

Redes neuronales para pronóstico de series de tiempo hidrológicas del Caribe colombiano

Neural networks for hydrological time series forecasting in the Colombian Caribbean

Brayan de la hoz, Olmar Canchano, Luis Coronado, Paola Sánchez Sanchez

Universidad Simón Bolívar, Barranquilla-Colombia

Open Access

Publicado:
1 julio de 2019

Correspondencia:
Psanchez9@unisimonbolivar.edu.co

Resumen

El modelado de series de tiempo hidrológicas ha recibido fuerte atención debido a la necesidad de estudiar el comportamiento de dichas series y predecir su comportamiento futuro. En esta medida las redes neuronales han demostrado ser una herramienta potente para el modelado de series de tiempo debido a sus capacidades de aprendizaje y adaptación que facilita la toma de decisiones asertivas. El objetivo de este artículo es desarrollar un modelo de redes neuronales y analizar su efectividad a la hora de hacer predicciones de la altitud de las olas medidas en seis diferentes puntos del caribe colombiano. La evidencia experimental demuestra que la mejor configuración de modelo de redes neuronales difiere para cada ubicación, no obstante, los niveles de error obtenidos en la predicción son muy precisos.

Palabras claves: Redes neuronales, Series de tiempo, Series de tiempo hidrológicas, Pronóstico.

Abstract

The hydrological time series modeling has received strong attention due to the need to study the behavior of these series and predict their future behavior. In this measure, neural networks have proven to be a powerful tool for modeling time series due to their learning and adaptation capabilities that facilitate assertive decision making. The objective of this article is to develop a model of neural networks and analyze their effectiveness in making predictions of the altitude of the waves measured in six different points of the Colombian Caribbean. Experimental evidence shows that the best configuration of the neural network model differs for each location, however, the error levels obtained in the prediction are very precise.

Keywords: Neural networks, Time series, Hydrological time series, Forecast.

Como citar (IEEE): B.De la hoz, O. Canchano, L. Coronado, P. Sánchez Sanchez, "Redes neuronales para pronóstico de series de tiempo hidrológicas del Caribe colombiano", *Investigación y Desarrollo en TIC*, vol. 10, no. 2, pp. 18-31., 2019

Introducción

Las Redes Neuronales Artificiales (RNA) son sistemas de procesamiento de la información cuya estructura y funcionamiento están inspirados en las redes neuronales biológicas. En los últimos años se ha masificado el uso de las RNA para la predicción de series temporales gracias a su capacidad generalizadora por el hecho de aprender a partir de ejemplos, tal y como lo hace el cerebro humano.

Las redes neuronales artificiales se utilizaron en varios modelos de pronóstico de entrada debido a sus capacidades de modelado no lineal y su naturaleza adaptativa. Las redes neuronales reconstruyen enlaces entre pares de entrada-salida para el sistema que se está modelando.

El pronóstico de series de tiempo es el proceso por el cual se obtiene información de datos de las de series de tiempo que se van a utilizar.

Una serie de tiempo es una sucesión de observaciones de un fenómeno determinado, ordenadas secuencialmente y registradas, usualmente, en un espacio de tiempo.

Cuando se trabaja con series temporales, una de las cosas más importantes es la de predecir los datos futuros de la serie, porque, a partir de los datos obtenidos en esas predicciones se proyectan los valores que tomara la variable determinada. Para llevar a cabo todas las actividades necesarias para realizar el artículo se construyó un algoritmo el cual obtendrá una serie de datos.

Muchas de las técnicas usadas hoy en día para modelar series de tiempo hidrológicas y para generar series sintéticas asumen relaciones lineales entre las variables. Los dos grupos de técnicas incluyen a los modelos de series de tiempo [1].

El modelado de series de tiempo ha recibido en las últimas décadas mucha atención, debido a la creciente necesidad de contar con herramientas que faciliten la toma de decisiones y permitan superar las limitaciones teóricas, conceptuales y prácticas que presentan muchas de las técnicas disponibles en la actualidad. [2]

El modelado de una serie consiste en construir sistemáticamente una representación matemática que permita capturar, total o parcialmente, el proceso generador de los datos; una vez se construye un modelo, es posible realizar el pronóstico de la serie para un horizonte determinado, es decir, estimar sus valores futuros [3, 4, 5, 28].

Los perceptrones multicapa (MLP, por su sigla en inglés) parecen ser la arquitectura de redes neuronales artificiales más utilizada para la predicción de series de tiempo no lineales [6]. Debido a que este es el componente principal de una red neuronal ya que proporciona la base para la mayoría de las aplicaciones de las redes neuronales.

El análisis de series de tiempo hidrológicas ha llevado a los investigadores a aprovechar los desarrollos de los últimos 30 a 40 años en estadística matemática, teoría de probabilidad y teoría de las comunicaciones. El resultado ha sido la evolución de un campo de especialización llamada hidrología estocástica

El objetivo de este artículo es desarrollar un modelo de redes neuronales aplicadas a series de tiempo hidrológicas del caribe colombiano, va dirigido a todos aquellos profesionales e investigadores, que están relacionados con las tendencias actuales y las principales direcciones de investigación sobre los pronósticos y la predicción de series de tiempo hidrológicas.

Metodología utilizada

Para lograr la ejecución de los procesos de predicción en el modelo propuesto, se debe contar con redes neuronales de tipo perceptron multicapa - MLP.

Las RNA de tipo MLP se encuentran entre las arquitecturas de red más poderosas y populares. Están formadas por una capa de entrada, un número arbitrario de capas ocultas, y una capa de salida (véase Figura 1). Cada una de las neuronas ocultas o de salida recibe una entrada de las neuronas de la capa previa (conexiones hacia atrás), pero no existen conexiones laterales entre las neuronas dentro de cada capa

Las redes neuronales son un marco computacional que consiste en un procesamiento simple de conexión masiva entre unidades. Estas unidades son un análogo a las neuronas en el cerebro humano, debido a la estructura altamente conectada [7], que es donde se hace el proceso de análisis y transformación de los datos de entrada, para conseguir una solución.

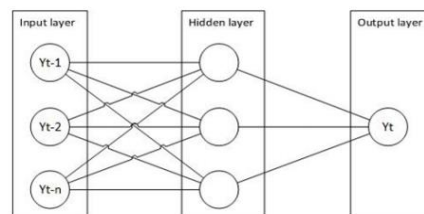


Figura 1. RNA de tipo MLP

La estrategia que se utiliza en éste artículo se basa en un enfoque sistémico orientado a la especificación de modelos de redes neuronales para el pronóstico de series de tiempo, la cual se basa en 4 fases: diseño estructural, diseño funcional, selección del mejor modelo y validación [8], los cuales se presentan en la Figura 2.



Figura 2: Metodología adoptada para la predicción de redes neuronales. Esquema adaptado de [8]

Diseño Estructural

Con el diseño estructural de la red se busca identificar los posibles modelos de red neuronal que se adecuan a las características de la serie de tiempo, esto es, la(s) arquitectura(s) de red adecuada(s) para modelar la serie y determinar la(s) configuración(es) o topología(s) de red que se usará(n). En este paso es necesario definir los siguientes elementos:

- Arquitectura de red neuronal
- Número de neuronas de entrada
- Número de capas ocultas
- Número de neuronas ocultas en cada capa
- Número de neuronas de salida
- Función de Transferencia o Activación en cada capa.

En este trabajo se desarrolla un proceso sistémico que permite la elección del número de neuronas de entrada, el número de neuronas ocultas en cada capa, y la estimación y entrenamiento de los parámetros de la red neuronal, por lo que muchos de los aspectos se abordaran de manera general, y se especificará sólo lo referente al artículo.

A continuación se aborda cada uno de estos aspectos [8].

Arquitectura de Red Neuronal

Para el pronóstico de series de tiempo se han usado diferentes arquitecturas de redes neuronales [29]; la elección inicial de la arquitectura de red neuronal, a menudo, está asociada con los siguientes factores:

- Grado de conocimiento del modelador acerca de la arquitectura
- Robustez: Algunas arquitecturas de red neuronal poseen estructuras complejas, que desfavorecen su elección y restan funcionalidad.
- Disponibilidad de paquetes informáticos o desarrollos de software
- Características de la serie de tiempo: Algunos modelos son más orientados a la representación de ciertas características de la serie, este es el caso de redes recurrentes las cuales han sido utilizadas en series de tiempo con dinámicas altamente volátiles [9].
- Credibilidad ante la comunidad científica
- Tendencias en la implementación: Algunas elecciones se orientan al modelo que está en auge o “de moda”, o la decisión misma de usar siempre los modelos clásicos reconocidos dentro del ámbito de la serie de tiempo; es así como, en finanzas se usan regularmente los ARNN [10].

Número de Neuronas de Entrada

El número de neuronas de entrada o variables de entrada es un parámetro necesario en la configuración de la red neuronal. Este paso implica definir cuantas y cuales entradas serán incluidas en el modelo de red neuronal.

Las entradas, en número, deben ser suficientes para permitir explicar el comportamiento de la serie, pero no excesivas, de tal forma que se evite el sobre-entrenamiento. Así mismo, deben incluirse aquellas variables de entrada cuya influencia sea relevante en el modelo [11]. En el caso general las entradas corresponden a variables causales que influyen la evolución de la serie, rezagos de dichas variables y rezagos de la serie misma; en este trabajo sólo se consideran como entradas valores rezagados de la serie.

Número de Capas Ocultas

En cuanto al número de capas ocultas, teóricamente una red neuronal con una capa oculta y un número suficiente de neuronas puede aproximar con cierta precisión cualquier función continua definida en un dominio compacto [12]. Algunos autores sugieren el uso de una capa oculta cuando la serie de tiempo es continua, y de dos sólo si existe algún tipo de discontinuidad [13], [14]. Sin embargo, otras investigaciones han mostrado que una red con dos capas ocultas puede resultar en una arquitectura más compacta y con una alta eficiencia que redes con una sola capa oculta [15], [16], [17]. En éste artículo se evalúa el desempeño de redes neuronales con una y dos capas ocultas.

Número de Neuronas Ocultas en Cada Capa

La elección del número adecuado de neuronas ocultas ha sido considerado un parámetro determinante en el éxito o fracaso de la red neuronal, toda vez que manifiesta efectos importantes en el nivel de ajuste del modelo a la serie. Un número de neuronas ocultas pequeño hace que la red no puede aprender adecuadamente las relaciones existentes en los datos, mientras que un gran número hace que la red memorice los datos con una pobre generalización y una poca utilidad para la predicción. Algunos autores proponen que el número de neuronas ocultas debe ir en función del número de variables de entrada, sin embargo, este criterio está a su vez relacionado con la extensión de la serie de tiempo y de los conjuntos de entrenamiento, validación y pronóstico.

Usualmente, para la selección del número de neuronas en cada capa oculta se siguen procedimientos ad hoc, donde se busca, bajo algún criterio de minimización del error y algún criterio de parada, encontrar el número adecuado de neuronas; sin embargo, este criterio es sólo matemático y se carece de un sustento conceptual y metodológico, que indique la conveniencia de usar tal procedimiento.

Dado que el valor de los pesos en cada neurona depende del grado de error existente entre el valor deseado y el pronosticado por la red, la selección del número óptimo de neuronas ocultas está directamente asociada al proceso de entrenamiento usado, por lo que es adecuado relacionar ambos procesos. En éste artículo se realiza un proceso iterativo incremental donde se prueban entre 1 y 25 neuronas por cada oculta.

Número de Neuronas de Salida

Decidir acerca del número de neuronas de salida es algo más sencillo, ya que hay razones de peso para utilizar siempre una sola neurona de salida [18]. Los modelos de redes neuronales con múltiples salidas producen resultados inferiores a los comparados con salidas simples. Si se atiende a múltiples salidas la red tenderá a concentrar sus esfuerzos en reducir los errores más grandes, los cuales a menudo se concentran en unas salidas, sin producir cambios favorables en las demás salidas de la red. La solución en este caso, es usar redes especializadas en cada salida.

Igualmente, el concentrar el trabajo de la red en una única salida, el número de parámetros necesarios para afinar el modelo final se reduce.

En el caso de problemas de predicción, la salida del modelo es la predicción del valor siguiente en la secuencia temporal (pronóstico un paso adelante); si bien, es posible generar múltiples salidas en la predicción (pronóstico multipaso), estas adolecen de las falencias anteriormente mencionadas.

Función de Transferencia o Activación en Cada Capa

Diferentes formas funcionales para la especificación de la función de transferencia o activación $g(u)$ han sido propuestas en la literatura. La clásica función sigmodia o logística ha sido usada exitosamente por muchos autores debido a su capacidad de aproximar comportamientos no lineales y continuos, características que son comunes en series reales [19], [16]. Igualmente han sido propuestas en la literatura otras formas funcionales, tales como función Gaussiana, Tangente Hiperbólica, Lorentziana, Onda plana y basadas en Onditas – Wavelets [20]. Funciones lineales son usadas como transferencia de la capa oculta a la capa de salida por la similitud que ellas presentan con los modelos de regresión.

[21] sugiere que si el comportamiento de los datos en promedio debería usarse como transferencia de la capa de entrada a la capa de salida la función sigmodia, mientras que si existen desviaciones del comportamiento promedio, la función tangente hiperbólica presenta un mejor desempeño. Las funciones de transferencia basadas en Onditas pueden exhibir una aproximación más adecuada toda vez que por sus propiedades de cambios de escala (dilatación) pueden distinguir características locales de una señal a varias escalas, y por translación permiten cubrir una amplia región de estudio [22], sin embargo su aplicación es computacionalmente más compleja que otras funciones, por lo que se sale del alcance de este artículo.

Diseño funcional

El diseño funcional de una red neuronal, se basa en la estimación de los parámetros adecuados y el entrenamiento del modelo, el cual es un proceso determinante en el desempeño de la red neuronal y que conlleva múltiples implicaciones en todas las etapas del modelado.

El objetivo del diseño funcional de una red neuronal es encontrar el conjunto de pesos de los parámetros que minimizan una función de error; es decir, aquellos pesos que hacen que el modelo tenga un mejor ajuste a la serie de tiempo. A menos que el modelo este sobreentrenado, el conjunto de pesos debería producir una buena generalización.

El entrenamiento de una red neuronal es un problema de minimización no lineal no restringido en el cual los pesos de la red son iterativamente modificados con el fin de minimizar el error entre la salida deseada y la obtenida.

En el diseño funcional es necesario definir los siguientes elementos [8]:

- Algoritmo de entrenamiento
- Número máximo de iteraciones del algoritmo (diferentes puntos de inicio) y método de inicialización
- Número máximo de épocas (entrenamiento de los pesos)

Algoritmo de Entrenamiento

Para las pruebas y desarrollo del presente artículo se hizo uso de dos algoritmos de entrenamiento que son:

- ADAM
- SGD

Número Máximo de Iteraciones del Algoritmo (Diferentes Puntos de Inicio) y Método de Inicialización

Este parámetro mide el número de reinicios necesarios para garantizar que los pesos iniciales de los parámetros sean válidos, la cual es vista como la posibilidad de alcanzar un óptimo global.

La elección del número de iteraciones del algoritmo, a menudo, es subjetiva, toda vez que en la literatura son escasos los estudios que presentan una forma sistemática de elegir este parámetro; siendo común que el número sea elegido más por criterios expertos.

Algunos estudios basados en métodos de inicialización de los pesos orientan acerca del número de iteraciones de los pesos iniciales necesarias, las cuales van en función de formas aleatorias, determinísticas y heurísticas, de recorrer el espacio de búsqueda [véase como ejemplos de estos los trabajos de [23], [24]]. El criterio comúnmente usado en la literatura para la inicialización de los pesos es elegir valores aleatorios provenientes de una distribución de probabilidad Uniforme o Normal.

Número Máximo de Épocas (Iteraciones de Entrenamiento)

Este parámetro mide el número de iteraciones de entrenamiento necesarias en el algoritmo para alcanzar un ajuste deseado, por lo tanto, está directamente relacionado con el criterio de parada. Un número pequeño de iteraciones no le permite al algoritmo aprender las características de la serie; mientras que, un número grande provoca sobre-entrenamiento; esto es, la red se dedica a memorizar particularidades presentes en los patrones, perdiendo su habilidad para generalizar.

Algunos autores como [25], y [26], proponen diferentes formas funcionales para elegir el número adecuado de iteraciones de entrenamiento, algunas en función del número de entradas del modelo.

Para el desarrollo de este artículo se probó con diferentes números de épocas (100, 500, 1000)

Selección del mejor modelo

Si bien, se ha resaltado la necesidad de contar con un algoritmo acertado frente a los criterios de reducción del error a medida que se adicionan parámetros, lo cual es un indicador claro de la eficiencia de este, en ocasiones es igualmente importante seleccionar entre los modelos posibles uno que equilibre el nivel de error y el número de parámetros del modelo.

El hecho de modificar el número de parámetros de un modelo de redes neuronales implica un cambio en la configuración de la red, es decir, se generan cada vez modelos diferentes de los cuales se requiere elegir el mejor. [8]

La selección del mejor modelo está orientada por elegir entre los modelos posibles y candidatos aquel que se cumpla mayoritariamente una serie de características, lo cual permite catalogarlo con el “mejor”. [27] define las siguientes características comparativas que debe cumplir un modelo para ser seleccionado como el modelo final:

- Ajuste: El error de ajuste del modelo debe ser el menor respecto a otros modelos candidatos.
- Generalización: El modelo debe tener la capacidad de proveer un buen ajuste para los datos por fuera de la muestra de calibración, por lo tanto ofrecer un error de pronóstico menor respecto al error de otros modelos candidatos.
- Parsimonia: El número de parámetros del modelo (entradas y neuronas) es menor respecto a otros modelos candidatos que exhiben un desempeño similar.

En la literatura se han propuesto diferentes criterios de selección de modelos que reúnen en algún grado las características anteriormente enunciadas. En este trabajo se realiza la selección basada en el error de ajuste, la cual se basa en elegir aquel modelo que presente una menor medida del error. Estos criterios requieren la división de la muestra de datos en al menos dos conjuntos: ajuste o entrenamiento, y prueba o pronóstico. El conjunto de entrenamiento se usa para estimar los parámetros adecuados del modelo, mientras que el conjunto de pronóstico para verificar la validez y utilidad del modelo.

El error utilizado en este trabajo fue el “Raíz del Error Cuadrático Medio (RMSE: Root Mean Squared Error)”

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}$$

Validación

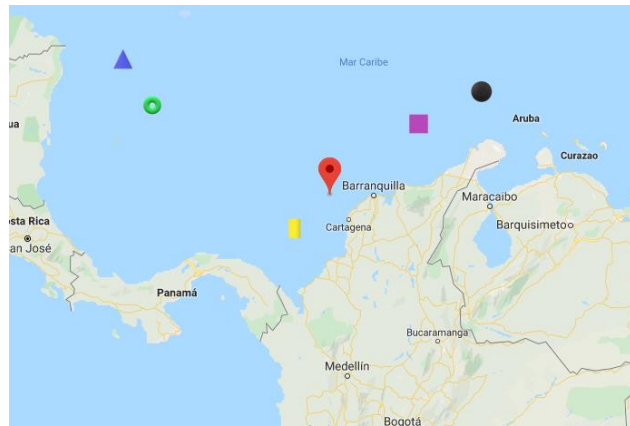
En la validación se analizan los errores o residuales del modelo con el fin de verificar si cumplen con los supuestos de normalidad y aleatoriedad. Este paso, tan común en los modelos estadísticos de series de tiempo, es poco usado en modelos de redes neuronales por las siguientes razones:

Los criterios son basados en características lineales de la serie, los cuales no se garantiza que deban cumplirse en series no lineales.

La comunidad científica se interesa más por la funcionalidad que ofrece un error de ajuste bajo, que por la estabilidad que pueda tener.[8]

Resultados

Las pruebas fueron realizadas a 6 bases de datos que hacen referencia a 6 boyas que describen la altura en olas del mar caribe (vea Figura 3). Cada una de las seis bases de datos cuenta con 745 datos.



Boya 1: Boya 2: Boya 3:
 Boya 4: Boya 5: Boya 6:

Figura 3. Distribución de datos de series de tiempo del Caribe Colombiano

Se llevaron a cabo más de 300 pruebas en ejecución del algoritmo adaptado a las bases de datos. En cada una de las pruebas se midió el porcentaje de error (RMSE) y se comparaban con las demás para ir descartando las que no son óptimas. El algoritmo sufre modificaciones durante cada prueba, esto, buscando la mejor respuesta posible [30-34].

Hay que decir que todas las pruebas fueron realizadas con los algoritmos de entrenamiento ADAM y SGD y modificaciones en las iteraciones o épocas (epochs) de 100, 500 y 1000 para cada una de las alteraciones o combinaciones que se muestran en las tablas.

Nº capas ocultas	1				2			
Nº neuronas	5	10	15	20	15,5	15,10	15,15	15,20
RMSE Train	81	81	80	80	87	86	86	86
RMSE Test	80	79	79	79	86	86	86	86

Tabla 1: Evaluación Boya 1

Después del desarrollo de pruebas a la boya 1 (véase Tabla 1) se concluye que la mejor arquitectura para la serie es con 500 épocas, una capa oculta con 15 neuronas y el algoritmo de entrenamiento ADAM (vea Figura 4).

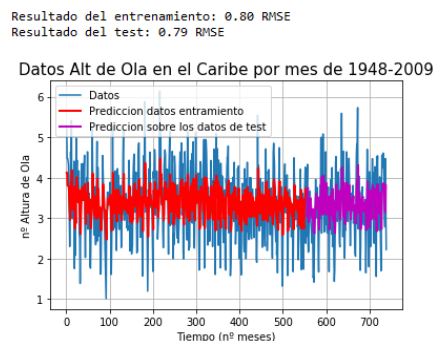


Figura 4. Pronostico de serie Boya 1

El resultado obtenido en las pruebas de la boya 2 (véase Tabla 2) nos refleja que la mejor arquitectura para esta serie en particular es con 500 épocas, una capa oculta que tiene 15 neuronas y el algoritmo de entrenamiento ADAM (Figura 5).

Nº capas ocultas	1				2			
Nº neuronas	5	10	15	20	15,5	15,10	15,15	15,20
RMSE Train	74	74	73	73	77	77	77	77
RMSE Test	74	74	73	73	77	77	77	77

Tabla 2: Evaluación Boya 2

Resultado del entrenamiento: 0.73 RMSE
 Resultado del test: 0.73 RMSE

Datos Alt de Ola en el Caribe por mes de 1948-2009

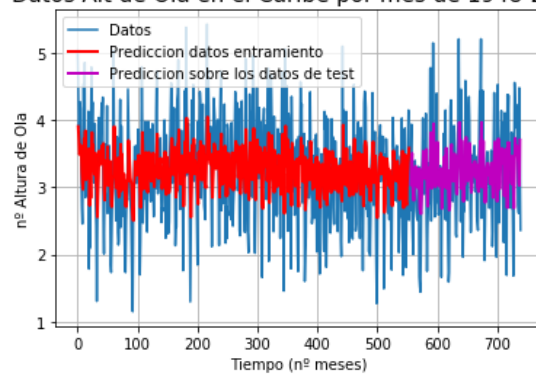


Figura 5. Pronostico de serie boya 2

Para la boya 3 se obtuvieron pruebas (Tabla 3) que arrojaron como mejor resultado para la arquitectura. Con 500 épocas, una capa oculta de 5 neuronas y el algoritmo de entrenamiento ADAM (vea la Figura 6)

Nº capas ocultas	1				2			
Nº neuronas	5	10	15	20	5,5	5,10	5,15	5,20
RMSE Train	69	71	69	69	87	80	83	86
RMSE Test	67	69	67	67	86	80	83	85

Tabla 3: Evaluación boya 3

Resultado del entrenamiento: 0.69 RMSE
 Resultado del test: 0.67 RMSE

Datos Alt de Ola en el Caribe por mes de 1948-2009

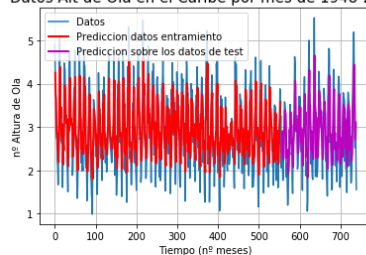


Figura 6. Pronostico de serie boya 3

Según los resultados arrojados por la Tabla 4 en el proceso de evaluación, nos muestra que la mejor arquitectura para la serie de la boya 4 es con 500 épocas, dos capas ocultas compuestas por 5 y 15 neuronas en la primera y segunda capa respectivamente la cual se presenta en la Figura 7.

N° capas ocultas	1				2			
N° neuronas	5	10	15	20	5,5	5,10	5,15	5,20
RMSE Train	80	78	78	78	77	77	76	76
RMSE Test	79	77	77	77	77	77	77	78

Tabla 4: Evaluación boya 4

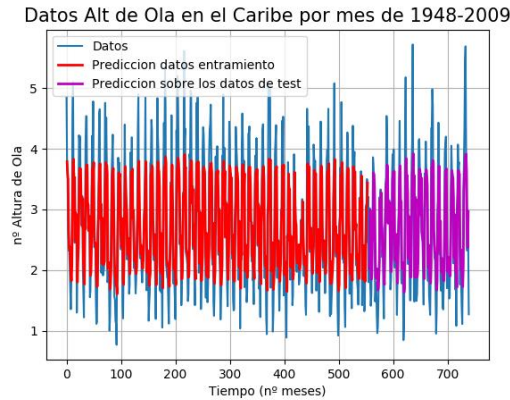


Figura 7. Pronostico de serie boya 4

Los mejores resultados para la boya 5 (Tabla 5) fueron de la arquitectura con 500 épocas, una capa oculta de 5 neuronas y el algoritmo de entrenamiento SGD como se muestra en la Figura 8.

N° capas ocultas	1				2			
N° neuronas	5	10	15	20	5,5	5,10	5,15	5,20
RMSE Train	52	52	52	52	52	52	53	52
RMSE Test	49	49	50	49	49	50	48	51

Tabla 5: Evaluación boya 5

Resultado del entrenamiento: 0.52 RMSE
Resultado del test: 0.49 RMSE

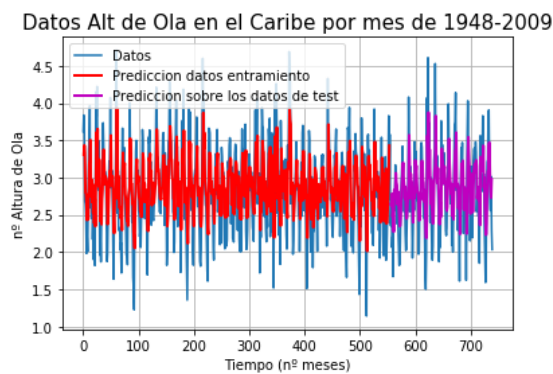


Figura 8. Pronostico de serie boya 5

La Tabla 6 nos muestra los diferentes resultados de las pruebas hechas a la boya 6. Escogiendo como mejor arquitectura con 500 épocas, una capa oculta con 10 neuronas y el algoritmo de entrenamiento SGD (véase Figura 9)

Nº capas ocultas	1				2			
Nº neuronas	5	10	15	20	10,5	10,10	10,15	10,20
RMSE Train	62	61	62	62	62	61	61	61
RMSE Test	69	58	58	58	59	58	58	58

Tabla 6: Evaluación boya 6

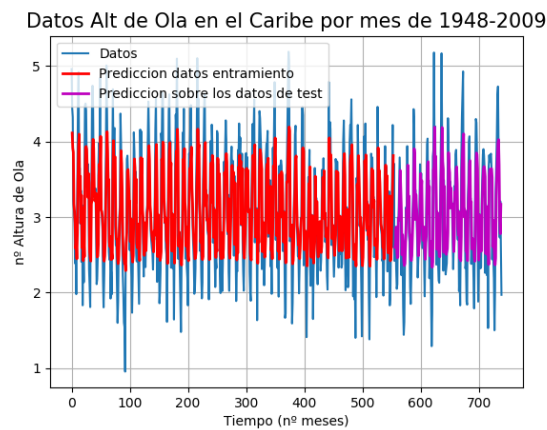


Figura 9. Pronostico de serie boya 6

Conclusiones

En este artículo se ha demostrado como las redes neuronales artificiales son utilizadas en las series de tiempo hidrológicas, el modelo que fue planteado se fue evaluando mediante pruebas realizadas durante el desarrollo de este mismo.

La aplicación experimental del algoritmo propuesto a las series de tiempo analizadas presento resultados satisfactorios. Además, se logro alcanzar niveles cercanos a cero por error, lo que demuestra que es posible alcanzar los niveles deseados de error de ajuste, incluso cuando el objetivo del error es cero.

Referencias bibliográficas

1. Hernandez S., Cesar A.; Pedraza M., Luis Fernando; Escobar Diaz, Andres Aplicaciones De Las Series De Tiempo En Modelos De Tráfico Para Una Red De Datos
2. Sánchez, Paola; Velásquez, Juan D. Problemas de investigación en la predicción de series de tiempo con redes neuronales artificiales Revista Avances en Sistemas e Informática, vol. 7, núm. 3, diciembre, 2010, pp. 67-73.
3. C. Montgomery, L. A. Johnson y J. S. Gardiner, Forecasting & Time Series Analysis, Segunda ed., Singapore: Mc-Graw-Hill, Inc., 1990.
4. R. Gençay y T. Liu, Nonlinear modeling and prediction with feedforward and recurrent networks, Physica D: Nonlinear Phenomena, vol. 108, nº 1-2, pp. 119-134 , 1997.

5. B. L. Bowerman, R. T. O'Connell y A. Koehler, *Forecasting, Time Series, and Regression: An Applied Approach*, Cuarta ed., Ohio: Cengage Learning Brooks Cole, 2006.
6. Hornik, K., Stinchcombe, M., y White, H. 1989. Multilayer feedforward Networks are Universal Approximators. En: *Neural Networks*, Vol. 2, No. 5, pp. 359–366.
7. Z. Tang and P. A. Fishwick, "Feedforward Neural Nets as Models for Time Series Forecasting," *ORSA J. Comput.*, vol. 5, no. 4, pp. 374–385, 1993.
8. Sánchez, Paola Andrea, José Rafael García González, Carlos Hernán Fajardo-Toro and Paloma María Teresa Martínez Sánchez. "Designing a Neural Network Model for Time Series Forecasting." *Theoretical and Applied Mathematics in International Business*. IGI Global, 2020. 259-284. doi:10.4018/978-1-5225-8458-2.ch012
9. R. Hafezi, J. Shahrabi, and E. Hadavandi, "A bat-neural network multi-agent system (BNNMAS) for stock price prediction: Case study of DAX stock price," *Appl. Soft Comput. J.*, vol. 29, pp. 196–210, 2015.
10. Castillo, E., Cobo, A., Gutiérrez, J.M., Pruneda, R.E.: *Introducción a las Redes Funcionales con Aplicaciones. Un Nuevo Paradigma Neuronal*. Editorial Paraninfo S.A. Madrid. España. pp.5-8; 8-16; 21-24, 30-34, 53-100. (1999)
11. S. Mishra and S. K. Patra, "Short term load forecasting using a novel recurrent neural network," *International Journal of Computational Intelligence: Theory and Practice*, vol. 4, no. 1, pp. 39-45, 2009.
12. P. Franses and D. Van Dijk, *Non-linear time series models in empirical finance*. UK: Cambridge University Press, 2000.
13. S. Crone and N. Kourentzes, "Input-variable Specification for Neural Networks - An Analysis of Forecasting low and high Time Series Frequency," *Proceedings of the International Joint Conference on Neural Networks, IJCNN'09*, p. inress, 2009.
14. I. Kaastra and M. Boyd, "Designing a neural network for forecasting financial and economic series," *Neurocomputing*, no. 10, pp. 215-236, 1996.
15. K. Hornik, "Approximation capability of multilayer feedforward networks," *Neural Networks*, vol. 4, pp. 251-257, 1991.
16. K. Hornik, M. Stinchcombe, and H. White, "Multilayer Feedforward Networks are Universal Approximators," *Neural Networks*, vol. 2, no. 5, pp. 359-366, 1989.
17. D. Srinivasan, A. C. Liew, and C. S. Chang, "A neural network short-term load forecaster," *Electric Power Systems Research*, vol. 28, p. 227–234, 1994.
18. X. Zhang, "Time series analysis and prediction by neural networks," *Optimization Methods and Software*, vol. 4, p. 151–170, 1994.
19. D. L. Chester, "Why two hidden layers are better than one?," *Proceedings of the International Joint Conference on Neural Networks*, p. 1265–1268, 1990.
20. I. Kaastra and M. Boyd, "Designing a neural network for forecasting financial and economic series," *Neurocomputing*, no. 10, pp. 215-236, 1996.

21. G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control: Signals and Systems*, vol. 2, p. 202–314, 1989.
22. J. Zhang, G. G. Walte, Y. Miao, and W. N. Wayne, "Wavelet neural networks for function learning," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, p. 1485–1497, 1995.
23. C. C. Klimasauskas, "Applying neural networks," in *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real World Performance*. Chicago: Probus, 1993, pp. 64-65.
24. L. Ait Gougam, M. Tribeche, and F. Mekideche-Chafa, "A systematic investigation of a neural network for function approximation," *Neural Networks*, vol. 21, pp. 1311-1317, 2008.
25. M. Lehtokangas, "Modelling with constructive backpropagation," *Neural Networks*, vol. 12, no. 4-5, p. NeuralNetworks, 1999.
26. M. Cottrell, B. Girard, Y. Girard, M. Mangeas, and C. Muller, "Neural modeling for time series: a statistical stepwise method for weight elimination," *IEEE Transactions on Neural Networks*, vol. 6, no. 6, p. 1355–1364, 1995.
27. D. L. Chester, "Why two hidden layers are better than one?," *Proceedings of the International Joint Conference on Neural Networks*, p. 1265–1268, 1990.
28. Sánchez, P. A., & García, J. R. (2017). A New Methodology for Neural Network Training Ensures Error Reduction in Time Series Forecasting. *Journal of Computational Science*, 13(7), 211–217. doi:10.3844/jcssp.2017.211.217
29. Sánchez, P. A., & García, J. R. (2018). Autoregressive Moving Average Recurrent Neural Networks Applied to the Modelling of Colombian Exchange Rate. *International Journal of Artificial Intelligence*, 16(2), 194–207.
30. A. Viloría and W. Quintero, "Gestión de información sat río Manzanares", *Investigación e Innovación en Ingenierías*, vol. 4, no. 2, p. 128, 2016. DOI: <https://doi.org/10.17081/invinno.4.2.2493>
31. A. Jiménez Coronado, "Estrategia de competitividad y emprendimiento, una revisión de la literatura", *Investigación e Innovación en Ingenierías*, vol. 4, no. 2, p. 104, 2016. DOI: <https://doi.org/10.17081/invinno.4.2.2492>
32. J. Guillot and A. Viloría, "Estudio Sobre el Abastecimiento Constante de Energía Eólica", *Investigación e Innovación en Ingenierías*, vol. 3, no. 2, 2015. DOI: <https://doi.org/10.17081/invinno.3.2.2028>
33. R. Piteres, M. Cabarcas., y H. Gaspar, "El recurso humano factor de competitividad en el sector salud", *Investigación e Innovación en Ingenierías*, vol. 6, n°. 1, pp. 93 - 101., 2018 DOI: <https://doi.org/10.17081/invinno.6.1.2778>
34. C.A. Gutiérrez, R. Almeida., y W. Romero, "Diseño de un modelo de migración a cloud computing para entidades públicas de salud", *Investigación e Innovación en Ingenierías*, vol. 6, n°. 1, pp. 10 - 26., 2018. DOI: <https://doi.org/10.17081/invinno.6.1.2772>