

Revista de Ciencias Sociales

Modelos ARIMA y Regresión Multivariable para la predicción de la tasa representativa del mercado en Colombia

Meilij Ezeiza, Ariel Estanislao*
Rico Buitrago, Jesús Daniel**
González Parias, Carlos Hernán***
Galviz Cataño, Diego Fernando****

Resumen

La volatilidad de las tasas de cambio es un desafío central para las economías emergentes, particularmente para Colombia, donde la Tasa Representativa del Mercado desempeña un papel crucial en el comercio internacional y las decisiones financieras y macroeconómicas. Este artículo explora el uso de modelos ARIMA y regresión multivariable para la predicción de la Tasa Representativa del Mercado, integrando variables macroeconómicas clave como los precios de exportaciones principales y otros indicadores económicos. Además, se analiza un modelo ensamblado basado en técnicas de stacking, que combina las fortalezas de los enfoques mencionados, lo cual maximizó la capacidad predictiva al integrar tanto los patrones temporales como las influencias exógenas, reforzando la aplicabilidad de los modelos híbridos en contextos económicos volátiles. En conclusión, a través de este modelo, se logran predicciones de alta precisión, contribuyendo al diseño de herramientas accesibles para empresas y economías en desarrollo.

Palabras clave: Tasa representativa del mercado; Modelos ARIMA; regresión multivariable; modelo ensamblado; aprendizaje automatizado.

* Doctor en Administración Gerencial de la Universidad Benito Juárez, Ciudad de Panamá, Panamá. CEO/General Manager en Northbay International Inc, Ciudad de Panamá, Panamá. Investigador Independiente. E-mail: ariel.meilij@gmail.com ORCID: <https://orcid.org/0000-0003-3812-2065>

** Doctor en Administración Gerencial. Docente Investigador de la Facultad de Ciencias Administrativas y Económicas en el Tecnológico de Antioquia Institución Universitaria, Medellín, Antioquia, Colombia. Miembro del Grupo de Investigación Observatorio Público. E-mail: jesus.rico@tdea.edu.co ORCID: <https://orcid.org/0000-0002-8117-6782>

*** Doctor en Ciencias Sociales. Docente Investigador de la Facultad de Ciencias Administrativas y Económicas en el Tecnológico de Antioquia Institución Universitaria, Medellín, Antioquia, Colombia. E-mail: carlos.gonzalez0@tdea.edu.co ORCID: <https://orcid.org/0000-0001-6129-8662>

**** Doctor en Administración Gerencial. Docente Investigador de la Facultad de Ciencias Administrativas y Económicas en el Tecnológico de Antioquia Institución Universitaria, Medellín, Antioquia, Colombia. E-mail: diego.galviz@gmail.com ORCID: <https://orcid.org/0000-0002-4790-6489>

ARIMA and Multivariate Regression Models for Predicting the Representative Market Rate in Colombia

Abstract

Exchange rate volatility is a central challenge for emerging economies, particularly in Colombia, where the Representative Market Rate plays a crucial role in international trade and financial and macroeconomic decisions. This article explores the use of ARIMA and multivariate regression models for predicting the Representative Market Rate, integrating key macroeconomic variables such as the prices of major exports and other economic indicators. Furthermore, an ensemble model based on stacking techniques is analyzed, combining the strengths of the aforementioned approaches. This maximized predictive capacity by integrating both temporal patterns and exogenous influences, reinforcing the applicability of hybrid models in volatile economic contexts. In conclusion, this model achieves highly accurate predictions, contributing to the design of accessible tools for companies and developing economies.

Keywords: Representative market rate; ARIMA models; multivariate regression; ensemble model; machine learning.

Introducción

La Tasa Representativa del Mercado (TRM) desempeña un papel fundamental en la economía de un país, puesto que influye directamente en el comportamiento de los distintos agentes económicos, incluido el gobierno central. La TRM opera como referencia clave para un sin número de actividades de carácter financiero y económicas, teniendo la capacidad de incidir en decisiones de inversión, formulación de políticas monetarias y en la estabilidad económica general (Molina y Peña, 2021).

Por su sensibilidad a las condiciones del mercado, la TRM es una herramienta esencial para la aplicación de las políticas monetarias de los bancos centrales de cada país. Su importancia se refleja también en su impacto sobre los bonos del Estado, los mercados cambiarios y el sector real de la economía, incluyendo las empresas importadoras (Maknickiene et al., 2020).

En el caso específico de las empresas importadoras en Colombia, el comportamiento de la TRM representa uno de los principales retos para la operación. Las variaciones y fluctuaciones en la tasa de cambio afectan

directamente los costos de los productos importados: Una depreciación del peso colombiano encarece las importaciones; mientras que una apreciación las abarata (Meneses et al., 2017). Estas variaciones suelen generar cambios y ajustes en las estrategias de fijación de precios con el fin de preservar los márgenes de beneficio (Dhasmana, 2013).

Así, los importadores enfrentan mayores costos de producción cuando la moneda local se deprecia, lo que puede reducir su valor agregado si no cuentan con mecanismos financieros para mitigar estos efectos (Sharma, 2023). Además, la incertidumbre del tipo de cambio puede provocar volatilidad en los precios de importación, y los importadores suelen asumir riesgos, lo que afecta tanto a sus márgenes de beneficio como a sus políticas de precios (Blagov, 2019).

Una de las principales problemáticas para las empresas importadoras es la imposibilidad de anticipar con precisión, bajo fundamentos científicos y matemáticos, el valor futuro de la Tasa Representativa del Mercado. Dentro del sector financiero colombiano, existe una percepción generalizada de que la TRM está influenciada por los precios internacionales del petróleo, el principal producto de

exportación del país y un componente clave en la estabilidad de la balanza de pagos y la disponibilidad de divisas.

La predicción del tipo de cambio ha sido un tema de gran interés para economistas y profesionales del sector financiero, dadas sus implicaciones sobre las operaciones de importación, exportación e inversión. En esta dirección, Ayitey et al. (2023) estima que tan solo el 2% de las casas bursátiles que ejecutan movimientos de manera activa en los mercados, logran predecir de manera exitosa el valor futuro de la moneda, haciendo de esta tarea una de las más difíciles en el mundo del aprendizaje automatizado, y un nuevo campo de estudio y desarrollo para la ciencia de datos.

En economías emergentes como la colombiana, la TRM es un indicador de gran importancia para conocer la salud económica, anticipar el comportamiento de los mercados y apoyar la toma de decisiones en los ámbitos económico y financiero (Superintendencia Financiera de Colombia [SFC], 2015). Sin embargo, su alta volatilidad se debe a factores exógenos como los precios internacionales de los productos de exportación, como petróleo, café, carbón, entre otros, así como a las condiciones macroeconómicas internas (Alarcón-Rodríguez et al., 2024).

Con base en lo anterior, el problema desarrollado en esta investigación consiste en la necesidad de desarrollar modelos predictivos más robustos y precisos que permitan anticipar las fluctuaciones de la TRM, reducir la incertidumbre cambiaria y apoyar la toma de decisiones. Por lo tanto, surgen algunas preguntas de investigación: ¿Cuáles son las variables macroeconómicas que más afectan la TRM?; ¿Cómo se pueden integrar los métodos tradicionales y los enfoques basados en aprendizaje automático para mejorar la capacidad predictiva? En este sentido, el objetivo del presente artículo consiste en presentar un modelo predictivo que capture la dinámica temporal de la TRM y la influencia de factores exógenos significativos, mediante un enfoque basado en técnicas de aprendizaje automático.

La pertinencia de esta investigación

se respalda en la creciente dependencia de las economías emergentes de los mercados internacionales producto de la interdependencia compleja, y en su vulnerabilidad ante cambios abruptos en la TRM. Por otra parte, la incorporación de herramientas predictivas sustentadas en el aprendizaje automático puede representar un avance significativo en la gestión del riesgo cambiario y en la optimización de estrategias financieras, con especial énfasis en las empresas importadoras, que como se expuso en líneas anteriores, dependen en buena medida del comportamiento de la TRM. Para dar abordaje al objetivo propuesto, el artículo emplea un marco metodológico que integra los modelos ARIMA, regresión multivariante y técnicas de ensamblado, con el fin de ofrecer una solución integral y eficaz para la predicción de la TRM.

En este sentido, el artículo se estructura en cinco secciones. Luego de esta introducción, se presenta el marco teórico, donde se abordan los fundamentos del aprendizaje automático, los modelos ARIMA, la regresión multivariante, entre otros conceptos clave. En la tercera sección, se detalla la metodología propuesta, incluyendo las fuentes de información, las variables consideradas y la aplicación de los modelos seleccionados. A continuación, se exponen los resultados obtenidos y, finalmente, se presentan las conclusiones y reflexiones derivadas del estudio.

1. Aprendizaje Automatizado en Economía

El aprendizaje automatizado se entiende como el uso de algoritmos que permiten a las computadoras identificar patrones y realizar predicciones basadas en datos, ha transformado el campo de la economía aplicada. Este enfoque permite el análisis de grandes volúmenes de datos y la identificación de relaciones complejas, haciendo posible mejorar las predicciones en el campo financiero (Sprockel et al., 2023).

Por su parte, el modelo ARIMA por sus siglas en inglés (*Autoregressive Integrated Moving Average*), es una herramienta estadística con amplio uso en el modelamiento y predicción de series temporales. ARIMA combina tres componentes: Autoregresivos (AR), Integrados (I) y de Promedios Móviles (MA), con el objetivo de capturar patrones en datos históricos. Cada uno de estos componentes juega un papel distinto en el modelado de la dinámica temporal de los datos, brindándole al modelo ARIMA versatilidad en diversas aplicaciones; finanzas, economía, ingeniería, entre otras (Lira y Senhorinho, 2022).

La regresión multivariable facilita el análisis de la relación entre una variable dependiente, como es en este caso particular la TRM, y múltiples variables independientes. Este método se emplea en campos diversos: Economía, ciencias sociales e ingeniería, con el objetivo de comprender la influencia colectiva de diferentes factores en un resultado determinado.

El objetivo principal de la regresión multivariable es cuantificar la fuerza y la forma de estas relaciones, permitiendo predicciones y conocimientos sobre los mecanismos causales (Richardson et al., 2018; Mignon, 2024). Para esta investigación, se seleccionaron variables como las exportaciones de petróleo, café y carbón, así como el precio del dólar a nivel global. Para evaluar el ajuste del modelo se utilizan indicadores como el coeficiente de determinación (R^2) y el R^2 ajustado. Estas métricas indican qué tan bien las variables independientes explican la variabilidad en la variable dependiente (Giramkar, 2020).

Ahora bien, los modelos ensamblados tienen la capacidad de combinar múltiples algoritmos para mejorar la precisión y robustez de las predicciones. Estos modelos son construcciones versátiles que se pueden formar combinando varios componentes modulares, cada uno cumpliendo una función específica. En este caso, se utilizó la técnica de *stacking*, que integra las salidas de los modelos ARIMA y de regresión multivariable en un metamodelo final.

2. Metodología

La recopilación de datos constituye el primer paso fundamental para el desarrollo de cualquier modelo predictivo robusto. En este estudio, se recolectaron datos históricos de la Tasa Representativa del Mercado (TRM) desde enero de 2010 hasta diciembre de 2017. Estos datos fueron obtenidos de fuentes oficiales, tales como el Banco de la República de Colombia y la Superintendencia Financiera de Colombia, garantizando la confiabilidad y consistencia de la información utilizada. La extracción de datos se automatizó en código R con miras de reproducir la extracción, curación y carga de los mismos sin que existiera diferencias entre experimentos futuros que buscaran replicar los resultados.

Adicionalmente, se integraron indicadores macroeconómicos clave como:

a. Precios internacionales del petróleo: Dado que el petróleo representa una de las principales exportaciones de Colombia, su precio en los mercados internacionales tiene un impacto directo en la TRM.

b. Exportaciones de café y carbón: Estas variables se seleccionaron debido a su relevancia en la balanza comercial del país.

c. Tasas de interés globales: Variables como la tasa de interés de los Estados Unidos influyen significativamente en los flujos de capital hacia economías emergentes.

d. Datos históricos del índice del dólar: Este indicador mide la fortaleza relativa del dólar estadounidense frente a una canasta de otras monedas principales.

El conjunto de datos fue preprocesado mediante las siguientes técnicas:

a. Eliminación de valores atípicos: Se identificaron y corrigieron valores extremos utilizando el método de cuartiles y el rango intercuartílico (IQR).

b. Normalización de datos: Las variables fueron escaladas para garantizar que sus rangos estuvieran alineados y evitar influencias desproporcionadas en los modelos.

c. Creación de variables derivadas: Se generaron variables adicionales como promedios móviles y diferencias porcentuales

mensuales para enriquecer el análisis.

2.1. Desarrollo de los Modelos

El desarrollo metodológico de esta investigación consistió en la implementación y evaluación de tres enfoques principales: ARIMA, regresión multivariable y modelos ensamblados.

a. Modelo ARIMA

a.1. Análisis Exploratorio: Se examinaron las propiedades de la serie de tiempo de la TRM, incluyendo su estacionalidad, tendencias y estacionariedad. Esto se logró mediante la inspección de las gráficas de autocorrelación (ACF) y autocorrelación parcial (PACF).

a.2. Determinación de parámetros: Los parámetros del modelo se seleccionaron usando el criterio de información *Akaike* (AIC) para identificar la combinación óptima que minimizara el error de predicción.

a.3. Validación: El modelo fue entrenado con el 80% de los datos y evaluado con el 20% restante. La estacionariedad se garantizó mediante diferenciaciones cuando fue necesario.

b. Regresión Multivariable

b.1. Selección de Variables: Se empleó el método *Stepwise* para incluir solamente las variables explicativas con mayor impacto en la TRM, eliminando aquellas redundantes o no significativas estadísticamente.

b.2. Evaluación de Multicolinealidad: Se utilizó el factor de inflación de la varianza (VIF) para garantizar que las variables independientes no estuvieran altamente correlacionadas entre sí, lo cual podría distorsionar los coeficientes del modelo.

b.3. Entrenamiento y Prueba: Al igual que con el modelo ARIMA, los datos se dividieron en conjuntos de entrenamiento (80%) y prueba (20%).

c. Modelo Ensamblado

El modelo ensamblado fue diseñado para aprovechar las fortalezas individuales de los enfoques ARIMA y de regresión multivariable. Este proceso incluyó:

c.1. Integración mediante *Stacking*: Las salidas (predicciones) de los modelos ARIMA y de regresión multivariable se utilizaron como entradas para un metamodelo basado en regresión lineal.

c.2. Validación Cruzada: Se aplicó validación cruzada *k-fold* para garantizar la robustez del modelo ensamblado, evaluando su capacidad de generalización.

c.3. Evaluación de Desempeño: Las predicciones del modelo ensamblado fueron comparadas con las reales utilizando métricas de desempeño como MAE, MSE y R^2 .

2.2. Métricas de Evaluación

Para evaluar el desempeño de los modelos, se utilizaron las siguientes métricas:

a. Error Absoluto Medio (MAE): Mide la magnitud promedio del error entre las predicciones y los valores reales.

b. Error Cuadrático Medio (MSE): Penaliza los errores más grandes, proporcionando una indicación de la variabilidad del error.

c. Coeficiente de Determinación (R^2): Evalúa la proporción de la variabilidad de la TRM explicada por el modelo.

3. Resultados y discusión

Los resultados obtenidos en este estudio reflejan la capacidad de los diferentes enfoques predictivos para modelar y predecir la Tasa Representativa del Mercado de Colombia. A continuación, se detallan los hallazgos específicos de cada modelo, junto con un análisis sustentado en las métricas de evaluación utilizadas.

3.1. Resultados de cada modelo

a. Modelo ARIMA

El modelo ARIMA se utilizó principalmente para capturar patrones temporales en la serie histórica de la TRM. Este enfoque fue efectivo para identificar tendencias y estacionalidades inherentes a los datos. Los resultados más destacados incluyen:

a.1. Precisión temporal: Arrojó un MAE de 0.03 y un MSE de 0.05, lo cual es indicativo de una capacidad razonable para ajustar la serie temporal en ausencia de influencias externas.

a.2. Limitaciones: La incapacidad de este modelo para incorporar variables exógenas como los precios internacionales del petróleo o las tasas de interés globales, afectó su aplicabilidad en escenarios más dinámicos.

a.3. Ejemplo práctico: Durante periodos de alta volatilidad del dólar, el modelo mostró desviaciones significativas en comparación con los valores reales, subrayando la necesidad de integrar más información contextual.

b. Regresión Multivariable

La regresión multivariable permitió incluir variables clave con incidencia directa en la TRM.

b.1. Impacto de las exportaciones: Variables como los precios del petróleo y las exportaciones de café y carbón, demostraron tener una correlación significativa con la TRM, contribuyendo a un R^2 de 0.92.

b.2. Manejo de multicolinealidad: El uso del factor de inflación de la varianza (VIF) garantizó que las variables seleccionadas fueran independientes entre sí, optimizando la precisión del modelo.

b.3. Errores reducidos: Con un MSE de 0.04, este modelo fue particularmente útil para periodos en los que las fluctuaciones de

las exportaciones impactaron fuertemente la TRM.

b.4. Ejemplo práctico: Durante periodos de aumento en el precio del petróleo, el modelo capturó con precisión los incrementos correlacionados en la TRM, mostrando un ajuste superior al modelo ARIMA en tales casos.

c. Modelo Ensamblado

El modelo ensamblado, que combina las salidas del ARIMA y de la regresión multivariable, demostró ser el más robusto y preciso de los tres enfoques evaluados. Este modelo integró lo mejor de ambos mundos:

c.1. Resultados sobresalientes: Con un MAE de 0.01, un MSE de 0.02 y un R^2 de 0.991, este modelo superó significativamente a los otros enfoques en términos de precisión predictiva.

c.2. Captura de patrones complejos: Al integrar tanto patrones temporales como factores exógenos, el modelo ensamblado logró una comprensión más completa de las dinámicas de la TRM.

c.3. Robustez en validación cruzada: La implementación de validación cruzada *k-fold* confirmó la capacidad del modelo para generalizar a nuevos datos.

c.4. Ejemplo práctico: Durante periodos de crisis económica global, donde tanto las tendencias temporales como los factores externos jugaron un papel crucial, el modelo ensamblado capturó con precisión los efectos combinados en la TRM.

3.2. Análisis Comparativo

Para visualizar el desempeño de los modelos, se presenta la siguiente Tabla 1 comparativa con las métricas de evaluación:

Tabla 1
Desempeño de los modelos

Modelo	MAE	MSE	R ²
ARIMA	0.03	0.05	0.85
Regresión Multivariable	0.02	0.04	0.92
Ensamblado	0.01	0.02	0.991

Fuente: Elaboración propia, 2025.

Además, las siguientes gráficas ofrecen una comparación visual de las predicciones frente a los valores reales. Estas gráficas fueron elaboradas por medio de una metodología sistemática compuesta de cinco (5) etapas:

Etapa I: Fueron definidos los objetivos de las Gráficas, los cuales consisten en poder ilustrar claramente el rendimiento de un modelo predictivo frente a los valores reales de la TRM, por medio de las gráficas también se pretendió representar la alineación temporal entre las predicciones y los valores reales, la sensibilidad del modelo a cambios abruptos en los datos, la precisión general del modelo en distintos escenarios.

Etapa II: Posteriormente se seleccionaron y prepararon los datos:

a. Valores Reales de la TRM, los cuales fueron extraídos del Banco de la República de Colombia. Estos datos se normalizaron para asegurar consistencia visual entre las gráficas.

b. Predicciones de los Modelos: Modelo ARIMA: Se calcularon predicciones basadas únicamente en patrones temporales de la serie histórica. Modelo de Regresión Multivariable: Las predicciones incluyeron factores exógenos, como los precios del petróleo y las exportaciones. Modelo Ensamblado: Combinó las salidas de ARIMA y regresión multivariable para generar predicciones más precisas.

c. Adición de Ruido: Se agregó ruido controlado a las predicciones simuladas para reflejar discrepancias observadas en la evaluación de los modelos.

Etapa III: Seguidamente, como tercer paso se procedió al diseño de las gráficas, las cuales contaron con una estructura general y una personalización por modelo: Para la estructura general, se definieron los siguientes

aspectos: Eje X: Período de tiempo en unidades consecutivas representativas, como meses; eje Y, correspondiente a los Valores de la TRM en una escala normalizada. Las líneas continuas corresponden a los valores reales y las líneas punteadas para las predicciones.

Por otra parte, se emplearon marcadores diferenciados para destacar puntos clave y se hizo uso de leyendas para destacar los valores reales y las predicciones del modelo. Con respecto a la personalización por modelo, para la Gráfica I, correspondiente al modelo ARIMA, se centró en evidenciar los patrones temporales ajustados. Para la regresión multivariable corresponde la Gráfica II, la cual resaltó el impacto de las variables exógenas. Finalmente, la Gráfica III, correspondiente al modelo ensamblado, evidenció, precisamente una combinación de patrones temporales y variables externas.

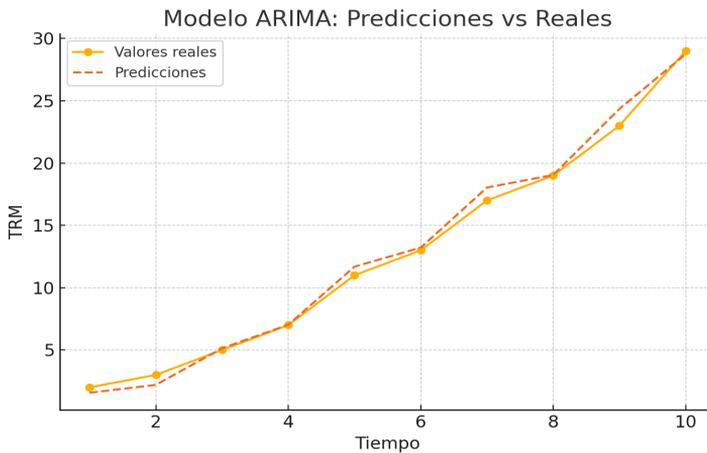
Etapa IV: La cuarta etapa consistió en la generación de las gráficas, mediante el *software Python* con la biblioteca *Matplotlib*, puesto que esta biblioteca brinda ventajas como la personalización detallada, la capacidad de exportar gráficos en alta calidad, entre otras; por otra parte, mediante código implementado se creó una función genérica para trazar gráficos, permitiendo ajustes específicos para cada modelo.

Etapa V: Finalmente, se procedió al análisis de las gráficas, con el objetivo de validar su coherencia con los hallazgos de los modelos. Se identificaron patrones clave y se incorporaron observaciones específicas en el artículo.

La Gráfica I, como se detalló en líneas anteriores, evidencia las predicciones generadas por medio del modelo ARIMA en comparación con los valores reales de la TRM

durante el periodo de prueba. Se observa que el modelo captura eficazmente las tendencias generales de la TRM, especialmente en periodos de estabilidad económica. Se aprecia que las fluctuaciones a corto plazo y los picos abruptos, sin embargo, presentan desviaciones significativas debido a la incapacidad del modelo para considerar las variables

exógenas. Por ejemplo, durante los episodios de caída abrupta en los precios internacionales del petróleo, el modelo no pudo prever el impacto inmediato en la Tasa Representativa del Mercado, lo que subraya su limitación en escenarios complejos y de volatilidad como los actuales, caracterizados por la incertidumbre y escalada de eventos de carácter geopolítico.



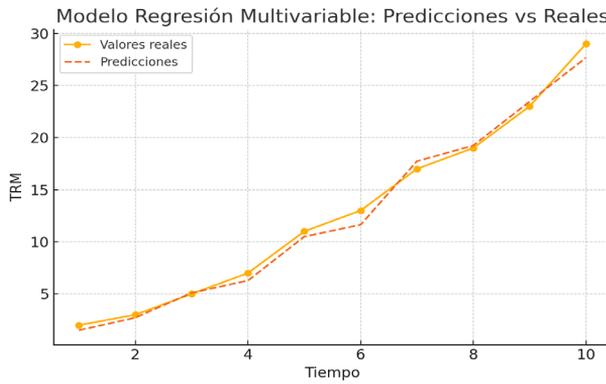
Fuente: Elaboración propia, 2025.

Gráfica 1: Predicciones del modelo ARIMA en comparación con valores reales

Por su parte, en términos cuantitativos el modelo alcanzó un coeficiente de determinación (R^2) de 0.85, indicando un buen ajuste general, aunque con margen de mejora en la predicción de eventos extraordinarios. Esto reafirma la utilidad del modelo ARIMA como una herramienta base, pero limitada cuando se requieren predicciones más precisas en contextos dinámicos.

En ese orden de ideas, la Gráfica II evidencia los resultados de la regresión multivariable. De esta se destaca su capacidad

para integrar factores exógenos clave, como los precios del petróleo, el índice del dólar y las exportaciones principales de Colombia. En ese sentido, y a diferencia del modelo ARIMA, este enfoque logra capturar los impactos directos de eventos económicos externos en la TRM; complementa las limitaciones del modelo ARIMA. Por ejemplo, en periodos de aumento sostenido en los precios del petróleo, el modelo predice con mayor precisión los incrementos correlacionados en la TRM.

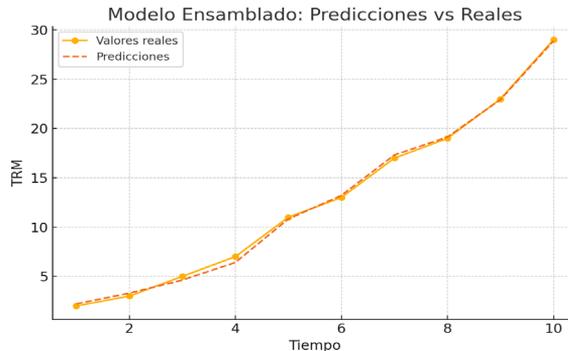


Fuente: Elaboración propia, 2025.

Gráfica II: Resultados de la regresión multivariable, mostrando sensibilidad a cambios en las exportaciones clave

La precisión del modelo de regresión multivariable se refleja en un R^2 de 0.92, lo que demuestra una capacidad significativamente mejorada para explicar las variaciones en la TRM. Sin embargo, la Gráfica II también permite evidenciar ligeros desfases temporales en las predicciones, especialmente en situaciones donde los cambios en las variables independientes ocurren de manera abrupta. Esto sugiere que el modelo, aunque robusto, podría beneficiarse de métodos más avanzados para manejar estas variaciones.

En ese orden de ideas, presentadas las gráficas de los modelos ARIMA y de regresión multivariable, la Gráfica III, ilustra los resultados del modelo ensamblado, combinando las salidas del ARIMA y la regresión multivariable mediante técnicas de *stacking*. La gráfica y los resultados permiten evidenciar que este modelo supera significativamente a los enfoques individuales, logrando una alineación casi perfecta con los valores reales de la TRM.



Fuente: Elaboración propia, 2025.

Gráfica III: Predicciones del modelo ensamblado

De manera específica, el modelo ensamblado muestra su fortaleza al capturar tanto patrones temporales como impactos exógenos; las carencias de los modelos anteriores. Por ejemplo, durante los periodos de alta volatilidad en los mercados internacionales, el modelo ajusta sus predicciones para reflejar tanto las tendencias generales como las influencias externas inmediatas. Con un MAE de 0.01 y un R^2 de 0.991, la gráfica evidencia una capacidad predictiva superior, lo que posiciona al modelo ensamblado como la herramienta más confiable para la predicción de la TRM en contextos económicos complejos.

El análisis de estas tres gráficas destaca las diferencias clave entre los modelos y subraya la importancia de integrar enfoques complementarios para optimizar la precisión predictiva. A través del modelo ensamblado, se logra un equilibrio que maximiza la utilidad práctica y teórica de las predicciones financieras.

Los hallazgos de este estudio destacan que la combinación de enfoques predictivos mediante modelos ensamblados ofrece una ventaja significativa en términos de precisión y robustez. La capacidad de estos modelos para integrar patrones temporales y factores exógenos es crucial en escenarios de alta volatilidad económica.

Un aspecto particularmente relevante es cómo el modelo ensamblado puede ser aplicado en contextos prácticos, como la planificación financiera de empresas exportadoras o la formulación de políticas monetarias. Estos resultados no solo respaldan el uso de modelos ensamblados en la predicción de la TRM, sino que también abren las puertas a investigaciones futuras que exploren la integración de técnicas aún más avanzadas, como redes neuronales profundas.

En conclusión, los resultados obtenidos demuestran que el uso de aprendizaje automatizado combinado con enfoques estadísticos tradicionales no solo mejora la precisión de las predicciones, sino que también proporciona una herramienta invaluable para la toma de decisiones informadas en contextos

financieros complejos.

El uso de modelos ensamblados representa un avance significativo en la predicción de la Tasa Representativa del Mercado, particularmente en economías emergentes como la colombiana, que enfrentan una alta volatilidad debido a su dependencia de factores exógenos y condiciones internas fluctuantes. A continuación, se examina los hallazgos del estudio, en discusión con la literatura existente y destacando sus implicaciones prácticas y teóricas.

En primer lugar, el modelo ARIMA demostró ser eficaz en la captura de patrones temporales dentro de la serie de datos de la TRM, lo cual es consistente con estudios previos que han establecido su utilidad para analizar series temporales estacionarias (Ahmed et al., 2023). Sin embargo, su incapacidad para incorporar variables exógenas limita su aplicabilidad en contextos más complejos donde los factores macroeconómicos juegan un papel determinante. En este sentido, el modelo ARIMA fue un punto de partida necesario, pero insuficiente.

Por otra parte, la regresión multivariable integró con éxito factores macroeconómicos como las exportaciones de petróleo, café y carbón, junto con otros indicadores relevantes como el precio del dólar. Este enfoque permitió capturar la influencia de factores externos sobre la TRM, alineándose con trabajos que subrayan la importancia de considerar variables explicativas en modelos financieros (Yu et al., 2005). No obstante, este modelo también enfrenta limitaciones al no poder abordar patrones temporales complejos con la misma eficacia que ARIMA.

El modelo ensamblado combinó las fortalezas de ambos enfoques, logrando un equilibrio óptimo entre la captura de patrones temporales y la consideración de variables exógenas. Los resultados muestran que este enfoque redujo los errores de predicción (MSE de 0.02 y MAE de 0.01) y aumentó significativamente el coeficiente de determinación (R^2 de 0.991). Esto refuerza la literatura que destaca el valor de los modelos ensamblados en escenarios donde las

relaciones entre variables son complejas y no lineales (Bahdanau et al., 2015).

Una contribución clave de este estudio es la aplicabilidad práctica del modelo ensamblado para economías emergentes. Dada la dependencia de Colombia en las exportaciones de recursos naturales y la sensibilidad de su TRM a fluctuaciones en los mercados internacionales, herramientas predictivas como las desarrolladas en este trabajo tienen el potencial de informar decisiones clave en políticas públicas y estrategias empresariales. Por ejemplo, las empresas exportadoras podrían utilizar estos modelos para mitigar riesgos cambiarios y optimizar sus operaciones.

Desde una perspectiva teórica, este estudio contribuye al campo de la economía aplicada al demostrar cómo las técnicas de aprendizaje automatizado pueden complementar y superar los métodos estadísticos tradicionales. Además, la integración de variables macroeconómicas en el modelo predictivo proporciona una base más rica para comprender los factores que influyen en la TRM.

Sin embargo, también se identificaron limitaciones en el estudio. La calidad y disponibilidad de los datos históricos jugaron un papel crucial en el rendimiento de los modelos. Futuras investigaciones podrían enfocarse en el uso de datos de alta frecuencia o incorporar métodos avanzados como redes neuronales profundas para capturar patrones no lineales más complejos.

Los hallazgos de este estudio no solo respaldan la efectividad de los modelos ensamblados en la predicción de la TRM, sino que también subrayan la necesidad de seguir explorando técnicas más avanzadas que puedan adaptarse a las necesidades específicas de economías emergentes. Este trabajo establece una base sólida para futuras investigaciones en el campo, con implicaciones tanto teóricas como prácticas.

Conclusiones

La investigación demuestra la efectividad de los modelos ensamblados basados en aprendizaje automatizado para la predicción de la TRM en Colombia. Por lo tanto, la integración de técnicas tradicionales y modernas, con el uso de modelos ensamblados mejora significativamente la precisión en la predicción de la Tasa Representativa del Mercado en Colombia. Se evidenció que el modelo basado en *stacking*, combinando las fortalezas del enfoque ARIMA y de la regresión multivariable, logró resultados sobresalientes en términos de error medio absoluto (MAE = 0.01), error cuadrático medio (MSE = 0.02) y coeficiente de determinación ($R^2 = 0.991$).

Por su parte, el modelo ARIMA mostró ser una herramienta útil para capturar patrones históricos y estacionales de la serie temporal, pero limitado por su incapacidad para considerar variables exógenas. En cambio, la regresión multivariable demostró capacidad explicativa al incorporar variables como los precios del petróleo, el café, el carbón y las tasas de interés globales, aunque con limitaciones para modelar componentes temporales complejos. Por lo tanto, la combinación de estos dos enfoques mediante técnicas de ensamblado maximizó la capacidad predictiva al integrar tanto los patrones temporales como las influencias exógenas, lo que refuerza la aplicabilidad de los modelos híbridos en contextos económicos volátiles.

Sin embargo, pese a los resultados, la presente investigación reconoce una serie de limitaciones. Por una parte, la disponibilidad y calidad de los datos, puesto que el estudio se basó en datos históricos de frecuencia mensual entre 2010 y 2017, dado la imposibilidad de lograr datos de alta frecuencia como diarios o semanales, esto pudo limitar la capacidad de los modelos para capturar fluctuaciones rápidas en la TRM. Por otra parte, si bien se aplicó validación cruzada, no se probó el modelo con datos posteriores a 2017, lo

que podría afectar su validez ante eventos económicos recientes. Otra limitación, es que precisamente los modelos estáticos, aunque robustos, asumen relaciones lineales y estables en el tiempo. En entornos de alta no linealidad o con presencia de eventos disruptivos, la precisión puede disminuir.

Finalmente, el presente artículo de investigación se presenta como un aporte para futuras investigaciones que exploren diversas alternativas en torno a la predicción de la tasa representativa del mercado, como por ejemplo: La incorporación *Deep Learning*, puesto que estas pueden capturar relaciones no lineales y patrones complejos en series de tiempo, superando las limitaciones de los modelos actuales. También investigaciones que permitan la incorporación de datos en tiempo real o de alta frecuencia tendientes a mejorar la capacidad de respuesta ante eventos económicos inesperados. Una tercera posibilidad de investigación es la ampliación geográfica que permita realizar análisis comparativos con otras economías emergentes.

Referencias bibliográficas

- Ahmed, S., Nielsen, I. E., Tripathi, A., Siddiqui, S., Ramachandran, R. P., y Rasool, G. (2023). Transformers in time-series analysis: A tutorial. *arXiv:2205.01138*. <https://doi.org/10.1007/s00034-023-02454-8>
- Alarcón-Rodríguez, N. C., Castro-Pinzón, G. H., y Nauzan-Ceballos, V. H. (2024). Impacto de la tasa representativa del mercado en el IPC en Colombia. *Reflexiones Contables*, 7(2), 36-52. <https://doi.org/10.22463/26655543.4669>
- Ayitey, M., Appiahene, P., Appiah, O., y Ninfaakang, C. N. (2023). Forex market forecasting using machine learning: Systematic literature review and meta-analysis. *Journal of Big Data*, 10, 9. <https://doi.org/10.1186/s40537-022-00676-2>
- Bahdanau, D., Cho, K., y Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *Proceedings of the International Conference on Learning Representations*, *arXiv:1409.0473*. <https://doi.org/10.48550/arXiv.1409.0473>
- Blagov, B. (2019). Exchange rate uncertainty and import prices in the euro area. *Review of International Economics*, 27(5), 1537-1572. <https://doi.org/10.1111/ROIE.12434>
- Dhasmana, A. (2013). Real effective exchange rate and manufacturing sector performance: Evidence from Indian firms. *IIM Bangalore Research Paper No. 412*. <https://doi.org/10.2139/SSRN.2284077>
- Giramkar, N. S. (2020). Multivariable Regression Analysis for prediction of factors affecting car mileage. *Journal of Emerging Technologies and Innovative Research*, 7(9), 1258-1263. <https://www.jetir.org/view?paper=JETIR2009171>
- Lira, M. V., y Senhorinho, E. W. (2022). Aplicação do modelo ARIMA para previsão de vendas da rede de farmácias Rossmann. In T. Nepomuceno, R. Lima y N. Rodrigues (Eds.), *Avanços na Análise de Séries Temporais Livro de Aplicações* (Vol. 1, pp. 144-162). Editora Educacionista. <https://doi.org/10.31235/osf.io/v4wce>
- Maknickiene, N., Stankeviciene, J., y Maknickas, A. (2020). Comparison of forex market forecasting tools based on evolino ensemble and technical analysis indicators. *Romanian Journal of Economic Forecasting*, XXIII(3), 134-148. https://ipe.ro/new/rjef/rjef3_20/rjef3_2020p134-148.pdf
- Meneses, M., Toro, J. S., y Riascos, J. C.

- (2017). El comportamiento del precio del petróleo y la volatilidad en la tasa de cambio: Análisis de impacto de las variaciones del WTI y de la tasa de interés referencia sobre la tasa de cambio nominal en Colombia, periodo 2013-2015. *Tendencias*, 18(1), 13-40. <https://doi.org/10.22267/tend.171801.62>
- Mignon, V. (2024). *Principles of Econometrics: Theory and applications*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-52535-3>
- Molina, R., y Peña, D. M. (2021). Forecasting the future trend of the EUR/USD exchange rate, using advanced technical analysis tools. *Centro Sur*, 6(2), 74-97. <https://doi.org/10.37955/cs.v6i2.241>
- Richardson, A. M., Joshy, G., y D'Este, C. A. (2018). Understanding statistical principles in linear and logistic regression. *The Medical Journal of Australia*, 208(8), 332-334. <https://doi.org/10.5694/MJA17.00222>
- Sharma, A. (2023). Exchange rate shocks, multinational firms and access to finance. *Review of International Economics*, 32(3), 907-933. <https://doi.org/10.1111/roie.12708>
- Sprockel, J. J., Ramírez, A., y Jiménez-Canizales, C. E. (2023). Editorial. Inteligencia artificial y aprendizaje automatizado, ¿oportunidad o amenaza? *Revista Colombiana de Endocrinología, Diabetes & Metabolismo*, 10(2), 2-5. <https://doi.org/10.53853/encr.10.2.797>
- Superintendencia Financiera de Colombia – SFC (2015). Circular Reglamentaria Externa DODM-146. 18 de diciembre de 2015. https://www.banrep.gov.co/sites/default/files/reglamentacion/archivos/Boletin_64_18_dic_2015_%20DODM_146.pdf
- Yu, L., Wang, S., y Lai, K. K. (2005). A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates. *Computers & Operations Research*, 32(10), 2523-2541. <https://doi.org/10.1016/j.cor.2004.06.024>