



Facultad de Ciencias Económicas y Empresariales  
Doble Grado en Administración y Dirección de Empresas y  
Análisis de Negocios/ Business Analytics (E2 Analytics)

# **TRABAJO DE FIN DE GRADO EN ANÁLISIS DE NEGOCIOS / BUSINESS ANALYTICS**

## **OPTIMIZACIÓN DE ROBOTS DE TRADING**

Autor: David Serrano González  
Director: Luis Ángel Calvo Pascual

MADRID | Junio de 2023

## Resumen

En un mundo en constante digitalización, los mercados financieros con su flujo incesante de datos han encontrado en las nuevas tecnologías de la información a sus mayores aliados. Los ordenadores, aportando su inigualable capacidad de cálculo, junto con diversos algoritmos y modelos de aprendizaje automático, están revolucionando la manera en la que estos datos son aprovechados. Esta combinación, está dando lugar a un mercado cada vez más automatizado en el que gradualmente los traders humanos están siendo desplazados en favor de sus contrapartes automatizadas, capaces de tomar decisiones rápidas sin los sesgos que un humano pudiera sufrir.

Este trabajo se sumerge en esta realidad, explorando el desarrollo y optimización de diversos robots de trading. A lo largo del trabajo se desarrollan tanto robots derivados de estrategias de trading convencionales, como otros de modelos avanzados de aprendizaje automático, al mismo tiempo que se complementa este desarrollo con un análisis en profundidad del funcionamiento de los mismos.

El análisis del funcionamiento de estos robots se ha llevado a cabo haciendo uso de múltiples valores contenidos en el índice S&P 500 durante un periodo de backtesting que abarca desde 2017 hasta octubre del 2022. Los resultados muestran variedad en el desempeño de los diversos robots, evidenciando no obstante la capacidad de generar estrategias de trading efectivas. Entre los robots desarrollados, destaca el robot que implementa una estrategia optimizada basada en medias y Bandas de Bollinger. Sin embargo, otro robot, el basado en Q-Learning despierta un especial interés debido a su potencial, dada su limitada volatilidad y su prometedora capacidad si se consigue una profunda comprensión de su funcionamiento.

Palabras clave: Trading algorítmico, Python, Bandas de Bollinger, CAPM, Q-Learning, S&P 500, Finanzas.

## **Abstract**

In a world in constant digitalization, the financial markets, with their incessant flow of data, have found their greatest allies in the new information technologies. Computers, with their unrivaled computing power, together with various algorithms and machine learning models, are revolutionizing the way in which this data is used. This combination is giving rise to an increasingly automated market in which human traders are gradually being displaced in favor of their automated counterparts, capable of making quick decisions without the biases that a human might suffer from.

This paper dives into this reality, exploring the development and optimization of various trading robots. Throughout the work, both robots derived from conventional trading strategies and advanced machine learning models are developed, while this development is complemented with an in-depth analysis of their performance.

The analysis of the performance of these robots has been carried out by making use of multiple stocks contained in the S&P 500 index during a backtesting period spanning from 2017 to October 2022. The results show variety in the performance of the various robots, evidencing nonetheless the ability to generate effective trading strategies. Among the robots developed, the robot that implements an optimized strategy based on Bollinger Bands and Averages stands out. However, another robot, the one based on Q-Learning arouses special interest due to its potential, given its limited volatility and its promising capability if a deep understanding of its operation is achieved.

Keywords: Algorithmic Trading, Python, Bollinger Bands, CAPM, Q-Learning, S&P 500, Finance.

# ÍNDICE

<b>CAPÍTULO 1: INTRODUCCIÓN</b> .....	<b>7</b>
<b>1.1 Contexto</b> .....	<b>7</b>
1.1.1 Interés en el tema propuesto .....	7
1.1.2 Robots de Trading y Machine Learning.....	8
1.1.3 El índice S&P 500 .....	9
<b>1.2 Marco Teórico - Antecedentes</b> .....	<b>10</b>
1.2.1 Evolución del Trading Algorítmico y Arbitraje de Alta Frecuencia.....	10
1.2.2 Algoritmos y técnicas más utilizados en el trading algorítmico .....	12
1.2.3 Trading Algorítmico en el S&P 500 .....	14
<b>1.3 Objetivos</b> .....	<b>14</b>
<b>1.4 Metodología</b> .....	<b>15</b>
<b>1.5 Estructura del Trabajo</b> .....	<b>16</b>
<b>CAPÍTULO 2: OPTIMIZACIÓN DEL ROBOT DE MEDIAS Y BANDAS</b> .....	<b>17</b>
<b>2.1 Robot de Trading Básico: Medias Móviles</b> .....	<b>17</b>
2.1.1 Presentación del robot de trading .....	17
2.1.2 Descripción en profundidad del funcionamiento del robot.....	19
2.1.3 Resultados del robot de trading .....	20
<b>2.2 Sofisticación del robot de trading: Bandas de Bollinger</b> .....	<b>24</b>
2.2.1 Presentación del robot de trading .....	24
2.2.2 Descripción en profundidad del funcionamiento del robot.....	25
2.2.3 Resultados del robot de trading .....	27
<b>2.3 Optimización del robot de trading</b> .....	<b>32</b>
2.3.1 Descripción de la optimización .....	32
2.3.2 Resultados de la optimización .....	34
<b>CAPÍTULO 3: OPTIMIZACIÓN DEL ROBOT DE Q-LEARNING</b> .....	<b>38</b>
<b>3.1 Presentación de robot de Q-Learning</b> .....	<b>38</b>
<b>3.2 Robot básico de Q-Learning</b> .....	<b>39</b>
3.2.1 Descripción en profundidad del robot de trading .....	39
3.2.2 Resultados del robot de trading .....	44
<b>3.3 Robot optimizado de Q-Learning</b> .....	<b>45</b>
3.3.1 Descripción en profundidad del robot de trading .....	45
3.3.2 Resultados del robot de trading .....	49
<b>CAPÍTULO 4: OPTIMIZACIÓN DEL ROBOT CAPM</b> .....	<b>52</b>
<b>4.1 Presentación de robot de CAPM</b> .....	<b>52</b>
<b>4.2 Robot basado en CAPM básico</b> .....	<b>53</b>
4.2.1 Descripción en profundidad del robot de trading .....	53
4.2.2 Resultados del robot de trading .....	57
<b>4.3 Optimización del robot CAPM</b> .....	<b>59</b>
4.3.1 Descripción en profundidad del robot de trading .....	59
<b>CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN</b> .....	<b>64</b>
<b>BIBLIOGRAFÍA</b> .....	<b>69</b>

# ÍNDICE DE FIGURAS

<i>Figura 1: Ilustración de la estrategia de trading de reversión de la media.....</i>	<i>18</i>
<i>Figura 2: Análisis de rentabilidad variando tamaño de la ventana temporal para el cálculo de la media móvil .....</i>	<i>22</i>
<i>Figura 3: Ilustración de la estrategia de venta con Bandas de Bollinger .....</i>	<i>26</i>
<i>Figura 4: Ilustración de la estrategia de compra con Bandas de Bollinger .....</i>	<i>26</i>
<i>Figura 5: Análisis de rentabilidad según la combinación de parámetros para el robot de Bandas de Bollinger .....</i>	<i>30</i>
<i>Figura 6: Diagrama de flujo del proceso algorítmico optimizado para el robot de medias y bandas.....</i>	<i>34</i>
<i>Figura 7: Análisis de rentabilidad según la combinación de parámetros para el robot optimizado de medias y bandas.....</i>	<i>35</i>
<i>Figura 8: Análisis de rentabilidad según la combinación de parámetros para el robot optimizado de medias y bandas, con foco en rango de valores óptimo .....</i>	<i>36</i>
<i>Figura 9: Ejemplo simplificado de la Tabla o Matriz Q.....</i>	<i>39</i>
<i>Figura 10: Histograma de rentabilidades de las distintas acciones haciendo uso del robot de Q-Learning optimizado .....</i>	<i>50</i>
<i>Figura 11: Gráfico de cajas de las rentabilidades de las distintas acciones haciendo uso del robot de Q-Learning optimizado .....</i>	<i>50</i>
<i>Figura 12: Diagrama de flujo del proceso algorítmico .....</i>	<i>56</i>
<i>Figura 13: Resultado de simulaciones variando umbral de inversión .....</i>	<i>59</i>
<i>Figura 14: Distribución de rentabilidades predichas del robot basado en CAPM.....</i>	<i>60</i>
<i>Figura 15: Valores de <math>\beta</math> de todas las empresas estudiadas, durante el periodo Enero 2000 – Octubre 2022 (ventanas de 21 días).....</i>	<i>63</i>
<i>Figura 16: Comportamiento del Robot de CAPM frente al índice S&amp;P 500 .....</i>	<i>64</i>
<i>Figura 17: Comportamiento de los robots a lo largo del periodo de backtesting.....</i>	<i>64</i>
<i>Figura 18: Comparativa de Ratio de Sharpe del mejor robot vs. mercado durante periodo de backtesting .....</i>	<i>67</i>

## ÍNDICE DE TABLAS

<i>Tabla 1: Análisis del comportamiento del robot de medias móviles.....</i>	<i>23</i>
<i>Tabla 2: Rentabilidad del robot de Bandas de Bollinger haciendo uso de las combinaciones de parámetros no incluidas en el análisis general .....</i>	<i>29</i>
<i>Tabla 3: Análisis del comportamiento del robot de Bandas de Bollinger.....</i>	<i>31</i>
<i>Tabla 4: Análisis del comportamiento del robot optimizado de medias y bandas .....</i>	<i>37</i>
<i>Tabla 5: Resultados del robot de Q-Learning básico .....</i>	<i>45</i>
<i>Tabla 6: Resultados del robot de Q-Learning optimizado .....</i>	<i>49</i>
<i>Tabla 7: Resultados de las simulaciones para FLT.....</i>	<i>51</i>
<i>Tabla 8: Resultados de las simulaciones para AMD.....</i>	<i>51</i>
<i>Tabla 9: Resultados del robot de trading basado en CAPM .....</i>	<i>57</i>
<i>Tabla 10: Resultados de entrenamiento de modelos de CAPM alternativos.....</i>	<i>61</i>
<i>Tabla 11: Resultados de validación de modelos de CAPM alternativos.....</i>	<i>62</i>
<i>Tabla 12: Comparativa de rentabilidad de los distintos robots estudiados .....</i>	<i>65</i>
<i>Tabla 13: Comparativa de volatilidad de los distintos robots estudiados.....</i>	<i>65</i>
<i>Tabla 14: Comparativa de Ratio de Sharpe de los distintos robots estudiados.....</i>	<i>66</i>

# **CAPÍTULO 1: INTRODUCCIÓN**

## **1.1 Contexto**

### **1.1.1 Interés en el tema propuesto**

Este estudio se adentra en el campo del trading algorítmico, una disciplina que combina los conocimientos financieros con las habilidades informáticas con el objetivo de diseñar e implementar sistemas de inversión automatizados en los mercados bursátiles (Chaboud, Chiquoine, Hjalmarsson, & Vega, 2014). Este campo suscita un gran interés debido a la combinación de dos campos que por sí mismos ya resultan complejos y desafiantes.

En primer lugar, los mercados financieros, a pesar de su longevidad, continúan planteando un desafío intelectual para los que se adentran en ellos. Los inversores individuales, a menudo atraídos por la posibilidad de obtener rendimientos importantes, se encuentran frecuentemente lidiando con pérdidas financieras debido a la complejidad inherente de los mercados (Barber, Lee, Liu, & Odean, 2009). Sin embargo, esta misma complejidad también puede representar una lucrativa oportunidad para aquellos con la capacidad necesaria para comprenderlos, como es el caso de los profesionales y empresas, que, haciendo uso de su conocimiento de las dinámicas y tendencias de los mercados, consiguen obtener rentabilidades superiores a las que se observan en el inversor individual (Shapira & Venezia, 2001).

En segundo lugar, el crecimiento exponencial de las tecnologías de inteligencia artificial y de los algoritmos predictivos que se ha observado en los años recientes resulta fascinante. Estos avances están facilitando la automatización de actividades que, hasta ahora, se pensaba que solo podían ser realizadas por seres humanos, dando lugar a una nueva era tecnológica en la que las máquinas superan a las personas en eficiencia, incluso en aspectos que antes se consideraban exclusivamente humanos (Haenlein & Kaplan, 2019).

Este progreso en el campo de los algoritmos predictivos y de la inteligencia artificial está permitiendo el desarrollo de modelos capaces de procesar y analizar grandes volúmenes de datos de diversas tipologías, permitiendo el desarrollo de estrategias de inversión automatizadas que en muchas ocasiones pueden superar el rendimiento humano, sobre todo en cuanto a tiempos de ejecución de las mismas (Chaboud, Chiquoine, Hjalmarsson,

& Vega, 2014). La conjunción del sector financiero, con su inherente generación constante de un volumen masivo de información, y estas tecnologías emergentes, están haciendo posible la generación de sistemas capaces de predecir el valor en bolsa de un amplio espectro de activos, optimizando de esta manera las decisiones de inversión.

Tan relevante es esta unión, que muchas de las compañías líderes en el sector financiero ya cuentan con departamentos específicamente destinados al desarrollo de este tipo de algoritmos, no solo para hacer un uso propio de ellos, sino también para ofrecérselos a sus clientes (J.P. Morgan, 2018).

### **1.1.2 Robots de Trading y Machine Learning**

El trading algorítmico, conocido también como trading basado en reglas y procesos, ha revolucionado el paisaje de los mercados financieros en las últimas décadas, llegando a ser responsable de la gran mayoría de operaciones llevadas a cabo en los mercados financieros internacionales (Hendershott, Jones, & Menkveld, 2011). En el núcleo de este cambio se encuentran los robots de trading: programas informáticos diseñados para ejecutar operaciones en los mercados financieros de una manera automática, rápida y precisa, según una serie de reglas y estrategias predefinidas en ellos.

Estos robots pueden realizar tareas como la identificación de oportunidades de trading, el envío de órdenes de compra y venta, y la gestión de las carteras de inversión de los traders. Todas estas tareas se realizan de forma autónoma, sin precisar de participación humana una vez que están puestos en funcionamiento. Esta nula intervención humana dota a estos robots de su principal ventaja frente al trading convencional: su capacidad para operar de manera continua sin verse afectados por las emociones humanas u otras limitaciones que pueden influir en la toma de decisiones de los traders humanos. Además, pueden procesar una cantidad extremadamente superior de información y tomar decisiones de inversión en milisegundos, mucho más rápido de lo que cualquier individuo podría hacerlo (Jones, 2013).

Estos robots, sin embargo, son tan buenos como los propios algoritmos y estrategias en los que se basan, es por tanto que el desarrollo de robots de trading se ha potenciado significativamente con la aplicación de tecnologías de *Machine Learning* (ML). ML,

también conocido como aprendizaje automático, proporciona a los ordenadores la habilidad para aprender y mejorar su rendimiento sobre la realización de una tarea, sin la necesidad de tener que ser programados explícitamente para ello (Samuel, 1959). Los algoritmos de ML, a raíz de los datos, consiguen identificar patrones y estructuras para con ello mejorar su desempeño en una tarea determinada, como pueden ser tareas de predicción numérica, también conocidas como tareas de regresión, tareas de clasificación, o tareas de agrupación o *clustering*, entre muchas otras.

En la industria del trading, los algoritmos de ML son utilizados para diseñar y mejorar las estrategias que siguen los sistemas de negociación, permitiéndoles aprender de los datos históricos de los mercados financieros y adaptarse a las nuevas condiciones del mercado.

### **1.1.3 El índice S&P 500**

El Standard & Poor's 500, comúnmente conocido como S&P 500, es uno de los índices bursátiles más importantes del mundo. Es un índice ponderado por capitalización de mercado en el que están contenidas 500 de las mayores empresas de Estados Unidos, representando aproximadamente el 70% del total de capitalización bursátil de todo el país (Blume & Edelen, 2004). Dada su amplia representación de la economía de los EE.UU., el S&P 500 es generalmente considerado como el mejor indicador de la salud del mercado de valores de EE.UU. y un barómetro confiable de la economía de estadounidense a nivel macro, así como un punto de referencia muy común para la comparación de rendimientos de inversión.

El S&P 500, gracias a su amplitud y diversidad, se presenta como un índice atractivo para su uso en el trading algorítmico, principalmente debido a la gran cantidad de datos disponibles para su análisis y explotación. Además, su relevancia económica coloca a sus componentes en el centro de una multitud de investigaciones y estudios, los cuales exploran diversos aspectos, incluyendo estudios sobre las variables macroeconómicas y su influencia sobre los precios del índice. Por ejemplo, Pruess y Hiranto (2019) destacan la fuerza del dólar americano, la prima de riesgo y los tipos de interés como factores determinantes en la formación de los precios del índice, mientras que Sirucek (2012), de manera similar, subraya la influencia de la inflación, el desempleo y los tipos de interés. Esta valiosa información, entre toda aquella disponible en la multitud de estudios

realizados, puede ser integrada en los algoritmos, aumentando así la eficacia de estos al dotar a sus creadores de sólidos puntos de partida para la creación de sus modelos.

## **1.2 Marco Teórico - Antecedentes**

### **1.2.1 Evolución del Trading Algorítmico y Arbitraje de Alta Frecuencia**

El trading algorítmico ha pasado por varias fases de desarrollo desde sus inicios hasta llegar a su estado actual. Los primeros sistemas de trading algorítmico surgieron a inicios de los años 90, década en la que según argumenta Marco Avellaneda (2011), debido a la “decimalización” de los mercados financieros (cambio de las cotizaciones del formato fracción a formato decimal), el acceso directo al mercado, la aparición de mercados 100% digitales, la reducción de las comisiones y la creación de nuevos mercados más allá del New York Stock Exchange, se produjo una verdadera explosión de los sistemas de trading algorítmico en las compañías financieras.

Estos sistemas surgen de la necesidad que tenían las grandes empresas de hacer operaciones que implicaban la compra o venta de enormes cantidades de acciones que, en muchas ocasiones, impactaban en el precio del mercado. Este impacto en el precio de los valores hacía que este tipo de operaciones resultasen más costosas para las compañías, por lo que se buscaron formas de dividir estas grandes órdenes en otras más pequeñas que se ejecutasen electrónicamente a lo largo de un periodo de tiempo más largo, para así no afectar a la liquidez del mercado (Avellaneda, 2011).

Como apuntan Hendershott, Jones y Menkveld (2011), antes, si un inversor necesitaba hacerse con una gran cantidad de acciones de una empresa en particular, necesitaría contratar a un bróker presente en la planta de negociación de una bolsa para que inteligentemente fuese comprando acciones a un precio competitivo. Sin embargo, según los mercados se fueron digitalizando, estos brókeres presentes en las principales bolsas fueron siendo sustituidos por programas informáticos que cumplían su misma tarea, pero de una forma más fácil y económica.

Ahora, años después de la digitalización de los mercados financieros, son pocos los brókeres que no ofrecen una gran variedad de algoritmos a sus clientes corporativos. Algoritmos que, monitorizando dinámicamente el mercado, consiguen ejecutar las

operaciones deseadas por los inversores al mejor precio y tiempo, es decir, optimizándolas (Hendershott, Jones, & Menkveld, 2011).

Por las propias características del trading algorítmico, como es la rapidez en las operaciones, una modalidad comúnmente empleada por estos robots es el llamado trading de alta frecuencia. Como explica Jones (2013) en su trabajo, este tipo de robots son algoritmos que se dedican a generar retornos a través de la realización de grandes cantidades de operaciones diarias a una velocidad extraordinariamente rápida. Uno de los tipos de trading de alta frecuencia más representativos es el conocido como arbitraje de alta frecuencia, en el que los inversores, por medio de robots extremadamente rápidos consiguen aprovechar diferencias en precios de un mismo activo en diferentes mercados. Este tipo de operación, además de generar una rentabilidad a un riesgo prácticamente nulo, consigue que los mercados sean más eficientes, permitiendo que los desajustes en precios no duren más de pocos milisegundos reflejados en los mercados internacionales (Jones, 2013).

El trading de alta frecuencia y su relación con una mejora de la eficiencia de los mercados ha sido estudiado por varios autores como Hendershott, Jones y Menkveld (2010) o Boehmer, Fong y Wu (2015), demostrando todos ellos esta relación positiva. Sin embargo, a raíz del “*flash crash*” del año 2010, una situación en la que los precios de múltiples índices bursátiles americanos colapsaron y rebotaron a un nivel de precios similar al previo en un periodo de tiempo de poco más de media hora, empezaron a surgir críticas contra estos algoritmos que realizan operaciones autónomamente, sosteniendo que este tipo de algoritmos aumentan la volatilidad. Sin embargo, por medio de trabajos como el de Brogaard, Hendershott y Riordan (2013), se ha demostrado que estas críticas no son ciertas y que estos robots de trading de alta frecuencia no solo no aumentan la volatilidad de precios en los mercados de valores, sino que la reducen.

Dicho esto, cabe señalar que el trading algorítmico de alta frecuencia, si bien fascinante, tiene sus limitaciones y desafíos, siendo el más evidente la necesidad de acceso a datos en tiempo real. Esto trae consigo importantes complejidades como la necesidad de desarrollo de redes de baja latencia y sistemas muy complejos que ejecuten las órdenes en tiempos muy reducidos (Donadio & Ghosh, 2019). Es por ello por lo que, en contraste, este trabajo se enfocará en estrategias de trading algorítmico con una frecuencia diaria de

recopilación de datos. Este enfoque resulta más accesible para los inversores individuales, debido a la disponibilidad de APIs públicas gratuitas como la de Yahoo Finance, que facilitan enormemente la recopilación y adquisición de datos, así como el posterior desarrollo y verificación de los robots creados.

### **1.2.2 Algoritmos y técnicas más utilizados en el trading algorítmico**

El trading algorítmico durante los últimos años ha ido evolucionando significativamente, evolucionando desde estrategias basadas en reglas predefinidas hacia algoritmos elaborados que hacen uso de técnicas de aprendizaje automático.

En esta evolución, el trading algorítmico se ha basado durante sus primeras etapas en estrategias basadas en el análisis técnico, es decir, en una metodología que hace uso de los precios históricos de los valores de mercado, así como de representaciones gráficas de estos, con el objetivo de predecir futuros movimientos en los precios (Achelis, 2001). Es más, estas estrategias de análisis de patrones en los precios históricos vienen siendo utilizadas desde mucho antes de la aparición de los mercados bursátiles, como presentan Lo y Hasanhodzic (2010) quienes defienden que técnicas similares han sido explotadas desde civilizaciones antiguas como la babilónica o la griega, quienes ya contaban con registros de precios y análisis de los sentimientos de los comerciantes, con el objetivo de obtener un beneficio de ello por medio del trading y la especulación.

A lo largo de los años, muchos inversores han sido críticos con la idea de que el análisis técnico pudiese reportar rentabilidades positivas a quienes hiciesen uso de ellos. Sin embargo, estudios como el presentado por Brock, Lakonishok y LeBaron (1992), demostraron que, por medio del uso de técnicas simples de análisis técnico, como son las medias móviles, consiguieron obtener rendimientos positivos directamente atribuibles al uso de estas técnicas.

Dentro de estas técnicas de análisis técnico, algunos de los indicadores y técnicas más utilizados incluyen las ya mencionadas medias móviles, las Bandas de Bollinger (Bollinger, 1992), el índice de fuerza relativa (RSI) presentado originalmente por Wilder (1978) o el oscilador estocástico presentado por Lane (1984), entre otros. Estos indicadores buscan identificar tendencias, puntos de reversión y condiciones de

sobrecompra o sobreventa en los mercados financieros. Sin embargo, estos enfoques requieren una selección y configuración manual previa de los distintos parámetros que componen el indicador, lo que puede resultar en rendimientos subóptimos si no se realiza correctamente.

En los últimos años, los algoritmos de aprendizaje automático se han convertido en una herramienta cada vez más popular en el trading algorítmico debido a su capacidad para aprender de gran cantidad de datos financieros y adaptarse a nuevas condiciones del mercado (Kubat, 2017). Existen muchos algoritmos de aprendizaje automático diferentes que son utilizados con el objetivo de predecir los movimientos de mercado o los precios futuros de las acciones, todos ellos han conseguido demostrar a lo largo de múltiples estudios un buen rendimiento en los mercados bursátiles. Entre los algoritmos más utilizados se incluyen las máquinas de vectores de soporte (SVM) (Cao & Tay, 2001), los árboles de decisión (Tsai & Hsiao, 2010) o las redes neuronales (Rajihy, Nermend, & Alsakaa, 2017). Además, más recientemente, a raíz del desarrollo del aprendizaje profundo (*Deep Learning*), que utiliza redes neuronales de múltiples capas para modelar relaciones no lineales complejas, se ha demostrado que este tipo de modelos más complejos resulta ser especialmente útil para modelar los mercados financieros, al presentar estas estructuras muy cambiantes y altamente no lineales (Chong, Han, & Park, 2017).

Por último, el aprendizaje por refuerzo (*Reinforcement Learning*), una subárea del aprendizaje automático en la que los agentes aprenden a tomar decisiones óptimas a través de la interacción con su entorno, ha surgido como un enfoque prometedor para el trading algorítmico. El aprendizaje por refuerzo se diferencia de otros enfoques de aprendizaje automático en que se centra en aprender una secuencia de acciones que maximizan una recompensa acumulativa, en lugar de simplemente predecir una variable objetivo basándose en las características de entrada. Esto, de acuerdo con lo presentado por Fischer (2018), hace que el aprendizaje por refuerzo sea especialmente adecuado para el trading algorítmico, donde se pueden alinear perfectamente el problema de aprendizaje automático con los objetivos del inversor por medio de la modificación de recompensas y penalizaciones, en función del comportamiento con el que se quiera dotar al robot.

### **1.2.3 Trading Algorítmico en el S&P 500**

Como ya fue presentado en el contexto del trabajo, el índice S&P 500 es uno de los índices bursátiles más reconocidos y seguidos en el mundo, representando a 500 de las empresas más importantes de Estados Unidos y, por tanto, del mundo. Debido a esta gran relevancia, el S&P 500 y sus empresas que lo componen han sido objeto de numerosos estudios en el campo del trading algorítmico (Mailgaha Kumbure, Lohrmann, & Luukka, 2022).

Haciendo un repaso a la literatura existente, podemos encontrar múltiples estudios que exploran la capacidad de distintos algoritmos para predecir movimientos en el valor del índice S&P 500 de forma correcta. Estudios como el de Lohrmann y Lukka (2018), que hace uso de Bosques Aleatorios (*Random Forest*), una técnica que combina múltiples árboles de decisión; o el de Rosillo, Giner y De la Fuente (2013), que hace uso de máquinas de vectores de soporte (SVM) en combinación con el índice VIX (que mide la volatilidad del mercado).

Otros estudios, como el presentado por Zhang (2009), también se centran en la predicción del índice S&P 500 a través de técnicas avanzadas de redes neuronales. Este trabajo, sin embargo, presenta una importante diferencia con los anteriores mencionados, y es que, este último, se centra en los valores bursátiles que componen el índice, algo similar a lo que será realizado en el presente trabajo.

## **1.3 Objetivos**

El propósito general de este trabajo es estudiar diferentes algoritmos y estrategias de trading algorítmico y su implementación en robots de trading, es decir, en programas informáticos que ejecutarán órdenes de compra o venta en los mercados financieros siguiendo unas reglas o criterios predefinidos. El objetivo será analizar el funcionamiento, resultado, complejidad de ejecución y sobre todo rendimiento financiero de los diferentes robots de trading. Los objetivos específicos serán los siguientes:

- Desarrollar e investigar distintas técnicas de predicción de valores en series temporales haciendo uso de técnicas provenientes del análisis técnico como son el uso de medias móviles y Bandas de Bollinger.

- Conocer en profundidad e implementar técnicas para la predicción de precios de mercado que hagan uso de algoritmos más complejos como son los que hacen uso de técnicas de aprendizaje por refuerzo.
- Desarrollar robots de trading en base a teorías usadas en la valoración de activos financieros, como es el modelo CAPM.
- Comparar el desempeño entre los diferentes robots de trading que implementan estas técnicas, utilizando como criterios de evaluación la rentabilidad, la volatilidad, el *Drawdown* máximo (MDD) y el Ratio de Sharpe.

## 1.4 Metodología

La metodología que se va a emplear en este trabajo combina elementos cuantitativos y cualitativos, aunque predomina el enfoque cuantitativo, ya que se basará en la programación de distintos robots de trading que implementarán estrategias, pero también se comentará sobre sus resultados y posibles puntos fuertes o débiles. Los pasos que se seguirán son los siguientes:

- Revisión de literatura sobre técnicas de trading algorítmico empleadas en la actualidad, así como de diferentes modelos de Machine Learning que se puedan usar en la predicción de valores de acciones. Para ello, se hará una búsqueda sistemática en bases de datos académicas y se seleccionarán los artículos más relevantes y recientes sobre el tema.
- Realización de las prácticas que propone el tutor de este trabajo, comprendiendo e indagando en los distintos modelos y robots que se proponen en ellas.
- Desarrollo en Python de los distintos robots de trading haciendo uso de las técnicas y algoritmos aprendidos. Para el desarrollo de estos robots, se hará uso de cuadernos de Jupyter, un tipo de archivo que facilitará el compartir los resultados con el tutor, ya que permiten presentar las salidas que se obtienen de la ejecución de los programas.

- Una vez estos robots sean desarrollados, tratarán de ser optimizados con el objetivo de conseguir un mejor desempeño.
- Se utilizarán datos de mercado provenientes principalmente de la base de datos de *Yahoo Finance*, haciendo uso de su API a través de la librería “yfinance” disponible en Python de manera gratuita. Se descargarán datos de los distintos valores que componen el índice S&P 500 a cierre de octubre de 2022, para el periodo contenido desde inicios del año 2000 hasta esta misma fecha.
- Una vez desarrollados y optimizados estos robots, su comportamiento y desempeño será comparado a través de diferentes indicadores estadísticos.

## **1.5 Estructura del Trabajo**

Con el fin de proporcionar una visión general de la organización del trabajo, en esta sección se presenta la estructura principal que seguirá el mismo.

En primer lugar, en el capítulo 2 del trabajo se responderá al primer objetivo concreto del trabajo, en el que se hará uso de robots de trading basados en medias móviles y Bandas de Bollinger. En este segundo capítulo del trabajo también se buscará la optimización de estos robots por medio del ajuste de parámetros y otras técnicas de optimización.

En el tercer capítulo de este trabajo se abordará el desarrollo y optimización de un robot basado en Q-Learning, un algoritmo de aprendizaje por refuerzo con el que se cumplirá el segundo objetivo específico de este trabajo.

Durante el cuarto capítulo, se desarrollará y se entrará a conocer un robot de trading que estará basado en una de las teorías más conocidas en el ámbito financiero: el modelo de valoración de activos de capital (CAPM).

Para la respuesta al cuarto objetivo específico del trabajo, a lo largo de los cuatro capítulos mencionados se irá haciendo un análisis del desempeño de los distintos robots con un criterio unificado. Finalmente, en las conclusiones del trabajo se abordarán de nuevo estos resultados para mostrar las diferencias en rendimiento de manera unificada, así como las diferencias con el mercado.

## **CAPÍTULO 2: OPTIMIZACIÓN DEL ROBOT DE MEDIAS Y BANDAS**

### **2.1 Robot de Trading Básico: Medias Móviles**

#### **2.1.1 Presentación del robot de trading**

Este análisis de distintos robots de trading se inicia con un primero cuyo comportamiento está guiado por la utilización de medias móviles, una modalidad considerada de las más elementales dentro del universo de los algoritmos de trading por la teoría que soporta su funcionamiento y por su facilidad de implementación. Como ha sido indicado previamente, el robot hará uso de los distintos valores recogidos en el índice bursátil S&P 500 durante su operativa.

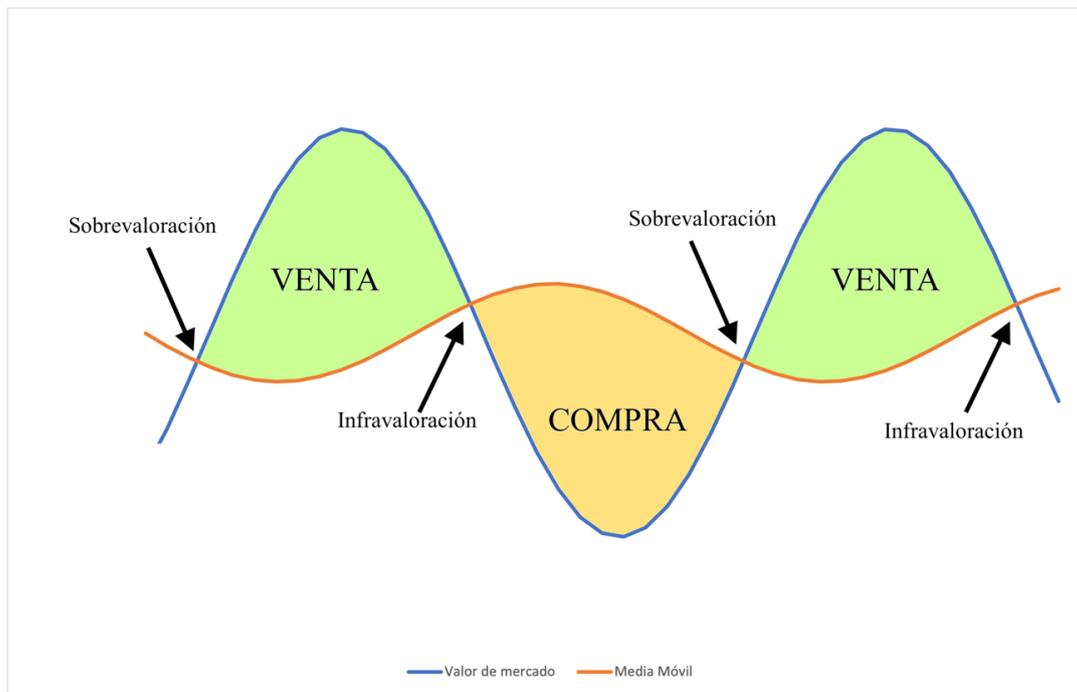
Las medias móviles son herramientas estadísticas ampliamente utilizadas en un gran abanico de campos. Estas permiten suavizar fluctuaciones a corto plazo y revelar tendencias subyacentes en los datos. En términos más simples, la media móvil es una medida del precio medio de un valor financiero durante un determinado periodo de tiempo (o ventana de tiempo). Es por este motivo, por su capacidad de identificar tendencias, por lo que es un indicador ampliamente utilizado en trading, ya que si, por ejemplo, el precio de un activo se encuentra por encima de su media móvil, indicaría que está a un precio superior al habitual, una situación que, sin una razón financiera detrás, sería un indicativo de una sobrevaloración del mismo y, por lo tanto, haría de la venta de ese activo algo razonable.

Este primer robot de trading, por tanto, emplea una estrategia basada en medias móviles, de manera que cuando la media móvil de un activo determinado (en este caso las acciones que constituyen el índice S&P 500) está por encima del precio actual del valor, el robot interpreta esta situación como un indicio de que el activo está siendo infravalorado por el mercado y, en consecuencia, genera una señal de compra. La razón tras esta estrategia radica en la hipótesis de que los valores tienden a igualar a su media móvil y, por tanto, si el precio actual de una acción es menor que su media móvil, el precio de la acción debería sufrir un “rebote” para volver al nivel de medio de precios.

De manera inversa, cuando la media móvil se sitúa por debajo del precio actual de mercado, el robot interpreta esta situación como una señal de que el activo está siendo sobrevalorado por el mercado y, por tanto, genera una señal de venta, siguiendo una lógica simétrica a la explicada para el caso de infravaloración y esperando por tanto una bajada

a corto plazo en el precio del activo para igualar a su media móvil. En la siguiente figura (Figura 1), se puede observar una ilustración simplificada que ejemplifica el funcionamiento de este algoritmo.

*Figura 1: Ilustración de la estrategia de trading de reversión de la media*



*Fuente: Elaboración propia*

Este tipo de estrategias, sin embargo, presentan algún riesgo asociado, y es que, se fundamentan puramente en la teoría de la reversión de la media, en la que se sostiene que el precio de un activo tiende a bajar (subir) después de alcanzar un máximo (mínimo), volviendo por tanto a un precio promedio (Exley, Mehta, & Smith, 2004). En esta aproximación, por tanto, se llevan a cabo operaciones de compra y venta sin tener en cuenta si estos cambios en el precio de los activos se fundamentan en una razón económica, como podría ser el anuncio de resultados negativos por parte de una compañía, o si son puras anomalías temporales que se espera que se corrijan a corto plazo.

### 2.1.2 Descripción en profundidad del funcionamiento del robot

Al igual que sucederá con todos los robots que se desarrollarán a lo largo de este trabajo, este robot de trading comienza su actividad con un capital inicial de 10 millones de dólares. Este capital será destinado a la inversión en los distintos activos que componen el índice bursátil S&P 500 a fecha de cierre de octubre de 2022, siendo dividido uniformemente entre todos estos valores y utilizado para operar sobre ellos durante el periodo de *backtesting*, que abarca desde el inicio de 2017 hasta finales de octubre de 2022. Estos valores, concretamente 505 valores, serán asignados inicialmente con un capital de 19.801,98 dólares y, por lo tanto, el rendimiento en uno de estos valores no afectará al rendimiento en el resto, ya que el capital dedicado a cada uno de ellos estará delimitado.

Este robot hará uso de datos de cierre de mercado diarios y, por lo tanto, realizará como máximo una operación al día por cada uno de los valores que componen el índice.

En cuanto al funcionamiento concreto del robot, cuando este detecte que el precio de un activo está por debajo de su media móvil, inmediatamente emitirá una orden de compra, destinando el 100% del capital reservado para ese valor a la adquisición de acciones. Por ejemplo, para la primera adquisición de cualquiera de los valores, el robot adquirirá acciones valoradas en 19.801,98 dólares.

De manera inversa, cuando la media móvil del precio de una acción sea inferior al precio actual de mercado, el robot generará una señal de venta. Cuando se encuentre ante esta situación, el robot tomará una posición corta, es decir, venderá el total del valor asignado a esa acción asumiendo un margen del 100% en la operación, o explicado de otra forma, venderá acciones por valor del capital disponible para ese determinado activo. Este comportamiento supera los requerimientos mínimos impuestos por la Reserva Federal (FED) de los Estados Unidos, que estipulan que el margen en este tipo de operaciones debe ser de al menos el 50% (Federal Reserve System, 1998).

Es importante destacar que este robot incorpora un mecanismo para el control del riesgo en las operaciones en las que se posiciona en corto. En particular, el robot está diseñado para responder de manera proactiva a las denominadas “llamadas de margen” o “*margin calls*”. Estas llamadas de margen existen como un requisito establecido por la FED para

que los inversores que utilizan apalancamiento en sus operaciones incrementen el capital comprometido en aquellas operaciones que están yendo en su contra. Es decir, cuando una posición en corto aumenta su valor de manera significativa, en torno al 53,8% de su valor inicial, debido al requisito de mantener un margen del 30% (Financial Industry Regulation Authority, 2022), el robot automáticamente cerrará la posición, limitando de este modo el riesgo de pérdidas potencialmente ilimitadas que son inherentes a las posiciones en corto. Este mecanismo es muy importante para que el robot mantenga un nivel de riesgo asumible, ya que, mientras que en una operación de compra las pérdidas están limitadas (el valor de una acción no puede ser inferior a cero), en las posiciones cortas el riesgo de pérdida puede ser, teóricamente, infinito, dado que el precio de una acción no tiene un límite superior, pudiendo resultar catastrófica una subida acelerada del precio de un activo mientras se tiene una posición en corto abierta para el mismo.

Por último, se ha llevado a cabo una serie de pruebas de funcionamiento del robot en las cuales se ha ido variando el tamaño de la ventana temporal en la configuración de la media móvil. En el siguiente apartado de este trabajo, se discutirá con mayor detalle cómo estas modificaciones en los parámetros del robot han afectado su rendimiento.

### **2.1.3 Resultados del robot de trading**

A continuación, se entrará en detalle en cómo se ha ido comportando el robot de trading al variar el tamaño de la ventana temporal para el cálculo de la media móvil. Para realizar este análisis, se pondrá el foco en un conjunto específico de métricas con las que se evaluará el funcionamiento del robot. Estos *KPIs*, o métricas seleccionadas, son las siguientes: Rendimiento (comparando rendimientos totales y anualizados), volatilidad anualizada, el máximo *drawdown* (MDD) y el Ratio de Sharpe. Es importante recalcar que estas métricas no solo proporcionan información sobre el rendimiento generado por el robot, sino que también permiten evaluar el nivel de riesgo asociado a su operativa.

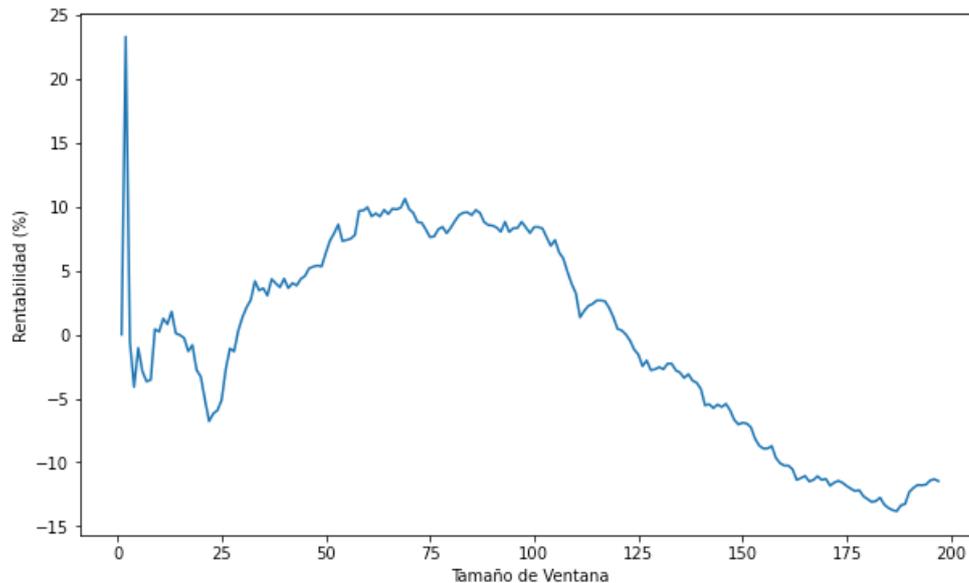
Estas métricas, además, serán utilizadas a lo largo de este trabajo para evaluar el funcionamiento de los distintos robots, por lo que antes de comenzar a utilizarlas, se va a proceder a hacer una descripción de la relevancia y funcionamiento de las mismas.

- Rendimiento: Se observarán los rendimientos totales y los rendimientos anualizados. A través del análisis del rendimiento del robot, se podrá observar la rentabilidad (positiva o negativa) que consiguen los diferentes robots a lo largo del periodo de *backtesting* estudiado. Para facilitar la comparación con inversiones de todo tipo, también se muestra la rentabilidad anualizada, algo común en el sector financiero en el que lo más común es presentar la rentabilidad de un proyecto en términos anuales.
- Volatilidad anualizada: La volatilidad, que indicará la dispersión de los rendimientos del robot de trading para los diferentes días, se utilizará como una medida del riesgo del mismo. Esta volatilidad es una medida común del riesgo, ya que, cuanto mayor volatilidad presenta un activo, más grandes serán las variaciones en el rendimiento del mismo, tanto en sentido positivo como negativo.
- Máximo *drawdown* (MDD): El MDD representa la pérdida máxima (en porcentaje) que habría experimentado el robot si hubiese empezado a operar durante un periodo de tiempo específico que abarca desde el periodo con el máximo capital hasta el punto con menor capital. Esta métrica también proporciona una idea del riesgo que está afrontando el robot, pero a diferencia de la volatilidad que indica la variación respecto a una media de precio, esta medida permite tener una visión sobre el riesgo de valores negativos, mostrando la máxima caída ocurrida durante un periodo estudiado. Por lo tanto, un robot que no sufriese ninguna pérdida durante el dicho periodo de tiempo tendría un MDD de 0%, mientras que un robot que pierde todo el capital durante ese mismo periodo mostraría un MDD del -100%.
- Ratio de Sharpe: Como última medida se hará uso del Ratio de Sharpe (Sharpe, 1994). Esta es una medida que permite evaluar la eficiencia de una estrategia de inversión en términos de rendimiento ajustado por riesgo. El Ratio de Sharpe, por tanto, hace uso de dos de las métricas mencionadas previamente en esta sección, la rentabilidad (o rendimiento) y la volatilidad. Este ratio se calcula como el rendimiento medio en exceso (es decir, el rendimiento por encima de la tasa libre de riesgo) dividido por la volatilidad de los rendimientos en exceso. Debido a esto,

un robot de trading que muestre un Ratio de Sharpe más alto indicará que está consiguiendo generar una mayor rentabilidad por unidad de riesgo asumida, algo ideal, ya que lo que se persigue al tomar decisiones de inversión es incrementar el rendimiento limitando lo máximo posible el riesgo asumido en ellas.

Antes de entrar en detalle en el análisis de cada una de estas métricas, se ha decidido hacer un análisis previo sobre qué tamaños de ventana para el cálculo de la media móvil proporcionan una mayor rentabilidad durante el periodo de *backtesting*. Los resultados sobre esta relación se muestran en la siguiente figura (Figura 2).

*Figura 2: Análisis de rentabilidad variando tamaño de la ventana temporal para el cálculo de la media móvil*



*Fuente: Elaboración propia*

De este análisis destacan principalmente 2 valores de tamaño de ventana temporal que presentan rentabilidades superiores sobre ventanas temporales de tamaños similares. Sorprendentemente, al hacer uso de una ventana temporal de 2 días, es decir, una que está compuesta por el día actual y el día anterior, es cuando mejores resultados obtiene el robot. Este resultado es sorprendente ya que, debido al pequeño tamaño de la ventana, este robot funciona simplemente comprando si el precio del día actual es menor que el precio de cierre del día anterior y vendiendo en el caso contrario. Yendo a valores más altos para el tamaño de esta ventana, más en línea con la teoría de la reversión de la media

a la que se venía haciendo referencia, se observa que un valor de 69 días, lo que equivaldría a poco más de tres meses de trading, también representa un máximo local, presentando un mejor rendimiento que otros valores cercanos a este. A pesar de ello, este valor está muy alejado del máximo absoluto encontrado en la ventana temporal de 2 días, obteniendo un rendimiento de aproximadamente la mitad.

En la siguiente tabla (Tabla 1), se puede observar cómo estos tamaños de ventana más “óptimos” se han comportado en las diferentes métricas estudiadas, así como ilustraciones acerca de la evolución del robot a lo largo de este tiempo, observando la cantidad invertida del capital que tenía en cada momento del tiempo.

*Tabla 1: Análisis del comportamiento del robot de medias móviles*

Ventana temporal	Indicadores de Rendimiento	Evolución
2 días	Rdto. Total: 23,29% Rdto. Anualizado: 3,66% Volatilidad Anual.: 14,84% MDD: -20,41% Ratio de Sharpe: 0,25	
69 días	Rdto. Total: 10,61% Rdto. Anualizado: 1,83% Volatilidad Anual.: 18,92% MDD: -37,08% Ratio de Sharpe: 0,14	

*Fuente: Elaboración propia*

Como había sido adelantado previamente, los resultados que se obtienen son sorprendentes, especialmente cuando se utiliza una ventana temporal de 2 días para el cálculo de la media móvil. Este hallazgo es interesante no solo por el hecho de que esta configuración proporciona la mayor rentabilidad entre todas las ventanas temporales

probadas, sino también porque presenta un riesgo significativamente menor en comparación con ventanas temporales más amplias. Este menor riesgo se evidencia tanto en términos de volatilidad, que es menor a la presentada en ventanas superiores, como en términos del máximo *drawdown* (MDD), siendo este para la ventana de 2 días aproximadamente la mitad del MDD observado para la ventana de 69 días. Esto se traduce en una aparente mayor resistencia a los periodos de crisis financieras, como la que se observó en el inicio de 2020 debido a la pandemia de COVID-19, la cual tuvo un menor impacto sobre este robot.

La combinación de una alta rentabilidad y un nivel moderado de riesgo convierten a esta configuración del robot en la mejor alternativa entre las estudiadas durante este primer robot, presentando un Ratio de Sharpe de 0,25, significativamente superior al obtenido para las otras ventanas temporales.

Además, profundizando en cómo la modificación de la ventana temporal afecta al funcionamiento operativo del robot, se puede observar que, a medida que la ventana temporal es más corta, la cantidad de operaciones realizadas por el robot aumenta, observando fluctuaciones más frecuentes y pronunciadas en la cantidad de capital invertido.

## **2.2 Sofisticación del robot de trading: Bandas de Bollinger**

### **2.2.1 Presentación del robot de trading**

El segundo robot que será analizado en este trabajo hará uso de unas herramientas de análisis técnico llamadas Bandas de Bollinger, ligeramente presentadas en la revisión bibliográfica de este trabajo, y que se tratan de una sofisticación de la herramienta presentada en el anterior robot, las medias móviles.

Las Bandas de Bollinger son un conjunto de tres líneas que se presentan en relación con los precios de mercado de un activo. Estas líneas son la ya conocida media móvil, una banda superior y otra banda inferior. Estas bandas superior e inferior se basan en la combinación de una media móvil simple, como las descritas en el primer robot de este trabajo, junto con la desviación estándar móvil del precio del activo. Estas bandas proporcionan información adicional con respecto a una media móvil simple, ya que

permiten tener una visión de cuáles pueden ser valores normales de un determinado activo basándonos en su volatilidad histórica.

Con todo ello, y al igual que ocurría en el robot anterior, estas bandas darán indicios de si un título está sobrevalorado o infravalorado, lo que llevará a su venta o compra consecuentemente. Con este robot se persigue una teoría muy similar a la utilizada en el anterior robot, la teoría de la reversión de la media, en la que se espera que un valor fluctúe alrededor de un precio medio.

En su trabajo original, Bollinger (1992) presenta las fórmulas para el cálculo de estas bandas. Estas se definen como la media móvil a la que se le suman o se le restan dos desviaciones estándar móviles, correspondiendo a la banda superior e inferior respectivamente. Sin embargo, en este trabajo, con el objetivo de obtener un robot lo más optimizado posible, se evaluarán tanto el número de desviaciones estándar que se utilizarán para el cálculo de las bandas, como el tamaño de la ventana temporal empleada para dichos cálculos.

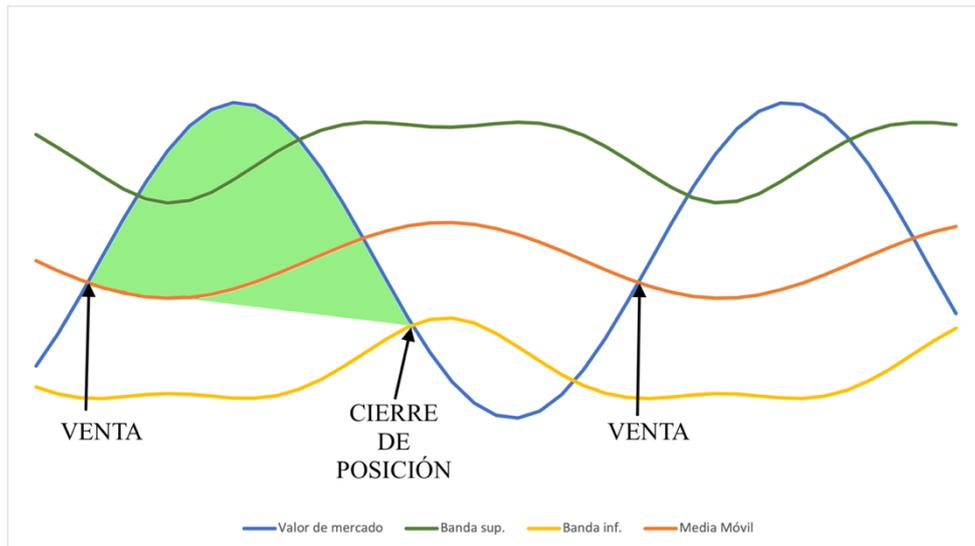
### **2.2.2 Descripción en profundidad del funcionamiento del robot**

Tal y como fue introducido en el robot basado en medias móviles, este robot también comenzará su operativa con un capital inicial de 10 millones de dólares, los cuales se utilizarán para la inversión en los diferentes valores que constituyen el índice bursátil S&P 500. Este robot, de igual manera que ocurría con su versión antecesora, distribuirá de manera equitativa el total del capital inicial entre todos los valores que componen el índice mencionado. De este modo, se asignará un capital inicial de 19.801,98 dólares a cada valor, manteniendo delimitado el capital para cada uno de ellos durante su operativa, y, por tanto, manteniendo los resultados entre los distintos valores independientes.

El robot se basará en la siguiente estrategia: cuando el precio actual de un valor bursátil esté por encima de la “banda media” (la media móvil), procederá a posicionarse en corto (venta) en el activo en cuestión, entendiendo una sobrevaloración de dicho activo. Una vez tenga esta posición de venta abierta, aguantará esta posición hasta que el valor decrezca por debajo de la banda de Bollinger inferior, consiguiendo aprovechar un potencial beneficio mayor al que se podría conseguir con la estrategia original de medias

móviles, en la que se vendería una vez se cruzase de nuevo la media móvil. Este funcionamiento descrito se muestra de manera simplificada en la siguiente Figura 3.

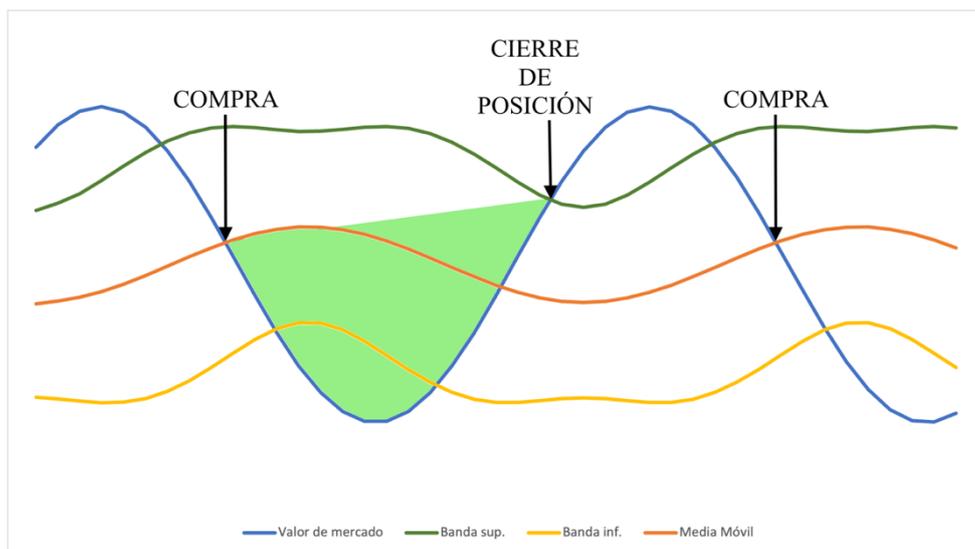
*Figura 3: Ilustración de la estrategia de venta con Bandas de Bollinger*



*Fuente: Elaboración propia*

En el caso contrario, en el que el precio del activo en cuestión se encuentre por debajo de la media móvil, el robot procederá a comprar, entendiéndolo una infravaloración del mismo. Simétricamente al caso anterior, cuando el precio del activo supere el límite de la banda superior, el robot cerrará la posición, como se ilustra en la Figura 4.

*Figura 4: Ilustración de la estrategia de compra con Bandas de Bollinger*



*Fuente: Elaboración propia*

Haciendo uso de la estrategia descrita, para el caso en el que el robot realice una orden de compra, este procederá a comprar acciones de esa compañía utilizando todo el capital disponible para dicho valor, mientras que, en el caso de tratarse de una orden de venta, el robot procederá a vender acciones por valor del 100% del capital disponible en ese momento, es decir, realizará una operación en corto usando un margen del 100%.

Al igual que se incorporó en el robot de medias móviles un mecanismo de gestión del riesgo basado en las llamadas de margen (*margin calls*), esta característica también estará disponible en este robot, protegiendo al mismo de la posibilidad de incurrir en pérdidas infinitas.

### **2.2.3 Resultados del robot de trading**

Con el fin de conseguir una optimización en el desempeño de este robot y lograr un rendimiento máximo, se ha realizado un análisis del funcionamiento del robot realizando variaciones tanto en el tamaño de la ventana temporal utilizada para el cálculo de la media móvil y las desviaciones estándar móviles, como en el número de desviaciones estándar usadas para el cálculo de las Bandas de Bollinger superior e inferior. Dos parámetros, por tanto, claves para el funcionamiento del robot, ya que guiarán por completo el comportamiento de este. Para la realización de este análisis de los parámetros, se han seleccionado 7 valores distintos como número de desviaciones típicas que serán utilizados para el cálculo de las bandas, cubriendo el rango desde 0,5 desviaciones hasta 3,5, incrementando en saltos de 0,5 los distintos valores estudiados, y por tanto haciendo que el límite para el cierre de posiciones sea cada vez más exigente. Para el número de días utilizados como ventana temporal para el cálculo de las medias móviles y desviaciones típicas móviles, se ha usado un rango desde 4 días hasta 100 días.

En este punto es necesario subrayar que, a pesar de que en el anterior robot de medias móviles el robot más equilibrado haya sido el que hacía uso de una ventana de 2 días, en esta ocasión no se podrá probar con ese valor por una cuestión matemática, y es que, a pesar de que el robot sí que podría comprar los valores al poder encontrarse su precio por debajo de la media, nunca podría llegar a cerrar una posición. Esto es debido a que es matemáticamente imposible que la suma entre la media de dos valores y su desviación típica multiplicada por un factor  $Z$ , lo que define a la Banda de Bollinger superior, sea

inferior a alguno de los dos valores contenidos en la ventana temporal, salvo que este factor  $\mathcal{Z}$  fuese inferior a 1. Ocurriendo lo mismo simétricamente para el escenario en el que el robot se posiciona en una operación en corto. Se puede observar esta afirmación a través de las siguientes ecuaciones:

Siendo la media entre dos valores  $x$  e  $y$ :

$$\mu = \frac{x + y}{2}$$

Y la desviación típica entre estos valores:

$$\sigma = \sqrt{\frac{(x - \mu)^2 + (y - \mu)^2}{2}}$$

Para cumplirse la condición de venta, en el caso de estar en una posición en largo (habiendo comprado), se tendría que dar que la suma entre la media y la desviación típica multiplicada por un número  $\mathcal{Z}$  de desviaciones estándar, fuese menor que el segundo valor (precio actual), o lo que es lo mismo, se tendría que cumplir la siguiente inecuación:

$$\frac{x + y}{2} + \mathcal{Z} * \sqrt{\frac{(x - \frac{x + y}{2})^2 + (y - \frac{x + y}{2})^2}{2}} < y$$

Sin embargo, esta inecuación solo presenta resultados para el caso en el que el número de desviaciones estándar aplicado ( $\mathcal{Z}$ ), fuese menor que 1 (o -1):

$$\begin{aligned} y > x \quad y \quad \mathcal{Z} < 1 \quad y \quad x \in \mathbb{R} \\ y < x \quad y \quad \mathcal{Z} < -1 \quad y \quad x \in \mathbb{R} \end{aligned}$$

Esta misma limitación, aunque presentando una ligera mayor complejidad, se puede observar también para el caso de una ventana temporal de 3 días. En esta situación, se evidencia que no sería posible cerrar la posición en ningún momento si el número de desviaciones estándar escogido fuese superior a 1.

Para este análisis de parámetros, solo uno de los 7 valores estudiados como número de desviaciones estándar cumple la inecuación para el caso de la ventana temporal de 2 días, el que hace uso de 0,5 desviaciones típicas; mientras que únicamente 2 casos cumplen la inecuación para la ventana temporal de 3 días (el de 0,5 y 1 desviación estándar). Con el objetivo de no distorsionar el análisis general con múltiples robots que no tienen un sentido práctico, al no poder cerrar posiciones, estos 3 casos son evaluados separadamente en la siguiente tabla (Tabla 2).

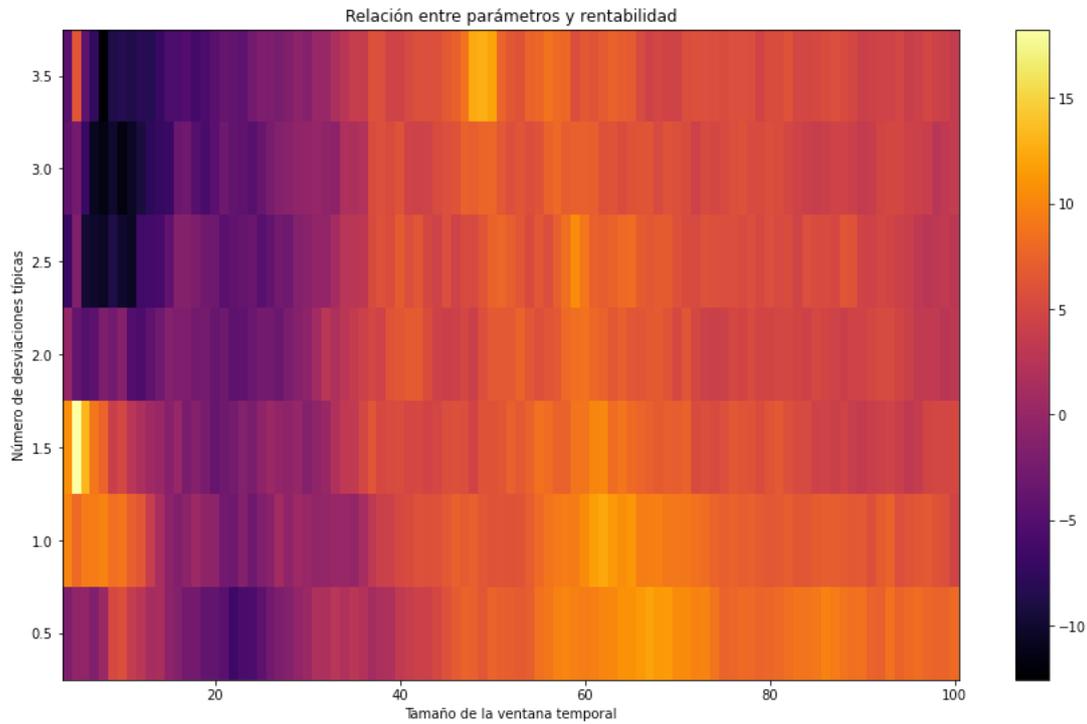
*Tabla 2: Rentabilidad del robot de Bandas de Bollinger haciendo uso de las combinaciones de parámetros no incluidas en el análisis general*

<b>Combinación de parámetros</b>	<b>Rentabilidad</b>
Ventana temporal: 2 días Desviaciones estándar: 0,5	Rentabilidad total: 12,87%
Ventana temporal: 3 días Desviaciones estándar: 0,5	Rentabilidad total: -4,02%
Ventana temporal: 3 días Desviaciones estándar: 1	Rentabilidad total: 7,30%

*Fuente: Elaboración propia*

A continuación, se muestra gráficamente los resultados en términos de rentabilidad que han resultado de la combinación de los distintos parámetros, mostrando un color más claro aquellas combinaciones que han resultado más rentables y un color más oscuro las que menos (Figura 5).

*Figura 5: Análisis de rentabilidad según la combinación de parámetros para el robot de Bandas de Bollinger*



*Fuente: Elaboración propia*

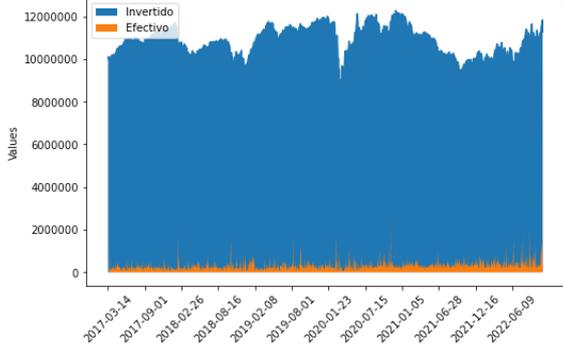
De la Figura 5 anterior, se puede extraer mucha información, gran parte de ella poco intuitiva si no se hubiera realizado este análisis. En primer lugar, se observa como el máximo de rentabilidad de este robot es alcanzado cuando se le aplica una ventana temporal de 5 días en combinación de 1,5 desviaciones estándar. Además, se pueden observar varias áreas de concentración de altas rentabilidades, como la observada alrededor del valor máximo mencionado u otras en selecciones muy variadas de los parámetros, como la encontrada para la combinación de 49 días de ventana temporal y 3,5 desviaciones típicas.

Es destacable como la interrelación entre ambos parámetros afecta en gran medida el desempeño del robot. Esta cuestión es fácilmente observable para valores de ventana temporal pequeños, menores a 15 días, en los que se encuentran tanto las rentabilidades más altas observadas como las más bajas, siendo el número de desviaciones típicas el único parámetro que varía. Además, es fácilmente observable la presencia de un valle de rentabilidad para tamaños de ventana medianos (alrededor de entre 15 y 40 días), presentando un rendimiento significativamente inferior al presentado por tamaños de ventana más pequeños y más grandes.

A continuación, y por medio de las métricas previamente seleccionadas, se realizará un análisis en profundidad sobre el funcionamiento de estos robots, escogiendo para este análisis a los 4 robots que más han destacado en cuanto a rentabilidad en los análisis preliminares. Los resultados en profundidad de estos robots se presentan en la siguiente tabla (Tabla 3).

*Tabla 3: Análisis del comportamiento del robot de Bandas de Bollinger*

Ventana temporal	Desv. Estándar	Indicadores de Rendimiento	Evolución
5 días	1,5	Rdto. Total: 18,26% Rdto. Anualizado: 2,93% Volatilidad Anual.: 18,19% MDD: -31,22% Ratio de Sharpe: 0,19	
6 días	1,5	Rdto. Total: 13,14% Rdto. Anualizado: 2,15% Volatilidad Anual.: 18,25% MDD: -33,20% Ratio de Sharpe: 0,15	
2 días	0,5	Rdto. Total: 12,87% Rdto. Anualizado: 2,10% Volatilidad Anual.: 15,98%% MDD: -22,86% Ratio de Sharpe: 0,15	

49 días	3,5	Rdto. Total: 12,80% Rdto. Anualizado: 2,16% Volatilidad Anual.: 19,22% MDD: -32,90% Ratio de Sharpe: 0,15	
---------	-----	---	--

*Fuente: Elaboración propia*

De los robots analizados, hay uno que sobresale claramente sobre el resto: el que hace uso de 1,5 desviaciones típicas en conjunto con 5 días de ventana temporal para el cálculo de las Bandas de Bollinger. Este robot no solo supera a los demás en términos de rentabilidad, sino que también conserva un nivel de riesgo similar al del resto, otorgándole, por tanto, del mejor Ratio de Sharpe. No obstante, es importante destacar que al igual que sucedía con el primer robot basado en medias móviles, el robot con el riesgo asociado más bajo, tanto en términos de volatilidad como de MDD, es el que hace uso de una ventana temporal de 2 días para el cálculo de las bandas.

A pesar de sobresalir entre los otros que hacen uso de Bandas de Bollinger, este robot no ha sido capaz de superar a su versión antecesora, el robot que solamente hacía uso de medias móviles. Este desempeño inferior de la versión más sofisticada del robot hace que se plantee la cuestión de si este robot es apto para todas las condiciones de mercado, tal como se está utilizando actualmente, o si se debería hacer uso del mismo en condiciones específicas, como en las situaciones que cumplan con la hipótesis de la reversión de la media. Esta cuestión se abordará en la siguiente sección de este capítulo.

## 2.3 Optimización del robot de trading

### 2.3.1 Descripción de la optimización

Con el objetivo de continuar con la optimización del primer robot de este trabajo, el robot basado en medias y bandas, se ha tomado como base una idea muy extendida en el campo del análisis técnico y que se fundamenta en un concepto fácilmente entendible. La idea defiende que este tipo de robots ofrecen un mejor funcionamiento cuando se encuentran ante condiciones de mercado que presentan una tendencia lateral, o lo que es lo mismo, en las que los precios se mantienen relativamente estables a lo largo del tiempo sin presentar pendientes pronunciadas en ninguna dirección (Li, Zhao, Gopalkrishnan, &

Hoi, 2012). Estas condiciones de mercado, por tanto, favorecerían el cumplimiento de la hipótesis de la reversión de la media, en la que se fundamenta este robot.

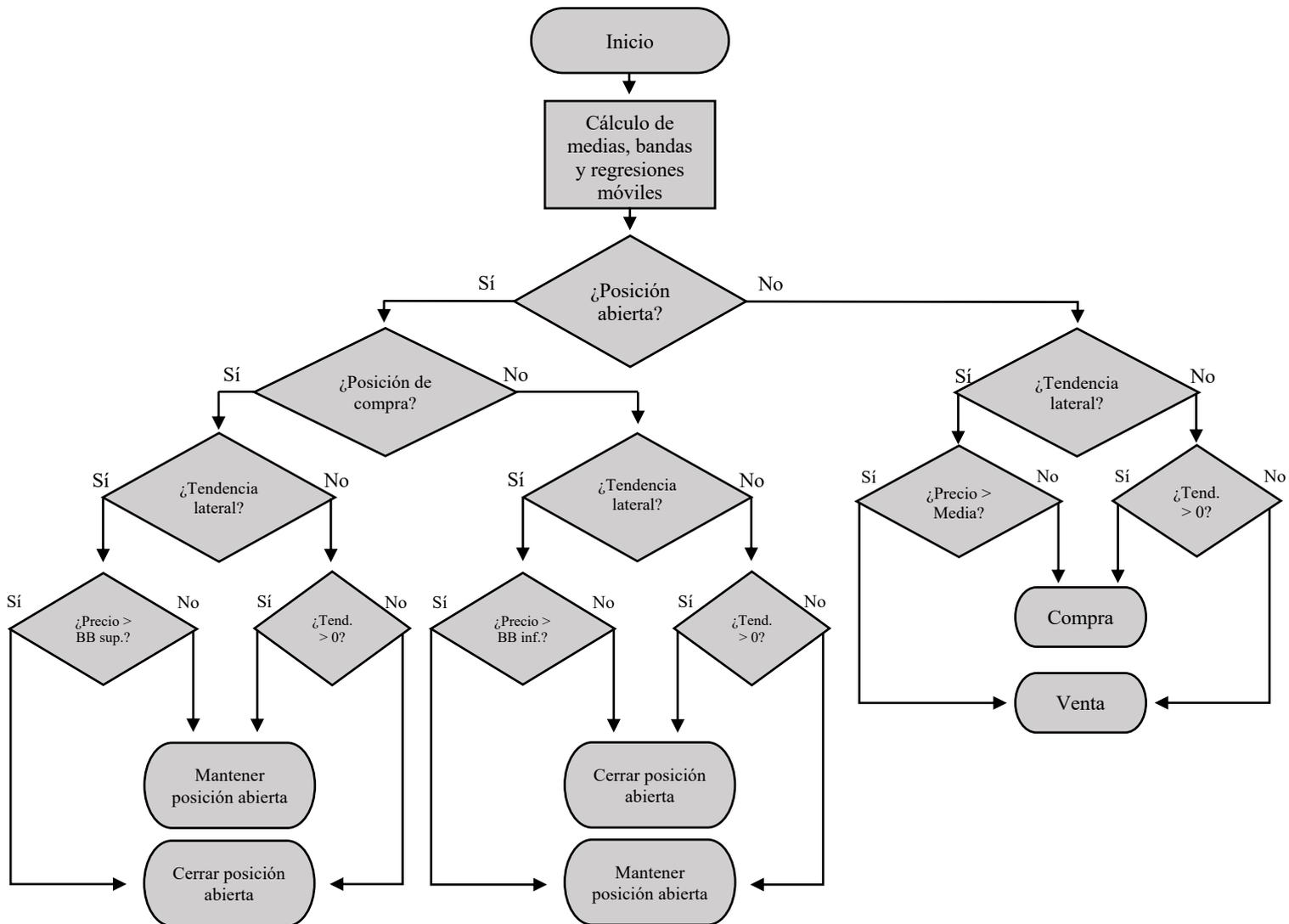
Para incorporar esta teoría en el robot de trading, se han realizado una serie de modificaciones en la lógica que sigue el mismo para tomar decisiones. En primer lugar, además de calcular las medias móviles y las Bandas de Bollinger de los diferentes activos, en esta ocasión, también se realizarán regresiones lineales móviles con una ventana de tiempo concreta, para determinar si el comportamiento del activo está siguiendo una tendencia lateral o no. A raíz de esta modificación, ha sido imprescindible añadir dos parámetros nuevos al robot, cuya combinación será optimizada más adelante:

- El tamaño de la ventana temporal para el cálculo de la regresión lineal móvil.
- Un umbral de pendiente por la que una tendencia dejará de considerarse tendencia lateral.

En base a estos nuevos parámetros, cuando el algoritmo detecte la presencia de tendencia lateral, el comportamiento que seguirá el robot será idéntico al descrito previamente en este capítulo, capitalizando sobre la hipótesis de la reversión de la media. Sin embargo, cuando el robot se encuentre ante una tendencia alcista, marcada por una pendiente en la regresión móvil superior al umbral definido, procederá a comprar, siendo esta la mejor estrategia en este tipo de condiciones. De manera simétrica, cuando el robot detecte la presencia de una tendencia bajista, procederá a la apertura de posiciones en corto.

Para facilitar la comprensión del algoritmo lógico que seguirá el robot, a continuación, se proporciona un diagrama que ilustra gráficamente su funcionamiento (Figura 6).

Figura 6: Diagrama de flujo del proceso algorítmico optimizado para el robot de medias y bandas



Fuente: Elaboración propia

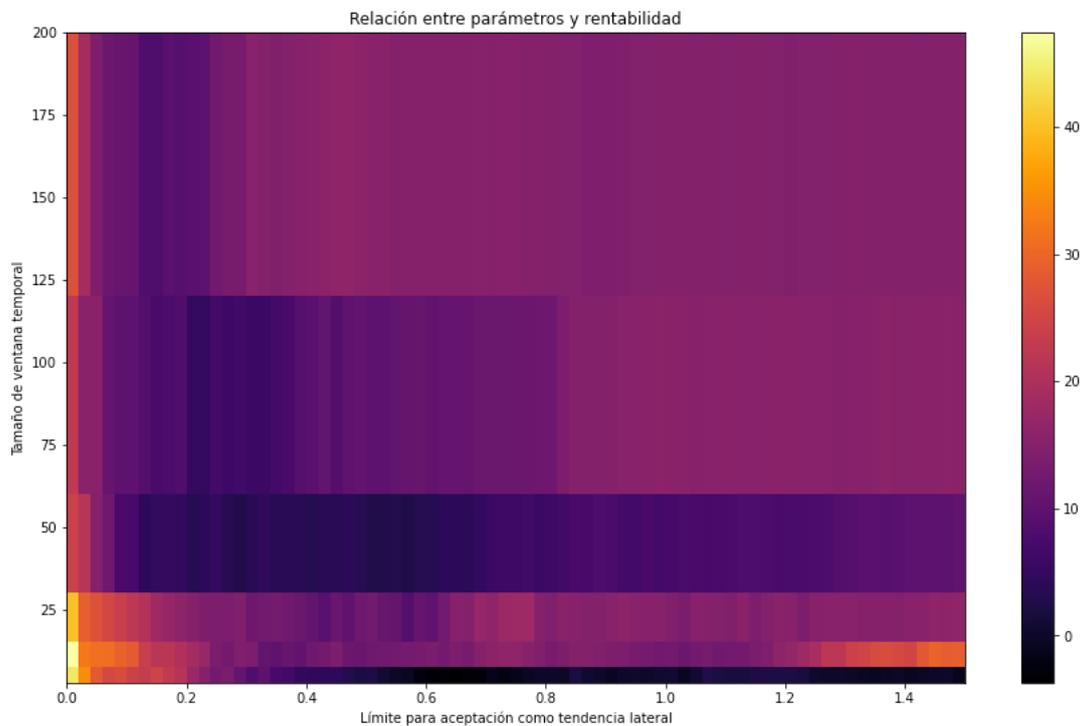
### 2.3.2 Resultados de la optimización

Para el desarrollo de esta versión optimizada del robot de trading basado en medias y bandas, se ha empleado como punto de partida el robot que presentó un mayor rendimiento entre los robots más sofisticados, los que hacían uso de las Bandas de Bollinger. En concreto se ha hecho uso del robot que utiliza 5 días para el cálculo de las

medias y bandas móviles y que emplea 1,5 desviaciones típicas para el cálculo de las bandas.

Partiendo de este modelo, se ha puesto foco en tratar de optimizar la combinación de los dos nuevos parámetros que proporcionen un mayor rendimiento. Para ello, se ha realizado un análisis utilizando seis valores diferentes de ventanas temporales (5, 10, 20, 40, 80 y 160 días) en combinación de 75 valores distintos para el umbral de pendiente, abarcando un rango desde 0,01 hasta 1,49 (Figura 7).

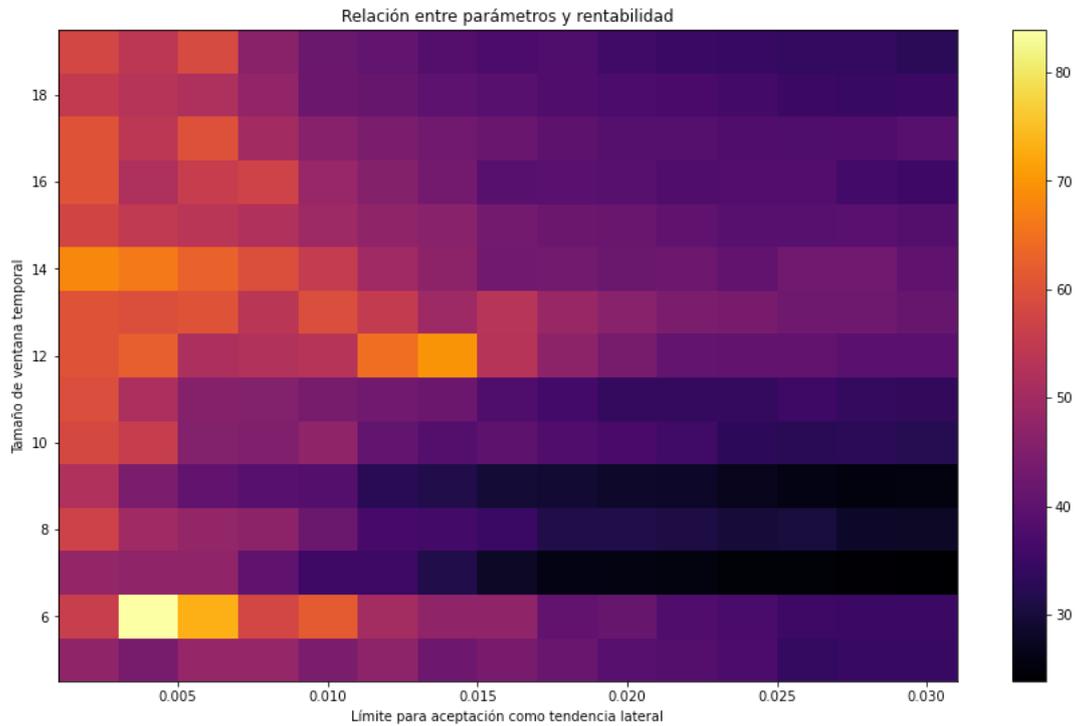
*Figura 7: Análisis de rentabilidad según la combinación de parámetros para el robot optimizado de medias y bandas*



*Fuente: Elaboración propia*

De todas las combinaciones probadas, aquellas que hacen uso de una ventana temporal baja y un umbral de pendiente bajos son las que mejores resultados arrojan. Debido a esto, a continuación se realizará un segundo análisis enfocado en este rango de valores con el objetivo de acercarse con mayor precisión al óptimo absoluto real (Figura 8).

*Figura 8: Análisis de rentabilidad según la combinación de parámetros para el robot optimizado de medias y bandas, con foco en rango de valores óptimo*



*Fuente: Elaboración propia*

La combinación de parámetros óptima identificada se concreta en una ventana temporal de 6 días para el cálculo de las regresiones junto con un umbral de 0,004 para determinar si la pendiente de las regresiones indica una tendencia lateral. Este robot arroja unos resultados sobresalientes, superando en más de 60 puntos porcentuales a los mejores robots vistos hasta el momento en términos de rentabilidad, además de mostrar un comportamiento mucho más estable y, por tanto, menos arriesgado. A continuación, se presenta una tabla a modo de resumen con los resultados del mismo (Tabla 4).

Tabla 4: Análisis del comportamiento del robot optimizado de medias y bandas

Ventana de regresión	Umbral tend. lateral	Indicadores de Rendimiento	Evolución
6 días	0,004	Rdto. Total: 83,94% Rdto. Anualizado: 11,07% Volatilidad Anual.: 11,49% MDD: -13,14% Ratio de Sharpe: 0,88	

Fuente: Elaboración propia

Este gran desempeño plantea la cuestión de si su éxito es fruto del uso de un umbral excesivamente bajo, que haga que todos los escenarios sean interpretados como tendencias negativas o positivas, abandonando por tanto la estrategia inicial basada en la reversión de la media. Sin embargo, tras una evaluación de esta hipótesis haciendo uso de un umbral de 0 (lo que implicaría la imposibilidad de hacer uso de la estrategia basada en la reversión de la media para concentrarse únicamente en una estrategia de seguimiento de la tendencia), se observa que la rentabilidad disminuye hasta 56,03%.

Por lo tanto, se ha logrado un robot que consigue incorporar y aprovechar las ventajas de ambas estrategias, ajustando su comportamiento en función de las condiciones de mercado detectadas, comportándose de acuerdo a la teoría de la reversión de la media cuando detecta que el activo presenta una tendencia lateral y adoptando posiciones en largo (corto) cuando detecta tendencias alcistas (bajistas).

## CAPÍTULO 3: OPTIMIZACIÓN DEL ROBOT DE Q-LEARNING

### 3.1 Presentación de robot de Q-Learning

En esta sección se presenta un robot de trading que utiliza una técnica de aprendizaje más innovadora y avanzada que las técnicas tradicionales de aprendizaje automático: el aprendizaje por refuerzo. Este enfoque permite al robot adaptarse y aprender una estrategia de negociación eficiente a partir de las interacciones con su entorno, el mercado financiero. Concretamente, este robot de trading emplea un algoritmo de aprendizaje por refuerzo llamado Q-Learning.

El Q-Learning es un algoritmo de aprendizaje por refuerzo que permite a un agente aprender a tomar decisiones óptimas en un entorno a través de la interacción con el mismo y la recepción de recompensas por sus acciones. Esta técnica, mucho más sofisticada y compleja que las presentadas hasta ahora, ha presentado durante este trabajo una limitación principal: los tiempos necesarios para el entrenamiento de los robots. En la primera aproximación, el proceso de entrenamiento demandó más de dos horas; mientras que, en la segunda, el tiempo se incrementó hasta cerca de seis horas. Es importante subrayar que estos tiempos de entrenamiento no garantizan necesariamente un rendimiento destacable, por lo que, en muchas ocasiones, estos entrenamientos hay que repetirlos múltiples veces y hacer numerosas pruebas y ajustes en ellos, antes de poder sacar conclusiones sobre un robot determinado, multiplicando así esta barrera de los excesivos tiempos de entrenamiento.

Como mencionaba, el entrenamiento de este tipo de modelos requiere de mucha capacidad computacional, aun así, durante estos últimos años se han desarrollado variantes aún más avanzadas del algoritmo de Q-Learning, como el *Deep* Q-Learning, que combina el aprendizaje por refuerzo con redes neuronales profundas para mejorar la capacidad de aprendizaje del agente. Aunque el *Deep* Q-Learning puede ofrecer resultados superiores en algunos casos, también presenta una mayor complejidad en términos de tiempo de entrenamiento y demanda de capacidad de cálculo. En este trabajo, se ha optado por utilizar el algoritmo “básico” de Q-Learning en lugar del *Deep* Q-Learning debido a estas consideraciones prácticas.

## 3.2 Robot básico de Q-Learning

### 3.2.1 Descripción en profundidad del robot de trading

El Q-Learning es un algoritmo de aprendizaje por refuerzo que permite a un agente, aprender a tomar decisiones óptimas en un entorno. En este caso, el agente será el robot de trading y el entorno será el mercado financiero. El agente realizará distintas acciones en este entorno, que se limitarán a la compra, la venta y la conservación de la posición, recibiendo recompensas basadas en el resultado de esas acciones. El objetivo de este algoritmo es optimizar la toma de decisiones mapeada en la llamada matriz “Q” (de calidad) que asigna un valor a cada acción para cada posible situación (estado), con el objetivo de que el agente sepa cuál es la mejor acción a tomar en el futuro. A continuación, se muestra un ejemplo simplificado de la matriz Q, con la que el robot de trading tomará las decisiones (Figura 9).

Figura 9: Ejemplo simplificado de la Tabla o Matriz Q

Acciones	0- Mantener	1- Comprar	2- Vender
Estados			
[-1, -1, -1, -1, -1]	20	3	170
⋮	⋮	⋮	⋮
[1, 1, 1, 1, 1]	22	180	5

Fuente: Elaboración propia

Este robot tomará decisiones en base a una política definida llamada  $\epsilon$ -greedy, en la que se entrará en detalles posteriormente. Bajo esta política, generalmente, el robot se basará en escoger la acción “a” dado un estado “s” que maximice el valor de  $Q(s,a)$ . Si se observa el ejemplo simplificado en la Figura 9 anterior y el robot se encontrase en el primer estado, este tomaría la acción 2, vender, ya que es la acción que mayor valor de Q presenta (170).

En esta primera aproximación más básica del robot de Q-Learning, los estados estarán basados exclusivamente en una variable de mercado: los precios de cierre. Concretamente estarán compuestos por las variaciones diarias de los precios de cierre de los pasados 5 días, que reflejarán la última semana de negociación del determinado activo. A su vez, como se mencionó anteriormente, las acciones se limitarán a comprar, vender o mantener la posición.

El proceso de aprendizaje se realiza mediante la actualización iterativa de los valores de la matriz Q utilizando una ecuación matemática específica que tiene en cuenta la recompensa recibida, el valor Q de la acción óptima en el estado siguiente y una tasa de aprendizaje. La ecuación permite al agente ajustar gradualmente los valores de la matriz Q, mejorando su conocimiento sobre qué acciones son más rentables en cada situación (Watkins, 1992).

La ecuación matemática que el robot usa para la actualización de la matriz Q y por tanto la optimización de las acciones es la siguiente:

$$\widehat{Q}(s, a) = Q(s, a) + \alpha * [R + \gamma * \max(Q(s', a')) - Q(s, a)]$$

Esta ecuación fue presentada por Bellman (1957) bajo la idea de que el valor óptimo de una determinada decisión puede ser determinado a partir de los valores óptimos de las próximas decisiones. Esta ecuación está formada por varios componentes, los cuales hay que entender para poder aplicar correctamente el algoritmo de Q-Learning:

- $Q(s, a)$  indica el valor de tomar la acción “a” en el estado actual “s”. Es la estimación actual de la recompensa futura esperada tomando la acción “a”.
- $\alpha$  indica la tasa de aprendizaje del algoritmo. Este parámetro controla a qué velocidad se actualizarán los nuevos valores de la matriz Q, implicando que si el valor de  $\alpha$  es cercano a 0, el aprendizaje se realizará de manera más lenta, y si este es cercano a 1 la actualización de valores en la matriz se realizará muy rápido, poniendo mucho peso en la nueva información que se le proporciona. Sin embargo, una tasa de aprendizaje excesivamente alta también puede ser problemática ya que el algoritmo podría llegar a no converger a una estrategia óptima.
- $R$  indica la recompensa inmediata que recibe el agente al tomar la acción “a” en el estado “s”, en este caso esta recompensa estará calculada en función de si la operación que está realizando el robot está reportando beneficios a través de las siguientes fórmulas:

$$base_t = \frac{Capital_{t+1} - Capital_t}{abs(Capital_t)}$$

$$R = \begin{cases} -1 * (-base)^{1,5}, & \text{si } base < 0 \\ base^{1,5}, & \text{si } base > 0 \\ 0, & \text{si } base = 0 \end{cases}$$

Nótese el uso del valor absoluto en el denominador para tratar casos en los que el robot se encuentre en situación de pérdidas. Además, gracias al planteamiento de las recompensas por medio de la anterior fórmula, las pérdidas muy grandes serán gravemente penalizadas por el algoritmo, mientras que las ganancias muy altas serán exponencialmente recompensadas.

- $\gamma$  indica el factor de descuento de las recompensas futuras, es decir, que si el valor de  $\gamma$  es bajo, implicará que el robot dará mucha más importancia a las recompensas inmediatas ( $R$ ), frente a las recompensas de los próximos estados. Sin embargo, cuanto más cercano a 1 sea este valor de descuento, el robot dará una importancia similar a las recompensas futuras en comparación a las inmediatas.
- El término  $max(Q(s', a'))$  representa el valor de tomar la mejor acción posible  $a'$  en el próximo estado,  $s'$ , según los valores presentes en su tabla Q actual.
- La parte de la ecuación de  $[R + \gamma * max(Q(s', a')) - Q(s, a)]$ , se la conoce como error de diferencia temporal (TD Error). Esta diferencia indica la discrepancia entre la estimación actual de la recompensa futura esperada  $Q(s,a)$  y la recompensa que finalmente experimenta ( $R$ ) sumada a la estimación del valor del próximo estado  $Q(s',a')$ , lo que ayudará a actualizar la estimación del agente acercándole a valores más próximos a la realidad.

Además de esta serie de parámetros contenidos en la ecuación que hará que se actualicen los distintos valores de la matriz Q, el robot de Q-Learning usará dos parámetros más que están relacionados a la política que seguirá el mismo. Para equilibrar la exploración de nuevas acciones y la explotación de las acciones ya conocidas, el agente seguirá una

política llamada  $\epsilon$ -greedy. Con probabilidad  $\epsilon$ , el agente elige una acción aleatoria para explorar, y con probabilidad  $1-\epsilon$ , elige la acción con el valor Q más alto para el estado dado, explotando su conocimiento actual. La política  $\epsilon$ -greedy ayuda al agente a evitar quedarse atrapado en decisiones subóptimas al garantizar que siempre haya posibilidad de explorar nuevas acciones (Sutton & Barto, 2015). Esta política implica hacer uso de los siguientes parámetros:

- El parámetro  $\epsilon$ , como ya ha sido mencionado, ajustará el nivel de exploración aleatoria del robot, rechazando tomar la acción presentada como óptima en su matriz Q, con una probabilidad  $\epsilon$ . Por tanto, cuanto mayor sea el valor de este parámetro, más aleatoria será la toma de decisiones.
- El segundo parámetro, muy en relación con el primero, se trata del  $\epsilon$ -Decay (o decaimiento de  $\epsilon$ ). Este parámetro se ajusta para que el valor de  $\epsilon$  no tenga que tomar un valor fijo a lo largo del entrenamiento. El valor de  $\epsilon$  se irá multiplicando en cada nuevo episodio por este parámetro, que será siempre menor que 1, reduciendo por tanto la probabilidad de aleatoriedad en la toma de decisiones a lo largo del proceso de entrenamiento del robot.

Haciendo uso de estos parámetros y tras la repetición de este proceso múltiples veces, el algoritmo de Q-Learning ayuda al agente a aprender qué acciones son las más rentables en cada situación, mejorando su desempeño a medida que el agente interactúa con el entorno y adquiere experiencia, optimizando los valores de la matriz Q.

Y es aquí, en la estructura de la matriz Q, donde se encuentra la principal limitación de este robot y donde un algoritmo más avanzado podría superar el rendimiento de este robot indudablemente. Este desafío surge debido a la forma en la que la matriz Q es utilizada: para que un estado tenga mapeada una acción optimizada en ella, el agente debe haber experimentado el mismo estado previamente (en este primer robot de Q-Learning serán variaciones del precio de una acción durante los últimos 5 días).

En este primer robot básico de Q-Learning, se ha decidido estructurar los estados de una manera particular con el objetivo de que fuese más fácil experimentar situaciones similares durante el proceso de entrenamiento del robot. Primero, el intervalo de cinco días de trading proporcionado al robot no se seleccionó al azar. Debido a que el robot necesita experimentar condiciones de mercado idénticas para aprender de una situación

específica, un intervalo de tiempo más extenso haría exponencialmente más difícil para el robot encontrar una serie de días con retornos iguales.

En segundo lugar, para estandarizar aún más los estados, se transforman los valores de los retornos de las acciones, convirtiendo los retornos diarios inferiores a -1% al valor -1 y a los mayores de 1% a 1, dejando los valores intermedios con el valor 0:

$$valor = \begin{cases} -1, & si \quad r_{diaria} \leq -0.01 \\ 0, & si \quad 0.01 > r_{diaria} > -0.01 \\ 1, & si \quad r_{diaria} \geq 0.01 \end{cases}$$

Este método de preprocesamiento de datos y limitación del intervalo de tiempo usado por el robot aumentará significativamente la probabilidad de encontrar estructuras de datos similares con el tiempo, limitando el número de estados posibles a 243. Sin embargo, también se debe reconocer que este proceso reduce tanto la cantidad como la calidad de información el robot tiene disponible, lo que podría complicar su capacidad para desarrollar una estrategia optimizada.

A pesar de la limitación explicada anteriormente sobre las matrices Q para el aprendizaje, se ha decidido crear un robot distinto para cada valor en el mercado, con el objetivo de que los agentes puedan aprender de las particularidades de los distintos valores y sus distintos comportamientos en bolsa, en lugar de tratar todos los valores como si se tratase de uno solo. Por lo tanto, se repartirá el capital inicial uniformemente entre los distintos valores, lo que supone que el rendimiento de un robot en un valor en particular no afectará al desempeño del resto de los robots.

Para esta primera aproximación del algoritmo de Q-Learning, estos robots tendrán que aprender de un conjunto de datos de entrenamiento, que abarcarán el periodo desde inicios del año 2000 hasta finales del año 2016, para posteriormente aplicar este conocimiento al periodo de *backtesting* (Desde inicio del 2017 hasta finales de octubre del 2022). El robot llevará a cabo un entrenamiento a lo largo de 1.000 episodios antes de comenzar a predecir durante el periodo de *backtesting*, lo que implica que el robot examinará el mismo conjunto de datos de entrenamiento 1.000 veces, optimizando su tabla de Q-Learning a partir de estos datos. En esta primera aproximación, los robots se limitarán a

realizar operaciones semanalmente, lo que reducirá los tiempos de entrenamiento, al no tener que actualizar los valores de su tabla Q con cada día de negociación.

En cuanto a lo que respecta al trading de acciones, cuando el agente de Q-Learning tome la acción de comprar, usará el 100% del efectivo que tenga disponible en ese instante para comprar acciones. Sin embargo, si el robot se encontraba en ese momento en una posición en corto, comprará las acciones que tenía en corto cerrando la posición, lo que le dejaría en una posición neutral (con el dinero en efectivo y sin acciones).

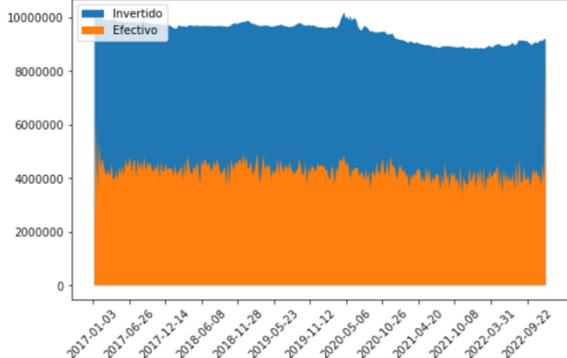
En el caso de que el robot tome la segunda acción, vender, podrán ocurrir varias cosas. En primer lugar, si el agente ya contaba con acciones en su poder, las venderá cerrando la posición y quedándose de nuevo en una posición neutral. Pero si el robot no contaba con acciones anteriormente, se pondrá en una posición en corto por un valor igual al que tiene disponible en efectivo en ese momento, vendiendo unas acciones que tendrá que recomprar en un futuro.

Además de las dos acciones mencionadas, el agente de Q-Learning también tiene la capacidad de tomar la decisión de mantener la posición actual. Esta tercera acción se ejecuta cuando el agente estima que mantener la posición actual, ya sea larga, corta o neutral es la estrategia óptima, por lo que, en esta ocasión, no realizaría ninguna operación adicional.

### **3.2.2 Resultados del robot de trading**

El robot de Q-Learning en esta primera aproximación, debido principalmente a las limitaciones en cuanto a la optimización de la tabla Q mencionadas previamente, no ha sido capaz de aprender una estrategia eficiente que le llevara a obtener un rendimiento positivo, como se muestra en la siguiente tabla (Tabla 5) con los resultados del rendimiento del robot durante el periodo de *backtesting*.

Tabla 5: Resultados del robot de Q-Learning básico

Indicadores de Rendimiento	Evolución
Rdto. Total: -8,27% Rdto. Anualizado: -1,47% Volatilidad: 3,17% MDD: -13,42% Ratio de Sharpe: -0,80	

Fuente: Elaboración Propia

Como se observa, el rendimiento del robot es notablemente bajo, obteniendo este una rentabilidad negativa. Este bajo rendimiento indica claramente como el robot no ha conseguido aprender una estrategia eficiente a lo largo del periodo de entrenamiento de 16 años. Esta primera aproximación deja claro, por tanto, que para hacer uso de un modelo de aprendizaje por refuerzo como el Q-Learning, no se puede alimentar el modelo únicamente con variaciones diarias de los precios, puesto que no le aportan un conocimiento replicable en el tiempo. Con el objetivo de optimizar este robot, y para demostrar que el uso de un algoritmo de Q-Learning puede llegar a ser eficiente en un entorno como el mercado financiero, a continuación se procederá a elaborar una versión mejorada de este robot de trading basado en Q-Learning.

### 3.3 Robot optimizado de Q-Learning

#### 3.3.1 Descripción en profundidad del robot de trading

En esta sección, se describirá el procedimiento implementado para mejorar el rendimiento del robot de trading basado en Q-Learning. Se han introducido tres mejoras fundamentales: la ampliación de la información proporcionada al robot, lo que conlleva una metodología diferente para la generación de estados; la implementación de trading diario, permitiendo una revisión diaria de las posiciones para cada valor; y la modificación del proceso de entrenamiento, con ajustes en la política que sigue el robot en el periodo de entrenamiento y en el número de episodios utilizados.

- Enriquecimiento de la información provista al robot: A diferencia del primer robot de Q-Learning desarrollado, y en vista de la necesidad de una mayor calidad en los datos de entrenamiento sugerida por los resultados obtenidos, se ha optado por enriquecer la información provista al robot en esta segunda aproximación más optimizada, en la que, en lugar de limitarse la información a la variación diaria de los precios, ahora se introducen tres parámetros que le aportarán al robot información variada:

- Ratio entre precio y media móvil a corto plazo. Esta métrica tiene como objetivo proporcionar información sobre el estado actual del precio de un activo en comparación con su desempeño en un período corto de tiempo, en lugar de simplemente ofrecer un vector con las variaciones diarias de los últimos días de negociación. En este caso, se ha establecido este período en 3 días. Este parámetro es atemporal, ya que es un ratio que indica la magnitud de las diferencias entre el precio de mercado y la media del activo. Esta atemporalidad del parámetro y, por tanto, su replicabilidad y comparabilidad en el tiempo, son características fundamentales que es necesario que tomen nuestras métricas para una correcta optimización de la tabla Q.
- Amplitud de las Bandas de Bollinger. Este parámetro aportará al robot un dato fundamental como es el de la volatilidad del mercado. Haciendo uso de 1 sola desviación típica móvil para el cálculo de estas bandas, el cálculo de este parámetro es muy sencillo, ya que se obtiene simplemente multiplicando por 2 veces la desviación típica. El período de tiempo utilizado para su cálculo será el mismo que el utilizado para la media móvil a corto plazo (3 días), además, para garantizar que este parámetro sea atemporal y replicable, como el anterior, se ha decidido ajustar este valor a la media móvil a corto plazo:

$$Amplitud\ BB_t = \frac{2 * \sigma_t}{\mu_t}$$

- Ratio de precio y media móvil a largo plazo. Este ratio, atemporal al igual que los otros dos, ofrece al robot una perspectiva más global del estado del precio de la acción en comparación con una ventana temporal mucho más amplia de lo que se habría podido proporcionar con el primer enfoque. Este parámetro le marcará al

robot la tendencia del precio del activo a lo largo de un periodo de tiempo de 1 año (252 días de negociación).

La introducción de estas tres métricas distintas conlleva una modificación en la generación de los estados que el robot usará para optimizar su tabla Q. En este caso, los diversos parámetros podrán asumir valores desde 0 hasta 9, determinados por rangos que se predefinirán durante el periodo de entrenamiento. Estos rangos se calculan mediante una discretización de los datos de las variables en base a los valores que ese parámetro adopte durante el periodo de entrenamiento.

Cada “sub-estado”, es decir, cada valor discretizado de cada una de las variables, se configurará de manera que comprenda el mismo número de observaciones durante el periodo de entrenamiento. Así, el 10% inferior de los valores se etiquetará con un 0, los valores que caigan entre el 10% y el 20% inferior se etiquetarán con un 1, y así sucesivamente. Durante el periodo de *backtesting*, estos estados se generarán utilizando los rangos ya definidos previamente durante el entrenamiento del modelo y si se diese el caso de encontrar un valor por debajo del anterior mínimo o por encima del anterior máximo, se les asignará directamente el valor 0 o 9, respectivamente.

Una vez obtenidos todos estos “sub-estados” para cada una de estas tres variables, se conformará el estado final para cada punto en el tiempo. Este estado final será una simple concatenación de los distintos “sub-estados”. Por ejemplo, una observación que tuviese un 8 para la métrica del ratio entre el precio y la media móvil a corto plazo, 6 para la amplitud de las Bandas de Bollinger, y 7 para el ratio entre el precio y la media móvil a largo plazo, se encontraría en el estado 867.

Este proceso genera un total de 1.000 estados posibles, que van desde el estado 000, donde todas las métricas estarían en valores cercanos a los mínimos, hasta el estado 999, que representa el caso contrario donde todas las métricas alcanzarían valores cercanos a sus máximos.

- Implementación de trading diario. A diferencia del primer robot, en el que la negociación de activos se limitaba a una operación por semana, en esta ocasión el robot tendrá la capacidad de hacer una operación al día. Esto tiene implicaciones en el tiempo

de entrenamiento, ya que lo multiplica en cinco, pero a su vez le proporciona al robot capacidad para adaptarse a las condiciones cambiantes del mercado a una mayor velocidad.

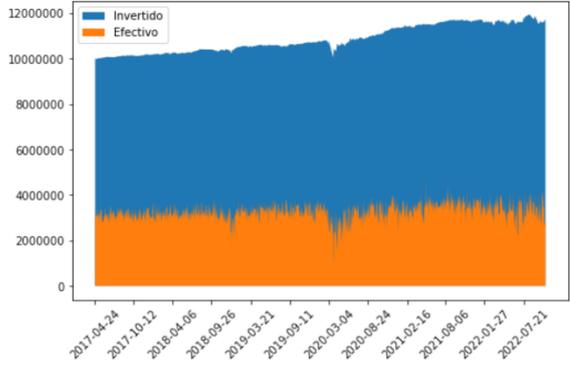
- Modificación del proceso de entrenamiento. Este apartado abarca tres conceptos importantes relacionados con el entrenamiento del robot:

- En primer lugar, se ha modificado ligeramente la política que seguirá a cabo el robot. Pese a seguir implementando una política  $\epsilon$ -greedy, en la que el robot eligirá el máximo valor de Q para el estado actual con una probabilidad de  $1 - \epsilon$ , se mantendrá fijo el valor de  $\epsilon$  en 0,3 durante todo el proceso de entrenamiento del robot, lo que ayudará al robot a no caer en estrategias subóptimas, animándole a explorar durante todo su periodo de entrenamiento.
- El periodo de entrenamiento se ha limitado a un lapso de 4 años, que comprende desde principios de 2013 (utilizando datos de 2012 para calcular la media móvil anual) hasta inicios de 2017. Este cambio temporal se fundamenta principalmente en un aumento de los tiempos de entrenamiento con respecto al robot de la primera aproximación. A pesar de esta reducción importante en los datos de entrenamiento, el robot sí que parece aprender en esta ocasión, lo que plantea la cuestión de si los datos de mercado de hace 20 años son una buena fuente de información válida para la predicción de precios en la actualidad.
- Por un motivo similar al anterior, el del aumento del tiempo de entrenamiento, se ha limitado el número de episodios para el entrenamiento de este robot a 100, un número que ha parecido suficiente para extraer conocimiento de los datos de entrenamiento, como se mostrará más adelante. A pesar de la limitación de estos parámetros de entrenamiento, el proceso de ajuste de estos robots ha llevado alrededor de 6 horas, un tiempo elevado que dificulta la prueba y el ajuste de los parámetros del robot.

### 3.3.2 Resultados del robot de trading

Un simple vistazo a la Tabla 6 evidencia claras diferencias entre esta segunda aproximación y la primera. Este segundo robot exhibe una notable mejora en rentabilidad, superando en más de 25 puntos porcentuales a su predecesor, manteniendo una volatilidad comparable a la anterior y mostrando una reducción significativa en el Máximo Drawdown (MDD). Todo ello le otorga un Ratio de Sharpe considerablemente más alto que el presentado en el modelo anterior.

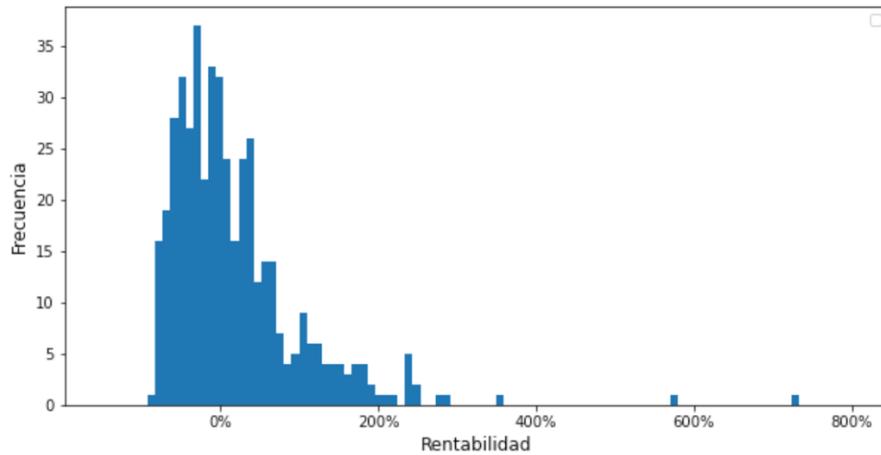
Tabla 6: Resultados del robot de Q-Learning optimizado

Indicadores de Rendimiento	Evolución
<p>Rdto. Total: 16,98%</p> <p>Rdto. Anualizado: 2,88%</p> <p>Volatilidad Anualizada: 4,00%</p> <p>MDD: -7,05%</p> <p>Ratio de Sharpe: 0,47</p>	

Fuente: Elaboración propia

A pesar de superar con creces al robot desarrollado en la aproximación inicial, este nuevo robot aún se encuentra lejos de los niveles de rendimiento que otros robots propuestos en este trabajo han logrado. Al examinar más en profundidad este rendimiento menos sobresaliente, se observa que la rentabilidad está fuertemente influenciada por rendimientos negativos en la mayoría de los robots entrenados, lo que indica que no se ha aprendido una estrategia óptima para muchos de los valores bursátiles. Sin embargo, una interpretación más matizada de estos resultados revela un lado positivo: para ciertas acciones, el rendimiento ha sido extremadamente favorable. De hecho, este rendimiento positivo es tan prominente que compensa con creces el efecto de que más de la mitad de los robots no logran valores positivos, demostrando que, para determinadas acciones, el robot ha logrado aprender una estrategia eficaz.

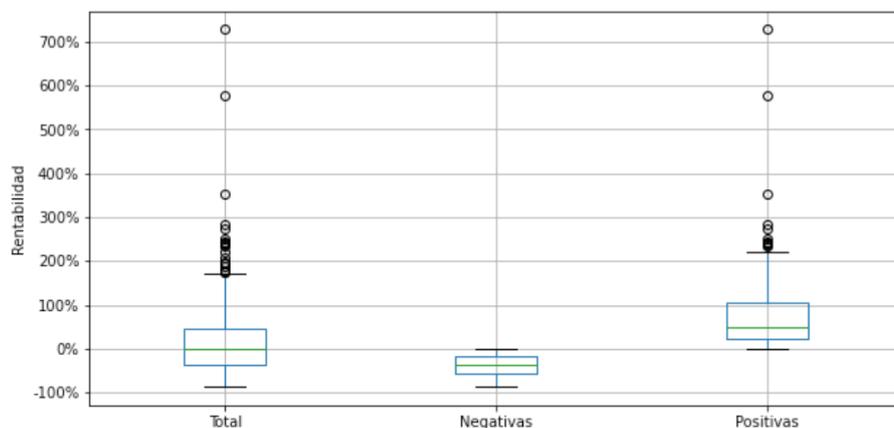
*Figura 10: Histograma de rentabilidades de las distintas acciones haciendo uso del robot de Q-Learning optimizado*



*Fuente: Elaboración propia*

Como se aprecia en la Figura 10 anterior, las rentabilidades exhiben un notable sesgo positivo. Estas rentabilidades presentan comportamientos muy diferentes cuando son negativas en comparación con cuando son positivas, tal como se observa en la Figura 11 siguiente. En el caso de rentabilidades negativas, donde el robot no logra generar una estrategia óptima, la variación de estas es limitada, presentando una desviación estándar del 22%. Sin embargo, en el caso de las rentabilidades positivas, esta variación es cerca de cuatro veces mayor (87%), lo que brinda la oportunidad de aprovechar rentabilidades que, en muchas ocasiones, superan el 100%.

*Figura 11: Gráfico de cajas de las rentabilidades de las distintas acciones haciendo uso del robot de Q-Learning optimizado*



*Fuente: Elaboración propia*

Al profundizar en aquellas acciones para las que el robot ha logrado desarrollar una estrategia eficiente, se observa que estos resultados notablemente satisfactorios no son producto del azar. Más bien, el robot a través de la tabla Q, logra aprender de los datos de entrenamiento y elaborar una estrategia competitiva. A continuación, se muestra cómo al repetir el proceso de entrenamiento 21 veces para los robots de dos acciones muy diferentes que arrojaron buenos resultados durante la prueba general de todos los robots, se obtienen resultados coherentes con lo obtenido inicialmente. Esto respalda la idea de que el robot logra obtener resultados destacables para algunas acciones específicas, pero no para otras.

Tabla 7: Resultados de las simulaciones para FLT

<b>FLEETCOR Technologies (FLT)</b>		
<i>Ejemplo iteración con rentabilidad = 207,42%</i>		
	<b>Rentabilidad original acción</b>	25,90%
	<b>Rentabilidad media robot</b>	169,47%
	<b>Mediana de rentabilidad del robot</b>	151,57%

Fuente: Elaboración propia

Tabla 8: Resultados de las simulaciones para AMD

<b>Advanced Micro Devices (AMD)</b>		
<i>Ejemplo iteración con rentabilidad = 1.770,13%</i>		
	<b>Rentabilidad original acción</b>	442,50%
	<b>Rentabilidad media robot</b>	641,94%
	<b>Mediana de rentabilidad del robot</b>	461,00%

Fuente: Elaboración propia

A modo de conclusión de este robot, es importante destacar el gran potencial que presenta el mismo. Sin embargo, a pesar de haber dedicado un esfuerzo a entender cuáles son las características inherentes a las distintas acciones que facilitan un aprendizaje óptimo y cuáles son las que lo dificultan, no se ha logrado identificar ninguna diferencia evidente.

A pesar de ello, sí que se ha podido observar una consistencia en el desempeño de determinados activos financieros frente a otros, los cuales, como en el caso de las dos acciones seleccionadas en el análisis previo, FLT y AMD, han mostrado rentabilidades positivas en cada una de las 44 simulaciones realizadas sobre ellas (21 simulaciones más la simulación general inicial por cada acción).

Estas observaciones presentan, como ha sido mencionado, un robot muy prometedor. Si se logran descifrar los factores que permiten a los robots aprender de una manera efectiva a partir de ciertos valores, y se hace uso de este conocimiento para utilizar el robot en los casos en los que una acción cumpla con esas condiciones, se podrían llegar a alcanzar rendimientos excepcionales.

## **CAPÍTULO 4: OPTIMIZACIÓN DEL ROBOT CAPM**

### **4.1 Presentación de robot de CAPM**

El Modelo de Valoración de Activos Financieros, más conocido en el mundo de las finanzas como CAPM (*Capital Asset Pricing Model*), es un modelo financiero ampliamente utilizado para determinar la rentabilidad esperada de un activo, teniendo en cuenta tanto el riesgo sistemático (*risk-free rate*), como el rendimiento esperado del mercado y su relación con el activo en cuestión (Fama & French, 2004). Este modelo plantea la relación directa entre riesgo y rentabilidad esperada, una relación coherente si se piensa en que los inversores exigirán un mayor retorno para hacer frente a un mayor riesgo.

$$R_i = R_f + \beta_i * (R_m - R_f) \quad (1)$$

La rentabilidad esperada de un activo ( $R_i$ ) será la suma del riesgo sistemático ( $R_f$ ) más el riesgo de mercado ( $R_m - R_f$ ) multiplicado por un factor de riesgo  $\beta_i$ , que determinará la

relación de un activo determinado con respecto al resto de mercado, por lo que un activo con  $\beta = 1$ , se comportará como la media del mercado, mientras que uno con  $\beta$  mayor que 1 se comportará en la misma dirección que el mercado, pero con movimientos más pronunciados, implicando un mayor riesgo. A su vez, un activo con  $\beta$  negativa, se comportará de manera inversa al resto del mercado, pudiendo ser un valor interesante a considerar cuando el resto del mercado está atravesando momentos complicados.

El modelo CAPM es ampliamente utilizado en el sector financiero, siendo utilizado en diversos ámbitos como en la estimación del coste de capital para proyectos de inversión o en la gestión de riesgos de mercado de una cartera de inversión.

En este trabajo, se propone el desarrollo de un robot de trading algorítmico basado en la teoría del CAPM, que estime la rentabilidad esperada de los activos financieros y tome decisiones de inversión en función de dicha estimación. El enfoque adoptado sigue el modelo de trabajo de Michael Hasler y Charles Martineau (2019), quienes demostraron como el CAPM puede predecir los retornos de los activos financieros a raíz de la rentabilidad esperada del mercado, que a su vez ha de ser predicha por una variación del modelo CAPM. El objetivo es desarrollar un robot de trading basado en la aportación de Hasler y Martineau, investigando si es posible mejorar la estimación de los parámetros utilizados por ellos por medio de otros algoritmos más sofisticados.

## **4.2 Robot basado en CAPM básico**

### **4.2.1 Descripción en profundidad del robot de trading**

El propósito de este algoritmo de trading, basado en la teoría del CAPM, es obtener valores de rentabilidad esperada para los diferentes valores que componen el índice S&P 500, con los que posteriormente tomar decisiones de inversión. Para lograr este objetivo, se seguirá el enfoque presentado por Hasler y Martineau (2019) además de hacer uso del trabajo de Zhou (2009) para calcular ciertos componentes necesarios para poder hacer uso de la fórmula del CAPM correctamente.

Como se mencionaba, antes de entrar directamente en la regresión del CAPM, será necesario realizar un proceso previo en el que, en primer lugar, se tendrá que obtener el

rendimiento esperado del mercado en el mes t+1. Para este cálculo, Hasler y Martineau proponen la siguiente regresión:

$$R_{m,t+1} = c_1 + c_2 * VRP_t + \varepsilon_{t+1} \quad (2)$$

Como se puede observar, esta ecuación consiste en una regresión lineal simple, donde  $VRP_t$  representa la Prima de Riesgo de la Varianza (*Variance Risk Premium*) para el periodo actual. La Prima de Riesgo de la Varianza (VRP) se refiere a la diferencia entre la volatilidad (varianza) implícita de un activo subyacente derivada de los precios de las opciones y la volatilidad (varianza) realizada o histórica del mismo activo para un intervalo de tiempo [t, t+1] (3). La VRP suele ser positiva, lo que significa que el mercado sobrestima sistemáticamente la volatilidad futura en comparación con la que se realiza posteriormente.

$$VRP_t = E_t^I(Var_{t,t+1}) - E_t^R(Var_{t,t+1}) \quad (3)$$

Dado que la VRP en un tiempo actual no es observable en la práctica ya que requiere datos futuros, Zhou (2009) propone una aproximación para su cálculo. Para calcular la varianza implícita (la volatilidad derivada del precio de las opciones), se utiliza la volatilidad implícita calculada por el CBOE en el índice VIX, que se elevará al cuadrado para obtener la varianza. Mientras que para calcular la varianza realizada, se emplea una aproximación discretizada ( $RV_{t,t+1}$ ) presentada por Zhou (2009). La fórmula para esta aproximación es la siguiente:

$$RV_{t,t+1} = \sum_{j=1}^n \left[ p_{t+\frac{j}{n}} - p_{t+\frac{j-1}{n}} \right]^2 \rightarrow Var_{t,t+1} \quad (4)$$

donde  $p_t$  representa el precio logarítmico de un activo en un instante dado (en este caso del índice S&P 500). Según Zhou (2009), al aumentar la frecuencia de los datos con los que se calcula este sumatorio hacia el infinito ( $n \rightarrow \infty$ ), se obtiene una aproximación muy precisa de la varianza realizada. Para este cálculo, en este trabajo se han obtenido datos minuto a minuto desde 2009 hasta el 31 de octubre de 2022, una información difícil de

obtener por la cantidad de datos y por la limitada disponibilidad, estando disponible en muy pocas plataformas que cobran altos precios para su consulta.

Una vez estimada la VRP, utilizando el último mes de datos para su cálculo, se procede a realizar la regresión planteada en la ecuación (2), con la que se calculará una predicción del retorno esperado del mercado  $E_t(r_{M,t+1})$  a un mes vista (21 días). Una vez obtenido el valor de retorno esperado del mercado, solo faltará el cálculo de las  $\beta$  de los distintos valores financieros para contar con todos los componentes necesarios para poder hacer uso del CAPM. Estas  $\beta$  se calculan como la covarianza entre el mercado y el activo en cuestión, dividida por la varianza realizada del mercado:

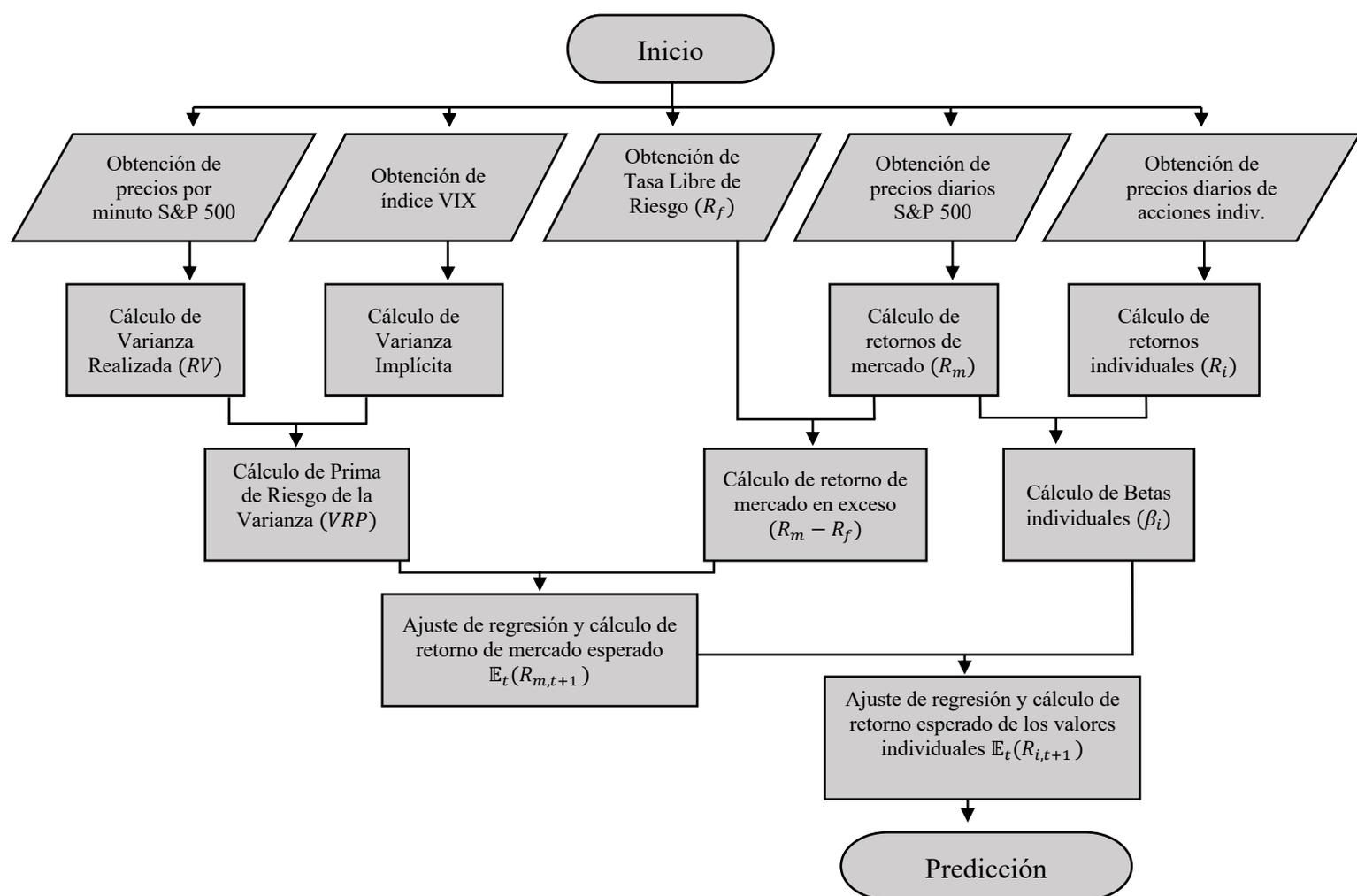
$$\beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)} \quad (5)$$

Con todos estos componentes, ya se puede proceder a realizar la regresión del CAPM y predecir el retorno esperado de los distintos activos que componen el índice S&P 500. Hay que destacar en este punto, que a pesar de hacer uso de la idea del modelo CAPM, para el cálculo final se hace uso de una pequeña variación de la regresión original (1), propuesta también por Hasler y Marineau (2015), que se muestra a continuación en la siguiente fórmula (6).

$$R_{i,t+1} = a + b * [\beta_{i,t} * E_t(R_{m,t+1})] + \varepsilon_{i,t} \quad (6)$$

A continuación, y para facilitar la comprensión del proceso algorítmico descrito a lo largo de esta sección, se muestra un diagrama de flujo con los distintos procesos necesarios que seguirá el robot para proporcionar el retorno esperado de un activo (Figura 12).

Figura 12: Diagrama de flujo del proceso algorítmico



Fuente: Elaboración propia

El algoritmo del robot se basará por tanto en la teoría del CAPM para tomar decisiones de inversión en función de las rentabilidades esperadas de los distintos activos que componen el S&P 500. El robot comenzará a invertir con un capital inicial de 10.000.000 de euros, que se distribuirá equitativamente entre todas las acciones, es decir, que al igual que ocurría con otros robots, el rendimiento del robot en un activo no afectará al rendimiento en otros activos, ya que el capital destinado a cada uno estará delimitado.

El robot, para tomar las decisiones de inversión, se basará en una serie de reglas bastante sencillas, rigiéndose principalmente por unos límites (o umbrales) que se establecerán de antemano y que afectarán a la manera de tomar decisiones de inversión. En primer lugar, si la rentabilidad esperada de un activo para el próximo mes supera el umbral superior

establecido manualmente, el robot procederá a comprar todas las acciones disponibles con el capital asignado a ese activo, siempre y cuando no haya ninguna posición abierta en dicho activo. De manera similar, si la rentabilidad esperada de un activo es inferior al umbral inferior que se establezca, el robot se pondrá en una posición en corto, vendiendo acciones por el valor del capital que tenga disponible en ese momento. En el siguiente apartado, se probará con distintos valores de umbrales para ver la diferencia en rendimiento en función de este parámetro.

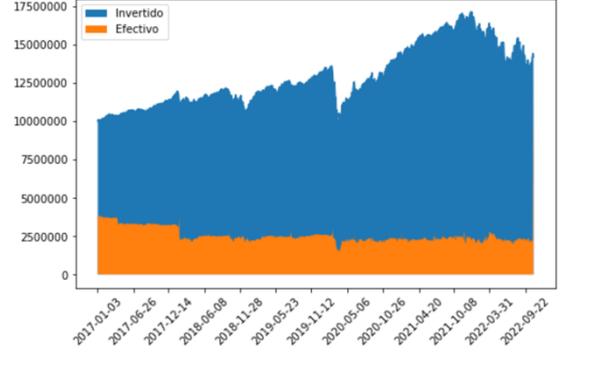
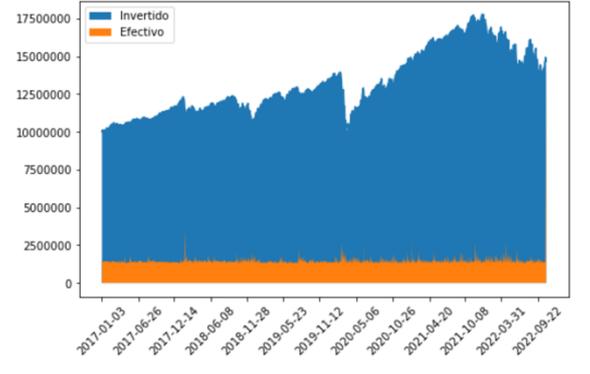
Para cerrar una operación de compra que esté abierta, el robot esperará a que la rentabilidad esperada para el próximo mes sea negativa, de igual manera que para cerrar una posición de venta, se esperará a que la rentabilidad futura esperada sea positiva. A su vez, el robot también tendrá en cuenta el problema del “riesgo ilimitado” al invertir en corto, por lo que cerrará la posición en caso de que la acción haya crecido por encima del límite establecido por la FED para realizar el llamado *margin call* (una subida del 53,8%).

#### 4.2.2 Resultados del robot de trading

Para mostrar la efectividad del robot de trading en función de los distintos umbrales de inversión, se utilizará una tabla que ilustre su rendimiento, junto con otros distintos parámetros y visualizaciones (Tabla 9).

Tabla 9: Resultados del robot de trading basado en CAPM

Límites	Indicadores de Rendimiento	Evolución
2,5% / -2,5%	Rdto. Total: 30,49% Rdto. Anualizado: 4,67% Volatilidad Anual.: 14,28% MDD: -28,76% Ratio de Sharpe: 0,32	

1% / -1%	Rdto. Total: 41,84% Rdto. Anualizado: 6,18% Volatilidad Anual.: 16,90% MDD: -33,67% Ratio de Sharpe: 0,38	
0,035% / -0,035%	Rdto. Total: 46,52% Rdto. Anualizado: 6,78% Volatilidad Anual.: 17,97% MDD: -35,34% Ratio de Sharpe: 0,40	

Fuente: Elaboración propia

A partir de estos resultados, se pueden extraer varias conclusiones sobre el funcionamiento del robot de trading basado en el modelo de valoración CAPM. En primer lugar, se observa que, al modificar los límites establecidos, el comportamiento del robot varía significativamente. Cuando se establecen límites más altos, el robot se vuelve más cauteloso en cuanto al riesgo, manteniendo una mayor cantidad de dinero en efectivo durante el periodo analizado, centrándose únicamente en activos con expectativas de rendimiento sólidas. Esto resulta en un menor *Máximo Drawdown* (MDD) y una menor volatilidad en comparación con los otros escenarios presentados.

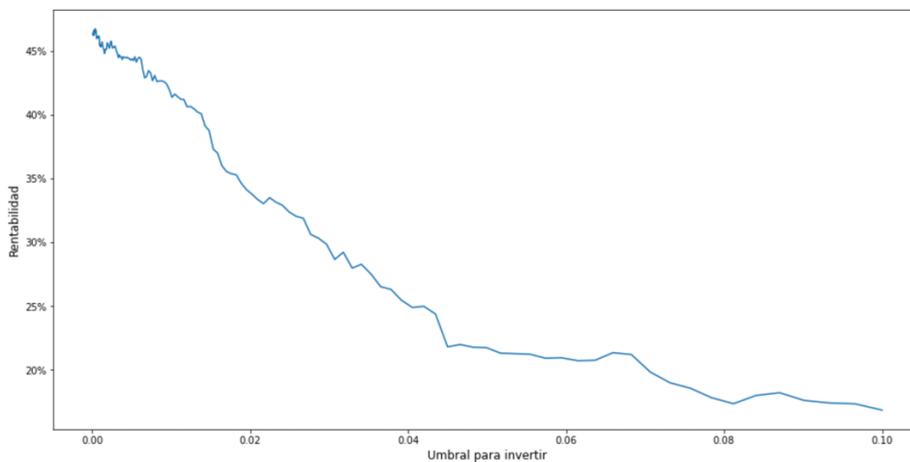
Por el lado opuesto, al disminuir los límites, el rendimiento del robot parece mejorar, obteniendo la mejor rentabilidad entre los escenarios probados en aquel con los límites en 0,035%, donde se obtiene la mayor rentabilidad por unidad de volatilidad (Ratio de Sharpe) y se obtiene una rentabilidad total máxima del 46,52%.

### 4.3 Optimización del robot CAPM

#### 4.3.1 Descripción en profundidad del robot de trading

Tras la evaluación inicial del robot de trading basado en el modelo CAPM, se ha observado una relación entre los umbrales de inversión establecidos y el rendimiento del mismo. En un intento de comprender mejor esta relación y comprobar su efecto en la rentabilidad del robot, se ha llevado a cabo una serie de simulaciones, concretamente 200, en las que se ha ido variando el umbral de inversión en una escala logarítmica, con un valor mínimo del 0.01% y un valor máximo del 10% (Figura 13).

*Figura 13: Resultado de simulaciones variando umbral de inversión*

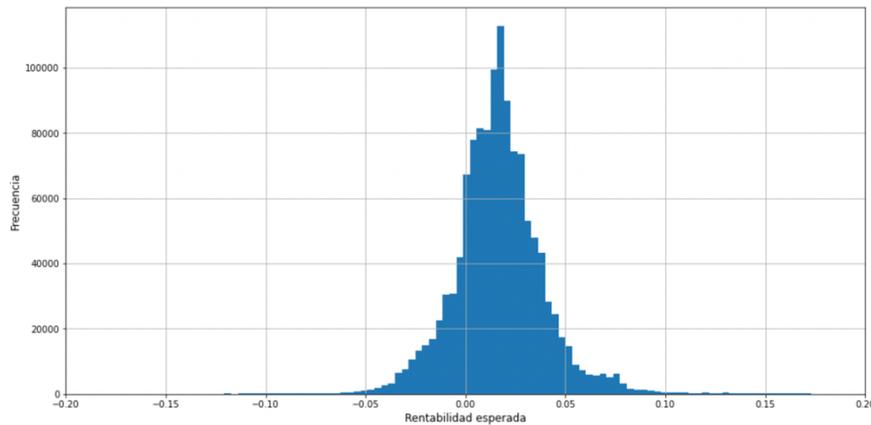


*Fuente: Elaboración propia*

Tras analizar estas simulaciones, se pueden confirmar las sospechas iniciales de que, efectivamente, el robot muestra un rendimiento superior cuando se le aplican límites bajos. Sin embargo, este comportamiento no se mantiene indefinidamente, ya que, a partir de bajar el umbral de inversión por debajo de un valor excesivamente bajo, concretamente 0,04%, el rendimiento del robot comienza a disminuir. Esto sugiere dos cosas, la primera, que el establecimiento de umbrales es efectivo para descartar predicciones que no sean muy fiables (aquellas que caen cercanas a 0); y la segunda, que el robot funciona de manera óptima cuando adopta una estrategia binaria, es decir, comprar ante la previsión de rentabilidades positivas y vender ante predicciones de rentabilidades negativas, sin entrar en consideraciones sobre la magnitud de estas previsiones, ya que no se consiguen rendimientos superiores al limitar las operaciones a aquellas que a priori deberían ser más rentables.

Además, si se observan las predicciones que hace el robot para el periodo de *backtesting*, se evidencia una clara tendencia del robot hacia la predicción de valores positivos, por lo que, en consecuencia, el funcionamiento del robot en la mayoría de las ocasiones se limitará a ejecutar órdenes de compra (Figura 14).

*Figura 14: Distribución de rentabilidades predichas del robot basado en CAPM*



*Fuente: Elaboración propia*

Esta tendencia a prever rentabilidades positivas y, por tanto, a comprar valores, nos demuestra como los modelos, en base a ser ajustados durante un periodo de entrenamiento marcado por una gran tendencia positiva en el mercado, han conseguido aprender de esta característica general, dejando un poco de lado las posibles peculiaridades de cada uno de los valores concretos.

Como ha sido mencionado anteriormente, para realizar el cálculo del retorno esperado del mercado se emplea una regresión lineal (2). Este modelo es un algoritmo simple que ajusta una fórmula lineal a los datos, lo que plantea la pregunta de si un algoritmo más avanzado podría proporcionar mejores resultados al poder ajustar datos que no presenten una estructura lineal. Para abordar esta cuestión, se probaron distintos modelos utilizando un conjunto de datos de entrenamiento y otro de validación, con el objetivo de evitar el sobreajuste (*overfitting*).

Se entrenaron cinco algoritmos diferentes en el conjunto de datos de entrenamiento: dos regresiones lineales, una utilizando los datos originales para establecer un punto de partida y otra utilizando datos transformados mediante el logaritmo; una *Support Vector*

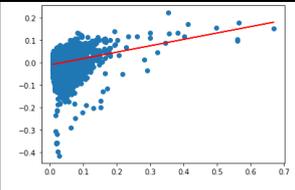
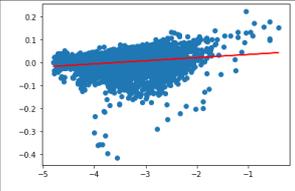
*Regression* (SVR) con kernel *Radial Basis Function* (RBF) aplicado a los datos originales, otro SVR con kernel RBF aplicado a los datos logarítmicos y, finalmente, un modelo XGBoost.

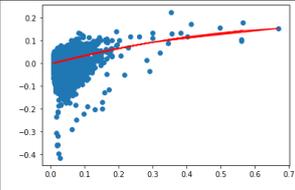
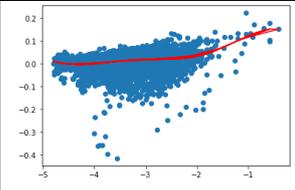
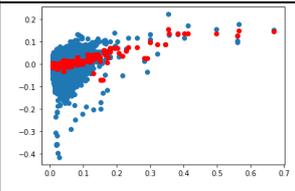
La *Support Vector Regression* (SVR) es una variante de las Máquinas de Vectores de Soporte (SVM), diseñada específicamente para abordar problemas de regresión, en vez de problemas de clasificación. El objetivo principal de este algoritmo es encontrar la mejor función que pueda predecir los valores de una variable continua a partir de los datos de entrada. Estos SVR hacen uso de kernels para transformar los datos, y para el caso del kernel *Radial Basis Function* (RBF), el kernel utilizado, los datos se transforman en un espacio de características de mayor dimensión, facilitando así la búsqueda de relaciones no lineales entre las variables.

Por el otro lado, el XGBoost es un algoritmo de aprendizaje supervisado basado en árboles de decisión. Este algoritmo utiliza una técnica llamada *Gradient Boosting*, la cual combina las predicciones de múltiples árboles de decisión simples, para construir un modelo más fuerte y preciso.

Los resultados de las predicciones realizadas por los diferentes algoritmos, así como una representación gráfica de sus predicciones en el conjunto de datos de entrenamiento, se presentan en la siguiente tabla:

*Tabla 10: Resultados de entrenamiento de modelos de CAPM alternativos*

Modelo	R2	MSE	Visualización
Regresión Lineal Simple	5,787%	0,002079	
Resesión Lineal (log)	3,734%	0,002124	

SVR RBF kernel	0,874%	0,002187	
SVR RBF kernel (log)	1,647%	0,002170	
XGBoost	11,486%	0,001954	

*Fuente: Elaboración propia*

Como se puede observar, pese a la mayor complejidad de algoritmos, solo uno (XGBoost) supera a la regresión lineal simple a la hora de ajustar los datos del conjunto de entrenamiento, por lo que solo se comparará el rendimiento entre ambos modelos sobre el conjunto de validación. Los resultados sobre el conjunto de validación pueden observarse en la siguiente tabla:

*Tabla 11: Resultados de validación de modelos de CAPM alternativos*

Modelo	R2	MSE
Regresión Lineal Simple	4,639%	0,044229
XGBoost	1,441%	0,044965

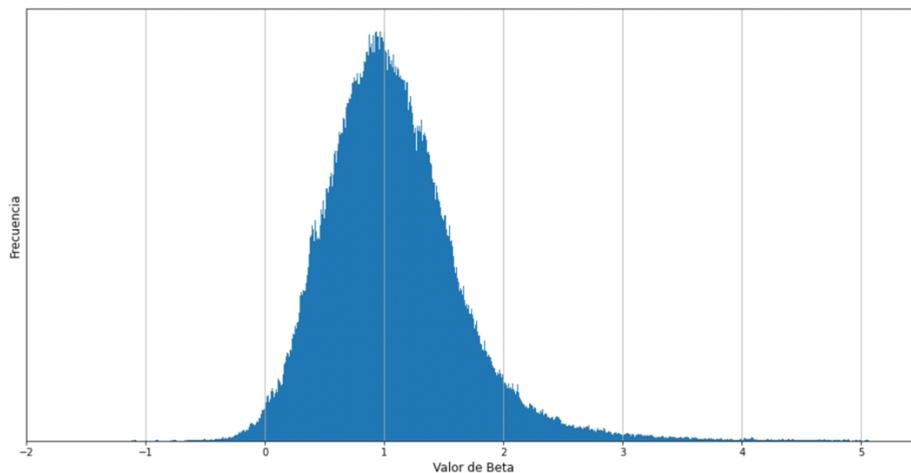
*Fuente: Elaboración propia*

Pese a presentar un rendimiento muy superior en el conjunto de entrenamiento, en el conjunto de validación el modelo XGBoost no consigue superar el nivel de predicción de la regresión lineal simple, lo que indica que posiblemente sufriese de un gran sobreajuste sobre los datos con los que se entrenó. Este sobreajuste se trata de algo esperable, teniendo en cuenta la poca profundidad en información que se le proporciona al modelo, estando toda contenida en una sola variable (VRP). Este planteamiento, como se ha observado, dificulta el uso de modelos más sofisticados, ya que, al buscar patrones y estructuras más complejas, es muy probable que se sobreajuste la información proporcionada a través de la única variable.

En consecuencia, y tras haber estudiado diferentes alternativas a las propuestas por Hasler y Martineau (2015) para el cálculo del retorno esperado del mercado, no ha sido posible obtener un mejor modelo al propuesto por los autores.

Esta limitación, por tanto, deja un robot con un comportamiento bastante generalista, el cual, ante un valor dado de VRP, es habitual que anticipe un crecimiento en el mercado. Por lo tanto, la decisión de tomar una posición en largo o en corto dependerá mayormente de la relación que exista entre el activo específico y el mercado. Una relación que se puede observar a través de la ya explicada  $\beta$ , la cual, en periodos de 21 días, como los aplicados en este robot, raramente exhibe una relación inversa (menor que 0) (Figura 15).

*Figura 15: Valores de  $\beta$  de todas las empresas estudiadas, durante el periodo Enero 2000 – Octubre 2022 (ventanas de 21 días)*



*Fuente: Elaboración propia*

Así pues, el robot presenta un comportamiento que va a menudo al unísono con el comportamiento del mercado, debido a que, por su estructura algorítmica simple, no tiene la capacidad de aprender peculiaridades de los datos. Como resultado, el robot se encuentra excesivamente influenciado por las tendencias a largo plazo del mercado en el que ha sido ajustado, que para el caso concreto del S&P 500, un índice destacado por su constante crecimiento a lo largo de los años, se observa una marcada tendencia hacia la apertura de posiciones en largo. Esto le hace asemejarse al comportamiento del propio índice, con la diferencia de que, mientras el índice pondera en función de la capitalización de mercado, este robot invierte en las diferentes acciones indiscriminadamente.

Figura 16: Comportamiento del Robot de CAPM frente al índice S&P 500

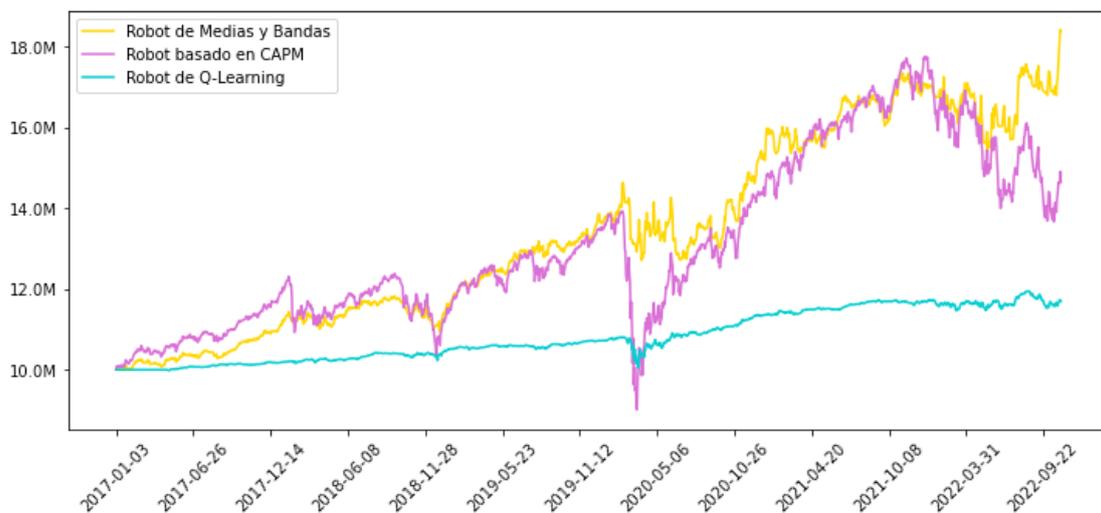


Fuente: Elaboración propia

## CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

En esta sección final, se proporciona una comparación general del desempeño de los distintos robots desarrollados a lo largo del trabajo. En la Figura 17 se muestra una primera visión de la comparación del funcionamiento de los distintos robots, mostrando la evolución del capital total disponible por los mismos a lo largo del periodo de *backtesting*.

Figura 17: Comportamiento de los robots a lo largo del periodo de *backtesting*



Fuente: Elaboración propia

De la figura previa se pueden extraer varias conclusiones, que se irán detallando a lo largo de esta sección. En primer lugar, se puede observar que, en términos de rentabilidad, hay dos robots que han demostrado un desempeño muy superior: el robot de medias y bandas y el robot basado en CAPM. El robot que ha hecho uso de Q-Learning, por el contrario, y pese a presentar resultados extraordinarios en algunos de los valores negociados, no ha conseguido obtener una rentabilidad asombrosa en promedio. En la siguiente tabla (Tabla 12), se muestran los valores concretos de esta variable para los distintos robots estudiados.

*Tabla 12: Comparativa de rentabilidad de los distintos robots estudiados*

<b>Robot</b>	<b>Rentabilidad total</b>
Robot de Medias y Bandas	83,94%
Robot basado en CAPM	46,52%
Robot de Q-Learning	16,98%

*Fuente: Elaboración propia*

En términos de riesgo, entendido éste como la volatilidad que han presentado los distintos robots durante su operativa, también se pueden extraer conclusiones a partir del análisis de la anterior Figura 17. A pesar de que el Robot de Q-Learning no era el referente en términos de rentabilidad, sí que lo es en términos de bajo riesgo, presentando un crecimiento lento pero muy estable a lo largo del tiempo. Por el contrario, el robot basado en medias y bandas y el basado en CAPM exhiben una volatilidad bastante más pronunciada, sobre todo este segundo, para el cual el fenómeno es fácilmente visible durante el periodo de inicios de la pandemia del Covid-19 (inicios del 2020), durante el cual el robot llegó a sufrir un *Máximo Drawdown* de -35,34%, frente a los -7,05% y -13,14% del Robot de Q-Learning y el Robot de Medias y Bandas respectivamente.

*Tabla 13: Comparativa de volatilidad de los distintos robots estudiados*

<b>Robot</b>	<b>Volatilidad anualizada</b>
Robot de Medias y Bandas	11,49%
Robot basado en CAPM	17,97%
Robot de Q-Learning	4,00%

*Fuente: Elaboración propia*

El Ratio de Sharpe, al combinar la rentabilidad junto a la volatilidad, proporciona un indicador razonable de la eficiencia y calidad de las estrategias implementadas por los robots (Tabla 14). Este indicador, tal y como se ha discutido a lo largo del trabajo, muestra una visión de la rentabilidad en exceso (retorno por encima de la tasa libre de riesgo) por unidad de volatilidad, relacionando, por tanto, retornos y riesgo. A través de este parámetro se evidencia que la rentabilidad no lo es todo para conseguir una buena estrategia, ya que, si esta no va acompañada de un riesgo moderado, la estrategia seguida no será óptima. El caso del robot de Q-Learning, el cual supera en Ratio de Sharpe al robot basado en CAPM muestra claramente la necesidad de mantener este equilibrio.

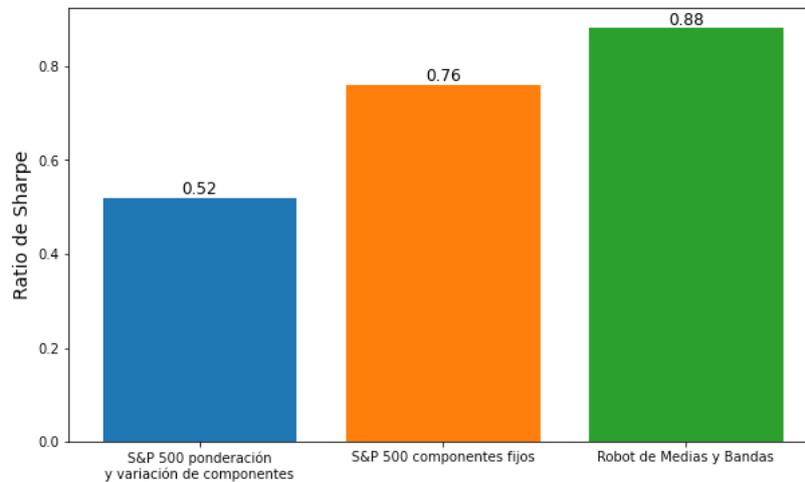
*Tabla 14: Comparativa de Ratio de Sharpe de los distintos robots estudiados*

<b>Robot</b>	<b>Ratio de Sharpe</b>
Robot de Medias y Bandas	0,88
Robot basado en CAPM	0,40
Robot de Q-Learning	0,47

*Fuente: Elaboración propia*

Indiscutiblemente, el robot de Medias y Bandas emerge como el robot con mejor desempeño entre los estudiados, casi duplicando en Ratio de Sharpe al robot que le sigue, el robot de Q-Learning. Este robot optimizado ha logrado adaptarse a las condiciones cambiantes del mercado, adaptando su comportamiento a la teoría de la reversión de la media cuando los valores presentan tendencias laterales y adoptando comportamientos alcistas o bajistas cuando el mercado se comporta de tal manera. Esta adaptabilidad, además de proporcionarle el mejor desempeño entre los robots estudiados, también le proporciona un desempeño superior a estrategias de inversión pasivas o basadas en el propio índice S&P 500 (Figura 18).

*Figura 18: Comparativa de Ratio de Sharpe del mejor robot vs. mercado durante periodo de backtesting*



*Fuente: Elaboración propia*

Fruto de la anterior comparación, cabe recalcar que la inclusión del índice S&P 500, mostrado como “S&P 500 ponderación y variación de componentes” en la Figura 18, tiene un carácter meramente informativo. Este índice cambia su composición y ponderación dinámicamente a lo largo del tiempo, mientras que, para este trabajo, los diferentes robots han utilizado una combinación fija de compañías con una distribución equitativa del capital. Es por ello por lo que una representación más fiel del mercado seleccionado se encuentra en la segunda barra de la Figura 18, mostrada como “S&P 500 componentes fijos”, en la que se consideran únicamente los mismos componentes utilizados en este trabajo, asignándoles una ponderación equitativa a cada uno de los valores.

Para futuras investigaciones se sugiere abordar esta cuestión en la que los robots desarrollados tengan la capacidad de variar tanto los componentes como la cantidad de capital utilizado en cada uno de ellos. Esta línea de investigación podría permitir abordar cuestiones que trascienden al propio trading, como por ejemplo la búsqueda de carteras equilibradas por medio del análisis de las interrelaciones entre los distintos activos.

Además, se propone como otra futura línea de investigación la profundización en el análisis del robot de Q-Learning. Este robot ha demostrado rendimientos excepcionales y consistentes para ciertos valores, pero no para todos. Sería de gran interés investigar en

profundidad las razones que llevan al robot a destacar en su desempeño sobre ciertas acciones y a no tener la capacidad de replicarlo sobre otras.

Como conclusión a este trabajo y con el objetivo de simplificar tanto la comprensión del mismo como el desarrollo de nuevas líneas de investigación, a continuación se proporciona un enlace a un repositorio de GitHub. Este contiene cuadernos de Jupyter con el código en Python de los distintos robots desarrollados y optimizados a lo largo de este trabajo:

[https://github.com/DavidSerranoE2A/TFG\\_OptimizacionDeRobotsDeTrading](https://github.com/DavidSerranoE2A/TFG_OptimizacionDeRobotsDeTrading)

## BIBLIOGRAFÍA

- Achelis, S. B. (2001). Technical Analysis from A to Z. Obtenido de [https://asamvn.com.vn/Uploads/files/ae498cfa-d6bd-4e09-8683-2bf04a52720b\\_pdf\\_.pdf](https://asamvn.com.vn/Uploads/files/ae498cfa-d6bd-4e09-8683-2bf04a52720b_pdf_.pdf)
- Avellaneda, M. (2011). Algorithmic and High-frequency trading: an overview. New York University & Finance Concepts LLC, Quant Congress USA. Obtenido de [https://c.mql5.com/forextd/forum/168/quant\\_congress\\_usa\\_2011\\_algo\\_trading\\_last.pdf](https://c.mql5.com/forextd/forum/168/quant_congress_usa_2011_algo_trading_last.pdf)
- Barber, B. M., Lee, Y.-T., Liu, Y.-J., & Odean, T. (2009). Just How Much Do Individual Investors Lose by Trading? *The Review of Financial Studies*, 22(2), 609-632. Obtenido de <https://doi.org/10.1093/rfs/hhn046>
- Bellman, R. (1957). Dynamic Programming. Obtenido de <https://gwern.net/doc/statistics/decision/1957-bellman-dynamicprogramming.pdf>
- Blume, M. E., & Edelen, R. M. (2004). S&P 500 Indexers, Tracking Error, and Liquidity. *The Journal of Portfolio Management*, 30(3), 37-46. Obtenido de [https://www.researchgate.net/publication/277432332\\_SP\\_500\\_Indexers\\_Tracking\\_Error\\_and\\_Liquidity](https://www.researchgate.net/publication/277432332_SP_500_Indexers_Tracking_Error_and_Liquidity)
- Boehmer, E., Fong, K., & Wu, J. (2012). International Evidence on Algorithmic Trading. *SSRN Electronic Journal*, 1-44. Obtenido de [https://www.researchgate.net/publication/264887675\\_International\\_Evidence\\_on\\_Algorithmic\\_Trading](https://www.researchgate.net/publication/264887675_International_Evidence_on_Algorithmic_Trading)
- Bollinger, J. (1992). Using Bollinger bands. *Stocks & Commodities*, 2(10). Obtenido de <https://c.mql5.com/forextd/forum/211/Using%20Bollinger%20Bands%20by%20John%20Bollinger.pdf>
- Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *The Journal of Finance*, 67(5), 1731-1764. Obtenido de <https://doi.org/10.2307/2328994>
- Brogaard, J., Hendershott, T., & Riordan, R. (2013). High Frequency Trading and Price Discovery. *Working Paper Series*, 1602, 1-55. Obtenido de <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1602.pdf>
- Cao, L., & Tay, F. E. (2001). Financial Forecasting Using Support Vector Machines. *Neural Computing & Applications*, 184-192. Obtenido de <https://doi.org/10.1007/s005210170010>

- Chaboud, A. P., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2014). Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. *The Journal of Finance*, 69(5), 2045-2083. Obtenido de [https://www.researchgate.net/publication/335010481\\_Rise\\_of\\_the\\_Machines\\_Algorithmic\\_Trading\\_in\\_the\\_Foreign\\_Exchange\\_Market](https://www.researchgate.net/publication/335010481_Rise_of_the_Machines_Algorithmic_Trading_in_the_Foreign_Exchange_Market)
- Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems With Applications*, 187-2015. Obtenido de <https://doi.org/10.1016/j.eswa.2017.04.030>
- Donadio, S., & Ghosh, S. (2019). *Learn Algorithmic Trading*. Birmingham: Packt Publishing Ltd.
- Exley, J., Mehta, S., & Smith, A. (2004). Mean Reversion. Obtenido de [https://www.researchgate.net/publication/264858032\\_MEAN\\_REVERSION](https://www.researchgate.net/publication/264858032_MEAN_REVERSION)
- Fama, E. F., & French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. Obtenido de <http://mba.tuck.dartmouth.edu/bespeneckbo/default/AFA611-Eckbo%20web%20site/AFA611-S6B-FamaFrench-CAPM-JEP04.pdf>
- Federal Reserve System. (1998). Subchapter A. En *Title 12—Banks and Banking* (pág. 17). Obtenido de <https://www.govinfo.gov/content/pkg/CFR-2021-title12-vol3/pdf/CFR-2021-title12-vol3-part220.pdf>
- Financial Industry Regulation Authority. (2022). 4210. Margin Requirements. En *FINRA Rules*. Obtenido de <https://www.finra.org/rules-guidance/rulebooks/finra-rules/4210>
- Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4), 5-14. Obtenido de <https://doi.org/10.1177/0008125619864925>
- Hasler, M., & Martineau, C. (2019). Does the CAPM Predict Returns? *Microeconomics: General Equilibrium & Disequilibrium Models of Financial Markets eJournal*. Obtenido de <https://www.rotman.utoronto.ca/-/media/Files/Faculty-Bios/Michael-Hasler/HMForecasting20190408pdf.pdf>
- Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does Algorithmic Trading Improve Liquidity? *The Journal of Finance*, 66(1), 1-33. Obtenido de <https://doi.org/10.1111/j.1540-6261.2010.01624.x>
- J.P. Morgan. (2018). *J.P. Morgan Private Bank Algorithmic Trading Guide (European*

- Markets). Obtenido de [https://privatebank.jpmorgan.com/content/dam/jpm-wm-aem/documents/en/other/multi-family-offices/EQ-ETS-Algorithmic-Trading-Guide-\(Europe-Markets\).pdf](https://privatebank.jpmorgan.com/content/dam/jpm-wm-aem/documents/en/other/multi-family-offices/EQ-ETS-Algorithmic-Trading-Guide-(Europe-Markets).pdf)
- Jones, C. M. (2013). What Do We Know About High-Frequency Trading? Columbia Business School Research Paper, 11(13). Obtenido de [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2236201](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2236201)
- Kubat, M. (2017). An Introduction to Machine Learning, Second Edition. Coral Gables: Springer. Obtenido de <https://www.ms.sapientia.ro/~lalo/concurs/Formafelismeres/An%20Introduction%20to%20Machine%20Learning%20-%20Miroslav%20Kubat%20-%200978-3-319-63913-0.pdf>
- Lane, G. C. (1984). Lane's Stochastics. *Stocks & Commodities*, 2(3), 87-90. Obtenido de [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiqLHU3pf\\_AhUgVqQEHVwnAcgQFnoECAgQAQ&url=https%3A%2F%2Fwww.forexfactory.com%2Fattachment%2Ffile%2F3499852&usg=AOvVaw0sb2KxmguEKXnzAntcZsg0](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiqLHU3pf_AhUgVqQEHVwnAcgQFnoECAgQAQ&url=https%3A%2F%2Fwww.forexfactory.com%2Fattachment%2Ffile%2F3499852&usg=AOvVaw0sb2KxmguEKXnzAntcZsg0).
- Li, B., Zhao, P., Gopalkrishnan, V., & Hoi, S. C. (2012). PAMR : passive aggressive mean reversion strategy for portfolio selection. *Machine Learning*, 87(2), 221-258. Obtenido de <https://doi.org/10.1007/s10994-012-5281-z>
- Lohrmann, C., & Luukka, P. (2019). Classification of intraday S&P500 returns with a Random Forest. *International Journal of Forecasting*, 35(1), 390-407. Obtenido de <https://doi.org/10.1016/j.ijforecast.2018.08.004>
- Mailgaha Kumbure, M., Lohrmann, C., & Luukka, P. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems With Applications*, 197, 1-41. Obtenido de <https://doi.org/10.1016/j.eswa.2022.116659>
- Pruess, P., & Hiranto, P. (2019). Multiple Regression: Impacted Variables to S&P500. *SSRN Electronic Journal*. Obtenido de Pruess, Pornprom and Hiranto, Pattaraporn, Multiple Regression: Impacted Variables to S&P500 (January 24, 2019). Available at SSRN: <https://ssrn.com/abstract=3321542> or <http://dx.doi.org/10.2139/ssrn.3321542>
- Rajihy, Y., Nermend, K., & Alsakaa, A. (2017). Back-propagation artificial neural networks in stock market forecasting. An application to the Warsaw Stock Exchange WIG20. *AESTIMATIO, THE IEB INTERNATIONAL JOURNAL OF FINANCE*, 88-99. Obtenido de

<https://dialnet.unirioja.es/servlet/articulo?codigo=6637821>

Rosillo, R., Giner, J., & De la Fuente, D. (2013). The effectiveness of the combined use of VIX and Support Vector Machines on the prediction of S&P 500. *Neural Computing and Applications*, 321-332. Obtenido de <https://doi.org/10.1007/s00521-013-1487-7>

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3, 210-229. Obtenido de <https://ieeexplore.ieee.org/abstract/document/5389202>

Shapira, Z., & Venezia, I. (2001). Patterns of behavior of professionally managed and independent investors. *Journal of Banking & Finance*, 25(8), 1573-1587. Obtenido de [https://doi.org/10.1016/S0378-4266\(00\)00139-4](https://doi.org/10.1016/S0378-4266(00)00139-4)

Sharpe, W. F. (1994). The Sharpe Ratio. *The Journal of Portfolio Management*, 21(1), 49-58. Obtenido de [https://www.pm-research.com/content/ijpormgmt/21/1/49#copy\\_link](https://www.pm-research.com/content/ijpormgmt/21/1/49#copy_link)

Sirucek, M. (2012). Macroeconomic Variables and Stock Market: US Review. *IJCSMS International Journal of Computer Science and Management Studies*, 12(3). Obtenido de [http://www.ijcsms.com/journals/Volume%2012,%20Issue%2003,%20September%202012\\_Martin%20Paper.pdf](http://www.ijcsms.com/journals/Volume%2012,%20Issue%2003,%20September%202012_Martin%20Paper.pdf)

Sutton, R. S., & Barto, A. G. (2015). *Reinforcement Learning: An Introduction*. Cambridge: MIT Press. Obtenido de <https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf>

Tsai, C.-F., & Hsiao, Y.-C. (2010). Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. *Decision Support Systems*, 50(1), 258-269. Obtenido de <http://dx.doi.org/10.1016/j.dss.2010.08.028>

Watkins, C. J. (1992). Q-Learning. *Machine Learning*, 8, 279-292. Obtenido de <http://www.gatsby.ucl.ac.uk/~dayan/papers/cjch.pdf>

Welles Wilder, J. (1978). *New Concepts in Technical Trading Systems*. Greensboro: Hunter Publishing Company.

Zhang, Y. (2009). Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. *Expert Systems With Applications*, 36(5), 8849-8854. Obtenido de <https://doi.org/10.1016/j.eswa.2008.11.028>