

# Red neuronal autorregresiva difusa tipo Sugeno con funciones de membresía triangular y trapezoidal: una aplicación al pronóstico de índices del mercado bursátil

## Sugeno Type Fuzzy Nonlinear Autoregressive Neural Networks with Triangular and Trapezoidal Membership Functions: An Application to Forecast the Stock Market Index

---

José Eduardo Medina Reyes\*

Judith Jazmin Castro Pérez\*\*

Agustín Ignacio Cabrera Llanos\*\*\*

Salvador Cruz Aké\*\*\*\*

(Fecha de recepción: 27 de octubre de 2019, Fecha de aceptación 26 de diciembre de 2019)

### RESUMEN

La presente investigación desarrolla una comparación entre la nueva Red Neuronal Autorregresiva no Lineal Difusa y la Red Neuronal Autorregresiva para evaluar el pronóstico de Índices bursátiles. Para ello se aplica la metodología a la rentabilidad de cuatro índices accionarios, IPC, IBEX 35, S&P 500 y el Nikkei 225 en formato diario

---

\* Escuela Superior de Economía,  
Instituto Politécnico Nacional (IPN), Ciudad de México, México  
eduardomedina\_94@yahoo.com, ORCID: 0000-0001-6320-9299

\*\* Escuela Superior de Economía,  
Instituto Politécnico Nacional (IPN), Ciudad de México, México  
castro.410210@gmail.com, ORCID: 0000-0002-3412-2079

\*\*\* Unidad Profesional Interdisciplinaria de Biotecnología,  
Instituto Politécnico Nacional (IPN), Ciudad de México, México  
aicllbuda@yahoo.com, ORCID: 0000-0003-3540-7313

\*\*\*\* Escuela Superior de Economía,  
Instituto Politécnico Nacional (IPN), Ciudad de México, México  
salvador.ake22@gmail.com, ORCID: 0000-0003-1452-377X

desde enero de 2015 hasta diciembre de 2018, adjuntando los primeros cinco días de enero de 2019 para pronóstico fuera de muestra. Se estimó una Red Neural Autorregresiva No Lineal con tres rezagos y con algoritmo de aprendizaje Bayesiano y la Red Neuronal Difusa fue estimada con tres rezagos y con el algoritmo Backpropagation. Los resultados muestran que los modelos propuestos generan un mejor pronóstico dentro y fuera de la muestra en comparación con la Red Neuronal Autorregresiva No Lineal. Lo anterior es consecuencia de que las redes neuronales pueden aprender de la dinámica de las series temporales y si se añade la teoría difusa, también pueden aprender de la incertidumbre inherente a las variables financieras, esta situación hace que el método propuesto sea mejor que la red neuronal tradicional.

Clasificación JEL: C01; C32; C45.

**Palabras clave:** Red neuronal difusa, función de pertenencia triangular, función de pertenencia trapezoidal, series de tiempo difusas.

#### ABSTRACT

*This article compares the results obtained when forecasting the Stock Market Index applying a proposed Fuzzy Nonlinear Autoregressive Neuronal Network with those obtained using the Autoregressive Neuronal Network. For this purpose, the methodology is applied to four stock indices, IPC, IBEX 35, S&P 500 and the Nikkei 225 using daily data from January 2015 to December 2018, the first five financial days of January 2019 are added to carry out a forecast outside the sample. A Nonlinear Autoregressive Neuronal Network with three lags and Bayesian learning algorithms and the Fuzzy Nonlinear Autoregressive Neuronal Networks with three lags and a Backpropagation algorithm were used to calculate a forecast. The results have shown that the models proposed generate better forecasts considering in-sample and out-sample tests than the Nonlinear Autoregressive Neuronal Network. It was shown that the neural networks can learn from the dynamics of time series, and if fuzzy theory is added, they can also learn from the uncertainty around financial variables. This indicates that method proposed yields better results than the traditional network method.*

*Classification JEL: C01; C32; C45*

**Keywords:** *Fuzzy Nonlinear Autoregressive Neuronal Network; Triangular Membership Function; Trapezoidal Membership Function; Fuzzy Time Series.*

## Introducción

La presente investigación desarrolla la comparación entre la eficiencia de una nueva propuesta de Red Neuronal Autorregresivas basada en teoría difusa y la Red Neuronal Autorregresiva tradicional o Red Neuronal Autorregresiva No Lineal, para generar pronóstico de los índices bursátiles. Para realizar el estudio se tomaron los siguientes índices: IPC; IBEX 35; S&P 500 y Nikkei 225; en un formato diario desde enero de 2015 hasta diciembre de 2018, anexando los cinco primeros días de enero de 2019 para efectuar el

análisis fuera de muestra. El principal argumento de este documento es que la dinámica de las series de tiempo financieras puede modelarse con un mejor ajuste si se otorgan grados de membresía a la volatilidad.

Desde que la teoría de Series de Tiempo Difusas fue propuesta por Song y Chissom (1993a) se han generado varias aplicaciones en el pronóstico de series asociadas a temas como aplicaciones a universidades, economía, finanzas y otros. Se asume que la serie de tiempo es un conjunto difuso y por lo tanto puede ser analizado mediante relaciones de lógica difusas y razonamiento aproximado. Este modelo se basa en la incertidumbre y el conocimiento impreciso que contienen los datos de las series temporales. Song y Chissom (1993b) presentaron la definición de serie temporal difusa a partir de procesos esencialmente dinámicos en los que las observaciones son valores lingüísticos.

Tseng *et al.* (2001) a partir del modelo de series de tiempo ARIMA y el modelo de regresión difusa desarrolló una nueva metodología llamada modelo ARIMA-Difuso (Fuzzy-ARIMA). Éste se aplicó para pronosticar el tipo de cambio del dólar taiwanés contra el dólar americano. La importancia de este modelo es que proporciona a los tomadores de decisiones los mejores y peores escenarios posibles.

Tsaur (2012) elaboró un modelo de series temporales difusas con Cadenas de Markov para realizar pronósticos; mediante este modelo las series temporales difusas se modelan y se analizan mediante lógica difusa y el pronóstico se determina mediante inferencia difusa.

Otras formas de analizar las series temporales difusas fueron revisadas por Singh (2017); las principales características de sus investigaciones son: la determinación de la longitud de los intervalos, el establecimiento de relaciones de lógica difusa entre diferentes factores y la *desdifusificación* en un modelo de hibridación. También realizó una revisión de los trabajos que se han desarrollado con esta metodología.

Dash *et al.* (2016), desarrollaron una nueva metodología en la que consideran a la volatilidad de las series financieras como un proceso difuso. A partir de la ecuación de varianza del modelo EGARCH se crea una función de membresía gaussiana para modelar la volatilidad, y se genera un pronóstico a partir de reconocer que la volatilidad de las series de tiempo sigue un comportamiento normal. Se encontró que este modelo híbrido mejora significativamente los resultados con respecto a los modelos de menor complejidad. Pal *et al.* (2017) desarrollan un modelo que utiliza la técnica de Algoritmo Genético para decidir el tamaño y el número de intervalos y relaciones difu-

sas en la Serie de Tiempo Difusa. Este trabajo mejora el pronóstico con respecto a investigaciones anteriores sobre el tema.

Pal *et al.* (2018), desarrollan un estudio sobre el pronóstico de series temporales sobre diversos temas, donde se destaca el análisis de redes neuronales que modifican sus pesos sinápticos bajo modelos difusos tipo-2. Los resultados numéricos indican que su modelo logra captar la incertidumbre de las diversas situaciones en que está inmersa la sociedad. Otros trabajos similares son (Yu y Huarng, 2010); (Xiao, 2017); (Han, Zhong, y Han., 2018); (Egrioglu, Aladag y Yolcu, 2013); (Souza y Torres, 2018).

Srinivasan *et al.* (2019) analizan el pronóstico de las series temporales del mercado bursátil utilizando las Series Temporales Difusas y el Algoritmo de Búsqueda Gravitacional, partiendo de reconocer la incertidumbre de la acción a partir de su precio de apertura, máximo, mínimo y cierre. La metodología aplicada para este análisis muestra una mejora en comparación con los resultados de otras metodologías de Series Temporales Difusas.

Medina-Reyes *et al.* (2019), desarrollaron una serie de modelos mediante lógica difusa y modelos de varianza condicional, en los que se destaca la incorporación de parámetros difusos gaussianos a la ecuación de varianza de los modelos. Se encontró que los modelos híbridos, generan mejor pronóstico que los modelos tradicionales de econometría financiera.

Los sistemas difusos se han destacado por mejorar la solución de diversos problemas en los que existe ambigüedad en la información, y han sido de gran ayuda para identificar mejor las características de eventos difíciles de observar. Por tanto, esta investigación tiene por objeto incorporar los sistemas difusos para mejorar el aprendizaje de las redes neuronales (Mohammadian, 2020), (Cox, 2019).

El objetivo de este trabajo es realizar una comparación entre las redes neuronales autorregresivas tradicionales y los métodos difusos propuestos, para ello, la hipótesis es que los modelos difusos reconocen mejor la volatilidad de los índices bursátiles que los modelos tradicionales. Con el fin de mostrar la aplicabilidad y eficacia del método propuesto, se genera el pronóstico de los índices bursátiles Nikkei 225, IBEX 35, IPC y S&P 500. El resultado muestra que la metodología propuesta obtiene mejores pronósticos que los modelos tradicionales (Red neuronal autorregresiva no lineal).

La estructura de este trabajo se organiza de la siguiente manera: en la primera sección se examinan los conceptos de series temporales difusas y redes adaptativas basadas en el modelo del sistema de inferencia difusa. En

la sección 2, se formulan y proponen el modelo híbrido de Series Temporales Difusas y Redes Neuronales No Lineales. En la sección 3, los modelos se aplican para pronosticar los índices bursátiles de EE. UU., España, México y Japón; y se comparan con los pronósticos obtenidos con la Red Neural Autorregresiva No Lineal. Por último, se presentan las conclusiones.

## 1. Revisión conceptual de los modelos de series temporales difusas y redes neuronales adaptativas basadas en sistemas de inferencia difusa

Las series de tiempo difusas se entienden según Song y Chissom (1994), como un proceso  $Y(t)$  ( $t = \dots, 0, 1, \dots$ ), un subconjunto de  $Z$  y el correspondiente discurso del conjunto difuso  $\mu_i(t)$  ( $i = 1, 2, \dots$ ), y tal que sea  $F(t)$  una colección de  $\mu_1(t), \mu_2(t), \dots$  funciones de membresía. Entonces,  $F(t)$  será conocida como una serie de tiempo difusa en  $Y(t)$  ( $t = \dots, 0, 1, \dots$ ). Donde el universo discurso Song y Chissom (1993a) es un conjunto difuso, de forma que su cota inferior y superior es delimitada por la información de la serie de tiempo.

$$U = [L_{bd}, U_{bd}] \quad (1)$$

Tomando a  $F(t-1) = A_i$  y  $F(t) = A_j$ , se dice que existe una relación de lógica difusa entre  $F(t-1)$  y  $F(t)$ , si  $A_i \rightarrow A_j$ . Por otro lado, las reglas IF-THEN de  $F(t)$  en función de  $F(t-1)$  se expresan como  $F(t) = F(t-1) \circ R(t, t-1)$  esto se conoce como modelo de primer orden de  $F(t)$  Song y Chissom (1993a).

También si  $R(t, t-1) = R(t-1, t-2) \forall t$  entonces  $F(t)$  será llamada serie de tiempo difusa en tiempo invariante, y en caso contrario, como serie de tiempo difusa en tiempo variante. De esta manera, si  $F(t)$  es función de  $F(t-1), F(t-2), \dots, F(t-n)$  se dice que es un modelo de orden superior  $F(t-n), \dots, F(t-2), F(t-1) \rightarrow F(t)$  Song y Chissom (1993b).

Hasta el momento se asume que la serie de tiempo es un conjunto difuso y toda la información para conocer su dinámica está en su historia. Sin embargo, la dificultad de modelar series financieras obliga a utilizar modelos de mayor complejidad y como respuesta a ello se usan los modelos lineales difusos en combinación con modelos autorregresivos.

Otras formas de analizar las series temporales difusas fueron examinadas por Singh (2017) y Yu (2005), donde las principales características de las investigaciones son: la determinación de la longitud de los intervalos, el establecimiento de relaciones lógicas difusas entre diferentes factores, y la *desfuzzificación* de los valores de las series temporales difusas.

A partir de Wu *et al.* (2000), se propone un modelo de redes neuronales, donde la premisa fundamental es que la información que alimenta la red es un conjunto difuso. En cada fase de la red, los nodos pueden expresarse a través de funciones de pertenencia. Estas últimas, buscan extraer el grado de pertenencia de la información sobre el conjunto de estudio.

### 1.1. Redes neuronales adaptivas basadas en sistemas de inferencia difusos

Según Jang (1993), es necesario considerar una variable  $x$  que genera otra variable de salida  $z$  a través de una red neuronal, a partir de la cual se forma un modelo difuso de primer orden del tipo Sugeno, en el que las reglas de *If-Then* se determinan de la siguiente manera:

$$R_1: \text{if } x \text{ is } A_1 \text{ then } f_1 = p_1x + r_1 \quad (2)$$

$$R_2: \text{if } x \text{ is } A_2 \text{ then } f_2 = p_2x + r_2$$

En la primera fase de la estructura de red neuronal, cada nodo representa una variable de tipo lingüístico. En otras palabras, existe un subconjunto difuso para cada nodo de entrada de la red neuronal. En la segunda fase, cada nodo estará representado por una función de pertenencia, de la siguiente forma:

$$\mu_{ij}(x_i) = e^{-\frac{(x_i - c_{ij})^2}{\sigma_j^2}}; \quad i = 1, \dots, r; j = 1, \dots, u. \quad (3)$$

Donde,  $\mu_{it}(x_i)$  es una función de membresía con elementos  $x_i$ ,  $\sigma_j^2$  es el ancho y  $c_{ij}$  el centro de la función. Las variables que alimentan la red son  $r$  y el número de funciones de membresía es  $u$ . En la terce-

ra fase, cada nodo representa una parte de la regla difusa  $I_j$  y en la  $j$  – ésima regla  $R_j$ , su salida viene dada por:

$$\varphi_j = e^{-\frac{\sum_{i=1}^r (x_i - c_{ij})^2}{\sigma_j^2}} \quad (4)$$

$X = (x_1, \dots, x_r)$  y  $C_j$  es el centro de la  $j$  – ésima representa la unidad de la función de base radial. En la cuarta fase, los perceptrones se denotan como nodos normalizados. El número de nodos normalizados  $N$  es igual al número de nodos  $R$ .

$$\theta_j = \frac{\varphi_j}{\sum_{k=1}^u \varphi_k} \quad (5)$$

En la quinta fase, cada nodo está representado por las variables de salida como la suma de las señales de entrada a esta fase, de tal manera que:

$$\theta_j = \frac{\varphi_j}{\sum_{k=1}^u \varphi_k} \quad (6)$$

Donde  $y$  es el valor de la variable de salida y  $w_{2k}$  es el peso de cada regla. Los pesos están dados por:

$$w_{2k} = k_{j0} + k_{j1}x_1 + \dots + k_{jr}x_r \quad (7)$$

En conclusión, este modelo nos permite modelar las series temporales difusas utilizando los componentes de una red neuronal. Se asume, que las funciones de pertenencia son de tipo gaussiano, lo que permite identificar la pertenencia del valor de entrada en el subconjunto difuso. Por lo tanto, se reconoce que es posible desarrollar el pronóstico del mercado de valores mediante la teoría difusa y las redes neuronales artificiales.

## 2. Formulación de la Red Neuronal Autorregresiva Difusa No Lineal Híbrida: NARNET Triangular difusa y NARNET trapezoidal difusa

Esta sección muestra la estructura teórica de la Red Neuronal Autorregresiva Difusa para dos casos. En el primero, la red neuronal con función de membresía triangular y su capacidad de generar predicciones de la volatilidad de las variables financieras. Y en segundo, la red neuronal con función de membresía trapezoidal y sus cualidades teóricas para generar estimaciones de series de alta volatilidad. Estos modelos se construyeron a partir de dos metodologías existentes, las “Series Temporales Difusas” y las “Redes Neuronales Autoregresivas”.

La red neuronal NARNET Triangular Difusa (Fuzzy Triangular NARNET)<sup>1</sup> es un modelo de lógica difusa de primer orden del tipo Sugeno, y sus reglas *If-Then* en la capa de entrada se determinan de la siguiente manera:

$$R_1: \text{if } y_{t-i} \text{ is } A_1 \text{ then } f_1 = w_{11}y_{t-1} + w_{12}y_{t-2} + \dots + w_{n1}y_{t-n} \quad (8)$$

$$R_2: \text{if } y_{t-i} \text{ is } A_2 \text{ then } f_2 = w_{21}y_{t-1} + w_{22}y_{t-2} + \dots + w_{n2}y_{t-n}$$

$$R_3: \text{if } y_{t-i} \text{ is } A_3 \text{ then } f_3 = w_{31}y_{t-1} + w_{32}y_{t-2} + \dots + w_{n3}y_{t-n}$$

donde  $R$  representa las reglas *If-Then* de la serie de tiempo borrosa,  $A$  es el subconjunto triangular difuso para cada función  $f$ , que denota cada perceptrón con sus respectivos pesos sinápticos  $w$ . La Gráfica 1, muestra la función de membresía asociada a cada regla *If-Then*  $R$  que corresponde a un subconjunto difuso  $A$ . En este caso, los subconjuntos difusos son los niveles de volatilidad y las reglas *If-Then* son las funciones de aprendizaje difuso de la red neuronal de la primera capa.

Las reglas *If-Then* en la capa oculta son:

$$R_1: \text{if } y_{t-i} \text{ is } A_1 \text{ then } f_l = w_{k1}f_1 + w_{k2}f_1 \quad (9)$$

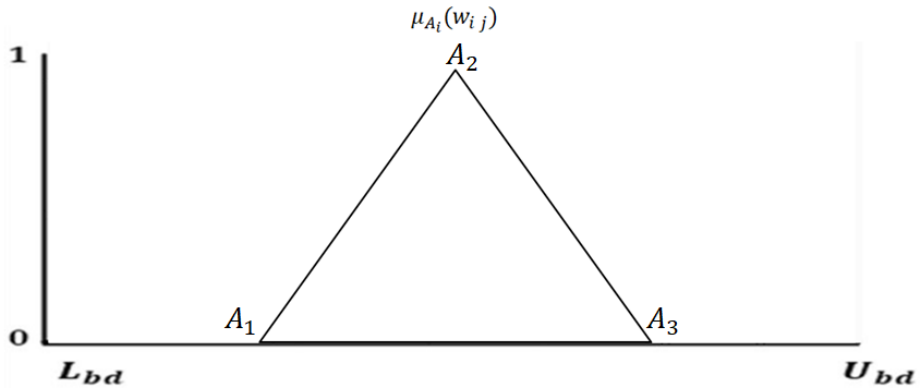
$$R_2: \text{if } y_{t-i} \text{ is } A_2 \text{ then } f_l = w_{m1}f_2 + w_{m2}f_2$$

$$R_3: \text{if } y_{t-i} \text{ is } A_3 \text{ then } f_l = w_{m1}f_3 + w_{m2}f_3$$

<sup>1</sup> NARNET por sus siglas en inglés Non Linear Autoregressive Neural Network, Red Neuronal Autorregresiva No Lineal Difusa Híbrida.



Gráfica 1. Función de Membresía Triangular



Fuente: elaboración propia en MatLab.

donde  $A$  es el subconjunto triangular difuso de cada  $f$  percepción difusa con sus respectivos pesos para la capa oculta de la red neuronal.

La ventaja de este método con respecto a la Red Neural Autoregresiva tradicional, es que el modelo propuesto aprende de la volatilidad de tres maneras: alta volatilidad, causada por buenas noticias; alta volatilidad, generada por malas noticias; y en baja volatilidad. Por ejemplo, la técnica propuesta genera un pronóstico con relación a la volatilidad de la serie financiera por medio de una función de membresía, lo que permite generar estimaciones según el grado de volatilidad.

La segunda red neuronal es la NARNET trapezoidal difusa (Fuzzy Trapezoidal NARNET). Se utiliza una red neuronal autorregresiva y un modelo difuso de primer orden del tipo Sugeno, para identificar sus componentes. Las reglas *If-Then* en la capa de entrada se determinan de la siguiente manera:

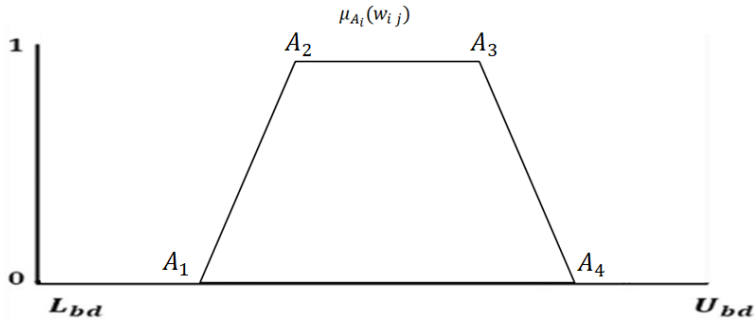
$$R_1: \text{if } y_{t-i} \text{ is } A_1 \text{ then } f_1 = w_{11}y_{t-1} + w_{12}y_{t-2} + \dots + w_{1n}y_{t-n} \quad (10)$$

$$R_2: \text{if } y_{t-i} \text{ is } A_2 \text{ then } f_2 = w_{21}y_{t-1} + w_{22}y_{t-2} + \dots + w_{2n}y_{t-n}$$

$$R_3: \text{if } y_{t-i} \text{ is } A_3 \text{ then } f_3 = w_{31}y_{t-1} + w_{32}y_{t-2} + \dots + w_{3n}y_{t-n}$$

$$R_4: \text{if } y_{t-i} \text{ is } A_4 \text{ then } f_4 = w_{41}y_{t-1} + w_{42}y_{t-2} + \dots + w_{4n}y_{t-n}$$

Gráfica 2. Función de Membresía Trapezoidal



Fuente: elaboración propia en MatLab.

donde  $R$  representa las reglas *If-Then* de la serie de tiempo difuso,  $A$  es el subconjunto difuso trapezoidal de cada función  $f$  asociado a un perceptrón difuso con sus respectivos pesos  $w$ . En este caso, se asume que la volatilidad puede capturarse a través de cuatro subconjuntos difusos.

En este sentido sea  $A$  el subconjunto difuso trapezoidal de cada función asociado a un perceptrón difuso con sus respectivos pesos  $w$  en la capa oculta de la red neuronal. Ahora las reglas de *If-Then* en la capa oculta son:

$$R_1: \text{if } y_{t-i} \text{ is } A_1 \text{ then } f_l = w_{k1}f_1 + w_{k2}f_1 \quad (11)$$

$$R_2: \text{if } y_{t-i} \text{ is } A_2 \text{ then } f_l = w_{m1}f_2 + w_{m2}f_2$$

$$R_3: \text{if } y_{t-i} \text{ is } A_3 \text{ then } f_l = w_{m1}f_3 + w_{m2}f_3$$

$$R_4: \text{if } y_{t-i} \text{ is } A_4 \text{ then } f_l = w_{m1}f_4 + w_{m2}f_4$$

En la primera fase de la red, cada nodo se concentra con la respectiva función de pertenencia del subconjunto difuso. Esto significa que  $y_t$  es la variable que alimenta cada nodo  $i$  y  $A_i$  es el subconjunto difuso denotado como valor lingüístico.

$$y_t = A_i$$

donde  $A_i$  es el conjunto difuso de  $y_t$  y expresa la serie temporal difusa. En específico, primero se usan las funciones de membresía de tipo triangular o trapezoidal y después, para mejorar la función de aprendizaje, la información se suaviza de la siguiente manera García *et al.* (2002):

$$z_{t-1} = \frac{y_{t-1} - \mu(y_{t-1} + y_{t-2} + \dots + y_{t-n})}{\sigma(y_{t-1} + y_{t-2} + \dots + y_{t-n})} \quad (12)$$

En la segunda fase, en la red neuronal, cada uno de los nodos está representado para una función determinada por las reglas *If-Then* (8) o (10). Esto significa que el nodo es la multiplicación de las diferentes señales y su suma es la salida.

$$f = w_1 z_{t-1} + w_2 z_{t-2} + \dots + w_n z_{t-n} \quad (13)$$

En la tercera fase, la salida del paso anterior se transforma a través de una función sigmoideal, así se tiene, la capa oculta se determina por las reglas *If-Then* (9) o (11).

$$f_s(f) = \frac{1}{1 + e^{-f}} \quad (14)$$

Por tanto, la capa oculta es denotada por la siguiente ecuación:

$$f_l = w_1 f_s + w_2 f_s \quad (15)$$

En el último nodo se realiza la suma de todas las señales de salida de la tercera fase y la variable de salida  $\hat{y}_t$  se representa como:

$$\hat{y}_t = f_l * \sigma(y_{t-1} + y_{t-2} + \dots + y_{t-n}) + \mu(y_{t-1} + y_{t-2} + \dots + y_{t-n}) \quad (16)$$

donde  $\hat{y}_t$  es el pronóstico de  $y_t$ ,  $\sigma$  es la varianza de los valores pasados y  $\mu$  representa la media de los valores pasados.

La función de costos se define por medio de la desviación media absoluta, tal como:

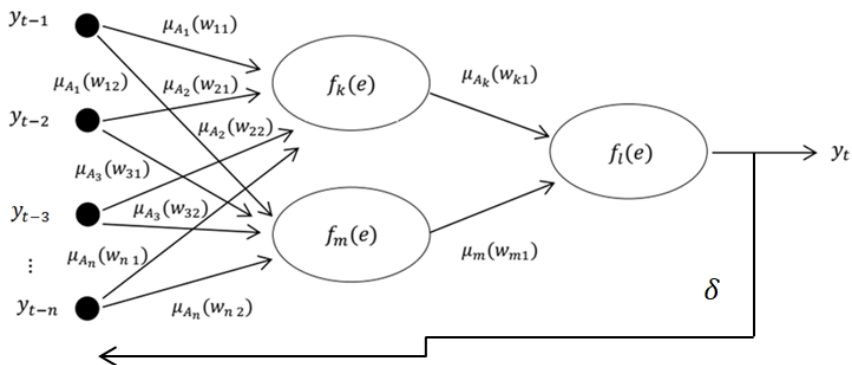
$$\epsilon = \frac{E|y_t - \hat{y}_t|}{n} \quad (17)$$

El algoritmo de aprendizaje de la red adaptativa basado en la inferencia difusa, está dado por las reglas *If-Then* y la observación de los parámetros de los modelos, lo que puede ser expresado por una combinación lineal de resultados generales expresado de la siguiente manera:

$$f = \bar{w}_1(p_1 y_t + r_1) + \bar{w}_2(p_2 y_t + r_2)$$

La principal ventaja de las redes neuronales difusas en relación con la red neuronal autorregresiva es que la red difusa permite decidir, además de los rezagos de las series temporales, la membresía de la volatilidad. En otras palabras, el método propuesto permite definir el aprendizaje de la red neuronal tanto en los patrones del pasado como en la función de membresía de la serie temporal.

Gráfica 3. Estructura del modelo Hybrid Fuzzy Nonlinear Autoregressive Neural Network



Fuente: elaboración propia.

### 3. Pronóstico de la Red Neuronal Autorregresiva Difusa Híbrida

Las Redes Neuronales Autorregresivas Difusas Híbridas se aplicaron a los datos diario de cuatro índices bursátiles: IPC, IBEX 35, S&P 500 y el Nikkei 225, desde enero de 2015 hasta diciembre de 2018, anexando los primeros cinco días de enero de 2019 para el análisis fuera de muestra (Yahoo!, 2019). Se estimó una Red Neural Autorregresiva No Lineal con tres rezagos, y un algoritmo de aprendizaje Bayesiano (NARNET), y se comparó con los resultados obtenidos con las redes neuronales difusas que se proponen en esta investigación.

El Cuadro 1, muestra los estadísticos generales de la rentabilidad de los índices bursátiles analizados, los resultados muestran que las cuatro variables tienen exceso de curtosis y distribución leptocúrtica, además la desviación estándar diaria está alrededor de 1% y la asimetría es negativa. De esta manera, las estadísticas indican que los valores extremos tienen una alta probabilidad de ocurrir, por lo que el pronóstico es aún más complicado de realizar. Por otro lado, la volatilidad en esta investigación se mide por la rentabilidad, y el modelo propuesto intenta estimar su comportamiento.

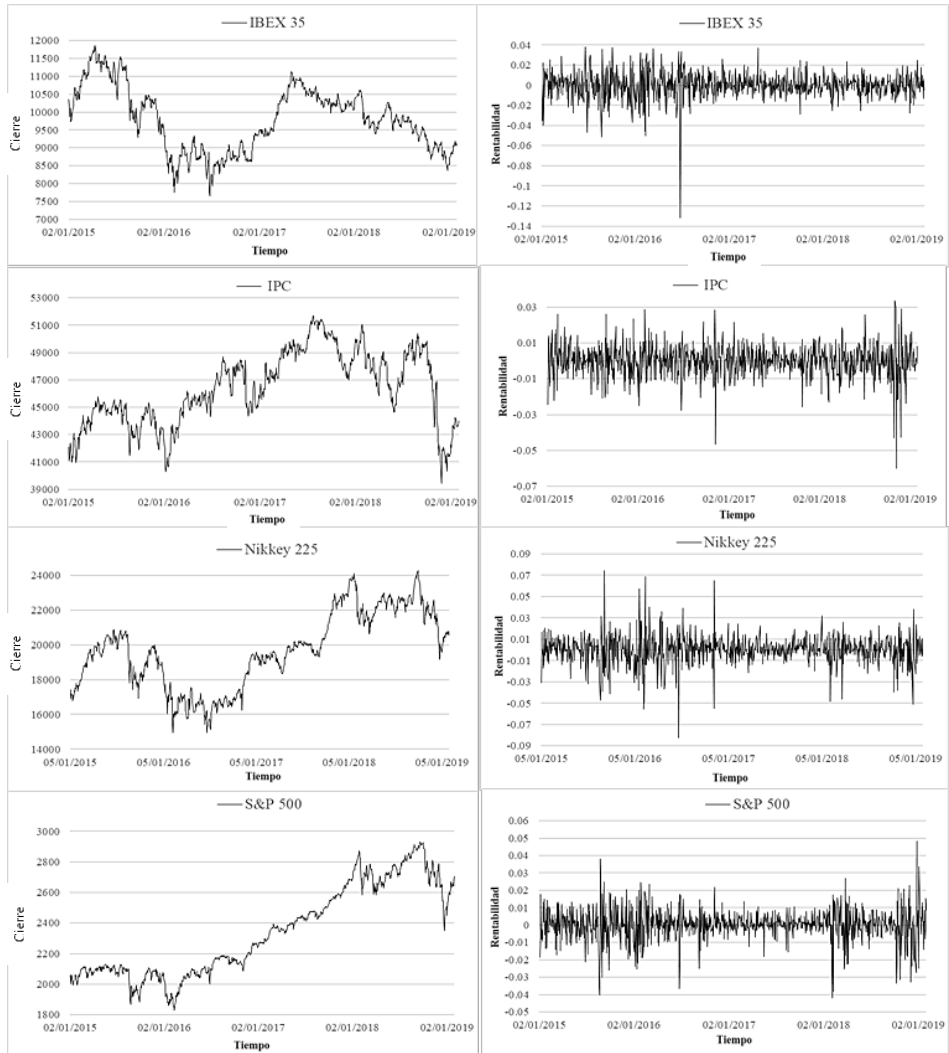
Los resultados se han compilado e ilustrado en una serie de gráficos y cuadros que se muestran en la Gráfica 4. La estructura de estos elementos se distribuye para tener en cuenta los índices bursátiles de México, Japón, EE. UU. y España. La razón para hacer esto, es identificar qué metodología puede modelar mejor la volatilidad de los diversos mercados. En este sentido, es importante reconocer los tres principales mercados del mundo, el europeo, el americano y el asiático (representados por un índice de cada continente).

Cuadro 1. Estadísticos generales de la rentabilidad de los índices bursátiles

Estadísticos \ índice Bursátil	IPC	IBEX 35	S&P 500	Nikkei 225
Media	0.0042%	-0.0129%	0.0266%	0.0177%
Desviación Estandar	0.86%	1.21%	0.87%	1.29%
Sesgo	-0.63	-1.39	-0.45	-0.34
Curtosis	4.6	14.29	3.88	5.58

Fuente: elaboración propia en Excel y MatLab con información de Yahoo Finance.

Gráfica 4. Cotización de cierre diario y rentabilidad de los índices bursátiles



Fuente: elaboración propia en Excel con información de Yahoo Finance.

La relación existente entre las variables económicas y los eventos sociales y políticos; causan cambios repentinos en el comportamiento de las variables financieras, y al mismo tiempo el incremento de la incertidumbre (medida por la volatilidad) ocasionando que el pronóstico sea cada vez más complicado (Rossi, 2013).

El Cuadro 2 ilustra los subconjuntos difusos de cada índice bursátil o las series temporales difusas asociadas a las variables. Las subsecciones (a, b, c y d) indican la función de pertenencia trapezoidal difusa del IPC,

Cuadro 2. Subconjunto Difusos

		IPC		Nikkei 225		
		$L_{BD}$	$U_{BD}$	$L_{BD}$	$U_{BD}$	
a)	A1	-0.06	-0.01	A1	-0.09	-0.01
	A2	-0.01	0	A2	-0.01	0
	A3	0	0.010	A3	0	0.010
	A4	0.01	0.04	A4	0.01	0.09
b)			IBEX 35		S&P 500	
			$L_{BD}$	$U_{BD}$	$L_{BD}$	$U_{BD}$
	A1	-0.06	-0.01	A1	-0.06	-0.01
	A2	-0.01	0	A2	-0.01	0
A3	0	0.010	A3	0	0.01	
A4	0.01	0.04	A4	0.01	0.05	
e)			IPC		Nikkei 225	
			$L_{BD}$	$U_{BD}$	$L_{BD}$	$U_{BD}$
	A1	-0.06	-0.01	A1	-0.07	-0.01
A2	-0.01	0.012	A2	-0.01	0.01	
A3	0.012	0.04	A3	0.01	0.02	
g)			IBEX 35		S&P 500	
			$L_{BD}$	$U_{BD}$	$L_{BD}$	$U_{BD}$
	A1	-0.06	-0.01	A1	-0.06	-0.01
A2	-0.01	0	A2	-0.01	0.005	
A3	0	0.02	A3	0.005	0.04	

Fuente: elaboración propia.

Nikkey 225, IBEX 35 y el S&P 500, respectivamente, y las subsecciones (e, f, g y h) indican la función de pertenencia triangular difusa de la serie temporal. Hay que destacar que los subconjuntos de mayor amplitud en sus subconjuntos difusos son los de mayor incertidumbre (como lo indican sus estadísticos).

En el Cuadro 2 se define la clasificación de la volatilidad asociada a los índices bursátiles. Por ejemplo, la función de membresía trapezoidal IPC tiene cuatro pasos de comportamiento (subsección a) el primero corresponde a los subconjuntos  $[-0.06, -0.01]$  asociado a volatilidad negativa (malas noticias),  $[-0.01, 0]$  negativo estable (insignificante pero noticias negativas),  $[0, 0.01]$  positivo estable (insignificante pero noticias positivas) y  $[0.01, 0.04]$  positivo (buenas noticias); este subconjunto representa el aprendizaje clasificado de la *NARNET Trapezoidal Difusa* de cuatro grados.

La subsección e define los subconjuntos difusos del índice Nikkei 225,  $[-0,07, -0,01]$  asociados a volatilidad negativa (malas noticias),  $[-0,01, 0,01]$  estable (noticias no significativas) y  $[0,01, 0,02]$  positivo (buenas noticias); este subconjunto muestra el aprendizaje clasificado de la *Fuzzy Triangular NARNET* de tres grados. La interpretación de los conjuntos difusos asociados a otros índices bursátiles es la misma, la única diferencia son las cotas superior e inferior de cada subconjunto difuso.

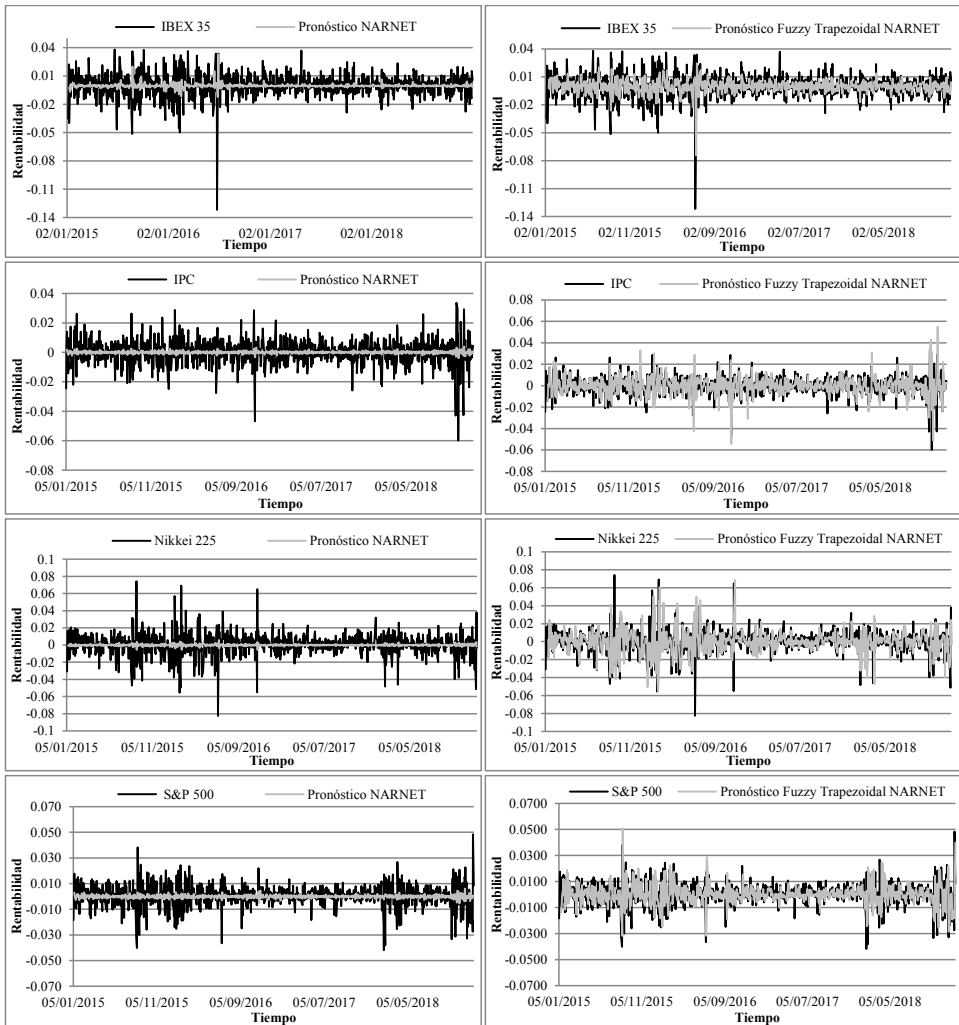
La Gráfica 5 muestra la comparación entre el pronóstico de la *NARNET Trapezoidal Difusa* (lado derecho) y la *NARNET tradicional* (lado izquierdo), en ambos casos, la línea gris claro representa el pronóstico y la línea negra la rentabilidad. En el caso del IBEX 35, el método tradicional señala, que la estimación no captaba la alta volatilidad de esta variable, por otro lado, la red difusa genera un mejor pronóstico; como consecuencia de que se acerca más al comportamiento real. Esta situación es similar en el índice bursátil IPC, S&P 500 y el Nikkei 225, condición que se puede corroborar en el Cuadro 2, el cual indica que el error medio absoluto de la *NARNET Trapezoidal Difusa* es 0.13% menos que el modelo NARNET para el S&P 500, la red difusa del IBEX 35 tiene un error menor de 0.22%, el pronóstico difuso del IPC tiene un error inferior al 0.16% del pronóstico tradicional y por último la *NARNET Trapezoidal Difusa* aplicada al Nikkei tiene un error diario de 0.21% menor que la *NARNET tradicional*.

La Gráfica 6 indica la comparación de la *NARNET Triangular Difusa* (lado izquierdo) y la *NARNET tradicional* (lado derecho), en ambos casos representada por la línea gris claro. La predicción del IBEX 35 señala que el método propuesto genera una mejor estimación que la red tradicional,



esto se puede ver en la Gráfica 6 en que la predicción difusa se ajusta mejor al índice bursátil español que el modelo no-difuso. Una situación similar se da con el Índice Mexicano, el Índice de Estados Unidos y el Índice de la Bolsa de Valores de Japón.

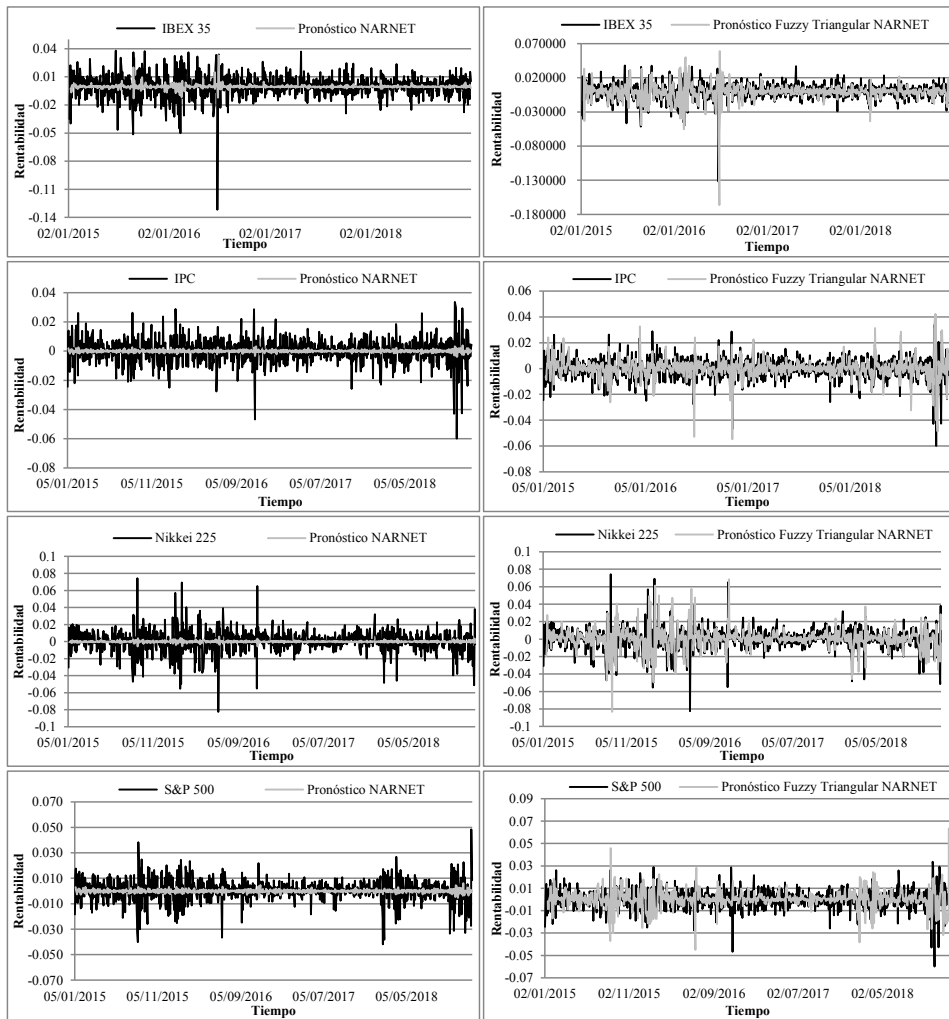
Gráfica 5. Comparación del pronóstico de los modelos NARNET Trapezoidal Difuso y NARNET tradicional



Fuente: elaboración propia en Excel y MatLab con información de Yahoo Finance.

Estos resultados muestran que el aprendizaje clasificado de las redes propuestas mejora significativamente la predicción de las variables de alta volatilidad, ya que al incorporar una función de membresía triangular al aprendizaje de las redes neuronales autorregresivas, se identifican tres

Gráfica 6. Comparación del pronóstico de los modelos NARNET Triangular Difusa y la NARNET tradicional



Fuente: elaboración propia en Excel y MatLab con información de Yahoo Finance.

grados de membresía de la volatilidad de las series temporales financieras, es decir, se obtienen los pesos sinápticos que modelan el comportamiento de la variable con un entorno positivo (buenas noticias), estable (noticias no significativas) y uno negativo (malas noticias); permitiendo así obtener mejores predicciones. Por otro lado, al incorporar la función de membresía trapezoidal, se clasifican cuatro grados de volatilidad, un entorno positivo (buenas noticias), estable positivo (noticias insignificantes pero positivas), estable negativo (noticias insignificantes pero negativas) y negativo (malas noticias); generando una mejor adaptación a la incertidumbre de la serie temporal.

La principal diferencia de los modelos propuestos es que depende de los grados de incertidumbre asociados a las series de tiempo financieras que las redes neuronales aprenderán mejor con una función triangular o trapezoidal. Por lo tanto, según el grado de volatilidad, se debe elegir la función de membresía que mejor se adapte a las condiciones de mercado estudiado.

El Cuadro 3 ilustra la comparación de la desviación media absoluta para los cuatro índices analizados, el principal resultado en relación con las redes difusas y la NARNET es que los modelos difusos tienen un error más bajo que el modelo tradicional, por ejemplo, el pronóstico difuso del S&P 500 tiene un 0.16% menor de error, el pronóstico difuso del IBEX 35 tiene un error 0.105 más bajo, en el pronóstico difuso del IPC el error es 0.14 más bajo que el de la red tradicional y el pronóstico difuso del Nikkei tiene un error 0.13% menor, para datos diarios en todos los casos.

Es importante reconocer que con la metodología propuesta se obtiene una mejora significativa en términos de reducción del error con respecto a los modelos tradicionales de redes neuronales y según (Medina-Reyes,

Cuadro 3. Comparación de la desviación media absoluta diaria

Índice Bursátil	<i>NARNET</i>	<i>NARNET</i>	<i>NARNET</i>
	<i>Trapezoidal Difusa</i>	<i>Triangular Difusa</i>	
IPC	0.4650%	0.4670%	0.6302%
IBEX 35	0.6146%	0.7392%	0.8378%
S&P 500	0.4626%	0.5814%	0.6288%
Nikkei 225	0.6749%	0.7594%	0.8870%

Fuente: elaboración propia en Excel.

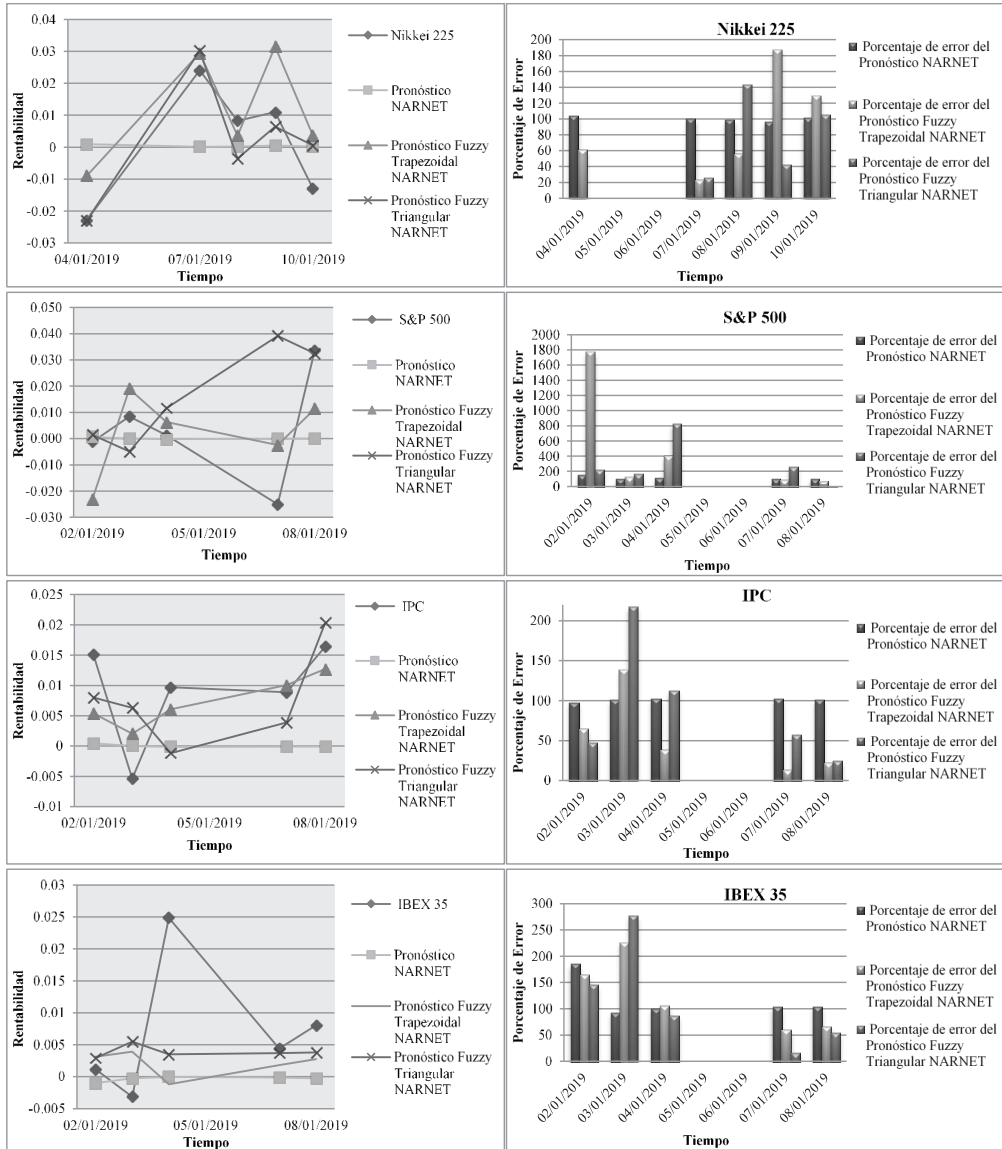
2019) esta mejora no sólo se da en las redes neuronales, sino también en los modelos lineales, (ARIMA) y no lineales (GARCH). Por lo tanto, los modelos propuestos modelan el comportamiento de las variables financieras mejor que los modelos y metodologías anteriores, mostrando que los modelos difusos son mejores que las metodologías mencionadas.

La Gráfica 7 muestra la comparación del modelo propuesto y la prueba de la red neuronal autorregresiva no lineal fuera de muestra durante los primeros cinco días cotizados por cada índice en 2019. En el lado derecho está el pronóstico del índice de acciones y en el lado izquierdo está el error porcentual de pronóstico, el pronóstico NARNET tradicional es la línea gris claro con marcadores cuadrados en los cuatro casos, se observa un pronóstico consistente en comparación con los valores reales (línea negra). Por otro lado, la línea gris claro es NARNET Trapezoidal Difusa, la cual genera un mejor pronóstico que la NARNET tradicional y en el caso del Nikkei 225 y el IPC, la línea gris claro tiene un comportamiento muy similar al valor real del índice bursátil. Es posible identificar que el pronóstico difuso se adapta mejor a las variaciones en la serie temporal de la muestra que la metodología tradicional de la red neuronal.

La Gráfica 7 también ilustra los resultados del pronóstico fuera de la muestra en escala de grises; la sección izquierda muestra el pronóstico de los 4 índices bursátiles. Se destaca que la predicción NARNET, denotada con la línea con marcadores cuadrados, tiende a ser una línea recta que pierde potencial de pronóstico a lo largo del tiempo. Por otra parte, un resultado destacable en la estimación fuera de muestra de los modelos basados en teoría difusa, indicados con las líneas con marcadores de triángulos y de equis, es que estos se adaptan mejor a la volatilidad de los mercados y generan buenos pronósticos en comparación con los modelos de redes bayesianas.

La *NARNET Triangular Difusa* genera una mejor estimación que el modelo tradicional en el Nikkei 225 y el IPC, pero para el S&P 500 y el IBEX 35 todos los modelos no tienen un buen pronóstico. Esto se puede apreciar en el porcentaje de error diario (gráfico de la izquierda) donde el error del índice de EE. UU. y España fue mayor que el de Japón y México. En este gráfico, algunos días no tienen valores debido a que para esos días el mercado no generó precios. La *NARNET Triangular Difusa* no generó error en el primer día de pronóstico en el Nikkey 225 y en los índices bursátiles fue el mejor modelo, pero en los días siguientes el error aumentó para los tres modelos. El S&P 500 tiene una alta volatilidad y los modelos no capturan su comportamiento,

Gráfica 7. Pronóstico fuera de muestra y porcentaje de error diario



Fuente: elaboración propia en Excel con información de Yahoo Finance.

pero el error disminuye en las predicciones de los siguientes días, la misma situación ocurre en el IBEX.

Por último, los modelos propuestos generan un mejor pronóstico dentro de muestra y fuera de muestra que la Red Neuronal Autorregresiva No Lineal. Por lo tanto se concluye que las redes neuronales pueden aprender de la dinámica de las series temporales y si se añade la teoría difusa también pueden aprender de la incertidumbre de las variables financieras. Esta situación hace que el método propuesto se considere mejor que la red tradicional.

## Conclusión

Los modelos tradicionales de redes neuronales son generalmente aceptados en la ciencia económica por su gran contribución al estudio y pronóstico de la no linealidad de las series temporales financieras; en la presente investigación, el análisis se realiza con la metodología de las redes neuronales autorregresivas basadas en teoría difusa que incorpora en su estructura las funciones de membresía triangular y trapezoidal. Los resultados sugieren que la metodología propuesta genera una mejor estimación de la volatilidad de los índices bursátiles IBEX 35, S&P 500, IPC y Nikkei 225 en comparación con las redes neuronales tradicionales.

La mejora en la estimación de los modelos presentados se atribuye a la incorporación de un elemento muy relevante dentro de la teoría difusa, es decir, la función de membresía. En estos métodos se considera que la función de membresía como **clasificadora** de la volatilidad juega un papel de suma importancia para el pronóstico de las series temporales, apoyado específicamente en la identificación de los grados de incertidumbre asociados a las diversas condiciones económicas, políticas y sociales que pueden impactar en el comportamiento del mercado de valores.

El principal resultado obtenido es que la aplicación de la teoría difusa dentro de la metodología tradicional de redes neuronales constituye una contribución relevante en el estudio de la no linealidad de las series de tiempo financieras, esto se atribuye a que los modelos propuestos reconocen mejor la incertidumbre de las variables con gran volatilidad en su comportamiento, lo anterior se debe a que la combinación de ambas metodologías ocasiona un aprendizaje clasificado de la incertidumbre existente en las series temporales para lograr así pronósticos más precisos.

Un aspecto importante por destacar es que, al incorporar una función de membresía triangular al aprendizaje de las redes neuronales autorregresivas, se identifican tres grados de membresía de la volatilidad en las series de tiempo financieras. Es decir, se obtienen los pesos sinápticos que modelan el comportamiento de la variable con un entorno positivo (buenas noticias), estable (noticias no significativas) y negativo (malas noticias); permitiendo así obtener mejores predicciones. Por otro lado, al incorporar la función de membresía trapezoidal, se clasifican cuatro grados de volatilidad, un entorno positivo (buenas noticias), estable positivo (noticias insignificantes pero positivas), estable negativo (noticias insignificantes pero negativas) y negativo (malas noticias); generando una mejor adaptación a la incertidumbre de las series temporales. La principal diferencia entre los modelos propuestos consiste en que, dependiendo de la incertidumbre asociada a una serie temporal financiera, las redes neuronales aprenderán mejor con una función triangular o trapezoidal.

Otro hallazgo es que al realizar la comparación con investigaciones precedentes sobre el tema, la incorporación de la función de membresía en las redes neuronales se adapta de forma diferente, porque en lugar de suponer que la serie temporal tiene una función de membresía, la función se utiliza como clasificadora de la volatilidad, lo que permite variar los perceptrones para mejorar el aprendizaje de la no linealidad de las series temporales financieras.

Estos modelos no sólo pueden predecir el índice del mercado de valores, si no también otras variables financieras como los tipos de cambio, el flujo de caja, los tipos de interés, las acciones y otras variables multidisciplinarias con la característica de alta volatilidad en su comportamiento.

## Referencias bibliográficas

- Cox, E. (2019). *The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, and Maintaining Fuzzy Systems*/Book and Disk.
- Dash, R., y Dash, P. (2016). "An Evolutionary Hybrid Fuzzy Computationally Efficient EGARCH Model for Volatility Prediction". *Applied Soft Computing*, pp. 40-60.
- Egrioglu, E., Aladag, C. H., y Yolcu, U. (2013). "Fuzzy Time Series Forecasting With a Novel Hybrid Approach Combining Fuzzy c-means and Neural Networks". *Expert Systems with Applications*, pp. 854-857.

- Egrioglu, E., Aladag, C. H., Yolcu, U., Basaran, M. A., y Uslu, V. R. (2009). "A New Hybrid Approach Based on SARIMA and Partial High Order Bivariate Fuzzy Time Series Forecasting Model". *Expert Systems with Applications*, vol. 36, núm 4, pp. 7424-7434.
- García, J., López, A. M., Romero, J., García, A. R., Camacho, C., Cantero, J. L., Salas, R. (2002). "Hojas de cálculo para la simulación de redes neuronales artificiales". *Questtió*, núm. 26, pp. 289-305.
- Han, M., Zhong, K., y Han., B. (2018). "Interval Type-2 Fuzzy Neural Networks for Chaotic Time Series Prediction: A Concise Overview". *IEEE Transactions on Cybernetics*, pp. 1-12.
- Jang, J. R. (1993). "ANFIS: Adaptive-Network-Based Fuzzy Inference System". *IEEE Transactions on Systems*, pp.665-685.
- Medina-Reyes, J. E. (2019). "Studies of Fuzzy Time Series Models: An Application to the Forecast Exchange Rate, Mexican Peso/ American Dollar". *Research-Gate. (IPN Ed.)*, Thesis for: Master of Science, Advisor: Salvador Cruz Aké; Agustín Ignacio Cabrera Llanos, DOI:10.13140/RG.2.2.21331.53288.
- Medina-Reyes, J. E., Cabrera-Llanos, A. I., y Cruz-Aké, S. (2019). "FUZZY GARCH and FUZZY EGARCH Models with Gaussian parameters for forecast the Foreign Exchange Market". *IX FIMEF International Financial Research Conference* (p. 14). Merida, Yucatán: Fundación de Investigación IMEF .
- Mohammadian, M. (2020). "Modeling, Control, and Prediction using Hierarchical Fuzzy Logic Systems: Design and Development. In *Robotic Systems: Concepts, Methodologies, Tools, and Applications*" (pp. 187-207). IGI Global. Pal, S. S., y Kar, S. (2017). "Fuzzy Time Series Model for Unequal Interval Length Using Genetic Algorithm". *Advances in Intelligent Systems and Computing*, pp. 205-216.
- Pal, S. S., y Kar, S. (2018). "A Hybridized Forecasting Method Based on Weight Adjustment of Neural Network Using Generalized Type-2 Fuzzy Set". *International Journal of Fuzzy Systems*, pp. 1-13.
- Rossi, G. D. (2013). "La Volatilidad en Mercados Financieros y de Commodities. Un Repaso de sus Causas y la Evidencia Reciente". *Invenio*, pp. 59-74.
- Singh, P. (2017). "A Brief Review of Modeling Approaches Based on Fuzzy Time Series". *International Journal of Machine Learning and Cybernetics*, 8(2), pp. 397-420.
- Song, Q., y Chissom, B. S. (1993a). "Fuzzy Time Series and its Models". *Fuzzy Sets and Systems*, núm. 54, pp. 269-277.
- Song, Q., y Chissom, B. S. (1993b). "Forecasting Enrollments With Fuzzy Time Series — Part I". *Fuzzy Sets and Systems*, núm. 54, pp. 1-9.



- Song, Q., y Chissom, B. S. (1994). "Forecasting Enrollments With Fuzzy Time Series Part II". *Fuzzy Sets and Systems*, núm. 62, pp. 1-8.
- Souza, P. V., y Torres, L. C. (2018). "Regularized Fuzzy Neural Network Based on Or Neuron for Time Series Forecasting". *Springer International Publishing AG*, part of Springer Nature, pp. 13-23.
- Srinivasan, y Lakshmi. (2019). "Stock Price Prediction Using Fuzzy Time-Series Population-Based Gravity Search Algorithm". *International Journal of Software Innovation*, pp. 50-64.
- Tsaur, R.-C. (2012). "A Fuzzy Time Series-Markov Chain Model with an Application to Forecast the Exchange Rate Between the Taiwan and US Dollar. *International Journal of Innovative Computing*". *Information, and Control*, núm. 8, pp. 4931-4942.
- Tseng, F.-M., y Tzeng, G.-H. (2002). "A Fuzzy Seasonal ARIMA Model for Forecasting. *Fuzzy Sets and Systems*, núm. 126, pp. 367-376.
- Tseng, F.-M., Tzeng, G.-H., Yu, H.-C., y Yuan, B. J. (2001). "Fuzzy ARIMA Model for Forecasting the Foreign Exchange Market". *Fuzzy Sets and Systems*, núm. 118, pp. 9-19.
- Wu, S., y Er, M. J. (2000). "Dynamic Fuzzy Neural Networks; A Novel Approach to Function Approximation". *IEEE Transactions on Systems, Man, and Cybernetic*
- Xiao, Q. (2017). "Time Series Prediction Using Bayesian Filtering Model and Fuzzy Neural Networks". *International Journal for Light and Electron Optics*, pp. 1-20.
- Yahoo (2019). *Yahoo! Finance*. Obtained from: <https://finance.yahoo.com/>.
- Yu, H.-K. (2005). "Weighted Fuzzy Time Series Models for TAIEX Forecasting". *Physica A: Statistical Mechanics and its Applications*, núm. 349, pp. 609-624.
- Yu, T. H.-K., y Huarng, K.-H. (2010). "A Neural Network-Based Fuzzy Time Series Model to Improve Forecasting". *Expert Systems with Applications*, pp. 3366-3372.