

Trabajo Final de Maestría

**Una revisión sistemática acerca de las metodologías para el pronóstico
de índices de mercado: su estado actual y tendencias futuras**

Erika Silvana Navales Cardona, Especialista en Ingeniería Financiera



Director de Tesis

Prof. Juan David Velásquez Henao, M.Sc., Ph.D.

Universidad Nacional de Colombia
Facultad Nacional de Minas
Escuela Ingeniería de la Organización
Programa de Maestría en Ingeniería – Ingeniería Administrativa

Noviembre 2012

A Frank,
por el ánimo que me brinda constantemente para alcanzar nuevas metas.

A mi familia: papá, mamá, hermanos y sobrinos,
a quienes agradezco su apoyo incondicional.

Agradecimientos

A todas las personas que hicieron posible el desarrollo de este trabajo.

RESUMEN

Contexto: La predicción de los índices de mercado se ha convertido en un tema importante y de creciente interés; en los últimos años se han desarrollado múltiples metodologías para tal fin, basadas generalmente en técnicas estadísticas, matemáticas o inteligencia computacional. Sin embargo, no hay estudios orientados a establecer el estado actual de las investigaciones relacionadas con el tema.

Objetivo: Este estudio pretende determinar el estado actual de las investigaciones sobre las metodologías más utilizadas para el pronóstico de índices de mercado.

Método: Se realizó una revisión sistemática de 64 artículos publicados desde 1994 hasta 2011, usando seis preguntas de investigación. Los artículos seleccionados en esta revisión se pueden encontrar comentados en el Anexo 1.

Resultados: El número de artículos finalmente seleccionados en el trabajo respecto al pronóstico de índices de mercado y el número de citas que han tenido, evidencian el crecimiento y la importancia que ha adquirido el tema durante los últimos años; adicionalmente, se evidencia la necesidad de establecer un marco consolidado acerca de las metodologías de pronóstico existentes. Según los resultados obtenidos, dentro de las técnicas más utilizadas para el pronóstico de índices de mercado están: la lógica difusa, utilizada desde el año 2001 y teniendo su mayor auge en el período 2010-2011; y las redes neuronales artificiales, que aunque tienen publicaciones asociadas desde el año 2000, son aplicadas en gran medida durante el año 2011. Adicionalmente, se han desarrollado métodos basados en máquinas de soporte vectorial, algoritmos adaptativos, algoritmos genéticos, entre otros. En cuanto a las técnicas usadas en las metodologías de pronóstico de volatilidad, se encuentran los modelos ARCH, GARCH y sus variaciones, utilizados desde el año 1994, teniendo sus principales investigaciones y publicaciones a partir del 2007 y un auge importante en el 2011.

Conclusión: Según la literatura analizada sobre el pronóstico de índices una de las prácticas usuales en los estudios analizados consiste en usar varios métodos para el pronóstico y seleccionar el que genere mayor precisión en los resultados, medidos a partir de las comparaciones del error de pronóstico en cada uno de ellos; y comparar modelos híbridos contra técnicas clásicas con el fin de determinar si los modelos híbridos contribuyen al mejoramiento de la precisión de los resultados de los modelos. Adicionalmente, tanto los valores extremos de la serie de tiempo de los índices de mercado, como la longitud de los intervalos de los datos utilizados para la predicción, afectan significativamente los resultados de los modelos. Es por esto, que dichos factores deben tenerse en cuenta a la hora de aplicar las metodologías de pronóstico de índices bursátiles.

Palabras clave: índices de mercado, pronóstico, modelado, volatilidad, dirección de los cambios.

CONTENIDO

	Pag
1 INTRODUCCIÓN.....	1
2 METODOLOGÍA DE INVESTIGACIÓN	5
2.1 Planificación de la revisión.....	6
2.1.1 Identificación de la necesidad de revisión.....	6
2.1.2 Definir un protocolo de revisión.....	6
2.2 Desarrollo de la revisión	7
2.2.1 Identificación de la investigación	7
2.2.2 Selección de estudios.....	7
2.2.3 Evaluación y calidad del estudio	7
2.2.4 Extracción de los datos	8
2.2.5 Síntesis de datos	8
2.3 Publicación de resultados.....	8
3 PROTOCOLO DE INVESTIGACIÓN.....	9
3.1 Proceso de búsqueda.....	9
3.2 Criterios de inclusión o exclusión.....	9
3.3 Criterios de calidad	10
3.4 Selección de estudios y análisis de datos.....	10
4 RESULTADOS OBTENIDOS.....	13
5 DISCUSIÓN.....	35
5.1 RQ1 ¿Cuáles son las metodologías de pronóstico de índices de mercado comúnmente usadas?	35

5.2	RQ2 ¿Cuáles son las metodologías de pronóstico de la dirección de los cambios de índices de mercado comúnmente usadas?	41
5.3	RQ3 ¿Cuáles son las metodologías de pronóstico de la volatilidad de índices de mercado comúnmente usadas?	43
5.4	RQ4 ¿Cuáles son las metodologías más utilizadas para modelar índices de mercado?.....	45
5.5	RQ5 ¿Cuáles son las metodologías más utilizadas para modelar la volatilidad de índices de mercado?.....	47
5.6	RQ6 ¿Cuáles son los principales factores que afectan la precisión de los modelos de pronóstico?	47
6	CONCLUSIONES.....	51
	ANEXO 1: Descripción artículos analizados	53
	BIBLIOGRAFÍA	83

LISTADO DE TABLAS

	Pag
Tabla 1. Artículos seleccionados	13

LISTADO DE FIGURAS

	Pag
Figura 1. Número de artículos por año de publicación.....	31
Figura 2. Número de citaciones de artículos por año de publicación.....	32
Figura 3. Proporción de artículos publicados por índice de mercado	33
Figura 4. Parámetro para la evaluación del método	34
Figura 5. Evolución de las técnicas empleadas para el pronóstico de índices de mercado	42
Figura 6. Evolución de las técnicas utilizadas para el pronóstico de volatilidad de índices de mercado	46

1 INTRODUCCIÓN

En esta investigación se utiliza la metodología de revisión sistemática de la literatura para establecer el estado actual sobre los métodos de pronóstico de índices de mercado. Los índices de mercado son variables que reflejan el comportamiento y las variaciones del mercado de valores y sus niveles de capitalización; indican la tendencia general de los precios de las acciones (Martinovic, Vukmirovic y Vukmirovic, 2005; Chi, Liu y Lau, 2009; Zhang y Zhang, 2009); son la base para proporcionar a los inversionistas criterios para la toma de decisiones y para que éstos se protejan de los riesgos potenciales del mercado; y representan oportunidades para los especuladores y arbitrajistas. Adicionalmente, son la base para proveer información a los potenciales inversionistas para definir la mejor forma de invertir su capital. Sin embargo, su pronóstico es difícil debido a la complejidad de los datos del mercado y del mercado de valores como tal (Lin, Haider y Chi-Chung, 2002; Khurram y Jassal, 2011; Sallehuddin, Shamsuddin y Hashim, 2008; Zhang y Zhang, 2009).

En los últimos años, el pronóstico de los precios de valores financieros y de los índices de mercado se ha convertido en un tema importante y de creciente interés (Lu, Lee y Chiu, 2009; Shen, Guo, Wu y Wu, 2011); en los últimos años se ha aumentado el número de artículos publicados, especialmente en los períodos 2001, 2008, 2009, 2010 y 2011, con 4, 7, 13, 15 y 16 artículos publicados respectivamente. Se han desarrollado múltiples metodologías para el pronóstico de estos índices, basadas generalmente en técnicas estadísticas, matemáticas o de

inteligencia computacional (Huang y Wu, 2008; Shen et al., 2011; Guresen et al., 2011).

Si bien es cierto que el tema ha tomado gran importancia, no existe información consolidada sobre las metodologías actuales y no existe claridad sobre cuál es el método de pronóstico más apropiado que genere resultados más precisos, según las características estadísticas que presenten las series de tiempo de los índices.

El objetivo principal de este trabajo es establecer el estado actual de las investigaciones sobre las metodologías de pronóstico de índices de mercado y documentar los resultados, a través de la respuesta de las siguientes preguntas de investigación:

RQ1: ¿Cuáles son las metodologías de pronóstico de índices de mercado comúnmente usadas?

RQ2: ¿Cuáles son las metodologías de pronóstico de la dirección de los cambios de índices de mercado comúnmente usadas?

RQ3: ¿Cuáles son las metodologías de pronóstico de la volatilidad de índice de mercado comúnmente usadas?

RQ4: ¿Cuáles son las metodologías más utilizadas para modelar índices de mercado?

RQ5: ¿Cuáles son las metodologías más utilizadas para modelar la volatilidad de índices de mercado?

RQ6: ¿Cuáles son los principales factores que afectan la precisión de los modelos de pronóstico?

Las preguntas de investigación RQ1, RQ2 y RQ3 tienen como fin mapear las principales metodologías enfocadas en el pronóstico del valor de los índices de mercado, de sus cambios de dirección y de la volatilidad de índices. Las preguntas

RQ4 y RQ5 tienen como fin establecer cuáles metodologías han sido comúnmente usadas para entender la evolución histórica (el modelado es diferente del pronóstico) de los índices y particularmente, de su volatilidad; esto tiene como fin, usualmente, el establecer la relevancia de diferentes variables que podrían afectar el mercado. La motivación para la pregunta RQ6 es determinar y consolidar cuáles serían los principales factores que podrían afectar los diferentes métodos de pronóstico con el fin de que dichos factores sean considerados en el desarrollo de nuevos métodos de modelado y de pronóstico que generen resultados más precisos.

Antes de dar respuesta a las preguntas anteriores, es importante establecer la diferencia entre los términos modelar y pronosticar. El modelado de series de tiempo se refiere al proceso metodológico de obtener el modelo de una serie de tiempo, tal que, dicho modelo capture todas las propiedades estadísticas de la serie estudiada y se cumplan los supuestos de los que se parte para formular dicho modelo; estos supuestos se verifican a través de las propiedades estadísticas de los residuos (tales como incorrelación, normalidad, varianza constante, ausencia de no linealidades remanentes, etc.) y de los parámetros del modelo (estadísticamente significativos, independientes del tiempo, etc.). Así, el modelo matemático se acepta únicamente si se comprueba estadísticamente que se cumplen todas las propiedades enunciadas. Un punto fundamental en el modelado es que se incorporan variables explicativas cuando éstas se encuentran disponibles. En el modelado, el modelo mismo es una teoría de cómo se comporta la serie de tiempo.

A diferencia del modelado, la predicción se centra en lograr la mayor precisión de los valores futuros, tal que el cumplir los supuestos estadísticos ya enunciados pierde importancia. En los modelos de predicción no se requieren fuertes supuestos del modelo y se puede asignar cualquier función sin una suposición a priori sobre las propiedades de los datos.

Para dar solución a las preguntas de investigación planteadas se utilizó la metodología de Revisión Sistemática de la Literatura adaptada por Kitchenham (2004). La principal finalidad de esta metodología es realizar una revisión metódica y rigurosa; es decir, adoptar un enfoque sistemático para evaluar y consolidar los resultados de la investigación de manera equilibrada y objetiva, con el fin de mapear el conocimiento existente a partir de las preguntas de investigación planteadas. Adicionalmente, le permite al investigador el encontrar vacíos actuales y temas de investigación futuros (Kitchenham, 2004; Kitchenham, 2007).

Uno de los resultados de la aplicación de la revisión sistemática al tema de investigación propuesto en este Trabajo Final de Maestría es que no se encontraron artículos previos con objetivos similares al propuesto.

Este trabajo se encuentra dirigido y es de interés principalmente para: profesionales, interesados en el pronóstico de índices de mercado; organizaciones e inversionistas, interesados en evitar o cubrir riesgos potenciales en el mercado; y en general, sirve como referencia para los investigadores y profesionales en sus estudios sobre temas financieros.

Este documento se encuentra organizado de la siguiente manera: en el Capítulo 2, se detalla la metodología de investigación utilizada; en el Capítulo 3, se desarrolla el protocolo de investigación; en el Capítulo 4, se presentan los resultados obtenidos; en el Capítulo 5, se solucionan las preguntas de investigación; y finalmente, en el Capítulo 6, se presentan las conclusiones.

2 METODOLOGÍA DE INVESTIGACIÓN

La metodología aplicada para dar respuesta a las preguntas de investigación planteadas es la revisión sistemática de literatura. A continuación se describe esta metodología.

Una revisión sistemática se define como una manera de identificar, evaluar e interpretar toda la investigación disponible relativa a una pregunta de investigación, área temática o fenómeno de interés (Kitchenham, 2004).

Debido a que la información relevante del tema de investigación está desagregada en diversas fuentes, se hace necesario agruparla en una fuente común, de manera completa, concisa y clara. El objetivo principal de utilizar esta metodología es resumir toda la información existente respecto al tema de investigación de manera rigurosa, imparcial y objetiva. Al mismo tiempo, encontrar vacíos en la investigación actual con el fin de determinar aspectos de investigación futuros (Kitchenham, 2004; Kitchenham, 2007).

La revisión sistemática se compone de tres etapas básicas: planificación de la revisión, desarrollo de la revisión, y publicación de resultados (Kitchenham, 2004). Estas están compuestas a su vez por sub-etapas que permiten guiar de mejor manera la revisión.

Si bien estas tres etapas que componen la metodología implican un mayor esfuerzo en comparación con las revisiones tradicionales, su principal ventaja

radica en que proporciona información sobre el tema de investigación a través de la mejor evidencia disponible en la literatura.

A continuación se presenta la descripción de las etapas y sub-etapas de la metodología de revisión sistemática de la literatura:

2.1 Planificación de la revisión

Esta etapa tiene como finalidad establecer los parámetros más importantes que deben ser considerados al realizar la revisión. Esto incluye la justificación de la revisión y la forma de realizar tanto la búsqueda como la revisión. Esta etapa se subdivide en las siguientes sub-etapas.

2.1.1 Identificación de la necesidad de revisión

La necesidad de una revisión sistemática surge de la necesidad de los investigadores para resumir la información existente sobre algún fenómeno de interés de manera exhaustiva e imparcial.

2.1.2 Definir un protocolo de revisión

El protocolo de revisión especifica los métodos que se utilizarán para llevar a cabo la revisión sistemática. En el protocolo se indican los antecedentes, las preguntas de investigación, la estrategia y los recursos que se utilizarán para realizar la búsqueda de las investigaciones, los criterios de inclusión y exclusión para la selección de las investigaciones, los criterios de calidad para la selección final de las investigaciones y la estrategia para la extracción de los datos.

En el protocolo de búsqueda se establecen y documentan las estrategias de búsqueda predefinidas, de manera que se pueda garantizar la exhaustividad de la misma, que se pueda encontrar la mayor cantidad de información relevante disponible y que los lectores puedan comprobar su rigor y exhaustividad.

Por su parte, los criterios de inclusión y exclusión definen y documentan los criterios que conforman la base para evaluar el potencial de cada estudio encontrado y seleccionar de manera objetiva los más adecuados.

2.2 Desarrollo de la revisión

Una vez se ha definido el protocolo para la investigación, es posible iniciar la revisión sistemática. Las sub-etapas se presentan a continuación.

2.2.1 Identificación de la investigación

El objetivo de una revisión sistemática es encontrar la mayor cantidad de información acerca de un tema de interés específico. En esta sub-etapa se contempla la generación de una estrategia de búsqueda, administración de la bibliografía y recuperación de documentos y documentación de la búsqueda con el fin de que sea transparente y pueda ser replicable.

2.2.2 Selección de estudios

Una vez obtenidas las investigaciones disponibles, es necesario evaluarlos y seleccionarlos. Es por esto que deben definirse criterios de inclusión y exclusión claros, que permitan clasificar las investigaciones o estudios de manera confiable.

2.2.3 Evaluación y calidad del estudio

Además de definir criterios de inclusión y exclusión, se considera importante definir criterios para evaluar la calidad de las investigaciones, con el fin de limitar los sesgos formarse una idea más precisa de las potenciales comparaciones y guiar la interpretación de los resultados. Y finalmente, seleccionar los artículos relevantes y que contengan la información necesaria para dar respuesta a las preguntas de investigación planteadas.

2.2.4 Extracción de los datos

Su objetivo es registrar con precisión la información de las investigaciones o estudios primarios. Para reducir la posibilidad de sesgo, es necesario que se defina la información y la forma de extracción de datos.

2.2.5 Síntesis de datos

Consiste en registrar la información extraída de los estudios primarios siguiendo una estrategia definida. Los datos pueden sintetizarse, considerando el enfoque que se pretende dar a la presentación del estado del arte o la identificación del fenómeno de interés.

2.3 Publicación de resultados

Esta etapa se refiere a la utilización de los resultados de la revisión sistemática. Es importante difundir los resultados obtenidos ya sea a través de la participación en conferencias, de la publicación de un artículo, entre otros.

3 PROTOCOLO DE INVESTIGACIÓN

A continuación se presenta el protocolo de investigación utilizado para realizar la revisión de la literatura:

3.1 Proceso de búsqueda

Como herramienta de búsqueda se utilizó el sistema de indexación Scopus. La búsqueda se realizó en abril de 2012 bajo las siguientes condiciones:

- *Áreas temáticas:* ciencias computacionales, economía, econometría, finanzas, matemáticas, física, negocios, administración y contabilidad.
- *Tipo de documentos:* artículos publicados en revistas científicas hasta diciembre de 2011.
- *Palabras clave:* (“forecast*” OR “predict*” OR “foresight” OR “modeling” OR “modeling”) AND (“market index*” OR “stock index*” OR “market indices” OR “stock indices”)

3.2 Criterios de inclusión o exclusión

Se realizó una revisión del título, resumen y palabras clave de los artículos recuperados en la búsqueda sistemática. Se incluyeron los artículos que cumplen, al menos, alguna de las siguientes condiciones:

- Pronostica un índice de mercado usando métodos formales como técnicas estadísticas, econométricas o de inteligencia artificial.
- Modela un índice de mercado.
- Describe factores que afectan la precisión de un modelo de pronóstico.

Se excluyeron los artículos que cumplen uno o más de los siguientes criterios:

- Pronostica el valor de una acción en particular.
- Su objetivo no es el modelamiento o el pronóstico de índices.
- Su alcance excluye el pronóstico o el modelado mediante metodologías estadísticas, econométricas o de inteligencia artificial.

3.3 Criterios de calidad

Partiendo de los artículos seleccionados según los criterios de inclusión y exclusión, se definieron los siguientes criterios de calidad con el fin de evaluar algunas características específicas de los artículos seleccionados que permitan medir su calidad:

- El método de predicción descrito en el artículo debe estar aplicado a un índice de mercado específico.
- El artículo debe contener un parámetro determinado que sirva para la evaluación del método de pronóstico propuesto

3.4 Selección de estudios y análisis de datos

Para cada artículo finalmente seleccionado se recuperó la siguiente información:

- Autores.
- Título.

- Año de publicación.
- Nombre de la revista científica.
- Número de citas.
- Resumen.
- Metodología de modelado o pronóstico utilizada.
- Índice de mercado donde se aplicó la metodología de pronóstico o modelado.
- Factor que afecta la precisión de un modelo de pronóstico.
- Parámetro para la evaluación de la metodología.

4 RESULTADOS OBTENIDOS

Usando la metodología de revisión y el protocolo de búsqueda descritos anteriormente, se recuperaron 150 artículos potenciales.

Después de aplicar los criterios de inclusión y exclusión se seleccionaron 64 artículos para dar respuesta a las preguntas de investigación planteadas. Los artículos seleccionados se presentan en la Tabla 1.

Tabla 1. Artículos seleccionados

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Effective lengths of intervals to improve forecasting in fuzzy time series	Huang, 2001	Fuzzy Sets and Systems	127	TAIEX	Diario	01/1996 - 01/1997	MSE	La longitud del intervalo afecta significativamente los resultados de la predicción de un índice de mercado cuando se usan modelos de series de tiempo difusas y se propone un método basado en distribución y en promedios de la longitud de la serie de tiempo para encontrar la longitud más eficiente del intervalo de tiempo de la serie de datos que debería considerarse para realizar la predicción.
Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high-frequency index returns	Blair, Poon y Taylor, 2001	Journal of Econometrics	110	S&P100	Diario	1987 - 1999	RMSE	Modelo ARCH que utiliza la variación de los rendimientos del índice de mercado durante el día.
Forecasting S&P 500 stock index futures with a hybrid AI system	Tsaih, Hsu y Lai, 1998	Decision Support Systems	81	S&P500	Diario	12/1983 - 12/1993	No está descrito	Modelo híbrido de redes neuronales artificiales que utiliza reglas definidas por expertos utilizadas para el entrenamiento de la red neuronal.
A type 2 fuzzy time series model for stock index forecasting	Huang y Yu, 2005	Physica A: Statistical Mechanics and its Applications	64	TAIEX	Diario	2000 - 2003	RMSE	Modelo de series de tiempo difusas tipo 2 en el que se incluyen observaciones adicionales para mejorar las relaciones difusas obtenidas de un modelo tipo 1 y así mejorar los resultados de pronóstico.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
The application of neural networks to forecast fuzzy time series	Huang y Yu, 2006	Physica A: Statistical Mechanics and its Applications	64	TAIEX	Diario	1991 - 2003	RMSE	Modelo básico con un enfoque de redes neuronales de alimentación hacia adelante y un modelo híbrido neuro-difuso con redes neuronales de propagación hacia atrás, que predice los patrones conocidos y desconocidos de la serie de tiempo del índice.
A refined fuzzy time-series model for forecasting	Yu, 2005	Physica A: Statistical Mechanics and its Applications	63	TAIEX y TAIFEX	Diario	01/2000 - 12/2000	RMSE	Modelo de series de tiempo difusas en el que se desarrolla el concepto de refinamiento como mecanismo para mejorar la precisión del modelo.
A bivariate fuzzy time series model to forecast the TAIEX	Yu y Huang, 2008	Expert Systems with Applications	40	TAIEX	Diario	01/1994 - 12/2004	RMSE	Modelos de redes neuronales de dos variables, se proponen tres modelos: modelo bivariente de redes neuronales, modelo bivariente de redes neuronales basado en un modelo de series de tiempo difusas y un método bivariente basado en un modelo de series de tiempo difusas con sustitutos.
Financial time series forecasting using independent component analysis and support vector regression	Lu, Lee y Chiu, 2009	Decision Support Systems	39	TAIEX	Diario	01/2003 - 02/2006	RMSE, NMSE, MAD, DS, CPY CD	Modelo de máquinas de soporte vectorial que utiliza un modelo de análisis de componentes independientes para seleccionar las entradas relevantes y mejorar los resultados del modelo.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network	Yudong y Lenan, 2009	Expert Systems with Applications	38	S&P500	Diario	10/1998 - 02/2008	RMSE	Modelo de redes neuronales artificiales con propagación hacia atrás, se utiliza quimiotactismo bacteriano para entrenar la red neuronal.
Flexible neural trees ensemble for stock index modeling	Chen, Yang y Abraham, 2007	Neurocomputing	25	NASDAQ-100 y NIFTY	Diario	01/1995 - 01/2002	RMSE, MAPE y MAP	Modelo de árboles neuronales flexibles para el modelado de índices de mercado, como parámetros de entrada se utilizan los precios de apertura, los precios de cierre y el precio máximo.
Multi-attribute fuzzy time series method based on fuzzy clustering	Cheng, Wang, 2008	Expert Systems with Applications	25	TAIEX	Diario	01/2004 - 12/2004	MSE y MAE	Modelo de series de tiempo difusas de múltiples atributos que aplica el algoritmo de agrupamiento Fuzzy C Mean para particionar los datos y para modificar los pronósticos lingüísticos, con el fin de mejorar la exactitud del modelo, el atributo principal lo constituye el precio de cierre diario del índice y los atributos secundarios se constituyen por el precio diario más alto y el precio diario más bajo del índice
A new approach of bivariate fuzzy time series analysis to the forecasting of a stock index	Hsu, Tse y Wu, 2003	International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems	24	TAIEX	Diario	12/2000 - 02/2001	MSE	Modelo de series de tiempo difusas bivariado.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques	Majhi, Panda, Majhi y Sahoo, 2009	Expert Systems with Applications	24	S&P500	Diario	01/1994 - 10/2006	RMSE	Modelo basado en optimización de forrajeo de bacterias BFO y optimización adaptativa de forrajeo de bacterias ABFO.
Fuzzy dual-factor time-series for stock index forecasting	Chu, Chen, Cheng y Huang, 2009	Expert Systems with Applications	20	TAIEX y NASDAQ	Diario	01/2004 - 12/2004	RMSE	Modelo de series de tiempo difusas de factor dual modificado que utiliza como entradas tanto el conjunto de datos del índice y como de los volúmenes transados para predecir el índice; adicionalmente, utiliza dos enfoques de predicción para procesar los dos factores por separado y una ecuación de doble factor para generar el pronóstico.
High-order fuzzy time-series based on multi-period adaptation model for forecasting stock markets	Chen, Cheng y Teoh, 2008	Physica A: Statistical Mechanics and its Applications	19	TAIEX	Diario	1991 - 1999	RMSE	Modelo de series de tiempo difusas en el cual se extraen las relaciones lógicas de alto orden para construir reglas que permitan pronosticar los datos y aumentar la capacidad de predicción del modelo.
Integrating GA-based time-scale feature extractions with SVMs for stock index forecasting	Huang y Wu, 2008	Expert Systems with Applications	15	CAC40, FTSE100, DAX30, MIB40, TSX60, S&P500 y NK225	Diario	01/2003 - 12/2004	RMSE, MSE, MAE y MAPE	Modelo híbrido de máquinas de soporte vectorial cuyas entradas se descomponen usando wavelets y se seleccionan usando algoritmos genéticos.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Using support vector machine with a hybrid feature selection method to the stock trend prediction	Lee, 2009	Expert Systems with Applications	15	NASDAQ	Diario	08/2001 - 08/2007	MSE	Modelo de máquinas de soporte vectorial que utiliza un método híbrido de selección de entradas relevantes con el fin de mejorar el pronóstico del modelo: se aplica el método F-Score para filtrar las entradas y el método F-SSFS para seleccionar las características óptimas de los datos de entrada.
Financial time-series analysis with rough sets	Yao y Herbert, 2009	Applied Soft Computing Journal	15	NZX	Diario	07/1991 - 04/2000	Diferencia entre el índice real y la proyección	Modelo de conjuntos rugosos.
Predictable variation and profitable trading of US equities: A trading simulation using neural networks	Motiwalla y Wahab, 2000	Computers and Operations Research	13	Nasdaq Composite, Nasdaq Industrial, New York Stock Exchange Composite, S&P100, Russell 300, Russell large cap 1000, S&P Mid-Cap, Value Line, Wilshire 5000 y S&P 500	Diario	01/1990 - 08/1998	Diferencia del índice real y el pronóstico	Modelo de redes neuronales artificiales perceptrón multicapa y se compara con un modelo paramétrico de regresión no lineal.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Periodo de datos utilizados	Parámetro para evaluar la metodología	Metodología
Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods	de Faria, Albuquerque, Gonzalez, Cavalcante y Albuquerque, 2009	Expert Systems with Applications	12	Ibovespa	Diario	09/1998 - 04/2007	RMSE	Se presentan dos modelos para el pronóstico de la dirección de los cambios de índices de mercado: un modelo basado en redes neuronales perceptrón multicapa y un modelo adaptativo de suavizado exponencial en el que los parámetros del modelo se van actualizando a lo largo de la predicción.
Volatility model based on multi-stock index for TAIEX forecasting	Cheng y Wei, 2009	Expert Systems with Applications	10	TAIEX	Diario	1997 - 2003	RMSE	Modelo adaptativo multivariable de pronóstico que utiliza la volatilidad de otros índices para predecir la volatilidad del índice objetivo; adicionalmente, utiliza técnicas de aprendizaje adaptativo para adaptar los parámetros lineales del modelo propuesto con el fin de alcanzar un rendimiento óptimo.
Classifying trend movements in the msci U.S.A.capital market index-A comparison of regression, arima and neural network methods	Wood y Dasgupta, 1996	Computers and Operations Research	8	MSCI	Diario	02/1982 - 11/1993	RMSE	Modelo de redes neuronales multi-variable (parte de la relación del índice de mercado objetivo con índices de otras bolsas de valores) con funciones de activación no lineales.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Development and performance evaluation of FLANN based model for forecasting of stock markets	Majhi, Panda y Sahoo, 2009	Expert Systems with Applications	7	Dow Jones y S&P500	Diario y mensual	01/1994 - 10/2006	MAPE	Modelo de redes neuronales con una sola neurona y una sola capa que utiliza una función trigonométrica y mínimos cuadrados con el fin de encontrar los pesos óptimos para formar el modelo.
Forecasting TAIEX based on fuzzy time series and particle swarm optimization	Kuo, Horng, Chen, Run, Kao, Chen, Lai y Lin, 2010	Expert Systems with Applications	6	TAIEX	Diario	08/1998 - 08/1998	MSE	Modelo híbrido de pronóstico de índices de mercado basado en modelos de series de tiempo difusas y optimización evolutiva por enjambre de partículas utilizada para entrenar las reglas difusas basada en los datos de entrada con el fin de mejorar la precisión del modelo.
Forecasting Taiwan's major stock indices by the Nash nonlinear grey Bernoulli model	Chen, Hsin y Wu, 2010	Expert Systems with Applications	5	TAIEX	Mensual	01/2008 - 09/2008	ARPE	Modelo de caja gris de Bernoulli compuesto por una ecuación diferencial no lineal de potencia n determinada por un programa iterativo que calcula el ARPE mínimo, se incorpora el concepto de equilibrio de Nash para agregar un parámetro adicional a la ecuación con el fin de mejorar la precisión del modelo.
Using artificial neural network models in stock market index prediction	Guresen, Kayakutlu y Daim, 2011	Expert Systems with Applications	5	NASDAQ	Diario	07/2008 - 06/2009	MSE y MAD	Evaluación de los modelos percepción multicapa, red neuronal artificial dinámica DAN2 y redes neuronales híbridas que utilizan modelos GARCH para seleccionar las entradas relevantes.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Forecasting the turning time of stock market based on Markov-Fourier grey model	Hsu, Liu, Yeh y Hung, 2009	Expert Systems with Applications	5	TAIEX	Diario	10/1996 - 08/2007	MRE	Método basado en teoría de los modelos de caja gris, se aplican series de Fourier para afinar los errores residuales producidos por el modelo y se utiliza el estado de Markov para producir los resultados intermedios generados por el modelo.
The impact of information timeliness on the predictability of stock and futures returns: An application of vector models	Ostermark y Hernesnie mi, 1995	European Journal of Operational Research	5	FOX	Diario	05/1988 - 08/1991	MAPE	Modelo basado en vectores autorregresivos con media móvil y variables exógenas VARMAX.
Information fusion in the context of stock index prediction	Siekman, Kruse, Gebhardt, Van, Overbeek y Cooke, 2001	International Journal of Intelligent Systems	5	DAX	Mensual	10/1998 - 10/1998	RMSE	Modelo neuro-difuso con una red neuronal de cuatro capas, dos de ellas son ocultas en las cuales se encuentra un motor de reglas basadas en el conocimiento de expertos que tienen un peso importante en la predicción del índice de mercado.
A neural network-based fuzzy time series model to improve forecasting	Yu y Huang, 2010	Expert Systems with Applications	5	TAIEX	Diario	2000 - 2004	RMSE	Modelo de series de tiempo difusas que aplica redes neuronales para el pronóstico de índices de mercado.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Predicting DAX trends from Dow Jones data by methods of the mathematical theory of democracy	Tangian, 2008	European Journal of Operational Research	4	DAX	Diario	No se indica	MSE	Modelo que aplica la teoría matemática de la democracia, la cual funciona según los índices de popularidad y universalidad que se utilizan para encontrar representantes socialmente óptimos; matemáticamente, el objetivo es encontrar datos óptimos que puedan representar el comportamiento de la serie de datos del índice.
Sequential combination of statistics, econometrics and Adaptive Neural-Fuzzy Interface for stock market prediction	Ansari, Kumar, Shukla, Dhar y Tiwari, 2010	Expert Systems with Applications	3	NASDAQ	Diario	12/2006 - 12/2008	MSE	Método híbrido que emplea análisis estadístico y econométrico para comprender la naturaleza exacta del conjunto de datos y la relación entre ellos con el fin de optimizar los parámetros de pronóstico, luego agrupa los datos y se emplea un modelo ANFIS para crear funciones de membresía difusas y entrenar la red neuronal.
TAIEX forecasting based on fuzzy time series and fuzzy variation groups	Chen y Chen, 2011	IEEE Transactions on Fuzzy Systems	3	TAIEX	Diario	01/1999 - 12/2004	RMSE	Método multivariante basado en modelos de series de tiempo difusas, se utilizan grupos de variación difusos con el fin de analizar en detalle los datos de entrenamiento para el pronóstico, en éstos el factor principal lo constituye el índice y los factores secundarios otros índices de mercado.
Fusion ANFIS models based on multi-stock volatility causality for TAIEX forecasting	Cheng, Wei y Chen, 2009	Neurocomputing	3	TAIEX	Diario	1997 - 2003	RMSE	Método ANFIS para la predicción de índices de mercado basado en la volatilidad de otros índices accionarios.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Wavelet-based detection of outliers in financial time series	Gran y Veiga, 2010	Computational Statistics and Data Analysis	3	Dow Jones, FTSE-100 y S&P500	No aplica	No se indica	No está descrito	Los valores extremos de las series de tiempo y la longitud de los intervalos usados para la predicción afectan los resultados de los modelos de pronóstico de índices bursátiles. Se propone un método basado en la transformada wavelet, para la detección de valores extremos en series de tiempo financieras y poder seleccionar las entradas que son realmente relevantes.
Examining the volatility of Taiwan Stock Index returns via a three-volatility-regime Markov-switching ARCH model	Li y Lin, 2003	Review of Quantitative Finance and Accounting	3	TAIEX	Diario	02/1981 - 09/1998	MSE, MAE y LE	Comparación del modelo SWARCH con los modelos convencionales ARCH y GARCH.
Forecasting S&P-100 stock index volatility: The role of volatility asymmetry and distributional assumption in GARCH models	Liu y Hung, 2010	Expert Systems with Applications	3	S&P100	Diario	10/1997 - 09/2003	MSE	Comparación de los modelos GJR.GARCH y EGARCH bajo diferentes distribuciones de error.
Adaptive-expectation based multi-attribute FTS model for forecasting TAIEX	Liu, Chen, Cheng y Chen, 2010	Computers and Mathematics with Applications	3	TAIEX	Diario	01/2000 - 12/2003	MAER	Modelo de series de tiempo difusas de múltiples atributos que aplica el algoritmo de agrupamiento Fuzzy C Mean para particionar los datos y para modificar los pronósticos lingüísticos, con el fin de mejorar la exactitud del modelo.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Periodo de datos utilizados	Parámetro para evaluar la metodología	Metodología
Modeling and forecasting volatility of the Malaysian and the Singaporean stock indices using asymmetric GARCH models and non-normal densities	Mohd y Shamiri, 2007	Malaysian Journal of Mathematical Sciences	3	KLCI e ITS	Diario	01/1991 - 12/2004	MSE, MAE, MAPE, TIC, AIC, BIC, APC, AJDR y AR	Comparación de los modelos GJR.GARCH y EGARCH bajo diferentes distribuciones de error.
Modeling and forecasting volatility of the Malaysian stock markets	Shamir e Isa, 2009	Journal of Mathematical Statistics	3	KLCI	Diario	01/1998 - 12/2008	MSE, MDSE, RMSE y MAE	Comparación de los modelos GARCH, EGARCH y NAGARCH bajo diferentes distribuciones de error.
Integrating ensemble of intelligent systems for modeling stock indices	Abraham y AuYeung, 2003	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	2	NASDAQ-100 y S&P CNX NIFTY	Diario	01/1995 - 01/2002	RMSE, MAP y MAPE	Se propone un modelo para modelar índices de mercado a partir de la integración de cuatro paradigmas inteligentes diferentes: redes neuronales, máquinas de soporte vectorial, sistema neuro-difuso y redes neuronales probabilísticas, ensamblados mediante dos enfoques: algoritmos genéticos e integración directa.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
A hybrid ANFIS model based on AR and volatility for TAIEX forecasting	Chang, Wei y Cheng, 2011	Applied Soft Computing Journal	2	TAIEX	Diario	1997 - 2003	RMSE	ANFIS
Fuzzy forecasting based on fuzzy-trend logical relationship groups	Chen y Wang, 2010	IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics	2	TAIEX	Diario	04/1992 - 10/1992	RMSE	Modelo de series de tiempo difusas en el cual se extraen las relaciones lógicas para construir las reglas que permitan pronosticar los datos y utiliza un algoritmo de agrupamiento automático con el fin de clasificar los datos históricos en intervalos de longitudes diferentes para encontrar la longitud del intervalo óptima.
Multi-variable fuzzy forecasting based on fuzzy clustering and fuzzy rule interpolation techniques	Shyi-Ming y Yu-Chuan, 2010	Information Sciences	2	TAIEX	Diario	11/2004 - 12/2004	RMSE	Modelo multivariable de pronóstico de series de tiempo difusas que aplica el algoritmo de agrupamiento Fuzzy C Mean para particionar los datos y para modificar los pronósticos lingüísticos, y se aplican técnicas de reglas de interpolación difusas para determinar el peso de cada regla con respecto a las observaciones de entrada, con el fin de mejorar la precisión del modelo; adicionalmente, se utiliza información de otros índices de mercado para pronosticar el índice objetivo.
A heuristic time-invariant model for fuzzy time series forecasting	Bai, Wong, Chu, Xia y Pan, 2011	Expert Systems with Applications	1	TAIFEX	Diario	1971 - 1989	MSE	Modelo de series de tiempo difusas que incluye un factor predictor para indicar la tendencia de aumento o disminución de la serie de tiempo del índice con el fin de capturar las tendencias de las observaciones con mayor precisión.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Periodo de datos utilizados	Parámetro para evaluar la metodología	Metodología
Deterministic regression model and visual basic code for optimal forecasting of financial time series	Balbas, Galperin, Galperin, 2008	Computers and Mathematical Applications	1	S&P500	Diario	01/2006 - 03/2008	Error absoluto y error relativo	Modelo de regresión determinista que incorpora ecuaciones diferenciales aplicadas a las series de tiempo del índice de mercado con el fin de proporcionar un patrón de comportamiento mucho más amplio que los métodos estadísticos o econométricos existentes, incorporando crecimientos polinomiales, exponenciales y logarítmicos que reflejan las tendencias de los datos; y aplica un algoritmo de regresión determinista que permite estimar los valores futuros de la serie de tiempo del índice.
Fuzzy adaptive decision-making for boundedly rational traders in speculative stock markets	Bekiros, 2010	European Journal of Operational Research	1	S&P500, KLCI, HANGSENG, SET100, NYSE, FTSE100, CAC40, Stock Exchange Weighted, Jakarta Stock Exchange Composite Y Straits Times	Diario	01/1990 - 03/2001	RMSE	Sistema híbrido neuro-difuso basado en la actualización dinámica de la adaptación y calibración de las reglas heurísticas difusas de aprendizaje.

Tabla 1. Continuación

Título	Estudio	Revista	#Citas	Índice	Unidad de medida	Periodo de datos utilizados	Parámetro para evaluar la metodología	Metodología
Neuro-genetic system for stock index prediction	Mandziuk y Jaruszewicz, 2011	Journal of Intelligent and Fuzzy Systems	1	DAX	Diario	04/2004 - 08/2004	MSE	Sistema neuro-genético con redes neuronales de alimentación hacia adelante con una capa oculta y algoritmos genéticos para seleccionar de manera eficaz las variables de entrada, con el fin de mejorar los resultados de pronóstico del modelo
A note on the predictability of Finnish stock market returns: Evidence from stock index futures markets	Martikainen y Puttonen, 1994	European Journal of Operational Research	1	FOX	Diario	05/1998 - 03/1990	MSE	Modelo GARCH para el pronóstico de volatilidad de índices de mercado a partir de series de tiempo de índices de futuros sobre acciones relacionados.
Forecasting changes in Korea Composite Stock Price Index (KOSPI) using association rules	Na y Sohn, 2011	Expert Systems with Applications	1	KOSPI	Diario	01/2006 - 12/2008	MSE	Modelo basado en técnicas de minería de datos para encontrar reglas de asociación para la predicción de la dirección de los cambios de índices de mercado, se utilizan series de tiempo de otros índices para predecir el índice objetivo
A generalized method for forecasting based on fuzzy time series	Qiu, Liu y Li, 2011	Expert Systems with Applications	1	Shanghai stock index	Diario	01/1997 - 12/2006	MSE	Modelo de series de tiempo difusas que aplica técnicas de ensamble con el fin de mejorar la selección de las entradas relevantes.
Shanghai composite index forecasting based on multi-stock index	Sun y Bao, 2010	Journal of Computational Information Systems	1	Shanghai Composite index	Diario	04/2009 - 04/2010	RMSE	Método que utiliza regresión lineal y combina varios factores para el pronóstico.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Forecasting stock index volatility	Bramante y Luigi, 2001	Applied Stochastic Models in Business and Industry	0	MIB30	Diario	01/1997 - 04/1999	MSE	Modelo ARCH que utiliza los rendimientos históricos de las acciones que componen el índice de mercado con el fin de pronosticar su volatilidad.
Fuzzy forecasting based on high-order fuzzy logical relationships and automatic clustering techniques	Chen y Tanuwijaya, 2011	Expert Systems with Applications	0	TAIFEX	Diario	08/1998 - 09/1998	MSE	Modelo de series de tiempo difusas en el cual se extraen las relaciones lógicas de alto orden para construir reglas que permitan pronosticar los datos y aumentar la precisión del modelo; adicionalmente, se utilizan técnicas automáticas de agrupamiento con el fin encontrar el intervalo óptimo de la serie de tiempo que se debe utilizar para el pronóstico y así mejorar la capacidad de predicción
Applied rough set logics for multi-criteria decision analysis in stock index volatility projection	Ho, 2011	Review of Pacific Basin Financial Markets and Policies	0	TAIEX	Diario	01/1999 - 12/2008	MSE	Modelo de conjuntos rugosos que aplica algoritmos de agrupamiento automático para seleccionar las entradas relevantes con el fin de mejorar los resultados del modelo.
A hybrid procedure for stock price prediction by integrating self-organizing map and genetic programming	Hsu, 2011	Expert Systems with Applications	0	TAIEX	Diario	01/1996 - 09/2009	RMSE, MAE y MAPE	Modelo híbrido de redes neuronales de mapas auto-organizados SOM que utiliza programación genética para optimizar los pesos de las conexiones de la red, con el fin de lograr predecir con mayor precisión los índices de mercado.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Periodo de datos utilizados	Parámetro para evaluar la metodología	Metodología
Forecasting stock indices with wavelet domain kernel partial least square regressions	Huang, 2011	Applied Soft Computing Journal	0	NASDAQ, NK225, TWSI y KOSPI	Diario	01/2003 - 12/2004	MSE, RMSE, MAE y MAPE	Modelo de pronóstico que aplica wavelets con el fin de identificar las características de la serie de tiempo del índice y se emplea regresión por mínimos cuadrados parciales para crear un sub-espacio más eficiente que mantiene la máxima covarianza entre entradas y salidas para llevar a cabo el pronóstico final.
An efficient CMAC neural network for stock index forecasting	Lu y Wu, 2011	Expert Systems with Applications	0	Nikkei 225 y TAIEX	Diario	12/2004 - 12/2008	RMSE	Modelo de redes neuronales de articulación cerebelosa.
Automatic stock market forecasting system based on extended language template model	Nan y Sun, 2011	Journal of Information and Computational Science	0	CSI 300	Diario	08/2006 - 11/2008	MSE	Modelo de redes neuronales y sistema de lenguaje natural con el fin de expresar la información utilizada para el pronóstico en lenguaje natural para usuarios no expertos.
S&P 500 index direction forecasting from 1976 to 2010: A fuzzy system approach	Souto-Maior, Borba y da Costa, 2011	International Journal of Digital Accounting Research	0	S&P500	Diario	01/1970 - 12/2010	MSE	Modelo de series de tiempo difusas en el que las reglas del motor de inferencia fueron definidas teniendo en cuenta el comportamiento histórico del índice.
A self-organized neuro-fuzzy system for stock market dynamics modeling and forecasting	Su, Chen y Yang, 2010	WSEAS Transactions on Information Science and Applications	0	TAIEX	Diario	04/1997 - 11/2008	RMSE	Modelo neuro-difuso con redes neuronales de mapas auto-organizados SOM de cinco capas mediante el uso de indicadores técnicos y se propone un modelo adicional de media móvil para seleccionar el conjunto de entrada para el modelo neuro-difuso.

Tabla 1. Continuación

Título	Estudio	Revista	#Citaciones	Índice	Unidad de medida	Período de datos utilizados	Parámetro para evaluar la metodología	Metodología
Stock market volatility and the forecasting performance of stock index futures	Wang, 2009	Journal of Forecasting	0	TAIFEX y SGX MSCI	Diario	01/1997 - 07/1998	MPE y MAPE	Modelo basado en el concepto de equilibrio general que utiliza un estimador ponderado de media móvil exponencial con el fin de mejorar la capacidad de pronóstico del modelo.
Forecasting stock indices with back propagation neural network	Wang, Wang, Zhang y Guo, 2011	Expert Systems with Applications	0	Shanghai Composite index	Diario	01/1993 - 12/2009	MAE, RMSE y MAPE	Modelo de redes neuronales de propagación hacia atrás cuyas entradas se descomponen usando wavelets.
Adaptive time-variant models for fuzzy-time-series forecasting	Wong, Bai y Chu, 2010	IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics	0	TAIEX	Diario	1971 - 2000	RMSE, AFER	Modelo de series de tiempo difusas adaptativo variante en el tiempo, el cual mediante el modelo de Wagner adapta automáticamente el tamaño de la serie de datos para especificar el modelo de pronóstico.

En la Figura 1 se muestra el número total de artículos recuperados por año de publicación. Después del 2008 inclusive, se ha publicado aproximadamente el 74% del total de artículos. Y en promedio, desde ese mismo período se publican aproximadamente 12 artículos por año.

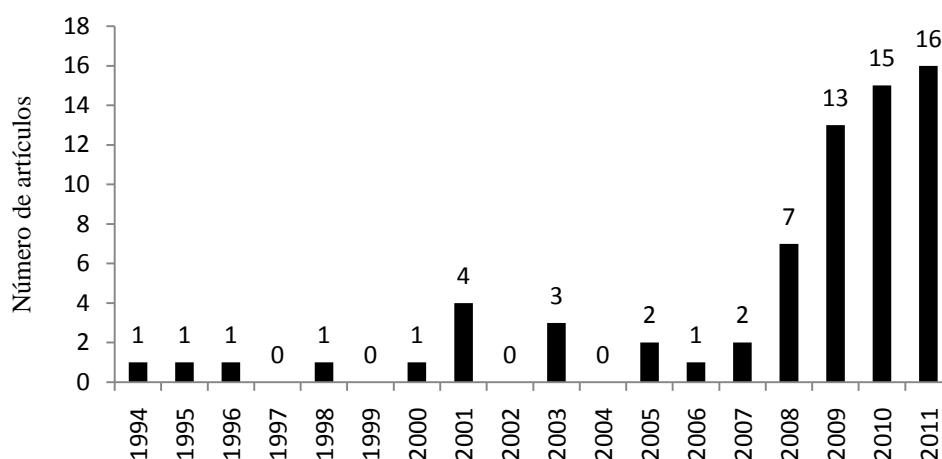


Figura 1 Número de artículos por año de publicación

Por otro lado, la Figura 2 presenta el número de citas que han tenido los artículos por año de publicación. Los períodos con mayor número de citas son 2001, 2005, 2008 y 2009, con el 70% de las citas totales aproximadamente, lo que demuestra que la mayoría de las investigaciones más importantes se han llevado a cabo durante estos períodos.

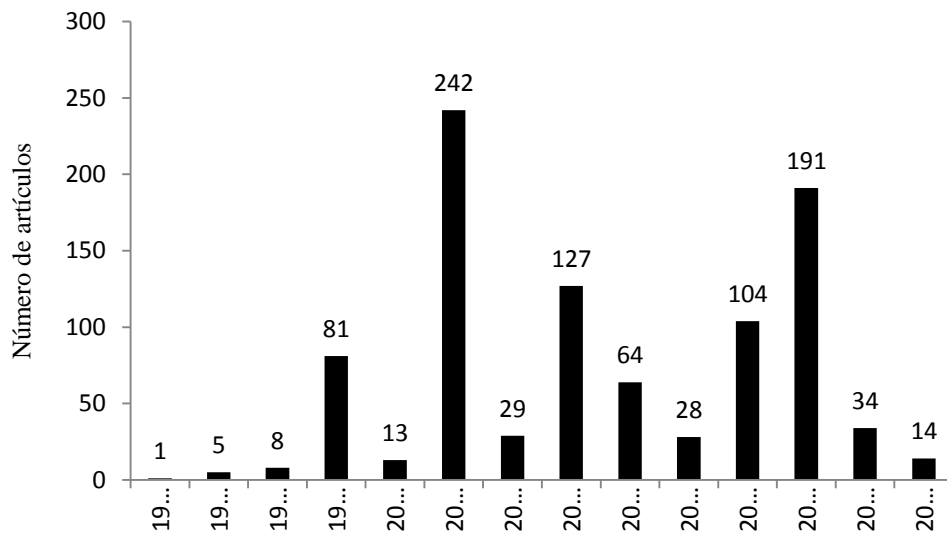


Figura 2. Número de citaciones de artículos por año de publicación

La Figura 1 y la Figura 2 demuestran el crecimiento y la importancia que ha adquirido el tema de pronóstico de índices de mercado durante los últimos años.

Por último, la Figura 3 muestra la proporción de artículos que aplican una metodología de pronóstico por cada índice de mercado. Los resultados muestran que los índices con mayores reportes de investigaciones son: TAIEX, NASDAQ y S&P500.

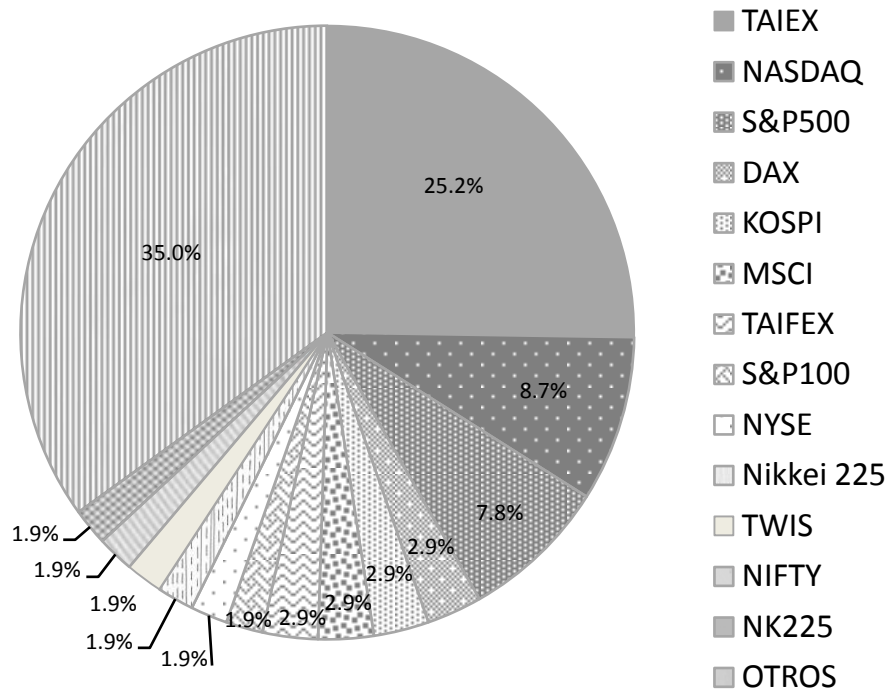


Figura 3. Proporción de artículos publicados por índice de mercado

Los resultados mostrados en las figuras anteriores, evidencian la necesidad de establecer un marco común y consolidado acerca de todas las metodologías de pronóstico de índices de mercado que se han definido.

Por otro lado, al evaluar los criterios definidos para evaluar la calidad de los artículos seleccionados, se tiene como resultado que todos los artículos predicen o modelan un índice de mercado específico, tal como lo muestra la Figura 3. Adicionalmente, el 97% de los artículos utilizan un parámetro para evaluar la metodología propuesta. La Figura 4 muestra los diferentes parámetros empleados para esta evaluación.

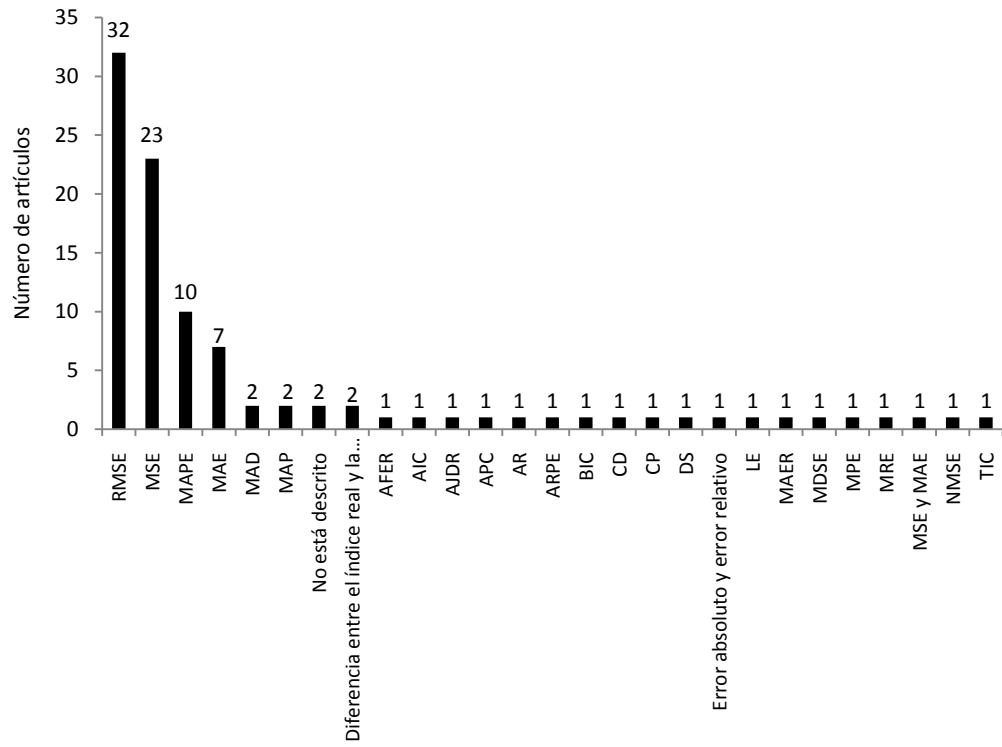


Figura 4. Parámetro para la evaluación del método

5 DISCUSIÓN

En esta sección se da respuesta a las preguntas de investigación. La información detallada de la contribución de cada trabajo finalmente seleccionado se presenta en el Anexo 1.

5.1 RQ1 ¿Cuáles son las metodologías de pronóstico de índices de mercado comúnmente usadas?

Según la literatura analizada sobre el pronóstico de índices de mercado, se puede concluir que no hay una técnica dominante y los estudios se caracterizan generalmente por:

- Usar varios métodos para el pronóstico y seleccionar el que genere mayor precisión en los resultados medidos a partir de las comparaciones del error de pronóstico en cada uno de ellos.
- Comparar modelos híbridos contra técnicas clásicas con el fin de determinar si los modelos híbridos contribuyen al mejoramiento de la precisión de los resultados de los modelos.

Los métodos revisados se presentan a continuación:

- (1) ANFIS: método híbrido que emplea análisis estadístico y econométrico para comprender la naturaleza exacta del conjunto de datos y la relación entre ellos con el fin de optimizar los parámetros de pronóstico, luego agrupa los datos y se emplea un modelo ANFIS para crear funciones de pertenencia difusas y entrenar la red neuronal (Ansari, Kumar, Shukla,

Dhar y Tiwari, 2010); método ANFIS para la predicción de índices de mercado basado en la volatilidad de otros índices accionarios (Cheng, Wei y Chen, 2009).

- (2) Método que utiliza regresión lineal y combina varios factores para el pronóstico de índices de mercado (Sun y Bao, 2010).
- (3) Modelos de series de tiempo difusas: modelo de series de tiempo difusas que incluye un factor predictor para indicar la tendencia de aumento o disminución de la serie de tiempo del índice con el fin de capturar las tendencias de las observaciones con mayor precisión (Bai, Wong, Chu, Xia y Pan, 2011);), método multivariate basado en modelos de series de tiempo difusas, se utilizan grupos de variación difusos con el fin de analizar en detalle los datos de entrenamiento para el pronóstico, en éstos el factor principal lo constituye el índice y los factores secundarios otros índices de mercado (Chen y Chen, 2011); modelo de series de tiempo difusas en el cual se extraen las relaciones lógicas de alto orden para construir reglas que permitan pronosticar los datos y aumentar la capacidad de predicción del modelo (Chen, Cheng y Teoh, 2008); modelo de series de tiempo difusas de múltiples atributos que aplica el algoritmo de agrupamiento Fuzzy C Mean para particionar los datos y para modificar los pronósticos lingüísticos, con el fin de mejorar la exactitud del modelo, el atributo principal lo constituye el precio de cierre diario del índice y los atributos secundarios se constituyen por el precio diario más alto y el precio diario más bajo del índice (Cheng, Cheng y Wang, 2008); modelo de series de tiempo difusas en el cual se extraen las relaciones lógicas de alto orden para construir reglas que permitan pronosticar los datos y aumentar la precisión del modelo; adicionalmente, se utilizan técnicas automáticas de agrupamiento con el fin encontrar el intervalo óptimo de la serie de tiempo que se debe utilizar para el pronóstico y así mejorar la capacidad de predicción (Chen y Tanuwijaya, 2011); modelo de series de tiempo difusas en el cual se extraen las relaciones lógicas para construir

las reglas que permitan pronosticar los datos y utiliza un algoritmo de agrupamiento automático con el fin de clasificar los datos históricos en intervalos de longitudes diferentes para encontrar la longitud del intervalo óptima (Chen y Wang, 2010); modelo de series de tiempo difusas de factor dual modificado que utiliza como entradas tanto el conjunto de datos del índice y como de los volúmenes transados para predecir el índice; adicionalmente, utiliza dos enfoques de predicción para procesar los dos factores por separado y una ecuación de doble factor para generar el pronóstico (Chu, Chen, Cheng y Huang, 2009); modelo bivariado de series de tiempo difusas (Hsu, Tse y Wu, 2003); modelo de series de tiempo difusas tipo 2 en el que se incluyen observaciones adicionales para mejorar las relaciones difusas obtenidas de un modelo tipo 1 y así mejorar los resultados de pronóstico (Huarng y Yu, 2005); modelo híbrido de pronóstico de índices de mercado basado en modelos de series de tiempo difusas y optimización evolutiva por enjambre de partículas utilizada para entrenar las reglas difusas basada en los datos de entrada con el fin de mejorar la precisión del modelo (Kuo, Horng, Chen, Run, Kao, Chen, Lai y Lin, 2010); modelo de series de tiempo difusas de múltiples atributos que aplica el algoritmo de agrupamiento Fuzzy C Mean para particionar los datos y para modificar los pronósticos lingüísticos, con el fin de mejorar la exactitud del modelo (Liu, Chen, Cheng y Chen, 2010); modelo multivariable de pronóstico de series de tiempo difusas que aplica el algoritmo de agrupamiento Fuzzy C Mean para particionar los datos y para modificar los pronósticos lingüísticos, y se aplican técnicas de reglas de interpolación difusas para determinar el peso de cada regla con respecto a las observaciones de entrada, con el fin de mejorar la precisión del modelo; adicionalmente, se utiliza información de otros índices de mercado para pronosticar el índice objetivo (Shyi-Ming y Yu-Chuan, 2010); modelo de series de tiempo difusas que aplica técnicas de ensamble con el fin de mejorar la selección de las entradas relevantes (Qiu, Liu y Li,

2011), modelo de series de tiempo difusas adaptativo variante en el tiempo, el cual mediante el modelo de Wagner adapta automáticamente el tamaño de la serie de datos para especificar el modelo de pronóstico (Wong, Bai, Chu, 2010), modelo de series de tiempo difusas en el que se desarrolla el concepto de refinamiento como mecanismo para mejorar la precisión del pronóstico (Yu, 2005); modelo de series de tiempo difusas que aplica redes neuronales para el pronóstico de índices de mercado (Yu y Huarng, 2010).

- (4) Modelos neuro-difusos: sistema neuro-genético con redes neuronales de alimentación hacia adelante con una capa oculta y algoritmos genéticos para seleccionar de manera eficaz las variables de entrada, con el fin de mejorar los resultados de pronóstico del modelo (Mandziuk y Jaruszewicz, 2011); modelo neuro-difuso con una red neuronal de cuatro capas, dos de ellas son ocultas en las cuales se encuentra un motor de reglas basadas en el conocimiento de expertos que tienen un peso importante en la predicción del índice de mercado (Siekman, Kruse, Gebhardt, Van Overbeek y Cooke, 2001); modelo neuro-difuso con redes neuronales de mapas auto-organizados SOM de cinco capas y se propone un modelo adicional de media móvil para seleccionar el conjunto de entrada para el modelo neuro-difuso (Su, Chen y Yang, 2010).
- (5) Máquinas de soporte vectorial: modelo de máquinas de soporte vectorial que utiliza un modelo de análisis de componentes independientes para seleccionar las entradas relevantes y mejorar los resultados del modelo (Lu, Lee y Chiu, 2009); modelo híbrido de máquinas de soporte vectorial cuyas entradas se descomponen usando wavelets y se seleccionan usando algoritmos genéticos (Huang y Wu, 2008).
- (6) Modelo basado en optimización de forrajeo de bacterias BFO y optimización adaptativa de forrajeo de bacterias ABFO (Majhi, Panda, Majhi y Sahoo, 2009).

- (7) Modelos de redes neuronales artificiales: evaluación de los modelos perceptrón multicapa, red neuronal artificial dinámica DAN2 y redes neuronales híbridas que utilizan modelos GARCH para seleccionar las entradas relevantes (Guresen, Kayakutlu y Daim, 2011), modelo híbrido de redes neuronales de mapas auto-organizados SOM que utiliza programación genética para optimizar los pesos de las conexiones de la red, con el fin de lograr predecir con mayor precisión los índices de mercado (Hsu, 2011); modelo básico con un enfoque de redes neuronales de alimentación hacia adelante y modelo híbrido neuro-difuso con redes neuronales de propagación hacia atrás, que predice los patrones conocidos y desconocidos de la serie de tiempo del índice (Huarng y Yu, 2006); modelo de redes neuronales de articulación cerebelosa (Lu y Wu, 2011), modelo de redes neuronales con una sola neurona y una sola capa que utiliza una función de trigonométrica (Majhi, Panda, Majhi y Sahoo, 2009); modelo de redes neuronales artificiales perceptrón multicapa y se compara con un modelo paramétrico de regresión no lineal (Motiwalla y Wahab, 2000), modelo de redes neuronales que utiliza un sistema de lenguaje natural con el fin de expresar la información utilizada para el pronóstico en lenguaje natural para usuarios no expertos (Nan y Sun, 2011); modelo de redes neuronales de propagación hacia atrás cuyas entradas se descomponen usando wavelets (Wang, Wang, Zhang y Guo, 2011); modelo de redes neuronales multi-variable (parte de la relación del índice de mercado objetivo con índices de otras bolsas de valores) con funciones de activación no lineales (Wood y Dasgupta, 1996); modelo de redes neuronales artificiales con propagación hacia atrás, se utiliza quimiotactismo bacteriano para entrenar la red neuronal (Yudong y Lenan, 2009); y finalmente, modelos de redes neuronales de dos variables, se proponen tres modelos: modelo bivalente de redes neuronales, modelo bivalente de redes neuronales basado en un modelo de series de tiempo

difusas y un método bivariante basado en un modelo de series de tiempo difusas con sustitutos (Yu y Huarng, 2008).

- (8) Modelo basado en vectores autorregresivos con media móvil y variables exógenas VARMAX (Ostermark y Hernesniemi, 1995).
- (9) Modelos matemáticos: modelo de regresión determinista que incorpora ecuaciones diferenciales aplicadas a las series de tiempo del índice de mercado con el fin de proporcionar un patrón de comportamiento mucho más amplio que los métodos estadísticos o econométricos existentes, incorporando crecimientos polinomiales, exponenciales y logarítmicos que reflejan las tendencias de los datos; y aplica un algoritmo de regresión determinista que permite estimar los valores futuros de la serie de tiempo del índice (Balbas, Balbas, Galperin y Galperin, 2008), modelo de caja gris de Bernoulli compuesto por una ecuación diferencial no lineal de potencia n determinada por un programa iterativo que calcula el ARPE mínimo, se incorpora el concepto de equilibrio de Nash para agregar un parámetro adicional a la ecuación con el fin de mejorar la precisión del modelo (Chen, Hsin y Wu, 2010); método basado en teoría de los modelos de caja gris, se aplican series de Fourier para afinar los errores residuales producidos por el modelo y se utiliza el estado de Markov para producir los resultados intermedios generados por el modelo (Hsu, Liu, Yeh y Hung, 2009); modelo de pronóstico que aplica wavelets con el fin de identificar las características de la serie de tiempo del índice y se emplea regresión por mínimos cuadrados parciales para crear un sub-espacio más eficiente que mantiene la máxima covarianza entre entradas y salidas para llevar a cabo el pronóstico final (Huang, 2011); modelo que aplica la teoría matemática de la democracia, la cual funciona según los índices de popularidad y universalidad que se utilizan para encontrar representantes socialmente óptimos; matemáticamente, el objetivo es encontrar datos óptimos que puedan representar el comportamiento de la serie de datos del índice (Tangian, 2008).

Las técnicas más utilizadas para los métodos desarrollados para el pronóstico de índices de mercado, los modelos de series de tiempo difusas, comenzaron a utilizarse a partir del año 2001. Sin embargo, las mayores investigaciones y publicaciones se desarrollaron en el periodo comprendido entre los años 2010 y 2011. Por otra parte, las redes neuronales artificiales, aunque hay investigaciones publicadas desde el año 2000, tienen su auge durante el año 2011. A continuación se presenta la evolución de cada una de las técnicas utilizadas para el pronóstico de índices de mercado (ver Figura 5).

5.2 RQ2 ¿Cuáles son las metodologías de pronóstico de la dirección de los cambios de índices de mercado comúnmente usadas?

Se han desarrollado algunos métodos para el pronóstico de la dirección de los cambios de índices de mercado; los más comúnmente usados son:

- (1) Modelos neuro-difusos: modelo híbrido neuro-difuso basado en la actualización dinámica de la adaptación y calibración de las reglas heurísticas difusas de aprendizaje (Bekiros, 2010).
- (2) ANFIS: modelo ANFIS (Chang, Wei y Cheng, 2011).
- (3) Máquinas de soporte vectorial: modelo de máquinas de soporte vectorial que utiliza un método híbrido de selección de entradas relevantes con el fin de mejorar el pronóstico del modelo: se aplica el método F-Score para filtrar las entradas y el método F-SSFS para seleccionar las características óptimas de los datos de entrada (Lee, 2009).
- (4) Modelos neuro-difusos: modelo híbrido neuro-difuso basado en la actualización dinámica de la adaptación y calibración de las reglas heurísticas difusas de aprendizaje (Bekiros, 2010).
- (5) ANFIS: modelo ANFIS (Chang, Wei y Cheng, 2011).

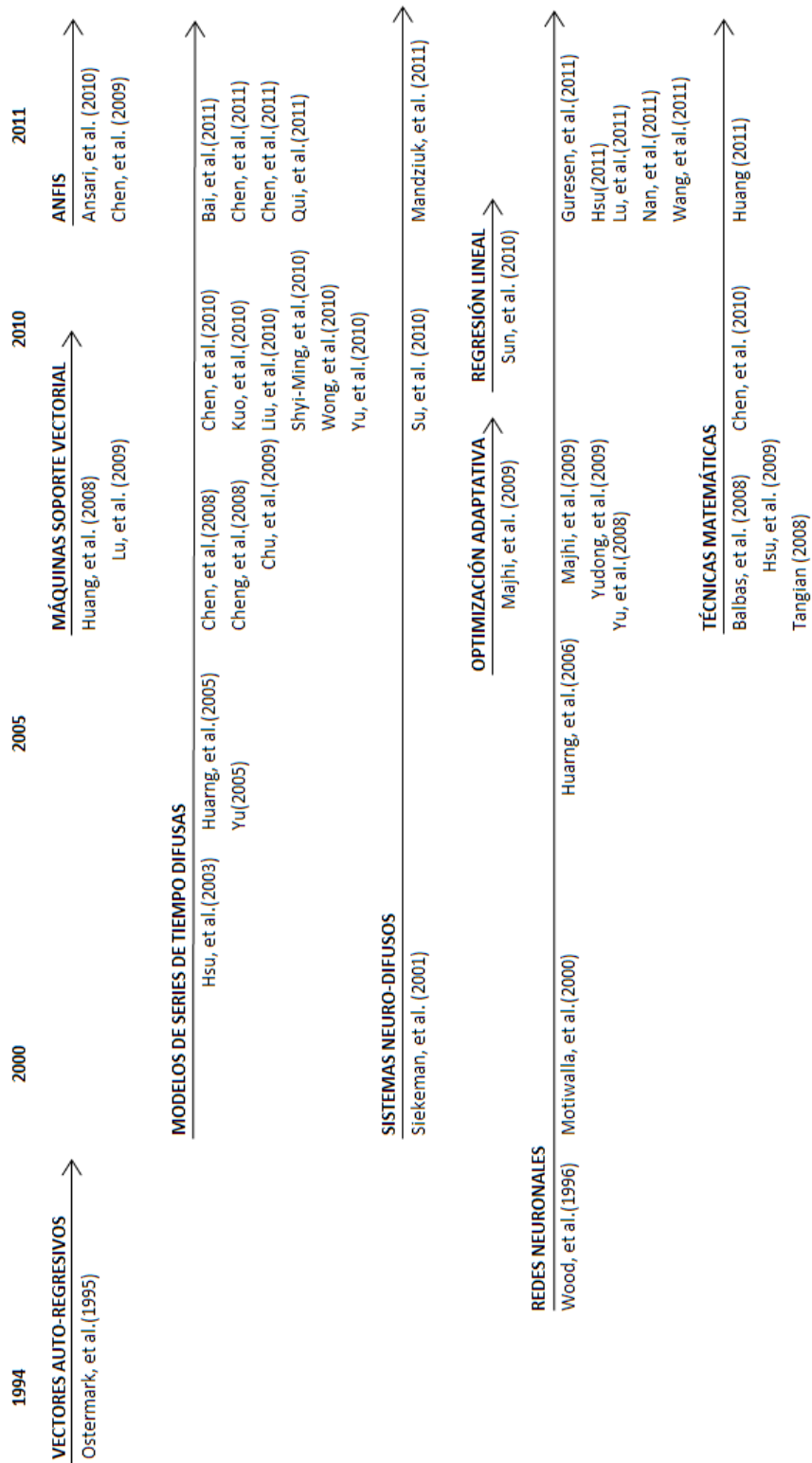


Figura 5. Evolución de las técnicas empleadas para el pronóstico de índices de mercado

- (6) Máquinas de soporte vectorial: modelo de máquinas de soporte vectorial que utiliza un método híbrido de selección de entradas relevantes con el fin de mejorar el pronóstico del modelo: se aplica el método F-Score para filtrar las entradas y el método F-SSFS para seleccionar las características óptimas de los datos de entrada (Lee, 2009).
- (7) Redes neuronales: se presentan dos modelos para el pronóstico de la dirección de los cambios de índices de mercado: un modelo basado en redes neuronales perceptrón multicapa y un modelo adaptativo de suavizado exponencial en el que los parámetros del modelo se van actualizando a lo largo de la predicción (de Faria, Albuquerque, Gonzalez, Cavalcante y Alburquerque, 2009); modelo híbrido de redes neuronales artificiales que utiliza reglas definidas por expertos utilizadas para el entrenamiento de la red neuronal (Tsaih, Hsu y Lay, 1998).
- (8) Reglas de asociación: modelo basado en técnicas de minería de datos para encontrar reglas de asociación para la predicción de la dirección de los cambios de índices de mercado, se utilizan series de tiempo de otros índices para predecir el índice objetivo (Na y Sohn, 2011).
- (9) modelo de series de tiempo difusas en el que las reglas del motor de inferencia fueron definidas teniendo en cuenta el comportamiento histórico del índice (Souto-Maior, Borba y da Costa, 2011).

5.3 RQ3 ¿Cuáles son las metodologías de pronóstico de la volatilidad de índices de mercado comúnmente usadas?

La práctica tradicional consiste en usar diferentes tipos de métodos. Los más comúnmente usados se basan en modelos ARCH y GARCH o variaciones de éstos. A continuación se presentan las metodologías para el pronóstico de la volatilidad de índices de mercado encontradas en la literatura:

- (1) ANFIS: modelo ANFIS para la predicción de índices de mercado basado en la volatilidad de otros índices accionarios (Cheng, Wei y Chen, 2009).
- (2) Conjuntos rugosos: modelo de conjuntos rugosos que aplica algoritmos de agrupamiento automático para seleccionar las entradas relevantes con el fin de mejorar los resultados del modelo (Ho, 2011); modelo de conjuntos rugosos (Yao y Herbert, 2009).
- (3) Modelo basado en el concepto de equilibrio general que utiliza un estimador ponderado de media móvil exponencial con el fin de de mejorar la capacidad de pronóstico del modelo (Wang, 2009).
- (4) Modelo adaptativo multivariable de pronóstico que utiliza la volatilidad de otros índices para predecir la volatilidad del índice objetivo; adicionalmente, utiliza técnicas de aprendizaje adaptativo para adaptar los parámetros lineales del modelo propuesto con el fin de alcanzar un rendimiento óptimo (Cheng y Wei, 2009).
- (5) Modelos ARCH: modelo ARCH que utiliza los rendimientos históricos de las acciones que componen el índice de mercado con el fin de pronosticar su volatilidad (Bramante y Luigi, 2001); modelo ARCH que utiliza la variación de los rendimientos del índice de mercado durante el día con el objetivo de predecir su volatilidad (Blair, Poon y Taylor, 2001).
- (6) Modelos GARCH y sus variaciones: comparación de los modelos GJR-GARCH y EGARCH bajo diferentes distribuciones de error (Liu y Hung, 2010); comparación de los modelos GARCH, GJR.GARCH y EGARCH bajo diferentes distribuciones de error (Mohd Nor y Shamiri, 2007); modelo GARCH para el pronóstico de volatilidad de índices de mercado a partir de series de tiempo de índices de futuros sobre acciones relacionados (Martikainen y Puttonen, 1994); comparación de los modelos GARCH, EGARCH y NAGARCH bajo diferentes distribuciones de error (Shamiri e Isa, 2009).

Los principales métodos para el pronóstico de volatilidad de índices están basados en modelos ARCH y GARCH. Los primeros se utilizaron principalmente en el año 2011, mientras que los segundos y sus variaciones se comenzaron a aplicar desde 1994, teniendo sus principales investigaciones y publicaciones a partir del 2007. La Figura 6 presenta la evolución en el tiempo de las técnicas que se han utilizado para el pronóstico de volatilidad de índices de mercado.

5.4 RQ4 ¿Cuáles son las metodologías más utilizadas para modelar índices de mercado?

Antes de dar solución a esta pregunta de investigación es importante establecer la diferencia entre modelado y pronóstico. El modelado de series de tiempo se refiere al proceso metodológico de obtener el modelo de una serie de tiempo, tal que, dicho modelo capture todas las propiedades estadísticas de la serie estudiada y se cumplan los supuestos de los que se parte para formular dicho modelo. A diferencia del modelado, la predicción se centra en lograr la mayor precisión de los valores futuros, tal que el cumplir los supuestos estadísticos ya enunciados pierde importancia. En los modelos de predicción no se requieren fuertes supuestos del modelo y se puede asignar cualquier función sin una suposición a priori sobre las propiedades de los datos.

En la literatura se encontraron diferentes métodos que modelan índices de mercado los cuales se presentan a continuación.

- (1) Paradigmas inteligentes: modelo para modelar índices de mercado a partir de la integración de cuatro paradigmas inteligentes diferentes redes neuronales, máquinas de soporte vectorial, sistema neuro-difuso y redes neuronales probabilísticas, ensamblados mediante dos enfoques: algoritmos genéticos e integración directa (Abraham y AuYeung, 2003).

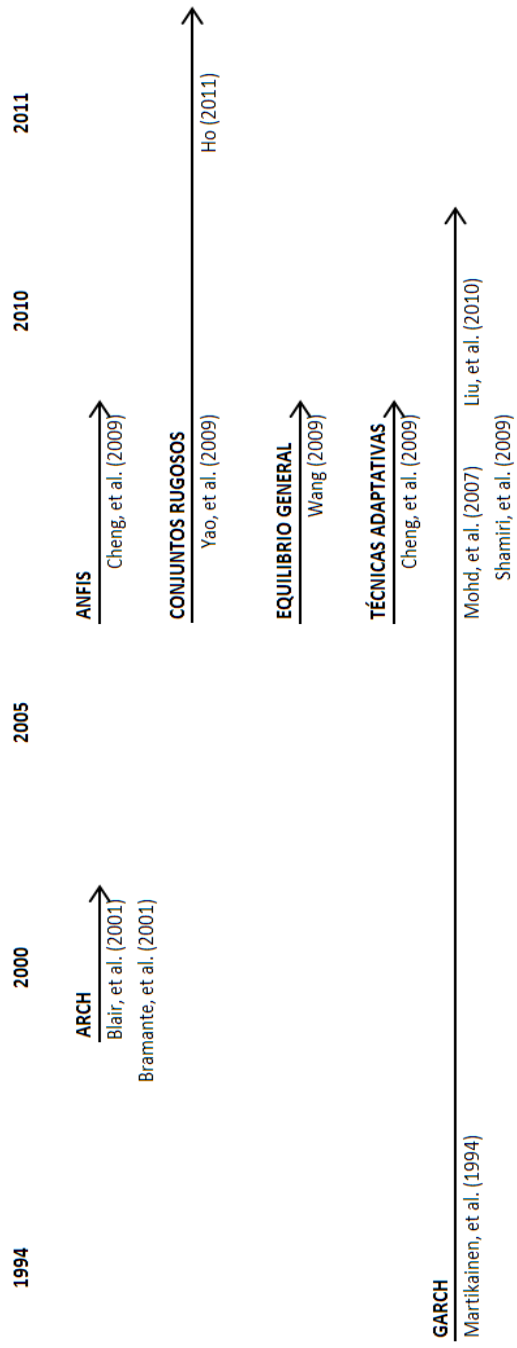


Figura 6. Evolución de las técnicas utilizadas para el pronóstico de volatilidad de índices de mercado

- (2) Paradigmas inteligentes: modelo para modelar índices de mercado a partir de la integración de cuatro paradigmas inteligentes diferentes redes neuronales, máquinas de soporte vectorial, sistema neuro-difuso y redes neuronales probabilísticas, ensamblados mediante dos enfoques: algoritmos genéticos e integración directa (Abraham y AuYeung, 2003).
- (3) Modelo de árboles neurales flexibles para el modelado de índices de mercado, como parámetros de entrada se utilizan los precios de apertura, los precios de cierre y el precio máximo (Chen, Yang y Abraham, 2007).

5.5 RQ5 ¿Cuáles son las metodologías más utilizadas para modelar la volatilidad de índices de mercado?

Según la literatura actual, el método comúnmente usado para modelar la volatilidad de índices de mercado consiste en variaciones de modelos ARCH: comparación del modelo SWARCH con los modelos convencionales ARCH y GARCH (Li y Lin, 2003).

5.6 RQ6 ¿Cuáles son los principales factores que afectan la precisión de los modelos de pronóstico?

A continuación se listan los factores que afectan la precisión de los modelos de pronóstico que han sido reportados en la literatura:

- (1) Uno de los factores que afecta los resultados de los modelos de predicción de series de tiempo difusas de manera significativa es la longitud de los intervalos de las series de tiempo utilizadas para el pronóstico (Chen y Tanuwijaya, 2011).

- (2) Los modelos de series de tiempo convencionales para el pronóstico de índices de mercado tienen tres problemas principales que afectan los resultados de los mismos: la mayoría de los métodos estadísticos se basan en suposiciones específicas acerca de las variables, se utiliza una sola variable para el pronóstico, y las reglas que se extraen de los modelos de redes neuronales artificiales no son fácilmente comprensibles (Cheng, Wei y Cheng, 2011).
- (3) Los valores extremos de las series de tiempo y la longitud de los intervalos usados para la predicción afectan los resultados de los modelos de pronóstico de índices bursátiles. Se propone un método basado en la transformada wavelet, para la detección de valores extremos en series de tiempo financieras y poder seleccionar las entradas que son realmente relevantes. (Gran y Veiga, 2010).
- (4) La longitud del intervalo afecta significativamente los resultados de la predicción de un índice de mercado cuando se usan modelos de series de tiempo difusas y se propone un método basado en distribución y en promedios de la longitud de la serie de tiempo para encontrar la longitud más eficiente del intervalo de tiempo de la serie de datos que debería considerarse para realizar la predicción (Huarng, 2001).
- (5) La mayoría de modelos de pronóstico de series de tiempo difusas utilizan una sola variable para la predicción de índices de mercado, este factor afecta la precisión de sus resultados (Huarng y Yu, 2005).
- (6) Los factores que afectan la precisión de los modelos de series de tiempo difusas para la predicción de índices son principalmente la longitud de los intervalos de datos usados para la predicción y la utilización de un solo atributo para el pronóstico (por ejemplo: precio de cierre) y no múltiples atributos (por ejemplo: precio mínimo, precio máximo, precio de apertura, entre otros.) (Liu, Chen, Cheng y Chen, 2010).

- (7) Los valores atípicos de las series de tiempo de los precios de las acciones afectan significativamente los resultados de los modelos de pronóstico (Wang, Wang, Zhang y Guo, 2011).
- (8) Los factores que impactan en gran medida la precisión de los modelos de pronóstico se centran principalmente en la partición del discurso, el contenido de las reglas de pronóstico y los métodos de defuzzificación (Wong, Bai y Chu, 2010).
- (9) La formulación de las relaciones difusas y la longitud de los intervalos de los datos utilizados para el pronóstico cuando se usan modelos de series de tiempo difusas para la predicción de índices, son factores críticos que afectan los resultados de pronóstico (Yu, 2005).

6 CONCLUSIONES

En este trabajo se presentan los resultados de un estudio de revisión sistemática de la literatura, realizado con el propósito de determinar el estado actual de las investigaciones relacionadas con el pronóstico de índices de mercado.

La metodología empleada: revisión sistemática de la literatura, permitió seleccionar 64 artículos relevantes para dar respuesta a las preguntas de investigación planteadas. Estos artículos fueron publicados a partir de 1994, aumentando significativamente la publicación a partir del 2001 y teniendo un importante auge en los años 2008 y 2009. Esto demuestra que la mayoría de las principales investigaciones se han llevado a cabo durante estos períodos. Adicionalmente, las publicaciones sobre el tema han venido aumentando en los últimos años lo que demuestra que la predicción de índices de mercado es relevante actualmente.

En general, puede concluirse que gran parte de las investigaciones están enfocadas en el pronóstico de los índices de mercado como tal. Sin embargo, se han propuesto metodologías para el pronóstico de volatilidad, pronóstico de la dirección de los cambios, modelamiento de los índices y modelamiento de la volatilidad de los índices de mercado.

Según la literatura analizada, se puede concluir que no hay una técnica dominante para el pronóstico de índices de mercado, sin embargo es común encontrar modelos híbridos de redes neuronales o de lógica difusa. Por otro lado, para el pronóstico de la volatilidad de los índices de mercado se siguen utilizando y aplicando los modelos ARCH, GARCH y sus variaciones.

Una de las prácticas usuales en los estudios analizados consiste en usar varios métodos para el pronóstico y seleccionar el que genere mayor precisión en los resultados, medidos a partir de las comparaciones del error de pronóstico en cada uno de ellos; y comparar modelos híbridos contra técnicas clásicas con el fin de determinar si los modelos híbridos contribuyen al mejoramiento de la precisión de los resultados de los modelos.

Por otro lado, según la información revisada de los artículos recuperados, tanto los valores extremos de la serie de tiempo de los índices de mercado, como la longitud de los intervalos de los datos utilizados para la predicción, afectan significativamente los resultados de los modelos. Es por esto, que dichos factores deben tenerse en cuenta a la hora de aplicar las metodologías de pronóstico de índices bursátiles.

Los resultados obtenidos presentan diferentes y numerosas metodologías para el pronóstico de índices de mercado y cada una de ellas muestra pequeños errores de pronóstico. Por esta razón, es difícil determinar cuál de ellas es la más adecuada para tal fin. Se hace necesario establecer trabajos de investigación posteriores para determinar cuál es la metodología más adecuada para el pronóstico de índices de mercado.

ANEXO 1: Descripción artículos analizados

1. Abraham A., AuYeung A. (2003). Integrating ensemble of intelligent systems for modeling stock indices. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol.2687, 774-781.

Se demuestra cómo el comportamiento caótico de los mercados de valores es muy bien representado por un modelo que integra cuatro paradigmas inteligentes diferentes: redes neuronales, máquinas de soporte vectorial, sistema neuro-difuso y redes neuronales probabilísticas, ensamblados mediante dos enfoques: algoritmos genéticos e integración directa. Este método es probado mediante su aplicación en la predicción de los índices Nasdaq-100 y S&P, para los cuales se tomaron los valores históricos de apertura para cada índice, precio mínimo, precio máximo y precio de cierre.

2. Ansari T., Kumar M., Shukla A., Dhar J. y Tiwari R. (2010). Sequential combination of statistics, econometrics and Adaptive Neural-Fuzzy Interface for stock market prediction. *Expert Systems with Applications*, Vol.37, 5116-5125.

Se propone un método híbrido para el pronóstico del índice de mercado NASDAQ. El método emplea análisis estadístico y econométrico para identificar la naturaleza exacta del conjunto de datos y la relación entre ellos con el objetivo de optimizar los parámetros de pronóstico; luego, utilizando un método de agrupamiento de datos se agrupan los datos y se utiliza un modelo ANFIS con el fin de crear las funciones de membresía difusas y entrenar la red neuronal. El método propuesto se compara con modelos econométricos de predicción y según los resultados experimentales se demuestra que los resultados del mismo son más precisos y presentan menor error de pronóstico.

3. Bai E., Wong W.K., Chu W.C., Xia M. y Pan F. (2011). A heuristic time-invariant model for fuzzy time series forecasting. *Expert Systems with Applications*, Vol.38, 2701-2707.

Se propone un modelo de series de tiempo difusas para el pronóstico del índice de mercado de la bolsa de valores de Taiwán TAIEX. El modelo incluye un factor predictor para indicar la tendencia de aumento o disminución de la serie de tiempo con el fin de capturar las tendencias de las observaciones con mayor precisión. Según los resultados experimentales, el modelo propuesto mejora los resultados de pronóstico en comparación con otros modelos de series de tiempo difusas.

4. Balbas A., Balbas B., Galperin I. y Galperin E. (2008). Deterministic regression model and visual basic code for optimal forecasting of financial time series. *Computers and Mathematics with Applications*, Vol.56, 2757-2771.

Se presenta un modelo de regresión determinista en el que se analiza la historia y la evolución actual de la serie de tiempo del índice S&P500 para realizar el pronóstico. El modelo se basa en ecuaciones diferenciales aplicadas a las series de tiempo del índice con el fin de proporcionar un patrón de comportamiento mucho más amplio que los métodos estadísticos o econométricos existentes, incorporando crecimientos polinomiales, exponenciales y logarítmicos que reflejan las tendencias de los datos; y se propone un algoritmo de regresión determinista que permite estimar los valores futuros de la serie de tiempo del índice. Los resultados empíricos confirman la validez y eficacia del modelo propuesto; adicionalmente los resultados son poco sensibles respecto a los posibles errores de entrada en la serie de datos.

5. Bekiros S.D. (2010). Fuzzy adaptive decision-making for boundedly rational traders in speculative stock markets. *European Journal of Operational Research*, Vol.202, 285-293.

Se propone un sistema híbrido neuro-difuso con el fin de predecir la dirección del mercado en diez de los índices bursátiles más importantes de Estados Unidos, Europa y el sudeste asiático: S&P500 (Usa), KLCI (Malasia), HangSeng (Hong Kong), SET100 (Tailandia), NYSE (Usa), FTSE100 (Reino Unido), CAC40 (Francia), Stock Exchange Weighted (Taiwán), Jakarta Stock Exchange Composite (Indonesia), Straits Times (Singapur). Se demuestra a través de un análisis empírico que el modelo permite obtener rendimientos significativamente más altos en comparación con los modelos de redes neuronales recurrentes, proporcionando información válida sobre los cambios de la dirección del índice de mercado para el siguiente día de negociación. El pronóstico óptimo se basa en la actualización dinámica de la adaptación y calibración de las reglas heurísticas difusas de aprendizaje.

6. Blair B.J., Poon S.-H. y Taylor S.J. (2001). Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics*, Vol.105, 5-26.

Se propone un modelo ARCH para el pronóstico del índice de mercado S&P100 a partir de la variación de los rendimientos del índice durante el día. Los resultados arrojan un RMSE pequeño en comparación con otros modelos similares.

7. Bramante R. y Luigi S. (2001). Forecasting stock index volatility. *Applied Stochastic Models in Business and Industry*, Vol.17, 19-26.

Se propone un modelo ARCH para la predicción del índice de mercado italiano MIB30 usando los rendimientos históricos de las acciones que componen el índice. Específicamente, el enfoque de los autores supone que la volatilidad del índice se puede explicar por la volatilidad de los retornos históricos de sus componentes relacionados.

8. Chang J.-R., Wei L.-Y. y Cheng C.-H. (2011). A hybrid ANFIS model based on AR and volatility for TAIEX forecasting. *Applied Soft Computing Journal*, Vol.11, 1388-1395.

Se propone modelo ANFIS para el pronóstico de los movimientos del índice de la bolsa de valores de Taiwán (TAIEX). Los resultados de la aplicación del modelo indican que el modelo propuesto tiene mayor precisión y menor error de pronóstico que otros modelos similares.

9. Chen C.I., Hsin P.H. y Wu C.S. (2010). Forecasting Taiwan's major stock indices by the Nash nonlinear grey Bernoulli model. *Expert Systems with Applications*, Vol.37, 7557-7562.

Se propone un modelo para el pronóstico de datos no lineales usando el modelo de caja gris de Bernoulli con el fin de predecir series de tiempo de índices de mercado, caracterizadas por presentar no linealidad en los datos. El modelo consiste en una ecuación diferencial no lineal de potencia n determinada por un programa iterativo que calcula el ARPE mínimo. Con el fin de mejorar la precisión del modelo, se incorpora el concepto de equilibrio de Nash que agrega un nuevo parámetro a la ecuación y mejora la capacidad de ajuste del modelo. Los resultados empíricos enfocados en el pronóstico del índice TAIEX, muestran que el modelo de caja gris de Bernoulli modificado con el concepto de equilibrio de Nash mejora la precisión del pronóstico.

10. Chen S.-M. y Chen C.-D. (2011). TAIEX forecasting based on fuzzy time series and fuzzy variation groups. *IEEE Transactions on Fuzzy Systems*, Vol.19, 1-12.

Se presenta un método basado en un modelo multivariante de series de tiempo difusas para predecir el índice de mercado TAIEX. Se utilizan grupos de variación difusas con el fin de analizar en detalle los datos de entrenamiento para el pronóstico. En estos grupos el factor principal es el TAIEX y los factores secundarios son los índices Dow Jones, Nasdaq y Mb1. El modelo es comparado con modelos univariantes y bivariantes de series de tiempo difusas. Los resultados experimentales demuestran que el RMSE del modelo propuesto es mejor que el de los modelos con los que fue comparado.

11. Chen S.-M. y Tanuwijaya K. (2011). Fuzzy forecasting based on high-order fuzzy logical relationships and automatic clustering techniques. *Expert Systems with Applications*, Vol.38, 15425-15437.

Se presenta un método basado en modelos de series de tiempo para el pronóstico del índice de mercado TAIEX, en el cual se extraen las relaciones lógicas de alto orden para construir reglas que permitan pronosticar los datos y aumentar la precisión del modelo. Adicionalmente, se utilizan técnicas automáticas de agrupamiento con el fin encontrar el intervalo óptimo de la serie de tiempo que se debe utilizar para el pronóstico y así mejorar la capacidad de predicción del modelo. Los resultados experimentales demuestran que este método consigue una mayor precisión de pronóstico que los métodos convencionales de series de tiempo difusas.

12. Chen S.-M. y Wang N.-Y. (2010). Fuzzy forecasting based on fuzzy-trend logical relationship groups. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol.40, 1343-1358.

Se propone un modelo de series de tiempo difusas con el fin de pronosticar el índice de mercado TAIEX. El modelo utiliza un algoritmo de agrupamiento automático con el fin de clasificar los datos históricos en intervalos de longitudes diferentes para encontrar la longitud del intervalo óptima, y se extraen las relaciones lógicas para construir las reglas que permitan pronosticar los datos. Según los resultados, el modelo propuesto obtiene una mayor tasa de precisión que los modelos de series de tiempo difusas

13. Chen T.-L., Cheng C.-H. y Teoh H.-J. (2008). High-order fuzzy time-series based on multi-period adaptation model for forecasting stock

markets. *Physica A: Statistical Mechanics and its Applications*, Vol.387, 876-888.

Se propone un modelo de series de tiempo difusas en el cual se extraen las relaciones lógicas de alto orden para construir reglas que permitan pronosticar los datos y aumentar la capacidad de predicción de los índices bursátiles, específicamente del índice de mercado TAIEX. El modelo propuesto es comparado con un modelo de series de tiempo tradicional. Los resultados empíricos demuestran que el modelo propuesto supera los modelos con los que fue comparado y reduce significativamente el RMSE.

14. Chen Y., Yang B., Abraham A. (2007). Flexible neural trees ensemble for stock index modeling. *Neurocomputing*, Vol.70, 697-703.

Se propone un modelo de árboles neurales flexibles, programación genética y técnicas de conjuntos, específicamente regresión polinomial ponderada local, para el modelado de los índices de mercado NASDAQ100 y NIFTY. Como parámetros de entrada se utilizan los precios de apertura, cierre y máximo para cada índice. Los resultados experimentales indican que los parámetros más importantes que afectan a los precios de las acciones son los precios de apertura y de cierre. Después de la evaluación del modelo, los autores concluyen que incluso en las fluctuaciones aparentemente al azar, existe una función determinística subyacente que está asociada directamente con los precios de apertura, cierre y máximo del índice de cualquier día, haciendo posible su modelado.

15. Cheng C.-H., Cheng G.-W. y Wang J.-W. (2008). Multi-attribute fuzzy time series method based on fuzzy clustering. *Expert Systems with Applications*, Vol.34, 1235-1242.

Se presenta un modelo de series de tiempo difusas de múltiples atributos que aplica el algoritmo de agrupamiento Fuzzy C Mean para particionar los datos y para modificar los pronósticos lingüísticos, con el fin de mejorar la exactitud del modelo. El atributo principal lo constituye el precio de cierre diario del índice y los atributos secundarios se constituyen por el precio diario más alto y el precio diario más bajo del índice. Los resultados experimentales demuestran que el método propuesto puede ser aplicado de manera efectiva para el pronóstico de índices de mercado.

16. Cheng C.-H. y Wei L.-Y. (2009). Volatility model based on multi-stock index for TAIEX forecasting. *Expert Systems with Applications*, Vol.36, 6187-6191.

Se propone un modelo adaptativo multivariable de pronóstico que utiliza la volatilidad de los índices Dow Jones y NASDAQ para predecir la volatilidad del índice de mercado TAIEX. Adicionalmente, utiliza técnicas de aprendizaje adaptativo para adaptar los parámetros lineales del modelo propuesto con el fin de alcanzar un rendimiento óptimo. Los resultados de las pruebas estadísticas demuestran que la volatilidad del Dow Jones y el NASDAQ afectan el TAIEX de manera significativa y que el modelo propuesto presenta errores de pronóstico menores que otros modelos similares.

17. Cheng C.-H., Wei L.-Y. y Chen Y.-S. (2009). Fusion ANFIS models based on multi-stock volatility causality for TAIEX forecasting. *Neurocomputing*, Vol.72, 3462-3468.

Se propone un modelo ANFIS para el pronóstico del índice de mercado de la bolsa de Taiwán TAIEX basado en la volatilidad de los índices

estadounidenses NASDAQ y Dow Jones. Adicionalmente, el modelo propuesto se compara con modelos convencionales de series de tiempo difusas. Los resultados experimentales indican que el modelo propuesto es superior a los métodos de series de tiempo difusas en términos del error cuadrático medio. Se evidencia también que los índices de mercado estadounidenses afectan significativamente el índice TAIEX.

18. Chu H.-H., Chen T.-L., Cheng C.-H., Huang C.-C. (2009). Fuzzy dual-factor time-series for stock index forecasting. *Expert Systems with Applications*, Vol.36, 165-171.

Se propone un modelo difuso de factor doble modificado para predecir series de tiempo de los índices de mercado TAIEX y NASDAQ. Este modelo considera el índice de la bolsa y un indicador de análisis técnico del volumen del comercio de acciones para predecir los índices de mercado. El modelo propuesto utiliza dos enfoques de predicción para procesar los dos factores por separado y una ecuación de predicción de doble factor para generar el pronóstico.

19. De Faria E.L., Albuquerque M.P., Gonzalez J.L., Cavalcante J.T.P. y Albuquerque M.P. (2009). Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. *Expert Systems with Applications*, Vol.36, 12506-12509.

Se proponen dos modelos para predecir el índice principal de la bolsa de valores Brasileña Ibovespa: un modelo basado en redes neuronales perceptrón multicapa y un modelo adaptativo de suavizado exponencial en el que los parámetros del modelo se van actualizando a lo largo de la predicción. El objetivo principal es comparar la capacidad de pronóstico de ambos métodos. El resultado de la comparación de ambos modelos

indica que se obtiene un menor error de pronóstico con el modelo de redes neuronales perceptrón multicapa.

20. Gran A. y Veiga H. (2010). Wavelet-based detection of outliers in financial time series. *Computational Statistics and Data Analysis*, Vol.54, 2580-2593.

Debido a que los valores extremos de las series de tiempo y la longitud de los intervalos usados para la predicción afectan los resultados de los modelos de pronóstico de índices bursátiles. Se propone un método basado en wavelets para la corrección del valor extremo en modelos de pronóstico de volatilidad de índices de mercado. Su objetivo es la detección y corrección del valor extremo que se puede aplicar a los residuos de diferentes modelos de volatilidad cuyos errores pueden seguir cualquier distribución conocida. Adicionalmente, el procedimiento de corrección de valores extremos permite intervalos más precisos de predicción. Para ilustrar el desempeño del modelo se utilizan los índices Dow Jones, FTSE-100 y el S&P 500.

21. Guresen E., Kayakutlu G. y Daim T.U. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, Vol.38, 10389-10397.

Se evalúa la eficacia de los modelos de redes neuronales utilizados para predecir el índice de mercado NASDAQ, que son conocidos por ser dinámicos y eficaces en la predicción del mercado de valores. Los modelos analizados son: el perceptrón multicapa, red neuronal artificial dinámica DAN2 y redes neuronales híbridas que utilizan modelos GARCH para seleccionar las entradas relevantes. La comparación de cada modelo se realiza bajo dos puntos: error cuadrático medio y la desviación

media absoluta. Los resultados muestran que el modelo clásico perceptrón multicapa supera los modelos DAN2 y redes neuronales híbridas que utilizan modelos GARCH con una pequeña diferencia.

22. Ho W.-R.J. (2011). Applied rough set logics for multi-criteria decision analysis in stock index volatility projection. *Review of Pacific Basin Financial Markets and Policies*, Vol.14, 715-735.

Su principal objetivo es utilizar las técnicas basadas en reglas de pronóstico para predecir la volatilidad de los índices de mercado. Se establece un prototipo basado en la teoría de conjuntos rugosos, algoritmo de Delphi y técnicas de análisis de agrupamiento o de conglomerados para seleccionar los atributos más importantes y pronosticar el índice bursátil de la bolsa de valores de Taiwán (TAIEX).

23. Hsu C.-M. (2011). A hybrid procedure for stock price prediction by integrating self-organizing map and genetic programming. *Expert Systems with Applications*, Vol.38, 14026-14036.

Se propone un modelo híbrido de redes neuronales de mapas auto-organizados SOM que utiliza programación genética para optimizar los pesos de las conexiones de la red, con el fin de lograr predecir con mayor precisión el índice de mercado TAIEX. Los resultados experimentales demuestran que el modelo propuesto puede ser considerado como una herramienta viable y efectiva para el pronóstico de índices bursátiles.

24. Hsu Y.-T., Liu M.-C., Yeh J. y Hung H.-F. (2009). Forecasting the turning time of stock market based on Markov-Fourier grey model. *Expert Systems with Applications*, Vol.36, 8597-8603.

Se presenta un método de predicción para el pronóstico del índice de mercado TAIEX. Este método está basado en la teoría de los modelos de caja gris, se aplican series de Fourier para afinar los errores residuales producidos por el modelo y se utiliza el estado de Markov para producir los resultados intermedios generados por el modelo. Los resultados experimentales demuestran que el enfoque propuesto tiene una exactitud de pronóstico mayor que otros métodos.

25. Hsu Y.-Y., Tse S.-M. y Wu B. (2003). A new approach of bivariate fuzzy time series analysis to the forecasting of a stock index. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol.11, 671-690.

Se propone un modelo de series de tiempo difusas bivariado para el pronóstico del índice TAIEX. Se aplica el modelo para los precios de cierre de las series de tiempo de los índices y el volumen de acciones negociadas. Este modelo se compara con el modelo bivariado ARMA. Los resultados evidencian que, según el MAFA (mean absolute forecasting accuracy), el modelo propuesto tiene errores de predicción menores que los modelos ARMA.

26. Huang S.-C. y Wu T.-K. (2008). Integrating GA-based time-scale feature extractions with SVMs for stock index forecasting. *Expert Systems with Applications*, Vol.35, 2080-2088.

Se propone un modelo híbrido de máquinas de soporte vectorial cuyas entradas se descomponen usando wavelets y se seleccionan usando algoritmos genéticos, para el pronóstico de los índices bursátiles CAC40, FTSE100, DAX30, MIB40, TSX60, S&P500 y NK225. Este modelo se compara con los modelos tradicionales de redes neuronales, máquinas de

soporte vectorial puras y modelos GARCH. Los errores de predicción del modelo propuesto son significativamente menores que los demás métodos con los que fue comparado.

27. Huang S.-C. (2011). Forecasting stock indices with wavelet domain kernel partial least square regressions. *Applied Soft Computing Journal*, Vol.11, 5433-5443.

Se propone un modelo para el pronóstico de los índices de mercado NASDAQ, NK225, TWSI y KOSPI, que aplica wavelets con el fin de identificar las características de la serie de tiempo del índice y se emplea regresión por mínimos cuadrados parciales para crear un sub-espacio más eficiente que mantiene la máxima covarianza entre entradas y salidas para llevar a cabo el pronóstico final. Los resultados empíricos demuestran que el modelo propuesto tiene menor error de pronóstico que modelos tradicionales de redes neuronales, máquinas de soporte vectorial y modelos GARCH.

28. Huarng K. (2001). Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets and Systems*, Vol.123, 387-394.

Utilizando el índice de mercado de la bolsa de valores de Taiwán (TAIEX), los autores demuestran que la longitud del intervalo afecta significativamente los resultados de la predicción de un índice de mercado cuando se usan series de tiempo difusas y propone un método basado en distribución y promedios de la longitud de la serie de tiempo para encontrar la longitud más eficiente del intervalo de tiempo que debería considerarse para realizar la predicción.

29. Huarng K., Yu H.-K. (2005). A type 2 fuzzy time series model for stock index forecasting. *Physica A: Statistical Mechanics and its Applications*, Vol.353, 445-462.

Los autores tienen como objetivo proponer un marco para el desarrollo de un modelo de series de tiempo difusas tipo 2 con el fin de mejorar los resultados de los pronósticos del índice TAIEX. Este modelo utiliza observaciones adicionales para mejorar las relaciones difusas obtenidas de un modelo tipo 1. Los resultados empíricos demuestran que el RMSE en el caso del modelo de series de tiempo difusas tipo 2 son más pequeños que los obtenidos para el modelo tipo 1, adicionalmente, puede concluirse que los modelos tipo 2 son más consistentes.

30. Huarng K. y Yu T.H.-K. (2006). The application of neural networks to forecast fuzzy time series. *Physica A: Statistical Mechanics and its Applications*, Vol.363, 481-491.

Se proponen dos modelos: un modelo básico con un enfoque de redes neuronales de alimentación hacia adelante y un modelo híbrido neuro-difuso con redes neuronales de propagación hacia atrás, que predice los patrones conocidos y desconocidos de la serie de tiempo del índice. A partir del análisis empírico, la red neuronal obtiene mejores resultados al pronosticar los patrones conocidos de la serie de tiempo del índice. Sin embargo, el modelo híbrido supera el modelo básico y los modelos de series de tiempo difusas convencionales al prever los patrones desconocidos.

31. Kuo I.-H., Horng S.-J., Chen Y.-H., Run R.-S., Kao T.-W., Chen R.-J., Lai J.-L. y Lin T.-L. (2010). Forecasting TAIEX based on fuzzy time series

and particle swarm optimization. *Expert Systems with Applications*, Vol.37, 1494-1502.

Se presenta un modelo híbrido para pronosticar el índice de mercado TAIEX. El método está basado en modelos de series de tiempo difusas y optimización evolutiva por enjambre de partículas. En este caso, la optimización por enjambre de partículas se utiliza para entrenar las reglas difusas de pronóstico basada en los datos de entrada. La principal diferencia con modelos anteriores que utilizan este mismo principio, se centra en el procedimiento para la creación de las reglas de predicción. Los resultados experimentales muestran que el modelo propuesto es más preciso y genera menores errores de pronóstico que los modelos con los que fue comparado.

32. Lee M.-C. (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Expert Systems with Applications*, Vol.36, 10896-10904.

Se propone un modelo de máquinas de soporte vectorial para predecir la tendencia del índice de mercado NASDAQ. Utiliza un método híbrido de selección de entradas relevantes con el fin de mejorar el pronóstico del modelo. En primer lugar se aplica el método F-Score para filtrar las entradas y en segundo lugar, se aplica el método F-SSFS para seleccionar las características óptimas de los datos de entrada. Este método combina la ventaja de los métodos de selección de entradas relevantes para seleccionar el subconjunto de datos óptimos de un conjunto de datos originales. Los resultados experimentales demuestran que el modelo propuesto tiene alto nivel de precisión y un mejor rendimiento que los modelos de máquinas de soporte vectorial.

33. Li M.-Y.L. y Lin H.-W.W. (2003). Examining the volatility of Taiwan Stock Index returns via a three-volatility-regime Markov-switching ARCH model. *Review of Quantitative Finance and Accounting*, Vol.21, 123-139.

Se comparan los modelos de volatilidad ARCH y GARCH con el modelo ARCH de Markov SWARCH para el pronóstico del TAIEX. Adicionalmente, los autores llevan a cabo pruebas de sensibilidad en los índices Dow Jones y Nikkei. Los resultados experimentales demuestran que el modelo SWARCH supera los modelos ARCH y GARCH y disminuye significativamente el error de pronóstico.

34. 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*, Vol.37, 4928-4934.

Se investiga la predicción de la volatilidad diaria del índice de mercado S&P100 mediante diferentes distribuciones de error en y variaciones del modelo GARCH. Los resultados indican que el modelo GJR GARCH genera pronósticos de volatilidad más precisos, seguido de cerca por el modelo EGARCH. Adicionalmente, se demuestra que la asimetría de los modelos es más importante que la distribución del error en el pronóstico de series de tiempo financieras.

35. Liu J.-W., Chen T.-L., Cheng C.-H. y Chen Y.-H. (2010). Adaptive-expectation based multi-attribute FTS model for forecasting TAIEX. *Computers and Mathematics with Applications*, Vol.59, 795-802.

Se propone un modelo de series de tiempo difusas de múltiples atributos para el pronóstico del índice de mercado TAIEX que aplica el algoritmo de agrupación Fuzzy C Mean para particionar los datos y para modificar

los pronósticos lingüísticos con el fin de mejorar la exactitud del modelo. Los autores indican que para tener una mejor precisión de un índice de mercado, es necesario incluir en el modelo cuatro atributos básicos: precio más alto, precio más bajo, precio de apertura y precio de cierre. Según los resultados, se evidencia que el modelo supera otros métodos de series de tiempo difusas.

36. Lu C.-J., Lee T.-S. y Chiu C.-C. (2009). Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems*, Vol.47, 115-125.

Se presenta un modelo para el pronóstico de los índices de mercado Nikkei 225 y TAIEX mediante la integración de dos modelos: un modelo de análisis de componentes independientes con el fin seleccionar las entradas relevantes y de mejorar los resultados del modelo de máquinas de soporte vectorial. Los resultados experimentales muestran que el modelo propuesto es efectivo al eliminar la información o las entradas que no son relevantes y puede producir menor error de pronóstico cuando se compara con un modelo de máquinas de soporte vectorial tradicional.

37. Lu C.-J. y Wu J.-Y. (2011). An efficient CMAC neural network for stock index forecasting. *Expert Systems with Applications*, Vol.38, 15194-15201.

Se propone un modelo de redes neuronales de articulación cerebelosa CMAC por sus siglas en inglés Cerebellar Model Articulation Controller Neural Network para el pronóstico de los índices de mercado Nikkei 225 y TAIEX. Es un modelo que presenta ventajas en la rapidez de aprendizaje, la capacidad de generalización y la resistencia al ruido. Este método fue comparado con modelos basados en máquinas de soporte vectorial y en

redes neuronales de propagación hacia atrás. Los resultados experimentales muestran que el método tiene mayores niveles de precisión que los métodos con los que fue comparado.

38. Majhi R., Panda G., Majhi B., Sahoo G. (2009). Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques. *Expert Systems with Applications*, Vol.36, 10097-10104.

Los autores proponen un modelo de predicción del índice bursátil S&P500, mediante el uso de técnicas de optimización de forrajeo de bacterias BFO y optimización adaptativa de forrajeo de bacterias ABFO. El rendimiento de la predicción a corto y largo plazo se compara con los obtenidos mediante otras técnicas de computación evolutiva tales como algoritmos genéticos y optimización por enjambre de partículas. Los resultados experimentales demuestran que el modelo propuesto presenta un menor error de pronóstico que los demás métodos con los que fue comparado.

39. Majhi R., Panda G. y Sahoo G. (2009). Development and performance evaluation of FLANN based model for forecasting of stock markets. *Expert Systems with Applications*, Vol.36, 6800-6808.

Se desarrolla un modelo de redes neuronales con una sola neurona y una sola capa para el pronóstico de los índices bursátiles Dow Jones y S&P500. El modelo utiliza una función de activación trigonométrica. Los resultados muestran que es un modelo simple y requiere poca capacidad computacional; adicionalmente, evidencian que la combinación de las técnicas empleadas genera errores pequeños de pronóstico para un día, un mes e incluso dos meses de predicción.

40. Mandziuk J. y Jaruszewicz M. (2011). Neuro-genetic system for stock index prediction. *Journal of Intelligent and Fuzzy Systems*. Vol.22, 93-123.

Se introduce y se evalúa de manera experimental un sistema neuro-genético basado redes neuronales artificiales de alimentación hacia adelante de una capa oculta y se aplican algoritmos genéticos para seleccionar de manera eficaz las variables de entrada; con el fin de predecir en el corto plazo el índice bursátil de la bolsa de Alemania DAX, a través de su relación con los índices de las bolsas de Tokio y Nueva York. Los resultados experimentales demuestran que es un modelo eficiente para el pronóstico de índices de mercado.

41. Martikainen T. y Puttonen V. (1994). A note on the predictability of Finnish stock market returns: Evidence from stock index futures markets. *European Journal of Operational Research*, Vol 73, 27-32.

Los autores pretenden contribuir con el mejoramiento del rendimiento de los modelos GARCH para el pronóstico de la volatilidad del índice bursátil FOX de Helsinki, a partir de la inclusión de series de tiempo de índices de futuros relacionados con el índice FOX. Los resultados demuestran el papel significativo que tienen las acciones de futuros sobre el índice de la bolsa de Helsinki.

42. Mohd Nor A.H.S. y Shamiri A. (2007). Modeling and forecasting volatility of the Malaysian and the Singaporean stock indices using asymmetric GARCH models and non-normal densities. *Malaysian Journal of Mathematical Sciences*, Vol.1, 83-102.

Se examinan y calculan los modelos de asimétricos de volatilidad: GARCH, el modelo de de Glosten, Jagannathan y Runkle GJR_GARCH, y el modelo exponencial EGARCH, con las distribuciones Normal, T-Student y distribución de error generalizada, para predecir dos índices asiáticos: KLCI de la bolsa de Malasia y ITS del mercado bursátil de Singapur. Los resultados de las estimaciones muestran que los modelos EGARCH y GJR-GARCH tienen mayor capacidad predictiva que el modelo simétrico GARCH. Por otro lado, se encontró que el modelo GJR-GARCH proporciona mejores predicciones fuera de la muestra para el mercado de valores de Malasia; mientras que el modelo EGARCH proporciona una mejor estimación para el mercado bursátil de Singapur.

43. Motiwalla L. y Wahab M. (2000). Predictable variation and profitable trading of US equities: A trading simulation using neural networks. *Computers and Operations Research*, Vol.27, 1111-1129.

Se propone un modelo de predicción para los índices de mercado de Estados Unidos basado en una red neuronal artificial perceptrón multicapa y se compara con un modelo paramétrico de regresión no lineal. Las pruebas se realizan con los índices: Nasdaq Composite, Nasdaq Industrial, New York Stock Exchange Composite, S&P100, Rusell 300, Rusell large-cap 1000, S&P Mid-Cap, Value Line, Wilshire 5000 y S&P 500. Los resultados experimentales sugieren que el modelo de redes neuronales genera resultados más ajustados que el modelo de regresión.

44. Na S.H. y Sohn S.Y. (2011). Forecasting changes in Korea Composite Stock Price Index (KOSPI) using association rules. *Expert Systems with Applications*, Vol.38, 9046-9049.

Los autores proponen la aplicación de técnicas de minería de datos para encontrar reglas de asociación para la predicción de cambios en el índice de mercado de Corea KOSPI basados en los datos de series de tiempo de varios índices de mercado de las bolsas del mundo relacionados entre sí. Los resultados experimentales demuestran que el índice de mercado KOSPI tiende a moverse en la misma dirección que los índices bursátiles de Estados Unidos y Europa. En contraste, tiende a moverse de manera opuesta a los índices bursátiles de países al Este de Asia, tales como Hong Kong y Japón. Además se concluye que el modelo presenta resultados precisos de pronóstico.

45. Nan X. y Sun X. (2011). Automatic stock market forecasting system based on extended language template model. *Journal of Information and Computational Science*, Vol.8, 112-118.

Se presenta un modelo basado en redes neuronales para el análisis técnico del mercado de valores y su aplicación a la predicción del índice de la bolsa de valores de Japón. Adicionalmente, se describe un sistema de lenguaje natural que aplica el modelo de plantillas de lenguajes LTM para expresar la información utilizada para la predicción del índice CSI 300 en lenguaje natural para usuarios no expertos. El modelo se compone por la red neuronal, una unidad de procesamiento, una unidad de procesamiento superior y una unidad de generación de lenguaje natural. Según los resultados experimentales, este método de entrenamiento propuesto ha contribuido a mejorar la precisión de la predicción de los índices del mercado de valores y ayuda a los usuarios no expertos a comprender el modelo de pronóstico mediante el uso de expresiones comprensibles en lenguaje natural.

46. Ostermark R. y Hernesniemi H. (1995). The impact of information timeliness on the predictability of stock and futures returns: *An application of vector models*. *European Journal of Operational Research*, Vol.85, 111-131.

Se estudia la conexión entre los cambios en el índice FOX y el precio de un futuro basado en el índice, utilizando un modelo basado en vectores autorregresivos con media móvil y variables exógenas VARMAX. Los resultados demuestran que el modelo propuesto mejora los cálculos de los parámetros de predicción y reduce el error estándar de las estimaciones realizadas.

47. Qiu W., Liu X. y Li H. (2011). A generalized method for forecasting based on fuzzy time series. *Expert Systems with Applications*, Vol.38, 10446-10453.

Se propone un modelo basado en series de tiempo difusas para predecir el índice de la bolsa de valores de Shanghai. Se aplican técnicas ensamble con el fin de mejorar la precisión de la selección de entradas relevantes. Los resultados experimentales demuestran que el modelo propuesto puede proporcionar mejores resultados que los modelos convencionales y que las técnicas de minería de datos mejoran la precisión del pronóstico.

48. Shamiri A. e Isa Z. (2009). Modeling and forecasting volatility of the Malaysian stock markets. *Journal of Mathematics and Statistics*, Vol.5, 234-240.

Se realiza una comparación de variaciones de los modelos de predicción de volatilidad: GARCH (GARCH, EGARCH y NAGARCH) bajo diferentes distribuciones de error para determinar cuál de ellos es más

efectivo para el pronóstico del índice del mercado bursátil de Malasia KLCI. Los resultados experimentales demuestran que el pronóstico exitoso de un modelo de volatilidad depende en mayor medida de la elección de la distribución del error que de la elección del modelo GARCH como tal.

49. Shyi-Ming y Yu-Chuan (2010). Multi-variable fuzzy forecasting based on fuzzy clustering and fuzzy rule interpolation techniques. *Information Sciences*, Vol.180, 4772-4783.

Se presenta un modelo multivariable de pronóstico de series de tiempo difusas para predecir el índice de la bolsa de Taiwán TAIEX. Se aplica el algoritmo de agrupamiento Fuzzy C Mean para particionar los datos y para modificar los pronósticos lingüísticos y se aplican técnicas de reglas de interpolación difusas para determinar el peso de cada regla con respecto a las observaciones de entrada, con el fin de mejorar la precisión del modelo. Se utilizan las series de tiempo de los índices Nasdaq, Dow Jones y M1b para el pronóstico del TAIEX. Los resultados experimentales demuestran que el modelo supera en precisión a otros modelos de series de tiempo difusas.

50. Siekmann S., Kruse R., Gebhardt J., Van Overbeek F. y Cooke R. (2001). Information fusion in the context of stock index prediction. *International Journal of Intelligent Systems*, Vol.1638, 363-373.

Los autores proponen un método neuro-difuso para la predicción del índice bursátil alemán DAX, utilizando la información proporcionada por varias fuentes de información diferentes: (1) formulación del conocimiento de expertos en forma de reglas, (2) bases de datos de series de tiempo financieras. Se muestra cómo fusionar estos dos diferentes tipos de

conocimiento mediante el uso de métodos neuro-difusos. El modelo contiene una red neuronal de cuatro capas, dos de ellas son ocultas en las cuales se encuentra un motor de reglas basadas en el conocimiento de expertos que tienen un peso importante dentro de la predicción del índice. Una particularidad de este método es que a diferencia de otros métodos de predicción, éste no utiliza como dato de entrada la información histórica del índice que se busca pronosticar, en su lugar hace uso de series de tiempo de otros índices que influyen significativamente en el índice objetivo. Los resultados experimentales demuestran la utilidad de estos nuevos conceptos y se presenta como ventaja la simplicidad del método utilizado y la disponibilidad de herramientas de software para implementarlo.

51. Souto-Maior C.D., Borba J.A. y da Costa N.C.A. (2011). S&P 500 index direction forecasting from 1976 to 2010: *A fuzzy system approach. International Journal of Digital Accounting Research*, Vol.11, 111-134.

Se propone un modelo para el pronóstico de la dirección del índice bursátil S&P500, basado modelos de series de tiempo difusas y en el método propuesto por O'Connor, Remues y Griggs el cual se demuestra que los datos presentan diferentes tendencias y comportamientos en todo el intervalo de la serie de datos. Una diferencia del modelo propuesto con otros modelos similares es que el modelo no devuelve una salida exacta sino una variable lingüística probabilística. Esta salida se puede utilizar con otro tipo de información económica o no económica, para apoyar la toma de decisiones de inversión. Las reglas del motor de inferencia fueron definidas tomando en cuenta el comportamiento histórico del índice. Los resultados empíricos demuestran que el modelo es preciso en el pronóstico y presenta un menor error en el pronóstico que otros modelos similares.

52. Su C.L., Chen C.J. y Yang S.M. (2010). A self-organized neuro-fuzzy system for stock market dynamics modeling and forecasting. *WSEAS Transactions on Information Science and Applications*, Vol.2, 733-745.

Se propone un modelo neuro-difuso con redes neuronales de mapas auto-organizados SOM de cinco capas mediante el uso de indicadores técnicos con el fin de predecir el índice TAIEX. El modelo es evaluado a partir de los siguientes indicadores: el oscilador estocástico %K (medida bruta para formular la idea de impulso del oscilador) y %D (media móvil de %K), media móvil ponderada que utiliza el volumen como factor de ponderación y facilidad de movimiento. Se propone un modelo modificado en el que se agrega un modelo de media móvil con el fin de seleccionar el conjunto de entrada para el modelo neuro-difuso. Los resultados experimentales muestran que el modelo es eficaz y exacto en los pronósticos y que el modelo adicional de media móvil obtiene mejores resultados de predicción.

53. Sun Y. y Bao S.X. (2010). Shanghai composite index forecasting based on multi-stock index. *Journal of Computational Information Systems*, Vol.6, 1111-1116.

Se construye un modelo que utiliza regresión lineal y combina varios factores para pronosticar el índice compuesto de Shanghai. El modelo encuentra el mejor parámetro de adaptación evaluado a través del RMSE, a partir de las diferencias de la volatilidad del índice de mercado objetivo y de la volatilidad de los índices Dow Jones, HK (Heng Seng) y el índice del mercado internacional del oro, con el fin de mejorar los resultados del modelo. Los resultados experimentales muestran que es un modelo simple y eficaz con errores de predicción pequeños.

54. Tangian A. (2008). Predicting DAX trends from Dow Jones data by methods of the mathematical theory of democracy. *European Journal of Operational Research*, Vol.185, 1632-1662.

Se propone un modelo para el pronóstico del índice DAX a partir de su relación con las acciones que componen el índice Dow Jones. El modelo aplica la teoría matemática de la democracia, la cual funciona según los índices de popularidad y universalidad que se utilizan para encontrar representantes socialmente óptimos. Matemáticamente, el objetivo es encontrar datos óptimos que puedan representar el comportamiento de la serie de datos del índice.

55. Tsaih R., Hsu Y. y Lai C.C. (1998). Forecasting S&P 500 stock index futures with a hybrid AI system. *Decision Support Systems*, Vol.23, 161-174.

Se define un modelo híbrido de redes neuronales artificiales para predecir la dirección de los cambios en los precios diarios de las acciones que componen el índice S&P500. El modelo incluye reglas definidas por expertos utilizadas para entrenar la red neuronal. Los resultados experimentales demuestran que el modelo proporciona resultados más aproximados que otros modelos de redes neuronales tradicionales.

56. Wang J. (2009). Stock market volatility and the forecasting performance of stock index futures. *Journal of Forecasting*, Vol.28, 277-292.

Se propone un modelo que integra el concepto de equilibrio general con el fin de pronosticar los índices de mercado TAIEX y SGX MSCI de la bolsa de Taiwán. Adicionalmente, se utiliza un estimador ponderado de media móvil exponencial con el fin de mejorar la capacidad de pronóstico del

modelo. Los resultados experimentales demuestran que el modelo genera pequeños errores de pronóstico y que el estimador de media móvil exponencial mejora la precisión de pronóstico del modelo en comparación con otros estimadores.

57. 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*, Vol.38, 14346-14355.

Se propone un nuevo enfoque para pronosticar los índices de mercado a través de un modelo de redes neuronales de propagación hacia atrás cuyas entradas se descomponen usando wavelets. Se utilizan los precios de cierre mensuales del índice Compuesto de Shanghai para predecirlo y probar el método. Para mostrar la ventaja de este nuevo enfoque para el pronóstico de índices bursátiles, la red neuronal se compara con el modelo de redes neuronales de propagación hacia atrás. Se demuestra que el modelo propuesto presenta una mayor precisión en los resultados de pronóstico y que la transformada wavelet es un mecanismo eficaz de pre-procesamiento de datos que puede ser combinada con cualquier tipo de modelo de pronóstico.

58. Wong W.-K., Bai E. y Chu A.W.-C. (2010). Adaptive time-variant models for fuzzy-time-series forecasting. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol.40, 1531-1542.

Se propone un método de series de tiempo difusas adaptativas variante en el tiempo para el pronóstico del índice de mercado de la bolsa de Taiwán. Este método utiliza el modelo de Wagner para adaptar automáticamente el tamaño de la serie de tiempo difusa con el fin de aumentar la precisión de la predicción en la fase de entrenamiento y utiliza el modelo de Singh para

ejecutar modelos adaptativos y reglas para generar valores los de pronóstico. Los resultados experimentales muestran que el modelo propuesto logra una mejora significativa en la precisión de la predicción con respecto a otros modelos de pronóstico de series de tiempo difusas.

59. Wood D. y Dasgupta B. (1996). Classifying trend movements in the msci U.S.A.capital market index-A comparison of regression, arima and neural network methods. *Computers and Operations Research*, Vol.23, 611-622.

Se describen los resultados de la aplicación de redes neuronales para predecir el índice MSCI de EE.UU a partir de su relación con los índices de la bolsa de Alemania, de Japón y de Reino Unido. El objetivo es poner a prueba la capacidad de una red neuronal de aprendizaje multi-variable con funciones de activación no lineales, con el fin de proporcionar información valiosa a un administrador de una cartera global, necesaria para evaluar oportunidades de inversión en los mercados de valores. Este modelo es comparado con modelos de regresión lineal múltiple y modelos ARIMA. De los resultados de la evaluación se puede concluir que el modelo propuesto es mejor que los modelos con los que fue comparado al producir menor error de pronóstico

60. Yao J. y Herbert J.P. (2009). Financial time-series analysis with rough sets. *Applied Soft Computing Journal*, Vol.9, 1000-1007.

Se propone un modelo de conjuntos rugosos para la predicción del índice de la bolsa de valores de Nueva Zelanda NZX. Este modelo es una técnica emergente para hacer frente a la vaguedad y la incertidumbre en los datos. Los resultados experimentales indican que el pronóstico de los índices bursátiles mediante este método es muy preciso y genera errores pequeños

de pronóstico en comparación con modelos de series de tiempo difusas y redes neuronales.

61. Yu H.-K. (2005). A refined fuzzy time-series model for forecasting. *Physica A: Statistical Mechanics and its Applications*, Vol. 346, 657-681.

Debido a que la longitud de los intervalos de las series de datos en los modelos de series de tiempo difusas afectan significativamente los resultados de pronóstico, se propone un modelo de series de tiempo difusas en el que se desarrolla el concepto de refinamiento como mecanismo para mejorar la precisión del pronóstico del índice de mercado TAIEX. Los resultados demuestran que es un método más efectivo para el pronóstico de índices bursátiles que los modelos de series de tiempo difusas.

62. Yu T.H.-K. y Huarng K.-H. (2008). A bivariate fuzzy time series model to forecast the TAIEX. *Expert Systems with Applications*, Vol.34, 2945-2952.

Este trabajo tiene como objetivo aplicar redes neuronales a los modelos de series de tiempo difusas con el fin de predecir los índices de mercado TAIEX y TAIFEX. Se proponen modelos de dos variables con el fin de mejorar el pronóstico: modelo bivalente de redes neuronales, modelo bivalente de redes neuronales basado en un modelo de series de tiempo difusas y un método bivalente basado en un modelo de series de tiempo difusas con sustitutos. Los resultados experimentales demuestran que el modelo bivalente neuro-difuso sustitutos presenta el menor error de pronóstico, seguido de modelo bivalente de redes neuronales tradicional, mientras que el modelo bivalente neuro-difuso obtuvo el error de pronóstico más alto.

63. Yu T.H.-K. y Huarng K.-H. (2010). A neural network-based fuzzy time series model to improve forecasting. *Expert Systems with Applications*, Vol.37, 3366-3372.

Se propone un modelo de series de tiempo difusas que aplica redes neuronales para el pronóstico del índice de mercado TAIEX de la bolsa de Taiwán, pretende mejorar los resultados de los modelos de series de tiempo difusas convencionales. Todos los grados de pertenencia de las observaciones son tenidos en cuenta. Los resultados experimentales medidos a través del RMSE demuestran que debido a la cobertura de más información, el modelo propuesto supera otros modelos de series de tiempo difusas.

64. 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*, Vol.36, 8849, 8854.

Se presenta un modelo basado en una red neuronal artificial con propagación hacia atrás y quimiotactismo bacteriano para entrenar la red neuronal; con el fin de predecir el índice de mercado S&P500 a corto plazo (1 día) y a largo plazo (15 días). Los resultados experimentales demuestran que el método propuesto tiene menor complejidad computacional que un modelo de redes neuronales de propagación hacia atrás, utiliza menor tiempo de entrenamiento y genera resultados más precisos.

BIBLIOGRAFÍA

- Chi, K. T., Liu, J., y Lau, F. C. . (2009). Winner-Take-All Correlation-Based Complex Networks for Modeling Stock Market and Degree-Based Indexes. *International Symposium on Nonlinear Theory and its Applications*.
- Guresen, E., Daim, T. U., y others. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*.
- Huang, S. C., y Wu, T. K. (2008). Integrating GA-based time-scale feature extractions with SVMs for stock index forecasting. *Expert Systems with Applications*, 35(4), 2080–2088.
- Khurram Jassal, M. (2011). The Effect of Optimization of Error Metrics.
- Kitchenham, B. A. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33, 2004.
- Kitchenham, B. A. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. *Technical Report EBSE-2007-01*.
- Lin, C. S., Haider, A. K. y Chi-Chung, H. (2002). Can the neuro fuzzy model predict stock indexes better than its rivals? *CIRJE-F-165*.
- Lu, C. J., Lee, T. S. y Chiu, C. C. (2009). Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems*, 47(2), 115–125.

- Martinovic, J., Vukmirovic, D. y Vukmirovic, J. (2005). MARKET INDICES AND THEIR APPLICATIONS. Presented at the The 7th Balkan Conference on Operational Research, Constata, Romania.
- Sallehuddin, R., Shamsuddin, S. M. y Hashim, S. Z. M. (2008). Hybridization Model of Linear and Nonlinear Time Series Data for Forecasting (pp. 597–602). IEEE. doi:10.1109/AMS.2008.142
- Shen, W., Guo, X., Wu, C. y Wu, D. (2011). Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm. *Knowledge-Based Systems*, 24(3), 378–385.
- Zhang, D. y Zhang, X. (2009). Study on Forecasting the Stock Market Trend Based on Stochastic Analysis Method. *International Journal of Business and Management*, 4.