

MODELIZACIÓN Y PREDICCIÓN DE SERIES DE TIEMPO FINANCIERAS UTILIZANDO REDES NEURONALES

HUGO ROBERTO BALACCO Y GUSTAVO GERMÁN MARADONA

RESUMEN

El objetivo perseguido en el presente trabajo lo constituye la modelización y predicción de series temporales financieras utilizando redes neuronales. Al respecto, se seleccionó una red neuronal total recurrente con dos capas ocultas, una capa para la función umbral lineal y otra para función arcotangente. Las series utilizadas fueron los índices Merval (Argentina) y Dow Jones (USA). Los resultados, obtenidos con información correspondiente al período 1995-2006, se exponen en comparación con técnicas alternativas y con resultados obtenidos por otros investigadores.

Clasificación JEL: C4

Palabras Clave: Redes Neuronales, Pronósticos, Tipos de Arquitectura, Funciones de Transferencia, Error Medio Absoluto.

ABSTRACT

The purpose of this work is to model and predict Financials Time Series by using neural networks. In order to achieve this aim, a recurrent total neural network with two hidden layers has been chosen; one layer for the linear threshold function and the other for the arctangent function. The series used in this research paper are the Merval index (Argentina) and the Dow Jones (USA). These results are based on information obtained over a period that goes from 1995 to 2006. The presentation will deal with the comparison of alternative techniques and the results obtained by other research workers.

Clasificación JEL: C4

Palabras Clave: Neural Network, Forecast, Architecture Types, Transfer Functions, Mean Absolute Error.

MODELIZACIÓN Y PREDICCIÓN DE SERIES DE TIEMPO FINANCIERAS UTILIZANDO REDES NEURONALES

HUGO ROBERTO BALACCO Y GUSTAVO GERMÁN MARADONA¹

I. Introducción

Durante las últimas décadas, la representación de relaciones económicas no lineales ha sido un tema de investigación muy importante (Van Djck, (1999); Granger y Teräsvirta, 1993), de tal forma que la econometría empírica se ha nutrido de muchos modelos y técnicas, cuyos orígenes se encuentran en la estadística y la inteligencia artificial. Particularmente, el uso de las redes neuronales artificiales ha venido ganando terreno, debido a su capacidad para representar relaciones desconocidas a partir de los datos mismos, en aquellos casos donde la economía no brinda suficientes elementos para establecer la forma funcional de dicha relación. Más aún, se han dedicado importantes esfuerzos para desarrollar aproximaciones metodológicas que permitan formular y formalizar esta clase de modelos a partir de criterios estadísticos y, así, abandonar los métodos ad hoc tradicionales –véase, por ejemplo a Anders y Korn (1999), Fukumizu (2003), entre otros-.

Igualmente, ha resultado importante la aplicación de estas metodologías en el modelado y predicción de series económicas y financieras reales con el fin de ganar una mayor experiencia en su uso, ya que, tal como lo señalan Granger y Teräsvirta (1993), no hay suficientes indicios que permitan determinar cuáles modelos podrían ser los más adecuados en la representación empírica de relaciones económicas.

El objetivo básico que se persigue en este trabajo es el de modelización, utilizando una red neuronal, de las series de tiempo financieras, correspondiente a los índices Merval y Dow-Jones, con la finalidad de evaluar su capacidad predictiva.

La principal característica asociada a una red neuronal es su flexibilidad a la hora de modelizar estructuras no lineales como presuntos procesos generadores de datos correspondientes a una serie temporal.

¹ Facultad de Ciencias Económicas, Universidad Nacional de Cuyo.

Por otra parte, la selección de una red neuronal es una tarea bastante exigente, pues implica la prueba de diferentes alternativas que combinan estructuras o arquitecturas disponibles con funciones inputs, de transferencias y algoritmos de aprendizaje. La forma más adecuada de evaluar los distintos modelos es a través de su capacidad, predictiva, no en términos de nivel, sino de dirección de cambio del output de referencia.

En este trabajo, finalmente, se seleccionó una red total recurrente con dos capas ocultas, una capa para la función umbral lineal y otra para la función arcotangente. Por otra parte, cada capa oculta consta de cuatro nodos o neuronas. La función input correspondiente es la función Dot y el algoritmo de aprendizaje, el algoritmo de Powell. El software utilizado fue el ThinksPro.²

Los resultados se evaluaron en términos de su capacidad de predicción según la dirección de cambio del output y se exponen comparándolos con evidencia similar obtenida por otros investigadores, como así también con la metodología tradicional de predicción correspondiente a modelos de tipo ARIMA. Por último, los resultados correspondientes al período junio de 1995 - mayo de 2000, se confrontan con los obtenidos en el período enero 2003 - marzo 2006.

II. Generalidades

Las redes neuronales son una clase de modelos no lineales inspirados en la arquitectura neuronal del cerebro. Nacieron en el campo de la inteligencia artificial como un intento de modelizar la capacidad de aprendizaje de los sistemas biológicos neuronales mediante la modelización de la estructura del cerebro.

En otras palabras, una red neuronal es una estructura de interconexiones, compuesta por una serie de capas, formadas por neuronas, de manera que cada una de ellas alimenta a todas las neuronas de la capa siguiente (la misma estructura que se observa en el cerebro). La neurona proporciona una salida o output que es el resultado que se obtiene desde una función, llamada función de activación, que básicamente consiste en la suma de los inputs respectivos multiplicados por ponderaciones o pesos. Estas ponderaciones se ajustan

² En el software ThinksPro, "*Neural Networks for Window*", Logical Designs, 2015, Olite Court, La Jolla, CA 92037, 1999, se especifica en forma completa todo lo relacionado con la arquitectura de una red neuronal, como así también lo relativo a funciones inputs, algoritmos de aprendizaje y demás aspectos inherentes a la modelización correspondiente.

conforme a un algoritmo de aprendizaje especificado con el que se pretende minimizar la función de coste calculada a partir del error cometido si se compara la salida de la red con el valor real.

El motivo básico para el uso de redes neuronales es la predicción de una variable y desde un conjunto de variables predeterminadas que pueden incluir valores rezagados de y .

En el campo de la predicción, generalmente, se toma como punto de partida, el modelo lineal general de regresión, de acuerdo a la siguiente expresión:

$$y_i = \beta_0 + \sum_{i=1}^k \beta_i x_i + \eta_i, \quad (1)$$

$$\eta_i \approx N(0, \sigma^2) \quad (2)$$

donde la variable η es la perturbación estocástica, normalmente distribuida con media cero y varianza σ^2 y $\{\beta\}$ es el conjunto de parámetros a estimar. El vector $\{\hat{\beta}\}$ representa el vector de parámetros estimados, y el conjunto de valores generados por el modelo se expresa como $\{\hat{y}\}$.

En el contexto econométrico convencional, existe una variedad de métodos para estimar $\{\hat{\beta}\}$, bajo supuestos alternativos acerca de la distribución del término de perturbación η .

El objetivo del proceso de estimación es el de encontrar un conjunto de parámetros para el modelo de regresión que minimice la suma de los residuos en la siguiente formulación estándar:

$$y_t = \beta_0 + \sum_{i=1}^k \beta_i x_{it} + \sum_{j=1}^{k^*} \delta_j y_{t-j} + \eta_t, \quad (3)$$

$$\eta_t \approx N(0,1), \quad (4)$$

en donde k^* representa el número de rezagos inherentes a la variable dependiente y ; y por lo tanto, un número extra de parámetros a estimar.

De esta forma, el modelo lineal de regresión o modelo autorregresivo, constituye un punto de partida razonable, por su simplicidad, para fines predictivos; sobre todo, en aquellos mercados en los que se observan sólo pequeños cambios simétricos, en la variable a predecir, alrededor de una tendencia de largo plazo.

Sin embargo, en principio, este modelo no sería adecuado para el pronóstico de series temporales correspondientes a mercados financieros, en los cuales se observan asimetrías de comportamiento, volatilidad, etc.; creando el campo propicio para un enfoque no lineal en el análisis de dichas series financieras. Por ejemplo, las denominadas “burbujas” donde suaves movimientos hacia arriba en precios de activos son seguidos de colapsos imprevistos. De esta manera, el modelo lineal no podría captar o predecir en forma satisfactoria en presencia de estos acentuados puntos de inflexión. Por esta razón, la atención se ha dirigido a técnicas de predicción no lineales. (McNelis (1997)).

En este contexto, una red neuronal constituye una alternativa frente a los procesos descriptos. Las razones para su uso son simples y directas. El objetivo básico es disponer de un método adecuado y relativamente sencillo de cara a la predicción. Se supone que una serie temporal ha sido generada por un modelo no lineal complejo y se pretende que el proceso de aprendizaje permita la selección de una estructura con el menor número de parámetros que sea posible.

Los atributos básicos de una red neuronal pueden ser divididos en arquitectura y en neurodinámica.

La arquitectura define la estructura de la red; es decir el número de neuronas artificiales y sus conexiones. En otras palabras, una red neuronal consiste en varias neuronas o elementos de proceso interconectados, con características propias como entradas, pesos sinápticos, umbral, activación y salida.

La neurodinámica de la red define sus propiedades; es decir, cómo la red aprende, recuerda, asocia y continuamente compara la nueva información con el conocimiento existente; clasifica nueva información y eventualmente, desarrolla nuevas clasificaciones si fuera posible.

Para implementar una red, se debe seleccionar el modelo o modelos matemáticos que se supone mejor describiría la información bajo análisis, (Velásquez Henao y González Rivera (2006)).

III. Arquitectura

Las redes neuronales se componen de un número considerable de neuronas interconectadas, formando diferentes arquitecturas, las cuales procesan información a través de canales bien definidos.

Se pueden distinguir redes de una sola capa oculta (singlelayer), y de dos o más capas, llamadas multicapa (multilayer). Normalmente, la capa en donde es suministrada la información a la red es la capa de entrada (input layer); mientras que la capa en donde el producto se obtiene es la capa de salida (output layer).

Si se tiene en cuenta cómo se alimentan las neuronas de una red, en relación a la información, se distinguen tres tipos de red que constituyen las de mayor uso en el ámbito de la economía y de los mercados financieros, las redes alimentadas hacia delante (feed forward), retroalimentadas y recurrentes (feedback), o totalmente recurrentes.

En las redes configuradas hacia delante cada capa es alimentada desde el vector externo (vector input). Además, cada capa recibe entradas (input), desde cada capa anterior a través de un esquema lineal. Por ejemplo, si una red tiene tres capas, la primera sólo recibe el vector externo. La segunda, recibe el vector externo y el producto de la primera. Por último, la tercer capa recibe el vector externo y los vectores de productos de la primera y segunda capa. (Olmedo y otros (2004)).

Como las aplicaciones de las redes neuronales para la predicción de series temporales se basan en un ajuste dinámico, las redes alimentadas hacia adelante acomodan esta dinámica incluyendo en el algoritmo una función de coste que relaciona los valores de salida de la red a partir de los valores de entrada y comparándolas con los valores deseados. Sin embargo, en este tipo de representación dinámica no se incluye la figura de retroalimentación o feedback interno. Esta retroalimentación es propia de las redes retroalimentadas en las que cada nodo se alimenta no sólo del vector externo, sino del producto de los otros nodos de la red.

Por último, y de uso más incipiente en el análisis de series de tiempo financieras, es la red neuronal recurrente o total recurrente, la que además de retroalimentación, permite conexiones dentro de una misma capa.

La Figura 1 muestra en forma simplificada las estructuras de las redes neuronales alimentadas hacia delante (feedforward), retroalimentadas (feedback), y recurrente o total recurrente.³

Las restricciones de las redes feedforward y feedback, quedan resueltas en la arquitectura recurrente, en la que cada nodo recibe como inputs, los inputs externos, como así también los outputs de cualquier otro nodo de la red. El nodo tiene también una interconexión desde su propio output.

Esto implica una secuencia temporal en el proceso de activación de las neuronas (nodos) de la red. Comenzando al tiempo $t=0$, la activación de las neuronas son llevadas a ese punto inicial. Al tiempo $t=1$, cada nodo es activado por un vector de inputs integrado por el input externo para $t=1$, y también con el output corriente, generado en $t=0$, en cada nodo.

Las arquitecturas descritas en ThinksPro permiten hasta 3 capas ocultas. Normalmente, en una Red Total Recurrente puede usarse una o dos capas ocultas. En realidad, conviene aclarar que cada capa oculta está en función directa con el número de funciones de transferencia que se usan.

En el presente capítulo se ha seleccionado una arquitectura total recurrente con dos capas ocultas, una para la función de transferencia umbral lineal y otra para la función de transferencia arcotangente. Además, cada capa oculta consta de cuatro nodos o neuronas cada una.⁴

A. Función input

Cada instancia de procesamiento o nodo recibe un vector que está constituido por un número determinado de inputs y en algunos casos, además, por una constante o sesgo.

Los inputs pueden ser externos o internos, que son los outputs de otros nodos, o ambos a la vez.

³ OLMEDA, I., "Avances Recientes en Predicción Bursátil", ed. Universidad de Alcalá. Alcalá de Henares, pág. 196, 1996.

⁴ El número de nodos o neuronas se estableció de acuerdo a la siguiente fórmula $(M+N)/2$, siendo M y N el número de inputs y de output respectivamente. Al respecto, ver Chan, M.C., Wong, C.C. y Lan, C.C. (2001). "Financial Time Series Forecasting by Neural Network using conjugate Gradient Learning Algorithm and Multiple Linear Regression Weight Initialization". Mimeo, pág. 4, 2001.

La función input combina un vector de inputs con otro vector de ponderaciones; por lo tanto el input neto f es una función de un vector de inputs x y de un vector de ponderaciones w .

En este trabajo, se ha seleccionado la función Dot, que es una suma ponderada de inputs más un sesgo.

$$f(x, w) = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n . \quad (5)$$

En esta función, el primer elemento de (5), es una constante especial llamada sesgo o constante, y los siguientes elementos son ponderaciones especiales llamadas sesgo de ponderación. Intuitivamente, la función Dot pondera las entradas (inputs), de acuerdo a la importancia relativa de cada una de ellas en el vector de entradas que accede a cada nodo de la red.

B. Función de transferencia

La función de transferencia describe las formas en que se procesa el input en los nodos. La no linealidad de la función de transferencia es la que otorga a la red neuronal la flexibilidad y, por lo tanto, las ventajas en modelización de series temporales financieras sobre las técnicas tradicionales de regresión. (Panda y Narasimhan (2006)).

Desde un punto de vista intuitivo la función tiene una interpretación particular; y es la de limitar o acotar el crecimiento de la función input, f , en un intervalo finito. Es decir, en ciertas oportunidades, el proceso de aprendizaje permite a la función de transferencia modificar ponderaciones, incluso cuando las magnitudes son relativamente grandes.

Sin embargo, las consideraciones de mayor consistencia técnica, son las que se refieren a la facultad de que la función de transferencia introduzca no linealidades en la red neuronal. Al respecto, es posible ver una red como representando una función de tipo englobadora compuesta por “funciones básicas”, que son las funciones de transferencia.

En este trabajo, tal como se indicó en la introducción, luego de una exigente tarea de prueba, utilizando la información disponible, se

seleccionaron las funciones umbral lineal (threshold lineal) y arcotangente, ambas acotadas entre 0 y 1. Las Figuras 2 y 3 reproducen estas funciones.⁵

IV. Resultados obtenidos

Tal como se indicó en la introducción, la idea central de este trabajo es investigar empíricamente si un modelo de predicción más completo podría mejorar la capacidad predictiva por encima de los métodos convencionales; por ejemplo, modelos de series temporales de tipo ARIMA que comprenden la modelización camino aleatorio.

Con esta finalidad, la arquitectura seleccionada se corresponde con la red neuronal total recurrente con dos funciones de transferencia, umbral lineal y arcotangente.⁶

El proceso de aprendizaje utilizado fue el método de Powell, descrito en forma detallada en el software ThinksPro, que resulta expresamente indicado para este tipo de red. Además, como criterio de convergencia se optó por el llamado error medio absoluto (EMA), cuyo uso se aconseja en el campo de las series económicas y financieras porque constituye una estadística robusta, desde que penaliza errores groseros (outliers). De acuerdo a este indicador, y en todos los casos de estimación de la red, la convergencia se alcanzó antes de las 20.000 iteraciones.⁷

En la literatura actual especializada, la evaluación de estas redes tiene lugar estrictamente con relación a su capacidad predictiva, no en términos de nivel sino de dirección de cambio del output correspondiente (Leung y otros (2000)).

En otras palabras, se intenta evaluar la red por los aciertos alcanzados al predecir si un índice se moverá hacia arriba o hacia abajo (aumentará o disminuirá) desde un punto temporal determinado. En este contexto, una parte de la información muestral se utiliza para estimar la red, y el resto como

⁵ El uso de las funciones de transferencia Sigmoid, Bipolar Sigmoid y Gaussiana reportaron resultados inferiores a los obtenidos con la función de transferencia arcotangente, tanto en relación con el Error Medio Absoluto (EMA), como en función de la capacidad predictiva en términos de la dirección de cambio.

⁶ Una función para cada capa oculta.

⁷ Los datos de las series temporales fueron reescalados, debido a que las funciones de transferencia utilizadas están acotadas entre cero y uno. La transformación consistió en restar a cada valor de la serie el valor mínimo de las mismas y dividirlo por el rango.

horizonte de pronóstico, para evaluar las predicciones en comparación con los valores observados.

En una primera instancia, se estimó la red en un período comprendido entre el primero de junio de 1995 y el 30 de abril del año 2000, es decir, 1212 observaciones.

El período comprendido entre el 1ro. de mayo del año 2000 y el 30 de octubre del mismo año se consideró como horizonte de pronóstico para evaluar la capacidad predictiva de la Red, 120 observaciones.

Los resultados de la evaluación se exponen en la Tabla 1. Los modelos seleccionados para ambos indicadores contienen cuatro rezagos ($t-1$, $t-2$, $t-3$, $t-4$), tanto para el índice Merval como para el DOW JONES.⁸

Como puede observarse, para ambos índices la cantidad de aciertos se aproxima al cincuenta por ciento.

A. La inclusión de la tasa de interés

Estudios relacionados con la predictibilidad de retornos y precios de activos, y sus determinantes, son numerosos en la literatura especializada, al considerar la posibilidad de comprobar la hipótesis semifuerte de mercados eficientes; información pública disponible. Al respecto, suelen mencionarse como principales inputs, la tasa de interés, el índice de precios al consumidor, índices de producción industrial, y por supuesto rezagos de retornos o precios.

Con relación a dichos inputs, y únicamente para el índice Merval, se utilizó la tasa de interés como un determinante adicional a los rezagos correspondientes a los índices de referencia. La tasa de interés finalmente seleccionada, corresponde a la tasa de interés para depósitos a treinta días del Banco de la Nación Argentina.⁹

El período de estimación y de pronóstico fueron los mismos que figuran en la Tabla 1, siendo el modelo más adecuado para el índice Merval el que contiene cuatro rezagos, ($t-1$, $t-2$, $t-3$ y $t-4$) para el componente autorregresivo y dos rezagos ($t-1$ y $t-2$), para la tasa de interés.

⁸ Modelos con mayor número de rezagos fueron estimados y evaluados, con seis y ocho rezagos, pero su capacidad predictiva no supera al modelo seleccionado.

⁹ En Leung, M.T., Daouk H. y Chen, A.S., "Forecasting stocks indices: a comparison of Classification and Level Estimation models, *International Journal of forecasting*", pág.176, 2000, se destaca: "Though other macroeconomic variables can be used as inputs, the majority of useful information for forecasting is subsumed by the interest rates and lagged returns".

En la Tabla 2 se exponen los resultados obtenidos.

Como se puede observar, la inclusión de la tasa de interés mejora la capacidad predictiva de la red. De todos modos, como se expondrá en el punto siguiente, en la literatura internacional comparada, se encuentran resultados similares a los logrados en este estudio.

B. Análisis comparativo

Leung y otros (2000), realizaron un estudio para evaluar la capacidad predictiva de una red neuronal alimentada hacia adelante, utilizando tres índices de precios financieros, S&P 500 (USA), UKFTSE 100 (Reino Unido) y Japan Nikkei 225 (Japón).

El estudio de referencia utiliza datos mensuales que cubren el período enero 1967 a diciembre de 1990 (228 meses de observación), mientras que el horizonte de pronóstico comienza en enero de 1991, hasta diciembre de 1995.

Los resultados se pueden resumir en la Tabla 3, en la que además de la performance predictiva de la red neuronal, con relación a los aciertos en un contexto de predicción de dirección de cambio, se exponen también, los porcentajes de aciertos correspondientes a otras modelizaciones más convencionales, como modelo adaptativo con suavización exponencial, VAR (Vectores autorregresivos), con filtro de Kalman y Función de Transferencia Multivariada.

Como puede apreciarse, el número de porcentaje de aciertos, correspondiente a la red neuronal multicapa, levemente supera los resultados obtenidos en este estudio con una red total recurrente.

Velásquez Henao y Gonzalez Rivera (2006), utilizan el tipo de cambio real efectivo del peso colombiano frente a las monedas de 18 países, calculado por el Fondo Monetario Internacional (ITCRFMI), para el período comprendido entre 1989:12 y 2003:10. El modelo utilizado es una red neuronal perceptrón multicapa, y alternativamente, un modelo autorregresivo de primer orden, AR(1). Los autores comparan la capacidad predictiva de ambos modelos, dentro del período muestral, a través de estadísticos de ajuste convencionales que se resumen en la Tabla 4.

Como se puede apreciar, el ajuste no revela resultados contundentes, a favor de la red, en términos de los estadísticos consignados.

Por otra parte, Chakradhara, P. y Narasimhan, V. (2006), realizaron un estudio para Bombay, con datos diarios comprendidos entre 1994:1 y 2004:12, utilizando como indicador del mercado un índice agregado de activos financieros (Bombay Stock Exchange Sensitive Index). Comparan la capacidad predictiva de una red neuronal alimentada hacia adelante (a simple hidden layer feedforward network), con un modelo camino aleatorio y otro lineal autorregresivo con dos rezagos. Las conclusiones de estos autores, en términos de estadísticos convencionales de ajuste, similares a los utilizados por Velásquez Heano y Gonzalez Rivera, ponen de manifiesto que los errores de pronósticos no difieren sustancialmente entre los tres modelos, lo que impide consagrar, en forma contundente, la superioridad de la red neuronal.

Por último, resulta conveniente consignar la performance alcanzada por una red neuronal con arquitectura alimentada hacia adelante. Esto es así, desde que estas redes son también utilizadas internacionalmente con propósitos de pronóstico. Las funciones de transferencia utilizadas fueron las mismas usadas en la red total recurrente; es decir, umbral lineal y arcotangente, y dos capas ocultas. Pero como el algoritmo de Powell es incompatible con este tipo de arquitectura, se utilizó, en su reemplazo, el algoritmo conjugado gradiente; y como función de ingreso se sustituyó la función Dot por una función con ponderaciones obtenidas sobre la base de regresión lineal general.

Para el índice Merval, y teniendo en cuenta los mismos procesos de estimación y pronóstico fuera de la muestra, utilizados en este capítulo, el porcentaje de aciertos de pronósticos de dirección de cambio alcanzó el cincuenta y uno por ciento.

Por otra parte, se destaca que si se utiliza la información disponible para pronosticar con los modelos AR(2) del Merval y AR(1) con tendencia del DOW JONES, respetando los mismos períodos de estimación y de pronóstico que se utilizaron para la red neuronal total recurrente, los resultados son apenas inferiores. En efecto, el porcentaje de aciertos para el índice Merval es del cincuenta y tres por ciento, y para el DOW JONES alcanza al cuarenta y ocho por ciento, el número de aciertos, en términos de dirección de cambio del input.¹⁰

¹⁰ En Pindyck, R., S., and Rubinfeld, D.L., “*Econometría. Modelos y Pronósticos*”, 4ª ed. Mc Graw Hill, 1998, Cap. 18, se expone la metodología de estimación, validación y pronóstico utilizando modelos ARIMA.

Además, tal como se indicó en la introducción de este capítulo, se exponen los resultados obtenidos con posterioridad a la crisis económica que tuvo lugar en el año 2001 y parte del año 2002 en la República Argentina. La estructura y especificación de la red neuronal es similar a la utilizada en la obtención de los resultados de la Tabla 2.

Estos resultados corresponden, en primer instancia, a un período de estimación comprendido entre el día 2 de Enero de 2003 y el 28 de Febrero de 2006, y de predicción entre el 1 de Marzo de 2006 y el 11 de Mayo de 2006.

En una segunda instancia se estima la red para el mismo período anterior, pero se acortan los períodos de pronóstico, los cuales se extienden desde el 16 de marzo de 2006 hasta el 11 de Mayo de 2006. Los resultados se exponen en la Tabla 7.

V. Conclusiones

Los resultados obtenidos en este trabajo, de alguna manera, justificarían los esfuerzos realizados y a efectuar en un futuro, con la finalidad de continuar con las investigaciones teórico-empíricas con el objeto de perfeccionar la modelización, con fines de predicción, utilizando redes neuronales.

Esto es así, ya que la performance predictiva de las redes neuronales expuesta en este trabajo, como así también en la literatura especializada comparada, fue siempre superior en términos generales, aunque en ciertas situaciones levemente a los métodos convencionales.

En particular, en este estudio, se destaca la predicción efectuada con el índice Merval, en el horizonte de pronóstico comprendido entre el 16-03-2006 y el 11-05-2006, que alcanzaría casi el sesenta y tres por ciento en términos de aciertos.

La existencia, en mercados financieros, de anomalías, no normalidades, burbujas, costos de entrada y de salida, costos de adquirir y procesar información, demoras en la reacción de los agentes, etc., no puede ponerse en duda. Por ello, modelizaciones no lineales, tendientes a considerar estos aspectos, y planteados como alternativa a las convencionales que engloban la hipótesis de camino aleatorio, podrían, en un futuro, aportar resultados diferenciales importantes en el contexto de la predicción en mercados financieros.

Referencias

Anders, U. and O. Korn (1999). "Model selection in neural networks". *Neural Networks*, Vol. 12: 309-323.

Chakradhara, P. and V. Narasimhan (2006). "Predicting Stock Returns: An Experiment of the Artificial Neural Network in Indian Stock Market". *South Asia Economic Journal*, Vol. 7: 205-218.

Fukumizu, K. (2003). "Likelihood ratio of unidentifiable models and multilayer neural networks". *Annals of Statistics*, Vol. 31 (3): 833-851.

Granger, C. W. and T. Terasvirta (1993). "*Modelling Nonlinear Economic Relationships*". Oxford: Oxford University Press, Oxford.

Leung, M. T., H. Daouk and A. S. Chen (2000). "Forecasting stock indices: a comparison of Classification and Level Estimation models". *International Journal of Forecasting*, 16: 173-190.

Mcnelis, P. D. (1997). "Neuronal Networks and Genetic Algorithms: Tools for Forecasting and Risk Analysis in Financial Markets". Mimeo Department of Economics, Georgetown University.

Olmeda, I. (1996). "*Avances Recientes en Predicción Bursátil*". Alcalá De Henares: Editorial Universidad de Alcalá.

Olmedo, E., J. M. Valderas, R. Mateos y R. Gimeno (2004). "Utilización de Redes Neuronales en la Caracterización, Modelización y Predicción de Series Temporales Económicas en un Entorno Complejo". Mimeo, Madrid-Sevilla, España.

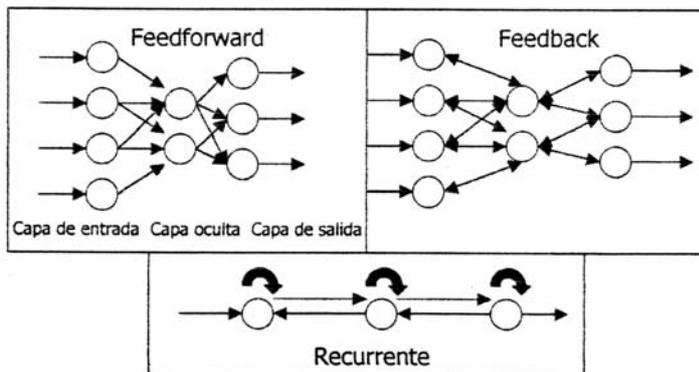
Panda, C. and V. Narasimhan (2006). "Predicting Stock Returns: An Experiment of the Artificial Neural Network in Indian Stock Market". *South Asia Economic Journal*, Vol. 7 (2): 205-218.

Pindyck, R. and D. L. Rubinfeld (1998). "*Econometría. Modelos y Pronósticos*". Mc Graw Hill, 4ª ed.

Van Djck, D. (1999). "Smooth transition models: extensions and outlier robust inference". Tesis de Ph. D. no publicada, Erasmus University, Rotterdam, Holanda.

Velasquez Henao, J. D. y L.M. Gonzalez Rivera (2006). "Modelado del Índice de Tipo de Cambio Colombiano usando Redes Neuronales Artificiales". Mimeo, Bogotá, Colombia.

Figura 1
Redes Neuronales Feedforward, Feedback y Recurrente



Fuente: OLMEDA, I., "Avances Recientes en Predicción Bursátil", ed. Universidad de Alcalá. Alcalá de Henares, pág. 196, 1996.

Figura 2
Función de Transferencia Umbral Lineal

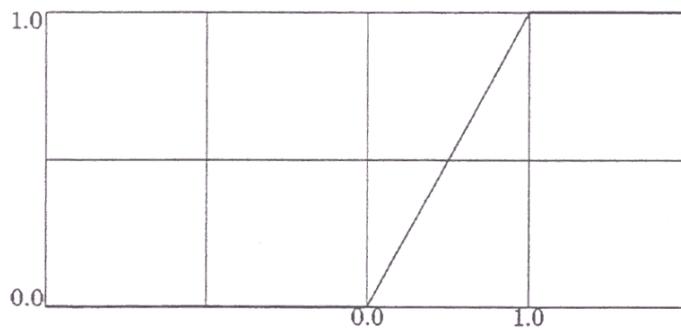
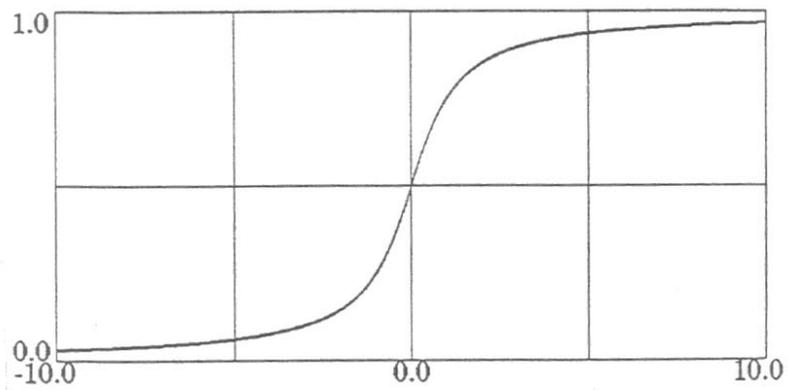


Figura 3
Función de Transferencia Arcotangente



Nota: la fórmula de referencia de la función arcotangente es: $t(f) = \arctan(f / \pi) + 0.5$.

Tabla 1
Porcentaje de aciertos en el pronóstico de la dirección de cambio
utilizando una Red Neuronal Total Recurrente

Índice	Período de estimación	Horizonte de pronóstico	EMA	Porcentaje de aciertos
MERVAL	01-06-1995 30-04-2000	01-05-2000 30-10-2000	0.0283	48,74
DOW- JONES	01-06-1995 30-04-2000	01-05-2000 30-10-2000	0.0166	48,70

Tabla 2
Porcentaje de aciertos en el pronóstico de la dirección de cambio
utilizando una Red Neuronal Total Recurrente, incluyendo la tasa de
interés como input adicional

Indice	Período de estimación	Horizonte de pronóstico	EMA	Porcentaje de aciertos
MERVAL	01-06-1995 30-04-2000	01-05-2000 30-10-2000	0.015611	54,62

Tabla 3
Capacidad predictiva de Modelos de Nivel (porcentaje de aciertos), en el
trabajo de Leung, Daouk y Chen (2000)

Modelo	Porcentaje de aciertos		
	Indice		
	S&P500	UKFTSE 100	Japan Nikkey 225
Adaptativo con suavización exponencial	48	55	63
VAR con Filtro de Kalman	53	53	58
Función de Transferencia Multivariada	53	56	58
Red Multilayered feed forward	63	50	60

Tabla 4
Estadísticos de ajuste de los modelos a la serie, en el trabajo de Velásquez
Henao y Gonzalez Rivera (2006)

Estadístico	MLP ($H=1$)	AR(1)
Error Cuadrático medio	13,890	14,500
Error medio absoluto	2,119	2,091
Raíz del error cuadrático geométrico	1,336	1,219
U-Theil	0,972	0,994
R^2	0,938	0,937

Tabla 5
Porcentaje de aciertos en el pronóstico de la dirección de cambio
utilizando una red alimentada hacia adelante, incluyendo la tasa de
interés como input adicional

Indice	Período de estimación	Horizonte de pronóstico	EMC	Porcentaje de aciertos
MERVAL	01-06-1995 30-04-2000	01-05-2000 30-10-2000	0,0401	51

Tabla 6
Porcentaje de aciertos en el pronóstico utilizando un modelo AR(2) para
el índice MERVAL y un modelo AR(1) con tendencia para el índice DOW
JONES

Indice	Período de estimación	Horizonte de pronóstico	Porcentaje de aciertos
MERVAL	01-06-1995 30-04-2000	01-05-2000 30-10-2000	53,00
DOW JONES	01-06-1995 30-04-2000	01-05-2000 30-10-2000	48,00

Tabla 7
Porcentaje de aciertos en el pronóstico de la dirección de cambio
utilizando una red neuronal total recurrente (período post-crisis)

Indice	Período de estimación	Horizonte de pronóstico	Porcentaje de aciertos	EMC
Merval	02-01-2003	01-03-2006	47,25	0,01366
	28-02-2006	11-05-2006		
	02-01-2003	16-03-2006	62,85	0,01666
	28-02-2006	11-05-2006		
	02-01-2003	30-03-2006	46,00	0,0284
	28-02-2006	11-05-2006		