

Rentabilidad y comportamiento del SPY: un estudio de inversión a largo plazo

Resumen

El siguiente estudio tiene como objetivo analizar la evolución del índice bursátil SPY (conocido como el ETF del S&P 500) a través de distintas visualizaciones. Dicho índice es considerado uno de los principales referentes del mercado financiero estadounidense. Se busca analizar y comprender su comportamiento en el tiempo y detectar patrones relevantes para así evaluar su posible previsibilidad. Además, se incluye una comparativa con el IBEX 35 como índice europeo de referencia.

Introducción

El S&P 500 es un índice bursátil compuesto por las 500 empresas más importantes de Estados Unidos, seleccionadas por su capitalización bursátil. Cuando una empresa pierde valor y es superada por otra que no está en el índice, es sustituida, lo que garantiza que el S&P 500 esté compuesto siempre por las compañías más representativas del mercado. Esta rotación natural contribuye a que el índice tenga una menor exposición al riesgo en comparación con otros activos. En el año 2000, el conjunto de estas 500 empresas representaba un valor de mercado de 11,7 billones de dólares, lo que equivalía al 68,7 % del total de las acciones cotizadas en Estados Unidos (Blume & Eden, 2004).

En este estudio se analiza el comportamiento del SPDR S&P 500 ETF (SPY), un fondo cotizado (*ETF*) que replica el rendimiento del índice S&P 500, permitiendo a los inversores comprar y vender participaciones que representan a todo el índice de forma sencilla.

La base de datos utilizada contiene la siguiente información para cada jornada bursátil y explicada por Choo, 2025:

- **Fecha:** Día, mes y año de la sesión.
- **Open:** Precio de apertura del índice.
- **High:** Precio máximo alcanzado durante la jornada.
- **Low:** Precio mínimo.
- **Close:** Precio de cierre.
- **Volumen:** Número de acciones intercambiadas durante la sesión.

Estos datos permiten representar el comportamiento diario del índice mediante gráficos de velas japonesas, donde cada vela muestra si la jornada fue alcista o bajista, con su cuerpo y mechas correspondientes.

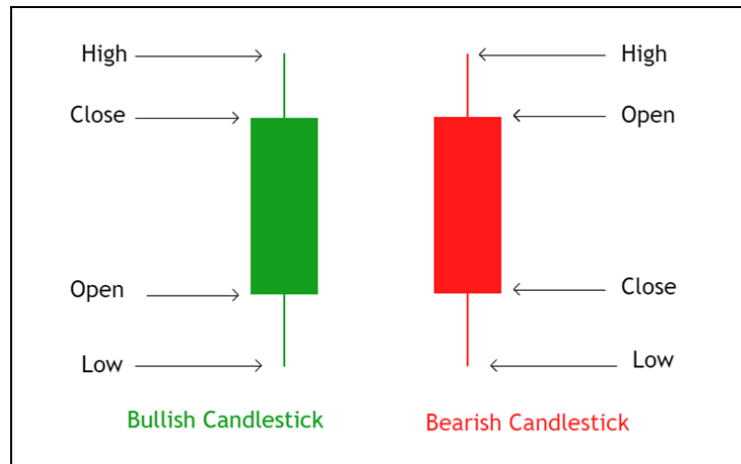


Figura 1: Candlestick Chart. Sbi securities

En la Figura 1 se puede apreciar donde pertenecen los valores mencionados previamente.

A través del análisis visual y estadístico de estos valores, el objetivo de este trabajo es estudiar la evolución y rentabilidad del SPY en distintos periodos temporales, así como explorar patrones que puedan tener valor para la toma de decisiones de inversión. Además se intenta prever qué valor tendrá en el futuro con simulaciones.

Metodologia

El estudio se ha realizado a partir de dos conjuntos de datos encontrados en *Kaggle*:

- **Spy 500:** Contiene 8150 registros y va desde 1993 hasta 2025.
- **Ibex 35:** Contiene 6826 registros y va desde 1994 hasta 2020.

Los dos conjuntos de datos contienen el día de la jornada bursátil con sus precios correspondientes a dicha fecha.

Primero, se han importado los conjuntos de datos en MySQL. Una vez hecha la conexión con Python y Microsoft Power BI se han creado las visualizaciones. En Python se han utilizado librerías como pandas, matplotlib, seaborn, plotly, sklearn y numpy. El proceso analítico ha sido el siguiente:

- **Análisis principal:** El objetivo de esta fase es tener una idea general del rendimiento del spy, junto a una comparación con el Ibex 35 para tener una referencia de un índice europeo. También se busca correlacionar el volumen con el precio de cierre y con la volatilidad para saber si es determinante para el comportamiento del mercado.
- **Estrategias de inversión:** Se plantea qué estrategia de trading puede ser la mejor, para ello se usan 4 tipos de estrategia: Comprar cada mes, comprar cada año, comprar cada 3 años, o una menos realista, comprar solo al inicio. Se invierte 7200\$ cada 3 años, 2400\$ cada año y 200\$ cada mes.

- **Análisis predictivo:** Se entrenan tres modelos distintos para intentar predecir una subida del 2% en los próximos tres días. Se usa un modelo lineal (Logistic Regression) y dos modelos no lineales (Random Forest y Gradient Boosting). Para ello se usarán 3 variables: La volatilidad y el cambio porcentual de precio y volumen del día actual con el anterior.
- **Análisis simulado:** A partir de datos pasados, se simula la posible dirección que tendrá el precio en el futuro. Se usan los 2 conjuntos de datos porque se conoce el precio actual del Ibex 35 y enriquece la visualización.

Resultados

Los resultados obtenidos con los análisis han sido los siguientes:

- No parece existir ninguna relación entre el volumen de mercado y los cambios en el precio. Hay días en los que hay poco volumen con cambio positivo del precio y otros donde hay mucho volumen pero poco cambio en el precio. Lo mismo ocurre con la volatilidad. La fecha, como día, mes o año, no influyen en el mercado.
- La mejor estrategia es invertir cada 3 años, seguida por la de invertir cada mes y finalmente la de invertir cada año.
- Aunque por muy poco, el modelo de Random Forest es el que mejor rendimiento obtiene. Aún así, los resultados obtenidos no dejan de ser ligeramente mejores al lanzamiento de una moneda. En la Figura 2 se puede apreciar que el modelo acierta más prediciendo que el precio va a subir pero falla más acertando que bajará, lo cual es positivo si interesa conocer cuándo comprar.

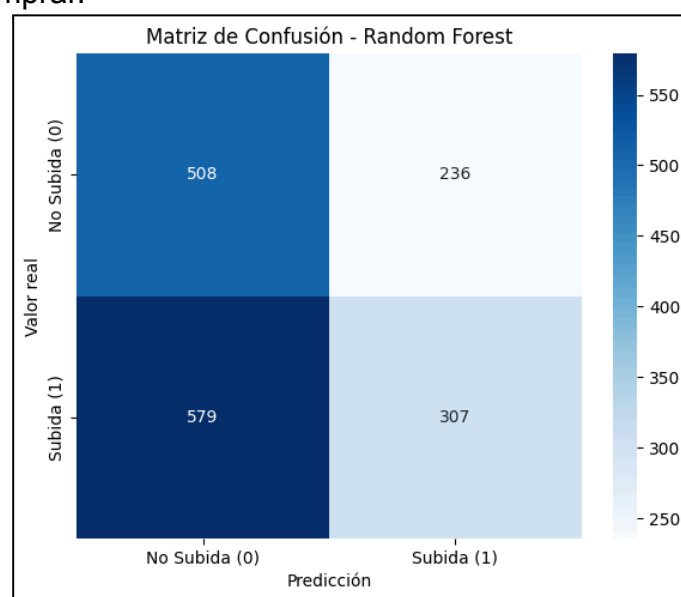


Figura 2: Matriz Confusión de Random Forest. Creación propia

- En la simulación de Montecarlo, a pesar de que los resultados no difieren mucho de la evolución histórica del SPY, deben interpretarse con cautela, ya que provienen de un proceso aleatorio.

Discusión

Los resultados encontrados reflejan las dificultades vistas en otros estudios de predicción bursátil. A pesar del uso de modelos variados los niveles de precisión obtenidos han sido moderados y apenas superiores a una predicción aleatoria. Esto concuerda con la hipótesis del mercado eficiente (Fama, 1965), según la cual toda la información disponible ya está reflejada en los precios, y por tanto no es posible obtener beneficios extraordinarios de forma consistente mediante modelos basados en datos históricos.

Sin embargo, los modelos han mostrado cierta capacidad para identificar patrones complejos en los datos, lo que sugiere que, aunque el mercado es altamente eficiente, pueden existir oportunidades de predicción parcial o condicional bajo ciertas circunstancias.

Por otro lado, la simulación de Monte Carlo permite ver la alta variabilidad de posibles trayectorias futuras del índice SPY. A pesar de basarse en supuestos estadísticos, este enfoque resulta útil para entender la incertidumbre a largo plazo y puede ser una herramienta práctica para la gestión del riesgo.

Además, la comparación entre diferentes estrategias de inversión (inversión inicial única, aportaciones anuales o mensuales) evidencia que el enfoque de "buy & hold" o las inversiones periódicas tienden a generar resultados más estables y positivos en el largo plazo, especialmente la mensual.

Finalmente, estos resultados abren la posibilidad de usar indicadores técnicos más avanzados o el uso de modelos más sofisticados, especialmente alguno que sea capaz de entender tendencias sociales, eventos políticos o incluso crisis económicas, ya que intervienen de forma directa con el precio del SPY y otros índices, es decir, tener en cuenta otros datos que no sean los pasados como el precio o volumen.

Conclusiones

Este trabajo ha permitido aplicar diferentes técnicas de análisis y predicción sobre el índice SPY, utilizando tanto modelos estadísticos como simulaciones. Se ha comprobado que, aunque existen patrones que los modelos pueden detectar parcialmente, la predicción precisa del comportamiento bursátil sigue siendo extremadamente difícil.

Por otro lado, las simulaciones de Monte Carlo han demostrado ser útiles para representar la incertidumbre futura de manera visual e intuitiva, y las estrategias de inversión pasiva o periódica han confirmado su eficacia como alternativas sólidas frente a intentos de "predecir el mercado".

En conjunto, este trabajo contribuye a reforzar la idea de que los mercados financieros, especialmente los índices ampliamente seguidos como el SPY, presentan un alto grado de

eficiencia, y que las decisiones de inversión deben considerar tanto la información histórica como el riesgo y la incertidumbre inherente al comportamiento futuro.

Referencias bibliográficas

Blume, M. E., & Edelen, R. M. (2004). S & P 500 Indexers, Tracking Errors, and Liquidity. *Journal of Portfolio Management*, 30(3), 37-46.

Cho, B. J. (2025). *SPY Stock Price Movement Technical Analysis Using Machine Learning* (Doctoral dissertation, CALIFORNIA STATE UNIVERSITY, NORTHRIDGE).

SBI Securities. (s. f.). *How to Read Candlestick Charts*. Recuperado 22 de julio de 2025, de <https://www.sbisecurities.in/blog/how-to-read-candlestick-charts>

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, 38(1), 34-105.