



Predicción de Movimientos en los Principales Índices Bursátiles mediante Random Forest

José Luis Ruiz Solivella

Curso 2024-2025

Universidad Miguel Hernández

Facultad de Ciencias Sociales y Jurídicas de Elche

Estadística Empresarial

Trabajo de Fin de Grado

Tutor: José Luis Sainz-Pardo Auñón

Índice

1.	Resumen	1
2.	Abstract	2
3.	Introducción	3
	3.1. El mercado de Valores y los índices bursátiles	
	3.2. Machine Learning y su uso en el mercado de valores	
	3.3. Objetivos del Proyecto	
4.	Marco Teórico	5
	4.1. Índices bursátiles analizados	
	4.2. Fundamentos de Machine Learning	
	4.2.1. Técnicas de Machine Learning	
	4.2.2. Algoritmos de Clasificación: Random Forest	
5.	Metodología	10
	5.1. Descripción de los datos utilizados	
	5.2. Variables	
	5.2.1 Variables técnicas básicas	
	5.2.2. Indicadores de volatilidad	
	5.2.3. Indicadores de Momentum	
	5.2.4. Indicadores de tendencia	
	5.2.5. Variable objetivo	
	5.3. Preprocesamiento de los datos y selección de variables	
	5.3.2. Alineación de las variables	
	5.3.1 Eliminación de datos faltantes	
	5.3.2. Alineación de las variables	
	5.3.3. División en Conjuntos de Entrenamiento y Prueba	
	5.3.4. Selección de Variables Relevantes	
6.	Desarrollo de los Modelos	24
	6.1. Implementación del Modelo Random Forest	
	6.2. Entrenamiento y Validación	
	6. 3. Métricas de Evaluación	
7.	Análisis de Resultados	26
	7.1. Métricas de Evaluación	
	7.2. Simulación de Inversión y análisis de resultados	
8.	Síntesis y estrategias de mejora	31
	8.1. Conclusiones	
	8.2. Propuestas de Mejora	
9.	Bibliografía y webs consultadas	34
10.	Código empleado y simulación completa	34

1. Resumen

Este trabajo se centra en la aplicación de técnicas de aprendizaje automático, específicamente el modelo Random Forest, para predecir los movimientos intradías en tres índices bursátiles: IBEX 35, DAX 40 y S&P 500. El objetivo del modelo es determinar si la variación del precio de cierre de un día es positiva o negativa respecto al precio de apertura, creando una herramienta que pueda replicar decisiones de inversión en las aperturas diarias de los mercados.

La metodología del estudio se fundamenta en el análisis técnico, construyendo un conjunto de variables predictoras basadas en indicadores ampliamente utilizados, como Bandas de Bollinger, MACD, RSI, ATR y medias móviles. Estas variables técnicas se calcularon utilizando datos históricos descargados de Yahoo Finance, abarcando el periodo de enero de 2023 a enero de 2025. Con el fin de optimizar la selección de variables predictoras, se utilizó la métrica “Mean Decrease Gini” para determinar que variables tienen mayor impacto en los modelos eliminando las variables que aporten menos información con el fin de mejorar los resultados sin comprometer la eficiencia del modelo.

La base de datos resultante se separó en dos subconjuntos. El primero el de train que abarcaba de junio de 2023 hasta junio de 2024 y el segundo el de prueba o test que abarcaba de julio de 2024 hasta enero de 2025 y que es donde se replicaría lo aprendido en el modelo mediante el entrenamiento para comprobar su eficiencia. Los hiperparámetros número de árboles de decisión (Ntree) y número de predictores por nodo (Mtry) se ajustaron mediante la prueba de diferentes configuraciones obteniendo unos valores de 160 para Ntree en los tres modelos. Por la parte de Mtry, se repitió 10 tanto en el modelo del DAX como en el del S&P500 mientras que para el IBEX se seleccionaron 5.

Los resultados obtenidos destacaron la eficiencia del modelo para predecir patrones alcistas de los índices siendo el mejor en líneas generales el DAX con una precisión

(Accuracy) del 57.69%. A destacar también en el IBEX35 su sensibilidad superior al 80%. Por su parte el modelo del S&P500 a pesar de no mostrarse tan eficiente como los otros dos sirvió como comparativo de estos.

Tras la evaluación de resultados se decidió realizar una simulación para contemplar la ganancia real que hubieran tenido los modelos en el periodo de prueba. Esta simulación se basó en comprar al inicio de la jornada bursátil siempre que el modelo lo predijera contemplando la salida de la operación en el cierre del mismo día. El capital de entrada sería acumulativo ya que tanto si se pierde como si se gana se debe contar con esta variación en el capital de entrada del siguiente día. Los resultados de esta simulación confirmaron la eficiencia del modelo tanto para el índice del IBEX como para el del DAX, consolidando a su vez al modelo de Random Forest como una herramienta óptima para el análisis bursátil.

2. Abstract

This work focuses on the application of machine learning techniques, specifically the Random Forest model, to predict intraday movements in three stock market indexes: IBEX 35, DAX 40 and S&P 500. The objective of the model is to determine whether the variation of the closing price of a day is positive or negative with respect to the opening price, creating a tool that can replicate investment decisions at daily market openings.

The methodology of the study is based on technical analysis, building a set of predictor variables based on widely used indicators, such as Bollinger Bands, MACD, RSI, ATR and moving averages. These technical variables were calculated using historical data downloaded from Yahoo Finance, covering the period from January 2023 to January 2025. In order to optimize the selection of predictor variables, the Mean Decrease Gini metric was used to determine which variables have the greatest impact on the models by eliminating variables that provide less information in order to improve the results without compromising the efficiency of the model.

The resulting database was separated into two subsets. The first was the training subset from June 2023 to June 2024 and the second was the test subset from July 2024 to January

2025, where the model's learning was replicated through training to check its efficiency. The hyperparameters number of decision trees (Ntree) and number of predictors per node (Mtry) were adjusted by testing different configurations, obtaining values of 160 for Ntree in the three models. For the Mtry part, 10 were repeated in both the DAX and S&P500 models while 5 were selected for the IBEX.

The results obtained highlighted the efficiency of the model in predicting bullish patterns of the indices, the DAX being the best in general terms with an accuracy (Accuracy) of 57.69%. The IBEX35 also stood out for its sensitivity of over 80%. On the other hand, the S&P500 model, although not as efficient as the other two, served as a comparison.

After the evaluation of results, it was decided to carry out a simulation to contemplate the real profit that the models would have made during the test period. This simulation was based on buying at the beginning of the trading day as long as the model predicted it, contemplating the exit of the operation at the close of the same day. The entry capital would be cumulative since both losses and gains should be taken into account in the next day's entry capital. The results of this simulation confirmed the efficiency of the model for both the IBEX and DAX indexes, consolidating the Random Forest model as an optimal tool for stock market analysis.

3. Introducción

3.1 El Mercado de Valores y los índices bursátiles

El mercado de valores o mercado bursátil es una rama del sistema financiero donde se realizan compras y ventas de activos entre los que destacan bonos, acciones o fondos. Es un pilar fundamental del sistema económico global permitiendo a las empresas conseguir capital para financiar sus actividades ofreciendo la oportunidad de obtener ganancias a través de la compra de acciones a sus inversores. Los índices bursátiles como el IBEX (España), el DAX 30(Alemania) y el S&P 500(Estados Unidos) representan un conjunto de las empresas más importantes de sus respectivos mercados nacionales y son indicadores clave de la salud económica.

Por otro lado, estos mercados financieros son complejos y volátiles, influenciados por factores económicos, políticos y emocionales. La predicción de su comportamiento ha sido un gran desafío constante para los inversores y analistas. La aparición de nuevas herramientas tecnológicas, como el aprendizaje automático, abre una nueva puerta hacia la elaboración más precisa de modelos analíticos que permitan anticipar estos movimientos, explorando patrones que no siempre son evidentes para los métodos tradicionales.

3.2 Machine Learning y su uso en el mercado de valores

El aprendizaje automático (Machine Learning, ML) es un subconjunto de la inteligencia artificial que permite a los sistemas aprender de un histórico de datos y realizar modelos predictivos sin ser programados explícitamente para cada caso. En el contexto financiero, el Machine Learning ha demostrado ser una herramienta valiosa para analizar grandes volúmenes de datos históricos y extraer relaciones complejas entre variables predictoras y resultados observados. En este proyecto, se emplean algoritmos de Machine Learning para predecir los movimientos de los índices bursátiles IBEX 35, DAX 30 y S&P 500.

3.3 Objetivo del Proyecto

El objetivo de este proyecto es diseñar, evaluar y comparar modelos predictivos capaces de determinar si al abrir los respectivos mercados el precio de cierre de ese día será superior o inferior al precio de apertura de ese día. Este enfoque es práctico, ya que replica escenarios en los que los inversores toman decisiones al inicio de la jornada bursátil basándose en predicciones que no dependen únicamente de variables donde la última información es el precio de cierre del día anterior, sino que, al darle el precio de apertura del día de la propia inversión y su correspondiente impacto en las variables utilizadas, el modelo posee una mayor precisión al estar actualizado. De esta manera los inversores tendrán un modelo que les ayude a tomar la decisión si entrar o no en el mercado ese día.

4. Marco Teórico

4.1 Índices bursátiles analizados

Un índice bursátil es un registro estadístico que miden la variación de un conjunto determinado de acciones que cotizan en un mercado, reflejando el comportamiento de la economía de un país, de un conjunto de empresas o de un sector determinado. Para la realización de este proyecto se han estudiado el comportamiento de los modelos propuestos sobre índices bursátiles nacionales. Estos índices sirven para evaluar el rendimiento general de la economía tanto de su respectivo país como de la economía global. Se ha escogido:

IBEX 35

El IBEX 35. Es el índice de referencia de nuestro país. Este compuesto por las 35 empresas del mercado nacional con mayor liquidez y mayor capitalización. Fue creado en 1992 por las Bolsas y Mercados Españoles (BME) y se actualiza periódicamente con la entrada y salida de empresas según su ganancia o pérdida de importancia económica reflejando así los cambios propios de la economía nacional.

Incluye empresas de diversos sectores como banca, energías, construcción e infraestructuras y telecomunicaciones proporcionando una representación amplia de la economía de España.

DAX 40

El DAX 40 es el principal índice bursátil de Alemania escogiéndose en parte para tener una comparativa del modelo aplicado al IBEX con otro modelo que trabaje con un índice europeo. Agrupa a las 40 (hasta 2021 eran las 30) principales empresas del mercado alemán que cotizan en la bolsa de Frankfurt. No sirve solo como barómetro de la economía germana, también es uno de los índices más representativos de Europa al ser Alemania uno de los motores económicos de Europa. Incluye diversos sectores clave como la automoción, tecnología o manufactura industrial.

S&P 500

El S&P 500 es junto al NASDAQ el principal representante del mercado estadounidense en lo que respecta a los índices bursátiles. Está integrado por las 500 empresas más grandes de Estados Unidos. Su amplia diversificación sectorial, así como su gran volumen de capitalización lo hacen ser uno de los índices más seguidos en el mercado de valores tanto por inversores como por analistas de todo el mundo representando un termómetro de la economía global. Contando con el S&P 500 en el proyecto se logra una comparación intercontinental de los resultados del modelo, evaluando si los patrones observados se replican en diferentes mercados.

El análisis simultáneo de estos índices ofrece una visión comparativa del comportamiento de los mercados permitiendo evaluar patrones comunes y divergencias en los movimientos del mercado.

4.2 Fundamentos de Machine Learning

El Machine Learning se basa en la creación de estructuras analíticas capaces de realizar predicciones utilizando un histórico de datos sin tener que ser programados los modelos de forma particular. En la actualidad el Machine Learning ha transformado la gran mayoría de sectores permitiendo la automatización de sistemas a través del estudio y modelación de patrones complejos que en el pasado eran muy difíciles de poder detectar o interpretar.

En lo que respecta al sector financiero, el ML ha tenido un gran impacto aportando a las empresas procesos más eficientes y análisis más profundos. La gran mayoría de entidades del sector han incorporado a sus empresas departamentos para abordar por completo esta tecnología además de ser muy habitual la colaboración de estas con las denominadas Fintech (empresas centradas en la combinación de tecnología y servicios financieros).

Por otra parte, en lo que compete a los inversores, el ML ha brindado de nuevas e innovadoras herramientas para la optimización de sus estrategias de inversión ayudando el fácil acceso a esta tecnología en la actualidad y su bajo costo, los inversores pueden

tanto realizar modelos que guíen por completo sus fondos o simplemente apoyarse en ellos para ayudar a realizar acciones premeditadas por ellos mismos. Se hace muy sencillo a día de hoy acceder a plataformas que te ayuden a realizar estos modelos siendo de ejemplo las páginas donde antes los inversores accedían simplemente para informarse de la variación de los precios.

En su inmensa mayoría estas han añadido nuevas donde secciones donde el usuario puede empezar a practicar con estrategias sencillas ya creadas o crear las suyas propias. Esto combinado con la gran cantidad de material al que se puede acceder por internet hace que el ML sea accesible para todo inversor que quiera intentar encontrar nuevas estrategias basadas en algoritmos predictivos.

Se puede comenzar con el desarrollo de algoritmos simples con el uso de una sola variable sencilla de comprender como puede ser la media móvil o algún indicador de totalidad e ir probando a añadir otras derivadas de las primeras para tener un modelo más completo que pueda obtener una mayor información.

4.2.1 Técnicas de Machine Learning

Generalmente se diferencian tres tipos diferentes de aprendizaje:

- **Aprendizaje supervisado:** Este algoritmo cuenta con un aprendizaje previo basado en un entrenamiento con datos etiquetados. Se proporcionan tanto las entradas (variables predictoras) como las salidas (resultados deseados o variable(s) a predecir).
- **Aprendizaje no supervisado:** El modelo no cuenta con un conocimiento o entrenamiento previo (los datos no están etiquetados) instando al modelo a que encuentre patrones o estructuras con el conjunto global de datos
- **Aprendizaje por refuerzo:** El modelo aprende a través de la interacción con el entorno es decir a partir de su propia experiencia, optimizando su comportamiento mediante recompensas o penalizaciones a base de prueba y error.

“El aprendizaje supervisado, también conocido como Machine Learning supervisado, se define por su uso de conjuntos de datos etiquetados para entrenar algoritmos que clasifiquen datos o predigan resultados con precisión. A medida que se introducen datos de entrada en el modelo, éste ajusta sus ponderaciones hasta que se ha ajustado adecuadamente. Esto ocurre como parte del proceso de validación cruzada para garantizar que el modelo evite el sobreajuste o el infra ajuste. El aprendizaje supervisado ayuda a las organizaciones a resolver una variedad de problemas del mundo real a escala, como clasificar el spam en una carpeta separada de su bandeja de entrada. Algunos métodos utilizados en el aprendizaje supervisado son las redes neuronales, el clasificador bayesiano ingenuo, la regresión lineal y logística, el bosque aleatorio (Random Forest) y la máquina de vectores de soporte (SVM).” (1)

Al no tener una referencia clara en lo que es la selección de variables se seleccionó el método de Random Forest ya que además de ser ampliamente reconocido por su capacidad para modelar relaciones no lineales y manejar conjuntos de datos complejos también ofrece diferentes métricas que muestran la contribución de cada una de las variables al modelo, pudiendo seleccionar las de mayor importancia realizando una criba de las variables iniciales y seleccionar las más importantes para el modelo final.

Al modelo se le da un histórico de datos etiquetados de un índice bursátil obteniendo este histórico de Yahoo Finance (aquí se encuentran métricas básicas que se los facilitan directamente como el volumen diario, el precio máximo y mínimo...). Tras esto, se añaden nuevas variables técnicas a la base de datos construidas en su mayoría a partir del precio de apertura (entre las que se encuentra la variable a predecir, el Rendimiento Open to Close) y otras variables derivadas de estas. Se realiza un entrenamiento del modelo supervisado con todas las variables tras lo que se hace una selección de variables con Mean Decrease Gini. Se crea el modelo final y se realiza una evaluación de resultados, así como una comparación de los diferentes modelos.

(1) Extraído de <https://www.ibm.com/es-es/topics/supervised-learning>

4.2.2 Algoritmos de Clasificación: Random Forest

Random Forest (RF o árboles de decisión) es una de las técnicas de Machine Learning más populares entre los analistas o científicos de datos debido a su versatilidad, fácil interpretación y a su fácil validación. La técnica se basa en la creación de árboles de decisión y que utiliza un método de bifurcación(ramas) para representar los posibles resultados al ejecutar una decisión. Cada uno de estos árboles es entrenado a través de una muestra aleatoria de los datos obtenido del conjunto de entrenamiento que se ha definido previamente. Para predecir la variable o variables objetivo se tiene en cuenta los resultados de todos los árboles haciendo así que el resultado sea más preciso. Estas características hacen que RF sea un modelo de clasificación y predicción eficiente con capacidad para manejar grandes volúmenes de datos como es el caso del estudio.

En el algoritmo de Random Forest hay que definir de forma previa al entrenamiento dos hiperparámetros básicos siendo estos el número de árboles (Ntree) y el Número de Predictores por nodo (Mtry).

Ntree representa el número de árboles de decisión generados. Cuanto mayor sea el número de árboles más robusto es el modelo aumentando así tanto su precisión como su estabilidad y su capacidad para patrones complejos. Por otra parte, el sobre incremento del número de árboles puede suponer problemas como la necesidad de aumentar el tiempo de entrenamiento o la posibilidad de sobre ajustar el modelo. Aunque el riesgo de sobreajuste en Random Forest es muy bajo con un número demasiado grande de árboles el modelo puede aprender en exceso los detalles del conjunto de datos del entrenamiento ocasionando problemas en la generación de nuevos datos.

Mtry controla el número de variables aleatorias que se consideran en cada división del árbol. Al igual que en Ntree se trata de maximizar el Mtry teniendo en cuenta que puede tener un excesivo número de predictores por nodo puede llevar a un sobreajuste y a una menor diversidad entre árboles.

Random Forest se utiliza en el análisis bursátil en diversos procesos como por ejemplo en la predicción de precios y variaciones tanto bajistas como alcistas además de para establecer relaciones entre diversas métricas o indicadores técnicos. En este proyecto se

construyeron tres modelos de RF aplicados a los índices bursátiles previamente mencionados. Los modelos utilizan el mismo periodo de entrenamiento, así como las mismas variables iniciales. Por otra parte, difieren en las variables del modelo final además de en número de Mtry seleccionados.

5. Metodología

5.1. Descripción de los datos utilizados

El modelo utiliza datos históricos obtenidos de la plataforma de Yahoo Finance. Esta es una fuente de confianza y altamente utilizada por inversores para estar al tanto de información financiera global. Los índices que se van a utilizar en el modelo son el IBEX35, el DAX40 y el S&P 500, siendo estos índices representativos de los mercados de valores de España, Alemania y Estados Unidos respectivamente.

El periodo de análisis abarca desde junio de 2023 hasta enero de 2025. Los datos se han dividido en dos conjuntos diferentes:

Conjunto de entrenamiento: desde junio de 2023 hasta junio de 2024.

Conjunto de prueba: desde julio de 2024 hasta enero de 2025.

El enfoque principal del estudio es predecir el movimiento de los índices desde la apertura (Open) hasta el cierre (Close) del mercado en un mismo día. Esta decisión permite centrarse en las variaciones intradía, que son fundamentales para operaciones de corto plazo.

Frecuencia de los datos: Se ha trabajado con datos diarios para cada índice. Cada observación incluye información sobre los precios de apertura, cierre, máximo y mínimo del día, así como el volumen diario de operaciones y la fecha correspondiente.

5.2. Variables

En este apartado se describen las variables utilizadas en los modelos para predecir el movimiento intradía de los índices IBEX 35, DAX 40 y S&P 500. Estas variables fueron seleccionadas por su capacidad para capturar diferentes aspectos de los movimientos del mercado, tales como tendencias, volatilidad y momentum, y se derivaron principalmente a partir de los precios de apertura (Open) de cada día, respetando el objetivo del trabajo: realizar predicciones basadas en datos disponibles en el momento de la apertura del mercado.

5.2.1 Variables técnicas básicas:

-Rendimiento diario:

El rendimiento diario es una variable técnica que calcula el cambio porcentual entre el precio de apertura de un día respecto al precio de apertura del día anterior. Es una medida fundamental ya que resume como han variado las expectativas de mercado entre el inicio de dos días consecutivos (nótese que en todo momento en este trabajo se hablan de días de mercados por lo que no se tienen en cuenta días donde los correspondientes mercados de valores no estén operativos tales como fines de semana o festivos nacionales de los países de los índices utilizados)

Para calcularla se resta al precio de apertura del día actual el precio de apertura del día anterior dividiendo este resultado por el precio de apertura del día anterior y multiplicando el resultado por cien.

Una sucesión de resultados positivos consecutivos de esta variable puede sugerir un indicativo de una tendencia alcista en el índice estudiado. Por contra, un rendimiento negativo regular puede sugerir una tendencia bajista.

-Media Móvil (MA):

La media móvil es un indicador estadístico que combina los precios de un activo (en este caso del índice bursátil correspondiente) de un rango específico de tiempo seleccionada para que se tenga en cuenta tanto el momento actual como los movimientos recientes del índice suavizando así las fluctuaciones del precio mostrando así unas tendencias

generales. Se suele utilizar para comprobar si el mercado esta alcista o bajista dependiendo si el precio es superior o inferior a la media móvil. Se establece un periodo de 20 días para la media móvil simple (aproximadamente el número de días por mes en el que opera el mercado) aplicada a los precios de apertura.

$$SMA = \frac{A_1 + A_2 + \dots + A_n}{n}$$

Dónde:

A_n = precio de un activo en el período n

n = número de periodos totales

Ilustración 1 Imagen obtenida de vectorcrypto.com

-Señal MA:

La señal MA es una variable discreta derivada de la variable anterior que muestra de forma más directa la tendencia del mercado. Si el precio de apertura del día esta por encima del valor de la media móvil la señal MA es 1 y si el precio de apertura es menor la señal MA es -1.

-Volumen promedio

Variable que ofrece directamente Yahoo Finance y que muestra el número de operaciones en un determinado rango de tiempo, 20 días en el modelo. Sirve como indicador de interés por parte de los inversores pudiendo mostrar una confirmación o cambio de tendencias.

5.2.2 Indicadores de volatilidad:

-Bandas de Bollinger (BB):

Las bandas de Bollinger son uno de los indicadores más populares utilizados en el análisis técnico de los mercados financieros introducidos por John Bollinger en 1980. Sirven para medir la volatilidad sobre una banda central que es la media móvil dando a los inversores información sobre si un activo esta sobre o infravalorado

Sobre esta primera banda se generan dos más, una superior y otra inferior. La superior se calcula sumando a la media móvil 2 veces el valor de su desviación estándar. Por su parte la inferior se genera restando 2 veces este valor a la media móvil. Generalmente se han utilizado para identificar rangos temporales de sobrecompra (cuando el precio sobrepasa la banda superior) o de sobreventa (cuando el precio está por debajo de la banda inferior). En el modelo se utilizan para evaluar si los precios diarios de apertura de los 3 índices se encuentran en extremos de volatilidad.

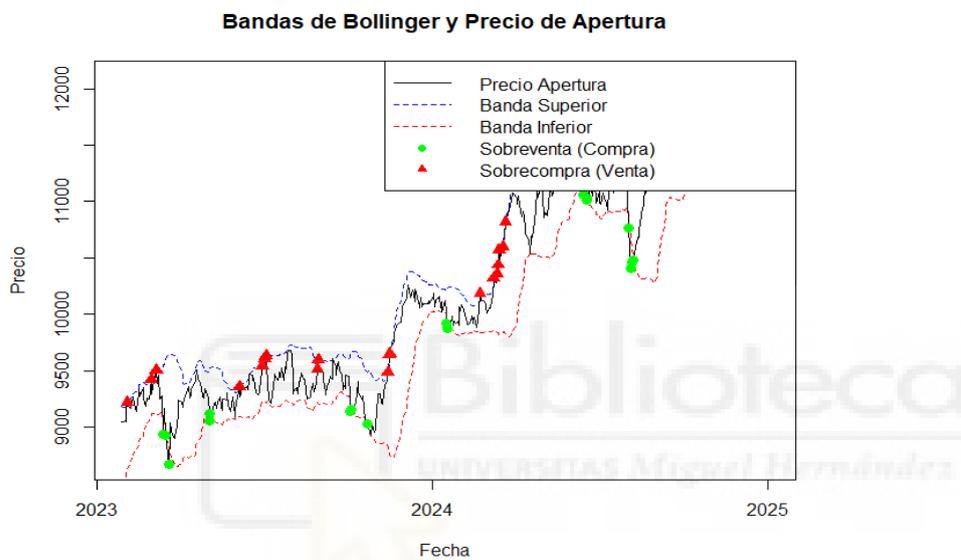


Ilustración 2 Ejemplo Bandas de Bollinger en IBEX. Fuente: Elaboración propia

En la Ilustración 2 se pueden observar en color azul y rojo las bandas de Bollinger superior e inferior respectivamente. Los indicadores verdes indican puntos donde el precio cae por debajo de la banda inferior reflejando oportunidades de entrada o compra al igual que los indicadores rojos marcan situaciones donde el precio es mayor a la banda superior sugiriendo oportunidades de venta o salidas.

$$Upper\ Band = Moving\ Average + Constant \sqrt{\frac{\sum_{i=1}^n (y_i - Moving\ Average)^2}{n}}$$

$$Lower\ Band = Moving\ Average - Constant \sqrt{\frac{\sum_{i=1}^n (y_i - Moving\ Average)^2}{n}}$$

Ilustración 3 . Fórmulas bandas de Bollinger. Fuente: kaabar-sofien.medium.com

-Señal BB

La señal de las Bandas de Bollinger Variable discreta derivada de la anterior que permite una interpretación directa ya que, si el precio de apertura es mayor que la banda superior la señal BB es 1, si es menor que la banda inferior el valor de la señal BB es -1 y en caso en el que se encuentre entre las dos bandas el valor es 0.

-Distancia BB Upper y BB Lower:

Las distancias superior o inferior de las Bandas de Bollinger son otras variables derivadas de las bandas de Bollinger que se centran en la posición del precio de apertura sobre las bandas, Una mayor proximidad a la banda superior puede sugerir al analista una posible resistencia mientras que una cercanía a la banda inferior puede suponer un nivel de soporte. Esta variable tiene una interpretación más compleja que la señal BB pero ofrece una visión más detallada y continua de la interacción del precio con las bandas permitiendo anticipar movimientos futuros al estudiar la intensidad con la que el precio se aproxima o se aleja de los límites del canal.

$$\text{Distancia BB Upper} = \frac{P_{\text{apertura}} - BB_{\text{superior}}}{P_{\text{apertura}}}$$

$$\text{Distancia BB Lower} = \frac{P_{\text{apertura}} - BB_{\text{inferior}}}{P_{\text{apertura}}}$$

-ATR (Average True Range):

El ATR es un indicador que se utiliza para medir la volatilidad del mercado desarrollado por el analista Welles Wilder en 1978, Su objetivo es medir únicamente la volatilidad y no las tendencias del mercado. Se basa en una media ajustada de los valores del rango real de un activo en un periodo de tiempo determinado permitiendo en el modelo determinar la volatilidad de los precios de aperturas del modelo y facilitando puntos de entrada y de salida del índice bursátil. Un ATR elevado indica variaciones significativas en el mercado, mientras que valores bajos reflejan condiciones más estables y predecibles. El rango de tiempo seleccionado en el modelo es el más común para ATR, 14 días de mercado.

$$TR = \max[(H - L), |H - C_P|, |L - C_P|]$$

$$ATR = \frac{1}{n} \sum_{i=1}^n TR_i$$

- TR_i : A particular True Range
- n : The time period employed
- H : Current High
- L : Current Low
- C_P : Previous Close

Ilustración 4. Fórmula ATR. Fuente: <https://stackoverflow.com/>

-ATR Ratio:

El ATR Ratio es un indicador derivado del anterior que ajusta el valor del ATR dividiéndolo por el precio de apertura. Esta adaptación estandariza la volatilidad con respecto al nivel de precio, lo que permite realizar comparaciones más directas entre activos o en diferentes momentos del tiempo. Es especialmente útil para analizar periodos con precios variables permitiendo medir la volatilidad de manera relativa, asegurando que los movimientos se evalúen en función del contexto del precio y no únicamente de la magnitud de las variaciones

$$ATR \text{ Ratio} = \frac{ATR}{\text{Precio Apertura}}$$

-Volatilidad semanal:

La volatilidad semanal es un indicador calculado a partir de la diferencia entre el precio apertura más alto y el más bajo de los últimos 5 días de mercado. Sirve para obtener información sobre la actividad reciente del activo permitiendo evaluar si se está analizando sobre una semana volátil o estable.

$$\text{Volatilidad Semanal} = \max(P_{\text{apertura}, t=5}) - \min(P_{\text{apertura}, t=5})$$

5.2.3 Indicadores de momentum:

-MACD (Moving Average Convergence Divergence):

El MACD es un indicador de impulso que mide la relación entre dos promedios móviles del precio (normalmente entre de 12 y 26 periodos) para evaluar la fuerza, dirección y duración de una tendencia. La diferencia entre el MACD y su línea de señal, media móvil de 9 periodos (llamada EMA, media móvil exponencial) muestra momentos de convergencia o divergencia con estas y según la proximidad de estas dos métricas se encuentran puntos de entrada o salida de la inversión., resalta posibles puntos de entrada o salida, identificando momentos de convergencia o divergencia.

- Diferencia MACD

Métrica derivada de la anterior variable obtenida de la resta de los dos promedios de 12 y 26 periodos

$$\text{Diferencia MACD} = EMA_{12} - EMA_{26}$$

-Señal MACD

La señal MACD es un indicador derivado de la MACD con el fin de simplificar su interpretación siendo una variable binaria. Se calcula la Señal MACD (no es la variable final) derivada de la EMA y por otra parte se cuenta con la línea MACD de la variable anterior. Si la línea está por encima de la señal MACD (primitiva) entonces la Señal MACD final es 1 y muestra un posible periodo alcista. Si por el contrario está por debajo el valor de la Señal MACD final es 0 e indica una posible tendencia bajista.

-RSI (Relative Strength Index):

El RSI, al igual que el MACD es un indicador de impulso que muestra si un activo está sobrecomprado y que posee un valor de 0 a 100. En un primer momento se calcula el promedio de variaciones al alza y el promedio de variaciones a la baja de un periodo (normalmente 14 días como en el modelo). Con esto se calcula el RS siendo la división del promedio alcista sobre el bajista y por último el $RSI = 100 - 100 / (1 + RS)$, obteniendo

así una relación entre las ganancias y pérdidas de los precios durante los últimos 14 periodos.

Si el RSI es superior a 70 se considera que el índice está sobrecomprado y propenso a una corrección. Por otra parte, si es menor a 30 indica una sobreventa. Al igual que el resto de variables, en el modelo el RSI se basa en los precios de apertura, lo que proporciona una perspectiva oportuna para anticipar posibles correcciones o continuaciones de tendencia.

$$RS = \frac{\text{Promedio de Subidas}}{\text{Promedio de Bajadas}}$$

$$RSI = 100 - \frac{100}{1 + RS}$$

-TRIX:

El TRIX es un indicador que sirve para filtrar los movimientos de los precios de un activo, descartando las variaciones insignificantes y centrándose en tendencias sostenibles y duraderas. Representa la tasa de cambio del EMA triple suavizado durante un periodo de 15 días. El TRIX es útil para evitar señales falsas mejorando la capacidad del modelo para identificar ciclos de largo plazo. Los valores positivos del TRIX indican una tendencia alcista y negativos una tendencia bajista.

$$EMA = EMA(\text{periodo}) \text{ de cierre}$$

$$EMA^2 = EMA(\text{periodo}) \text{ de } EMA$$

$$EMA^3 = EMA(\text{periodo}) \text{ de } EMA^2$$

$$TRIX_i = \frac{EMA^3_i - EMA^3_{i-1}}{EMA^3_{i-1}} * 100$$

5.2.4 Indicadores de tendencia:

-EMA (Exponential Moving Average):

La EMA es una media móvil exponencial que otorga más importancia a los precios recientes haciendo que el modelo tenga una versión real y actualizada de las variaciones del precio del activo usando en este caso un periodo de 20 días de mercado. Al centrarse en las aperturas, la EMA ayuda a mejorar la capacidad de respuesta del modelo frente a los cambios del mercado.

$$EMA_t = \partial * P_t + (1 - \partial) * EMA_{t-1}$$

-Tendencia semanal:

Indicador que mida la variación de los precios de apertura de forma semanal teniendo así un contexto mayor que la tendencia diaria.

$$Tendencia\ Semanal = P_{apertura, t} - P_{apertura, t-5}$$

5.2.5 Variable objetivo:

-Resultado siguiente:

Se trata de la variable objetivo y la base del modelo. Es binaria e indica si el precio de cierre es mayor que el precio de apertura para un día. Si hay una subida respecto al inicio de la operativa diaria la variable tendrá valor de 1 y si por contrario el precio ha bajado durante la jornada el valor es 0, evaluando de esta manera si las señales generadas permiten anticipar correctamente la variación positiva o negativa durante la sesión de mercado.

5.3. Preprocesamiento de los Datos y selección de variables

El preprocesamiento de datos tiene como objetivo garantizar la validez de los datos utilizados en el proyecto asegurándose con esta etapa de que no se trabaje sobre valores nulos, perdidos o atípicos provenientes de un fallo en la carga de la base de datos, así como incongruencias en las fechas y métricas (que los valores reflejen un día distinto por ejemplo)

5.3.1. Eliminación de Datos Faltantes

Con el fin de evitar errores en el entrenamiento del modelo, así como en modelo predictivo final se realiza un barrido buscando observaciones vacías o nulas. Como algunas variables requieren periodos previos ya que se calculan sobre periodos extensos de días en un primer momento se cargan datos del 1 de enero de 2023 hacia adelante para asegurarse que llegado el inicio del periodo de entrenamiento todas las variables estarían correctamente calculadas. Además, se realiza una búsqueda y eliminación de valores nulos para asegurarse por completo del correcto funcionamiento del modelo. Estos fueron eliminados usando la función `na.omit()`, garantizando que el conjunto de datos estuviera completo.

5.3.2. Alineación de las Variables

Siguiendo con el punto anterior se realiza una alineación temporal de las variables para asegurar que las observaciones que hemos descargado de Yahoo Finance hagan referencia a los días correspondientes. Se añade la función `merge()` con el fin de eliminar los días en los que faltara alguna observación y dejando solo las fechas en las que la información fuera completa.

5.3.3. División en Conjuntos de Entrenamiento y Prueba

El conjunto de datos se dividió en dos partes:

- **Entrenamiento:** Desde junio de 2023 hasta junio de 2024.
- **Prueba:** Desde julio de 2024 hasta enero de 2025.

Se escoge un periodo amplio para el entrenamiento, pero en la medida de lo posible reciente para no entrar en ciclos pasados. Se escoge un periodo de prueba de siete meses para que se pueda tener en cuenta diferentes tendencias del mercado y no aclimatarse a una única.

5.3.4. Selección de Variables Relevantes

En primera instancia se crea el modelo de Random Forest utilizando todas las variables descritas anteriormente para después hacer una selección de estas escogiendo las más determinantes para el modelo. Existen diferentes métodos para realizar esta selección en Random Forest. En este caso se usa una evaluación de la importancia de las variables a través de Mean Decrease Gini. Esta métrica actúa como una métrica que mide la contribución de cada variable a la homogeneidad de los nodos y ramas en el bosque aleatorio resultante. Cuanto mayor sea la puntuación de Gini de la disminución media, mayor será la importancia de la variable en el modelo.

$$Gini = 1 - \sum p_i^2$$

Donde p_i es la proporción de elementos de la clase i en la partición.

Cuando un nodo de un árbol se divide usando una variable X_i , el índice de Gini se reduce. El MDG mide el descenso promedio de este índice a lo largo de todos los árboles y nodos donde X_i fue utilizado.

Las variables con un valor bajo tienen menor impacto y se eliminan con el fin de simplificar el modelo.

Se escogió un umbral de selección de variables para todos los modelos de 7.5. Las variables con un valor superior al umbral fueron seleccionadas para el entrenamiento final del modelo, mientras que el restantes fueron descartadas.

A continuación, se muestran las variables seleccionadas para cada índice:

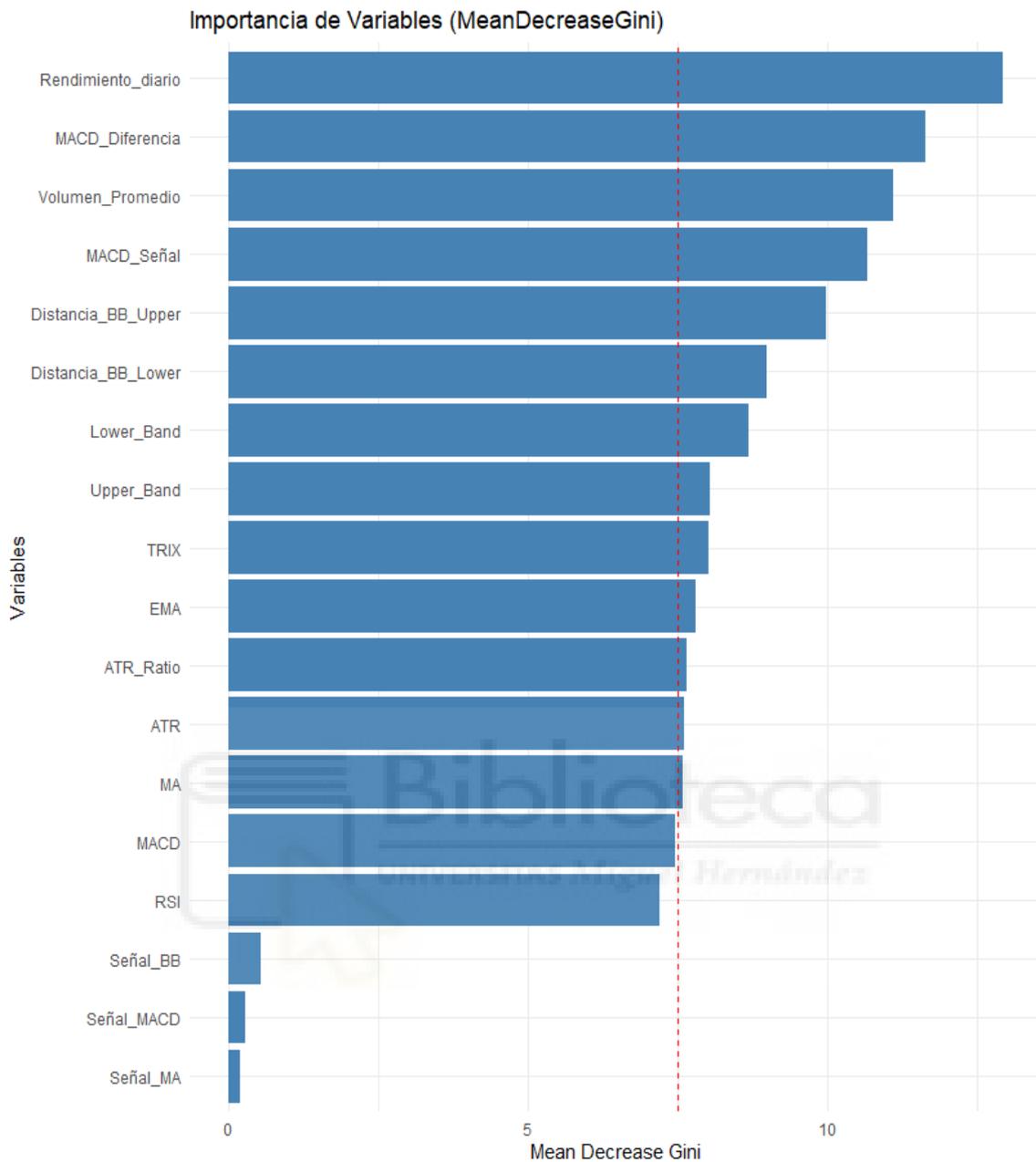


Ilustración 5 Importancia de las variables en IBEX. Fuente: Elaboración propia

Para el modelo del IBEX se han seleccionado: Rendimiento_diario, MA, Upper_Band, Lower_Band, MACD_Señal, MACD_Diferencia, TRIX, ATR, Volumen_Promedio, Distancia_BB_Upper, Distancia_BB_Lower, ATR_Ratio, EMA. En total pasan el umbral 12 variables.

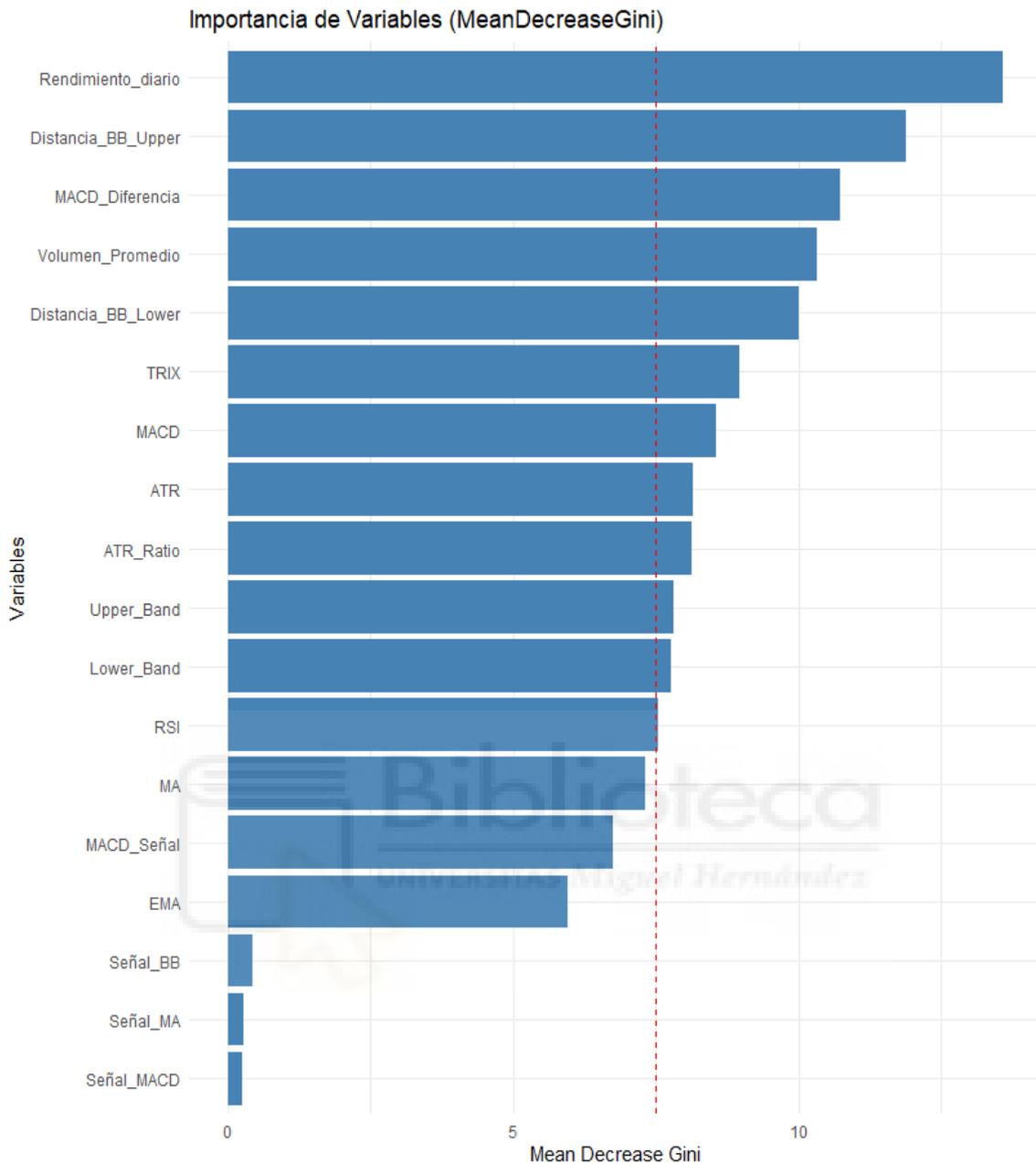


Ilustración 6 Importancia de las variables en DAX. Fuente: Elaboración propia

Para el modelo del DAX se han seleccionado: Rendimiento_diario, Upper_Band, Lower_Band, MACD, MACD_Diferencia, TRIX, RSI, ATR, Volumen_Promedio, Distancia_BB_Upper, Distancia_BB_Lower, ATR_Ratio.

Para este índice pasan el umbral un total de 13 variables.

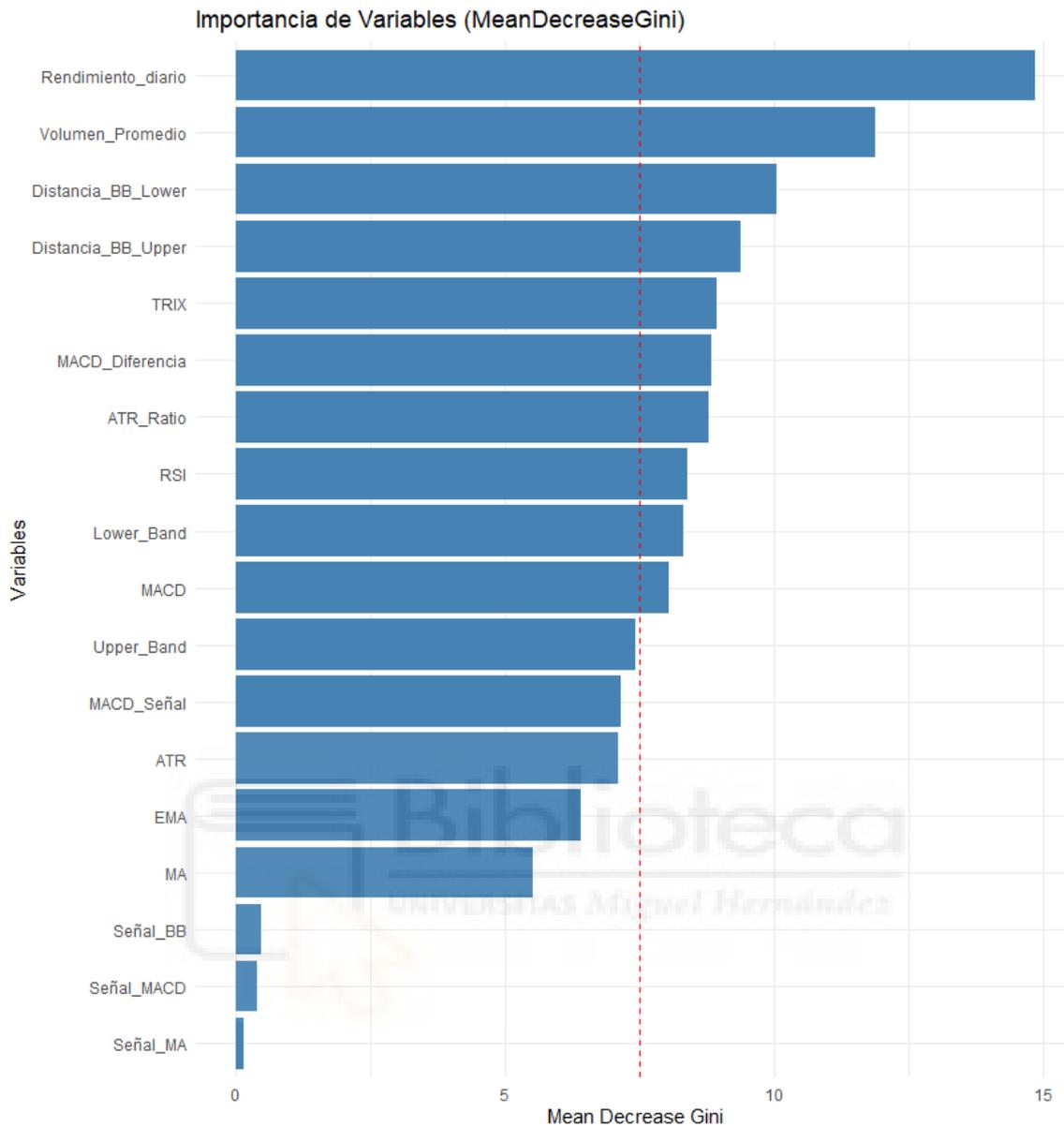


Ilustración 7 Importancia de las variables en SP500. Fuente: Elaboración propia

Para el modelo del SP500 se han seleccionado: Rendimiento_diario, Lower_Band, MACD, MACD_Diferencia, TRIX, RSI, Volumen_Promedio, Distancia_BB_Upper, Distancia_BB_Lower, ATR_Ratio.

Para el S&P500 se encuentran 10 variables cuya importancia en Mean Decrease Gini es mayor a 7.5.

6. Desarrollo del modelo

6.1. Implementación del Modelo Random Forest

El modelo de Random Forest se ha llevado a cabo en R Studio, un entorno de desarrollo integrado (IDE) para el lenguaje de programación R centrado en la computación estadística. Se han utilizado diferentes librerías para este modelo como:

- Quantmod para descargar los datos del mercado y calcular indicadores técnicos.
- randomForest siendo la librería del propio modelo predictivo además de ser usada para el cálculo de la importancia de las variables
- caret para la evaluación y validación de los modelos
- dplyr usada para la manipulación y limpieza de datos
- ggplot2 centrada en gráficos avanzados
- tidyr transformación y restructuración de datos
- TTR para el cálculo de variables técnicas como MACD, RSI o ATR
- Zoo para la manipulación avanzada de series temporales.

Es habitual el uso de Random Forest en R para la creación de modelos predictivos aplicados al mercado de valores siendo el RF uno de los mejores métodos de aprendizaje supervisado para el análisis de series temporales extensas.

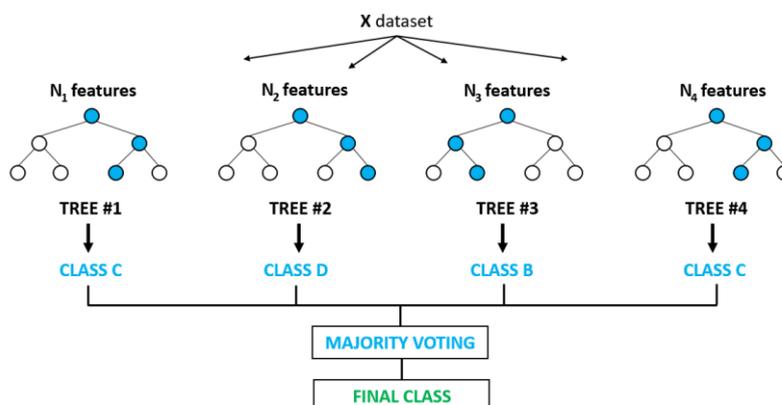


Ilustración 8 Modelo básico Random Forest. Fuente: rpubs.com

En primer lugar, se descargaron los datos históricos diarios de los 3 índices estudiados (IBEX35, DAX40 y S&P 500) desde enero de 2023 hasta enero de 2025 desde la fuente de Yahoo Finance obteniendo el precio de apertura (Open), precio de cierre (Close), volumen de transacciones (Volume), máximo diario (High) y mínimo diario (Low).

Tras ello se crearon indicadores técnicos además de indicadores derivados de estos primeros que resultarían en las 18 variables iniciales.

Se realiza una limpieza de datos para asegurar la consistencia de las observaciones eliminando con `merge()` y `na.omit()` posibles registros nulos o no completos que pudieran mal influir en el modelo además de realizar una alineación de las variables con las fechas correspondientes.

Se separa en dos subconjuntos siendo el primero el entrenamiento(`train`) y el segundo el conjunto de prueba donde se analizará la efectividad del modelo(`test`), además se realiza en este punto la selección de variables importantes para cada modelo.

6.2. Entrenamiento y Validación

Una vez creado los dos conjuntos se entreno el modelo utilizando el primero.

El proceso de entrenamiento se centró en optimizar el modelo Random Forest a través de la manipulación de los dos hiperparámetros fundamentales:

Ntree: Se usaron valores desde 100 hasta 220 árboles alcanzando en 160 el valor óptimo para el modelo en los tres índices.

Mtry: Se usaron los valores 5, 8 y 10 encontrando el primer valor el óptimo para el modelo del IBEX y el último para los modelos del DAX y del S&P50.

Resultado del entrenamiento:

Índice	Ntree	Mtry	Accuracy(%)	Sensitivity(%)	Specificity(%)
IBEX	160	5	51,52	83,33	29,49
DAX	160	10	57,69	63,64	56,33
S&P500	160	10	51,05	37,88	62,34

Ilustración 9 Tabla de elaboración propia

6.3. Métricas de Evaluación

Para evaluar el rendimiento del modelo, se utilizaron las siguientes métricas clave:

Accuracy (Precisión Global): Medición del porcentaje de predicciones correctas sobre el total de observaciones.

Sensitivity (Sensibilidad): Capacidad del modelo para predecir de forma correcta los días de subida. Es importante para los modelos que se basen en estrategias alcistas.

Specificity (Especificidad): Capacidad del modelo para predecir de forma correcta los días de bajada. Es importante para los modelos que se basen en estrategias bajistas.

Estas métricas son habituales en los modelos de Random Forest y permiten evaluar de forma la capacidad predictiva del modelo.

7. Análisis de Resultados

7.1. Métricas de Evaluación

Tras la finalización del modelo se evalúa el desempeño del modelo Random Forest para los tres índices, obteniéndose los siguientes resultados:

DAX 40:

-Precisión: 57.69%.

-Sensibilidad: 63.64%.

-Especificidad: 53.33%.

En este índice el modelo obtiene un muy buen resultado con un accuracy por encima del 50% y mostrando una gran eficacia para predecir subidas con una sensibilidad cercana al 64%. Por su parte su predicción de bajadas no es destacable.

IBEX 35:

-Precisión: 51.52%.

-Sensibilidad: 83.33%.

-Especificidad: 29.49%.

En este índice el modelo presenta una precisión mediocre que no supera el 52% aunque con una sensibilidad mayor al 80% muestra una eficacia excepcional para la predicción de subidas. Por otro lado, es ineficiente para la predicción de bajadas. La gran diferencia entre sensibilidad y especificidad desequilibrio sugiere que el modelo tiende a sobreestimar las subidas pudiendo generar esto decisiones erróneas en mercados bajistas.

S&P 500:

-Precisión: 50%.

-Sensibilidad: 38.33%.

-Especificidad: 60%.

Los resultados del S&P 500 son más discretos en comparación con los de los otros dos índices, destacando únicamente una aceptable especificidad del 60% para la predicción de bajadas, pero siendo claramente poco preciso en la detección de subidas.

Interpretación de los Resultados:

- **DAX 40:** Presenta los mejores resultados, teniendo un rendimiento sólido y equilibrado, aunque destacando claramente su faceta para la predicción de subidas.
- **IBEX 35:** Su baja especificidad hace que el modelo solo pueda ser considerado para la predicción de subidas en la que presenta el mayor porcentaje obtenido de los 3 modelos estudiados.

- **S&P 500:** Es el modelo con el rendimiento menos destacado, solo defendido por su aceptable especificad no se tiene en cuenta en la elección de la decisión estratégica oportuna, aunque se realizará de todas formas a modo de comparativa.

Siendo los resultados de las métricas claramente mejores en la sensibilidad se concluye que el modelo estratégico a desarrollar debe ser una estrategia aplicada únicamente a la simulación de predicciones de subidas.

7.2. Simulación de Inversión y análisis de resultados

Una vez decidido que el modelo se va a aplicar para intentar detectar movimientos de subidas en Open to Close de las jornadas bursátiles se realiza una simulación para cada índice en la que se empieza con un capital inicial de 10.000 euros. Se invertirá tras el open de cada día siempre que el modelo prediga que el índice va a subir (1). El capital es acumulado, es decir si el primer día que se invierte se pierden 100 el siguiente día se invierten 9.900 y así de forma continua. Cuando la predicción del modelo sea 0(día bajista) no se invertirá.

Resultados de la Simulación:

Se presentan también a continuación los gráficos generados en R que resumen los resultados para cada índice mostrando el beneficio acumulado de cada modelo.

1. DAX 40:

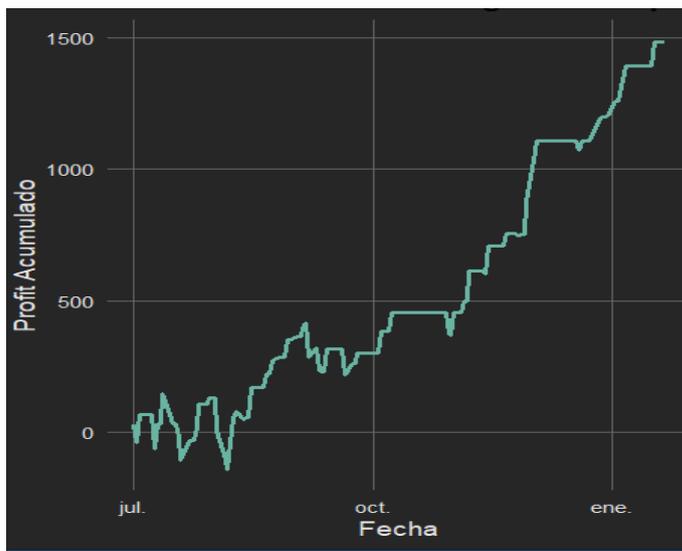


Ilustración 10 Profit acumulado del modelo en DAX

A día 21 de enero de 2025 el profit acumulado en el modelo de DAX es +1485.48

El modelo predijo subidas en 32 días del total de días con mercado abierto desde el día 1 de julio de 2024. De esos 62 días se cumplió en 42 días.

El modelo tuvo un acierto del 67.74 % para la predicción de subidas en el periodo estudiado. Otra virtud a destacar es que la pérdida diaria más grande fue de 130 y el beneficio mínimo acumulado que se dio en toda la simulación fue de -140 siendo un dato muy bajo que en la vida real no hubiera expulsado a inversores como si puede pasar si durante el desarrollo de la inversión se dan pérdidas significativas haciendo esto que el inversor saque el capital antes de que el modelo consiga funcionar.

Los resultados son muy buenos teniendo un porcentaje de beneficio sobre el capital inicial de un 14.85% en apenas 6 meses lo que demuestra la gran efectividad del modelo y anima a llevarlo a la práctica de forma real. Para concluir añadir que la ganancia máxima diaria fue de 192, dato que indica que no hay ninguna ganancia anómala que afecte a la lectura modelo.

2. IBEX 35:

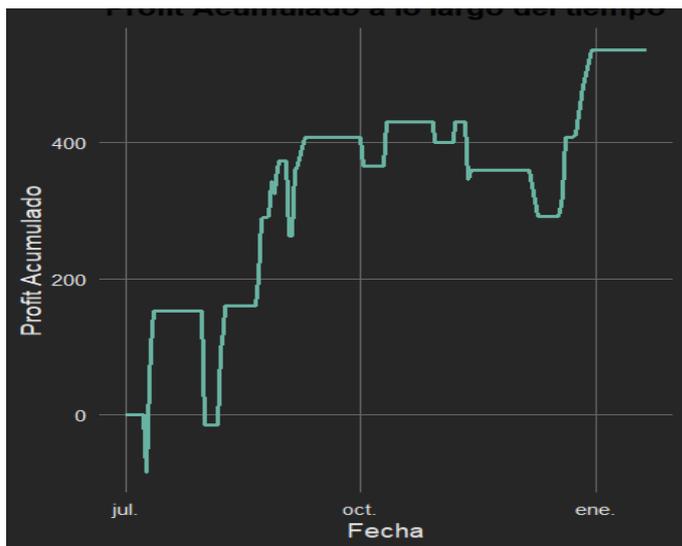


Ilustración 11 Profit acumulado del modelo en IBEX

A día 21 de enero de 2025, el profit acumulado en el modelo del IBEX es de +536.6. El modelo predijo subidas en 32 días del total de días con mercado abierto desde el 1 de julio de 2024. De esos 32 días, la predicción se cumplió en 23 días obteniendo un acierto del 71.88%.

Si bien los resultados del beneficio acumulado se alejan bastante de los del DAX se comprueba que tiene una gran efectividad sabiendo además que presenta casi la mitad de días de operativa que el anterior modelo y teniendo una tasa de acierto mayor. El resultado final es la obtención de un +5.37% de beneficio sobre el capital inicial aportado y al igual que el anterior anima probar el modelo con capital real.

3. S&P 500:

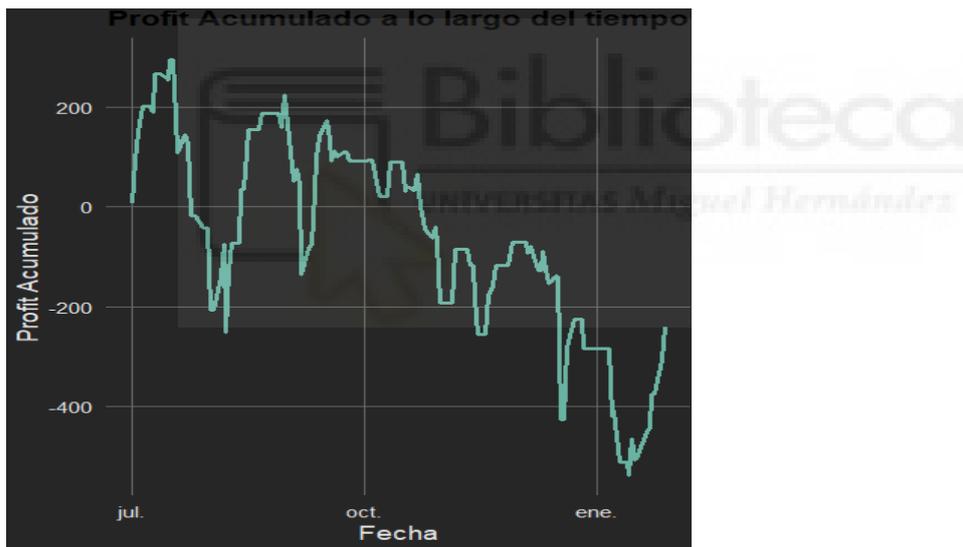


Ilustración 12 Profit acumulado del modelo en S&P 500

A día 21 de enero de 2025, el beneficio acumulado del modelo S&P500 es de -505.5.

El modelo predijo subidas en 91 días del total de días con mercado abierto desde el 1 de julio de 2024. De esos 91 días, la predicción se cumplió en 50 días teniendo un acierto del 54.95%. Destaca por otro lado que llegó a estar en un comienzo en +200 en 2 periodos diferentes y que la máxima pérdida acumulada se dio en enero de 2025 con -537.

Aunque el porcentaje de aciertos no está del todo mal, la pérdida de capital se entiende como unas bajadas mucho más fuertes en los días en los que el modelo predijo que iba a subir respecto a las subidas en los días en los que el modelo las predecía ya que si bien el modelo acierta más de la mitad de los días en los que predice subidas el beneficio final para este índice es de -5.05%, siendo incapaz de generar valor real para los inversores.

Aunque el beneficio acumulado no es realmente bajo teniendo en cuenta la volatilidad de este índice, la comparación con los otros dos modelos guía al inversor a centrarse en los otros dos índices y crea la necesidad de realizar cambios en este modelo.

Estos resultados destacan la importancia de evaluar cada modelo dentro de su contexto, indicando que, para el S&P 500, sería necesario revisar las variables empleadas, ajustar los parámetros del modelo, o explorar métodos alternativos que se adapten mejor a la naturaleza de este índice. Además, siempre hay que tener en cuenta que los contextos político económico y social es realmente influyente en los mercados de valores y en este caso durante la creación del estudio se estaban dando las elecciones presidenciales de Estados Unidos de noviembre de 2024 siendo los mercados nacionales realmente volátiles en este periodo.

Es importante destacar que estos rendimientos están condicionado al periodo analizado y podría variar en diferentes contextos de mercado.

8. Síntesis y estrategias de mejora

8.1 Conclusiones

Tras completar el modelo y evaluar los resultados se reafirma la eficacia de Random Forest como herramienta en el análisis bursátil. A pesar de no estar directamente relacionados los mercados de estos índices, se han creado 3 modelos diferentes a partir de uno inicial sin necesidad de entrar en detalle en la economía nacional que ligada a cada uno de estos demostrando los

resultados que el análisis técnico sigue siendo una metodología válida y relevante. Simplemente con una selección de variables numerosa encuentra (que realmente el propio modelo se encarga de revelar la importancia de cada una de estas) e inspirándose en diferentes modelos ya creados que abundan en internet se consigue un modelo capaz de predecir con una eficiencia respetable la variación Open to Close de los índices bursátiles que se considere ya que simplemente con cambiar el ticker de estos en el código se pueden crear centenares de modelos.

Esta facilidad que permite la tecnología no hace más que poner en valor a los analistas bursátiles que empezaron su camino sin estos avances y que fueron pioneros en un mundo que, en su momento, estaba reservado únicamente a las grandes entidades financieras.

Este proyecto contribuye a su vez a desmentir la creencia actual de que la variación del precio de los activos está únicamente ligada a las noticias económicas o políticas y que el análisis técnico queda relegado a un papel muy secundario.

Es posible que la entrada cada vez mayor en el mundo del trading de nuevos estafadores o “scammers” los cuales venden “estrategias infalibles” basadas en su análisis técnico que carece de principios y fundamentos haya contribuido a esta concepción errónea. Sin embargo, cualquier persona con una base algo sólida en análisis de datos y conceptos financieros como el autor de este trabajo puede desmentir fácilmente estas prácticas engañosas.

8.2 Propuestas de Mejora

Con el fin de mejorar el modelo creado en este proyecto se proponen las siguientes estrategias:

- Ampliar el periodo de análisis. Los resultados de los modelos para IBEX y DAX son muy buenos, para corroborar la eficiencia del modelo se deberían destear en diferentes periodos de tiempo para eliminar la idea de que simplemente se haya analizado un periodo alcista y de ahí los resultados.
- Añadir gastos a los movimientos. Se podría tener en cuenta las comisiones de entradas y salidas en la operativa para tener un escenario más realista.
- Modificar umbrales de Gini en el modelo de S&P500. En el caso del modelo aplicado a este índice se comete una redundancia ya que el umbral de la selección de variables solo lo superan 10 y después se utiliza un Mtry del mismo número. Esto puede ser uno de los causantes de los malos resultados de este.
- Selección de nuevos índices. El modelo se basa en un código ya creado y que se debe modificar mínimamente cuando se quiera probar su precisión en otros índices bursátiles-
- Extrapolar el modelo a otro tipo de activos como acciones. El modelo puede ser útil para la predicción de variaciones de precio en valores de acciones individuales, las cuales son obviamente mucho más abundantes.
- Uso de otros modelos más avanzados. XGboost o el propio Deep Learning pueden ser considerados como proyectos futuros partiendo de la base del modelo de Random Forest

En conjunto, estas propuestas buscan mejorar la efectividad y aplicabilidad del modelo, permitiendo a los inversores tomar decisiones más informadas y maximizando las oportunidades de mercado identificadas. Estas estrategias no solo fortalecerían el modelo

actual, sino que también establecerían una base sólida para futuros estudios en predicción de mercados financieros

9. Bibliografía y webs consultadas

<https://www.ibm.com/es-es/topics/random-forest>

<https://la.mathworks.com/discovery/machine-learning.html>

<https://www.grapheverywhere.com/machine-learning-que-es-tipos-ejemplos-y-como-implementarlo/>

Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest (<https://ieeexplore.ieee.org/abstract/document/7883053>)

10. Código empleado y Simulación completa

Código empleado en R Studio (en específico el referente al modelo del IBEX)

```
# Paquetes necesarios
```

```
library(quantmod)
```

```
library(randomForest)
```

```
library(caret)
```

```
library(dplyr)
```

```
library(ggplot2)
```

```
library(tidyr)
```

```
# 1. Cargar datos del IBEX desde Yahoo Finance
```

```
getSymbols("^IBEX", src = "yahoo", from = "2023-01-01", to = Sys.Date(), auto.assign = TRUE)
```

2. Calcular variables técnicas basadas en precios de apertura

```
IBEX$Rendimiento_diario <- Delt(Op(IBEX), lag(Op(IBEX))) * 100
```

```
IBEX$MA <- SMA(Op(IBEX), n = 20)
```

```
IBEX$Señal_MA <- ifelse(Op(IBEX) > IBEX$MA, 1, -1)
```

```
bb <- BBands(Op(IBEX))
```

```
IBEX$Upper_Band <- bb$up
```

```
IBEX$Lower_Band <- bb$dn
```

```
IBEX$Señal_BB <- ifelse(Op(IBEX) > bb$up, 1, ifelse(Op(IBEX) < bb$dn, -1, 0))
```

```
macd <- MACD(Op(IBEX))
```

```
IBEX$MACD <- macd$macd
```

```
IBEX$MACD_Señal <- macd$signal
```

```
IBEX$MACD_Diferencia <- macd$macd - macd$signal
```

```
IBEX$Señal_MACD <- ifelse(macd$macd > macd$signal, 1, 0)
```

```
IBEX$TRIX <- Delt(EMA(Op(IBEX), n = 15), k = 1)
```

```
IBEX$RSI <- RSI(Op(IBEX), n = 14)
```

```
IBEX$ATR <- ATR(HLC(IBEX), n = 14)$atr
```

```
IBEX$Volumen_Promedio <- rollapply(Vo(IBEX), width = 20, FUN = mean, fill = NA,  
align = "right")
```

```
IBEX$Distancia_BB_Upper <- (Op(IBEX) - IBEX$Upper_Band) / Op(IBEX)
```

```
IBEX$Distancia_BB_Lower <- (Op(IBEX) - IBEX$Lower_Band) / Op(IBEX)
```

```

IBEX$ATR_Ratio <- IBEX$ATR / Op(IBEX)

IBEX$EMA <- EMA(Op(IBEX), n = 20)

IBEX$Rendimiento_OC <- Delt(Op(IBEX), Cl(IBEX)) * 100
IBEX$Resultado <- ifelse(IBEX$Rendimiento_OC > 0, 1, 0)

# Filtrar datos sin NA en Señal_BB
valores_no_na <- which(!is.na(IBEX$Señal_BB))

par(bg = "white")

# Gráfico de Bandas de Bollinger y Precio de Apertura
plot(index(IBEX[valores_no_na]), Op(IBEX[valores_no_na]), type = "l", col = "black",
lwd = 1.5,
      ylab = "Precio", xlab = "Fecha", main = "Bandas de Bollinger y Precio de Apertura")

# Añadir Bandas de Bollinger
lines(index(IBEX[valores_no_na]), IBEX$Upper_Band[valores_no_na], col = "blue",
lty = 2, lwd = 1.2)

lines(index(IBEX[valores_no_na]), IBEX$Lower_Band[valores_no_na], col = "red", lty
= 2, lwd = 1.2)

# Añadir Precio de Apertura
lines(index(IBEX[valores_no_na]), Op(IBEX[valores_no_na]), col = "black", lwd = 1.5)

```

```
# Añadir Señales de Compra (Sobreventa, banda inferior)
```

```
points(index(IBEX[valores_no_na][IBEX$Señal_BB[valores_no_na] == -1]),
```

```
  Op(IBEX[valores_no_na][IBEX$Señal_BB[valores_no_na] == -1]),
```

```
  col = "green", pch = 16, cex = 1.3)
```

```
# Añadir Señales de Venta (Sobrecompra, banda superior)
```

```
points(index(IBEX[valores_no_na][IBEX$Señal_BB[valores_no_na] == 1]),
```

```
  Op(IBEX[valores_no_na][IBEX$Señal_BB[valores_no_na] == 1]),
```

```
  col = "red", pch = 17, cex = 1.3)
```

```
# Añadir leyenda
```

```
legend("topright", legend = c("Precio Apertura", "Banda Superior", "Banda Inferior",  
"Sobreventa (Compra)", "Sobrecompra (Venta)"),
```

```
  col = c("black", "blue", "red", "green", "red"), lty = c(1, 2, 2, NA, NA),
```

```
  pch = c(NA, NA, NA, 16, 17), lwd = c(1.5, 1.2, 1.2, NA, NA),
```

```
  bg = "white", box.lwd = 1.2)
```

```
# 3. Crear dataset y eliminar NAs
```

```
IBEX_alineado <- na.omit(merge(
```

```
  IBEX$Rendimiento_diario, IBEX$MA, IBEX$Señal_MA, IBEX$Upper_Band,  
  IBEX$Lower_Band,
```

```

IBEX$Señal_BB, IBEX$MACD, IBEX$MACD_Señal, IBEX$MACD_Diferencia,
IBEX$Señal_MACD,

IBEX$TRIX,      IBEX$RSI,      IBEX$ATR,      IBEX$Volumen_Promedio,
IBEX$Distancia_BB_Upper,

IBEX$Distancia_BB_Lower, IBEX$ATR_Ratio, IBEX$EMA, IBEX$Resultado
))

```

4. Dividir los datos

```
trainData <- IBEX_alineado["2023-06/2024-06"]
```

```
testData <- IBEX_alineado["2024-07/"]
```

```
trainData <- as.data.frame(trainData)
```

```
testData <- as.data.frame(testData)
```

```
trainData$Resultado <- as.factor(trainData$Resultado)
```

```
testData$Resultado <- as.factor(testData$Resultado)
```

5. Entrenar modelo Random Forest

```
set.seed(123)
```

```
modelo_rf <- randomForest(Resultado ~ ., data = trainData, ntree = 500, mtry = 10,
importance = TRUE)
```

```
importancia_variables <- importance(modelo_rf)
```

```
importancia_df <- as.data.frame(importancia_variables)
```

```
importancia_df$Variable <- rownames(importancia_df)
```

```
# 6. Gráfico de importancia de variables
```

```
umbral <- 7.5
```

```
p1 <- ggplot(importancia_df, aes(x = reorder(Variable, MeanDecreaseGini), y =  
MeanDecreaseGini)) +
```

```
  geom_bar(stat = "identity", fill = "steelblue") +
```

```
  geom_hline(yintercept = umbral, color = "red", linetype = "dashed") +
```

```
  labs(title = "Importancia de Variables (MeanDecreaseGini)", x = "Variables", y = "Mean  
Decrease Gini") +
```

```
  coord_flip() +
```

```
  theme_minimal()
```

```
print(p1)
```



```
# 7. Selección de variables
```

```
variables_seleccionadas <- rownames(importancia_variables[importancia_variables[,  
"MeanDecreaseGini"] > umbral, ])
```

```
trainData_filtrado <- trainData[, c(variables_seleccionadas, "Resultado")]
```

```
testData_filtrado <- testData[, c(variables_seleccionadas, "Resultado")]
```

```
# Mensaje sobre variables seleccionadas
```

```
cat("Se han seleccionado las variables cuya Mean Decrease Gini es mayor a", umbral,  
".\n")
```

```
cat("Las variables seleccionadas han sido:\n")
```

```
cat(paste(variables_seleccionadas, collapse = ", "), "\n")
```

8. Reentrenar el modelo con variables seleccionadas

```
set.seed(123)
```

```
modelo_rf_filtrado <- randomForest(Resultado ~ ., data = trainData_filtrado, ntree = 160,  
mtry = 5, importance = TRUE)
```

```
predicciones <- predict(modelo_rf_filtrado, testData_filtrado)
```

Calcular matriz de confusión y métricas

```
conf_matrix <- confusionMatrix(predicciones, testData_filtrado$Resultado)
```

Mostrar métricas importantes

```
cat("\n--- Métricas de evaluación del modelo ---\n")
```

```
cat("Accuracy:", round(conf_matrix$overall['Accuracy'] * 100, 2), "%\n")
```

```
cat("Sensitivity:", round(conf_matrix$byClass['Sensitivity'] * 100, 2), "%\n")
```

```
cat("Specificity:", round(conf_matrix$byClass['Specificity'] * 100, 2), "%\n")
```

```
cat("Kappa:", round(conf_matrix$overall['Kappa'], 2), "\n")
```

```
cat("-----\n\n")
```

9. Simulación de inversión

```
rendimiento_oc_real <- IBEX$Rendimiento_OC["2024-07/"]
```

```
rendimiento_oc_real <- head(rendimiento_oc_real, length(predicciones))
```

```
simulacion <- data.frame(  

```

```
Fecha = as.Date(row.names(testData_filtrado)),
```

```

Prediccion_Modelo = as.numeric(as.character(predicciones)),
Realidad = as.numeric(as.character(testData_filtrado$Resultado)),
Rendimiento_OC = rendimiento_oc_real,
Capital = 10000
)

for (i in 1:nrow(simulacion)) {
  if (i == 1) {
    simulacion$Profit[i] <- ifelse(simulacion$Prediccion_Modelo[i] == 1,
      simulacion$Capital[i] * simulacion$Rendimiento_OC[i] / 100,
      0)
  } else {
    simulacion$Capital[i] <- simulacion$Capital[i - 1] + simulacion$Profit[i - 1]
    simulacion$Profit[i] <- ifelse(simulacion$Prediccion_Modelo[i] == 1,
      simulacion$Capital[i] * simulacion$Rendimiento_OC[i] / 100,
      0)
  }
}

# Calcular Profit Acumulado
simulacion <- simulacion %>%
  mutate(Profit_Acumulado = cumsum(Profit))

# Redondear valores a 2 decimales

```

```

simulacion <- simulacion %>%
  mutate(
    Capital = round(Capital, 2),
    Profit = round(Profit, 2),
    Profit_Acumulado = round(Profit_Acumulado, 2)
  )

```

```

simulacion <- simulacion %>%
  mutate(Profit_Acumulado = cumsum(Profit))

```

```

print(simulacion)

```



```

# 10. Calcular resumen de predicciones y aciertos

```

```

predicciones_subida <- simulacion[simulacion$Prediccion_Modelo == 1, ]
aciertos_subida <- sum(predicciones_subida$Prediccion_Modelo ==
predicciones_subida$Realidad)

```

```

total_predicciones_subida <- nrow(predicciones_subida)

```

```

if (total_predicciones_subida > 0) {
  porcentaje_acierto_subida <- (aciertos_subida / total_predicciones_subida) * 100
} else {
  porcentaje_acierto_subida <- 0
}

```

```

# Mostrar el resumen

```

```

cat("El modelo predijo subidas en", total_predicciones_subida, "días.\n")

cat("Se cumplió en", aciertos_subida, "días.\n")

cat("Acierto del", round(porcentaje_acierto_subida, 2), "% cuando el modelo predice
subida.\n")

# Gráfico de Profit Acumulado con estilo oscuro

p2 <- ggplot(simulacion, aes(x = Fecha, y = Profit_Acumulado)) +

  geom_line(color = "#69b3a2", size = 1.2) + # Línea de profit acumulado

  labs(

    title = "Profit Acumulado a lo largo del tiempo",

    x = "Fecha",

    y = "Profit Acumulado"

  ) +

  theme_minimal(base_size = 14) +

  theme(

    panel.grid.major = element_line(color = "gray40"),

    panel.grid.minor = element_blank(),

    plot.title = element_text(hjust = 0.5, face = "bold", size = 16),

    axis.text = element_text(color = "gray90"),

    axis.title = element_text(color = "gray90"),

    panel.background = element_rect(fill = "gray15", color = "gray15"),

    plot.background = element_rect(fill = "gray15", color = "gray15")

  )

# Mostrar el gráfico

print(p2)

```

Simulación completa del DAX

Fecha	Prediccion_Modelo	Realidad	Capital	Profit	Profit_Acumulado
01/07/2024	1	1	10000	29,78	29,78
02/07/2024	1	0	10029.78	-67,35	-37,57
03/07/2024	1	1	9962.43	105,65	68,08
04/07/2024	0	1	10068.08	0	68,08
05/07/2024	0	1	10068.08	0	68,08
08/07/2024	1	0	10068.08	-1,21	66,87
09/07/2024	1	0	10066.87	-128,16	-61,29
10/07/2024	1	1	9938.71	92,9	31,61
11/07/2024	0	1	10031.61	0	31,61
12/07/2024	1	1	10031.61	115,66	147,27
15/07/2024	1	0	10147.27	-78,65	68,62
16/07/2024	1	0	10068.62	-37,05	31,57
17/07/2024	0	0	10031.57	0	31,57
18/07/2024	1	0	10031.57	-46,01	-14,44
19/07/2024	1	0	9985.57	-91,97	-106,41
22/07/2024	1	1	9893.6	73,76	-32,65
23/07/2024	0	1	9967.36	0	-32,65
24/07/2024	0	0	9967.36	0	-32,65
25/07/2024	1	1	9967.36	48,89	16,24
26/07/2024	1	1	10016.25	89	105,24
29/07/2024	0	0	10105.25	0	105,24
30/07/2024	1	1	10105.25	26	131,24
31/07/2024	0	1	10131.25	0	131,24
01/08/2024	0	0	10131.25	0	131,24
02/08/2024	1	0	10131.25	-126,68	4,56
05/08/2024	1	0	10004.56	-106,52	-101,96
06/08/2024	1	0	9898.05	-38,77	-140,73
07/08/2024	1	1	9859.27	99,79	-40,94
08/08/2024	1	1	9959.06	93,44	52,5
09/08/2024	1	1	10052.51	28,43	80,93
12/08/2024	1	0	10080.94	-32,63	48,3
13/08/2024	1	1	10048.31	6,99	55,29
14/08/2024	0	1	10055.3	0	55,29
15/08/2024	1	1	10055.3	114,9	170,19
16/08/2024	0	1	10170.2	0	170,19
19/08/2024	0	1	10170.2	0	170,19
20/08/2024	0	0	10170.2	0	170,19
21/08/2024	1	1	10170.2	49,95	220,14
22/08/2024	0	1	10220.15	0	220,14
23/08/2024	1	1	10220.15	51,95	272,09

26/08/2024	1	1	10272.1	14,99	287,08
27/08/2024	0	1	10287.08	0	287,08
28/08/2024	0	1	10287.08	0	287,08
29/08/2024	1	1	10287.08	66,46	353,54
30/08/2024	0	1	10353.54	0	353,54
02/09/2024	1	1	10353.54	9,48	363,02
03/09/2024	0	0	10363.02	0	363,02
04/09/2024	1	1	10363.02	33,48	396,5
05/09/2024	1	1	10396.5	20,95	417,45
06/09/2024	1	0	10417.45	-130,94	286,51
09/09/2024	1	1	10286.51	35,84	322,35
10/09/2024	1	0	10322.35	-84,74	237,61
11/09/2024	1	0	10237.61	-6,87	230,74
12/09/2024	0	0	10230.75	0	230,74
13/09/2024	1	1	10230.75	84,5	315,24
16/09/2024	0	1	10315.25	0	315,24
17/09/2024	0	1	10315.25	0	315,24
18/09/2024	0	0	10315.25	0	315,24
19/09/2024	0	1	10315.25	0	315,24
20/09/2024	1	0	10315.25	-95,98	219,26
23/09/2024	1	1	10219.27	43,49	262,75
24/09/2024	0	1	10262.76	0	262,75
25/09/2024	1	1	10262.76	39,52	302,27
26/09/2024	0	1	10302.29	0	302,27
27/09/2024	0	1	10302.29	0	302,27
30/09/2024	0	0	10302.29	0	302,27
01/10/2024	0	0	10302.29	0	302,27
02/10/2024	0	0	10302.29	0	302,27
03/10/2024	0	0	10302.29	0	302,27
04/10/2024	1	1	10302.29	82,6	384,87
07/10/2024	0	0	10384.88	0	384,87
08/10/2024	1	1	10384.88	70,61	455,48
09/10/2024	0	1	10455.5	0	455,48
10/10/2024	0	0	10455.5	0	455,48
11/10/2024	0	1	10455.5	0	455,48
14/10/2024	0	1	10455.5	0	455,48
15/10/2024	0	0	10455.5	0	455,48
16/10/2024	0	0	10455.5	0	455,48
17/10/2024	0	1	10455.5	0	455,48
18/10/2024	0	1	10455.5	0	455,48
21/10/2024	0	0	10455.5	0	455,48
22/10/2024	0	0	10455.5	0	455,48
23/10/2024	0	0	10455.5	0	455,48
24/10/2024	0	1	10455.5	0	455,48
25/10/2024	0	1	10455.5	0	455,48

28/10/2024	0	1	10455.5	0	455,48
29/10/2024	0	0	10455.5	0	455,48
30/10/2024	1	0	10455.5	-78,72	376,76
31/10/2024	1	0	10376.78	-10,4	366,36
01/11/2024	1	1	10366.38	87,4	453,76
04/11/2024	0	0	10453.78	0	453,76
05/11/2024	1	1	10453.78	46,96	500,72
06/11/2024	0	0	10500.74	0	500,72
07/11/2024	1	1	10500.74	112,53	613,25
08/11/2024	0	0	10613.27	0	613,25
11/11/2024	0	1	10613.27	0	613,25
12/11/2024	0	0	10613.27	0	613,25
13/11/2024	1	0	10613.27	-11,81	601,44
14/11/2024	1	1	10601.46	106,45	707,89
15/11/2024	0	1	10707.91	0	707,89
18/11/2024	0	0	10707.91	0	707,89
19/11/2024	0	0	10707.91	0	707,89
20/11/2024	0	0	10707.91	0	707,89
21/11/2024	1	1	10707.91	47,89	755,78
22/11/2024	0	1	10755.8	0	755,78
25/11/2024	0	0	10755.8	0	755,78
26/11/2024	1	0	10755.8	-7,34	748,44
27/11/2024	1	1	10748.46	2,22	750,66
28/11/2024	0	1	10750.68	0	750,66
29/11/2024	1	1	10750.68	125,28	875,94
02/12/2024	1	1	10875.96	192,93	1068,87
03/12/2024	1	1	11068.89	38,61	1107,48
04/12/2024	0	1	11107.5	0	1107,48
05/12/2024	0	1	11107.5	0	1107,48
06/12/2024	0	1	11107.5	0	1107,48
09/12/2024	0	0	11107.5	0	1107,48
10/12/2024	0	1	11107.5	0	1107,48
11/12/2024	0	1	11107.5	0	1107,48
12/12/2024	0	0	11107.5	0	1107,48
13/12/2024	0	0	11107.5	0	1107,48
16/12/2024	0	0	11107.5	0	1107,48
17/12/2024	0	0	11107.5	0	1107,48
18/12/2024	0	0	11107.5	0	1107,48
19/12/2024	1	0	11107.5	-34,48	1073
20/12/2024	1	1	11073.01	34,57	1107,57
23/12/2024	1	1	11107.58	0,45	1108,02
27/12/2024	1	1	11108.03	85,5	1193,52
30/12/2024	1	1	11193.52	9,1	1202,62
02/01/2025	1	1	11202.62	57,12	1259,74
03/01/2025	0	0	11259.75	0	1259,74

06/01/2025	1	1	11259.75	131,59	1391,33
07/01/2025	0	1	11391.33	0	1391,33
08/01/2025	0	1	11391.33	0	1391,33
09/01/2025	0	1	11391.33	0	1391,33
10/01/2025	0	0	11391.33	0	1391,33
13/01/2025	0	0	11391.33	0	1391,33
14/01/2025	0	1	11391.33	0	1391,33
15/01/2025	0	1	11391.33	0	1391,33
16/01/2025	0	0	11391.33	0	1391,33
17/01/2025	1	1	11391.33	94,15	1485,48
20/01/2025	0	1	11485.48	0	1485,48
21/01/2025	0	1	11485.48	0	1485,48

Simulación completa del IBEX

Fecha	Prediccion_Modelo	Realidad	Capital	Profit	Profit_Acumulado
01/07/2024	0	0	10000	0	0
02/07/2024	0	0	10000	0	0
03/07/2024	0	1	10000	0	0
04/07/2024	0	1	10000	0	0
05/07/2024	0	0	10000	0	0
08/07/2024	0	1	10000	0	0
09/07/2024	1	0	10000	-83,8	-83,8
10/07/2024	1	1	9916.2	132.23	48.43
11/07/2024	1	1	10048.44	49.81	98.24
12/07/2024	1	1	10098.25	54.69	152.93
15/07/2024	0	0	10152.94	0	152.93
16/07/2024	0	1	10152.94	0	152.93
17/07/2024	0	1	10152.94	0	152.93
18/07/2024	0	1	10152.94	0	152.93
19/07/2024	0	0	10152.94	0	152.93
22/07/2024	0	0	10152.94	0	152.93
23/07/2024	0	1	10152.94	0	152.93
24/07/2024	0	1	10152.94	0	152.93
25/07/2024	0	1	10152.94	0	152.93
26/07/2024	0	1	10152.94	0	152.93
29/07/2024	0	0	10152.94	0	152.93
30/07/2024	0	1	10152.94	0	152.93
31/07/2024	0	0	10152.94	0	152.93
01/08/2024	1	0	10152.94	-167,25	-14,32

02/08/2024	0	0	9985.69	0	-14,32
05/08/2024	0	1	9985.69	0	-14,32
06/08/2024	0	0	9985.69	0	-14,32
07/08/2024	1	1	9985.69	109.63	95.31
08/08/2024	1	1	10095.32	28.67	123.98
09/08/2024	1	1	10123.99	35.62	159.6
12/08/2024	0	0	10159.61	0	159.6
13/08/2024	0	1	10159.61	0	159.6
14/08/2024	0	0	10159.61	0	159.6
15/08/2024	0	1	10159.61	0	159.6
16/08/2024	0	1	10159.61	0	159.6
19/08/2024	0	1	10159.61	0	159.6
20/08/2024	0	0	10159.61	0	159.6
21/08/2024	0	1	10159.61	0	159.6
22/08/2024	1	1	10159.61	41.97	201.57
23/08/2024	1	1	10201.58	86.75	288.32
26/08/2024	1	1	10288.33	3.93	292.25
27/08/2024	1	1	10292.26	50.13	342.38
28/08/2024	1	0	10342.39	-17,49	324.89
29/08/2024	1	1	10324.9	22.04	346.93
30/08/2024	1	1	10346.94	26.57	373.5
02/09/2024	1	0	10373.51	0	373.5
03/09/2024	1	0	10373.51	-110,91	262.59
04/09/2024	0	1	10262.59	0	262.59
05/09/2024	1	1	10262.59	98.06	360.65
06/09/2024	0	0	10360.65	0	360.65
09/09/2024	1	1	10360.65	46.81	407.46
10/09/2024	0	0	10407.46	0	407.46
11/09/2024	0	1	10407.46	0	407.46
12/09/2024	0	1	10407.46	0	407.46
13/09/2024	0	1	10407.46	0	407.46
16/09/2024	0	1	10407.46	0	407.46
17/09/2024	0	1	10407.46	0	407.46
18/09/2024	0	0	10407.46	0	407.46
19/09/2024	0	0	10407.46	0	407.46
20/09/2024	0	1	10407.46	0	407.46
23/09/2024	0	1	10407.46	0	407.46
24/09/2024	0	0	10407.46	0	407.46
25/09/2024	0	1	10407.46	0	407.46
26/09/2024	0	1	10407.46	0	407.46
27/09/2024	0	1	10407.46	0	407.46
30/09/2024	0	0	10407.46	0	407.46
01/10/2024	0	0	10407.46	0	407.46
02/10/2024	1	0	10407.46	-41,7	365.76
03/10/2024	0	1	10365.77	0	365.76

04/10/2024	0	1	10365.77	0	365.76
07/10/2024	0	1	10365.77	0	365.76
08/10/2024	0	1	10365.77	0	365.76
09/10/2024	0	1	10365.77	0	365.76
10/10/2024	0	0	10365.77	0	365.76
11/10/2024	1	1	10365.77	64.43	430.19
14/10/2024	0	1	10430.2	0	430.19
15/10/2024	0	1	10430.2	0	430.19
16/10/2024	0	1	10430.2	0	430.19
17/10/2024	0	0	10430.2	0	430.19
18/10/2024	0	1	10430.2	0	430.19
21/10/2024	0	0	10430.2	0	430.19
22/10/2024	0	0	10430.2	0	430.19
23/10/2024	0	1	10430.2	0	430.19
24/10/2024	0	0	10430.2	0	430.19
25/10/2024	0	0	10430.2	0	430.19
28/10/2024	0	1	10430.2	0	430.19
29/10/2024	0	0	10430.2	0	430.19
30/10/2024	1	0	10430.2	-29,74	400.45
31/10/2024	0	1	10400.46	0	400.45
01/11/2024	0	1	10400.46	0	400.45
04/11/2024	0	0	10400.46	0	400.45
05/11/2024	0	1	10400.46	0	400.45
06/11/2024	0	0	10400.46	0	400.45
07/11/2024	1	1	10400.46	29.57	430.02
08/11/2024	0	0	10430.03	0	430.02
11/11/2024	0	0	10430.03	0	430.02
12/11/2024	1	0	10430.03	-84,85	345.67
13/11/2024	1	1	10345.68	14.48	360.15
14/11/2024	0	1	10360.16	0	360.15
15/11/2024	0	1	10360.16	0	360.15
18/11/2024	0	1	10360.16	0	360.15
19/11/2024	0	0	10360.16	0	360.15
20/11/2024	0	0	10360.16	0	360.15
21/11/2024	0	1	10360.16	0	360.15
22/11/2024	0	1	10360.16	0	360.15
25/11/2024	0	0	10360.16	0	360.15
26/11/2024	0	0	10360.16	0	360.15
27/11/2024	0	0	10360.16	0	360.15
28/11/2024	0	0	10360.16	0	360.15
29/11/2024	0	1	10360.16	0	360.15
02/12/2024	0	1	10360.16	0	360.15
03/12/2024	0	1	10360.16	0	360.15
04/12/2024	0	1	10360.16	0	360.15
05/12/2024	0	1	10360.16	0	360.15

06/12/2024	0	0	10360.16	0	360.15
09/12/2024	1	0	10360.16	-68,46	291.69
10/12/2024	0	0	10291.7	0	291.69
11/12/2024	0	0	10291.7	0	291.69
12/12/2024	0	0	10291.7	0	291.69
13/12/2024	0	0	10291.7	0	291.69
16/12/2024	0	1	10291.7	0	291.69
17/12/2024	0	0	10291.7	0	291.69
18/12/2024	1	1	10291.7	16.41	308.1
19/12/2024	1	1	10308.11	15.25	323.35
20/12/2024	1	1	10323.36	84.96	408.31
23/12/2024	0	0	10408.32	0	408.31
24/12/2024	1	1	10408.32	3.54	411.85
27/12/2024	1	1	10411.86	71.45	483.3
30/12/2024	1	1	10483.31	53.34	536.64
31/12/2024	0	1	10536.65	0	536.64
02/01/2025	0	1	10536.65	0	536.64
03/01/2025	0	0	10536.65	0	536.64
06/01/2025	0	1	10536.65	0	536.64
07/01/2025	0	1	10536.65	0	536.64
08/01/2025	0	0	10536.65	0	536.64
09/01/2025	0	1	10536.65	0	536.64
10/01/2025	0	0	10536.65	0	536.64
13/01/2025	0	1	10536.65	0	536.64
14/01/2025	0	1	10536.65	0	536.64
15/01/2025	0	1	10536.65	0	536.64
16/01/2025	0	0	10536.65	0	536.64
17/01/2025	0	1	10536.65	0	536.64
20/01/2025	0	1	10536.65	0	536.64
21/01/2025	0	1	10536.65	0	536.64

Simulación completa del S&P500

Fecha	Prediccion_Modelo	Realidad	Capital	Profit	Profit_Acumulado
01/07/2024	1	1	10000	7,33	7,33
02/07/2024	1	1	10007.33	86,43	93,76
03/07/2024	1	1	10093.76	54,21	147,97
05/07/2024	1	1	10147.97	53,65	201,62
08/07/2024	1	1	10201.62	0,18	201,8
09/07/2024	1	0	10201.8	-13,26	188,54
10/07/2024	1	1	10188.54	77,72	266,26
11/07/2024	0	0	10266.26	0	266,26
12/07/2024	0	1	10266.26	0	266,26
15/07/2024	1	0	10266.26	-12,64	253,62
16/07/2024	1	1	10253.62	41,98	295,6
17/07/2024	0	0	10295.61	0	295,6
18/07/2024	1	0	10295.61	-117,43	178,17
19/07/2024	1	0	10178.18	-70,45	107,72
22/07/2024	1	1	10107.73	36,22	143,94
23/07/2024	1	0	10143.95	-17,42	126,52
24/07/2024	1	0	10126.53	-144,77	-18,25
25/07/2024	0	0	9981.76	0	-18,25
26/07/2024	0	1	9981.76	0	-18,25
29/07/2024	1	0	9981.76	-23,71	-41,96
30/07/2024	0	0	9958.05	0	-41,96
31/07/2024	0	1	9958.05	0	-41,96
01/08/2024	1	0	9958.05	-163,92	-205,88
02/08/2024	0	0	9794.13	0	-205,88

05/08/2024	1	1	9794.13	66,91	-138,97
06/08/2024	1	1	9861.03	63,66	-75,31
07/08/2024	1	0	9924.69	-175,56	-250,87
08/08/2024	1	1	9749.13	123,87	-127
09/08/2024	1	1	9873.01	54,8	-72,2
12/08/2024	0	0	9927.81	0	-72,2
13/08/2024	1	1	9927.81	106,07	33,87
14/08/2024	0	1	10033.88	0	33,87
15/08/2024	1	1	10033.88	76,77	110,64
16/08/2024	1	1	10110.66	43,42	154,06
19/08/2024	0	1	10154.07	0	154,06
20/08/2024	0	0	10154.07	0	154,06
21/08/2024	1	1	10154.07	32,19	186,25
22/08/2024	0	0	10186.26	0	186,25
23/08/2024	0	1	10186.26	0	186,25
26/08/2024	0	0	10186.26	0	186,25
27/08/2024	0	1	10186.26	0	186,25
28/08/2024	0	0	10186.26	0	186,25
29/08/2024	1	0	10186.26	-27,87	158,38
30/08/2024	1	1	10158.39	64,54	222,92
03/09/2024	1	0	10222.93	-172,62	50,3
04/09/2024	1	1	10050.32	24,44	74,74
05/09/2024	1	0	10074.76	-30,42	44,32
06/09/2024	1	0	10044.33	-180,39	-136,07
09/09/2024	1	1	9863.94	52,53	-83,54
10/09/2024	1	1	9916.47	9,05	-74,49
11/09/2024	1	1	9925.51	104,21	29,72
12/09/2024	1	1	10029.73	69,08	98,8
13/09/2024	1	1	10098.81	40,88	139,68
16/09/2024	1	1	10139.69	32,29	171,97
17/09/2024	1	0	10171.98	-37,64	134,33
18/09/2024	1	0	10134.33	-42,07	92,26
19/09/2024	1	1	10092.26	19,49	111,75
20/09/2024	1	0	10111.75	-12,56	99,19
23/09/2024	1	1	10099.19	11,79	110,98
24/09/2024	0	1	10110.98	0	110,98
25/09/2024	1	0	10110.98	-20,09	90,89
26/09/2024	0	0	10090.9	0	90,89
27/09/2024	0	0	10090.9	0	90,89
30/09/2024	0	1	10090.9	0	90,89
01/10/2024	0	0	10090.9	0	90,89
02/10/2024	0	1	10090.9	0	90,89
03/10/2024	1	1	10090.9	3,1	93,99
04/10/2024	0	1	10094	0	93,99
07/10/2024	1	0	10094	-73,64	20,35

08/10/2024	0	1	10020.35	0	20,35
09/10/2024	0	1	10020.35	0	20,35
10/10/2024	0	1	10020.35	0	20,35
11/10/2024	1	1	10020.35	69,3	89,65
14/10/2024	0	1	10089.65	0	89,65
15/10/2024	0	0	10089.65	0	89,65
16/10/2024	0	1	10089.65	0	89,65
17/10/2024	1	0	10089.65	-58,64	31,01
18/10/2024	1	1	10031.01	8,97	39,98
21/10/2024	1	0	10039.98	-6,58	33,4
22/10/2024	1	1	10033.4	31,82	65,22
23/10/2024	1	0	10065.22	-63,97	1,25
24/10/2024	1	0	10001.26	-13,65	-12,4
25/10/2024	1	0	9987.61	-31,93	-44,33
28/10/2024	1	0	9955.67	-17,77	-62,1
29/10/2024	1	1	9937.91	22,61	-39,49
30/10/2024	1	0	9960.52	-32,41	-71,9
31/10/2024	1	0	9928.11	-120,14	-192,04
01/11/2024	0	1	9807.96	0	-192,04
04/11/2024	0	0	9807.96	0	-192,04
05/11/2024	0	1	9807.96	0	-192,04
06/11/2024	1	1	9807.96	107,28	-84,76
07/11/2024	0	1	9915.24	0	-84,76
08/11/2024	0	1	9915.24	0	-84,76
11/11/2024	0	0	9915.24	0	-84,76
12/11/2024	1	0	9915.24	-32,39	-117,15
13/11/2024	1	0	9882.85	-0,61	-117,76
14/11/2024	1	0	9882.24	-66,84	-184,6
15/11/2024	1	0	9815.41	-70	-254,6
18/11/2024	0	1	9745.4	0	-254,6
19/11/2024	1	1	9745.4	77,91	-176,69
20/11/2024	1	1	9823.32	4,6	-172,09
21/11/2024	1	1	9827.92	13,45	-158,64
22/11/2024	1	1	9841.37	41,36	-117,28
25/11/2024	0	0	9882.72	0	-117,28
26/11/2024	0	1	9882.72	0	-117,28
27/11/2024	0	0	9882.72	0	-117,28
29/11/2024	1	1	9882.72	46,75	-70,53
02/12/2024	0	1	9929.47	0	-70,53
03/12/2024	0	1	9929.47	0	-70,53
04/12/2024	0	1	9929.47	0	-70,53
05/12/2024	1	0	9929.47	-22,7	-93,23
06/12/2024	1	1	9906.77	14,48	-78,75
09/12/2024	1	0	9921.25	-49,19	-127,94
10/12/2024	0	0	9872.06	0	-127,94

11/12/2024	1	1	9872.06	39,16	-88,78
12/12/2024	1	0	9911.22	-37,59	-126,37
13/12/2024	1	0	9873.63	-27,79	-154,16
16/12/2024	1	1	9845.84	16,71	-137,45
17/12/2024	1	0	9862.55	-3,16	-140,61
18/12/2024	1	0	9859.39	-286,1	-426,71
19/12/2024	0	0	9573.29	0	-426,71
20/12/2024	1	1	9573.29	145,6	-281,11
23/12/2024	1	1	9718.89	55,33	-225,78
24/12/2024	0	1	9774.22	0	-225,78
26/12/2024	0	1	9774.22	0	-225,78
27/12/2024	1	0	9774.22	-57,49	-283,27
30/12/2024	0	0	9716.72	0	-283,27
31/12/2024	0	0	9716.72	0	-283,27
02/01/2025	0	0	9716.72	0	-283,27
03/01/2025	0	1	9716.72	0	-283,27
06/01/2025	0	0	9716.72	0	-283,27
07/01/2025	1	0	9716.72	-136,56	-419,83
08/01/2025	1	1	9580.16	12,3	-407,53
10/01/2025	1	0	9592.47	-103,1	-510,63
13/01/2025	0	1	9489.37	0	-510,63
14/01/2025	1	0	9489.37	-26,5	-537,13
15/01/2025	1	1	9462.87	71,63	-465,5
16/01/2025	1	0	9534.5	-42	-507,5
17/01/2025	1	1	9492.5	2	-505,5
21/01/2025	1	1	9494.5	55,44	-450,06