

Métodos para predecir índices bursátiles

Methods for Predicting Stock Indexes

*Martha Cecilia García**

*Aura María Jalal***

*Luis Alfonso Garzón****

*Jorge Mario López*****

Recibido: 17/05/13

Aprobado: 29/10/13

* Ingeniería Industrial (e), Universidad de Córdoba, Colombia. [mgapupo@gmail.com].

** Ingeniería Industrial (e), Universidad de Córdoba, Colombia.
[auramariajalal@gmail.com].

*** Magíster en Ingeniería. Profesor de la Universidad de Córdoba, Colombia.
[lgarzon2003@yahoo.es].

**** Magíster en Ingeniería Industrial. Profesor de la Universidad de Córdoba, Colombia.
[jotalopez@correo.unicordoba.edu.co].

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA

AURA MARÍA JALAL

LUIS ALFONSO GARZÓN

JORGE MARIO LÓPEZ

Resumen

Este artículo presenta una revisión bibliográfica acerca de los métodos que se han utilizado en las últimas dos décadas para predecir Índices Bursátiles. Los métodos estudiados van desde aquellos que logran capturar las características lineales presentes en los índices de bolsa, pasando por los que se enfocan en las características no lineales y finalmente métodos híbridos que son más robustos, pues capturan características lineales y no lineales. Además, se incluyen aquellos métodos que utilizan variables macroeconómicas para predecir los índices de diferentes Bolsas de Valores en el mundo.

Palabras clave

Bolsa de Valores, índice, pronósticos.

Abstract

This paper presents a literature review on methods that have been used in the last two decades to predict Stock Market Indexes. Methods studied range from those enabling to grab the linear characteristics present in the stock market indexes, going through those that focus on non-linear features and finally hybrid methods that are more robust, since they capture linear and non-linear features. In addition, this research includes methods that use macroeconomic variables to predict indexes from different stock exchanges around the world.

Key Words

Stock Exchange, Index, Forecasts.

Clasificación JEL: E17; E19; E27

1. Introducción

Desde hace tiempo los economistas han estudiado e intentado comprender los movimientos de los precios en la bolsa de valores, debido a que las inversiones en bolsa están sujetas a riesgos, los rendimientos son variables y su existencia es incierta. La predicción de la bolsa de valores es un tema de interés, en particular para quienes invierten en ella. Sería muy provechoso poder predecir la tendencia y, si fuera posible, el precio de las acciones, ya que con esta información los inversionistas podrían realizar movimientos apropiados y así ganar dinero.

Por lo tanto, predecir un índice de la bolsa de valores representa un gran reto y en las últimas dos décadas ello ha sido objeto de muchos estudios dadas las aplicaciones comerciales que tiene. Numerosos métodos han sido propuestos para brindar predicciones más precisas a los inversores. Algunos de estos estudios han utilizado modelos autorregresivos, promedios móviles, Arima (Reddy, 2010), regresión múltiple (Chang, Yeung y Yip, 2000), algoritmo genético (Kim y Han, 2000), redes neuronales artificiales (Chena, Leung y Daouk, 2003), suavizado exponencial, métodos lineales y no lineales (Zemke, 1998), entre otras. En este artículo se hace una revisión bibliográfica acerca de los principales métodos que se han utilizado para predecir índices bursátiles.

2. Métodos comunes para predecir índices bursátiles

Los modelos lineales, por la fácil interpretación de sus elementos, tienen una considerable ventaja sobre otros, lo que ha hecho que sean utilizados en un sinnúmero de aplicaciones; una de ellas ha sido la predicción de series de tiempo financieras. Según Alonso y García (2009), a principio de la década de 1980 los modelos empleados para modelar la media de un activo eran los Arima o Arimax, pero entrada la década de los noventa con los modelos ARCH y Garch se le da más importancia a la volatilidad que a la media de los datos.

En 1999 Franses y Ghijssels confirmaron que los modelos Garch sirven para predecir la volatilidad del mercado de valores, pero es frecuente que los residuos estimados de estos modelos tengan exceso de curtosis y dichos modelos al parecer no capturan todas las características de los datos.

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA

AURA MARÍA JALAL

LUIS ALFONSO GARZÓN

JORGE MARIO LÓPEZ

Chang, Yeung y Yip (2000) indicaron cómo cada mes muchos inversores esperan conocer los resultados de indicadores económicos como la tasa de empleo o el índice de precios del consumidor, ya que cada uno es una medida de alguna faceta de la economía, entonces relacionaron cada indicador mediante un gráfico con el índice de bolsa para entender cuán bien se relacionaban las variables con la bolsa de valores. Los gráficos revelaron que los indicadores que parecían tener una relación consistente eran el índice de precios del consumidor, los ingresos personales, índice de diez indicadores líderes, manufactura, venta y comercio.

En 2003 los estudios al respecto continúan: Lasfer, Melnik & Thomas documentan el comportamiento del precio de las acciones a corto plazo después de un período de estrés del mercado de valores, concentrándose en el comportamiento de los precios diarios del mercado por medio de los índices de 39 bolsas de valores. La investigación demostró que los desempeños anormales después del estrés son significativamente mayores para los mercados emergentes y el tamaño de los saltos luego del periodo de estrés está relacionado con liquidez del mercado.

Pai y Lin (2005) plantean que el modelo autorregresivo integrado de media móvil (Arima), que fue introducido por Box y Jenkins, ha sido uno de los enfoques más populares utilizado en la predicción. En un modelo Arima se establece el supuesto de que el valor futuro de la variable es una combinación lineal de valores y errores pasados; sin embargo, cualquier grupo de características no lineales lo limita.

En 2005 Bhardwaj y Swanson sugirieron un nuevo modelo denominado Arfima, que hace estimaciones usando una variedad de procedimientos estándares que proporcionan significativamente mejores predicciones que AR, MA, ARMA, Garch y modelos relacionados, con base en el análisis de la media de los errores cuadráticos del pronóstico (MSFE) y en el uso de pruebas de precisión de predicción.

Liu y Hung (2010) realizaron un estudio en el cual, usando la volatilidad diaria del índice S&P-100 de Estados Unidos, compararon los modelos Garch-N, Garch-t, Garch-HT y Garch-SGT con modelos del tipo asimétrico como GJR-Garch y Egarch. Los resultados obtenidos muestran que estos últimos mejoran la volatilidad del pronóstico y demuestran que el componente asimétrico es más importante que la especificación de la distribución cuando hay presencia de asimetría, leptocurtosis y efectos Leverage.

3. Métodos no lineales de predicción del índice de la bolsa

La era actual de la exploración de datos está caracterizada por el creciente uso de modelos no lineales de predicción. La teoría económica destaca diversas fuentes potenciales para la presencia de no linealidades y ciclos en los precios de los activos financieros.

Komo, Chang y KO (1994) utilizan modelos estadísticos no paramétricos y no lineales debido a que muchas relaciones importantes en el área de las finanzas tienen este tipo de relaciones. Las redes neuronales artificiales poseen la propiedad de capturar las características no lineales de los índices de bolsa y han demostrado que pueden ser entrenadas con una cantidad suficiente de información para identificar dichas relaciones no lineales entre los valores de entrada y salida.

La aplicación de las redes neuronales a la predicción de series financieras está creciendo en los últimos años, ya que parece ser un método eficaz y presenta multitud de oportunidades. Zemke (1999) plantea que los mercados con menor volumen de operaciones son más fáciles de predecir, utiliza como referente la Bolsa de Valores de Varsovia y plantea las técnicas Machine Learning (ML) para analizarla. La tarea es predecir el valor del índice a través de decisiones binarias, es decir, pronosticar si el valor del índice WIG en la semana de operaciones se encuentra arriba o abajo del valor actual. Las cuatro técnicas ML utilizadas son: predicción con redes neuronales, clasificador bayesiano, *K-nearest neighbor* y *K-nearest neighbor prediction scrutinized*. Con estas técnicas se concluyó que mediante los métodos de *K-nearest neighbor*, con 64% de efectividad, y redes neuronales, se tienen mejores predicciones debido a que la bolsa de valores está dominada por la no linealidad de los datos.

En el mismo año, Lee y Jo desarrollaron un intérprete gráfico para predecir el mercado, al que llamaron *Candlestick Chart Analysis Expert System*. Este sistema tiene patrones y normas que pueden predecir futuros movimientos de precios de acciones. Los patrones definidos se clasifican en cinco grupos con respecto a su significado: bajando, subiendo, neutral, continuando la tendencia y patrones donde se invierte la tendencia. Los resultados experimentales obtenidos por Lee y Jo revelaron que el modelo tenía un porcentaje de aciertos promedio de 72%, lo cual ayuda a los inversores a obtener mayores beneficios de su inversión en acciones.

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA

AURA MARÍA JALAL

LUIS ALFONSO GARZÓN

JORGE MARIO LÓPEZ

Chen, Leung y Daouk (2003) modelan y predicen el cambio en el índice de la Bolsa de Valores de Taiwán a través de otro modelo de red neuronal artificial denominado red neuronal probabilística (PNN); este método utiliza datos históricos del índice de la Bolsa de Valores de Taiwán, y los resultados muestran que las estrategias de inversión basadas en él obtienen mayores rendimientos que otras estrategias como los métodos generalizados de momentos (GMM). La superioridad de este modelo se debe a que es capaz de identificar valores atípicos y datos erróneos.

Toro, Molina y Garcés (2006) presentaron un estudio comparativo entre la predicción de precios en bolsa de valores utilizando redes neuronales y neurodifusas. Se muestra una metodología aplicable a la predicción del comportamiento de cualquier tipo de acción, basada en el cálculo preliminar de la correlación entre el precio y otras variables de mercado. El pronóstico del valor de las acciones en la bolsa de valores incide en los procesos de toma de decisiones de las empresas u otros agentes del mercado afectados por su comportamiento. Los resultados obtenidos fueron satisfactorios en cuanto al porcentaje de error del valor estimado frente al valor real. Las técnicas inteligentes se muestran como una herramienta interesante que está mostrando resultados promisorios y abre un abanico de posibilidades por explorar en el campo del pronóstico.

Chaigusin, Chirathamjaree y Clayden (2008) muestran un caso particular en el cual la predicción del índice de la Bolsa de Tailandia se hace a través de redes neuronales del tipo *backpropagation*, y los datos de entrada usados corresponden al índice SET, Dow Jones, Nikkei, Hang Seng, el precio del oro, la tasa mínima de préstamo y los tipos de cambio del thai baht y del dólar, demostrándose mediante la investigación que los movimientos del índice de Bolsa de Tailandia son sensibles a las anteriores entradas.

Zhu, Xu, Wang y Li (2008) realizaron una investigación que indica que los modelos de redes neuronales con volúmenes de negociación conducen a mejoras, en distinta medida, en la previsión de desempeño en diferentes horizontes de tiempo. Los resultados empíricos indican que los volúmenes de negociación llevan a moderadas mejoras en el rendimiento del índice bursátil a mediano y largo plazo.

Alonso y García (2009) emplean datos de alta frecuencia para encontrar un modelo estadístico que permita disminuir el grado de incertidumbre sobre el comportamiento del futuro inmediato (próximos diez minutos) de operadores del mercado accionario

colombiano, específicamente del IGBC; para ello se usan los modelos Garch en media (Garch-M), que tienen en cuenta los efectos Leverage, día de la semana, hora y día-hora.

En Lu, Chang, Chen, Chiu y Lee (2009) la predicción del índice de la Bolsa de Shanghai B-Share es utilizado para comparar las metodologías MARS (*multivariate adaptive regression splines*), BPN (*backpropagation neural network*), SVR (*support vector regression*) y MLR (*multiple linear regression*). Los resultados obtenidos mostraron que el modelo MARS provee mejores predicciones en términos de error y precisión que los demás, con un porcentaje de error de 1.15% y una precisión direccional (DA) del 82%.

Guresen, Kayakutlu y Daim (2011) observaron que en la mayoría de los casos los modelos de redes neuronales permiten obtener mejores resultados que otros métodos. Los datos utilizados corresponden a los valores diarios del índice Nasdaq desde el 7 de octubre de 2008 hasta el 26 de junio de 2009. Los modelos que se compararon en este trabajo son *multi-layer perceptron* (MLP), *dynamic artificial neural network* (DAN2) y un modelo híbrido de redes neuronales y Garch. El rendimiento de los modelos se evaluó a través del error cuadrado medio y la desviación absoluta media, obteniéndose que el clásico modelo MLP supera a los demás al obtener mejores resultados.

4. Predicción de la bolsa a través de modelos híbridos

El término híbrido hace referencia a la combinación de dos o más elementos. En la predicción de la bolsa de valores muchos autores han realizado combinaciones de métodos o modelos de predicción buscando incorporar en el método híbrido las ventajas de cada uno de los modelos anteriores.

Kim y Han (2000) expresan que existe una gran cantidad de estudios que incluyen modelos de redes neuronales artificiales (ANN), sin embargo las ANN tienen limitaciones en el aprendizaje de patrones debido a que los datos de la bolsa de valores tienen dimensionalidad compleja y gran ruido. Mediante el algoritmo genético (GA) es posible seleccionar el tipo de red neuronal de tal forma que se pueda optimizar la característica relevante del subconjunto. GA se enfoca, en este trabajo, a discretizar los datos y así simplificar el proceso de aprendizaje para las ANN y reducir el ruido y los datos redundantes.

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA

AURA MARÍA JALAL

LUIS ALFONSO GARZÓN

JORGE MARIO LÓPEZ

Pai y Lin (2005) proponen una metodología híbrida entre Arima y el modelo *support vector machines*, este último, una nueva técnica de red neural que se ha aplicado con éxito en la resolución de problemas de estimación de regresión no lineal. Dicha metodología es aplicada en problemas de predicción de los precios de las acciones logrando resultados prometedores. El modelo presentado mejora en gran medida el rendimiento de predicción de precios de las acciones del modelo Arima y el modelo SVM individualmente, pues tanto teórica como empíricamente, hibridando dos modelos diferentes reduce los errores de predicción.

Parisi, A.; Parisi, F. y Díaz, D. (2006) utilizan modelos multivariados dinámicos construidos a partir de algoritmos genéticos y modelos de redes neuronales para predecir el signo de las variaciones semanales de los índices bursátiles asiáticos. Se comparó el signo de la proyección con el de la variación observada en cada i -ésima semana, obteniendo que los modelos multivariados con algoritmos genéticos son más robustos y permiten obtener mejores índices de rentabilidades que los modelos construidos con redes neuronales, así que por medio de esta técnica el inversionista puede tomar posiciones en activos más volátiles, comprándolos cuando se encuentren en la parte inferior del ciclo básico de evolución de precios y vendiéndolos al ubicarse en la parte superior.

Roh (2007) efectuó un estudio en el que usa los test ADF (*Augmented Dickey-Fuller*) para verificar la estabilidad y el test ARCH LM para verificar heterocedasticidad de los datos del índice de Corea Kospi 200. Luego, se hace una comparación del poder de predicción de diferentes modelos individuales tales como EWNA, Garch, Egarch y ANN, y modelos híbridos, obteniéndose como resultado que el error absoluto medio fue menor para los híbridos NN-Garch y NN-Egarch, lo cual permite concluir que los modelos ANN junto a modelos de series de tiempos pueden mejorar el poder de predicción en términos de dirección y precisión.

Yudong y Lenan (2009), por su parte, plantean que las redes neuronales *Back Propagation* se han utilizado para predecir índices; en este trabajo se incorporan con *Improved Bacterial Chemotaxis Optimization* (IBCO), los datos usados son los del índice S&P 500. El error cuadrado medio (MSE) disminuye bastante en el entrenamiento de la red y le toma menos tiempo que el modelo individual BP. El modelo IBCO-BP ofrece menor complejidad computacional, mayor precisión en el pronóstico y menos tiempo de entrenamiento.

El-Henawy, Kamal, Abdelbary y Abas (2010) usan redes neuronales *Multi-Layer Perceptron* para predecir el índice KOSPI 200 (*Korea stock price index 200*) en un periodo de once años. Para mejorar su arquitectura, parámetros, y aumentar la precisión de los pronósticos y disminuir el tiempo de entrenamientos, son utilizados tres algoritmos: *Simulated Annealing* (SA), *Genetic Algorithm* (GA) y un híbrido que combina los dos anteriores. Los resultados muestran que el mejor algoritmo es SA, el cual supera a la aproximación híbrida en un 30% de precisión, y 40% al algoritmo genético. SA necesita 7 minutos de entrenamiento de la red, GA 73 minutos y el enfoque híbrido 98 minutos.

Boyacioglu y Avci (2010) confirman que los rendimientos del mercado bursátil se pueden predecir a través de rendimientos pasados y variables macroeconómicas y financieras. Predecir los rendimientos de las acciones es una tarea compleja debido a la gran cantidad de factores que intervienen en la bolsa de valores. Esto hace que la serie de los precios de las acciones sea dinámica, no lineal, complicada y caótica. En este trabajo se utiliza el modelo Anfis, el cual combina la teoría de redes neuronales y la lógica difusa utilizando variables que comprenden otros índices de bolsa y variables macroeconómicas, el índice de precios del consumidor, la producción industrial, la tasa de cambio del dólar, etc. El modelo obtenido posee un R-cuadrado de 98%, lo que significa que Anfis predice con un alto grado de certeza el índice de la Bolsa de Valores de Estambul.

Hadavandi, Shavandi y Ghanbari (2010) presentan un enfoque integrado basado en sistemas difusos genéticos (GFS) y redes neuronales artificiales (RNA) en la construcción de sistemas expertos para el pronóstico del precio de las acciones, logrando precisión en las predicciones realizadas. Se aplicó el modelo en datos de precios de acciones del sector de las tecnologías de la información (TI) como Dell e IBM y el sector aéreo como las compañías Ryanair. Los resultados mostraron que la precisión de la predicción de CGFS supera al resto de los enfoques con respecto a la evaluación MAPE y CGFS, por lo tanto es una herramienta de previsión adecuada para precios de las acciones.

Wang, Cheng y Hsu (2010) incorporan la cadena de Markov en el modelo estocástico difuso y utiliza los datos del índice de la Bolsa de Taiwán. Dicho modelo es capaz de considerar simultáneamente las tasas de cambio, el aumento de las probabilidades y caídas de los índices bursátiles. De los 330 ensayos realizados durante los tres meses de duración del experimento el modelo resultó ser significativamente mejor en 298 de ellos.

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA

AURA MARÍA JALAL

LUIS ALFONSO GARZÓN

JORGE MARIO LÓPEZ

Yu y Huarng (2010) buscaron aplicar redes neuronales para implementar un nuevo modelo difuso de series de tiempo que mejore los pronósticos e incluya los diversos grados de pertenencia en el establecimiento de relaciones difusas que ayudan en la captura de las relaciones más adecuadas. Estas relaciones difusas se utilizan para pronosticar el índice de la Bolsa de Taiwán.

Según Cheng, Chen y Wei (2010), dos inconvenientes se han encontrado en muchos de los modelos de predicción pasados: 1) los supuestos estadísticos sobre variables que se requieren para los modelos de series de tiempo, tales como el ARMA y ARCH para modelos de predicción de ecuaciones matemáticas, y estas no son fáciles de entender por los inversores en acciones; 2) las normas extraídas de algunos algoritmos de inteligencia artificial (AI), tales como las redes neuronales (NN), no se realizan con facilidad.

Para superar estos inconvenientes se propone un modelo de pronóstico híbrido, utilizando indicadores multitécnicos para predecir las tendencias de precios de acciones. Además, incluye cuatro procedimientos propuestos en el modelo híbrido con el fin de proporcionar normas eficaces para la previsión, que se desarrolló a partir de las reglas extraídas de la teoría de los conjuntos aproximados (RST): 1) seleccionar indicadores técnicos esenciales; 2) utilizar el enfoque de distribución de probabilidad acumulativa (CDPA); 3) emplear un algoritmo RST para extraer reglas lingüísticas del conjunto de datos del indicador técnico lingüístico; 4) utilizar algoritmos genéticos (AG) para refinar las reglas extraídos a efectos de obtener mejor precisión de las previsiones. La eficacia del modelo propuesto se verifica con dos tipos de evaluaciones de desempeño: la precisión y retorno de valores, y mediante el uso de un período de seis años del índice de Taiwán Taiex como el conjunto de datos de prueba. Los resultados experimentales muestran que el modelo propuesto es superior a los modelos de previsión, teoría de los conjuntos aproximados (RST) y algoritmo genético en términos de precisión, y las evaluaciones de rendimiento de las acciones han revelado que los beneficios producidos por el modelo propuesto son más altos que los modelos comprar y mantener, RST y GAS.

Enke, Graue y Mehdiyev (2011) presentan un sistema de predicción del mercado de valores a tres etapas. En la primera fase, el análisis de regresión múltiple se aplica para definir las variables económicas y financieras que tienen una fuerte relación con la salida. En la segunda fase se implementa la evolución diferencial con base en el *Fuzzy Clustering* tipo 2 para crear un modelo de predicción. En la tercera fase se utiliza una

red Fuzzy neural tipo-2 para realizar el razonamiento de la predicción del futuro precio de las acciones. Los resultados de la red de simulación del modelo propuesto indican que supera a los modelos tradicionales de previsión de precios del mercado de valores.

Shen, Guo, Wub y Wu (2011) seleccionaron una red neuronal del tipo *Radial Basis Function* (RBF) para predecir índices bursátiles de la Bolsa de Shanghai. Con la finalidad de optimizar la red e incrementar la eficiencia de las predicciones, se utiliza el algoritmo *Artificial Fish Swarm Algorithm* (AFSA). Los resultados de la RBF optimizada con AFSA, algoritmo genético y *Particle Swarm Optimization* (PSO) son tan buenos como los pronósticos hechos con Arima, *Back Propagation* (BP) o *Support Vector Machine* (SVM), presentando el RBF con AFSA no con la mayor precisión, pero al ser un nuevo algoritmo inteligente es exitoso el hecho de que incrementa el resultado de las predicciones de la red original *Radial Basis Function* (RBF).

Wang J.-Z., Wang, Zhang y Guo (2011) exponen que debido a la variedad de factores que afectan a la bolsa de valores, se propone la utilización del algoritmo *Wavelet De-noising-based Back Propagation* (WDBP); los datos utilizados corresponden a los del índice de la Bolsa de Shanghai desde enero de 1993 a diciembre de 2009. El modelo WDBP fue comparado con un modelo BP (*Back Propagation*), mejorando el primero la precisión de las predicciones.

Para Aboueldahab y Fakhreldin (2011) la predicción de la bolsa de valores es uno de los asuntos más importantes en el campo financiero, por tanto proponen un modelo denominado *Hybrid Genetic Algorithm / Particle Swarm Optimization* (GA/PSO) con *perturbation term*, el cual será evaluado a través de índices como Nasdaq100, Dow Jones y el S & P500. La red neuronal usada como predictor es la *Sigmoid Diagonal Recurrent Neural Network* (SDRNN), debido a que su arquitectura disminuye el error e incrementa la precisión en muchas aplicaciones. La nueva perturbación, agregada al modelo híbrido permite que todas las partículas realicen la búsqueda global en todo el espacio de búsqueda para encontrar nuevas regiones con mejor desempeño.

Asadi, Hadavandi, Mehmanpazir y Nakhostin (2012) proponen un modelo híbrido que es una combinación de métodos de preprocesamiento de datos, algoritmos genéticos y algoritmos Levenberg-Marquardt (LM), usados para el aprendizaje de la alimentación de las redes neuronales en la predicción del índice de acciones. También usan la informa-

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA

AURA MARÍA JALAL

LUIS ALFONSO GARZÓN

JORGE MARIO LÓPEZ

ción de los métodos de preprocesamiento tales como la transformación de los datos y la selección de las variables de entrada para mejorar la precisión del modelo. La competencia del método planteado se prueba con su aplicación para predecir algunos índices bursátiles utilizados en la literatura. Los resultados muestran que el método propuesto es capaz de hacer frente a las fluctuaciones de los valores de la bolsa y también se obtiene una buena precisión de la predicción. Por lo tanto, se puede utilizar para modelar relaciones complejas entre entradas y salidas o para encontrar patrones de datos mientras se realiza la predicción financiera.

Wang, Wang, Zhang Guo (2012) presentan el modelo híbrido (PHM), que es una combinación de los modelos Arima, *exponential smoothing model* (ESM) y BPNN, que permite capturar las características lineales y no lineales en una serie de tiempo. Para probar el modelo se utilizaron los datos mensuales de los índices SZII de China y el *Dow Jones Industrial Average Index* (DJIAI-USA). Se utiliza la predicción direccional (AD) para evaluar la precisión de las predicciones. Los resultados obtenidos demuestran que el modelo híbrido PHM provee mejores resultados en términos de error y precisión que otros modelos como Arima, ESM y BPNN, debido a su robustez.

Dai, Wu y Lu (2012) plantean un modelo de predicción de series de tiempo mediante la combinación de análisis no lineales de componentes independientes y redes neuronales para predecir los índices de las bolsas de valores asiáticas. Se utilizan como ejemplos representativos el índice de cierre de la Bolsa de Valores de Nikkei 225 y de Shanghai *B-share*. Los resultados experimentales muestran que el modelo de predicción propuesto no solo mejora la precisión de la predicción del enfoque de redes neuronales, sino que supera los tres métodos de comparación con precisiones mayores a 80%.

5. Predicción de la bolsa usando variables económicas

Muchos analistas opinan que los movimientos que se dan en la bolsa de valores están estrechamente ligados al comportamiento de la economía del país donde ella opera, entonces hay algunas variables que pueden llegar a incidir en el mercado público de valores: el crecimiento económico, las tasas de interés, el precio del dólar, la producción nacional, los precios del petróleo, los comportamientos de la inflación, el desempleo, entre otras variables de la economía macro de un país.

Kwon y Shin (1998) publican un estudio en el que investigaron si la actividad económica en Corea puede explicar los rendimientos del mercado de valores, para ello usaron una prueba de cointegración y otra de causalidad de Granger a partir de un vector *Error Correction Model*. Este estudio encuentra que los índices bursátiles están cointegrados con un conjunto de variables macroeconómicas: índice de producción, tipo de cambio, balanza comercial y la oferta monetaria que ofrece relación directa de equilibrio a largo plazo con cada índice bursátil.

Gjerde y Sættem (1999) investigan en qué medida los resultados de relaciones entre los rendimientos de las acciones y factores macroeconómicos de los principales mercados son válidos en una economía pequeña y abierta como Noruega, mediante la utilización del vector autorregresivo multivariado (VAR) en los datos de la Bolsa de Valores de Noruega. Los resultados confirman que cambios reales en los tipos de interés afectan tanto los rendimientos de las acciones como a la inflación, y el mercado de valores responde con precisión a los cambios en los precios del petróleo.

López y Vásquez (2002) estudian el riesgo sistemático en una muestra de activos que cotizan en la Bolsa Mexicana de Valores, mediante un modelo que intenta capturar el riesgo derivado de la influencia de variables macroeconómicas que son del conocimiento público. A través del método de extracción de componentes principales (ACP) se selecciona un subconjunto de variables macroeconómicas que puedan representar el riesgo sistemático de los activos mexicanos. Una vez seleccionadas esas variables, se analiza una muestra de 31 acciones que se cotizan en la Bolsa Mexicana de Valores ajustando un modelo $Egarchx(1,1)$ que incluye las variables económicas seleccionadas por medio del ACP en la estimación de los parámetros de la ecuación del rendimiento y la estructura supuesta de los residuales. Se concluye que durante el periodo analizado dichas variables exhibieron influencia en los rendimientos, por lo cual consideramos que la evidencia proporcionada por el estudio sugiere que son explicativas del riesgo sistemático valorado por el mercado mexicano.

Maysami, Howe y Hamzah (2004) evidencian que hay una creciente literatura que muestra la fuerte influencia y relación entre las variables macroeconómicas y el mercado de valores, sobre todo en países industrializados, particularmente en los asiáticos. Entonces, usan un modelo llamado *Vector Error Correction Model* (VECM), el cual produce estimadores más eficientes, ya que permite información completa que produce cointegración en un sistema de ecuaciones sin que variables específicas se normalicen.

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA

AURA MARÍA JALAL

LUIS ALFONSO GARZÓN

JORGE MARIO LÓPEZ

Pierdzioch, Dopke y Hartmann (2008) comparan pronósticos de bolsas de valores volátiles basados en datos macroeconómicos en tiempo real y datos macroeconómicos revisados. Para esto se utilizan dos criterios estadísticos, uno basado en la utilidad y otro en las opciones. El principal resultado es que el valor estadístico y económico de las predicciones basadas en datos macroeconómicos en tiempo real es comparable al valor estadístico y económico de las predicciones basadas en datos macroeconómicos revisados.

Cai, Chou y Li (2009) publican un estudio en el cual se investigan las correlaciones dinámicas entre algunos índices bursátiles internacionales y la inflación. Se utiliza un modelo de series de tiempo autorregresivo denominado (DSTCC-CARR), que revela cómo las correlaciones son significativamente variables en el tiempo y la evolución de estas están relacionadas con las fluctuaciones cíclicas de las tasas de inflación y la volatilidad del mercado. Los índices internacionales utilizados son los de Francia, Alemania, Rusia, Hong Kong, Japón y Estados Unidos, y los niveles de correlación fueron significativos en 95%. Las correlaciones más altas se presentan cuando los países se encuentran en una fase contractiva y las más bajas cuando estos están en una fase expansiva.

Chen (2009) sugiere que las variables macroeconómicas sirven para predecir las recesiones de la Bolsa de Valores de Estados Unidos (el índice de precios Standard & Poor's & P 500), especialmente los diferenciales de rendimiento de curva y las tasas de inflación son los indicadores más útiles; esto se concluyó después de utilizar métodos paramétricos y no paramétricos para identificar los periodos de recesión en el mercado de valores y considerar tanto dentro como fuera de la muestra, pruebas de capacidad predictiva de las variables.

Majhi, Panda, Majhi y Sahoo (2009) proponen predecir índices bursátiles usando *Adaptive Bacterial Foraging Optimization* (ABFO) y BFO, la estructura utilizada en estos modelos de predicción es un combinador lineal simple. Los nuevos modelos son computacionalmente más eficientes, con predicción más precisa, y muestran una convergencia más rápida en comparación con otros modelos de computación evolutivos tales como el algoritmo genético.

Gençtürk, Çelik y Binici (2012) encontraron que no solo es importante la relación entre las variables macroeconómicas y la Bolsa de Valores, sino también que los vínculos

causales entre las relaciones a corto y largo plazo son importantes. Las variables macroeconómicas como la tasa de interés, el tipo de cambio y el índice de precios al consumidor afectan las decisiones del inversor. Se utilizó un modelo VECM, con datos de la Bolsa de Valores de Turquía, con el cual se obtuvo que no hay relación a largo plazo entre el índice de la Bolsa de Turquía y el índice de precios del consumidor, la tasa de interés y la tasa de cambio del dólar, pero sí entre el índice y la producción industrial, lo cual lleva a concluir que la producción industrial es un factor fuerte para predecir la Bolsa de Valores de Turquía.

Caldas y Pires (2012) proporcionaron evidencias empíricas acerca de la influencia de las variables macroeconómicas y el riesgo país en el principal índice del mercado de valores brasileño (Ibovespa). La evidencia práctica se obtuvo con la aplicación de mínimos cuadrados ordinarios (OLS), del método de generalización de momentos (GMM) y de sistemas GMM. Los resultados muestran que la política monetaria y la gestión de la deuda pública, así como la credibilidad y la reputación, afectan el riesgo país y el desempeño del mercado de valores brasileño.

6. Conclusiones

Son muchos los modelos y métodos que han sido utilizados en las dos últimas décadas para predecir índices de las bolsas de valores en los países de mundo. Los primeros modelos fueron autorregresivos, conocidos como Arima, que simplemente capturaban las características lineales de la bolsa y su ventaja radica en el hecho de que no requieren distintas series de datos, implicando esto un ahorro en la especificación e identificación del modelo en el sentido de la econometría tradicional (Domínguez y Zambrano, 2011), pero debido a la gran cantidad de relaciones no lineales entre los datos fue necesario implementar modelos no lineales y no paramétricos. Luego, para mejorar la precisión de los pronósticos se introdujeron modelos híbridos que agruparan las ventajas de los dos o más modelos involucrados.

Particularmente, el descubrimiento de la influencia de las variables macroeconómicas sobre algunas bolsas de valores del mundo ha impulsado el diseño de modelos que utilizan dichas variables como datos de entrada para predecir índices bursátiles, resultando muy eficaces.

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA

AURA MARÍA JALAL

LUIS ALFONSO GARZÓN

JORGE MARIO LÓPEZ

Esto no quiere decir que modelos autorregresivos como Arima o Garch son obsoletos, sino que por su estructura no pueden capturar las características no lineales presentes en las series de tiempo financieras, como las series de índices bursátiles, haciéndose necesario combinar estos métodos con otros que puedan modelar dichas características y mejorar las predicciones.

Este recorrido ha dejado como resultado una amplia lista de métodos de predicción del comportamiento de la bolsa o acciones de una compañía individualmente. Es necesario seguir proponiendo y adaptando los métodos de predicción a necesidades, a contextos, a características de los índices de bolsa, para lograr predicciones más precisas, minimizando el riesgo del inversor.

Tabla 1.

Modelos de predicción de índices de bolsa (lineal, no lineal e híbrido)

Autor	Año de publicación	Mercado estudiado	Tipo de modelo	Técnica	Resultado obtenido
Yale Chang, Carl Yeung y Chris Yip	2000	Índice TCB 500	Lineal	Regresión múltiple	R ² = 0,945
Geetesh Bhardwaj, Norman R. Swanson	2005	Predicción de índices bursátiles	Lineal	Arfima	El modelo Arfima supera el 50% de las veces a otros modelos (AR, ARMA, Arima, Garch)
Philip Hans Franses y Hendrik Ghijlsels	1999	Bolsa de Valores de Amsterdam (AEX), Alemania (DAX), España (MADR) e Italia (MIL)	No lineal	Garch con <i>additive outliers</i> (AO)	MSE: para índice AEX (entre 0,115 y 0,376); para índice DAX (entre 0,171 y 1,240); para índice MADR (entre 0,119 y 1,160); para índice MIL (entre 0,208 y 0,621) Nota: Para el periodo 1987- 1994

Darmadi Komo, Chein- I y Hanseok Ko	1994	Índice industrial Dow Jones	No lineal	Redes neuronales (<i>Radial Basis Function</i> , RBF) y <i>Back-propagation</i>)	Predicciones hasta del 90%
Stefan Zemke	1999	Bolsa de Valores de Varsovia	No lineal	Redes neuronales, <i>Naive Bayes classifier</i> , <i>K-nearest neighbor</i> y algoritmo genético	K- nearest neighbor presenta mejor precisión: 64%
K. H. Lee y G. S. Jo	1999	Bolsa de Valores de Corea	No lineal	<i>Candlestick Chart Analysis Expert System</i>	Aciertos promedios: 72%
An-Sing Chen, Mark T. Leung, Hazem Daouk	2003	Índice de Bolsa de Taiwán	No lineal	Redes neuronales probabilísticas	Predecir la dirección del exceso de rentabilidad más del 50% de las veces Mayor grado de acierto en la red neuronal con desviación para cinco días de patrones de entrenamiento de 2,13%, para diez días de 3,40%, para quince días de 3,72%, para veinte días de 3,29%, para treinta días de 3,81% y para cuarenta días de 3,56%
Eliana Toro, Alexander Molina, Alejandro Garcés	2006	Elementos de la economía colombiana y su Bolsa de Valores	No lineal e híbrido	Red neuronal y red neuronal difusa	

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA
 AURA MARÍA JALAL
 LUIS ALFONSO GARZÓN
 JORGE MARIO LÓPEZ

Suchira Chaigusin, Chaipayorn ChirathamjaRee, Judy Clayden	2008	Bolsa de Valores de Tailandia	No lineal	redes neuronales	MAPE: 1,26594
					Error de predicción absoluto promedio (MAPE): 9,31
Julio César Alonso y Juan Carlos García	2009	Índice General de la Bolsa de Colombia	No lineal	Garch-M	Nota: el promedio de los MAPE de los 18 modelos Garch-M utilizados.
Lu, Chang, Chen, Chiu, y Lee	2009	Bolsa de Valores de Shangai	No lineal	MARS (<i>Multivariate adaptive regression splines</i>)	Error: 1.15% y precisión direccional (DA) del 82%.
Guresen, Kayakutlu y Daim	2011	Índice Nasdaq	No lineal	Redes neuronales (MLP)	Desviación absoluta media: 2,324% (<i>Training</i>) y 2,516% (<i>test</i>)
Kyung-Jae Kim, Ingoo Han	2000	Índice de la Bolsa de Valores de Korea	Híbrido	Redes neuronales con algoritmo genético	Rendimiento promedio de predicción: 65,79% (<i>Training</i>) y 61,70% (<i>Holdout</i>)
Ping-Feng Pai y Chih-Sheng Lin	2005	Precios de las acciones de empresas estadounidenses	Híbrido	Arima- <i>Support Vector Machine</i>	MAPE (modelo híbrido) < MAPE (modelos individuales Arima y SVM)
Antonio Parisi, Franco Parisi y David Díaz	2006	Índices de la Bolsa de Valores de Asia	Híbrido	Modelos multivariados obtenidos con algoritmo genético	El porcentaje de predicción de signos (PPS): mayor 50%

Tae Hyup Roh	2007	Índice Kospi	Híbrido	Red neuronal-Modelos de series de tiempos (NN-Egarch)	Proporción de aciertos: 60,63%
Zhang Yudong, Wu Lenan	2009	S&P 500	Híbrido	<i>Bacterial chemotaxis optimization (BCO) y Back Propagation</i>	Tiempo computacional entre 15 y 17 minutos
Melek Acar Boyacioglu, Derya Avci	2010	Bolsa de Valores de Estambul	Híbrido	Anfis	Error: 0,0068
Esmail Hadavandi, Hassan Shavandi, Arash Ghanbari	2010	Empresas: IBM y Dell Corporations del sector de las IT, British Airlines y Ryanair Airlines, del sector de aerolíneas	Híbrido	<i>Clustering-genetic fuzzy system (CGFS)</i>	MAPE: IBM (0,63), DELL (0,534), British Airlines (1,413) y Ryanair Airlines (1,241)
Yi-Fan Wang, Shih-min Cheng, Mei-Hua Hsu	2010	Bolsa de Valores de Taiwán	Híbrido	Estocástico difuso que incorpora cadenas de Markov	Desviaciones entre 0.0025% -1,03%.
Tiffany Hui-Kuang Yu, Kun-Huang Hwang	2010	Bolsa de Valores de Taiwán	Híbrido	Modelo de series de tiempo difuso que usa redes neuronales	Error cuadrado medio: 441,8
I. M. El-Henawy, A. H. Kamal, H. A. Abdalbary, y A. R. Abas	2010	Índice Kospi 200	Híbrido	Redes neuronales (MLP) y SA (<i>simulated annealing</i>)	<i>Testing error:</i> 0,4897 y <i>Training error:</i> 0,4886

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA
 AURA MARÍA JALAL
 LUIS ALFONSO GARZÓN
 JORGE MARIO LÓPEZ

Ching-Hsue Cheng, Tai-Liang Chen, Liang-Ying Wei	2010	Bolsa de Valores de Taiwán	Híbrido	Rough Sets Theory (RST) y Algoritmo Genético	Precisión promedio: 0,601
Tarek Abouel-dahab y Mahumod Fakhreldin	2011	Índices: DJIA, Nasdaq 100 y S & P 500	Híbrido	Genetic Algorithm / Particle Swarm Optimization (GA/ PSO) con perturbación	MAPE: <i>Training</i> : 0,10% (DJIA), 0,097% (Nasdaq 100) y 0,073% (S & P 500). <i>Testing</i> : 0,10% (DJIA), 0,11% (Nasdaq 100) y 0,077 % (S & P 500)
David Enke, Manfred Grauer, Nijat Mehdiyevb	2011	Índice S & P 500	Híbrido	Multiple Regression, Fuzzy Type-2 Clustering and Neural Networks	RMSE: 0,909
Wei Shen, Xiaopen Guo, Chao Wu, Desheng Wu	2011	Bolsa de Valores de Shangai	Híbrido	Radial Basis Function (RBF) optimizada con Fish Swarm Algorithm (ASFI)	Error: 0,5395%
Ju-Jie Wang, Jian-Zhou Wang, Zhe-George Zhang, Shu-Po Guo	2011	Índice de Bolsa de Shangai	Híbrido	Wavelet De-noising-based Back Propagation (WDBP)	MAPE: BP1 (0,4364), BP2 (0,2584), BP3 (0,4831), BP4 (0,1948), BP5 (0,3056) y BP6 (0,4276)

Wensheng Dai, Jui-Yu Wu, Chi-Jie Lu	2012	Nikkei 225 y Shanghai B-share	Híbrido	<i>Nonlinear independent component analysis</i> (NLICA) y redes neuronales (BPN)	Para Nikkei 225 (MAPE= 0,242% y DS = 85,69%). Para Shanghai B-share (MAPE = 0,95% y DS = 80,50%)
Shahrokh Asadi, Esmaeil Hadavandi, Farhad Mehmanpazir, Mohammad Masoud Nakhostin	2012	Taiwan Stock Exchange index (TSE) 0.51, Tehran Stock Exchange Prices Index (Tepix), Index of top 50 companies, Industry index, Index of Financial Group, Dow Jones Industrial Average Index Series, Nasdaq Index Series	Híbrido	Métodos de procesamiento de datos, algoritmo Levenberg–Marquardt (LM) y algoritmo genético	MAPE: Taiwan Stock Exchange index (TSE) 0,51, Tehran Stock Exchange Prices Index (Tepix) 0,50, Index of top 50 companies 0,76, Industry index 0,89, Index of Financial Group 0,66, Dow Jones Industrial Average Index Series 1,41, Nasdaq Index Series 0,13
Ju-Jie Wang, Jian-Zhou Wang, Zhe-George Zhang Shu-Po Guo	2012	Índices: SZII y DJIAI	Híbrido	<i>Exponential Smoothing Mode</i> (ESM), Arima y redes neuronales (BP)	MAPE: SZII (26,06%) y Djiai (30,53%)

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA

AURA MARÍA JALAL

LUIS ALFONSO GARZÓN

JORGE MARIO LÓPEZ

Tabla 2.

Descripción de modelos de predicción de índices de bolsa

Técnica	Descripción	Tipo de modelo
Regresión múltiple	Este método permite estudiar y modelar el efecto que algunas variables tienen sobre otra, estimando parámetros para cada variable con el fin de construir un modelo que describa el fenómeno.	Lineal
<i>Arfima Autoregressive Fractionally Integrated Moving Average</i>	Un método referencia en el análisis de series temporales con dependencia a largo plazo. Estos tipos de modelos permiten ajustar la dinámica a largo plazo, problema comúnmente asociado a los modelos ARMA (<i>AutoRegressive Moving Average</i>), proporcionando una descripción más parsimoniosa de las series económicas o financieras.	Lineal
Garch Modelos generalizados autorregresivos condicionalmente heterocedásticos	Los modelos Garch, que extiende la clase de los modelos ARCH, fue introducida por Bollerslev (1991). Creados para el alojamiento de dependencias temporales de econometría financiera, se convirtieron en una herramienta popular para modelado de la volatilidad y la previsión. Sin embargo, a pesar del éxito del modelo Garch, ha sido criticado por no capturar volatilidad asimétrica	No Lineal
Egarch o Garch exponencial	Los modelos Garch no tienen en cuenta un comportamiento asimétrico propio de la volatilidad de los retornos de las series financieras, que es el efecto de apalancamiento. Para incluir el efecto asimétrico que tiene el cambio de los precios de un activo en su volatilidad, Nelson (1991) propone el modelo Egarch (modelo exponencial generalizado, autorregresivo, condicionalmente heterocedástico). Modela el efecto de asimetría al considerar una función g de las innovaciones z_t , las cuales son variables i.i.d. de media cero, que involucra tanto el valor de la innovación z_t como su magnitud expresada por medio de $ z_t - E(z_t)$.	No Lineal
Garch-M o Garch <i>in mean</i>	Este modelo fue desarrollado por Engle, Lilien y Robbins (1987) para modelar la volatilidad de los retornos de acciones. En este modelo se relacionan los retornos con un término de la volatilidad esperada.	No lineal

Garch con additive outliers (AO)	<p>En el modelo Garch se observa con frecuencia que los residuos estimados de este modelo aún tienen exceso de curtosis. Una de las posibles causas es que hay ciertas observaciones de los retornos, llamados <i>outliers</i> aditivos [AO], no son capturados por un modelo Garch estándar. Tenga en cuenta que los AO despreciados también conducen a estimaciones de los parámetros sesgadas y a pronósticos sesgados. Por lo tanto, dejarlos de lado al pronosticar retornos financieros puede dar lugar a previsiones de volatilidad sesgadas, por ello en este modelo se corrigen estas observaciones.</p>	No lineal
Redes neuronales (<i>Radial Basis Function</i> , RBF) y <i>Backpropagation</i>)	<p>Las redes de neuronas artificiales son un paradigma de aprendizaje y procesamiento automático inspirado en la forma como funciona el sistema nervioso de los animales. Las redes de base radial son un tipo de redes de neuronas artificiales que calculan la salida de la función en función de la distancia a un punto denominado centro. Los parámetros de esta red neuronal son pesos, media y la varianza, y se utilizó la función de Gauss como la función de transferencia. Estos parámetros describen más o menos la importancia relativa otorgada a los valores actuales y pasados de los datos de entrada</p>	No lineal
Métodos de <i>K-nearest neighbor</i>	<p>Este es un método de clasificación no paramétrico, que estima el valor de la función de densidad de probabilidad o directamente la probabilidad <i>a posteriori</i> de que un elemento X pertenezca a una determinada clase a partir de la información proporcionada por el conjunto de prototipos. En el proceso de aprendizaje no se hace ninguna suposición acerca de la distribución de las variables predictoras.</p>	No lineal
<i>Candlestick Chart Analysis Expert System</i>	<p>Es una técnica de inteligencia artificial que pretende explicar el estado del mercado. Los patrones usados se clasifican en cinco grupos con respecto a su significado: bajando, subiendo, neutral, continuando la tendencia y patrones donde se invierte la tendencia.</p>	No lineal

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA
 AURA MARÍA JALAL
 LUIS ALFONSO GARZÓN
 JORGE MARIO LÓPEZ

Red neuronal probabilística (PNN)	<p>Las redes neuronales probabilísticas (RNP) pertenecen a la familia de redes neuronales con función base radial. Las redes neuronales probabilísticas están compuestas por cuatro capas: una capa de entrada, una capa de patrones, una por cada vector, una capa sumatoria. La red neuronal probabilística (PNN) se utiliza para predecir la dirección de rendimiento del índice después de que se entrena con datos históricos.</p>	No lineal
Algoritmo genético (GA)	<p>Los algoritmos genéticos consisten en una función matemática o una rutina que simula el proceso evolutivo de las especies, teniendo como objetivo encontrar soluciones a problemas específicos de maximización o minimización. Un algoritmo genético recibe como entrada una generación de posibles soluciones para el problema en cuestión, y arroja como salida los especímenes más aptos por generación (es decir, las mejores soluciones), para que estos se reproduzcan y generen mejores descendientes, los que a su vez deberían tener características superiores que las generaciones pasadas.</p>	No lineal
Redes neuro-difusas	<p>Los algoritmos genéticos trabajan con códigos que representan las posibles soluciones al problema. Por ello, es necesario establecer una codificación para todo el rango de posibles soluciones antes de comenzar a trabajar con el algoritmo.</p> <p>Estas redes neuronales son aplicables al comportamiento de las acciones en la bolsa. La lógica difusa permite procesar información cuando se manejan términos inexactos e imprecisos, como funciona el cuerpo humano.</p>	No lineal e híbrido
MARS <i>Multivariate adaptive regression splines</i>	<p>Es una técnica de regresión no lineal y no paramétrica, un procedimiento flexible que modela las relaciones “casi aditivas” o implican interacciones con menos variables. En comparación con los modelos de RVS y BPN, la mayor ventaja de MARS es que puede identificar importantes variables independientes a través de las funciones de base construidas cuando se consideran muchas variables posibles. Por otra parte, MARS no necesita largo proceso de formación y por lo tanto puede ahorrar tiempo de modelado.</p>	No lineal

Redes neuronales MLP	Es una red neuronal artificial formada por múltiples capas, esto le permite resolver problemas que no son linealmente separables, lo cual es la principal limitación del perceptrón (también llamado perceptrón simple). Dos importantes características de MLP son: su procesamiento de elementos no lineales, los cuales tienen una no linealidad que debe ser suavizada, y su masiva interconectividad. Estas redes son normalmente entrenadas con el algoritmo <i>Back Propagation</i> .	No lineal
Arima y SVM <i>Support vector machines</i>	Es una técnica de red neural que combina el modelo Arima, el cual supone que el valor futuro de una variable es resultado de la combinación de sus valores y errores pasados, y el SVM que busca minimizar el límite superior del error generalizado, en lugar del error empírico como otras redes neuronales, y genera un modelo de regresión a partir de un conjunto de funciones lineales.	Híbrido
Modelos multivariados dinámicos contruidos a partir de algoritmos genéticos y de los modelos de redes neuronales	Se hace una combinación de las características de estas dos técnicas de la siguiente manera: el objetivo del algoritmo genético es encontrar un modelo multivariado dinámico que maximice el porcentaje de predicción de signo de las variaciones de los índices bursátiles en estudio. Por su parte en el modelo de red neuronal la variable de salida está dada por la variación del índice en determinado período.	Híbrido
Modelo IBCO-BP <i>Improved Bacterial Chemotaxis Optimization - Back Propagation</i>	La estructura del modelo es básicamente un BP adaptativo cuyas ponderaciones se actualizan con la herramienta IBCO, la cual se inspira en el comportamiento de alimentación bacteriana, resulta de una analogía con la forma en bacterias, reaccionan a quimioatrayentes en los gradientes de concentración.	Híbrido
Anfis <i>Adaptive neuro fuzzy inference system</i>	Integra las redes neuronales y principios de lógica difusa, tiene potencial para aprovechar los beneficios de ambos en un solo marco. El Anfis consiste en reglas si-entonces y datos de entrada que produce una salida. La lógica difusa se adapta mejor al mundo real e incluso puede comprender y funcionar con expresiones del lenguaje humano. La clave de esta adaptación al lenguaje se basa en comprender los cuantificadores de cualidad para nuestras inferencias, esto resulta útil en la predicción de índices bursátiles.	Híbrido

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA
 AURA MARÍA JALAL
 LUIS ALFONSO GARZÓN
 JORGE MARIO LÓPEZ

CGFS <i>Clustering-genetic fuzzy system</i>	<p>Enfoque integrado basado en sistemas difusos genéticos (GFS) y redes neuronales artificiales (RNA) para la construcción de sistemas expertos en el pronóstico del precio de las acciones.</p>	Híbrido
Cadena de Markov y modelo estocástico difuso	<p>La cadena de Markov es un tipo especial de proceso estocástico discreto en el que la probabilidad de que ocurra un evento depende del evento inmediatamente anterior, tiene memoria, recuerda el último evento y esto condiciona las posibilidades de los eventos futuros. En este método se incorpora la cadena de Markov en el modelo estocástico difuso y dicho modelo es capaz de considerar simultáneamente las tasas de cambio, el aumento de las probabilidades y caídas de los índices bursátiles en las predicciones.</p>	Híbrido
Modelo de series de tiempo difuso que usa redes neuronales	<p>Las redes neuronales han sido populares debido a sus capacidades en el manejo de relaciones no lineales. Este método las combina con un modelo de serie temporal en el que se establecen distintos grados de participación que ayudan en la captura de las relaciones con más propiedad, mejorando el pronóstico.</p> <p>El recocido simulado (SA) es una metaheurística probabilística para la optimización global en problemas de localizar una buena aproximación en determinada función, en un gran espacio de búsqueda.</p>	Híbrido
Redes neuronales y recocido simulado SA	<p>Este modelo utiliza algoritmos de búsqueda en lugar de <i>Backpropagation</i> para el entrenamiento de la red neuronal mediante la búsqueda de parámetros óptimos que incluye pesos y el número de unidades ocultas en la capa oculta para aumentar la precisión de la predicción y evitar inconvenientes comunes de <i>Backpropagation</i>.</p>	Híbrido
<i>Rough sets theory and genetic algorithms</i>	<p>Modelo de pronóstico híbrido, utilizando indicadores multitécnicos para predecir las tendencias de precios de acciones. Incluye cuatro procedimientos propuestos: 1) seleccionar indicadores técnicos esenciales; 2) utilizar el enfoque de distribución de probabilidad acumulativa (CDPA); 3) emplear un algoritmo RST para extraer reglas lingüísticas del conjunto de datos del indicador técnico lingüístico; 4) utilizar algoritmos genéticos (AG) para refinar las reglas extraídas con la finalidad de obtener mejor precisión de las previsiones.</p>	Híbrido

<i>Algorithm / Particle Swarm Optimization (GA/PSO) con perturbación</i>	<p>El PSO es un método computacional que optimiza un problema iterativamente al tratar de mejorar una solución candidata con respecto a una determinada medida de la calidad. Este híbrido está inspirado en el mecanismo biológico congregación pasiva para superar el problema de la restricción de la búsqueda local en los modelos estándar híbridos (GA/PSO). El término de perturbación se basa en la cooperación entre las diferentes partículas en la determinación de nuevos puestos en lugar de depender de las partículas, pensamiento egoísta que permite a todas las partículas realizar la búsqueda global en todo el espacio de búsqueda para encontrar nuevas regiones con mejor rendimiento.</p>	Híbrido
<i>Multiple Regression, Fuzzy Type-2 Clustering y Neural Networks</i>	<p>En esta técnica se realiza un análisis de regresión múltiple para definir las variables económicas que tienen una fuerte relación con la salida. Luego se implementa <i>Fuzzy Type-2 Clustering</i> para crear un modelo de predicción, y finalmente se utiliza la red <i>Fuzzy neural tipo-2</i> para realizar el razonamiento de la predicción del futuro precio de las acciones.</p>	Híbrido
<i>Radial Basis Function (RBF) y Artificial Fish Swarm Algorithm (AFSA)</i>	<p>El <i>Artificial Fish Swarm Algorithm</i> se inspira en el comportamiento social natural de los peces. Ha sido ampliamente utilizado en la predicción de corto plazo para su propia adaptación y características de autoaprendizaje. Sin embargo, al ser una red típica de la inteligencia artificial, tiene limitaciones con respecto a la velocidad de convergencia y precisión de las predicciones.</p>	Híbrido
<i>Algoritmo Wavelet De-noising-based Back Propagation (WDBP)</i>	<p>Para hacer frente a este problema los investigadores han introducido algoritmos inteligentes a los fines de optimizar RBF, en este caso la técnica se combinó con <i>Radial Basis Function</i>.</p> <p>En este modelo los datos originales se descomponen primero en múltiples capas a través de la transformación Wavelet. Cada capa tiene un componente de señal de baja y de alta frecuencia.</p>	Híbrido

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA
 AURA MARÍA JALAL
 LUIS ALFONSO GARZÓN
 JORGE MARIO LÓPEZ

<i>Nonlinear independent component analysis and neural network</i> NLICA	<p>La Nlica es una técnica novedosa e importante de extracción para encontrar fuentes independientes de datos mixtos no lineales observados, donde hay disponibles mecanismos para mezclar información no relevante. En el método propuesto, primero se utiliza la NLICA en transformar el espacio de entrada compuesta de los datos de serie de tiempo originales en el espacio de características, que consiste en componentes independientes que representan la información subyacente de los datos originales. Entonces, los IC sirven de variables de entrada de la red neuronal para construir el modelo de predicción.</p>	Híbrido
Algoritmo genético y algoritmos Levenberg-Marquardt (LM)	<p>Modelo híbrido que combina métodos de preprocesamiento de datos, algoritmos genéticos y algoritmos Levenberg-Marquardt (LM) usados para el aprendizaje de la alimentación de las redes neuronales en la predicción del índice de acciones. También usa la información de los métodos de preprocesamiento tales como la transformación de los datos y la selección de las variables de entrada, para la mejora de la precisión del modelo.</p>	Híbrido
Modelo híbrido PHM: suavizado exponencial (ESM), Arima y BPNN	<p>El suavizado exponencial es un método relativamente simple pero robusto en la predicción de series de tiempo. Es aplicable a series de tiempo homocedásticas y heterocedásticas. Un híbrido entre los tres modelos permite aprovechar las bondades de cada uno.</p>	Híbrido

Fuente: definiciones tomadas de los artículos usados en la revisión bibliográfica.

Cada nuevo investigador deberá indagar sobre métodos usados para pronosticar los índices bursátiles de la bolsa donde va invertir y, por su parte, los académicos deben seguir trabajando para analizar y modelar el comportamiento de las bolsas de valores y determinar las relaciones que estas tienen con factores macroeconómicos, a fin de que se conozcan los efectos de esas relaciones y se propongan metodologías, con base en lo encontrado, que obtengan mejores predicciones. Tal es el caso de un estudio que se presentará más adelante.

Referencias

- KOMO, D.; CHANG, C.-I. Y KO, H. (1994). Neural Network Technology For Stock Market Index Prediction. *International Symposium on Speech, Image Processing and Neural Networks* (pp. 543-546). Hong Kong: IEEE.
- LASFER, M.; MELNIK, A. Y THOMAS, D. (2003). Short-term reaction of stock markets in stressful circumstances. *Journal of Banking & Finance*, 1959-77.
- LIU, H.-C. Y HUNG, J.-C. (2010). Forecasting S&P-100 stock index volatility: The role of volatility asymmetry and distributional assumption in Garch models. *Expert Systems with Applications*, 4928-34.
- ZEMKE, S. (1999). Nonlinear index prediction. *Physica A: Statistical Mechanics and its Applications*, 177-83.
- ABOUELDAHAB, T. Y FAKHRELDIN, M. (2011). Prediction of Stock Market Indices using Hybrid Genetic Algorithm/ Particle Swarm Optimization with Perturbation Term. *International Conference on swarm intelligence*. Cergy, France.
- ALONSO, J. Y GARCIA, J. (2009). ¿Qué tan buenos son los patrones del IGBC para predecir su comportamiento? *Universidad Icesi*, 13-36.
- ASADI, S.; HADAVANDI, E.; MEHMANPAZIR, F. Y NAKHOSTIN, M. (2012). Hybridization of evolutionary Levenberg-Marquardt neural networks and data pre-processing for stock market prediction. *Knowledge-Based Systems*, 245-58.
- BHARDWAJ, G. Y SWANSON, N. (2006). An empirical investigation of the usefulness of Arfima models for predicting macroeconomic and financial time series. *Journal of Econometrics*, 539-78.
- BOYACIOGLU, M. A. Y AVCI, D. (2010). An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange. *Expert Systems with Applications*, 7908-12.
- CAI, Y.; CHOU, R. Y LI, D. (2009). Explaining international stock correlations with CPI fluctuations and market volatility. *Journal of Banking & Finance*, 2026-35.
- CALDAS MONTES, G. Y PIRES TIBERTO, B. (2012). Macroeconomic Environment, country risk and stock market performance: Evidence for Brazil. *Elsevier Science*, 1666-78.

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA
AURA MARÍA JALAL
LUIS ALFONSO GARZÓN
JORGE MARIO LÓPEZ

- CHAIGUSIN, S.; CHIRATHAMJAREE, C. Y CLAYDEN, J. (2008). The Use of Neural Networks in the Prediction of the Stock Exchange of Thailand (SET) index. *Computational Intelligence for Modelling Control & Automation*. IEEE.
- CHANG, Y.; YEUNG, C. Y YIP, C. (2000). *Analysis of the influence of economic indicators on stock prices using multiple regression*.
- CHEN, A.-S.; LEUNG, M. Y DAOUK, H. (2003). Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index. *Computers & Operations Research*, 901-23.
- CHEN, S.-S. (2009). Predicting the bear stock market: Macroeconomic variables as leading indicators. *Journal of Banking & Finance*, 211-23.
- CHENG, C.-H.; CHEN, T.-L. Y WEI, L.-Y. (2010). A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting. *Information Sciences*, 1610-29.
- CLEMENTS, M.; FRANCES, P. Y SWANSON, N. (2004). Forecasting economic and financial time-series with non-linear models. *International Journal of Forecasting*, 169-83.
- DAI, W.; WU, J.-Y. Y LU, C.-J. (2012). Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes. *Expert Systems with Applications*, 4444-52.
- DOMÍNGUEZ GUIÓN, R. Y ZAMBRANO REYES, A. (2011). Pronóstico con modelos Arima para los casos del índice de precios y cotizaciones (IPC) y la Acción de América Móvil (AM). *Memoria del XXI Coloquio Mexicano de Economía Matemática y Econometría*.
- EL-HENAWY, I.; KAMAL, A.; ABDELBAR, H. Y ABAS, A. (2010). Predicting Stock Index Using Neural Network Combined with Evolutionary Computation Methods. *IEEE*, 1-6.
- ENKE, D.; GRAUER, M. Y MEHDIYEV, N. (2011). Stock Market Prediction with Multiple Regression, Fuzzy Type-2 Clustering and Neural Networks. *Procedia Computer Science*, 201-06.
- FRANCES, P. Y GHIJSELS, H. (1999). Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting*, 1-9.
- GENÇTÜRK, M.; ÇELİK, I. Y BINICI, Ö. (2012). Causal relations among stock returns and macroeconomic variables in a small and open economy. *African Journal of Business Management*, 6177-82.

- GJERDE, Ø. Y SÆTTEM, F. (1999). Causal relations among stock returns and macroeconomic. *Journal of International Financial Market*, 61-74.
- GURESEN, E.; KAYAKUTLU, G. Y DAIM, T. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, 10389-97.
- HADAVANDI, E.; SHAVANDI, H. Y GHANBARI, A. (2010). Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting. *Knowledge-Based Systems*, 800-08.
- KIM, K.-J. Y HAN, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications*, 125-32.
- KWON, C. Y SHIN, T. (1998). Cointegration and causality between macroeconomic variables and stock market returns. *Global Finance Journal*, 71-81.
- LEE, K. Y JO, G. (1999). Expert system for predicting stock market timing using a candlestick chart. *Expert Systems with Applications*, 357-64.
- LÓPEZ HERRERA, F. Y VÁSQUEZ TÉLLEZ, F. (2002). Variables macroeconómicas y un modelo multifactorial para la Bolsa Mexicana de Valores: análisis empírico sobre una muestra de activos. *Academia. Revista Latinoamericana de investigación*, 5-28.
- LU, C.-J.; CHANG, C.-H.; CHEN, C.-Y.; CHIU, C.-C. Y LEE, T.-S. (2009). Stock Index Prediction: A Comparison of MARS, BPN and SVR in an Emerging Market. *Industrial Engineering and Engineering Management*. IEEE.
- MAJHI, R.; PANDA, G.; MAJHI, B. Y SAHOO, G. (2009). Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques. *Expert Systems with Applications, Elsevier*, 10097-104.
- MAYSAMI, HOWE Y HAMZAH (2004). Relationship between Macroeconomic Variables and Stock Market Indices: Cointegration Evidence from Stock Exchange of Singapore's All-S Sector Indices. *Journal Pengurusan*, 47-77.
- PAI, P.-F. Y LIN, C.-S. (2005). A hybrid Arima and support vector machines model in stock price forecasting. *Omega, Elsevier*, 497-505.
- PARISI, A.; PARISI, F. Y DÍAZ, D. (2006). Modelos de algoritmo genético y redes neuronales en la predicción de índices bursátiles asiáticos. *Cuadernos de economía*, 251-84.

Métodos para predecir índices bursátiles

MARTHA CECILIA GARCÍA
AURA MARÍA JALAL
LUIS ALFONSO GARZÓN
JORGE MARIO LÓPEZ

- PIERDZIOCH, C.; DOPKE, J. Y HARTMANN, D. (2008). "Forecasting stock market volatility with macroeconomic variables in real time". *Journal of Economics and Business*, 256-76.
- REDDY, B. (2010). Prediction of Stock Market Indices – Using SAS. *IEEE*.
- ROH, T. (2007). Forecasting the volatility of stock price index. *Expert Systems with Applications*, 916-22.
- SHEN, W.; GUO, X.; WUB, C. Y WU, D. (2011). Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm. *Knowledge-Based Systems*, 378-85.
- TORO OCAMPO, E. M.; MOLINA CABRERA, A. Y GARCÉS RUIZ, A. (2006). Pronóstico de bolsa de valores empleando técnicas inteligentes. *Tecnura*.
- WANG, J.-Z.; WANG, J.-J.; ZHANG, Z.-G. Y GUO, S.-P. (2011). Forecasting stock indices with back propagation neural network. *Expert Systems with Applications*, 14346-55.
- WANG, J.-J.; WANG, J.-Z.; ZHANG, Z.-G. Y GUO, S.-P. (2012). Stock index forecasting based on a hybrid model. *Omega*, 758-66.
- WANG, Y.-F.; CHENG, S. Y HSU, M.-H. (2010). Incorporating the Markov chain concept into fuzzy stochastic prediction of stock indexes. *Applied Soft Computing*, 613-17.
- YU, T.-K. Y HUARNG, K.-H. (2010). A neural network-based fuzzy time series model to improve forecasting. *Expert Systems with Applications*, 3366-72.
- YUDONG, Z. Y LENAN, W. (2009). Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. *Expert Systems with Applications*, 8849-54.
- ZHU, X.; XU, L.; WANG, H. Y LI, H. (2008). Predicting stock index increments by neural networks: The role of trading volume under different horizons. *Expert Systems with Applications*, 3043-54.