



# Predicción de tendencias del EUR/USD usando modelos de predicción para series temporales

Juan José Pérez Avendaño, Andrés López Echeverri

August 2025

## 1. Resumen

La predicción de tendencias en los mercados financieros constituye un desafío relevante tanto para la investigación académica como para la toma de decisiones en inversión y trading algorítmico. En este trabajo se aborda el par EUR/USD, el más líquido del mercado de divisas, empleando datos históricos en dos escalas temporales: diaria (Yahoo Finance) y horaria (Binance). Se construyó un conjunto de features basados en indicadores técnicos de momentum, tendencia, volatilidad y volumen, sobre los cuales se entrenó el algoritmo XGBoost, optimizando hiperparámetros para maximizar la capacidad predictiva. Los resultados muestran un accuracy del 65 % en datos diarios y del 60 % en datos horarios, confirmando que en horizontes más amplios el modelo logra un mejor equilibrio en la detección de tendencias alcistas y bajistas. Adicionalmente, mediante pruebas de hipótesis se valida estadísticamente que el rendimiento obtenido supera significativamente al azar, lo que respalda la existencia de patrones predictivos en las series temporales analizadas. Estos hallazgos sugieren que modelos de machine learning aplicados a divisas pueden constituir una herramienta robusta para el análisis de mercados, con potencial de extenderse a otros pares y marcos temporales.

## 2. Introducción

Los mercados financieros constituyen un sistema complejo en el que confluyen factores económicos, políticos y sociales, generando dinámicas altamente volátiles y difíciles de anticipar. Dentro de este contexto, el mercado de divisas (Foreign Exchange Market, ForEx) ocupa un lugar central al ser el mayor y más líquido del mundo, con un volumen diario que supera los trillones de dólares. Entre los distintos pares de divisas, el euro frente al dólar estadounidense (EUR/USD) destaca por su relevancia global y su uso como referencia en múltiples transacciones internacionales.

El interés en desarrollar herramientas que permitan analizar y, en la medida de lo posible, predecir los movimientos de este par no solo radica en sus aplicaciones prácticas en inversión y gestión de riesgos, sino también en el reto metodológico que representa. La naturaleza no estacionaria de los precios, su alta sensibilidad a noticias y eventos externos, y la presencia de ruido hacen que los modelos tradicionales de predicción resulten insuficientes en muchos casos.

En este trabajo se plantea un enfoque basado en técnicas de machine learning para la clasificación de tendencias en el par EUR/USD, considerando diferentes horizontes temporales. El objetivo es evaluar el potencial predictivo de un modelo de gradient boosting (XGBoost) sobre datos diarios

y horarios, y analizar cómo varía el desempeño según la granularidad temporal.

Se plantea un flujo completo de trabajo: desde la recopilación y preprocesamiento de datos históricos, hasta la construcción de features relevantes a partir de indicadores técnicos como retornos, medias móviles, volatilidad, RSI, MACD y Williams %R, entre otros.

Principalmente se utilizó el método XGBoost, un algoritmo de gradient boosting ampliamente reconocido por su capacidad de manejar relaciones no lineales, procesar datos tabulares de alta dimensionalidad y ofrecer gran rendimiento incluso en contextos con ruido. Además, se incorporan estrategias de optimización de hiperparámetros tanto del modelo como de los indicadores técnicos empleados, buscando maximizar la capacidad predictiva.

Finalmente, se incluye una prueba de hipótesis estadística que compara los resultados de los modelos con un experimento aleatorio (lanzamiento de moneda), con el fin de validar que el rendimiento obtenido no es atribuible al azar.

Este enfoque busca no solo alcanzar altos niveles de accuracy, sino también construir un sistema sólido que pueda adaptarse a diferentes marcos temporales y ofrecer un punto de partida para extenderse a otros pares de divisas y contextos financieros.

### 3. Marco Teórico

#### Series de tiempo y modelos de ML aplicados a las finanzas

Podemos definir una serie de tiempo como una sucesión de observaciones recogidas en intervalos regulares de tiempo, cuyo análisis tiene la intención de identificar patrones, tendencias y dinámicas subyacentes. En el caso de los mercados financieros, las series de tiempo contienen información sobre la evolución de variables como el precio de cierre, el volumen transado o la volatilidad de un activo. Ver [1, 2].

En general, las series de tiempo presentan cuatro componentes fundamentales. Vease [3]:

- **Tendencia:** dirección general del comportamiento de la serie en un intervalo de tiempo determinado.
- **Estacionalidad:** patrones que se repiten de manera periódica en intervalos de tiempo específicos (no suele presentarse en los mercados financieros).
- **Ciclo:** comprende las fluctuaciones presentes en periodos largos y se asocian a condiciones macroeconómicas.
- **Variabilidad:** se refiere a la impredecibilidad en el corto plazo, puede deberse a la especulación entre los actores que interactúan de manera directa con los mercados, noticias económicas o políticas, o la alta volatilidad del mercado.

Un aspecto crucial en el análisis de series financieras es la no estacionariedad. Una serie es estacionaria si sus propiedades estadísticas, como la media y la varianza, permanecen constantes en el tiempo. Sin embargo, los precios financieros suelen mostrar heterocedasticidad (la dispersión de los datos alrededor de la línea de regresión no es uniforme), cambios abruptos y dependencia temporal, lo que complica el modelado y requiere transformaciones (como diferenciación o estandarización) antes de aplicar técnicas de predicción [2].

El uso de técnicas de *machine learning* en mercados financieros responde a la necesidad de capturar patrones ocultos y no lineales en los datos, que los modelos estadísticos tradicionales no siempre logran representar de manera adecuada. Estos métodos se pueden aplicar tanto en la predicción de precios futuros como en la clasificación de tendencias, la detección de anomalías y la gestión de riesgos. Vease [4, 5].

En particular, la tarea de predicción de tendencias puede abordarse como un problema de clasificación, donde el objetivo es asignar una etiqueta que represente el movimiento del mercado en un horizonte temporal determinado. Para este trabajo, las tendencias se codifican con dos valores:

- 0: tendencia bajista
- 1: tendencia alcista

Este enfoque presenta ventajas frente a la predicción de precios absolutos, ya que reduce la influencia de valores extremos y se centra en la dirección del movimiento, lo cual es más relevante para la toma de decisiones en contextos de inversión o trading; sin embargo, no tener en cuenta una clasificación para el comportamiento lateralizado de la serie temporal, que se presenta cuando no hay una tendencia clara, puede resultar en un sesgo hacia una de las dos clases y a una mala funcionalidad en estos mercados laterales Ver [6].

## 4. Metodología

### 4.1. Datos y temporalidades

Los datos fueron extraídos de dos plataformas completamente diferentes: Yahoo Finance y Binance [7, 8], utilizando sus respectivas APIs públicas. La primera tiene la limitante de que los datos pueden extraerse con una temporalidad mínima de un día y las características iniciales son: precio de apertura y cierre, y precio más alto y bajo; pero con la ventaja de que posee datos desde inicios de diciembre del año 2003, aunque únicamente se utilizaron desde el 01/01/2010 hasta el más reciente registrado. Por otro lado, la segunda plataforma, al estar plenamente enfocada en el trading de criptomonedas, solo nos permite extraer datos de EUR/USDT, divisa que representa la comparación entre el euro y el tether, que es la stablecoin vinculada al dólar estadounidense y tiene un comportamiento similar al EUR/USD pero con mayor volatilidad; la ventaja es que permite extraer los datos en temporalidades menores (una hora en nuestro caso), además de estar compuestos de una mayor cantidad de características iniciales, como el volumen y el número de operaciones; en este caso se utilizaron datos en el rango [01/03/2020 , 01/10/2025].

### 4.2. Características construidas

Para los datos de Yahoo Finance se contruyeron inicialmente las siguientes features con el fin de mejorar la clasificación:

- **Indicadores de momentum:** se utilizan para medir la fuerza y dirección del movimiento del precio en distintos horizontes temporales. Ayudan a identificar si el mercado mantiene una tendencia o está perdiendo impulso.
  - Retorno porcentual diario (**Return\_1d**): para capturar la dirección y cambio reciente del precio.
  - Retorno porcentual en los últimos 5 días (**Return\_5d**): para capturar tendencias a corto plazo (semanal, pues en ForEx los mercados financieros se cierran los fines de semana).
  - Retorno porcentual en los últimos 10 días (**Return\_10d**): para identificar movimientos sostenidos o señales de reversión a mediano plazo.
- **Indicadores de tendencia:** Permiten suavizar las fluctuaciones del precio para detectar la dirección general del mercado y posibles señales de cambio.
  - Media móvil de 5 días (**MA\_5**): representa la tendencia de muy corto plazo.
  - Media móvil de 10 días (**MA\_10**): refleja una tendencia de corto a mediano plazo y se puede usar junto con **MA\_5** para identificar cruces de medias.
- **Osciladores técnicos:** diseñados para detectar condiciones de sobrecompra o sobreventa, lo que puede anticipar reversiones de tendencia.

- Índice de fuerza relativa (**RSI**): mide la magnitud de las ganancias recientes frente a las pérdidas para detectar si un activo está sobrecomprado o sobrevendido.
- Williams %R (**WILLR**): indicador similar al RSI pero más sensible a los cambios rápidos de precio.
- **Indicadores de volatilidad:** analizan la variabilidad del precio y su posición relativa dentro de un rango, ayudando a identificar momentos de contracción o expansión del mercado.
  - Porcentaje B de las Bandas de Bollinger (**BB\_percent\_B**): indica la posición del precio dentro del canal de Bollinger, útil para detectar extremos del mercado.
- **Ponderación de observaciones:**
  - Peso ajustado por volatilidad (**Weight**): diseñado para reducir el impacto de observaciones con grandes movimientos diarios, que pueden introducir ruido en el modelo.

Sin embargo, tras entrenar el modelo, se aplicó un algoritmo para decidir qué combinación de estas influía de mejor manera en el desempeño del mismo y se encontró que los features que conducían al mejor performance del modelo en la temporalidad de un día son:

- **Return\_1d** (Indicador de momentum)
- **Return\_5d** (Indicador de momentum)
- **Return\_10d** (Indicador de momentum)
- **MA\_5** (Indicador de tendencia)
- **MA\_10** (Indicador de tendencia)
- **RSI** (Oscilador técnico)
- **WILLR** (Oscilador técnico)
- **BB\_percent\_B** (Indicador de volatilidad)

Por otro lado, para los datos de Binance las características construidas inicialmente fueron:

- **Indicadores de momentum:**
  - **RSI\_6**: índice de fuerza relativa con ventana de 6 períodos; mide el impulso reciente del precio en horizontes muy cortos.
  - **RSI\_14**: índice de fuerza relativa con ventana estándar de 14 períodos; útil para detectar condiciones de sobrecompra o sobreventa.
  - **MACD**: diferencia entre dos EMAs (6 y 19); indicador de momentum que señala cruces y cambios en la fuerza de la tendencia.
  - **MACD\_signal**: línea de señal del MACD (EMA de 4 períodos sobre el MACD); usada para generar señales de compra/venta.
  - **MACD\_hist**: histograma que muestra la diferencia entre el MACD y su línea de señal; indica el impulso creciente o decreciente.
  - **ROC\_6**: tasa de cambio del precio de cierre a 6 períodos; indica la velocidad y dirección del cambio de precio.
- **Indicadores de tendencia:**
  - **EMA\_6**: media móvil exponencial de 6 períodos; responde más rápidamente a los cambios recientes en el precio.
  - **EMA\_24**: media móvil exponencial de 24 períodos; captura tendencias más estables y de mediano plazo.

■ **Indicadores de volatilidad:**

- ATR\_6: rango verdadero promedio con ventana de 6 períodos; mide la volatilidad del precio considerando gaps y rangos diarios.
- BB\_percent\_B: porcentaje B de las Bandas de Bollinger con ventana de 12 períodos; muestra la posición relativa del precio dentro de las bandas.
- STD\_3: desviación estándar del precio de cierre en una ventana de 3 períodos; mide la volatilidad a muy corto plazo.
- STD\_12: desviación estándar en una ventana de 12 períodos; mide la volatilidad a corto-mediano plazo.

■ **Osciladores técnicos:**

- WILLR\_12: indicador Williams %R con ventana de 12 períodos; señala sobrecompra o sobreventa comparando el cierre con el rango alto-bajo reciente.

■ **Indicadores basados en volumen:**

- OBV: volumen en balance; acumula volumen según si el cierre fue superior o inferior al anterior. Relaciona movimiento de precio con presión de volumen.

Tras aplicar el algoritmo para encontrar las features más influyentes en el performance del modelo para los datos con temporalidad de una hora se redujo a los siguientes:

- Close (Precio)
- ROC\_6 (Indicador de momentum)
- Open (Precio)
- EMA\_6 (Indicador de tendencia)
- ATR\_6 (Indicador de volatilidad)
- STD\_12 (Indicador de volatilidad)
- High (Precio)
- STD\_3 (Indicador de volatilidad)
- MACD (Indicador de momentum)
- Low (Precio)
- OBV (Indicador basado en volumen)

En este caso se definieron ventanas de 24 horas para la clasificación previa de los datos, de manera que se pretende estimar la tendencia en la siguiente hora en base a los datos de las 24 horas anteriores.

Para ambos casos (temporalidad diaria y horaria) se utilizó el algoritmo XGBoost, un método de boosting basado en árboles de decisión reconocido por su precisión, eficiencia y capacidad para manejar relaciones no lineales en los datos. La elección de este modelo se debe a su robustez frente al ruido, su flexibilidad para integrar distintos tipos de features y su idoneidad en tareas de clasificación.

El entrenamiento incluyó la optimización de hiperparámetros clave, como la profundidad máxima de los árboles, la tasa de aprendizaje, el número de iteraciones y parámetros de regularización, lo que permitió obtener un equilibrio entre ajuste y generalización en ambas temporalidades.

## 5. Análisis y Resultados

Los modelos entrenados mostraron un desempeño superior al azar en ambos conjuntos de datos, validando la idea inicial de la posibilidad de extraer información predictiva de divisas utilizando modelos de machine learning a partir de indicadores técnicos.

### 5.1. Yahoo Finance: Resultados con datos diarios

El modelo entrenado con datos diarios de Yahoo Finance alcanzó un accuracy del 65 % como se puede observar en la Tabla 1, además muestra un balance adecuado entre clases. Para la clase 0 (tendencia bajista), la precisión fue de 0.68, lo que indica que cuando el modelo predice una caída acierta casi 7 de cada 10 ocasiones. El recall, en cambio, se situó en 0.59, reflejando que de todas las caídas reales el modelo logra identificar unas 6 de cada 10. Así, el F1-score de 0.63 muestra un buen equilibrio general para esta clase.

En cuanto a la clase 1 (la tendencia alcista), la precisión fue medianamente menor, de 0.62, pero con un recall significativamente más alto, de 0.70, lo que muestra que el modelo es capaz de detectar la mayoría de los días en los que efectivamente el precio subió, aunque comete más errores de tipo falsos positivos que para la clase 0 (aún así con buena precisión). Así, para esta clase el modelo tiene un F1-score de 0.66, un mejor balance aún que el obtenido con la clase 0.

Finalmente, tanto el promedio macro como el ponderado alcanzaron 0.65 en todas las métricas, lo que evidencia un rendimiento equilibrado y consistente entre ambas clases.

	Precision	Recall	F1-score	Support
0	0.68	0.59	0.63	414
1	0.62	0.70	0.66	396
Accuracy			0.65	810
Macro avg	0.65	0.65	0.65	810
Weighted avg	0.65	0.65	0.65	810

Cuadro 1: Reporte de clasificación datos Yahoo Finance

La Figura 1 muestra la evolución de las ganancias acumuladas sobre un dataset de prueba durante varios días, aplicando el modelo generado. Se puede notar un crecimiento estable y sostenido durante todos los días de prueba, lo que muestra un funcionamiento ideal para el modelo, sin mejor comportamiento en una época que en otra.

### 5.2. Binance: Resultados con datos horarios

En el caso de los resultados obtenidos con los datos horarios de Binance, el modelo alcanzó un accuracy global del 60 %, lo cual confirma que, aunque la capacidad predictiva disminuye frente al escenario con datos diarios, sigue existiendo una buena predicción superior al simple azar.

Se observa que para la clase 0 la precisión alcanzó un valor de 0.58, lo que significa que cuando el modelo predice una caída acertó en un 58 % de los casos. Sin embargo, el recall fue más bajo, de 0.49, indicando que de todas las caídas reales, el modelo solo fue capaz de detectar la mitad aproximadamente. Esto refleja una dificultad importante para identificar de manera consistente los movimientos bajistas en marcos intradía, que se puede notar en su F1-score de 0.54, casi 0.5.

Por el contrario, para la clase 1 el modelo mostró un mejor desempeño, con una precisión de 0.61, y más relevante aún, un recall de 0.69, señalando que el modelo logra detectar cerca del 70 % de las jornadas en las que efectivamente el precio subió. Así, esta clase tiene un F1-score de 0.65 que confirma este mejor equilibrio en comparación con la clase bajista.

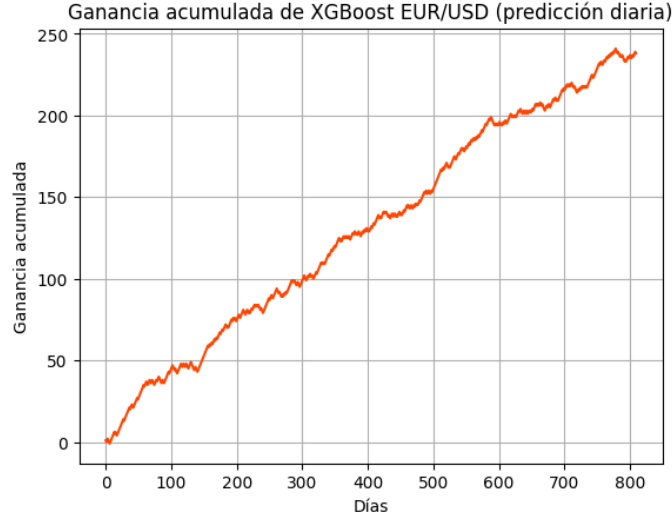


Figura 1: Ganancia acumulada del modelo.

En términos generales, tanto el promedio macro como el ponderado se sitúan alrededor de 0.59-0.60, lo cual refleja un rendimiento aceptablemente balanceado entre ambas clases, aunque con una clara desventaja en la predicción de tendencias bajistas. Estos resultados sugieren que, en escalas horarias, la mayor volatilidad y el ruido del mercado dificultan la identificación de patrones robustos, limitando la capacidad del modelo para capturar señales bajistas, aunque sí logra aprovechar con mayor efectividad las señales alcistas.

	Precision	Recall	F1-score	Support
0	0.58	0.49	0.54	4117
1	0.61	0.69	0.65	4674
Accuracy			0.60	8791
Macro avg	0.60	0.59	0.59	8791
Weighted avg	0.60	0.60	0.59	8791

Cuadro 2: Reporte de clasificación datos Binance

La comparación entre ambos escenarios evidencia que el modelo entrenado con datos diarios de Yahoo Finance ofrece un rendimiento superior al basado en datos horarios de Binance, con una diferencia de aproximadamente cinco puntos porcentuales en el accuracy (65 % frente a 60 %). Se observa que en temporalidades diarias el modelo logra un balance más estable entre ambas clases, detectando tanto subidas como bajadas con niveles similares de precisión y recall. En cambio, en temporalidades horarias, el rendimiento se ve condicionado por la dificultad para identificar caídas, mientras que la detección de movimientos alcistas resulta más efectiva. Esta diferencia sugiere que, a mayor granularidad temporal, el ruido y la volatilidad del mercado diluyen las señales predictivas, limitando la capacidad del algoritmo para extraer patrones consistentes. En conjunto, los resultados refuerzan la idea de que horizontes temporales más amplios permiten construir modelos más robustos y con mayor capacidad de generalización.

### 5.3. Prueba de hipótesis

Con el fin de validar estadísticamente que los resultados obtenidos no se deben al azar, se realizó una prueba de hipótesis contrastando el accuracy alcanzado por los modelos frente al esperado en un experimento aleatorio equivalente al lanzamiento de una moneda con probabilidad  $p=0.5$ .

La hipótesis nula ( $H_0$ ) plantea que el modelo no ofrece ventaja predictiva respecto al azar, es decir, que su accuracy esperado es del 50 %. La hipótesis alternativa ( $H_1$ ) sostiene que el modelo

alcanza un accuracy significativamente mayor a este valor.

Para llevar a cabo la prueba se simuló la distribución del accuracy bajo la hipótesis nula, generando múltiples experimentos de clasificación aleatoria con el mismo número de observaciones que en nuestros conjuntos de prueba. Como se aprecia en las Figuras 2-3, la distribución se concentra en torno al 50 %, con una dispersión relativamente estrecha producto del tamaño de las muestras.

En el caso de Yahoo Finance, el modelo alcanzó 524 aciertos, un valor muy alejado de la región central de la distribución bajo  $H_0$ . El valor-p asociado a esta observación fue de  $2,48 \times 10^{-17}$ , lo que representa evidencia estadísticamente suficiente en contra de la hipótesis nula, bajo cualquier nivel de significancia razonable.

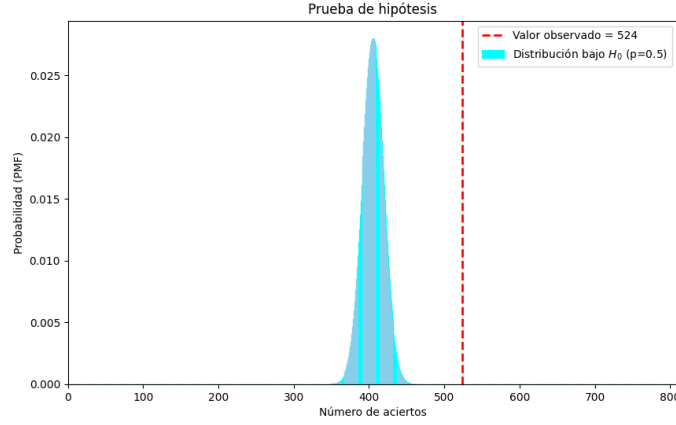


Figura 2: Distribución de la probabilidad de aciertos bajo la hipótesis nula ( $H_0$ ). La línea punteada roja indica el número de aciertos observados por el modelo que usó los datos de Yahoo Finance.

Por su parte, para el caso de Binance, el modelo obtuvo 5256 aciertos, también ubicado claramente lejos de la distribución de referencia. El valor-p correspondiente fue de  $2,95 \times 10^{-79}$ , lo cual permite de nuevo rechazar contundentemente la hipótesis nula.

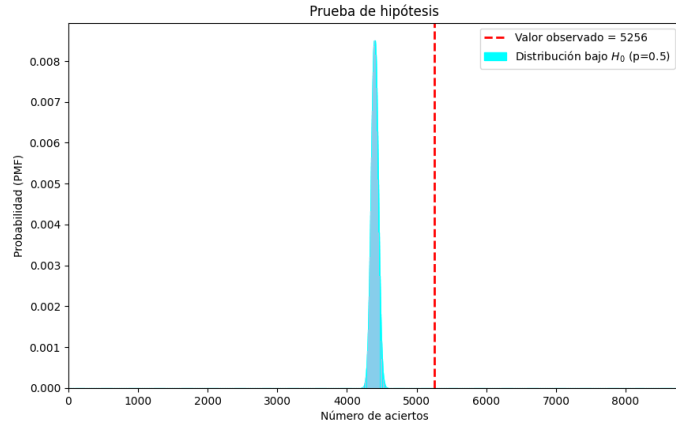


Figura 3: Distribución de la probabilidad de aciertos bajo la hipótesis nula ( $H_0$ ). La línea punteada roja indica el número de aciertos observados por el modelo que usó los datos de Yahoo Finance.

En conjunto, los resultados muestran que el desempeño del modelo supera significativamente lo que cabría esperar de una estrategia aleatoria, confirmando que existe capacidad predictiva real en los patrones aprendidos por el modelo.



## 5.4. Validación del modelo usando datos diferentes

Para el caso de Binance se realizó la respectiva validación utilizando datos del año 2025, fuera del rango considerado para el entrenamiento y prueba del mismo. En el cuadro 3 se muestra el desempeño obtenido, en el cual notamos que el accuracy global baja a 58% y que particularmente las estimaciones de la tendencia bajista son significativamente peores comparadas con las conseguidas con los datos de test, mientras que para la tendencia alcista son bastante similares; esto puede deberse a que al no considerar el comportamiento lateralizado como una característica en sí misma, se presenta un sesgo hacia una de las dos definidas que en este caso aparece a favor de la tendencia alcista.

	Precision	Recall	F1-score	Support
0	0.51	0.17	0.26	2260
1	0.59	0.88	0.71	3077
<b>Accuracy</b>			0.58	5337
<b>Macro avg</b>	0.55	0.53	0.48	5337
<b>Weighted avg</b>	0.56	0.58	0.52	5337

Cuadro 3: Reporte de clasificación (datos de 2025)

## 5.5. Importancia de los features

Además del rendimiento global, resulta relevante identificar cuáles fueron las variables que más contribuyeron a las predicciones del modelo. En anexos, la Figura 4 presenta el ranking de importancia de los indicadores técnicos utilizados como features para el caso de los datos diarios de Yahoo Finance. Se observa que las variables con mayor peso fueron los retornos en diferentes ventanas temporales (1, 5 y 10 días), junto con el indicador *Williams %R*. Estos resultados sugieren que la información contenida en los cambios porcentuales recientes del precio tiene un papel determinante en la predicción, complementada por la sensibilidad a niveles de sobrecompra o sobreventa.

Asimismo, otros indicadores como las Bandas de Bollinger y el RSI también mostraron una contribución significativa, aunque algo menor en comparación con los retornos. Por último, las medias móviles de 5 y 10 periodos tuvieron una importancia más baja, lo cual indica que en temporalidades diarias las señales de tendencia más largas tienen un peso reducido frente a los movimientos recientes del precio.

En el caso de los datos horarios de Binance, la Figura 5 muestra un escenario diferente. El indicador *On-Balance Volume* (OBV) resalta como la característica dominante, con una importancia muy superior al resto (90 frente a valores menores a 30 en las demás variables). Esto sugiere que, en temporalidades intradía, el volumen acumulado de negociación constituye la señal más determinante para el modelo.

En conjunto, estos resultados contrastan con los obtenidos con datos diarios de Yahoo Finance: mientras que en escalas largas las predicciones se apoyan principalmente en retornos pasados y osciladores, en escalas cortas la dinámica del volumen gana un papel preponderante. Esto refuerza la idea de que la granularidad temporal condiciona el tipo de señales más útiles para el modelo.

## 6. Conclusiones

Este trabajo exploró la predicción de tendencias en el par EUR/USD mediante modelos de *machine learning*, aplicados a datos en dos escalas temporales: diaria y horaria. Los resultados permiten extraer varias conclusiones principales:

1. **Capacidad predictiva real:** Los modelos entrenados con XGBoost lograron desempeños significativamente superiores al azar, lo que fue corroborado mediante pruebas de hipótesis

estadísticas. Esto confirma que, aun en un mercado altamente volátil, es posible identificar patrones útiles en las series temporales de precios.

2. **Mejor rendimiento en temporalidad diaria:** El modelo con datos de Yahoo Finance alcanzó un *accuracy* del 65 %, superando en cinco puntos porcentuales al obtenido con datos horarios de Binance (60 %). Esto evidencia que en marcos de tiempo más amplios el ruido del mercado es menor y los patrones resultan más estables, lo que se traduce en un mejor equilibrio en la predicción de tendencias alcistas y bajistas.
3. **Limitaciones en datos horarios:** Aunque el modelo intradía ofrece resultados aceptables, muestra una menor capacidad para identificar caídas, lo que refleja la dificultad de capturar movimientos bajistas en contextos de alta volatilidad.
4. **Robustez del enfoque:** La estrategia de construcción de *features* a partir de indicadores técnicos resultó adecuada para capturar información relevante, mientras que la optimización de hiperparámetros contribuyó a mejorar el desempeño y la generalización de los modelos.
5. **Perspectivas futuras:** Este trabajo abre la puerta a extender el análisis hacia otros pares de divisas, incorporar técnicas más avanzadas como redes neuronales recurrentes (LSTM, TFT) y evaluar marcos temporales adicionales. Asimismo, sería valioso explorar la inclusión de variables macroeconómicas y noticias financieras como *features* complementarias para enriquecer el poder predictivo.

En conjunto, se demuestra que los modelos de *machine learning* pueden constituir una herramienta útil y robusta para el análisis de divisas, aportando información valiosa para la gestión de inversiones y el diseño de estrategias de trading algorítmico.

## 7. Apéndices

### Códigos:

- Código predicción con Yahoo Finance
- Código predicción con Binance

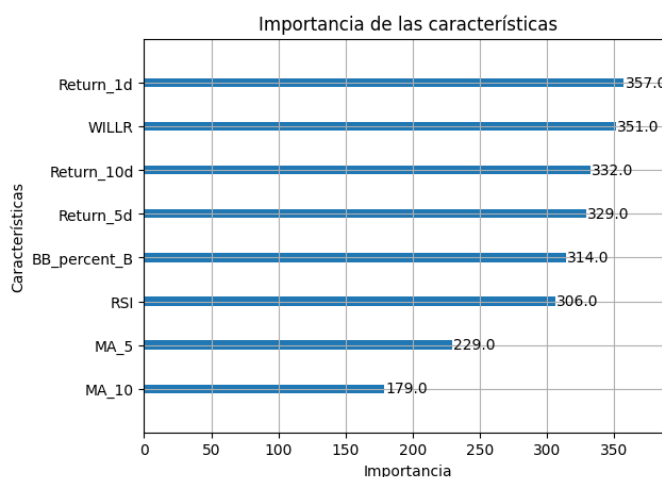


Figura 4: Importancia de cada feature en el entrenamiento con datos de Yahoo Finance.

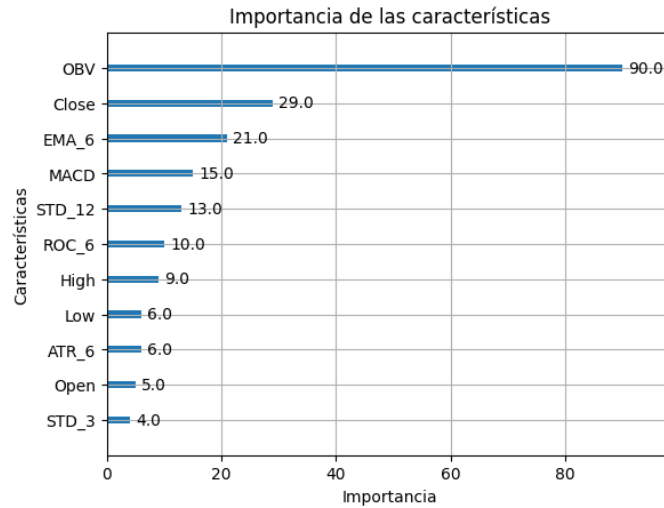


Figura 5: Importancia de cada feature en el entrenamiento con datos de Binance.

## Referencias

- [1] J. D. Hamilton, *Time Series Analysis*. Princeton University Press, 1994.
- [2] R. S. Tsay, *Analysis of Financial Time Series*. Wiley, 2010.
- [3] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. Wiley, 2015.
- [4] B. M. Henrique, V. A. Sobreiro, and H. Kimura, “Literature review: Machine learning techniques applied to financial market prediction,” *Expert Systems with Applications*, vol. 124, pp. 226–251, 2019.
- [5] A. Gupta and B. Dhingra, “Stock market prediction using machine learning algorithms,” *International Journal of Computer Applications*, vol. 70, no. 19, 2012.
- [6] T. Fischer and C. Krauss, “Deep learning with long short-term memory networks for financial market predictions,” *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, 2018.
- [7] Binance, “Binance: Cryptocurrency exchange.” <https://www.binance.com/>, 2025. Accedido: 16-ago-2025.
- [8] Yahoo Finance, “Yahoo finance.” <https://finance.yahoo.com/>, 2025. Accedido: 16-ago-2025.