

PRONOSTICO DEL INDICE GENERAL DE LA BOLSA DE VALORES DE COLOMBIA USANDO REDES NEURONALES

PREDICT OF THE GENERAL INDEX OF THE BAG OF VALUES DE COLOMBIA USING NETS NEURONALES

RESUMEN

Una de las formas de medir el comportamiento de la economía de un país, es a través de las bolsas de valores locales. En Colombia el comportamiento económico de la Bolsa de valores se mide a través de tres indicadores principalmente: el Colcap, el Col20 y el Igbc, este último es el más antiguo y más utilizado por los agentes que intervienen en el mercado bursátil. En el presente documento se expone como pronosticar el comportamiento del índice bursátil Igbc a través de la metodología de las redes neuronales, brindando al inversionista la posibilidad de proyectar el comportamiento del mercado bursátil colombiano de forma eficiente.

PALABRAS CLAVES: redes neuronales, índice bursátil, rentabilidad, topología, tendencia, mercado accionario.

ABSTRACT

One in the ways of measuring the behavior of the economy of a country, is through the local stock exchanges. In Colombia the economic behavior of the Stock exchange is measured mainly through three indicators: the Colcap, the Col20 and the Igbc, this last one is the oldest and more used by the agents that intervene in the market market. Presently document is exposed as predicting the behavior of the market index Igbc through the methodology of the nets neuronales, offering the investor the possibility to project the behavior of the market market Colombian in an efficient way.

KEYWORDS: *nets neuronales, market index, profitability, topology, tendency, stock market.*

1. INTRODUCCIÓN

Los inversionistas de mercados financieros se enfrentan a los diferentes movimientos de los precios de los activos financieros ocasionados por agentes externos como los precios del petróleo y la tasa de cambio (TRM) del peso frente al dólar, y a situaciones incontrolables de la economía de Colombia como son el riesgo país riesgo soberano, la política fiscal y monetaria, generando caos e incertidumbre a los inversionistas nacionales y extranjeros. Se han desarrollado diferentes trabajos investigativos que orienten al inversor y le faciliten la toma de decisiones en mercados altamente especulativos como son los de las economías emergentes. Entre los trabajos que se apoya la presente investigación, vale la pena resaltar los trabajos de Markowitz (1952), con el modelo de varianza/ covarianza para determinar la frontera eficiente, los modelos teóricos presentados por Black (1972), Merton (1973) y más tarde por Levy y

Sarnat (1982), Elton (1995) y Gruber (1997), el resumen de elaborado por Alexander, Sharp y Balley (2003), conduce a que el portafolio óptimo se puede encontrar utilizando principalmente la línea de mercado de capitales para seleccionar en la frontera eficiente la conformación de portafolios de títulos financieros que generen la prima por riesgo máxima.

La mayoría de los operadores en los mercados bursátiles buscan a través del tiempo encontrar algún método o criterio que les facilite predecir de la manera más cercana posible la evolución futura de las acciones, los bonos, los contratos derivativos y el mercado de divisas, en la medida que ese punto se logre se obtendrían altos márgenes de ganancia y disminuir el riesgo y las pérdidas.

En el presente trabajo, se presenta una metodología que facilita la predicción del índice Igbc a través de las redes neuronales

EDUARDO ARTURO CRUZ T

Ingeniero Industrial, M. Sc.
Profesor Asistente
Universidad Tecnológica de
Pereira
ecruz@utp.edu.co

JORGE HERNAN RESTREPO

Ingeniero Industrial, M. Sc.
Profesor Asistente
Universidad Tecnológica de
Pereira
jhrestrepoco@utp.edu.co

PEDRO MEDINA VARELA

Ingeniero Mecánico, M. Sc.
Profesor Auxiliar
Universidad Tecnológica de
Pereira
pemedin@utp.edu.co

**Grupo de Investigación
Administración Económica y
Financiera.**

2. LAS REDES NEURONALES

El origen de esta tecnología puede datarse en las primeras décadas del siglo XX. El primer modelo de este tipo fue un dispositivo hidráulico diseñado por Russell[1] en 1913, aunque fue en la década de los cuarenta cuando se crearon modelos que tuvieron gran repercusión y que son considerados hoy en día como los precursores de las actuales redes neuronales artificiales (ANN).

En 1943, Warren McCulloch y Walter Pitts [2] construyeron el primer modelo matemático de ANN. Aunque este modelo causó gran impacto inicial, adolecía de un factor clave, la capacidad de aprendizaje. Posteriormente, en 1949, Donald Hebb a partir de sus estudios sobre la neurona y el aprendizaje, desarrolló un algoritmo de aprendizaje al que se le dio su nombre: Aprendizaje Hebbiano

Las redes neuronales se componen de muchas neuronas interconectadas de cierta manera (ver figura 1), formando diferentes arquitecturas (denominadas topología de la red), las cuales procesan información sobre la base de un proceso bien definido. Se pueden distinguir redes de una sola capa (single layer) y multicapa (multilayer). Normalmente, la capa donde la información es suministrada a la red es la capa de entrada (o input layer), la capa donde la salida es obtenida es la capa de salida (o output layer) y las capas entre la capa de entrada y la de salida son las capas ocultas (o hidden layers). Puede haber una o más capas ocultas, las cuales se llaman así porque su salida no es directamente observable [3].

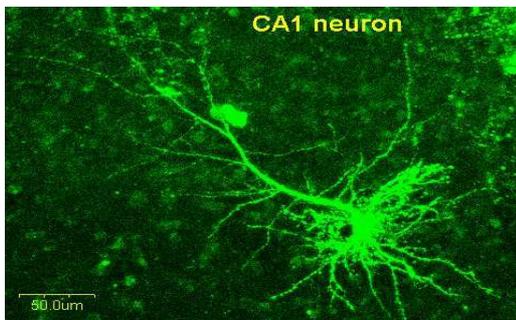


Figura 1. Imagen de una Neurona

Las redes neuronales artificiales (RNA) presentan las siguientes características. Ver figura 2.

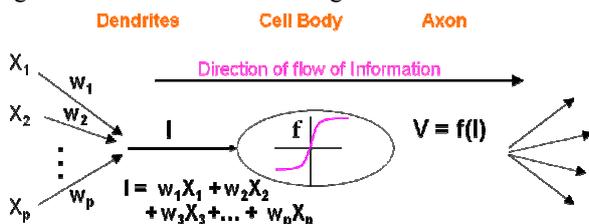


Figura 2. Estructura de una red neuronal artificial.

- Recibe diferentes entradas $X_1, X_2 \dots X_3$ de otras neuronas o ambientes.
- Las entradas alimentan la red a través de conexiones y pesos.
- El total de entradas es igual a la suma de los pesos de todas las entradas y fuentes.
- La función de transferencia convierte las entradas en salidas
- Las salidas van a otras neuronas o ambientes.

La forma en que se enlazan las diferentes capas que conforman una neurona se muestra a continuación. Ver figura 3.

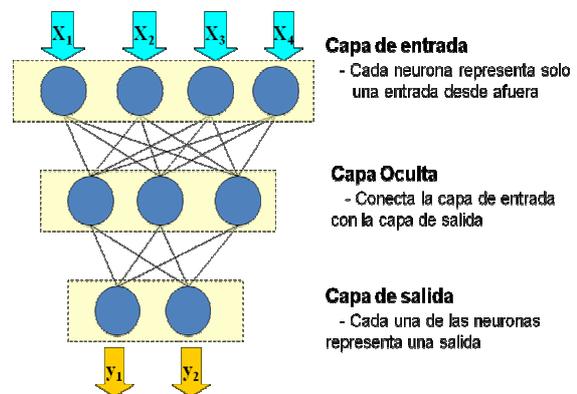


Figura 3. Capas que conforman una neurona.

2.1 RNA COMO MODELO NO LINEAL

Una Red de neuronas artificiales (RNA), sin capas ocultas es similar a un modelo lineal generalizado. Mientras que las RNA con alimentación hacia adelante son un subconjunto de una clase más grande de modelos de regresión y discriminación no lineales. Los modelos económicos lineales no son capaces de capturar patrones no lineales y tendencias en las relaciones de la mayoría de las variables de predicción econométricas usadas en economía y finanzas.

2.2 MODELO MATEMATICO DE LAS REDES NEURONALES

Los elementos que constituyen a la red neuronal con una capa oculta para este caso son los siguientes:

- s es una función de valores reales, conocida como la función de transferencia.
- $\bar{o}^{(0)}$ es la capa de entrada
 $o^{(0)} = \mathbf{x} = (x_1, \dots, x_n)^T$
 $\bar{o}^{(1)}$ es la capa oculta, el vector extendido de
 $o^{(1)} = (o_1^{(1)}, \dots, o_k^{(1)})^T$

- $\mathbf{o}^{(2)} = (o_1, \dots, o_m)^T$ es la capa de salida, considerado como el vector que aproxima al valor deseado $\mathbf{t} = (t_1, \dots, t_m)^T$.
- $\mathbf{W}^{(1)}$ es una matriz de tamaño $(n + 1) \times k$ cuyos valores $W_{ij}^{(1)}$ son los pesos de la conexión entre las unidades $\bar{o}_i^{(0)}$ y $o_j^{(1)}$.
- $\mathbf{W}^{(2)}$ es una matriz de tamaño $(k + 1) \times m$ cuyos valores $W_{ij}^{(2)}$ son los pesos de la conexión entre las unidades $\bar{o}_i^{(1)}$ y $o_j^{(2)}$.

De estos elementos, únicamente las matrices $\mathbf{W}^{(l)}$ son consideradas como los parámetros de la red, ya que los valores $\bar{o}^{(l)}$ son el resultado de cálculos que dependen de las matrices de pesos, del valor de entrada $\bar{\mathbf{x}}$ y de la función de transferencia s .

La función de transferencia s considerada es conocida como función sigmoideal, y esta definida como

$$s(u) = \frac{1}{1 + \exp(-u)}$$

Esta función además de ser diferenciable, tiene la particularidad de que su derivada se puede expresar en términos de sí misma:

$$\frac{ds(u)}{du} = s(u)(1 - s(u)).$$

3. PROCESO METODOLOGICO APLICADO EN LA INVESTIGACION

Se seleccionan las acciones con alta bursatilidad durante un determinado periodo de tiempo y que integran la conformación de la canasta de acciones más negociadas del mercado.

Determinar el rendimiento de un activo para un solo período es: $R_i = \frac{P_t - P_{t-1}}{P_{t-1}}$ donde, R_i es la rentabilidad

del activo i , P_t es el precio de mercado del activo en el período t y P_{t-1} es el precio en el período inmediatamente anterior.

Seleccionar cual es la topología de la red neuronal (se utiliza el software Neurosolutions), pertinente al problema de la variación de los rendimientos de las acciones.

Generar los precios estimados con la red neuronal y compararlos con los datos originales.

Pronosticar las variaciones de rentabilidad o pérdida de las acciones seleccionadas.

Pronosticar la variación del índice Igbc con las redes neuronales.

Comparar los resultados obtenidos con los datos reales del Igbc.

4. DESARROLLO DE LA METODOLOGIA PROPUESTA

Se toman los precios de las acciones de Banco de Bogotá, Bic, Grupo aval, Compañía de chocolates, Cementos Argos como datos de entrada y el índice IBC como dato de salida ver figura 4. [6]

	A	B	C	D	E	F	G
1	fc_oper	bog	bic	aval	cho	arg	ibc
2	02/01/2007	8950.16	17120.00	23203.65	16094.40	8950.16	10979.87
3	03/01/2007	8976.14	17500.00	23439.94	16151.97	8976.14	11041.81
4	04/01/2007	8954.35	17634.83	23683.71	16221.43	8954.35	11015.92
5	05/01/2007	8800.47	16864.19	23142.85	15934.35	8800.47	10722.11
6	09/01/2007	8374.94	16639.30	21916.72	15335.88	8374.94	10203.40
7	10/01/2007	8150.70	16370.73	21552.44	15204.82	8150.70	10560.05
8	11/01/2007	8602.25	16908.29	22970.86	15852.57	8602.25	10739.54
9	12/01/2007	8693.26	16789.66	23154.22	15786.67	8693.26	10783.22
10	15/01/2007	8850.90	16860.00	23409.42	15850.85	8850.90	10884.07
11	16/01/2007	8738.00	16942.29	23103.22	15779.12	8738.00	10823.39
12	17/01/2007	8717.33	16774.99	22995.54	15685.42	8717.33	10853.44

Figura 4. Cotizaciones de las acciones y el índice IBC.

$$R_i = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Se calculan las variaciones ver figura 5.

	A	B	C	D	E	F	G
1	fc_oper	bog	bic	aval	cho	arg	ibc
2	03/01/2007	0.2902%	2.2196%	1.0183%	0.3577%	0.2902%	0,5642%
3	04/01/2007	-0.2428%	0.7705%	1.0400%	0.4300%	-0.2428%	-0,2345%
4	05/01/2007	-1.7184%	-4.3700%	-2.2837%	-1.7697%	-1.7184%	-2,6671%
5	09/01/2007	-4.8353%	-1.3336%	-5.2981%	-3.7558%	-4.8353%	-4,8378%
6	10/01/2007	-2.6776%	-1.6141%	-1.6621%	-0.8546%	-2.6776%	3,4954%
7	11/01/2007	5.5400%	3.2837%	6.5812%	4.2601%	5.5400%	1,6997%
8	12/01/2007	1.0580%	-0.7016%	0.7983%	-0.4156%	1.0580%	0,4067%
9	15/01/2007	1.8133%	0.4190%	1.1022%	0.4065%	1.8133%	0,9352%
10	16/01/2007	-1.2756%	0.4881%	-1.3080%	-0.4525%	-1.2756%	-0,5576%
11	17/01/2007	-0.2365%	-0.9874%	-0.4661%	-0.5938%	-0.2365%	0,2777%
12	18/01/2007	1.7194%	0.8164%	0.8347%	0.0123%	1.7194%	-0,1011%

Figura 5 Variaciones de los precios.

Se determinan los rangos de entrada en el software neurosolutions [4], en este caso es una red de cinco entradas de datos conformada por las variaciones de los precios de las acciones; ver figura 6.

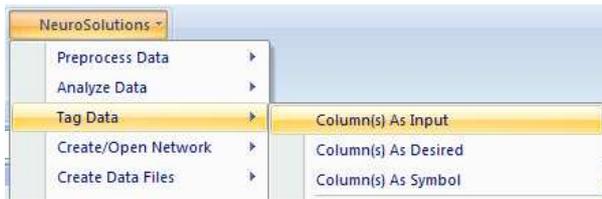


Figura 6. Definición de datos de entrada

Se establecen los datos de salida en este caso es el rango de datos del índice de la bolsa de valores Igb, es una red neuronal con una salida; ver figura 7.

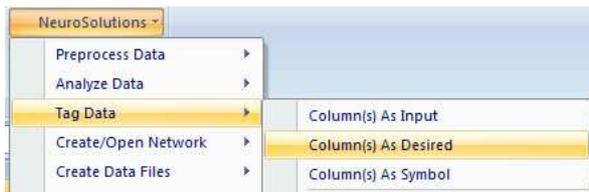


Figura 7. Definición de datos de salida

Para concluir la definición de datos de entrada se incluye el rango de datos para el entrenamiento, el rango de datos para evaluación, el rango de datos para validación y el rango de datos para pronóstico-producción. Ver figura 8.

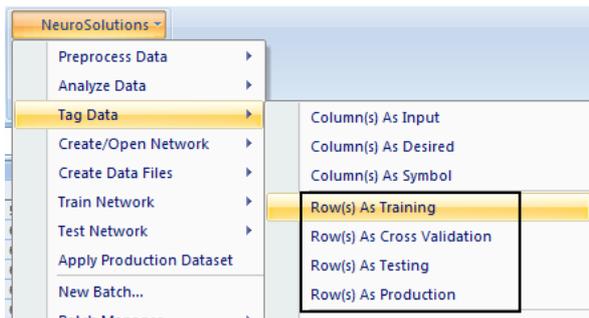


Figura 8. Definición de los rangos de datos para la red

Con la incorporación del rango de datos de validación se establece un porcentaje de dichos datos para su validación; ver figura 9.

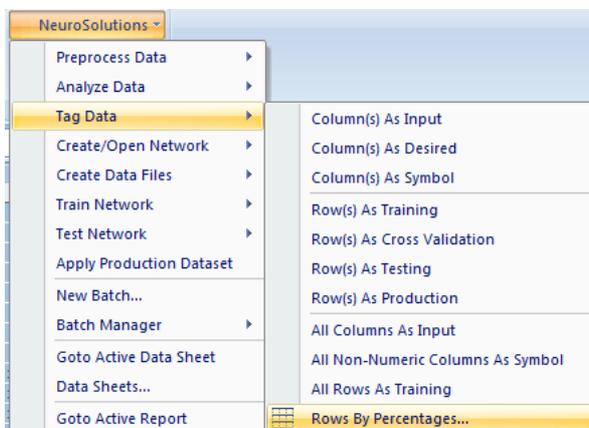


Figura 9. Seleccionar el porcentaje de datos a validar

Se selecciona el porcentaje de datos a tomar del rango de entrenamiento de la red, el porcentaje de validación y el porcentaje de prueba. Ver figura 10.

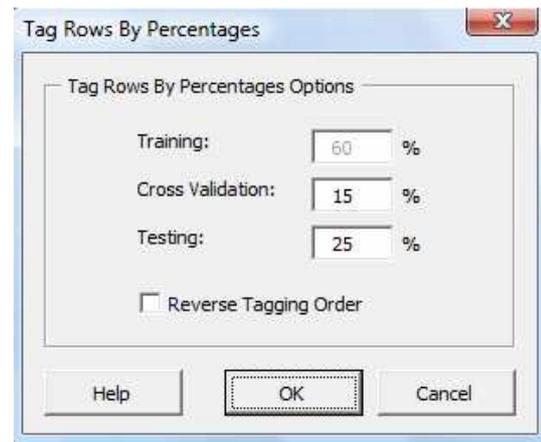


Figura 10. Prueba y validación de los datos.

Con la definición de los rangos de datos de entrada, salida, entrenamiento, prueba, validación y producción, se determina la mejor estructura de la red neuronal. Ver figura 11.

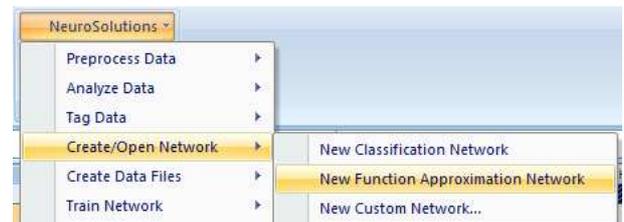


Figura 11. Creación de la red neuronal

La estructura de la red generada corresponde a la familia de las redes de backpropagation, Ver figura 12.



Figura 12. Red neuronal artificial determinada.

Se pasa a la etapa de entrenar la red y se parametriza. Ver figura 13.

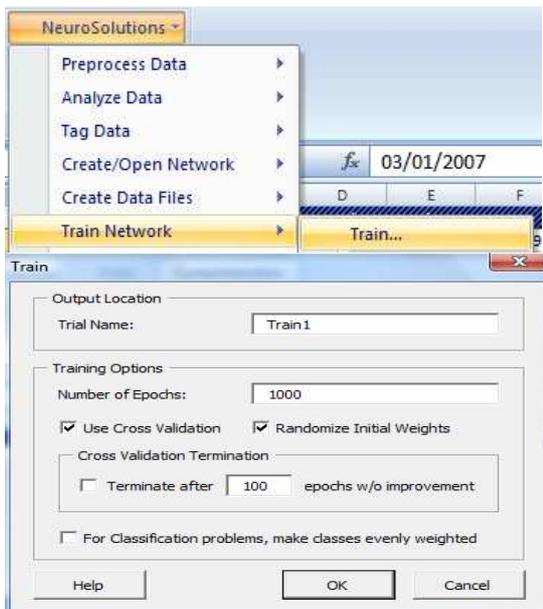


Figura 13. Parámetros para entrenar la RNA

Se realiza el entrenamiento de la red con mil periodos y su validación, calculándose el margen de error. Ver figura 14.

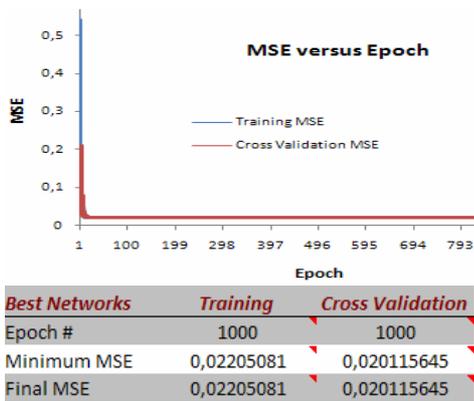
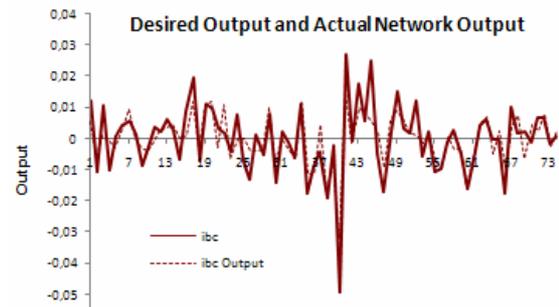


Figura 14. Validación del entrenamiento de RNA

El proceso de test o prueba de las RNA tiene dos aplicaciones uno es el de clasificación de la información y la otra aplicación es el pronóstico, en este caso se realiza un proceso de pronóstico o regresión, obteniéndose un coeficiente de correlación de 85.6%. Ver figura 15.



Performance	ibc
MSE	3,54373E-05
NMSE	0,290759666
MAE	0,004581679
Min Abs Error	5,19093E-05
Max Abs Error	0,019370325
r	0,856465296

Figura 15. Prueba de los datos de salida (IBC).

Para la realización del pronóstico o proyección del IBC, se debe proyectar individualmente las variaciones de precios de cada acción, dado que los datos no presentan un comportamiento tendencial o de crecimiento se realiza un análisis de datos no estacionales con la herramienta Crystal Ball [5]. A continuación se presenta como ejemplo el análisis para la acción del Banco de Bogotá incluyendo la línea de tendencia, el mismo ejercicio se realiza para las demás acciones. Ver figura 16.



Figura 16. Variación de los precios de Banco de Bogotá

La proyección de las variaciones de los precios de cada acción se aprecia en la figura 17.

	A	B	C	D	E	F	G
1	fc_oper	bog	bic	aval	cho	arg	ibc
179	25/09/2007	0,005689	0,00429	0,002794	0,011826	0,005689	0,010574
180	26/09/2007	0,013897	-0,002638	0,010233	0,010801	0,013897	0,001919
181	27/09/2007	-0,019319	0,020376	-0,009306	0,002681	-0,019319	0,002579
182	28/09/2007	0,009711	0,027446	-0,003001	0,003754	0,009711	-0,001187
183	01/10/2007	-0,004548	0,005141	0,010286	0,013643	-0,004548	0,006868
184	02/10/2007	0,009179	0,007559	0,020601	0,00274	0,009179	0,0069
185	03/10/2007	-0,003405	-0,011513	0,001668	-0,003925	-0,003405	-0,001578
186	04/10/2007	0,005307	0,015384	-0,005169	0,011095	0,005307	0,001098
187	05/10/2007	0,003694	0,00381	0,0057	0,003303	0,003694	
188	06/10/2007	0,001866	0,00256	0,000733	0,003491	0,001866	
189	07/10/2007	0,003622	0,007252	0,000421	0,005963	0,003622	
190	08/10/2007	0,003061	0,004541	0,002285	0,004252	0,003061	
191	09/10/2007	0,002849	0,004784	0,001147	0,004569	0,002849	

Figura 17. Variaciones proyectadas de las acciones

Con base en los datos obtenidos de las variaciones de los precios de las acciones, se procede a determinar las variaciones del índice IBC para los cinco días siguientes, con el software neurosolutions®.

El siguiente paso del proceso es generar el pronóstico para los siguientes cinco días del Ibc, con base en las proyecciones de las variaciones de las acciones. Ver figura 18.

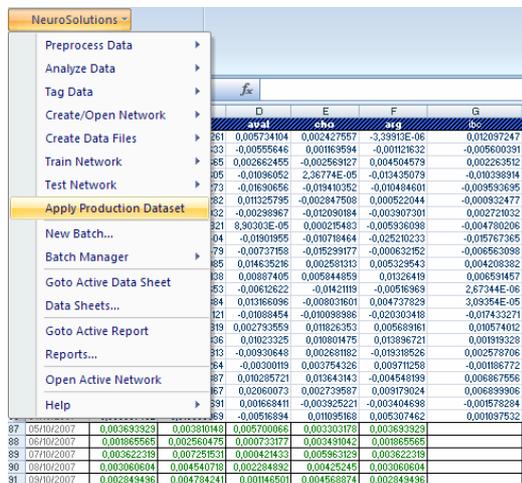


Figura 18. Carga de los datos para el pronóstico

Los datos proyectados del igb se presentan a continuación, también se incluyen los datos reales del índice para comparar los resultados obtenidos. Ver figura 19.

	A	G	H
1	fc_oper	igb	igb real
187	05/10/2007	0,00446351	0,0002540715
188	06/10/2007	0,001829147	0,0007621951
189	07/10/2007	0,001919098	0,0002540457
190	08/10/2007	0,002747745	0,0002540392
191	09/10/2007	0,002167958	0,0002540328

Figura 19. Datos comparativos del Igb estimado y real

Se complementa el análisis de los datos con un gráfico lineal que manifiesta la volatilidad de los datos, propio en etapas de turbulencia principalmente en economías emergentes, a pesar de ello las redes neuronales presentan un método de precisión bastante aceptable a la información real del la variación del Ibc. Ver figura 20.

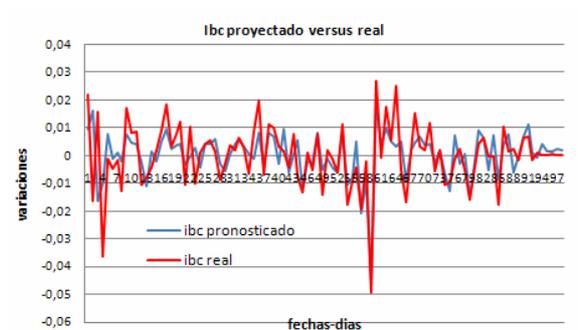


Figura 20. Gráfico comparativo del Ibc real y el Ibc proyectado.

5. CONCLUSIONES Y RECOMENDACIONES

Las redes neuronales artificiales tienen una aplicación práctica en el mercado bursátil sobre los modelos estadísticos tradicionales porque no dependen de supuestos teóricos sobre los que se basan las técnicas estadísticas (normalidad, homoscedasticidad, independencia, etc).

Los resultados y las pruebas de bondad y ajuste representados por R², indican que el modelo basado en RNA es mejor que los modelos tradicionales de tendencia o crecimiento y en muchos casos a los modelos Arima y Garch.

El modelo propuesto se puede utilizar como parte integral de la operación financiera ya que se cuenta con herramientas computacionales especializadas para RNA y su aplicación en la toma de decisiones financieras en el manejo de la tesorería y gestión de riesgos financieros y de mercado.

El cálculo del Igb depende de las acciones que conforman el indicador, y cada acción participa de acuerdo al volumen y precios transados en cada rueda de negociación. Para mejorar los resultados se recomienda involucrar todas las acciones que participan en la canasta del índice.

Es recomendable para el pronóstico del Igb proyectado, métodos de pronóstico más específicos para cada una de las acciones que componen el índice y el vector de entrada de la red neuronal.

6. BIBLIOGRAFÍA

- [1] RUSSELL, S., A Practical Device to Simulate the Working of Nervous Discharges, en “Journal of nimal Behaviour”, Vol. 3, (1913), citado por OLMEDA, Ignacio, op. cit., pág. 185.
- [2] McCULLOCH, W.S. y PITTS, W., *A Logical Calculus of the Ideas Immanent in Nervous Activity*, en “Bulletin of Mathematical Biophysics”, Vol. 5, citado por OLMEDA, Ignacio, op. cit., pág. 185.
- [3] KOHONEN, Teuvo, *Self-Organization and Associative Memory*, third edition (Berlin, Springer-Verlag, 1989).
- [4] Neurosolutions® Software
- [5] Crystal Ball® Software
- [6] Sitio Web superintendencia financiera de Colombia: <http://superfinanciera.gov.co>