

Pronóstico del tipo de cambio USD/MXN con redes neuronales de retropropagación

Francisco D. Meneses-Bautista, Matías Alvarado

Centro de Investigación y de Estudios Avanzados del IPN,
Departamento de Computación,
Ciudad de México, México

fmeneses@computacion.cs.cinvestav.mx, matias@cs.cinvestav.mx

Resumen. El análisis de las series de tiempo permite caracterizar un fenómeno e incluso predecir, con un cierto grado de precisión, su comportamiento a futuro. La posibilidad de anticipar los movimientos de los mercados resulta sumamente atractiva para los responsables de la toma de decisiones, tanto en la iniciativa privada como en el sector público. En esta investigación se aborda la regresión de series de tiempo del tipo de cambio dólar estadounidense/peso mexicano, empleando un modelo generado por redes neuronales artificiales de retropropagación. Los resultados obtenidos ratifican empíricamente las ventajas de la utilización de las redes neuronales en el análisis y pronóstico de series de tiempo financieras, producto de las capacidades de aproximación de funciones y generalización que presentan dichas redes.

Palabras clave: pronóstico de tipos de cambio, redes neuronales artificiales, series de tiempo.

Forecasting USD/MXN Exchange Rate with Backpropagation Neural Networks

Abstract. The analysis of time series allows to characterize a phenomenon and even to predict, under a certain degree of precision, its future behavior. The possibility of anticipating market movements is extremely attractive for decision-makers, both in the private sector and in the public sector. This research addresses the regression of time series of the US dollar / Mexican peso exchange rate, using a model generated by backpropagation artificial neural networks. The results obtained ratify empirically the advantages of the use of neural networks in the analysis and forecast of financial time series, as a result of the capabilities of function approximation and generalization that these networks present.

Keywords: exchange rates forecasting, artificial neural networks, time series.

1. Introducción

Conocer con antelación el estado futuro de ciertas variables financieras permite obtener el mayor beneficio en un entorno de volatilidad. En las áreas de la economía y las finanzas, los especialistas han usado tradicionalmente diferentes herramientas estadísticas y econométricas de análisis de series de tiempo, con la finalidad de comprender mejor la mecánica inherente a fenómenos económicos complejos. Se han desarrollado métodos analíticos que pretenden explicar la manera en la que la interacción entre las variables independientes da forma al comportamiento observado (análisis fundamental). Sin embargo, también existen otros métodos que no buscan explicar las relaciones causales entre los factores y el fenómeno que producen, sino que sitúan el foco de interés en las propias series de tiempo y su gráfica (análisis técnico). Desde esta última perspectiva, la serie de tiempo contiene la suficiente información sobre cómo cambia el fenómeno a través del tiempo, de tal forma que es posible utilizarla para realizar regresiones y pronósticos. Son muy utilizados los modelos *ARIMA* (en especial, la metodología de Box-Jenkins [1]), los modelos de heterocedasticidad condicional autoregresiva *ARCH* y *GARCH*, y el modelo del *Paseo Aleatorio* [2].

Sin embargo, en las últimas décadas, el auge de las tecnologías de la información ha permitido desarrollar sistemas autoadaptativos de cómputo, como las redes neuronales artificiales, que son capaces de hallar, de manera automática, las correlaciones subyacentes en medio de grandes conjuntos de datos aparentemente inconexos. Esta habilidad es de gran valor para el análisis de regresión, puesto que permite determinar en qué grado influye un conjunto de variables independientes en el comportamiento de un fenómeno dado. Además, sometiendo a la red neuronal a un proceso de entrenamiento adecuado, es posible obtener un modelo de cómputo implícito que no sólo reproduzca los datos de la serie de tiempo, sino que también pueda ser utilizado para extrapolar puntos futuros de la serie. En este trabajo, se emplearon redes neuronales entrenadas con el algoritmo de retropropagación para analizar y pronosticar el comportamiento del tipo de cambio dólar estadounidense/peso mexicano (USD/MXN). El objetivo de los experimentos fue evaluar la capacidad predictiva de las redes neuronales de retropropagación, probando diferentes arquitecturas y conjuntos de datos: series de tiempo del tipo de cambio del peso mexicano respecto a varias divisas, tomadas en distintos periodos históricos.

El campo de regresión de series de tiempo con redes neuronales ha sido abordado en varias investigaciones y desde una amplia variedad de perspectivas. Zhang et al. [8] realizaron experimentos sobre el pronóstico del tipo de cambio GBP/USD con redes neuronales. Analizan cómo responden diferentes diseños de arquitecturas, cambiando tanto el tamaño de los vectores de entrada, como el conjunto de objetos de entrenamiento. Al final, concluyen que el número de nodos de entrada tiene un impacto importante en su modelo y que el índice de desempeño de la red (ya sea la raíz del error cuadrático medio *RMSE* o el porcentaje medio de error absoluto *MAPE*) sirve también para determinar el tamaño del conjunto de objetos de entrenamiento usado. Al final determinan que la red neuronal tiene una mayor capacidad predictiva que el modelo del

Paseo Aleatorio, sobre todo, cuando se trata de periodos mensuales, puesto que en periodos más prolongados, ambos métodos obtienen resultados muy similares.

En el trabajo de Yao et al. [7], se consigue hacer una aproximación de tipos de cambio con tasas de acierto superiores al 70 por ciento, empleando redes neuronales de retropropagación (*BPNN*). Los autores desarrollan un modelo para pronosticar las variaciones en los tipos de cambio, que incorpora algunas variables de tipo cualitativo. Además, complementan su investigación implementando un simulador de inversiones.

En años más recientes, Saini et al. [6] realizaron una comparación entre la capacidad de predicción de tipos de cambio que tienen las redes *BPNN* y redes neuronales recurrentes (*RNN*, *Recurrent Neural Network*). Calcularon independientemente la regresión de cuatro series de tiempo, de divisas distintas. Los experimentos se hicieron sobre una red *BPNN* y una *RNN*, ambas con arquitectura fija. Los parámetros del experimento fueron el factor de aprendizaje, la cantidad de épocas de entrenamiento y el *momentum*. Sus resultados confirman la habilidad de ambos tipos de redes neuronales para aproximar satisfactoriamente las series de tiempo. Sin embargo, las redes *RNN* tuvieron, en general, un mejor desempeño en sus experimentos.

En el trabajo de Galeshchuk [4] se analiza el efecto que se genera sobre la predicción al utilizar series de tiempo con diferentes intervalos de muestreo. El modelo de pronóstico utilizado en sus experimentos es un perceptrón multicapa, y se alimenta con series de tiempo con muestras de frecuencia diaria, mensual y cuatrimestral. Los resultados empíricos obtenidos indican que la red entrega una mejor predicción cuando el intervalo solicitado es a corto plazo.

2. Pronóstico con redes neuronales

Una red neuronal artificial es un modelo de cómputo autoadaptativo y bioinspirado. En ella, se trata de imitar la gran cantidad de interconexiones que presentan las células nerviosas, de manera que el arreglo global de células individuales simples alcanza niveles de procesamiento muy elevados, brindándole al organismo la capacidad de enfrentarse a problemas muy complejos, con gran eficacia y precisión.

Las neuronas artificiales son las unidades básicas de procesamiento en las redes neuronales artificiales. Cada neurona artificial efectúa una evaluación de los valores que recibe como entrada, a través de una *función de activación*. Suele elegirse como función de activación alguna que sea de tipo no lineal y continua. Las neuronas artificiales se unen entre sí por medio de enlaces sinápticos, ponderados por un conjunto de *pesos*. Se llama *capa* a la agrupación de neuronas que se encuentran en el mismo nivel de profundidad en la red (Figura 1). Las señales de salida de una capa alimentan a todas las neuronas de la capa siguiente. La función de salida de la red neuronal es una composición de las funciones de activación de todas las neuronas en la red.

El problema del *aprendizaje* de la red neuronal consiste en encontrar un conjunto adecuado de pesos, que permita a la red producir la salida esperada.

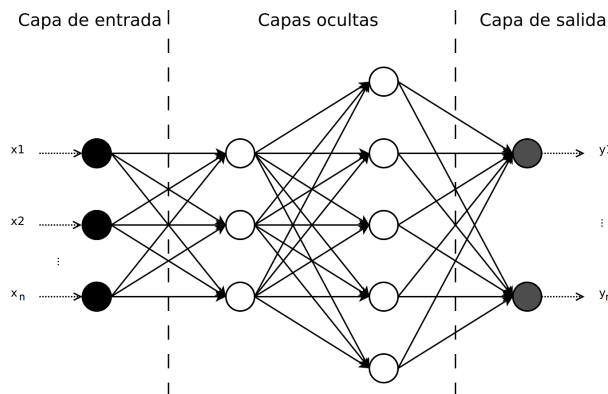


Fig. 1. Red neuronal artificial de cuatro capas.

En el *entrenamiento* por retropropagación, se busca minimizar el error global de la red, expresado como la diferencia entre la salida producida por la red y la salida esperada, ante un conjunto conocido de patrones de entrada. El proceso de retropropagación calcula la aportación al error global proveniente de cada neurona en la capa de salida, reajusta el conjunto de pesos de la capa de salida de acuerdo a un *factor de aprendizaje* y al gradiente de la función de error, y continúa con la capa anterior (aquella que está directamente conectada con la capa de salida). Este procedimiento sigue con cada capa de la red hasta llegar a la capa de entrada. Tras cada *época* del entrenamiento, el error de la red va reduciéndose hasta llegar al mínimo. Después de un entrenamiento exitoso, la red neuronal es capaz de reproducir las salidas esperadas con un nivel de error despreciable, de manera que puede ser empleada como un aproximador universal de funciones [3,5].

Las habilidades de las redes neuronales de generalización y de aproximación de funciones son muy importantes para las tareas de regresión y pronóstico con series de tiempo. La red neuronal entrenada puede concebirse como un modelo de regresión implícito, que reproduce las observaciones de la serie de tiempo al ser alimentado con un conjunto de variables independientes. Entonces, para utilizar este modelo en el pronóstico, se debe tomar como entrada valores de las variables independientes hasta el momento t y entrenar a la red para que produzca el valor de la serie de tiempo en el momento $t + 1$ como salida.

3. Desarrollo experimental

3.1. Variables independientes y conjunto de datos

Sea $F_t = f(\mathbf{x}_t)$ el valor de la función f en el momento t , y \mathbf{x}_t el vector cuyas componentes $(x_t^1, x_t^2, \dots, x_t^n)$ representan al conjunto de las n variables independientes que describen al fenómeno, en el momento t .

El modelo propuesto en esta investigación establece que la función pronóstico g , calculada mediante redes neuronales, determina las correlaciones que permiten inferir el siguiente valor de la serie de tiempo (F_{t+1}), dadas las variables independientes $x_t^1, x_t^2, \dots, x_t^n$, junto al valor de la función F_t . El valor siguiente de la serie de tiempo F_{t+1} , obtenido de la función g así definida, queda expresado como

$$F_{t+1} = g(\mathbf{x}_t, F_t) = g(x_t^1, x_t^2, \dots, x_t^n, F_t).$$

La decisión sobre cuáles variables integrarán los datos de entrada puede realizarse a través de un análisis causal o estadístico. Sin embargo, el enfoque adoptado en este trabajo consistió en probar diferentes combinaciones de series de tiempo como variables independientes. Al finalizar el entrenamiento de las redes neuronales, puede calcularse un índice de error y asociarlo a cada conjunto de variables independientes. De esta manera, el mejor conjunto para realizar el pronóstico es aquél que produce el menor grado de error durante la validación.

Los datos empleados en los experimentos fueron obtenidos del sitio web del Banco de México¹. En el sitio se hayan disponibles valores diarios del tipo de cambio peso/dólar y los precios de las divisas que integran la canasta del DEG (Derecho Especial de Giro), a saber, el dólar estadounidense, el euro, la libra esterlina, el yen japonés y el yuan chino. Para el desarrollo de los experimentos fueron elegidas como variables independientes cuatro de las divisas mencionadas: el dólar estadounidense, el euro, la libra esterlina y el yen japonés. El yuan chino fue descartado debido a la escasez de los datos disponibles². A partir de estas cuatro series de tiempo, se generaron varios conjuntos de vectores de entrenamiento. Los objetos que integran los conjuntos de entrenamiento pueden describirse como vectores de tipo:

$$\mathbf{x}_t = (x_t^{\text{USD/MXN}}, x_t^{\text{EUR/MXN}}, x_t^{\text{GBP/MXN}}, x_t^{\text{JPY/MXN}}), \quad \text{con } \mathbf{x}_t \in \mathbb{R}^4,$$

donde

- $x_t^{\text{USD/MXN}} \in \mathbb{R}$, es el valor del tipo de cambio dólar estadounidense / peso mexicano, en el instante t ,
- $x_t^{\text{EUR/MXN}} \in \mathbb{R}$, es el valor del tipo de cambio euro / peso mexicano, en el instante t ,
- $x_t^{\text{GBP/MXN}} \in \mathbb{R}$, es el valor del tipo de cambio libra esterlina / peso mexicano, en el instante t ,
- $x_t^{\text{JPY/MXN}} \in \mathbb{R}$, es el valor del tipo de cambio yen japonés / peso mexicano, en el instante t .

Se probaron cinco diferentes conjuntos de objetos de entrenamiento. El primero de ellos está formado por los vectores descritos, con cuatro componentes. Los otros conjuntos de objetos de entrenamiento contienen vectores cuyas componentes son una combinación de sólo tres de las cuatro divisas elegidas anteriormente

¹ <http://www.banxico.org.mx>

² Sólo se tienen registros del yuan chino desde octubre de 2016 (año en que se integró al DEG).

(Tabla 1). La hipótesis inicial sobre los conjuntos de entrenamiento fue que, dada una red neuronal con una arquitectura determinada, los datos de las cuatro series de tiempo permiten obtener los mejores resultados en el pronóstico, ya que al descartar cualquiera de las series, se afecta el modelo calculado por la red.

Tabla 1. Componentes de los vectores en los conjuntos de entrenamiento.

Conjunto	VARIABLES INDEPENDIENTES
00deyl	USD/MXN, EUR/MXN, GBP/MXN, JPY/MXN
01eyl	EUR/MXN, GBP/MXN, JPY/MXN
02dy1	USD/MXN, GBP/MXN, JPY/MXN
03del	USD/MXN, EUR/MXN, JPY/MXN
04dey	USD/MXN, EUR/MXN, GBP/MXN

El conjunto de objetivos para el entrenamiento está constituido por series de tiempo con valores $y_t = x_{t+1}^{\text{USD/MXN}} \in \mathbb{R}$, es decir, el tipo de cambio dólar estadounidense/peso mexicano en el instante $t + 1$.

Los datos de las series de tiempo empleadas se componen de muestras diarias que corresponden al periodo entre el 2 de Enero de 2015 y el 24 de Marzo de 2017. De la totalidad de estos datos, se formaron tres subconjuntos de entrenamiento con distinta cantidad de objetos: el primero con datos en un periodo semestral; el segundo, anual y el último, a dos años (Tabla 2). Para la validación (pronóstico) se eligieron, en los tres casos, series de tiempo con 30 muestras (equivalentes, aproximadamente, a un periodo de mes y medio), a partir del día inmediatamente posterior al último objeto en cada serie de entrenamiento.

Tabla 2. Detalle de los periodos elegidos para entrenamiento y pronóstico.

Periodo	Entrenamiento			Pronóstico		
	Desde	Hasta	Objetos	Desde	Hasta	Objetos
Semestral	01-Jul-2016	31-Ene-2017	149	01-Feb-2017	15-Mar-2017	30
Anual	04-Ene-2016	30-Dic-2016	252	02-Ene-2017	13-Feb-2017	30
Dos años	02-Ene-2015	08-Feb-2017	530	09-Feb-2017	23-Mar-2017	30

Los valores de las series de tiempo recibieron un procesamiento previo para mejorar el desempeño durante el entrenamiento de las redes neuronales. Se realizó un escalamiento en todos los valores de las series, de tal manera que a cada muestra se le hace corresponder un valor real en el intervalo $[0, 1]$, muy

apropiado para usarse en las redes de retropropagación con función de activación sigmoïdal logística.

En el caso particular del yen japonés, la serie de tiempo fue sometida, en primer lugar, a un escalamiento, por un factor de 100, debido que el valor nominal del yen, desde el año 2015, ha estado alrededor de los 15 centavos de peso mexicano. Posteriormente, todas las series de tiempo fueron transformadas de acuerdo al procedimiento que se describe a continuación.

Sea F_t el valor de una serie de tiempo en el momento t , y sean F_{\min} y F_{\max} dos valores constantes, extremos del intervalo $[F_{\min}, F_{\max}]$, y tales que F_{\min} es menor o igual que todas las observaciones que componen a las series de tiempo y, de manera similar, F_{\max} es mayor o igual que todos las muestras de dichas series. Entonces, el valor escalado $F_{N(t)}$ que representa al valor F_t de la serie de tiempo original, puede ser expresado de esta manera:

$$F_{N(t)} = \frac{F_t - F_{\min}}{F_{\max} - F_{\min}}, \quad 0 \leq F_{N(t)} \leq 1.$$

Los valores elegidos para delimitar el intervalo $[F_{\min}, F_{\max}]$, anteriormente descrito, fueron $F_{\min} = 10.00$ y $F_{\max} = 30.00$. El valor de F_{\min} se determinó al observar que el valor mínimo en las series de tiempo corresponde al precio que presentó el yen el día 6 de Marzo de 2015, con un valor nominal de 0.1221 pesos, que sujeto al escalamiento inicial de la serie de tiempo del yen, equivale a 12.12 pesos. El valor F_{\min} elegido brinda una holgura de 2 pesos, aproximadamente. En el caso del valor F_{\max} elegido, éste supera por casi 3 pesos al valor máximo presente en las series de tiempo, correspondiente al valor nominal de la libra el día 12 de Febrero de 2016, cuando se cotizó en 27.6716 pesos. El intervalo holgado escogido hace que sea posible realizar pronósticos que rebasen los límites mínimo y máximo de los valores de las series de tiempo.

3.2. Diseño de la red neuronal de retropropagación

Denótese como N a la función de pronóstico obtenida por la red neuronal después del entrenamiento. Entonces, la definición de la función N en términos de la funciones de regresión f y de pronóstico g , anteriormente mencionadas, es

$$N(\mathbf{x}_t) = g(\mathbf{x}_t, F_t) = f(\mathbf{x}_{t+1}) = y_t,$$

que, expresada en términos de las cuatro variables independientes utilizadas, resulta

$$N(\mathbf{x}_t) = N(x_t^{\text{EUR/MXN}}, x_t^{\text{GBP/MXN}}, x_t^{\text{JPY/MXN}}, x_t^{\text{USD/MXN}}) = y_t = x_{t+1}^{\text{USD/MXN}}.$$

Con el fin de encontrar una arquitectura que modelara apropiadamente la función N , se ensayaron redes con distinta cantidad de capas y neuronas (Figura 2). Para todas las arquitecturas de red probadas, el número de entradas fue tres o cuatro, dependiendo del conjunto de variables independientes utilizado. La capa de salida siempre fue construida con una única neurona.

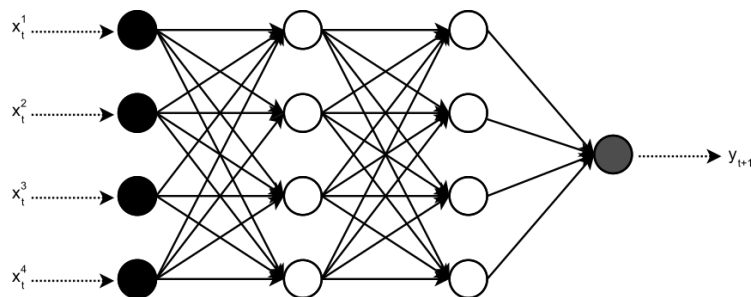


Fig. 2. Red de retropropagación de arquitectura a4441.

Se probaron en total doce arquitecturas para cada caso de pronóstico, con una o dos capas ocultas, y cada capa contó con dos, cuatro u ocho neuronas (Tabla 3). Se empleó la función sigmoial logística como función de activación en todas las neuronas de la red ($\text{logsig}(x) = \frac{1}{1+e^{-x}}$).

Tabla 3. Arquitecturas probadas en los experimentos.

Arquitecturas		Neuronas por capa oculta	
3 entradas	4 entradas	1ª capa	2ª capa
a321	a421	2	0
a341	a441	4	0
a381	a421	8	0
a3221	a4221	2	2
a3241	a4241	2	4
a3281	a4281	2	8
a3421	a4421	4	2
a3441	a4441	4	4
a3481	a4481	4	8
a3821	a4821	8	2
a3841	a4841	8	4
a3881	a4881	8	8

3.3. Entrenamiento

El proceso de entrenamiento fue de tipo on-line: para cada patrón de entrenamiento evaluado por la red, se calculó el error cuadrático medio (MSE) como estimador del error global de la red, y con este índice de error, se efectuó el ajuste de pesos por retropropagación. El factor de aprendizaje fue fijado a $\eta = 0.4$. El proceso fue restringido a 100,000 épocas de entrenamiento, y el error mínimo

deseado fue de $RMSE=0.001$ (raíz del error cuadrático medio acumulado en una época). Se registró el grado de error obtenido por cada arquitectura, tanto en la regresión de la serie de tiempo de entrenamiento, como en la de pronóstico.

4. Resultados

En la Tabla 4 se muestran los resultados generales de las etapas de entrenamiento y de regresión. Para ambas etapas, se presenta el $RMSE$ promedio obtenido, por todas las arquitecturas y periodos de entrenamiento. Cabe mencionar que el error mínimo deseado para el entrenamiento no fue alcanzado por ninguna arquitectura (el $RMSE$ mínimo obtenido fue de 0.00453514, con la arquitectura *a4881* (4 entradas, 8 neuronas en la primera capa oculta, 8 neuronas en la segunda capa oculta y una salida), con el conjunto de variables *00deyl* (dólar-euro-yen-libra) y con datos de periodo anual).

Tabla 4. RMSE promedio del entrenamiento y la regresión.

	Variables Entrenamiento	Regresión
00deyl	0.0061432136	0.0096681008
01eyl	0.0092615111	0.0168266333
02dyl	0.0063488836	0.0094228061
03del	0.0062851569	0.0097144117
04dey	0.0063465556	0.0101616786

De los resultados obtenidos en la etapa de pronóstico fue posible identificar dos casos distintos: el pronóstico de 30 muestras, con un mayor grado de error, y el pronóstico de 1 muestra, con el menor grado de error. En la Tabla 5 se muestra el $RMSE$ en el pronóstico, empleando las series de tiempo de validación completas (con las 30 muestras), para cada conjunto de variables utilizado. También se incluyen los valores mínimo, máximo y promedio en cada caso. En la Tabla 6 se pueden observar los mismos datos, esta vez, para el pronóstico de la primera muestra en las series de tiempo de validación, es decir, el pronóstico al siguiente día).

4.1. Análisis

De los datos mostrados en la Tabla 5, se puede destacar la siguiente información sobre el pronóstico de 30 muestras:

- El mejor pronóstico se presentó al usar las variables *02dyl* (dólar-yen-libra), la arquitectura *a341* (3 entradas, 4 neuronas en una capa oculta y 1 salida), con un periodo de un año en las series de entrenamiento y obteniendo un $RMSE$ de 0.00842705 (Figura 3).

Tabla 5. RMSE del pronóstico de 30 muestras.

RMSE por cada periodo

Variables	Menor RMSE			Mayor RMSE		
	Arquitectura	Periodo	RMSE	Arquitectura	Periodo	RMSE
00deyl	a481	1y	0.00846566	a4881	6m	0.0139388
01eyl	a341	2y	0.0146082	a341	1y	0.0789643
02dyl	a341	1y	0.00842705	a381	2y	0.0129769
03del	a3841	1y	0.00853617	a3821	2y	0.0164238
04dey	a341	1y	0.00908241	a3881	6m	0.013041

RMSE promedio de todos los periodos

Variables	Menor RMSE		Mayor RMSE	
	Arquitectura	RMSE	Arquitectura	RMSE
00deyl	a4821	0.0124279333	a4221	0.0106122233
01eyl	a3281	0.0178648667	a341	0.0382287333
02dyl	a3221	0.0106841667	a3821	0.0115879667
03del	a341	0.0106401867	a3821	0.0123929067
04dey	a3241	0.0098830933	a3821	0.0110833733

- El pronóstico con el mayor grado de error fue producido por la red con arquitectura *a341* (3 entradas, 4 neuronas en una capa oculta y 1 salida), con el conjunto de variables *01eyl* (euro-yen-libra), con datos de entrenamiento de un año, y con *RMSE* de 0.0789643 (Figura 4).
- Considerando el valor promedio de los tres periodos de entrenamiento, el mejor caso de pronóstico se obtiene de la red con arquitectura *a3241* (3 entradas, 2 neuronas en la primera capa oculta, 4 neuronas en la segunda capa oculta y 1 salida), usando como el conjunto de variables *04dey* (dólar-euro-yen), con valor *RMSE* de 0.0098830933.

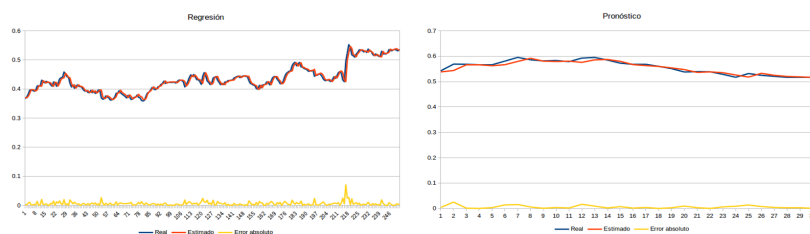


Fig. 3. Regresión y pronóstico a341(02dyl).

Tabla 6. RMSE del pronóstico de 1 muestra (al día siguiente).

RMSE por cada periodo

Variables	Menor RMSE			Mayor RMSE		
	Arquitectura	Periodo	RMSE	Arquitectura	Periodo	RMSE
00deyl	a4881	2y	0.000442617	a4821	6m	0.0139143
01eyl	a3221	2y	0.000981482	a3841	1y	0.037046
02dyl	a3841	2y	0.000979935	a3221	1y	0.025883
03del	a3441	2y	0.00257343	a3821	6m	0.00879175
04dey	a3421	2y	0.003174	a3841	6m	0.0156557

RMSE promedio de todos los periodos

Variables	Menor RMSE		Mayor RMSE	
	Arquitectura	RMSE	Arquitectura	RMSE
00deyl	a4881	0.0061538857	a4221	0.0076484267
01eyl	a3441	0.00536915	a3881	0.0164226433
02dyl	a341	0.00756548	a3221	0.0122008333
03del	a3481	0.0052086567	a3821	0.00672563
04dey	a321	0.0079934933	a3821	0.0091191733

Por otra parte, en la Tabla 6, para el caso del pronóstico a 1 día, se observa que:

- El mejor pronóstico se presentó al usar las variables *00deyl* (dólar-euro-yen-libra), la arquitectura *a4881* (4 entradas, 8 neuronas en la primera capa oculta, 8 neuronas en la segunda capa oculta y 1 salida), con un periodo de 2 años en las series de entrenamiento y obteniendo un *RMSE* de 0.000442617 (Figura 5).
- El pronóstico con el mayor grado de error fue producido por la red con arquitectura *a3841* (3 entradas, 8 neuronas en la primera capa oculta, 4 neuronas en la segunda capa oculta y 1 salida), con el conjunto de variables

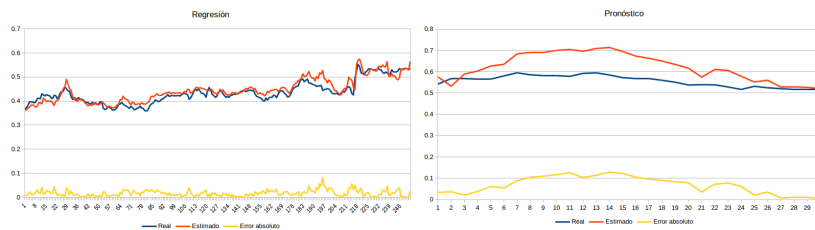


Fig. 4. Regresión y pronóstico a341(01eyl).

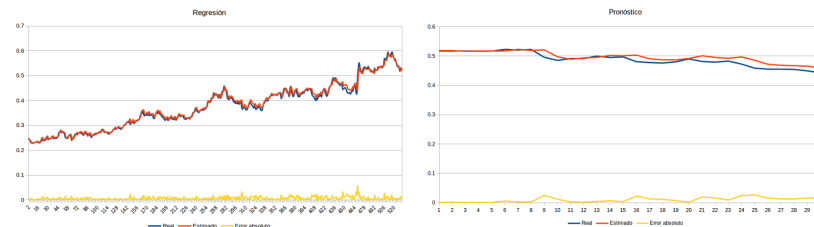


Fig. 5. Regresión y pronóstico a4881(00deyl).

01eyl (euro-yen-libra), con datos de entrenamiento de un año, y con $RMSE$ de 0.037046 (Figura 6).

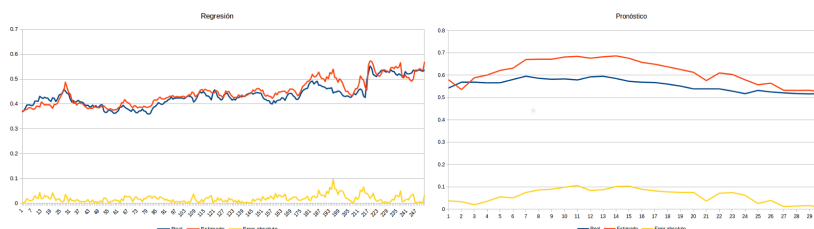


Fig. 6. Regresión y pronóstico a3841(01eyl).

- Considerando el valor promedio de los tres periodos de entrenamiento, el mejor pronóstico se obtiene de la red con arquitectura $a3481$ (3 entradas, 4 neuronas en la primera capa oculta, 8 neuronas en la segunda y 1 salida), usando como el conjunto de variables $03del$ (dólar-euro-libra), con valor $RMSE$ de 0.0052086567.

Al evaluar el pronóstico a 30 muestras calculado por las diferentes redes, se encontró que la mejor predicción fue producida por la arquitectura $a341$ (3 entradas, 4 neuronas en una capa oculta y 1 salida). Sin embargo, el grado de error obtenido es equivalente, en pesos mexicanos, a una variación de $(F_{\max} - F_{\min})(RMSE) = (20.0)(0.00842705) = 0.168541$, es decir, aproximadamente 17 centavos de peso mexicano. Este elevado grado de error puede ser atribuido al modelo de pronóstico propuesto, el cual está diseñado para pronosticar el valor de la variable USD/MXN únicamente al día siguiente. Además, es necesario notar que la misma arquitectura reportó el mayor grado de error con el conjunto de variables $01eyl$ (euro-yen-libra), por lo que, con estos datos, no se puede llegar a una afirmación concluyente sobre la capacidad predictiva de esta arquitectura.

Un razonamiento similar se sigue al evaluar los conjuntos de entrenamiento. Los resultados obtenidos sugieren que el mejor conjunto de variables independientes es el conjunto $02dyl$ (dólar-yen-libra), con datos dentro de un periodo

anual. Esta evidencia empírica rechaza, inicialmente, la hipótesis establecida en la sección 3.1. Sin embargo, al observar los resultados obtenidos, no se puede identificar una tendencia relevante que permita argumentar firmemente a favor o en contra de ese conjunto de variables, dado que, usando la arquitectura equivalente *a441* (4 entradas, 4 neuronas en una capa oculta y 1 salida) y el mismo periodo anual en los datos de entrenamiento, pero con el conjunto de datos *00deyl* (dólar-euro-yen-libra), se obtiene un *RMSE* de 0.00846566, es decir, una diferencia de 0.00003861, equivalente a 0.07722 centavos de peso mexicano.

En cuanto al pronóstico a 1 día, es notable la disminución en el grado de error obtenido, con respecto al pronóstico a 30 días, pasando de 0.00842705 a 0.000442617, en los mejores pronósticos para cada caso. Esta disminución del grado de error puede explicarse por la manera en que fueron entrenadas las redes neuronales: la serie de tiempo modelada por la red corresponde al valor del tipo de cambio dólar estadounidense/peso mexicano al día siguiente. Además, al comparar el *RMSE* promedio obtenido, puede afirmarse que el pronóstico a 1 día (con *RMSE* promedio de 0.0052086567) supera en calidad al pronóstico de 30 observaciones (con *RMSE* promedio de 0.0098830933).

Los resultados del pronóstico a 1 día indican que la arquitectura con mejores capacidades predictivas es *a4881* (4 entradas, 8 neuronas en la primera capa oculta, 8 neuronas en la segunda capa oculta y 1 salida), que es también, la arquitectura más compleja de las que se probaron. El *RMSE* de 0.000442617 (equivalente a un error absoluto de 0.885234 centavos de peso mexicano), obtenido por esta arquitectura, supera al siguiente mejor pronóstico, que se consiguió con las variables *02dyl* (dólar-yen-libra) y la arquitectura *a3841*, con datos en periodo de 2 años y *RMSE* de 0.000979935. La diferencia entre los dos pronósticos fue de 0.000537318, es decir, más del doble del *RMSE* obtenido en el primer caso. Al comparar los conjuntos de entrenamiento, el mejor pronóstico se consiguió con el conjunto *00deyl* (dólar-euro-yen-libra), apoyando la hipótesis de la sección 3.1.

Del análisis de los pronósticos con mayor grado de error, es importante hacer notar que, tanto para el caso de una muestra como para 30, el conjunto de datos *01eyl* (euro-yen-libra), con datos en periodo de un año, se ve asociado con esos pronósticos de baja calidad. Dicho conjunto es el único que no incluye la variable USD/MXN. Este resultado puede indicar que la información que aporta el valor del tipo de cambio USD/MXN en un determinado día es esencial para calcular el pronóstico al día siguiente y, por lo tanto, no debe ser descartada.

5. Conclusiones

Los experimentos realizados prueban la habilidad que tienen las redes neuronales de retropropagación para modelar y reproducir de manera adecuada la serie de tiempo que describe la variación del tipo de cambio USD/MXN. En el pronóstico, las redes neuronales calcularon predicciones de mejor calidad en el corto plazo (al día siguiente) que en el caso de 30 muestras a futuro. Este deterioro en la calidad de los pronósticos puede haberse debido al modelo empleado, que sólo se enfoca en realizar la predicción al día siguiente. Los datos

empleados durante el entrenamiento de las redes fueron adaptados para concordar completamente con dicho modelo, y esto puede representar una dificultad para que se alcance un nivel de generalización aceptable que permita predecir correctamente varias muestras. Será necesario continuar con las investigaciones para determinar si una extensión en el modelo puede mejorar la capacidad de predicción en plazos mayores.

Como parte del análisis de las variables empleadas, los resultados del pronóstico a un día apoyan de decisión de utilizar de los cuatro tipos de cambio en el modelo, sin descartar ninguno, así como la importancia de incluir el mismo tipo de cambio peso/dólar para mejorar la predicción. Sin embargo, este resultado está limitado a las variables elegidas en esta investigación y sería interesante analizar el efecto del uso de un conjunto ampliado de variables independientes, que no necesariamente se restrinja a tipos de cambio.

El pronóstico a un día alcanzó un error absoluto menor a un centavo, en el mejor de los casos. Si bien, se puede considerar que es un grado de error aceptable, es necesario evaluar el modelo con más periodos de prueba, con la finalidad de observar si la calidad del pronóstico se mantiene a través del tiempo, ante las condiciones cambiantes que se presentan cotidianamente en los mercados de divisas.

Agradecimientos. Agradecemos al Consejo Nacional de Ciencia y tecnología y al Centro de Investigación y de Estudios Avanzados del IPN por el apoyo recibido para la realización de este trabajo de investigación.

Referencias

1. Box, G., Jenkins, G.: Time series analysis: forecasting and control. Holden-Day series in time series analysis, Holden-Day (1970)
2. Brockwell, P.J., Davis, R.A.: Introduction to time series and forecasting, pp. 5–12. Springer International Publishing, 3rd edn. (2016)
3. Cybenko, G.: Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems* 2(4), 303–314 (1989)
4. Galeshchuk, S.: Neural networks performance in exchange rate prediction. *Neurocomputing* 172, 446–452 (2016)
5. Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. *Neural Networks* 2(5), 359 – 366 (1989)
6. Saini, S.S., Parkhe, O., Khadtare, T.: Analysis of feedforward and recurrent neural network in forecasting foreign exchange rate. *Imperial Journal of Interdisciplinary Research* 2(6) (2016)
7. Yao, J., Tan, C.L.: A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing* 34(1-4), 79 – 98 (2000)
8. Zhang, G., Hu, M.Y.: Neural network forecasting of the british pound/us dollar exchange rate. *Omega* 26(4), 495–506 (1998)