activo de refugio, junto con un incremento en la adopción e interés institucional que contribuyó a su apreciación en 2021.

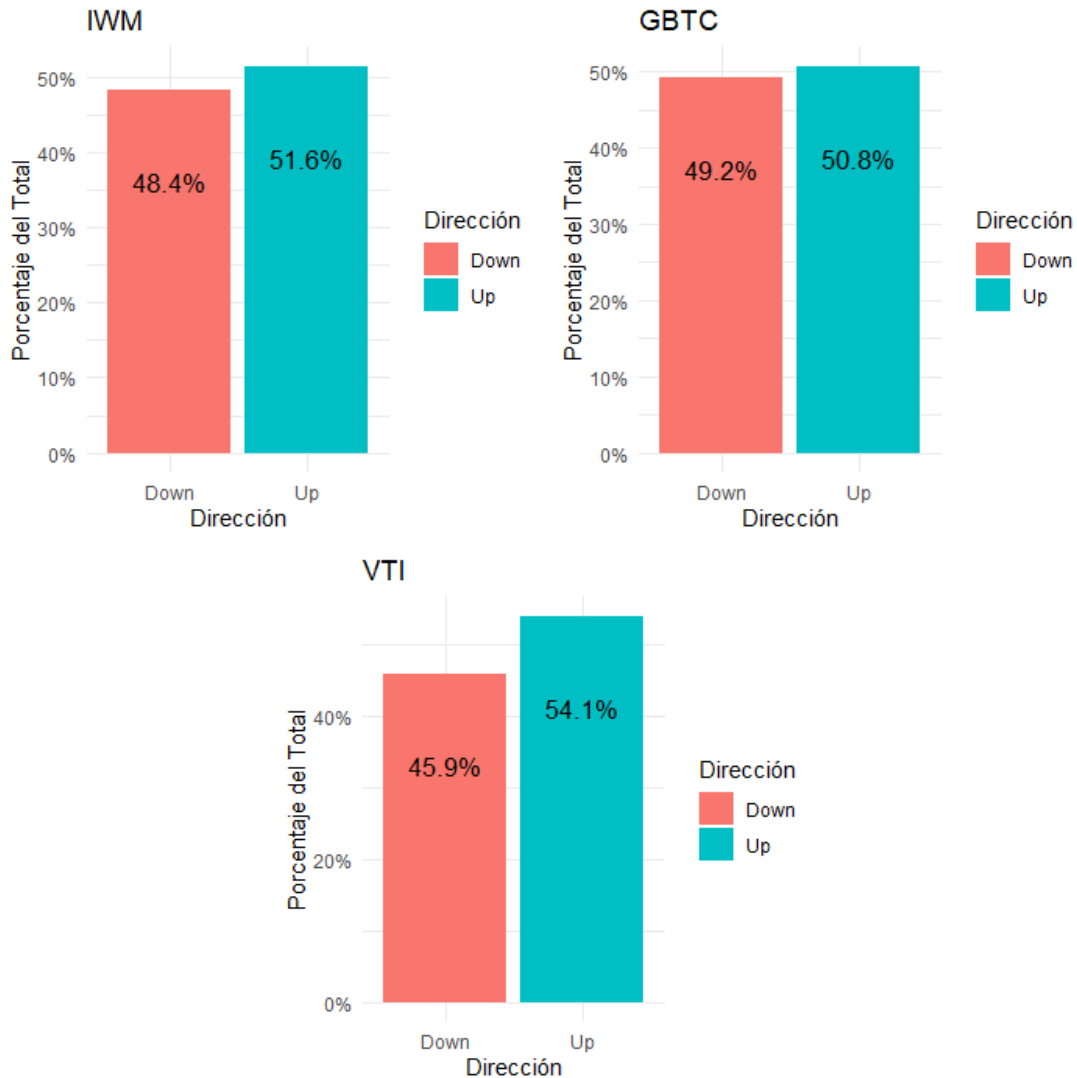
Posterior a este periodo de inestabilidad, los mercados comenzaron a recuperarse y se observó una notable tendencia ascendente en los precios de los ETFs vinculados a acciones, evidenciada a partir de finales de 2020 y durante 2021. Este ascenso refleja la recuperación económica y la confianza renovada en los

mercados financieros, impulsada por intervenciones gubernamentales, avances en vacunación y una adaptación exitosa a las nuevas condiciones operativas por parte de empresas y consumidores. Además, el notable incremento en el precio del ETF GBTC durante 2021 ilustra el creciente interés y adopción del Bitcoin como un activo de inversión legítimo, reforzando su reputación como un refugio seguro frente a la incertidumbre económica y la inflación. Este período de recuperación resalta la resiliencia del mercado y la capacidad de los activos financieros para adaptarse y prosperar incluso frente a desafíos económicos globales significativos.

Continuando con el análisis, se hizo énfasis en las tendencias temporales examinando la variabilidad porcentual del cambio en los precios, para establecer una base sólida para nuestro modelo predictivo. Este enfoque cuantitativo permitió capturar las dinámicas de subidas y bajadas en el mercado, ofreciendo una perspectiva clara sobre la volatilidad y el comportamiento general de los ETFs. Este paso inicial fue crucial para la construcción y afinación de nuestro modelo, permitiendo desarrollar un contexto sobre el comportamiento de los precios.

**Figura 3.** Porcentaje de días en que los precios subieron o bajaron





Como se indica en la figura 5, los ETFs relacionados con acciones mostraron una tendencia ascendente, superando los días de declive. Esto sugiere que la recuperación del mercado postpandemia influyó en el aumento predominante. Dicho esto, se plantea la cuestión de que la capacidad del modelo predictivo de identificar tendencias alcistas, sin datos específicos sobre el impacto de la pandemia, indicaría una alta precisión en su desempeño.

El análisis descriptivo de las tendencias históricas de los ETFs ha revelado patrones significativos y cambios en el comportamiento del mercado financiero durante el período estudiado. Se pudo visualizar y cuantificar estos movimientos, ofreciendo insights valiosos sobre la naturaleza fluctuante de los mercados. La resiliencia del mercado financiero, evidente en la recuperación postpandemia, subraya la capacidad de adaptación de los inversores y la economía en general. Esta primera parte proporciona una base sólida para

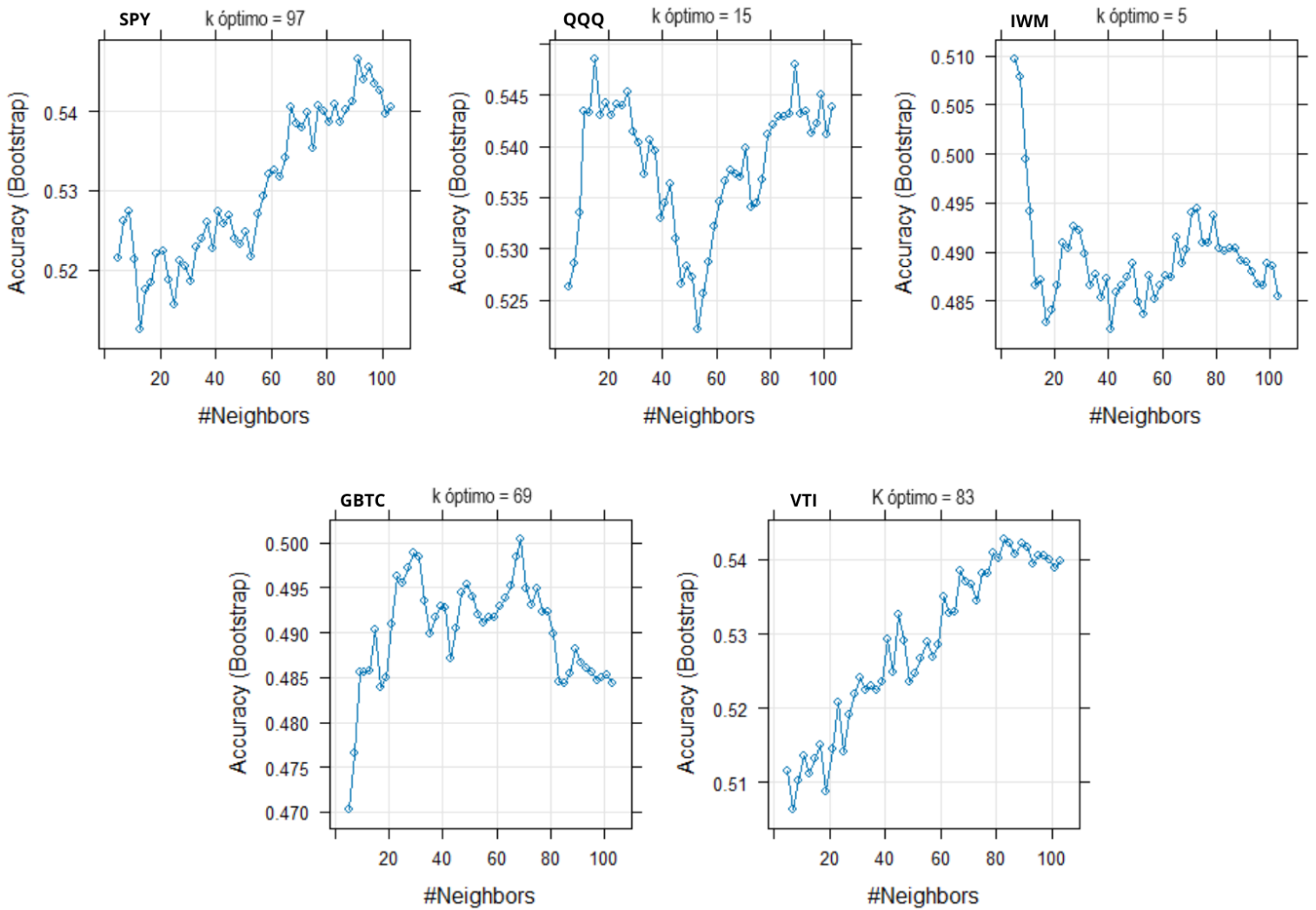
el desarrollo del modelo predictivo, y también destaca las dinámicas del mercado para tomar decisiones de inversión informadas.

## **7.2 Construcción del modelo predictivo y aplicación método KNN**

Tras concluir el análisis descriptivo del comportamiento histórico de los ETFs seleccionados, se avanzó hacia la construcción de un modelo predictivo. Inicialmente, se dividieron los datos en conjuntos de entrenamiento y prueba, asignando el 75% para el primero. Esta división se basó en índices aleatorios, preparando el terreno para el desarrollo del modelo. Durante la fase de entrenamiento, se ajustó uno de los parámetros del modelo para que este usara 50 valores distintos para el parámetro "k", y se repitió este mismo proceso con todos los ETF permitiendo identificar el "k" óptimo para cada uno de ellos, optimizando la precisión del modelo predictivo.

Como se puede evidenciar en la *figura 6* la precisión del modelo KNN para diferentes ETFs (SPY, QQQ, IWM, GBTC, VTI) en función del número de vecinos (k) varía significativamente. Por ejemplo, IWM alcanza su máxima precisión con solo 5 vecinos, mientras que VTI requiere 83 vecinos para optimizar su precisión. La variabilidad en los k óptimos sugiere diferencias en las dinámicas de mercado subyacentes de cada ETF. Este análisis revela la importancia de ajustar cuidadosamente el modelo KNN a las características específicas de cada serie temporal de precios, destacando cómo la elección de k influye directamente en la capacidad del modelo para predecir tendencias de manera precisa.

**Figura 4.** Valores de los K óptimos para los ETF



Una vez entrenado, el modelo predictivo se sometió a pruebas para evaluar su capacidad de predecir si los precios de los ETFs subirían o bajarían. Este proceso implicó calcular la probabilidad estimada por el modelo para cada dirección del mercado. Completado el testeo, el modelo quedó listo para ser implementado en la práctica, permitiéndonos comenzar a observar y analizar sus predicciones en tiempo real y evaluar su eficacia en la anticipación de movimientos del mercado.

**Figura 5.** Resultados matrices de confusión

<b>SPY</b>	Referencia	
Predicción	Baja	Sube
Baja	34	29
Sube	82	106

<b>QQQ</b>	Referencia	
Predicción	Baja	Sube
Baja	46	46
Sube	67	92

<b>IWM</b>	Referencia	
Predicción	Baja	Sube
Baja	57	58
Sube	64	71

<b>GBTC</b>	Referencia	
Predicción	Baja	Sube
Baja	49	50
Sube	74	77

<b>VTI</b>	Referencia	
Predicción	Baja	Sube
Baja	14	19
Sube	101	117

Después de correr el modelo se pudieron observar las matrices de confusión, estas muestran los resultados de predicción para los ETFs SPY, QQQ, IWM, GBTC, y VTI usando el modelo KNN. La diagonal principal de cada matriz indico los aciertos: predicciones correctas de bajadas (Baja-Baja) y subidas (Sube-Sube). Las cifras fuera de esta diagonal representan errores: predicciones de subida cuando hubo bajada y viceversa. Por ejemplo, para SPY, se predijo correctamente 34 bajadas y 106 subidas, pero se erró en 82 y 29 ocasiones respectivamente.

La observación de las matrices de confusión reveló que el modelo KNN mostró una mayor precisión al predecir aumentos en el valor de los activos financieros estudiados. Este patrón sugiere una tendencia del modelo a captar con más efectividad las condiciones que propician un ascenso en los precios de los ETFs. Tras identificar esta característica, el análisis se enfocó en explorar más a fondo los resultados del modelo. Cada ETF presentó un comportamiento relativamente parecido, reflejado en la cantidad de predicciones correctas e incorrectas para movimientos alcistas y bajistas.

**Figura 6.** Resultados del modelo predictivo

	SPY	QQQ	IWM	GBTC	VTI
Precisión	55.78%	54.98%	51.20%	50.40%	52.19%
Kappa	0.0809	0.075	0.0215	0.0047	-0.019
Clase positiva	"Baja"	"Baja"	"Baja"	"Baja"	"Baja"
Sensibilidad	29.31%	40.71%	47.11%	39.84%	12.17%
Especificidad	78.52%	66.67%	55.04%	60.63%	86.03%
Precisión balanceada	53.91%	53.69%	51.07%	50.23%	49.10%

Siguiendo con los resultados, se observó que la precisión de los modelos para todos los ETFs superó el 50%, lo cual sugiere una potencial utilidad de estos modelos de IA en decisiones de inversión. El coeficiente Kappa, que evalúa la concordancia de las predicciones con los resultados reales ajustando por azar, indica en casos como el del ETF SPY (Kappa de 0.0809) que el modelo ofrece una mejora marginal sobre las predicciones aleatorias, marcando un inicio prometedor hacia la afinación y mejora de este modelo predictivo.

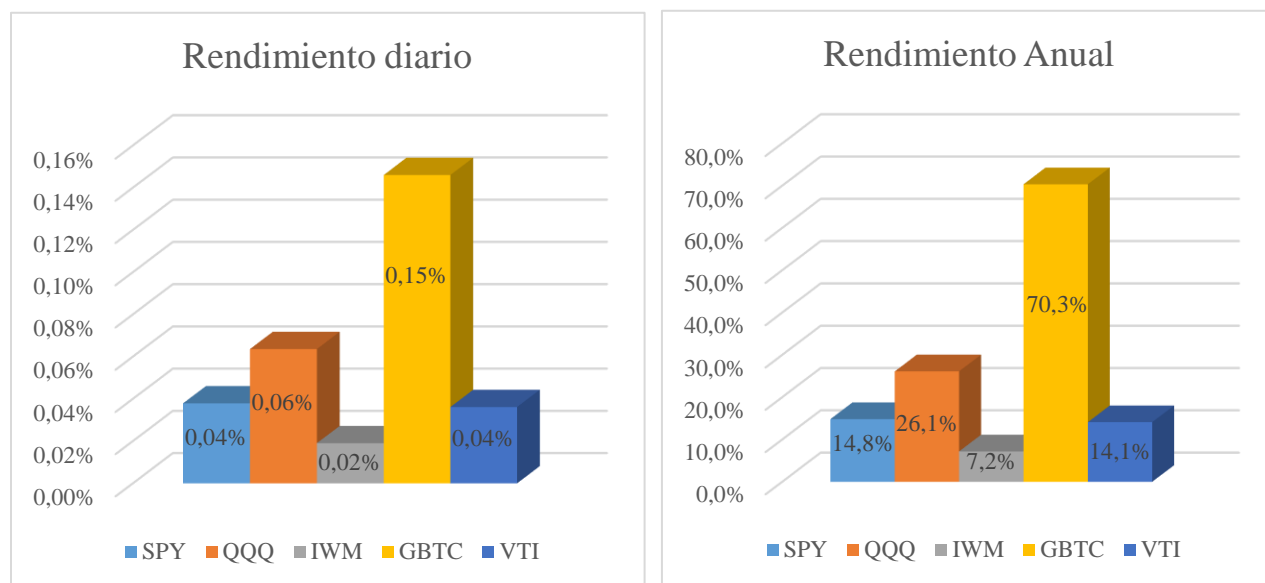
Con respecto al ETF VTI, el valor de kappa negativo sugiere que realizar predicciones aleatorias podría ser más efectivo que usar el modelo actual. Esto implica la necesidad de optimizar el modelo, ya sea mediante la selección de características más relevantes, el ajuste de parámetros, o la implementación de técnicas de balanceo de clases, para mejorar la concordancia medida por kappa.

El análisis reveló que la sensibilidad, es decir, la habilidad del modelo para identificar correctamente los días de “baja” es generalmente baja, destacando el VTI con solo un 12.17%. Por otro lado, la especificidad, o la capacidad de reconocer días de subida, resultó ser alta, con el VTI alcanzando un 86.03%. Esta tendencia sugiere que los modelos son más precisos al predecir aumentos que disminuciones en el valor de los ETFs. La precisión balanceada indica un rendimiento moderado en la predicción de ambas direcciones del mercado, mostrando que el modelo tiene una eficacia similar en identificar subidas y bajadas.

### **7.3 Construir un portafolio maximizando los rendimientos y minimizando el riesgo.**

Iniciando la fase final de esta investigación, se abordó la construcción de un portafolio óptimo. Este proceso comenzó con el cálculo de los rendimientos diarios y anuales de cada ETF seleccionado, estableciendo un contexto para evaluar su potencial de inversión.

**Figura 7.** Rendimientos ETF (SPY, QQQ, IWM, GBTC, VTI)



La fase inicial de preparación para la optimización del portafolio subrayó la relevancia de considerar la diversidad sectorial de los ETFs seleccionados. Para calcular los rendimientos diarios y anuales, se utilizó la fórmula tradicional del logaritmo natural del precio de hoy menos el logaritmo natural del precio de ayer. Los resultados, mostrados en la figura 7, reflejan una rentabilidad diaria para SPY del 0.04% (anual del 14.8%), QQQ del 0.06% (anual del 26.1%), IWM del 0.02% (anual del 7.2%), GBTC del 0.15% (anual del 70.3%), y VTI del 0.04% (anual del 14.1%). La evaluación de los rendimientos y las volatilidades específicas de cada ETF resaltó diferencias significativas, enfatizando la importancia de una estrategia de diversificación cuidadosa. Este método reafirmó el objetivo de diversificación marcado desde el inicio de la investigación y también estableció una base para proceder a una optimización de portafolio que busque un equilibrio óptimo entre rendimiento y riesgo.

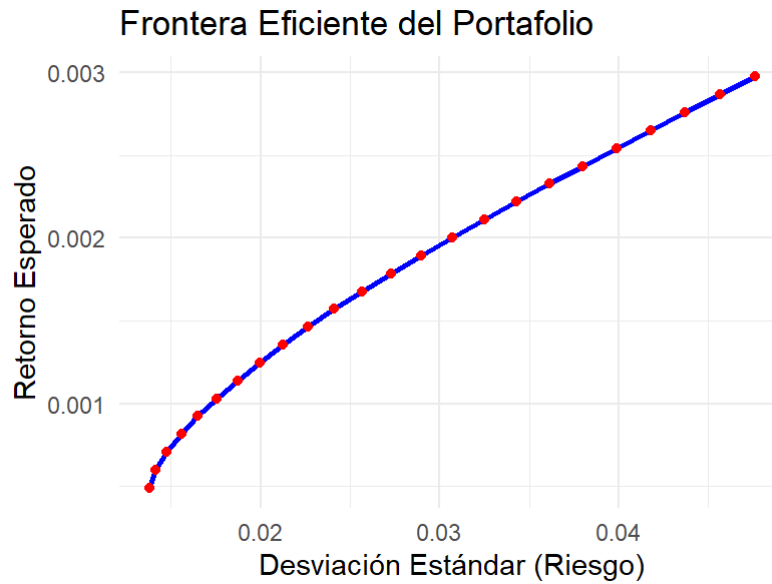
Antes de optimizar, se calculó la frontera eficiente sin restricciones para comprender mejor los retornos esperados de distintos portafolios a un nivel de riesgo específico. En este análisis específico, se identificaron

24 puntos en la frontera eficiente, cuya distribución de pesos por ETF se muestra en la figura 8. Posteriormente, se trazó la curva de la frontera eficiente, que se puede visualizar en la figura 9.

**Figura 8.** Pesos de los activos en la frontera eficiente.

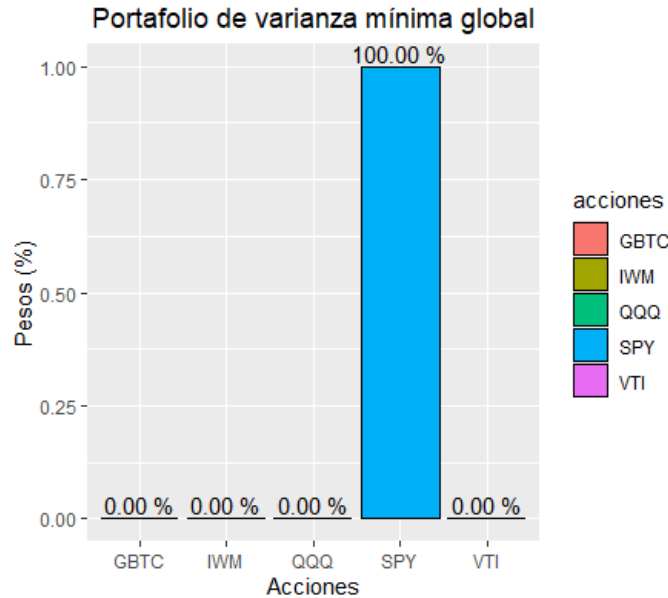
Punto	SPY	QQQ	IWM	GBTC	VTI
1	99,00%	0,00%	0,00%	0,00%	0,00%
2	94,80%	0,00%	0,00%	4,20%	0,00%
3	90,60%	0,00%	0,00%	8,40%	0,00%
4	79,10%	8,20%	0,00%	11,70%	0,00%
5	63,60%	20,70%	0,00%	14,70%	0,00%
6	48,10%	33,30%	0,00%	17,60%	0,00%
7	32,60%	45,80%	0,00%	20,60%	0,00%
8	17,10%	58,40%	0,00%	23,50%	0,00%
9	1,70%	70,90%	0,00%	26,40%	0,00%
10	0,00%	68,10%	0,00%	30,90%	0,00%
11	0,00%	63,40%	0,00%	35,60%	0,00%
12	0,00%	58,80%	0,00%	40,20%	0,00%
13	0,00%	56,30%	0,00%	44,30%	0,00%
14	0,00%	52,10%	0,00%	48,90%	0,00%
15	0,00%	47,50%	0,00%	53,50%	0,00%
16	0,00%	42,80%	0,00%	58,20%	0,00%
17	0,00%	38,20%	0,00%	62,80%	0,00%
18	0,00%	33,50%	0,00%	67,50%	0,00%
19	0,00%	28,90%	0,00%	72,10%	0,00%
20	0,00%	24,20%	0,00%	76,80%	0,00%
21	0,00%	19,60%	0,00%	81,40%	0,00%
22	0,00%	14,90%	0,00%	86,10%	0,00%
23	0,00%	10,30%	0,00%	90,70%	0,00%
24	0,00%	5,60%	0,00%	95,40%	0,00%

**Figura 9.** Frontera eficiente del portafolio.



Después de la fase de cálculo inicial, se llevó a cabo la optimización de los portafolios. Los resultados mostraron que, para obtener el portafolio con el menor riesgo, sería necesario concentrar la inversión en el activo menos riesgoso, lo cual resultaría en un portafolio de mínima varianza. Sin embargo, esto entra en contradicción con el principio de diversificación que propone la teoría moderna de portafolios de Markowitz. Esta situación pone de relieve el dilema entre minimizar el riesgo y maximizar la diversificación, un equilibrio esencial para la gestión eficiente de inversiones.

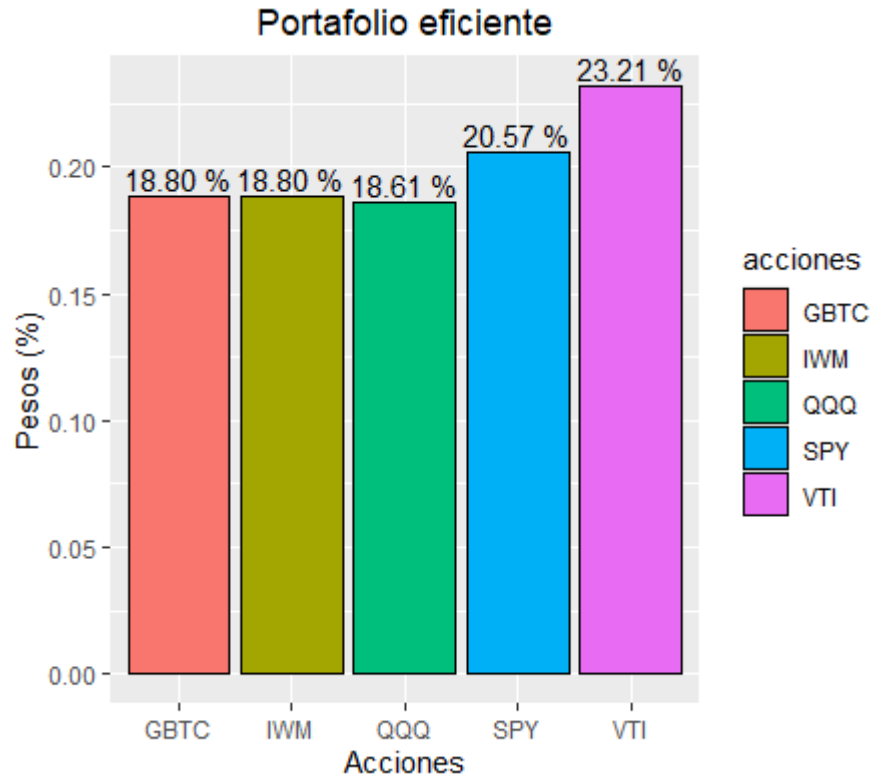
**Figura 10.** Portafolio sin diversificación



A continuación, para abordar el desafío de equilibrar riesgo y diversificación, se implementó una restricción adicional en la optimización del portafolio, estableciendo un mínimo del 5% y un máximo del 30% en la participación de cada activo. Esto garantiza que, durante la re-optimización, se explore una solución de menor varianza que involucre a todos los activos, alineándose más estrechamente con los principios de diversificación de la teoría moderna de portafolios.

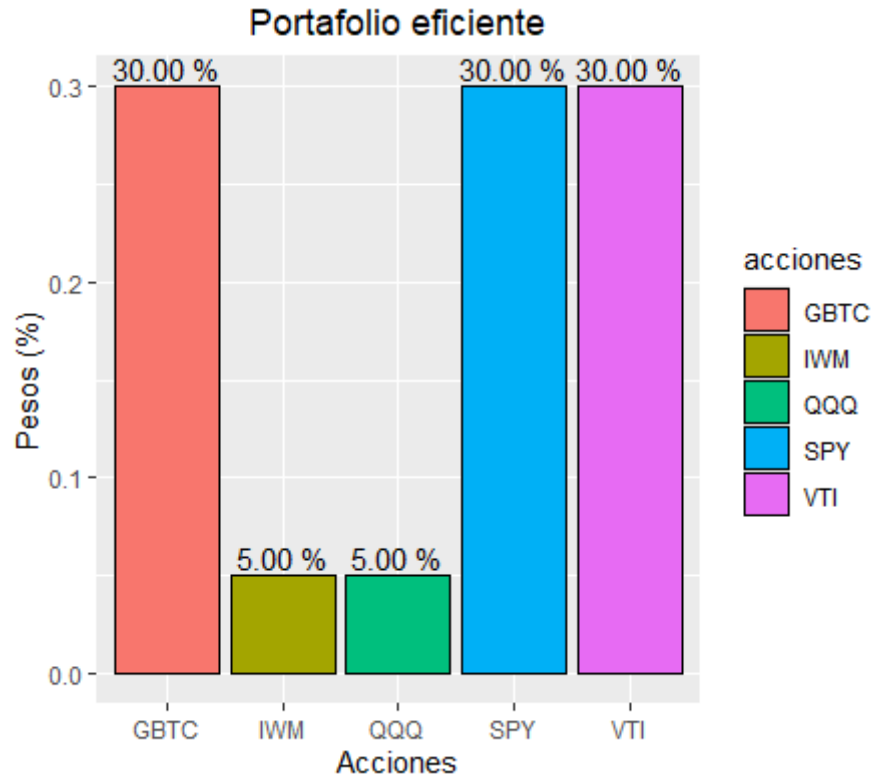
El resultado fue un portafolio equilibrado, con inversiones distribuidas entre VTI (23.21%), SPY (20.57%), IWM (18.80%), GBTC (18.80%) y QQQ (18.61%). A pesar de incluir activos como el GBTC, conocidos por su volatilidad, el portafolio logró una desviación estándar del 0.44%, indicando un riesgo reducido. Este enfoque demuestra cómo la diversificación y la selección estratégica de activos pueden conducir a un portafolio de bajo riesgo, incluso con inversión en activos volátiles.

**Figura 11.** Portafolio eficiente con menor varianza



Por otro lado, se buscó construir y optimizar un portafolio orientado a maximizar los rendimientos, aceptando para ello un mayor nivel de riesgo. Esta estrategia contrasta con la búsqueda de minimizar la volatilidad y refleja un enfoque más agresivo en la gestión de inversiones. La distribución resultante asignó mayor peso a GBTC, SPY y VTI (30% cada uno), mientras que IWM y QQQ representaron el 5% cada uno. Este enfoque refleja una estrategia agresiva, optando por activos de mayor rendimiento. Como resultado, el rendimiento anual esperado para este portafolio alcanzó el 17.31% efectivo anual, evidenciando cómo la priorización de activos de alto rendimiento puede elevar significativamente la rentabilidad esperada del portafolio.

**Figura 12.** Portafolio eficiente máximo rendimiento



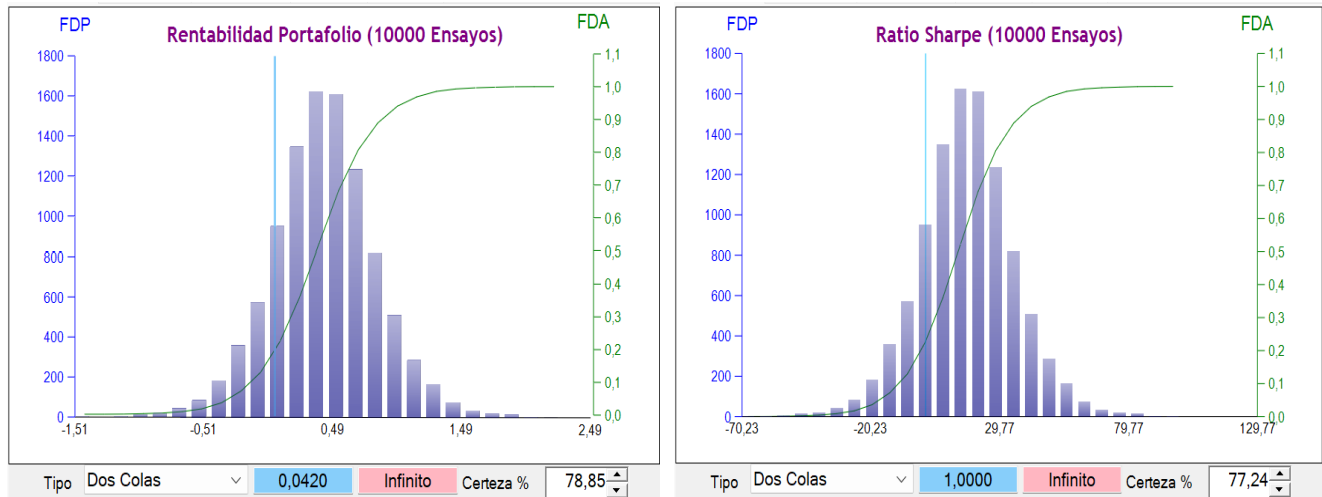
Finalmente, se examinó la relación entre el portafolio de mercado y la línea tangente a la frontera eficiente, también conocida como la línea del mercado de capitales (CML). Este análisis se centró en identificar el punto donde el portafolio de mercado tangente maximiza el Ratio de Sharpe, ofreciendo el mejor equilibrio posible entre retorno esperado y riesgo adicional. Al posicionar el portafolio en esta línea tangente, se aprovecha de manera óptima la relación entre rendimiento y volatilidad, conformando así un enfoque estratégico esencial para cualquier inversionista que busque eficiencia en la gestión de sus inversiones.

Para este análisis, se empleó el simulador de riesgos de Excel, utilizando como parámetro la tasa libre de riesgo los bonos del Tesoro de Estados Unidos a 10 años, que fue del 4.6% en abril de 2024. Se realizaron 10,000 simulaciones para obtener resultados robustos. Posteriormente, se optimizó el portafolio para maximizar el Ratio de Sharpe, respetando las restricciones previas de mantener un mínimo del 5% y un máximo del 30% en la participación de cada activo. Los resultados de esta optimización conformaron un portafolio con las siguientes distribuciones: SPY (30%), QQQ (30%), IWM (5%), GBTC (30%), y VTI (5%). Este portafolio alcanzó un Ratio de Sharpe de 13.70, con una varianza de solo 0.05% y una rentabilidad anual del 34.45%.

Como paso final, se llevó a cabo un análisis de sensibilidad utilizando la simulación Montecarlo para evaluar el comportamiento del portafolio frente a posibles cambios en las rentabilidades de sus activos. Se

realizaron 10,000 simulaciones para estimar la probabilidad de que el retorno del portafolio fuera mayor a la tasa libre de riesgo y de que el Ratio de Sharpe superara el valor de 1, incluso ante variaciones en las rentabilidades anuales de los precios de los activos. Este enfoque proporciona una comprensión más profunda del comportamiento del portafolio bajo diferentes escenarios económicos y de mercado.

**Figura 13.** Resultados de la simulación Montecarlo (10.000 escenarios)



Los resultados de esta simulación, ilustrados en la Figura 13, indican que existe un 78.85% de probabilidad de que la rentabilidad del portafolio supere la tasa libre de riesgo (4.6%), incluso ante fluctuaciones significativas en los precios de los activos. Esto sugiere que hay una alta probabilidad de que el portafolio mantenga su rentabilidad bajo diversas condiciones de mercado. En cuanto al Ratio de Sharpe, los datos de la simulación muestran que hay un 77.24% de probabilidad de que el Ratio de Sharpe del portafolio sea superior a 1. Esto implica que, a pesar de posibles cambios en los precios de los activos, también hay una alta probabilidad de que el portafolio continúe ofreciendo un rendimiento ajustado por riesgo favorable.

## 8. Conclusiones

En este estudio, se evaluó la integración de métodos de aprendizaje supervisado y la Teoría de Portafolio de Markowitz en la gestión de inversiones en ETFs, específicamente enfocándose en la minimización del riesgo y maximización de los rendimientos. Los hallazgos clave indican que la aplicación del método KNN (K-Nearest Neighbors) en la clasificación de tendencias de precios para diversos ETFs resulta en una variabilidad de precisión, lo que sugiere una correlación significativa entre la selección de parámetros específicos y la eficacia predictiva del modelo. Además, se observó que el impacto de la pandemia del COVID-19 en los mercados financieros ha tenido un efecto notable en el comportamiento de los ETFs,

mostrando una recuperación económica marcada después de la disminución inicial de precios. Este comportamiento refleja la resiliencia y adaptabilidad del mercado frente a choques económicos externos.

Se integraron las predicciones de movimiento de precios de ETFs obtenidas a través del método KNN con la optimización de portafolios del modelo de Markowitz. Ajustando los pesos del portafolio según las probabilidades de movimientos ascendentes o descendentes predichas, se propone una metodología que no solo optimiza los retornos en función del riesgo, sino que también anticipa cambios en el mercado. Este enfoque podría ser de gran interés para reguladores e inversores, ya que proporciona una base sólida para la toma de decisiones informada y la gestión de riesgos en entornos financieros dinámicos. A futuro, se podría usar las predicciones del método KNN como una base para asignar pesos en el modelo de Markowitz, implementando una estrategia de rebalanceo que ajuste periódicamente los pesos del portafolio basado en nuevas predicciones del método KNN, entre otras estrategias.

La sinergia entre el método KNN para la predicción de tendencias de precios y la teoría de portafolios de Markowitz se manifiesta en una estrategia de inversión adaptativa y orientada al futuro. Utilizando el método KNN, se logró prever las direcciones de los precios de los ETFs con una precisión mejorada, lo cual es crucial para la asignación dinámica de activos recomendada por Markowitz. Esta predicción precisa permitiría a los inversores y gestores de fondos ajustar los pesos de los activos en sus portafolios no solo en respuesta a los cambios del mercado ya ocurridos, sino también en anticipación a movimientos futuros potenciales. Por ende, este enfoque proactivo asegura que la diversificación del portafolio no solo se basa en condiciones históricas o actuales, sino que también incorpora expectativas de mercado, maximizando así el rendimiento ajustado por riesgo. Finalmente, la integración de estas técnicas avanza la gestión de portafolios desde un modelo estático a uno dinámico, donde la reacción rápida y la anticipación se convierten en elementos clave para el éxito en la inversión.

La investigación también destaca la importancia de la diversificación en la gestión de portafolios de ETFs, conforme a la teoría de portafolio de Markowitz. Se demostró que la construcción de portafolios óptimos, que equilibran el rendimiento y el riesgo, mejora significativamente los resultados de inversión. Este enfoque confirma la necesidad de una selección estratégica de activos y la aplicación de métodos analíticos avanzados para la optimización de carteras. A pesar de los desafíos encontrados en la predicción precisa de tendencias del mercado, los resultados sugieren que el uso de modelos de inteligencia artificial y teorías de inversión bien desarrolladas puede ofrecer una ventaja en la gestión de inversiones, particularmente en tiempos de incertidumbre económica.

Por otro lado, los hallazgos de este estudio se alinean con la literatura existente, destacando la eficacia de combinar la inteligencia artificial, específicamente a través del método KNN, con la Teoría de Portafolio de Markowitz en la gestión de ETFs. Este enfoque refleja investigaciones previas que resaltan la importancia de la inteligencia artificial en la predicción de movimientos del mercado financiero y la optimización de portafolios. Consistente con la literatura que examina la adaptabilidad y eficacia de los ETFs como herramientas de inversión, los resultados confirman la capacidad de estos métodos para mejorar la toma de decisiones en la gestión de inversiones. Este estudio respalda los hallazgos existentes sobre las aplicaciones de la inteligencia artificial y la teoría de portafolios proporcionando una visión contemporánea relevante ante situaciones de mercado fluctuantes.

También tiene implicaciones significativas tanto prácticas como teóricas en el campo de las finanzas y la gestión de inversiones. Desde una perspectiva teórica, este trabajo subraya la aplicabilidad y relevancia de combinar modelos avanzados de inteligencia artificial, como el método KNN, con principios establecidos de la Teoría de Portafolio de Markowitz. En el ámbito práctico, los resultados ofrecen estrategias concretas para inversores y gestores de fondos, destacando la importancia de la diversificación de portafolio y la utilización de la inteligencia artificial para mejorar la precisión de las predicciones de mercado. Esto sugiere que las herramientas basadas en IA pueden ser instrumentales en la identificación de oportunidades de inversión y la mitigación de riesgos, particularmente en periodos de incertidumbre económica como los provocados por la pandemia del COVID-19.

Con respecto a las limitaciones de esta investigación, una de ellas es el ajuste del modelo, específicamente el método KNN utilizado para la clasificación de tendencias de precios en ETFs. Aunque el KNN demostró ser útil, la selección del número óptimo de vecinos y la métrica de distancia son críticas y pueden afectar significativamente la precisión del modelo. Esta sensibilidad a los parámetros sugiere que los resultados podrían variar con diferentes configuraciones, limitando la generalización de los hallazgos. De la misma forma, la selección del paquete CARET en R para la implementación de modelos de inteligencia artificial. Aunque CARET es una herramienta versátil para la modelización predictiva, no abarca todos los posibles métodos de clasificación ni todas las funcionalidades de ajuste de modelos avanzados disponibles. La elección de este paquete puede restringir la exploración de técnicas más novedosas o personalizadas en la inteligencia artificial.

Finalmente, como sugerencia para futuras investigaciones se podría plantear un enfoque distinto al uso de algoritmos de clasificación, explorando otros modelos de inteligencia artificial que podrían proporcionar diferentes perspectivas o mejorar la precisión de las predicciones. La exploración de otras técnicas,

incluyendo redes neuronales, máquinas de soporte vectorial o algoritmos de regresión avanzada, podría ofrecer una comprensión más completa y mejorar la capacidad de predicción que estos modelos.

## **9. Recomendaciones**

Con base a los resultados obtenidos en este estudio, se recomienda a los profesionales financieros integrar herramientas de inteligencia artificial, como el método KNN, en sus estrategias de inversión para mejorar la predicción de tendencias de mercado y la clasificación de activos. La implementación de estos algoritmos puede proporcionar perspectivas detalladas y mejorar significativamente la toma de decisiones en la gestión de carteras.

Además, se enfatiza la necesidad de una evaluación y actualización periódica de los modelos de IA utilizados en la gestión de inversiones. Los profesionales deben asegurarse de que sus modelos estén constantemente alineados con las condiciones actuales del mercado y adaptarlos en respuesta a nuevos datos y cambios en el entorno económico. Esto implica no solo el ajuste de parámetros existentes, sino también la exploración de nuevas metodologías y la inclusión de conjuntos de datos más recientes para mantener la precisión y relevancia de las predicciones.

Ahora bien, con respecto a las recomendaciones para futuras investigaciones en el campo de la gestión de inversiones y la inteligencia artificial, se recomienda explorar una variedad más amplia de modelos de IA y algoritmos de clasificación para mejorar la precisión en la predicción de tendencias del mercado y la eficacia en la gestión de portafolios. Es vital investigar cómo diferentes configuraciones y enfoques, como redes neuronales profundas o máquinas de vector de soporte, pueden compararse o complementarse con los métodos utilizados en este estudio, como el KNN.

## Anexo 1: Valores de Hiperparámetros Óptimos y Herramientas Utilizadas

### Valores de Hiperparámetros Óptimos

Los valores óptimos de los hiperparámetros obtenidos durante el proceso de optimización se presentan a continuación para cada uno de los ETFs analizados:

<b>Modelo</b>	<b>Hiperparámetro</b>	<b>Valor Optimo</b>
KNN (SPY)	k	34
KNN (QQQ)	k	41
KNN (IWM)	k	5
KNN (GBTC)	k	56
KNN (VTI)	k	83

### Software y Librerías Utilizadas

Para la implementación y optimización de los modelos, así como para la visualización y manipulación de datos, se utilizaron las siguientes herramientas y librerías:

#### Entorno de Desarrollo:

**RStudio:** Utilizado como entorno de desarrollo integrado (IDE) para escribir y ejecutar scripts en R.

#### Librerías de R:

**caret:** Versión 6.0-86. Utilizada para la construcción y validación de modelos predictivos, así como para la búsqueda en cuadrícula y validación cruzada.

**PortfolioAnalytics:** Versión 1.2.0. Empleada para la optimización de portafolios según la teoría de Markowitz.

**ggplot2:** Versión 3.3.3. Utilizada para la visualización de datos mediante gráficos.

**dplyr:** Versión 1.0.4. Utilizada para la manipulación y transformación de datos.

## **Procedimiento de Optimización**

El proceso de optimización de hiperparámetros se llevó a cabo de la siguiente manera:

**Definición de Hiperparámetros:** Se identificaron los hiperparámetros clave para el método KNN, específicamente el número de vecinos ( $k$ ) y la métrica de distancia.

**Búsqueda en Cuadrícula (Grid Search):** Se empleó la técnica de búsqueda en cuadrícula para evaluar todas las combinaciones posibles de los valores predefinidos de los hiperparámetros.

**Validación Cruzada:** Se aplicó una validación cruzada de 10-fold para evaluar el rendimiento de cada combinación de hiperparámetros, asegurando que los modelos fueran evaluados de manera robusta.

**Selección de Valores Óptimos:** Se seleccionaron los valores de hiperparámetros que proporcionaron el mejor rendimiento en términos de precisión de clasificación para cada ETF.

## Referencias bibliográficas:

**Bollen, J., Mao, H., & Zeng, X. (2011).** Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8.

**Burrell, P.R., & Folarin, B.O. (1997).** The Impact of Neural Networks in Finance. *Neural Computing & Applications*, 6, 193-200.

**Cao, L., & Tay, F. E. H. (2001).** Financial Forecasting Using Support Vector Machines. *Neural Comput & Applic*, 10, 184–192.

**Daniels, H., & Kamp, B. (1999).** Application of MLP Networks to Bond Rating and House Pricing. *Neural Computing & Applications*, 8, 226–234.

**Deng, Y., Bao, F., Kong, Y., Ren, Z., & Dai, Q. (2017).** Deep direct reinforcement learning for financial signal representation and trading. *IEEE transactions on neural networks and learning systems*, 28(3), 653-664.

**Dixon, M., Klabjan, D., & Bang, J. H. (2020).** Classification-based financial markets prediction using deep neural networks. *Algorithmic Finance*, 8(3-4), 147-160.

**Fama, E. F. (1970).** Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383-417.

**Fama, E. F., & French, K. R. (2004).** The Capital Asset Pricing Model: Theory and evidence. *Journal of Economic Perspectives*, 18(3), 25-46.

**Garcia, D. (2013).** Sentiment during recessions. *The Journal of Finance*, 68(3), 1267-1300.

**Huang, W., Nakamori, Y., & Wang, S.-Y. (2005).** Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32(10), 2513-2522.

**Kamalov, F. (2020).** Forecasting significant stock price changes using neural networks. *Neural Computing and Applications*, 32, 17655–17667.

**Kamalov, F., Smail, L., & Gurrib, I. (2021).** Stock price forecast with deep learning.

**Kumar, M., & Thenmozhi, M. (2006).** Forecasting stock index movement: A comparison of support vector machines and random forest. Indian Institute of Capital Markets 9th Capital Markets Conference Paper.

**Lettau, M., & Madhavan, A. (2018).** Exchange-traded funds 101 for economists. *Journal of Economic Perspectives*, 32(1), 135-154.

**Ma, Y., Han, R., & Wang, W. (2020).** Optimización de carteras con predicción de rendimiento utilizando aprendizaje profundo y aprendizaje automático. *Expert Systems With Applications*, 165, 113973.

**Markowitz, H. (1952).** Portfolio Selection. *The Journal of Finance*, 7(1), 77-91.

**Serrano, W. (2021).** The random neural network in price predictions. *Neural Computing and Applications*, 34, 855–873.

**Sharpe, W. F. (1964).** Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425-442.

**Silva de Oliveira, A., Ceretta, P. S., & Albrecht, P. (2023).** Performance comparison of multifractal techniques and artificial neural networks in the construction of investment portfolios. *Finance Research Letters*, 55, 103814.

**Statista. (2023).** Percentage of U.S. Adults Investing in the Stock Market. Recuperado de <https://www.statista.com/statistics/270034/percentage-of-us-adults-to-have-money-invested-in-the-stock-market/>

**Tetlock, P. C. (2007).** Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139-1168.

**Uysal, A. S., Li, X., & Mulvey, J. M. (2021).** End-to-End Risk Budgeting Portfolio Optimization with Neural Networks. Department of Operations Research and Financial Engineering, Princeton University.

**Zhang, L., Agarwal, A., & Prasad, V. (2019).** A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS one*, 14(7), e0219650.