

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA EN INFORMÁTICA

**PRONOSTICO NEURONAL AUTOREGRESIVO
PARA LA TENDENCIA DE ACCIONES
PERTENECIENTES AL IPSA, UTILIZANDO
APRENDIZAJE EVOLUTIVO**

LEONARDO ESTEBAN GONZÁLEZ ROJAS

INFORME FINAL DEL PROYECTO
PARA OPTAR AL TÍTULO PROFESIONAL DE
INGENIERO CIVIL EN INFORMÁTICA

SEPTIEMBRE 2014

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA EN INFORMÁTICA

**PRONOSTICO NEURONAL AUTOREGRESIVO
PARA LA TENDENCIA DE ACCIONES
PERTENECIENTES AL IPSA, UTILIZANDO
APRENDIZAJE EVOLUTIVO**

POR

LEONARDO ESTEBAN GONZÁLEZ ROJAS

CARRERA

INGENIERÍA CIVIL INFORMÁTICA

PROFESOR GUÍA

NIBALDO RODRIGUEZ AGURTO

PROFESOR CORREFERENTE

JOSÉ MIGUEL RUBIO LEÓN

SEPTIEMBRE 2014

Gracias a mi familia por su apoyo incondicional en todos estos años.

Resumen

La capacidad para predecir la dirección del precio y rentabilidad de las acciones y divisas es de vital importancia para los inversionistas, esto con el fin de maximizar sus ganancias.

Dado lo anterior, este trabajo propone un modelo basado en redes neuronales artificiales y algoritmos evolutivos para el pronóstico del precio y la rentabilidad de tres acciones transadas en la Bolsa de Comercio de Santiago. Para la selección de las acciones se realizó un estudio de las acciones pertenecientes al IPSA. El estudio se llevó a cabo mediante el coeficiente de Husrt, lo que entrego el grado de persistencia de las acciones transadas en la Bolsa de Comercio de Santiago.

Los resultados experimentales muestran que pronosticar la dirección de la rentabilidad trae mejores resultados que pronosticar el precio de la acción, alcanzo la rentabilidad un 75% de acierto en la dirección de la curva.

Palabras Claves: ANN, Husrt, BSA, IPSA, Acciones, Precio, Rentabilidad

Abstract

The ability to predict the direction of price and profitability of stock markets and currencies is vital for investors, this in order to maximize profits.

Given the above, this work proposes an approach based on artificial neural networks and evolutionary algorithms for the prediction of the price and profitability of three stocks traded on the Santiago Stock Exchange model. For the selection of actions a study of the shares belonging to IPSA was performed. The study was conducted by the coefficient of Husrt, which give the degree of persistence of the shares traded on the Santiago Stock Exchange.

Experimental results show that forecast the direction of profitability brings better results to predict the stock price, reached profitability 75% accuracy in the direction of the curve.

Keywords: ANN, Husrt, BSA, IPSA, Stock Markets, Price, Profitability

I. Contenido

I.	Contenido	i
II.	Lista de Ilustraciones	iii
III.	Lista de Tablas	iv
IV.	Lista de Abreviaturas	v
1.	Introducción	1
2.	Objetivos del Proyecto	3
2.1.	Objetivo General	3
2.2.	Objetivos Específicos	3
3.	Trabajos Relacionados	4
4.	Teoría de Mercado Eficiente y Exponente de Husrt	6
4.1.	Análisis de Rango Reescalado (R/S)	7
5.	Pronóstico con Redes Neuronales	10
5.1.	Red Neuronal Biológica	10
5.2.	Red Neuronal Artificial	11
5.3.	Fundamentos Matemáticos de Redes Neuronales Artificiales	11
5.4.	Redes Neuronales Artificiales Feed-Forward	12
5.5.	Función de Activación de Base Radial	13
5.6.	Método de aprendizaje de una Red Neuronal Artificial	14
5.6.1.	Método Directo de Pronóstico	14
5.6.2.	Método Iterativo de Pronóstico	15
5.7.	Pronostico con Redes Neuronales Artificiales	15
6.	Enfoques de Aprendizaje de una Red Neuronal Artificial	16
6.1.	Algoritmo BSA	16
6.1.1.	Inicialización	17
6.1.2.	Selección-I	17
6.1.3.	Mutación	17
6.1.4.	Cruza	18
6.1.5.	Selección-II	19
7.	Pronosticador Neuronal Propuesto	20

7.1.	Proceso de Aprendizaje Pronosticador Neuronal	21
7.2.	Métricas de Rendimiento	22
8.	Análisis y Discusión de Resultados.	24
8.1.	Análisis del exponente de Husrt	24
8.2.	Descripción de los Datos de Entrada	24
8.3.	Pronosticador Neuronal utilizando aprendizaje BSA	25
8.4.	Número de Entradas al Pronosticador Neuronal.....	25
8.5.	Rendimiento del Pronosticador Neuronal.....	26
8.6.	Pronosticador Neuronal en Producción	30
9.	Conclusiones	32
10.	Referencias Utilizadas	34

II. Lista de Ilustraciones

Ilustración 1 Neurona Biológica.....	10
Ilustración 2 Neurona Artificial.....	11
Ilustración 3 Red Neuronal Artificial	12
Ilustración 4 Familia de Funciones de Base Radial.....	13
Ilustración 5 Red Neuronal Artificial Propuesta	20
Ilustración 6 Proceso de Aprendizaje y Pruebas del Pronosticador Neuronal	22
Ilustración 7 Entradas al Pronosticador Neuronal	25
Ilustración 8 Precio Real vs Precio Estimado - EMBONOR - B	27
Ilustración 9 Correlación - EMBONOR – B	27
Ilustración 10 Precio Real vs Precio Estimado - CHILE	28
Ilustración 11 Correlación - CHILE	28
Ilustración 12 Precio Real vs Precio Estimado - BCI.....	29
Ilustración 13 Correlación - BCI	29
Ilustración 14 Pronostico del Precio - EMBONOR-B.....	30
Ilustración 15 Pronostico de la Rentabilidad - EMBONOR-B.....	30
Ilustración 16 Pronostico del Precio - CHILE.....	31
Ilustración 17 Pronostico de la Rentabilidad - CHILE.....	31
Ilustración 18 Pronostico del Precio - BCI.....	31
Ilustración 19 Pronostico de la Rentabilidad - BCI.....	31

III. Lista de Tablas

Tabla 1 Algoritmo BSA.....	16
Tabla 2 Cruza I - BSA	18
Tabla 3 Cruza-II BSA.....	19
Tabla 4 Selección-II BSA.....	19
Tabla 5 Resultados Pronosticador Neuronal	26

IV. Lista de Abreviaturas

ANN: Redes Neuronales Artificiales

ANN-FF: Redes Neuronales Artificiales Feed-Forward

B&H: Comprar y Mantener.

BSA: Backtracking Search Optimization Algorithm

DAC: Dirección de la Curva

IPSA: Índice de Precio Selectivo de Acciones

MAE: Error Absoluto Medio

MAPE: Error Porcentual Absoluto Medio

MILA: Mercado Integrado Latino Americano

MPL: Redes Neuronales Artificiales Multi-Capa

MSE: Error Cuadrático Medio

RMSE: Raíz Error Cuadrático Medio

1. Introducción

Los mercados financieros son aquellos donde concurren los entes deficitarios y superavitarios de capital, los primeros con el objetivo de obtener financiación y los segundos invertir sus excedentes de recursos. Las acciones sobre empresas son los títulos más transados comúnmente en un mercado de valores, el cual a su vez es un componente fundamental de los mercados financieros. El pronóstico del precio de las acciones es por lo tanto de alto interés para los inversionistas ya que le indicará las señales de compra o venta de estos títulos con el fin de maximizar sus beneficios. Sin embargo, no es una tarea fácil dada la cantidad de variables macroeconómicas y microeconómicas que determinan su valor[1].

Cabe señalar que una acción es un instrumento de renta variable, donde la acción es parte alícuota del capital de una empresa, al poseer una acción se tiene derechos económicos y políticos sobre la empresa, y de ser un tipo de acción preferente, se puede recibir dividendos antes que las acciones comunes.

Los mercados mundiales están en constante cambio, debido a una infinidad de factores, tales como, ofertas y demandas de productos, servicios, commodities, etc. Esta volatilidad genera gran incertidumbre respecto a los precios de las acciones.

En el contexto chileno, el mercado de valores ha tomado gran importancia por su crecimiento sostenido durante los últimos nueve años dada la recuperación de la confianza inversionista. Esto además ha permitido la integración con otras bolsas en el llamado *Mercado Integrado Latino Americano* (MILA) que agrupa a las bolsas de valores de Colombia, Perú y Chile.

En particular el mercado de capitales chileno se organiza en tres grandes sectores clasificados de acuerdo al regulador encargado de su supervisión, estos son las administradoras de fondos de pensiones los bancos, y el mercado de los valores y seguros.

Sobre este último punto, la Bolsa de Comercio de Santiago, es el lugar donde se transan los diferentes instrumentos tales como, las acciones, opciones de suscripción de acciones, los llamados instrumentos de renta fija e intermediación financiera. Donde la Superintendencia de Valores y Seguros es la entidad encargada de la supervisión de este mercado.

Ahora bien, para estudiar el mercado de capitales es necesario contar con dos tipos de análisis, estos son el Análisis Fundamental y el Análisis Técnico.

El análisis fundamental estudia en profundidad todos los acontecimientos de origen político, social y económico poniendo el énfasis en cómo y por qué estos han afectado históricamente a la cotización de las distintas divisas alrededor del mundo. De esta forma, se pretende entender el estado actual de la realidad y anticipar la probable evolución de una cotización de acuerdo a las circunstancias socioeconómicas y Macroeconómicas que se están desarrollando en la actualidad o de aquellas que se prevé va a ocurrir. Donde los elementos más relevantes del Análisis Fundamental, son el Crecimiento, Balanza de Pagos, Impacto de la Inflación, Empleo y Mercado de valores.

En cambio el análisis técnico examina datos pasados acerca de precios y volúmenes para prever futuros movimientos de precios. Este tipo de análisis se centra en la formación de cuadros y fórmulas para capturar tendencias importantes y menores, identificar oportunidades de compra/ venta evaluando el alcance de los vuelcos favorables del mercado. Dependiendo de su límite de tiempo, es posible utilizar el análisis técnico en una base dentro del día (5 minutos, 15 minutos, por hora) o en una base semanal o mensual.

Cuando se habla de análisis técnico, no es posible dejar de lado la teoría Dow. Esta es una de las teorías más antiguas utilizadas para predecir los movimientos bursátiles. Su principal postulado establece que los precios reflejan toda la información que afecta la oferta y demanda del mercado, descontando las variaciones producto de noticias, datos o fenómenos de tipo natural.

Para estudiar el comportamiento de las acciones o divisas, el análisis se basa por lo general en 5 aspectos claves, estos son el precio de apertura, menor precio, mayor precio, precio de cierre y volumen de transacciones. El horizonte de tiempo para las variables antes señaladas puede ser inter-día (días, semanas, años) o intra-día (segundos, minutos, horas).

Actualmente, en el análisis técnico, se puede encontrar diversas técnicas las cuales estudian el comportamiento histórico de las acciones y divisas. En este marco destacan un sinnúmero de modelos paramétricos y no paramétricos para predecir el precio o rentabilidad, donde los modelos no paramétricos han tomado gran protagonismo, sin considerar el alto interés de diversos investigadores en el mundo[10].

Profundizando en los modelos no paramétricos encontramos las Redes Neuronales Artificiales, que básicamente es un mapeo de la organización y funcionamiento cerebral humano. La Red Neuronal Artificial es una estructura compuesta de un número de unidades interconectadas, donde cada unidad posee una característica de entrada/salida dada por una función de activación. Con el entrenamiento adecuado una red Neuronal Artificial es capaz de capturar comportamientos que a simple vista resultan totalmente aleatorios.

Considerando lo anterior, en este trabajo se propone un modelo de pronóstico basado en Redes Neuronales Artificiales. Este modelo busca dar solución a la problemática de miles de inversores del mercado accionario, pronosticando la tendencia de tres acciones transadas en la Bolsa de Comercio de Santiago.

2. Objetivos del Proyecto

A continuación se detallan el objetivo general, los objetivos específicos y el plan de trabajo.

2.1. Objetivo General

Desarrollar un Modelo Neuronal con aprendizaje basado en algoritmos evolutivos, para pronosticar la tendencia del precio y rentabilidad de tres acciones del IPSA, seleccionadas a través del exponente de Hurst.

2.2. Objetivos Específicos

- Explicar el proceso de aprendizaje neuronal, utilizando algoritmos evolutivos
- Calcular el grado de persistencia de las acciones pertenecientes al IPSA mediante el exponente de Hurst.
- Implementar y evaluar la exactitud del pronosticador propuesto utilizando métricas residuales y direccionales.

3. Trabajos Relacionados

Al realizar una revisión de trabajos relacionados sobre pronóstico de divisas y acciones, es posible observar que unas de las técnicas matemáticas utilizadas para abordar el estudio del comportamiento de mercados financieros han sido los modelos multivariantes y univariantes, pero presentan deficiencias cuando se trata de realizar predicciones fuera de la muestra[2].

Estudios posteriores han demostrado que la presencia de dinámicas no lineales podría implicar la posibilidad de realizar predicciones más precisas que aquellas proporcionadas por un modelo estocástico lineal y, en concreto, por el modelo Random Walk[3]. Otros autores han utilizado métodos estocásticos basados en cadenas de Markov con el fin de afrontar los estados aleatorios en el comportamiento de estos mercados[4]. A pesar de no entregar un valor concreto de pronóstico, su ventaja se da en que permite predecir cambios posibles en los estados de los precios en términos de probabilidad de ocurrencia. Su aplicación al mercado financiero chino mostró la utilidad del método en términos de disponer un criterio adicional a los pronósticos entregados por otros métodos basados en series de tiempo.

Ante el comportamiento no lineal de este tipo de variables económicas, desde la década de los 90 se han propuesto nuevos métodos basados en redes neuronales artificiales. Su principal característica de permitir establecer relaciones lineales y no lineales entre las entradas y salidas de un sistema, ha hecho posible mostrar su aplicabilidad en mercados de alta volatilidad, cuyas variables obedecen a comportamientos no lineales en diversas áreas de la ingeniería y en los mercados de electricidad[5].

Una primera revisión donde se muestra un conjunto amplio de aplicaciones exitosas de las redes neuronales a las finanzas se presenta en Trippi y Turban[6]. En este libro se destacan trabajos de análisis de reportes contables, predicción de quiebras, análisis de riesgo, pronóstico de divisas y estrategias de negociación de índices financieros entre otros. En una publicación más reciente Li and Ma[7], presentan una revisión actualizada de estas aplicaciones a predicciones en los mercados de valores, derivados, divisas y crisis financieras. De estas revisiones se resalta la superioridad en el desempeño de las redes neuronales con respecto a los métodos econométricos y otros modelos lineales.

En la literatura científica se encuentran también modelos que combinan redes neuronales con algoritmos genéticos. Hao[8] propone una red neuro-genética para pronosticar el precio de las acciones en el corto plazo en la bolsa de valores de Shenzhen (China), la cual combina la capacidad de búsqueda de los algoritmos genéticos con el fin de determinar los pesos óptimos de la red neuronal. El modelo propuesto arrojó muy buenos resultados en el pronóstico de los cuatro días siguientes, sin embargo, el error era mayor aumentando el número de días o al tratar de abordar el problema con datos semanales.

En el plano local, Antonino, Rebolledo y Cornejo[9], realizaron una predicción del IPSA, utilizando Lógica Difusa, lo que trajo buenos resultados en rentabilidad, superando a la estrategia *Comprar y Mantener* (B&H).

Ahora bien, la justificación de usar redes neuronales viene dado por una revisión de 100 publicaciones científicas dedicadas al pronóstico de precios en los mercados de valores de diferentes partes del mundo, usando redes neuronales y redes neuro-difusas, este estudio es

presentado por Atsalakis y Valavanis[10]. Todos estos trabajos demuestran la superioridad de estas técnicas de computación inteligente con respecto a los modelos convencionales en cuanto a una mejor precisión en el pronóstico. Sin embargo queda en evidencia la dificultad en la definición en la estructura del modelo pues en la mayoría de los casos se realizó por prueba y error.

Profundizando en trabajos relacionados sobre Redes Neuronales Artificiales, como herramienta de pronóstico en el mercado de valores. Podemos encontrar dos enfoques para abordar esta problemática, estos son el pronóstico del precio de la acción, y el pronóstico de la rentabilidad de la acción.

En relación al pronóstico del precio de la acción, Caoa, Leggioa y Schniederjans[11], lograron predecir el movimiento del precio de las acciones del mercado de Shanghái utilizando redes neuronales de una variable o univariados y modelos lineales que fueron comparados con pronósticos lineales extraídos de literatura de pronóstico financiero. Se pudo comprobar que las redes neuronales son una gran herramienta para la predicción de mercados emergentes. Caoa, Leggioa y Schniederjans[11] comprobaron que un modelo simple de una variable tiene mejores resultados que un modelo multivariado.

En el 2009, Pei-Chann Chang, Chen-Hao Liu, Jun-Lin Lin, Chin-Yuan Fan, Celeste S.P. Ng.[12] lograron pronosticar el movimiento que tomaría una acción combinando ventanas de tiempo dinámico, razonamiento basado en casos y redes neuronales, utilizaron una ventana dinámica basada en casos con backpropagation para minimizar un falso positivo de comprar o vender una acción. De esta forma, pudieron ver el precio de una acción cualquiera sin importar la relevancia que tenga la acción en el mercado.

Por otro lado, en el caso de Thira Chavarnakul, David Enke[13], quienes desarrollaron un modelo que explora la rentabilidad del mercado de acciones usando una red neuronal que pueda asistir la toma de decisiones de los indicadores VAMA y EMV. Esta es una red neuronal de regresión generalizada (GRNN), esta red es utilizada para pronosticar el precio futuro de las acciones. Para que el modelo neuronal pueda predecir los precios de esta forma también debe predecir el valor de EMV. Con esto se espera que los corredores de bolsa puedan entrar y salir antes de que el movimiento de la curva de las acciones cambio su curso.

Estudios recientes muestran que hay un significativo casualidad no lineal bidireccional entre las ganancias con la compra de acciones y el volumen de operaciones. En relación a este punto, Xiaotian Zhu , Hong Wang, Li Xu , Huaizu Li[14] avalan esta afirmación y muestran que el uso de redes neuronales ayuda a pronosticar el movimiento de acciones a corto medio y largo plazo. En su estudio, Xiaotian Zhu , Hong Wang, Li Xu , Huaizu Li [14], muestran combinaciones diferentes de índices y componentes del volumen de operaciones. La red neuronal se entrena con ganancias de acciones y volumen de ella de NASDAQ, DIJA y STI.

4. Teoría de Mercado Eficiente y Exponente de Hurst

La hipótesis de los mercados eficientes afirma que los mercados financieros son eficientes en relación a la información que manejan, esto quiere decir que los precios de los activos (acciones) ya tienen interiorizada toda información. La hipótesis de los mercados eficientes afirma que es imposible superar consistentemente al mercado mediante el uso de cualquier información que el mercado ya conoce.

La hipótesis de los mercados eficientes fue desarrollada por el profesor Eugene Fama [15] de la Escuela de Negocios de la Universidad de Chicago. Esta hipótesis fue ampliamente aceptada hasta la década de 1990, cuando surgieron los economistas de las finanzas conductuales. De acuerdo a la hipótesis de los mercados eficientes las acciones siempre se tranzan en su valor justo, haciendo imposible que los inversionistas compren acciones subvaloradas o vendan acciones con precios inflados, por lo que sería imposible ganarle al mercado a través de alguna estrategia de selección de acciones como el análisis técnico, y la única forma de en la que un inversionista puede obtener mayores rendimientos es mediante la compra de las acciones que presenten un mayor nivel de riesgo.

La hipótesis de los mercados eficientes emergió como una prominente teoría a mediados de 1960. En 1965 Eugene Fama publicó su disertación sobre la hipótesis del paseo aleatorio (Random Walk) y Samuelson publicó una versión de la hipótesis de los mercados eficientes. En 1970 Fama publicó un documento donde se refería tanto a la teoría como a la evidencia de la hipótesis, el documento complementaba la teoría dándole un enfoque más completo y refinado, incluyendo las definiciones de tres formas de eficiencia de mercado: Débil, Semi-fuerte y Fuerte.

Existen tres formas en la hipótesis de los mercados eficientes, cada uno de los cuales tienen diferentes implicancias en los mercados.

Donde la forma débil de eficiencia consiste en que no pueden obtenerse retornos sobrenaturales mediante el uso de estrategias de inversión basadas en los precios históricos de las acciones. La Forma Semi-fuerte de eficiencia implica que los precios de las acciones se ajustan a disposición del público mediante nueva información que se internaliza muy rápidamente y en una manera imparcial. Y por último, en la forma fuerte de eficiencia, los precios de las acciones reflejan toda la información, pública y privada, y no puede generar retornos superiores.

De ser así, sería altamente complejo hacer pronóstico sobre las acciones. Para esto agregaremos a este análisis el exponente de Hurst, para demostrar de manera empírica la volatilidad o no volatilidad de las acciones.

El coeficiente o exponente de Hurst es una medida de independencia de las series de tiempo que fue estudiada inicialmente por el científico británico Harold Edwin Hurst como elemento para distinguir series fractales. Hurst descubrió que muchos fenómenos naturales exhiben un comportamiento que puede ser caracterizado por un proceso aleatorio sesgado, en el cual existe “memoria de largo plazo” entre las observaciones, es decir, que los eventos de un periodo influyen en todos los siguientes.

Posteriormente, las ideas de Hurst fueron retomadas por Benoit B. Mandelbrot quien generalizó su trabajo y lo llamó análisis de rango reescalado (R/S), definido como un método estadístico utilizado para evaluar la ocurrencia de eventos poco comunes, dando origen a una

herramienta ideal para procesos físicos y financieros, aunque no se limita solamente a este tipo de eventos, por lo que puede ser usado en cualquier serie de tiempo.

Si $H=0.5$ (ruido blanco) implica un proceso completamente aleatorio e independiente, con ausencia de correlación entre los incrementos de la señal. Correlación = 0.

Si $0.5 < H \leq 1$ implica series de tiempo que muestran procesos persistentes o correlacionados (un periodo de crecimiento es seguido de otro análogo) y presentan un aspecto suave. Una serie persistente está caracterizada por efectos de memoria de largo plazo. Teóricamente lo que suceda hoy impactará en el futuro por siempre, todos los cambios diarios están correlacionados con todos los cambios futuros. Se ha encontrado que las series persistentes son las más comunes en la naturaleza. Un $H=1$ indicaría un comportamiento determinístico. A este comportamiento también se le atribuye el nombre de Ruido Negro, ya que se relaciona con procesos cíclicos de largo plazo, como niveles de ríos, cambios de precios en la bolsa de valores, entre otros.

Si $0 \leq H < 0.5$ corresponde a un comportamiento de anti-persistencia o anti-correlacional en la serie de tiempo (un periodo de crecimiento es seguido de otro de decrecimiento) que se caracteriza por un mayor contenido de alta frecuencia. Un incremento en los sucesos del pasado, supone un descenso en los sucesos futuros y viceversa. Un sistema anti-persistente tiende a regresar constantemente al lugar de procedencia y tienen la particularidad de ser señales muy irregulares. A este tipo de comportamientos se le conoce también como Ruido Rosa y se caracterizan por que abundan en la naturaleza y se encuentran relacionados con procesos de turbulencia.

En 1975 Mandelbrot demostró que el valor del exponente de Hurst (H) se encuentra acotado entre 0 y 1, explicando que este fenómeno ocurría debido a un síntoma de invariancia al cambio de escala, que es la característica clave de las series de tiempo fractales.

4.1. Análisis de Rango Reescalado (R/S)

Una de las formas de encontrar el parámetro o exponente de Hurst, es mediante el análisis de Rango Reescalado (R/S), el cual es un valor numérico que hace posible determinar la auto-correlación en una serie de datos.

A continuación se describe la metodología desarrollada por Hurst:

- a. Se inicia con una serie de tiempo de tamaño M , pero como interesa los rendimientos logarítmicos entonces la serie original se reduce a una nueva serie de tiempo de tamaño $N=M$ donde cada rendimiento está definido por:

$$N_i = \log\left(\frac{M_{i+1}}{M_i}\right), i = 1, 2, 3, \dots, N \quad (1)$$

- b. Este periodo de tiempo N se divide en A subperiodos contiguos de longitud n , tal que $A * n = N$. Se nombra cada uno de los subperiodos o subgrupos i_a , con $a = 1, 2, 3, \dots, A$. Cada elemento en i_a , es etiquetado $N_{k,a}$, tal que $k =$

1,2,3, ..., n. Para cada subperiodo i_a de longitud n , el valor promedio está definido por la siguiente expresión:

$$e_a = \frac{1}{n} \sum_{k=1}^n N_{k,a} \quad (2)$$

- c. Las diferencias de cada elemento $N_{k,a}$ con respecto a la media e_a para cada subperiodo i_a se van sumando para obtener la serie de tiempo acumulada ($X_{k,a}$), definida como:

$$X_{k,a} = \sum_{i=1}^k (N_{i,a} - e_a), \quad k = 1,2,3, \dots, n \quad (3)$$

- d. El rango RI_a se define como la diferencia entre el valor máximo y el valor mínimo de $X_{k,a}$ para cada subperiodo i_a :

$$RI_a = M_{ax}(X_{k,a}) - M_{in}(X_{k,a}), \quad \text{donde } 1 \leq k \leq n \quad (4)$$

- e. Se calcula la desviación estándar muestral SI_a de la forma tradicional para cada periodo i_a .

$$SI_a = \left(\frac{1}{n} \sum_{k=1}^n (N_{k,a} - e_a)^2 \right)^{1/2} \quad (5)$$

- f. Para cada periodo I_a , su rango RI_a se normaliza dividiendo por su desviación estándar muestral SI_a correspondiente. Por lo tanto el rango reescalado para cada subperiodo I_a es igual a RI_a/SI_a . Como tenemos periodos continuos de longitud, entonces tomamos el valor promedio R/S para periodos de longitud y que está definido como:

$$(R/S)_n = \left(\frac{1}{A} \sum_{a=1}^A (RI_a/SI_a) \right)^{1/2} \quad (6)$$

- g. La longitud n o el tamaño del subperiodo se incrementa al siguiente valor posible de tal forma que $\frac{(M-1)}{n}$ sea un valor entero. Iniciamos con el valor más

pequeño de acuerdo a la condición anterior y se repiten los pasos del 1 al 6 hasta $n = \frac{(M-1)}{2}$ utilizando siempre la serie completa. Posteriormente aplicamos una regresión de mínimos cuadrados de $\log(R/S)n$ contra $\log(n)$. La ordenada al origen es el $\log(c)$ y la pendiente de la ecuación es la estimación del exponente Hurst H.

5. Pronóstico con Redes Neuronales

La forma común de realizar análisis técnico es estudiando gráfico, ya sea velas japonesas, análisis de soportes y resistencias, etc. Pero basado en las investigaciones antes señalas sobre métodos no paramétricos en la predicción de los mercados accionario, se propone abordar desde una arista distinta, utilizando Redes Neuronales Artificiales como herramienta de pronóstico.

A continuación se presenta el origen, adaptación y utilización de las Redes Neuronales para construir el Modelo de Pronóstico Neuronal.

5.1. Red Neuronal Biológica

El objetivo de las redes neuronales de tipo biológico consiste en desarrollar un elemento sintáctico que permita verificar las hipótesis correspondientes a los demás sistemas biológicos. Es decir, las redes neuronales de tipo biológico deben recibir y procesar información de otros sistemas biológicos y devolver una respuesta de acción efectiva.

La mayor parte de las neuronas posee una estructura arbórea, como se muestra en, Ilustración 1, formada en su mayor parte por dendritas que, conectadas a otras neuronas, se encargan de recibir los estímulos de entrada neuronales mediante uniones denominadas sinapsis. Algunas neuronas tiene una estructura que las comunica con miles de neuronas más, mientras otras sólo puede comunicarse con unas cuantas a su alrededor.

El cerebro humano promedio cuenta con aproximadamente mil millones de neuronas. Asimismo, durante las sinapsis cada una de estas neuronas recibe en promedio alrededor de 1000 estímulos de entrada y genera alrededor de 1000 estímulos de salida. En este sentido, la principal ventaja del cerebro humano promedio recae en su conectividad, interpretada como la capacidad del mismo para realizar diferentes procedimientos lógicos a la vez. Sin embargo, su principal debilidad recae en la velocidad de procesamiento de la información, dando paso a la idea de *Redes Neuronales Artificiales* (ANN).

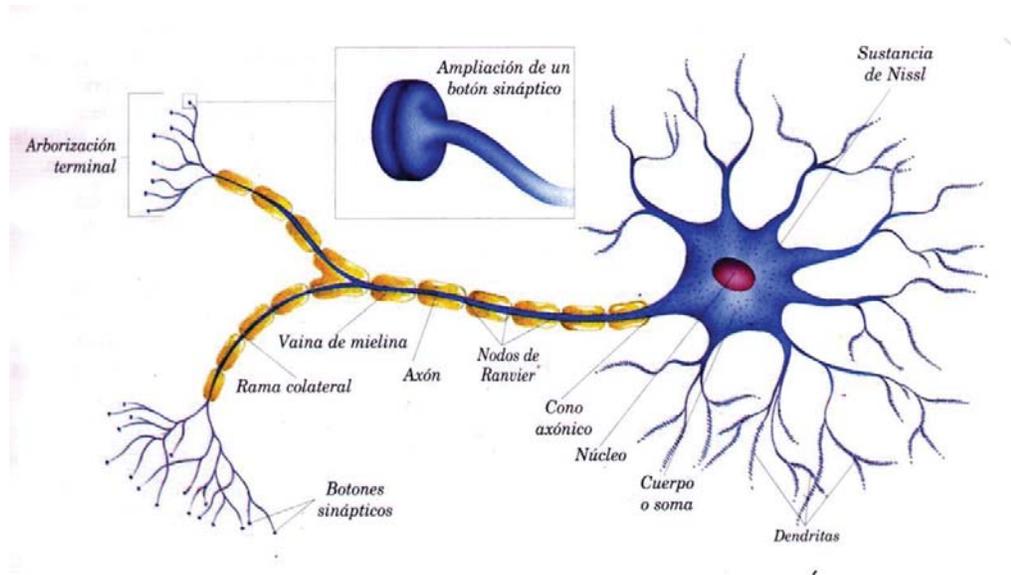


Ilustración 1 Neurona Biológica

5.2. Red Neuronal Artificial

Las Redes Neuronales Artificiales se constituyen entonces como una simulación abstracta de los sistemas nerviosos biológicos, como se muestra en Ilustración 2, formado por un conjunto de unidades llamadas neuronas o nodos conectados unos con otros. Las conexiones de estos nodos se asemejan a las dendritas y axones de los sistemas nerviosos biológicos.

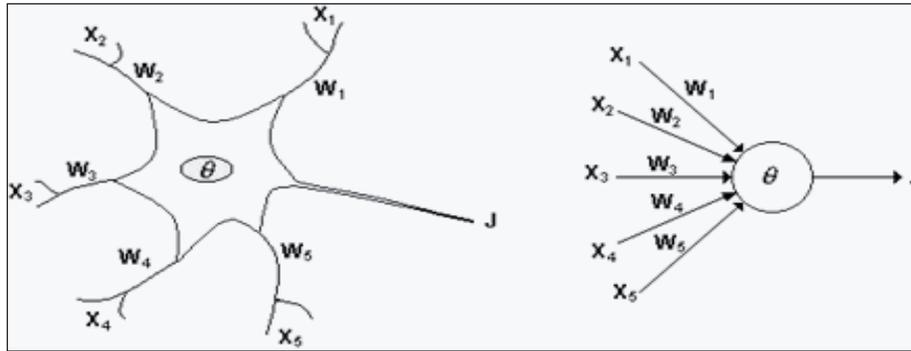


Ilustración 2 Neurona Artificial

Warren McCulloch y Walter Pitts[16] desarrollaron en 1943 el primer modelo de red neuronal en términos de modelo sistemático de actividad nerviosa, este primer modelo sirvió de ejemplo para los modelos posteriores, llegando actualmente a modelos que simplifican extremadamente el funcionamiento del cerebro, pero reteniendo algunas características esenciales biológicas:

- Las neuronas no pueden encontrarse más que en dos estados posibles, activos o en reposo.
- Están interconectadas mediante sinapsis que pueden ser modificadas por aprendizaje.
- El estado de una neurona a cada instante es determinado por el de otras, información que es transmitida por las sinapsis.

5.3. Fundamentos Matemáticos de Redes Neuronales Artificiales

La topología o arquitectura de una red neuronal consiste en la organización y disposición de las neuronas que la conforman. Estas estructuras o modelos conexionistas se encuentran fuertemente ligados al algoritmo de aprendizaje utilizado en su elaboración y se pueden clasificar de diferentes formas dependiendo del criterio de uso.

La propiedad más importante de las redes neuronales artificiales es su capacidad de aprender a partir de un conjunto de patrones de entrenamientos, es decir, es capaz de encontrar un modelo que ajuste los datos. El proceso de aprendizaje también conocido como entrenamiento de la red, puede ser supervisado o no supervisado.

El aprendizaje supervisado consiste en entrenar la red a partir de un conjunto de datos o patrones de entrenamiento compuesto por patrones de entrada y salida. El objetivo del algoritmo de aprendizaje es ajustar los pesos de la red, de manera tal que la salida generada por la ANN sea lo más cercanamente posible a la verdadera salida dada una cierta entrada. Es

decir, la red neuronal trata de encontrar un modelo al proceso desconocido que generó la salida. Este aprendizaje se llama supervisado pues se conoce el patrón de salida el cual hace el papel de supervisor de la red.

En cambio en el aprendizaje no supervisado se presenta sólo un conjunto de patrones a la ANN, y el objetivo del algoritmo de aprendizaje es ajustar los pesos de la red de manera tal que la red encuentre alguna estructura o configuración presente en los datos.

De acuerdo a su estructura las redes neuronales pueden clasificarse en los siguientes tipos: Redes Recurrentes, que son explicadas de gran manera por Chalup y Blair[17], y *Redes Neuronales Artificiales Feed-Forward* (ANN-FF).

5.4. Redes Neuronales Artificiales Feed-Forward

Son el tipo de arquitectura más simple construida en base a las neuronas descritas, usualmente se les denomina Redes Perceptrón[21]. Una red de capa única consiste en un arreglo de neuronas, de tamaño igual a la cantidad de salidas que tenga el sistema bajo análisis. Se denomina Perceptrón a una red de capa única formada por unidades lineales con umbral. Cuando dicha estructura se extiende a un mayor número de capas, se generan las *Redes Multi-Capa* (MPL), como es posible ver en la Ilustración 3. En ellas se distinguen una capa de entrada, una capa de salida, y algunas capas escondidas, de acuerdo a su posición dentro de la red. En estos modelos, el flujo de información es generalmente síncrono, es decir todos los estados de las neuronas se actualizan al mismo tiempo, o siguiendo alguna secuencia determinística. De este modo, las MPL son modelos estáticos.

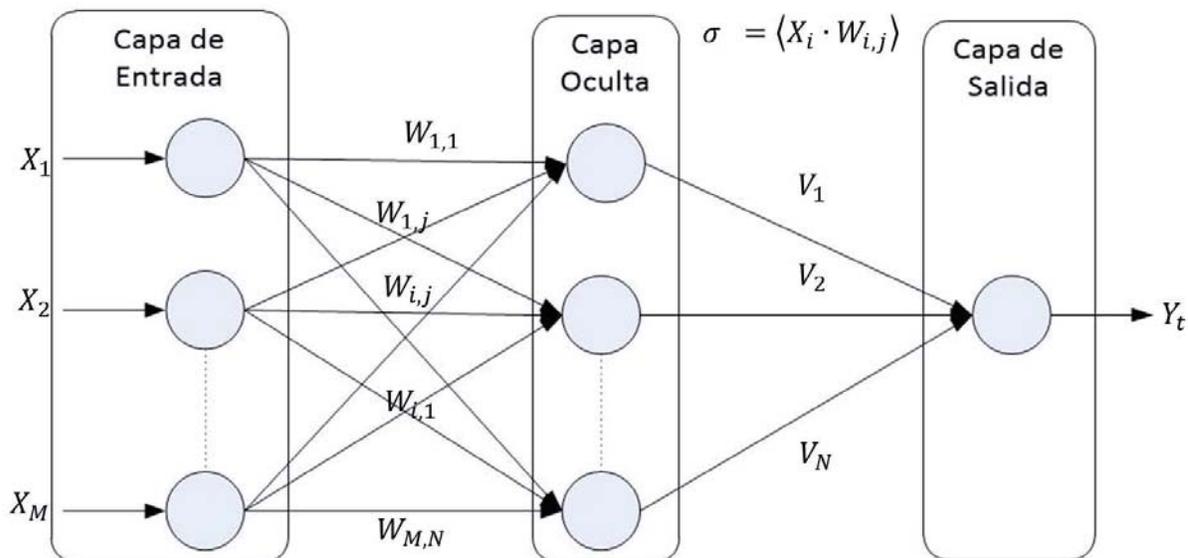


Ilustración 3 Red Neuronal Artificial

De esta forma, ANN-FF[18] con una capa oculta, que es el tipo de red que será empleada en el Pronosticador Neuronal se define como:

$$Y_t(x) = \sum_{j=0}^n V_j \sigma \left(\sum_{i=1}^m \langle W_{ij} \cdot X_i \rangle \right) \quad (7)$$

Donde m es el numero de entradas, n es el numero de nodos ocultos, W_{ij}, V_j son los pesos de conexión, X_i son los datos de entrada, $\langle W_{ij} \cdot X_i \rangle$ es el producto punto de W_{ij} y X_i , y σ es la función de activación de la ANN-FF.

5.5. Función de Activación de Base Radial

Para que una Red Neuronal Artificial pueda realizar el proceso de pronóstico, es necesario que se establezca la función de activación o función de transferencia, esta tiene la misión de transformar las entradas de ANN para su posterior salida.

Las *Redes Neuronales con Función Activación Base Radial* (RBFNN), han recibido gran atención en la literatura, por sus diversas aplicaciones, por ejemplo en pronóstico, debido a esto son producto de investigación continua, generando diversas mejoras. Donde la más utilizada es *Estándar de Gauss*(SRBF), pero existen otras variantes como, Thin-Plane Spline (TPSRBF), Multicuadrática (MRBF), Inversa Multicuadrática (IMRBF) y Cauchy (CRBF), que son profundizadas por Fernández-Navarro[19]. Una comparación entre las diferentes RBFNNs es posible ver en la Ilustración 5.

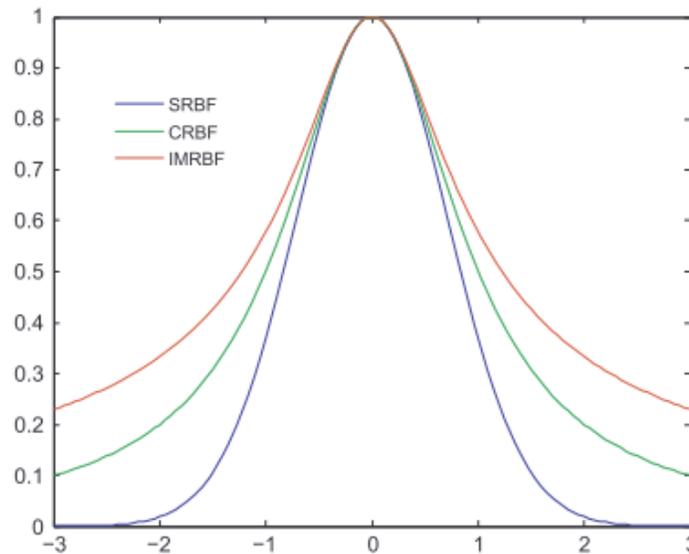


Ilustración 4 Familia de Funciones de Base Radial

Ahora bien, RBFNNs requieren menos tiempo de computacional para el aprendizaje y también tienen una topología más compacta. En las últimas investigaciones se ha optado por reemplazar la función de activación Sigmoidal, por una función de activación de Base Radial.

Una *Función Base Radial* (RBF) es una función que se ha construido en un criterio de distancia con respecto a un centro. Sea el número de nodos en la entrada p y la capa oculta sea m . Para un vector de entrada $x = (X_1, X_2, X_3, \dots, X_p)$, la salida de RBFNN es $f_{RBFNN}(x)$. El modelo viene definido por:

$$f_{RBFNN}(x) = \sum_{i=1}^m \beta_i \cdot \varphi_i(d_i(x)) \quad (8)$$

Donde $\varphi_i(d_i(x))$ es mapeado no lineal de la capa de entrada a la capa oculta, a $\beta = (\beta_0, \beta_1, \beta_2, \dots, X_m)$, es el vector incluye los pesos de la conexión entre la capa oculta y la capa de salida. La función $d_i(x)$ es definida como:

$$d_i(x) = \frac{\|x - c_i\|^2}{\theta_i^2} \quad (9)$$

Donde θ_i es el parámetro escalar que define el ancho del i-esimo unidad radial, $\|\cdot\|$ representa la norma euclidiana y $c_i = [c_1, c_2, \dots, c_p]$ es el i-simo centro de la RBFNN. El SRBF es la función Gaussiana, esta viene dada por:

$$\varphi_{SRBF_i}(d_i(x)) = e^{-d_i(x)} \quad (10)$$

5.6. Método de aprendizaje de una Red Neuronal Artificial

Hasta el momento las Redes Neuronales Artificiales estudiadas son capaces de pronosticar un solo valor en razón de sus entradas. Al considerar un horizonte de pronóstico para las Redes Neuronales Artificiales mayor a uno, la literatura científica muestra dos grandes métodos para alcanzar este objetivo, expresado en detalle por Marcellino and James[20], estos son el Método Directo y el Método Iterativo.

5.6.1. Método Directo de Pronóstico

El método directo consiste en generar n modelos diferentes para n días de horizonte de pronóstico. En el caso de las redes neuronales artificiales, esto se ve reflejado en los pesos sinápticos. Cabe señalar que este método considera las mismas entradas para cada modelo.

La desventaja de este método radica en este mismo punto, debido a que la Red Neuronal debe aprender a pronosticar días con distintos intervalos de tiempo.

El método directo viene dado por:

$$\hat{y}_t = f(x_1, \dots, x_m, w_{dim}^1) \quad (11)$$

$$\hat{y}_{t+n} = f(x_1, \dots, x_m, w_{dim}^{n+1}) \quad (12)$$

Donde x_1, \dots, x_m son las entradas al modelo, $w_{dim}^1 \dots w_{dim}^n$ son los pesos sinápticos para cada modelo y el $\hat{y}_t, \dots, \hat{y}_{t+n}$ pronóstico del modelo.

5.6.2. Método Iterativo de Pronóstico

El método iterativo consiste en generar 1 modelos diferentes para n días de horizonte de pronóstico, sin alterar los pesos sinápticos del modelo. Cabe señalar que este método considera la reinyección del pronóstico previo en el modelo.

La desventaja de este método radica en este mismo punto, debido a que la Red Neuronal se ve afectada por la propagación del error de los días pronosticados.

El método iterativo viene dado por:

$$\hat{y}_t = f(x_1, \dots, x_m, w_{dim}) \quad (13)$$

$$\hat{y}_{t+n} = f(x_{n+1}, \dots, x_m, \hat{y}_t, \dots, \hat{y}_{t+n-1}, w_{dim}) \quad (14)$$

Donde x_1, \dots, x_m son las entradas al modelo, w_{dim} pesos sinápticos del modelo, y $\hat{y}_t, \dots, \hat{y}_{t+n}$ el pronóstico del modelo.

5.7. Pronostico con Redes Neuronales Artificiales

Una de las mayores aplicaciones de las redes neuronales es el pronóstico. Las redes neuronales proveen una atractiva herramienta alternativa para los investigadores ya que estas poseen varias características que las hacen valiosas para las tareas de predicción. Primero, al contrario de los métodos tradicionales estadísticos, las redes neuronales son métodos adaptativos basado en datos, donde a veces se tiene muy poco conocimiento del problema bajo estudio. Ellas aprenden en base a ejemplos capturando relaciones funcionales sutiles entre los datos aun cuando estas relaciones sean desconocidas o difíciles de describir. Además son muy adecuadas para aquellos problemas cuyas soluciones requieren un conocimiento difícil de especificar, pero para los cuales existen suficientes datos u observaciones.

Segundo, las redes neuronales pueden generalizar, luego de la fase de entrenamiento pueden correctamente inferir el futuro de la población aun cuando los datos contengan información ruidosa. Como el pronóstico se realiza vía la predicción del comportamiento futuro en base al comportamiento pasado, el pronóstico es una aplicación ideal para las redes neuronales al menos en principio.

Por último, las redes neuronales son aproximadores funcionales universales. La experiencia de trabajos previos ha demostrado que una red puede aproximar cualquier función continua a una precisión deseada. Las redes neuronales tienen una estructura más flexible que los métodos estadísticos tradicionales. Cualquier modelo de pronóstico asume que existe una relación fundamental entre las variables de entrada (valores pasados de una serie de tiempo en este caso u otras variables relevantes) y las salidas (valores futuros). Frecuentemente los métodos estadísticos tradicionales tienen limitaciones a la hora de realizar las estimaciones de estas relaciones fundamentales debido a la complejidad de los problemas reales.

6. Enfoques de Aprendizaje de una Red Neuronal Artificial

Últimamente se encuentra en la literatura científica, diversos enfoques para el aprendizaje de ANN, escapando del clásico enfoque *Backpropagation*(BP), explicado en detalle por Haykin[21]. Además aparece el uso de algoritmos evolutivos tales como *Particle Swarm Optimization*(PSO), *Quantum Particle Swarm Optimization* (QPSO), y últimamente el Algoritmo Firefly profundizado por Nandy[22].

Aunque existen tantas Meta Heurísticas como modelos de ANN, todas comparten ciertas propiedades, que por lo general son una abstracción de un comportamiento que ocurre en la naturaleza. Pero al margen de esta situación, existen ciertos comportamientos recurrentes, como el imitar enjambres de abejas, o colonias de hormigas. Donde por lo general, las partículas van moviéndose para encontrar una solución local, que después será solución del sistema.

El problema que presentan, en general las Meta heurísticas, son el alto tiempo computacional en el proceso de aprendizaje. Constituyendo el hecho de un modelo lento, pero con un buen nivel de convergencia.

A continuación se explica el algoritmo BSA, implementado en el Pronosticador Neuronal.

6.1. Algoritmo BSA

En el año 2013 Pinar Civicioglu[23], propuso el algoritmo BSA (Backtracking Search Optimization Algorithm), este es un Algoritmo Evolutivo (EA) iterativo basado en poblaciones, diseñado para ser un minimizador global. BSA puede ser explicado dividiendo sus funciones en 5 procesos, siguiendo la estructura de los EAs. Estos pasos son; Inicialización, Selección-I, mutación, cruza y selección-II.

En Tabla 2, se muestran la estructura general de BSA.

Tabla 1 Algoritmo BSA

<ul style="list-style-type: none">• Inicialización• Repetir<ul style="list-style-type: none">○ Selección I○ Generación de Población de Prueba<ul style="list-style-type: none">▪ Mutación▪ Cruza○ Selección II• Hasta que las condiciones de detención sean alcanzadas

6.1.1. Inicialización

La inicialización de BSA, se realiza utilizando la siguiente ecuación:

$$P_{ij} \sim U(\text{low}_j, \text{up}_j) \quad (15)$$

Para $i = 1, \dots, N$ y $j = 1, \dots, D$, donde N y D son el tamaño de la población y la dimensión del problema respectivamente. U es la distribución uniforme y P_i es el identificador individual de la población P .

6.1.2. Selección-I

Esta etapa del BSA determina la población histórica $oldP$ para ser usada en el cálculo de la dirección de búsqueda. La inicialización de la población histórica viene dada por:

$$oldP_{ij} \sim U(\text{low}_j, \text{up}_j) \quad (16)$$

BSA tiene la opción de redefinir $oldP$ al comienzo de cada iteración a través de la regla del “if then” mostrada a continuación:

$$\text{if } a < b \text{ then } oldP := P|a, b \sim U(0,1) \quad (17)$$

Donde $:=$ es la operación de actualización. La ecuación anterior asegura que BSA designa una población perteneciente a una anterior, seleccionando al azar una generación como la población histórica y recordando esta población histórica hasta que se cambie. Por lo tanto, la BSA tiene memoria.

Después se cambia el orden de los individuos en $oldP$, mediante una función de intercambio aleatorio, como se muestra en la siguiente ecuación:

$$oldP := \text{permuting}(oldP) \quad (18)$$

6.1.3. Mutación

El proceso de mutación de BSA genera la forma inicial de la población de prueba de *Mutant* usando la siguiente ecuación.

$$Mutant = P + F \cdot (oldP - P) \quad (19)$$

En la ecuación anterior, F controla la amplitud de la matriz de búsqueda-dirección $(oldP - P)$. Porque la población histórica es usada en el cálculo de la matriz búsqueda-dirección. BSA genera una población de prueba, tomando ventaja parcial de sus experiencias de la generación anterior.

6.1.4. Cruza

El proceso de cruza de BSA genera la forma final de la población de prueba T . El valor inicial de la población de prueba es $Mutant$, como conjunto en el proceso de mutación. Los individuos de prueba con mejor valor fitness para la optimización del problema son usados en la evolución de los individuos de la población objetivo. El proceso de cruza BSA tiene dos etapas. La primera etapa calcula una matriz de valores enteros binarios (map) de tamaño $N \cdot D$ que indican los individuos de T a ser manipulados usando la relevancia individual de P . Si $map_{n,m} = 1$, donde $n \in \{1, \dots, N\}$ y $m \in \{1, \dots, D\}$. Donde T es actualizado con $T_{n,m} := P_{n,m}$.

En la siguiente tabla se muestra la estrategia de cruza.

Tabla 2 Cruza I - BSA

Entrada: Mutant, mixrate, N, D
Salida: T //Población de Prueba
<ul style="list-style-type: none"> • $map_{(1:N,1:D)} = 1$ //matriz de Nx D de valores 1 • Si $a < b$ $a, b \sim U(0,1)$ entonces <ul style="list-style-type: none"> ○ Desde $i = 1$ hasta N <ul style="list-style-type: none"> ▪ $map_{i,u_{(1:[mixrate \cdot rnd \cdot D])}} = 0$ $u = \text{permutar}(\langle 1,2,3, \dots, D \rangle)$ ○ Fin • Sino <ul style="list-style-type: none"> ○ Desde $i = 1$ hasta N <ul style="list-style-type: none"> ▪ $map_{i,randi(D)} = 0$ ○ Fin • Fin • $T = Mutant$ • Desde $i = 1$ hasta N <ul style="list-style-type: none"> ○ Desde $j = 1$ hasta D <ul style="list-style-type: none"> ▪ Si $map_{i,j} = 1$ entonces $T_{i,j} = P_{i,j}$ ○ Fin • Fin

En el algoritmo anterior, en la línea 3, se indica la función tope , definida como $rnd \sim U(0,1)$. La estrategia de cruza de BSA es bastante diferente de otras estrategias usadas en Evolución Diferencial (DE) y sus variantes. La tasa de mezcla (mixrate) en el proceso de cruza BSA controla el número de elementos de los individuos que mutarían en una prueba mediante $[\text{mixrate} \cdot \text{rnd} \cdot D]$.

Las estrategias predefinidas son el uso aleatorio para definir el map de BSA. La primera estrategia es usar mixrate . La segunda estrategia permite solamente elegir un individuo aleatoriamente para mutar en cada prueba. El proceso de cruza BSA es más complejo que los utilizados en DE.

Algunos individuos obtenidos de la poblacion de prueba, al final de el proceso de cruza BSA puede sobrepasar el espacio de busqueda permitido, limitando la estrategia de mutación de BSA. Los indivisuos mas alla del epacio de buisqueda son regenerados usando el siguiente algoritmo.

Tabla 3 Cruza-II BSA

Entrada: T, Búsqueda limite espacial (i.e., low_j, up_j)

Salida: T

- **Desde $i = 1$ hasta N**
 - **Desde $j = 1$ hasta D**
 - **Si $(T_{i,j} < low_j \text{ o } T_{i,j} > up_j)$ entonces**
 - $T_{i,j} = rnd * (up_j - low_j) + low_j$
 - **Fin**
 - **Fin**
- **Fin**

6.1.5. Selección-II

En la etapa de Selección-II de BSA, el T_i s que tienen mejores valores de fitness que el correspondiente P_i s es usado para actualizar el P_i s basado sobre una selección ambiciosa. Si el mejor individuo de P (P_{best}) tiene un mejor valor fitness que el mínimo valor global obtenido hasta aquí por BSA. La estructura de BSA es bastante simple, por lo que es muy fácil de adaptar para diferentes problemas de optimización, entre estos las Redes Neuronales Artificiales.

A continuacion, Tabla 5, se muestra el algoritmo Selección-II de BSA.

Tabla 4 Selección-II BSA

- $fitnessT = ObjFnc(T)$
- **Desde $i = 1$ hasta N**
 - **Si $fitnessT_i < fitnessP_i$ entonces**
 - $fitnessP_i = fitnessT_i$
 - $P_i = T_i$
 - **Fin**
- **Fin**
- $fitnessP_{best} = \min(fitnessP) \mid best \in \{1,2,3, \dots, N\}$
- **Si $fitnessP_{best} < globalminimum$ entonces**
 - $globalminimum = fitnessP_{best}$
 - $globalminimizer = P_{best}$
 - *// Exportar $globalminimum$ y $globalminimizer$*
- **Fin**

7. Pronosticador Neuronal Propuesto

El modelo propuesto toma parte del trabajo expuesto por Asadi [24], aunque no se considera el análisis utilizando Algoritmos Genéticos para la selección de variables en la entrada de ANN. Entonces, el Pronosticador Neuronal es una ANN-FF de tres capas, una capa de entrada, una capa oculta y la capa de salida. El motivo para tal elección respecto al número de capas oculta, es que está comprobado que al aumentar el número de capas intermedias, no hay un aumento significativo en la calidad del pronóstico, peor aún, en algunos casos empeora la ANN, hay que considerar también que aumenta su complejidad de ANN.

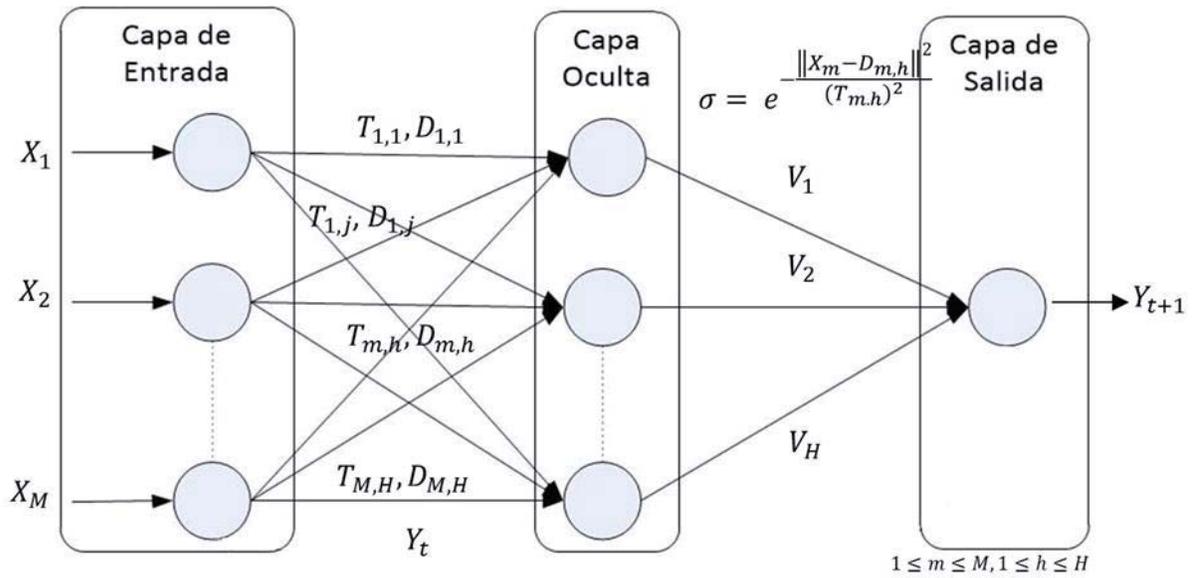


Ilustración 5 Red Neuronal Artificial Propuesta

Para la función de activación se consideró la función Gaussiana (SRBF) que es parte de la familia de funciones de base radial explicadas en el capítulo 5. Cabe señalar que la función Fitness es GCV, con la cual es posible calibrar el número de entradas al modelo.

$$fitness = GCV(Y_t, Y) \quad (20)$$

Debido a que el Pronosticador Neuronal propuesto es un modelo univariado, los datos de entrada considerados son el precio de cierre de cada acción.

Para el aprendizaje del Pronosticador Neuronal, se utilizó un enfoque evolutivo, es por este motivo que el algoritmo elegido fue BSA, el cual es explicado en el Capítulo 6.

$$W' = BSA(W), \text{ donde } W \text{ es } T, D, V \quad (21)$$

En relación al horizonte de Tiempo del pronóstico se utilizó un enfoque híbrido entre el método Directo y Método Iterativo. Este consiste en calibrar el Pronosticador Neuronal para cada día de pronóstico, lo que se ve reflejado en los pesos sinápticos. Además es utilizado el día pronosticado previo como entrada para el día $t+1$.

$$\hat{Y}_t(X_1, \dots, X_M) = \sum_{h=1}^H V_h^1 \left(\sum_{m=1}^M \sigma(X_1, \dots, X_m, T_{m,h}^1, D_{m,h}^1) \right) \quad (22)$$

$$\begin{aligned} \hat{Y}_{t+n}(X_{n+1}, \dots, X_M, \hat{Y}_t, \dots, \hat{Y}_{t+n-1}) = \\ \sum_{h=1}^H V_h^{n+1} \left(\sum_{m=1}^M \sigma(X_{n+1}, \dots, X_m, \hat{Y}_t, \dots, \hat{Y}_{t+n-1}, T_{m,h}^{n+1}, D_{m,h}^{n+1}) \right) \end{aligned} \quad (23)$$

Donde X_1, \dots, X_m son entradas al modelo, $T_{m,h}^0, \dots, T_{m,h}^n, D_{m,h}^0, \dots, D_{m,h}^n, V_h^0, \dots, V_h^n$ pesos sinápticos del modelo, y $\hat{Y}_t, \dots, \hat{Y}_{t+n}$ pronóstico del modelo.

7.1. Proceso de Aprendizaje Pronosticador Neuronal

A continuación se detallan las etapas del Proceso de Aprendizaje del Pronosticador Neuronal.

ETAPA I: Pre-Procesamiento

En esta etapa se utiliza el siguiente método de normalización, Ecuación, la cual consiste en dividir cada precio de cierre de la acción por el máximo valor de la serie, de esta forma la serie normalizada oscila entre 0 y 1.

$$X_{new}^* = \left(\frac{X}{X_{max}} \right) \quad (24)$$

Donde X es un valor a normalizar de la serie y X_{max} es el valor máximo de la serie.

Los datos de entrada al Pronosticador Neuronal, son normalizados por la sencilla razón de eliminar algún ruido, de esta forma se escalan los atributo a un conjunto de valores apropiados para este caso.

ETAPA II: Entrenamiento utilizando BSA

En esta fase, se reciben los valores normalizados, como se muestra en la Ilustración 6, posteriormente, se inicializan los pesos de manera aleatoria. Comenzando el proceso de entrenamiento del Pronosticador Neuronal.

ETAPA III: Pruebas de Rendimiento

Con el Pronosticador Neuronal calibrado, se evalúa el rendimiento para obtener el mejor modelo para cada acción estudiada.

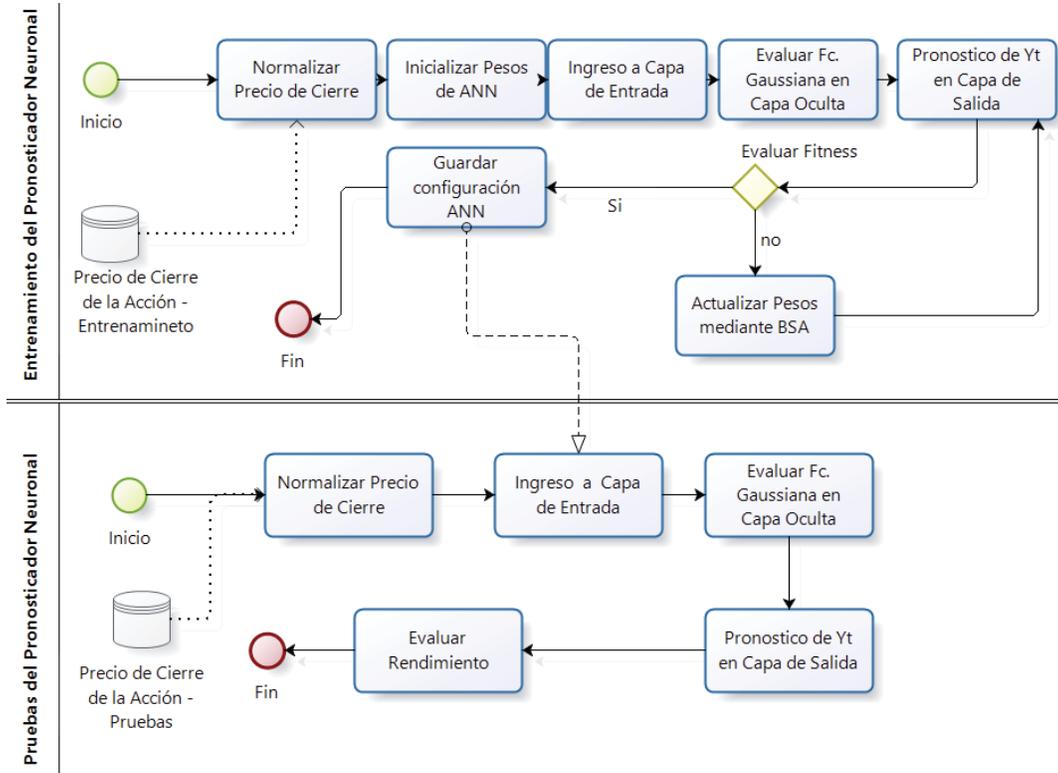


Ilustración 6 Proceso de Aprendizaje y Pruebas del Pronosticador Neuronal

7.2. Métricas de Rendimiento

Para evaluar el rendimiento del Pronosticador Neuronal se utilizaron las siguientes métricas. Estas son, *Error Cuadrático Medio (MSE)*, *Raíz Error Cuadrático Medio (RMSE)*, *Error Absoluto Medio (MAE)*, *Error Porcentual Absoluto Medio (MAPE)*, *Coficiente de Correlación (R)*, *Coficiente de Determinación (R²)*, *Validación Cruzada Generalizada (GCV)* y *Dirección de la curva (DAC)*.

MSE viene dada por:

$$MSE = \frac{1}{n} \sum_{t=1}^n (y'_t - y_t)^2 \quad (25)$$

RMSE viene dada por:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y'_t - y_t)^2} \quad (26)$$

MAE es definido por:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y'_t - y_t}{y_t} \right| \quad (27)$$

MAPE es definido por:

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{y'_t - y_t}{y_t} \right| \quad (28)$$

R es definido por:

$$R = \frac{cov(y_t, y'_t)}{\sigma(y_t) \cdot \sigma(y'_t)} \quad (29)$$

R² es definido por:

$$R^2 = \frac{\sum_{t=1}^n (y'_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (30)$$

GCV es definido por:

$$GCV = \frac{MAPE}{\left(1 - \frac{P}{N}\right)^2} \quad (31)$$

DAC es definido por:

$$DAC = \frac{\sum_{t=1}^N \text{signo}(y_t - y_{t-1}) \equiv \text{signo}(y'_t - y'_{t-1})}{N} \quad (32)$$

Donde n es el número de observaciones consideradas, P el número de parámetros, y_t es el precio real, y'_t es el precio estimado por el modelo, e \bar{y} es la media del precio real.

8. Análisis y Discusión de Resultados.

En este capítulo, son expuestos los resultados obtenidos al evaluar Pronosticador Neuronal, frente a 3 acciones transadas en la Bolsa de Comercio de Santiago.

8.1. Análisis del exponente de Hurst

Para dimensionar la volatilidad del mercado accionario, recurrimos al Exponente de Hurst, calculado mediante el análisis de Rango Re escalado, descrito en el Capítulo 3. Donde evaluamos las 40 acciones pertenecientes al IPSA [25], como es posible ver en la siguiente Ilustración.

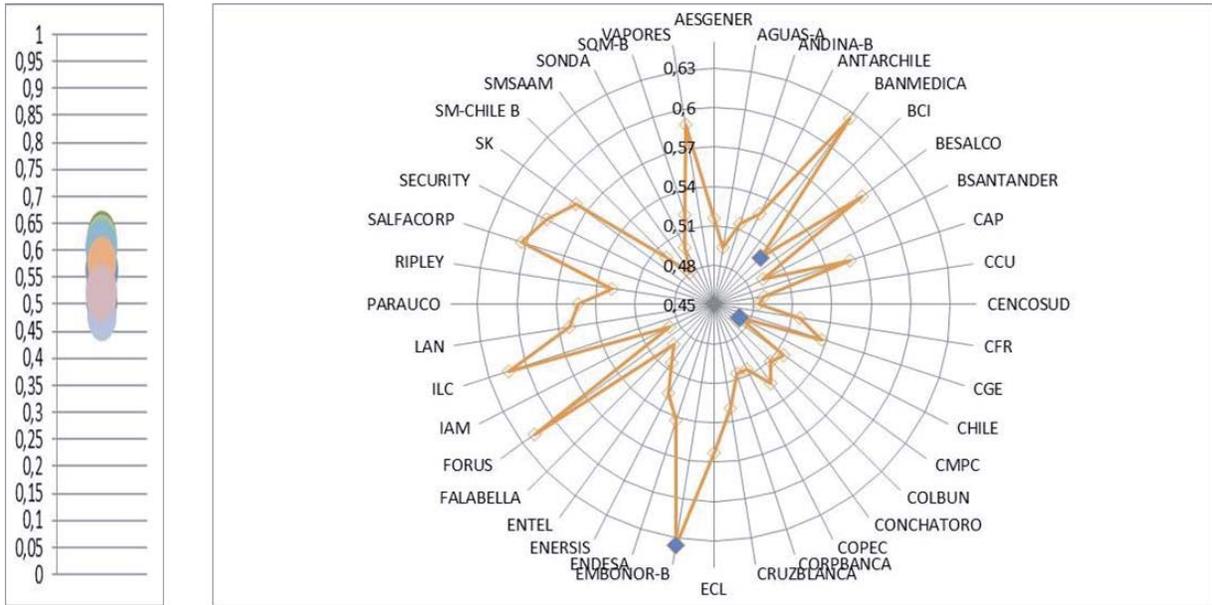


Ilustración 7 Exponente de Hurst - Acciones pertenecientes al IPSA

En esta investigación se tomó la acción de mayor coeficiente Coca-Cola (EMBONOR-B), que posee un $H=0,63548$, dado que posee la mayor correlación y muestra un comportamiento persistente. También se seleccionó la acción de BCI donde $H= 0,50026$, el cual es el mejor ejemplo de ruido blanco del total de acciones del IPSA. Finalmente la acción de Banco de Chile (CHILE) tiene un $H=0,47187$ lo que implica que posee un comportamiento de anti-persistencia.

De esta forma, se toman acciones representativas de cada fenómeno para analizar el rendimiento del Pronosticador Neuronal.

8.2. Descripción de los Datos de Entrada

En esta investigación fueron utilizados el precio de cierre diario de las acciones antes señaladas (EMBONOR-B, BCI, CHILE). Estos datos fueron obtenidos desde la Bolsa de Comercio de Santiago.

Para los tres papeles accionarios, fue considerado el mismo horizonte de tiempo. Este es el precio de cierre diario desde el 06 de junio de 2004 hasta el 01 de junio de 2014.

Por lo tanto, se tienen 2485 datos para cada acción. Separando estos datos para entrenamiento y pruebas, donde el 80% de los datos son para entrenamiento y el 20% es para pruebas.

8.3. Pronosticador Neuronal utilizando aprendizaje BSA

El Pronosticador Neuronal propuesto es entrenado bajo el algoritmo evolutivo BSA. Este algoritmo requiere de los siguientes parámetros ($Iter = 1000$; $P = 10$; $low = -1$; $up = 1$; $= 10$), para el correcto funcionamiento.

Donde, $Iter$ es el máximo de épocas, P el tamaño de la población, low y up , son cotas del dominio de la solución y Sim es el número de simulaciones. Además como estrategia de generación de escala se utiliza STANDARD brownian-walk, esto es $F = 3 * rndn$, donde $rndn \sim N(0,1)$ (N es la distribución Normal estándar).

Los nodos ocultos fueron calculados en base al tamaño de la muestra de entrenamiento, de la siguiente forma $\log_2 Tr$, lo que dio por resultado $H = 10$, donde Tr es el tamaño de la muestra y H el número de nodos ocultos.

8.4. Número de Entradas al Pronosticador Neuronal

Para obtener buenos resultados en el pronóstico, es necesario que las entradas al modelo sean las requeridas por cada acción, esto es que las entradas al Pronosticador Neuronal deben contener la información de alta frecuencia para el pronóstico, debido a que el modelo capturo en los pesos sinápticos la baja frecuencia de la acción.

Para esto, se probó con un rango de 6 a 10 días para cada acción, para obtener los números de días de entrada al modelo óptimos para el Pronosticador Neuronal.

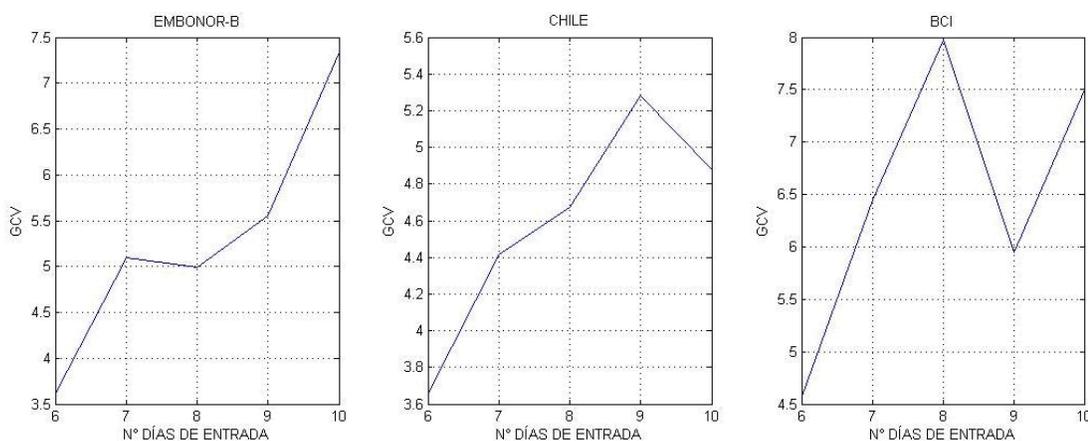


Ilustración 8 Entradas al Pronosticador Neuronal

Como se muestra en la Ilustración anterior, para las tres acciones estudiadas, la mejor configuración de días de entrada al modelo es 6 días.

8.5. Rendimiento del Pronosticador Neuronal

A continuación se detallan los resultados obtenidos en la etapa de pruebas para cada acción. Se consideran 5 días de pronóstico para cada acción. Evaluando en forma individual el modelo por cada día pronosticado.

Tabla 5 Resultados Pronosticador Neuronal

Acción	Día	MSE	RMSE	MAE	MAPE	R	R ²	GCV	DAC
EMBONOR-B	1	0.000056	0.010562	0.008862	2.4154%	92.2814%	86.2130%	4.8328%	39.0144 %
	2	0.000037	0.008605	0.006618	1.7221 %	88.6791%	78.6398%	3.4456%	42.2998 %
	3	0.000068	0.011632	0.009124	2.4053 %	68.5302%	46.9639%	4.8126%	44.9692 %
	4	0.000084	0.012941	0.009871	2.6175 %	66.1974%	43.8210%	5.2371%	43.5318%
	5	0.000049	0.009946	0.007786	2.0870 %	71.4033%	84.5004%	4.1758%	43.9425 %
CHILE	1	0.000087	0.013223	0.010622	2.7107%	88.6791%	78.6398%	6.2005%	40.9465 %
	2	0.000053	0.010301	0.008620	2.3345%	81.7483%	66.8278%	5.3400%	40.9465 %
	3	0.000210	0.020492	0.015394	4.4631%	87.3131%	76.2357%	10.2090%	44.4444 %
	4	0.000161	0.017930	0.013556	3.4721%	63.2323%	39.9829%	7.9423%	44.6502%
	5	0.000243	0.022053	0.016822	4.4265%	35.3227%	12.4769%	10.1253%	43.6214%
BCI	1	0.000033	0.008120	0.006508	2.0742 %	97.0801 %	94.2455%	4.1500%	43.9425 %
	2	0.000077	0.012415	0.010102	3.1332 %	91.9591 %	84.5648 %	6.2689%	47.4333 %
	3	0.000382	0.027657	0.025252	7.4266%	91.2303%	83.2297 %	14.8592%	47.0226 %
	4	0.000311	0.024937	0.018648	5.6897%	56.3197 %	31.7191%	11.3840%	44,3532 %
	5	0.000155	0.017634	0.014175	4.5985 %	84.2642%	71.0047 %	9.2007%	47.0226 %

Es importante analizar que para las tres acciones el MSE es de muy pequeño, del orden de 10^{-4} , por otro lado si analizamos el MAPE, este está entre el 1,3% y el 7%, considerando una media de 3,41% y una desviación estándar de 0,0018. Demostrando que el nivel del Pronosticador Neuronal es excelente si es analizado por las métricas antes mencionadas. Pero al analizar la dirección de la curva, el panorama es desolador. En general esta métrica esta entre el 40% y 50% para las tres acciones. Esto se explica por un retardo o desfase en capturar el movimiento de cada acción. Esto es posible apreciar en las siguientes Ilustraciones.

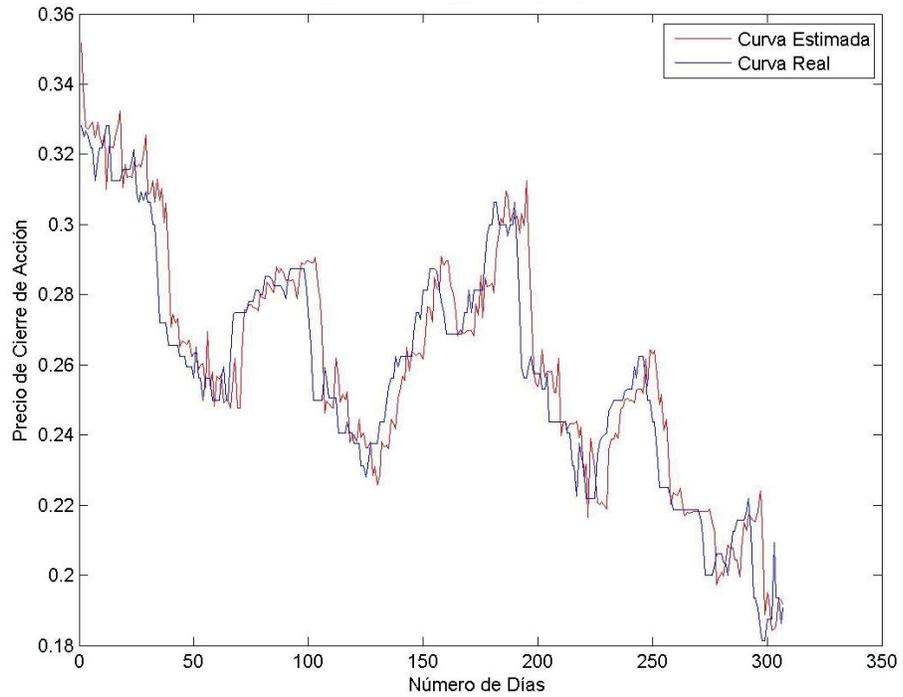


Ilustración 9 Precio Real vs Precio Estimado - EMBONOR - B

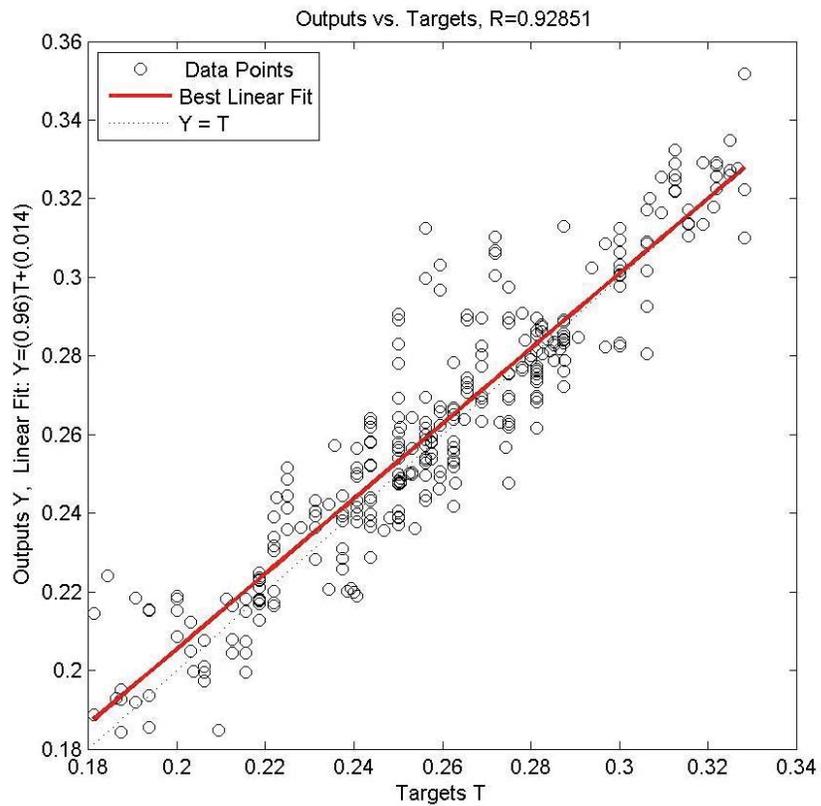


Ilustración 10 Correlación - EMBONOR - B

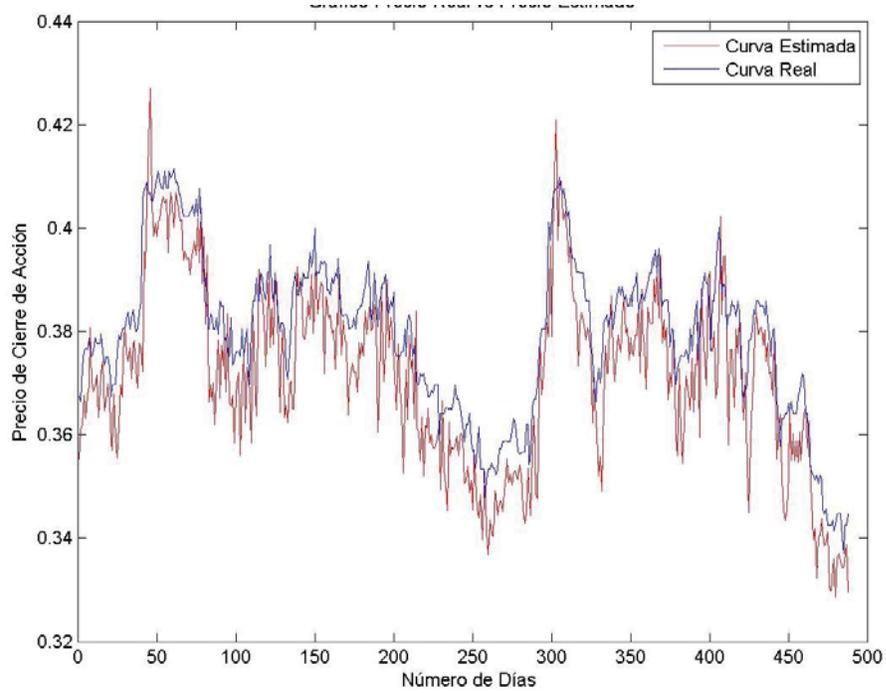


Ilustración 11 Precio Real vs Precio Estimado - CHILE

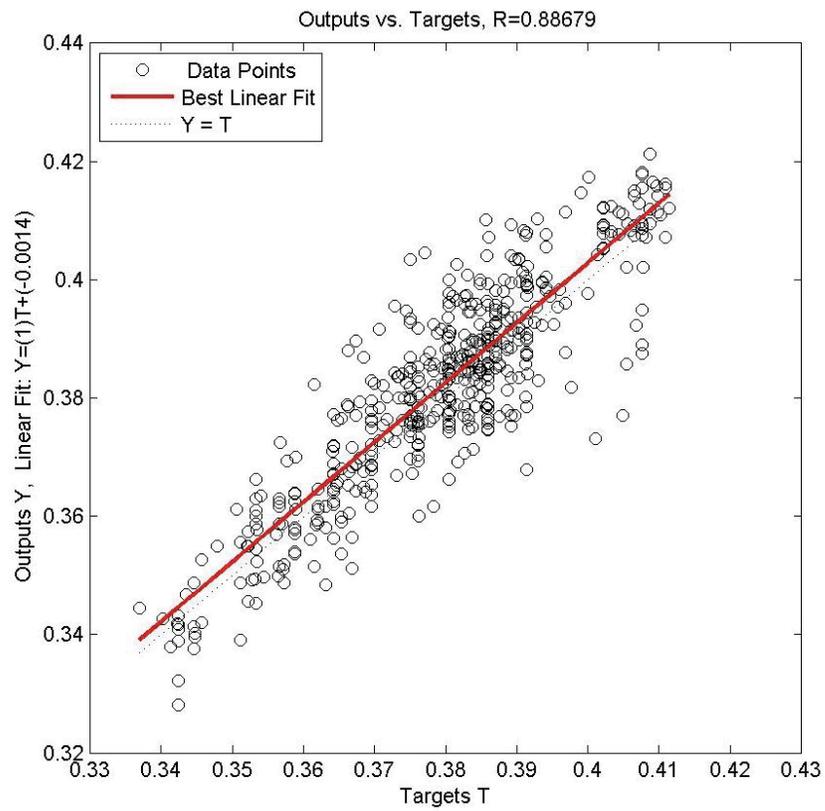


Ilustración 12 Correlación - CHILE

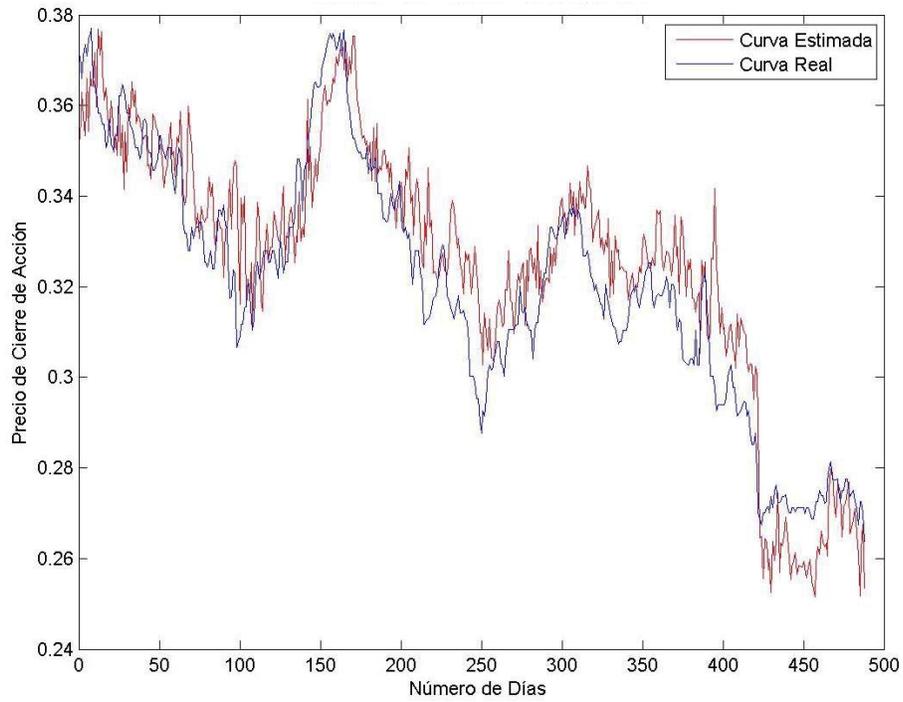


Ilustración 13 Precio Real vs Precio Estimado - BCI

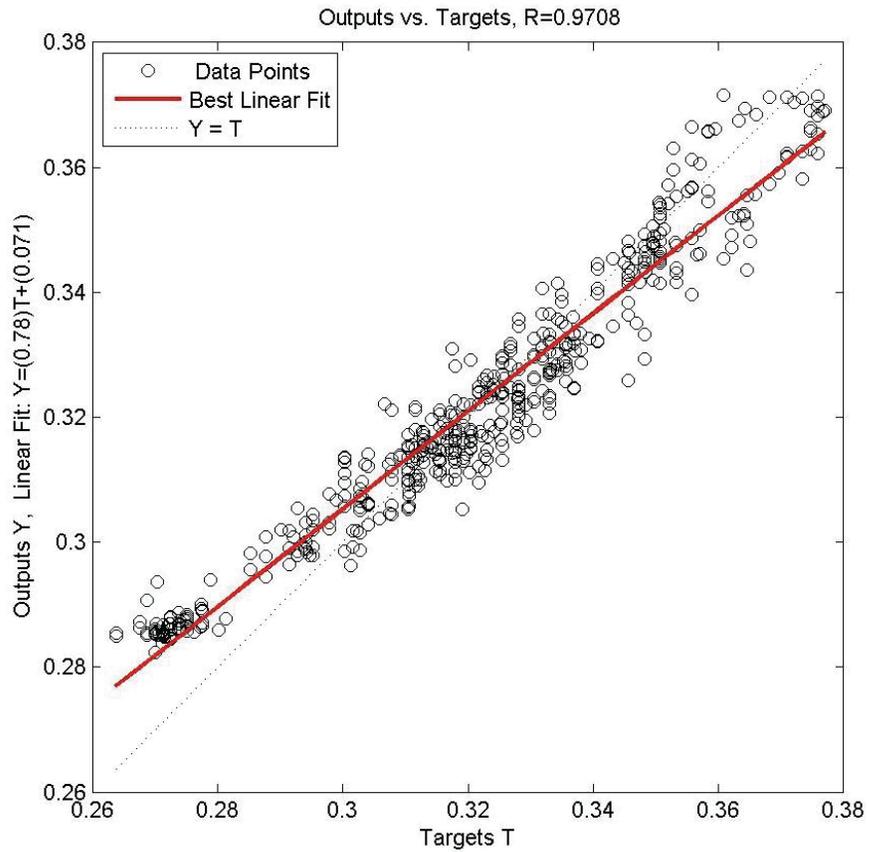


Ilustración 14 Correlación - BCI

8.6. Pronosticador Neuronal en Producción

En esta etapa se prueba el Pronosticador Neuronal con datos fuera de la muestra de entrenamiento y pruebas.

Donde utilizando el mejor modelo encontrado por cada acción, se pronostica el comportamiento de la acción para los 5 días futuros.

En el caso de EMBONOR-B el comportamiento fue el siguiente. Si bien el Pronosticador Neuronal no se ajusta el precio real de la acción, si es capaz de captar la tendencia para la semana. De igual forma la rentabilidad calculada es positiva para toda la semana, siendo más mesurada que la rentabilidad real. Es interesante que la dirección de la curva sea del 75%, debido que la rentabilidad si bien va a la baja, sigue siendo mayor que cero.

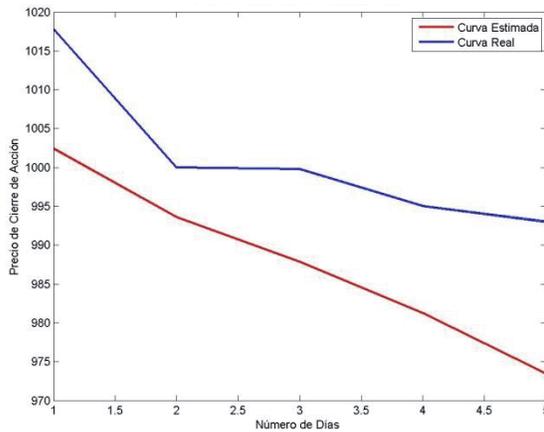


Ilustración 15 Pronóstico del Precio - EMBONOR-B

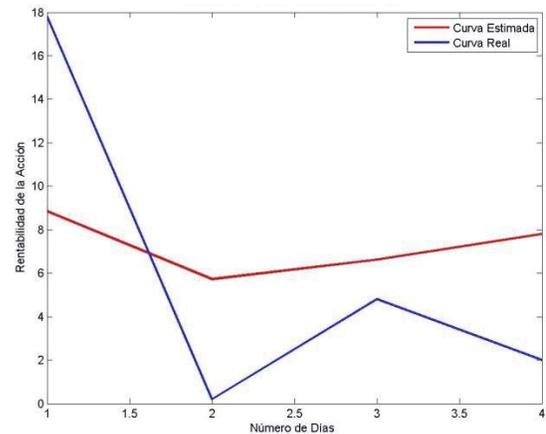


Ilustración 16 Pronóstico de la Rentabilidad - EMBONOR-B

Al realizar el pronóstico sobre la acción de banco de Chile los resultados son similares en relación a EMBONOR-B, si vemos que el precio y la rentabilidad proyectados no coinciden con el precio y la rentabilidad real. Pero la tendencia del precio de la acción es capturada por el modelo, el cual cuenta con una dirección de la curva del 75%, esto quiere decir que el Pronosticador Neuronal captura las mínimas alzas y bajas del precio de la acción, que tiene carácter decimal (centavos), pero amplificando estas alzas o bajas en algunos pesos. Esto es posible apreciarlo en las Ilustraciones 16 y 17.

De igual forma al pronosticar los siguientes 5 días de la acción de BCI, el modelo estimo el precio de la acción para el primer día en \$33.487 a diferencia de los \$32.276 que correspondía al primer día. De ahí en más, el Pronosticador Neuronal obtuvo un 100% en la dirección de la curva cómo es posible ver en las Ilustraciones 18 y 19.

Cabe señalar que el MAPE para cada acción (EMBONOR-B, CHILE y BCI) fue 1.35%, 2.10%, 2.63% respectivamente.

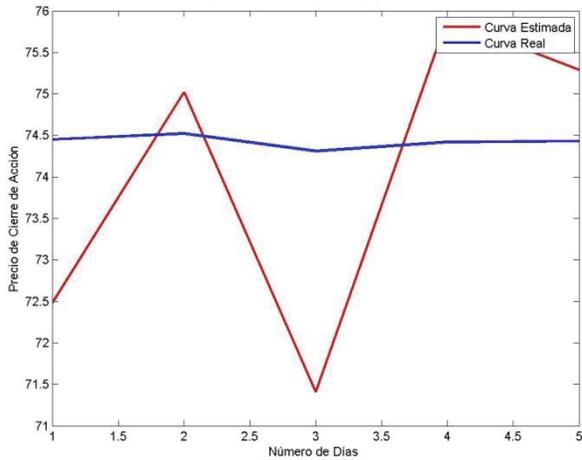


Ilustración 17 Pronostico del Precio - CHILE

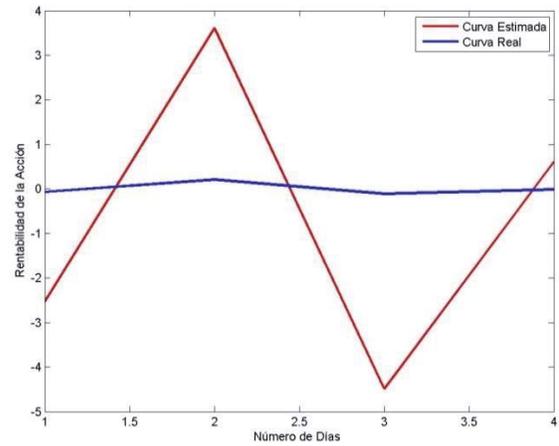


Ilustración 18 Pronostico de la Rentabilidad - CHILE

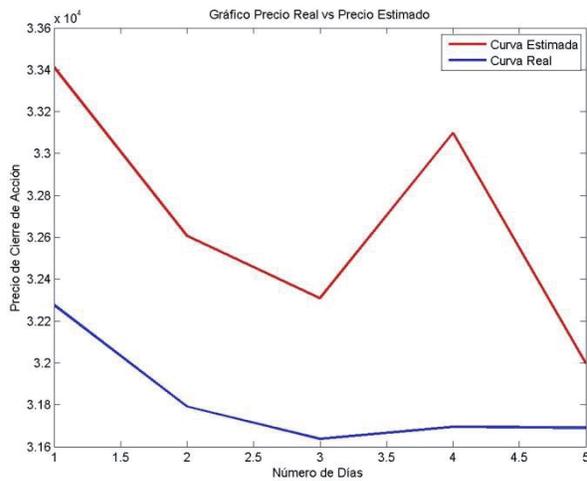


Ilustración 19 Pronostico del Precio - BCI

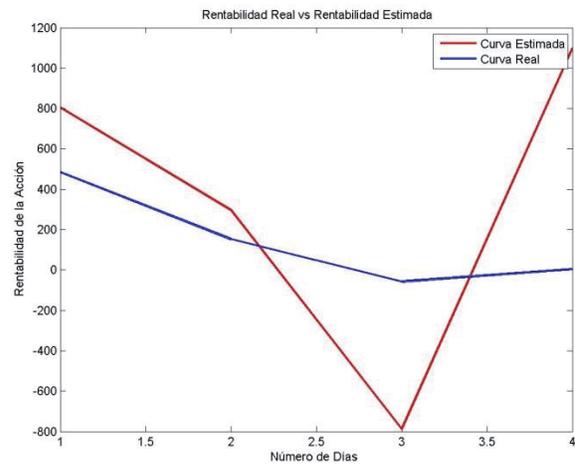


Ilustración 20 Pronostico de la Rentabilidad - BCI

Los precios de cierre que entraron al modelo corresponden al precio de cierre del día 5 de junio del 2014 hasta el 12 de Junio de 2014, los días pronosticados van del día 13 de junio hasta el 20 de junio. Considerar que las acciones solo se transan días hábiles.

9. Conclusiones

Como punto de partida, en esta investigación, se realizó un análisis sobre las acciones pertenecientes al IPSA mediante el exponente de Hurst. Con la finalidad de tener una prueba empírica de la volatilidad del mercado accionario, y poder seleccionar la mejor acción para pronosticar mediante este índice. Donde el papel accionario de EMBONOR-B es el más predecible debido a su índice de Hurst ($H=63,5\%$). También se seleccionó 2 acciones más, estas son CHILE y BCI, particularmente estas dos acciones cumplen el supuesto contrario, son acciones altamente impredecibles debido a su índice de Hurst, $H=47,1\%$ y $H=50\%$ respectivamente.

Gracias a lo anterior, es posible afirmar que el mercado accionario chileno, representado por las acciones correspondientes al IPSA, presenta una forma débil de eficiencia, debido a que las acciones transadas en la Bolsa de Comercio de Santiago, no internaliza rápidamente la información del mercado de capitales chileno. Donde solo el 12,5% de las acciones supera el 60% en el índice de Hurst, dificultando la predictibilidad del mercado de accionario chileno.

Respecto a los datos de entrada, se comprobó que el precio de cierre contiene información necesaria para el pronóstico, donde GCV varía de 3,5% a 4,5% en las acciones estudiadas. Es interesante que para las tres acciones estudiadas el número de entradas corresponda a 6 días, y por consiguiente al aumentar el número de días empeore la predictibilidad de las acciones, alcanzando un GCV promedio de 6,9%.

Un punto importante en el aprendizaje del modelo viene dado por la elección del algoritmo genético BSA, como meta heurística de aprendizaje. BSA obtuvo un tiempo de cómputo en el entrenamiento de 20 minutos promedio por modelo, alcanzando pesos sinápticos con un MSE promedio del orden de 10^{-4} para las tres acciones estudiadas. Por otro lado, BSA muestra un alto grado de correlación en la convergencia con el índice de Hurst, a mayor índice de Hurst más rápido converge el algoritmo de aprendizaje, en el caso de EMBONOR-B converge cerca de las 290 iteraciones, en cambio las acciones de CHILE y BCI convergen alrededor de las 450 iteraciones.

Considerando lo anterior, los resultados obtenidos por el Pronosticador Neuronal en la fase de prueba arrojan muy buenos resultados en las métricas estadísticas, pero en cambio al analizar la dirección de la curva (DAC), no existe una correlación entre ellas. En el caso de la acción de CHILE el MAPE fluctúa entre 2.0742 % y 4.4631%, de forma similar en el caso de las otras dos acciones, obteniendo un MAPE promedio de 3,43%, en las tres acciones. Pero la dirección de la curva no supera el 50% en ninguna de las tres acciones. Esto demuestra que los supuestos estadísticos no tienen consigo una significancia en la rentabilidad a largo plazo.

En la fase de producción del modelo, quedo en evidencia que pronosticar el precio de la acción o el pronóstico de la rentabilidad es altamente complejo. Pero si es posible capturar la tendencia de la acción para el horizonte de 5 días. Por lo tanto, es posible dar una recomendación al inversionista si la acción va al alza, se mantiene en el orden o simplemente va a la baja. El comportamiento señalado aplica para los tres casos estudiados, donde no importa lo caótico que pueda ser una acción, para el Pronosticador Neuronal es factible capturar la tendencia del precio de las acciones, obteniendo las tres acciones sobre un 75% en la dirección de la curva.

Ahora bien, si se analiza el precio pronosticado de las acciones es posible apreciar una fluctuación amplificada del comportamiento real de cada acción. Pero al añadir bandas de confianza de \$5 es posible encajonar la fluctuación.

Existen muchos aspectos a mejorar en relación al pronóstico de acciones, más aún en el caso chileno, donde quedó demostrado el nivel de caos de las acciones en la Bolsa de Comercio de Santiago. Principalmente aumentar el horizonte de pronósticos a más de 5 días, ayuda a realizar una mejor recomendación a largo plazo.

Finalmente, al analizar los objetivos propuestos, es posible ver que se han cumplido a cabalidad en este proyecto. Donde capturar la tendencia de las acciones siempre es una tarea titánica, más aún pronosticar el precio exacto o la rentabilidad exacta. Pero el pronóstico de la tendencia de una acción mediante la dirección de la curva es factible para el Pronosticador Neuronal propuesto.

10. Referencias Utilizadas

- [1] E. Hadavandi, H. Shavandi, and A. Ghanbari, "Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting," *Knowledge-Based Systems*, vol. 23, no. 8, pp. 800–808, 2010.
- [2] R. A. Meese and A. K. Rose, "An Empirical Assessment of Non-Linearities in Models of Exchange Rate Determination," *The Review of Economic Studies*, vol. 58, no. 3, pp. 603–619, 1991.
- [3] O. C. Ibe, "8 - Random Walk," in *Markov Processes for Stochastic Modeling (Second Edition)*, Second Edition., Oxford: Elsevier, 2013, pp. 205–261.
- [4] D. Zhang and X. Zhang, "Study on Forecasting the Stock Market Trend Based on Stochastic Analysis Method," *International Journal of Business and Management*, vol. 4, no. 6, pp. 163–170, 2009.
- [5] V. L. Paucar and M. J. Rider, "Artificial neural networks for solving the power flow problem in electric power systems," *Electric Power Systems Research*, vol. 62, no. 2, pp. 139–144, 2002.
- [6] A. Timmerman, "Neural networks in finance and investing. Using artificial intelligence to improve realworld performance: Robert R. Trippi and Efraim Turban (eds.) 1996, Irwin Professional Publishing Co., Burr Ridge IL, 1996, 70.00 821 pp., {ISBN} 1-55738-919-5," *International Journal of Forecasting*, vol. 13, no. 1, pp. 144–146, 1997.
- [7] Y. Li and W. Ma, "Applications of Artificial Neural Networks in Financial Economics: A Survey," in *Computational Intelligence and Design (ISCID), 2010 International Symposium on*, 2010, vol. 1, pp. 211–214.
- [8] H.-N. Hao, "Short-term forecasting of stock price based on genetic-neural network," in *Natural Computation (ICNC), 2010 Sixth International Conference on*, 2010, vol. 4, pp. 1838–1841.
- [9] E. Antonino; Rebolledo Julio; Cornejo Parisi F., "Modelos de Lógica y Lógica Borrosa en la Predicción del IPSA," *Revista de Estudios de Administración*, vol. 13, no. 1, pp. 37–66, 2006.
- [10] G. S. Atsalakis and K. P. Valavanis, "Surveying stock market forecasting techniques – Part II: Soft computing methods," *Expert Systems with Applications*, vol. 36, no. 3, Part 2, pp. 5932–5941, 2009.
- [11] Q. Cao, K. B. Leggio, and M. J. Schniederjans, "A comparison between Fama and French's model and artificial neural networks in predicting the Chinese stock market," *Computers & Operations Research*, vol. 32, no. 10, pp. 2499–2512, 2005.
- [12] P.-C. Chang, C.-H. Liu, J.-L. Lin, C.-Y. Fan, and C. S. P. Ng, "A neural network with a case based dynamic window for stock trading prediction," *Expert Systems with Applications*, vol. 36, no. 3, Part 2, pp. 6889–6898, 2009.
- [13] T. Chavarnakul and D. Enke, "Intelligent technical analysis based equivolume charting for stock trading using neural networks," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1004–1017, 2008.

- [14] X. Zhu, H. Wang, L. Xu, and H. Li, "Predicting stock index increments by neural networks: The role of trading volume under different horizons," *Expert Systems with Applications*, vol. 34, no. 4, pp. 3043–3054, 2008.
- [15] E. F. Fama, "EFFICIENT CAPITAL MARKETS: A REVIEW OF THEORY AND EMPIRICAL WORK*," *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.
- [16] R. Moreno-Díaz and A. Moreno-Díaz, "On the legacy of W.S. McCulloch," *Biosystems*, vol. 88, no. 3, pp. 185–190, 2007.
- [17] S. K. Chalup and A. D. Blair, "Incremental training of first order recurrent neural networks to predict a context-sensitive language," *Neural Networks*, vol. 16, no. 7, pp. 955–972, 2003.
- [18] G. A. Anastassiou, "Multivariate sigmoidal neural network approximation," *Neural Networks*, vol. 24, no. 4, pp. 378–386, 2011.
- [19] F. Fernández-Navarro, C. Hervás-Martínez, P. A. Gutiérrez, J. M. Peña-Barragán, and F. López-Granados, "Parameter estimation of q-Gaussian Radial Basis Functions Neural Networks with a Hybrid Algorithm for binary classification," *Neurocomputing*, vol. 75, no. 1, pp. 123–134, 2012.
- [20] M. Marcellino, J. H. Stock, and M. W. Watson, "A comparison of direct and iterated multistep {AR} methods for forecasting macroeconomic time series," *Journal of Econometrics*, vol. 135, no. 1–2, pp. 499–526, 2006.
- [21] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 1st ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1994.
- [22] S. Nandy, P. P. Sarkar, and A. Das, "Analysis of a Nature Inspired Firefly Algorithm based Back-propagation Neural Network Training," *CoRR*, vol. abs/1206.5360, 2012.
- [23] P. Civicioglu, "Backtracking Search Optimization Algorithm for numerical optimization problems," *Applied Mathematics and Computation*, vol. 219, no. 15, pp. 8121–8144, 2013.
- [24] S. Asadi, E. Hadavandi, F. Mehmanpazir, and M. M. Nakhostin, "Hybridization of evolutionary Levenberg–Marquardt neural networks and data pre-processing for stock market prediction," *Knowledge-Based Systems*, vol. 35, no. 0, pp. 245–258, 2012.
- [25] G. de Planificación y Desarrollo - Bolsa Comercio de Santiago, "Índices Bursátiles: Metodología de selección y cálculo," p. 10, 2011.