

Aplicación de Redes Neuronales de Elman y Modelos GARCH en el Pronóstico del Precio de Energía Eléctrica en Colombia

Application of Elman Neural Networks and GARCH Models in Forecasting Electricity Prices in Colombia

William Arley Rincón Gómez^a
williamrincon@usantotomas.edu.co

Daniel Esteban Gómez Rodríguez^b
daniel.gomezr@usantotomas.edu.co

Resumen

Este paper presenta el estudio del comportamiento de los precios de la energía eléctrica en el mercado no regulado en Colombia a través de dos metodologías diferentes. Primero, se utilizaron los modelos de heteroscedasticidad condicional autorregresiva generalizada (GARCH) junto con algunas de sus variaciones y, por otro lado, se realizó el análisis de la variable de estudio desde la perspectiva de las redes neuronales recurrentes (RNN), específicamente las redes neuronales recurrentes de Elman. El precio de la energía eléctrica es bastante volátil debido a que existen múltiples factores que influyen en él, como el comportamiento climático—fenómenos del Niño (no llueve) y de la Niña (muchísima lluvia)—, la demanda de los consumidores y los costos de producción, entre otros (Francisco, 2013). Es en este punto donde toman relevancia las metodologías que permiten modelar variables con estas características de volatilidad (Tschora et al., 2022). El objetivo del presente artículo fue realizar el modelamiento del precio de la energía eléctrica y su volatilidad en el mercado no regulado. Se utilizaron y compararon modelos pertenecientes a la familia de los modelos GARCH, así como metodologías de aprendizaje profundo, como las redes neuronales recurrentes (RNN), con el fin de observar si este tipo de modelos se ajustaban y pronosticaban de manera adecuada dicha variable de estudio. Se demostró que, aunque ambos modelos son apropiados para modelar el precio de energía eléctrica no regulada, el modelo de Elman presentó mejores resultados en términos de errores más bajos en los pronósticos.

Palabras clave: Precios, Modelos GARCH, RNN, Energía, Mercado no regulado..

Abstract

This article presented the study of electricity price behavior in the unregulated market in Colombia through two different methodologies. First, generalized autoregressive conditional heteroskedasticity (GARCH) models were used along with some of their variations, and on the other hand, the analysis of the study variable was carried out from the perspective of recurrent neural networks (RNN), specifically Elman recurrent neural networks. Electricity price is quite volatile due to multiple factors that influence it, such as climate behavior, consumer demand, or production costs, among others. It is at this point where methodologies that allow modeling variables with these volatility characteristics become relevant. Therefore, this work proposed to model electricity price and its volatility in the unregulated market. Models belonging to the GARCH family were used and compared, as well as deep learning methodologies such as recurrent neural networks (RNN), in order to observe whether these types of models adequately adjusted and forecasted this study variable. It was demonstrated that, although both models are appropriate for modeling unregulated electricity prices, the Elman model presented better results in terms of lower forecast errors.

Keywords: Prices, GARCH models, RNN, Energy, Unregulated market..

^aUniversidad Santo Tomás, Bogotá-Colombia

^bUniversidad Santo Tomás, Bogotá-Colombia

1. Introducción

La energía eléctrica constituye un factor determinante del desarrollo económico, la productividad industrial y la calidad de vida de la población. Por esta razón, comprender el comportamiento del precio de la electricidad y sus patrones de volatilidad resultó fundamental para la toma de decisiones informadas en tres niveles: consumidores y hogares, quienes deben gestionar su presupuesto energético; el sector industrial, que requiere certidumbre de costos para su competitividad; y los responsables de política pública, encargados de garantizar la seguridad energética y la estabilidad tarifaria del país.

Para el estudio de series temporales, se consideraron en primera instancia los modelos autorregresivos de medias móviles (ARMA) descritos en la metodología planteada por Box and Jenkins (1976). Estos modelos fueron de gran utilidad en el análisis de series de tiempo. Sin embargo, no tuvieron en cuenta la volatilidad de las variables en el modelo, lo cual constituyó una limitación importante, ya que esta metodología asumía que la varianza de las series era constante a lo largo del tiempo (homocedasticidad), un supuesto que no se cumplía en muchas áreas de investigación.

Debido a esto fue necesario implementar alternativas que consideraran la volatilidad como un proceso estocástico con el fin de modelar y analizar su comportamiento histórico. En principio surgieron los modelos de heteroscedasticidad condicional autorregresiva (ARCH) propuestos por Engle (1982), los cuales tuvieron en cuenta la dependencia de los errores pasados para la estimación de la varianza condicional. Posteriormente, Bollerslev (1986) propuso una generalización de dichos modelos conocidos como los modelos GARCH, los cuales además de utilizar los errores pasados utilizaron las volatilidades pasadas para modelar la varianza condicional a lo largo del tiempo.

En aras de complementar estos modelos surgió el modelo exponencial autorregresivo condicional heterocédastico generalizado (EGARCH) propuesto por Nelson (McAleer, 2014); dichos modelos además de tener en cuenta las consideraciones de los modelos GARCH dieron importancia a comportamientos asimétricos que pudieran presentar las series de estudio. Si bien estos modelos enunciados anteriormente presentaron diferencias en su construcción, tuvieron en común que fueron ampliamente utilizados en la modelación de series de tiempo con características de volatilidad no constante a lo largo del tiempo.

En el campo de pronósticos de precios de energía eléctrica, documentos actuales como el de Billé et al. (2023) utilizaron para fines de pronóstico, de manera conjunta con los modelos de la familia GARCH, variables exógenas, mostrando que el uso de variables exógenas como predicciones del consumo y las fuentes de energía renovables mejoraron la precisión del pronóstico de los precios de la electricidad.

Por otro lado, se encontró la metodología de las redes neuronales artificiales (ANN, por su sigla en inglés), la cual fue ampliamente utilizada en las últimas décadas (Fan et al., 2021), debido a los constantes avances en computación. Las ANN están basadas en la forma en que el cerebro humano y las redes neuronales biológicas procesan información. Básicamente, consistieron en unidades de procesamiento llamadas nodos, que estuvieron conectados por valores numéricos llamados pesos, los cuales fueron ajustados para minimizar errores en las estimaciones.

Existieron diferentes tipos de ANN que se pudieron utilizar según el contexto del problema y el tipo de datos disponibles. Para el objetivo de este artículo, se emplearon las RNN; como se describió en Arana (2021), esta estructura de red fue adecuada para el modelado de datos secuenciales, como las series de tiempo. Específicamente, se utilizó la red neuronal recurrente de Elman en este documento, con el propósito de modelar los precios de energía no regulados y comparar sus estimaciones con las obtenidas mediante la metodología de los modelos GARCH.

2. Metodología

Una de las metodologías que permitió pronosticar valores futuros de una serie de tiempo basándose en valores pasados fue el proceso autorregresivo integrado de medias móviles (ARIMA), también conocido como método univariado de Box y Jenkins (Tsay, 2010). Implementar esta metodología tuvo ciertas restricciones, siendo la más importante que las series fueran estacionarias (Tsay, 2010). Se entendió que una serie era estacionaria

si su media, su varianza y su covarianza eran constantes a lo largo del tiempo. Si la serie fue estacionaria, se procedió a utilizar la metodología de Box y Jenkins propuesta en Box and Jenkins (1976), la cual consistió en la identificación del modelo, la estimación de parámetros, la verificación de supuestos y la decisión de si el modelo era óptimo o no; en caso de no serlo, se repitió el proceso planteando otro modelo.

Obtención y preprocesamiento de datos

Se obtuvieron los datos de XM la cual es una entidad que se encarga de operar el sistema interconectado y administrar el mercado de energía mayorista en Colombia. De su sitio Web se extrajo la información histórica de los precios promedio de energía transada mensualmente en el mercado no regulado en el periodo de tiempo comprendido entre enero del 2006 y octubre de 2022.

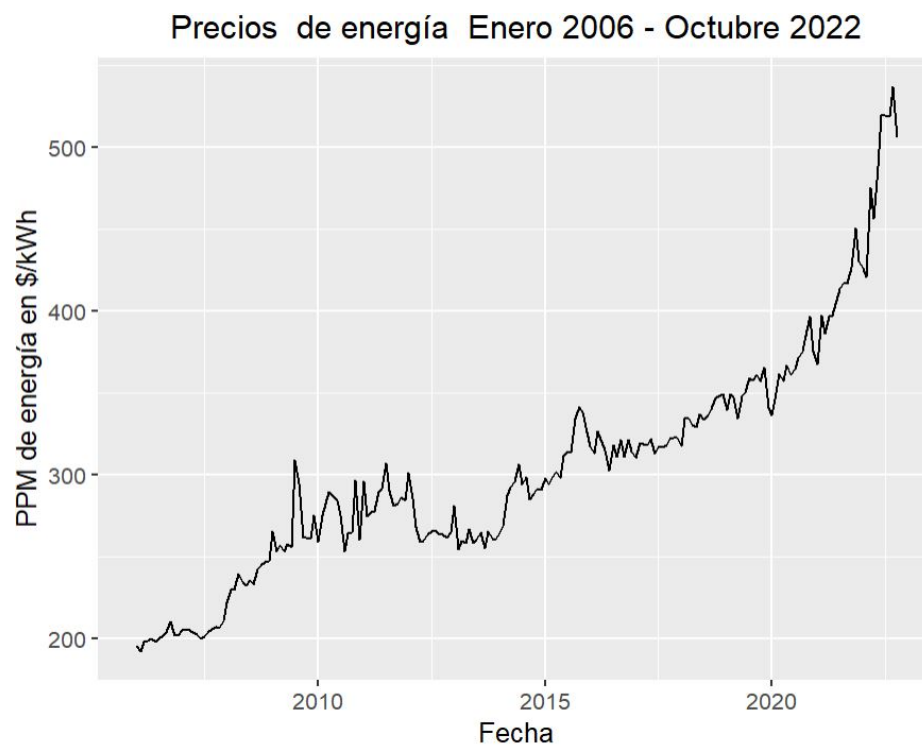


Figura 1: Elaboración propia. Fuente de datos: Portal XM

La Figura 1 es el gráfico de precios promedio mensuales de energía eléctrica en mercado no regulado, se puede observar un comportamiento bastante variable a lo largo del tiempo, con tendencia a la alza lo cual está acorde con el crecimiento de los indicadores económicos del país, se identifica una tendencia alcista marcada a partir del año 2020, superior a la tendencia de los años 2015 a 2020, también se observa una relativa estabilidad en los precios en el periodo de 2010 a 2015, como constante se puede observar la alta volatilidad de la serie, lo cual ya había sido mencionado en los primeros párrafos del presente artículo.

Prueba de estacionariedad para la serie de precios

Los análisis de los modelos de la familia GARCH, requieren que la serie que se está analizando sea estacionaria. Para determinarlo se utiliza la prueba de hipótesis (1) de raíz unitaria de Dickey-Fuller con el siguiente sistema de hipótesis (Tsay, 2010).

$$\begin{cases} H_0 : \text{La serie no es estacionaria} \\ H_a : \text{La serie es estacionaria} \end{cases} \quad (1)$$

Se obtuvo un valor p igual a 0.99, por lo que se puede inferir al no rechazar la hipótesis nula que la serie de precios no es estacionaria. Como la serie de precios no es estacionaria, es necesario realizar una transformación para obtener estacionariedad en la serie.

Transformación de los datos

Se calcularon los retornos logarítmicos Figura 2, para la serie de precios con el fin de obtener estacionariedad en la serie y poder aplicar los modelos propuestos. Para el cálculo de los retornos se utilizó la siguiente fórmula:

$$r_t = \ln \left(\frac{p_t}{p_{t-1}} \right) \quad (2)$$

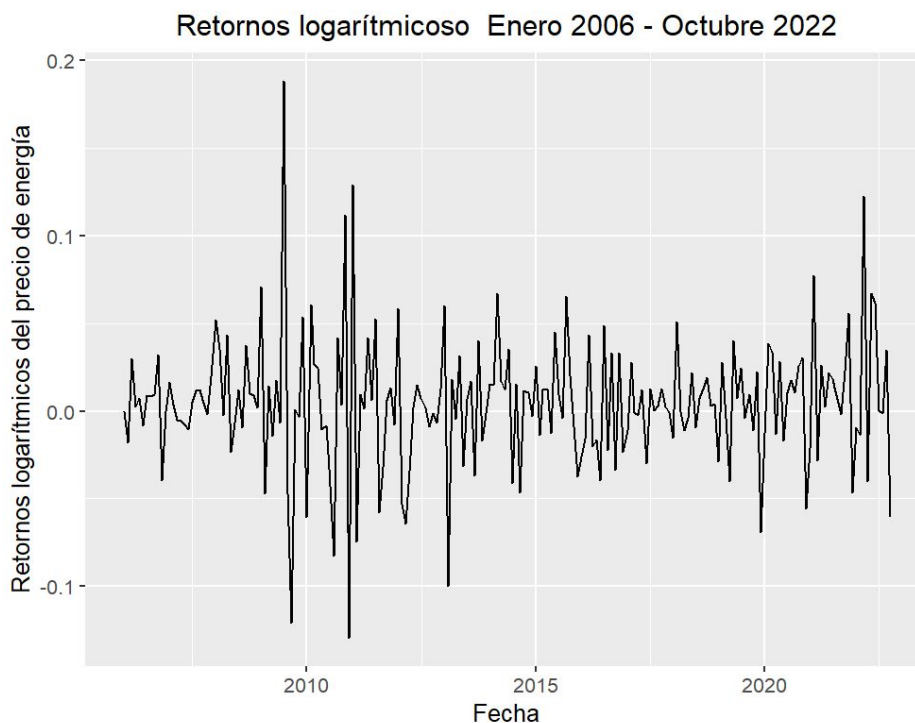


Figura 2: Retornos Logarítmicos. Fuente: Elaboración propia.

Se realizó la gráfica de los retornos logarítmicos de los precios de energía eléctrica transados mensualmente, donde es posible observar en la Figura 2, algunos conglomerados de volatilidad, sobre todo en los primeros meses de estudio, con un salto de volatilidad muy destacado, en octubre y noviembre del año 2009, y una volatilidad marcada a finales del año 2022.

Prueba de estacionariedad para la serie de retornos

Luego de aplicar la prueba de hipótesis (1), a la serie de retornos logarítmicos de los precios de energía eléctrica, se obtuvo un valor p de 0.01, por lo que, a un nivel de significancia del 5 %, se puede inferir que

la serie es estacionaria. La estimación de los modelos se realizará para la serie de los retornos obtenidos, pues cumple con el supuesto de estacionariedad, necesario para aplicar los modelos de Box y Jenkins (Tsay, 2010).

Elección del modelo ARMA

Parte de la metodología planteada Box y Jenkins para los modelos autorregresivos (AR) es descrita por Mora and Losada (2023), en donde se dice que un modelo es autorregresivo si la variable y_t , es explicada por observaciones de ella misma en periodos anteriores añadiéndose un término de error. Expresándose de la siguiente manera:

$$y_t = \phi_0 + \phi_1 y_{t-1} + a_t \quad (3)$$

Donde el termino a_t es ruido blanco con media cero y varianza constante. La expresión (3) hace referencia a un modelo AR(1), de este modo la generalización para un modelo AR(p) está dada en la ecuación (4) como:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + a_t \quad (4)$$

En donde el ruido blanco se modela como (5)

$$a_t \sim RB(0, \sigma_a^2) \quad (5)$$

Los modelos de medias móviles (MA) son aquellos que explican el comportamiento de alguna variable y_t en un periodo de tiempo t por medio de un término independiente y una sucesión de errores sucedidos con anterioridad y ponderados convenientemente (Thomaz et al., 2020). El modelo MA(1) o modelo de medias móviles de orden uno se puede expresar con la ecuación (6), de la siguiente manera:

$$y_t = \mu + \theta_1 a_{t-1} + a_t \quad (6)$$

El cual puede generalizarse a (7) como un modelo MA(q) o modelo de medias móviles de orden q de la siguiente manera:

$$y_t = \theta_0 + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q} + a_t \quad (7)$$

Teniendo en cuenta que a_t se modela como la ecuación (5) es decir ruido blanco, con media cero y varianza constante. Cabe destacar que la aproximación de procesos estocásticos por medios de modelos AR(p) y MA(q) está condicionado a aquellos procesos que cumplan las condiciones de estacionariedad, ya mencionadas.

Se puede considerar a los modelos ARMA como la generalización de los modelos AR y MA en donde se incluye una parte autorregresiva y otra parte de medias móviles para modelar los valores de la serie. El modelo denominado ARMA (p,q) se describe con la ecuación (8) como:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q} \quad (8)$$

Con a_t modelado como la ecuación (5). Para la elección del modelo de media condicional, se realiza el análisis gráfico de los retornos logarítmicos y se acompañan de las pruebas estadísticas pertinentes.

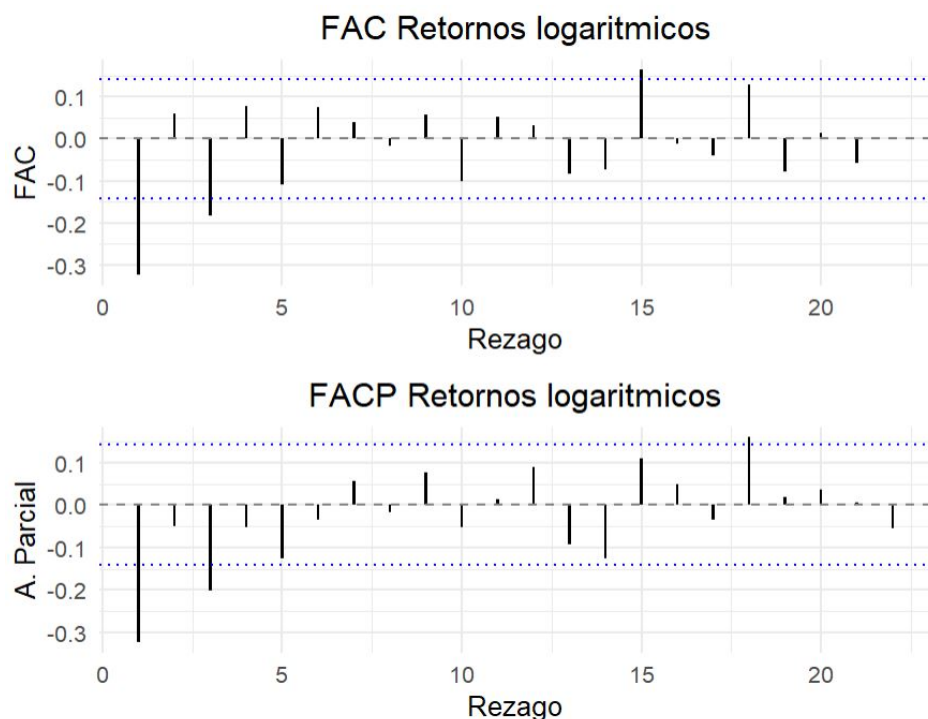


Figura 3: Gráfico de la función de autocorrelación y autocorrelación parcial. Fuente: Elaboración propia.

De acuerdo con la función de autocorrelación y autocorrelación parcial de los retornos logarítmicos, se puede observar Figura 3, que hay autocorrelaciones significativas en el primer y tercer rezago, lo que sugiere posibles órdenes ARMA (1,0), (1,1), (3,0), (0,1), (0,3), entre otros. Después de realizar un análisis de los supuestos de los residuos de los modelos propuestos, y teniendo en cuenta las pruebas de hipótesis adecuadas, se obtuvieron los siguientes resultados mostrados en Tabla 1, para diferentes órdenes ARMA (p,q).

Tabla 1: Análisis de supuestos. Fuente: Elaboración propia.

p,q	Signif.	Media 0	p-valor	No correl.	p-valor	Normalidad	p-valor
0,1	✓	✓	0,992	✓	0,479	X	4,60E-07
1,0	✓	✓	0,990	✓	0,351	X	9,77E-08
1,1	✓	✓	0,996	✓	0,608	X	4,57E-07
0,3	X	✓	0,997	✓	0,915	X	1,66E-07
3,0	X	✓	0,993	✓	0,842	X	2,80E-07
3,3	X	✓	0,992	✓	0,801	X	7,88E-07
3,1	X	✓	0,994	✓	0,957	X	1,97E-07
1,3	X	✓	0,995	✓	0,968	X	4,87E-07

En términos de los supuestos, todos los modelos se comportaron de manera similar; sin embargo, los modelos donde se incluyen tres retardos, ya sea en la parte de medias móviles o autorregresiva, presentaron parámetros que no son significativos. Por lo tanto, los modelos definitivos postulados para modelar la media son el ARMA (0,1), ARMA (1,0) y ARMA (1,1). Posteriormente, se compararon en la Tabla 2, se presentan los términos del criterio de Akaike (AIC) y el criterio de información bayesiana (BIC) para escoger el modelo definitivo para la media.

Tabla 2: Criterios de información. Fuente: Elaboración propia.

Modelo	AIC	BIC
ARMA(0,1)	-737,94*	-728,19*
ARMA(1,0)	-734,89	-725,13
ARMA(1,1)	-737,69	-724,68

El mejor modelo para la media en términos del AIC y BIC es el modelo ARMA(0,1), el cual presenta menores valores en los resultados de los dos criterios.

Elección del modelo de heterocedasticidad condicional

Para entender las generalidades de los modelos de heterocedasticidad condicional es importante conocer primero el concepto de volatilidad. Según Rossi (2013), la volatilidad es un concepto que se refiere a la inestabilidad o variabilidad en los precios de alguna serie financiera, además tiene una característica particular y es que no es directamente observable, por lo que general se utilizan medidas de dispersión como la desviación estándar para tener un estimado de dicha variable, también es posible detectarla en una serie financiera por medio de lo que se conocen como conglomerados de volatilidad los cuales son periodos de bastante variación en lapsos de tiempo determinados. Es posible expresar las series de tiempo por medio de un modelo lineal ecuación (9) y la suma de dos componentes de esta forma.

$$y_t = \mu_t + a_t = E(y_t|I_{t-1}) + a_t \quad (9)$$

Donde I_{t-1} representa toda la información de la serie hasta el periodo $t - 1$. En los modelos estudiados anteriormente el objetivo principal era modelar $\mu_t = E(y_t|I_{t-1})$ asumiendo que a_t es un proceso de ruido blanco tal que $a_t \sim RB(0, \sigma_\epsilon^2)$. Dichos modelos asumían que la varianza de a_t era constante a través del tiempo, pero este supuesto cambia en los modelos de heterocedasticidad condicional ya que estos asumen que el segundo momento condicional o varianza condicional depende del tiempo como se representa en la siguiente ecuación (10), es decir:

$$\sigma_t^2 = \text{Var}(y_t|I_{t-1}) = E((y_t - \mu_t)^2|I_{t-1}) = E(a_t^2|I_{t-1}) \quad (10)$$

Es decir, esta metodología modela la volatilidad a través de la varianza condicional.

Modelo Autorregresivo Condicional Heterocédastico (ARCH)

Los modelos ARCH propuestos por Engle (1982), son modelos que tienen en cuenta la dependencia de los errores pasados ver ecuaciones, para la estimación de la varianza condicional, los modelos tienen la siguiente estructura (11):

$$\begin{aligned} r_t &= \mu_t + a_t \\ a_t &= \sigma_t^2 \epsilon_t \\ \text{Con } \epsilon_t &\sim iid(0, 1) \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 \end{aligned} \quad (11)$$

Como se describe en Tsay (2010), r_t representa los retornos de la serie de tiempo, μ_t la media condicional modelada generalmente por un modelo ARMA, a_t se refiere al término de error, y ϵ_t es una secuencia de variables aleatorias idénticamente distribuidas con media cero y varianza uno, σ_t^2 representa la varianza condicional modelada a través de los errores pasados al cuadrado. Los parámetros del modelo de volatilidad

deben cumplir ciertas restricciones como con relación a sus parámetros, estas son, que $\alpha_0 > 0$, $\alpha_i \geq 0$ y $\sum_{i=1}^m \alpha_i < 1$.

De acuerdo con los gráficos de autocorrelación y autocorrelación parcial presentados en la Figura 4, se puede observar la presencia de autocorrelaciones significativas en ciertos rezagos. Esto puede indicar la existencia de heterocedasticidad condicional o el efecto ARCH. Sin embargo, es necesario realizar pruebas de hipótesis adicionales para corroborar esta afirmación.

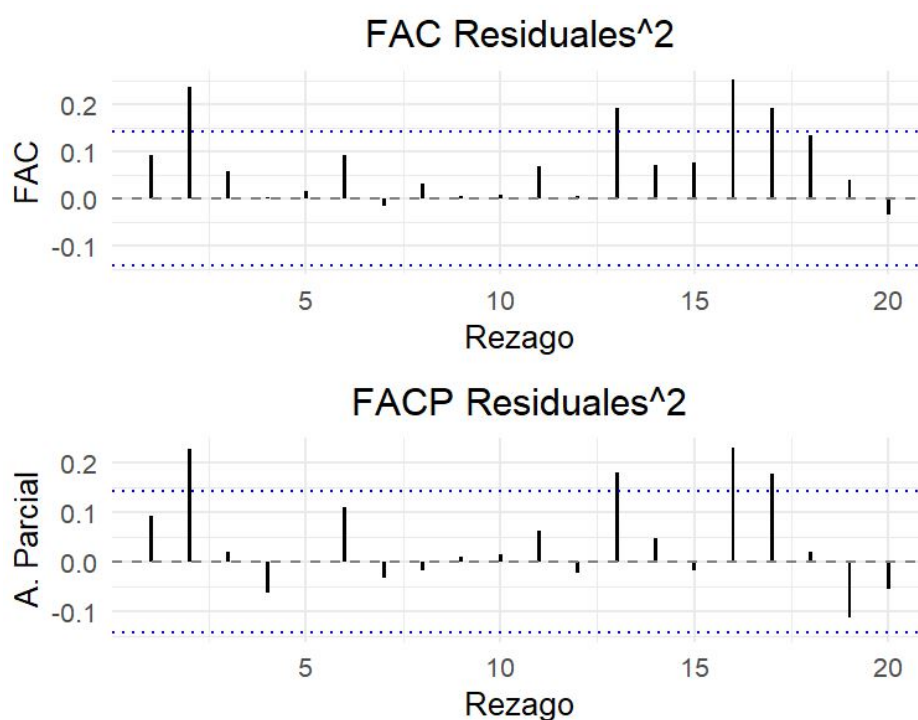


Figura 4: Gráfico FAC y PAF de los residuales al cuadrado. Fuente: Elaboración propia.

La hipótesis para la prueba de Ljung-Box (Tsay, 2010), la hipótesis es nula que todas las autocorrelaciones son cero contra la alternativa que al menos una es diferente de cero (Tsay, 2010).

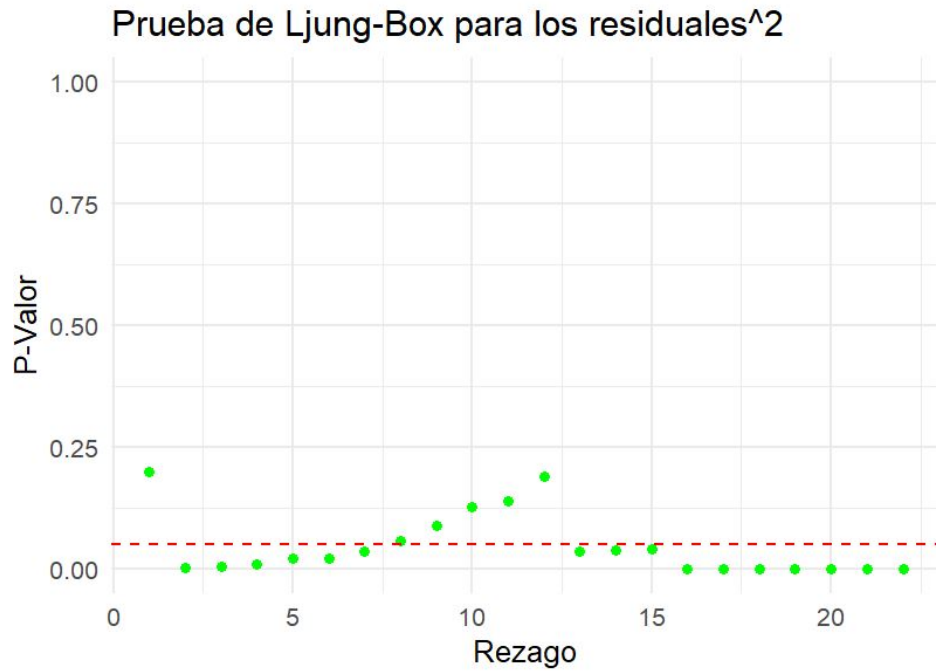


Figura 5: Gráfico de los p valores para diferentes rezagos. Fuente: Elaboración propia.

Los resultados de la prueba arrojan valores p menores a 0.05 en diferentes rezagos, lo que permite concluir que existen autocorrelaciones significativas en los residuos cuadrados del modelo. Se puede observar en la Figura 5, diferentes valores p con diferentes rezagos para el cálculo de la prueba. Los valores de p para esta prueba, permiten verificar la presencia de efecto ARCH.

Modelo Exponencial Autorregresivo Condicional Heterocédastico Generalizado

Los modelos GARCH son bastante útiles para modelar la volatilidad en las series de tiempo, pero tienen algunas limitaciones. Por esta razón Nelson (Tsay, 2010), propone los modelos de heteroscedasticidad condicional autorregresiva exponencial generalizada (EGARCH) con el objetivo de tener en cuenta en la estimación los efectos asimétricos de los choques positivos y negativos en la serie de datos. El modelo está dado en la ecuación (12). La expresión $g(\epsilon_t)$ puede ser reescrita de la siguiente manera para observar el efecto de asimetría producido por los choques aleatorios, ya sean negativos o positivos. Cabe destacar que ϵ_t es una secuencia de variables aleatorias idénticamente distribuidas con media cero y varianza uno.

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^m \alpha_i \ln(\sigma_{t-i}^2) + \sum_{j=1}^s \beta_j g(\epsilon_{t-j}) \quad (12)$$

Con $g(\epsilon_t) = \theta \epsilon_t + \gamma[|\epsilon_t| - E(|\epsilon_t|)]$

$$g(\epsilon_t) = \begin{cases} (\theta + \gamma)\epsilon_t - \gamma E(|\epsilon_t|) & \text{Si } \epsilon_t \geq 0 \\ (\theta - \gamma)\epsilon_t - \gamma E(|\epsilon_t|) & \text{Si } \epsilon_t < 0 \end{cases}$$

Si ϵ_t se distribuye normal su valor esperado está dado en la ecuación (13).

$$E(|\epsilon_t|) = \sqrt{\frac{2}{\pi}} \quad (13)$$

Si ϵ_t se distribuye T-student entonces su valor esperado (14) se presenta por la ecuación:

$$E(|\epsilon_t|) = \frac{2\sqrt{v-2}\Gamma(\frac{v+1}{2})}{(v-1)\Gamma(\frac{v}{2})\sqrt{\pi}} \quad (14)$$

En un principio, se propusieron tres modelos de volatilidad, los cuales fueron un ARMA(0,1)-GARCH(1,1), ARMA(0,1)-MGARCH(1,1) y un ARMA(0,1)-EGARCH(1,1). A los modelos propuestos se les realizó un análisis de supuestos de los residuales estandarizados y se obtuvieron los siguientes resultados. Se examinó la hipótesis de media cero en los residuales estandarizados t-student descritos anteriormente. También se realizó la prueba de no correlación de Ljung-Box para los residuales estandarizados y los residuales estandarizados al cuadrado. Además, se obtuvieron los siguientes resultados en la Tabla 3 y 4.

Tabla 3: Validación de los supuestos

Modelo	Media 0	p-valor	No Correl. a_t/σ_t	p-valor
ARMA(0,1)-GARCH(1,1)	✓	0,7968	✓	0,8291
ARMA(0,1)-MGARCH(1,1)	✓	0,8972	✓	0,812
ARMA(0,1)-EGARCH(1,1)	✓	0,692	✓	0,839

Tabla 4: Validación de los supuestos (continuación)

Modelo	No Correl. a_t^2/σ_t^2	p-valor	Normalidad	p-valor
ARMA(0,1)-GARCH(1,1)	✓	0,9883	X	2,29E-08
ARMA(0,1)-MGARCH(1,1)	✓	0,9944	X	2,23E-08
ARMA(0,1)-EGARCH(1,1)	✓	0,451	X	0,04575

Los tres modelos planteados presentaron (ver Tabla 4) un buen ajuste con respecto a los supuestos de media cero y no correlación en los residuales estandarizados y los residuales estandarizados al cuadrado; sin embargo, el supuesto de normalidad no se cumple en ninguno de los tres modelos.

Elección del modelo de acuerdo al error medio absoluto (MAPE) y la raíz del error cuadrático medio ECM

Para el cálculo del MAPE y ECM y la prueba de los modelos en estudio, se realizan predicciones para los diez períodos posteriores que se habían dejado fuera de la estimación con el fin de evaluar la capacidad de predicción de los modelos estimados. Dado que los modelos pronostican la media de los retornos, se utilizó la fórmula (15) para calcular una estimación de los precios futuros. Despejando p_t de la ecuación (2) de los retornos logarítmicos, se obtiene (15), escrita en la siguiente expresión.

$$p_t = p_{t-1}e^{r_t} \quad (15)$$

Se utilizó la expresión anterior para el pronóstico de los precios futuros usando cada uno de los modelos y posteriormente de cálculo el MAPE y ECM para cada uno de ellos. A continuación, se presenta en la Tabla 5, los resultados obtenidos. Se puede observar en la Tabla 5, que el modelo que presenta un menor MAPE es el ARMA(0,1)-EGARCH(1,1).

Tabla 5: Modelo utilizando la distribución T_v . Fuente: elaboración propia.

MODELO	MAPE	Raíz ECM
ARMA(0,1)-GARCH(1,1)	9,18 %	52,78
ARMA(0,1) - GARCH-M(1,1)	9,38 %	54,00
ARMA(0,1)-EGARCH(1,1)	9,07 %	52,09

Si bien el modelo ARMA(0,1)-EGARCH(1,1) es el que ofrece mejores estimaciones, se realiza la prueba de estimación presentada en la Tabla 6 validada con los mismos datos, utilizando una distribución t-student, considerando la posible presencia de colas pesadas en la distribución de los residuos, con el fin de observar si la estimación del modelo con dicha distribución mejora las predicciones y se concluye que el modelo utilizando la distribución t con v-grados de libertad T_v mejoras las predicciones en términos MAPE y ECM por lo que será el modelo definitivo.

Tabla 6: Estimación con $\epsilon_t \sim T_v$. Fuente: elaboración propia.

Modelo	MAPE	Raíz ECM
ARMA(0,1)-EGARCH(1,1)	8.722 %	50,0268

Verificación del modelo ARMA(0,1)-EGARCH(1,1)

Media cero. En primer lugar, se realizó una prueba t de Student para verificar que los residuos del modelo estimado tienen media cero. Se utilizó el siguiente sistema de hipótesis (16).

$$\begin{cases} H_0 : \mu = 0 \\ H_a : \mu \neq 0 \end{cases} \quad (16)$$

En la prueba se obtuvo un valor p de 0,7659, por lo que se puede inferir que la media de los residuos del modelo es igual a 0.

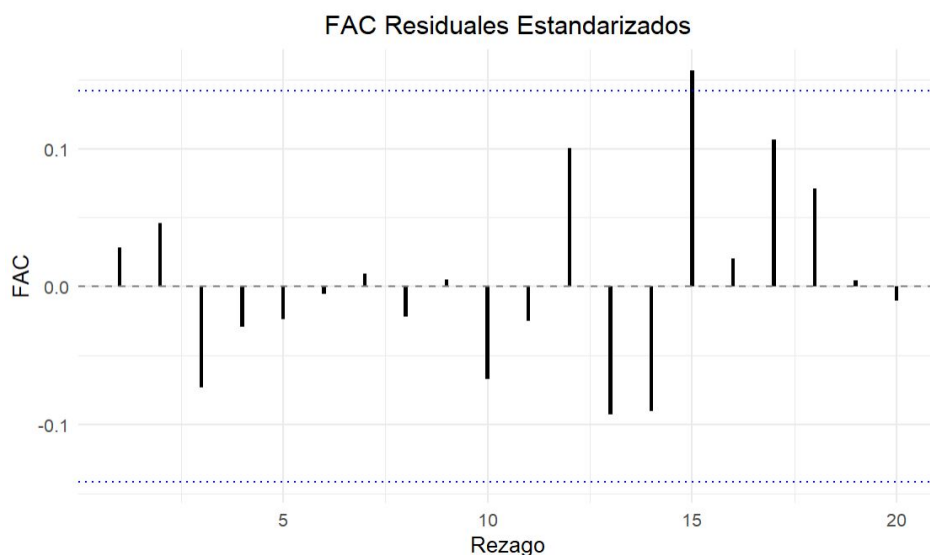


Figura 6: Función de autocorrelación de los residuos. Fuente: Elaboración propia.

Se graficó la función de autocorrelación para los residuos estandarizados Figura 6, donde se puede observar que no hay presencia de autocorrelación. Por lo tanto, se valida el modelo para la media.

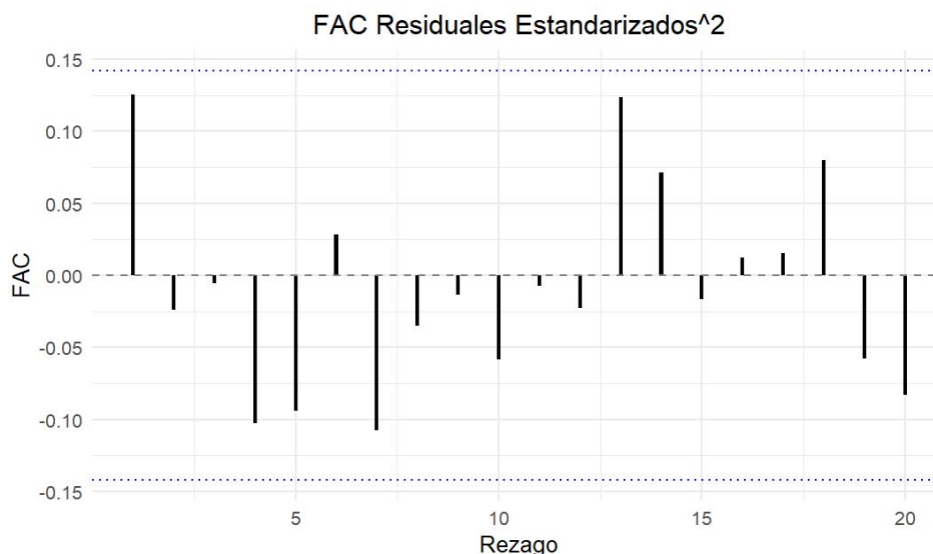


Figura 7: Función de autocorrelación de Residuales al cuadrado. Fuente: Elaboración propia.

Se graficó la función de autocorrelación para los residuales estandarizados al cuadrado Figura 7, y se observó que no se presentan autocorrelaciones significativas. Por lo tanto, se valida el modelo para la varianza. Si bien estos gráficos de autocorrelación son un buen indicio para inferir que no hay autocorrelaciones significativas en los residuos estandarizados y al cuadrado, también se calculó la prueba de Ljung-Box para confirmarlo. Se pudo observar que para todos los rezagos se obtienen valores p mayores al 5 %, por lo que se infiere que las autocorrelaciones para los residuos estandarizados al cuadrado son iguales a 0. Esto valida definitivamente la ecuación para la varianza. Se pudo observar que para todos los rezagos se obtienen valores p mayores al 5 %, por lo que se infiere que las autocorrelaciones de los residuos estandarizados son iguales a 0. Esto valida definitivamente la ecuación para la media.

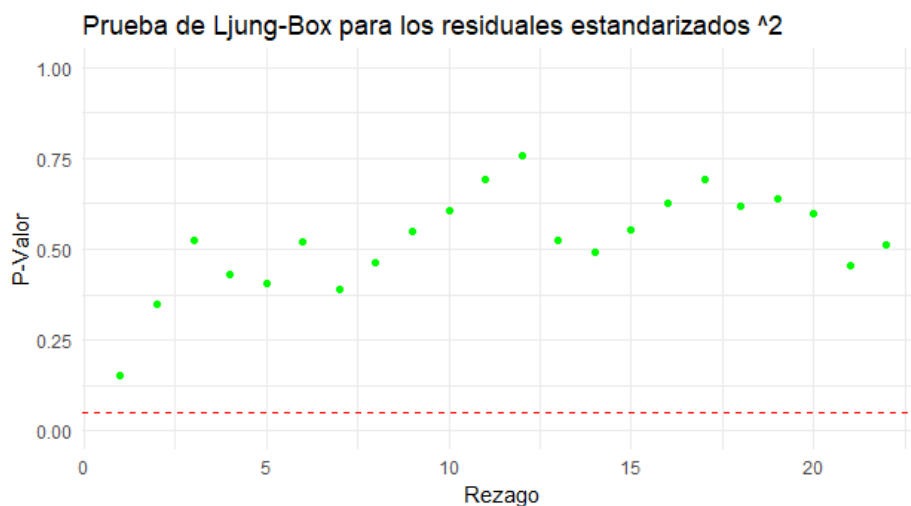


Figura 8: Gráfico de los p valores para diferentes rezagos. Fuente: Elaboración propia.

Modelo ARMA(0,1)- EGARCH(1,1)

La Tabla 7 presenta los valores de los parámetros para el modelo seleccionado, el proceso de construcción del modelo es presentado en las ecuación (17) a la (22).

Tabla 7: Parámetros del modelo estimado. Fuente: Elaboración propia.

Parámetro	Estimado	P-Valor
μ	0,0057	0
θ_1	-0,2941	0
α_0	-0,0629	0
δ_1	0,4404	0
β_1	0,9896	0
γ_1	-0,2162	0
v	2,3706	0

$$r_t = 0,0057 - 0,2941a_{t-1} + a_t; \quad \text{con } a_t = \sigma_t \epsilon_t \quad (17)$$

$$\epsilon_t \sim T(2, 3706)$$

$$\ln \sigma_t^2 = -0,0629 + 0,9896 \ln(\sigma_{t-1}^2) + 0,4404 \left| \frac{\epsilon_{t-1}}{\sigma_{t-1}} \right| - 0,2162 \frac{\epsilon_{t-1}}{\sigma_{t-1}} - E(|\epsilon_{t-1}|) \quad (18)$$

$$E(|\epsilon_{t-1}|) = \frac{2\sqrt{2,3706 - 2}\Gamma(\frac{2,3706+1}{2})}{(2,3706 - 1)\Gamma(\frac{2,3706}{2})\sqrt{\pi}} = 0,4924 \quad (19)$$

Calculando el valor de $E(|\epsilon_{t-1}|)$, en la ecuación 19, sumando algunas constantes y expandiendo el término de valor absoluto, se obtiene la siguiente expresión que permite visualizar el efecto de la asimetría en la estimación de la volatilidad.

$$\ln \sigma_t^2 = 0,0435 + 0,9896 \ln \sigma_t^2 + \begin{cases} (0,4404 + -0,2162)\epsilon_{t-1} & \text{si } \epsilon_{t-1} \geq 0 \\ (0,4404 - -0,2162)\epsilon_{t-1} & \text{si } \epsilon_{t-1} < 0 \end{cases} \quad (20)$$

La ecuación (20) se puede reescribir como la ecuación (21).

$$\ln \sigma_t^2 = 0,0435 + 0,9896 \ln \sigma_t^2 + \begin{cases} 0,2242\epsilon_{t-1} & \text{si } \epsilon_{t-1} \geq 0 \\ 0,6566\epsilon_{t-1} & \text{si } \epsilon_{t-1} < 0 \end{cases} \quad (21)$$

Para determinar el efecto de asimetría en la estimación de la σ_t^2 se asume un choque positivo de $\epsilon_{t-1} = 1$ y un choque negativo $\epsilon_{t-1} = -1$, y se compararon de la siguiente manera utilizando la ecuación (22).

$$\frac{\sigma_t^2(\text{si } \epsilon_{t-1} = 1)}{\sigma_t^2(\text{si } \epsilon_{t-1} = -1)} \quad (22)$$

Lo anterior se puede expresar de la siguiente manera.

$$\frac{e^{0,0435+0,9896 \ln \sigma_t^2+0,2242(1)}}{e^{0,0435+0,9896 \ln \sigma_t^2+0,6566(-1)}} \\ e^{0,2242} e^{-(-0,6566)} \approx 2,41$$

De acuerdo al resultado, se puede decir que el impacto en la volatilidad que tiene un choque positivo es aproximadamente 2,41 veces más grande que el impacto de un choque negativo de la misma magnitud.

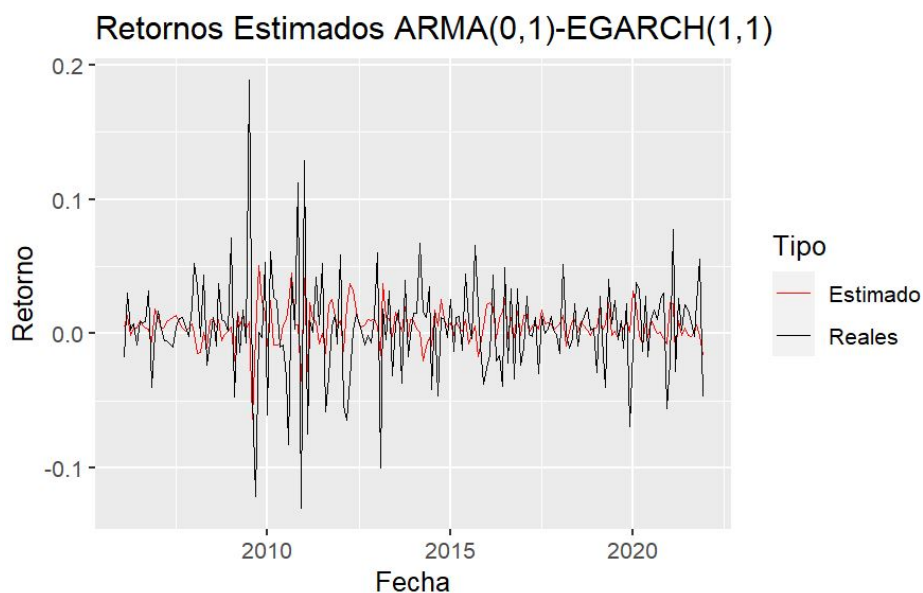


Figura 9: Retornos estimados por el modelo Vs retornos. Fuente: Elaboración propia.

Se puede observar que el modelo EGARCH pronostica bien en términos generales la media de los retornos; sin embargo, no logra captar algunos picos de los retornos reales cuando hay altos picos de volatilidad. La gráfica de volatilidad estimada indica que en los primeros periodos de estudio se presentó una volatilidad mucho más alta que en los demás periodos.

Redes neuronales Artificiales

Las redes neuronales artificiales son modelos de aprendizaje profundo inspirados en el funcionamiento de las redes neuronales biológicas. Básicamente, para una variable de entrada, una red neuronal lleva a cabo una secuencia de pasos de procesamiento para posteriormente devolver una salida. La arquitectura de una red neuronal está compuesta por capas y nodos. Generalmente, las redes neuronales cuentan con una capa de entrada, capas ocultas y una capa de salida. En cada una de las capas se encuentran nodos o neuronas interconectadas entre sí; dichas conexiones reciben el nombre de pesos, los cuales son optimizados para mejorar el proceso de estimación de una o más variables de estudio (Yamazaki et al., 2022).

Es posible clasificar las redes neuronales artificiales en dos tipos distintos. las redes neuronales artificiales de una sola capa, también conocidas como perceptrones, y las redes neuronales artificiales de múltiples capas (Gu et al., 2018). A continuación, en la Figura 10, se muestra una ilustración.

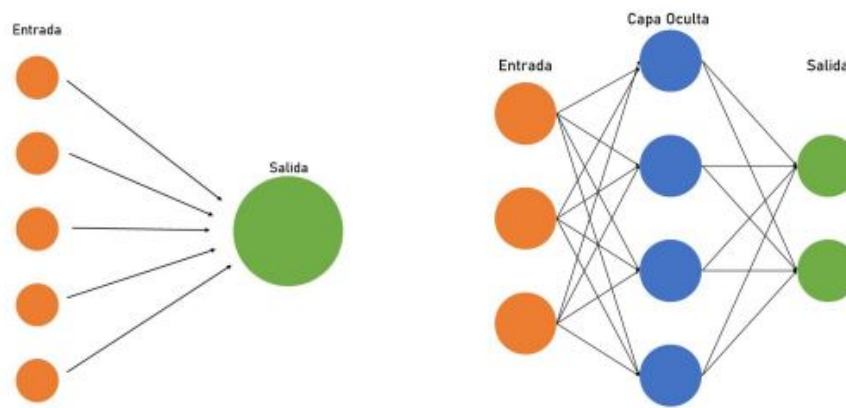


Figura 10: Red neuronal de múltiples capas. Fuente: Mendoza Gómez (2021)

Funciones de Activación

Son funciones que cumplen con el objetivo de restringir el rango de salida de las neuronas a un intervalo determinado. Estas funciones pueden ser lineales o no lineales, y su elección dependerá del tipo de problema de estudio. De acuerdo con el rango de la función de activación que se haya escogido, es necesario realizar transformaciones en los datos de entrada para que se ajusten a estos rangos (Ripley, 1996). Las funciones de activación más utilizadas son las siguientes.

Función Sigmoide: El rango de esta función está comprendido entre (0-1). Cuando los valores de entrada tienden a infinito, el rango tiende a uno, y cuando tienden a menos infinito, el rango tiende a cero, ecuación (23).

$$z(x) = \frac{1}{1 + e^{-x}} \quad (23)$$

Función Tangente Hiperbólica: Esta función tiene un comportamiento similar a la función sigmoide; sin embargo, el rango de salida estará comprendido en el intervalo (-1, 1), ver ecuación (24). En esta función, los valores negativos obtendrán un valor cercano a -1, los valores cercanos a 0 obtendrán un valor de salida cercano a cero, y los valores positivos obtendrán salidas cercanas a 1.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (24)$$

Función Identidad: Se caracteriza por ser una función matemática simple ecuación (25) que toma un valor de entrada y devuelve el mismo valor como salida.

$$f(x) = x \quad (25)$$

Componentes y funcionamiento general de las redes neuronales

En Ripley (1996) se describen los componentes de las redes neuronales de la siguiente manera. Un conjunto de entradas x_i junto con unos pesos sinápticos w_{ij} . Una regla de propagación g_i definida a partir de las variables de entrada x_i y los pesos sinápticos w_{ij} . Por lo general, la regla de propagación viene dada por una combinación lineal entre los pesos y las variables de entrada ecuación (26) de la siguiente forma.

$$g(x_1, x_2, \dots, x_n, w_{i1}, \dots, w_{in}) = \sum_j w_{ij} x_j + b \quad (26)$$

Donde b se conoce como un término de sesgo. Una función de activación, la cual representa la salida de la neurona. Denotando esta salida como y_i , se tiene lo siguiente.

$$y_i = f(g(x_1, x_2, \dots, x_n, w_{i1}, \dots, w_{in})) = f\left(\sum_j w_{ij}x_j + b\right) \quad (27)$$

Entrenamiento de la Red

Como se describe en Alamilla Jiménez (2021) para el entrenamiento de la red, se lleva a cabo un proceso de optimización conocido como descenso de gradiente, el cual se basa en minimizar una función de costo o pérdida. Existen varias funciones de costo, sin embargo, una de las más utilizadas ecuación (28) es el Error Cuadrático Medio, dado por la siguiente expresión.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (28)$$

Donde n es la cantidad de datos de entrada de entrenamiento, y_i es el valor observado y $\hat{y}_i$ es el valor predicho por el modelo. El objetivo principal del algoritmo del descenso de gradiente con MSE es hallar los valores óptimos para los pesos w y los sesgos b de la red neuronal para minimizar el Error Cuadrático Medio (MSE). Esto se logra mediante actualizaciones iterativas a los pesos y sesgos utilizando las derivadas parciales del MSE con respecto a cada parámetro y controlando el tamaño de los pasos de actualización a través de la tasa de aprendizaje α . En resumen, el algoritmo busca ajustar los pesos y sesgos de manera que las predicciones del modelo se acerquen lo más posible a los valores reales, reduciendo así el error cuadrático medio (Heydari et al., 2020). La actualización de los pesos en cada iteración del algoritmo del descenso de gradiente con MSE se puede resumir de la siguiente manera:

$$w_{t-1} = w_t - \alpha \frac{\partial MSE}{\partial w} \quad (29)$$

$$b_{t-1} = b_t - \alpha \frac{\partial MSE}{\partial b} \quad (30)$$

Redes Neuronales Recurrentes RNN

Las redes neuronales recurrentes (RNN) son una arquitectura especial de redes neuronales artificiales diseñadas para procesar secuencias de datos que muestran una dependencia temporal entre sus partes. A diferencia de las redes neuronales tradicionales, las RNN tienen conexiones recurrentes que les permiten mantener una especie de “memoria” interna o estado oculto, lo que las hace ideales para tareas que involucran datos secuenciales como texto, audio, video o series temporales (Mosso Vázquez et al., 2023).

La característica distintiva de las RNN es la presencia de ciclos en su arquitectura, lo que permite que la salida de una capa se retroalimente como entrada en la siguiente iteración, manteniendo así información relevante a lo largo del tiempo. Esta capacidad de recordar información previa en la secuencia es lo que las hace poderosas para tareas que requieren comprender el contexto y las dependencias temporales entre los datos (Lyu and Liu, 2021). Existen diferentes tipos de redes neuronales recurrentes como las redes neuronales LSTM, las redes neuronales GRU (Gated Recurrent Unit), las redes neuronales de Jordan o las redes neuronales de Elman. En este trabajo se estudiarán específicamente las redes neuronales de Elman debido a su capacidad de captar dependencias temporales a mediano y corto plazo.

Redes Neuronales Recurrentes de Elman

Fueron presentadas por Elman (1990). Este tipo de red recurrente se caracteriza por contar con cuatro capas: una capa de entrada, una capa oculta, una capa de salida y una capa adicional denominada capa de contexto (ilustrado en la Figura 11), la cual guarda la salida de la capa oculta de un paso anterior para utilizarla en el cálculo de la salida oculta del paso siguiente. En Guo et al. (2022) se describe matemáticamente una red neuronal de Elman de forma general de la siguiente manera.

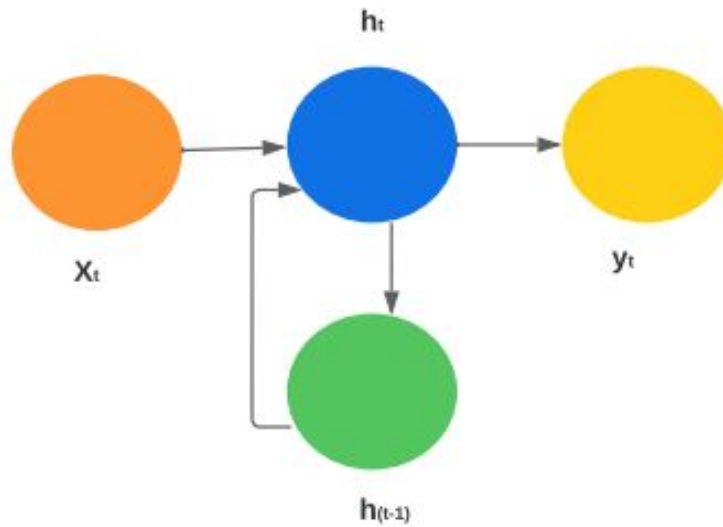


Figura 11: Red neuronal recurrente. Fuente: Elaboración propia

$$h_t = f(W_{ih}x_t + W_{hh}h_{t-1} + b_h) \quad (31)$$

$$y_t = f(W_{ho}h_t + b_o) \quad (32)$$

Donde x_t es el vector de entrada, h_t el vector de salida de la capa oculta, y_t el vector de salida, W_{ih} es la matriz de pesos entre capa de entrada y capa oculta, W_{hh} la matriz de pesos recurrentes entre capa oculta y sí misma, W_{ho} la matriz de pesos entre capa oculta y capa de salida, b_h el vector de sesgo para la capa oculta, b_o el vector de sesgo para la capa de salida y f es la función de activación determinada.

Estimación de las redes neuronales de Elman

Para la estimación de las redes neuronales de Elman, se utilizó una ventana de 12 meses para rezagar las variable de estudio. Además de esto, se empleó la función sigmoide como función de activación, y se variaron el número de capas ocultas, nodos tasas de aprendizaje para determinar qué arquitectura de la red estima mejor la variable de estudio. Posteriormente, se realizó el cálculo del MAPE y el ECM para los valores estimados y los valores reales intramuestra. Además, se calcularon estas medidas tanto para los

valores pronosticados como para los valores reales de prueba, diez periodos adelante, con el fin de observar la capacidad de pronóstico de los modelos estimados. Se realizaron estimaciones utilizando variables explicativas para el modelado de las redes neuronales. Además, se realizaron estimaciones sin variables explicativas y se compararon los resultados obtenidos en ambos casos. Las variables explicativas que se utilizaron fueron las siguientes. Demanda Mensual: la cantidad de giga watts consumidos en este mercado durante el mes. Delta del IPC: Cambio porcentuales en el índice de precios al consumidor IPC presentados mensualmente. Vale destacar que se utilizaron 1,2,3,4 y 5 capas ocultas combinadas con tasas de aprendizaje de 0.01, 0.03 y 0.05.

Transformación de las variables. Debido a que se utilizó la función de activación sigmoide, fue necesario escalar las variables explicativas Demanda Mensual y Delta IPC, así como la variable respuesta Precio de energía, en escala [0-1] utilizando la transformación escrita en la ecuación (35)

$$\frac{y - \text{mín } y}{\text{máx } y - \text{mín } y} \quad (33)$$

Se obtuvieron los siguientes resultados tabla 8.

Tabla 8: Validación de redes de Elman estimadas sin las variables explicativas. Fuente: Elaboración propia.

Capas	Nodos	Tasa	ECM Train	MAPE Train	ECM Test	MAPE Test
1	12	0,01	10,79	2,78 %	31,75	5,55 %
2	12	0,01	12,32	3,30 %	27,02	4,94 %
3	12	0,01	17,13	4,17 %	26,29	4,68 %
4	12	0,01	22,90	5,97 %	25,50	4,67 %
5	12	0,01	34,59	8,83 %	22,60	4,35 %
1	12	0,03	12,08	3,10 %	23,75	4,39 %
2	12	0,03	15,40	4,06 %	21,22	4,19 %
3	12	0,03	20,54	5,16 %	23,98	4,47 %
4	12	0,03	34,86	9,35 %	21,80	4,25 %
5	12	0,03	42,57	11,86 %	24,20	4,58 %
1	12	0,05	13,77	3,44 %	21,60	4,13 %
2	12	0,05	18,56	4,58 %***	20,06	3,89 %***
3	12	0,05	24,06	5,96 %	24,13	4,50 %
4	12	0,05	27,68	6,82 %	26,77	4,82 %
5	12	0,05	35,58	9,40 %	24,51	4,66 %

Después de realizar las estimaciones de las redes neuronales Elman con las diferentes arquitecturas planteadas, utilizando solo una variable, se logró observar que la red que presenta mejor desempeño es la que cuenta con dos capas ocultas, con doce nodos cada una, y una tasa de aprendizaje de 0,05. A continuación, se muestra en la Figura 12, una representación gráfica de la serie de datos estimada por la red de Elman, junto con la serie de datos reales, donde se puede observar que la red neuronal estima de manera adecuada la variable.

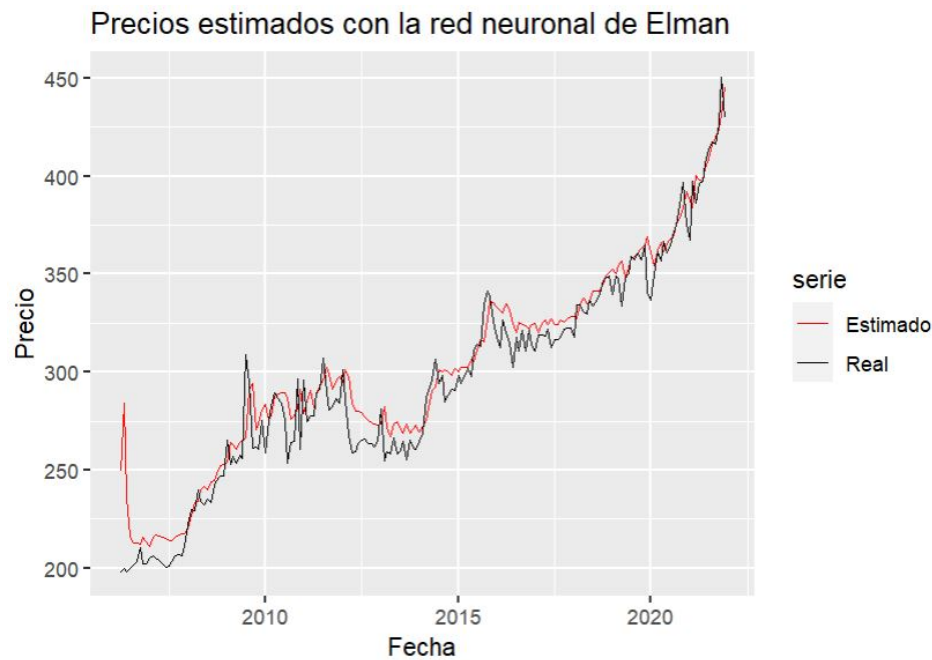


Figura 12: Elaboración propia (fuente de datos: Portal xm)

Resultados con las variables explicativas

Tabla 9: Validación de redes de Elman estimadas (multivariada)

Capas	Nodos	Tasa	ECM Train	MAPE Train	ECM Test	MAPE Test
1	36	0,01	12,73	3,43 %	49,56	6,89 %
2	36	0,01	16,21	4,38 %	38,66	6,83 %
3	36	0,01	21,84	6,03 %	33,03	5,66 %
4	36	0,01	47,52	13,05 %	81,96	14,06 %
5	36	0,01	39,78	9,09 %	37,75	6,84 %
1	36	0,03	13,71**	3,73 %**	22,21**	3,66 %**
2	36	0,03	232,07	76,54 %	133,47	22,64 %
3	36	0,03	35,90	8,33 %	107,62	18,56 %
4	36	0,03	39,82	8,31 %	26,14	4,47 %
5	36	0,03	35,41	8,01 %	68,62	12,21 %
1	36	0,05	14,58	3,67 %	32,37	5,48 %
2	36	0,05	461,77	151,74 %	337,03	69,17 %
3	36	0,05	36,86	6,37 %	120,45	20,49 %
4	36	0,05	44,01	9,74 %	105,57	20,35 %
5	36	0,05	34,07	6,69 %	29,26	5,39 %

Después de realizar las estimaciones de las redes neuronales Elman con las diferentes arquitecturas planteadas, utilizando las variables Demanda y Delta-IPC, se logró observar que la red que presenta mejor desempeño es la que cuenta con una capa oculta, con 36 nodos cada una, y una tasa de aprendizaje de 0,03.

A continuación, se muestra Figura 13, una representación de la serie de datos estimada por la red de Elman, junto con la serie de datos reales, donde se puede observar que la red neuronal estima de manera adecuada la variable.

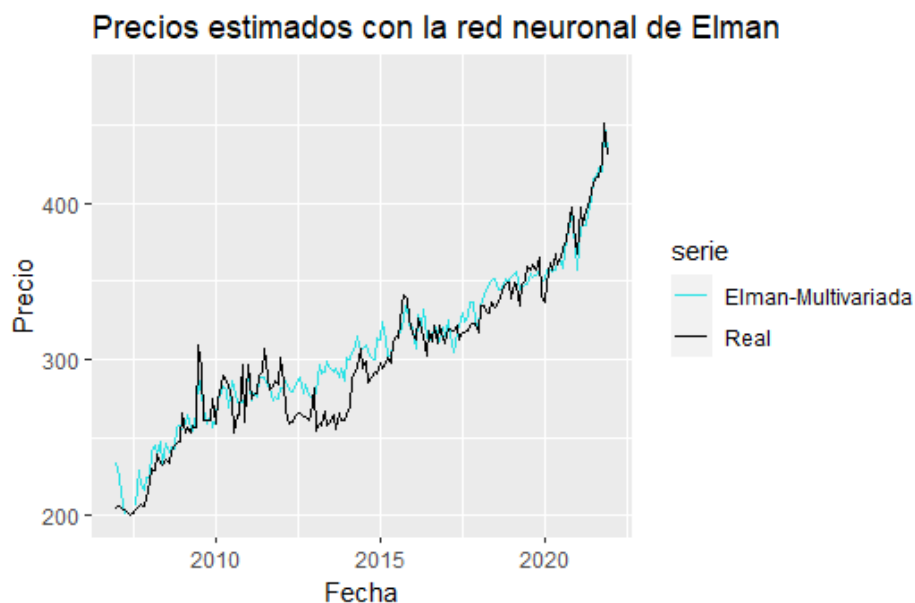


Figura 13: Elaboración propia (fuente de datos: Portal xm)

3. Resultados

Comparación de las dos metodologías

La siguiente Tabla 10, muestra los pronósticos diez períodos adelante por cada una de las metodologías estudiadas.

Tabla 10: Pronósticos 10 periodos adelante

Reales	EGARCH	ELMAN-Univariado	Elman -Multivariado
426,13	437,18	446,96	405,96
420,40	439,69	452,10	433,81
475,17	442,21	462,04	421,35
456,40	444,75	471,11	433,31
488,09	447,30	479,21	470,51
519,09	449,87	492,42	511,90
519,34	452,45	502,91	499,38
518,57	455,05	509,71	514,23
537,01	457,66	520,41	535,25
505,57	460,29	533,54	493,98

A continuación en la figura 14, se muestra una representación gráfica de los pronósticos obtenidos por cada una de las metodologías, donde se puede observar que la red de Elman univariada y multivariada presentan un mejor desempeño que el modelo ARMA(0,1)-EGARCH(1,1).

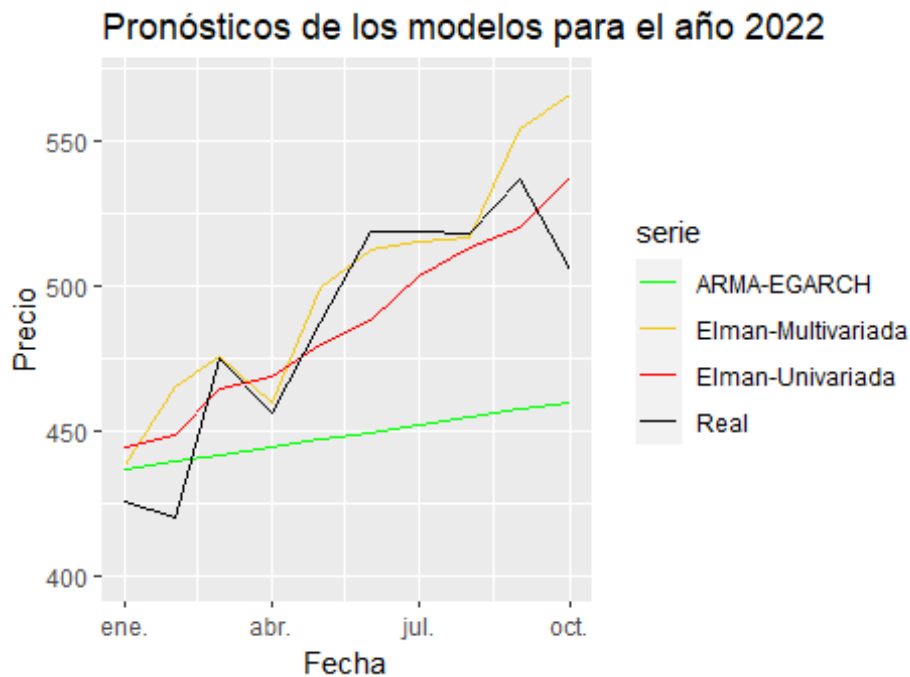


Figura 14: Elaboración propia (fuente de datos: Portal xm)

Posteriormente, se compararon los pronósticos obtenidos con cada metodología en términos del MAPE y ECM, donde se obtuvieron los siguientes resultados.

Tabla 11: Validación de modelos finales

Modelo	MAPE	ECM
ARMA(0,1)-EGARCH(1,1)	8,722 %	50,0268
Red de Elman univariada	3,89 %	20,06
Red de Elman multivariada	3,66 %	22,21

Se observa que en términos del MAPE y ECM los modelos de redes neuronales presentan un mucho mejor ajuste que el modelo ARMA(0,1)-EGARCH(1,1) planteado, por lo que se puede decir que estos modelos de aprendizaje profundo mejoran de manera significativa las estimaciones en este contexto de estudio.

4. Conclusiones

- El modelo de la familia GARCH que presentó mejor ajuste en términos de los criterios de información y medidas de error de predicción es el ARMA(0,1)-EGARCH(1,1).
- Se encontró que los choques positivos generaron un mayor impacto en la volatilidad estimada por el modelo.
- El modelo de Elman univariado que presentó mejor ajuste tiene una arquitectura de dos capas ocultas, con doce nodos en cada capa, y una tasa de aprendizaje de 0,05.
- El modelo de Elman multivariado que presentó mejor ajuste tiene una arquitectura de una capa oculta, con 36 nodos en cada capa, y una tasa de aprendizaje de 0,03.

- Los modelos de redes neuronales de Elman se comportaron bastante bien, tanto intramuestra como en las predicciones, y mejoran significativamente las estimaciones respecto al modelo ARMA(0,1)-EGARCH(1,1) planteado.
- La inclusión de variables explicativas mejoró significativamente las estimaciones, tanto intramuestra como en las predicciones.

Recibido: Agosto de 2025
Aceptado: Diciembre de 2025

Referencias

- E. Alamilla Jiménez. Redes neuronales y su aplicación en la clasificación de patrones, 2021.
- C. Arana. Redes neuronales recurrentes: Análisis de los modelos especializados en datos secuenciales. Documento de trabajo 753, CEMA, 2021.
- Anna Gloria Billé, Angelica Gianfreda, Filippo Del Grosso, and Francesco Ravazzolo. Forecasting electricity prices with expert, linear, and nonlinear models. *International Journal of Forecasting*, 39(2):570–586, 2023.
- Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.
- George EP Box and Gwilym M Jenkins. *Time series analysis: Forecasting and control*. Holden-Day, rev. ed. edition, 1976.
- Jeffrey L Elman. Finding structure in time. *Cognitive Science*, 14(2):179–211, 1990.
- Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.
- Jianqing Fan, Cong Ma, and Yiqiao Zhong. A selective overview of deep learning. *Statistical Science*, 36(2):264–290, 2021.
- A. R. L. Francisco. Determinantes del precio de la energía eléctrica en el mercado de energía de colombia. Master's thesis, Escuela Colombiana de Ingeniería Julio Garavito, 2013. URL <https://repositorio.escuelaing.edu.co/bitstream/001/939/1/Gonzalez%20Garzn%2C%20Cristhian%20Camilo%20-%202019.pdf>.
- Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, and Tsuhan Chen. Recent advances in convolutional neural networks. *Pattern Recognition*, 77:354–377, 2018.
- Y. Guo, D. Yang, Y. Zhang, L. Wang, and K. Wang. Online estimation of soh for lithium-ion battery based on ssa-elman neural network. *Protection and Control of Modern Power Systems*, 7(1):40, 2022.
- A. Heydari, M. Majidi Nezhad, E. Pirshayan, D. Astiaso Garcia, F. Keynia, and L. De Santoli. Short-term electricity price and load forecasting in isolated power grids based on composite neural network and gravitational search optimization algorithm. *Applied Energy*, 277:115503, 2020.
- S. Lyu and J. Liu. Convolutional recurrent neural networks for text classification. *Journal of Database Management*, 32(4):78–92, 2021.
- A. C. Malagón Duque, A. Sellamén Garzón, and V. Nova Rodríguez. Relación entre las fintech y el marketing financiero. *Semestre Económico*, 24(57):57, 2022.
- Michael McAleer. Asymmetry and leverage in conditional volatility models. *Econometrics*, 2(3):145–150, 2014.

- O. Mendoza Gómez. Diseño y aplicación de una red neuronal artificial para categorización e identificación de imágenes histopatológicas. Master's thesis, Benemérita Universidad Autónoma de Puebla, 2021. Repositorio Institucional BUAP.
- C. P. Mora and J. L. Losada. Análisis de series temporales en estaciones meteorológicas para la predicción de la precipitación en la ciudad de manizales, colombia. *Revista de Climatología*, 23:58–70, 2023.
- A. Mosso Vázquez, S. Martínez Juárez, R. Cruz Barbosa, and S. Castro Fernández. Una revisión de los fundamentos de aprendizaje profundo – el modelo de red neuronal xor. *Programación Matemática y Software*, 15(1):47–56, 2023.
- C. A. Pulzara Mora and J. D. Losada Losada. Análisis de series temporales en estaciones meteorológicas para la predicción de la precipitación en la ciudad de manizales, colombia. *Revista de Climatología*, 23: 58–70, 2023.
- Brian D Ripley. *Pattern recognition and neural networks*. Cambridge University Press, 1996.
- G. D. Rossi. La volatilidad en mercados financieros y de commodities. *Escritos de Política Económica*, 16: 1–42, 2013.
- P. S. Thomaz, V. L. D. de Mattos, L. R. Nakamura, A. C. Konrath, and G. D. S. Nunes. Modelos garch em ações financeiras: Um estudo de caso. *Exacta*, 18(3):470–486, 2020.
- Ruey S Tsay. *Analysis of financial time series*. Wiley, 3 edition, 2010.
- L. Tschora, E. Pierre, M. Plantevit, and C. Robardet. Electricity price forecasting on the day-ahead market using machine learning. *Applied Energy*, 313:118752, 2022.
- K. Yamazaki, V. K. Vo-Ho, D. Bulsara, and N. Le. Spiking neural networks and their applications: A review. *Brain Sciences*, 12(7):863, 2022.