



# Comparación de Métodos de Aprendizaje de Máquina en el Análisis de Series Temporales Para la Predicción de Tasas de Cambio

Stevens Restrepo Vallejo

Tesis de grado

Asesora:  
Paula Maria Almonacid Hurtado

Universidad EAFIT  
Escuela de Ciencias Aplicadas e Ingeniería  
Maestría en Ciencia de Datos y Analítica  
Medellín  
2024

# 1 RESUMEN

El estudio de los mercados financieros a nivel global constituye un terreno de investigación intrincado, caracterizado por su alta competitividad y volatilidad. El análisis de las tasas de cambio es un punto focal para inversores y firmas, quienes buscan maximizar su rentabilidad mientras minimizan sus riesgos. A pesar de que actualmente hay técnicas diversas para estimar fluctuaciones del precio de las tasas de cambio, persiste la complejidad en el mercado para el análisis preciso y confiable de estos datos debido a su naturaleza estocástica y la influencia que tienen factores político-económicos sobre su comportamiento. En este estudio, se aborda la predicción del precio de algunas de las tasas de cambio más relevantes en este mercado. Se comparan métodos de aprendizaje de máquina que han mostrado rendimiento sobresaliente en la literatura en la predicción de series temporales y se evalúa su rendimiento respecto a un modelo lineal de base; principalmente se usan Modelos de Bosques Aleatorios, Redes Neuronales tipo Long Term Short Memory (LSTM) y un modelo híbrido de Redes Neuronales Convolucionales (CNN) combinadas con LSTM. Además, se explora la robustez de este tipo de modelos en presencia de datos atípicos con el fin de mitigar el riesgo asociado a predicciones con comportamientos altamente variables en los datos. Se busca entonces desarrollar un marco analítico adaptable que permita a inversores y analistas financieros anticipar los movimientos del mercado, mejorando su capacidad para tomar decisiones fundamentadas en datos.

**Palabras clave:** Predicción de Tasas de Cambio, Aprendizaje de Máquina, Long Short Term Memory, Convolutional Neural Network

## Abstract

The study of global financial markets represents a complex field of research, characterized by high competitiveness and volatility. The analysis of exchange rates serves as a focal point for investors and firms aiming to maximize profitability while minimizing risks. Although various techniques currently exist for estimating exchange rate price changes, the inherent stochastic nature of the market, coupled with the influence of political-economic factors, continues to pose significant challenges for precise and reliable data analysis. This study addresses the prediction of the prices of some of the most significant exchange rates in this market. Machine learning methods, which have demonstrated outstanding performance in the literature on time series forecasting, are compared and evaluated against a baseline linear model. The study primarily employs Random Forest models, Long Short-Term Memory (LSTM) neural networks, and a hybrid model combining Convolutional Neural Networks (CNNs) with LSTMs. Additionally, the robustness of these models is explored in the presence of outliers, with the aim of mitigating the risks associated with predictions involving highly variable data behaviors. The goal is to develop an adaptable analytical framework that enables investors and financial analysts to anticipate market movements, thereby enhancing their ability to make data-driven, informed decisions.

**Keywords:** Exchange Rate Prediction, Machine Learning, Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN)

## 2 PLANTEAMIENTO DEL PROBLEMA

El análisis de los mercados financieros a nivel global constituye un campo de estudio complejo y altamente dinámico. En particular, el estudio de las tasas de cambio emerge como un área de alto interés, particularmente para inversores y firmas, dado su impacto directo en las estrategias de inversión y en la rentabilidad de las operaciones financieras. Sin embargo, dicho análisis es constantemente desafiante por la volatilidad de los mercados y la complejidad de los factores que influyen en las fluctuaciones de las tasas de cambio.

A pesar del amplio abanico de técnicas disponibles para prever las variaciones futuras en las tasas de cambio, que incluye métodos estadísticos tradicionales y otros enfoques como el análisis técnico y el análisis fundamental, persiste una dificultad significativa para generar predicciones precisas y confiables. Esta dificultad se atribuye, en parte, a la naturaleza estocástica que tienen los mercados financieros y a la presencia de factores externos impredecibles que pueden influir en las tendencias del mercado.

Estimar los precios de tasas de cambio de forma precisa es fundamental para inversores y analistas financieros en la toma de decisiones estratégicas y el manejo de sus portafolios de inversión. Aún teniendo en cuenta esto, un desafío que se encuentra de forma frecuente en el campo es la presencia y el manejo de datos atípicos en los conjuntos de datos financieros estudiados, que potencialmente pueden tener tanta influencia sobre los datos que pueden distorsionar las estimaciones realizadas por modelos predictivos tradicionales altamente sensibles a este tipo de datos.

Habiendo mencionado lo anterior, es necesario no solamente contar con algoritmos en capacidad de hacer estimaciones del precio de tasas de cambio, sino que también se espera poder desarrollar estos modelos en el marco de enfoques más robustos y eficientes que puedan identificar y manejar estas anomalías, de forma que se pueda mejorar la confiabilidad de las estimaciones y reducir el riesgo asociado a transacciones con tasas de cambio.

La elección y evaluación de algoritmos y técnicas de aprendizaje de máquina apropiados es crucial para desarrollar modelos predictivos precisos y confiables. La comparación de diferentes algoritmos de aprendizaje de máquina, como Bosques Aleatorios, Redes Long Term Short Memory (LSTM) y modelos híbridos, puede proporcionar datos relevantes respecto a qué modelos tienen un mejor ajuste ante diferentes tasas de cambio.

## 3 JUSTIFICACIÓN

En pleno auge de la cuarta revolución industrial, la inteligencia artificial y el uso de técnicas computacionales, matemáticas y estadísticas para la ejecución y optimización de tareas que típicamente han sido delegadas al análisis y gestión humano, ha surgido la necesidad de mantenerse actualizado en las tendencias de cada industria para permanecer competitivo y vigente respecto a otros negocios y profesionales, esto no ha sido ajeno al área financiera, ya que en este campo es donde más se han transformado los métodos de la industria para analizar y seguir los movimientos de activos financieros, predecirlos y optimizar las ganancias netas mientras se mitigan riesgos.

Además, contar con información adicional generada por herramientas de análisis y predicción, les permite a los inversores tomar decisiones informadas sobre la compra, venta o retención de activos financieros; sin embargo, la volatilidad inherente al mercado cambiario, junto con el efecto de factores macroeconómicos y geopolíticos, dificulta la predicción precisa de los precios de las tasas de cambio. En este sentido, el uso de técnicas de aprendizaje de máquina se presenta como una herramienta que potencialmente puede mejorar la capacidad de anticiparse a los mercados de los profesionales, ya que, de esta forma, pueden realizar análisis más grandes y con volúmenes de datos históricos mayores, y a su vez, identificar patrones complejos que pueden ser difíciles de detectar mediante métodos tradicionales de predicción sobre series de tiempo o incluso más, sobre técnicas de análisis fundamental usadas en la industria.

Sumado a los retos mencionados anteriormente, el mercado cambiario está a menudo expuesto a la aparición de comportamientos anómalos que pueden distorsionar la capacidad predictiva de los modelos y afectar considerablemente la precisión de las estimaciones. Por lo tanto, es fundamental desarrollar también enfoques robustos capaces de identificar y reducir la influencia de este tipo de datos, permitiendo así una mayor fiabilidad de los algoritmos usados en la toma de decisiones financieras.

Finalmente, el hecho de poder contar con un marco analítico adaptable para realizar estimaciones del precio de una tasa de cambio potencialmente permite a los inversores y analistas financieros tener una ventaja competitiva, mientras que disminuye el riesgo de operar en estos contextos volátiles al mejorar la capacidad de prever los movimientos de estos mercados y tomar determinaciones fundamentadas en datos.

## **4 OBJETIVOS**

A continuación se listarán los objetivos de este proyecto:

### **4.1 OBJETIVO GENERAL**

Desarrollar un marco analítico adaptable a diferentes tasas de cambio del mercado que esté basado en técnicas de aprendizaje de máquina para realizar estimaciones más precisas y que puedan mejorar la toma de decisiones de inversores y analistas financieros.

### **4.2 OBJETIVOS ESPECÍFICOS**

1. Identificar y recopilar conjuntos de datos relevantes para diferentes activos financieros del mercado cambiario mientras se evalúa su calidad e integridad.
2. Realizar limpieza y pre-procesamiento de los datos que serán usados en los modelos de aprendizaje de máquina que serán evaluados.
3. Evaluar la robustez y precisión de los modelos ante diferentes condiciones del mercado, escenarios adversos y la presencia de datos atípicos.

4. Interpretar los resultados obtenidos del modelo para proporcionar recomendaciones prácticas a inversores y analistas financieros.

## 5 ESTADO DEL ARTE

Diversos estudios se han realizado en el campo de la predicción de tasas de cambio y otros activos financieros haciendo uso de algoritmos de aprendizaje de máquina, (Soni & Tewari, 2022) realizó un estado del arte respecto a este tema en donde menciona que diversos algoritmos se han usado para la predicción de activos como el precio de acciones y tasas de cambio, y en su revisión divide estos modelos en cuatro grupos principales:

- Técnicas tradicionales de aprendizaje de máquina entre los que destacan el algoritmo de Bosques Aleatorios, Máquinas de Soporte Vectorial, Naive Bayes y K-Nearest Neighbors y en donde según el autor, para conjuntos de datos grandes la mejor precisión identificada se logró usando Bosques Aleatorios. y con conjuntos pequeños, usando Naive Bayes.
- Redes Neuronales y Redes Neuronales Profundas; en esta categoría el algoritmo de LSTM fue el que mejores resultados obtuvo en todos los experimentos llevados a cabo en la literatura.
- Análisis de series de tiempo con algoritmos como ARIMA (Autoregressive Integrated Moving Average)
- Otras aproximaciones basadas en grafos

Dentro de las aproximaciones que se revisaron en (Soni & Tewari, 2022) y (Panda & Pattnaik, 2022) , se encontraron algunos modelos híbridos que hacían combinaciones entre algunos tipos de redes neuronales y otros algoritmos basados en árboles u otras redes neuronales. Estas aproximaciones híbridas en general mostraron tener buenos resultados (superiores a otros modelos tradicionales) y gran robustez estadística ante la presencia de datos atípicos.

Consistente con los estudios encontrados en (Soni & Tewari, 2022), el modelo propuesto por (Hochreiter & Schmidhuber, 1997) de Long Term Short Memory (LSTM) ha sido ampliamente usado en la literatura de predicción de tasas de cambio como en (Yao, Gu, Chang, & Zhao, 2022), (Singh, Gutta, & Hadaegh, 2021) y (Ding & Qin, 2020) en donde frecuentemente se muestran mejores resultados haciendo uso de este modelo respecto a otras técnicas paramétricas y no paramétricas tradicionalmente usadas en el contexto del aprendizaje de máquina para la predicción de tasas de cambio.

También se encontraron algunos otros modelos que usaban datos alternos además de la serie temporal como en (Soni & Tewari, 2022) y en (Xueling, Xiong, & Yucong, 2023) en donde se usaron técnicas de análisis de sentimientos de las principales noticias de cada día en foros económicos para analizar las tendencias del mercado junto a la serie de tiempo. Esta metodología tuvo una buena precisión en las predicciones realizadas, y que fue superior a otros algoritmos contrastados. Además, (Shahbazi & Byun, 2022) evaluó métodos de

redes neuronales basados en indicadores técnicos de criptomonedas usando modelos de ensamble como el algoritmo de gradiente extremo de (Friedman, 2001) para clasificación de tendencias en tasas de cambio en donde el modelo XGboost mostró ligeras mejoras por sobre modelos CNN y LSTM.

Por su parte, (Alonso Monsalve, Suárez Cetrulo, Cervantes, & Quintana, 2020) usó un modelo híbrido de Redes Neuronales Convolucionales y LSTM para evaluar la tasa de cambio de criptomonedas en transacciones de alta frecuencia (predicción de lapsos pequeños de generalmente 1 o 2 minutos) con pequeñas variaciones, pero alto número de transacciones para un mayor retorno. Según los resultados de su experimento, este modelo híbrido fue significativamente mejor que el resto de los modelos usados en el estudio. De la misma forma, este fue usado en (Banu et al., 2021) destacando que su uso tiene mejores resultados que usar CNN's y LSTM's por separado.

Finalmente, los modelos como la Red Neuronal Robusta de Laguerre propuesta por (Nanda, Vyas, & Vamshidhar, 2018) muestran desempeños similares en entrenamiento a los de los modelos lineales tradicionales como ARMA o ARIMA; sin embargo, estos últimos tienen una gran caída de rendimiento en conjuntos de prueba, en los que el modelo robusto tiene superioridad contundente, mostrando la importancia del manejo de datos atípicos por medio de enfoques no paramétricos para este tipo de casos. (Amat, Michalski, & Stoltz, 2018) también hace diversos chequeos de robustez estadística de modelos tradicionales de aprendizaje de máquina a partir de métodos de remuestreo ordenado de datos (tomar muestras de diferentes secuencias de datos) y evalúa que el error obtenido en estos sea uniforme en el tiempo, en lugar de solo tener unos pocos ejemplos adecuadamente estimados.

## 6 MARCO TEÓRICO

El mercado de tasas de cambio es uno de los rubros financieros más grandes del mundo, que según (Wang, Wang, Li, & Wang, 2021) y (Xueling et al., 2023) a su vez influyen sobre otros sectores comerciales cuando hay cambios importantes en este rubro; además, se ven afectados significativamente los flujos de capital, el manejo de portafolios de inversión e incluso las actividades económicas de los países. Adicionalmente (Sarangi, Chawla, Ghosh, Singh, & Singh, 2022) y (Liu, Hou, & Liu, 2017) mencionan que este mercado al ser uno de los más transados diariamente en el mundo, mueve alrededor de cinco trillones de dólares al día. En términos sencillos, la tasa de cambio es la cantidad que se necesita de una moneda para adquirir otra y dado que este tipo de transacciones involucra la mayor parte de divisas a nivel mundial, afectará no solo la capacidad adquisitiva individual de las personas para comercializar bienes que se producen en otros países y regiones geográficas, sino también la capacidad de las naciones para abastecerse de los recursos necesarios en el mantenimiento y suministro de servicios básicos para su pueblo.

(Majhi, Rout, Majhi, Panda, & Fleming, 2012) menciona que la predicción de diversas tasas de cambio está influenciada por muchos factores, tanto económicos, políticos y hasta psicológicos, por lo que la predicción de estos valores suele ser una tarea compleja a través del uso de métodos estadísticos convencionales. Así mismo, este autor también resalta

que las aproximaciones estadísticas clásicas asumen una relación lineal en los datos, lo que no es verdad en el caso de la transacción de tasas de cambio, y que ha demostrado estar influenciada por tantas variables; en general, los modelos estadísticos actuales no pueden explicar el comportamiento de este tipo de series temporales de una manera lineal y precisa.

Al respecto (Wei, Kee, & Musa, 2023) resalta que tradicionalmente se han empleado métodos estadísticos como modelos lineales de tipo auto-regresivos y de media móvil (ARMA) que han logrado capturar patrones estacionales en los datos, haciendo referencia a la identificación de tendencias en los mercados a tener alzas o bajas durante ciertas temporadas del año, sin embargo, el mismo autor menciona que este tipo de modelos es sensible a anomalías como cambios en el calendario, temporadas de festividades e incluso el efecto de los fines de semana.

Por otra parte, los modelos auto-regresivos, pese a que pueden capturar ciertos cambios en los patrones estacionales que siguen los datos del precio de una divisa frente a otra en el tiempo (lo cual será considerado como series de tiempo a lo largo de este artículo), siguen siendo altamente sensibles a cambios fuertes en el comportamiento del precio que no pueden ser explicados por una relación lineal en los datos.

Dadas las limitaciones que suponen los modelos estadísticos paramétricos tradicionales y la gran relevancia que tiene para inversores, analistas financieros y, en general, para las naciones el poder anticiparse a las fluctuaciones en el precio de las tasas de cambio, han surgido aproximaciones no paramétricas (que no asumen una relación lineal en los datos o una distribución específica de los mismos) para superar el rendimiento de los modelos clásicos, y estos han logrado demostrar un mejor rendimiento en las estimaciones del precio futuro de una tasa de cambio. (Wang et al., 2021)

(Naidoo & Du, 2022) menciona que en las series de tiempo financieras, los datos atípicos son todos aquellos que no siguen la trayectoria normal de los datos, son aleatorios por naturaleza y no pueden ser predichos. Es en este contexto donde (Peña & Yohai, 2022) menciona la importancia de usar métodos de estimación robustos aplicados a series de tiempo, puesto que la presencia de datos atípicos puede alterar significativamente las estimaciones que un modelo hace del precio de una tasa de cambio, y este mismo autor menciona que hay tres condiciones en las que pueden presentarse este tipo de datos altamente desviados del comportamiento observado:

1. Eventos inusuales presentes en la serie de tiempo.
2. Datos que no fueron imputados bajo la misma escala o método de recolección usados en los otros datos
3. Datos que fueron imputados en un periodo de tiempo con diferentes condiciones respecto a otros (Aquí podríamos considerar datos como los del comportamiento de los mercados durante la pandemia COVID-19)

Respecto al abordaje de métodos no paramétricos y robustos para la predicción de tasas de cambio, diversos autores han trabajado sobre algoritmos de aprendizaje de máquina,

como una alternativa en la que el algoritmo usado puede aprender sobre los patrones que encuentra en los datos y hacer inferencias extraídas de las características de los mismos para mejorar su desempeño.

Los principales modelos de aprendizaje de máquina que se abordarán en este artículo son los de (Breiman, 2001) con el algoritmo de Bosques Aleatorios, el modelo de (Hochreiter & Schmidhuber, 1997) de Long Term Short Memory (LSTM) y la aproximación híbrida de Redes Neuronales Convolucionales (CNN por sus siglas en inglés) desarrollados por (Lecun, Bottou, Bengio, & Haffner, 1998) y extendido por (Wang et al., 2021) en donde se usó el modelo CNN de base para la extracción de características relevantes y que, posteriormente, se usaron como datos de entrada para un modelo LSTM con rendimiento mejorado. A continuación se explica a detalle los algoritmos mencionados:

## 6.1 RANDOM FOREST

El algoritmo desarrollado por (Breiman, 2001) parte del uso del algoritmo de árboles de decisión y, según el autor, se mostraron mejoras significativas en la precisión al generar un conjunto de árboles que realizan una selección aleatoria de características del conjunto de datos, para posteriormente realizar una votación por la clase mayoritaria elegida en el número de estimaciones totales.

Inicialmente (Breiman, 2001) formula este procedimiento como un clasificador; sin embargo, en la formulación que ofrece en su tesis, también se puede usar como mecanismo de regresión no paramétrico generando árboles dependientes de un vector  $\Theta$  tal que el árbol predictor  $h(x, \Theta)$  tome valores numéricos en lugar de etiquetas de clase, y finalmente, calcula un valor medio sobre el número de árboles (estimadores) usados.

Así entonces, tenemos que:

$$\hat{y}(x) = \frac{1}{B} \sum_{b=1}^B h_b(x)$$

donde:

- $\hat{y}(x)$  es la predicción final para la entrada  $x$ .
- $B$  es el número total de árboles en el bosque (Estimadores).
- $h_b(x)$  es la predicción del árbol  $b$  para la entrada  $x$ .

En vista de que este algoritmo selecciona características de forma aleatoria para cada estimador usado y de esta forma realiza las predicciones que finalmente son promediadas, se considera este como un algoritmo robusto respecto a la presencia de datos atípicos y que puede funcionar muy bien en tareas de predicción de tasas de cambio.

## 6.2 REDES LONG SHORT TERM MEMORY (LSTM)

Dentro del campo del aprendizaje de máquina, uno de los algoritmos que más se han popularizado son las Redes Neuronales Profundas (DNN) y Recurrentes (RNN), algunas de estas según (Hochreiter & Schmidhuber, 1997) emplean el mecanismo de retro-propagación que tiene como finalidad ajustar los pesos de las conexiones que se generan en la red para minimizar el error de sus predicciones, pero que en el caso de las series de tiempo financieras muy extensas, dicho error tiende a ya sea inflarse o a desvanecerse, lo que genera estimaciones muy desviadas de los precios reales observados para predecir en el corto plazo.

Para esto, (Hochreiter & Schmidhuber, 1997) plantea una solución a través de una nueva red recurrente que se combina con un algoritmo de aprendizaje basado en el gradiente descendiente capaz de conectar intervalos de tiempo incluso en casos de datos con mucho ruido y secuencias incomprensibles sin el problema de la pérdida de capacidades predictivas a corto plazo.

La arquitectura del algoritmo LSTM consta de varios elementos:

1. **Puerta de Olvido:** Decide qué información del estado de la celda anterior ( $C_{t-1}$ ) se debe olvidar.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

2. **Puerta de Entrada:** Determina qué nueva información se almacenará en el estado de la celda.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

3. **Candidatos a la Memoria de la Celda:** Genera nuevas posibles informaciones para agregar al estado de la celda.

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

4. **Actualización del Estado de la Celda:** Actualiza el estado de la celda combinando el estado anterior y los nuevos candidatos.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

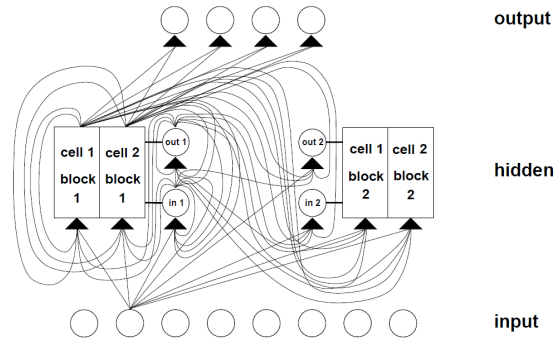


Figura 1: Arquitectura de una Red Neuronal Convolutiva en (Hochreiter & Schmidhuber, 1997)

5. **Puerta de Salida:** Decide qué parte del estado de la celda se usará como salida.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

6. **Actualización del Estado Oculto:** Calcula la salida de la celda y la usa como estado oculto para el siguiente paso temporal.

$$h_t = o_t * \tanh(C_t)$$

donde:

- $\sigma$  representa la función sigmoide.
- $\tanh$  es la función tangente hiperbólica.
- $*$  denota el producto elemento a elemento.
- $W_f, W_i, W_C, W_o$  son los pesos de las puertas.
- $b_f, b_i, b_C, b_o$  son los sesgos de las puertas.
- $h_{t-1}$  es el estado oculto en el tiempo  $t - 1$ .
- $x_t$  es la entrada en el tiempo  $t$ .
- $C_{t-1}$  es el estado de la celda en el tiempo  $t - 1$ .

Se ha mostrado en diversos estudios como este tipo de redes supera en creces las predicciones que puede realizar una Red Neuronal Recurrente en la aplicación a series temporales financieras.

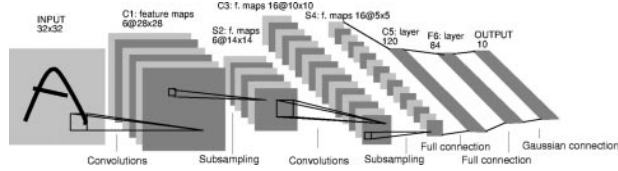


Figura 2: Arquitectura de una Red Neuronal Convolutiva en (Lecun et al., 1998)

### 6.3 MODELO HÍBRIDO (CNN) Y LSTM

Por su parte, (Lecun et al., 1998) estableció el algoritmo de Redes Neuronales de Convulsión (CNN) enfocándolas inicialmente al área de visión por computadora en donde este tipo de redes demostraron un rendimiento superior a los otros modelos empleados en esta área. Después, (Wang et al., 2021) implementaron la CNN como punto de partida para la extracción de características relevantes para la predicción de tasas de cambio a través de la capa convulsional de la red y la capa de pooling que reduce la dimensión de dichas características y que posteriormente se tomaron para realizar las predicciones del modelo LSTM.

La operación de convulsión en una dimensión está definida como:

$$s(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a) \cdot w(t - a)$$

donde:

- $s(t)$  es la salida de la convulsión en el tiempo  $t$ .
- $x(a)$  es la entrada.
- $w(t - a)$  es el filtro (Kernel).

Capas en una CNN:

- **Capas de convulsión:** Extraen características locales aplicando filtros a la entrada.
- **Capas de Pooling:** Reducen la dimensionalidad de las características extraídas.

$$\text{Max Pooling: } P(i, j) = \max_{0 \leq m < M, 0 \leq n < N} X(i + m, j + n)$$

Por otra parte, la segunda opción es usar el método de cálculo usando un estadístico de tendencia central como la media en lugar del valor máximo y este se expresa así:

$$\text{Mean Pooling: } P(i, j) = \frac{1}{M \cdot N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(i + m, j + n)$$

- **Capas de Activación:** Típicamente aplican una función lineal, como ReLU:

$$\text{ReLU: } f(x) = \max(0, x)$$

- **Capas Completamente Conectadas:** Después del paso por las capas de convolución y pooling, las características se llevan a un vector de una dimensión y se pasan a través de capas completamente conectadas para realizar la clasificación final.

## 7 METODOLOGÍA

Se seguirá el esquema del Proceso Estándar Cruzado de la Industria para el Desarrollo de Aplicaciones de Aprendizaje Automático con Garantía de Calidad (CRISP-ML(Q) por sus siglas en inglés) documentado en CRISP-ML(Q) (n.d.) y que cuenta con seis fases para la implementación del desarrollo de modelos y aplicaciones de aprendizaje de máquina:

1. Entendimiento del negocio y de los datos
2. Ingeniería de datos
3. Ingeniería de modelos de aprendizaje de máquina
4. Aseguramiento de la calidad para aplicaciones de aprendizaje de máquina
5. Despliegue
6. Monitoreo y mantenimiento

### 7.1 OBTENCIÓN Y ANÁLISIS DE DATOS

Se recopilan e identifican algunos de los pares de divisas más relevantes en el mercado cambiario desde Yahoo Finance, de los que posteriormente se realiza un análisis exhaustivo de su integridad y calidad a partir del uso del test de Dickey-Fuller para ver la estacionaridad y estacionalidad en los datos.

### 7.2 INGENIERÍA DE DATOS

Se identificarán en este paso las características relevantes para la tarea de predicción, datos atípicos e información faltante en las series de tiempo, así como también se filtrarán aquellas características irrelevantes para el análisis de la serie temporal; así mismo, se realizarán las transformaciones que se consideren pertinentes sobre los datos, como estandarización o normalización de las series; o de ser requerido, transformaciones para convertir los datos en estacionarios.

### 7.3 INGENIERÍA DE ALGORITMOS DE APRENDIZAJE

Se aplicarán los modelos de aprendizaje de máquina mencionados anteriormente, y se ensamblará el procedimiento predictivo del modelo híbrido sugerido de Redes Neuronales

Convolutionales para la extracción de características más relevantes que servirán de entrada para el modelo de Long Short Term Memory. Además, los algoritmos que se van a emplear se empaquetarán en pipelines para el entrenamiento de modelos, de modo que se pueda generar un paso a paso capaz de replicar el entrenamiento de cada algoritmo de forma estándar... En esta misma fase se asegurará que el modelo empleado sea igualmente reproducible por otros investigadores a partir de la referenciación de la metadata usada durante la etapa de implementación y el uso de semillas aleatorias fijas.

## **7.4 EVALUACIÓN DE MÉTRICAS DE DESEMPEÑO**

Para esta fase se evaluarán las métricas de desempeño del modelo sobre conjuntos de prueba de forma que se pueda medir cuantitativamente el error generado entre las estimaciones de los modelos y el precio real observado en las series de tiempo para los diferentes activos financieros. Así mismo, se evaluará el sesgo y varianza en el entrenamiento de cada modelo respecto a los conjuntos de prueba y la robustez estadística de los mismos en presencia de datos atípicos en la serie temporal. Todos los procedimientos de medición realizados serán documentados y analizados en los resultados de este trabajo.

## **7.5 DESPLIEGUE, MONITOREO Y MANTENIMIENTO**

El despliegue del modelo se realizará a través de un espacio en HuggingFace de forma que los analistas puedan interactuar con la herramienta desde sus computadoras, la herramienta será de ayuda en la proyección de tasas de cambio para algún par de divisas (no incluye criptomonedas) y en el manejo de portafolios de inversión en tasas de cambio. El monitoreo y mantenimiento de la herramienta no aplica para este caso dado que por cada interacción de los analistas con la herramienta se generará un nuevo modelo predictivo adaptado a los datos del día en el que se ejecute y de acuerdo al tipo de modelo elegido para el entrenamiento.

# **8 RESULTADOS**

## **8.1 OBTENCIÓN Y ANÁLISIS DE DATOS**

Durante la fase de obtención de datos, el flujo solicita al usuario inputar los dos pares de divisas que quieren ser evaluados y la cantidad de años de datos que se quieren introducir al modelo (así se evita el uso de fechas fijas e incompatibilidad en los caracteres introducidos por el usuario final). La consulta de los datos se hace a través de la librería de Yahoo Finance de donde se recupera una tabla con los precios de apertura, cierre, precio más alto y más bajo, cierre ajustado y el volumen de transacciones del par de divisas. De dicha tabla se extrae el precio de cierre únicamente, que en adelante será la serie temporal que se busca predecir y que se representa en la figura 3.

Posteriormente, en la fase de análisis de datos se realizó un test de Dickey-Fuller para determinar si la serie temporal era estacionaria o no, que para los pares de divisas evaluados



Figura 3: Evolución de la serie temporal USD-COP

arrojaron p-valores por encima del umbral de tolerancia (0.05) de este trabajo, lo que implica que no hay suficiente evidencia para rechazar la hipótesis nula (no se puede concluir que la serie sea estacionaria).

En vista de que la mayoría de pares de divisas evaluados inicialmente no eran estacionarios, se decidió aplicar una transformación a los datos de diferenciación para convertir las series en estacionarias, como se observa en los gráficos de Autocorrelación (ACF) y de Autocorrelación Parcial en la figura 4.

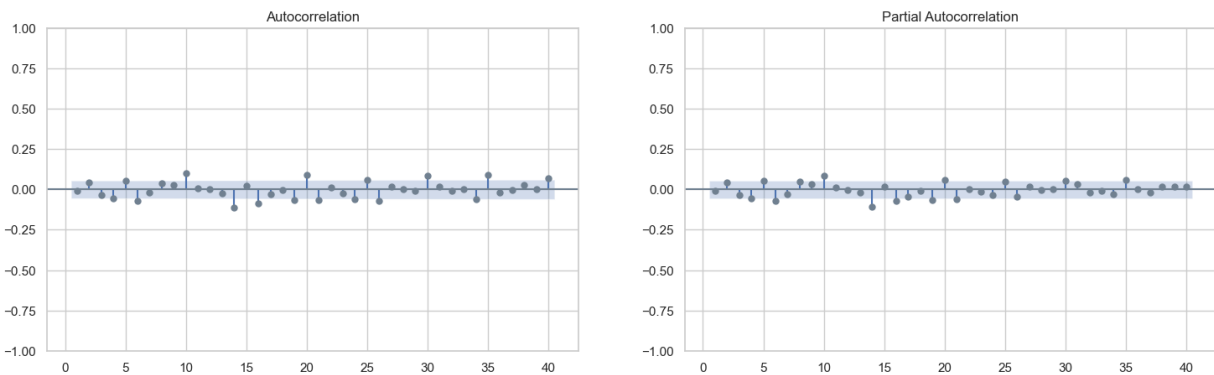


Figura 4: Grafica de Autocorrelación (ACF) y Autocorrelación Parcial (PACF) sobre la serie USD-COP

Se observó un patrón estacionario cada cinco días en la serie temporal que puede corresponder al calendario semanal de actividades de la bolsa de valores que opera de lunes a viernes (ver figuras 5 y 6). Este valor será relevante posteriormente para determinar el número de rezagos que se agregaron en las transformaciones de los datos.

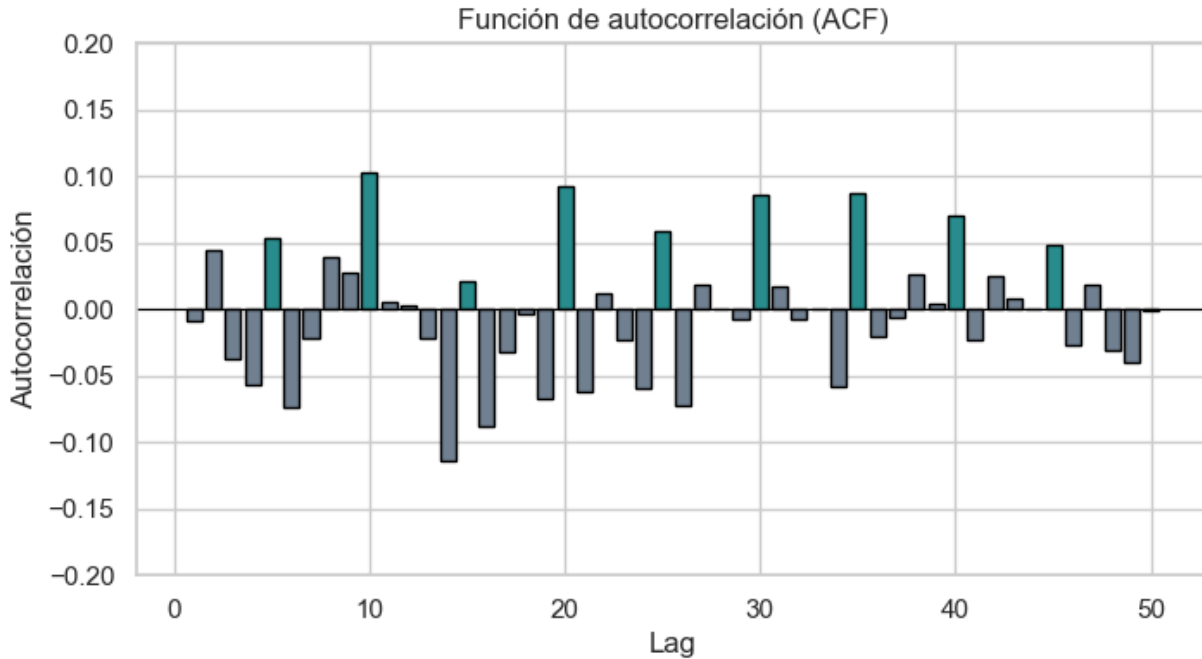


Figura 5: Gráfico de autocorrelación (ACF) en la serie USD-COP

Adicionalmente, se calculó la volatilidad del par de divisas a partir de los rendimientos diarios de los datos financieros. Primero, se calcularon los cambios porcentuales diarios (cambio del dato actual respecto al previo en cada día) y, para el valor faltante que se genera en el primer dato (pues este no tiene un dato previo para hacer el cálculo del cambio porcentual) se reemplazó por el promedio de los rendimientos generales; estos valores se ordenaron de menor a mayor y se estableció un valor de significancia  $\alpha$  del 0.05 para la identificación de datos atípicos basado en el cambio del rendimiento del precio. Como se ve en la figura 7, el rendimiento de los pares de divisas sigue un comportamiento aproximadamente normal. Por otra parte, el valor  $\alpha$  se dividió en partes iguales en ambas colas de la distribución para determinar que el rendimiento de un día se pueda considerar como atípico cuando los rendimientos son muy negativos o muy positivos (inferiores al 0.025 o superiores al 0.975).

Tras realizar este análisis, se determinó que esta variable podría ser incorporada como variable exógena a la serie temporal para proporcionar información valiosa a los algoritmos sobre el comportamiento del precio basado en la volatilidad de los días anteriores.

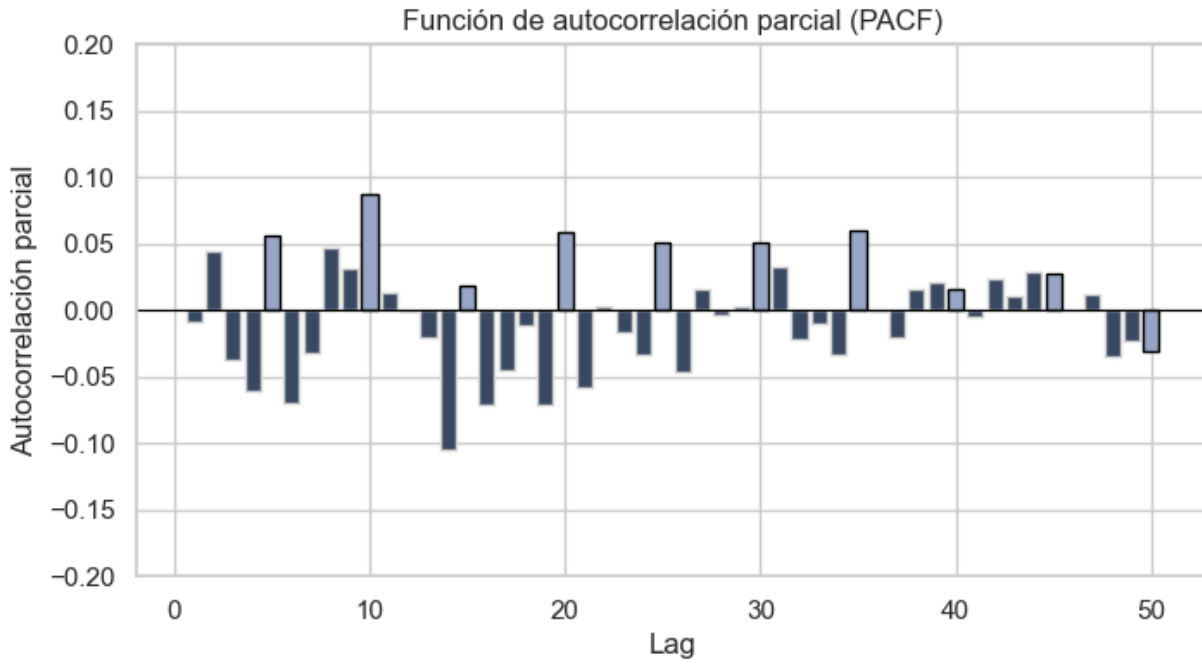


Figura 6: Gráfico de autocorrelación parcial (PACF) en la serie USD-COP

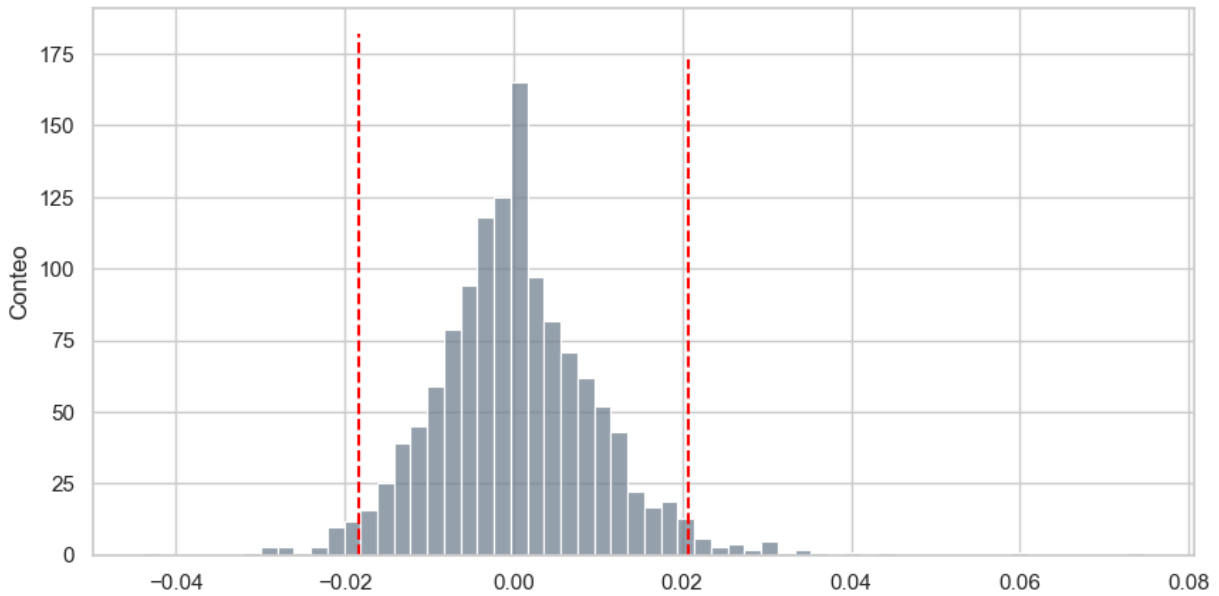


Figura 7: Distribución de los rendimientos de la serie USD-COP e identificación de datos atípicos

## 8.2 INGENIERÍA DE CARACTERÍSTICAS

En la fase de ingeniería de características, se agregó un conjunto de variables basadas en características rezagadas de la misma serie temporal, es decir, se crearon variables nuevas desde la serie temporal univariada con un total de cinco rezagos (basado en los resultados del test de Dickey-Fuller y el análisis del ACF y PACF previamente abordados), lo que convirtió la serie temporal en un conjunto de datos con el precio de cierre, sus últimos cinco valores rezagados y la variable exógena del cálculo del rendimiento respecto al día inmediatamente anterior y en donde los algoritmos podrían evaluar la influencia de atípicos sobre el comportamiento del precio del par de divisas.

Cuadro 1: Conjunto de Datos Transformado

Precio de Cierre	t-1	t-2	t-3	t-4	t-5	Rendimiento
------------------	-----	-----	-----	-----	-----	-------------

donde  $T - n$  se refiere a cada uno de los valores rezagados del precio de cierre en días anteriores en  $n = 1, 2, \dots, N$  y  $N$  es el número total de rezagos que se definieron.

Posteriormente se hace una partición de los datos usando un valor de 0.9 como punto de corte que será usado para entrenamiento y validación de los modelos, y el 0.1 restante serán usados como datos de prueba desconocidos para el modelo entrenado y a partir de estos determinar las métricas de cada uno de los modelos que se evalúan. Cabe resaltar que cuando se trata de series de tiempo, la partición debe mantener el orden secuencial de los datos y no ser aleatoria, por lo que hacer una segunda partición para dividir los datos de entrenamiento, en subconjuntos de entrenamiento y validación permitirá introducir una ventana de tiempo de datos desconocidos inmediatamente previos al conjunto de prueba, asegurando así que los resultados en las predicciones no esté basado en tener todos los datos disponibles, sino más bien en la apropiada captura de los patrones propios de la serie temporal del par de divisas.

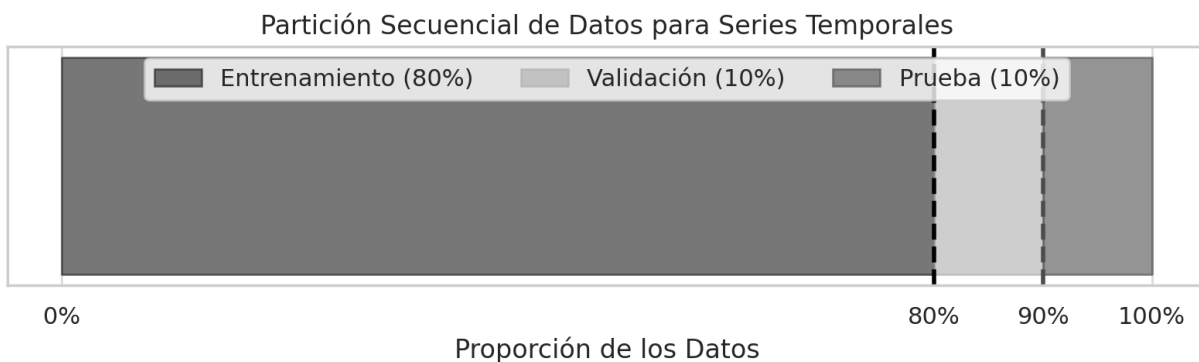


Figura 8: Partición del conjunto de datos

Adicionalmente, los datos que fueron transformados con diferenciación no se les aplicaron transformaciones adicionales, mientras que un segundo conjunto que no fue transformado

con el método de diferenciación, se estandarizó usando StandardScaler de la librería SciKit Learn para centrar los datos y reducir su escala.

Finalmente, se creó un flujo de entrenamiento de modelos que incorpora el escalado de datos, la transformación de los mismos para agregar características rezagadas y la variable exógena y hacer el entrenamiento de cada uno de los algoritmos de aprendizaje de máquina.

### 8.3 INGENIERÍA DE ALGORITMOS DE APRENDIZAJE DE MÁQUINA

Para los algoritmos de aprendizaje de máquina, se usó un modelo lineal de base Linear Regression desde la librería SciKit Learn, este no se le añadió ningún parámetro, ni se ajustó nada de su configuración por defecto. También se usó un modelo de Bosques Aleatorios (Random Forest Regressor) desde la misma librería, cuyo único parámetro fue random state para fijar una semilla aleatoria y asegurar la reproducibilidad de los resultados de este trabajo.

Con el algoritmo de base, se hizo una aproximación inicial para evaluar el desempeño de los modelos al usar datos estacionarios transformados con diferenciación y sin usar la misma. Los resultados se resumen en el cuadro 2.

Cuadro 2: Comparación del MSE entre Datos Estacionarios y No Estacionarios

<b>Condición de los Datos</b>	<b>MSE</b>
No Estacionarios	2.76296
Estacionarios	300.78629

A partir del cuadro anterior, se pudo concluir que la transformación con diferenciación en los datos no aporta a la reducción de la métrica de error obtenida por el modelo, razón por la que se opta por no transformar los datos para convertirlos en estacionarios para el resto de modelos de aprendizaje de máquina.

En cuanto a los algoritmos basados en redes neuronales, se construyeron las clases para el modelo LSTM y para el modelo híbrido que usa CNN-LSTM, cuyas arquitecturas usan los parámetros descritos en el cuadro 3 y 4.

Cuadro 3: Parámetros del Modelo LSTM

<b>Parámetro</b>	<b>Valor</b>
Capa LSTM (unidades)	256
Activación LSTM	ReLU
Capa de salida	Densa (1 unidad)
Optimizador	Adam
Función de pérdida	MSE
Épocas	20
Tamaño del lote	60

Cuadro 4: Parámetros del Modelo híbrido

<b>Parámetro</b>	<b>Valor</b>
Capa Conv1D (filtros)	256
Tamaño del kernel	1
Activación Conv1D	ReLU
Padding Conv1D	Causal
Capa LSTM (unidades)	256
Activación LSTM	ReLU
Capa de salida	Densa (1 unidad)
Optimizador	Adam
Función de pérdida	MSE
Épocas	20
Tamaño del lote	60

Se puede observar en el cuadro 4 que la parte LSTM incorporada al modelo híbrido es igual a la de la arquitectura del modelo LSTM; de esta forma, se evaluó la influencia que tiene la capa convolucional de extracción de características sobre el resultado obtenido en la métrica de error.

## 8.4 EVALUACIÓN DE MÉTRICAS DE DESEMPEÑO

Una vez desarrollado el flujo de entrenamiento de modelos; de forma iterativa se evaluaron cuatro de las divisas (Euro ['EUR'], Dolar Estado Unidense ['USD'], Dolar Canadiense ['CAD'] y Libra Esterlina ['GBP']) más transadas en el mercado y más relevantes a nivel global y de forma iterativa se formaron doce pares de divisas que fueron evaluadas a partir de cada uno de los modelos en periodos de dos hasta diez años en un flujo de ciclos anidados por cada uno de los algoritmos de aprendizaje de máquina, de esta forma se obtuvo la métrica de error sobre el conjunto de prueba de cada uno de los modelos, y sobre cada par de divisas.

Se construyó así entonces una tabla de resultados con 432 métricas de error para cada modelo entrenado con series temporales desde dos hasta diez años de extensión, y por cada par de divisas evaluado. Lo anterior se resume en el siguiente cuadro en donde se promedió el resultado del entrenamiento de todos los lapsos para cada modelo y en cada par de divisas.

En el cuadro 5 se puede observar que todos los modelos entrenados alcanzaron métricas de error cercanas a cero, sin embargo, resaltan especialmente el modelo de base y el modelo híbrido. por reducir en mayor escala el error para la mayoría de pares de divisas evaluadas.

Adicionalmente, en la figura 9 se puede observar que los modelos en general tienden a mejorar conforme incrementa el número de años de datos de entrenamiento proporcionados, en esta figura resalta que el modelo Random Forest necesita de más información que el resto de modelos para poder converger a un error bajo mientras que con pocos datos, tiende a tener un desempeño más deficiente que el resto de modelos.

Cuadro 5: Resultados promedio del MSE por Modelo y Par de Divisas

Par de Divisas	Híbrido	LSTM	Regresión Lineal	Random Forest
CAD/EUR	9.561111e-08	2.730889e-07	4.766667e-09	0.000003
CAD/GBP	2.783656e-06	3.551200e-06	1.900000e-08	0.000053
CAD/USD	2.954556e-07	5.303667e-07	1.006667e-08	0.000010
EUR/CAD	9.774333e-07	1.470089e-06	1.897778e-08	0.000017
EUR/GBP	2.618222e-07	5.289556e-07	2.187778e-08	0.000007
EUR/USD	9.115444e-07	2.498856e-06	1.133333e-08	0.000004
GBP/CAD	2.871556e-06	2.731793e-05	1.609222e-07	0.000524
GBP/EUR	1.342778e-07	8.733778e-07	1.304444e-08	0.000014
GBP/USD	2.123533e-06	5.391089e-06	6.388889e-08	0.000022
USD/CAD	5.168556e-07	2.280311e-06	3.182222e-08	0.000037
USD/EUR	6.472667e-07	1.649922e-06	7.755556e-09	0.000003
USD/GBP	1.006933e-06	1.730800e-06	2.687778e-08	0.000007

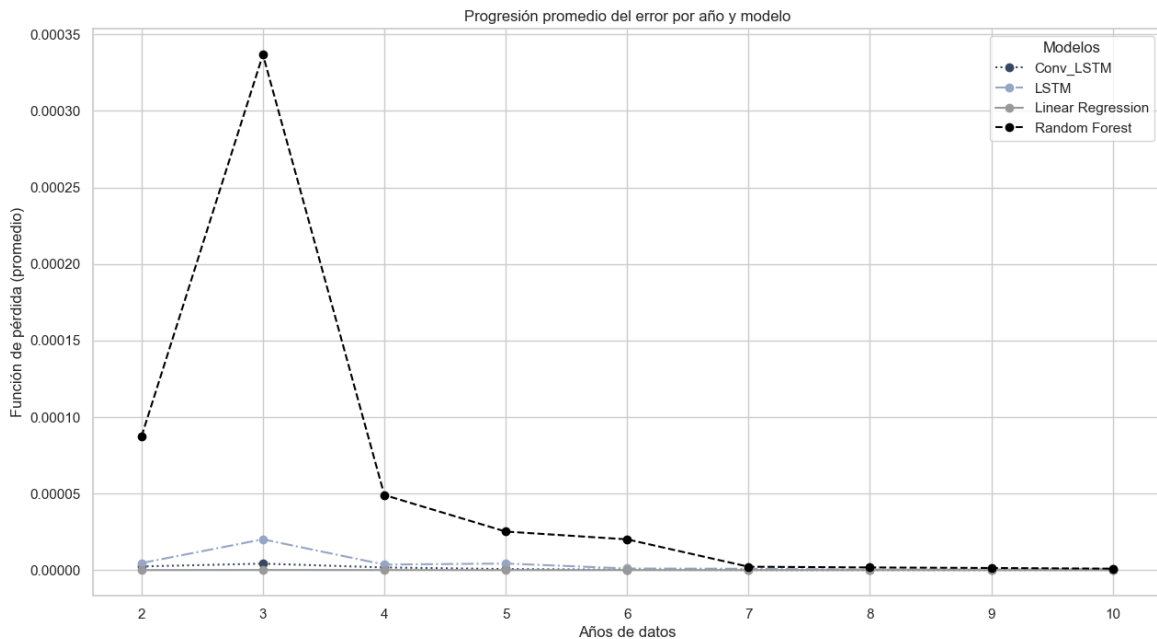


Figura 9: Comparación del rendimiento de los modelos a lo largo de los años

Por otra parte, se observa también que el modelo híbrido muestra resultados consistentes y gracias a su arquitectura logra capturar mejor los patrones temporales a largo plazo respecto al modelo LSTM. Incluso, llegando a igualar al modelo de regresión lineal de base que siempre mantuvo un buen desempeño. Se resalta que en conjuntos de datos con muchos años de información, tiene mejores resultados el modelo híbrido que el modelo de base.

Se evaluó la varianza de la función de pérdida evaluada sobre cada uno de los modelos y se observó que el modelo base y el modelo híbrido fueron los más consistentes casi sin variaciones en la métrica de desempeño. Esto se puede observar en las figuras 10 y 11, donde nuevamente se destaca que el modelo Random Forest es el que mayor varianza tiene

por lo que es el modelo menos confiable para este tipo de tarea entre los modelos evaluados.

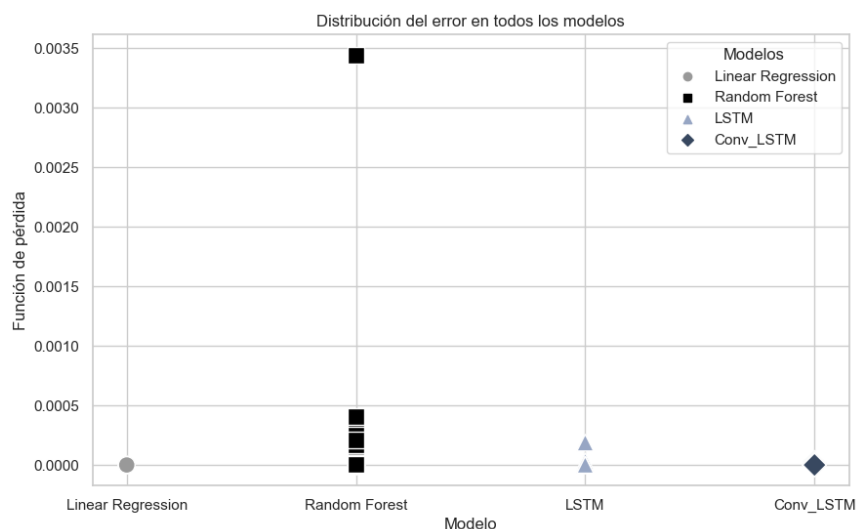


Figura 10: Varianza del MSE por modelo

Particularmente en la figura 10 se ve que la escala del modelo base y del modelo híbrido son las que tienen la función de pérdida MSE más reducidas, sin embargo, es de resaltar que todos los modelos se desempeñan bien cuando la serie temporal tiene suficiente información.

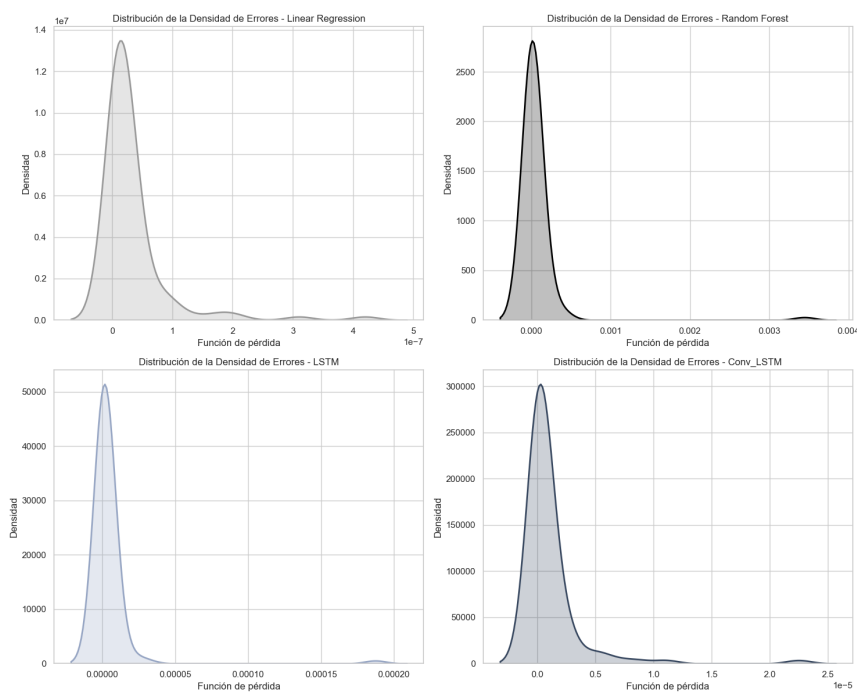


Figura 11: Distribución de densidad del MSE

Por último, respecto a la robustez de los modelos de aprendizaje de máquina, se observó de forma gráfica que los cuatro modelos se ajustan adecuadamente a los cambios en los

patrones del precio cuando hay presencia de datos atípicos; de hecho, en el análisis, esta información es la que suele mostrar cambios en la tendencia de los datos o bajas y subidas repentinas en el comportamiento del precio. Añadir esta característica exógena mejoró considerablemente el rendimiento de todos los algoritmos de aprendizaje de máquina.

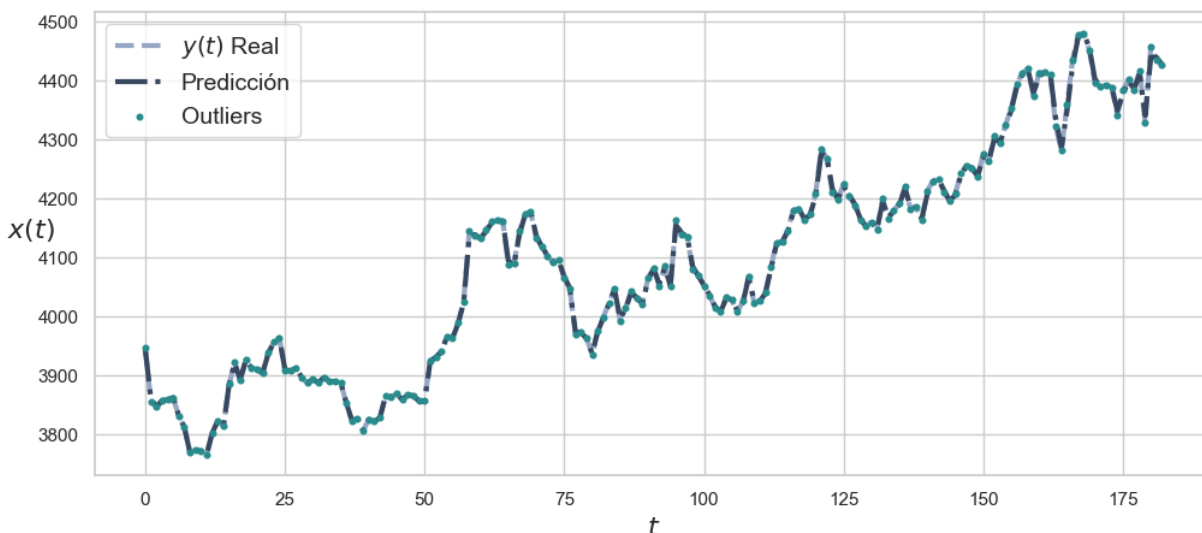


Figura 12: Predicción de la serie con regresión lineal y sensibilidad a datos atípicos USD-COP

En la figura 12 se observa la capacidad predictiva del algoritmo lineal de base, que al tener agregada la variable exógena y las cinco características rezagadas robustece el modelo y hace que pueda capturar con mayor facilidad los patrones de la serie temporal, incluso en presencia de datos atípicos (Outliers). De la misma forma, los otros algoritmos evaluados tienen un rendimiento similar y solo el algoritmo de Bosques Aleatorios muestra mayor dificultad para capturar dichos patrones cuando la cantidad de datos es aún muy reducida.

## 9 CONCLUSIONES

De forma general, los modelos estudiados reflejaron funciones de pérdida bajas en la mayor parte de escenarios; se observó que el algoritmo Random Forest tiene mayor varianza en la distribución de la función de pérdida respecto al resto de modelos cuando generaliza sobre datos desconocidos. Sin embargo, dicha varianza disminuye y converge hacia cero cuando se contaba con por lo menos siete años de datos para hacer el entrenamiento del modelo.

Cuando se evaluaron las gráficas ACF y PACF, se encontró una estacionalidad cercana a los 5 días en la mayor parte de divisas del mercado de tasas de cambio, que se puede explicar por las actividades inherentes a la bolsa de valores y sus transacciones entre lunes y viernes. Se podría pensar que semanalmente se siguen patrones transaccionales de divisas que pudieron ser capturados por los algoritmos de aprendizaje de máquina evaluados.

La distribución de la densidad de los errores para la regresión lineal de base muestra ser la

mejor con una función de pérdida cercana a  $1e-7$ , mientras LSTM y el Conv-LSTM tienen un rendimiento similar en escala de  $1e-4$  y  $1e-5$ , respectivamente... Además, el algoritmo de Random Forest Regressor muestra un error en escala de  $1e-03$ , aún cuando ya se mencionó que el rendimiento es el más bajo entre los métodos evaluados.

Pese a que el algoritmo híbrido Conv-LSTM incorpora la misma arquitectura que la red LSTM, se destaca que agregar una capa convolucional disminuye la función de pérdida mientras mejora la capacidad predictiva, esto se debe a la extracción de características que hace la parte convolucional al aplicar diferentes filtros a los datos para generar nuevas representaciones de los mismos y esto se mueve como entrada a la red LSTM posteriormente; el modelo LSTM tiende a tener mayor varianza que el híbrido, aunque también reduzca de forma significativa el error en las predicciones. Sin embargo, el modelo híbrido tiene mejores métricas incluso conforme el tamaño de la serie de tiempo crece.

Las características rezagadas y la variable exógena agregadas a la serie temporal para esta tarea predictiva, tienen un papel muy relevante en la mejora del rendimiento de los algoritmos evaluados. Esta configuración obtuvo mejores resultados en el algoritmo lineal de base, que, por sus supuestos matemáticos, suele ser mucho más sensible a la presencia de datos atípicos, afectando la precisión. Sin embargo, la variable exógena incorporada que hizo las veces de indicador de retrocesos e incrementos en el valor de cierre, redujo de forma significativa dicha sensibilidad. Por otra parte, se observó que no fue necesario aplicar transformaciones a las series temporales para hacerlas estacionarias mediante el uso de métodos como la diferenciación. Los mejores resultados se obtuvieron solamente estandarizando los datos y agregando las características rezagadas y una variable exógena a los modelos, dicha variable consiste en el cálculo de los rendimientos del día actual de la serie respecto al día inmediatamente anterior, esta variable se incorporó como el cambio porcentual entre cada dato en la secuencia.

Valdría la pena continuar explorando el comportamiento de estos algoritmos al introducir otras variables exógenas como el día de la semana o el mes en el que se hacen las transacciones; incluso sería valioso intentar incorporar predicciones desde otros modelos como entrada de este; como hacer uso de un modelo GARCH para predecir la volatilidad en momentos futuros e incorporar dicha predicción a los modelos, de la misma manera, sería relevante incorporar pronósticos del precio del par de divisas por intervalos de confianza que pueden representar mayor confiabilidad para el analista financiero o evaluar el modelo con base a los valores de la Tasa Representativa del Mercado (TRM).

Finalmente, es importante mencionar que la herramienta puede hacer predicciones confiables; sin embargo, por la naturaleza estocástica del mercado cambiario, conforme se extiende el rango de las predicciones se comienza a acumular errores sobre las predicciones siguientes, por lo que generar proyecciones a largo plazo podría llevar a errores más grandes.

## 10 ANEXOS

### 10.1 REPOSITORIO EN GITHUB

Todos los archivos de este proyecto fueron cargados a un repositorio en GitHub en donde podrá revisarse la descripción del proyecto y requerimientos. Para ingresar al repositorio podrá hacerse a través del siguiente link: [https://github.com/StevensRV/TASAS\\_DE\\_CAMBIO](https://github.com/StevensRV/TASAS_DE_CAMBIO)

### 10.2 MODELO DESPLEGADO

El modelo se desplegó en un espacio de HuggingFace, para facilidad de uso de analistas financieros y otras personas que puedan estar interesadas; en el espacio se puede evaluar cualquier par de tasas de cambio en un periodo mínimo de 2 años (sujeto a la disponibilidad de datos en la API Yahoo Finance). El ingreso a dicho espacio es a través del siguiente enlace: <https://huggingface.co/spaces/StevensRV93/EAFIT>

## REFERENCIAS

- Alonso Monsalve, S., Suárez Cetrulo, A., Cervantes, A., & Quintana, D. (2020). Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators. *Expert Systems With Applications*. doi: <https://doi.org/10.1016/j.eswa.2020.113250>
- Amat, C., Michalski, T., & Stoltz, G. (2018). Fundamentals and exchange rate forecastability with simple machine learning methods. *Journal of International Money and Finance*. doi: <https://doi.org/10.1016/j.jimonfin.2018.06.003>
- Banu, J., Rajeshwari, S., Kallimani, J., Vasanthi, S., Buttar, A., Sangeetha, M., & Bhargava, S. (2021). Modeling of hyperparameter tuned hybrid cnn and lstm for prediction model. *Intelligent Automation & Soft Computing*. doi: <https://doi.org/10.32604/iasc.2022.024176>
- Breiman, L. (2001). *Random forests*. Berkeley, California: University of California.
- CRISP-ML(Q). (n.d.). *The ml lifecycle process*. <https://ml-ops.org/content/crisp-ml>. (Accessed: 2024-05-28)
- Ding, G., & Qin, L. (2020). Study on the prediction of stock price associated network model of lstm. *International Journal of Machine Learning and Cybernetics*, 1307-1318. doi: <https://doi.org/10.1007/s13042-019-01041-1>
- Friedman, J. (2001). *Greedy function approximation: A gradient boosting machine*. (IMS 1999 Reitz Lecture)
- Hochreiter, S., & Schmidhuber, J. (1997). Long short term memory. *Neural Computation*. doi: <https://doi.org/10.1162/neco.1997.9.8.1735>
- Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11).
- Liu, C., Hou, W., & Liu, D. (2017). Foreign exchange rates forecasting with convolutional neural network. *Neural Process Lett*. doi: <https://doi.org/10.1007/s11063-017-9629-z>

-z

- Majhi, B., Rout, M., Majhi, R., Panda, G., & Fleming, P. J. (2012). New robust forecasting models for exchange rates prediction. *Expert Systems with Applications*. doi: <https://doi.org/10.1016/j.eswa.2012.05.017>
- Naidoo, V., & Du, S. (2022). A deep learning method for the detection and compensation of outlier events in stock data. *Electronics*, 11(21). Retrieved from <https://www.mdpi.com/2079-9292/11/21/3465> doi: <https://doi.org/10.3390/electronics11213465>
- Nanda, S., Vyas, R., & Vamshidhar, k. (2018). Forecasting foreign exchange rate using robust laguerree neural network. In *3rd international conference and workshops on recent advances and innovations in engineering*.
- Panda, M., & Pattnaik, P. (2022). Forecasting foreign currency exchange rate using convolutional neural network. *International Journal of Advanced Computer Science and Applications*. (Recuperado el 26 de Mayo de 2024, de [https://thesai.org/Downloads/Volume13No2/Paper\\_72-Forecasting\\_Foreign\\_Currency\\_Exchange\\_Rate.pdf](https://thesai.org/Downloads/Volume13No2/Paper_72-Forecasting_Foreign_Currency_Exchange_Rate.pdf))
- Peña, D., & Yohai, V. (2022). A review of outlier detection and robust estimation methods for high dimensional time series data. *Econometrics and Statistics*. doi: <https://doi.org/10.1016/j.ecosta.2023.02.001>
- Sarangi, P., Chawla, M., Ghosh, P., Singh, S., & Singh, P. (2022). Forex trend analysis using machine learning techniques: Inr vs usd currency exchange rate using ann-ga hybrid approach. *Materials Today: Proceedings*. doi: <https://doi.org/10.1016/j.matpr.2020.10.960>
- Shahbazi, Z., & Byun, Y.-C. (2022). Knowledge discovery on cryptocurrency exchange rate prediction using machine learning pipelines. *Sensors*. doi: <https://doi.org/10.3390/s22051740>
- Singh, S., Gutta, S., & Hadaegh, A. (2021). Stock prediction using machine learning. *Department of Computer Science and Information System California State University*. doi: <https://doi.org/10.37394/232018.2021.9.17>
- Soni, P., & Tewari, D. K., Y. (2022). Machine learning approaches in stock price prediction: A systematic review. *Journal of Physics: Conference Series*, 11. doi: <https://doi.org/10.1088/1742-6596/2161/1/012065>
- Wang, J., Wang, X., Li, J., & Wang, H. (2021). A prediction model of cnn-tlstm for usd/cny exchange rate prediction. *1School of Information Science and Engineering, Hebei University of Science and Technology*. doi: <https://doi.org/10.1109/ACCESS.2021.3080459>
- Wei, M., Kee, O. L., & Musa, S. (2023). Seasonal versus non-seasonal trends in stock market malaysia. *E3S Web of Conferences* 389, 09039. doi: <https://doi.org/10.1051/e3sconf/202338909039>
- Xueling, L., Xiong, X., & Yucong, S. (2023). Exchange rate market trend prediction based on sentiment analysis. *Computers and Electrical Engineering*. doi: <https://doi.org/10.1016/j.compeleceng.2023.108901>
- Yao, W., Gu, Y., Chang, S., & Zhao, J. L. (2022). Stock price analysis and forecasting. *SPIEDigitalLibrary.org/conference-proceedings-of-spie*. doi: <https://doi.org/10.1117/12.2662176>