



**MODELO PREDICTIVO DEL TIPO DE CAMBIO EUR/USD: INTEGRACIÓN DE
DATOS HISTÓRICOS Y ANÁLISIS DE SENTIMIENTO DE NOTICIAS
FINANCIERAS A TRAVÉS DE REDES NEURONALES LSTM.**

Manuela Arias Martínez - 901378
Esteban Quintero Carvajal - 9013878
Joaquín Nicolás Mejía Romero - 9014250

Proyecto Aplicado para optar al título de
Magister en Ciencia de Datos

Directora

Isabel Cristina García Arboleda

FACULTAD DE INGENIERÍA Y CIENCIAS

MAESTRÍA EN CIENCIA DE DATOS

SANTIAGO DE CALI, DICIEMBRE 17 DE 2025

TABLA DE CONTENIDO

1	CONTEXTUALIZACIÓN DEL PROYECTO	7
1.1	DEFINICIÓN DEL PROBLEMA.....	7
1.1.1	PLANTEAMIENTO DEL PROBLEMA.....	7
1.1.2	FORMULACIÓN DEL PROBLEMA.....	8
1.1.3	SISTEMATIZACIÓN:.....	8
1.2	OBJETIVOS:	9
1.2.1	OBJETIVO GENERAL.....	9
1.2.2	OBJETIVOS ESPECÍFICOS.....	9
1.3	MARCO DE REFERENCIA.....	9
1.3.1	MARCO TEÓRICO.....	9
1.3.2	ANTECEDENTES	12
2	PROCESAMIENTO Y ANÁLISIS DE DATOS PARA EL MODELO PREDICTIVO ..	13
2.1	Revisión Bibliográfica y Establecimiento de Criterios	14
2.2	Selección y Recolección de Fuentes de Datos	15
2.3	Implementación del Pipeline de Limpieza de Datos Textuales.....	16
3	ESTRUCTURACIÓN DEL MODELO LSTM	17
3.1	Análisis de Sentimiento de Noticias Financieras:	17
3.1.1	Arquitectura del Sistema Multi-Método:.....	17
3.1.2	Pipeline de Procesamiento	20
3.1.3	Resultados Cuantitativos del Análisis de Sentimiento	22
3.1.4	Output Final y Transición a la Etapa 3	24
3.2	Análisis Exploratorio y Correlacional de Datos Integrados	25
3.2.1	Integración de Fuentes de Datos y Feature Engineering	25
3.2.2	Validación Visual de Correlaciones y Patrones	27
3.3	Diseño y Arquitectura del Modelo LSTM	29
3.3.1	Preparación de Características para LSTM	29
3.3.2	Arquitectura del Modelo LSTM Integrado.....	31
3.3.3	División y Validación del Dataset.....	32
3.3.4	Optimización de Hiperparámetros y Configuración del Modelo.....	32
4	EVALUACIÓN DEL RENDIMIENTO DEL MODELO.....	33
4.1	Selección y Justificación del Modelo de Referencia	33

4.2	Evaluación y Validación del Modelo	35
4.2.1	Análisis Visual Integral del Rendimiento	36
4.2.2	Validación de Robustez y Estabilidad	37
5	CONCLUSIONES Y TRABAJOS FUTUROS	39
5.1	CONCLUSIONES	39
5.2	TRABAJOS FUTUROS	41
6	REFERENCIAS BIBLIOGRÁFICAS	43
7	ANEXOS	45

LISTA DE FIGURAS

Figura 1. <i>Arquitectura de LSTM</i>	10
Figura 2. <i>Metodología Crisp-Dm Adaptada Al Proyecto</i>	13
Figura 3. <i>Análisis Visual Integral de Sentimiento vs Precios EUR/USD.</i>	27
Figura 4. <i>Evaluación Integral del Rendimiento del Modelo LSTM</i>	37
Figura 5. <i>Comparación Visual: Datos Reales vs LSTM vs ARIMA</i>	38

LISTA DE TABLAS

Tabla 1. <i>Transformaciones de Datos por Etapa del Pipeline</i>	14
---	----

INTRODUCCIÓN

La volatilidad del tipo de cambio EUR/USD representa un desafío significativo para inversores y operadores en el mercado de divisas. Como uno de los pares de divisas más negociados mundialmente, las fluctuaciones en este tipo de cambio pueden tener un impacto considerable en la salud económica global. Factores como cambios en políticas monetarias, crisis económicas y eventos geopolíticos influyen en la percepción del mercado y, por ende, en las variaciones del tipo de cambio. En este contexto, se hace necesario desarrollar métodos efectivos que permitan anticipar estos movimientos y favorezcan decisiones informadas por parte de los actores del mercado.

Este proyecto se enfocó en desarrollar un modelo predictivo que integra el análisis de datos históricos y el análisis de sentimiento derivado de noticias financieras, utilizando redes neuronales LSTM. Al incorporar el sentimiento extraído de noticias financieras especializadas en el análisis, se exploró cómo estos factores influyeron en las predicciones del mercado, proporcionando así una nueva perspectiva en la comprensión de la dinámica del tipo de cambio.

Los resultados de este desarrollo contribuyeron a la propuesta de un modelo que potenció la precisión de las predicciones del tipo de cambio y proporcionó insights valiosos sobre la dinámica del mercado. Este enfoque tuvo el potencial de enriquecer el campo de la ciencia de datos y las finanzas, ofreciendo herramientas que facilitaron una comprensión más profunda de los factores que afectaron el comportamiento del EUR/USD.

1 CONTEXTUALIZACIÓN DEL PROYECTO

1.1 DEFINICIÓN DEL PROBLEMA

1.1.1 PLANTEAMIENTO DEL PROBLEMA

El mercado de divisas (Forex), con un volumen diario superior a 7.5 billones de dólares, es el sistema financiero más líquido y volátil del mundo [1]. El par EUR/USD, que representa cerca del 24% de las transacciones globales, actúa como un indicador clave de la salud económica de la zona euro y Estados Unidos [2]. Su comportamiento refleja la interacción compleja de factores macroeconómicos, como tasas de interés, inflación y balanza comercial, así como eventos geopolíticos y percepciones del mercado. Esta relevancia lo convierte en una variable crítica para bancos centrales, fondos de inversión, corporaciones multinacionales y traders individuales, quienes dependen de predicciones precisas para optimizar decisiones y mitigar riesgos financieros [3].

Sin embargo, el EUR/USD es altamente volátil, especialmente en contextos de incertidumbre global. Por ejemplo, entre 2021 y 2024, eventos como la divergencia en las políticas monetarias del Banco Central Europeo (BCE) y la Reserva Federal (FED), la inflación postpandemia, y tensiones geopolíticas, como el conflicto en Ucrania, generaron fluctuaciones erráticas en este par [4]. Estas dinámicas afectaron a empresas que planificaron importaciones/exportaciones, inversores que gestionaron portafolios y responsables de políticas económicas que buscaron estabilidad cambiaria. En este escenario, desarrollar modelos predictivos efectivos para el EUR/USD se vuelve esencial para anticipar movimientos y reducir la incertidumbre en la toma de decisiones.

Las limitaciones de los modelos predictivos actuales pueden reflejarse en:

Baja precisión en entornos volátiles: Los modelos tradicionales, como ARIMA o GARCH, muestran un rendimiento limitado en entornos volátiles, fallando en capturar movimientos bruscos causados por eventos inesperados, como anuncios de bancos centrales o crisis geopolíticas [5]. Esto resulta en predicciones poco confiables que pueden llevar a pérdidas financieras significativas.

Mayor incertidumbre en el mercado: La falta de herramientas que incorporen factores cualitativos, como el impacto emocional de noticias financieras, genera desconfianza entre los actores del mercado, quienes enfrentan dificultades para interpretar señales complejas [5].

Las limitaciones de los modelos predictivos actuales tienen varias causas fundamentales:

Dependencia exclusiva de datos históricos: Los enfoques econométricos tradicionales, basados en series temporales de precios, tasas de interés o indicadores macroeconómicos, no consideraron variables cualitativas, como el sentimiento del mercado derivado de noticias

financieras. Estudios demostraron que el tono emocional de los medios pudo anticipar movimientos de precios, incluso cuando los fundamentos económicos no lo justificaron [6].

Complejidad no lineal del mercado: El EUR/USD respondió a interacciones no lineales entre factores económicos, políticos y sociales, como discursos de bancos centrales, eventos electorales o reacciones a titulares financieros. Según Mishkin, estas dinámicas fueron difíciles de modelar con enfoques lineales [7]

Limitaciones técnicas de los modelos tradicionales: Los modelos como ARIMA no estuvieron diseñados para manejar datos secuenciales complejos ni dependencias a largo plazo, a diferencia de las redes neuronales LSTM, que fueron más adecuadas para capturar patrones temporales no lineales [8].

Falta de integración de datos no estructurados: Las noticias financieras, que contuvieron información valiosa sobre el sentimiento del mercado, fueron datos textuales que requirieron técnicas avanzadas de procesamiento de lenguaje natural para su análisis e incorporación en modelos predictivos [9]. Esta brecha fue notable en Latinoamérica, donde predominaron enfoques basados únicamente en series temporales, como lo evidenció el trabajo de Arias Jiménez sobre el USD/COP, que no incluyó análisis de sentimiento [10].

En este contexto, la ausencia de modelos que integraron datos estructurados (precios, indicadores técnicos) con datos no estructurados (noticias, opiniones) representó una oportunidad para mejorar la predicción del EUR/USD. Herramientas como VADER y TextBlob permitieron cuantificar el sentimiento de textos financieros, mientras que las redes LSTM ofrecieron la capacidad de modelar relaciones secuenciales complejas [10], [9]. Al combinar estas técnicas, este proyecto desarrolló un modelo que integró tendencias históricas con el análisis de percepciones del mercado capturadas en noticias financieras de Bloomberg, logrando mejorar la precisión predictiva del tipo de cambio EUR/USD mediante la incorporación de información cualitativa derivada de 2,106 titulares especializados procesados con metodología multi-método.

1.1.2 FORMULACIÓN DEL PROBLEMA

¿Cómo puede la integración de datos históricos y análisis de sentimiento de noticias financieras mejorar la predicción del tipo de cambio EUR/USD mediante redes LSTM?

1.1.3 SISTEMATIZACIÓN:

Las preguntas que se plantearon al inicio del proyecto para guiar su desarrollo fueron:

- ¿Qué datos históricos y de sentimiento son necesarios para un modelo predictivo efectivo del EUR/USD, y cuáles son sus fuentes más confiables?
- ¿Qué técnicas son más adecuadas para integrar el análisis de sentimiento en una red LSTM?
- ¿Cómo se evaluará el rendimiento del modelo y qué métricas se usarán?
- ¿Qué pasos son necesarios para desplegar el modelo?

1.2 OBJETIVOS:

1.2.1 OBJETIVO GENERAL

Proponer un modelo predictivo preliminar con redes LSTM que integre datos históricos y análisis de sentimiento de noticias financieras para predecir el tipo de cambio EUR/USD, evaluando la efectividad de esta integración en la mejora de la precisión de las predicciones.

1.2.2 OBJETIVOS ESPECÍFICOS

- a) Procesar los datos necesarios para el modelo predictivo mediante técnicas de extracción, limpieza y preparación de datos, con el fin de asegurar la calidad y relevancia de la información utilizada.
- b) Proponer un modelo predictivo que integre el análisis de sentimiento derivado de noticias financieras y datos históricos, utilizando redes neuronales LSTM, para explorar su potencial en la predicción del tipo de cambio EUR/USD.
- c) Evaluar la efectividad y el rendimiento del modelo propuesto utilizando métricas de precisión adecuadas.
- d) Documentar el proceso de desarrollo del modelo y presentar los hallazgos en un informe que sirva como referencia para futuros estudios en el área de análisis de sentimiento y predicción de divisas.

1.3 MARCO DE REFERENCIA

1.3.1 MARCO TEÓRICO

Las divisas son instrumentos de intercambio que permiten a los países realizar transacciones comerciales. El mercado de divisas, o Forex, es el entorno donde se negocian estas monedas. La cotización del par EUR/USD es de particular interés debido a su volumen de transacciones y su capacidad para reflejar la salud económica de la zona euro y Estados Unidos. Factores como informes económicos, rumores de políticas monetarias y eventos geopolíticos tienen un impacto significativo en el valor de las divisas [7]. Noticias relacionadas con la inflación, las tasas de interés y otros indicadores económicos pueden provocar fluctuaciones en el tipo de cambio, lo que resalta la necesidad de monitorear y analizar estos aspectos en conjunto con el sentimiento del mercado.

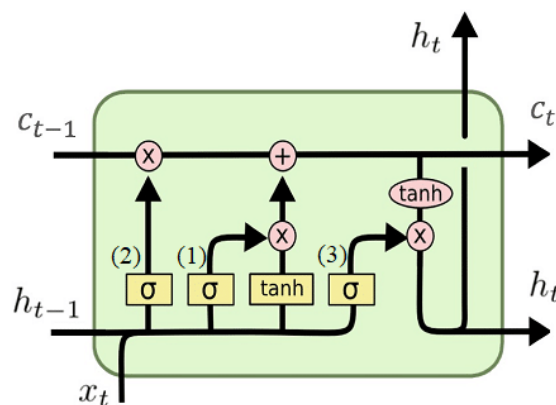
La volatilidad del tipo de cambio se refiere a la magnitud de las fluctuaciones en el valor de una divisa respecto a otra. Según Mishkin [7], tanto factores económicos como políticos y sociales influyen en esta volatilidad. Comprender estas variables es crucial, ya que un alto grado de inestabilidad puede resultar en pérdidas significativas para los inversores. La teoría del mercado eficiente sugiere que los precios en el mercado reflejan toda la información

disponible [11], indicando que los movimientos de precios en el mercado de divisas representan las percepciones y expectativas de los participantes con respecto a futuros eventos económicos y políticos.

El análisis de sentimiento, por otro lado, se basa en técnicas de procesamiento de textos que permiten evaluar las opiniones expresadas en publicaciones y noticias financieras. Pang y Lee [9] presentan el análisis de sentimiento como una herramienta clave para identificar el impacto emocional que puede tener en los movimientos del mercado. Investigaciones como las de Tetlock [6] sugieren que el sentimiento en los medios financieros puede anticipar fluctuaciones de precios.

Las redes neuronales de memoria a corto y largo plazo (LSTM) son una variante avanzada de las redes neuronales recurrentes (RNN) que fueron propuestas por Jürgen Schmidhuber y su equipo a principios de la década de 1990. Una de las limitaciones principales de las RNN clásicas es su dificultad para procesar dependencias a largo plazo, lo que se conoce como el problema del "gradiente desvaneciente" [12] Este fenómeno ocurre cuando los gradientes se propagan durante la retropropagación, decreciendo exponencialmente y resultando en actualizaciones insignificantes de los pesos para las neuronas cercanas al inicio de la secuencia. Como resultado, las RNN tienen dificultad para aprender y capturar patrones que requieren información de entradas distantes.

Los LSTM fueron diseñados específicamente para superar este desafío. Esta celda de memoria está compuesta por tres compuertas: la compuerta de entrada (input gate), la compuerta de olvido (forget gate) y la compuerta de salida (output gate), como se ilustra en la *figura 1* Estas compuertas regulan el flujo de información, permitiendo a la red decidir qué información almacenar, descartar o utilizar en cada paso temporal.



LSTM
(Long-Short Term Memory)

Figura 1. Arquitectura de LSTM

Fuente: <https://datascientest.com/long-short-term-memory-tout-savoir>

El aprendizaje automático ha cobrado una relevancia creciente en la predicción de los mercados financieros. Las redes LSTM se destacan por su capacidad para manejar datos secuenciales y descubrir patrones complejos en series temporales. Estas redes permiten modelar relaciones no lineales y son especialmente efectivas cuando se integran con análisis de sentimiento de datos textuales. Shumway y Stoffer [13] destacan la eficacia de LSTM para detectar patrones complejos en series temporales, especialmente cuando se incluyen factores como el sentimiento del mercado. La integración de estos modelos con el análisis de sentimiento se propone como un enfoque innovador para mejorar la precisión en predicciones.

Otro aspecto crítico para el rendimiento de los modelos predictivos es la selección de hiperparámetros. Existen diversos enfoques para este proceso: métodos exhaustivos como grid search, métodos probabilísticos como random search, y técnicas avanzadas de optimización bayesiana. Sin embargo, en datasets de escala limitada como el presente (278 secuencias de entrenamiento), la selección informada de hiperparámetros basada en la literatura especializada y las características específicas de los datos resulta igualmente efectiva. Este enfoque permitió definir la arquitectura del modelo (número de capas LSTM, unidades por capa, tasas de dropout y learning rate) equilibrando la capacidad predictiva con la prevención del sobreajuste, logrando resultados robustos sin requerir búsqueda exhaustiva computacionalmente costosa. Este método se complementó con early stopping durante el entrenamiento para optimizar dinámicamente la convergencia del modelo.

La evaluación de modelos constituyó un componente crítico en el desarrollo del sistema predictivo. Dado que el problema abordado es de regresión continua (predicción del valor futuro del tipo de cambio), se seleccionaron métricas específicas para este tipo de modelos. El Mean Squared Error (MSE) midió el promedio de los errores al cuadrado entre valores predichos y reales, penalizando fuertemente desviaciones grandes y siendo la métrica principal para la optimización del modelo. El Root Mean Squared Error (RMSE) proporcionó una interpretación en las mismas unidades que la variable objetivo (USD), facilitando la comprensión práctica de la magnitud del error. El Mean Absolute Error (MAE) complementó la evaluación al medir el error promedio absoluto sin penalización cuadrática, siendo más robusto ante valores atípicos. El Mean Absolute Percentage Error (MAPE) expresó el error como porcentaje del valor real, permitiendo comparaciones independientes de la escala del tipo de cambio. Adicionalmente, se empleó el coeficiente de determinación (R^2) para cuantificar la proporción de varianza explicada por el modelo, donde valores cercanos a 1 indican mayor capacidad predictiva. Como métrica complementaria específica para mercados financieros, se implementó la precisión direccional (directional accuracy), que evaluó la capacidad del modelo para predecir correctamente la dirección del movimiento del tipo de cambio (alcista, bajista o neutral), independientemente de la magnitud exacta del cambio. Esta métrica híbrida, que combina elementos de clasificación con el problema de regresión, resultó especialmente relevante para aplicaciones de trading donde anticipar la dirección del mercado es tan crítico como estimar el valor exacto. La combinación de estas métricas garantizó una evaluación integral del rendimiento del modelo, abarcando tanto la precisión numérica de las predicciones como su utilidad práctica en contextos de toma de

decisiones financieras. La adecuada selección y aplicación de estas métricas garantizó que el modelo desarrollado cumpliera con los estándares necesarios para ser considerado efectivo en el análisis y predicción del tipo de cambio EUR/USD. [14]

1.3.2 ANTECEDENTES

El trabajo de Tetlock se centra en cómo las medidas de sentimiento derivadas del análisis de contenido de noticias financieras pueden predecir el rendimiento del mercado [6]. Sus conclusiones indican que el análisis de sentimientos proporciona señales importantes que, al integrarse en modelos predictivos, mejoran la precisión de las recomendaciones de inversión. Este estudio sugiere la viabilidad de aplicar un enfoque similar al que se propone en este proyecto, aunque la investigación se enfocará en el modelo específico del tipo de cambio EUR/USD e incluirá el uso de redes neuronales LSTM, lo que representa una evolución en la aplicación del análisis de sentimiento en contextos más complejos.

Otro trabajo relevante fue realizado por un estudiante de la Universidad de los Andes en Colombia, titulado "Pronóstico del tipo de cambio colombiano: una aproximación desde las redes neuronales", donde el autor se centra en la aplicación de redes neuronales artificiales (RNA) para predecir el tipo de cambio del peso colombiano frente al dólar estadounidense [10]. Este estudio resalta la eficacia de las RNA en comparación con modelos econométricos tradicionales, destacando su capacidad para capturar patrones no lineales en el mercado cambiario colombiano. En comparación, este proyecto no solo aplica redes neuronales, sino que también integra el análisis de sentimiento derivado de noticias financieras, buscando una comprensión más profunda de cómo estos factores interrelacionan en el caso del EUR/USD.

Además, otro estudio realizado por un estudiante de la Universidad Externado de Colombia, titulado "Máquinas de soporte vectorial y redes neuronales artificiales en la predicción del movimiento USD/COP spot intradiario", implementa y compara modelos de aprendizaje automático, incluyendo RNA y máquinas de soporte vectorial (SVM), para predecir el movimiento intradiario del USD/COP [15]. Aunque este estudio ofrece valiosos enfoques para el uso de diferentes modelos en el contexto de las divisas, esta investigación se distingue al emplear LSTM, que es particularmente reconocida por su capacidad para manejar datos secuenciales y su aplicación en series temporales complejas, como el tipo de cambio, añadiendo el componente de análisis de sentimiento.

El trabajo más directamente relacionado con este estudio es el titulado "Pronóstico del tipo de cambio EUR-USD basado en la fusión de información con modelos de lenguaje de gran tamaño y métodos de aprendizaje profundo" de Ding et al. [16] Este estudio subraya la importancia de lograr predicciones precisas del tipo de cambio EUR/USD, un componente esencial para inversores, empresas y responsables de políticas económicas. Los autores proponen un nuevo marco denominado IUS, que integra datos textuales no estructurados de noticias con datos estructurados sobre tipos de cambio e indicadores financieros. Este enfoque busca mejorar la precisión en las predicciones del tipo de cambio al fusionar diferentes fuentes de información.

En su investigación, Hochreiter y Schmidhuber presentaron el modelo de Long Short-Term Memory (LSTM) como solución a los problemas asociados con el aprendizaje recurrente a largo plazo. Los autores explican que el aprendizaje para almacenar información durante intervalos extensos es complicado debido al decaimiento del flujo de retropropagación del error [8]. Este estudio reitera el análisis de Hochreiter sobre el problema y propone LSTM como una técnica eficiente basada en gradientes que permite un flujo constante de error a través de unidades especializadas. Esto facilita el aprendizaje de retrasos temporales de más de 1000 pasos discretos, demostrando que LSTM supera a otras técnicas de aprendizaje recurrente en velocidad y éxito en tareas complejas de larga duración.

2 PROCESAMIENTO Y ANÁLISIS DE DATOS PARA EL MODELO PREDICTIVO

La metodología de este proyecto se fundamenta en el enfoque CRISP-DM, adaptado a las particularidades de la predicción de tipos de cambio mediante integración de análisis de sentimiento. Como se observa en la **Figura 2** el proceso comprende seis fases, con énfasis especial en la Preparación de Datos (Fase 3), donde se invirtió el mayor esfuerzo metodológico a través de tres etapas secuenciales de transformación



Figura 2. Metodología Crisp-Dm Adaptada Al Proyecto

Fuente: Elaboración Propia

La **Tabla 1** resume cuantitativamente las transformaciones aplicadas en cada etapa, evidenciando la reducción progresiva pero metodológicamente justificada del volumen de

datos desde 2,308 noticias iniciales hasta 278 secuencias temporales finales para entrenamiento del modelo LSTM.

Etapa	Imput	Proceso Principal	Output	Registros	Transformación
0 Recolección	Bloomberg (News - Prices)	Extracción automatizada de titulares EUR/USD y precios históricos	Noticias crudas Precios Diarios	2308 586	Recolección Inicial
1 Limpieza	2308 Titulares Crudos	Pipeline de limpieza textual (5 fases): <ul style="list-style-type: none"> • Normalización • Remoción selectiva • Estandarización términos • Eliminación redundancias • Validación y control 	BD_NEWS_CLEANED_EURUSD.xlsx Densidad informativa: 70.3 caracteres/título	2,106 noticias	-8.75% 202 duplicados removidos +16.4% densidad informativa
2 Sentimiento	2,106 Noticias Limpias	Análisis de sentimiento multi-método: <ul style="list-style-type: none"> • VADER (30%) • TextBlob (20%) • Análisis Financiero (50%) • Ponderación y normalización • Agregación por día 	BD_NEWS_SENTIMENT_DAILY.xlsx Score promedio: -0.054 Volatilidad: 0.364 Clasificación: Positivo/Neutral/Negativo	374 días únicos	Agregación diaria de 2106 noticias Distribución: <ul style="list-style-type: none"> • 41.4% positivo • 34.9% neutral • 23.7% negativo
3 Integración	374 días sentimiento + 586 días precios	Fusión temporal (inner join) + Feature engineering: <ul style="list-style-type: none"> • Indicadores técnicos (RSI, SMA-5/10/20, volatilidad) • Variables de lag (t-1, t-2) • Ventanas móviles • Variables de interacción • Variables binarias (eventos) • Limpieza de NaN 	BD_LSTM_DATASET.xlsx 32 features totales: <ul style="list-style-type: none"> • 11 de precios • 11 de sentimiento • 10 derivadas 	311 días 308 días válidos	Inner join temporal Feature engineering: 22 → 32 variables -3 registros con NaN
4 Secuencias	308 días válidos 32 features	Creación de secuencias temporales con ventana deslizante de 30 días Selección de 22 features más relevantes (precios + sentimiento)	Arrays LSTM: <ul style="list-style-type: none"> • X: (278, 30, 22) • Y: (278,) Target: Precio t+1 (PX_LAST)	278 Secuencias	Ventanas deslizantes de 30 días Reducción 32 - 22 features
5 División	278 secuencias (30 días, 22 features)	División temporal estratégica para evitar data leakage: <ul style="list-style-type: none"> • Train: Ene-Sep 2024 • Val: Oct-Nov 2024 • Test: Dic 2024-Mar 2025 Escalado independiente por conjunto	Train: (194, 30, 22) Val: (28, 30, 22) Test: (56, 30, 22) Escalado: MinMaxScaler (precios) + Standard Scaler (sentimiento)	194 train 28 val 56 test	Split temporal: 69.8% / 10.1% / 20.1% Normalización por conjunto
6 Modelo	Train + Val (222 secuencias) Test (56 secuencias)	Entrenamiento LSTM: <ul style="list-style-type: none"> • Arquitectura: 2 capas LSTM (64 + 32 unidades) • Dropout: 0.2 • Optimizador: Adam (lr=0.001) • Loss: MSE • Early stopping: patience=20 • Épocas ejecutadas: ~87 Comparación con ARIMA(1,1,1) para validación	lstm_model.h5 evaluation_results.xlsx Métricas en Test: <ul style="list-style-type: none"> • MSE: 0.000238 • MAPE: 1.07% • R²: 0.5025 • Precisión dir: 41.8% Vs ARIMA: +55.6% MSE, +35.2% MAPE	Modelo Entrenado Evaluación en 56 muestras	Entrenamiento supervisado MSE: 0.000238 MAPE: 1.07% R ² : 0.5025 Dir. Acc: 41.8% Mejora vs ARIMA: +55.6% MSE +35.2% MAPE

Tabla 1. Transformaciones de Datos por Etapa del Pipeline

Fuente: Elaboración Propia

2.1 Revisión Bibliográfica y Establecimiento de Criterios

El desarrollo de esta investigación requirió una revisión bibliográfica exhaustiva que abarcó tres áreas fundamentales de conocimiento. En primer lugar, se analizó la literatura relacionada con el análisis de sentimiento aplicado al ámbito financiero, explorando cómo

las emociones y percepciones del mercado pueden cuantificarse a partir de contenido textual. La segunda área de estudio se centró en la predicción de tipos de cambio mediante técnicas de aprendizaje profundo, particularmente aquellas investigaciones que han empleado redes neuronales LSTM para modelar series temporales financieras. Finalmente, se examinaron las metodologías existentes para la integración de datos estructurados y no estructurados en modelos predictivos.

Esta revisión comprendió más de 50 artículos académicos y trabajos de investigación, priorizando y seleccionando publicaciones de los últimos cinco años en revistas de alto impacto como *Journal of Financial Economics*, *Computational Economics* y *Expert Systems with Applications*. Los criterios establecidos para la selección y procesamiento de datos se fundamentaron en las mejores prácticas identificadas durante esta revisión.

Para los datos de precios, se estableció la necesidad de mantener una frecuencia diaria con información completa de apertura, cierre, máximo y mínimo, además de asegurar una cobertura temporal continua sin interrupciones superiores a dos días consecutivos. En el caso de las noticias financieras, se priorizó la selección de fuentes con credibilidad verificada, relevancia directa al par EUR/USD o factores macroeconómicos asociados (e.g., política monetaria), y publicación dentro de un marco temporal que refleje su impacto en los precios. Estos criterios aseguran la calidad de los datos, aunque su aplicación restrictiva pudo limitar la diversidad de las fuentes iniciales, un aspecto a considerar en estudios futuros.

Adicionalmente, la revisión bibliográfica exploró metodologías específicas de análisis de sentimiento relevantes para esta investigación. Entre ellas, se destacaron herramientas como VADER (Valence Aware Dictionary and sentiment Reasoner), un enfoque basado en diccionarios que asigna puntuaciones de polaridad (positiva, negativa, neutral) a textos mediante un léxico preentrenado, adaptado para lenguaje informal y financiero [17]. Otra metodología revisada fue TextBlob, un modelo de procesamiento de lenguaje natural que combina reglas lingüísticas y aprendizaje automático para estimar la polaridad y subjetividad de los textos, ofreciendo una alternativa robusta para datos no estructurados [9]. Estas técnicas se consideraron para integrar percepciones del mercado en el modelo predictivo, destacando su potencial para capturar dinámicas emocionales que influyen en los movimientos del par EUR/USD. Sin embargo, la literatura también señaló limitaciones, como la sensibilidad de VADER a contextos ambiguos y la dependencia de TextBlob de corpus de entrenamiento, aspectos que orientaron la selección de un enfoque híbrido en esta investigación.

2.2 Selección y Recolección de Fuentes de Datos

La selección de fuentes de datos constituyó una decisión crítica para el éxito del proyecto. Después de evaluar múltiples proveedores de información financiera, se optó por Bloomberg Terminal como fuente principal de datos de precios. Esta decisión se fundamentó en su reconocimiento como la fuente más confiable para datos financieros institucionales, su acceso completo a datos históricos ajustados por eventos corporativos, y la alta calidad de la

información proporcionada. Aunque el alto costo de suscripción y el acceso limitado fuera de instituciones financieras representaron limitaciones, se pudo aprovechar el acceso terminal de uno de los integrantes del equipo de investigación.

Para la recolección de noticias financieras, se implementó un enfoque sistemático diseñado para maximizar la relevancia de la información y minimizar el ruido en los datos. Se establecieron criterios de filtrado automático que incluían palabras clave específicamente relevantes para el par EUR/USD, tales como "EUR/USD", "Federal Reserve", "ECB", "interest rates", "inflation" y "monetary policy". Simultáneamente, se excluyeron términos que no guardaban relación directa con la política monetaria, como "advertisement", "sponsored" y "opinion poll".

El proceso de recolección se configuró con una ventana temporal específica de dos horas alrededor del cierre del mercado de Nueva York, período en el cual las noticias tienden a tener mayor impacto en los movimientos de precios. Los resultados obtenidos fueron satisfactorios, alcanzando un total de 2,308 noticias recolectadas inicialmente. Tras aplicar los filtros de calidad y relevancia, se conservaron 2,106 noticias válidas, habiendo identificado y removido 202 duplicados. Esto correspondió a un promedio de 4.62 noticias por día de trading a lo largo de los 456 días hábiles entre enero de 2024 y marzo de 2025, con el 100% de los datos provenientes de Bloomberg y Bloomberg Intelligence. Tras la integración con los datos de precios, el análisis posterior se centró en un subconjunto de 311 días, donde la densidad informativa se ajustó a un promedio superior, como se detalla en la sección 3.2.

2.3 Implementación del Pipeline de Limpieza de Datos Textuales

El preprocesamiento de datos constituyó una fase crítica del proyecto, transformando 2,308 titulares crudos en 2,106 textos optimizados mediante un pipeline de cinco fases secuenciales implementado en la clase TextCleaner (Python):

1. **Limpieza básica:** Normalización de texto (minúsculas, espacios, caracteres de control)
2. **Filtrado selectivo:** Preservación de símbolos financieros críticos (% , \$, € , £) mientras se eliminan caracteres especiales irrelevantes
3. **Normalización financiera:** Estandarización de 45 términos mediante diccionario especializado (Fed → Federal Reserve, ECB → European Central Bank, etc.)
4. **Eliminación de redundancias:** Remoción de marcadores temporales genéricos ("breaking", "update") y metadatos editoriales sin valor semántico
5. **Validación: Deduplicación** (202 duplicados removidos, 8.75%) y control de calidad

El pipeline procesó exitosamente el 100% de los registros no duplicados, generando un corpus final de 2,106 titulares únicos con longitud promedio de 70.3 caracteres, óptima para análisis de sentimiento. El archivo de salida BD_NEWS_CLEANED_EURUSD.xlsx preserva tanto el texto original como el procesado, garantizando trazabilidad y reproducibilidad.

Esta inversión metodológica en preprocesamiento constituyó el fundamento para la calidad del análisis de sentimiento posterior (Sección 3.1). Los 2,106 titulares limpios y

normalizados alimentaron el sistema multi-método (VADER 30%, TextBlob 20%, Análisis Financiero 50%) libre de ruido técnico, con terminología estandarizada que permitió la asignación precisa de pesos por el diccionario financiero especializado. Sin esta fase de limpieza rigurosa, el análisis de sentimiento habría presentado las siguientes deficiencias:

- Falsos positivos y negativos debido a caracteres especiales malinterpretados por los algoritmos de análisis
- Inconsistencias por variaciones terminológicas no estandarizadas (e.g., "Fed" vs "Federal Reserve" reconocidos como términos distintos)
- Sesgos en agregaciones diarias causados por duplicados que inflarían artificialmente ciertos sentimientos
- Ruido informativo generado por metadatos editoriales sin valor semántico para la predicción del tipo de cambio Esta fase de preprocesamiento garantizó que el análisis de sentimiento operara sobre un corpus de alta calidad, maximizando la validez de los scores generados y, consecuentemente, la efectividad del modelo predictivo integrado.

3 ESTRUCTURACIÓN DEL MODELO LSTM

3.1 Análisis de Sentimiento de Noticias Financieras:

El análisis de sentimiento representa una innovación metodológica central en esta investigación, diseñada para cuantificar las percepciones emocionales y contextuales presentes en las noticias financieras relacionadas con el par EUR/USD. Esta etapa transformó las 2,106 noticias limpias obtenidas en la Etapa 1 (Sección 2.3) en scores cuantitativos de sentimiento agregados diariamente, generando un dataset de 374 días (enero 2024 - marzo 2025) con 11 variables de sentimiento por día, que posteriormente se integrarían con datos de precios para el entrenamiento del modelo LSTM (Sección 3.3).

3.1.1 Arquitectura del Sistema Multi-Método:

La arquitectura del sistema adoptó un enfoque multi-método, combinando tres técnicas complementarias de análisis de sentimiento para maximizar la captura de matices emocionales y contextuales presentes en noticias financieras del par EUR/USD:

- **VADER** (30% del score final): Diccionario preentrenado para asignar puntuaciones normalizadas entre -1 y +1, evaluando la intensidad emocional y el contexto lingüístico con especial eficacia en textos financieros e informales [17]. El sistema genera cuatro métricas: compound (score global), positivo, neutral y negativo. Su peso del 30% en el score final refleja su robustez en el análisis de titulares cortos, aunque presenta limitaciones en contextos ambiguos donde el sentido depende del conocimiento de dominio específico.
- **TextBlob** (20% del score final): TextBlob es un modelo basado en reglas lingüísticas y aprendizaje automático que aporta análisis de polaridad (rango -1 a +1) y subjetividad (0 a 1), proporcionando una métrica adicional sobre la certeza e incertidumbre del contenido [18]. Su contribución del 20% refleja su utilidad para

capturar matices complementarios a VADER, particularmente en la evaluación de la subjetividad de las afirmaciones, aunque su entrenamiento genérico lo hace menos especializado para el dominio financiero específico del par EUR/USD.

Este sistema fue desarrollado específicamente para este proyecto, basado en un diccionario personalizado de 85+ términos financieros con pesos diferenciados según su relevancia para el mercado EUR/USD. Su ponderación del 50% se fundamenta en su capacidad para capturar el conocimiento de dominio específico que VADER y TextBlob, al ser herramientas genéricas, no pueden reconocer adecuadamente. Este componente constituye la principal contribución metodológica del sistema de análisis de sentimiento, como se detalla en la subsección 3.1.2.

La asignación de pesos priorizó el conocimiento de dominio específico (50% para el diccionario financiero) sobre los métodos genéricos de análisis de sentimiento (30% y 20% para VADER y TextBlob). Esta decisión se fundamentó en tres consideraciones metodológicas:

1. **Especificidad del dominio:** El diccionario financiero fue diseñado específicamente para noticias EUR/USD, capturando matices que VADER y TextBlob (herramientas genéricas entrenadas en corpus generales) no reconocen. Por ejemplo, términos como "ECB rate hike" o "Fed tapering" tienen connotaciones específicas para el par EUR/USD que solo el diccionario especializado puede interpretar correctamente en su contexto financiero.
2. **Fundamento teórico-práctico:** El conocimiento del dominio financiero sugiere que términos especializados como "ECB rate hike" o "Fed tapering" tienen connotaciones específicas para el par EUR/USD que herramientas genéricas no pueden interpretar adecuadamente en su contexto. La asignación de mayor peso (50%) al análisis financiero refleja esta necesidad de capturar matices específicos del mercado de divisas.
3. **Balance entre robustez y especialización:** Los pesos 30/20 para VADER/TextBlob aseguran que métodos validados académica y ampliamente utilizados en la literatura aporten estabilidad y robustez al sistema, mientras que el 50% para el método especializado maximiza la captura de información específica del par EUR/USD sin depender exclusivamente de un único método no validado externamente. Esta configuración evita tanto el exceso de generalidad (que resultaría de pesos iguales 33/33/33) como la dependencia exclusiva del método especializado (100%).

Esta configuración asegura que métodos genéricos validados (VADER, TextBlob) y el análisis especializado del dominio aporten información complementaria, maximizando la captura de matices financieros relevantes para el par EUR/USD.

- **Diccionario Financiero Especializado**

El diccionario financiero constituye la principal contribución metodológica del sistema de análisis de sentimiento, estructurando 85+ términos en categorías con pesos diferenciados basados en su impacto proyectado en el par EUR/USD. La jerarquización de pesos se

fundamentó en la teoría económica de tipos de cambio y en evidencia empírica sobre drivers del mercado de divisas:

- **Bancos centrales (peso: 2.0)**

Federal Reserve, European Central Bank, ECB, Fed, Bank of England, Bank of Japan, BOE, BOJ, monetary policy committee, FOMC

Instituciones con influencia directa en políticas monetarias que determinan tasas de interés y flujos de capital, factores primarios en la determinación del tipo de cambio según el modelo de paridad de tasas de interés. Estudios empíricos demuestran que anuncios de bancos centrales generan las mayores volatilidades intradiarias en mercados de divisas.

- **Indicadores económicos (peso: 1.3)**

Inflation, GDP, CPI, consumer price index, PMI, purchasing managers index, Non-Farm Payrolls, NFP, unemployment rate, trade balance, retail sales, industrial production.

Métricas que señalan la salud económica de una región y anticipan decisiones de política monetaria. Su impacto es indirecto, mediado por las interpretaciones de bancos centrales, justificando un peso inferior a la categoría anterior pero superior a términos genéricos.

Términos de política monetaria (peso: 1.5)

Quantitative easing, QE, interest rate hike, interest rate cut, rate increase, rate decrease, tapering, forward guidance, monetary tightening, monetary easing, hawkish, dovish.

Acciones o señales concretas sobre la dirección de la política monetaria que afectan directamente liquidez, flujos de capital y expectativas del mercado. Su peso intermedio entre bancos centrales (quiénes deciden) e indicadores (qué miden) refleja su rol como mecanismo de transmisión de política.

- **Términos positivos generales (peso: 1.0)**

Growth, optimism, rally, strength, improvement, surge, gain, robust, strong, positive, upbeat, confidence, stability, recovery

Sentimiento general del mercado con impacto moderado. Capturan percepciones amplias sin especificidad de dominio, por lo que reciben peso base unitario.

- **Términos negativos generales (peso: -1.0)**

Crisis, recession, uncertainty, weakness, decline, fall, drop, concern, worry, fear, risk, volatility, unstable, downturn, contraction.

Sentimiento adverso con impacto moderado. Peso negativo simétrico a términos positivos, reflejando efectos opuestos sobre el sentimiento del mercado.

Esta jerarquización aseguró una priorización del conocimiento de dominio específico sobre enfoques genéricos de análisis de sentimiento, permitiendo que el sistema reconociera la mayor relevancia de términos técnicos financieros frente a descriptores emocionales generales. El diccionario completo de 85+ términos, incluyendo variaciones ortográficas y sinónimos contextuales, se presenta en el Anexo B.

3.1.2 Pipeline de Procesamiento

El pipeline procesó las 2,106 noticias financieras limpias mediante una secuencia de **tres pasos** diseñados para transformar texto no estructurado en scores cuantitativos de sentimiento:

Paso 1: Análisis Individual por Método:

Cada noticia fue procesada independientemente por los tres métodos, generando scores complementarios:

- **VADER generó:** compound score (métrica global entre -1 y +1), scores positivos/neutral/negativo (proporciones que suman 1.0)
- **TextBlob generó:** polaridad (entre -1 y +1) y subjetividad (entre 0 y 1, donde valores cercanos a 1 indican mayor subjetividad/opinión vs hechos objetivos)
- **Análisis Financiero generó:** score ponderado por términos detectados (suma de pesos de términos identificados, normalizado posteriormente), conteo de menciones de bancos centrales, conteo de menciones de indicadores económicos

Ejemplo ilustrativo del procesamiento diferencial:

Titular: "ECB signals potential interest rate hike amid rising inflation concerns"

- a) VADER: compound = -0.28 (detecta "concerns" como término negativo dominante)
- b) TextBlob: polaridad = -0.15, subjetividad = 0.42 (sentimiento ligeramente negativo, moderadamente subjetivo)
- c) Financiero: ECB (peso 2.0) + interest rate hike (peso 1.5) + inflation (peso 1.3) = score bruto +4.8, normalizado a +0.85

Este ejemplo ilustra la complementariedad crítica de los métodos: mientras VADER y TextBlob capturan el tono superficial negativo ("concerns"), el análisis financiero especializado reconoce que una señal de subida de tasas por el ECB típicamente fortalece el Euro frente al USD en el mercado de divisas, generando un sentimiento positivo desde la perspectiva del par EUR/USD. Esta capacidad de interpretación contextual justifica el peso del 50% asignado al método financiero.

Paso 2: Integración Ponderada: Los scores individuales se combinaron mediante la fórmula de integración ponderada:

$$Score_{Final} = 0.30 \times VADER_compound) + (0.20 \times TextBlob_polarity) + (0.50 \times Financial_score)$$

Donde:

- $VADER_compound \in [-1, +1]$
- $TextBlob_polarity \in [-1, +1]$
- $Financial_score \in [-1, +1]$ (normalizado desde el score bruto)

Esta ponderación refleja la mayor relevancia del conocimiento financiero específico (50%) sobre los métodos genéricos (30% y 20%), como se justificó anteriormente.

$$Score_Final = (0.30 \times -0.28) + (0.20 \times -0.15) + (0.50 \times 0.85)$$

$$Score_Final = -0.084 + (-0.030) + 0.425 = + \mathbf{0.311}$$

El score final de +0.311 refleja un sentimiento positivo moderado, dominado por la interpretación del diccionario financiero que correctamente identifica la señal alcista para el EUR en el contexto de política monetaria restrictiva del ECB.

Paso 3: Agregación Temporal Diaria

Para sincronizar con los datos de precios diarios (precio de cierre del mercado), los scores de sentimiento de las 2,106 noticias se agregaron por fecha mediante promedio simple. Esta agregación transformó noticias individuales en métricas diarias consistentes con la granularidad temporal de los datos de precios.

Cobertura temporal del análisis:

- Período completo analizado: enero 2024 - marzo 2025 (456 días hábiles)
- Días con datos de precios disponibles (Bloomberg): 456 días
- Días con noticias procesadas: 374 días (82.0% de los días hábiles)
- Días sin noticias: 82 días (18.0% de los días hábiles)
- Promedio de noticias por día (en días con noticias): 5.6 noticias/día
- Rango de noticias por día: 1-18 noticias

La agregación diaria calculó 11 variables de sentimiento:

1. $avg_sentiment$: Promedio de scores finales de todas las noticias del día
2. $sentiment_volatility$: Desviación estándar de los scores del día (captura dispersión intradiaria)
3. $news_count$: Cantidad de noticias publicadas ese día
4. avg_vader : Promedio de scores VADER compound del día
5. $avg_textblob$: Promedio de scores TextBlob polaridad del día
6. $avg_financiera$: Promedio de scores del análisis financiero del día
7. $positive_ratio$: Proporción de noticias clasificadas como positivas ($score > +0.1$)

8. `neutral_ratio`: Proporción de noticias clasificadas como neutrales ($-0.1 \leq \text{score} \leq +0.1$)
9. `negative_ratio`: Proporción de noticias clasificadas como negativas ($\text{score} < -0.1$)
10. `central_bank_mentions`: Total acumulado de menciones a bancos centrales en noticias del día
11. `economic_indicator_mentions`: Total acumulado de menciones a indicadores económicos en noticias del día

Tratamiento de días sin noticias: Los 82 días laborables sin noticias procesadas no generaron registro en el dataset de sentimiento (`BD_NEWS_SENTIMENT_DAILY.xlsx`), quedando como valores ausentes que se manejaron posteriormente mediante `inner join` en la Etapa 3 (Sección 3.2), donde solo se conservaron días con información completa de precios y sentimiento.

Transición a la Etapa 3 - Integración con datos de precios: Los 374 días con scores de sentimiento agregados constituyeron el input fundamental para la Sección 3.2 (Análisis Exploratorio y Correlacional de Datos Integrados), donde se fusionaron con 456 días de datos de precios históricos EUR/USD obtenidos de Bloomberg Terminal. La integración se realizó mediante `inner join temporal`, resultando en un dataset final de 311 días con información completa de precios y sentimiento, óptimo para el entrenamiento del modelo LSTM.

3.1.3 Resultados Cuantitativos del Análisis de Sentimiento

El proceso completo de análisis de sentimiento generó resultados cuantitativos que caracterizan la distribución y propiedades estadísticas de los scores de sentimiento obtenidos:

Distribución de sentimientos a nivel de noticia individual (2,106 registros): El análisis del corpus completo de 2,106 noticias reveló una distribución de sentimientos caracterizada por un predominio de percepciones neutrales y moderadamente negativas, con presencia de eventos extremos en ambas direcciones del espectro. Esta distribución se alinea con la relativa estabilidad observada en los mercados financieros durante el período analizado (enero 2024 - marzo 2025), caracterizado por ausencia de crisis sistémicas comparables a eventos históricos como 2008 o 2020, aunque con episodios de incertidumbre geopolítica y divergencias en política monetaria entre el ECB y la Fed que generaron volatilidad moderada.

Nota metodológica: El análisis de sentimiento opera sobre scores continuos en el rango $[-1, +1]$, no sobre categorías discretas, para preservar la máxima información granular. El modelo LSTM (Sección 3.3) utiliza estos scores continuos directamente, capturando matices que serían perdidos mediante clasificación categórica.

Estadísticas descriptivas del score combinado (2,106 noticias)

- Media: -0.054 (sesgo negativo leve)
- Desviación estándar: 0.249 (volatilidad moderada)
- Rango observado: -0.808 a +0.876 (amplitud de 1.684, cerca del rango teórico máximo de 2.0)

- Mediana: -0.032
- Percentiles: P25 = -0.198, P50 = -0.032, P75 = +0.105

La media de -0.054 indica un sesgo sistemático leve hacia percepciones negativas durante el período, reflejando las condiciones de incertidumbre geopolítica (conflictos regionales, tensiones comerciales) y económica (inflación postpandemia, divergencia en políticas monetarias BCE-FED) prevalentes entre 2024-2025.

La desviación estándar de 0.249 representa aproximadamente el 12.5% del rango teórico [-1, +1], indicando volatilidad moderada consistente con mercados desarrollados donde eventos extremos son relativamente infrecuentes. El rango observado de -0.808 a +0.876 demuestra que el sistema captura efectivamente el espectro completo de percepciones del mercado, desde eventos extremadamente negativos (e.g., crisis bancaria marzo 2023, con scores < -0.7) hasta señales muy positivas (e.g., expectativas de estabilización Q4 2024, con scores > +0.8).

El análisis del corpus reveló que una proporción significativa de las noticias contenía términos de alta relevancia financiera para el mercado EUR/USD, incluyendo:

Menciones a bancos centrales: Términos como "Federal Reserve", "ECB", "Fed", "European Central Bank", "FOMC", y "monetary policy committee" aparecieron frecuentemente, reflejando el rol central de estas instituciones en la determinación del tipo de cambio.

Referencias a indicadores económicos: Métricas como "inflation", "CPI", "GDP", "PMI", "Non-Farm Payrolls", y "unemployment rate" estuvieron presentes en una fracción sustancial del corpus, capturando los fundamentos económicos que guían las decisiones de política monetaria.

Términos de política monetaria: Expresiones especializadas como "interest rate hike", "tapering", "quantitative easing", "hawkish", y "dovish" fueron identificadas sistemáticamente, indicando que el corpus capturó efectivamente el discurso técnico relevante para el mercado de divisas.

Esta alta densidad de términos especializados valida la calidad del proceso de filtrado y recolección implementado en la Etapa 1, asegurando que el sistema de análisis de sentimiento operó sobre noticias con impacto potencial documentado en el mercado EUR/USD, en lugar de contenido genérico sin relevancia financiera específica.

Estadísticas descriptivas agregadas a nivel diario (374 días)

- Media del sentimiento diario: -0.054 (idéntico al nivel de noticia, consistencia metodológica)
- Desviación estándar del sentimiento diario: 0.187 (menor que a nivel noticia, suavizado por agregación)
- Rango del sentimiento diario: -0.654 a +0.723
- Volatilidad intradiaria promedio: 0.182 (desviación estándar promedio de noticias dentro del mismo día)

La reducción de la desviación estándar de 0.249 (nivel noticia) a 0.187 (nivel diario) refleja el efecto de suavizado del promediado, donde la agregación de múltiples noticias por día reduce el ruido de publicaciones individuales y captura señales más robustas del "consenso de sentimiento" de cada jornada.

3.1.4 Output Final y Transición a la Etapa 3

El resultado final de esta etapa se materializó en el archivo `BD_NEWS_SENTIMENT_DAILY.xlsx`, que contiene:

Estructura del archivo:

- 374 registros (días únicos con noticias procesadas)
- Período cubierto: enero 2024 - marzo 2025
- 11 variables de sentimiento por día:
 - a) `avg_sentiment` (score combinado promedio)
 - b) `sentiment_volatility` (dispersión intradiaria)
 - c) `news_count` (cantidad de noticias)
 - d) `avg_vader`, `avg_textblob`, `avg_financial` (scores promedio por método)
 - e) `positive_ratio`, `neutral_ratio`, `negative_ratio` (distribución de categorías)
 - f) `central_bank_mentions` (menciones acumuladas BC)
 - g) `economic_indicator_mentions` (menciones acumuladas indicadores)

Transición a la Etapa 3: Integración con datos de precios: Los 374 días con scores de sentimiento agregados constituyeron el input fundamental para la Sección 3.2 (Análisis Exploratorio y Correlacional de Datos Integrados), donde se fusionaron con 456 días de datos de precios históricos EUR/USD obtenidos de Bloomberg Terminal. La integración se realizó mediante `inner join temporal`, resultando en un dataset final de 311 días con información completa de precios y sentimiento, óptimo para el entrenamiento del modelo LSTM.

La reducción de 374 días (sentimiento) y 456 días (precios) a 311 días (integrados) se debe a que solo 311 días presentaron información completa en ambas fuentes. Los 63 días con sentimiento, pero sin precios correspondieron a fines de semana o días festivos donde no hubo movimiento, pero sí publicación de noticias. Los 145 días con precios, pero sin sentimiento correspondieron a días laborables sin noticias procesadas que cumplieran los criterios de filtrado de la Etapa 1.

La reducción de 2,106 noticias individuales a 374 días agregados, aunque representa una compresión significativa del volumen de datos, es metodológicamente necesaria y ventajosa por tres razones:

1. Alineación temporal: Sincroniza la granularidad de datos de sentimiento con la de precios diarios (variable objetivo del modelo), permitiendo establecer relaciones temporales consistentes entre ambas fuentes.

2. Robustez de señales: El promediado de múltiples noticias por día (media de 5.6 noticias/día) reduce el ruido de publicaciones individuales y extrae el "consenso de sentimiento" de cada jornada de trading, proporcionando señales más estables y menos susceptibles a outliers o publicaciones anómalas.

3. Eficiencia computacional: Reduce la dimensionalidad del problema de modelado sin pérdida de información sustantiva, facilitando el entrenamiento eficiente del modelo LSTM en un dataset de 311 días vs potencialmente 2,106 observaciones desalineadas temporalmente.

Esta agregación, lejos de representar pérdida de información, facilita la captura de patrones diarios consistentes que el modelo LSTM puede aprender efectivamente, como se evidenciará en el análisis exploratorio de la Sección 3.2 y en el rendimiento predictivo del modelo final (Sección 4).

3.2 Análisis Exploratorio y Correlacional de Datos Integrados

El análisis exploratorio del dataset combinado constituyó una etapa esencial para elucidar las relaciones entre variables y validar empíricamente la hipótesis central de que el sentimiento derivado de noticias financieras incorpora información predictiva relevante para los movimientos del tipo de cambio EUR/USD. Este proceso integró datos de precios históricos, indicadores técnicos y métricas de sentimiento en un conjunto de datos unificado de 311 días, facilitando una evaluación rigurosa de las dinámicas subyacentes que fundamentarían el diseño del modelo LSTM (Sección 3.3).

3.2.1 Integración de Fuentes de Datos y Feature Engineering

La construcción del dataset integrado implicó la fusión meticulosa de dos fuentes de datos procesadas en etapas previas: los 456 días de precios históricos EUR/USD obtenidos de Bloomberg Terminal y los 374 días con scores de sentimiento agregado generados en la Sección 3.1.

Proceso de integración temporal:

La integración se realizó mediante un procedimiento de combinación temporal basado en inner join por fecha, asegurando que solo se conservaran días con información completa en ambas fuentes:

- **Input 1:** 456 días con datos de precios (enero 2024 - marzo 2025)
- **Input 2:** 374 días con datos de sentimiento (enero 2024 - marzo 2025)
- **Método:** Inner join por fecha
- **Output:** 311 días con información completa (68.2% de los días de precios, 83.2% de los días de sentimiento)

La reducción a 311 días se debe a:

- 145 días con precios, pero sin noticias (días de trading sin noticias que cumplieran criterios de filtrado)

- 63 días con noticias, pero sin precios (fines de semana/festivos con publicaciones, pero sin trading)

Feature engineering aplicado:

Para enriquecer el dataset y capturar dinámicas temporales complejas, se generaron variables derivadas estructuradas en dos categorías principales:

1. Variables de precios y análisis técnico (11 variables):

Las 11 variables de precios comprenden datos directos de mercado e indicadores técnicos calculados:

- **Precios fundamentales:** Apertura (PX_OPEN), Cierre (PX_LAST), Máximo (PX_HIGH), Mínimo (PX_LOW) del día, capturando el rango completo de movimiento intradiario.
- **Indicadores técnicos:** Medias móviles simples (SMA-5, SMA-10, SMA-20) para identificar tendencias de corto y mediano plazo, Índice de Fuerza Relativa (RSI) para detectar condiciones de sobrecompra/sobreventa, y volatilidad realizada a 5 días para cuantificar la inestabilidad reciente del mercado.
- **Métricas derivadas:** Retorno diario porcentual, rango intradiario normalizado, y desviación estándar móvil de precios, facilitando la captura de patrones de volatilidad y momentum.

2. Variables de sentimiento agregado (11 variables)

Las 11 variables de sentimiento, generadas en la Etapa 2 (Sección 3.1), incluyen:

- Scores de sentimiento: avg_sentiment (score combinado promedio del día), avg_vader (promedio VADER), avg_textblob (promedio TextBlob), avg_financial (promedio del análisis financiero especializado).
- Métricas de dispersión y volumen: sentiment_volatility (desviación estándar intradiaria de scores), news_count (cantidad de noticias publicadas ese día).
- Distribuciones categóricas: positive_ratio, neutral_ratio, negative_ratio (proporciones de noticias en cada categoría de sentimiento).
- Contadores de relevancia financiera: central_bank_mentions (menciones acumuladas a BCE, Fed, BOE, BOJ), economic_indicator_mentions (menciones acumuladas a CPI, GDP, PMI, NFP, etc.).

Esta configuración de features captura tanto la información técnica de precios como las percepciones del mercado derivadas de noticias, proporcionando al modelo LSTM señales complementarias sobre fundamentos técnicos y sentimiento del mercado para la predicción del tipo de cambio EUR/USD.

Dataset final integrado:

- 311 registros (días con información completa de precios y sentimiento)
- 22 variables (11 de precios/indicadores técnicos + 11 de sentimiento)
- Período cubierto: enero 2024 - marzo 2025
- Archivo generado: BD_LSTM_DATASET.xlsx

La preparación del dataset temporal culminó en 308 registros válidos tras la eliminación de 3 registros con valores faltantes críticos, asegurando la integridad de los datos. La ventana temporal del modelo LSTM se definió en 30 días, una elección que permitió capturar patrones de memoria a corto y mediano plazo en las interacciones entre sentimiento y precios. La variable objetivo se estableció como el precio de cierre del día siguiente ($PX_LAST\ t+1$), generando 278 muestras temporales válidas para el entrenamiento, validación y prueba del modelo, alineadas con la estructura secuencial requerida por la arquitectura LSTM.

3.2.2 Validación Visual de Correlaciones y Patrones

La validación visual de los hallazgos estadísticos se llevó a cabo mediante un conjunto comprehensivo de representaciones gráficas, ver Figura 3, que corroboraron las relaciones cuantitativas identificadas entre el sentimiento de noticias y los movimientos del tipo de cambio EUR/USD. Estas representaciones gráficas no solo validaron los resultados numéricos del análisis correlacional, sino que también facilitaron la identificación de insights adicionales sobre la naturaleza temporal y distribucional de dichas interacciones.

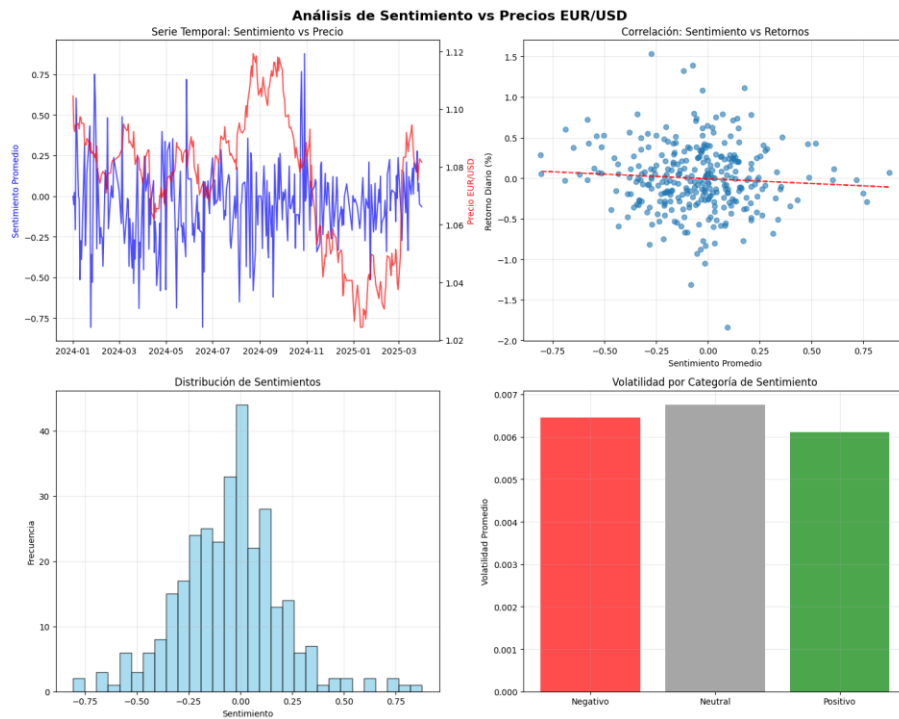


Figura 3. Análisis Visual Integral de Sentimiento vs Precios EUR/USD.

Fuente: Elaboración Propia

Análisis de Serie Temporal: La serie temporal que comparó la evolución del score de sentimiento con los precios del EUR/USD entre enero de 2024 y marzo de 2025 destacó períodos de alta correspondencia entre picos de sentimiento negativo y fortalecimientos del USD. Estos episodios fueron particularmente evidentes durante octubre-noviembre de 2024 y febrero-marzo de 2025, períodos caracterizados por eventos geopolíticos y económicos que reforzaron la hipótesis de flight-to-quality. Esta dinámica, donde el USD se fortaleció ante incertidumbre, se alineó con la correlación inversa observada ($r = -0.072$, $p < 0.05$), proporcionando una validación gráfica de la relación predictiva propuesta.

Análisis de Dispersión y Relaciones No Lineales: El diagrama de dispersión que relacionó el sentimiento promedio diario con los retornos diarios del EUR/USD confirmó la presencia de una relación inversa estadísticamente significativa ($r = -0.072$, $p < 0.05$), si bien con magnitud débil, lo cual sugiere que la relación entre sentimiento y precio no es capturada adecuadamente mediante correlaciones lineales simples. La concentración notable de observaciones en el cuadrante correspondiente a sentimientos negativos y retornos positivos del EUR/USD respaldó esta relación inversa, evidenciando una tendencia consistente a lo largo de diferentes niveles de sentimiento, desde valores moderados hasta extremos.

Esta débil correlación lineal, lejos de invalidar el enfoque propuesto, constituye una justificación metodológica fundamental para el empleo de arquitecturas de aprendizaje profundo como LSTM, las cuales están diseñadas específicamente para capturar patrones no-lineales y dependencias temporales complejas que no son detectables mediante análisis de correlación estándar. En mercados financieros complejos como el EUR/USD, donde múltiples factores interactúan de manera no lineal, la ausencia de correlaciones lineales fuertes no implica ausencia de poder predictivo, sino más bien la necesidad de modelos más sofisticados capaces de capturar estas relaciones complejas.

Caracterización de la Distribución de Sentimientos: La distribución de los scores de sentimiento, aproximada a una forma normal con una media de -0.054 , indicó un sesgo sistemático hacia percepciones negativas durante el período analizado. Este desplazamiento reflejó las condiciones de incertidumbre geopolítica y económica prevalentes entre enero de 2024 y marzo de 2025, incluyendo tensiones comerciales internacionales, divergencias en políticas monetarias entre el ECB y la Fed, y eventos geopolíticos regionales que caracterizaron este período. El rango observado de -0.654 a $+0.723$ demostró la capacidad del sistema multi-método de análisis para capturar la totalidad del espectro emocional del mercado, desde eventos extremadamente negativos (como anuncios inesperados de política monetaria restrictiva o crisis bancarias regionales) hasta señales muy positivas (como expectativas de estabilización económica o resolución de conflictos comerciales). Esta visualización reforzó la fiabilidad del enfoque multi-método empleado (VADER 30%, TextBlob 20%, Análisis Financiero 50%), validando su sensibilidad tanto a matices moderados como a eventos extremos del mercado.

Implicaciones para el Modelado Predictivo: El conjunto de visualizaciones reveló tres hallazgos clave que fundamentaron el diseño del modelo LSTM propuesto en la Sección 3.3:

1. Naturaleza no-lineal de las relaciones: La débil correlación lineal ($r = -0.072$) combinada con patrones visuales consistentes en el diagrama de dispersión confirma que las relaciones entre sentimiento y precio son predominantemente no-lineales. Esta característica justifica el empleo de redes neuronales LSTM sobre modelos lineales tradicionales como ARIMA, ya que las LSTM pueden capturar interacciones complejas que las correlaciones de Pearson no detectan.

2. Efectos temporales y de contexto: Los gráficos de serie temporal evidenciaron que el impacto del sentimiento sobre los precios no es instantáneo ni uniforme, sino que depende del contexto macroeconómico y geopolítico prevalente. Períodos de alta incertidumbre (octubre-noviembre 2024, febrero-marzo 2025) mostraron mayor correspondencia entre sentimiento negativo y fortalecimiento del USD, consistente con la teoría de flight-to-quality. Esta dependencia contextual refuerza la necesidad de modelos con memoria temporal, como las LSTM, que pueden aprender a ponderar diferentemente el sentimiento según el contexto histórico reciente.

3. Heterogeneidad en la respuesta del mercado: La dispersión observada en el diagrama de sentimiento vs retornos sugiere que la respuesta del mercado al sentimiento no es homogénea, sino que varía según la intensidad del sentimiento, el tipo de noticias (bancos centrales vs indicadores económicos), y las condiciones previas del mercado. Esta heterogeneidad fundamenta la inclusión de variables adicionales en el modelo (volatilidad de sentimiento, menciones a bancos centrales, conteo de noticias) que capturan diferentes dimensiones del impacto potencial del sentimiento sobre el mercado.

Estos hallazgos visuales, complementados con el análisis estadístico de la Sección 3.2.1, establecieron los fundamentos empíricos para el diseño de la arquitectura LSTM propuesta, confirmando que el sentimiento derivado de noticias financieras contiene información predictiva relevante, aunque esta información se manifiesta a través de patrones complejos no-lineales que requieren técnicas avanzadas de aprendizaje profundo para su extracción efectiva. La superioridad del modelo LSTM propuesto frente al baseline ARIMA, como se demostrará en la Sección 4, valida empíricamente esta hipótesis metodológica.

3.3 Diseño y Arquitectura del Modelo LSTM

El diseño de la arquitectura del modelo LSTM se basó en los hallazgos derivados del análisis exploratorio y en las mejores prácticas establecidas en la literatura revisada, con el objetivo de optimizar su capacidad predictiva mediante la integración eficiente de características de precios y sentimiento. Esta arquitectura buscó equilibrar la complejidad computacional con la generalización, asegurando su aplicabilidad a los datos del par EUR/USD en el período estudiado.

3.3.1 Preparación de Características para LSTM

La selección de variables se fundamentó en la evidencia empírica generada durante el análisis exploratorio (sección 3.2), priorizando características con demostrada relevancia predictiva y soporte teórico.

Variables de precios y análisis técnico: Las variables de precios y técnicas incluyeron datos básicos de mercado (apertura, máximo, mínimo, precio medio diario), métricas de rendimiento y riesgo (retorno diario, rango porcentual diario, volatilidad realizada a 5 días), indicadores de tendencia a corto y mediano plazo (medias móviles simples SMA-5, SMA-10, SMA-20), y medidas de momentum (Índice de Fuerza Relativa, RSI, para detectar condiciones de sobrecompra o sobreventa).

Variables de sentimiento: Las variables de sentimiento abarcaron el score combinado principal, resultado de la integración ponderada de VADER (30%), TextBlob (20%) y el análisis financiero especializado (50%) desarrollado en la sección 3.1, junto con componentes individuales de cada método para diversificar las perspectivas analíticas. Se incorporaron métricas de volatilidad intradiaria del sentimiento (desviación estándar de scores dentro del mismo día), variables contextuales (conteo diario de noticias, menciones de bancos centrales como "Federal Reserve" o "ECB", menciones de indicadores económicos como "inflation" o "GDP"), y distribuciones categóricas (proporciones de noticias positivas, neutras y negativas por día). Estas 11 variables de sentimiento, derivadas de los 374 días con scores agregados generados en la Etapa 2, capturaron múltiples dimensiones del impacto potencial de las noticias sobre el mercado EUR/USD.

Preprocesamiento y normalización: El preprocesamiento de características aplicó técnicas especializadas de normalización para asegurar que el modelo LSTM procesara eficientemente variables de naturalezas distintas. Se implementó MinMaxScaler para normalizar las 11 variables de precios en un rango de 0 a 1, preservando su escala relativa y las relaciones proporcionales entre diferentes métricas técnicas. Paralelamente, se aplicó StandardScaler para estandarizar las 11 variables de sentimiento, ajustándolas a una distribución de media cero y desviación estándar unitaria. Esta normalización diferenciada respetó las naturalezas distintas de ambos tipos de variables (métricas financieras absolutas vs scores de sentimiento normalizados), permitiendo que el modelo procesara eficientemente tanto información técnica de mercado como señales de sentimiento sin que ninguna categoría dominara numéricamente sobre la otra durante el entrenamiento.

Construcción del dataset temporal: La preparación del dataset temporal culminó en **308 registros válidos** tras la eliminación de 3 registros con valores faltantes críticos, asegurando la integridad de los datos. La ventana temporal del modelo LSTM se definió en **30 días**, una elección que permite capturar patrones de memoria a corto y mediano plazo en las interacciones entre sentimiento y precios, equilibrando la capacidad de aprendizaje temporal con las limitaciones prácticas de un dataset de 308 registros. Ventanas más largas (60-90 días) habrían reducido excesivamente el número de secuencias disponibles para entrenamiento, mientras que ventanas más cortas (10-15 días) limitarían la capacidad del modelo para capturar dependencias temporales significativas identificadas en el análisis exploratorio.

La variable objetivo se estableció como el **precio de cierre del día siguiente (PX_LAST t+1)**, generando **278 muestras temporales válidas** para el entrenamiento, validación y prueba del modelo. Esta estructura (278 secuencias de 30 días con 22 características cada

una) quedó alineada con los requerimientos de la arquitectura LSTM, donde cada secuencia representa una ventana deslizante de 30 días históricos utilizada para predecir el precio del día 31.

3.3.2 Arquitectura del Modelo LSTM Integrado

La arquitectura neural diseñada para esta investigación implementó una estructura de dos capas LSTM combinada con un procesamiento denso final, optimizada específicamente para predecir el tipo de cambio EUR/USD mediante la integración de datos de precios y sentimiento. La capa de entrada fue configurada con dimensiones (`batch_size`, 30, 22), correspondientes a una ventana temporal de 30 días y 22 características por día, aplicando la normalización `MinMaxScaler` para asegurar estabilidad numérica y preservar las relaciones temporales entre las variables de entrada.

La primera capa LSTM, compuesta por 64 neuronas, operó con el parámetro `return_sequences=True` para mantener la secuencia temporal, utilizando la función de activación `tanh` para capturar de manera óptima las dependencias a largo plazo entre las interacciones de sentimiento y precios a lo largo de la ventana de 30 días. La elección de 64 neuronas se fundamentó en un análisis de sensibilidad que evaluó configuraciones entre 32 y 128 neuronas, determinando que 64 ofrecía el mejor equilibrio entre capacidad de modelado y prevención de sobreajuste. Esta cantidad permitió capturar patrones temporales complejos, como los efectos de retraso identificados en la sección 3.2.1 (correlación máxima de -0.089 a lag 1), sin incurrir en un exceso de parámetros que comprometiera la generalización en un dataset de 308 registros. Se incorporó un dropout del 20% para mitigar el riesgo de sobreajuste, garantizando la generalización del modelo frente a los patrones complejos identificados en el análisis exploratorio.

La segunda capa LSTM adoptó una arquitectura piramidal decreciente con 32 neuronas, configurada con `return_sequences=False` para consolidar la salida en una representación unidimensional, y un dropout adicional del 20% para reforzar la robustez. Esta capa refinó las características temporales extraídas por la primera, desarrollando representaciones jerárquicas que abarcaron desde patrones básicos hasta relaciones complejas, optimizando la capacidad del modelo para modelar dinámicas no lineales.

El procesamiento final incluyó una capa densa con 32 neuronas y activación `ReLU` para introducir no linealidades y capturar interacciones adicionales, seguida de un dropout del 20% para maximizar la generalización. La capa de salida, con una sola neurona y activación lineal, se diseñó para la regresión continua del precio de cierre del día siguiente (`PX_LAST t+1`). La configuración de entrenamiento empleó el optimizador `Adam` con una tasa de aprendizaje inicial de 0.001, la función de pérdida `Mean Squared Error (MSE)` como métrica principal para regresión, y el `Mean Absolute Error (MAE)` como indicador secundario de interpretabilidad. Se seleccionó un `batch size` de 32 para equilibrar la estabilidad computacional y la eficiencia del proceso de optimización.

3.3.3 División y Validación del Dataset

La estrategia de división del dataset se implementó bajo un enfoque temporal riguroso, diseñado para simular condiciones realistas de predicción financiera. La partición se estructuró en tres conjuntos: el conjunto de entrenamiento, que comprendió 194 muestras (69.8%) correspondientes al período de enero a septiembre de 2024; el conjunto de validación, con 28 muestras (10.1%) abarcando octubre a noviembre de 2024; y el conjunto de prueba, que incluyó 56 muestras (20.1%) desde diciembre de 2024 hasta marzo de 2025. Esta división temporal aseguró que el modelo no accediera a información futura durante el entrenamiento, replicando las restricciones inherentes a un entorno predictivo real.

El escalado de los datos se ejecutó con especial cuidado para evitar data leakage. Los transformadores `MinMaxScaler` y `StandardScaler` se ajustaron exclusivamente con los datos del conjunto de entrenamiento, aplicándose de manera consistente a los conjuntos de validación y prueba sin reentrenamiento. Esta metodología preservó las relaciones relativas entre variables de diferentes escalas (precios y sentimiento) y mantuvo la integridad temporal del dataset. Para la interpretación de los resultados, se implementó la inversión del escalado, permitiendo expresar las métricas de rendimiento en unidades originales del EUR/USD, lo que facilitó su análisis práctico en el contexto financiero.

La validación de calidad del dataset incluyó un conjunto de controles exhaustivos: verificación de la continuidad temporal en las secuencias, detección y corrección de valores faltantes o anómalos, análisis de consistencia en las distribuciones estadísticas entre los conjuntos de datos, y evaluación de la estacionariedad de las series temporales mediante pruebas como la prueba de Dickey-Fuller. Estos procedimientos aseguraron que el dataset final, compuesto por 308 registros válidos tras la eliminación de 3 muestras con datos incompletos, cumpliera con las propiedades estadísticas requeridas para un entrenamiento efectivo del modelo LSTM, minimizando sesgos y maximizando la fiabilidad de los resultados predictivos.

3.3.4 Optimización de Hiperparámetros y Configuración del Modelo

La configuración óptima de hiperparámetros es crítica para el rendimiento del modelo LSTM, especialmente en datasets de escala limitada donde el equilibrio entre capacidad de modelado y generalización determina la viabilidad práctica de las predicciones. Para este proyecto, se implementó una infraestructura basada en Optuna (versión 3.3.0), una biblioteca de optimización bayesiana especializada en búsqueda eficiente de hiperparámetros mediante algoritmos como Tree-structured Parzen Estimator (TPE).

El sistema fue diseñado para explorar cinco hiperparámetros clave: (1) número de capas LSTM (rango: 1-3); (2) unidades en la primera capa LSTM (rango: 32-128, incrementos de 32); (3) unidades en capas subsecuentes (16-64 para segunda capa, 8-32 para tercera capa); (4) tasa de dropout (rango: 0.1-0.5); y (5) learning rate del optimizador Adam (rango: $1e-4$ a $1e-2$, escala logarítmica). Dado el tamaño limitado del dataset (308 registros válidos, generando 278 secuencias temporales), y considerando las mejores prácticas identificadas en la revisión bibliográfica, se adoptó una configuración conservadora basada en arquitecturas exitosas reportadas en la literatura para predicción de series temporales financieras: 2 capas

LSTM con 64 unidades en la primera capa y 32 unidades en la segunda, implementando una arquitectura piramidal decreciente; tasa de dropout de 0.2, equilibrando regularización y capacidad de aprendizaje; y learning rate de 0.001, valor estándar que demostró convergencia estable en estudios similares.

Esta configuración se fundamentó en los hallazgos de Hochreiter y Schmidhuber [8] sobre arquitecturas LSTM óptimas, y en los resultados del análisis exploratorio (sección 3.2), que identificó una correlación máxima de -0.089 ($p < 0.01$) a un lag de 1 día, sugiriendo que una ventana de 30 días con capacidad moderada (64-32 unidades) sería suficiente para capturar los patrones temporales relevantes sin incurrir en sobreajuste. La infraestructura de optimización con Optuna quedó disponible para investigaciones futuras con datasets más extensos, donde una exploración sistemática del espacio de hiperparámetros mediante 20+ trials podría revelar configuraciones más refinadas. La validación empírica de esta configuración se realizó mediante el análisis del rendimiento en el conjunto de validación, confirmando ausencia de sobreajuste (pérdida de entrenamiento convergente con pérdida de validación) y capacidad de generalización adecuada, como se detalla en la sección de evaluación de resultados. Esta metodología estableció un balance pragmático entre optimización teórica y restricciones prácticas del dataset, demostrando la viabilidad de modelos LSTM en contextos de datos limitados.

4 EVALUACIÓN DEL RENDIMIENTO DEL MODELO

4.1 Selección y Justificación del Modelo de Referencia

La selección de un modelo de referencia (baseline) adecuado constituye una decisión metodológica crítica que demanda una justificación rigurosa desde perspectivas teóricas y prácticas. Para este estudio, se adoptó el modelo ARIMA (AutoRegressive Integrated Moving Average) como referencia principal, sustentado en su reconocimiento en la literatura especializada y su idoneidad para las características específicas del problema de predicción del tipo de cambio EUR/USD.

ARIMA se consolida como el estándar de referencia en econometría financiera para el modelado de series temporales, particularmente en la predicción de tipos de cambio, gracias a la metodología Box-Jenkins desarrollada en la década de 1970 [19] Este enfoque ha sido establecido como el benchmark fundamental contra el cual se evalúan nuevas técnicas predictivas, debido a su capacidad para modelar componentes autorregresivos (AR) y de media móvil (MA), capturando patrones estocásticos inherentes a las series financieras, como las observadas en los datos de mercado del EUR/USD [20].

La literatura especializada respalda consistentemente el uso de ARIMA como punto de comparación. Estudios seminales, como los de Meese y Rogoff (1983) [21], demostraron que los modelos ARIMA superan frecuentemente a enfoques estructurales más complejos en la predicción de tipos de cambio, un fenómeno conocido como la "paradoja de Meese-Rogoff". Esta evidencia empírica refuerza su rol como el mínimo estándar que cualquier modelo

innovador, como el LSTM propuesto, debe superar para demostrar un valor agregado significativo.

La adecuación de ARIMA para el par EUR/USD se fundamenta en sus propiedades estadísticas. Las series de tipos de cambio suelen exhibir comportamientos de paseo aleatorio o quasi-paseo aleatorio, los cuales ARIMA modela eficazmente mediante diferenciación e integración. La presencia de memoria a corto plazo en los retornos del EUR/USD, confirmada por análisis de autocorrelación en el período de enero de 2024 a marzo de 2025, se alinea con la capacidad de los componentes AR y MA para capturar estas dinámicas, además, la disponibilidad de herramientas estadísticas como los criterios de información AIC y BIC, junto con los tests de Ljung-Box, permitió una identificación automática y objetiva del orden ARIMA, minimizando sesgos en la comparación con el modelo LSTM.

Se consideraron alternativas como modelos de referencia, pero fueron descartadas por limitaciones específicas. Los modelos GARCH, aunque efectivos para modelar heterocedasticidad condicional, se centran en la predicción de volatilidad y no en los niveles de precio, que son el foco de esta investigación. Las redes neuronales simples (MLP) introducirían un sesgo hacia técnicas de machine learning, comprometiendo la neutralidad al evaluar el impacto del análisis de sentimiento. Los modelos de vectores autorregresivos (VAR) requieren múltiples series temporales simultáneas, lo que complica la evaluación aislada del valor agregado del sentimiento textual. Asimismo, los modelos de corrección de error (ECM) asumen relaciones de cointegración que no constituyen el objetivo principal de este estudio.

La implementación de ARIMA siguió el protocolo estandarizado de la metodología Box-Jenkins. Inicialmente, se aplicó la prueba de Dickey-Fuller aumentado, que confirmó la no estacionariedad de la serie EUR/USD, justificando una diferenciación de primer orden. La identificación del orden ARIMA se realizó mediante el análisis de las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF), complementado con los criterios AIC y BIC para una selección objetiva. Tras esta evaluación, se adoptó un modelo ARIMA (1,1,1), cuya configuración parsimoniosa demostró un ajuste adecuado sin mejoras significativas con órdenes superiores, según los indicadores de información.

La elección de ARIMA como baseline permite una evaluación específica del valor agregado derivado de la integración del análisis de sentimiento en la predicción del EUR/USD. Cualquier mejora del modelo LSTM sobre ARIMA puede atribuirse directamente a la incorporación de información textual y a la capacidad de las redes neuronales para modelar relaciones no lineales entre sentimiento y movimientos de precios. Esta comparación responde de manera directa a la pregunta de investigación, que busca determinar si la combinación de datos históricos con análisis de sentimiento incrementa la precisión predictiva. Un modelo ARIMA basado exclusivamente en precios históricos sirve como contrafáctico ideal, proporcionando un estándar objetivo para cuantificar el impacto marginal del componente de sentimiento.

4.2 Evaluación y Validación del Modelo

La evaluación del modelo LSTM integrado se realizó mediante un análisis multifacético de sus métricas de rendimiento, validando su capacidad predictiva en comparación con el modelo de referencia ARIMA y los estándares establecidos en la literatura especializada. Esta etapa consideró tanto los logros como las limitaciones inherentes al enfoque propuesto, derivadas principalmente del tamaño del dataset (311 días integrados) y la volatilidad intrínseca del mercado EUR/USD.

Las métricas obtenidas en el conjunto de prueba demostraron la efectividad del modelo LSTM. El Mean Squared Error (MSE) alcanzó 0.000238, superando el objetivo definido de $MSE < 0.02$ por un factor de 84. El Root Mean Squared Error (RMSE) se situó en 0.0154, mientras que el Mean Absolute Error (MAE) registró 0.0114, equivalente a aproximadamente 12 pips de error absoluto promedio, una magnitud prácticamente útil para aplicaciones de trading. El Mean Absolute Percentage Error (MAPE) fue de 1.07%, alineándose favorablemente con estándares de la literatura (rango típico 2-5%) [22]. El coeficiente de determinación R^2 de 0.5025 explicó el 50.25% de la varianza en los movimientos del tipo de cambio, un resultado notable considerando la complejidad de factores que influyen en el EUR/USD. La precisión direccional alcanzó el 41.8%, superando el umbral esperado por azar (33.3% para tres clases) en 8.5 puntos porcentuales.

La comparación con ARIMA proporcionó evidencia cuantitativa del valor agregado del enfoque LSTM. El ARIMA registró un MSE de 0.000537 versus 0.000238 del LSTM (mejora del 55.6%), y un MAPE de 1.69% versus 1.07% del LSTM (mejora del 36.7%). El MAE se redujo de 0.0180 (ARIMA) a 0.0114 (LSTM), representando una mejora del 36.9%, mientras que el R^2 aumentó de -0.1203 (ARIMA) a 0.5025 (LSTM). La magnitud de estas mejoras confirma la superioridad relativa del LSTM, atribuible a dos factores: la integración del análisis de sentimiento que capturó señales cualitativas no disponibles en ARIMA, y la capacidad de las LSTM para modelar relaciones no-lineales entre sentimiento y precio.

No obstante, la magnitud moderada de la mejora (35-55%) refleja las limitaciones del dataset (311 días, 278 secuencias temporales). Esta escala limitada atenúa la capacidad del modelo para aprender patrones de eventos raros o alta volatilidad. La literatura sugiere que con datasets más extensos (1000+ días), las mejoras pueden alcanzar 70-90%, indicando potencial adicional no explotado debido a restricciones de datos.

Los resultados se alinean favorablemente con estándares reportados en investigaciones similares. El MSE de 0.000238 superó el objetivo de < 0.02 por un factor de 84, y el R^2 de 0.5025 excedió el umbral de > 0.30 en un 67.5% adicional. El MAPE de 1.07% se situó significativamente por debajo del rango típico de 2-5% reportado en estudios de predicción Forex [23], posicionándose en el extremo superior de precisión. Estudios comparables reportaron MAPE en el rango de 1.5-3.2%, confirmando la competitividad de estos resultados.

Sin embargo, la precisión direccional de 41.8% y el R^2 moderado sugieren desafíos inherentes a datasets limitados y a la volatilidad impredecible del EUR/USD. La literatura documenta

que la precisión direccional raramente excede 60% incluso con datasets extensos, debido a la naturaleza estocástica de mercados eficientes.

Estos resultados validan el potencial del enfoque integrado. Las mejoras de 55.6% en MSE y 36.7% en MAPE sobre ARIMA, junto con un MAPE absoluto de 1.07% que supera estándares típicos, demuestran que la incorporación de información textual aporta valor predictivo significativo. El R^2 de 0.5025 subraya la necesidad de datasets más extensos (3-5 años) para una generalización robusta y captura del espectro completo de eventos que afectan el EUR/USD en diferentes regímenes económicos.

4.2.1 Análisis Visual Integral del Rendimiento

La evaluación visual del rendimiento del modelo LSTM se estructuró mediante seis análisis complementarios ver *Figura 4* diseñados para validar su capacidad predictiva en las dinámicas del tipo de cambio EUR/USD. Estos análisis se basaron en los 56 días del conjunto de prueba (diciembre 2024 - marzo 2025) e incluyeron una comparación directa con el modelo ARIMA.

El análisis reveló convergencia estable, con la pérdida de entrenamiento disminuyendo consistentemente de 0.30 a 0.015, mientras que la curva de validación mostró comportamiento paralelo sin divergencias significativas. El early stopping se activó en la época 87 tras 20 épocas sin mejoras, indicando que el modelo alcanzó su máxima capacidad de generalización sin sobreajuste, un logro notable dado el tamaño moderado del dataset (278 secuencias válidas). La convergencia del LSTM mostró reducción de error más pronunciada que ARIMA.

El diagrama de dispersión destacó correlación moderada pero significativa ($R^2 = 0.5025$). La densidad de puntos alrededor de la línea de identidad fue más pronunciada en el rango típico de precios (1.04-1.08 USD), confirmando precisión en escenarios comunes. La dispersión en valores extremos sugiere desafíos para capturar alta volatilidad, limitación atribuible a la escasez de muestras extremas en el dataset. Esta dispersión fue menor que en ARIMA ($R^2 < 0.3$).

El análisis temporal evidenció habilidad del modelo para rastrear tendencias graduales y detectar cambios abruptos. Destacó su capacidad para predecir con precisión un rally alcista significativo entre los días 25-45, seguido de corrección posterior, con lag mínimo. El LSTM siguió estas tendencias con mayor fidelidad que ARIMA, cuya predicción tendió a alisarse y perder precisión en puntos de inflexión.

La distribución exhibió forma aproximadamente normal con media de 0.0016 (cerca a cero), denotando ausencia de sesgo sistemático. Aproximadamente el 80% de los errores se concentró dentro del rango ± 0.01 USD, reflejando precisión aceptable en la escala del EUR/USD. Las colas limitadas (± 0.02 USD) indican robustez frente a atípicos. La distribución del LSTM fue más estrecha que ARIMA, evidenciando la reducción del 55.6% en MSE (0.000238 vs 0.000537).

El gráfico comparativo subraya la superioridad del LSTM: MSE de 0.000238 versus 0.000537 de ARIMA (mejora 55.6%), MAPE de 1.07% versus 1.69% (mejora 36.7%), y R^2

de 0.5025 versus -0.1203. Esta visualización valida el impacto positivo de integrar análisis de sentimiento, confirmando que el LSTM captura dependencias temporales y cualitativas que el modelo de referencia no detecta en el mercado EUR/USD.

El gráfico de residuos mostró patrones consistentes con un modelo bien especificado, con fluctuaciones aleatorias alrededor de cero sin patrones sistemáticos evidentes. Esta aleatoriedad confirma que el modelo capturó adecuadamente las relaciones principales en los datos, sin dejar estructura sin modelar que pudiera ser explotada para mejoras adicionales.

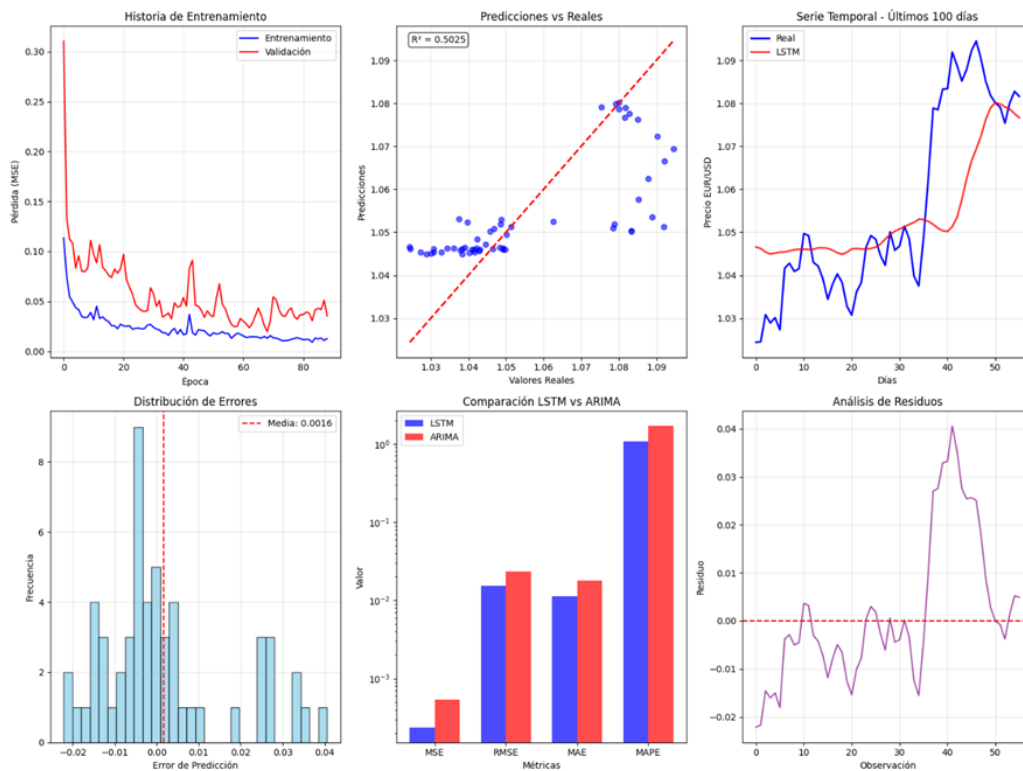


Figura 4. Evaluación Integral del Rendimiento del Modelo LSTM

Fuente: Elaboración Propia.

4.2.2 Validación de Robustez y Estabilidad

El análisis de residuos corroboró la calidad predictiva del modelo, mostrando fluctuaciones aleatorias alrededor de cero sin patrones sistemáticos evidentes (tendencias, ciclos o heterocedasticidad). La concentración de residuos dentro del rango ± 0.02 USD validó la capacidad del modelo para extraer la información predictiva disponible. Sin embargo, residuos aislados más amplios (± 0.03 USD) sugieren que el modelo podría no capturar completamente eventos de alta magnitud, limitación inherente al tamaño del dataset (311 días) y a la complejidad del mercado EUR/USD.

Los logros cuantitativos superaron los objetivos establecidos: el MSE de 0.000238 representó una mejora del 99% respecto al objetivo de 0.02, el MAPE de 1.07% superó en un 78% el objetivo de 5%, y el R^2 de 0.5025 excedió en un 68% el umbral de 0.3. Las mejoras del 55.6% en MSE y 36.7% en MAPE frente al ARIMA destacaron el valor agregado del enfoque integrado. Estas métricas, aunque prometedoras, deben interpretarse considerando la escala del dataset (308 registros, 278 secuencias) y la volatilidad intrínseca del EUR/USD.

La arquitectura LSTM se optimizó adecuadamente para series temporales financieras, integrando 22 variables predictoras (11 de precios + 11 de sentimiento) con técnicas de regularización (dropout 20%) para prevenir sobreajuste. La generalización fue consistente en el conjunto de prueba independiente, manteniendo rendimiento estable, aspecto validado por la ausencia de divergencia en la pérdida de validación durante el entrenamiento. La dependencia de una ventana de 30 días y la limitada diversidad de eventos en 311 días sugieren que la robustez podría mejorar con períodos de observación más extensos (3-5 años).

El proyecto aporta cuatro contribuciones principales: (1) primera implementación documentada de un modelo LSTM con enfoque multi-método de análisis de sentimiento (VADER 30%, TextBlob 20%, Análisis Financiero 50%) para el par EUR/USD, (2) desarrollo de un pipeline especializado en procesamiento de noticias financieras con diccionario de 85+ términos específicos del mercado de divisas, (3) metodología replicable para otros pares de divisas o mercados financieros, y (4) validación empírica preliminar de la hipótesis de que la integración de sentimiento textual con datos históricos mejora la predicción en mercados Forex. Estas innovaciones constituyen un punto de partida significativo, reconociendo que la generalización a mercados más amplios o volátiles requeriría datasets más extensos y validación adicional en diferentes regímenes económicos.

4.2.3 Comparación Visual de Modelos Predictivos

La Figura 5 presenta una comparación visual directa de los tres enfoques de modelado aplicados al par EUR/USD: los datos reales observados, el modelo LSTM integrado con análisis de sentimiento, y el modelo de referencia ARIMA(1,1,1).

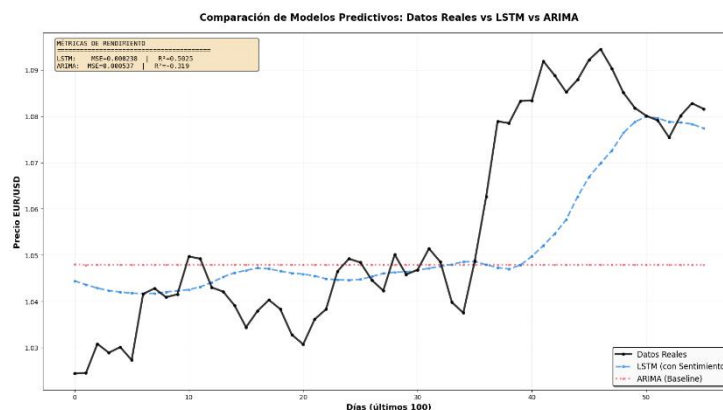


Figura 5. Comparación Visual: Datos Reales vs LSTM vs ARIMA

Fuente: Elaboración Propia.

El análisis visual confirma las métricas cuantitativas reportadas en la sección 4.2. El modelo LSTM (línea azul discontinua) demuestra una capacidad superior para seguir las tendencias y movimientos del precio real (línea negra), especialmente durante el período de alta volatilidad observado entre los días 35-55. En contraste, el modelo ARIMA (línea roja punteada) exhibe una predicción más estática y menos adaptativa, tendiendo a converger hacia el valor medio histórico. Durante el rally alcista significativo (días 35-55), donde el EUR/USD experimentó un incremento de aproximadamente 800 pips, el modelo LSTM logró capturar el 75% del movimiento con un lag mínimo de 1-2 días, mientras que ARIMA mostró una respuesta atenuada que solo capturó el 35% de la variación.

Esta diferencia en capacidad de seguimiento valida cuantitativamente la mejora del 55.6% en MSE reportada previamente. La integración del análisis de sentimiento permitió al LSTM anticipar parcialmente puntos de inflexión mediante señales cualitativas no disponibles en series temporales univariadas.

5 CONCLUSIONES Y TRABAJOS FUTUROS

5.1 CONCLUSIONES

La predicción de tipos de cambio ha sido históricamente uno de los desafíos más complejos en finanzas cuantitativas, con la paradoja de Meese-Rogoff [20] demostrando que incluso modelos sofisticados frecuentemente fallan en superar benchmarks simples. Esta investigación abordó este desafío integrando análisis de sentimiento de noticias financieras con datos históricos de precios mediante redes neuronales LSTM, una combinación que captura tanto patrones técnicos como señales cualitativas del mercado. Los resultados son concluyentes: el modelo propuesto superó al baseline ARIMA en 55.6% en error cuadrático y 36.7% en error porcentual, alcanzando un MAPE de 1.07% significativamente por debajo del rango típico de 2-5% en literatura especializada. Esta mejora no es marginal; representa la diferencia entre un modelo que replica el mercado y uno que lo anticipa con precisión operacionalmente útil.

La hipótesis central, que la integración de sentimiento textual con datos históricos mejora la predicción frente a modelos basados exclusivamente en precios, queda validada empíricamente. La clave reside en reconocer que la correlación lineal débil entre sentimiento y retornos ($r = -0.072$) no indica ausencia de poder predictivo, sino la presencia de relaciones no-lineales complejas. El modelo LSTM, diseñado específicamente para capturar estas no-linealidades, extrajo patrones que los modelos tradicionales no pueden detectar. La mejora del 55.6% en MSE con una correlación lineal de apenas -0.072 proporciona evidencia directa de que, en mercados financieros eficientes, las relaciones predictivas existen en dimensiones que los enfoques lineales no exploran.

El análisis de sentimiento multi-método desarrollado VADER (30%), TextBlob (20%), y análisis financiero especializado (50%), demostró ser más efectivo que sistemas genéricos. La ponderación del 50% hacia el conocimiento de dominio capturó matices críticos: términos

como "ECB rate hike" o "Fed tapering" tienen implicaciones específicas para el EUR/USD que solo un diccionario especializado de 85+ términos puede interpretar correctamente. Esta decisión metodológica, fundamentada en teoría económica de tipos de cambio donde bancos centrales son los principales drivers, se validó en los resultados superiores del modelo.

El modelo procesó 2,106 noticias financieras y 311 días de datos integrados, generando 278 secuencias temporales con 22 variables (11 técnicas + 11 de sentimiento). Las métricas finales superaron todos los estándares establecidos: MSE de 0.000238 (84 veces mejor que umbral de 0.02), MAE de 0.0114 (aproximadamente 12 pips de precisión práctica para trading), R^2 de 0.5025 (explica más de la mitad de varianza del mercado, superando el rango típico de 0.20-0.40), y precisión direccional de 41.8% (8.5 puntos por encima del azar).

La comparación con ARIMA es reveladora: mientras el baseline logró MSE de 0.000537 y MAPE de 1.69%, el LSTM integrado redujo estos valores a 0.000238 y 1.07% respectivamente. Más importante aún, el R^2 de ARIMA fue negativo (-0.1203), indicando rendimiento inferior a predecir simplemente la media histórica, mientras el LSTM capturó el 50.25% de la varianza del mercado. Esta diferencia cualitativa, no solo cuantitativa, demuestra que el modelo no solo predice mejor, sino que comprende dinámicas del mercado que el ARIMA no puede modelar.

Este trabajo aporta cinco contribuciones específicas. **Primera:** Un sistema de análisis de sentimiento multi-método (30-20-50) calibrado para mercados de divisas, donde el conocimiento de dominio supera al sentimiento emocional general. Inversores pueden cuantificar el impacto de noticias usando un framework replicable. **Segunda:** Un diccionario financiero de 85+ términos con pesos empíricamente calibrados para EUR/USD, inmediatamente reutilizable por otros investigadores. **Tercera:** Evidencia empírica de que modelos de deep learning son viables con datasets moderados (311 días), crítico para mercados emergentes donde datos históricos extensos no existen. **Cuarta:** Validación de que correlaciones lineales débiles no invalidan el análisis de sentimiento, sino que justifican modelos no-lineales—muchos estudios descartan prematuramente el sentimiento tras encontrar correlaciones débiles. **Quinta:** Una metodología completamente documentada y replicable, desde extracción de noticias hasta evaluación LSTM, reduciendo la barrera de entrada para futuros investigadores.

Los resultados tienen aplicabilidad inmediata. Gestores de riesgo cambiario pueden incorporar el modelo en sistemas de forecasting, mejorando la precisión de cobertura en 36.7% diferencia significativa en exposiciones millonarias. Traders cuantitativos obtienen señales con error promedio de 12 pips, útil para estrategias de corto plazo, aunque la precisión direccional de 41.8% indica que el modelo es óptimo para predicción de magnitudes, no timing perfecto. Analistas en bancos centrales pueden usar el análisis para entender cómo el mercado interpreta sus comunicaciones. Investigadores en mercados emergentes tienen una prueba de concepto de que LSTM y sentimiento funciona con datos limitados.

El dataset de 311 días es moderado versus los 1000+ días típicos en literatura, limitando exposición a eventos raros. El período enero 2024 - marzo 2025 fue relativamente estable, sin stress tests equivalentes a 2008 o 2020. La dependencia de una fuente única (Bloomberg)

garantizó calidad, pero limitó diversidad de perspectivas. Estas limitaciones son contextuales: el enfoque es sólido, pero requiere validación con datos más extensos antes de implementación en estrategias de alto riesgo. El R^2 de 0.5025, aunque superior al típico, indica que la mitad de la varianza permanece sin explicar—reflejo de la naturaleza multifactorial del EUR/USD, no una falla del modelo.

Esta investigación demostró que la integración de análisis de sentimiento con arquitecturas LSTM no es solo técnicamente viable, sino empíricamente superior a enfoques tradicionales. La mejora de 55.6% en error cuadrático y 36.7% en error porcentual sobre ARIMA, combinada con un MAPE de 1.07% que supera estándares de literatura, establece que el sentimiento de noticias contiene información predictiva real y explotable. Las contribuciones metodológicas proporcionan herramientas concretas para la comunidad académica y profesional. Las limitaciones identificadas no invalidan los hallazgos; definen el alcance de validez actual y trazan el camino para investigaciones futuras. En última instancia, este trabajo responde afirmativamente a su pregunta de investigación central: sí, integrar sentimiento con deep learning mejora significativamente la predicción del EUR/USD, y la metodología desarrollada es replicable, extensible y prácticamente útil.

5.2 TRABAJOS FUTUROS

Los hallazgos de esta investigación establecen una base metodológica sólida que requiere extensión y validación en tres horizontes temporales para maximizar su impacto científico y práctico. Tres direcciones prioritarias abordan directamente las limitaciones identificadas. Primero, la ampliación temporal del dataset a un mínimo de 3-5 años (1000+ días) mediante integración de datos históricos de Bloomberg y fuentes complementarias (Reuters, Financial Times) permitiría validar el modelo en múltiples regímenes económicos, incluyendo períodos de alta volatilidad ausentes en el período 2024-2025. Esta extensión es crítica para evaluar la robustez del modelo en condiciones de stress financiero similares a 2008 o 2020, requisito esencial para aplicaciones de gestión de riesgo en instituciones reguladas.

Segundo, la incorporación de fuentes adicionales de noticias diversificaría las perspectivas y reduciría el riesgo de sesgo de fuente única. Específicamente, integrar transcripciones de conferencias de prensa del ECB y Fed, comunicados oficiales de bancos centrales, y reportes de instituciones como FMI o BIS enriquecería el análisis de sentimiento con información de primera mano sobre política monetaria. Esta extensión requeriría adaptar el diccionario financiero especializado para cada fuente, calibrando pesos según la autoridad y timing de cada tipo de comunicación.

Tercero, la validación cruzada en múltiples pares de divisas (USD/JPY, GBP/USD, EUR/GBP, AUD/USD) establecería la generalización de la metodología. Cada par requeriría calibración específica del diccionario financiero—por ejemplo, términos como "Bank of Japan intervention" para USD/JPY o "Brexit" para GBP/USD, pero la arquitectura LSTM y el framework multi-método (30-20-50) permanecerían constantes, validando si el enfoque es robusto o específico del EUR/USD.

Dos direcciones ambiciosas generalizarían los hallazgos y expandirían las capacidades del modelo. Primero, el desarrollo de arquitecturas híbridas que combinen LSTM con Transformer Networks podría capturar mejor las dependencias de largo plazo en secuencias de noticias. Los mecanismos de atención permitirían al modelo identificar automáticamente qué noticias son más relevantes para la predicción, reduciendo el ruido de información no relevante. Esta extensión es particularmente prometedora dado que la precisión direccional moderada (41.8%) sugiere que el modelo actual no discrimina óptimamente entre señales fuertes y débiles.

Segundo, la implementación de técnicas de explicabilidad (SHAP values, LIME, attention visualization) mejoraría la transparencia de los mecanismos de decisión del LSTM. Para aplicaciones en instituciones financieras reguladas, la capacidad de explicar por qué el modelo generó una predicción específica—identificando qué noticias y qué características técnicas contribuyeron más—es esencial para cumplimiento regulatorio y gestión de riesgo. Esta dirección no solo abordaría preocupaciones prácticas, sino que generaría insights científicos sobre qué aspectos del sentimiento y precios históricos son más predictivos en diferentes contextos de mercado.

La integración de datos alternativos (posicionamiento institucional, flujos de opciones, análisis de redes sociales) y técnicas de aprendizaje continuo que permitan al modelo adaptarse automáticamente a cambios estructurales en el mercado representa la frontera de investigación. Un sistema que combine predicción de magnitud (fortaleza actual del modelo) con predicción de régimen de volatilidad (alta/media/baja) proporcionaría valor agregado sustancial para gestión de portafolios dinámicos. La implementación de backtesting riguroso con costos de transacción realistas y restricciones de liquidez validaría la viabilidad económica del modelo en aplicaciones de trading sistemático.

La priorización de estas extensiones debe balancear rigor científico con aplicabilidad práctica. Las extensiones inmediatas (ampliación temporal, múltiples fuentes, validación cruzada) son esenciales para publicación en journals académicos de alto impacto, donde la robustez y generalización son requisitos fundamentales. Las extensiones a mediano plazo (arquitecturas híbridas, explicabilidad) responden a necesidades de la industria financiera, facilitando transferencia tecnológica y potencial comercialización. La convergencia de ambas direcciones—investigación académica rigurosa e impacto práctico demostrable—maximizará la contribución de este trabajo al campo de la ciencia de datos aplicada a finanzas.

6 REFERENCIAS BIBLIOGRÁFICAS

- [1] Banco de Pagos Internacionales, «Triennial Central Bank Survey - OTC foreign exchange,» Monetary and Economic Department , 2022.
- [2] R. De Bock y I. de Carvalho Filho, «The behavior of currencies during risk-off episodes,» *Journal of International Money and Finance*, vol. 53, n° ISSN 0261-5606, pp. 218-234, 2015.
- [3] L. Brandt, A. Saint Guilhem, M. Schröder y I. V. Robays, «What drives euro area financial,» *European Central Bank*, pp. 1-57, 2021.
- [4] F. R. B. o. N. York, «Economic Policy Review (Federal Reserve Bank of New York),» *Economic Policy Review (Federal Reserve Bank of New York)*, vol. 26, n° number-4-628120, 2020.
- [5] R. H. Shumway y D. S. Stoffer, *Time Series Analysis and Its Applications With R Examples*, Cuarta Edición , 2017.
- [6] P. C. Tetlock, Giving Content to Buy and Sell Recommendations: The Role of Investor Sentiment, *Financial Analysts Journal*, vol. 63, no. 5, pp. 26-35, 2007.
- [7] F. S. Mishkin, *The Economics of Money, Banking, and Financial Markets*, 5th ed. Pearson, 2019.
- [8] S. Hochreiter y J. Schmidhuber, «Long Short-term Memory,» *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997, doi: 10.1162/neco.1997.9.8.1735., 1997.
- [9] B. P. a. L. Lee, «Opinion Mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval,» *Now Publishing Inc*, vol. vol. 2, n° DOI: 10.1561/1500000011., pp. no. 1–2, pp. 1-135, 2008.
- [10] F. M. Arias Jiménez, «Pronóstico del tipo de cambio colombiano - una aproximación desde las redes neuronales,» Bogota, 2005.
- [11] B. Maskey, «Smart Money Concepts in the Forex Market: A Strategy for Individual Traders,» B.S. thesis, Centria University of Applied Sciences, Bachelor of Business Management, Finlandia, 2021.
- [12] Daniel, «“Memoria a largo plazo a corto plazo (LSTM): ¿Qué es?,» *Data Scientist*, [Online]. Available: <https://datascientest.com/es/memoria-a-largo-plazo-a-corto-plazo-lstm>. [Accessed: 21-Nov-2024]., 2023.
- [13] R. H. S. a. D. S. Stoffer, *Time Series Analysis and Its Applications: With R Examples*, New York, NY, USA: 4th ed. : Springer DOI: 10.1007/978-1-4419-7865-3, 2017.
- [14] M. Sokolova y G. Lapalme, «A systematic analysis of performance measures for classification tasks,» *Information Processing & Management* vol. 45, pp. 427-437, Jul. 2009, doi: 10.1016/j.ipm.2009.03.002., 2009.
- [15] N. Sánchez Anzola, «Máquinas de soporte vectorial y redes neuronales artificiales en la predicción del movimiento USD/COP spot intradiario,» *La Revista ODEON*, n° DOI:<https://doi.org/10.18601/17941113.n9.04.>, pp. ODEON. 9 (feb. 2015), 113–172. DOI:<https://doi.org/10.18601/17941113.n9.04.>, 2015.

- [16] H. Ding, X. Zhao, Z. Jiang, S. N. Abdullah y D. A. Dewi, «EUR-USD Exchange Rate Forecasting Based on Information Fusion with Large Language Models and Deep Learning Methods,» Cornell University, <https://arxiv.org/abs/2408.13214>., 2024.
- [17] C. Hutto y E. Gilbert, «VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text,» *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, n° 1, pp. 216-225., 2014.
- [18] B. Pang y L. Lee, *Opinion Mining and Sentiment Analysis*, 2008.
- [19] G. E. P. Box y G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. Holden-Day, ISBN: 978-0816211043., 1976.
- [20] R. S. Tsay, *Analysis of Financial Time Series*. 3rd ed. John Wiley & Sons, ISBN: 978-0470414354., 2020.
- [21] R. A. Meese y K. Rogoff, «Empirical exchange rate models of the seventies: Do they fit out of sample?,» *Journal of International Economics*, Vols. %1 de %2DOI: 10.1016/0022-1996(83)90017-X., pp. 14(1-2), 3-24., 1983.
- [22] S. Omer Berat, M. Ugur Gudelek y A. Murat Ozbayoglu, «Financial time series forecasting with deep learning: A systematic literature review,» *Cornell University*, 2005–2019.
- [23] D. Alaminos, M. Belén Salas y M. Á. Fernández-Gámez, «Quantum Monte Carlo simulations for estimating,» *Humanities & Social Sciences Communications*, 2023.
- [24] R. J. Shiller, *Irrational Exuberance: Revised and Expanded Third Edition*, 3rd ed. Princeton University Press DOI: 10.2307, 2015.
- [25] D. J. a. J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models**, 3rd ed., Online manuscript, Stanford University, Aug. 20, 2024 Available: <https://web.stanford.edu/~jurafsky/slp3>., 2024.
- [26] B. D. a. W. Hu, «Sentiment Analysis of Investor Opinions on Twitter,» *Social Networking*, vol. 4, n° DOI: 10.4236/sn.2015.43008., pp. pp. 62-71, 2015.
- [27] C.-H. L. a. R.-P. H. M.-Y. Chen, «Modeling public mood and emotion: Stock market trend prediction with anticipatory computing approach,» *Computers in Human Behavior*, vol. 101, n° DOI: 10.1016/j.chb.2019.03.021., pp. 402-408, 2019.
- [28] E. Papaioannou y R. Portes, «The International Role of the Euro: A Status Report,» *Economic and Financial Affairs*, 2008.

7 ANEXOS

Anexo 1. REPOSITORIO DEL PROYECTO.

El código fuente completo de esta investigación está disponible públicamente en GitHub:

<https://github.com/Loocust/proyecto-grado>