



Análisis comparativo de pronósticos del IPC obtenidos mediante modelos GARCH y redes neuronales

Aldo Alejandro Vidal-Castillo
Instituto Politécnico Nacional, México
aldo.vidcas@gmail.com

Francisco Venegas-Martínez
Instituto Politécnico Nacional, México
fvenegas1111@yahoo.com.mx

Francisco Ortiz-Arango
Universidad Panamericana, México
fortizar62@gmail.com

Recibido el 6 de septiembre del 2021; aceptado el 31 de octubre del 2021; disponible en línea el 29 de noviembre del 2021

Resumen: En el presente trabajo realizamos un análisis comparativo de un modelo GARCH frente a un modelo de redes neuronales para el pronóstico del Índice de Precios y Cotizaciones (IPC) de la Bolsa Mexicana de Valores (BMV) tomando en cuenta cuatro aspectos: 1) contrastar los fundamentos teórico-económicos de los modelos GARCH y las redes neuronales; 2) enlistar los recursos necesarios para la implementación de los dos modelos; 3) analizar el horizonte temporal de los pronósticos; y 4) establecer la precisión de los pronósticos obtenidos con los dos modelos con respecto a los datos observados, dentro y fuera de la muestra. Los resultados empíricos obtenidos sugieren que el modelo de red neuronal es mejor que el modelo GARCH para el pronóstico del IPC. El resultado logrado también se sostiene en periodos de alta y baja volatilidad; siendo la implementación de nuestra red neuronal más sencilla.

Palabras clave: Métodos de predicción y pronóstico; mercados financieros; redes neuronales; modelos econométricos.

Clasificación JEL: C53, G10, C45, B23.

Comparative analysis of the IPC forecasts with GARCH models and neural networks

Abstract: In this paper, we perform a comparative analysis of a GARCH model against a neural network model for forecasting the Mexican Stock Market Index (IPC, for its acronym in Spanish) of the Mexican Stock Market (BMV, for its acronym in Spanish), taking into account four aspects: 1) contrasting the theoretical-economic fundamentals GARCH models and neural networks; 2) listing the necessary resources for the implementation of the two models; 3) analyzing the time horizon of the forecasts; and 4) establishing the accuracy of the obtained forecasts from the two models to the actual data, inside and outside the sample. The empirical results suggest that the neural network model is better than the GARCH model to forecast the IPC. This last result holds during high and low volatility periods, making the implementation of our neural network easier.

Keywords: Prediction and forecasting methods; financial markets; neural networks; econometric models.

JEL Codes: C53, G10, C45, B23.



1. Introducción

Un tema relevante en los mercados financieros es la predicción de rendimientos considerando dos cuestiones primordiales, la precisión y la facilidad del cálculo. La precisión de cálculo es el nivel de apego a la realidad y la facilidad de cálculo es la viabilidad y sencillez para calcularlo. Esto es así puesto que frecuentemente no se cuenta con mucho tiempo para tomar decisiones. Las decisiones al igual que sus resultados dependen en gran medida de la manera en que son descritos y realizados los pronósticos (McNelis, 2005 y Lorenzo y Ruiz, 2011). La evolución de los modelos de pronósticos financieros es de suma importancia, y este trabajo de investigación se concentra en dos cuestiones: la primera, teórica, para ayudar a comprender un poco más la naturaleza de los fenómenos financieros (Gómez *et al.*, 2001), y la segunda, práctica, para apoyar en la toma de decisiones (Argáez *et al.*, 2014).

Entre la literatura sobre pronósticos de variables financieras relevantes a través de redes neuronales (RN) vale la pena mencionar a Cruz y Morales (2011) que encontraron que para el pronóstico del futuro del CETE a 91 días en el MexDer un modelo basado Redes Neuronales (RN) pronostica mejor que el modelo ARIMA, aunque este último es más sencillo. Otro estudio de Ortiz *et al* (2013) muestra las ventajas de las Redes Neuronales Diferenciales (RND) sobre los modelos econométricos sobre todo en el caso de sistemas dinámicos con comportamientos no lineales; siendo este el caso más común en los mercados financieros. Asimismo, García (2002) desarrolla un algoritmo mediante una RN supervisada que calcula el precio de una opción de compra de una compañía telefónica utilizando datos históricos del mercado. Los autores generan una red capaz de calcular los precios de varios contratos con una precisión del 98%. Una ventaja de este modelo es que no utiliza la distribución normal. En el trabajo de Ortiz Arango (2016) se realiza un pronóstico del tipo de cambio euro/dólar utilizando RND llegando a la conclusión de que este tipo de red proporciona una reducción significativa en el periodo de aprendizaje, además de que proporciona pronósticos en la primera semana con un error relativo de .002%

En cuanto a los modelos GARCH, Lorenzo (2011) modela los rendimientos bursátiles mexicanos mediante modelos TGARCH y EGARCH, llegando a la conclusión que mediante los modelos GARCH se perciben los efectos de apalancamiento. Así mismo, el autor destaca que los modelos EGARCH describen adecuadamente los rendimientos agregados del mercado bursátil, mientras que los modelos TGARCH lo hacen a nivel desagregado. Asimismo, Gómez *et al.* (2011) realizan una comparación entre un modelo GARCH y una RN multicapa para el pronóstico del IPC, y muestran que la RN logra captar de manera más adecuada el comportamiento de la serie de tiempo, pero el modelo GARCH presenta un mejor ajuste dentro y fuera de la muestra. El autor concluye que el método de redes utilizado no es el más adecuado y propone que para trabajos futuros se utilicen métodos de optimización evolutiva, lo cual permitiría encontrar una arquitectura de red óptima para cada tipo de serie.

El objetivo de esta investigación es desarrollar una herramienta que, mediante el análisis cualitativo y cuantitativo, proporcione un mejor pronóstico del Índice de Precios y Cotizaciones (IPC). La pregunta de investigación sería entonces ¿Qué ventajas presenta una RN frente a un modelo GARCH, y bajo qué condiciones presenta las mismas? La hipótesis (aunque los dos



modelos son muy diferentes y cada uno tiene sus fundamentos) es que los pronósticos del IPC, hechos a partir del modelo de redes neuronales son más eficientes frente a los obtenidos por los modelos GARCH, en el sentido de que son más precisos y con menor costo, pero en el corto plazo. Mientras que los modelos GARCH son menos precisos, pero estables en el largo plazo, además de que sus parámetros tienen una interpretación económica que proporciona más información acerca de la serie de tiempo.

La presente investigación está dividida en tres secciones. En la sección 1 se presentan las generalidades de los mercados bursátiles en México y del IPC. En la sección 2 se presentan las estructuras de cada modelo. Y en la sección 3 se presenta la implementación de cada uno de los modelos (RN y GARCH), para luego hacer los pronósticos y la comparación con los datos reales. Por último, en la sección 4 se exponen las conclusiones.

2. Mercado bursátil en México y pronósticos

En esta sección se presenta brevemente al mercado bursátil en México y su principal indicador, el IPC, así como sus características más relevantes. Posteriormente se revisan los conceptos de series de tiempo y pronóstico. Primero se presenta el modelo GARCH y luego los fundamentos económicos para la utilización del modelo de RN.

El desempeño del mercado financiero en la economía es de esencial importancia debido a que proporciona los canales para financiar la inversión en nuevo capital que permite el crecimiento económico (Parkin, 2014). La importancia del estudio de los mercados bursátiles radica en que su evolución refleja en cierta medida las expectativas del desempeño económico (Guzmán, 2007) ya que se ha observado que el ciclo bursátil y el ciclo económico se relacionan directamente. Aún más, hay trabajos que sugieren que los niveles de precio en las bolsas se anticipan a los de la economía (Gómez *et al.*, 2011). Es decir, el mercado bursátil se podría ver como un condensado de información relevante para la última.

2.1 Generalidades del IPC

La Bolsa Mexicana de Valores (BMV) es una entidad financiera que ofrece servicios para la operación y venta de acciones en México con el objetivo de facilitar las transacciones y fomentar el desarrollo, expansión y competitividad del mercado. “Las empresas que requieren recursos para financiar su operación o proyectos de expansión, pueden obtenerlo a través del mercado bursátil, mediante la emisión de valores que son puestos a disposición de los inversionistas e intercambiados (comprados y vendidos) por las casas de bolsa o socios de la Bolsa Mexicana” (BMV, 2017).

El índice de precios y cotizaciones (IPC) es un indicador diseñado para medir el comportamiento de las 35 emisoras más importantes listadas en la BMV. El IPC busca medir el rendimiento de las acciones de mayor valor de capitalización, rotación y liquidez listadas en la Bolsa Mexicana de Valores (BMV). El IPC es representativo del mercado, fácilmente replicable y refleja la evolución del mercado accionario tomando como referencia la variación de precios de una



muestra de los títulos que se negocian. El mismo es calculado diariamente por la empresa Standard & Poors Dow Jones (S&P DJI)¹.

Recientemente, el cálculo del IPC tuvo diversos cambios, debido a acuerdos desde mayo del 2015 entre la BMV y S&P DJI para la operación y licenciamiento de sus índices. Es decir, ahora ambas empresas desarrollan o hacen mejoras en los índices para satisfacer las necesidades cambiantes del mercado y así tratar de aumentar la capacidad de cobertura global de la BMV.² A partir del 5 de junio de 2017, el IPC cambio de nombre a S&P/BMV IPC y a partir del 18 de septiembre de 2017 se aplicó una nueva metodología (Más fondos, 2017). Con estos acuerdos con S&P, se tiene la ventaja de una mayor difusión en los mercados internacionales, así como también hacerlos más confiables al adoptar estándares operativos internacionales con un nuevo comité de índices integrado por directivos de las dos instituciones, haciendo al mercado mexicano más atractivo para la inversión extranjera.

2.2 El concepto de pronóstico en el estudio económico

El concepto de pronóstico según Gómez *et al.* (2001) identifica los siguientes elementos que determinan un pronóstico en la Tabla 1:

Tabla 1 Características del pronóstico

Elemento	Interpretación	Clases
Objeto	Lo que se quiere pronosticar.	-Resultado de un evento (el evento ocurrirá, pero se desconoce su resultado). -Tiempo del evento (el evento ocurrirá y se conoce el resultado, pero no cuándo). -Serie temporal (proyección a futuro).
Tipo	Cómo va a ser el pronóstico del objeto determinado.	-Puntual (un solo número). -Intervalo (un intervalo de posibles valores dentro del cual estaría el valor real). -Densidad (la distribución de probabilidad del valor futuro en cuestión).
Exactitud	Nivel de certeza del pronóstico en base a la medición de sus errores.	-Error promedio (sesgo). -Varianza del error (dispersión). -Error cuadrático promedio (exactitud general).
Comparación		Función de verosimilitud

Fuente: Elaboración propia en base en Gómez (2011; p. 11).

De acuerdo con Rosen (1990), la predictibilidad es uno de los componentes fundamentales de la ciencia junto con la reproducibilidad. “En física, la previsibilidad significa que, entre los

¹ El cómo se calcula el IPC se puede consultar en Venegas-Martínez (2008).

² S&P Dow Jones Indexes es el proveedor más importante a nivel mundial de datos, investigación y conceptos acerca de índices, teniendo entre ellos el S&P500 y el Dow Jones Industrial Average (S&P Dow Jones Index, 2017b). Es importante también mencionar que las emisoras son ponderadas en función del Valor de Capitalización de Mercado ajustado al Capital Flotante, sujeto a los requisitos establecidos en la sección Criterios de Elegibilidad y Construcción del Índice de acuerdo con S&P DJI (2017b).



fenómenos investigados se puede encontrar el orden y, a partir de ese orden, se pueden formular leyes que predicen los resultados de nuevos experimentos. Entonces se pueden desarrollar teorías para explicar las leyes. Asimismo, es necesario entender cuáles variables dirigen o predicen otras variables cuando muchas de estas interactúan (McNelis, 2005; 2).

2.4 Características de las series de tiempo para el uso de modelos GARCH

Se denota la serie de tiempo de una variable como y_t . La variable y_t es observada para $t = 1, 2, \dots, n$. Mientras que se supone que las condiciones iniciales o los valores previos $y_0, y_{-1}, \dots, y_{1-p}$ estarán disponibles cuando sea necesario. Denotamos por Ω_{t-1} la historia o el conjunto de información al tiempo $t - 1$, la cual contiene toda la información disponible que puede ser empleada para el pronóstico de valores futuros $y_t, y_{t+1}, y_{t+2}, \dots$

Para propósitos de predicción, un modelo satisfactorio es uno simple que describa el comportamiento de una variable (o conjunto de variables) en términos de valores pasados, sin la ventaja de una teoría bien desarrollada. Se ha observado que los grandes y complejos modelos macroeconómicos de ecuaciones simultáneas de la década de los sesentas frecuentemente daban peores resultados predictivos que los modelos de series temporales univariantes (Greene 1998).

El trabajo de Box y Jenkins (1971), la metodología BJ o ARIMA³, fue precursor de los nuevos métodos de pronóstico (incluido el modelo GARCH que se puede ver como una evolución del modelo ARIMA), los cuales se centran “en el análisis de las propiedades probabilísticas de las series de tiempo” (Gujarati y Porter, 2010). Los modelos ARIMA son llamados modelos no teóricos ya que no se derivan de algún programa de investigación o paradigma en la Economía. Por lo que los fundamentos de la utilización del modelo GARCH como modelo de pronóstico son más bien empíricos y se basan en las características propias de las series de tiempo, tal como se verá a continuación. Por ejemplo, Franses y Dijk (2003) muestran evidencia de que la volatilidad de los activos financieros no es constante, sino más bien los periodos relativamente volátiles se alternan con periodos más tranquilos, por lo que se descartan modelos de varianza constante.

Se podría esperar que mucha de la información en los rendimientos está incluida en los más recientes, sin necesidad de incluir otras variables. Otra razón para restringir el análisis a modelos de una sola variable es la concentración en pronósticos de corto plazo. Además, los datos de las series de tiempo financieras muestran frecuentemente un comportamiento asimétrico. Las series temporales financieras, usualmente, no son lineales y tienen las siguientes características:

1. Grandes rendimientos (en términos absolutos) ocurren con más frecuencia de lo que se espera bajo el supuesto de que los datos se distribuyen normalmente.
2. Los rendimientos absolutos grandes tienden a aparecer en grupos, cúmulos de volatilidad (lo que indica la posible presencia de volatilidad variable en el tiempo).
3. Los rendimientos negativos grandes en valor absoluto aparecen más a menudo que los grandes positivos en los mercados bursátiles. Mientras que puede ser al revés para los tipos de cambio.

³ Para más información consulte (Gómez, 2001; Cap.3)



4. Los períodos volátiles suelen estar precedidos por grandes rendimientos negativos.

Uno de los hechos estilizados más destacados de los rendimientos de los activos financieros es que su volatilidad cambia con el tiempo. En particular, los períodos de grandes movimientos de precios se alternan con períodos durante los cuales los precios apenas cambian. Esta característica comúnmente se conoce como agrupamiento de volatilidad.

3. Estructura de los modelos de pronóstico

En la presente sección se muestra la estructura matemática y funcionamiento de cada uno de los modelos a utilizar. En primer lugar, se revisan los modelos de la familia de volatilidad auto regresiva condicionada con heteroscedasticidad (GARCH por sus siglas en inglés), los supuestos y características que los diferencian y que al mismo tiempo los hacen propicios para la predicción, dependiendo de las características de las series de tiempo. En segundo lugar, lo referente al modelo de redes neuronales desde su concepción como analogía biológica para ser un modelo de pronóstico.

3.1. Modelos GARCH

Uno de los motivos principales para el uso de los modelos de la familia GARCH es que las series financieras tienen colas pesadas en sus distribuciones, lo cual se debe a que existen saltos, cambios inesperados y significativos en el mercado, y una alta correlación que presenta la volatilidad de los datos con volatilidades pasadas (Argáez *et al.*, 2014).

3.1.1 Modelo ARCH

A diferencia de los modelos de regresión lineal, donde la principal suposición es que la varianza del cuadrado del error siempre es la misma, la mayoría de los datos de series de tiempo muestran heteroscedasticidad, esto es, las varianzas del error no son iguales. Entonces surge el problema de construir un modelo que cumpla con esta característica donde se obtengan estimaciones y modelos válidos de coeficientes para la varianza del término de error, el llamado modelo autorregresivo generalizado de heteroscedasticidad (ARCH), el primer desarrollo de los modelos de la familia GARCH. Fabozzi (2007; pp. 279-284) presenta el modelo ARCH de la siguiente manera. Una gran cantidad de fenómenos financieros observados confirman la volatilidad agrupada que se refiere a la tendencia de que a grandes alteraciones en los precios de los activos le siguen grandes cambios y a pequeñas alteraciones le siguen pequeños cambios. Es decir, la varianza en los datos financieros no es constante como lo proponen los modelos de regresión lineal, sino que tienen un comportamiento condicional, además de que existe una dependencia temporal y que ni siquiera están cerca de ser independientes e idénticamente distribuidos. Los modelos de la familia ARCH proporcionan un marco de referencia para el estudio de la volatilidad agrupada, modelan el comportamiento subyacente variable en el tiempo, pudiendo con esto, desarrollar modelos de pronóstico. Estos modelos fueron introducidos por Engle en 1982 y Bollerslev en 1987 con la siguiente forma funcional para p rezagos:⁴

⁴De acuerdo con Greene (1999). Considere un modelo de la forma

$$\sigma_t^2 = \varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_p \varepsilon_{t-p}^2, \varepsilon_t = y_t - \sum_{k=1}^K \beta_k \cdot x_{k,t}, \varepsilon_t \sim N(0, \sigma_t^2) \quad (1)$$

El modelo ARCH proporciona un primer y útil marco de referencia para modelar la volatilidad ya que captura parcialmente el fenómeno de que a grandes retornos le seguirán grandes retornos.

3.1.2 Modelo GARCH

El modelo GARCH(p, q) o modelo autorregresivo generalizado de heterocedasticidad (GARCH) condicional se define como sigue:

$$\sigma_t^2 = \alpha_0 + \sum_{p=1}^P \alpha_p \varepsilon_{t-p}^2 + \sum_{q=1}^Q \delta_q \sigma_{t-q}^2 \quad (2)$$

Obsérvese que la varianza condicional se define mediante un proceso ARMA(p, q) en las innovaciones ε_t^2 . La diferencia aquí es que la media de la variable aleatoria de interés y_t está descrita completamente por un modelo de regresión heterocedástico pero ordinario.

La varianza condicional sin embargo evoluciona en el tiempo dependiendo del valor de los parámetros p y q en un modelo GARCH (p, q) donde p se refiere al orden de la parte autorregresiva. Bollerslev demostró que la virtud de este método es que un modelo GARCH con un número pequeño de términos se comporta igual o mejor que un modelo ARCH con muchos términos. Es conveniente expresar a la varianza condicional en términos del operador de rezagos:

$$\sigma_t^2 = \alpha_0 + D(L) \sigma_t^2 + A(L) \varepsilon_t^2 \quad (3)$$

La condición de estacionariedad es que las raíces de la ecuación característica $1 - D(z) = 0$ deben estar fuera del círculo unitario, pero también se puede poner que $A(1) + D(1) < 1$ para efectos más prácticos. Considere ahora el modelo GARCH en media, la suposición principal es que la varianza del término de distribución aleatoria afecta directamente a la media de la variable dependiente y evoluciona a través del tiempo como una función de su propio valor pasado y el error de predicción pasado al cuadrado, siguiendo a McNelis (2005) se tiene que:

$$\sigma_t^2 = \delta_0 + \delta_1 \sigma_{t-1}^2 + \delta_2 \varepsilon_{t-1}^2, \varepsilon_t \sim N(0, \sigma_t^2), y_t = \alpha + \beta \sigma_t + \varepsilon_t \quad (4)$$

Donde y_t es la variable objetivo, α la tasa de apreciación y ε_t el termino de perturbación distribuido normal con media 0 y varianza σ_t^2 , β es el efecto sobre el rendimiento del activo mientras que δ_i definen la evolución de la varianza condicional.

El atractivo del modelo GARCH es que toma en cuenta la no linealidad en el proceso con la varianza condicional como una transformación no lineal de los valores y errores pasados. Además de que su afectación a la variable dependiente se justifica como un factor de riesgo que aumenta el rendimiento. Como en el caso ARCH, el modelo GARCH con distribución normal condicional de retornos implica distribuciones incondicionales que tienen colas más pesadas que la distribución normal, así el modelo GARGH permite modelar volatilidad acumulada y colas pesadas o incondicionales (Argáez *et al.*, 2014). La relación entre la curtosis de ε_t , la volatilidad

acumulada y la no normalidad condicional (innovaciones no gaussianas) es explicada a través de la volatilidad acumulada y la no normalidad condicional que pueden inducir leptocurtosis, típicamente observada en las series financieras. La desventaja del modelo es que la minimización de las funciones log-verosimilitud es a menudo difícil de alcanzar.

3.1.3 Modelo TGARCH

Argáez *et al.* (2014) explica de una manera sencilla el modelo GARCH por umbrales (TGARCH). Cuando se analiza una serie de tiempo se puede observar que las malas noticias tienen mayor repercusión que las buenas noticias. Más concretamente, algunos mercados financieros presentan el efecto de apalancamiento, es decir, la volatilidad decrece cuando los rendimientos se incrementan. De esta manera, el modelo GARCH tiene la desventaja de que la varianza condicional no puede modelar esta relación asimétrica. Entonces tenemos que la varianza condicional se puede modelar de la siguiente manera:

$$\sigma_t^2 = \omega + \sum_{i=1}^q (\alpha_i + \gamma_i d_{t-i}) \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (5)$$

con α_i , γ_i y β_j son parámetros no negativos que satisfacen las condiciones de los modelos GARCH y d_{t-i} es la variable indicadora o de umbral. En el modelo TGARCH(1,1) se tiene

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 d_{t-1} \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (6)$$

donde $d_{t-1} = 1$ si $\varepsilon_{t-1} < 0$ (malas noticias) y $d_{t-1} = 0$ si $\varepsilon_{t-1} \geq 0$ (buenas noticias). Además d_{t-1} usa un umbral igual a 0 para clasificar los impactos pasados, y γ_i es un parámetro de apalancamiento, cuando $\gamma_i=0$ tenemos el modelo GARCH(1,1).

3.2. Redes neuronales

Los métodos de RN o Redes Neuronales Artificiales (RNA) vienen de la ciencia del cerebro, de la teoría cognitiva y neurofisiología, ofreciendo una alternativa para los modelos de pronóstico, clasificación y evaluación de riesgos (McNelis, 2005)

McNelis (2005) expresa que cuando se realizan pronósticos de los mercados financieros, se hace sobre las expectativas o pronósticos de otros, es decir hay un proceso de aprendizaje continuo que va adaptando creencias o errores del pasado. Un punto importante que hay que aclarar, es que la RN a diferencia de otros modelos no lineales, es no paramétrico. No hay alguna hipótesis en específico acerca de los valores de los parámetros a ser estimados, ni de la forma funcional del modelo, es más, la mayoría de las veces no se pueden interpretar los coeficientes estimados en la RN como en los modelos econométricos.

El sector financiero de los mercados emergentes es un campo fértil de aplicación de las RN, además de los mercados con gran innovación y cambio, por dos razones: la primera es que los datos que se obtienen son ruidosos debido a la velocidad con que las noticias son difundidas y por lo tanto existen asimetrías y no linealidades que no se pueden dejar pasar; la otra razón es que los participantes de estos mercados están en un proceso de aprendizaje de prueba y error.



3.1.1. Analogía biológica

Medsker *et al.* (1996) explican que el cerebro humano está compuesto por células llamadas neuronas que conservan información, éstas a su vez trabajan en grupos llamados redes, cada grupo contiene un gran número de neuronas interconectadas. Las neuronas tienen cada una un núcleo (órgano de cómputo) donde se procesa la información que llega de las dendritas (canal de recepción de información, proporcionan la información de entrada), luego la neurona da una señal de salida mediante el axón (canal de salida) que está conectado a su vez con otras dendritas de otra neurona. Las señales pueden ser transmitidas sin cambio o pueden ser transmitidas mediante sinapsis; una sinapsis es capaz de incrementar o disminuir la fuerza de conexión causando excitación o inhibición a la neurona posterior. Con base en los conceptos anteriores se desarrolla una RNA, un modelo que simula una red neuronal biológica, de procesamiento paralelo masivo, con elementos de proceso llamadas neuronas, interconectadas entre sí en una arquitectura de red a diferencia, por ejemplo, de una computadora convencional. Esta última tiene una arquitectura Von Neumann, con un elemento de proceso que ejecuta instrucciones complejas en serie. La neurona artificial es análoga a la neurona biológica, ésta recibe una señal de entrada (como los impulsos electromecánicos que reciben las dendritas de otras neuronas), la procesa y da una señal de salida (como la señal de salida que manda la neurona biológica sobre su axón). Dichas señales artificiales de salida pueden variar similarmente al cambio que ocurre en la sinapsis. (Medsker *et al.*, 1996). Florez y Fernández (2008) explican lo anterior de manera concreta, la RNA trata de reproducir el funcionamiento del cerebro humano caracterizado por el aprendizaje a través de la experiencia (concepto fundamental de las RNA) y la extracción de conocimiento genérico a partir de un conjunto de datos. Para una RNA, una neurona artificial consiste en un nodo Y con sus enlaces asociados. Luego y_i son las señales de entrada y w_i sus pesos asociados. Las señales de entrada llegan mediante los enlaces, así el valor del nodo Y es la suma de todas las señales de entradas ponderadas por sus respectivos pesos, es decir $Y = \sum_i y_i w_i$.

3.2.2 Estructura general de las RNA

Hawley (1996) menciona que un sistema neuronal artificial, consiste de 3 componentes:

1. **Topología de la red:** Es el conjunto de nodos conectados mediante enlaces y la disposición de las conexiones. Si un nodo está conectado directamente a otro se dice que son vecinos, si está conectado a muchos otros, se dice que está conectado en paralelo. Además, una de las principales características, es que el flujo de información no es unidireccional.
2. **Método de activación de propagación:** El valor del nodo Y se compara con el nivel de activación de umbral, si lo cumple el nodo da una señal de salida. Este nivel de activación depende de los valores previos de activación del nodo, así como de los niveles de activación de sus vecinos. El nivel de activación es comunicado a los vecinos a cada instante, que lo usan para actualizar sus propios niveles de activación. Si el nivel de un nodo tiende a incrementar el nivel de un nodo vecino se designa como un enlace excitador y se le asigna un peso positivo, si pasa lo contrario, entonces es un enlace inhibitorio y se le asigna un peso negativo.

3. **Mecanismo de entrenamiento:** Puede ser pensado como un proceso de recompensas y castigos. Cuando el sistema responde correctamente a una entrada la recompensa consiste en el fortalecimiento de la matriz actual de pesos nodales. Si el sistema responde incorrectamente, el castigo sería un reajuste en la matriz de pesos nodales, así el sistema respondería de una manera diferente cuando en el futuro se encuentre con entradas similares. Un sistema neuronal artificial puede ser entrenado mediante una metodología supervisada o no supervisada, como se verá más adelante. McNelis (2005) lo explica de la siguiente forma: una RNA relaciona un conjunto de variables de entrada $\{x_i\} \quad i=1, \dots, k$ con una o un conjunto de variables de salida $\{y_j\} \quad j=1, \dots, k^*$ haciendo uso de una o más capas ocultas, en las cuales las variables de entrada (su suma ponderada) son aplastadas o transformadas por una función especial, una transformación logística o logsigmoideal. Todos estos elementos están unidos entre sí mediante conectores llamados sinapsis.

El uso de las RNA para modelar el proceso de decisión está basado en dos principios que Rustichini (*et al.*, 2002) define de la siguiente manera:

1. **Principio de segregación funcional:** “No todas las funciones en el cerebro son realizadas por el cerebro como un todo”
2. **Principio de integración funcional:** “Redes de diferentes regiones son activadas por diferentes funciones, superponiéndose a las regiones usadas en redes diferentes”

Esto es, el cerebro es un modelo óptimo, en cuanto a recursos se refiere, ya que, al hacer un procesamiento, no todo el cerebro se pondrá a funcionar, sólo aquellas regiones específicas, que de alguna manera reconozcan patrones. A continuación, se describen las ecuaciones que definen al perceptrón multicapa. Sean $x_i, i=1, \dots, i^*$ variables de entrada, $n_k, k=1, \dots, k^*$ neuronas y y_t la variable de salida:

$$n_{k,t} = w_{k,0} + \sum_{i=1}^{i^*} w_{k,i} x_{i,t} \quad (7)$$

$$N_{k,t} = L(n_{k,t}) = \frac{1}{1+e^{-n_{k,t}}} \quad (8)$$

$$y_t = \gamma_0 + \sum_{k=1}^{k^*} \gamma_k N_{k,t} \quad (9)$$

donde (7) representa a la neurona k al tiempo t y es igual a la combinación lineal de las entradas $x_{i,t}$ ponderadas por su peso $w_{k,i}$ mas una constante $w_{k,0}$, (8) es la transformación de la neurona k al tiempo t bajo la función sigmoideal y (9) es la variable de salida (pronóstico) al tiempo t , dado como una combinación lineal de $N_{k,t}$ por una coeficiente γ_k , más una constante γ_0 . Existen diversas funciones de transformación, entre las cuales destaca la función tangente hiperbólica:

$$N_{k,t} = T(n_{k,t}) = \frac{e^{n_{k,t}} - e^{-n_{k,t}}}{e^{n_{k,t}} + e^{-n_{k,t}}} \quad (10)$$

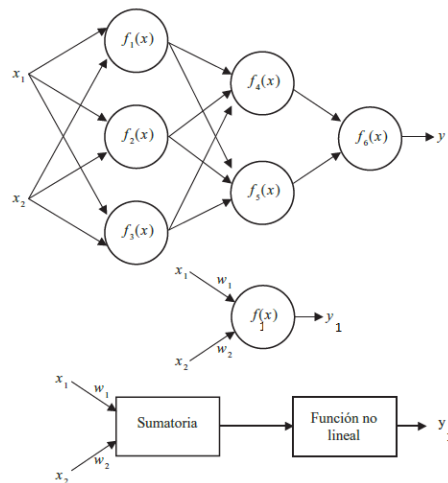
También se cuenta con la función gaussiana:

$$N_{k,t} = \Phi (n_{k,t}) = \int_{-\infty}^{n_{k,t}} \sqrt{\frac{1}{2\pi}} e^{-(1/2)n_{k,t}^2} \quad (11)$$

3.2.3 Aprendizaje y algoritmo backpropagation

Para desarrollar el algoritmo de la RNA que se utilizará para el pronóstico se debe tener en cuenta lo siguiente. El proceso de aprendizaje o entrenamiento consta de tres etapas según Medsker *et al.*, (1996): 1) Calcular las salidas, 2) comparar las salidas con las respuestas deseadas y 3) Ajustar los pesos y repetir el proceso. El proceso de aprendizaje comienza poniendo pesos aleatorios, se requiere minimizar la diferencia ($\Delta = \hat{y} - y$) entre la salida y la salida esperada, esto se logra al cambiar incrementalmente los pesos. El procesamiento de información con una RNA consiste en analizar patrones de actividades con información aprendida almacenada en los pesos de las conexiones. Diferentes RNA calculan el error en diferentes formas dependiendo del algoritmo de aprendizaje que se use. El entrenamiento de una red está dividido en dos categorías, casos de prueba y casos de entrenamiento.

Figura 1. Ejemplo de componentes y salida de una neurona.



Fuente: (Ponce, 2010, Pag. 225).

En la Figura 1 se aprecia que la primera neurona está compuesta de dos unidades, donde la primera suma los productos de las entradas por sus respectivos pesos, y la segunda unidad contiene la función de activación, su salida corresponde a la salida de la neurona 1, es decir:

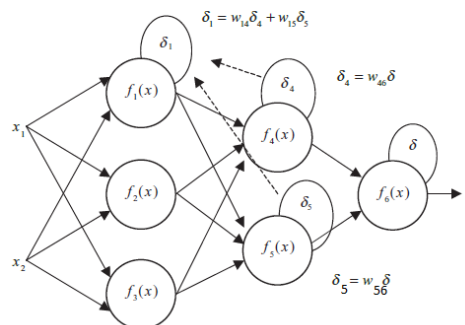
$$y_1 = f_1 (w_{x_1,1}x_1 + w_{x_2,1}x_2) \quad (12)$$

De acuerdo con Valencia *et al.* (2006), “El funcionamiento de la red backpropagation consiste en el aprendizaje de un conjunto predefinido de pares de entradas-salidas dados como ejemplo: primero se aplica un patrón de entrada como estímulo para la primera capa de las neuronas de la red, se va propagando a través de todas las capas superiores hasta generar una salida, se compara el resultado en las neuronas de salida con la salida que se desea obtener y se calcula un valor de error para cada neurona de salida. A continuación, estos errores se transmiten hacia atrás,

partiendo de la capa de salida hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total. Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se presente el mismo patrón, la salida esté más cercana a la deseada”. Ponce (2016, Cap. 3) lo explica de manera explícita. El perceptrón multicapa representa el aprendizaje supervisado donde se emplean funciones de activación elementales de forma binaria. La retropropagación del error es un entrenamiento supervisado que se emplea para redes multicapa, donde se ajusta el valor de los pesos en función del error generado. Esto permite tener un método de optimización que se encuentra al definir el gradiente del error y minimizarlo con respecto a los parámetros de la red neuronal⁵. Considere una red de 3 capas ocultas con 2 entradas y una salida, se observa a continuación, a manera de ejemplo lo que pasa en la primera neurona:

La salida de la red es comparada con el valor objetivo deseado, la diferencia se denomina error de la señal (δ). El algoritmo de retropropagación propaga el error de regreso a todas las neuronas, cuya salida fue la entrada de la última neurona, posteriormente el error se va propagando a las neuronas de capas anteriores considerando los pesos de las conexiones como se muestra en la Figura 2.

Figura 2. Ejemplo del algoritmo de retropropagación para una neurona



Fuente: Elaboración propia con base en (Ponce, 2010; 226).

3.2.4 Tipos de redes

De acuerdo con McNelis (2005) existen diversos tipos de redes, la primera y más general es el perceptrón multicapa visto anteriormente. A continuación, se presentan la red con varias capas ocultas y la red recurrente Elman.

En este caso, la diferencia con la Figura 1, es que después de la función de transformación de la primer capa oculta $N_{k,t}$, se seguiría otra neurona de la siguiente forma:

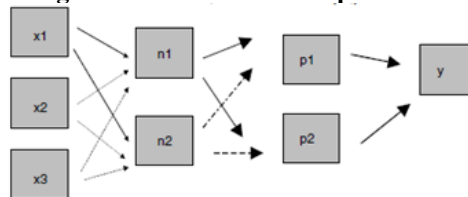
⁵ Para su desarrollo formal consulte (Ponce,2016, 229-231)

$$p_{l,t} = \rho_{l,0} + \sum_{k=1}^{k^*} \rho_{l,k} N_{k,t} \quad (13)$$

Es decir, la neurona l al tiempo t es igual a la combinación lineal de las salidas de la primera capa oculta ponderadas por sus pesos $\rho_{l,k}$ más una constante $\rho_{l,0}$. Después esta neurona entrara a la función de transformación:

$$P_{l,t} = \frac{1}{1+e^{-p_{l,t}}} \quad (14)$$

Figura 2. Red con varias capas ocultas.



Fuente: (McNelis,2005).

Lo anterior da como salida o pronóstico:

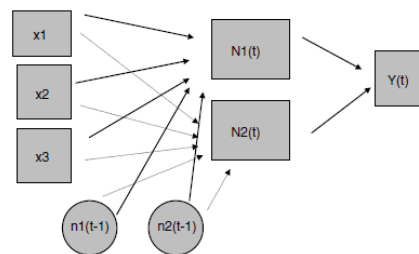
$$y_t = \gamma_0 + \sum_{k=1}^{k^*} \gamma_k P_{l,t} \quad (15)$$

Este proceso es iterativo dependiendo del número de capas ocultas, lo cual incrementa en gran medida el número de parámetros a ser estimados.

Por otro lado, la red recurrente Elman permite a las neuronas depender, no solamente de las variables de entrada sino también de sus valores anteriores; véase la Figura 4.

$$y_t = \beta_0 + \sum_{i=1}^{i^*} \beta_i x_{i,t} + \varepsilon_t + \sum_{j=1}^q v_j \widehat{\varepsilon}_{t-j}, \quad \widehat{\varepsilon}_{t-j} = y_{t-j} - \widehat{y}_{t-j} \quad (16)$$

Figura 4 Red recurrente Elman.



Fuente: (McNelis, 2005).

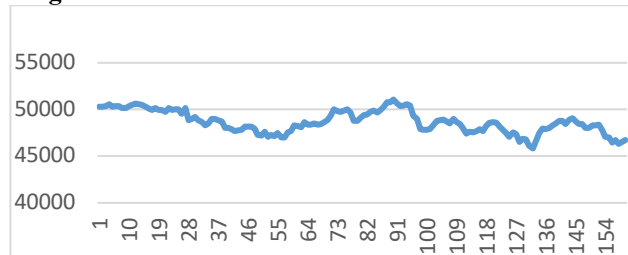
4. Implementación de los modelos y obtención de los pronósticos

En esta última sección se considera el S&P/BMV IPC para una muestra de 160 datos en un periodo de poca volatilidad (18 de septiembre de 2017 a 11 de mayo de 2018), donde 150 datos se usaron para ajustar los modelos y hacer el pronóstico de los últimos 10 datos. Para la segunda parte se tomó el IPC para una muestra de 2010 datos en un periodo de alta volatilidad (3 de enero de 2005 a 21 de diciembre de 2012), donde 2000 datos se usaron para ajustar el modelo y hacer el pronóstico de los últimos 10 datos. La fuente es Yahoo Finance con los datos IPC^{MXX} al cierre. Se presentan los diferentes y mejores modelos estimados (respecto a las características de los datos), por un lado, el modelo GARCH (apéndice A) y por otro lado el modelo de redes neuronales. Usando para el modelo GARCH el software Eviews9 y para el modelo de redes neuronales el software Matlab 2017 usando la herramienta nntool.

4.1 Modelo GARCH

El primer modelo que se presenta es para la muestra de 160 datos del S&P/BMV IPC al cierre del 18 de septiembre de 2017 al 11 de mayo de 2018, periodo en el que se aplicó la nueva metodología del S&P/BMV (Figura 5).

Figura 5. S&P/BMV IPC al cierre 18-09-17 a 11-05-18



Fuente: Elaboración propia con datos de Yahoo Finance.

Se trabajó con los primeros 150 datos (05/09/18 a 26/04/18) para intentar hacer el pronóstico de los últimos 10 datos (27/04/18 a 11/05/18). Para comenzar con el análisis, se hicieron pruebas para ver si la serie es estacionaria:

1) Prueba Dickey-Fuller aumentada con tendencia e intercepto.

Tabla 2. Prueba Dickey-Fuller de la primera serie de datos

Hipótesis nula: CLOSE tiene una raíz unitaria		
	Estadístico <i>t</i>	Prob.
Estadístico de la prueba Dickey-Fuller Aumentada	-2.1721	0.5012
Valores críticos de la prueba	nivel 1%	-4.0208
	nivel 5%	-3.4402
	nivel 10%	-3.1445

Fuente: Elaboración propia.

En la Tabla 2 podemos observar que la probabilidad cae en zona de no rechazo. Por lo tanto, no se rechaza la hipótesis nula de que la serie sea no estacionaria.

2) Prueba Phillips-Perron con tendencia e intercepto.

Tabla 3. Prueba Phillips-Perron de la primera serie de datos

Hipótesis nula: CLOSE tiene una raíz unitaria			
		Estadístico Ajustado	<i>t</i> Prob.
Estadístico de la prueba Phillips-Perron		-2.3708	0.3933
Valores críticos de la prueba	nivel 1%	-4.0208	
	nivel 5%	-3.4402	
	nivel 10%	-3.1445	

Fuente: Elaboración propia.

Se puede observar, en la Tabla 3, nuevamente que cae en zona de no rechazo. Por lo tanto, no se rechaza la hipótesis nula de que la serie sea no estacionaria.

3) Prueba KPSS con tendencia e intercepto.

Tabla 4. Prueba KPSS de la primera serie de datos

Hipótesis nula: CLOSE es estacionaria		
		Estadístico LM
Estadístico de la prueba KPSS		0.1179
Valores críticos asintóticos	nivel 1%	0.216
	nivel 5%	0.146
	nivel 10%	0.119

Fuente: Elaboración propia.

Podemos observar en el Cuadro 4 que como el estadístico LM es menor que el valor crítico al 5% no se puede rechazar la hipótesis nula de que la serie es estacionaria. Debido a que esta prueba es mejor que las anteriores, se aceptará que la serie es estacionaria y se trabajará con la serie a nivel en lugar de con los rendimientos o los rendimientos logarítmicos de la serie.

De acuerdo al correlograma y pruebas a residuos, se obtiene, por el momento, el mejor modelo ARMA para la serie y queda de la siguiente manera en la Tabla 5.

Se puede observar en la Tabla 5 que cada parámetro es estadísticamente significativo para la ecuación principal y para la ecuación de la varianza. Vemos que el parámetro de R^2 indica un muy buen grado de ajuste de más del 90%, el parámetro Prob(F -statistic) nos indica que los parámetros en conjunto son estadísticamente significativos.

Tabla 5. Modelo ARMA de la primera serie de datos

Variable	Coficiente	Error est.	Est <i>t</i>	Prob.
c	48853.61	413.1679		0.0000
AR(2)	0.874633	0.041926	20.8611	0.0000
MA(1)	0.983883	0.030154		0.0000
SIGMACuad	119127.0	12219.34		0.0000
			9.74905	
<i>R</i> -cuadrada	0.907705	Media var dep		48774.23
<i>R</i> cuadrada ajustada	0.905809	D.E. var dep		1139.905
E.S. de regresión	349.8440	Crit. info Akaike		14.59639
Sum cuadrada resid	17869056	Crit. de Schwarz		14.67668
Log verosimilitud	-1090.730	Crit. de Hannan-Quinn		14.6290
Estad <i>F</i>	478.6287	Estad Durbin-Watson		1.753116
Prob(Estad <i>F</i>)	0.000000			
Raíces AR invert				
	.94	-.94		
Raíces MA invert				
	-.98			

Variable dependiente: CLOSE; Método: ARMA Máxima Verosimilitud.
Muestra: 1 150; Observaciones incluidas: 150;
Convergencia lograda después de 9 iteraciones.
Coeficiente de covarianza calculado utilizando el producto externo de los gradientes.
Fuente: Elaboración propia.

Tabla 6. Correlograma de los residuos de la primera serie de datos.

		1	0.114	0.114	2.0018	
		2	0.074	0.062	2.8472	
		3	0.066	0.052	3.5201	0.061
		4	-0.047	-0.066	3.8705	0.144
		5	0.041	0.047	4.1364	0.247
		6	-0.108	-0.117	5.9989	0.199
		7	-0.047	-0.021	6.3519	0.273
		8	-0.017	-0.005	6.3990	0.380
		9	-0.081	-0.057	7.4581	0.383
		10	-0.053	-0.048	7.9076	0.443
		11	0.025	0.055	8.0128	0.533
		12	-0.080	-0.089	9.0723	0.525
		13	0.079	0.091	10.110	0.521
		14	0.107	0.096	12.030	0.443
		15	0.019	-0.010	12.093	0.520
		16	0.083	0.036	13.260	0.506
		17	0.042	0.044	13.565	0.559
		18	0.095	0.061	15.108	0.517
		19	-0.024	-0.059	15.209	0.580
		20	0.007	0.041	15.217	0.647
		21	0.054	0.044	15.726	0.675

Fuente: Elaboración propia.

Aunque se podrían mejorar aún más los parámetros del criterio de Schwarz, el criterio de Hannah-Quinn y el estadístico Durbin Watson. A continuación, vemos en el correlograma de los residuos que se controló la función de autocorrelación y de autocorrelación parcial (Tabla 6).

A continuación, se muestra el correlograma de los residuos al cuadrado en la Tabla 7. Posteriormente se presenta en el Tabla 8, la prueba de independencia BDS a los residuos.

Cuadro 7. Correlograma de los residuos al cuadrado de la primera serie de datos

		1	0.009	0.009	0.0125	0.911
		2	0.097	0.097	1.4615	0.482
		3	-0.008	-0.009	1.4707	0.689
		4	-0.050	-0.060	1.8672	0.760
		5	-0.038	-0.036	2.0954	0.836
		6	0.041	0.054	2.3655	0.883
		7	-0.003	0.003	2.3670	0.937
		8	-0.043	-0.057	2.6662	0.954
		9	-0.070	-0.074	3.4515	0.944
		10	0.030	0.046	3.5938	0.964
		11	-0.002	0.016	3.5947	0.980
		12	-0.073	-0.093	4.4773	0.973
		13	-0.071	-0.087	5.3096	0.968
		14	-0.047	-0.024	5.6735	0.974
		15	-0.023	0.007	5.7602	0.984
		16	-0.042	-0.054	6.0558	0.987
		17	-0.056	-0.086	6.5900	0.988
		18	-0.024	-0.019	6.6858	0.993
		19	-0.125	-0.104	9.4037	0.966
		20	0.104	0.100	11.312	0.938
		21	0.009	-0.002	11.327	0.956

Fuente: Elaboración propia.

Tabla 8. Prueba BDS de los residuos de la primera serie de datos.

Prueba BDS para RESID

Muestra: 1 160

Observaciones incluidas: 160

	Estad. BDS	Err.estan	Estad. z	Prob.
2	0.009532	0.006520	1.462050	0.1437
3	0.013764	0.010407	1.322566	0.1860
4	0.012217	0.012447	0.981528	0.3263
5	0.013616	0.013030	1.045008	0.2960
6	0.010663	0.012620	0.844867	0.3982

Fuente: Elaboración propia.

Debido a los resultados de la Tabla 8, no se puede rechazar que la serie sea iid. Por último, en la Tabla 9, se ve en la prueba de heterocedasticidad para verificar que existan efectos ARCH:

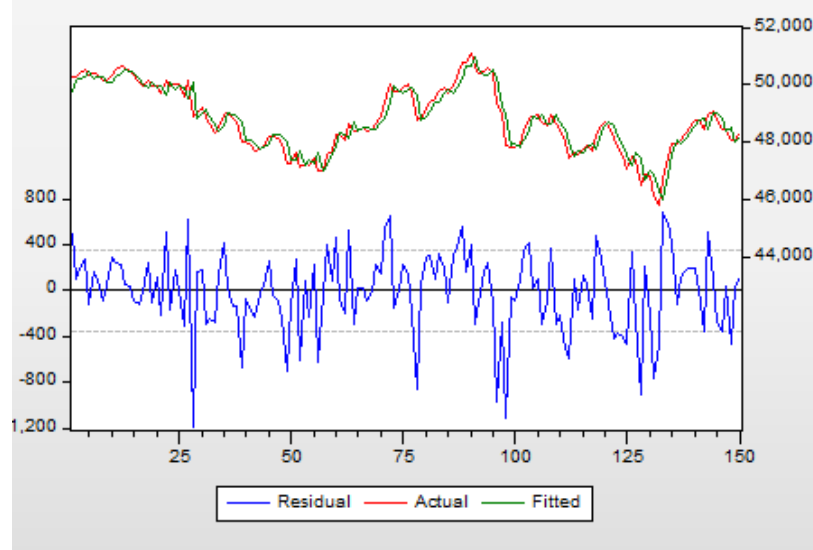
**Tabla 9. Prueba LM ARCH de los residuos de la primera serie de datos
 Prueba de Heterocedasticidad: ARCH**

Estad. F	0.012104	Prob.F(1,147)	0.9125
Obs R-cuadrada	0.012267	Prob.Chi cuadrada(1)	0.9118

Fuente: Elaboración propia.

Por lo cual no se puede rechazar la hipótesis nula de que no hay autocorrelación en los residuales al cuadrado. Con las pruebas anteriores se puede observar que no hay cabida para realizar un modelo GARCH, basta con el modelo ARMA, dado que no se observan relaciones de dependencia en los residuos. Posteriormente se presenta la gráfica de los datos originales vs el modelo ajustado y otra gráfica con los residuos. De donde se puede observar que el modelo tiene un buen ajuste, pero aún hay picos en los residuos que no se controlan, aunque no sugieren cúmulos de volatilidad (Figura 6).

Figura 6. Datos observados vs ajustados y residuos del modelo ARMA estimado

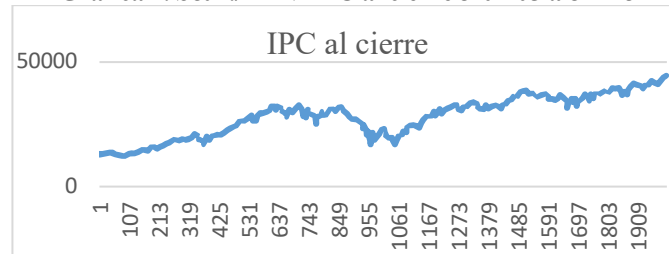


Fuente: Elaboración propia.

El segundo modelo que se presenta es para la muestra de 2010 datos del IPC al cierre del 3 de enero de 2005 al 8 de enero de 2013 (Tabla 7).

Se trabajó con los primeros 2000 datos (3-01-05 a 21-12-12) para intentar hacer el pronóstico de los últimos 10 datos (24-12-12 a 8-1-13). Primero las pruebas de raíz unitaria, la prueba Dickey-Fuller aumentada con tendencia e intercepto

Gráfica 7. S&P./BMV IPC al cierre 3-01-05 a 8-1-13



Fuente: Elaboración propia con datos de Yahoo Finance.

Tabla 10. Prueba Dickey-Fuller de la segunda serie de datos
Hipótesis Nula: CLOSE tiene una raíz unitaria

	Estad. t ajustado	Prob.*
Estad. De la prueba Dickey-Fuller	-2.119948	0.5338
Valores críticos	nivel 1%	-3.962639
	nivel 5%	-3.412058
	nivel 10%	-3.127941

Fuente: Elaboración propia.

En la Tabla 10 se puede ver que no se rechaza la hipótesis nula de que la serie tenga una raíz unitaria y por lo tanto sea no estacionaria.

2) La prueba Phillips-Perron con tendencia e intercepto

Tabla 11. Prueba Phillips-Perron de la segunda serie de datos.

	Estad. t ajustado	Prob.*
Estad. De la prueba Phillips-Perron	-2.076896	0.5579
Valores críticos	nivel 1%	-3.962639
	nivel 5%	-3.412058
	nivel 10%	-3.127941

Fuente: Elaboración propia.

En la Tabla 11 se puede observar que no se rechaza que sea una serie no estacionaria

Siendo esta última prueba la más robusta, podemos observar que se rechaza la hipótesis nula de que la serie es estacionaria (Tabla 12). Después de varios intentos de hacer la serie estacionaria en diferencias, con rendimientos o rendimientos logarítmicos, pero sin conseguir un buen grado de ajuste del modelo se decidió trabajar el modelo GARCH sobre la tendencia del logaritmo del IPC. Después del análisis de los correlogramas y análisis de residuos se obtuvo, hasta ahora, el mejor modelo GARCH. Primero se presenta en la Tabla 13 el modelo ARMA.

3) La prueba KPSS con tendencia e intercepto

Tabla 12. Prueba KPSS de la segunda serie de datos.

		Estad. LM.
Estad. de la prueba KPSS		0.356 163
Valores críticos asintóticos	nivel 1%	0.216000
	nivel 5%	0.146000
	nivel 10%	0.119000

Fuente: Elaboración propia.

**Tabla 13. Modelo ARMA de la segunda serie de datos.
Variable dependiente: LNCLOSE**

Variable	Coefficiente	Err estd	Estad. t	Prob.
C	9.629729	0.143506	67.10337	0.0000
TENDENCIA	0.000543	0.000130	4.169065	0.0000
AR(1)	0.995792	0.001766	563.8934	0.0000
MA(1)	0.087122	0.016106	5.409270	0.0000
SIGMASQ	0.000200	3.38E-06	59.26012	0.0000
R-cuadrada	0.998157	Media var dependiente	10.19490	
R-squared ajust	0.998153	E.S. var dependiente	0.329852	
E.S. de regression	0.014176	Criterio de Akaike	-5.669583	
Sum resid cuad	0.400903	Criterio de Schwarz	-5.655581	
Log verosimilitud	5674.583	Criterio de Hannan-Quinn	-5.664442	
Estad F	270080.8	Estad. Durbin-Watson	2.004151	
Prob(Estad F)	0.000000			

Fuente: Elaboración propia. Método: ARMA Máxima verosimilitud
Muestra: 1 2000, Obs incluidas: 2000, Convergencia lograda después de 6 iteraciones
Coeficiente de covarianza calculado utilizando el producto externo de los gradientes

En la Tabla 14 podemos ver que los parámetros son estadísticamente significativos, además de un buen grado de ajuste y de parámetros de Akaike, Schwarz y Hannan-Quinn. Posteriormente observamos el correlograma de los residuos.

En la Tabla 14 se puede observar que se controló la función de autocorrelación y de autocorrelación parcial. A continuación, se muestra el correlograma de los residuos al cuadrado.

Asimismo, en la Tabla 15 se puede observar que existen estructuras de dependencia. Por lo que hacemos una prueba de independencia BDS a los residuos.

Tabla 14. Correlograma de los residuos después del modelo ARMA de la segunda serie de datos.

		1	-0.002	-0.002	0.0114	
		2	-0.036	-0.036	2.6227	
		3	-0.033	-0.033	4.8269	0.028
		4	-0.019	-0.020	5.5459	0.062
		5	-0.004	-0.007	5.5787	0.134
		6	0.022	0.019	6.5152	0.164
		7	-0.079	-0.081	19.197	0.002
		8	0.014	0.014	19.583	0.003
		9	-0.023	-0.029	20.692	0.004
		10	0.009	0.006	20.861	0.008
		11	0.012	0.009	21.176	0.012
		12	0.056	0.054	27.461	0.002
		13	0.076	0.081	39.079	0.000
		14	-0.010	-0.012	39.287	0.000
		15	-0.022	-0.008	40.255	0.000
		16	0.043	0.046	43.997	0.000
		17	-0.040	-0.036	47.152	0.000
		18	-0.018	-0.017	47.787	0.000
		19	-0.006	-0.002	47.871	0.000
		20	0.027	0.038	49.366	0.000
		21	-0.030	-0.034	51.166	0.000

Fuente: Elaboración propia.

Cuadro 15. Correlograma de los residuos al cuadrado del modelo ARMA de la segunda serie

		1	0.143	0.143	40.923	0.000
		2	0.220	0.203	137.71	0.000
		3	0.142	0.093	178.08	0.000
		4	0.213	0.155	269.21	0.000
		5	0.266	0.205	411.01	0.000
		6	0.230	0.138	517.38	0.000
		7	0.240	0.133	633.34	0.000
		8	0.171	0.050	691.99	0.000
		9	0.228	0.097	796.28	0.000
		10	0.220	0.087	893.82	0.000
		11	0.299	0.162	1073.4	0.000
		12	0.219	0.076	1170.4	0.000
		13	0.157	-0.015	1220.0	0.000
		14	0.151	-0.023	1266.0	0.000
		15	0.139	-0.048	1304.9	0.000
		16	0.200	0.002	1385.8	0.000
		17	0.224	0.056	1487.4	0.000
		18	0.244	0.091	1607.5	0.000
		19	0.116	-0.043	1634.7	0.000
		20	0.102	-0.079	1655.7	0.000
		21	0.153	-0.014	1703.0	0.000

Fuente: Elaboración propia.



**Tabla 16. Prueba BDS de los residuos al después del modelo ARMA de la segunda serie
Prueba BDS a Residuales**

Dimension	Estad. BDS	Err estd	Estad z	Prob.
2	0.015192	0.002094	7.253337	0.0000
3	0.036662	0.003321	11.03771	0.0000
4	0.055026	0.003947	13.93964	0.0000
5	0.067563	0.004107	16.45242	0.0000
6	0.074631	0.003953	18.88010	0.0000

Fuente: Elaboración propia.

Con base en la Tabla 16 se rechaza la hipótesis nula de que los residuos son iid. Ahora probamos si existe heterocedasticidad en los residuos y presentamos en la Tabla 17

**Tabla 17. LM test ARCH de los residuos al después del modelo ARMA de la segunda serie
Prueba de heterocedasticidad: ARCH**

Estad. F	4.165403	Prob.F(1,1997)	0.0000
Obs'R-cuadrada	40.84382	Prob.Chi-Cuadrada(1)	0.0000

Fuente: Elaboración propia.

Con base en la Tabla 17 se rechaza la hipótesis nula de que no hay autocorrelación en los residuales al cuadrado. Con las pruebas anteriores se puede justificar el uso de un modelo GARCH. Quedando de la manera expuesta en la Tabla 18.

En la Tabla 18 se puede ver que cada uno de los parámetros es estadísticamente significativo en la ecuación principal, con parámetros de intercepto, de tendencia, autorregresivos de orden 1, y media móvil de orden 1. Para la ecuación de la varianza también se observa que los parámetros son estadísticamente significativos. Un buen grado de ajuste R^2 y los parámetros de Akaike, Schwarz y Hannan-Quinn. Luego vemos que se controló el correlograma de los residuos al cuadrado en la Tabla 19.

**Tabla 18. Modelo ARMA GARCH de la segunda serie de datos
Variable dependiente: LNCLOSE; Método: ML – ARCH**

Variable	Coefficiente	Err.estd	Estad. z	Prob.
C	9.959879	0.174621	57.03727	0.0000
TENDENCIA	0.000370	0.000110	3.372394	0.0007
AR(1)	0.995825	0.001702	584.9735	0.0000
MA(1)	0.060300	0.025313	2.382231	0.0172
Ecuación z de la varianza				
C	2.72E-06	5.84E-07	4.669009	0.0000
RESID(-1) ²	0.099606	0.010143	9.820046	0.0000
GARCH(-1)	0.887232	0.010407	85.25058	0.0000
R-cuadrada	0.998152	Media var dependiente		10.19526
R-cuadrada ajustada	0.998149	D.S..var dependiente		0.329541
E.S. de regresión	0.014177	Criterio de AKaike		-6.010071
Sum resid cuadrados	0.400978	Criterio de Schwarz		-5.990460
Logverosimilitud	6014.066	Criterio de Hannan-Quinn		-6.002870
Estad. Durbin-Watson	1.953404			

Fuente: Elaboración propia.

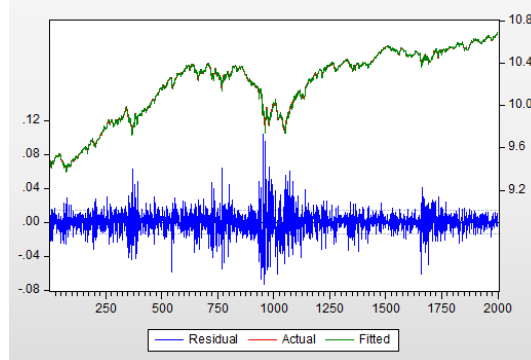
Tabla 19. Correlograma de los residuos al cuadrado ARMA GARCH de la segunda serie

1	-0.001	-0.001	0.0027	0.959
2	0.017	0.017	0.6045	0.739
3	-0.015	-0.015	1.0406	0.791
4	0.029	0.029	2.7161	0.606
5	-0.002	-0.001	2.7237	0.742
6	-0.022	-0.023	3.7171	0.715
7	0.014	0.015	4.1235	0.765
8	-0.009	-0.009	4.2998	0.829
9	0.006	0.005	4.3723	0.885
10	0.032	0.034	6.3861	0.782
11	-0.034	-0.035	8.6455	0.655
12	0.002	0.001	8.6513	0.732
13	-0.006	-0.004	8.7257	0.793
14	-0.006	-0.010	8.8015	0.844
15	0.020	0.023	9.6049	0.844
16	-0.011	-0.010	9.8291	0.875
17	0.002	-0.001	9.8416	0.910
18	-0.000	0.003	9.8416	0.937
19	0.009	0.005	9.9904	0.953
20	-0.020	-0.021	10.835	0.950
21	0.010	0.013	11.019	0.962

Fuente: Elaboración propia.

A continuación, se presenta la Figura 8 de los datos originales vs el modelo ajustado y otra gráfica con los residuos. De donde se puede observar que el modelo tiene un buen ajuste, pero aún hay picos en los residuos que no se controlan, y se alcanzan a apreciar los cúmulos de volatilidad:

Figura 8. Datos observados vs ajustados y residuos de la primera serie de datos ARMA-GARCH



Fuente: Elaboración propia.

4.2 Modelo de red neuronal estático

Lo que se intenta hacer al aplicar el modelo de red neuronal a una serie de tiempo, es lo siguiente: Tenemos N observaciones que son los datos reales y_t (en este caso los rendimientos logarítmicos del IPC) para $t=1,2,\dots,N$. De éstos tomamos k datos con $1 \leq k < N$ que serán metidos a la red para pronosticar el dato $k+1$, en la etapa de entrenamiento. Así, tenemos que el rendimiento logarítmico, el cuál será nuestro pronóstico k , satisface $\ln(y_k/y_{k-1}) = p_k$. Entonces el pronóstico del IPC es $y_k = y_{k-1}e^{p_k}$. Para la primera serie de modelos se trabajó con los primeros 150 datos para intentar hacer el pronóstico de los últimos 10 datos. Se realizaron los siguientes modelos, y se eligió al de menor error cuadrático medio.

Cuadro 20. Modelos con su error cuadrático medio

<i>Modelo.</i>	<i>ECM (muestra).</i>
4-4-1	6.0743E-05
6-6-1	5.69E-05
8-8-1	5.1973E-05

Por lo tanto, se trabajó con un modelo 8-8-1. En la Figura 9 se presentan los parámetros de la red neuronal para la primera serie de datos y en la Figura 10 se muestra el pronóstico del rendimiento logarítmico del IPC de la red neuronal de la primera serie de datos con las siguientes ponderaciones:

Weight to layer w(1,1)

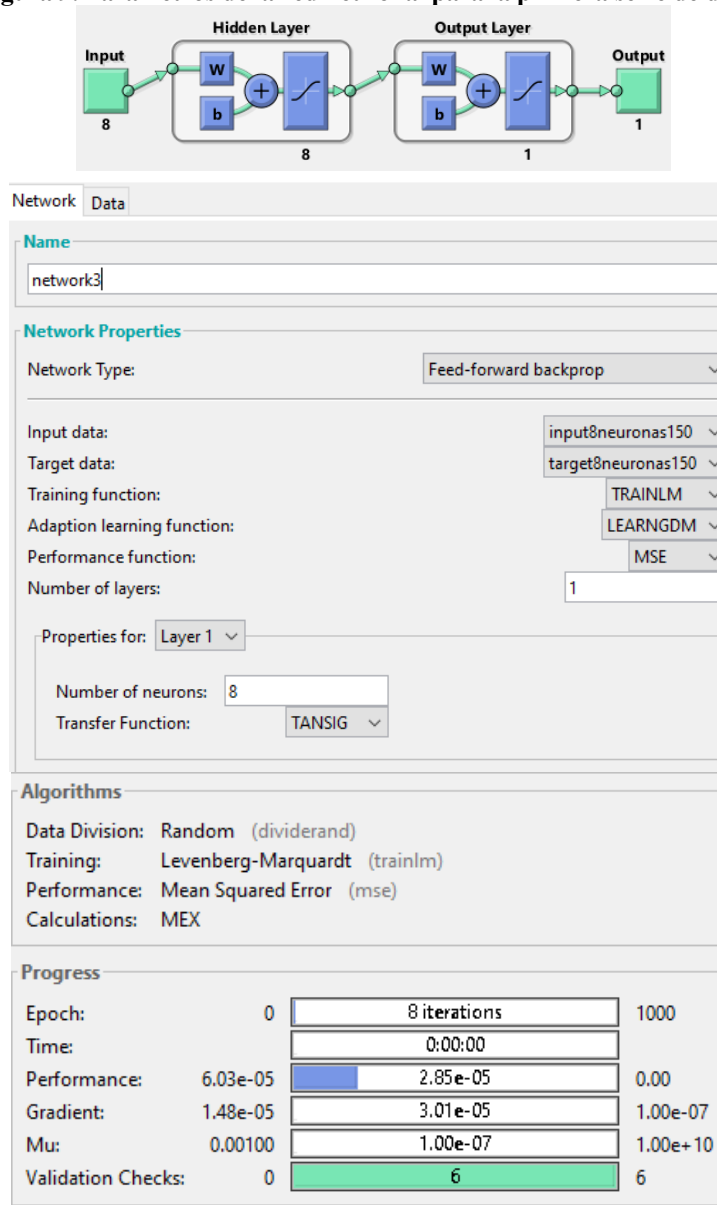
[0.6386 -0.25557 -0.88148 0.46219 0.94427 0.25665 -0.67373 -0.64351;
0.12312 -0.60212 0.73036 0.66923 -0.35906 -0.12847 0.82117 0.91427;
0.95192 0.3905 -0.68216 -0.12835 -0.17522 -0.32063 -0.3464 0.94275;
-0.9076 0.46154 0.12547 -0.53873 0.61783 -0.60247 0.93312 -0.57778;
-1.6353 0.89808 -0.55367 -0.66107 -0.34276 0.14761 -0.054284 0.43311;
-0.93663 0.92578 0.7712 0.2072 0.50287 -0.53639 0.56136 -0.18188;
-0.58352 0.69915 -0.018646 -0.62245 0.019221 -0.59978 0.47458 0.58379;
-0.4544 -0.64236 -0.1947 -0.97986 0.52129 0.62726 -0.4729 0.55919]

Weight to layer w(2,1)

[-0.25999 0.10574 0.4645 -0.071616 0.23511 0.012437 -0.17339 -0.22902]

Bias⁶ to layer 1 b(1)
[-2.0792; -1.4499; -1.6109; 0.23568; -0.20847; -0.84643; -1.0345; -2.0357]
Bias to layer 2 b(2)
[0.13372]

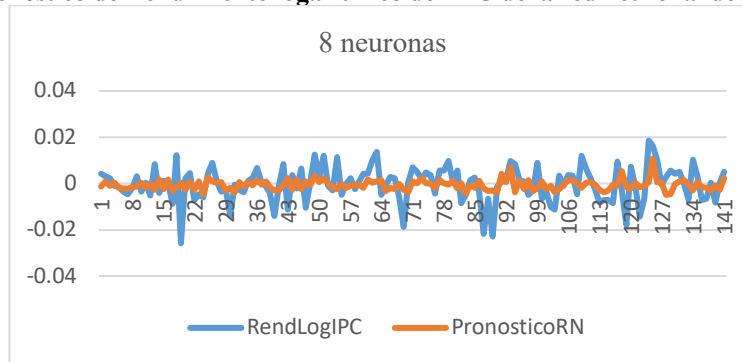
Figura 9. Parámetros de la red neuronal para la primera serie de datos



Fuente: Elaboración propia con Matlab.

⁶ Para más información consulte (Quora, 2016).

Figura 10. Pronóstico del rendimiento logarítmico del IPC de la red neuronal de la primera serie



Fuente: elaboración propia.

Para la segunda serie de modelos se trabajó con los primeros 2000 datos para intentar hacer el pronóstico de los últimos 10 datos. Se realizaron los siguientes modelos, y se eligió al que tuviera menor error cuadrático medio.

Tabla 21. Modelos con su error cuadrático medio

Modelo.	ECM (muestra).
4-4-1	0.00020005
6-6-1	0.00019932
8-8-1	0.00020074

Por lo tanto, se trabajó con un modelo 6-6-1. En la Figura 11 se presentan los parámetros de la red neuronal para la primera serie de datos y en la Figura 12 se muestra el pronóstico del rendimiento logarítmico del IPC de la red neuronal de la primera serie de datos con las ponderaciones:

Weight to layer w(1,1)

[-0.63916 1.0996 0.64865 0.44857 1.1786 -0.75109;
-1.4472 0.2649 0.49478 -0.11174 -1.0045 0.90037;
-0.76803 1.2565 -0.54429 -0.39811 -0.21915 0.36373;
-0.2428 -0.1518 0.97721 -0.79176 1.0383 -0.59075;
-0.58894 1.4682 1.1162 -0.14517 -0.16992 0.11258;
-0.518 -1.035 0.89029 0.29833 0.91483 -0.29163]

Weight to layer w(2,1)

[-0.27154 0.26168 0.075874 0.00016418 0.14556 0.23569]

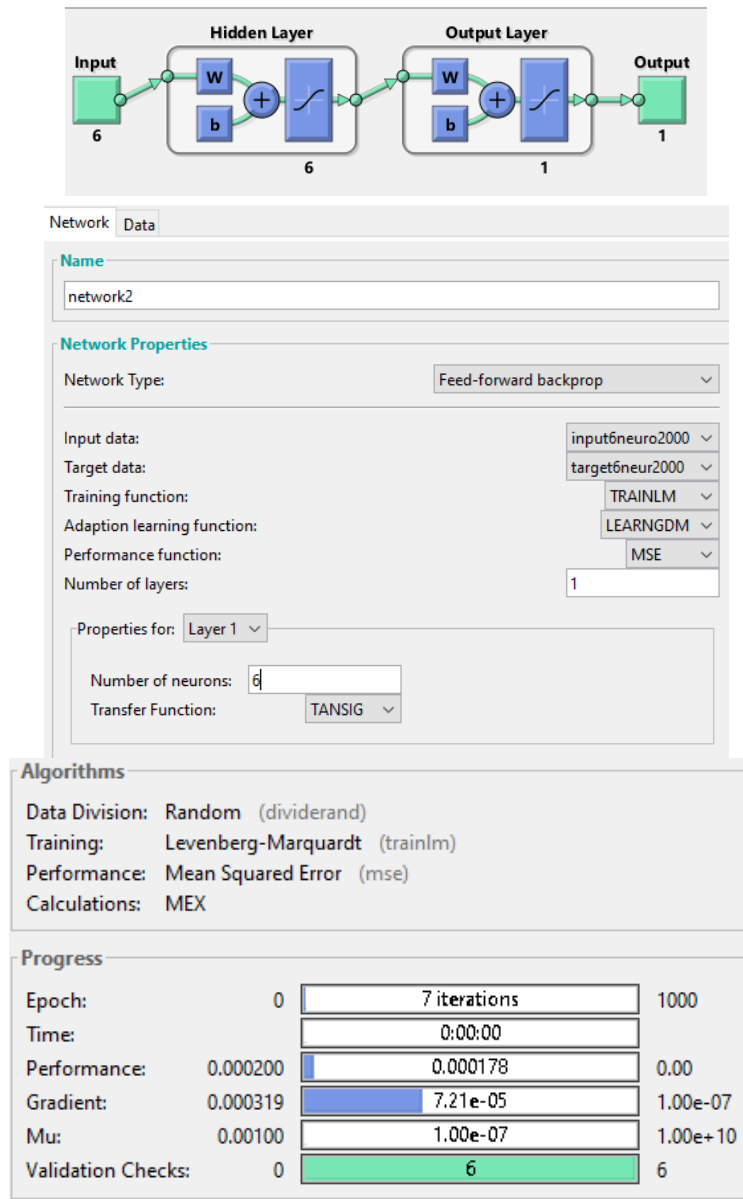
Bias to layer 1 b(1)

[1.8327; 2.0209; 0.64997; -0.61999; -1.8492; -1.918]

Bias to layer b(2)

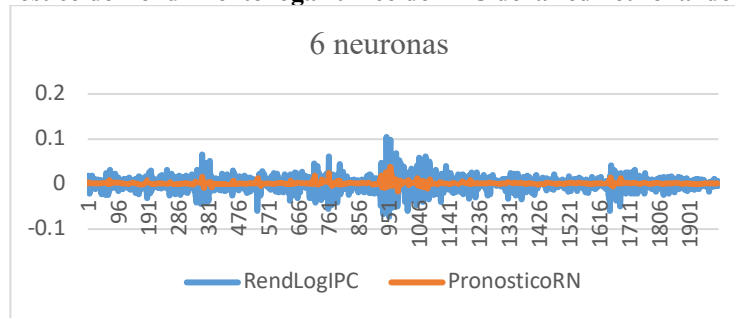
[0.13468]

Figura 11. Parámetros de la red neuronal para la primera serie de datos



Fuente: Elaboración propia con Matlab.

Figura 12. Pronóstico del rendimiento logarítmico del IPC de la red neuronal de la primera serie

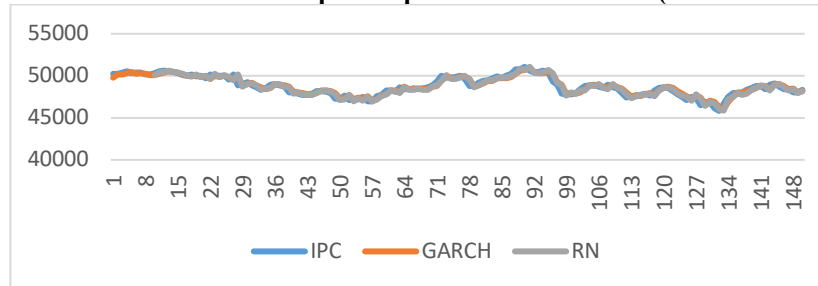


Fuente: elaboración propia.

4.3 Análisis y comparación de resultados

A continuación, se presenta la Figura 13 de los modelos vs los datos observados en la primera serie (dentro de la muestra).

Figura 13. IPC vs GARCH vs RN para la primera serie de datos (dentro de la muestra)



Fuente: elaboración propia.

Obtenemos los errores medios cuadráticos de la primera serie de datos (dentro de la muestra).

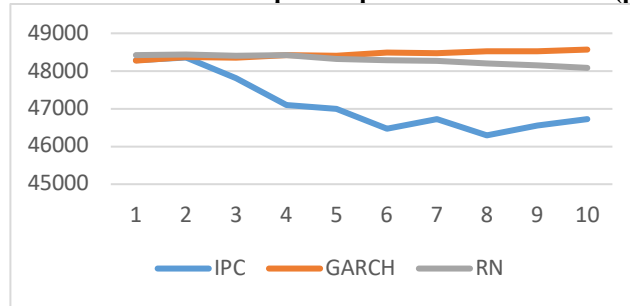
Tabla 22. Modelos con su error cuadrático medio

Muestra	
ECM_GARCH	119127.042
ECM_RN	122280.944

Fuente: Elaboración propia

Se presenta la Figura 14 de los modelos vs los datos observados en la primera serie (pronósticos).

Figura 14. IPC vs GARCH vs RN para la primera serie de datos (pronósticos)



Fuente: Elaboración propia.

Ahora se muestran los errores medios cuadráticos de la primera serie de datos (pronósticos).

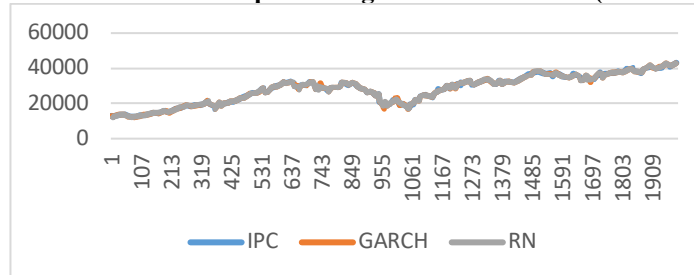
Tabla 23. Modelos con su error cuadrático medio

Pronóstico	
ECM_GARCH	2340470.32
ECM_RN	1756086.91

Fuente: Elaboración propia

Se presenta la Figura 15 de los modelos vs los datos observados en la segunda serie (dentro de la muestra).

Figura 15. IPC vs GARCH vs RN para la segunda serie de datos (dentro de la muestra)



Fuente: elaboración propia.

Obtenemos los errores medios cuadráticos de la segunda serie de datos (dentro de la muestra).

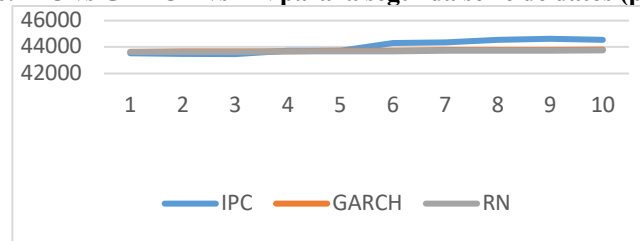
Cuadro 24. Modelos con su error cuadrático medio

Muestra	
ECM_GARCH	132542.645
ECM_RN	133633.077

Fuente: Elaboración propia

A continuación, se presenta en la Gráfica 16 de los modelos vs los datos observados en la segunda serie (pronósticos).

Figura 16. IPC vs GARCH vs RN para la segunda serie de datos (pronósticos)



Fuente: elaboración propia.

Por último, se obtenemos los errores medios cuadráticos de la segunda serie de datos (pronósticos).

Tabla 24. Modelos con su error cuadrático medio

Pronóstico	
ECM_GARCH	269141.574
ECM_RN	304953.001

Fuente: Elaboración propia

5. Conclusiones

En primer lugar, de la parte del análisis teórico de cada uno de los modelos se puede destacar que el modelo GARCH es un modelo atóxico (económicamente hablando), y su uso solamente está justificado empíricamente por las características propias de una serie financiera, tal como las colas pesadas (las cuales no coinciden con las características de una distribución normal) y los cúmulos de volatilidad (debidos a que el proceso tiene memoria). El modelo GARCH, al ser un modelo ARIMA y modelar al mismo tiempo la volatilidad con otro modelo ARIMA puede detectar estos cambios y guardarlos en memoria. Al mismo tiempo los parámetros obtenidos poseen sentido económico (volatilidad). El modelo de RN es un enfoque distinto, ya que simula el funcionamiento básico de una neurona.

En segundo lugar, de la parte de la implementación de los modelos se pueden señalar que para la primera serie de pocos datos y baja volatilidad, a diferencia de trabajos anteriores, basta con hacer un modelo ARIMA, se vio en las pruebas estadísticas que no era necesario hacer un modelo GARCH pues los residuos se comportan como variables aleatorias independientes idénticamente distribuidas y la volatilidad se puede asumir como constante.

Al comparar los modelos, GARCH con RN, se observa que dentro de la muestra el modelo GARCH tuvo un menor error, pero fuera de la muestra, en la parte de los pronósticos la RN tuvo menor error. Esto debido a que el modelo GARCH propuesto aumentará en sus pronósticos debido al aumento en su volatilidad, mientras que la RN puede aumentar o disminuir en sus pronósticos. En ambos casos sobrestimaron los valores reales del IPC.



Para la segunda serie de muchos datos y alta volatilidad, en la parte del modelo GARCH, se hicieron pruebas anteriores para hacer la serie estacionaria en diferencias, pero se obtenía un grado de ajuste muy bajo, es decir no se podía explicar buena parte de la serie; por lo tanto, se optó por hacer la serie de los logaritmos del IPC estacionaria en tendencia, lo cual se vio reflejado en un mejor grado de ajuste y se pudieron meter parámetros autorregresivos y de media móvil. Al hacer las pruebas estadísticas se observó que existían relaciones de dependencia en los residuos lo cual sugiere el uso del modelo GARCH para la modelación de la volatilidad. Además, se observó que dichos residuos no se distribuyen normalmente, pero al hacer cambios en la distribución no se notó una diferencia significativa, por lo que se optó por seguir con la distribución normal.

En cuanto a la RN, se observa una cuestión interesante, se podría suponer que al aumentar el número de neuronas mejoraría el ajuste, pero no fue así, ya que se observa un error menor del modelo de 6 neuronas frente al modelo de 8 neuronas. Otro punto importante es que la RN pudo capturar la volatilidad del periodo de crisis y volver a una baja volatilidad después de dicho periodo. Dentro de la muestra el error del modelo GARCH fue menor que el error de la RN, así como también en los pronósticos. Además, ambos modelos subestimaron los pronósticos frente a los datos reales. Con lo que se puede concluir que el modelo de RN es una buena alternativa al modelo GARCH en varios aspectos: Precisión y horizonte temporal: Aunque los ajustes y pronósticos del modelo GARCH fueron mejores, la RN no estuvo tan alejada y logró capturar los cambios significativos en el corto y largo plazo en escenarios de alta y baja volatilidad; Facilidad de implementación: En cuanto al manejo de los programas donde se realizaron los modelos no hay mucha diferencia. La diferencia radica en que para la elaboración del modelo GARCH se necesitan conocimientos específicos de econometría. Mientras que para la implementación del modelo de RN basta con una manipulación básica de los datos. Además, conviene señalar que el modelo de RN que se implementó es básico (perceptrón multicapa con una capa oculta).

Como posibilidad de trabajo futuro se podría observar como mejora la precisión de la RN con un modelo más sofisticado y automatizado, tal vez con más capas ocultas (mayor precisión), que sea multivariable (sentido económico) y con algoritmos genéticos (capacidad de clasificación evolutiva).

Referencias

- Argáez, J. *et al.* (2014). Un paseo por el modelo GARCH y sus variantes. Facultad de Matemáticas, Universidad Autónoma de Yucatán, artículo febrero de 2014. Consultado el 20-03-2017 en: <http://intranet.matematicas.uady.mx/journal/descargar.php?id=59>
- Bolsa Mexicana de Valores (2016a). *ACTINVER Casa de bolsa reporta resultados del tercer trimestre 2016*. Documento consultado el 20-03-2017 en: https://www.bmv.com.mx/docs-pub/infincab/cabinfin_705595_2016-03_2.pdf
- Bolsa Mexicana de Valores (2016b). *IPC Metodología*. Octubre 2016. Documento consultado el 10-05-17 en: http://www.bmv.com.mx/docs-pub/INDICES/CTEN_INNM/metodologia_bmv_ipc.pdf
- Bolsa Mexicana de Valores (2017). *Grupo BMV. Acerca de*. Página consultada el 10-05-2017 en: <https://www.bmv.com.mx/es/grupo-bmv/acerca-de>



- Box, G. and Jenkins G. (1971). Time Series Analysis, Forecasting and Control. *Journal of the Royal Statistical Society. Series A (General)* Vol. 134, No. 3. pp. 450-453
- Cruz, I y Morales, A. (2011). *Pronósticos utilizando redes neuronales artificiales y modelos ARIMA: El caso de futuros del CETE a 91 días en el MexDer*. En Avances recientes en valuación de activos y administración de riesgos. Ed. Universidad Panamericana, México. 438 pp.
- Cruz-Aké, S. (2018). Algoritmo para obtener el mejor modelo GARCH [Material de clase]. *Econometría II*. Instituto Politécnico Nacional. México, D.F.
- Devore, Jay (2008). *Probabilidad y estadística para ingeniería y ciencias*. Ed. Cengage Learning Editores, México D.F.
- Fabozzi, F. J. (2007). *Financial econometrics. From basics to advanced modeling techniques*. Ed. John Wiley & Sons. US.
- Florez, R. y Fernández, J. (2008). Las redes neuronales artificiales. Fundamentos teóricos y aplicaciones prácticas. Ed. Netbiblo. España.
- Franses, P. y Dijk, Dick van (2003). Nonlinear time series models in empirical finance. Ed. Cambridge University Pres. US.
- García-Estévez, P. (2002). Aplicaciones de las Redes Neuronales en las Finanzas. En Documentos de trabajo de la facultad de ciencias económicas y empresariales no. 2. Ed. Universidad Complutense de Madrid, Facultad de ciencias económicas y empresariales, Madrid, España.
- Gómez-Ramos, E. L., F. Venegas-Martínez y H. Allier-Campuzano (2011). 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*, Vol. 6, No. 32, pp. 3-22.
- Greene, W. (1999). Análisis econométrico. Tercera edición. Ed. Prentice Hall. España, Madrid.
- Gúzman, María de la Paz et. al. (2007) El futuro del índice de precios y cotizaciones de la Bolsa Mexicana de Valores. Análisis económico Vol. XXII no. 49. Universidad Autónoma Metropolitana Unidad Azcapotzalco, México.
- Hawley, Delvin D. et. al. (1996). *Artificial neural systems: A new tool for financial decision-making*. En Neural Networks in finance and investing. Using artificial intelligence to improve real-world performance. Ed. Irwin Professional Publishing.
- Lorenzo, Arturo y Ruiz, Antonio (2011). *Modelación de rendimientos bursátiles mexicanos mediante los modelos TGARCH y EGARCH: Un estudio econométrico para 30 acciones y el índice de precios y cotizaciones*. Munich Personal RePEc Archive, MPRA Paper no. 36782 Febrero de 2012. Consultado el 20-03-2017 en: <http://mpra.ub.uni-muenchen.de/36872/>
- Más fondos (2017) El IPC es S&P/BMV IPC, lo que cambia la metodología del cálculo. Página consultada el 20-09-2017 en: <http://blog.masfondos.mx/el-ipc-ahora-es-sp/bmv-ipc-lo-que-cambia-la-metodologia-del-calculo>
- McNelis, P. (2005). *Neural networks in finance: Gaining predictive edge in the market*. Ed. Elsevier academic press, UK.
- Medsker, Larry et. al. (1996). *Neural networks fundamentals for financial analysts*. En Neural Networks in finance and investing. Using artificial intelligence to improve real-world performance. Ed. Irwin Professional Publishing.
- MexDer (2018a). Presentación. Página consultada el 17-05.18 en: <http://www.mexder.com.mx/wb3/wb/MEX/presentacion>
- MexDer (2018b). Contratos de futuro. Página consultada el 29-10-18 en:



http://www.mexder.com.mx/wb3/wb/MEX/contratos_futuro

- Ortiz, F. *et al.* (2013). *Seguimiento y pronóstico de tres acciones que cotizan en la BMV empleando redes neuronales diferenciales*. En Avances recientes en valuación de activos y administración de riesgos. Ed. Universidad Panamericana, México. p.322
- Ortiz-Arango, F. A. I. Cabrera-Llanos y F. Venegas-Martínez (2016). Euro Exchange Rate Forecasting with Differential Neural Networks with an Extended Tracking Procedure. *International Journal of Pure and Applied Mathematics*. Vol. 107, No. 1, pp. 87-109.
- Parkin, M. (2014). Economía. Ed. Pearson Educación, México p.568.
- Ponce, P. (2010). *Inteligencia artificial con aplicaciones a la ingeniería*. Primera edición. Ed. Alfaomega. México
- Quora (2016). What is bias in artificial neural network? Página consultada el 29-10-18 en: <https://www.quora.com/What-is-bias-in-artificial-neural-network>
- Rosen, J. (1990). Fundamental manifestations of symmetry in physics. *Foundations in Physics*, Vol. 20 no.3 ed. Springer
- Rustichini, Aldo *et al.* (2002). A brain imaging study of procedural choice. Working paper. Department of economics. University of Minnesota.
- Standard and Poor's Dow Jones Indices (2017a). S&P/BMV Índices metodología. Agosto, Ciudad de México.
- Standard and Poor's Dow Jones Indices (2017b). S&P Dow Jones Índices emite el siguiente recordatorio de cambios a la metodología de los S&P/BMV Índices. Agosto, Ciudad de México.
- Standard and Poor's Dow Jones Indices (2017c). S&P Dow Jones Índices anuncia resultados del cambio de muestra para el S&P/BMV IPC. Septiembre, Ciudad de México.
- Valencia M., Yañez C. y Sánchez L. (2006). Algoritmo backpropagation para redes neuronales: conceptos y aplicaciones. Centro de investigación en Computación. Instituto Politécnico Nacional. Manuscrito.
- Venegas-Martínez, F. (2008). Riesgos financieros y económicos: Productos derivados y decisiones económicas bajo incertidumbre. Segunda edición. Editorial Cengage Learning. México.