

Comparación de los métodos de series de tiempo y redes neuronales

Comparison of time series methods and neural networks

César Menacho Ch.¹

Resumen

El propósito de la investigación es hacer un estudio comparativo de los métodos estadísticos clásicos aplicados al análisis de series de tiempo; regresión polinomial de segundo grado, promedio móvil, suavización exponencial simple y suavización exponencial doble con los modelos de redes neuronales artificiales multicapa backpropagación. Los métodos propuestos son aplicados a ocho series de tiempo obtenidas del portal del Instituto Nacional de Estadística e Informática (INEI). Se comparan los métodos, usando las medidas para la medición del error de pronóstico MAPE, MAD y MSE. Los resultados mostraron que los modelos de redes neuronales tuvieron menores valores de MAPE en las ocho series y menores valores de MAD y MSE en cuatro de las series de tiempo analizadas.

Palabras clave: medición del error; redes neuronales; regresión polinomial; series de tiempo; suavización exponencial.

Abstract

The purpose of the research is to make a comparative study of classical statistical methods applied to time series analysis, second-degree polynomial regression, moving average, exponential smoothing, simple and double exponential smoothing models backpropagation multilayer neural networks. The proposed methods are applied to eight time series obtained from the INEI portal. Methods are compared, using the measurements for measuring forecast error MAPE, MAD and MSE. The results showed that neural network models had lower MAPE values in the eight series, while lower MAD and MSE in four of the analyzed time series.

Keywords: measurement error; neural networks; polynomial regression; time series; exponential smoothing.

1. Introducción

El análisis de series de tiempo tiene como propósito estudiar el comportamiento de una variable cuyos datos son registrados cronológicamente en el tiempo y predecir sus valores futuros para apoyar la toma de decisiones en una organización. Los métodos estadísticos clásicos, que son usados para realizar el análisis de series de tiempo, son: regresión lineal y no lineal, suavización exponencial, promedios móviles y ARIMA.

Los modelos de redes neuronales artificiales (RNA) están siendo aplicados, en las últimas décadas, como métodos alternativos a los problemas de predicción y clasificación de datos porque muestran, entre otros, varias características distintivas que las hacen atractivas para la tarea de predicción; asimismo, son métodos auto-adaptativos basados en pocas suposiciones a priori acerca del problema en estudio. Las RNA tienen la capacidad de la generalización de datos a partir de una muestra, infiriendo los resultados y presentándolos; incluso si los datos contienen información faltante y ruido. Destacamos que los RNA son aproximaciones de

funciones universales, demostrando que una red puede aproximar cualquier función continua o no lineal con una precisión deseada.

En el análisis de series de tiempo, la aplicación de las RNA son consideradas como técnicas de regresión no paramétrica y no lineal, demostrando un amplio potencial en el pronóstico de las series debido a sus capacidades de adaptabilidad, generalización, aprendizaje y la posibilidad de representar relaciones no lineales; en contraposición de los modelos de Box-Jenkins o ARIMA, que suponen que las series de tiempo son generadas a partir de procesos lineales (Zhang *et al.*, 1998).

Sánchez y Velásquez (2010) mencionan las características que justifican y favorecen la aplicación de las RNA en el análisis de series de tiempo:

El proceso generador de los datos de las series de tiempo muestra un comportamiento desconocido y difícil de identificar, lo cual limita la capacidad de los modelos paramétricos para representar adecuadamente este comportamiento. Las redes neuronales son modelos auto adaptativos que no requieren supuestos a priori acerca del

¹ Profesor Principal del Departamento de Estadística Informática. Universidad Nacional Agraria La Molina, Lima, Perú. Email: cmenacho@lamolina.edu.pe

problema en estudio, características altamente deseables en casos donde el mecanismo generador de datos es desconocido e inestable.

Las series de tiempo evidencian comportamiento inestable. Las RNA tienen una capacidad de aprendizaje y generalización que les permite aprender comportamientos complejos directamente desde los datos y, a partir del conocimiento adquirido, inferir correctamente la evolución de la serie.

Las relaciones entre la serie de tiempo y sus variables explicativas son, generalmente, desconocidas. La característica de función de aproximación de las RNA les permite identificar dependencias ocultas de orden no lineal.

Las series de tiempo exhiben patrones de comportamiento tales como tendencias, ciclos y estacionalidades; pero también irregularidades como cambios estructurales, datos atípicos, etc. La flexibilidad para construir topologías de los modelos de redes neuronales (número de nodos de entrada, capa oculta y diferentes funciones de activación), les facilita identificar y pronosticar los patrones de comportamiento en las series.

Las series de tiempo provienen de múltiples campos del conocimiento y los datos son medidos en diferentes frecuencias de tiempo. La flexibilidad en la forma funcional que pueden adquirir y el carácter de aproximador universal, permite a las redes neuronales representar comportamientos complejos, sin importar el campo del conocimiento de donde provengan los datos ni la frecuencia de tiempo.

En la literatura se encuentra un creciente interés por la aplicación de los modelos de redes neuronales en el análisis de las series de tiempo. Sin embargo, existen estudios que muestran inconsistencias acerca del uso, tales como: modelo neuronal incorrecto, número inadecuado de nodos de entrada, nodos en la capa oculta, inadecuada función de transferencia. Las redes neuronales artificiales son aplicadas con éxito en el pronóstico de series de tiempo en diferentes campos del conocimiento, tales como biología, finanzas, economía, energía, medicina, meteorología y turismo.

Montaño *et al.* (2011) describen y comparan los modelos de redes neuronales artificiales, perceptron multicapa (PM), funciones de base radial (FBR), red neuronal de regresión generalizada (RNRG) y redes neuronales recurrentes (RNR) para la predicción de series de tiempo. Los resultados indicaron errores menores del 10 % en los cuatro modelos y con un alto grado de ajuste de la capacidad predictiva.

Adhikari *et al.* (2013), con la finalidad de predecir el enjambre de partículas de optimización en soluciones químicas, usan los modelos de redes neuronales feedforward (FANN) y elman (EANN) con el algoritmo de retro propagación y los compara con tres métodos clásicos de análisis de series de tiempo: estacional auto regresivo integrado de media móvil (SARIMA), Holt-Winters (HW) y soporte de máquina vectorial (SVM). La aplicación de los modelos propuestos a tres series de tiempo muestran que las predicciones obtenidas con

las redes neuronales FANN y EANN son más precisas (menores errores) que las obtenidas con los modelos SARIMA, HW y SVM.

Stern (1996) usa una arquitectura de un modelo de red neuronal con cinco unidades en la capa de entrada, cuatro unidades en la capa oculta y una en la salida. El modelo de red neuronal multicapa es ajustado a un modelo auto regresivo de orden 2 (AR(2)). Los resultados experimentales con 5000 conjuntos de datos usados para el entrenamiento mostraron que las redes neuronales tienen menores errores estándar que el modelo AR(2).

Nojek *et al.* (2003) presentan un estudio para pronosticar los datos históricos de las ventas de la Empresa Shell CAPSA usando los métodos estadísticos clásicos (promedios móviles, regresión, suavización exponencial) y cuatro topologías de modelos de red neuronal backpropagación (con diferentes número de rezagos, unidades en la capa de entrada, oculta y de salida) y considerando dos fases (una hacia adelante y una hacia atrás). Para evaluar si los métodos estadísticos son mejores que el modelo de red neuronal, se plantea la prueba de signos de Kendall, que es calculada a partir de las diferencias en valor absoluto entre los datos históricos y los pronósticos con el método estadístico y con el modelo red neuronal. Los resultados indican, con un nivel de confianza del 82,3 %, que las redes neuronales obtuvieron pronósticos más precisos en comparación con los métodos estadísticos en las siete series de tiempo analizadas.

El objetivo del presente trabajo es hacer un estudio comparativo de los modelos estadísticos clásicos para analizar series de tiempo, regresión polinomial, suavización exponencial, promedio móvil con las redes neuronales artificiales. Para la aplicación se consideran ocho series de tiempo tomadas del INE. Para comparar los métodos propuestos se usan las técnicas para la medición del error de pronóstico: Error porcentual absoluto medio (MAPE), Desviación absoluta media (MAD) y Error cuadrático medio (MSE).

2. Materiales y métodos

Materiales y Equipo

Los materiales y equipos que se utilizarán en este trabajo de investigación son:

- Una computadora con un procesador Intel Core i7. CPU 3.4 GHz. RAM 4 GB.
- Programa de análisis estadístico MINITAB para el análisis de las series de tiempo.
- El programa WEKA (Waikato Environment for Knowledge Analysis) es un software libre distribuido bajo licencia GNU-GPL y escrito en Java; desarrollado por la Universidad de Waikato de Nueva Zelanda. WEKA es un conjunto de algoritmos de aprendizaje automático para las tareas de minería de datos (Weka, 2010). Este programa se usará para el aprendizaje con redes neuronales.

Descripción de las series de tiempo

Para el estudio comparativo se utilizan ocho series de

tiempo tomados del portal del INEI para el período 2009 – 2012.

Serie 1. Producción en la construcción.

Variable: índice de la producción en la construcción (Base 1994=100.0).

Fuente: Instituto Nacional de Estadística e Informática.

Serie 2. Producción de electricidad.

Variable: Producción de electricidad (Gigawatt/hora).

Fuente: Ministerio de Energía y Minas.

Serie 3. Fabricación de papel y productos de papel.

Variable: índice de fabricación de papel y productos (Base 1994=100.0).

Fuente: Ministerio de la Producción - Viceministerio de MYPE e Industria.

Serie 4. Fabricación de productos textiles.

Variable: índice de producción textil (Base 1994=100.0).

Fuente: Ministerio de la Producción - Viceministerio de MYPE e Industria.

Serie 5. Producción minera e hidrocarburos.

Variable: índice de producción minera e hidrocarburos (Base 1994=100.0).

Fuente: Ministerio de Energía y Minas – PERUPETRO.

Serie 6. Producción de madera y productos de madera.

Variable: índice de producción de madera (Base 1994=100.0).

Fuente: Ministerio de la Producción - Viceministerio de MYPE e Industria.

Serie 7. Fabricación de prendas de vestir.

Variable: índice de prendas de vestir (Base: 1994=100.0).

Fuente: Ministerio de la Producción - Viceministerio de MYPE e Industria.

Serie 8. Producción sector fabril.

Variable: índice de producción del sector fabril (Base 1994=100.0).

Fuente: Ministerio de la Producción - Viceministerio de MYPE e Industria.

Metodología

Los métodos estadísticos de regresión polinomial de grado 2, promedio móvil, suavización exponencial simple y suavización exponencial doble y el modelo de red neuronal multicapa son aplicados a ocho series de tiempo.

Modelo de red neuronal para series de tiempo

En la literatura se proponen diferentes estructuras de redes neuronales para analizar y obtener pronósticos de las series de tiempo. El modelo de red neuronal multicapa (RNM) es el más utilizado, comprende un conjunto de

nodos (unidades de procesamiento) los cuales están organizados y distribuidos en tres capas: una capa de entrada, al menos una capa oculta y una capa de salida (Zhang *et al.*, 1998).

Alon *et al.* (2001) realizan una comparación de las redes neuronales artificiales con los métodos de suavización exponencial Winter, modelos de Box-Jenkins ARIMA y regresión multivariada para analizar series de tiempo sobre ventas al por menor. Los resultados indican que en promedio las redes neuronales obtienen pronósticos más precisos que los métodos tradicionales, siguiendo los modelos ARIMA. El modelo de red neuronal permite identificar patrones de tendencia y estacionalidad no lineales en las series de tiempo analizadas.

Singh y Kumar (2010) muestran un método para diseñar pronósticos usando redes neuronales para series de tiempo de los precios de stock de mercado. Los pronósticos conseguidos con la red neuronal son comparados con la técnica de regresión múltiple. La capacidad predictiva para ambos métodos se evalúa con técnicas de medición del error (MAPE, MSE y RMSE). Los resultados muestran que las redes neuronales, cuando son entrenadas con suficientes datos de entrada y seleccionando una topología adecuada, arrojan predicciones de los precios de stock con mayor precisión que la técnica estadística.

Un modelo RNM con p nodos de entradas, h nodos en la capa oculta y un nodo de salida puede ser usado para analizar una serie de tiempo. Una serie de tiempo y_t ; $t = 1, 2, 3, \dots, n$ se puede modelar con una RNM con p variables de entradas y_{t-i} ($i = 1, 2, \dots, p$) y relacionarla con una variable de salida y_t . El modelo RNM para el análisis de series de tiempo se indica en la siguiente expresión:

$$y_t = \alpha_0 + \sum_{j=1}^h \alpha_j G \left(\beta_{0j} + \sum_{i=1}^p \beta_{ij} y_{t-i} \right) + e_t \quad (1)$$

Donde:

y_t Valor observado de la serie de tiempo en el instante t . Representa el valor para la variable de salida de la RNM.

y_{t-i} Son los valores rezagados de la serie de tiempo en el instante t . Representan los valores de las p variables de entrada en la RNM.

α_j Representan los pesos de la capa oculta a la capa de salida. Siendo α_0 el respectivo sesgo.

β_{ij} Son coeficientes que representan los pesos de la capa de entrada a la capa oculta. Siendo β_0 el respectivo sesgo.

G Representa la función de activación o transferencia de la capa de entrada, la cual determina la salida de la capa oculta. Las funciones de activación que generalmente son usadas son la función logística para la capa de entrada y la función identidad para la capa de salida.

h Número de neuronas en la capa oculta.

p Número de neuronas en la capa de entrada. Su valor determina el número de rezagos con que se analizará la serie de tiempo.

e_t Son los errores aleatorios de modelo, los cuales se asumen que son independientes e idénticamente distribuidos con media cero y variancia constante.

La estructura de la red neuronal con p nodos de entradas, h nodos en la capa oculta y un nodo de salida, asociada a la expresión (1) se muestra en la **Fig. 1**.

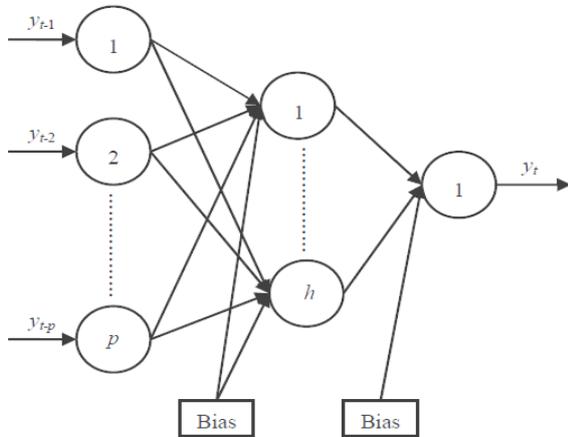


Figura 1. Modelo de red neuronal para series de tiempo

Características de la topología de la red neuronal multicapa:

- Tipo de conexión. Se considera la red neuronal con conexión hacia adelante; esto es, todas las neuronas se propagan hacia adelante a través de las capas de la red. No se consideran conexiones hacia atrás, laterales o autorecurrentes (una neurona de salida se aplica como su propia entrada).
- Capa de entrada. Es la capa que recibe los datos históricos de la serie de tiempo. El número de nodos o

variables de entrada (número de rezagos), se determinará por el número de autocorrelaciones significativas de la serie de tiempo. Si existen p autocorrelaciones significativas, las variables de entradas corresponden a los valores de las posiciones de la serie de tiempo: $y_{t-1}, y_{t-2}, y_{t-3}, \dots, y_{t-p}$.

- Capa oculta. Se considera una sola capa oculta, con el número de nodos igual al número de rezagos menos uno.
- Capa de salida. Conformada por un nodo en la capa de salida. La variable en la capa de salida corresponde al valor y_t de la serie de tiempo.
- Función de activación. La función de activación o transferencia usada para la capa de entrada es la Simoidal y para la capa de salida la identidad.

3. Resultados y discusión

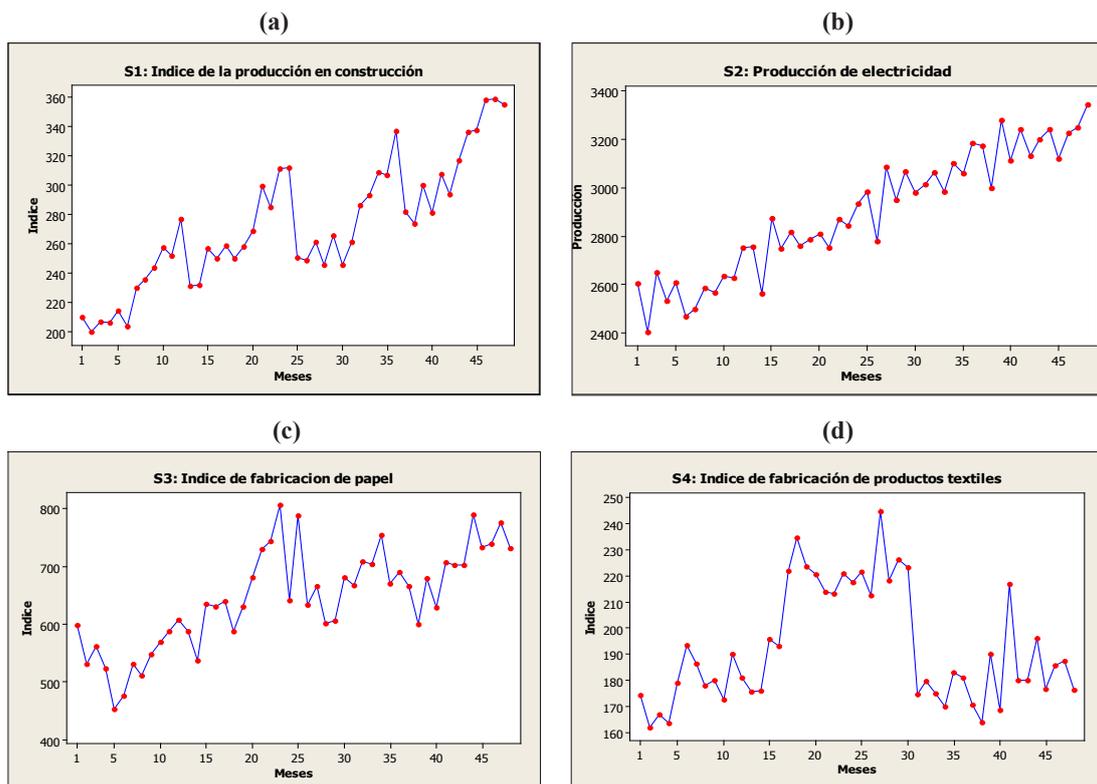
Para realizar el estudio comparativo de los métodos estadísticos clásicos para el análisis de series de tiempo con el modelo red neuronal, se presentan los ploteos de las ocho series de tiempo en la **Fig. 2** (a-h).

De los gráficos de las series de tiempo, se observa un comportamiento con tendencia y variación muy marcadas.

Modelos estadísticos para series de tiempo

Las series de tiempo son ajustadas a los siguientes cinco modelos estadísticos:

- M1: Polinomial de segundo grado.
- M2: Promedio móvil, centrado con tres términos.
- M3: Suavización exponencial simple, con constante de suavización 0,2.



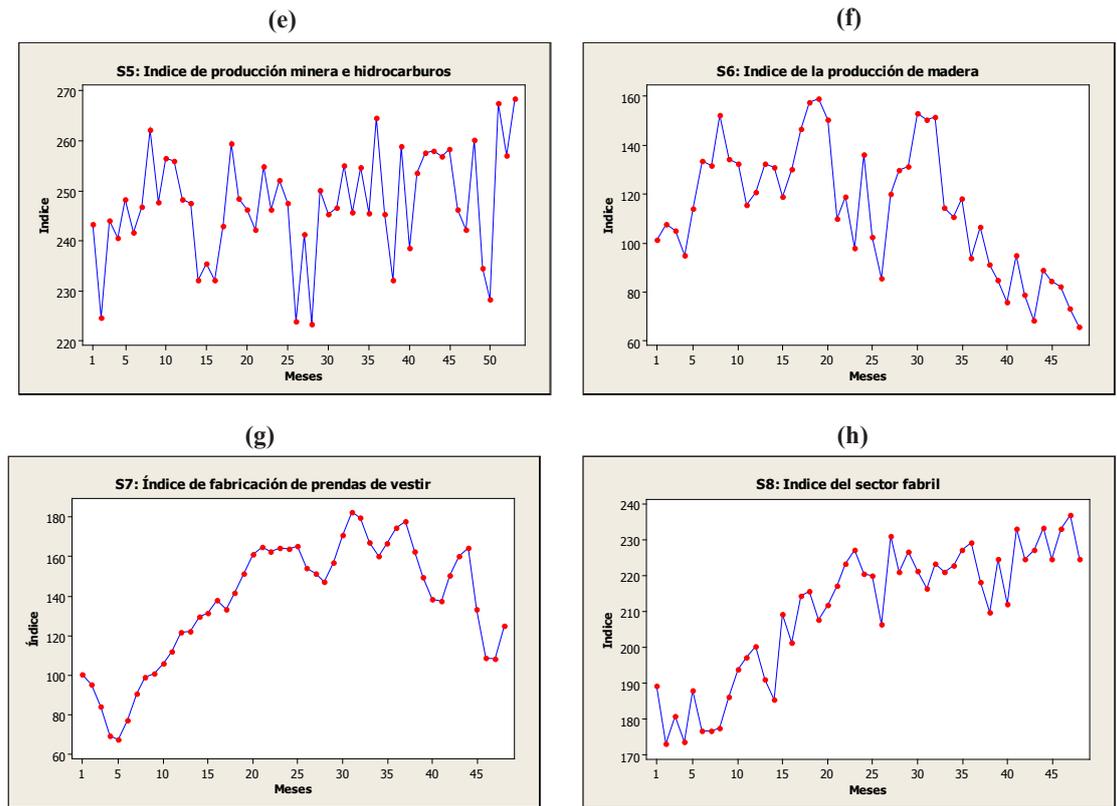


Figura 2. Gráficos de las ocho series de tiempo analizadas

M4: Suavización exponencial doble, con constantes de suavización 0,2 y 0,2.

Modelos de redes neuronales

Con la finalidad de determinar el número de nodos o variables de entrada del modelo de red neuronal, se obtuvo la función de autocorrelación para cada una de las ocho series (Anexo). Los análisis de las funciones de autocorrelación identifican entre dos y cuatro autocorrelaciones significativas. Por lo tanto, el número de nodos de entrada consideradas para los modelos de redes neuronales son dos y cuatro ($p=2$ y 4 rezagos). La topología de los modelos de red neuronal es definida de esta forma:

M5: Modelo de red neuronal, perceptron multicapa, función de activación logística (sigmoideal), con cuatro nodos de entrada ($p=4$ rezagos) y tres nodos en la capa oculta ($h=3$).

M6: Modelo de red neuronal, perceptron multicapa, función de activación logística (sigmoideal), con dos nodos de entrada ($p=2$ rezagos) y dos nodos en la capa oculta ($h=2$).

En las Figs. 3 y 4 se presentan las topologías de los modelos de red neuronal para los modelos M5 y M6, respectivamente.

Comparación de los modelos estadísticos y redes neuronales

Para realizar las comparaciones entre los modelos estadísticos y redes neuronales se usarán las medidas de medición de error MAPE, MAD y MSE. En las Tablas del 1 al 3, se muestran las comparaciones de las medidas

de error calculadas para los seis modelos ajustados para cada una de las ocho series de tiempo.

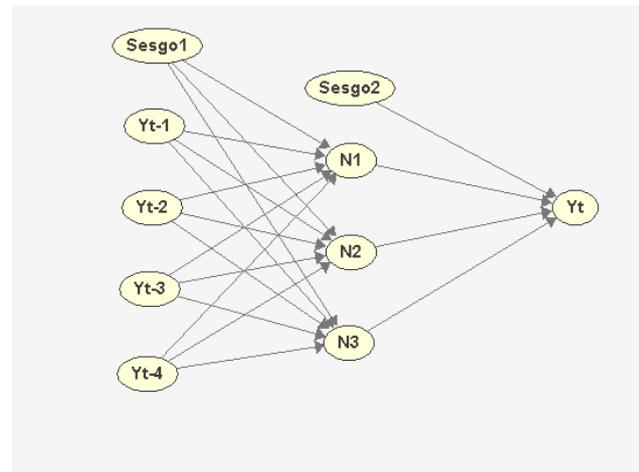


Figura 3. Red neuronal para el modelo M5

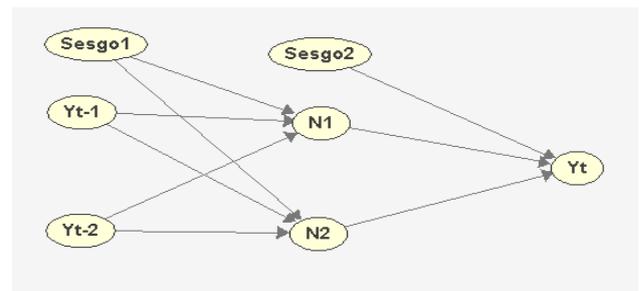


Figura 4. Red neuronal para el modelo M6

En la Tabla 1, se aprecian los modelos de redes neuronales (M5 y M6) que presentan los menores valores de MAPE en las ocho series de tiempo analizadas.

En la Tabla 2, se observa que a excepción del modelo M2, las redes neuronales M5 y M6 muestran menores valores de MAD en comparación con los métodos estadísticos (M1, M3 y M4).

En la Tabla 3, los modelos de redes neuronales (M5 y M6) resultan con menores MSE que los métodos estadísticos (M1, M3 y M4); salvo el método M2 en que son menores.

Finalmente, se han calculado los pronósticos para el mes de enero del 2013 para cada una de las series de tiempo con los seis modelos propuestos. Los pronósticos y los valores reales se presentan en la Tabla 4.

Tabla 1. Comparación de MAPE de modelos estadísticos y redes neuronales

Modelo	S1	S2	S3	S4	S5	S6	S7	S8
M1	6,45	2,10	6,34	7,59	2,96	11,72	8,70	2,89
M2	4,01	170	4,40	3,85	2,22	8,79	5,80	3,40
M3	7,34	3,20	7,40	6,83	3,24	1547	12,40	3,83
M4	6,98	2,40	7,90	7,99	3,41	16,51	12,10	3,53
M5	0,05	0,03	0,07	0,05	0,03	0,07	0,03	0,03
M6	0,06	0,03	0,06	0,05	0,03	0,12	0,05	0,03

Tabla 2. Comparación de MAD de modelos estadísticos y redes neuronales

Modelo	S1	S2	S3	S4	S5	S6	S7	S8
M1	17,6	60,8	40,4	14,6	7,2	13,0	10,5	5,9
M2	11,0	47,9	28,2	7,5	5,5	9,6	7,4	7,1
M3	20,9	94,0	48,3	13,2	7,9	16,6	15,8	8,1
M4	19,1	67,6	49,8	15,4	8,3	18,3	14,5	7,3
M5	15,2	82,2	44,0	9,2	7,4	7,5	4,4	5,8
M6	15,3	85,5	39,8	10,0	7,1	13,2	6,2	6,0

Tabla 3. Comparación de MSE de modelos estadísticos y redes neuronales

Modelo	S1	S2	S3	S4	S5	S6	S7	S8
M1	451,1	5700,4	2893,7	299,7	86,8	272,0	170,5	56,19
M2	190,8	3788,2	1149,0	103,0	47,9	134,0	79,5	72,87
M3	633,0	12398,0	3424,6	313,7	99,2	379,0	336,1	92,6
M4	537,5	7208,3	3696,6	412,9	112,5	464,9	290,6	75,5
M5	468,9	9915,4	2974,7	184,9	92,9	109,7	35,9	52,6
M6	393,0	11102,7	2330,9	179,8	83,8	249,5	68,2	56,4

Tabla 4. Comparación de pronósticos para el mes de enero del 2013

Modelo	S1	S2	S3	S4	S5	S6	S7	S8
M1	339,8	3298,8	720,4	162,8	255,1	57,6	120,9	226,1
M2	357,0	3273,2	748,6	183,1	264,3	73,7	113,9	23,5
M3	333,7	3217,0	728,1	183,7	254,3	80,7	134,1	227,9
M4	356,0	3280,8	749,9	179,5	254,7	63,1	118,6	230,2
M5	315,0	3156,0	706,0	186,0	249,0	75,0	147,8	228,1
M6	338,0	3105,0	715,9	180,5	250,1	80,0	1380	227,0
Valor real	333,7	3389,2	755,3	155,8	234,6	54,0	130,1	224,9

4. Conclusiones y recomendaciones

Las conclusiones y recomendaciones de la presente investigación son que los modelos de redes neuronales multicapa permiten analizar y obtener pronósticos precisos por sus capacidades auto-adaptativas, generalización, aprendizaje y representación de funciones no lineales en las series de tiempo con patrón de tendencia fluctuante en el tiempo. En la investigación, se consideraron dos topologías de RNM para las ocho series de tiempo. El modelo M5, con cuatro variables de entrada (4 rezagos) y tres nodos en la capa oculta. El modelo M6, con dos variables de entrada (2 rezagos) y dos nodos en la capa oculta. En los dos modelos se consideró la función logística en la capa de entrada y la identidad en la capa de salida.

Con respecto a los modelos de RNM M5 y M6, se observó que resultaron con menores valores de MAPE en las ocho series de tiempo, en comparación con los métodos estadísticos clásicos M1, M2, M3 y M4. Asimismo, los modelos M5 y M6 mostraron menores valores de MAD y MSE que los modelos M1, M3 y M4. Asimismo, los pronósticos obtenidos con los modelos M5 y M6 fueron más precisos en cuatro de las series de tiempo: S1, S5, S7 y S8.

En el modelamiento de las series de tiempo con redes neuronales, se debe considerar para una especificación correcta de la topología, aspectos como: número de rezagos, número de nodos de entrada, número de nodos en la capa oculta y tipo de función de transferencia.

Como recomendación, proponemos experimentar con otras clases de topologías de redes neuronales tales como redes recurrentes, redes dinámicas y redes de base radial, entre otras.

5. Literatura citada

Adhikari, Agrawal y Kant. 2013. *PSO based Neural Networks vs. Traditional Statistical Models for Seasonal Time Series Forecasting*. 3rd IEEE International Advance Computing Conference (IACC).

Alon, I.; Qi Min y Sadowski, R.J. 2001. Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services*, 8: 147-156.

Montaño, J.J.; Palmer, A. y Muñoz, P. 2011. Artificial neural networks applied to forecasting time series. *Psicothema*, 23(2): 322-329.

Nojek S.; Britos P.; Rossi B. y García Martínez, R. 2003. Pronóstico de Ventas: Comparación de predicción basada en redes Neuronales versus Método Estadístico. *Reportes Técnicos en Ingeniería del Software*, 5(1):1-12.

Sánchez, P. y Velásquez, J.D. 2010. Problemas de Investigación en la Predicción de Series de Tiempo con Redes Neuronales Artificiales. *Revista Avances en Sistemas e Informática*, 7(3): 67-73.

Singh, V.K. y Kumar, B. A. (2010). An Analysis of the Performance of Artificial Neural Network Technique for Stock Market Forecasting. (IJCSSE) *International Journal on Computer Science and Engineering*. 2(6):2104-2109.

Stern, H. 1996. Neural Network in Applied Statistics. *TECHNOMETRICS*, 38(3): 205-214.

Zhang, G.; Patuwo, B.E. y Hu, M.Y. 1998. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14:35-62.

Weka. 2010. Disponible en <http://www.cs.waikato.ac.nz/ml/weka/>

ANEXOS

Función de autocorrelaciones

