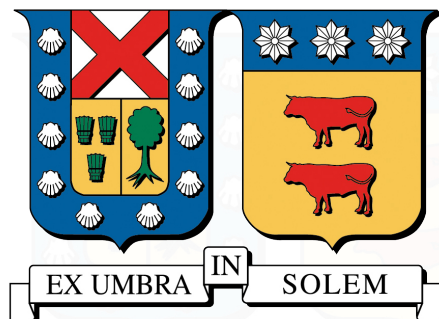


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INDUSTRIAS
VALPARAÍSO - CHILE



**PRONÓSTICO DE VOLATILIDAD DEL MERCADO ACCIONARIO CHILENO
MEDIANTE UN SISTEMA DE ENSAMBLE BASADO EN MODELOS HÍBRIDOS
ANN-EGARCH**

CAMILA FRANCISCA TORRES GONZÁLEZ

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL INDUSTRIAL

PROFESOR GUÍA : DR. WERNER KRISTJANPOLLER R.
PROFESOR CORREFERENTE : SR. KEVIN MICHELL V.

DICIEMBRE 2017

RESUMEN EJECUTIVO

En la presente memoria se aborda la temática del desarrollo de modelos de pronósticos precisos para la predección de la volatilidad financiera.

El propósito del estudio consiste en el diseño e implementación de un sistema de ensamble basado en modelos híbridos ANN-EGARCH mediante técnicas de programación para pronosticar la volatilidad del mercado accionario chileno. Dicho sistema se compone de tres modelos ANN-EGARCH bajo distintos supuestos distributivos que captan la normalidad, la simetría y el exceso de curtosis, y una red neuronal adicional que permite complementar los resultados individuales y otorgar un pronóstico final.

Los resultados demuestran que los sistemas de ensamble son efectivos en la predicción de volatilidad del índice bursátil IPSA del mercado accionario chileno, mejorando el desempeño predictivo de modelos de pronóstico populares tales como el modelo GARCH y los mismos modelos híbridos ANN-EGARCH que lo componen. Estos resultados son robustos y consistentes para diferentes arquitecturas de las redes neuronales que componen el sistema de ensamble y diferentes horizontes de pronóstico empleados.

Palabras Clave: Mercado de Valores, IPSA, Sistema de Ensamble, Redes Neuronales Artificiales, Modelos EGARCH, Pronóstico de Volatilidad

ABSTRACT

This thesis studies the development of accurate forecast models for the prediction of financial volatility.

The use of the study consists of the design and implementation of an assembly system based on hybrid ANN-EGARCH models through programming techniques to forecast the volatility of the Chilean stock market. This system consists of three ANN-EGARCH models under different distributive assumptions that capture normality, symmetry and excess of kurtosis, and an additional neural network that allows complementing the individual results and giving a final forecast.

The results show that the assembly systems are effective in predicting the volatility of the IPSA stock index of the Chilean stock market, improving the predictive performance of popular forecast models such as the GARCH model and the same ANN-EGARCH hybrid models that composes it. These results are robust and consistent for different neural network architectures that make up the assembly system and different prediction horizons used.

Keywords: Stock Market, IPSA, Ensemble System, Artificial Neural Networks, EGARCH Models, Volatility Forecast

ÍNDICE DE CONTENIDOS

| | |
|---|-----------|
| 1. INTRODUCCIÓN | 9 |
| 2. PROBLEMA DE INVESTIGACIÓN | 11 |
| 3. OBJETIVOS | 13 |
| 3.1. Objetivo Principal | 13 |
| 3.2. Objetivos Específicos | 13 |
| 4. ALCANCE | 14 |
| 5. MARCO TEÓRICO | 15 |
| 5.1. Redes Neuronales Artificiales | 15 |
| 5.2. Pronóstico de Series de Tiempo Financieras | 16 |
| 5.3. Modelos GARCH | 17 |
| 5.4. Modelos Híbridos | 20 |
| 5.5. Sistemas de Ensamble | 21 |
| 5.6. Índices Bursátiles y el Mercado Accionario Chileno | 23 |
| 6. METODOLOGÍA | 25 |
| 6.1. Series de Tiempo Financieras | 25 |
| 6.1.1. Estacionariedad | 26 |
| 6.1.1.1. Prueba de Dickey-Fuller Aumentada | 27 |
| 6.1.2. Ergodicidad | 28 |

| | |
|--|-----------|
| 6.2. Pronóstico mediante modelos de la familia GARCH | 28 |
| 6.2.1. Modelo de Media Condicional | 29 |
| 6.2.1.1. Modelo Autorregresivo | 29 |
| 6.2.1.2. Modelo de Media Móvil | 30 |
| 6.2.1.3. Modelo Autorregresivo y de Media Móvil | 30 |
| 6.2.2. Familia de Modelos GARCH | 30 |
| 6.2.2.1. Modelos ARCH | 31 |
| 6.2.2.2. Modelos GARCH | 32 |
| 6.2.2.3. Modelos EGARCH | 33 |
| 6.2.3. Distribuciones del Error | 33 |
| 6.3. Redes Neuronales Artificiales | 35 |
| 6.3.1. La Neurona Biológica | 35 |
| 6.3.2. Unidad de Procesamiento Elemental | 36 |
| 6.3.3. Perceptrón Simple | 38 |
| 6.3.4. Perceptrón Multicapa | 39 |
| 6.3.5. Entrenamiento de la Red Neuronal | 42 |
| 6.4. Sistema de Ensamble | 43 |
| 6.5. Volatilidad Realizada | 44 |
| 6.6. Medidas de Desempeño | 44 |
| 7. TRATAMIENTO DE DATOS | 46 |
| 7.1. Precios de Cierre | 46 |
| 7.2. Rentabilidad | 47 |
| 7.3. Volatilidad Realizada | 49 |
| 8. RESULTADOS | 51 |
| 8.1. Caso Base | 51 |
| 8.2. Análisis de Sensibilidad | 55 |

9. CONCLUSIONES Y RECOMENDACIONES **57**

REFERENCIAS **60**



ÍNDICE DE TABLAS

| | |
|---|----|
| 7.1. Estadísticos descriptivos del IPSA | 47 |
| 7.2. Estadísticos descriptivos de los retornos del IPSA | 48 |
| 7.3. Estadísticos descriptivos de la volatilidad realizada del IPSA a 21 días | 49 |
| 8.1. Resultados de desempeño del modelo de pronóstico GARCH (21 días) | 51 |
| 8.2. Especificaciones de parámetros de los modelos EGARCH seleccionados | 52 |
| 8.3. Especificación de las ecuaciones de los modelos EGARCH empleados | 52 |
| 8.4. Resultados de desempeño de los modelos de pronóstico EGARCH (21 días) | 53 |
| 8.5. Resultados de desempeño de los modelos de pronóstico EGARCH-ANN y sistema de ensamble a 21 días | 53 |
| 8.6. Resultados de desempeño del modelo de pronóstico GARCH (14 y 28 días) | 54 |
| 8.7. Resultados de desempeño de los modelos de pronóstico EGARCH, modelos híbridos y sistema de ensamble a 14 y 28 días bajo distintas distribuciones de error | 54 |
| 8.8. Resultados de desempeño para los modelos de volatilidad de sistemas de ensamble para diferentes números de capas y neuronas por cada en cada red neuronal | 56 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| 6.1. Representación simple de una neurona biológica | 36 |
| 6.2. Representación de la neurona artificial de McCulloch y Pitts | 37 |
| 6.3. Representación de un perceptrón simple | 39 |
| 6.4. Representación de una red feedforward | 40 |
| 6.5. Representación gráfica de cada red ANN-EGARCH utilizada | 41 |
| 6.6. Representación gráfica de la red final del sistema de ensamble utilizada | 41 |
| 6.7. Diseño genérico de un modelo híbrido ANN-EGARCH para la predicción de volatilidad | 43 |
| 6.8. Sistema de ensamble basado en modelos híbridos ANN-EGARCH | 44 |
| 7.1. Comportamiento del IPSA en el periodo 2009-2016 | 47 |
| 7.2. Comportamiento de la Rentabilidad del IPSA durante el periodo 2009-2016 | 48 |
| 7.3. Volatilidad Realizada del IPSA durante el 2010-2016 | 50 |
| 8.1. Pronóstico de Volatilidad del Sistema de Ensamble vs. Volatilidad Realizada a 21 días | 54 |

1. INTRODUCCIÓN

La volatilidad es una característica inherente de las series de tiempo financieras y se utiliza principalmente como una medida de riesgo en instrumentos de inversión, de manera que cuanto más volátil sea, mayor riesgo tendrá el inversionista de no recibir los rendimientos esperados.

Los mercados financieros se consideran sistemas dinámicos complejos y no lineales (Bosarge, 1993) que pueden verse influidos por diversos factores políticos, económicos y las expectativas de los operadores, que pueden aumentar aún más su comportamiento volátil, conocido como volatilidad. Una clara prueba de ello es la fuerte caída del IPSA ocurrida luego de los sorprendentes resultados de la elección presidencial de noviembre del 2017, en medio de la creciente incertidumbre política de cara a la segunda vuelta.

Pronosticar la volatilidad constituye un paso esencial en muchos procesos de toma de decisiones financieras y planificación futura, ya que proporciona datos concretos para que los participantes del mercado puedan decidir de la forma más informada y precisa posible actividades tales como la fijación de precios de activos financieros, administración y optimización de portafolios de inversiones, caracterización de la dinámica de los mercados y la gestión de riesgo. Además, dada la progresiva apertura de los mercados bursátiles alrededor del mundo, estos se han transformado en alternativas de inversión accesibles no solo para inversores e instituciones financieras, sino también para gente común y corriente.

De esta forma, la elaboración de la presente memoria se justifica en la gran atención que recibe el término de volatilidad dentro mercado accionario chileno y los mercados accionarios alrededor del mundo, donde se requiere una medida de riesgo al momento de realizar transacciones.

Dado el nivel de importancia que presenta la predicción de volatilidad en los mercados financieros, esta temática ha retenido el interés de una gran cantidad de docentes e investigadores, especialmente desde la introducción de los

primeros modelos de la familia GARCH, hace más de 35 años. Desde entonces se ha desarrollado y aplicado una gran cantidad de modelos de volatilidad para predecir el comportamiento de mercados accionarios, donde destaca la utilización de enfoques híbridos como los modelos ANN-GARCH durante las últimas décadas, los cuales mejoran significativamente el desempeño de los modelos GARCH en la literatura financiera. Sin embargo, el desempeño de los modelos híbridos es perfectible, existiendo más oportunidades para mejorar los modelos existentes y aumentar aún más el desempeño predictivo. Como un intento de mejorar la precisión de los pronósticos, algunos investigadores han propuesto la utilización de sistemas de ensamble en los cuales se combinan distintos modelos predictivos como los ya mencionados, con la finalidad de obtener información complementaria de los modelos que lo componen, mejorando el desempeño predictivo de cada modelo individual.

En la presente memoria se estudia el pronóstico de la volatilidad de los retornos del IPSA, el cual constituye el principal indicador del mercado bursátil chileno. Los pronósticos se obtienen mediante la implementación de un sistema de ensamble basado en modelos híbridos EGARCH y de redes neuronales artificiales (ANN-EGARCH) con distintos supuestos distributivos, buscando mejorar el desempeño de modelos de pronóstico populares como los modelos de la familia GARCH y modelos híbridos individuales.

Los resultados obtenidos son consistentes y robustos, y demuestran que el uso de sistemas de ensamble es efectivo en la previsión de volatilidad del IPSA, mejorando el desempeño de los modelos de pronóstico ampliamente utilizados para diferentes horizontes de pronósticos y diversas arquitecturas de las redes neuronales que conforman el sistema de ensamble analizado.

La presente memoria se organiza de la siguiente forma: la primera sección contiene una introducción al tema seguida por la descripción del problema de investigación, los objetivos en base a la problemática planteada, y el alcance del estudio. Posteriormente, se presenta una revisión de la literatura para mostrar el mercado teórico de esta investigación. Posteriormente se define la metodología utilizada, se analiza la estadística descriptiva de las series de tiempo empleadas, para finalizar con los resultados obtenidos y las conclusiones y recomendaciones realizadas.

2. PROBLEMA DE INVESTIGACIÓN

La volatilidad es una característica inherente a las series de tiempo financieras, y dado que cambia constantemente con el tiempo, los modelos de series de tiempo tradicionales que suponen varianza homoscedástica no son adecuados para modelar series de tiempo financieras.

Los modelos de la familia GARCH son populares y ampliamente utilizados en aplicaciones financieras para modelar y pronosticar la volatilidad. Estos asumen una estructura de correlación lineal de los datos y se limitan a distribuciones normales de variables y error. La evidencia empírica sugiere que otra de las características de la volatilidad financiera es su respuesta asimétrica frente a las innovaciones pasadas, comportamiento que es captado por modelos asimétricos de volatilidad como el EGARCH (Donaldson y Kamstra, 1997). Sin embargo, las series temporales financieras presentan estructuras no lineales que los modelos de la familia GARCH no pueden capturar.

Para mejorar la capacidad de estos modelos, la literatura ha propuesto combinarlos con redes neuronales artificiales por medio de modelos híbridos, en particular con redes feedforward, que corresponde a modelos no lineales flexibles y con la capacidad de describir funciones de distintas clases, dadas su propiedad de aproximadores universales (Hornik et al., 1989).

Numerosos trabajos recientes han demostrado que los modelos híbridos GARCH-ANN funcionan mejor que los procesos individuales de la familia GARCH para pronosticar la volatilidad de los mercados financieros. Aun así, la precisión de pronóstico del modelo híbrido GARCH-ANN no es perfecto. Además, como se ilustra en la revisión de la literatura en este documento, hay casos en los que la capacidad de pronóstico de volatilidad financiera ha sido mejorada a través del uso de sistemas de ensamble.

Por otro lado, el supuesto distributivo más utilizado en los modelos GARCH es que los errores presentan una distribución normal, sin embargo, generalmente las series de tiempo financieras exhiben excesiva asimetría y

curtosis. Por esta razón al combinar la predicción de modelos de pronóstico con los mismos datos, pero variando el supuesto distributivo del error, se logra una mejor predicción gracias a los resultados complementarios de los modelos individuales.

De esta forma en la presente memoria se propone el diseño e implementación de un sistema de ensamble basado en modelos ANN-EGARCH para pronosticar la volatilidad de los retornos en un mercado emergente como el chileno, constituyendo un enfoque para mejorar la capacidad predictiva de modelos tradicionales como GARCH o modelos híbridos. El sistema conjunto consta de dos componentes: el primero de ellos consta de tres modelos híbridos que realizan pronósticos de volatilidad basados en supuestos distributivos de los errores normal, t-Student y GED, de forma que cada modelo híbrido es capaz de capturar la normalidad, el exceso de asimetría y el exceso de curtosis en los datos para lograr la complementariedad. Particularmente se elige el modelo EGARCH ya que puede modelar variaciones asimétricas en los datos que el clásico modelo GARCH no puede capturar. El segundo componente del sistema de ensamble corresponde a una red neuronal adicional a la cual se ingresan los resultados de los tres modelos híbridos, con la finalidad de combinarlos y proporcionar el pronóstico de volatilidad final.

Para efectos de contrastar el desempeño de los modelos EGARCH, modelos híbridos y el sistema de ensamble, estos se comparan con los pronósticos de volatilidad de un modelo GARCH(1,1), el cual es ampliamente utilizado en la literatura financiera. La comparación de los modelos se evalúa en términos de su cercanía a la volatilidad realizada, utilizada como proxy de la verdadera volatilidad, calculando el error absoluto medio (MAE) y la raíz del error cuadrático medio (RMSE), que son populares en la medición de rendimiento de modelos predictivos.

En cuanto a la repercusión de la problemática presentada, esta afecta principalmente a los especuladores e inversionistas que realizan operaciones de compra y venta de activos financieros en el mercados accionario chileno, para quienes resulta crucial contar con pronósticos adecuados de la volatilidad, debido a que un exceso de volatilidad dificulta la planificación financiera y podría significar enormes pérdidas y/o ganancias.

En palabras de [Zamora, L. and Guarda, D. and Lay, N. \(2015\)](#), se debe tener en cuenta que, “como participantes del mercado bursátil, ya sea en el rol de especulador, arbitrajista o coberturista, siempre se busca la mejor manera de predecir el comportamiento de los activos financieros, de manera de saber cómo comportarse hoy, para así obtener beneficios extraordinarios a futuro”.

3. OBJETIVOS

3.1. Objetivo Principal

Definir y diseñar un sistema de ensamble basado en modelos híbridos ANN-EGARCH con diferentes supuestos distributivos mediante técnicas de programación para pronosticar la volatilidad del mercado accionario chileno.

3.2. Objetivos Específicos

Investigar e identificar tipos de modelos utilizados con anterioridad para el pronóstico de volatilidad en aplicaciones financieras, con la finalidad de formar una base de conocimiento respecto a la variedad de modelaciones realizables.

Formular, programar y entrenar modelos matemáticos basados en ANN óptimos para la predicción, alimentados con pronósticos de modelos EGARCH, basados en datos históricos del mercado accionario chileno.

Determinar el potencial de los sistemas de ensamble en la mejora de las predicciones de volatilidad obtenidas con modelos ampliamente utilizados en aplicaciones financieras.

4. ALCANCE

Como se señala en los objetivos, el trabajo realizado en la presente memoria busca aportar evidencias respecto a la potencia del uso de sistemas de ensamble basados en modelos híbridos ANN-EGARCH como herramienta para mejorar el pronóstico de la volatilidad del mercado de valores.

Para tal efecto, utilizando datos históricos del precio de cierre diario del IPSA, se realizan estimaciones y pronósticos de modelos EGARCH calculados mediante el uso de ventanas móviles bajo distintos supuestos distributivos de los errores. Con los resultados obtenidos se construyen tres modelos híbridos ANN-EGARCH, que permiten capturar normalidad, exceso de asimetría y el exceso de curtosis en los datos según el supuesto distributivo utilizado, cuyos pronósticos son combinados por medio de una red neuronal adicional que genera el pronóstico final del sistema de ensamble, para lograr la complementariedad. La capacidad predictiva de cada modelo por separado y del sistema de ensamble se evalúa en base a medidas de desempeño que determinan la cercanía entre los pronósticos generados y la volatilidad realizada.

Para desarrollar el sistema de ensamble de predicción se utilizan dos herramientas. La estimación y elaboración de los pronósticos de los modelos EGARCH se realiza mediante el programa Eviews 9, mientras que el diseño, entrenamiento y simulación de las redes neuronales artificiales se realizan por medio del programa MATLAB.

Considerando que el uso de sistemas de ensamble basados en modelos híbridos para pronosticar la volatilidad financiera y el uso complementario de distintos supuestos distributivos constituye un enfoque poco explorado en la literatura financiera, el tipo de estudio realizado en la presente memoria se considera de tipo exploratorio.

5. MARCO TEÓRICO

5.1. Redes Neuronales Artificiales

Dentro del campo financiero, las redes neuronales han sido utilizadas en una amplia variedad de aplicaciones, tales como la predicción del comportamiento de acciones en el mercado de valores, predicción de tipos de interés, análisis de carteras, evaluación y proceso de créditos, calificación de obligaciones, análisis de mercados de capitales, modelos de riesgos para inversiones, negociación de commodities e índices de futuro de renta variable y predicción de la volatilidad de opciones (Jiménez, 2000).

Revisiones de aplicaciones de redes neuronales en finanzas pueden encontrarse en las investigaciones de Trippi y Turban (1992) y Li y Ma (2010), donde se resalta la superioridad del desempeño de las redes neuronales con respecto a los métodos econométricos y otros modelos lineales. En este contexto, el éxito de las redes neuronales puede atribuirse a su capacidad de aproximar arbitrariamente bien una gran cantidad de funciones, mientras que sólo se requiere un pequeño número de parámetros para ser estimados en relación con el tamaño de la muestra, tal como señalan Rumelhart et al. (1986), Hecht-Nielsen (1988), Gybenko (1989), Carroll y Dickinson (1989), Funahashi (1989), Hornik et al. (1989) y White (1992) en sus respectivas investigaciones.

El desarrollo de los primeros modelos matemáticos de redes neuronales detallados data desde hace más de sesenta años, con los trabajos de McCulloch y Pitts (1943), Hebb (1949), Rosenblatt (1958), Widrow y Hoff (1960) y Posch (1968), entre otros. Sin embargo, cuando se pretende analizar un conjunto de datos con relaciones complejas y condiciones adicionales, el nivel de procesamiento requerido se eleva considerablemente, situación por la cual los primeros intentos por utilizar este tipo de modelos fueron abandonados debido a la limitada capacidad de los sistemas de cómputo disponibles en la época. El interés por el uso de redes neuronales se retoma durante la década de los

ochenta, de mano de los avances en la potencia de cálculo de los nuevos computadores y un mayor conocimiento sobre el funcionamiento del cerebro humano. Así, los estudios realizados en este campo por [Hopfield \(1982\)](#), [Hopfield \(1984\)](#), [Hopfield y Tank \(1986\)](#), [Rumelhart et al. \(1986\)](#), [Sejnowski y Rosenberg \(1988\)](#), [Feldman y Ballard \(1982\)](#), [Grossberg \(1987\)](#), entre otros, permitieron el resurgimiento de esta nueva área del conocimiento.

[Hutchinson et al. \(1994\)](#) desarrollan un modelo de redes neuronales para estimar los precios del índice S&P 500, mientras que [Komo et al. \(1994\)](#) pronostican el Dow Jones Industrial Index. [Ormoneit y Neuneier \(1996\)](#) utilizan un perceptrón multicapa y una red neuronal de estimación de densidad para pronosticar la volatilidad del índice DAX alemán, mientras que [Villada et al. \(2012\)](#) proponen un modelo basado en redes neuronales artificiales para el pronóstico de los precios de dos de las principales acciones transadas en mercado de valores colombiano. Otros autores que desarrollaron trabajos de predicción de índices bursátiles por medio de redes neuronales son [Zhu et al. \(2008\)](#) y [Guresen et al. \(2011\)](#).

Desde la primera modelación matemática de una neurona artificial realizada por [McCulloch y Pitts](#) en 1943, se ha generado una gran variedad de modelos de redes neuronales, entre los que destaca la arquitectura conocida como redes de propagación hacia atrás, o “Back Propagation”, que ha sido ampliamente utilizada en aplicaciones financieras por su habilidad para entrenar perceptrones multicapa y generalizar resultados para un amplio número de problemas ([Del Carpio, 2005](#)), destacando su uso en el pronóstico de precios y rentabilidades accionarias ([Chaigusin et al., 2008](#); [Wang et al., 2011](#)).

Luego de realizar un recorrido general respecto a las redes neuronales y sus principales aplicaciones, resulta necesario definir el contexto en el cual se realizan las transacciones bursátiles: el mercado de capitales, además de conocer las principales características de las series de tiempo financieras y sus métodos de predicción, con el fin de comprender de qué manera el empleo de redes neuronales conforma una mejora en el pronóstico de estas series.

5.2. Pronóstico de Series de Tiempo Financieras

Según la definición entregada por el [Ministerio de Hacienda Gobierno de Chile](#), “el mercado de capitales es aquél en que se coordinan la oferta y la demanda de capital financiero en sus más variadas formas (dinero y valores o activos financieros), surgiendo intermediarios especializados, regulados y controlados. Es decir, es el conjunto de regulaciones,

instituciones, prácticas e individuos que forman una infraestructura tal que permite a los oferentes de recursos vender dichos recursos a los demandantes de estos.” Los activos más transados en el mercado de capitales son las acciones, las cuales corresponden a títulos que acreditan un porcentaje de propiedad de una empresa, y es en este contexto en el que surge la importancia por parte de los inversores de poder predecir los movimientos futuros que tendrán las acciones tanto en su precio como en su rentabilidad, con la finalidad última de maximizar sus beneficios.

Una de las variables financieras más utilizada en predicción corresponde a los precios de cierre accionarios, sin embargo, una mejor opción es utilizar la rentabilidad de una acción, ya que es un indicador por excelencia de la performance de una inversión. Otro punto a favor para la utilización de la rentabilidad como variable financiera es el hecho de que al tomar la rentabilidad logarítmica de una acción, la serie obtenida habitualmente es estacionaria, principio en el cual se amparan la gran mayoría de los modelos de predicción.

En particular, la predicción de series de tiempo financieras presenta un desafío mayor que otro tipo de series, debido principalmente a la aparente aleatoriedad con que las variables financieras toman valores, razón por la cual la predicción financiera cuestiona la hipótesis de los mercados eficientes (Fama, 1970). Además, estas series de tiempo por lo general son muy ruidosas, presentan un comportamiento no lineal, y grandes cúmulos de volatilidad (heteroscedasticidad) que no son considerados en los modelos estadísticos tradicionales de predicción.

Teniendo en cuenta el contexto del mundo accionario y las principales características de las series de tiempo financieras, se procede a analizar las principales investigaciones que se han realizado hasta la fecha en el ámbito de la predicción de volatilidad financiera, donde los modelos de tipo GARCH y de redes neuronales artificiales, ya sea en forma individual o conjunta por medio de modelos híbridos, han sido los favoritos y más utilizados en la literatura.

5.3. Modelos GARCH

Engle propone en 1982 el modelo de pronóstico ARCH, que deja de lado la suposición de que la volatilidad es una constante. ARCH (Autorregresive Conditional Heteroscedasticity) es un modelo estocástico en el cual además del valor de la media, se expresa el valor de la varianza condicional como función lineal del cuadrado de los valores rezagados de la media del modelo, el cual fue utilizado por Engle (1982) para pronosticar la media y varianza de la

inflación en Reino Unido.

Posteriormente el modelo ARCH fue mejorado a través del modelo GARCH (Generalized Autoregressive Conditional Heterocedasticity) propuesto por [Bollerslev \(1986\)](#), el cual expresa la varianza condicional en base al valor de las medias rezagadas y de las varianzas rezagadas, de forma que las varianzas rezagadas actúan como un término suavizado en la predicción. [Bollerslev \(1986\)](#) realiza un caso aplicado en el cual modela el crecimiento del deflactor del PIB de Estados Unidos a través de modelos ARCH y GARCH, donde se obtiene que el modelo GARCH, además de proporcionar un mejor ajuste, presenta una cantidad de rezagos más razonables que el modelo ARCH.

La volatilidad financiera se comporta de forma distinta ante del signo de las perturbaciones de periodos rezagados, y los modelos ARCH y GARCH no tienen la capacidad de modelar el efecto de estas asimetrías ya que utilizan los cuadrados de los rezagos de la media. Debido a lo anterior, surgen varios modelos variantes del GARCH que sí logran modelar la asimetría de la volatilidad, entre los cuales se encuentran los modelos EGARCH ([Nelson, 1991](#)), TGARCH ([Zakoian, 1994](#)) y PGARCH ([Ding et al., 1993](#)).

Destaca especialmente la capacidad predictiva del modelo EGARCH, en el cual [Nelson \(1991\)](#) modifica la desviación estándar condicional a una función lineal por tramos, permitiendo de esta forma considerar las diferentes reacciones de la volatilidad ante los diferentes signos de los rezagos, cuyo ejemplo empírico se realiza para pronosticar el índice bursátil francés CAC 40.

Numerosos investigadores han desarrollado estudios sobre los modelos ARCH y la diversa familia de variantes y extensiones aplicadas a series de tiempo financieras, entre los que destacan [Engle y Bollerslev \(1986\)](#), [Bera y Higgins \(1993\)](#), [Bollerslev et al. \(1992\)](#) y [Bollerslev et al. \(1994\)](#).

Autores como [Taylor \(1986\)](#) y [Akgiray \(1989\)](#) presentan las primeras pruebas del poder predictivo de los modelos GARCH, mientras que en las siguientes décadas se desarrolló una gran cantidad de investigaciones que probaron el poder predictivo del modelo GARCH frente a otros métodos de series de tiempo para pronosticar la volatilidad de índices bursátiles en diversos mercados financieros. De dichas investigaciones, varias concluyen que el modelo EGARCH entrega los mejores pronósticos ([Pagan y Schwert, 1990](#); [Cao y Tsay, 1992](#); [Cumby et al., 1993](#); [Emenike, 2010](#); [Awartani y Corradi, 2005](#); [Ludlow y Mota, 2006](#); [Floros, 2008](#); [Yalama y Sevil, 2008](#); [Casas y Cepeda, 2008](#); [Miron y Tudor, 2010](#); [Su, 2010](#); [Liu y Hung, 2010](#); [Abdelaal, 2011](#); [Ezzat, 2012](#)).

En cuanto a mercados emergentes, puede mencionarse el trabajo de [Hernández \(2009\)](#), quien modela el comportamiento histórico del Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores por medio de modelos de la familia GARCH, concluyendo que el modelo EGARCH(1,1) posee la mejor capacidad predictiva. [Camargo et al. \(2012\)](#) obtienen resultados similares al modelar la volatilidad del Índice General de la Bolsa de Valores de Colombia (IGBC).

[Miron y Tudor \(2010\)](#) comparan diferentes modelos asimétricos de la familia GARCH (EGARCH, PGARCH y TGARCH) bajo distribuciones de error normal, t-Student y GED, encontrando que los modelos con errores normales no son capaces de capturar completamente la leptocursis en series de retornos de acciones de EE.UU y de Rumania, mientras que los modelos con errores t-Student y GED proporcionan una mejor descripción de la volatilidad condicional. Además, se encontró que las estimaciones de volatilidad dadas por el modelo EGARCH exhiben generalmente errores de pronóstico más bajos, y por lo tanto, son más precisas que las estimaciones dadas por los otros modelos GARCH asimétricos.

Como estrategia de control del riesgo y para la valoración de derivados financieros sobre acciones e índices, [Yalama y Sevil \(2008\)](#) emplean diferentes clases de modelos GARCH con el fin de pronosticar la volatilidad diaria de los principales índices de los mercados de 10 países diferentes, obteniéndose un mejor desempeño de los modelos de volatilidad asimétrica con respecto a los modelos históricos. Más recientemente, [Xiao \(2016\)](#) modelan y pronostican la volatilidad del índice Russell 3000 utilizando varios modelos GARCH, obteniendo los mejores resultados con los modelos EGARCH(1,1) y TARCH(1,1). [Ali y Talukdar \(2017\)](#) examinan el agrupamiento de volatilidad, la persistencia de la volatilidad y el efecto de apalancamiento en seis índices sectoriales de la Bolsa Nacional de Valores de La India por medio de modelos de la familia GARCH, mientras que [Rajvanshi et al. \(2017\)](#) analizan el poder predictivo de los modelos GARCH para pronosticar la volatilidad de los retornos del índice Nifty 50.

Habiendo realizado un recorrido por los principales modelos de la familia GARCH y sus aplicaciones en el ámbito financiero, a continuación se presenta la utilización de modelos híbridos que combinan modelos de la familia GARCH y redes neuronales artificiales, aplicados al pronóstico de volatilidad financiera.

5.4. Modelos Híbridos

[White \(1988\)](#) proporciona el primer artículo en el que se aplican modelos de redes neuronales al ámbito financiero, utilizando técnicas de modelación y aprendizaje de redes neuronales artificiales para buscar y decodificar patrones no lineales en el movimiento de los retornos de las acciones de IBM. A partir de entonces surge una gran variedad de estudios en los cuales se intenta mejorar los resultados de los modelos de la familia GARCH para pronosticar la volatilidad de los retornos financieros por medio de redes neuronales.

En palabras de [Nazarian et al. \(2013\)](#), “en los últimos años, la aplicación de modelos híbridos se ha vuelto muy popular en la predicción de variables económicas, especialmente en los mercados financieros, ya que, en primer lugar, estos modelos tienen una estructura no lineal que es capaz de identificar, modelar y predecir procesos lineales y no lineales. En segundo lugar, los mercados financieros tienen un comportamiento complicado y muy volátil debido a la claridad de la información, la liquidez y también la existencia de especuladores e inversionistas con diferentes decisiones. Esto lleva a una disminución de la precisión de los modelos lineales, preparando el terreno para el uso de modelos no lineales, especialmente los modelos híbridos”. En este contexto, [Donaldson y Kamstra \(1997\)](#) introducen un modelo GARCH basado en la literatura de las redes neuronales para predecir los retornos de cuatro de las principales índices de mercados internacionales: S&P500, NIKKEI, FTSE y TSEC, observando que el modelo híbrido captura efectos de la volatilidad que los modelos GARCH, EGARCH y GJR-GARCH pasan por alto.

[Roh \(2007\)](#) propone la utilización de modelos híbridos de redes neuronales y modelos de volatilidad GARCH y EWMA para predecir la volatilidad del índice de Corea Kospi 200 desde dos puntos de vista: desviación y dirección, obteniéndose los mejores resultados con el modelo ANN-EGARCH. [Bildirici y Ersin \(2009\)](#) desarrollan un modelo ANN-APGARCH para mejorar la capacidad de pronóstico del modelo APGARCH al evaluar la volatilidad de los retornos diarios del mercado accionario de Estambul. Por su parte, [Hajizadeh et al. \(2012\)](#) utilizan un modelo híbrido para predecir la volatilidad de los retornos del índice S&P500 obteniendo que de entre los modelos GARCH, EGARCH y GJR-GARCH, el que mejor estima la volatilidad es un modelo de tipo EGARCH, el cual se introduce como input para alimentar a una red neuronal en conjunto con otras variables de entrada, simulando los datos basados en el modelo EGARCH con las variables seleccionadas, para después entrenar el modelo de redes neuronales con la

volatilidad realizada. En su investigación, [Kristjanpoller et al. \(2014\)](#) desarrollan un modelo híbrido ANN-GARCH aplicado a mercados financieros emergentes para el pronóstico de la volatilidad de los retornos de los índices bursátiles latinoamericanos de Brasil, Chile y México, obteniéndose que los modelos híbridos pueden mejorar la capacidad de predicción de los modelos GARCH, además se ser consistentes y robustos en sus resultados.

Estos trabajos han demostrado que los modelos híbridos ANN-GARCH tienen mejores resultados que los modelos de la familia GARCH de forma individual en la predicción de la volatilidad de los mercados financieros.

En resumen, existe una gran cantidad de trabajos que demuestran la superioridad de los modelos híbridos GARCH-ANN sobre los modelos de la familia GARCH individuales para pronosticar la volatilidad de mercados financieros. A pesar de lo anterior, la precisión de los pronósticos de modelos híbridos ANN-GARCH no es perfecta, introduciéndose la idea de combinar varios modelos híbridos en un sistema de ensamble, como se presenta en la siguiente sección.

5.5. Sistemas de Ensamble

En los últimos años, los investigadores han propuesto métodos para mejorar la capacidad predictiva de los modelos de volatilidad, entre los cuales se encuentra la utilización de sistemas de ensamble, que corresponden a un método que combina modelos predictivos como los presentados anteriormente formando conjuntos y obteniendo un pronóstico final en base a la combinación de los pronósticos de cada modelo individual. La idea principal del uso de sistemas de ensamble es que la combinación de los diferentes modelos individuales puede mejorar el desempeño predictivo de cada modelo por separado.

Hay muchas razones aceptadas para explicar por qué la combinación de pronósticos es efectiva: en primer lugar, cada predicción puede ser obtenida usando diferentes fuentes de información o cada modelo se basa en diferentes supuestos sobre el tipo de relación en los datos. Además, cada modelo no está en la capacidad de identificar el proceso real de generación de los datos, pero captura información complementaria en relación con otros modelos alternativos ([Terui y Van Dijk, 2002](#)).

[Orozco y Velásquez \(2013\)](#) presentan un sistema compuesto basado en una serie de modelos GARCH anclados

a un conjunto de datos de series de tiempo que emplean diferentes funciones de pérdida para capturar diferentes características de la volatilidad. Dicho modelo combina las predicciones individuales mediante el uso de la media aritmética simple o de una red neuronal, y se emplea para predecir la rentabilidad mensual del índice S&P500. En este contexto, [Lahmiri \(2014\)](#) desarrolla un sistema de ensemble basado en cuatro redes neuronales con diferentes arquitecturas para predecir los retornos del índice NASDAQ. [Gonzalez et al. \(2015\)](#) presenta el diseño de un sistema de ensemble basado en máquinas de vectores de soporte (SVM) para pronosticar la dirección del movimiento del índice de la Bolsa de Valores de Sao Paulo, utilizando un algoritmo genético para seleccionar y optimizar los parámetros de cada uno de los componentes de dicho sistema, mientras que [Ling et al. \(2014\)](#) implementan un sistema de ensemble basado en el ajuste de modelos híbridos RBF (redes neuronales de función de base radial) para predecir la volatilidad del precio de las acciones en el mercado bursátil de China.

[Lahmiri y Boukadoum \(2015b\)](#) presenta un sistema de ensemble para la predicción de retornos del mercado bursátil, donde la transformada continua de wavelet (CWT) es utilizada para analizar la serie de retornos, mientras que tres redes neuronales CWT-ANN con diferentes funciones wavelet son usadas para procesar los coeficientes basados en la CWT, determinar los pesos óptimos del sistema de ensemble y proporcionar pronósticos que son combinados por medio de una última red que entrega los pronósticos finales. Dicho sistema se probó en tres mercados bursátiles asiáticos, obteniéndose que los sistemas de ensemble superan a cada modelo CWT-ANN individual.

Por su parte, la presente memoria se basa en el trabajo de [Lahmiri y Boukadoum \(2015a\)](#), quienes diseñan e implementan un sistema de ensemble para pronosticar la volatilidad de los mercados bursátiles, compuesto de tres modelos híbridos EGARCH-ANN cada uno con una distribución del error diferente (normal, t-student y GED), cuyos pronósticos son combinados por una red neuronal adicional que otorga el pronóstico final del sistema.

Habiendo incursionado por las principales investigaciones realizadas en el ámbito del pronóstico de volatilidades financieras, se procede a adentrarse más en el foco de este estudio, a saber, el mercado accionario chileno, definiendo el índice que se utilizará en el modelo de predicción propuesto: el IPSA. Además, se analizan los principales estudios realizados en el marco del mercado accionario chileno y las características que este posee.

5.6. Índices Bursátiles y el Mercado Accionario Chileno

Según la definición entregada por la [Bolsa de Comercio de Santiago](#), los índices bursátiles “son un promedio de las cotizaciones de una cartera de acciones representativas de un mercado, y se calculan como un promedio ponderado por capitalización de mercado. Los índices entregan una idea general del comportamiento del mercado, y constituyen la principal referencia (benchmark) para comparar el rendimiento de distintos instrumentos financieros”.

Existen índices prácticamente para cada uno de los sectores de la economía y del mercado de acciones, destacando a nivel mundial índices como el Dow Jones Industrial Average, el FTSE, el S&P500 y el Nikkei 225.

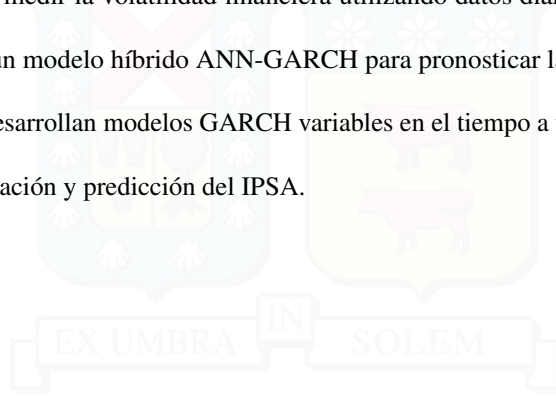
El Índice de Precio Selectivo de Acciones (IPSA), creado en 1977, constituye el principal índice bursátil chileno, y es elaborado por la Bolsa de Comercio de Santiago considerando la rentabilidad de las 40 acciones chilenas con mayor presencia bursátil, con la finalidad de reflejar las variaciones de precio de los títulos más activos del mercado en un contexto de corto plazo. “La capitalización bursátil de las compañías que componen el IPSA equivale a USD 211.774 MM, lo que representa a la fecha un 70, 53 % de la capitalización bursátil de todas las empresas listadas en Bolsa. La lista de empresas componentes es estable y las acciones se transan activamente, lo que implica una tasa de recambio de las empresas y un alto nivel de liquidez” ([Bolsa de Comercio de Santiago](#)).

El mercado accionario chileno es un sistema dinámico altamente no lineal, que puede verse afectado por una diversidad de factores, como por ejemplo las tasas de interés, la inflación, el entorno económico y factores políticos ([Romero-Meza et al., 2007](#)), constituyendo el tercer mercado latinoamericano con mayor volumen de transacciones ([Kristjanpoller y Caballero, 2012](#)).

A pesar de tratarse de un mercado emergente, destacan varios estudios realizados sobre el mercado accionario chileno, como las investigaciones que evidencian la presencia de anomalías de calendario en el mercado chileno ([Marshall y Walker, 2000](#); [Kristjanpoller, 2009](#); [Kristjanpoller y Muñoz, 2012](#)). Existe una gran cantidad de estudios que examinan la relación entre retornos accionarios y volúmenes transados en el mercado chileno ([Gregoire, 1985](#); [Urrutia, 1994](#); [Saatcioglu y Starks, 1998](#); [Parisi y Acevedo, 2001](#); [Parisi y Vasquez, 2000](#); [Romero-Meza et al., 2007](#)). [Aranda y Jaramillo \(2010\)](#) encuentran evidencia significativa de existencia de no linealidades en ambas series financieras, mientras que [Kristjanpoller y Caballero \(2012\)](#) evidencian estadísticamente la existencia de influencia significativa entre los

volúmenes transados y la asimetría del mercado accionario chileno.

En el ámbito de predicciones de series financieras del mercado chileno, [Parisi et al. \(2002\)](#) realizan un modelo de redes neuronales para evaluar la capacidad predictiva del signo de las variaciones semanales del IPSA, [Alfaro y Silva \(2008\)](#) examinan diversas formas de medir la volatilidad financiera utilizando datos diarios del IPSA, mientras que [Kristjanpoller et al. \(2014\)](#) realizan un modelo híbrido ANN-GARCH para pronosticar la volatilidad de los retornos accionarios, y [Ferreira et al. \(2017\)](#) desarrollan modelos GARCH variables en el tiempo a través de una representación del estado espacio aplicado a la estimación y predicción del IPSA.



6. METODOLOGÍA

6.1. Series de Tiempo Financieras

Una serie de tiempo corresponde a una secuencia cronológica de observaciones de una variable en particular (Bowerman y O'Connell, 1993), cuyo análisis se realiza con el motivo de describir y explicar el comportamiento de la variable en el tiempo, predecir los valores que adoptará en el futuro o controlar un rango específico en el cual se espera que la variable tome valores. Respecto a los componentes clásicos de una serie de tiempo, pueden mencionarse la tendencia o cambio de media, la presencia de ciclos representada por movimientos hacia arriba o hacia abajo alrededor de la línea de tendencia, variaciones estacionales que corresponden a patrones históricos que se repiten cada cierto tiempo, y fluctuaciones irregulares que no siguen un patrón reconocible (Gujarati, 2004). Por lo tanto, una serie de tiempo financiera puede definirse como una sucesión estadística de los valores de una variable financiera.

La serie de tiempo financiera analizada en la presente memoria corresponde a los precios de cierre diarios del IPSA, la cual se denota como P_t , y su unidad corresponde a pesos chilenos. Dado que para hacer una inferencia estadística y predecir el comportamiento futuro de una serie de tiempo es necesario que esta sea estacionaria y posea ergodicidad, se trabaja con la rentabilidad logarítmica de los precios r_t , como se muestra en la Ecuación 6.1.

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (6.1)$$

Considerando que al igual que los modelos ARMA y EGARCH utilizados en la presente memoria, la mayoría de los modelos de predicción se amparan en los conceptos de estacionariedad y ergodicidad, se procede a desarrollar ambos conceptos.

6.1.1. Estacionariedad

En términos generales, se dice que un proceso estocástico es estacionario si sus propiedades no dependen del tiempo. Un ejemplo común de una serie no estacionaria es la caminata aleatoria:

$$Y_t = Y_{t-1} + \varepsilon_t \quad (6.2)$$

Donde ε_t es un término de error estacionario. Sin embargo, la caminata aleatoria es una serie que se vuelve estacionaria en su primera diferencia, como se muestra en la [Ecuación 6.3](#).

$$Y_t - Y_{t-1} = \Delta Y_t = \varepsilon_t \quad (6.3)$$

Una serie que se vuelve estacionaria al diferenciarse se denomina integrada, y se representa como $I(d)$, donde d es la cantidad de veces que la serie debe diferenciarse hasta volverse estacionaria (orden de integración), o en otras palabras, la cantidad de raíces unitarias contenidas en la serie.

Existen dos tipos de definiciones de estacionariedad. Un proceso fuerte o estrictamente estacionario es un proceso estocástico cuya distribución de probabilidad en un instante de tiempo es la misma que para todos los instantes de tiempo. En consecuencia, todos los momentos de un proceso estocástico estrictamente estacionario son independientes del tiempo. Cabe mencionar que la estacionariedad fuerte es difícil de probar puesto que debe contrastarse por medio de distribuciones conjuntas para cualquier selección de variables del proceso. Por otro lado, la definición más relajada de estacionariedad, denominada estacionariedad débil, es más fácil de probar.

Se dice que un proceso estocástico es débilmente estacionario si su primer y segundo momento (media y varianza respectivamente) son constantes en el tiempo y el valor de la covarianza entre los periodos depende sólo de la distancia o rezago entre estos dos periodos, y no del tiempo en el cual se calculó la varianza ([Gujarati, 2004](#)).

Es importante mencionar que todo proceso estrictamente estacionario y con momentos de segundo orden finitos es débilmente estacionario, pero el recíproco no es cierto, debido a que la estacionariedad débil no impone restricción alguna sobre la distribución de las variables ni los momentos de orden superior a dos. Sin embargo, dada la imposibilidad

de probar la estacionariedad fuerte y al igual que en la mayoría de las aplicaciones prácticas, en la presente memoria se considera suficiente probar la estacionariedad débil de las series de tiempo utilizadas.

Una serie de tiempo estocástica Y_t es un proceso estacionario débil si:

$$\begin{aligned}\mathbb{E}(Y_t) &= \mu \\ \text{var}(Y_t) &= \mathbb{E}(Y_t - \mu)^2 = \sigma^2 \\ \gamma_k &= \mathbb{E}[(Y_t - \mu)(Y_{t+k} - \mu)]\end{aligned}\tag{6.4}$$

Donde γ_k , la covarianza en el rezago k , es la covarianza entre los valores de Y_t y Y_{t+k} , es decir, entre dos valores de Y separados por k periodos. De esta forma, si la serie es estacionaria, su media, varianza y su autocovarianza (en los diferentes rezagos) permanecen iguales sin importar el momento en el cual se midan.

6.1.1.1. Prueba de Dickey-Fuller Aumentada

Para evaluar la hipótesis de estacionariedad de las series de precios de cierre y retornos del índice IPSA debe testearse la presencia de raíz unitaria, para lo cual se emplea la prueba de Dickey-Fuller Aumentada (ADF). La hipótesis nula de la prueba es que la serie es una caminata aleatoria, y por lo tanto, el rechazo de esta hipótesis permite concluir que la variable de interés no contiene raíz unitaria.

El contraste ADF se basa en una regresión univariada de la primera diferencia de la variable sobre su valor rezagado, incluyendo diferencias rezagadas de la serie de tiempo para controlar por correlación serial de orden mayor a uno y otros términos opcionales como una constante y tendencia temporal. Específicamente asumiendo que la serie a probar, Y_t , sigue un proceso AR(p), el resultado de la prueba se obtiene al estimar la [Ecuación 6.5](#).

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \delta_1 \Delta Y_{t-1} + \dots + \delta_{p-1} \Delta Y_{t-p+1} + \varepsilon_t\tag{6.5}$$

La hipótesis nula y la hipótesis alternativa de la prueba ADF pueden escribirse como:

$$\begin{aligned}H_0: \gamma &= 0 \\H_1: \gamma &< 0\end{aligned}\tag{6.6}$$

Luego, el rechazo de esta hipótesis nula indica que la serie y_t no es una caminata aleatoria.

6.1.2. Ergodicidad

Un proceso estocástico se dice ergódico si su media y varianza pueden deducirse a partir de una muestra lo suficientemente larga de una realización del mismo.

La ergodicidad permite que al promediar una serie de tiempo, se esté continuamente añadiendo información nueva y relevante al promedio. Si una serie de tiempo cumple con los supuestos de estacionariedad y ergodicidad, entonces es posible formar buenos estimadores de las cantidades de interés promediando a través del tiempo.

Esta propiedad es muy difícil de comprobar en la práctica, ya que además de requerir un proceso estacionario débil, hay que asumir que una realización del proceso es un representante adecuado del proceso global.

6.2. Pronóstico mediante modelos de la familia GARCH

La familia de modelos estadísticos GARCH es muy popular y ampliamente utilizada en la predicción de volatilidades financieras, y esencialmente se conforma por modelos paramétricos que asumen una estructura de correlación lineal en los datos, que comúnmente se limitan a distribuciones estacionarias y normales en sus variables y errores (Lahmiri y Boukadoum, 2015a). El gran avance de estos modelos de volatilidad reside en la posibilidad de estimar las ponderaciones mediante datos históricos a pesar de que la volatilidad verdadera no se hubiese observado nunca (Engle, 2004). Una de las razones por las que estos modelos han sido ampliamente utilizados en la literatura financiera es porque capturan la agrupación de la volatilidad.

En los modelos GARCH, simultáneamente se presentan tres distintas especificaciones: la media condicional, la varianza condicional, y finalmente, la distribución condicional del error. Comúnmente se utilizan modelos ARMA

para describir la ecuación de la media condicional, mientras que la distribución condicional del error se formula bajo el supuesto de normalidad, aunque también se utiliza la distribución t de Student y la distribución generalizada del error (GED).

En la presente memoria se desarrollan tres modelos EGARCH para predecir la volatilidad de los retornos del IPSA, cada uno con un modelo ARMA como ecuación de la media condicional y con distintas distribuciones condicionales del error: Normal, t-Student y GED, que posteriormente se ingresan como entrada a los modelos híbridos que componen el sistema de ensamble. A continuación se expone cada una de las especificaciones utilizadas.

6.2.1. Modelo de Media Condicional

Para modelar la media condicional de la serie de retornos del IPSA se selecciona la mejor opción entre Modelos Autorregresivos puros, Modelos de Media Móvil y Modelos Autorregresivos de Media Móvil, los cuales fueron propuestos por [Box y Jenkins \(1978\)](#) para la predicción de series de tiempo.

6.2.1.1. Modelo Autorregresivo

En un proceso autorregresivo de orden n , denominado $AR(n)$, la observación actual Y_t es generada por un promedio ponderado de observaciones pasadas que se remontan a n periodos, junto con una perturbación aleatoria en el periodo actual. De esta forma el modelo solo considera los valores rezagados de la variable a pronosticar. La [Ecuación 6.7](#) describe este proceso.

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_n Y_{t-n} + u_t \quad (6.7)$$

Donde a_0 es un término constante que se relaciona con la media del proceso estocástico y u_t es error de ruido blanco.

6.2.1.2. Modelo de Media Móvil

En el proceso de media móvil de orden m , denominado $MA(m)$, cada observación Y_t es generada por un promedio ponderado de perturbaciones aleatorias que se remontan a m periodos anteriores. Este modelo de media móvil supone que las perturbaciones aleatorias están distribuidas en forma independiente a lo largo del tiempo, es decir, son generadas por un proceso de ruido blanco. La ecuación que describe este proceso es:

$$Y_t = \mu + b_0 u_t + b_1 u_{t-1} + b_2 u_{t-2} + \dots + b_q u_{t-m} \quad (6.8)$$

Donde $\mu = \mathbb{E}(Y)$.

6.2.1.3. Modelo Autorregresivo y de Media Móvil

Algunos procesos aleatorios estacionarios no pueden modelarse como procesos AR o MA puros, ya que tienen cualidades de ambos tipos de procesos. En dichos casos es posible utilizar un modelo mixto de n términos autorregresivos y m términos de media móvil, denominado proceso $ARMA(n, m)$, el cual se representa por medio de la [Ecuación 6.9](#).

$$Y_t = a_1 Y_{t-1} + \dots + a_p Y_{t-n} + c + u_t + b_1 u_{t-1} + b_q u_{t-m} \quad (6.9)$$

Además, las series de tiempo financieras por lo general no son estacionarias en su nivel, por lo cual es necesario diferenciar la serie una determinada cantidad de veces hasta hacerla estacionaria. Si se incorporan las d diferenciaciones de la serie al modelo $ARMA(n, m)$ se dice que la serie original es del tipo $ARIMA(n, d, m)$, es decir, es una serie de tiempo autorregresiva de media móvil.

6.2.2. Familia de Modelos GARCH

Los modelos de la familia GARCH formulan la varianza condicional de los rendimientos u otras variantes de ella mediante el procedimiento de máxima verosimilitud. Además, debido a la forma en que se construyen los modelos, la varianza condicional del periodo t se conoce en el periodo $t - 1$, por lo que el pronóstico de un paso adelante está

fácilmente disponible, mientras que los pronósticos que están más de un paso por delante se pueden formular por medio de un procedimiento iterativo (Poon y Granger, 2003).

Un punto interesante a tener en cuenta de estos modelos es que guardan muchas similitudes con los modelos ARMA, por lo que muchos de los métodos propuestos para su especificación no son más que aplicaciones de la metodología Box-Jenkins (Box y Jenkins, 1978).

Para comprender como surge el modelo EGARCH y las ventajas que conlleva su formulación, primero se presentan los modelos ARCH y GARCH, a partir de los cuales surge la variante del GARCH Exponencial (EGARCH).

6.2.2.1. Modelos ARCH

Los modelos ARCH (Engle, 1982) sugieren una forma de heterocedasticidad en la cual la varianza condicional del error de predicción depende del tamaño de las perturbaciones precedentes. De esta forma, un proceso de Heterocedasticidad Condicional Autorregresiva de orden p , denominado ARCH(p), está caracterizado por la Ecuación 6.10.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 \quad (6.10)$$

Donde es necesario imponer que $\alpha_0 > 0$ y $\alpha_i \geq 0 \forall i \geq 1$ para garantizar que la varianza sea positiva, y $\sum_{i=1}^p \alpha_i \leq 1$ para que sea estacionaria. Para construir la ecuación de la varianza condicional, es necesario disponer de una serie de residuos o términos de perturbación, los cuales están definidos por una ecuación de la media condicional como los modelos ARMA presentados con anterioridad. En muchas de las aplicaciones del modelo lineal ARCH(p) es necesario introducir un elevado número de retardos (p), para que la modelización resulte adecuada, lo cual puede plantear problemas, debido a la posible violación de la mencionada restricción de no negatividad.

En palabras de Engle (2004), “el gran avance de este modelo reside en la posibilidad de estimar las ponderaciones mediante datos históricos a pesar de que la volatilidad verdadera no se hubiese observado nunca”.

6.2.2.2. Modelos GARCH

Una estructura de retardos alternativa y más flexible, que exige un número de variables retardadas mucho menor, es a menudo proporcionada por el modelo ARCH generalizado o GARCH(p, q) (Bollerslev, 1986), cuya expresión está dada por la Ecuación 6.11.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (6.11)$$

Donde p corresponde al orden del promedio móvil ARCH y q es el orden del autorregresivo GARCH. El modelo GARCH representa que la varianza condicional en el momento t depende de tres factores: una constante, la información disponible relacionada con la volatilidad pasada (ε_{t-i}^2), y las predicciones de varianza para los periodos anteriores a t (σ_{t-j}^2). Los parámetros del modelo están sujetos a las restricciones $p \geq 0$, $q > 0$, $\alpha_0 > 0$, $\alpha_i \geq 0$ y $\beta_i \geq 0$ para asegurar que la varianza condicional de GARCH(p, q) sea siempre positiva. Además, para probar estacionariedad en sentido débil y garantizar una esperanza finita del proceso de varianza debe cumplirse que:

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1 \quad (6.12)$$

De esta forma, la suma de los términos ARCH y GARCH cuantifica el grado de persistencia de volatilidad en el largo plazo, condición suficiente para asegurar un proceso estacionario y ergódico.

Este modelo supera los problemas de estimación del ARCH(q) en cuanto a parámetros requeridos, ya que con pocos coeficientes es capaz de captar el comportamiento dinámico de la volatilidad (Robles, 2002). Incluso, las aplicaciones empíricas han puesto de manifiesto que es suficiente con la estimación de un modelo GARCH(1, 1) para una gran cantidad de variables financieras. Sin embargo, este modelo trata los choques positivos y negativos simétricamente a través de los errores cuadráticos rezagados, razón por la cual existen otras importantes generalizaciones del modelo que incorporan el efecto de apalancamiento financiero, donde destacan modelos tales como el GARCH exponencial, GRJ-GARCH, APARCH, entre otros.

6.2.2.3. Modelos EGARCH

El modelo EGARCH, desarrollado por Nelson (1991) propone un enfoque que a diferencia del modelo GARCH, no impone restricciones de no negatividad en los parámetros y permite estimar convenientemente el efecto de apalancamiento que existe en el mercado financiero.

EGARCH es un modelo de volatilidad asimétrica ampliamente utilizado en la literatura financiera, en el cual la variable dependiente corresponde al logaritmo de la varianza condicional, donde la volatilidad no responde de la misma forma ante perturbaciones positivas que ante perturbaciones negativas. La especificación del modelo EGARCH(p, q), se presenta en la Ecuación 6.13.

$$\log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^p \alpha_i \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| + \sum_{j=1}^q \beta_j \log(\sigma_{t-j}^2) + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \quad (6.13)$$

Donde γ captura el comportamiento asimétrico de los choques negativos y positivos en la volatilidad condicional, cuyo efecto de apalancamiento es exponencial más que cuadrático.

En la presente memoria la estimación y pronóstico de los modelos ARMA y EGARCH utilizados para predecir la media y volatilidad condicional de los retornos del IPSA es realizada por medio de una ventana móvil de 252 días, que corresponde aproximadamente a un año de transacciones. De esta forma, la primera muestra empleada para obtener el primer pronóstico un paso adelante utiliza los primeros 252 datos de la serie de retornos.

La elección del modelo EGARCH en lugar del clásico GARCH se debe a su capacidad de modelar variaciones asimétricas de los datos, y su eficacia en la realización de pronósticos del mercado de valores. Igualmente, se utiliza un modelo GARCH(1,1) para realizar comparaciones con los pronósticos de los modelos que conforman el sistema de ensamble, dada su amplia utilización en aplicaciones financieras.

6.2.3. Distribuciones del Error

Considerando que la estimación de los parámetros de la familia GARCH puede realizarse a partir de procedimientos de máxima verosimilitud, estos modelos requieren la especificación de la función de densidad, lo cual tiene

efectos en el proceso de estimación.

En la estimación empírica de las volatilidades basadas en EGARCH, se considera como supuesto que el término de error sigue tres distribuciones condicionales diferentes, incluyendo la distribución normal, t-Student y el error generalizado (GED) de [Box y Tiao \(1975\)](#). La función de densidad de probabilidad (PDF) de la distribución normal para la variable aleatoria z es bien conocida, y está dado por:

$$F(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) \quad (6.14)$$

Donde μ y σ son la media y la varianza, respectivamente. La función de densidad de probabilidad de la distribución t-Student se expresa por la [Ecuación 6.15](#).

$$F(z, \mu, \sigma, \nu) = \frac{C}{\sigma^2} \left(1 + \frac{(z-\mu)^2}{\sigma^2(\nu-2)}\right)^{-\frac{(\nu+1)}{2}} \quad (6.15)$$

Donde μ es la media, σ es la varianza (a menudo desconocida cuando se usa el PDF de Student), ν es el grado de libertad, y el parámetro C está dado por la [Ecuación 6.16](#).

$$C = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}(\nu-2)\Gamma\left(\frac{\nu}{2}\right)} \quad (6.16)$$

Donde $\Gamma(\bullet)$ es la función gamma. El parámetro ν se utiliza para controlar el espesor y la longitud significativa de las colas de la función de densidad, afectado así la asimetría de la función de densidad de probabilidad t-Student. Cuando ν es mayor que dos, el PDF tiene cero asimetría, independientemente de los datos. Finalmente, la función de densidad del GED se expresa de la siguiente manera:

$$F(z_t, \mu, \sigma, \beta) = \frac{\beta}{2\sigma\Gamma(1/\beta)} \exp\left(-\left(\frac{|z_t - \mu|}{\sigma}\right)^\beta\right) \quad (6.17)$$

Donde μ y σ son la media y la varianza de los datos, y β es un parámetro que controla la curtosis, que es prácticamente constante en $+\infty$, 0 y -2 para la distribución leptocúrtica, mesocúrtica y platicúrtica, respectivamente.

Por lo tanto, GED es menos sensible a los efectos de curtosis dentro de estos tipos de distribución.

6.3. Redes Neuronales Artificiales

Si bien no existe una definición clara, puede decirse que las redes neuronales artificiales son una aproximación poco abstracta de la rama del conocimiento de la Inteligencia Artificial, y representan modelos simplificados que intentan imitar la estructura y el funcionamiento del pensamiento humano respecto a sus procesos de aprendizaje y reconocimiento, teniendo como unidad básica las neuronales artificiales. En palabras de [Aleksander y Morton \(1990\)](#), “una red neuronal es un procesador masivo distribuido paralelamente, que tiene una propensión natural por el almacenamiento de conocimiento experimental haciéndolo disponible para su uso. Esto se asemeja al funcionamiento del cerebro en dos aspectos:

1. La red adquiere el conocimiento por medio de un proceso de aprendizaje.
2. La fuerza de las interconexiones neuronales se determina por pesos sinápticos que se utilizan para almacenar conocimiento.”

De esta definición es posible identificar que las redes neuronales artificiales intentan simular la técnica de procesamiento en paralelo de la información y la adaptabilidad al entorno que permite modificar los pesos sinápticos en base al aprendizaje propio de los sistemas nerviosos biológicos, a los cuales [Martin del Brio y Sanz \(1997\)](#) agregan la característica de la distribución amplia de la información por toda la red neuronal.

6.3.1. La Neurona Biológica

A fines del siglo XIX, el ganador del premio Nobel [Ramón y Cajal \(1888\)](#) descubrió por medio del estudio de fibras nerviosas, que las neuronas son las unidades estructurales y funcionales del sistema nervioso. [Ramón y Cajal](#) defendió la teoría de que las neuronas se interconectaban entre sí de forma paralela actuando como entidades discretas, no formando un circuito cerrado como el sistema sanguíneo. Afirmó, en base a sus observaciones, que entre las prolongaciones de las neuronas no se establecía un contacto directo, quedando un espacio libre a través del cual el impulso nervioso "saltaba" de una célula a la siguiente.

El cerebro es un sistema altamente complejo. Se calcula que hay aproximadamente 10 billones de neuronas en la corteza cerebral y cada una de ellas puede recibir unas 10.000 entradas y enviar a su vez su salida a varios cientos de neuronas. La conexión entre neuronas se lleva a cabo mediante un proceso denominado sinapsis, en el cual no hay una conexión física directa, sino que existe una distancia de $2 \mu\text{m}$ de separación. Además, se trata de conexiones unidireccionales, en las que la transmisión de la información se hace de forma eléctrica en el interior de la neurona y de forma química entre neuronas, gracias a sustancias específicas llamadas neurotransmisores.

En la [Figura 6.1](#) se presenta un esquema simplificado de una neurona biológica, en el cual se distinguen tres de sus principales componentes: el cuerpo o soma, las dendritas y el axón. Las dendritas reciben las señales de las neuronas adyacentes y las transmiten al cuerpo en forma de potencial eléctrico. Estas señales eléctricas son integradas por el cuerpo celular (soma). Si el potencial eléctrico es superior a un valor umbral, el soma genera un corto impulso eléctrico. Este impulso se transmite por el axón, que es una fibra nerviosa con una longitud que varía entre unos milímetros y varios metros. El axón se ramifica y dirige el impulso a varias neuronas vía sinapsis ([Díaz et al., 1996](#)).

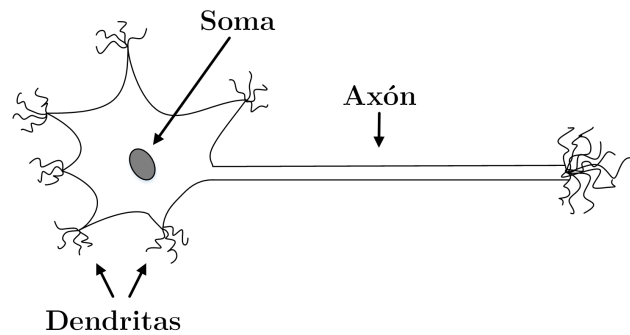


Figura 6.1: Representación simple de una neurona biológica
(Fuente: Elaboración Propia)

Cabe mencionar que existen muchos tipos diferentes de neuronas según el número de ramificaciones de sus dendritas, la longitud del axón y otros detalles de su estructura. Sin embargo, todas ellas operan bajo los mismos principios básicos expuestos.

6.3.2. Unidad de Procesamiento Elemental

El primer modelo matemático de una neurona artificial fue propuesto en el año 1943 por el neurólogo y cibernético Warren McCulloch y el matemático Walter Pitts, quienes modelaron una red neuronal simple mediante

circuitos eléctricos con el fin de llevar a cabo tareas básicas. Cabe destacar que la unidad elemental de este modelo, una neurona sencilla, se utiliza hasta la actualidad y sirve como referencia en el campo de redes neuronales.

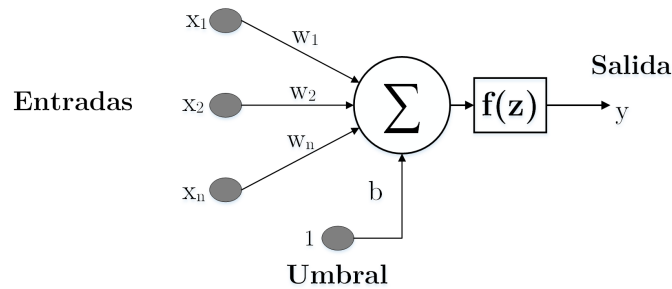


Figura 6.2: Representación de la neurona artificial de McCulloch y Pitts

(Fuente: Elaboración Propia)

Como puede observarse en la [Figura 6.2](#), el modelo neuronal de McCulloch y Pitts es una representación abstracta y sencilla, donde una neurona consta de un conjunto de entradas $x = (x_1, x_2, \dots, x_n)^T$ y una única respuesta o salida y . Cada entrada x_i tiene asociado un ponderador o peso sináptico w_i perteneciente al vector de pesos $W = (w_1, w_2, \dots, w_n)^T$, el cual equivale a las conexiones sinápticas de una neurona biológica. La respuesta de la neurona se obtiene a partir de la aplicación de la función de transferencia $f(z)$ al producto punto entre el vector de entradas y el vector de pesos, a lo que se suma el umbral de activación representado por b . El umbral se asemeja a los pesos sinápticos, con la diferencia de que su entrada corresponde a una constante 1. Lo anterior se expresa en la [Ecuación 6.18](#).

$$y = f(W \cdot x + w_0) = f(z) \quad (6.18)$$

Es importante notar que tanto los pesos sinápticos como el umbral son parámetros escalares ajustables de la neurona artificial. Así, la idea central de las redes neuronales artificiales es que dichos parámetros pueden ajustarse por medio de un proceso de entrenamiento, de forma tal que la red exhiba algún comportamiento deseado.

La función de transferencia propuesta por McCulloch y Pitts posee una salida binaria ± 1 conocida como función signo o función de todo o nada, que está dada por la [Ecuación 6.19](#).

$$f(z) = \text{sgn}(z) = \begin{cases} 1 & \text{si } z \geq 0 \\ -1 & \text{si } z < 0 \end{cases} \quad (6.19)$$

Una sola neurona artificial no tiene la capacidad de realizar un proceso lógico, sin embargo, un conjunto de ellas agrupadas en redes, como se presenta a continuación a través del perceptrón, es capaz de realizar trabajos de cálculo lógico.

6.3.3. Perceptrón Simple

En 1958, el psicólogo [Rosenblatt](#) desarrolló la primera red neuronal artificial, la cual consta de un modelo simple de neurona basado en la neurona artificial de McCulloch y Pitts y en una regla de aprendizaje basada en la corrección del error, denominado Perceptrón, cuya característica más importante es su capacidad de aprender a reconocer patrones.

La estructura del perceptrón se inspira en las primeras etapas de procesamiento de los sistemas sensoriales de los animales, en los cuales la información va atravesando sucesivas capas de neuronas, las cuales realizan un procesamiento que va aumentando su nivel progresivamente.

Como se presenta en la [Figura 6.3](#), el perceptrón simple es un modelo neuronal unidireccional, constituido por una capa de entrada de neuronas que reciben los patrones de entrada a reconocer o clasificar y una capa de neuronas de salida que se ocupa de clasificar a los patrones de entrada en dos clases, según que la salida de la misma sea 1 (activada) o 0 (desactivada).

Siguiendo con la notación presentada en el modelo de McCulloch y Pitts, en un perceptrón simple cada uno de los elementos del vector de entradas x están conectados con cada una de las neuronas de la capa de salida a través de la matriz de pesos W , y cada una de las neuronas de la capa de salida genera su respuesta a partir de la aplicación de la función de transferencia al producto punto entre entradas y pesos, más el umbral b . Finalmente, los resultados o salidas de la capa neuronal en conjunto forman el vector de salidas y .

Las neuronas de la capa de entrada no realizan ningún cómputo, sino que únicamente se encargan de enviar la información a las neuronas de salida, de forma que la red tiene la capacidad de discriminar conjuntos de datos a través de las funciones lógicas AND, OR y NOT. A pesar de la importancia del desarrollo de este tipo de red neuronal artificial, dicho modelo presenta muchas limitaciones, como por ejemplo, no es capaz de resolver problemas no lineales, como por ejemplo la función lógica XOR ([Minsky y Papert, 1969](#)).

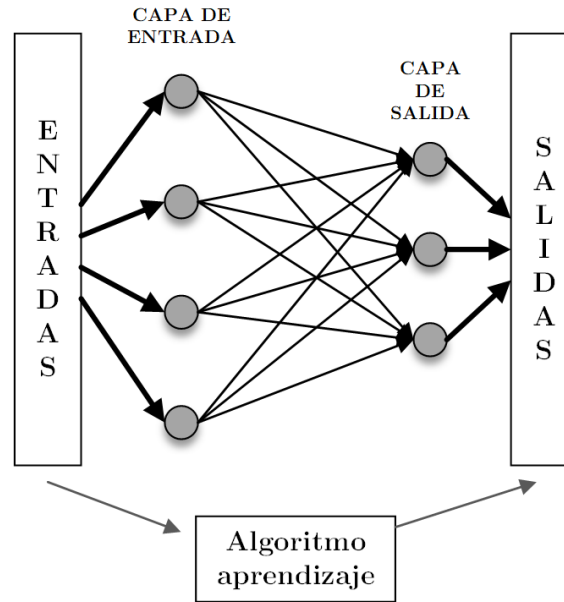


Figura 6.3: Representación de un Perceptrón Simple
(Fuente: Elaboración Propia)

La combinación de varios perceptrones simples es capaz de resolver ciertos problemas no lineales, sin embargo a principios de los años sesenta no existía un mecanismo automático para adaptar los pesos de las capas ocultas. Tuvieron que pasar unos años hasta que se propusiera la regla de aprendizaje de retropropagación del error que permitió trabajar con múltiples capas y funciones de activación no lineales, demostrándose que el perceptrón multicapa es un aproximador universal.

6.3.4. Perceptrón Multicapa

El perceptrón multicapa o MLP (Multi-Layer Perceptron) es una arquitectura de red neuronal muy popular, la cual corresponde a una generalización del perceptrón simple que surgió como consecuencia de las limitaciones de dicha arquitectura en lo referente al problema de la separabilidad no lineal.

El perceptrón multicapa, descrito formalmente por [Werbos \(1974\)](#), es una red con alimentación hacia delante (feedforward) compuesta por neuronas agrupadas en varias capas de diferentes niveles, usualmente tres (entrada, oculta y salida) donde cada neurona posee su respectivo nivel de umbral. Adicionalmente, se utiliza funciones sigmoideas como función de transferencia en las neuronas de la capa oculta, mientras que las funciones de la capa de salida pueden ser lineales o sigmoideas, dependiendo del tipo de salida que se requiera. La [Figura 6.4](#) muestra la arquitectura de una

red feedforward, como el perceptrón multicapa, por medio del empleo de un grafo.

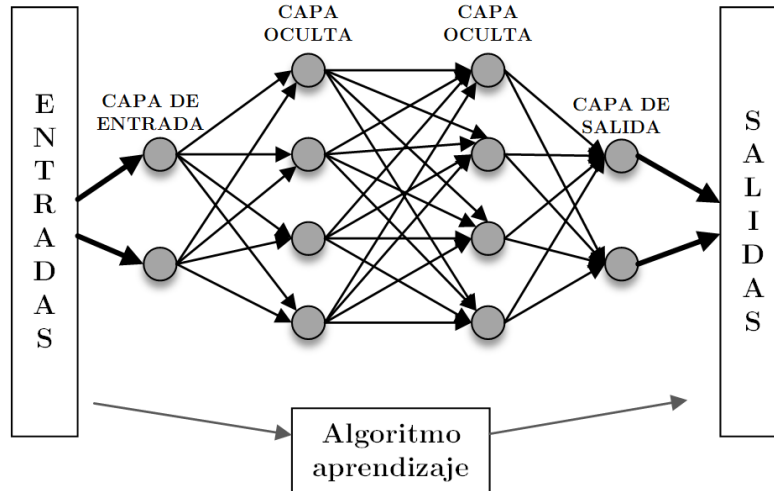


Figura 6.4: Representación de una red feedforward

(Fuente: Elaboración Propia en base a Jiménez (2000))

En el perceptrón multicapa, cada una de las capas de procesamiento tiene asociados una matriz de pesos W , un vector de umbrales b y un vector de salidas y .

No existe un método o regla que determine el número óptimo de capas ocultas y neuronas en cada una de ellas para resolver un problema dado, por lo que generalmente se determina mediante ensayo y error partiendo desde una arquitectura ya entrenada para luego realizar cambios aumentando y disminuyendo la cantidad de neuronas ocultas y capas hasta conseguir una arquitectura que se ajuste de mejor forma a la solución dada.

Los tres modelos híbridos construidos en la presente memoria se establecen introduciendo el resultado de los modelos de tipo EGARCH seleccionados en redes neuronales artificiales, conformando los modelos denominados ANN-EGARCH. En otras palabras, el resultado de los modelos de tipo EGARCH se considera como una variable de entrada en redes neuronales para aumentar el rendimiento de la predicción de la volatilidad en de los retornos del IPSA. Todas las redes neuronales utilizadas corresponden a perceptrones multicapa donde la función de transferencia en cada neurona de la capa oculta es la función sigmoideal (Ecuación 6.20) y en la capa de salida es la función lineal (Ecuación 6.21).

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6.20)$$

$$f(x) = x \quad (6.21)$$

En el caso base analizado en la presente memoria las redes híbridas ANN-EGARCH con los distintos supuestos distributivos del error constan de una neurona en la capa de entrada y dos capas ocultas con diez neuronas, mientras que la red que combina los pronósticos de las redes híbridas para entregar el pronóstico final del sistema de ensamble consta de tres neuronas en la capa de entrada y dos capas ocultas con diez neuronas. En todas las redes neuronales, la capa de salida tiene una neurona que entrega el valor pronosticado de la volatilidad.

La arquitectura de las redes híbridas y la red final del sistema de ensamble se presentan en la [Figura 6.5](#) y la [Figura 6.6](#), respectivamente.

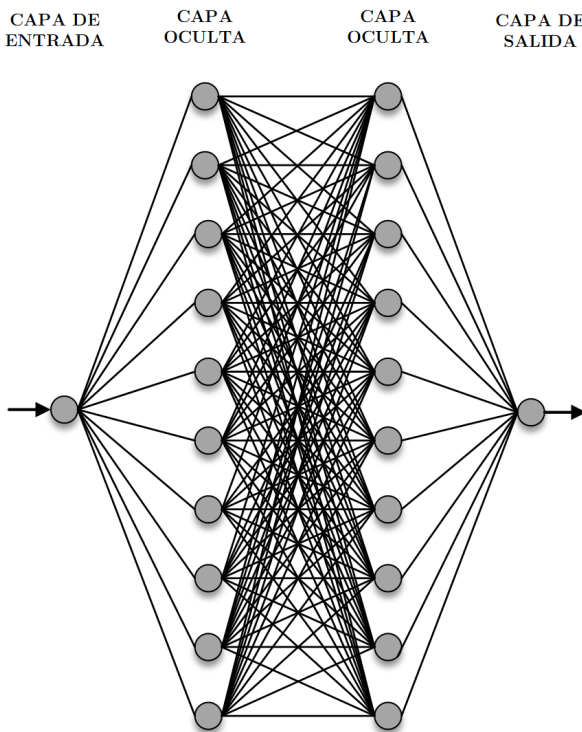


Figura 6.5: Representación gráfica de la red ANN-EGARCH utilizada

(Fuente: Elaboración Propia)

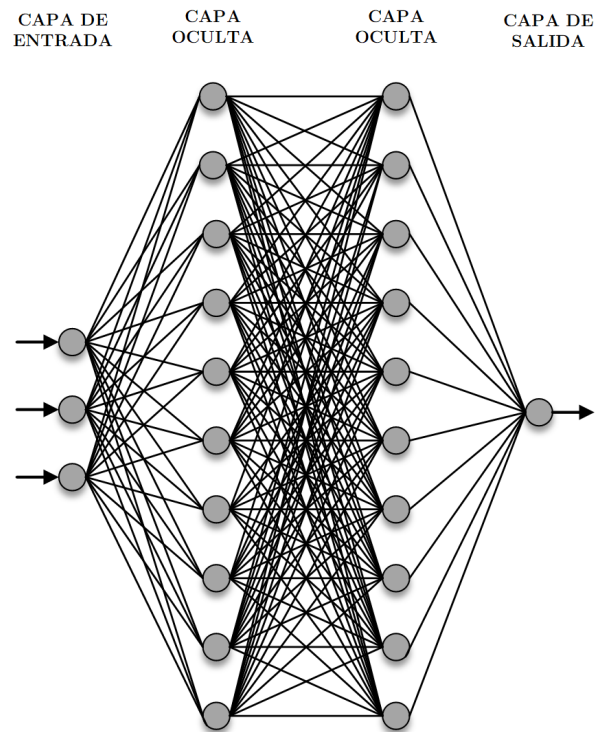


Figura 6.6: Representación gráfica de la red final del sistema de ensamble utilizada

(Fuente: Elaboración Propia)

6.3.5. Entrenamiento de la Red Neuronal

Para que una red neuronal descubra las conexiones no lineales entre dos conjuntos de datos, adaptando y modificando todos los parámetros de la red, es necesario entrenarla, lo cual se logra por medio de la utilización reiterada de un algoritmo, denominado de aprendizaje.

En el proceso de entrenamiento de redes multicapa, una práctica general consiste en dividir inicialmente los datos en tres submuestras. La primera corresponde a la submuestra de entrenamiento, la cual se utiliza para el cálculo y actualización de los pesos y umbrales de la red. La segunda submuestra es de validación, la cual se utiliza para verificar que el ajuste de la red sea apropiado, monitoreando el error de la red. Finalmente, la submuestra de prueba se utiliza para conocer el desempeño de la red fuera de la muestra.

De esta forma, el proceso de entrenamiento consiste en presentar un conjunto correspondiente al 70 % de los datos seleccionados de forma aleatoria y modificar iterativamente los pesos hasta obtener un error mínimo deseado en los resultados. Otro conjunto también aleatorio, con el 15 % de los datos es utilizado para la validación permitiendo interrumpir el entrenamiento cuando el error de validación empiece a incrementarse y evitar así el sobre entrenamiento. Finalmente, se testea la estructura de red obtenida con el 15 % de los datos restantes no incluidos en el proceso de entrenamiento.

Una de las características más importantes del perceptrón multicapa es que utiliza como función de aprendizaje la retropropagación de errores hacia atrás o backpropagation de [Werbos \(1974\)](#), atribuible también a [Parker \(1982\)](#) y [LeCun \(1985\)](#). Este constituye un algoritmo de aprendizaje supervisado, el cual parte de una muestra de ejemplo y calcula el error de salida. A partir de dicho error actualiza los pesos de salida y, posteriormente, propaga hacia atrás los errores a través de las sinapsis. De este modo, se obtienen las señales de error que corresponden a las capas ocultas y se pueden actualizar sus sinapsis.

Existen diferentes variantes de la regla de retropropagación hacia atrás que se utilizan en función del problema que se quiera resolver. En el caso de la presente memoria se opta por utilizar el algoritmo de Levenberg Marquardt ([Marquardt, 1963](#)), el cual tiene una considerable eficiencia relativa y que ha encontrado una amplia aplicación, ya que en general ha demostrado tener una convergencia más rápida, es decir, requiere menor cantidad de iteraciones para llegar

al nivel de error especificado (Qi, 1999; Villada et al., 2012).

En una red neuronal entrenada con el algoritmo Levenberg-Marquardt los pesos y umbrales se ajustan de la siguiente forma:

$$\Delta w_k = -(J_k^T J_k + \mu I)^{-1} J_k e_k \quad (6.22)$$

Donde J es la matriz jacobiana (primeras derivadas) de los pesos, I es la matriz identidad, μ es un parámetro adaptativo de aprendizaje ajustado en 0,001, y e es un vector de errores de la red. Los valores iniciales para los pesos y umbrales de la red son establecidos al azar dentro del rango $[-1,1]$.

6.4. Sistema de Ensamble

Siguiendo la metodología propuesta por Lahmiri y Boukadoum (2015a), el diseño de un sistema de ensamble consta de dos pasos. En primer lugar, se forman varios modelos híbridos, que en este caso corresponden a redes neuronales artificiales que tienen como entrada los resultados de los modelos EGARCH desarrollados previamente con distintos supuestos distributivos. A continuación, los resultados de predicción de los modelos híbridos se combinan, de forma que cada ANN-EGARCH en el sistema de ensamble es entrenada usando los ejemplos de entrenamiento, y sus salidas se combinan para proporcionar la salida final del conjunto de las redes.

La Figura 6.4 presenta un diagrama de flujo genérico del modelo híbrido ANN-EGARCH utilizado en la presente memoria para la elaboración de los tres modelos con diversos supuestos distributivos. Por otro lado, en la Figura 6.8 se presenta el sistema de ensamble empleado, donde los tres modelos ANN-EGARCH se construyen asumiendo una distribución diferente de error para cada uno: normal, t-student y GED. Para combinar los pronósticos de los tres modelos ANN-EGARCH se opta por utilizar una red adicional que aprenda los tres modelos, para así obtener el valor final de la predicción de la volatilidad del mercado bursátil chileno.

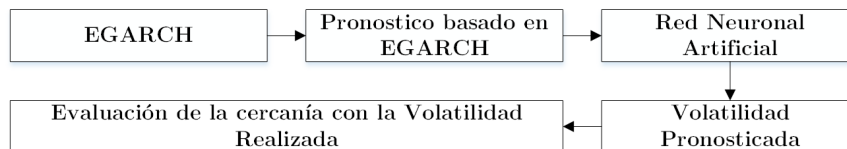


Figura 6.7: Diseño genérico de un modelo híbrido ANN-EGARCH para la predicción de volatilidad

(Fuente: Elaboración Propia en base a Lahmiri y Boukadoum (2015a))

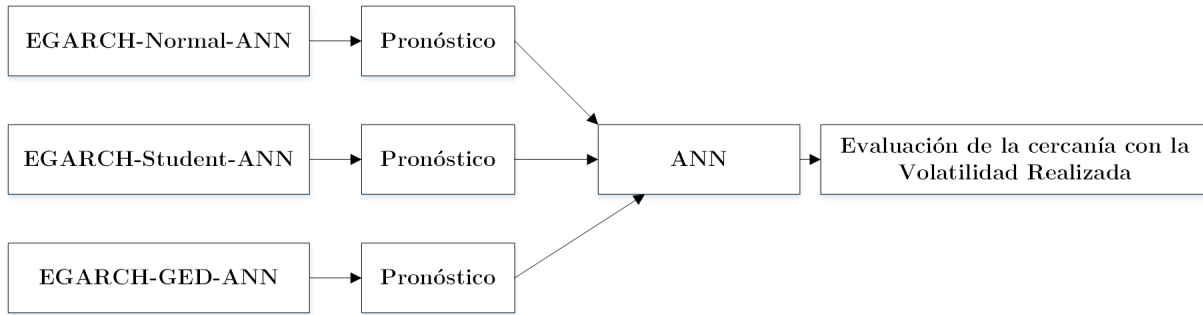


Figura 6.8: Sistema de ensamble basado en modelos híbridos ANN-EGARCH

(Fuente: Elaboración Propia en base a Lahmiri y Boukadoum (2015a))

6.5. Volatilidad Realizada

La volatilidad realizada es un indicador ex-post que mide la dispersión de los retornos de un activo en un periodo de tiempo, y se utiliza en esta memoria tanto para la selección de los parámetros como evaluar el desempeño de los pronósticos de cada modelo desarrollado. La volatilidad realizada, como se muestra en la Ecuación 6.23, se calcula como la varianza muestral de los retornos logarítmicos respecto a una media de D días al futuro.

$$VR_t = \frac{1}{D} \sum_{i=t+1}^{t+D} (r_i - \bar{r}_i)^2 \quad (6.23)$$

Donde VR_t es la volatilidad realizada de los rendimientos del IPSA, D es el ancho de la ventana de cálculo y \bar{r}_i representa el promedio de los retornos en la ventana considerada.

La ventana de cálculo utilizada para el caso base es de 21 días (Kristjanpoller et al., 2014), lo que corresponde aproximadamente a un mes de transacciones.

6.6. Medidas de Desempeño

El rendimiento predictivo de los modelos EGARCH, ANN-EGARCH y el sistema de ensamble se mide evaluando la cercanía de los pronósticos con la volatilidad realizada, mediante el cálculo del error absoluto medio (MAE) y la raíz

cuadrática media del error (*RMS E*), cuyas formulas se presentan en las siguientes ecuaciones:

$$MAE = \frac{1}{m} \sum_{i=1}^m |\hat{\sigma}_i^2 - VR_i| \quad (6.24)$$

$$RMS E = \sqrt{\frac{1}{m} \sum_{i=1}^m (\hat{\sigma}_i^2 - VR_i)^2} \quad (6.25)$$

Donde VR_i corresponde a la volatilidad realizada del i -ésimo periodo, $\hat{\sigma}_i^2$ es el i -ésimo valor predicho de la volatilidad, y m el total de observaciones en los datos de prueba. De esta forma, el modelo con mejor rendimiento predictivo corresponde al que presente los valores más bajos de estos indicadores estadísticos.

7. TRATAMIENTO DE DATOS

Los datos utilizados se extraen de la base de datos históricos de la Bolsa de Comercio de Santiago, y corresponde a los precios de cierre diarios del IPSA entre el 2 de enero del 2009 y el 30 de diciembre del 2016. Cabe mencionar que base de datos es ampliamente accesible para los individuos en general, ya es gratuita y se encuentra disponible en internet.

A continuación se presenta un análisis del comportamiento de la serie de precios, rentabilidades del IPSA y la volatilidad realizada por medio de sus principales estadísticos descriptivos.

7.1. Precios de Cierre

A partir de los estadísticos descriptivos presentados en la [Tabla 7.1](#) es posible observar que la serie de precios de cierre presenta una asimetría pequeña y negativa, mientras que la curtosis señala la presencia de una distribución más apuntada y con colas menos anchas que la normal. El valor más alto alcanzado por la serie durante el periodo analizado corresponde a 5.040,974 y fue alcanzado el 05/01/2011, mientras que el valor más bajo fue 2.359,997 alcanzado el 06/03/2009.

La prueba de Dickey Fuller Aumentada (ADF) es rechazada, lo cual permite concluir que la serie de precios de cierre del IPSA no es estacionaria, y por lo tanto no puede ser utilizada para modelar y pronosticar.

En general, en el gráfico de la serie presentado en la [Figura 7.1](#) se observan tres periodos de comportamiento. El primero caracterizado por una tendencia al alza del precio del IPSA, el cual abarca desde el 02/01/2009 hasta el 06/03/2009. El segundo periodo va desde el 06/01/2011 hasta el 20/01/2016 y se caracteriza por la baja en el valor de la serie. Finalmente, el tercer periodo presenta un alza sostenida en el precio del IPSA desde el 21/01/2016 hasta el último

Tabla 7.1: Estadísticos descriptivos de la serie de Precios de Cierre del IPSA
(Fuente: Elaboración Propia)

| Estadístico Descriptivo | IPSA |
|-------------------------|-----------|
| Media | 3.975,033 |
| Mediana | 3.960,997 |
| Máximo | 5.040,974 |
| Mínimo | 2.359,997 |
| Desviación Estándar | 508,1578 |
| Asimetría | -0,672972 |
| Curtosis | 4,104339 |
| Jarque Bera | 252,3418 |
| Observaciones | 1.998,0 |
| ADF(con intercepto) | -2,957992 |

Nota: La Prueba de Normalidad utilizada es la prueba de Jarque-Bera, que tiene una $\chi^2(q)$ con 2 grados de libertad bajo la hipótesis nula del error normalmente distribuido. El de valor crítico al 5 % es, por tanto, 5,99. La prueba de Dickey Fuller Aumentada se realiza considerando una constante y cuatro diferencias rezagadas de la variable, cuyo valor crítico al 1 % es de -3,43.

Precios de Cierre IPSA

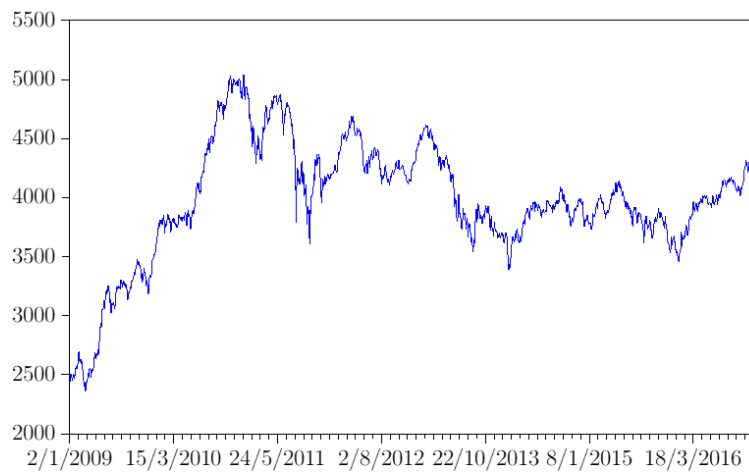


Figura 7.1: Comportamiento del IPSA en el periodo 2009-2016

(Fuente: Elaboración Propia en base a la serie de precios históricos de la [Bolsa de Comercio de Santiago](#))

dato de la muestra (30/12/2016).

7.2. Rentabilidad

En la [Figura 7.2](#) se aprecia el comportamiento de los retornos del IPSA, calculados como la primera diferencia logarítmica del IPSA, de forma que con la diferenciación se elimina la tendencia de la serie, mientras que la aplicación

del logaritmo natural se encarga de la variabilidad. En ella se aprecian periodos agrupados de alta volatilidad, los cuales coinciden con los periodos descendientes en los valores de los precios de cierre del índice. Esta tendencia de la volatilidad a ser superior en mercados con tendencia a la baja corresponde al efecto de volatilidad asimétrica descrito por Nelson (1991) con su modelo EGARCH.

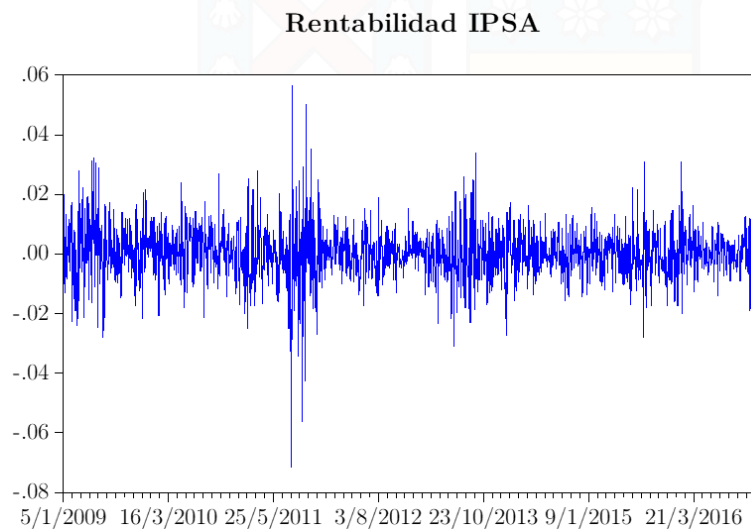


Figura 7.2: Comportamiento de la Rentabilidad del IPSA durante el periodo 2009-2016

(Fuente: Elaboración Propia)

Tabla 7.2: Estadísticos descriptivos de la serie de retornos del IPSA

(Fuente: Elaboración Propia)

| Estadístico Descriptivo | IPSA |
|-------------------------|-----------|
| Media | 0,027 % |
| Mediana | 0,035 % |
| Máximo | 5,669 % |
| Mínimo | -7,173 % |
| Desviación Estándar | 0,884 % |
| Asimetría | -0,219903 |
| Curtosis | 8,507374 |
| Jarque Bera | 2.539,901 |
| Observaciones | 1.997 |
| ADF(con intercepto) | -21,21763 |

Nota: La Prueba de Normalidad utilizada es la prueba de Jarque-Bera, que tiene una $\chi^2(q)$ con 2 grados de libertad bajo la hipótesis nula del error normalmente distribuido. El de valor crítico al 5 % es, por tanto, 5,99. La prueba de Dickey Fuller Aumentada se realiza considerando una constante y cuatro diferencias rezagadas de la variable, cuyo valor crítico al 1 % es de -3,43.

A partir de los estadísticos descriptivos presentados en la [Tabla 7.2](#) se aprecia que la rentabilidad media es positiva y cercana a cero durante el periodo considerado, situándose en el 0,0267 %, mientras que la desviación estándar

se aproxima a 1 (0,884 %), ambos valores típicos para series estacionarias, ya que poseen una media cercana a 0 con pequeñas variaciones. La curtosis, medida que analiza el grado de concentración de valores de la variable alrededor de la zona central de la distribución de frecuencias, presenta una distribución leptocúrtica, caracterizada por ser más apuntada y con colas menos anchas que la distribución normal, lo que constituye una característica típica de los rendimientos diarios.

La asimetría de la serie es pequeña y negativa, indicando que la “cola” a la izquierda de la media es levemente más larga que la de la derecha.

El test de Jarque-Bera rechaza al 5 % de significancia la hipótesis nula de que los datos pertenecen a una distribución normal.

La estacionariedad de la serie es analizada por la prueba de Dickey Fuller Aumentada (ADF), la cual resulta significativa al 1 %, rechazando la hipótesis nula de que la serie no es caminata aleatoria. Por lo tanto, puede concluirse que la serie de retornos del IPSA es estacionaria, y por lo tanto se puede utilizar para modelar y pronosticar.

7.3. Volatilidad Realizada

La [Tabla 7.3](#) presenta las principales estadísticas descriptivas para la serie de volatilidad realizada a 21 días. A partir de los datos, se determina que la serie presenta una mayor concentración de los datos en torno a la media (leptocúrtica) y asimétricamente positiva, lo cual se debe a que las varianzas necesariamente son valores mayores a cero.

Tabla 7.3: Estadísticos descriptivos de la serie de Volatilidad Realizada del IPSA a 21 días
(Fuente: Elaboración Propia)

| Estadístico Descriptivo | IPSA |
|-------------------------|-----------|
| Media | 0,756 % |
| Mediana | 0,663 % |
| Máximo | 2,579 % |
| Mínimo | 0,266 % |
| Desviación Estándar | 0,363 % |
| Asimetría | -0,672972 |
| Curtosis | 10,880490 |

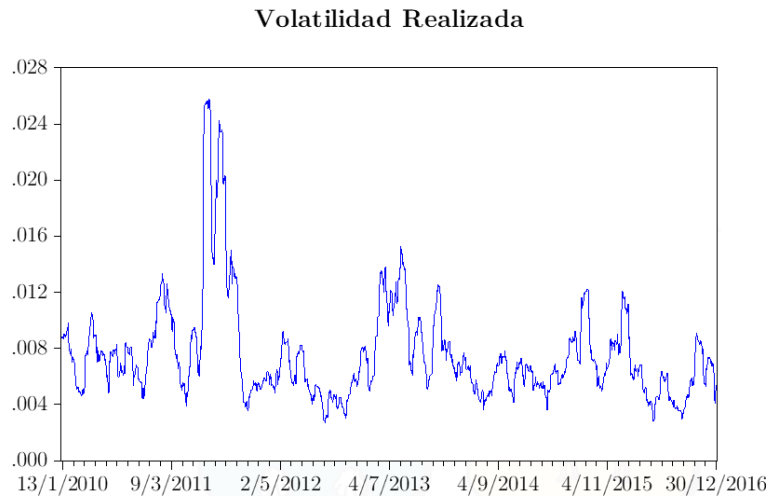


Figura 7.3: Volatilidad Realizada del IPSA durante el 2010-2016

(Fuente: Elaboración Propia)

A partir del gráfico de la [Figura 7.3](#), se observa que la volatilidad realizada presenta las mayores distorsiones a finales del año 2011, lo cual coincide con el comportamiento de las series de precios de cierre y retornos del índice bursátil.

8. RESULTADOS

8.1. Caso Base

En primer lugar, se realizan pronósticos mediante un modelo GARCH para utilizarlos como punto de referencia respecto a los pronósticos de los modelos EGARCH con los distintas distribuciones de error, los modelos híbridos EGARCH-ANN y el sistema de ensamble propiamente tal. Cabe mencionar que específicamente se utiliza un modelo GARCH(1,1) con distribución normal de los errores y una ecuación de media AR(1), dada su simplicidad y amplia utilización en aplicaciones financieras para la predicción de volatilidad. Estas especificaciones se muestran en la [Ecuación 8.1](#).

$$r_t = c + a_1 r_{t-1} \quad (8.1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

En la [Tabla 8.1](#) se presentan los resultados obtenidos de las medidas de desempeño para el pronóstico de volatilidad a 21 días del modelo GARCH.

Tabla 8.1: Resultados de desempeño del modelo de pronóstico GARCH (21 días)
(Fuente: Elaboración Propia)

| | 21 días | |
|------------|----------|----------|
| Modelo | RMSE | MAE |
| GARCH(1,1) | 0,003718 | 0,002421 |

Teniendo dichos resultados, a partir de los 875 modelos con distintas especificaciones de parámetros para las

ecuaciones de media y volatilidad condicional, y en base a la cercanía de la volatilidad pronosticada con la volatilidad realizada de los retornos del IPSA a 21 días, medido por el RMSE y MAE, se obtienen los tres modelos EGARCH cuyos pronósticos se emplean posteriormente como entradas para las redes neuronales, cuyas especificaciones se presentan en la [Tabla 8.2](#), mientras que la [Tabla 8.3](#) contiene las especificaciones de las ecuaciones de los modelos seleccionados.

Tabla 8.2: Especificaciones de parámetros de los modelos EGARCH seleccionados
(Fuente: Elaboración Propia)

| Distribución del Error | AR(n) | EGARCH(p,q) |
|------------------------|-------|-------------|
| Normal | 1 | 1,1 |
| T-Student | 4 | 1,1 |
| GED | 1 | 1,1 |

Tabla 8.3: Especificación de las ecuaciones de los modelos EGARCH empleados
(Fuente: Elaboración Propia)

| Distribución del Error | Especificación del Modelo |
|------------------------|---|
| Normal | $r_t = c + a_1 r_{t-1}$ $\log(\sigma_t^2) = \alpha_0 + \alpha_i \left \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right + \beta_1 \log(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ |
| T-Student | $r_t = c + a_1 r_{t-1} + a_2 r_{t-2} + a_3 r_{t-3} + a_4 r_{t-4}$ $\log(\sigma_t^2) = \alpha_0 + \alpha_i \left \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right + \beta_1 \log(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ |
| GED | $r_t = c + a_1 r_{t-1}$ $\log(\sigma_t^2) = \alpha_0 + \alpha_i \left \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right + \beta_1 \log(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ |

Cabe mencionar que los modelos para las distribuciones t-Student y GED presentan un parámetro adicional asociado, denominados parámetro t-Student y parámetro GED respectivamente. En cuanto al parámetro γ , este resultó negativo en más del 80 % de las estimaciones de ventana móvil para cada una de las distribuciones de error, lo cual conlleva como implicancia que en general, las innovaciones negativas (malas noticias) tienen un mayor impacto sobre la volatilidad condicional futura que las innovaciones positivas (buenas noticias) de la misma magnitud. Adicionalmente, alrededor del 98 % de los parámetros β_1 son menores que 1 en su valor absoluto, alcanzando las condiciones de proceso estacionado y ergódico de los tres modelos, condición necesaria y suficiente sustentada por [Nelson \(1991\)](#).

Al analizar los parámetros estimados tanto en la especificación de la media como de la volatilidad de los tres modelos, se tiene que en general los estimadores son estadísticamente significativos en los niveles convencionalmente analizados.

Los resultados de las medidas de desempeño del pronóstico de los modelos EGARCH a 21 días bajo los distintos supuestos distributivos utilizados se presentan en la [Tabla 8.4](#), donde se aprecia un mejor desempeño predictivo para los tres modelos de pronóstico en relación al modelo GARCH, obteniéndose disminuciones tanto en el RMSE como en el MAE para todos los modelos, mientras que el EGARCH-GED presentó el mejor desempeño, reduciendo el RMSE en un 10,99 % y el MAE en un 5,98 %.

Tabla 8.4: Resultados de desempeño de los modelos de pronóstico EGARCH (21 días)
(Fuente: Elaboración Propia)

| Modelo | RMSE | Var (%) | MAE | Var (%) |
|------------------|----------|----------|----------|----------|
| EGARCH-Normal | 0,003335 | -10,30 % | 0,002302 | -4,90 % |
| EGARCH-GED | 0,003309 | -10,99 % | 0,002276 | -5,98 % |
| EGARCH-t-Student | 0,003326 | -10,54 % | 0,002282 | -5,75 % |

Las medidas de desempeño RMSE y MAE de cada modelo EGARCH-ANN individual y el sistema de ensamble para el caso base con redes neuronales de 2 capas ocultas con 10 neuronas por capa se presentan en la [Tabla 8.5](#). Se aprecian importantes disminuciones de ambas medidas de desempeño respecto a los pronósticos del modelo GARCH, las cuales son superiores a las mejoras en el desempeño de los tres modelos EGARCH, mientras que el desempeño del sistema de ensamble supera a todos los modelos híbridos EGARCH-ANN individuales, independiente de los supuestos distributivos, logrando una reducción del RMSE en un 15,91 % y del MAE en un 13,99 %.

Tabla 8.5: Resultados de desempeño de los modelo de pronóstico EGARCH-ANN y sistema de ensamble a 21 días (2 capas, 10 neuronas por capa)
(Fuente: Elaboración Propia)

| Modelo | RMSE | Var (%) | MAE | Var (%) |
|----------------------|----------|----------|----------|----------|
| EGARCH-Normal-ANN | 0,003207 | -13,74 % | 0,002134 | -11,87 % |
| EGARCH-GED-ANN | 0,003208 | -13,71 % | 0,002167 | -10,47 % |
| EGARCH-t-Student-ANN | 0,003211 | -13,62 % | 0,002167 | -10,49 % |
| Sistema de Ensamble | 0,003126 | -15,91 % | 0,002082 | -13,99 % |

La [Figura 8.1](#) muestra un gráfico con los pronósticos de volatilidad del sistema de ensamble y la volatilidad realizada del IPSA a 21 días durante el periodo 2010-2016.

Con la finalidad de obtener resultados más consistentes, se realiza el mismo procedimiento cambiando el horizonte de pronóstico utilizado. De esta forma, la [Tabla 8.6](#) muestra las medidas de desempeño del modelo GARCH(1,1) para el pronóstico de volatilidad del IPSA a 14 y 28 días.

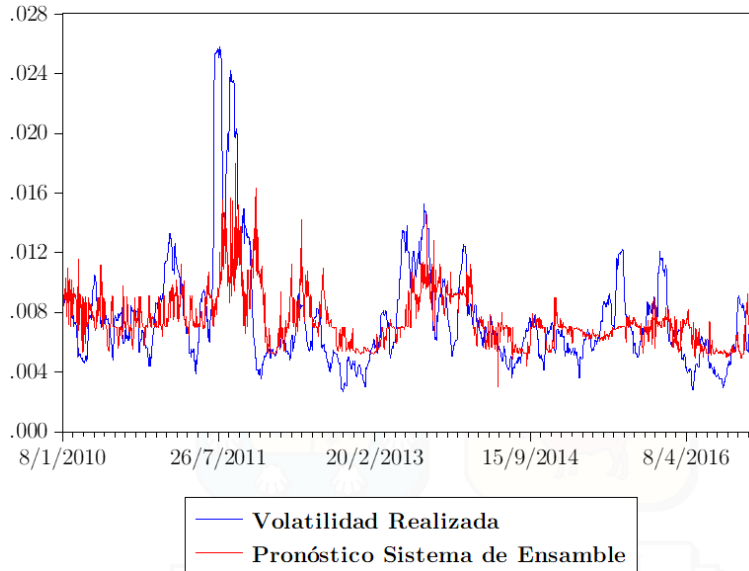


Figura 8.1: Pronóstico de Volatilidad del Sistema de Ensamble

(Fuente: Elaboración Propia)

Tabla 8.6: Resultados de desempeño del modelo de pronóstico GARCH (14 y 28 días)

(Fuente: Elaboración Propia)

| Modelo | 14 días | | 28 días | |
|------------|----------|----------|----------|----------|
| | RMSE | MAE | RMSE | MAE |
| GARCH(1,1) | 0,003600 | 0,002357 | 0,003600 | 0,002357 |

Tabla 8.7: Resultados de desempeño de los modelos de pronóstico EGARCH, modelos híbridos y sistema de ensamble a 14 y 28 días bajo distintas distribuciones de error

(Fuente: Elaboración Propia)

| Modelo | 14 días | | | | 28 días | | | |
|----------------------|----------|----------|----------|----------|----------|----------|----------|----------|
| | RMSE | Var (%) | MAE | Var (%) | RMSE | Var (%) | MAE | Var (%) |
| EGARCH-Normal | 0,003461 | -10,33 % | 0,002417 | -4,99 % | 0,003239 | -10,01 % | 0,002193 | -6,95 % |
| EGARCH-GED | 0,003439 | -10,90 % | 0,002400 | -5,68 % | 0,003212 | -10,78 % | 0,002167 | -8,07 % |
| EGARCH-t-Student | 0,003469 | -10,11 % | 0,002418 | -4,97 % | 0,003224 | -10,44 % | 0,002176 | -7,68 % |
| EGARCH-Normal-ANN | 0,003402 | -11,85 % | 0,002265 | -10,96 % | 0,003104 | -13,77 % | 0,002077 | -11,86 % |
| EGARCH-GED-ANN | 0,003385 | -12,30 % | 0,002287 | -10,13 % | 0,003096 | -13,99 % | 0,002078 | -11,82 % |
| EGARCH-t-Student-ANN | 0,003392 | -12,10 % | 0,002260 | -11,19 % | 0,003101 | -13,86 % | 0,002098 | -10,98 % |
| Sistema de Ensamble | 0,003318 | -14,02 % | 0,002213 | -13,01 % | 0,003038 | -15,60 % | 0,002040 | -13,46 % |

En la [Tabla 8.7](#) se presentan medidas de desempeño para el pronóstico de los modelos EGARCH, modelos híbridos y el sistema de ensamble a 14 y 28 días, donde al igual que para el caso de los pronósticos a 21 días, los sistemas de ensamble presentan el mejor resultado, con reducciones del RMSE en un 14,02 % y del MAE en un 13,01 %

para el horizonte de pronóstico a 14 días, mientras que el sistema de ensamble reduce el RMSE en un 15,60 % y el MAE en un 13,49 %.

8.2. Análisis de Sensibilidad

Luego de obtener los pronósticos de volatilidad por medio del sistema de ensamble para el caso base de redes neuronales con 2 capas ocultas y 10 neuronas por capa, se procede a buscar la mejor arquitectura de dicho sistema mediante la realización de pronósticos utilizando diferentes combinaciones de capas y neuronas por capa en cada una de las cuatro redes neuronales que conforman el sistema de ensamble. Las combinaciones utilizadas para esta tarea contemplan entre una y cinco capas ocultas con 2, 4, 6, 8 y 10 neuronas por capa.

Para determinar el mejor o los mejores modelos de pronóstico de volatilidad de los 25 modelos con diferentes arquitecturas se utiliza el procedimiento de Conjunto de Modelos de Confianza (MCS) (Hansen et al., 2003, 2011), el cual corresponde a un conjunto de modelos de pronóstico que está construido de manera tal que contiene el "mejor" modelo de pronóstico, dado un nivel de confianza determinado.

Los resultados se presentan en la [Tabla 8.8](#). Para el caso de la predicción de la volatilidad de los retornos del IPSA a 21 días, los resultados indican que la mejor arquitectura de las cuatro redes neuronales que conforman el sistema de ensamble es de 5 capas ocultas con 10 neuronas por capa, reduciendo el RMSE en un 17,64 % respecto al pronóstico del GARCH(1,1), lo cual representa una mejora significativa al 5 % según la prueba del procedimiento MCS. Para la volatilidad a 14 días, la mejor configuración de las redes neuronales del sistema de ensamble es de 1 capa oculta con 8 neuronas, que disminuye el RMSE en un 15,05 % respecto al modelo GARCH(1,1) a 14 días, y es significativo al 5 % según la prueba del procedimiento MCS. Finalmente, el mejor modelo para predecir la volatilidad de los retornos del IPSA a 28 días es el sistema de ensamble con 3 capas ocultas y 6 neuronas por capa en cada una de las cuatro redes neuronales que lo componen. Dicho modelo disminuye el RMSE en un 15,65 % y al igual que para los modelos de pronóstico a 21 y 14 días, presenta una mejora significativa al 5 % en base a la prueba del procedimiento MCS.

Tabla 8.8: Resultados de desempeño para los modelos de volatilidad de sistemas de ensamble para diferentes números de capas y neuronas por cada en cada red neuronal

(Fuente: Elaboración Propia)

| N° | Capa | Neuronas | 21 días | | 14 días | | 28 días | |
|----|------|----------|------------------|---------|-----------------|---------|-----------------|---------|
| | | | RMSE | valor p | RMSE | valor p | RMSE | valor p |
| 1 | 1 | 2 | 0,0031807 | | 0,003323 | ** | 0,003082 | *** |
| 2 | 1 | 4 | 0,0032083 | | 0,003324 | ** | 0,003107 | |
| 3 | 1 | 6 | 0,0032248 | | 0,003450 | *** | 0,003081 | *** |
| 4 | 1 | 8 | 0,0031741 | *** | 0,003278 | ** | 0,003097 | |
| 5 | 1 | 10 | 0,0031760 | *** | 0,003296 | ** | 0,003066 | *** |
| 6 | 2 | 2 | 0,0032188 | ** | 0,003314 | ** | 0,003081 | *** |
| 7 | 2 | 4 | 0,0032487 | ** | 0,003358 | *** | 0,003104 | ** |
| 8 | 2 | 6 | 0,0032716 | | 0,003288 | ** | 0,003072 | ** |
| 9 | 2 | 8 | 0,0032155 | | 0,003337 | ** | 0,003042 | ** |
| 10 | 2 | 10 | 0,0031263 | ** | 0,003318 | ** | 0,003038 | ** |
| 11 | 3 | 2 | 0,0031907 | | 0,003784 | | 0,003087 | *** |
| 12 | 3 | 4 | 0,0031859 | *** | 0,003293 | ** | 0,003082 | ** |
| 13 | 3 | 6 | 0,0031846 | | 0,003294 | ** | 0,003036 | ** |
| 14 | 3 | 8 | 0,0032071 | | 0,003285 | ** | 0,003055 | ** |
| 15 | 3 | 10 | 0,0031686 | *** | 0,003285 | ** | 0,003072 | ** |
| 16 | 4 | 2 | 0,0036284 | | 0,003783 | | 0,003081 | ** |
| 17 | 4 | 4 | 0,0031898 | | 0,003281 | ** | 0,003095 | *** |
| 18 | 4 | 6 | 0,0031764 | *** | 0,003289 | ** | 0,003072 | ** |
| 19 | 4 | 8 | 0,0031293 | ** | 0,003415 | ** | 0,003038 | ** |
| 20 | 4 | 10 | 0,0031625 | *** | 0,003344 | ** | 0,003054 | ** |
| 21 | 5 | 2 | 0,0036323 | | 0,003325 | ** | 0,003096 | *** |
| 22 | 5 | 4 | 0,0031754 | | 0,003328 | ** | 0,003090 | *** |
| 23 | 5 | 6 | 0,0031762 | *** | 0,003328 | ** | 0,003081 | * |
| 24 | 5 | 8 | 0,0031491 | *** | 0,003319 | ** | 0,003055 | ** |
| 25 | 5 | 10 | 0,0030619 | ** | 0,003330 | ** | 0,003102 | ** |

Nota: ***, ** y * corresponden a la significancia estadística del 1 %, 5 % y 10 % relacionada con el valor p del MCS de Hansen et al. (2003, 2011).

9. CONCLUSIONES Y RECOMENDACIONES

Un indicador fundamental para la cuantificación de riesgos de mercado es la volatilidad o variabilidad de los rendimientos de un activo financiero, la cual puede no ser directamente observable, por lo que se hace necesario encontrar métodos de pronóstico que proporcionen datos concretos para que los participantes del mercado puedan tomar decisiones financieras de la forma más informada y precisa posible.

Se ha demostrado que los modelos híbridos ANN-EGARCH son efectivos para pronosticar la volatilidad del mercado de valores. Sin embargo, la elección de supuestos distributivos relacionados con la estimación del modelo EGARCH afecta en la precisión del pronóstico de los modelos híbridos al capturar distintas características de los datos como son la normalidad, asimetría y exceso de curtosis, de forma tal que en conjunto, el sistema de ensamble es capaz de capturar las tres características de los datos. Por esta razón, en la presente memoria se implementa un sistema de ensamble para pronosticar la volatilidad, el cual está conformado por tres modelos ANN-EGARCH que poseen diferentes supuestos distributivos y una red neuronal adicional que combina los pronósticos de cada modelo híbrido y proporciona el pronóstico final. El modelo se aplica al mercado de bursátil chileno, utilizando específicamente los retornos del índice bursátil IPSA, considerando los datos de la serie de tiempo de los precios de cierre del IPSA desde el 2 de enero del 2009 al 30 de diciembre del 2016 como base para el desarrollo del estudio.

Respecto a la estimación por medio de los modelos EGARCH que son utilizados como entradas para los modelos híbridos, la significancia de los parámetros de asimetría sugiere la existencia del efecto de apalancamiento, es decir, que la volatilidad de los rendimientos del IPSA se ven afectados, en gran medida, por los efectos de las malas noticias. Además, al comparar los pronósticos de dichos modelos con el ampliamente utilizado modelo GARCH(1,1) para distintos horizontes de pronóstico, se aprecia la superioridad del modelo EGARCH, obteniendo disminuciones del orden del 10 % del RMSE respecto al modelo GARCH, situación que coincide con los resultados de las investigaciones

analizadas en el marco teórico. Por otro lado, destaca que en la selección de los parámetros óptimos, la especificación que permitió realizar los mejores pronósticos de volatilidad para cada uno de los modelos con distintos supuestos distributivos empleados es el modelo EGARCH(1,1), que constituye una especificación simple y con la menor cantidad de parámetros posibles de todas las estimaciones realizadas al variar los parámetros p y q en un rango de 1 a 5.

En cuanto a la utilización de los modelos híbridos ANN-EGARCH, es posible establecer que los modelos basados en redes neuronales son una buena herramienta de predicción dado que son flexibles, robustos y adaptables a muchas situaciones, y a su capacidad de encontrar patrones no lineales en series de tiempo. Tal como señala gran parte de la literatura financiera, el uso de modelos híbridos mejoró considerablemente el desempeño de los pronósticos, obteniendo disminuciones del orden del 13 % del RMSE respecto al modelo GARCH(1,1) para los distintos horizontes de pronóstico utilizados.

Los principales resultados de la presente memoria demuestran que el sistema de ensamble compuesto de modelos híbridos ANN-EGARCH con diferentes supuestos distributivos y una red neuronal adicional que proporciona el pronóstico final, mejora notoriamente las predicciones del modelo GARCH en un 15,91 % y supera el desempeño de cada uno de los tres modelos híbridos ANN-EGARCH por separado en la predicción de la volatilidad de los retornos del IPSA cuando se usa un horizonte de pronóstico de 21 días. Dichos resultados se repiten al el horizonte de pronóstico a 14 y 28 días, obteniéndose disminuciones del 14,02 % y 15,62 % del RMSE, respectivamente.

Desde el punto de vista de la efectividad en relación a los modelos híbridos que lo componen, la precisión de predicción del sistema de ensamble propuesto medido por el RMSE para el pronóstico de volatilidad a 21 días genera mejoras en el desempeño del 2,59 %, 2,62 % y 2,73 % respecto a las tres alternativas híbridas de ANN-EGARCH, situación que se replica al utilizar horizontes de pronóstico de 14 y 28 días, donde se obtienen mejoras del desempeño del mismo orden de magnitud. La mejora en la capacidad predictiva del sistema de ensamble respecto a los modelos híbridos es similar a la obtenida al comparar los modelos EGARCH de cada supuesto distributivo con su respectiva red híbrida ANN-EGARCH, lo cual valida la importancia del progreso que significa la utilización de sistemas de ensamble en la predicción de volatilidad.

Cuando se utiliza la mejor arquitectura para las redes neuronales que conforman el sistema de ensamble sobre los parámetros, los resultados tienden a ser los mismos. Los resultados indican que la mejor de las arquitectura consideradas

para predecir la volatilidad a 21 días es de 5 capas ocultas y 10 neuronas por capa, para la volatilidad a 14 días, es de 1 capa oculta y 8 neuronas, y para la volatilidad de 28 días, es de 3 capas ocultas y 6 neuronas por capa. En todos los casos mencionados, la mejor arquitectura presenta un desempeño superior en comparación al caso base de 2 capas ocultas y 10 neuronas por capa en cada una de las cuatro redes que conforman el sistema de ensamble, y resultan significativas al 5 % según el procedimiento MCS.

En resumen, los resultados son consistentes y robustos, y permiten concluir que el uso de sistemas de ensamble es efectivo en la previsión de volatilidad del IPSA, mejorando el desempeño de los modelos de pronóstico ampliamente utilizados para diferentes horizontes de pronósticos y diversas arquitecturas de las redes neuronales que conforman el sistema de ensamble analizado.

Dentro del ámbito de las inversiones, los resultados presentados aquí pueden resultar útiles para diferentes agentes para administrar mejor sus inversiones en el mercado accionario chileno, lo cual se vuelve más relevante aún dada la apertura de los mercados bursátiles alrededor del mundo, que se han transformado en alternativas de inversión cada vez más accesibles para la gente común y corriente.

Teniendo en cuenta que la presente memoria tiene entre sus objetivos determinar el potencial de los sistemas de ensamble en la mejora de los pronósticos de modelos de volatilidad ampliamente utilizados en aplicaciones financieras, la única serie de tiempo empleada corresponde a los retornos logarítmicos del IPSA. Sin embargo, luego de demostrar la mejora en el desempeño de los sistemas de ensamble para predecir la volatilidad financiera, surge la motivación de realizar investigaciones futuras como por ejemplo la implementación un sistema de ensamble en el cual se incorporen diversas variables financieras adicionales como entradas a las redes neuronales, tales como el volumen de transacciones, precios de cierre, retornos del índice, o incluso la volatilidad realizada rezagada en el tiempo, combinar otro tipo de modelos de pronóstico por medio del sistema de ensamble, como por ejemplo otros tipos de redes neuronales artificiales o máquinas de soporte vectorial, o por otro lado utilizar sistemas de ensamble aplicados al pronóstico de la volatilidad de commodities.



REFERENCIAS

- Abdelaal, M. (2011). Modeling and forecasting time varying stock return volatility in the egyptian stock market. [5.3](#)
- Akgiray, V. (1989). Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *Journal of business*, (pp. 55–80). [5.3](#)
- Aleksander, I. y Morton, H. (1990). *An introduction to neural computing*. Londres, Inglaterra: Chapman and Hall. [6.3](#)
- Alfaro, R y Silva, C. G. (2008). Volatilidad de índices accionarios: el caso del ipsa. *Cuadernos de economía*, 45(132), 217–233. [5.6](#)
- Ali, I. y Talukdar, A.K. (2017). Stock market volatility and returns: A study of six sectorial índices of national stock exchange in india. *International Journal of Research in Management Social Science*, 5(3), 47–55. [5.3](#)
- Aranda, R. y Jaramillo, P. (2010). Dinámica no lineal en el mercado accionario chileno: evidencia de retornos y volúmenes transados. *Economía Chilena*, 13(3), 67–94. [5.6](#)
- Awartani, B. y Corradi, V. (2005). Predicting the volatility of the s&p-500 stock index via garch models: the role of asymmetries. *International Journal of Forecasting*, 21(1), 167–183. [5.3](#)
- Bera, A. y Higgins, M. (1993). Arch models: properties, estimation and testing. *Journal of economic surveys*, 7(4), 305–366. [5.3](#)
- Bildirici, M. y Ersin, Ö.Ö. (2009). Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in istanbul stock exchange. *Expert Systems with Applications*, 36(4), 7355–7362. [5.4](#)
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327. [5.3](#), [6.2.2.2](#)

- Bollerslev, T.; Chou, R.; y Kroner, K. (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of econometrics*, 52(1-2), 5–59. 5.3
- Bollerslev, T.; Engle, R.; y Nelson, D. (1994). Arch models. *Handbook of econometrics*, 4, 2959–3038. 5.3
- Bolsa de Comercio de Santiago (2016). Mercados renta variable. <http://www.bolsadesantiago.com/mercado/Paginas/indicesbursatiles.aspx>. [Consultado el 17 de Octubre de 2016]. 5.6, 7.1
- Bosarge, W. E. (1993). Adaptive processes to exploit the nonlinear structure of financial markets. *Neural Networks in Finance and Investing. Probus Publishing*, 1(1), 371–402. 1
- Bowerman, B. y O’Connell, R. (1993). Forecasting and time series: An applied approach. 6.1
- Box, G. y Jenkins, G. (1978). *Time series analysis: forecasting and control*. San Francisco, EEUU: Holden Day. 6.2.1, 6.2.2
- Box, G. y Tiao, G. (1975). Bayesian inference in statistical analysis. 26(12), 223–239. 6.2.3
- Camargo, E.; Fontanilla, A.; y Luna, R. (2012). Modelación de la volatilidad y pronóstico del índice general de la bolsa de valores de colombia (igbc). *Clío América*, 6(12), 223–239. 5.3
- Cao, C. y Tsay, R. (1992). Nonlinear time-series analysis of stock volatilities. *Journal of Applied Econometrics*, 7(1). 5.3
- Carroll, S.M. y Dickinson, B.W. (1989). Construction of neural nets using the radon transform. In *IJCNN Proc. I* (pp. 607–611). 5.1
- Casas, M. y Cepeda, E. (2008). Modelos arch, garch y egarch: aplicaciones a series financieras. *Cuadernos de economía*, 27(48), 287–319. 5.3
- Chaigusin, S.; Chirathamjaree, C.; y Clayden, J. (2008). The use of neural networks in the prediction of the stock exchange of thailand (SET) index. In *Computational Intelligence for Modelling Control & Automation, 2008 International Conference on* (pp. 670–673).: IEEE. 5.1
- Cumby, R.; Figlewski, S.; y Hasbrouck, J. (1993). Forecasting volatilities and correlations with egarch models. *The Journal of Derivatives*, 1(2), 51–63. 5.3

- Del Carpio, J. (2005). Las redes neuronales artificiales en las finanzas. *Industrial Data*, 8(2), 28–32. [5.1](#)
- Díaz, A.o; Glover, F.; Ghaziri, H.; González, J.L.; Laguna, M.; Moscato, P.; y Tseng, F. (1996). Optimización heurística y redes neuronales. [6.3.1](#)
- Ding, Z.; Granger, C.; y Engle, R. (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1), 83–106. [5.3](#)
- Donaldson, R. y Kamstra, M. (1997). An artificial neural network-garch model for international stock return volatility. *Journal of Empirical Finance*, 4(1), 17–46. [2](#), [5.4](#)
- Emenike, K. (2010). Modelling stock returns volatility in nigeria using garch models. [5.3](#)
- Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 50(4), 987–1007. [5.3](#), [6.2.2.1](#)
- Engle, R. (2004). Riesgo y volatilidad: Modelos econométricos y práctica financiera. *Revista Asturiana de Economía RAE*, 1(31), 221–252. [6.2](#), [6.2.2.1](#)
- Engle, R. y Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric reviews*, 5(1), 1–50. [5.3](#)
- Ezzat, H. (2012). The application of garch and egarch in modeling the volatility of daily stock returns during massive shocks: The empirical case of egypt. [5.3](#)
- Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383–417. [5.2](#)
- Feldman, J. y Ballard, D. (1982). Connectionist models and their properties. *Cognitive science*, 6(3), 205–254. [5.1](#)
- Ferreira, G.; Navarrete, J.; Rodríguez-Cortés, F.; y Mateu, J. (2017). Estimation and prediction of time-varying GARCH models through a state-space representation: a computational approach. *Journal of Statistical Computation and Simulation*, 87(12), 2430–2449. [5.6](#)
- Floros, C. (2008). Modelling volatility using garch models: evidence from egypt and israel. *Middle Eastern Finance and Economics*, (2), 31–41. [5.3](#)

- Funahashi, K. (1989). On the approximate realization of continuous mappings by neural networks. *Neural networks*, 2(3), 183–192. [5.1](#)
- Gonzalez, R.; Padilha, C.; y Barone, Dante A. (2015). Ensemble system based on genetic algorithm for stock market forecasting. In *Evolutionary Computation (CEC), 2015 IEEE Congress on* (pp. 3102–3108).: IEEE. [5.5](#)
- Gregoire, J. (1985). El ajuste de los precios accionarios a la información: resultados empíricos. *Paradigmas en administración*, 7(1), 121–131. [5.6](#)
- Grossberg, S. (1987). *The adaptive brain I: Cognition, learning, reinforcement, and rhythm*, volume 42. Elsevier. [5.1](#)
- Gujarati, D. (2004). Econometría. *Editorial Mac Graw Hill*,. [6.1](#), [6.1.1](#)
- Guresen, E.; Kayakutlu, G.; y Daim, T. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, 38(8), 10389–10397. [5.1](#)
- Gybenko, G (1989). Approximation by superposition of sigmoidal functions. *Mathematics of Control, Signals and Systems*, 2(4), 303–314. [5.1](#)
- Hajizadeh, E.; Seifi, A.; Zarandi, M.; y Turksen, I. (2012). A hybrid modeling approach for forecasting the volatility of S&P 500 index return. *Expert Systems with Applications*, 39(1), 431–436. [5.4](#)
- Hansen, P.; Lunde, A.; y Nason, J. (2003). Choosing the best volatility models: the model confidence set approach. *Oxford Bulletin of Economics and Statistics*, 65(s1), 839–861. [8.2](#), [8.8](#)
- Hansen, P.; Lunde, A.; y Nason, J. (2011). The model confidence set. *Econometrica*, 79(2), 453–497. [8.2](#), [8.8](#)
- Hebb, D.O. (1949). *The organization of behavior: A neuropsychological approach*. John Wiley & Sons. [5.1](#)
- Hecht-Nielsen, R. (1988). Theory of the backpropagation neural network. *Neural Networks*, 1(1), 445–448. [5.1](#)
- Hernández, S. (2009). Pronóstico y volatilidad del ipyc de la bolsa mexicana de valores. *Revista Observatorio Calasanz*, 1(1), 25–36. [5.3](#)
- Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8), 2554–2558. [5.1](#)

- Hopfield, J.J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the national academy of sciences*, 81(10), 3088–3092. [5.1](#)
- Hopfield, J.J. y Tank, D. (1986). Computing with neural circuits- a model. *Science*, 233(4764), 625–633. [5.1](#)
- Hornik, K.; Stinchcombe, M.; y White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5), 359–366. [2](#), [5.1](#)
- Hutchinson, J.; Lo, A.; y Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance*, 49(3), 851–889. [5.1](#)
- Jiménez, J. and Ruiz, R. (2000). Las redes neuronales en su aplicación a las finanzas. 54(1), 251–257. [5.1](#), [6.4](#)
- Komo, D.; Chang, C.; y Ko, H. (1994). Neural network technology for stock market index prediction. In *Speech, Image Processing and Neural Networks, 1994. Proceedings, ISSIPNN'94., 1994 International Symposium on* (pp. 543–546).: IEEE. [5.1](#)
- Kristjanpoller, W. (2009). An Analysis of the Day-of-the-Week Effect in Latin American Stock Markets. *Lecturas de Economía*, 71(1), 189–207. [5.6](#)
- Kristjanpoller, W. y Caballero, V. (2012). Volumen y asimetría en los principales mercados accionarios latinoamericanos. *Lecturas de Economía*, 76(1), 119–141. [5.6](#)
- Kristjanpoller, W.; Fadic, A.; y Minutolo, M. (2014). Volatility forecast using hybrid neural network models. *Expert Systems with Applications*, 41(5), 2437–2442. [5.4](#), [5.6](#), [6.5](#)
- Kristjanpoller, W. y Muñoz, R. (2012). Análisis del Efecto Día de Semana en los principales mercados accionarios latinoamericanos: una aproximación mediante el criterio de Dominancia Estocástica. *Estudios de economía*, 39(1), 5–26. [5.6](#)
- Lahmiri, S. (2014). Intelligent ensemble systems for modeling nasdaq microstructure: A comparative study. In *IAPR Workshop on Artificial Neural Networks in Pattern Recognition* (pp. 240–251).: Springer. [5.5](#)
- Lahmiri, S. y Boukadoum, M. (2015a). An ensemble system based on hybrid egarch-ann with different distributional assumptions to predict s&p 500 intraday volatility. *Fluctuation and Noise Letters*, 14(1), 1–10. [5.5](#), [6.2](#), [6.4](#), [6.7](#), [6.8](#)

- Lahmiri, S. y Boukadoum, M. (2015b). Intelligent ensemble forecasting system of stock market fluctuations based on symmetric and asymmetric wavelet functions. *Fluctuation and Noise Letters*, 14(4), 1550033. 5.5
- LeCun, Y. (1985). Une procédure d'apprentissage pour réseau a seuil asymmetrique (a learning scheme for asymmetric threshold networks). In *Proceedings of Cognitiva 85, Paris, France*. 6.3.5
- Li, Y. y Ma, W. (2010). Applications of artificial neural networks in financial economics: a survey. In *Computational Intelligence and Design (ISCID), 2010 International Symposium on*, volume 1 (pp. 211–214): IEEE. 5.1
- Ling, Y.; Yue, B.; y Zhang, H. (2014). A new wrapped ensemble approach for financial forecast. *Journal of Intelligent Systems*, 23(1), 21–32. 5.5
- Liu, H. y Hung, J. (2010). Forecasting s&p-100 stock index volatility: The role of volatility asymmetry and distributional assumption in garch models. *Expert Systems with Applications*, 37(7), 4928–4934. 5.3
- Lu, X.; Que, D.; y Cao, G. (2016). Volatility Forecast Based on the Hybrid Artificial Neural Network and GARCH-type Models. *Procedia Computer Science*, 91(1), 1044–1049.
- Ludlow, J. y Mota, B. (2006). Volatilidad del ipc, nasdaq y s&p500: un modelo garch multivariado. *Análisis económico*, 21(48), 215–227. 5.3
- Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2), 431–441. 6.3.5
- Marshall, P. y Walker, E. (2000). Day-of-the-Week and Size Effects in emerging markets: Evidence from Chile. 15(2), 89–108. 5.6
- Martin del Brio, B. y Sanz, A. (1997). *Neural Networks and Fuzzy Systems, Theory and Practice*. Madrid, España: RA-MA Press, Madrid. 6.3
- McCulloch, W. y Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115–133. 5.1, 6.2
- Ministerio de Hacienda Gobierno de Chile (2016). ¿qué es el mercado de capitales chileno? <http://www.hacienda.cl/mercado-de-capitales/preguntas-frecuentes/que-es-el-mercado-de-capitales-chileno.html>. [Consultado el 17 de Octubre de 2016]. 5.2

- Minsky, M. y Papert, S. (1969). *Perceptrons: An introduction to computational geometry*. MIT press. [6.3.3](#)
- Miron, D. y Tudor, C. (2010). Asymmetric conditional volatility models: Empirical estimation and comparison of forecasting accuracy. *Romanian Journal of Economic Forecasting*, 13(3), 74–92. [5.3](#)
- Nazarian, R.; A., Gandali; Naderi, E.; y Amiri, A. (2013). Forecasting stock market volatility: A forecast combination approach. [5.4](#)
- Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 59(2), 347–370. [5.3](#), [6.2.2.3](#), [7.2](#), [8.1](#)
- Ornoneit, D. y Neuneier, R. (1996). Experiments in predicting the german stock index dax with density estimating neural networks. In *Computational intelligence for financial engineering, 1996., proceedings of the ieee/iafe 1996 conference on* (pp. 66–71): IEEE. [5.1](#)
- Orozco, J. y Velásquez, J. (2013). A new forecasting combination system for predicting volatility. *Innovar*, 23(50), 5–16. [5.5](#)
- Pagan, A. y Schwert, G. (1990). Alternative models for conditional stock volatility. *Journal of econometrics*, 45(1), 267–290. [5.3](#)
- Parisi, A.; Parisi, F.; y Guerrero, J.L. (2002). Evaluación de modelos de Redes neuronales de predicción del signo de la variación del IPSA. *Estudios de Administración*, 9(1), 67–103. [5.6](#)
- Parisi, F. y Acevedo, C. (2001). Volume and autocovariance in short-horizon stock returns: evidence from 1992 to 1998 in Chile. *International Review of Financial Analysis*, 10(3), 275–285. [5.6](#)
- Parisi, F. y Vasquez, A. (2000). Simple technical trading rules of stock returns: evidence from 1987 to 1998 in Chile. *Emerging Markets Review*, 1(2), 152–164. [5.6](#)
- Parker, D. (1982). Learning logic. invention report s81-64, file 1, office of technology licensing. *October, Stanford University*. [6.3.5](#)
- Poon, S. y Granger, C. (2003). Forecasting volatility in financial markets: A review. *Journal of economic literature*, 41(2), 478–539. [6.2.2](#)

- Posch, T.E. (1968). *Models of the Generation and Processing of Signals by Nerve Cells: A Categorically Indexed Abridged Bibliography*. Electronic Sciences Laboratory, University of Southern California. [5.1](#)
- Qi, M. (1999). Nonlinear predictability of stock returns using financial and economic variables. *Journal of Business & Economic Statistics*, 17(4), 419–429. [6.3.5](#)
- Rajvanshi, V.; Santra, A.; y Basu, S. (2017). Implied volatility and predictability of garch models. [5.3](#)
- Ramón y Cajal, S. (1888). Estructura de los centros nerviosos de las aves. [6.3.1](#)
- Robles, M.D. (2002). Medidas de Volatilidad. *31*(114), 1073–1110. [6.2.2.2](#)
- Roh, T. (2007). Forecasting the volatility of stock price index. *Expert Systems with Applications*, 33(4), 916–922. [5.4](#)
- Romero-Meza, R.; Bonilla, C.; y Hinich, M. (2007). Nonlinear event detection in the Chilean stock market. *Applied Economics Letters*, 14(13), 987–991. [5.6](#)
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386. [5.1](#), [6.3.3](#)
- Rumelhart, D. E.; Hinton, G. E.; y McClelland, J. L. (1986). A general framework for parallel distributed processing. *Parallel distributed processing: Explorations in the microstructure of cognition*, 1, 45–76. [5.1](#)
- Saatcioglu, K. y Starks, L. (1998). The stock price–volume relationship in emerging stock markets: the case of Latin America. *International Journal of forecasting*, 14(2), 215–225. [5.6](#)
- Sejnowski, T. y Rosenberg, C. (1988). *NETtalk: A parallel network that learns to read aloud*. MIT Press. [5.1](#)
- Su, C. (2010). Application of egarch model to estimate financial volatility of daily returns: The empirical case of china. [5.3](#)
- Taylor, S (1986). Modeling financial time series john wiley & sons. *Great Britain*. [5.3](#)
- Terui, N. y Van Dijk, H. (2002). Combined forecasts from linear and nonlinear time series models. *International Journal of Forecasting*, 18(3), 421–438. [5.5](#)

- Trippi, R. R. y Turban, E. (1992). *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real-World Performance*. Nueva York, Inglaterra: Edición Revisada, McGraw Hill. 5.1
- Urrutia, J. (1994). Time series properties of four latin american equity markets: Argentina, brazil, chile and mexico. *Estudios de Administración*, 2(1), 1–8. 5.6
- Villada, F.; Muñoz, N.; y García, E. (2012). Aplicación de las redes neuronales al pronóstico de precios en el mercado de valores. *Información tecnológica*, 23(4), 11–20. 5.1, 6.3.5
- Wang, J.; Wang, J.; Zhang, Z.; y Guo, S. (2011). Forecasting stock indices with back propagation neural network. *Expert Systems with Applications*, 38(11), 14346–14355. 5.1
- Werbos, P. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. *Doctoral Dissertation, Applied Mathematics, Harvard University, MA*. 6.3.4, 6.3.5
- White, H. (1988). Economic prediction using neural networks: The case of ibm daily stock returns. 5.4
- White, H. (1992). Nonparametric estimation of conditional quantiles using neural networks. In *Proceedings of 23rd symposium on the interface, computer science and statistics* (pp. 190–199). 5.1
- Widrow, B. y Hoff, M. (1960). *Adaptive switching circuits*. Technical report, Stanford Univ. CA Stanford Electronics Labs. 5.1
- Xiao, B. (2016). Predicting the volatility of the russell 3000 stock index. *International Journal of Financial Research*, 7(4), 18–28. 5.3
- Yalama, A. y Sevil, G. (2008). Forecasting world stock markets volatility. *International Research Journal of Finance and Economics*, 15, 159–174. 5.3
- Zakoian, J.M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, 18(5), 931–955. 5.3
- Zamora, L. and Guarda, D. and Lay, N. (2015). Patrones de Comportamiento Bursátil: Una Radiografía del Mercado Bursátil Chileno. http://www.faceauv.cl/cladea2015/images/stories/track2/CLADEA_2015_submission_88__2_.pdf. [Consultado el 3 de Octubre de 2016]. 2

Zhu, X.; Wang, H.; Xu, L.; y Li, H. (2008). Predicting stock index increments by neural networks: The role of trading volume under different horizons. *Expert Systems with Applications*, 34(4), 3043–3054. [5.1](#)

