

Pronósticos mediante redes neuronales artificiales y modelos ARIMA: el caso de los cetes en México

ARTURO MORALES CASTRO

IVÁN CRUZ TORRES*

La importancia de conocer el comportamiento de determinado activo financiero o instrumento derivado en el futuro permite tomar previsiones para evitar riesgos innecesarios o tener la oportunidad de obtener mayores beneficios. Las técnicas de pronóstico de series de tiempo con métodos estadísticos tradicionales posibilitan llevar a cabo esta labor. La metodología Box-Jenkins permite obtener buenas aproximaciones, ya que el proceso es iterativo y conduce al mejor modelo de acuerdo con la experiencia del investigador.

En general, hay dos grupos de modelos para pronosticar series de tiempo aplicadas a las finanzas: por un lado, los modelos estadísticos tradicionales como los ARIMA (siglas en inglés de modelos autorregresivos integrados de medias móviles); los modelos de heterosedasticidad condicional autorregresiva generalizada (GARCH, por sus siglas en inglés), y los modelos mixtos; por el otro, los modelos basados en áreas nuevas como la inteligencia artificial, las redes neuronales y los algoritmos genéticos.

El objetivo de la presente investigación es pronosticar la tendencia del Certificado de Tesorería (cete) con vencimiento en 91 días en el mercado mexicano de derivados (Mexder) mediante el uso de redes neuronales artificiales (RNA). De acuerdo con la mayor parte de la bibliografía, se espera un mejor desempeño de las RNA en comparación con los resultados del modelo ARIMA, que se utiliza como referencia. La diferencia es que las redes neuronales son capaces de identificar relaciones no lineales que, como se ha demostrado en

* División de Estudios Profesionales, Área de Finanzas, Facultad de Contaduría y Administración, Universidad Nacional Autónoma de México (UNAM) <amcastro@correo.fca.unam.mx>, y Facultad de Contaduría y Administración, UNAM <ivanct@yahoo.com>, respectivamente.

la mayoría de las investigaciones, se presentan en los mercados financieros.

La originalidad de esta investigación consiste en utilizar las RNA para predecir la tendencia (y no el precio) de los instrumentos financieros derivados. Además, al comparar este trabajo con los que hay para el caso de México, se utiliza una RNA propia (diseñada y construida), lo cual es nuevo entre los análisis sobre el tema.

REVISIÓN DE LA BIBLIOGRAFÍA

Las redes neuronales artificiales (RNA) son modelos matemáticos que funcionan como los sistemas nerviosos biológicos. De esta manera se intenta contar con un sistema que imite el funcionamiento natural de las neuronas; es decir, que pueda pensar y aprender como el cerebro humano, como un conjunto interconectado de neuronas.

La primera interpretación matemática de las redes neuronales fue hecha en 1865, cuando Mach, físico y filósofo austriaco, intentó explicar la interacción recíproca e inhibitoria de los elementos que rodean la retina.¹ Los trabajos sobre el desarrollo de modelos matemáticos detallados comenzaron desde hace más de 50 años, con los estudios de McCulloch y Pitts, Hebb, y Rosenblatt.² Sin embargo, cuando se utilizó un conjunto de datos con relaciones complejas y condiciones adicionales, la capacidad de procesamiento requerida se elevó de manera considerable, lo cual ocasionó que los primeros intentos por utilizar este tipo de modelos fueran abandonados debido a la limitada capacidad de los sistemas de cómputo disponibles en la época.

Debido a los avances en la potencia de cálculo de las nuevas computadoras y a un mayor conocimiento del funcionamiento del cerebro humano, en el decenio de los ochenta se retomó el interés por estos modelos. Este nuevo interés se debió al desarrollo de algoritmos y nuevas topologías de red, nuevas técnicas de implantación de circuitos analógicos y por algunas demostraciones

intrigantes por parte de Sejnowski y Hopfield,³ así como por la creciente fascinación por el funcionamiento del cerebro humano.

Las aplicaciones de las RNA en los negocios se dividen en dos categorías principales: la clasificación y el modelado.⁴ La clasificación se refiere a la manera de discriminar entre observaciones con características comunes en diferentes grupos; por ejemplo, la predicción de fallas corporativas, la asignación de créditos y la clasificación de bonos. El modelado consiste en simular el comportamiento o las reacciones de una entidad o variable ante un estímulo externo y está basado en observaciones pasadas; por ejemplo, predecir los movimientos de los precios de las acciones o fluctuaciones en el tipo de cambio.

El primer artículo sobre redes neuronales que considera información financiera fue realizado por White, quien estudió la predicción de los precios de las acciones con un modelo de red neuronal.⁵ Posteriormente la predicción de la cotización de las acciones, del tipo de cambio y de otras variables económicas y financieras ha sido uno de los temas más atractivos para los investigadores y analistas financieros.

Otros autores que han incursionado en la aplicación de las redes neuronales artificiales en las finanzas son: Dutta y Shekar; Bosarge; Trippi; Hawley *et al.*; Salchenberger, Cinar y Lash, y Tam y Kiang.⁶

Por su parte, Hill, O'Connor y Remus comparan el desempeño de las predicciones de series de tiempo basadas en redes neuronales con métodos estadísticos tradicionales (incluyendo el suavizamiento exponencial y

1. M. Peel y N. Wilson, "Neural Network Simulation: A New Approach to Risk Assessment and Business Forecasting", *Management Research News*, vol. 19, núm. 6, 1996, pp. 50-54.

2. W.S. McCulloch y W. Pitts, "A Logical Calculus of the Ideas Imminent in Nervous Activity", *Bulletin of Mathematical Biophysics*, vol. 5, 1943, pp. 115-133; D.O. Hebb, *The Organization of Behavior*, John Wiley & Sons, Nueva York, 1949, y R. Rosenblatt, *Principles of Neurodynamics*, Spartan Books, Nueva York, 1959.

3. T. Sejnowski y C.R. Rosenberg, "NETalk: A Parallel Network that Learns to Read Aloud", Technical Report JHU/EECS-86/01, Johns Hopkins University, 1986, y J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proc. Natl. Acad. Sci.*, vol. 79, abril de 1982, pp. 2554-2558.

4. M. Peel y N. Wilson, *op. cit.*

5. H. White, "Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns", IEEE International Conference on Neural Networks, 1988, pp. 451-459.

6. S. Dutta y S. Shekar, "Bond Rating: A Non-conservative Application of Neural Networks", Proceedings of the IEEE International Conference on Neural Networks, 1988, pp. 443-450; W. Bosarge, *A Non Random Walk Down Wall Street*, Princeton University Press, Princeton, 1989, p. 45; R. Trippi, "Intelligent Systems for Investment Decision Making", en F. Fabozzi (ed.), *Managing Institutional Assets*, Harper and Row, Nueva York, 1990, p. 251; D. Hawley *et al.*, "Artificial Neural Systems: A New Tool for Financial Decision-making", *Financial Analysts Journal*, núm. 23, pp. 63-72; L. Salchenberger, E. Cinar y N. Lash, "Neural Networks: A New Tool for Predicting Thrift Failures", *Decision Sciences*, vol. 23, 1992, pp. 899-916, y K.Y. Tam y M.Y. Kiang, "Predicting Bank Failures: A Neural Network Approach", *Management Science*, vol. 38, núm. 7, 1992, pp. 926-947.

Box-Jenkins), y encontraron que, cuando se utilizaron datos trimestrales y mensuales, el modelo con redes neuronales tuvo predicciones significativamente mejores que los estadísticos tradicionales.⁷ Sin embargo, con datos anuales, ambos modelos fueron comparables. El estudio se llevó a cabo como si fuera parte de la conocida competición de Makridakis.⁸

Shang-Wu usa redes neuronales para realizar pronósticos y arbitraje sobre el futuro del índice accionario Nikkei.⁹ De acuerdo con los resultados empíricos, las redes neuronales superaron al modelo ARIMA en las predicciones y también proporcionaron al operador mayores ganancias por arbitraje que los modelos tradicionales, aun cuando se observa un menor intervalo de tiempo redituable.

Zhang y Berardi, quienes utilizan redes neuronales para predecir el tipo de cambio entre la libra británica y el dólar estadounidense,¹⁰ de manera específica usan un método de partición sistemático para construir conjuntos de redes neuronales. Encuentran que la aproximación básica conjunta creada con arquitecturas de red entrenadas con diferentes pesos iniciales aleatorios no es eficaz para mejorar la precisión de la predicción y que los modelos en conjunto de diferentes estructuras neuronales pueden mejorar de manera consistente las predicciones de una única red. Sus resultados también muestran que las redes conjuntas basadas en diferentes particiones de datos son más eficientes que aquellas desarrolladas con el conjunto de datos de entrenamiento completo en la predicción fuera de muestra.

Medeiros, Veiga y Pedreira comparan diferentes alternativas para modelar y predecir series de tiempo del tipo de cambio mensual.¹¹ Los modelos que proponen son autorregresivos de transición suave de neurocoeficientes, lineales autorregresivos y de caminata aleatoria. Encuentran que la no linealidad sólo es relevante en algunos periodos de la serie, en particular al inicio y al

final de la muestra, y que los modelos no lineales se desempeñan mejor sólo en los casos en que la linealidad está distribuida de manera uniforme. También concluyen que no hay diferencias significativas respecto a un modelo lineal actual.

Collantes compara las metodologías Box-Jenkins, ARIMA y la función de transferencia en dos aplicaciones: la primera sobre la serie de tiempo del número de nacimientos mensuales ocurridos en España en el periodo de enero de 1960 a diciembre de 1999 y la segunda sobre dos series de tiempo, el gasto de publicidad mensual y el número de ventas mensuales.¹² Se concluye que la metodología ARIMA se puede utilizar como herramienta de procesamiento previo de datos, considerando como entradas los retrasos involucrados en el modelo proporcionado por esa metodología. Asimismo, que los resultados de la predicción, tanto dentro como fuera de muestra por parte de la red neuronal propuesta, fueron superiores a los de las otras metodologías.

Haefke y Helmenstein analizan una estrategia de negocio que explota la diferencia en información que implican los principios de construcción de índices de mercado de diferentes acciones.¹³ Se trata de ganar ventaja competitiva sobre otros participantes de mercado, usando redes neuronales para predicciones de un día y generar señales de compra y venta de acuerdo con la regla de negociación. Para ilustrar cómo trabaja el sistema, se aplica al índice de comercio austriaco (ATX). Sus resultados muestran que las redes seleccionadas se desempeñan bien respecto a R^2 , y a RMSE y MAPE (siglas en inglés de raíz cuadrada de error medio y error medio absoluto porcentual, respectivamente). Encuentran una relación positiva y significativa con los datos dentro de muestra y que, en algunos casos, las correlaciones fuera de muestra no son significativas o tienen el signo incorrecto.

Por su parte, García diseñó un modelo no lineal para el análisis y la predicción de la serie de tiempo del precio externo del café colombiano utilizando redes neuronales artificiales y lo comparó con un modelo clásico de predicción lineal ARIMA.¹⁴ La red usa una capa oculta, y

7. T. Hill, M. O'Connor y W. Remus, "Neural Network Models for Time Series Forecast", *Management Science*, vol. 42, núm. 7, 1996, pp. 1082-1092.

8. S. Makridakis *et al.*, "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition", *Journal of Forecasting*, vol. 1, 1982, pp. 111-153.

9. Y. Shang-Wu, "Forecasting and Arbitrage of the Nikkei Stock Index Futures: An Application of Backpropagation Networks", *Asia-Pacific Financial Markets*, vol. 6, diciembre de 1999, pp. 341-354.

10. G.P. Zhang y V.L. Berardi, "Time Series Forecasting with Neural Network Ensembles: An Application for Exchange Rate Prediction", *Journal of the Operational Research Society*, vol. 52, 2001, pp. 652-664.

11. M. Medeiros, A. Veiga y C. Pedreira, "Modeling Exchange Rates: Smooth Transitions, Neural Networks, and Linear Models", *IEEE Transactions on Neural Networks*, vol. 12, núm. 4, julio de 2001, pp. 755-764.

12. J.V. Collantes, *Predicción con redes neuronales: comparación con las metodologías de Box y Jenkins*, tesis de maestría, Universidad de los Andes, 2001.

13. C. Haefke y C. Helmenstein, "Index Forecasting and Model Selection", *International Journal of Intelligent Systems in Accounting, Finance and Management*, vol. 11, núm. 2, abril-junio de 2004, pp. 119-135.

14. I. García, "Análisis y predicción de la serie de tiempo del precio externo del café colombiano utilizando redes neuronales artificiales", *Revista de la Facultad de Ciencias*, vol. 8, Pontificia Universidad Javeriana, Bogotá, 2003, pp. 45-50.

el algoritmo de retropropagación y una función sigmoidea como función de activación. Los resultados obtenidos indican que la varianza del error del modelo RNA es menor que la del modelo ARIMA aproximadamente en 22%, por lo que el modelo de redes neuronales se considera mejor.

Stansell y Eakins usan redes neuronales para predecir la dirección de cambio en índices de acciones de 19 sectores sobre un intervalo de cinco meses de 2001 y de tres de 2002.¹⁵ Los resultados fueron evaluados en términos de la dirección de cambio y del número de puntos ganados en el índice del sector. Concluyen que un inversionista puede obtener ventaja de la habilidad de predicción de las redes neuronales utilizando los datos económicos públicamente disponibles.

En lo que se refiere a la predictibilidad en las finanzas para México, se encontraron sólo dos investigaciones con RNA: una de Berumen y otra de Pérez.¹⁶ Sin embargo, estos investigadores utilizan un programa comercial de computadora o bien el desarrollo de un programa en lenguaje C++ para hacer predicciones del precio de las acciones, por lo que su objetivo no es conocer la predicción de la tendencia de los instrumentos financieros derivados, fin que persigue esta investigación.

DATOS Y METODOLOGÍA

Para construir un modelo ARIMA que se aproxime de manera aceptable a las características de una serie de tiempo, se utiliza la metodología Box-Jenkins.

Los datos pertenecen a la serie de cotizaciones de cierre del futuro de los certificados de tesorería (cetes) a 91 días en el mercado mexicano de derivados (Mexder) del 1 de noviembre de 2002 al 12 de abril de 2007 y que vencen en diciembre de 2008. Cada contrato de futuro sobre cetes a 91 días ampara una cantidad de 10 000 cetes, equivalente a un valor nominal de 100 000 pesos. El símbolo o clave de pizarra del contrato de futuro es

En general, el proceso de encontrar un modelo basado en redes neuronales es más complejo y delicado que el desarrollo de un modelo ARIMA



15. S. Stansell y S. Eakins, "Forecasting the Direction of Change in Sector Stock Indexes: An Application of Neural Networks", *Journal of Asset Management*, vol. 5, núm. 1, junio de 2004, pp. 37-48.

16. G. Berumen, *La teoría del caos y las redes neuronales aplicadas a las finanzas*, tesis de maestría, Universidad La Salle, México, 1998, y G. Pérez, *Las redes neuronales artificiales: una herramienta cibernética para la estimación de precios en el mercado de productos derivados*, tesis de licenciatura, Universidad La Salle, México, 1998.

CE91 DC08. La celebración del contrato de futuro de cetes a 91 días en Mexder tendrá como unidad de cotización de la tasa futura la tasa porcentual anualizada, expresada en porcentaje, con dos dígitos después del punto decimal. Los días en los que no hubo negociación en el mercado no se consideran; de esta manera, se obtienen 1 115 datos.

De acuerdo con Kaastra y Boyd, suavizar tanto los datos de entrada como los de salida, usando ya sea promedios móviles simples o exponenciales, puede ser la mejor manera de predecir la tendencia de una serie, en lugar de intentar predecir los cambios en precios alrededor de la tendencia, en cuyo caso lo mejor es modelar las series de tiempo financieras por medio de modelos ARCH (siglas en inglés de heterosedasticidad condicional autorregresiva).¹⁷ En este trabajo, se realiza un suavizado exponencial de los datos, de manera que se obtengan pronósticos sobre la tendencia de la serie en lugar de los valores originales, tanto para el modelo ARIMA como para el modelo con RNA. Como la serie tiene cierta tendencia, es necesario trabajar en diferencias; tomando logaritmos se reduce, además, la dispersión de la serie. El resultado de la serie así calculada se llamará rendimiento logarítmico o RCT.

OBTENCIÓN DEL MODELO DE REDES NEURONALES PARA LA SERIE CETES DEL MEXDER

Para introducir la serie de tiempo en la red neuronal es necesario adecuar los datos de tal modo que se puedan manipular con facilidad. Para ello se copian los datos en un vector columna A como el siguiente:

$$A = \begin{bmatrix} a_{t-n} \\ \dots \\ a_{t-3} \\ a_{t-2} \\ a_{t-1} \\ a_t \end{bmatrix}$$

A partir del vector anterior se genera la matriz de retrasos B_r que contiene los elementos de retraso que serán utilizados como entradas a la red neuronal artificial. La primera columna se considerará como el objetivo, mientras que las demás serán las correspondientes a los retrasos hasta un máximo de τ , que será fijado de acuerdo con los criterios mencionados en la determinación de las entradas.



$$B_r = \begin{bmatrix} a_{t-n} & 0 & 0 & \dots & 0 \\ a_{t-n+1} & a_{t-n} & 0 & \dots & 0 \\ a_{t-n+2} & a_{t-n+1} & a_{t-n} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{t-2} & a_{t-3} & a_{t-4} & \dots & a_{t-r-2} \\ a_{t-1} & a_{t-2} & a_{t-3} & \dots & a_{t-r-1} \\ a_t & a_{t-1} & a_{t-2} & \dots & a_{t-r} \end{bmatrix}$$

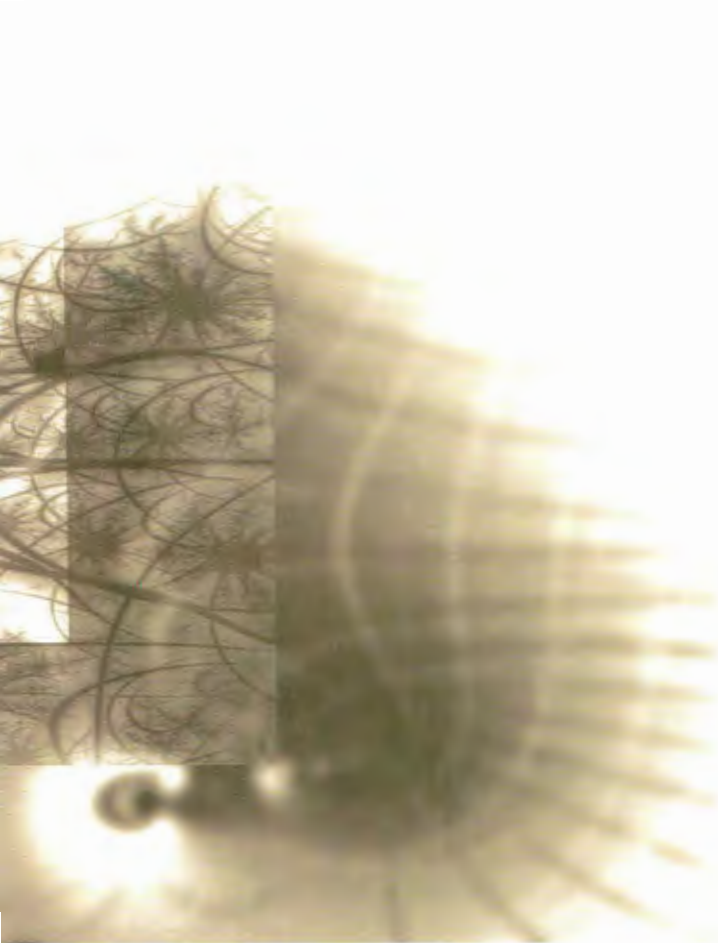
donde:

n es el número total de elementos de la serie, y r es el número de retrasos.

Como lo recomiendan Kallman y Wasny —quienes la consideran una función de transferencia ideal—, los datos de la serie se normalizan o transforman a valores en el intervalo $[-1, 1]$ para usar la función tangente hiperbólica cuyo codominio se encuentra en el mismo intervalo.¹⁸ Al terminar la predicción, los datos obtenidos se rescalan a su valor original invirtiendo la ecuación anterior. Los valores de la serie de tiempo se dividen en dos conjuntos de datos:

17. I. Kaastra y M. Boyd, "Designing a Neural Network for Forecasting Financial and Economic Time Series", *Neurocomputing*, núm. 10, 1996, pp. 215-236.

18. B.L. Kalman y S.C. Kwasny, "Why Tanh? Choosing a Sigmoidal Function", International Joint Conference on Neural Networks, Baltimore, Estados Unidos, 1992.



1) Un conjunto de entrenamiento, que constituye 80% de los datos de la serie que se seleccionan de manera consecutiva y ordenada. Este conjunto de datos es el que se utiliza para el entrenamiento de la RNA.

2) Un conjunto de prueba, formado por el conjunto de 20% de los datos restantes de la serie, una vez seleccionados los patrones de entrenamiento. Este conjunto de datos se utiliza para evaluar la capacidad de generalización o predicción de la red.

Los conjuntos de entrenamiento y prueba se obtienen a partir de la matriz B . En principio se eliminan los primeros renglones r de la matriz. A los elementos que no tienen retrasos se les asigna cero (dado que ya no hay más datos anteriores). Después, la matriz restante se divide en dos partes, una que contiene 80% de los renglones y otra el 20% restante, y que corresponden a los conjuntos de entrenamiento y prueba, respectivamente. De esta manera, para seleccionar un retraso específico y considerarlo en el entrenamiento basta con incluir la columna correspondiente a tal retraso. La primera columna será el objetivo por perseguir; es decir, lo que la red deberá aprender con base en los retrasos proporcionados.

Se propone una topología de red con las siguientes características:

Una red multicapa unidireccional o perceptrón multicapa (MLP); es decir, una red con conexiones que pa-

san las salidas en una sola dirección a neuronas en la siguiente capa; una red totalmente interconectada, en la que la salida de cada neurona es conectada a cada una de las otras neuronas de la siguiente capa; número de entradas: p ; número de capas ocultas: 1; número de neuronas en la capa oculta: q ; la función de transferencia o activación de las neuronas de la capa oculta: tangente hiperbólica, y la función de transferencia o activación de la capa de salida: lineal.

Como no hay una regla para especificar qué entradas son las mejores para hacer la predicción, se consideran dos opciones para determinar las entradas de la red neuronal. Se sugiere usar 10 retrasos y, como lo sugieren Faraway y Chatfield,¹⁹ analizar el peso de cada entrada para seleccionar las de mayor peso y comparar el error generado tanto en el ajuste como en la predicción con las entradas seleccionadas. Una vez determinado el modelo ARIMA, se eligen como entradas los retrasos correspondientes de este modelo (coeficientes estadísticamente significativos).

Al igual que en la determinación de las entradas, no hay una regla general para la obtención de los parámetros de la red que se desempeñe mejor en cada situación. Se realiza entonces un procedimiento automático de búsqueda, por medio de un *barrido* de los parámetros que determinan el comportamiento de la red. Para obtener los parámetros de la red neuronal que arrojan el menor error tanto en el ajuste como en la predicción, se realizan tres ciclos anidados que evalúan diversas redes en las que se establecen como parámetros de variación: el porcentaje de error máximo permitido de 0.01 a 0.1, con incrementos de 0.01; el número de neuronas en la capa oculta de 1 a 10 con incrementos de 1, y el número de retrasos también de 1 a 10 con incrementos de 1.

Debido a la gran cantidad de recursos de cómputo requeridos —sobre todo cuando el porcentaje de tolerancia al error es pequeño—, se evalúan dos redes por cada una de las combinaciones de parámetros cuando el error permitido va de 0.01 a 0.04, 50 redes cuando el error permitido es de 0.05, y 200 redes cuando el error permitido va de 0.06 a 0.1, para no evaluar sólo una red y evitar obtener una red que quede atrapada en un mínimo local que no permita decidir si la red es adecuada o no con los parámetros establecidos.

19. J. Faraway y C. Chatfield, "Time Series Forecasting with Neural Networks: A Comparative Study Using the Airline Data", *Applied Statistics*, vol. 47, núm. 2, 1989, pp. 231-250.

Mediante el ciclo de cálculos mencionado, se elige la red que presenta el menor error promedio de la suma del cuadrado de los errores, tanto en el ajuste o etapa de entrenamiento (SSE-E) como en la etapa de predicción (SSE-P). La red obtenida de este modo tiene seis neuronas ocultas, seis retrasos y un porcentaje de error permitido de 0.06.

Se propone utilizar el algoritmo de *retropropagación*, que es por mucho el más popular y más usado por los investigadores en redes neuronales artificiales, además de estar implantado en casi todos los paquetes de programas comerciales de computadora. Su popularidad se debe a la simplicidad en el diseño e implantación. La red neuronal se entrena en particular con el algoritmo Levenberg-Marquardt de Matlab, que es el método más rápido para entrenar redes de hasta varios cientos de pesos.

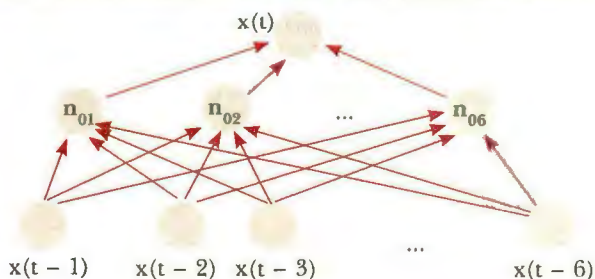
La elección de los pesos iniciales es importante y se recomienda probar con diferentes conjuntos de valores iniciales para obtener los mejores resultados. Los pesos iniciales se generan de manera aleatoria 50 veces y se selecciona el modelo que tenga menor promedio entre la suma de cuadrados de los errores de ajuste y predicción.²⁰

En redes neuronales, la función de error más común por minimizar es la suma de los cuadrados de los errores. Por medio de la ecuación de predicción definida por la red neuronal, se obtiene el valor de la predicción en el tiempo $t + 1$. Se calcula entonces el error de generalización.

A partir de los errores de entrenamiento y generalización se comparan las redes obtenidas y se selecciona aquella en la que ambos valores son mínimos. La arquitectura del modelo de red calculada se puede observar en el diagrama.

20. *Ibid.*

ARQUITECTURA DE LA RED NEURONAL ARTIFICIAL QUE PROPORCIONA LOS MEJORES RESULTADOS



PRINCIPALES RESULTADOS EMPÍRICOS Y HALLAZGOS PRELIMINARES

Una vez seleccionado el modelo, se puede pasar a la etapa de predicción; la gráfica del pronóstico se muestra en la gráfica 1.

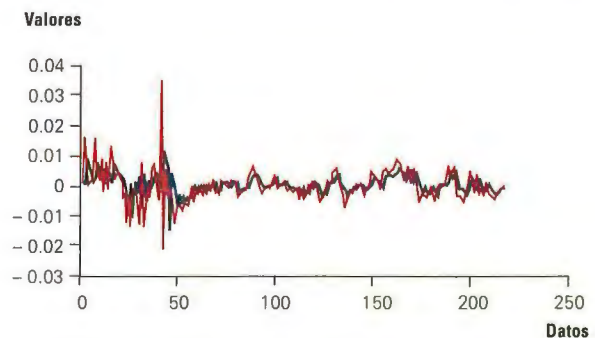
Los pronósticos de las series realizadas mediante modelos ARIMA y RNA en niveles se muestran en la gráfica 2. Se puede observar que el modelo obtenido por RNA pronostica mejor los datos en las primeras 60 observaciones; mientras que, después, el modelo ARIMA sigue más de cerca la tendencia original. Sin embargo, la forma de la serie obtenida por RNA es más parecida a la original, aunque sobrevalúa los datos.

Se evalúa también si los valores generados por ambos modelos pueden reproducir los primeros cuatro momentos de la distribución de los valores de la serie suavizada de manera exponencial sobre los 223 datos pronosticados.

De la comparación entre los momentos de la distribución de los valores observados y los correspondientes pronósticos (véase el cuadro 1), se desprende que la media, la mediana y la desviación estándar de los valores del pronóstico realizado con el modelo ARIMA se encuentran más cerca de los valores correspondientes a la distribución de los valores observados para la serie original, sobre todo cuando se da la sobrevaluación de los datos a partir del dato pronosticado número 60. Sin embargo, el modelo se desempeña mejor en los datos del periodo más cercano de la predicción (véase el cuadro 2).

G R Á F I C A 1

PRONÓSTICO HECHO MEDIANTE EL MODELO DE REDES NEURONALES ARTIFICIALES SOBRE LA SERIE DE RENDIMIENTO LOGARÍTMICO



Predicción, rojo = original, azul = simulación

La bondad de ajuste y capacidad de predicción de los modelos se puede apreciar en el cuadro 3.

CONCLUSIONES

A partir del desarrollo de ambos modelos, de los resultados obtenidos y de las pruebas de hipótesis es posible concluir lo siguiente:

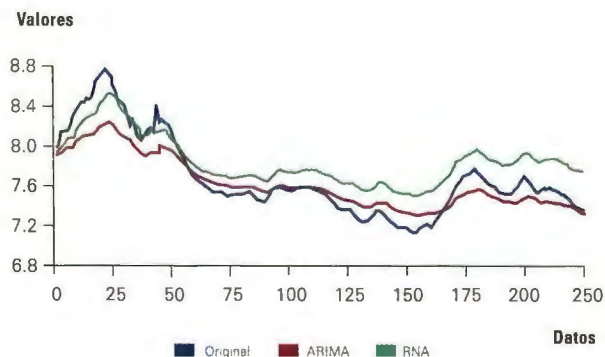
Hay varios modelos de predicción de series de tiempo, algunos más complejos que los desarrollados en esta tesis, como la familia de modelos ARCH. Sin embargo, para realizar un estudio comparativo entre distintas técnicas es necesario hacerlo sobre las mismas bases; es decir, en este estudio se usó una configuración de red neuronal equivalente a un modelo autorregresivo (AR) no lineal, razón por la cual se utilizó un modelo ARIMA como referencia.

Siempre es bueno contar con un modelo que realice predicciones más precisas, como el modelo propuesto, ya que conocer el comportamiento futuro, o al menos contar con una aproximación más cercana a la real, brinda la posibilidad de tomar mejores decisiones en un ambiente incierto como el financiero.

Los resultados y las pruebas de hipótesis, tanto de la bondad de ajuste representada por R^2 como las pruebas estadísticas de Akaike y Schwarz, indican que el modelo basado en RNA es mejor que el modelo ARIMA calculado y usado como referencia dentro de la muestra.

G R Á F I C A 2

COMPARACIÓN ENTRE PRONÓSTICOS MEDIANTE ARIMA Y RNA EN NIVELES¹



1. ARIMA: modelos autorregresivos integrados de medias móviles; RNA: redes neuronales artificiales.

C U A D R O 1

CARACTERÍSTICAS DE LAS DISTRIBUCIONES DE LA SERIE DEL CETE A 91 DÍAS EN EL MEXDER Y SUS SERIES PRONOSTICADAS POR ARIMA Y RNA¹

	Cete	Cete ARIMA	Cete RNA
Media	7.686204	7.665720	7.856051
Mediana	7.565988	7.590327	7.788938
Valor máximo	8.778021	8.337034	8.553452
Valor mínimo	7.139742	7.322352	7.506110
Desviación estándar	0.400236	0.277075	0.249297
Sesgo	1.037646	0.910380	0.952238
Curtosis	3.195469	2.613966	3.275114
Observaciones	223	223	223

1. Cete: Certificado de la Tesorería; Mexder: Mercado Mexicano de Derivados; ARIMA: modelos autorregresivos integrados de medias móviles; RNA: redes neuronales artificiales.

C U A D R O 2

CARACTERÍSTICAS DE LA SERIE DEL CETE A 91 DÍAS EN EL MEXDER Y SUS SERIES PRONOSTICADAS POR ARIMA Y RNA EN LOS PRIMEROS 60 DÍAS¹

	Cete	Cete ARIMA	Cete RNA
Media	8.256858	8.075380	8.189274
Mediana	8.227721	8.058860	8.177933
Valor máximo	8.778021	8.337034	8.557499
Valor mínimo	7.659780	7.748821	7.792983
Desviación estándar	0.282291	0.147812	0.200125
Sesgo	-0.082628	-0.159937	-0.010964
Curtosis	2.498411	2.579405	2.377994
Observaciones	60	60	60

1. Cete: Certificado de la Tesorería; Mexder: Mercado Mexicano de Derivados; ARIMA: modelos autorregresivos integrados de medias móviles; RNA: redes neuronales artificiales.

Fuera de la muestra, también se logró una ligera mejora expresada por los estadísticos raíz cuadrada de error medio (RMSE, por sus siglas en inglés) y U de Theil; sin embargo, el error medio absoluto porcentual (MAPE, por sus siglas en inglés) resultó mayor al planteado en la hipótesis, por lo que se considera que es mejor modelo para pronósticos aunque no de manera absoluta.

Se puede decir entonces que el modelo de RNA cumple su objetivo de generar mejores pronósticos de la tendencia del futuro del cete de 91 días con vencimiento en

BONDAD DE AJUSTE Y CAPACIDAD DE LOS MODELOS PARA EFECTUAR PRONÓSTICOS¹

Modelo	R ²	Akaike	Schwarz	RMSE	MAPE	U de Theil
ARIMA	0.193335	-6.958082	-6.941917	0.004842	186.9167	0.675264
RNA	0.201745	-6.777200	-6.477920	0.004642	204.8512	0.597587

1. ARIMA: modelos autorregresivos integrados de medias móviles; RNA: redes neuronales artificiales; RMSE: raíz cuadrada de error medio (*root mean square errors*); MAPE: error medio absoluto porcentual (*mean absolute percentage error*).

diciembre de 2008, en comparación con el modelo ARIMA usado como referencia. Sin embargo, lo anterior sólo es cierto en los primeros datos del intervalo de predicción (60 días), ya que después el modelo arroja valores que sobrevalúan a los originales e incluso el ARIMA proporciona mejores resultados.

El modelo propuesto puede ser usado como una herramienta adicional a las ya disponibles, mas no como sustituto, pues el modelo lineal es más sencillo de generar y eventualmente obtiene buenas aproximaciones.

También, de acuerdo con los resultados obtenidos y debido a que es un modelo autorregresivo capaz de identificar relaciones no lineales en los datos de la serie, se puede decir que la serie del futuro del cete a 91 días con fecha de vencimiento en diciembre de 2008 presenta relaciones no lineales identificadas por el modelo propuesto. Sin embargo, estas relaciones no lineales no fueron de gran magnitud, puesto que la red no pudo superar de manera muy significativa los resultados del modelo lineal, en cuyo caso se esperarían resultados mejores a los obtenidos.

En general, el proceso de encontrar un modelo basado en redes neuronales es más complejo y delicado que el desarrollo de un modelo ARIMA, ya que implica el cálculo de más parámetros, y la estabilización de la red requiere mayores recursos de cómputo y tiempo, pues el cálculo de sus parámetros es, de forma inherente, a prueba y error.

El modelo propuesto es importante porque se puede utilizar también como parte integral de la opera-

ción financiera de las empresas, ya que sus aplicaciones son amplias, desde la planeación y administración de la tesorería, los pronósticos de ventas, hasta la identificación de problemas financieros y, por supuesto, la valuación de instrumentos derivados en la cobertura de riesgos.

Como trabajo futuro se puede recomendar el uso de un modelo de redes neuronales que tenga la configuración de un modelo ARCH, de manera que pueda captar mejor la volatilidad presente en las series financieras. Aunque el modelo propuesto también puede ser comparado de este modo, no se efectuaría bajo las mismas consideraciones.

Dado que el modelo de RNA diseñado es no lineal, sería conveniente utilizar un modelo mixto que incluya la parte lineal de manera simultánea, para aprovechar las ventajas de ambos enfoques en un mismo modelo.

Otra aplicación interesante sería la evaluación de varias RNA aplicadas a otros mercados financieros, como el de divisas y el de capitales, además de otros instrumentos dentro del mercado de derivados. ◀CE

