



<https://creativecommons.org/licenses/by/4.0/>

PREDICCIÓN DE ACTIVOS FINANCIEROS USANDO MODELOS ARIMA Y REDES NEURONALES AUTORREGRESIVAS

Comparative Analysis of Forecasts for Prices and Returns of Financial Assets

JOHAN ANDRÉS URIBE ESCUDERO¹,
LIZ JOHANA RAMÍREZ CASTAÑO²,
ROMARIO ADEMIR CONTO LÓPEZ³

Recibido: 8 de enero de 2024. Aceptado: 30 de enero de 2025

DOI: <https://doi.org/10.21017/rimci.1128>

RESUMEN

En este trabajo son aplicados diferentes métodos de pronóstico para predecir los precios y rendimientos de las acciones para dos de las principales empresas que transan en la bolsa de valores de Colombia: Bancolombia y Ecopetrol. Para el estudio son consideradas técnicas de pronóstico basadas en modelos estadísticos y de aprendizaje automático, utilizando métricas de evaluación como el RMSE, MAE y MAPE. Los modelos se han seleccionado debido a su amplia aplicación en el campo de las finanzas y su capacidad para capturar diferentes características de las series de tiempo. Para los precios de Bancolombia y Ecopetrol, el modelo híbrido que combina ARIMA y redes neuronales proporcionó los mejores resultados. Sin embargo, para las predicciones de los rendimientos el modelo de red neuronal resultó ser el más efectivo en ambos casos. Esto sugiere que la dinámica de cada acción y las características específicas de los datos afectan el desempeño de los modelos.

Palabras clave: precios de activos; rendimientos de activos; ARIMA; ARMA; GARCH; redes neuronales autorregresivas.

ABSTRACT

In this study, various forecasting methods are employed to predict the prices and returns of stocks for two of the primary companies listed on the Colombian stock exchange: Bancolombia and Ecopetrol. The study considers forecasting techniques based on statistical models and machine learning, using evaluation metrics such as RMSE, MAE, and MAPE. The models were selected due to their widespread application in the field of finance and their ability to capture different characteristics of time series data, for the stock prices of Bancolombia and Ecopetrol, the hybrid model combining ARIMA and neural networks provided the best results. However, for predicting the returns of both stocks, the neural network model proved to be the most effective. This suggests that the dynamics of each stock and the specific characteristics of the data influence the performance of the models.

Key words: asset prices; asset returns; ARIMA; ARMA; GARCH; Autoregressive neural networks.

-
- 1 Ingeniero Financiero; Tecnólogo en análisis de costos y presupuestos, Instituto Tecnológico Metropolitano, Medellín, Colombia. ORCID: <https://orcid.org/0000-0002-2228-0061> Correo electrónico: johanuribe245977@correo.itm.edu.co
 - 2 Estudiante de Ingeniería Financiera; Tecnóloga en análisis de costos y presupuestos, Instituto Tecnológico Metropolitano, Medellín, Colombia. ORCID: <https://orcid.org/0000-0002-3817-7616> Correo electrónico: lizramirez105996@correo.itm.edu.co
 - 3 Ingeniero Administrador; Magíster en Estadística; Ph(c) en Ingeniería Industria y Organizaciones, Universidad Nacional de Colombia, Medellín, Colombia. Docente ocasional, Departamento de finanzas, Instituto Tecnológico Metropolitano, Medellín, Colombia. ORCID: <https://orcid.org/0000-0002-9944-137X> Correo electrónico: romarioconto@itm.edu.co

I. INTRODUCCIÓN

DESDE UN punto de vista dinámico, los mercados financieros se han consolidado como una herramienta para los empresarios que buscan maximizar sus beneficios a través del mercado bursátil, no solo porque ofrecen oportunidades para la inversión y el crecimiento de capitales, sino que también les permite acceder a financiamiento, invertir en nuevos proyectos y mejorar sus operaciones. Sin embargo, es fundamental mencionar que se enfrentan al riesgo inherente de experimentar volatilidades tanto altas como bajas. Esta naturaleza fluctuante de los mercados responde a una variedad de factores económicos, como cambios en las tasas de interés, políticas monetarias, eventos geopolíticos y tendencias del mercado.

Estas variaciones pueden impactar significativamente el rendimiento de las inversiones y, por ende, los beneficios que esperan recibir al momento de la inversión. Para mitigar estos riesgos, es crucial que desarrollen estrategias tales como, la diversificación del portafolio, el análisis de los datos y la gestión activa de los riesgos.

Las principales dificultades al tomar decisiones de inversión en el mercado bursátil radican en la anticipación y predicción de las rentabilidades futuras de los precios de las acciones. Esto se debe al alto grado de incertidumbre generado por las fluctuaciones del mercado, las variables económicas y otros factores imprevistos que pueden afectar el rendimiento de las inversiones. Además, se considera la volatilidad y la covarianza, que son fundamentales para medir los riesgos asociados a la compra y venta de activos, así como para determinar su rentabilidad potencial[1].

Por lo tanto, el pronóstico de precios de las acciones desempeña un papel importante en los mercados financieros, dado que, al realizar estimaciones eficaces, se pueden obtener considerables beneficios financieros para los inversores. En todas estas operaciones existe un riesgo de pérdida debido a una equivocada toma de decisiones; por ende, es imprescindible emplear técnicas que permitan prever de la manera más adecuada posible el comportamiento futuro de las acciones[2].

En la literatura actual existen diferentes métodos de pronóstico aplicados a precios de acciones,

como lo son los modelos Autorregresivos de Medias Móviles Integrados (ARIMA) dentro de la metodología Box-Jenkins y más recientemente el uso de las redes neuronales artificiales autorregresivas dentro de las técnicas de *Machine Learning*. Así mismo, para series de rendimientos están disponibles los modelos autorregresivos de medias móviles (ARMA), para filtrar la media, en conjunto con el modelo Autorregresivo Condicional Heteroscedástico Generalizado (GARCH), el cual modela la varianza de los datos[3].

El modelo ARIMA, por otro lado, amplía el ARMA al incluir un componente integrado (I), que diferencia la serie temporal para lograr estacionariedad, eliminando tendencias, lo que permite un análisis más preciso de series temporales no estacionarias[3].

En el caso colombiano, las grandes empresas con grandes expectativas de rentabilidad están logrando un mayor acercamiento e impacto en los mercados financieros a nivel global con el objetivo de volverse más atractivas para los inversionistas. Los grupos empresariales en Colombia han sido actores clave desde la segunda mitad del siglo XX, cuando adoptaron una nueva estructura organizativa que permitió la internalización del capital proporcionado por una nueva legislación financiera promovida por el Banco Mundial[4].

Por lo anterior, este trabajo tiene como objetivo modelar y predecir los precios y rendimientos de dos empresas colombianas, Ecopetrol y Bancolombia, ambas grandes empresas que representan una porción significativa de los ingresos del país y que tienen una alta capitalización de mercado, es decir, el valor total de sus acciones es significativamente mayor en comparación con otras compañías que cotizan en la Bolsa de Valores de Colombia (BVC). Convirtiéndolas en líderes en el mercado bursátil, además de que estas empresas hacen parte del índice COLCAP, que mide el desempeño de las acciones más líquidas y representativas en la BVC, factores que les confieren una posición de liderazgo y las convierten en referencias obligadas para inversionistas en Colombia y en el extranjero.

Para la predicción de precios se utilizarán los modelos ARIMA, redes neuronales autorregresivas y un modelo híbrido resultante de la combi-

nación de estos dos mientras que para la predicción de rendimientos se emplearán los modelos ARMA-GARCH y ARMA con redes neuronales.

Esta investigación aborda la compleja dinámica que enfrentan las grandes empresas colombianas en la planificación de sus inversiones en el mercado financiero, una tarea esencial en un entorno económico cada vez más volátil. Al centrarse en modelos de predicción, como ARIMA, ARMA, GARCH y redes neuronales artificiales, este estudio no solo contribuye a mejorar la precisión en las proyecciones de precios y rendimientos financieros, sino que también ofrece un recurso práctico para la toma de decisiones estratégicas en el sector empresarial.

Al aplicar y comparar distintos enfoques cuantitativos de pronóstico, este estudio aporta un conocimiento valioso sobre la capacidad de estos modelos para reducir la incertidumbre en las inversiones y, en consecuencia, contribuye al desarrollo de metodologías adaptadas a los mercados emergentes.

Este artículo se presenta de la siguiente manera: en la sección 2 se muestra el marco referencial en el cual se describen los modelos de pronósticos y las medidas de evaluación, además del estado del arte. La sección 3 describe la metodología utilizada en este trabajo, la sección 4 presenta la descripción de los datos y en la sección 5 se muestran los resultados obtenidos.

II. MARCO REFERENCIAL

2.1 Marco teórico

La presencia de Bancolombia y Ecopetrol en la Bolsa de Valores ha sido fundamental para transformar el mercado de capitales colombiano, liderando el camino en la internacionalización y estableciendo nuevos referentes de desempeño. Al ser las primeras empresas colombianas en cotizar en la Bolsa de Valores de Nueva York, no solo ampliaron sus horizontes, sino que también abrieron camino a otras empresas nacionales. En este contexto, la precisión en los pronósticos de precios y rendimientos se vuelve crucial para tomar decisiones de inversión estratégicas y maximizar los retornos en un entorno cada vez más competi-

tivo. A continuación, se expondrán los modelos de pronóstico que serán utilizados:

2.1.1 Modelo ARMA

Los modelos ARMA son una extensión de los modelos AR y MA, combinando sus características para ofrecer una mayor flexibilidad en el modelado de series temporales. Un proceso ARMA expresa el valor actual de una serie como una combinación lineal de sus valores y errores pasados. Los parámetros ϕ y θ controlan la influencia de estos términos en la predicción, permitiendo capturar tanto patrones deterministas como estocásticos en los datos. La capacidad del modelo ARMA para modelar la dependencia entre observaciones pasadas y presentes los convierte en una elección popular para analizar y predecir series temporales en campos tan diversos como la economía, la meteorología y la ingeniería[5]. Su ecuación es la siguiente:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (1)$$

Donde, $\theta_1, \dots, \theta_p$ son los coeficientes AR, $\theta_1, \dots, \theta_q$ son los coeficientes MA, Y_t es la serie temporal, μ es una constante y ε_t es un término de error.

2.1.2 Modelo ARIMA

Es una herramienta usada para predecir el futuro de datos que varían con el tiempo. Para que este modelo funcione se necesita que los datos tengan un comportamiento relativamente estable a lo largo del tiempo, lo que se denomina estacionariedad. Si los datos no son estacionarios, se pueden aplicar transformaciones como la diferenciación o el logaritmo para estabilizarlos[6].

Un modelo ARIMA (p, d, q) , es aquel donde μ es la media, p es el orden autorregresivo, d es el número de diferencias no estacionales requeridas para la estacionariedad y q es el orden de medias móviles. Su ecuación es la siguiente[6]:

$$\Delta^d y_t = \mu + \phi_1 \Delta^d y_{t-1} + \dots + \phi_p \Delta^d y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2)$$

Con $\theta_1, \dots, \theta_p$ coeficientes AR, $\theta_1, \dots, \theta_q$ son los coeficientes MA, y_t es la serie temporal, Δ^d es el operador diferencia de orden d y ε_t es un término de error.

2.1.3 Modelo GARCH

Es un modelo autorregresivo condicional heteroscedástico generalizado que estima la varianza condicional de una serie de tiempo. La varianza condicional en un momento dado depende de un promedio ponderado de las varianzas condicionales pasadas y de los cuadrados de los errores pasados. El parámetro p en el modelo GARCH (p, q) determina el número de períodos pasados que se consideran en este promedio ponderado, permitiendo capturar diferentes patrones de persistencia en la volatilidad.

El modelo GARCH es un modelo generalizado porque combina componentes autorregresivos y de media móvil. Es autorregresivo dado que la varianza condicional en un periodo depende de las varianzas pasadas, es condicional porque la varianza se modela dado el conjunto de información disponible hasta ese momento y es heteroscedástico debido a que la varianza no es constante a lo largo del tiempo[7]. Su ecuación es la siguiente:

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (3)$$

Considerando que $\varepsilon_t = \sigma_t \eta_t$, $\eta_t \sim iidN(0,1)$; σ_t y η_t son independientes. $\omega > 0$; $\alpha_i \geq 0$, $i = 1, 2, \dots, p$; $\beta_j \geq 0$, $j = 1, 2, \dots, q$; (Garantiza que la varianza sea positiva), $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ garantiza que exista varianza incondicional. Donde α_i son los coeficientes que ponderan los errores al cuadrado, β_j son los coeficientes que ponderan las varianzas pasadas. σ_t^2 es la varianza en el período t , ω es una constante y ε_t es un término de error.

2.1.4 Redes Neuronales Autorregresivas (ANN)

Las redes neuronales son modelos computacionales que imitan las capacidades de procesamiento de información del cerebro humano, logrando el reconocimiento de patrones complejos y el aprendizaje a través de extensas redes de neuronas. Sin programación explícita, estos modelos pueden extraer de forma autónoma representaciones y características útiles de conjuntos de datos masivos. En el sector financiero, las redes neuronales, como los precios de las acciones, los tipos de cambio de divisas y las tendencias del mercado, se utilizan ampliamente para pronosticar el mercado de valores[8].

La Fig. 1 muestra una red de propagación hacia delante con dos capas ocultas. Esta conecta capas de neuronas unidireccionalmente, multiplicando sus salidas por pesos específicos para cada conexión.

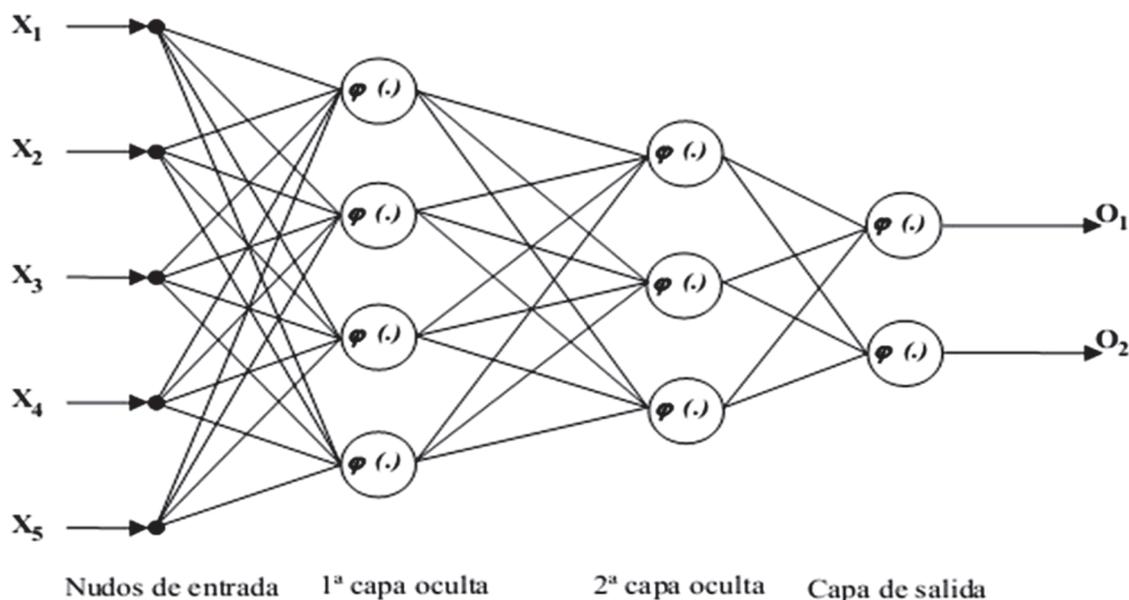


Fig. 1. Esquema de una red neuronal artificial[9].

2.2 Medidas de evaluación

Para la evaluación de modelos predictivos tres métricas destacan por su utilidad: el MAPE (Mean Absolute Porcentaje Error), el MAE (Mean Absolute Error) y el RMSE (Root Mean Squared Error). Las cuales permiten cuantificar la precisión de las predicciones y comparar diferentes modelos, tal que:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

$$MAPE = \frac{100}{n} * \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{|y_i|} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

Con n es el número total de observaciones, y_i es el valor real y \hat{y}_i es el valor predicho por el modelo.

III. ESTADO DEL ARTE

La capacidad de pronosticar series de tiempo es fundamental en diversos ámbitos, desde la toma de decisiones empresariales hasta la gestión de riesgos financieros. La elección del modelo más adecuado depende de las características específicas de cada aplicación, de los datos disponibles y del horizonte de predicción, lo que exige una cuidadosa evaluación de cada caso particular.

Por ejemplo, Em um estudio demostraron la viabilidad de emplear modelos ARIMA, SARIMA y Holt-Winters para pronosticar con precisión la demanda de energía eléctrica en municipios colombianos. Al evaluar la precisión de los modelos mediante métricas como ECM, MAE, RMSE y MAPE, los autores concluyeron que la metodología de Box y Jenkins es altamente efectiva para capturar los patrones inherentes a estas series temporales, superando modelos más simples[10].

Em otro estudio comparativo entre modelos de pronóstico GARCH y redes neuronales artificiales para la predicción de los índices accionarios de México y Estados Unidos. Sus hallazgos revelan que la red neuronal capta de forma adecuada el comportamiento de la serie de tiempo, pero es el modelo GARCH que tiene un mejor ajuste a las series dentro y fuera de la muestra[11].

Villada y otros[8] utilizaron un modelo de redes neuronales para pronosticar precios en el mercado de valores, probando varias estructuras con una capa oculta. Inicialmente, el número de neuronas fue igual al promedio entre entradas y salidas, incrementándose gradualmente hasta encontrar la mejor configuración. El uso de redes neuronales demostró ser efectivo para predecir el precio de dos acciones clave en la bolsa de Colombia, destacando su simplicidad y bajos errores de predicción tanto dentro como fuera de la muestra, validando su aplicabilidad en mercados emergentes.

Lecca y otros[12] emplearon modelos ARCH para modelar la volatilidad de los precios de cierre del mercado bursátil en la Bolsa de Valores de Lima. Los resultados confirmaron la presencia de heterocedasticidad condicional y efectos de apalancamiento. Entre los modelos evaluados, el TAR(1,1) resultó ser el más adecuado, capturando de manera eficiente la persistencia de la volatilidad y los impactos asimétricos de los shocks. Este estudio contribuye a la literatura existente al validar la aplicabilidad de los modelos GARCH en mercados emergentes.

Taylor y Conto[13] evaluaron diversos modelos para pronosticar de manera eficiente el comportamiento de la Unidad de Valor Real (UVR) a largo plazo, una unidad clave en el sistema financiero colombiano. Para ello se utilizaron métodos estadísticos y de Machine Learning. Los resultados demostraron que los modelos SARIMA y Holt-Winters capturan de manera efectiva los patrones estacionales y de tendencia de la UVR. El modelo Holt-Winters multiplicativo se destacó en predicciones a largo plazo, mientras que el modelo SARIMA ofreció mayor precisión en horizontes de pronóstico más cortos.

Finalmente, Zapata y otros[14] utilizan una red neuronal artificial, un suavizamiento Holt y un modelo ARIMA para predecir las tasas de cambio USD/COP y EUR/COP, obteniéndose mejores predicciones con la red neuronal. Estos hallazgos contribuyen a la literatura existente sobre el modelado de series de tiempo financieras y proporcionan una base para futuras investigaciones, reduciendo la incertidumbre y apoyando la toma de decisiones.

IV. METODOLOGÍA

La investigación es de tipo cuantitativa, basada en el análisis de los datos históricos de precios y rendimientos de las acciones de Ecopetrol y Bancolombia, y en el uso de métodos cuantitativos de pronóstico para la construcción de predicciones objetivas. Para el estudio son empleados datos históricos de los precios de cierre obtenidos de la plataforma Investing.com. El periodo de muestreo abarca tres años comprendidos entre agosto de 2021 y septiembre del 2024.

Son implementados modelos ARIMA, redes neuronales autorregresivas y un modelo híbrido resultante de la combinación de estos dos para la predicción de los precios, y para los rendimientos serán usados los modelos ARMA-GARCH y ARMA-Redes neuronales. Estos modelos fueron elegidos por su capacidad para capturar patrones temporales presentes en las series de tiempo financieras consideradas.

Para la aplicación de los modelos en cada una de las series de tiempo se lleva a cabo la estrategia de validación cruzada, donde el conjunto de datos será dividido en dos partes. La primera parte es considerada para el ajuste de los modelos y la segunda para validar su capacidad de pronóstico futuro. En ambos escenarios serán calculadas las medidas de capacidad de ajuste MAPE, MAE y RMSE, con las cuales se determinará el modelo que mejor se adapta a la estructura histórica de los precios y rendimientos. Los modelos son implementados utilizando en el software R.

V. DESCRIPCIÓN DE LOS DATOS

El periodo de estudio abarca tres años, desde agosto de 2021 hasta septiembre de 2024. En la tabla I se describen las series de tiempo de interés y su respectiva unidad de medida.

La tabla II muestra estadísticas descriptivas acerca de los precios de las acciones. Se puede ver que los precios de Bancolombia son considerablemente más fluctuantes que los de Ecopetrol, lo que conlleva un riesgo elevado para los inversores. Con la tendencia central, se puede ver que el precio medio de Bancolombia es notablemente superior, lo que evidencia su magnitud e importancia en el mercado. Las dos distribuciones muestran una asimetría positiva, con la de Ecopetrol siendo la más marcada. Bancolombia posee la desviación estándar más alta, lo que señala un riesgo elevado vinculado a su inversión.

La tabla III presenta un análisis estadístico de los rendimientos de Bancolombia y Ecopetrol. Al comparar ambos activos, se observa que la media de los rendimientos de Bancolombia es de 0.000195, un valor positivo pero cercano a cero. Esto indica que, en promedio, los rendimientos de Bancolombia han sido ligeramente positivos durante el período analizado. La media de los rendimientos de Ecopetrol es de -0.000427, un valor negativo y pequeño en términos absolutos. Esto sugiere que, en promedio, los rendimientos de Ecopetrol fueron ligeramente negativos durante el mismo período.

Tabla I. Descripción de las series de tiempo.

Variable	Descripción	Unidad
Precios de Bancolombia	Precios de cierre diarios de la acción de Bancolombia	Peso colombiano (COP)
Precios de Ecopetrol	Precios de cierre diarios de la acción de Ecopetrol	Peso colombiano (COP)
Rendimientos de Bancolombia	Rendimientos logarítmicos diarios del precio de la acción de Bancolombia	Adimensional
Rendimientos de Ecopetrol	Rendimientos logarítmicos diarios del precio de la acción de Ecopetrol	Adimensional

Fuente: elaboración propia.

Tabla II. Medidas descriptivas de los precios Bancolombia y Ecopetrol.

Estadísticas descriptivas	Precios de Bancolombia	Precios de Ecopetrol
Número de observaciones	752	752
Valores faltantes	0	0
Mínimo	27200	1840
Máximo	45320	3610
1. Cuartil	32680	2250
3. Cuartil	37972.5	2646.25
Media	35341.56915	2479.174202
Mediana	34800	2375
Varianza	14480676.36	127160.0189
Desviación estándar	3805.34839	356.595035
Asimetría	0.43423	1.061926
Curtosis	-0.53985	0.79992

Fuente: elaboración propia.

Tabla III. Medidas descriptivas de los rendimientos financieros de Bancolombia y Ecopetrol.

Estadísticas descriptivas	Rendimientos de Bancolombia	Rendimientos de Ecopetrol
Número de observaciones	751	751
Valores faltantes	0	0
Mínimo	-0.105361	-0.146354
Máximo	0.09531	0.106342
1. Cuartil	-0.012453	-0.010977
3. Cuartil	0.012034	0.012528
Media	0.000195	-0.000427
Mediana	0	0
Varianza	0.000553	0.000556
Desviación estándar	0.023517	0.023587
Asimetría	-0.027841	-0.763395
Curtosis	1.993137	4.472589

Fuente: elaboración propia.

Ambos activos muestran desviaciones estándar similares, lo que indica que han experimentado una volatilidad comparable durante el periodo analizado. Cabe destacar que valores más altos de desviación estándar suelen reflejar un mayor nivel de riesgo o incertidumbre en los rendimientos. En cuanto a la curtosis, los rendimientos de Bancolombia presentan un valor de 1.993137, inferior a 3, lo que indica que su distribución es platicúrtica. Esto significa que tiene colas más delgadas y menos eventos extremos que una distribución normal. Por otro lado, la curtosis de los rendimientos de Ecopetrol es de 4.472589, superior a 3, lo que sugiere una distribución leptocúrtica. Esto implica que Ecopetrol ha experimentado más eventos extremos de lo esperado en una distribución normal.

5.1 Análisis de precios y rendimientos de Bancolombia

La Fig. 2 detalla la serie temporal de los precios de cierre de las acciones de Bancolombia, mostrando su comportamiento a lo largo del período estudiado. La serie presenta cambios en su tendencia a largo plazo. En su ACF se nota un decrecimiento lento de los coeficientes de autocorrelación, confirmando la tendencia presente.

La Fig. 3 muestra los rendimientos de Bancolombia fluctúan en torno a un valor cercano a 0, con episodios de alta volatilidad reflejadas tanto en picos positivos como negativos. Asimismo, se observa que no existe una autocorrelación significativa en los rendimientos, ya que las barras posteriores al rezago 0 están muy cerca de 0 y dentro de los intervalos de confianza. En términos simples, los rendimientos actuales no guardan relación con los rendimientos anteriores.

5.2 Análisis de precios y rendimientos de Ecopetrol

La Fig. 4 presenta la serie temporal de los precios de cierre de las acciones de Ecopetrol, revelando su comportamiento a lo largo del período estudiado. Se percibe una variación en la tendencia a largo plazo con periodos de crecimiento y otros de decrecimiento. En su ACF se nota un decrecimiento lento en su autocorrelación temporal, reflejando una fuerte tendencia.

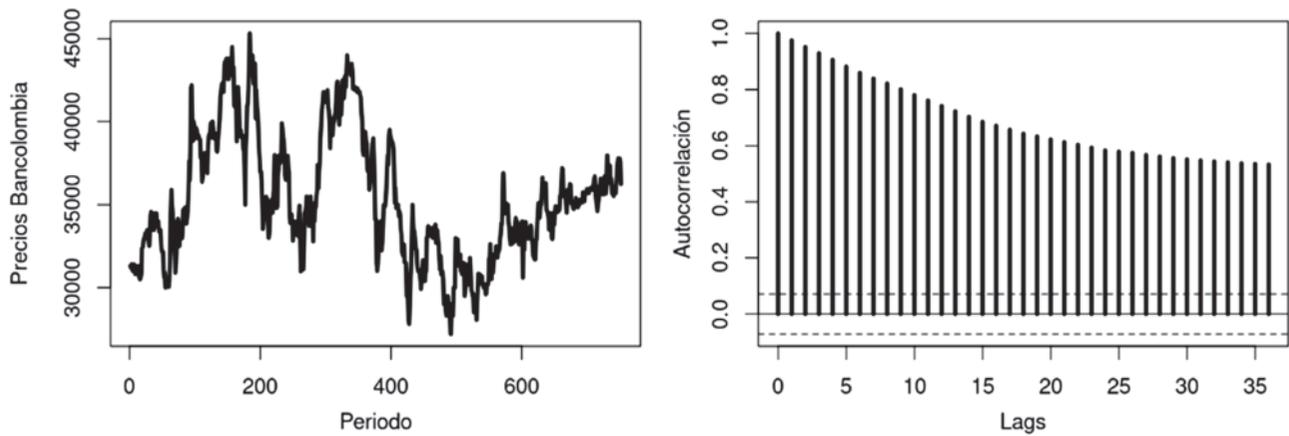


Fig. 2. Precios de Bancolombia y su ACF. Fuente: Elaboración propia con software R.

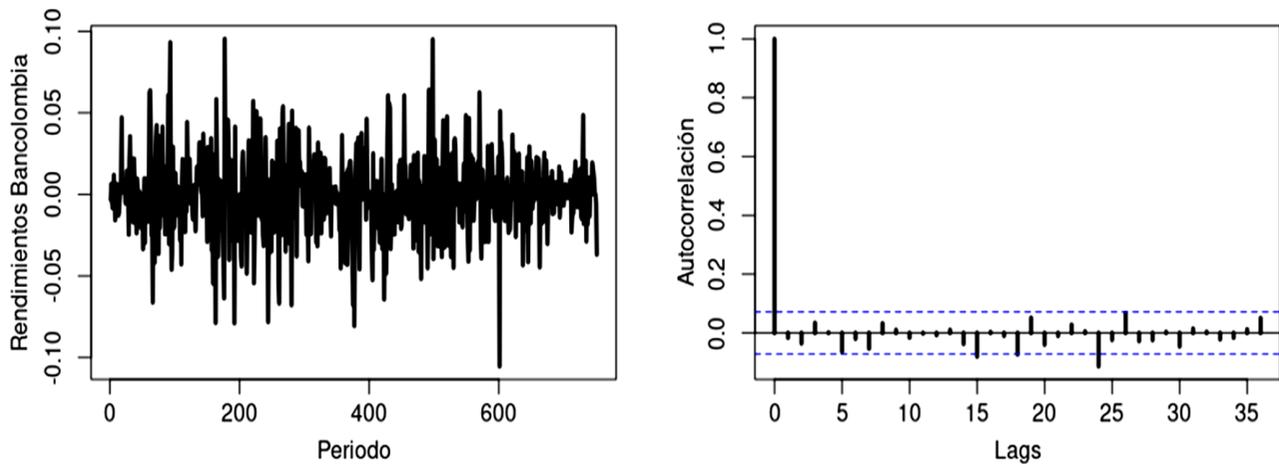


Fig. 3. Rendimientos de Bancolombia. Fuente: Elaboración propia con software R.

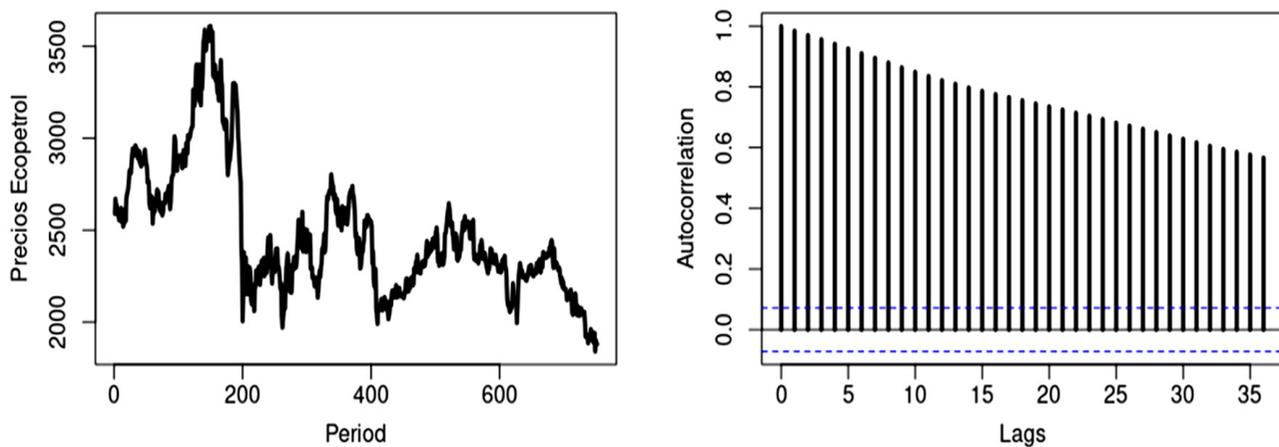


Fig. 4. Precios de Ecopetrol y su ACF. Fuente: Elaboración propia con software R.

La Fig. 5 parece tener volatilidad extrema en ciertos puntos (pico hacia abajo cerca del período 200), la gráfica de autocorrelación muestra que no hay correlación significativa entre los rendimientos en diferentes retardos, lo que sugiere que los rendimientos de Ecopetrol siguen un comportamiento de camino aleatorio.

VI. RESULTADOS

En esta sección se presentan los resultados obtenidos luego de aplicar la metodología propuesta a los datos de series temporales de precios y rendimientos de Bancolombia y Ecopetrol, descritos en la sección.

6.1 Análisis de métodos de pronóstico aplicados a los precios de la acción de Bancolombia.

Se procede a implementar los métodos de pronóstico seleccionados para las series de precios de las acciones, los cuales son: ARIMA, Redes neuronales autorregresivas y la combinación de estos dos. Lo anterior, teniendo en cuenta la estrategia de validación cruzada, donde no son considerados los últimos 15 días de precios conocidos para el ajuste de los modelos.

Primero, se procede a analizar los precios de Bancolombia, donde de acuerdo con la figura 2 se observa que tiene una tendencia cambiante en el tiempo, por lo que un modelo ARIMA resulta ser adecuada en este caso. Para la implementación de dicho modelo, se requiere que la serie sea estacionaria en media, por lo que se procede a diferen-

ciarla, obteniéndose la Fig. 6. En esta es posible ver que la serie de la primera diferencia no presenta autocorrelaciones significativas ni en ACF, ni en la PACF, por lo que la serie original del precio de la acción de Bancolombia sigue un modelo de caminata aleatoria ARIMA (0,1,0).

Como segundo paso, se procede a estimar una Red neuronal autorregresiva para la serie del precio de la acción de Bancolombia, esto con ayuda de la función `nnetar()` de la librería `forecast`, la cual permite un ajuste automático de la mejor red asociada a los datos. Se obtuvo una Red neuronal (1,1), es decir con un orden autorregresivo de uno (1) y una (1) neurona en la capa oculta.

Finalmente, se implementa un modelo híbrido que combina ambos modelos ARIMA + Red neuronal, tal que, el ARIMA modela la tendencia presente en los datos y la Red neuronal modela los residuales del ARIMA en busca de patrones que no fueron explicados, incluyendo relaciones no lineales. En este caso, se implementa un ARIMA (0,1,0) + Red neuronal (1,1).

En la tabla 4, es posible observar de manera comparativa el MAPE, RMSE y MAE para el ajuste de los tres métodos implementados, donde es posible concluir que el mejor ajuste lo tiene el modelo híbrido ARIMA (0,1,0) + Red (1,1), debido a que sus medidas de evaluación son menores en todos los casos.

Con los tres métodos ajustados se procedió a predecir el comportamiento del precio de la acción de Bancolombia para los siguientes 15 días y

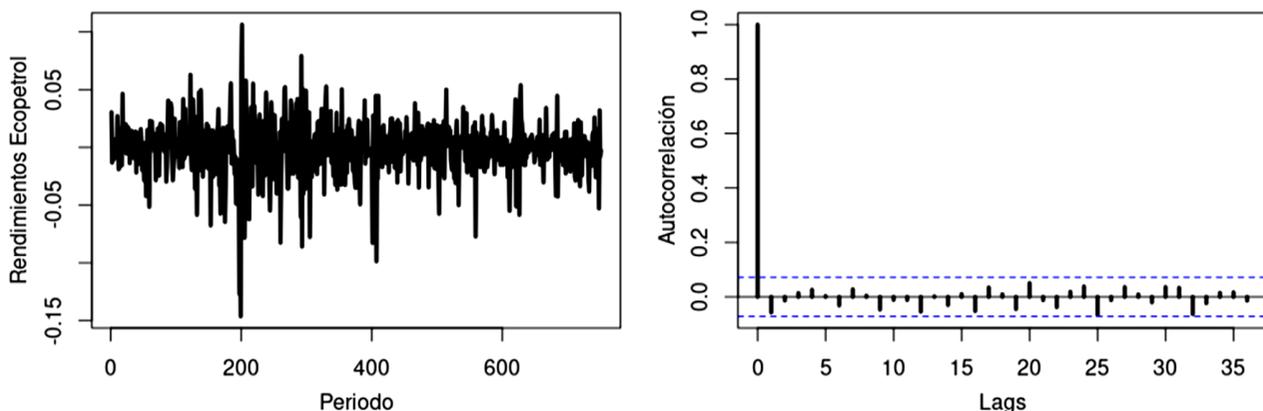


Fig.5. Rendimientos de Ecopetrol. Fuente: Elaboración propia con software R.

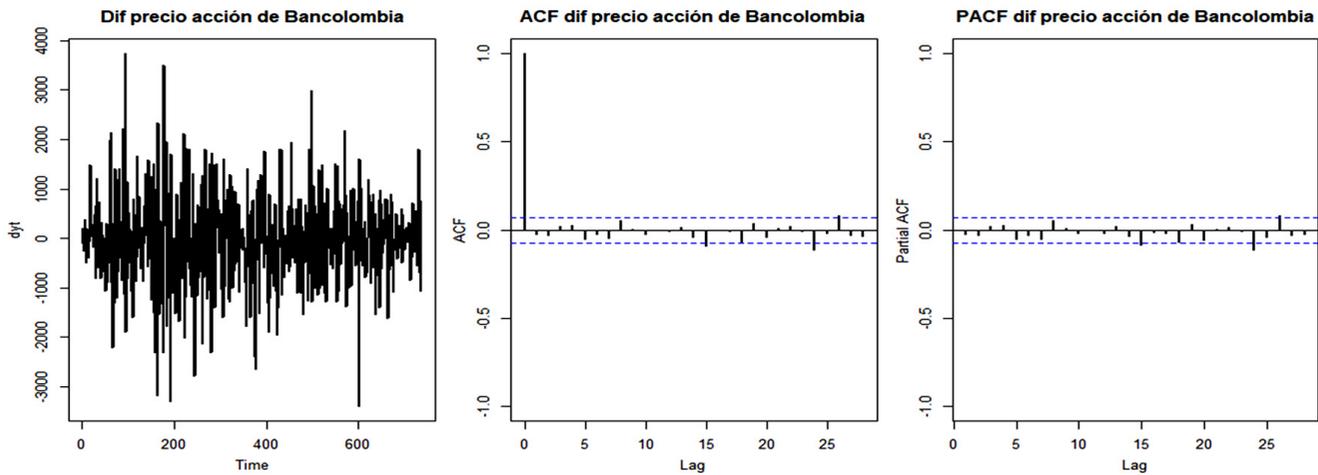


Fig. 6. Gráfico, ACF y PACF de la primera diferencia de los precios de la acción de Bancolombia.
Fuente: Elaboración propia con software R.

Tabla IV. Medidas de evaluación para el ajuste de los métodos con los precios de Bancolombia.

Método	MAPE	MAE	RMSE
ARIMA (0,1,0)	1.689	594.778	833.076
Red neuronal (1,1)	1.697	596.149	825.91
ARIMA (0,1,0) + Red (1,1)	1.682	591.656	825.579

Fuente: elaboración propia con software R.

comparar estos resultados con los verdaderos valores observados para dichos días, que no fueron considerados al momento de estimar los métodos. La tabla 5 muestra el MAPE, RMSE y MAE para el período de pronóstico futuro de 15 días con los tres métodos, donde es posible ver que el modelo híbrido ARIMA (0,1,0) + Red (1,1) es aquel con menores medidas, indicando que realiza pronósticos más eficientes.

Tabla V. Medidas de evaluación para el pronóstico futuro a 15 días con los precios de Bancolombia.

Método	MAPE	MAE	RMSE
ARIMA (0,1,0)	2.255	826.667	884.775
Red neuronal (1,1)	2.254	826.234	883.710
ARIMA (0,1,0) + Red (1,1)	2.253	825.814	884.643

Fuente: elaboración propia con software R.

La Fig. 7 muestra el gráfico comparativo de pronósticos con los tres métodos implementados para los siguientes 15 días de la acción de Bancolombia. En esta, es posible ver que se obtienen pronósticos estacionarios en media y cercanos a los verdaderos valores observados de la serie.

6.2 Análisis de métodos de pronóstico aplicados a los precios de la acción de Ecopetrol.

Se implementan tres modelos de pronóstico para las series de precios de las acciones, ARIMA, Redes neuronales autorregresivas y una mezcla de ambas. Para evaluar la capacidad predictiva de los modelos se reservaron los últimos 15 días de datos para su validación.

El primer análisis se centra en los precios de Ecopetrol. Según la Figura 3, se observa una tendencia cambiante a lo largo del tiempo, lo que sugiere que un modelo ARIMA es adecuado para este caso. Para implementar este modelo, es necesario que la serie sea estacionaria en media, por lo que se procede a diferenciarla, como se ilustra en la Figura 8. Esta figura señala que la serie de la diferencia inicial no muestra autocorrelaciones relevantes tanto en la función de autocorrelación (ACF) como en la función de autocorrelación parcial (PACF). Como resultado, se concluye que la serie original del precio de la acción de Ecopetrol sigue un modelo de caminata aleatoria ARIMA (0,1,0).

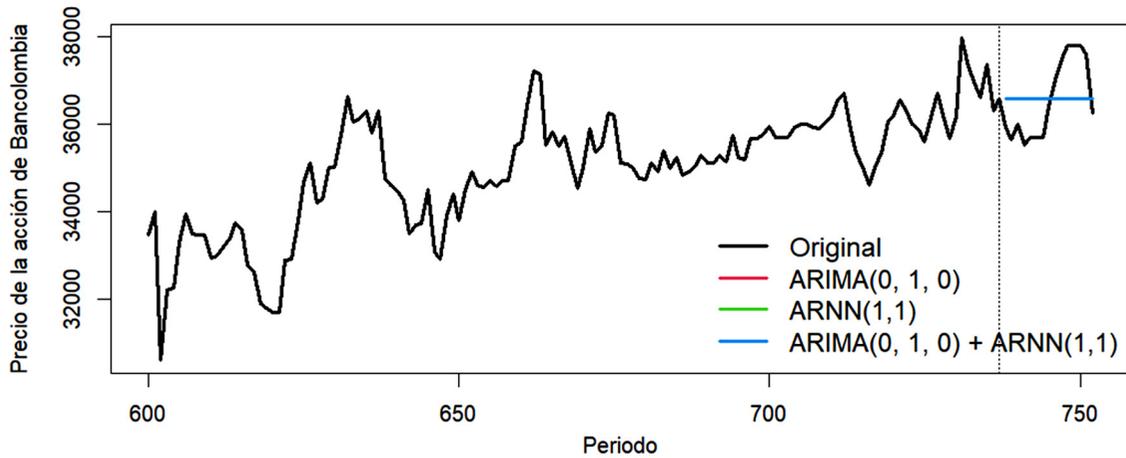


Fig. 7. Gráfico comparativo de pronósticos futuros con los tres métodos para los precios de Bancolombia.
 Fuente: Elaboración propia con software R.

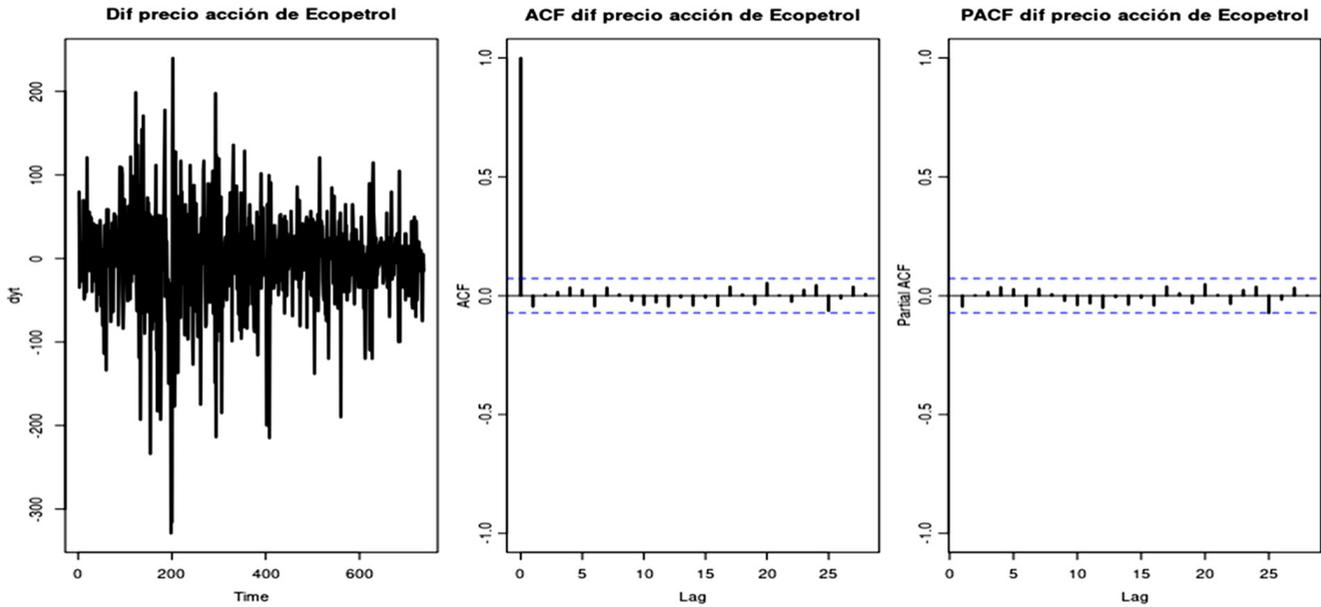


Fig. 8. Gráfico, ACF y PACF de la primera diferencia de los precios de la acción de Ecopetrol.
 Fuente: Elaboración propia con software R.

Como paso adicional, se lleva a cabo la estimación de una Red neuronal autorregresiva para la serie del precio de la acción de Ecopetrol. Se logró una Red neuronal (1,1), lo que significa que existe un orden autorregresivo de uno (1) y una (1) neurona en la capa oculta.

Finalmente, se aplica un modelo híbrido que fusiona los dos modelos ARIMA + Red neuronal, de manera que el ARIMA representa la tendencia presente en los datos y la Red neuronal analiza los

residuos del ARIMA en busca de patrones que no se han descrito, incluyendo relaciones no lineales. En esta situación, se aplica un ARIMA (0,1,0) + Red neuronal (1,1) para la implementación (Tabla VI).

Con los tres métodos ajustados se procedió a predecir el comportamiento del precio de la acción de Ecopetrol para los siguientes 15 días y comparar estos resultados con los verdaderos valores observados que no fueron considerados al momento de estimar los métodos. La tabla VII muestra el

Tabla VI. Medidas de evaluación para el ajuste de los métodos con los precios de Ecopetrol.

	MAPE	MAE	RMSE
ARIMA (0,1,0)	1.665	41.305	58.835
Red neuronal (1,1)	1.66984	41.43732	58.4008
ARIMA (0,1,0) + Red (1,1)	1.658	41.154	58.414

Fuente: elaboración propia en el programa R.

Tabla VII. Medidas de evaluación para el pronóstico futuro a 15 días con los precios de Ecopetrol.

Método	MAPE	MAE	RMSE
ARIMA (0,1,0)	30.90307	25	30.90307
Red neuronal (1,1)	2.361899	44.95587	50.27335
ARIMA (0,1,0) + Red (1,1)	1.307159	24.83977	30.78735

Fuente: elaboración propia en el programa R.

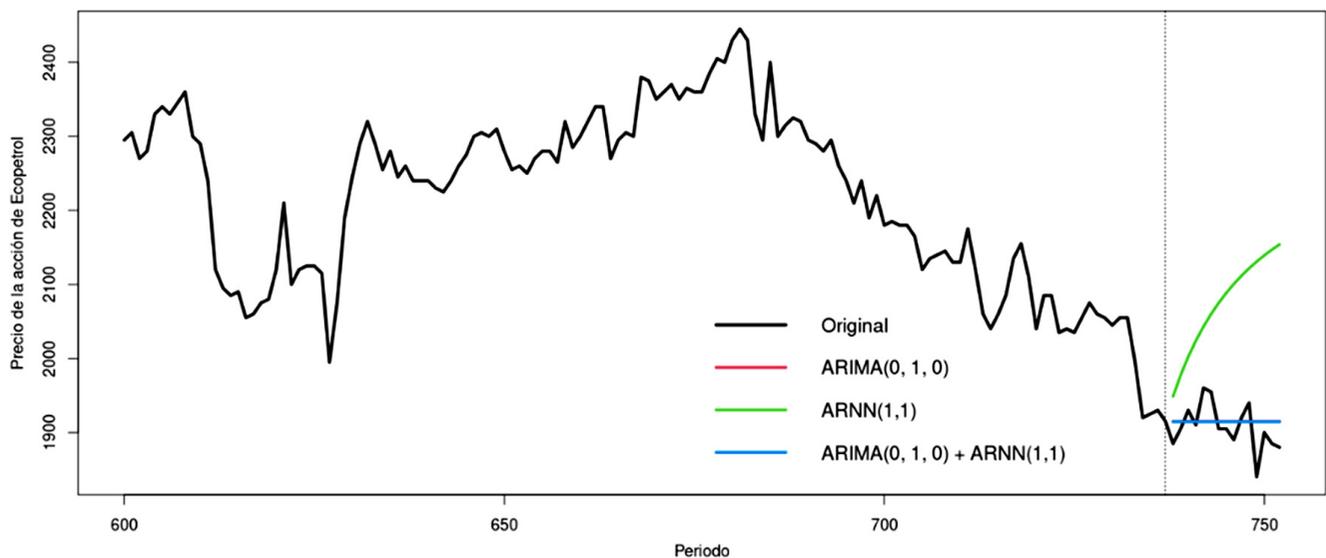
MAPE, MAE y RMSE para el período de pronóstico futuro de 15 días con los tres métodos, donde es posible ver que el modelo híbrido ARIMA (0,1,0) + Red (1,1) es aquel con menores medidas, indicando que realiza pronósticos más eficientes.

La Fig. 9 muestra el gráfico comparativo de pronósticos con los tres métodos implementados para los siguientes 15 días de la acción de Ecopetrol. En esta, es posible ver que el modelo híbrido ARIMA + ARNN puede ser el mejor predictor en este contexto, ya que parece ajustarse más adecuadamente a las fluctuaciones futuras y evita sobreestimación del ARNN.

6.3 Análisis de métodos de pronóstico aplicados a rendimientos de la acción de Bancolombia.

Para predecir los rendimientos de la acción de Bancolombia se implementan modelos ARMA con errores modelados de dos maneras: mediante un GARCH (1,1) y mediante una red neuronal autorregresiva. Esto con la finalidad de que el modelo ARMA capte la estructura de correlación de primer orden y los otros métodos modelen la posible heterocedasticidad o correlación de segundo orden presente en los datos. Para el análisis se recurrió nuevamente a la estrategia de validación cruzada, donde fueron dejados los últimos 15 datos para testear la capacidad predictiva de los modelos.

De acuerdo con la Fig. 2, la serie de rendimientos de Bancolombia no posee autocorrelaciones de primer orden significativas, por lo que no se tendrían coeficientes ARMA que ajustar. Luego de

**Fig. 9.** Gráfico comparativo de pronósticos futuros con los tres métodos para los precios de Ecopetrol.

Fuente: Elaboración propia con software R.

esto, se procede a analizar si la serie posee varianza condicional heteroscedástica mediante el análisis de los rendimientos al cuadrado.

La Fig. 10 muestra la ACF de los rendimientos al cuadrado, donde es posible notar que se presentan autocorrelaciones significativas, indicando que dichos rendimientos poseen volatilidad no constante, por lo que se procede a ajustar un modelo GARCH (1,1) y una red neuronal para dichos rendimientos.

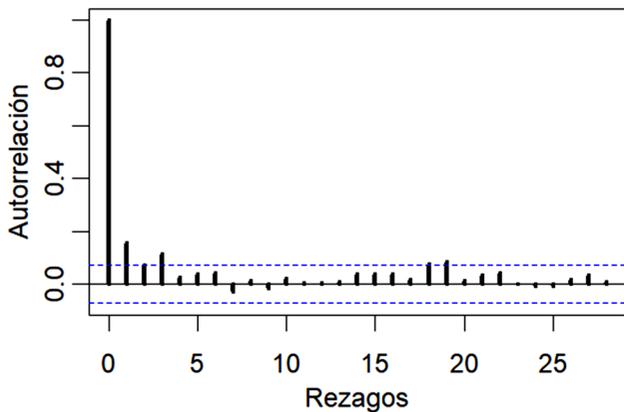


Fig. 10. ACF de los rendimientos de Bancolombia. Fuente: Elaboración propia con software R.

La Tabla VIII muestra los coeficientes estimados del modelo GARCH (1,1) y sus valores p, donde este último indica que todos son significativos y al hacer la prueba ACF para los errores del modelo GARCH, esta arroja que no existe autocorrelación de segundo orden, por lo que la varianza de dichos errores es constante.

Tabla VIII. Coeficientes estimados del modelo GARCH (1,1)

Coefficiente	Estimado	Valor-p
μ	-0.00050928	0.529825
ω	0.00010410	0.002209
α_1	0.17893117	0.000105
β_1	0.64606731	0.000000

Fuente: elaboración propia en el programa R.

Adicionalmente, se buscó la red neuronal que más se adecuaba al comportamiento histórico de

los rendimientos de Bancolombia y se obtuvo una Red neuronal (1,1), es decir, que los rendimientos son modelados mediante una autorregresión de orden 1 y una neurona en la capa oculta.

La Tabla IX muestra cómo los modelos GARCH (1,1) y de Red Neuronal (1,1) se ajustan a los datos históricos de los rendimientos. Ambos modelos presentan valores de MAE y RMSE muy cercanos, lo que sugiere que ambos capturan de manera similar la variabilidad de los datos históricos. Sin embargo, la Red Neuronal presenta ligeramente mejores resultados en ambos indicadores. Como se puede observar en la Tabla X, el modelo de red neuronal (1,1) muestra un MAE de 0.010296, lo que resulta menor al MAE del modelo GARCH (1,1) que tiene un valor de 0.010427. Esto indica que los pronósticos del modelo de red neuronal son más exactos en términos absolutos. El RMSE para la red neuronal es 0.01406, ligeramente superior al RMSE del GARCH que es 0.014052. Lo anterior indica que la red neuronal tiene un MAE mejor y su desempeño en términos de RMSE es comparable y ligeramente menos favorable. En este caso, no es posible calcular el MAPE dado que se tienen rendimientos iguales a cero y al hacer su cálculo se obtiene como resultado infinito. En general, el análisis muestra que el modelo de red neuronal (1,1) tiene un mejor desempeño en términos de MAE, lo que sugiere una mayor precisión en las predicciones.

Tabla IX. Medidas de evaluación para el ajuste de los métodos para los rendimientos de Bancolombia.

Método	MAE	RMSE
GARCH (1,1)	0.01694658	0.02366544
Red neuronal (1,1)	0.01688318	0.02341626

Fuente: elaboración propia en el programa R.

Tabla X. Medidas de evaluación para el pronóstico futuro a 15 días para los rendimientos de Bancolombia.

Método	MAE	RMSE
GARCH (1,1)	0.01042744	0.01405231
Red neuronal (1,1)	0.01029611	0.01406099

Fuente: elaboración propia en el programa R.

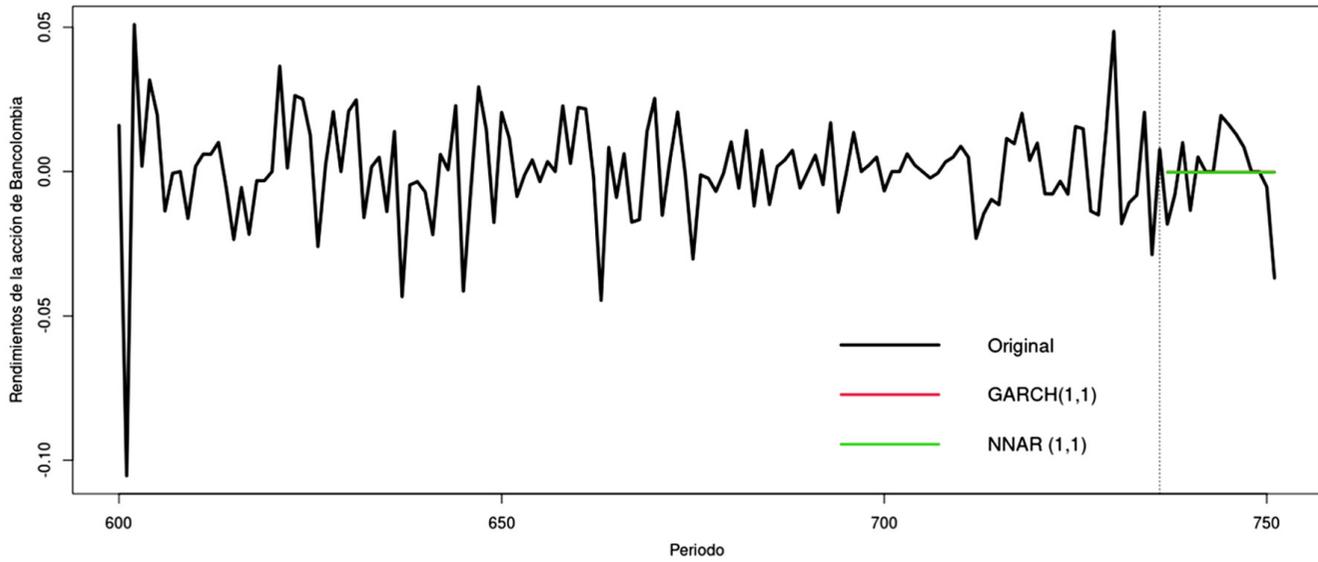


Fig. 11. Gráfico comparativo de pronósticos futuros con los dos métodos para los rendimientos de Bancolombia. Fuente: Elaboración propia con software R.

En la Fig. 11 se muestran los pronósticos a 15 días realizados con ambos métodos para los rendimientos considerados, en la cual se puede ver que estos son similares entre sí y siguen la estacionariedad en media presente en la serie.

6.4 Análisis de métodos de pronóstico aplicados a rendimientos de la acción de Ecopetrol.

El comportamiento de los rendimientos de la acción de Ecopetrol es tenido en cuenta, y su volatilidad es tratada mediante el uso de dos enfoques: los modelos ARMA-GARCH (1,1) y ARMA con errores modelados por redes neuronales autorregresivas. Un primer análisis de autocorrelación de los rendimientos efectuado en este caso (ver Fig. 4) no tuvo unos coeficientes significativos, es decir, no fue posible utilizar un modelo ARMA.

La Fig. 12 presenta la ACF de los cuadrados de los rendimientos, mostrando que existen autocorrelaciones significativas, lo que señala la existencia de heterocedasticidad. En este sentido, se optó por un modelo GARCH (1,1) y por redes neuronales que son capaces de captar la volatilidad condicional y ajustarse a relaciones no lineales en los datos. Con el fin de controlar esas variables, se aplicó una estrategia de validación cruzada en la que no se consideran los últimos 15 días para el ajuste, con el objetivo de usarlos validar la capacidad de pronóstico futuro.

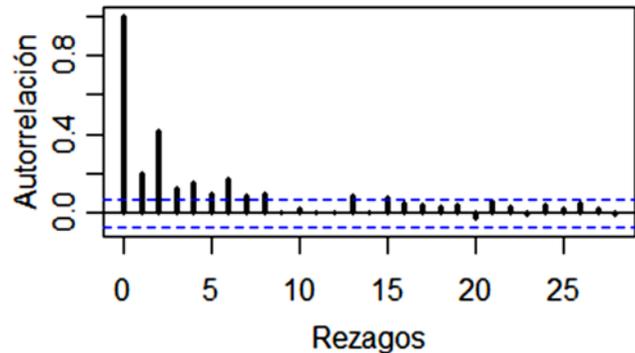


Fig. 12. ACF de los rendimientos de Ecopetrol. Fuente: Elaboración propia con software R.

La Tabla XI muestra los coeficientes estimados del modelo GARCH, los cuales son significativos dado que sus valores-p son menores al 5%. Este resultado muestra que rendimientos tienen un patrón de volatilidad condicional, implicando que los shocks antiguos influyen en algún grado sobre la volatilidad futura.

Al ajustar la red neuronal, se encontró como mejor opción para modelar el comportamiento histórico una Red Neuronal (1,1), con una neurona en la capa oculta y considerando una autorregresión de un período.

Los resultados presentados en la Tabla XII indican que tanto el modelo GARCH (1,1) como la Red Neuronal (1,1) son capaces de modelar los

Tabla XI. Coeficientes estimados del modelo GARCH (1,1)

Coeficiente	Estimado	Valor-p
μ	-0.00015492	0.83707
ω	0.00004182	0.02647
α_1	0.11081576	0.00033
β_1	0.81192587	0.00000

Fuente: elaboración propia en el programa R.

Tabla XII. Medidas de evaluación para el ajuste de los métodos con los rendimientos de Ecopetrol.

Método	MAE	RMSE
GARCH (1,1)	0.01662612	0.02363168
Red neuronal (1,1)	0.01647688	0.02335188

Fuente: elaboración propia en el programa R.

rendimientos de Ecopetrol. Sin embargo, la mayor flexibilidad de la Red Neuronal le permite capturar patrones no lineales y relaciones más complejas en los datos, lo que se traduce en un ajuste ligeramente mejor, según lo evidenciado por los valores de las métricas de evaluación. El modelo de red neuronal, según la Tabla XIII, presenta un MAE ligeramente inferior (0.015562) en com-

Tabla XIII. Medidas de evaluación para el pronóstico futuro a 15 días con los métodos para los rendimientos de Ecopetrol.

Método	MAE	RMSE
GARCH (1,1)	0.015570	0.020489
Red neuronal (1,1)	0.015562	0.020484

Fuente: elaboración propia en el programa R.

paración con el modelo GARCH (0.015570), lo que sugiere que la red neuronal tiene un rendimiento ligeramente superior en términos de error absoluto medio. Adicionalmente, la red neuronal usada muestra un RMSE más bajo (0.020484) en comparación con el GARCH (0.020489), lo que refuerza la conclusión de que dicha red proporciona predicciones más precisas. Ambos modelos presentan resultados competitivos, pero la Red Neuronal (1,1) supera al GARCH (1,1) en términos de ambas métricas (MAE y RMSE). Esto sugiere que, en este contexto específico, las redes neuronales pueden ser más efectivas para la predicción.

En la Fig. 13 se pueden ver los pronósticos para los siguientes 15 días realizados con ambos métodos para la serie de rendimiento de Ecopetrol, pudiendo concluir que en ambos casos se obtiene predicciones que siguen el patrón estacionario en media, prefiriéndose el uso de la red neuronal en este caso.

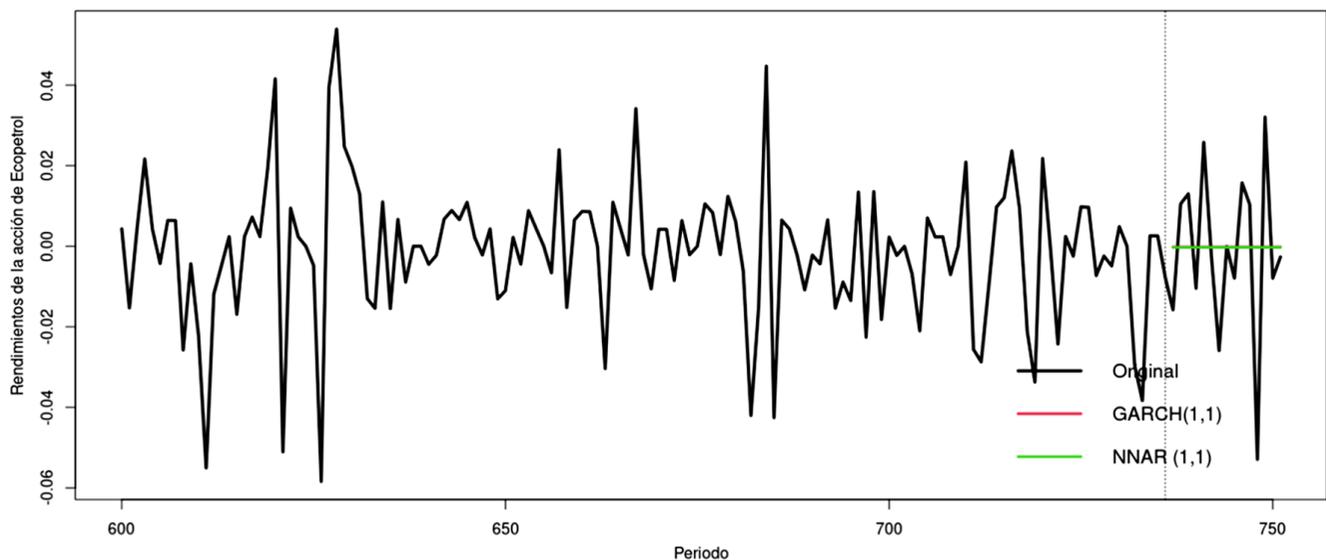


Fig. 13. Gráfico comparativo de pronósticos futuros con los dos métodos para los rendimientos de Ecopetrol.

Fuente: Elaboración propia con software R.

VII. CONCLUSIONES

Este estudio evaluó la capacidad predictiva de diversos modelos para pronosticar los precios de las acciones de Bancolombia y Ecopetrol. Se compararon modelos ARIMA tradicionales, redes neuronales autorregresivas y modelos híbridos que combinan ambas técnicas. Los resultados mostraron que los modelos híbridos superaron a los modelos individuales al combinar la capacidad de los ARIMA para capturar patrones lineales y la flexibilidad de las redes neuronales para modelar relaciones no lineales. Fue posible llegar a esta conclusión con ayuda de los criterios MAE, MAPE y RMSE.

Adicionalmente, fueron implementados dos modelos de pronósticos para los rendimientos de las acciones de Bancolombia y Ecopetrol. En este caso, se usó un modelo GARCH y una red neuronal autorregresiva con la finalidad de considerar la heterocedasticidad en los rendimientos de las acciones, pues al utilizar este tipo de modelos es posible capturar la variabilidad no constante en el tiempo que presentan estas series. El modelo ganador fue la red neuronal ya que proporciona predicciones más precisas de acuerdo con los criterios del MAE y RMSE.

Los hallazgos de esta investigación tienen importantes implicaciones para los profesionales de las finanzas y la toma de decisiones en el mercado de valores al proporcionar herramientas para predecir los movimientos de los precios y rendimientos de las acciones. Este estudio contribuye a reducir la incertidumbre ya que los modelos híbridos desarrollados ofrecen una mayor precisión en los pronósticos, lo que permite a los inversionistas disminuir el riesgo percibido en sus inversiones. Al contar con pronósticos más confiables, es posible diseñar estrategias de inversión más sólidas y tomar decisiones más informadas sobre cuándo comprar, vender o mantener activos financieros, y esto sirve como base para el desarrollo de nuevos productos financieros y demás estrategias que capitalicen las oportunidades identificadas en el mercado.

Como trabajo futuro, se propone aplicar el análisis realizado a otras series de precios y rendimientos utilizando metodologías de *Deep Learning*, como modelos de redes neuronales de memoria

a largo plazo (LSTM). Estas técnicas podrían proporcionar resultados que resulten ser eficientes para capturar la complejidad de los fenómenos financieros.

REFERENCIAS

- [1] W. D. Lopera Hernández, Comparación de metodologías basadas en una red neuronal artificial y un modelo GARCH para el pronóstico de la volatilidad del precio de las acciones cotizadas en la BVC, Tesis de Maestría, Universidad Nacional de Colombia, Medellín, Colombia, 2023. [En línea]. Disponible en: <https://repositorio.unal.edu.co/handle/unal/84122>
- [2] N. F. Carrillo Sandoval, J. A. Gómez Rodríguez y J. C. Guerrero González, *Los factores Delta y Gamma en las opciones financieras, como medida de riesgo en el mercado accionario colombiano*, Tesis de pregrado, Universidad Piloto de Colombia, Bogotá, Colombia, 2015.
- [3] D. Božanić, A. Randelović, M. Radovanović, y D. , «A hybrid LBWA-IR-MAIRCA multi-criteria decision-making model for determination of constructive elements of weapons,» *Facta Universitatis, Series: Mechanical Engineering*, vol. 18, no. 3, pp. 399-418, 2020, doi: 10.22190/FUME200528033B.
- [4] B. Rodriguez-Satizabal, «Only one way to raise capital? Colombian business groups and the dawn of internal markets,» in *Varieties of Capitalism Over Time*, Routledge, 2022, pp. 133-154, doi: 10.1080/00076791.2020.1796973.
- [5] A. Fournies, *Modelos ARMA y Box and Jenkins*, Universidad Técnica Federico Santa María, Chile, 2015. doi: 10.13140/RG.2.1.2907.0883.
- [6] E. Uriel y A. Peiró, *Introducción al análisis de series temporales*. Madrid, España: Paraninfo, 2000.
- [7] P. Rodó, «Agrupaciones de volatilidad,» *Economipedia*, 2019. [En línea]. Disponible en: <https://economipedia.com/definiciones/agrupaciones-de-volatilidad.html>Economipedia
- [8] W. Bao, Y. Cao, Y. Yang, H. Che, J. Huang y S. Wen, «Data-driven stock forecasting models based on neural networks: A review,» *Information Fusion*, Art. no. 102616, 2024, doi: 10.1016/j.inffus.2024.102616.
- [9] F. Villada, N. Muñoz y E. García, «Aplicación de las Redes Neuronales al Pronóstico de Precios en el Mercado de Valores,» *Información Tecnológica*, vol. 23, no. 4, pp. 11-20, 2012, doi: 10.4067/S0718-07642012000400003.

- [10] J. Murillo S., Á. Trejos y P. Carvajal Olaya, «Estudio del pronóstico de la demanda de energía eléctrica, utilizando modelos de series de tiempo,» *Scientia et Technica*, vol. 3, núm. 23, pp. 37-42, 2003.
- [11] E. L. Gómez-Ramos, F. Venegas-Martínez y H. Allier-Campuzano, «Análisis comparativo entre modelos GARCH y redes neuronales en el pronóstico de los índices bursátiles IPC y Dow Jones,» *Eseconomía: Revista de Estudios Económicos, Tecnológicos y Sociales del Mundo Contemporáneo*, vol. 6, núm. 32, pp. 3-22, 2011.
- [12] E. Raffo Lecca, L. Ráez Guevara y C. Quispe Atúncar, «Aplicación de la metodología GARCH al precio de cierre en la Bolsa de Valores de Lima,» *Industrial Data*, vol. 15, núm. 2, pp. 96-105, 2012, doi: 10.15381/idata.v15i2.6377.
- [13] E. L. Taylor-Conto y R. A. Conto-López, «Comportamiento de la UVR en el largo plazo,» *Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA*, vol. 24, núm. 1, pp. 21-34, 2023, doi: 10.24310/recta.24.1.2023.19860.
- [14] L. Y. Zapata Rodríguez, M. Medina Jiménez y R. A. Conto López, «Pronóstico de tasas USD/COP y EUR/COP mediante diferentes técnicas,» *Revista Ingeniería, Matemáticas y Ciencias de la Información*, vol. 11, núm. 22, pp. 55-62, jul.-dic. 2024, doi: 10.21017/rimci.1077.

