



UNIVERSIDAD  
POLITECNICA  
DE VALENCIA



**TRABAJO FIN DE MÁSTER EN DIRECCIÓN FINANCIERA Y FISCAL**

# **“PREDICCIÓN DEL ÍNDICE BURSÁTIL IBEX 35 MEDIANTE SELECCIÓN DE INPUTS CON REDES NEURONALES”**

**AUTORA: LAURA FERNÁNDEZ BENLLOCH**

**TUTOR: JAVIER OLIVER MUNCHARAZ**

**Curso Académico: 2015/2016**

# ÍNDICE:

I. INTRODUCCIÓN Y OBJETIVOS.....	4
CAPÍTULO I: INTRODUCCIÓN Y OBJETIVOS. ....	5
II. LAS REDES NEURONALES ARTIFICIALES.....	8
CAPÍTULO II: LAS REDES NEURONALES ARTIFICIALES .....	9
2.1.: CONCEPTO.....	9
2.2.: CLASIFICACIÓN DE REDES NEURONALES.....	15
2.3.: CONSTRUCCIÓN DE REDES NEURONALES.....	20
2.4.: PROCESO DE APRENDIZAJE DE RNA .....	24
2.5.: REDES NEURONALES BACKPROPAGATION .....	26
III. ESTUDIO COMPARATIVO DE LA APLICACIÓN DE LA RED NEURONAL.....	32
CAPÍTULO III: ESTUDIO COMPARATIVO DE LA APLICACIÓN DE LA RED NEURONAL .....	33
IV. CONCLUSIONES.....	52
CAPÍTULO IV: CONCLUSIONES.....	53
V. BIBLIOGRAFÍA.....	55
CAPÍTULO V: BIBLIOGRAFÍA .....	56

## ÍNDICE DE FIGURAS Y TABLAS:

<i>Ilustración 1: Estructura de la red Perceptrón</i>	16
<i>Ilustración 2 Estructura de la red Adaline</i>	17
<i>Ilustración 3 Esquema de estructura de una red neuronal</i>	20
<i>Ilustración 4 Función de Activación escalón</i>	22
<i>Ilustración 5 Función de Activación Lineal o Identidad</i>	22
<i>Ilustración 6 Función de Activación lineal-mixta</i>	22
<i>Ilustración 7 Función de Activación Sigmoidal</i>	23
<i>Tabla 1 Tabla de correlaciones de todos los inputs.</i>	34
<i>Tabla 2 Tabla de correlaciones medias móviles.</i>	35
<i>Tabla 3 Tabla de correlaciones retardos del IBEX.</i>	36
<i>Tabla 4 Tabla de correlaciones retardos del Futuro del IBEX.</i>	36
<i>Tabla 5 Tabla de correlaciones retardos del DAX.</i>	37
<i>Tabla 6 Tabla de correlaciones retardos del Futuro del Petróleo.</i>	37
<i>Tabla 7 Tabla de correlaciones retardos de Standard and Poors.</i>	38
<i>Tabla 8 Tabla de correlaciones retardos del FTSE.</i>	38
<i>Tabla 9 Tabla de correlaciones retardos del Dow Jones Industrial.</i>	39
<i>Tabla 10 Tabla de correlaciones retardos del Dow Jones Transportes.</i>	39
<i>Tabla 11 Tabla de correlaciones retardos del EURUSD.</i>	40
<i>Tabla 12 Tabla de correlaciones de medias móviles</i>	43
<i>Tabla 13 Tabla de correlaciones de los retardos R1</i>	44
<i>Tabla 14 Tabla de correlaciones de los retardos R2</i>	45
<i>Tabla 15 Tabla de correlaciones de los retardos R3</i>	46
<i>Tabla 16 Tabla de correlaciones de los retardos R4</i>	47
<i>Tabla 17 Tabla de correlaciones de los retardos R5</i>	48
<i>Tabla 18 Error de predicción con todos los inputs</i>	50
<i>Tabla 19 Error de predicción con los inputs seleccionados</i>	50
<i>Gráfica 1 Contribución de las medias móviles de los inputs.</i>	43
<i>Gráfica 2 Contribución de los retardos 1 de los inputs.</i>	44
<i>Gráfica 3 Contribución de los retardos 2 de los inputs.</i>	45
<i>Gráfica 4 Contribución de los retardos 3 de los inputs.</i>	46
<i>Gráfica 5 Contribución de los retardos 4 de los inputs.</i>	47
<i>Gráfica 6 Contribución de los retardos 5 de los inputs.</i>	48

# I. INTRODUCCIÓN Y OBJETIVOS

## CAPÍTULO I: INTRODUCCIÓN Y OBJETIVOS.

Durante los últimos años, ha habido diversos cambios debido a una gran expansión en los mercados financieros y bursátiles, como consecuencia de dos factores importantes, una fuerte globalización a nivel internacional de los mercados y el desarrollo tecnológico de las telecomunicaciones. Se ha observado una mayor afluencia de inversores posibilitando la adquisición de un mayor campo de inversiones debido a que pueden invertir en cualquier mercado (a nivel nacional e internacional), como por ejemplo la aparición de los mercados de derivados en el año 1982, aumentando las posibilidades de obtener una mayor ganancia y a su vez reduciendo costes de maniobra.

Los inversores se han encontrado con un mercado amplio, con múltiples operaciones de inversión, que anteriormente no existían, una forma más rápida de realizar sus operaciones y menos costosa. Ante dichos acontecimientos, muchos economistas han estudiado e intentado comprender los movimientos de los precios en la bolsa de valores, debido a los riesgos a los que se encuentra sujeto el mercado y la incertidumbre que presenta. Se han propuesto diversos métodos para brindar predicciones más precisas para los inversores y de esta manera reducir el mayor riesgo posible. Gracias a las nuevas tecnologías, se han podido desarrollar una serie de sistemas de predicción que permitirán reducir el riesgo de los inversores en sus operaciones y aumentará por tanto los movimientos en los mercados financieros.

El análisis técnico y/o fundamental tiene como objetivo la toma de decisión de compra/venta de un activo financiero. No realizan una predicción exacta del precio. Mientras que el análisis técnico se ciñe a la evolución de los precios, mediante los gráficos de cotizaciones pretende predecir la evolución de los mismos. Por otro lado, el análisis fundamental trata de obtener un valor teórico al que debería cotizar la empresa teniendo en cuenta sus datos contables, financieros, así como sus expectativas. Existen otros modelos más cuantitativos que intentan también predecir bien la evolución de los precios, su dirección y o variabilidad (volatilidad) de los mismos. Existen diferentes modelos basados en la Econometría con modelos ARIMA y de la familia GARCH, otros utilizan las máquinas de vector soporte, redes neuronales, etc...

Dada la necesidad de reducir la incertidumbre o el ruido a la hora de predecir un valor para poder tomar una decisión con el menor riesgo posible, existen numerosas investigaciones y estudios de modelos de predicción que se han desarrollado a lo largo de estos años y que han servido de ejemplo para otros investigadores economistas para encontrar el modelo más eficaz para ello, entre ellos

se encuentran modelos autorregresivos con promedios móviles (Arima)<sup>1</sup>; de regresión múltiple; de algoritmo genético<sup>2</sup>; redes neuronales artificiales<sup>3</sup>; suavizado exponencial; métodos lineales y no lineales, entre otros (Ariyo *et al* Oluyinka *et al* Korede, 2014).

El avance de las tecnologías informáticas se caracteriza por inspirarse en los sistemas biológicos para resolver problemas, ofreciendo soluciones cada día de más fácil implementación. A su vez, ha surgido un conjunto de metodologías que comprenden la lógica borrosa<sup>4</sup>, basada en lo relativo de lo observado como posición diferencial; las redes neuronales, utilizadas en la investigación desarrollada en el presente trabajo; el razonamiento aproximado, basado en el razonamiento humano; los algoritmos genéticos<sup>5</sup>, que mediante una serie de pasos organizados describe el proceso que se debe seguir para dar solución a un problema específico; la teoría del caos<sup>6</sup>, que explica que el resultado de algo depende de distintas variables y que es imposible de predecir; y la teoría del aprendizaje<sup>7</sup>, que pretenden describir los procesos mediante los cuales tanto los seres humanos, como los animales aprenden.

Actualmente, las Redes Neuronales Artificiales son las de mayor impacto gracias a su aplicación práctica. Es utilizada principalmente por profesionales dedicados a la estadística y el análisis de datos, los cuales incorporan las redes neuronales al conjunto de herramientas estadísticas orientadas a clasificar patrones y estimar variables continuas.

El presente trabajo tiene como objetivo la modelización del precio del índice Ibex-35 mediante una red neuronal de tipo Backpropagation, en el que se realizará una preselección de los inputs. Dicha preselección se estima necesaria dado que, a pesar de la capacidad de generalización de un problema de la red Backpropagation, la incorporación de inputs altamente correlacionados entre sí puede provocar ruido en el proceso de aprendizaje de la red. Para evitar este hecho, se pretende realizar una preselección de aquellos inputs que contribuyan y tengan un peso específico mayor en la predicción del índice Ibex 35. Dicha preselección se realizará también mediante una red neuronal Backpropagation pretendiendo conseguir una reducción del ruido de

---

<sup>1</sup> Desarrollado a finales de los sesenta del siglo XX. Box y Jenkins en el año 1976. Time Series Analysis: Forecasting and Control, 2nd. ed. Oakland, CA: Holden-Day.

<sup>2</sup> Investigación desarrollada por John Henry Holland, fue una de las líneas más prometedoras de la inteligencia artificial, en el año 1970.

<sup>3</sup> Alan Turing en 1936 fue el precursor del estudio del cerebro humano asimilándolo a la computación. En 1943, Warren McCulloch y Walter Pitts, propusieron el primer modelo matemático de red neuronal.

<sup>4</sup> Fue formulada en 1965 por el ingeniero y matemático Lotfi Zadeh.

<sup>5</sup> Desarrollada a partir de la inteligencia artificial de John Henry Holland. D. E. Goldberg. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA. 1989. Genetic Algorithms in Search, Optimization and Machine Learning.

<sup>6</sup> Surge cuando en 1963, Edward Lorenz dio a conocer su modelo climático.

<sup>7</sup> Desarrollado por diversos autores, entre los que destacan Jean Piaget, Lev Vygostsky, Frederic Skinner y Jerome Bruner.

posibles interacciones entre dichos inputs, mediante el método de redes neuronales de tipo Backpropagation. En este proyecto justificaremos si la metodología aplicada en las redes neuronales Backpropagation demuestra su teoría y mejora los resultados obtenidos haciendo una comparativa con la aplicación del método de redes neuronales sin preselección, o no.

En el capítulo 2 del presente trabajo se presenta, una descripción de las redes neuronales artificiales y su clasificación, así como la construcción de las mismas. Además, se estudiará los mecanismos y tipos de aprendizaje de las redes y se hará una revisión de los principales modelos desarrollados. Para finalizar este capítulo, se describe en profundidad la red neuronal Backpropagation, su estructura y proceso de aprendizaje, que se utiliza en el estudio sobre la predicción de índice Ibex-35.

En el capítulo 3 presenta un posible modelo de predicción aplicado al índice bursátil Ibex-35 mediante el uso de una red neuronal de tipo Backpropagation. Para ello, se realiza una preselección de los inputs que formarán parte de la red neuronal. Se analiza los errores de predicción cometidos con el modelo comparándolo con una red neuronal en la que no se haya realizado dicha preselección. Dichas predicciones se realizan diariamente obteniendo un análisis estadístico de los errores cometidos por ambos modelos y verificar la hipótesis de partida de la necesidad de realizar una preselección de los inputs.

Por último, en el capítulo 4, se establecen las principales conclusiones obtenidas en la comparativa de dicho estudio. Se analiza también la mejora que supondría la puesta en práctica de dicho modelo en la toma de decisiones a la hora de invertir en el mercado de valores dado que la preselección de los inputs puede tener un efecto importante en la eficacia y eficiencia en el proceso de aprendizaje de la red que provocan mejoras en las predicciones frente a una red neuronal en la que no se realiza dicha preselección.

Finalmente se acompaña la bibliografía consultada para la elaboración del presente trabajo, así como un anexo con las tablas y gráficos resultantes del proceso realizado tanto en la preselección de los inputs, entrenamiento de las redes, como en los resultados obtenidos en la predicción del índice Ibex-35.

## **II. LAS REDES NEURONALES ARTIFICIALES.**

## CAPÍTULO II: LAS REDES NEURONALES ARTIFICIALES

### 2.1.: CONCEPTO

Las redes neuronales artificiales se definen como una serie de algoritmos matemáticos con la función de modelar la forma de procesamiento de la información en sistemas nerviosos biológicos. Tienen su aplicación en diversos campos científicos y principalmente en la economía y las finanzas donde se emplea como método para hacer predicciones de series temporales.

Las redes neuronales artificiales surgen en las primeras décadas del siglo XX. En finanzas, se destaca la aplicación de las redes neuronales en un trabajo realizado por White en 1988, en el que estudió el precio diario de las acciones de IBM mediante un modelo de red neuronal. A partir de este trabajo, se extendió la aplicación de las redes neuronales a las finanzas principalmente en trabajos relacionados con las cotizaciones bursátiles y tipos de cambio, entre otras variables económicas y financieras. Las RNA pueden programarse para su ejecución en paralelo y trabajar incluso en tiempo real (Vanstone *et al* Finnie, 2008).

A continuación, se describe, de forma resumida, la importancia y evolución histórica de las redes neuronales:

**1936- Alan Turing.** Estudió el cerebro humano en base a los fundamentos matemáticos sobre las redes neuronales de Warren McCulloch, neurofisiólogo, y Walter Pitts, matemático, aplicando sus conocimientos a la computación. Posteriormente Warren McCulloch y Walter Pitts, lanzaron una teoría en 1943 sobre el funcionamiento y modo de trabajo de las neuronas, que fue muy importante para el desarrollo de futuros estudios e investigaciones.

**1949- Donald Hebb.** Desarrolló los procesos de aprendizaje de las neuronas basándose en una teoría sobre cómo se desarrollaba el proceso de aprendizaje. Se puede afirmar que se convirtió en la teoría base de la mayoría de las funciones de aprendizaje de las redes neuronales.

**1950- Karl Lashley.** Tras varias pruebas y ensayos, descubrió que la información del cerebro humano no era almacenada en forma centralizada en el cerebro, sino que era distribuida por todo él.

**1954- Marvin Minsky.** Realizó las primeras prácticas de redes neuronales junto con Dean Edmons aplicando modelos y teorías de McCulloch y Pitts, incorporando el aprendizaje Hebbiano de las redes neuronales.

**1956- Congreso de Dartmouth.** Acontecimiento considerado como el nacimiento de la inteligencia artificial desde su celebración.

**1957- Frank Rosenblatt.** Investigó y desarrolló lo que conocemos como Perceptrón. El Perceptrón es la red neuronal más antigua teniendo en la actualidad varias aplicaciones como reconocedor de patrones. No obstante, según los estudios avanzados, el modelo presentaba limitaciones al ser incapaz de realizar clasificaciones de clases no separables linealmente.

**1960- Bernard Widrow/Marcial Hoff.** Juntos desarrollaron el modelo Adaline (Adaptative Linear Elements), que es la primera aplicación de una red neuronal a un problema real consiguiendo la eliminación de los ecos en las líneas telefónicas.

**1969- Marvin Minsky/Seymour Papert.** Publicaron el libro titulado "Perceptrons", donde desarrollan y defienden la idea de que el Perceptrón, realmente no era capaz de resolver problemas relativamente fáciles como el aprendizaje de una función no lineal, funciones primordiales en el campo de la computación y en muchos problemas reales.

**1974- Paul Werbos.** Desarrolla la idea base del algoritmo de aprendizaje de propagación hacia atrás de las redes neuronales, las más conocidas como Backpropagation.

**1977- Stephen Grossberg.** Planteó la Teoría de Resonancia Adaptada. Es una arquitectura de red que se diferencia de todas las planteadas hasta la fecha y es capaz de simular la memoria a largo y corto plazo.

**1985- John Hopfield.** Escribe su libro "Computación neuronal de decisiones en problemas de optimización". Su publicación supuso el renacimiento de las redes neuronales, donde describe un sistema basado en sistemas olfativos, obteniendo un modelo construido por elementos interconectados que buscan un estado interno de mínima energía.

**1986- David Rumelhart/G. Hinton.** Profundizaron en la investigación de las redes neuronales Backpropagation, debido al avance de aquella época, nuevas empresas, nuevos productos y grandes avances en el software y hardware.

En la actualidad, se han desarrollado diferentes estudios basados en las RNAs para la aplicación de estudios financieros. A continuación, se describen algunos estudios.

**1994- Komo, Chang y KO.** Se han centrado en la aplicación de modelos estadísticos no paramétricos y no lineales debido a la aplicación tan importante que

tiene en el área de las finanzas. Las redes neuronales artificiales, según su estudio, poseen la propiedad de capturar las características no lineales de los índices de bolsa y se ha demostrado su eficacia de entrenamiento con cantidades de información con el objetivo de identificar dichas relaciones no lineales entre los valores de entrada y salida. Se ha notado un aumento en la aplicación de las redes neuronales para la predicción de series financieras. Según se ha demostrado es un método totalmente eficaz y con muchas aplicaciones (García *et al* Jalal *et al* Garzón *et al* López, 2013).

**1999-Zemke.** Propuso usar mercados con menor volumen de operaciones al ser más fáciles de predecir con técnicas Machine Learning (ML) para analizarla y usando como referente la Bolsa de Valores de Varsovia. Para ello la primera tarea es la predicción del valor del índice binario, o lo que es lo mismo, pronosticar la posición del valor del índice WIG si ha subido o bajado. Las cuatro técnicas ML utilizadas son: Redes neuronales; clasificador bayesiano; K-nearest neighbor; K-nearest neighbor prediction scrutinized. La conclusión fue que mediante la técnica de K-nearest neighbor se consiguió una efectividad del 64% y con redes neuronales se consiguieron mejores predicciones ya que la bolsa de valores se rige por la no linealidad de sus datos. (García *et al* Jalal *et al* Garzón *et al* López, 2013).

**2000- Kim y Han.** Teorizan que existen multitud de estudios en los que usan modelos de redes neuronales artificiales (ANN), no obstante, estos modelos quedan limitados en el aprendizaje de patrones por la complejidad y gran ruido que posee la bolsa de valores. Por medio de un algoritmo genético podemos optimizar la característica relevante de un subconjunto en una red neuronal. El algoritmo genético (GA) consigue simplificar el proceso de aprendizaje para las ANN y gracias a ello conseguimos reducir el ruido de la red neuronal (García *et al* Jalal *et al* Garzón *et al* López, 2013).

**2003- Chen, Leung y Daouk.** Mediante un modelo de red neuronal artificial, denominado red neuronal probabilística (PNN), modelan y predicen el cambio en el índice de la Bolsa de Valores de Taiwán. Los resultados demuestran que con la estrategia de inversiones basadas en dicha red neuronal, utilizando los datos históricos del propio índice, se obtienen mayores rendimientos en relación al resto de estrategias planteadas durante esos años. La ventaja de este modelo es la capacidad de identificar valores atípicos y datos erróneos (García *et al* Jalal *et al* Garzón *et al* López, 2013).

**2003- Jorge de Andrés Sánchez.** Publica el artículo “Dos aplicaciones empíricas de las redes -neuronales artificiales a la clasificación y la predicción financiera en el mercado español”, centrado como su título indica, en los problemas de clasificación y de predicción financiera de las Redes Neuronales Artificiales. Jorge de Andrés Sánchez defiende que las RNAs son superiores a los métodos clásicos de la época, principalmente en la predicción de la quiebra. Dicha predicción, junto a la estimación de los rendimientos de carteras de valores, las hacen ventajosas frente al resto de métodos multivariantes.

**2006 - Alexander Molina Cabrera/ Alejandro Garcés Ruiz/ Eliana Mirdedy Toro O Campo.** Autores del artículo “Pronóstico de bolsa de valores empleando técnicas inteligentes”, estudiaron la importancia de crear un análisis de datos disponibles en primer lugar, antes de utilizar las herramientas de predicción, y así poder calibrar estos datos para ganar tiempo o quitar datos redundantes que provocan retraso en la respuesta. Así pues, se consigue una mejor respuesta con la información disponible, dándonos como error, en vez de cero, un número pequeño que se pudiera memorizar y permitiese no generalizar.

**2008- Chaigusin, Chirathamjaree y Clayden.** Aplican el estudio, mediante redes neuronales del tipo Backpropagation, para la predicción del índice de la Bolsa de Tailandia. Utilizan como datos de entrada: el índice SET, Dow Jones Nikkei, Hang Seng, el precio del oro, la tasa mínima de préstamo y los tipos de cambio del Thai Baht y del dólar. Los resultados de la investigación indican que los movimientos del índice de la Bolsa de Tailandia son sensibles a entradas anteriores (Younghui *et al* Dongmei *et al* Weihui, 2013).

**2008- Zhu, Xu, Wang y Li.** Realizaron una investigación que indica que los modelos de redes neuronales con volúmenes de negociación conducen a mejoras, en función de la previsión del desempeño en diferentes horizontes temporales. Los resultados empíricos indican que los volúmenes de negociación mejoran el rendimiento del índice bursátil a medio y largo plazo (García *et al* Jalal *et al* Garzón *et al* López, 2013).

**2009 - Luis Ayala Jiménez, Sebastián, Letelier González, Pablo Zagal Morgado.** Presentaron el Modelo de Redes Neuronales para la Predicción de la Variación del Valor de la Acción de First Solar. Mediante esta investigación quieren determinar cómo

las redes neuronales pueden ser usadas en las finanzas aplicadas, es decir, el apoyo de este tipo de modelos a la predicción en bases a identificación de patrones en los precios de los instrumentos financieros. Además, quieren demostrar como con los modelos de Inteligencia Artificial es posible entregar parámetros de predicción que la econometría clásica no es capaz dado que esta se sustenta en un supuesto muy fuerte y lejano a la realidad financiera, el cual dice que los errores a priori tienen una distribución normal.

Los resultados obtenidos comprobaron lo anteriormente comentado, esto confirma que es posible obtener rendimientos superiores al mercado, mediante herramientas que utilicen técnicas de redes neuronales. Realizaron un estudio de predicción de forma diaria, que era algo nuevo en este campo. Además, demostraron que no se cumple la hipótesis de Random Walk (camino aleatorio), es decir, existe evidencia significativa que para las series financieras hay capacidad predictiva. El hecho de que los precios no sigan un camino aleatorio da espacio para realizar predicciones.

**2009 -Eduardo Arturo Cruz, Jorge Hernan Restrepo, Pedro Medina Varela.** Publican “Pronóstico del índice general de la bolsa de valores de Colombia usando redes neuronales” explicando que una de las formas de medir el comportamiento de la economía de un país, es a través de las bolsas de valores locales. De esta forma en Colombia el comportamiento económico de la Bolsa de valores se mide a través de tres indicadores principalmente: el Colcap, el Col20 y el IGBC, este último es el más antiguo y más utilizado por los agentes que intervienen en el mercado bursátil por ello, escogieron esos tres inputs para su RNA con el objetivo de predecir el comportamiento del mercado bursátil colombiano.

Nos demuestran de nuevo que las RNA son mejores que los modelos tradicionales de tendencia o crecimiento e incluso en muchos casos mejor que los modelos Arima y Garch.

**2011- Guresen, Kayakutlu y Daim.** Observaron que en la mayoría de los casos los modelos de redes neuronales permiten obtener mejores resultados que otros métodos. Los datos utilizados corresponden a los valores diarios del índice Nasdaq desde el 7 de octubre de 2008 hasta el 26 de junio de 2009. Los modelos que se compararon en este trabajo son multi-layer perceptron (MLP), dynamic artificial neural network (DAN2) y un modelo híbrido de redes neuronales y Garch. El rendimiento de los modelos se evaluó a través del error cuadrado medio y la desviación absoluta media y como resultado se obtuvo que el clásico modelo MLP supera a los demás modelos (García *et al* Jalal *et al* Garzón *et al* López, 2013).

**2011- Guo, Shen, Wub y Wu.** Para predecir los índices bursátiles de la bolsa de Shanghai escogieron una red neuronal del tipo Radial Function. Para aumentar la eficiencia y optimización de la red usaron el Artificial Fish Swarm Algorithm (AFSA). El resultado de estos dos algoritmos con el algoritmo genético y Particle Swarm Optimization (PSO) fueron tan buenos como los propuestos por Arima, Backpropagation (BP) o Support Vector Machine (SVM), no con la misma precisión, pero al ser novedoso, fue exitoso al incrementar parte del resultado en las predicciones de la red original Radial Basis Function (RBF) (García *et al* Jalal *et al* Garzón *et al* López, 2013).

**2011- Wang J.-Z., Wang, Zhang y Guo.** Exponen que debido a la variedad de factores que afectan a la Bolsa de Valores, se propone la utilización del algoritmo Wavelet De-noising-based Backpropagation (WDBP); los datos utilizados en el estudio corresponden a los del índice de la Bolsa de Shanghai desde enero de 1993 a diciembre de 2009. El modelo WDBP fue comparado con un modelo de red neuronal Backpropagation, mejorando el primero la precisión de las predicciones (García *et al* Jalal *et al* Garzón *et al* López, 2013).

A continuación, se comentan las características principales de las RNA.

- **Aprendizaje Adaptativo:** Capacidad de aprender a realizar tareas basadas en un entrenamiento. No se le indican las reglas para dar una solución, sino que extrae sus propias reglas a partir de los ejemplos de aprendizaje, modifican su comportamiento en función de la experiencia.
- **Autoorganización:** Una red puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje.
- **Tolerancia a Fallos:** La destrucción parcial de una red conduce a una degradación de su estructura; sin embargo, algunas capacidades de la red se pueden retener.
- **Operación en tiempo real:** Los computadores neuronales pueden ser realizados en paralelo, y se diseñan y fabrican máquinas con hardware especial para obtener tal capacidad.
- **Fácil construcción en CI:** Se puede obtener chips especializados para redes neuronales que mejoren su capacidad en ciertas tareas. Ello facilitará la integración modular de los sistemas existentes.

## 2.2.: CLASIFICACIÓN DE REDES NEURONALES

En este apartado, se clasifican las redes neuronales dependiendo de su Topología o su Algoritmo de aprendizaje.

- Según su topología se clasifican:
  - **Monocapa**; son redes con una sola capa que para unirse las neuronas crean conexiones laterales para conectar con otras neuronas de su capa. Las redes más representativas son la red de Hopfield, la red BRAIN-STATE-IN-A-BOX o memoria asociativa y las máquinas estocásticas de Botzmann y Cauchy. Son frecuentemente utilizadas en circuitos eléctricos debido a que son adecuadas para ser implementadas mediante hardware, usando matrices de diodos que representan las conexiones de las neuronas.
  - **Multicapa**; están formadas por varias capas de neuronas. A su vez se clasifican según la manera en la que se conexionan sus capas.
    - Redes con conexiones hacia adelante: Las capas están ordenadas por el orden en que reciben la señal desde la entrada hasta la salida y están unidas en ese orden, son las llamadas feedforward. Como ejemplos de estas redes destacan: Perceptron, Adaline, Madaline, Backpropagation y los modelos LQV y TMP de Kohonen.
    - Redes con conexiones hacia atrás: Aparte del orden normal de las capas, algunas de ellas también están unidas desde la salida hasta la entrada en el orden inverso en que viajan las señales de información, son las llamadas feedback. Como ejemplos de estas redes destacan: las redes ART, Bidirectional Associative Memory (BAM) y Cognitron.
- Según su algoritmo de aprendizaje se clasifican:
  - **Aprendizaje supervisado en Redes neuronales**; requiere de un agente externo o supervisor que determina, mediante un entrenamiento, las respuestas que debe dar la RNA a partir de la entrada de datos.
  - **Aprendizaje no supervisado en Redes neuronales**; se basa en que no es necesario un supervisor para ajustar los pesos de las conexiones de las neuronas.

Ambos tipos de aprendizaje se desarrollan con mayor profundidad en el capítulo 2.4: PROCESO DE APRENDIZAJE DE RNA.

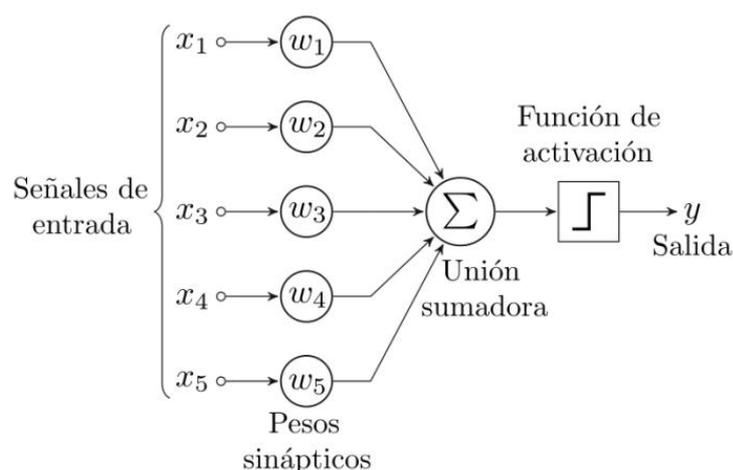
Se va a realizar un paseo histórico viendo como fueron evolucionando las redes neuronales desde sus orígenes hasta la actualidad para entender su proceso.

Durante el año 1943 se desarrolló la primera red neuronal por Warren McCulloch y Walter Pitts que consistía en el conjunto de señales de entrada, multiplicadas por unos valores de pesos aleatorios. La entrada era comparada con un patrón preestablecido para determinar la salida de la red. Durante el desarrollo de los sistemas de inteligencia artificial, se encontró gran similitud entre su comportamiento y el de los sistemas biológicos.

Posteriormente, el psicólogo Frank Rosenblatt en el año 1957, creía que la conectividad en las redes biológicas era aleatoria, y se oponía a la teoría de McCulloch Pitts. Rosenblatt opinaba que la herramienta de análisis más apropiada era la teoría de probabilidades. Así pues, se desarrolló el primer modelo de Perceptrón, desarrollado en un ambiente biológico imitando el funcionamiento del ojo humano, ya que respondía a señales ópticas. El Perceptrón era inicialmente un dispositivo de aprendizaje, en su configuración inicial no estaba la capacidad de distinguir patrones de entrada complejos, pero si era capaz mediante un proceso de aprendizaje adquirir dicha capacidad y responder no solo a patrones concretos sino también patrones aleatorios, aunque se demostró que su precisión disminuía conforme aumentaba el número de patrones a aprender.

En 1969 Marvin Minsky y Seymour Papert publicaron su libro: "Perceptrons: An introduction to Computational Geometry", centrado en las limitaciones del Perceptrón y muchos pensaban que sería el final de las redes neuronales. Por muy básico que fuese el Perceptrón, ha servido como base a partir de los cuáles se han desarrollado otros modelos de la red neuronal, como la red Adaline y las redes multicapa (MINSKY, 1987).

La estructura de red del Perceptrón es la siguiente:



**Ilustración 1: Estructura de la red Perceptrón**

El Perceptrón únicamente presenta una neurona de salida haciendo la suma ponderada de las entradas, resta el umbral y el resultado pasa a una función de transferencia de tipo escalón. Al constar de una sola capa de entrada y otra de salida, tiene una capacidad limitada de representación, sólo es factible aplicarlo a patrones sencillos. También se desarrolló el Perceptrón multicapa, es una red con alimentación hacia adelante, compuesta de varias capas de neuronas entre la entrada y la salida de la misma y permite establecer regiones de decisión más complejas.

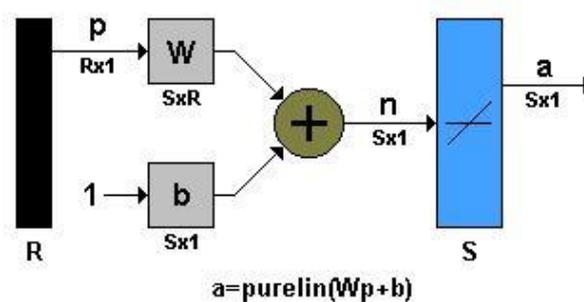
A la vez que se desarrollaba el Perceptrón, Bernard Widrow y su estudiante Marcian Hoff introdujeron el modelo de la red Adaline y su regla de aprendizaje llamada algoritmo LMS (Least Mean Square). La red Adaline presenta mejoras en la función de transferencia, siendo ésta de tipo lineal. Además, el algoritmo LMS es mucho más potente que la regla de aprendizaje del Perceptrón.

El elemento de procesamiento en la red Adaline realiza la suma de los productos de los vectores de entrada y de pesos, y aplica una función de salida para obtener un único valor de salida, si el sumatorio da un resultado de salida +1 es positiva o -1 es negativa. La función de salida es igual a la función de activación, por eso es tan importante el resultado de la misma. En términos generales la salida de la red está dada por:

$$a = W^T p$$

Existe un procedimiento para modificar los pesos con el objetivo de hacer posible que proporcione el valor de salida correcto para la entrada dada.

La estructura de la red tipo Adaline es la siguiente:



**Ilustración 2 Estructura de la red Adaline**

La red Adaline ha sido aplicada en diferentes campos principalmente en el procesamiento de señales. Comenzaron los estudios con las comunicaciones electrónicas diseñando filtros analógicos empleando circuitos RLC (Resistencias,

Inductores, Condensadores) para eliminar el ruido en las señales utilizadas en las comunicaciones. La aplicación de estas redes sigue siendo muy importante en la industria para eliminar el ruido en señales portadoras de información, es decir en la radio. La información que se transmite se puede codificar en forma de una señal analógica que reproduce exactamente las frecuencias y las amplitudes del sonido original. Además, aunque las características de la señal cambiasen, la red Adaline puede adaptarse automáticamente porque toma la entrada y la salida deseada, y se ajusta a sí misma para ser capaz de llevar a cabo la transformación deseada. La herramienta matemática más utilizada para el diseño de filtros es la Serie de Fourier, que describe la naturaleza de las señales periódicas en el dominio frecuencial.

Ante el desarrollo del Perceptrón y la red Adaline, diseñados para entrenar redes de una sola capa, hizo que se planteara el avance de redes multicapa para superar la desventaja que presentaban los anteriores modelos. Los primeros resultados se desarrollaron en 1974 por Paul Werbos pero no fue hasta mediados de los años 80 cuando las redes multicapa fueron aceptadas con el descubrimiento del algoritmo Backpropagation o algoritmo de propagación redescubierto por varios investigadores como David Rumelhart, Geoffrey Hinton y Ornal Williams, David Parker y Yann Le Cun. Precisamente la red Backpropagation es la herramienta principal en la que se basa este proyecto y que va a demostrar más utilidades de las que actualmente ya existen y que posteriormente desarrollaremos en profundidad.

Una de las características principales de las redes Backpropagation y uno de los avances fue la reducción del tiempo requerido por un procesador secuencial para determinar la correspondencia entre unos patrones dados. Estas redes son un tipo de red de aprendizaje de tipo supervisado y básicamente funciona aplicando un patrón de entrada a la red, este se propaga desde la primera capa a través de las capas superiores de la red hasta generar una salida. Se compara la señal resultado de la salida con la salida deseada y se calcula el margen de error de cada una de las salidas. A diferencia del resto de redes, las Backpropagation propagan hacia atrás las salidas de error hacia todas las neuronas de la capa oculta que contribuyen a la salida. Sin embargo, la neurona de la capa oculta solo recibe una fracción de la señal total del error, basándose en la contribución relativa que haya aportado cada neurona a la salida original. Así de forma reiterativa hasta que las neuronas de la red hayan recibido una señal de error que describa su contribución relativa al error total. Posteriormente, se actualizan los pesos de conexión de cada neurona para obtener un resultado y clasificar los patrones de entrenamiento correctamente.

En resumen, las redes Backpropagation y su importancia se encuentran en que durante el proceso, a medida que se entrena la red, las neuronas de las capas intermedias se organizan aprendiendo a reconocer distintas características. Una vez tengan un patrón arbitrario de entrada que contenga ruido, las neuronas de la capa

oculta de la red responden con una salida activa si la nueva entrada tiene un patrón más semejante a la característica que las neuronas individuales hayan aprendido a reconocer durante su entrenamiento. Más adelante se explica con más detenimiento este tipo de redes.

Por último, las redes se clasifican también según el tipo de aprendizaje que presenten, se distinguen entre aprendizaje supervisado y no supervisado. El aprendizaje supervisado requiere de un agente externo que puede controlar de tres formas: por corrección de error, por refuerzo y aprendizaje de tipo estocástico. Por otro lado, el aprendizaje no supervisado basando en la no necesidad de tener un supervisor para ajustar los pesos de las conexiones de las neuronas distingue dos formas principales: hebbiano y competitivo y cooperativo. Más adelante, en este proyecto, son descritas ya que tienen un papel fundamental en el desarrollo de las redes neuronales.

### 2.3.: CONSTRUCCIÓN DE REDES NEURONALES

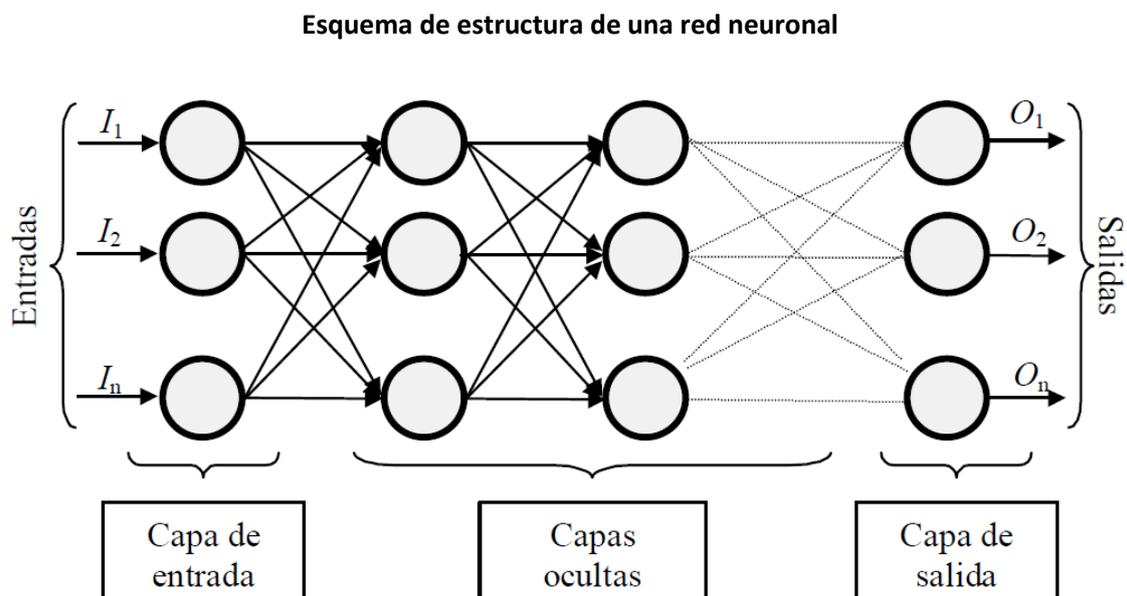
Si partimos de la definición de las redes neuronales, sabemos que las RNA intentan simular el comportamiento de las neuronas; por tanto, serán éstas las principales unidades de proceso, de forma que como mínimo tendrán un canal de entrada de la información y uno de salida. La red neuronal tendrá una o más neuronas distribuidas a lo largo de una estructura definida para resolver el problema planteado.

Según la función que desempeñan en la estructura de la red existen tres tipos de neuronas:

- Las que reciben los estímulos externos y las que captan la información de entrada a la red.
- Las que se encargan del procesamiento de la información y aprendizaje de la misma, llamadas neuronas ocultas.
- Las que, una vez realizado el proceso de aprendizaje, muestran la información resultante dando respuesta al problema planteado.

Destacamos los tres tipos principales: el modelo del Perceptrón, el modelo de la red Adaline y las redes Backpropagation, desarrollados anteriormente.

Definimos capa o nivel a un conjunto de neuronas con la misma fuente de entrada de datos cuyas salidas se encaminan a una misma dirección.

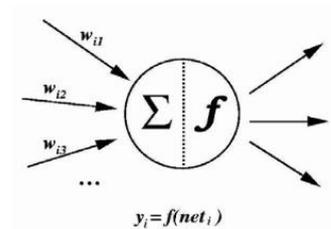


**Ilustración 3 Esquema de estructura de una red neuronal**

Fuente: Investigación del Grupo de Investigación Aplicada a la Ingeniería Química. "Redes Neuronales: Conceptos Básicos y Aplicaciones". Autor: Damián Jorge Match.

Se recibe una serie de inputs desde otras unidades o de una fuente externa de datos, donde cada input tiene un peso asociado ( $w$ ), que se va modificando en el proceso de aprendizaje. Cada unidad (elemento básico de computación) aplica una función dada  $f$  de la suma de los inputs ponderadas mediante los pesos.

$$y_i = \sum_j w_{ij} y_j$$

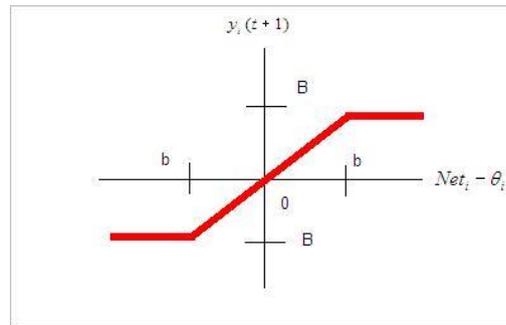


Así pues, a partir de los inputs, calculamos los pesos que van a tener cada una de las combinaciones de los mismos en cuanto a las interconexiones que forman las diferentes neuronas que son además bidireccionales, formando una única neurona que posteriormente nos dará una información de salida (output), obteniendo así un resultado. Esa neurona generada a partir de la ponderación de los pesos, tendrá en cuenta lo siguiente: por una parte, la información de las neuronas anteriores; por otra parte, la Función de Activación, que estima el nuevo estado de activación de una neurona en un momento posterior teniendo en cuenta el estado de activación anterior y la entrada total calculada ( $y_i$ ); y por último, la Función de Actividad o Transferencia (Del Carpio, 2005).

Se dice que una neurona se activará o no según unos parámetros establecidos. Si se activa, entra en funcionamiento la función de transferencia (se activará según la información y si supera un cierto número llamado umbral). Cuando una neurona se encuentra en un estado activo se representa por 1, mientras que si se encuentra en un estado pasivo se representa con un 0.

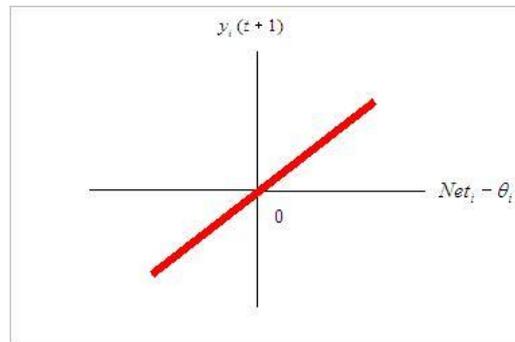
La función de Actividad o Transferencia, sirve para combinar las entradas de información de la neurona con su estado de activación actual. Encontramos diferentes tipos de función de activación de las neuronas:

- **Función de activación escalón:** Se utiliza cuando las neuronas tienen un estado de activación binario (0, 1). La neurona se activará cuando el valor neto de entrada ( $y_i$ ) sea superior o igual al umbral de la neurona y no se activará cuando no lo supere.



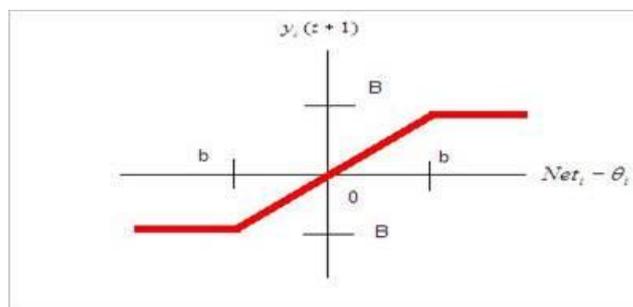
**Ilustración 4 Función de Activación escalón**

- **Función de activación lineal o identidad:** Puede obtener cualquier número real y el estado de activación coincide con el valor de la entrada neta ( $y_i$ ).



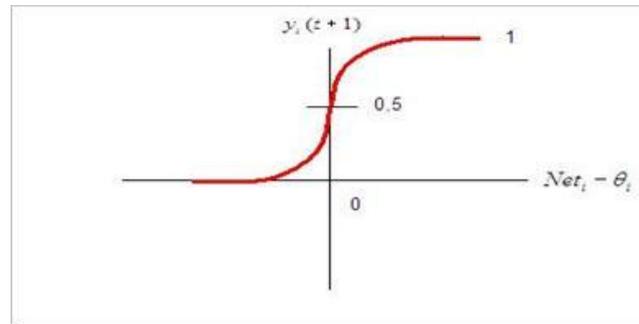
**Ilustración 5 Función de Activación Lineal o Identidad**

- **Función de activación lineal-mixta:** Se trata de una variante de la función de activación lineal. En este caso los valores de estado que puede obtener la neurona se encuentran fijados dentro de un intervalo de números reales.



**Ilustración 6 Función de Activación lineal-mixta**

- **Función de activación sigmoideal**: Se trata de una función de activación continua dentro del intervalo del eje real, con un valor de salida continuo entre 0 y 1.



### Ilustración 7 Función de Activación Sigmoideal

Por otra parte, el último componente que una neurona necesita es la función de salida y su objetivo supone la determinación del valor de salida que transfiere a otras neuronas que se encuentran vinculadas con ella. Hay que tener en cuenta si la neurona se encuentra activada o no y que la información que pasa a la función serán datos continuos, binarios, etc. dependiendo del valor de la función de transferencia. Las funciones de salida más comunes son:

- **Ninguna**: la salida es la misma que la entrada. Llamada también función identidad.
- **Binaria**: 
$$\begin{cases} 1 & \text{si } act_i \geq \xi_i \\ 0 & \text{de lo contrario} \end{cases}, \text{ donde } \xi_i \text{ es el umbral.}$$

## 2.4.: PROCESO DE APRENDIZAJE DE RNA

El proceso de aprendizaje dependerá directamente de los estados de activación de cada neurona. Durante el aprendizaje se van ajustando de forma adecuada e internamente todos y cada uno de los pesos asociados a cada rama para obtener la salida deseada, así podrá responder por sí sola a situaciones diferentes a las aprendidas. Las neuronas crearán una nueva conexión cuando el valor de su peso sea mayor que cero y se destruirá la misma cuando el peso tome valor cero.

El proceso de aprendizaje finalizará cuando los pesos de cada conexión de neuronas sean constantes y no tengan variación. Se destacan así dos tipos principales de aprendizaje, el supervisado y el no supervisado que se desarrollan a continuación.

### A) Aprendizaje Supervisado:

Requiere de un agente externo o supervisor que determina, mediante un entrenamiento, las respuestas que debe dar la RNA a partir de la entrada de datos. Su aplicación destaca en la bioinformática, en el manejo de base de datos de marketing, previsión de estados financieros fraudulentos, recuperación de información y reconocimiento de voz, patrones y spam, entre otros. Este entrenamiento controlado puede darse de tres formas diferentes:

- Aprendizaje por corrección de error: Consiste en que el supervisor va ajustando los pesos de las conexiones de las neuronas en función de las diferencias que se van produciendo entre los valores que se desean obtener y los que realmente ha obtenido la RNA. Uno de los algoritmos más utilizados es el de la regla delta generalizada o error Backpropagation, que es el empleado para el estudio de este proyecto.
- Aprendizaje por refuerzo: Se basa en que el supervisor indica a la RNA una señal de refuerzo si la salida obtenida es satisfactoria (+1) o en caso contrario (-1) penalizando dicha salida. En función de estas señales los pesos de las neuronas irán variando. Es un aprendizaje mucho más lento, ya que se basa en que no se dispone de la información completa referente al comportamiento deseado de la red ante una entrada de datos.
- Aprendizaje estocástico: Se basa en provocar cambios aleatorios en el valor de los pesos de las conexiones de las neuronas para, posteriormente, realizar una evaluación de los efectos producidos a partir del objetivo deseado y de una distribución de probabilidad determinada. Una vez realizado un cambio en los pesos, se determina si el comportamiento de la red es próximo al deseado.

## **B) Aprendizaje No Supervisado:**

Se basa en que no es necesario un supervisor para ajustar los pesos de las conexiones de las neuronas. Es un aprendizaje autosupervisado y la red no tiene información referente a si las salidas generadas son o no correctas. Dado que no se le indica a la red qué salida es correcta, en función del tipo de estructura y del algoritmo de aprendizaje utilizado variará la interpretación de las salidas que proporcione. Las aplicaciones más típicas de este tipo de aprendizaje son: aplicación de reconocimiento facial, compresión de datos, creación de mapas topográficos y desarrollo de aplicaciones que simplifiquen el aprendizaje a otras redes.

Se pueden distinguir dos tipos principales de aprendizaje: no supervisado o autosupervisado.

- **Aprendizaje Hebbiano:** Aprendizaje propuesto por Hebb en 1949 que decía que: *“Cuando un axón de una celda A está suficientemente cerca como para conseguir excitar una celda B y repetida o persistentemente toma parte de su activación, algún proceso de crecimiento o cambio metabólico tiene lugar en una o ambas celdas, de tal forma que la eficiencia de A, cuando la celda A activa B, aumenta”*. Hebb definía las celdas como un conjunto de neuronas muy interconectadas con una estructura muy compleja y la eficiencia se identifica con la intensidad de las conexiones de las neuronas, es decir, el peso (Hebb, 1949).  
Este aprendizaje consiste en el ajuste de pesos de las conexiones de las neuronas conforme a la correlación existente entre los valores de activación (salidas) de las dos neuronas conectadas.
- **Aprendizaje competitivo y cooperativo:** Las neuronas, en este tipo de aprendizaje, compiten y cooperan entre ellas para conseguir un objetivo dado. Ante la información de entrada en la red, busca que las neuronas de salida de la misma obtengan su máximo valor de respuesta posible. (Gómez *et al* Fernández *et al* López *et al* Díaz-Marta, 2010)

## 2.5.: REDES NEURONALES BACKPROPAGATION

El funcionamiento de la red Backpropagation consiste en el aprendizaje de un conjunto predefinido de pares de entradas-salidas. Lo interesante de este proceso es que consiste en que a medida que la red es entrenada las neuronas de capas intermedias van organizándose a sí mismas de tal modo que las neuronas van aprendiendo las distintas características del espacio total de entrada.

Una vez entrenadas, cuando se les presente un patrón arbitrario de entrada que tenga ruido o esté incompleto, las neuronas de la capa oculta de la red responderán con una salida activa si la nueva entrada contiene un patrón que se asemeje a aquella característica que las neuronas individuales que habrían aprendido a reconocer en el entrenamiento.

Las redes neuronales Backpropagation no son solo usadas para índices bursátiles, sino que tienen infinitas aplicaciones, algunos estudios y aplicaciones se muestran a continuación:

La predicción de incendios forestales como el usado en Whitecourt Forest, Canada, realizado por C. Vega Garcia, P.M Woodward S.J. Titus y W.L. Adamowicz, donde se utilizaron las RNAs para predecir el número de incendios intencionados, cuyo estudio se realizó en el área Whitecourt, para ello hicieron dos modelos uno de regresión logarítmica versus su homólogo en RNAs, demostrando que era mejor la red neuronal, pero nos muestra la importancia de la selección de Inputs.

Otro ejemplo sería lo que nos plantea Daniel Humberto Cortes sobre el uso de Redes Neuronales para analizar Modelos Mecánicos de Tejidos Biológicos, donde explica que dada la complejidad de las soluciones analíticas y computacionales de algunos modelos mecánicos del tejidos biológicos, se utilizaron redes neuronales para identificar la fuerza de indentación a partir de las propiedades elásticas y geométricas de una capa de cartílago articular modelada como un material transversalmente isótropo y para hallar el esfuerzo nominal en los ligamentos modelados como un material viscoelástico sometido a deformaciones finitas. Las soluciones analíticas de los dos problemas se utilizaron para crear el conjunto de datos necesarios para el entrenamiento de redes neuronales tipo base radial, Backpropagation y Anfis. La verificación de su funcionamiento mostró que algunas de las redes aproximan la solución analítica con menos de un 2% de error (HUMBERTO, 2002).

Las redes Backpropagation de dos capas y un número intermedio de neuronas en la capa de entrada (14-20) mostraron el mejor comportamiento. Con redes neuronales debidamente entrenadas es posible obtener resultados de problemas complejos en unas pocas décimas de segundo, lo cual las convierten en una herramienta numérica que se compara favorablemente con otras.

Como conclusión a su estudio nos comenta lo eficiente y rápida para los análisis biomecánicos pueden ser las redes neuronales, es más una vez entrenada, podría ser usada por muchos investigadores, sin necesidad de resolver problemas complejos o métodos numéricos sofisticados, tiene a su vez otra ventaja y es su poco espacio en memoria y tiempos de ejecución muy cortos, una ventajas más que añadir sería que no requieren de un modelo teórico ya que pueden ser entrenadas con resultados experimentales. También nos comenta que, de las redes usadas, la Backpropagation fue la más indicada, debido al tiempo de entrenamiento, y porqué el error obtenido fue el menor de las redes (Nazemi *et al* Abbasi *et al* Omid, 2015).

Una desventaja de las redes Backpropagation es que, durante el entrenamiento, no siempre se obtienen los mismos pesos con la misma arquitectura y lote de entrenamiento, ya que los pesos de la red son inicializados de forma aleatoria. Por el contrario, con las redes tipo anfis y base radial siempre se producen los mismos resultados durante el entrenamiento.

Otro estudio que se hizo con el uso de redes neuronales Backpropagation fue el de Haniyas, Curtis, Thalassinos (2007) que aplicaron la red Backpropagation para predecir el mercado de valores de Grecia. Prediction with Neural Networks: The Athens Stock Exchange Price Indicator. Según este estudio la característica que tienen todos los indicadores de valores en común es la incerteza, la cual está relacionada con su corto o largo estado futuro, esta propiedad no es deseable para un inversor, pero es inevitable, lo único que se puede hacer es intentar reducir su incerteza. En este estudio nos enseñan como la red Backpropagation puede predecir de una forma adecuada el indicador de bolsa de valores de Atenas, usando inputs de 3 días, predijeron 9 días, con un error mínimo de  $MSE=0.0024$  (Freitas *et al* De souza *et al* De Almeida, 2009).

Detallamos a continuación brevemente todo el proceso de aprendizaje y entrenamiento de la red. Las seis fases en las que se divide el proceso comienzan con la asignación de pesos a las neuronas y termina en la última fase, cuando la red se detiene. Sin olvidar los pasos que caracterizan a la red Backpropagation durante el proceso que irá reiterando de forma recurrente la modificación de los pesos mientras la red va aprendiendo.

### **Fase 1**

Inicialmente, se dan valores a los pesos de las neuronas. Generalmente serán valores pequeños y obtenidos de forma aleatoria.

## Fase 2

Se presenta a la red el vector o patrón de entrada  $X_p : X_{p1}, X_{p2}, \dots, X_{pN}$ , y se especifican los datos de salida deseados  $d_1, d_2, \dots, d_M$  (estamos ante una red con aprendizaje supervisado). En caso de que la red Backpropagation se utilice para obtener clasificaciones de los datos, las salidas deseadas serán cero excepto la de la clase a la que pertenezca el patrón de entrada.

## Fase 3

Se procede a calcular la salida de la red (primera fase del algoritmo hacia adelante) presentándose todas las entradas y calculando cada salida de cada capa hasta llegar a la salida de la red  $y_1, y_2, \dots, y_M$ .

Para lograr este objetivo, primero se calculan las entradas netas para las neuronas de las capas ocultas.

$$net_{pj}^h = \sum_{i=1}^N w_{ij}^h x_{pi} + \theta_j^h$$

donde el índice  $h$  representa la magnitud de la capa oculta (hidden): el subíndice  $p$  representa al  $p$ -ésimo vector de entrenamiento: la  $j$  es la  $j$ -ésima neurona oculta. El resto de parámetros ya son conocidos.

A continuación, se han de calcular las salidas de las neuronas ocultas en función de las entradas netas correspondientes.

$$y_{pj} = f_j^h (net_{pj}^h)$$

Por último, se realizan los mismos cálculos (entrada y salida), pero para las neuronas que se encuentran en la capa de salida (capa  $o$ , output).

$$net_{pk}^o = \sum_{j=1}^L w_{kj}^o y_{pj} + \theta_k^o$$

$$y_{pk} = f_k^o (net_{pk}^o)$$

#### Fase 4

Una vez finalizado el proceso hacia adelante, comienza el proceso hacia atrás. Para ello se han de calcular los errores cometidos por todas las neuronas de la red. En el caso de ser una neurona de la capa de salida, el valor de la delta será:

$$\delta_{pk}^o = (d_{pk} - y_{pk}) f_k^o(\text{net}_{pk}^o)$$

Como ya se ha indicado en otras ocasiones la función  $f$  debe ser derivable, por lo que la función de activación no puede ser de tipo escalón. Por tanto, se pueden utilizar bien la función lineal que como vimos es  $f_k(\text{net}_{jk}) = \text{net}_{jk}$  o bien la función de activación sigmoideal  $f_k(\text{net}_{jk}) = \frac{1}{1+e^{-\text{net}_{jk}}}$ .

La elección de una función u otra depende de los datos de salida de la red que se deseen. Si queremos que las neuronas de salida sean binarias se utilizará la función sigmoideal, ya que esta función es derivable y estable. En cualquier otro caso se pueden utilizar cualquiera de las dos indistintamente.

Si la neurona sobre la que se quiere calcular el error, neurona  $j$ , no es una neurona de salida, es decir, es una neurona de capa intermedia, entonces el valor de la delta que se obtiene será:

$$\delta_{pj}^h = f_j^h(\text{net}_{pj}^h) \sum_k \delta_{pk}^o w_{kj}^o$$

donde el error de las neuronas en las capas ocultas depende de todos los errores de las neuronas de la capa de salida (proceso de propagación hacia atrás).

#### Fase 5

Una vez determinados los errores de todas las neuronas, se procede al ajuste de los pesos de cada conexión de cada neurona, empezando desde la capa de salida hacia atrás. Los pesos se ajustarán para reducir la delta generalizada, es decir, la disminución del error.

En el caso de los pesos de las neuronas de la capa de salida se ajustarán así:

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \Delta w_{kj}^o(t+1)$$

$$\Delta w_{kj}^o(t+1) = \alpha \delta_{pk}^o y_{pj}$$

Para los pesos de las neuronas de las capas intermedias (ocultas) se ajustarán:

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \Delta w_{ji}^h(t+1)$$

$$\Delta w_{ji}^h(t+1) = \alpha \delta_{pj}^h x_{pi}$$

Para acelerar el proceso de aprendizaje se puede incorporar el término *momento*. Para las neuronas de salida el momento será igual a

$$\beta(w_{kj}^o(t) - w_{kj}^o(t-1))$$

Para el caso de neuronas de capas intermedias u ocultas:

$$\beta(w_{ji}^h(t) - w_{ji}^h(t-1))$$

## Fase 6

Todo este proceso se repite hasta que el término de error se reduce a la mínima para cada patrón aprendido.

$$E_P = \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2$$

Aplicando las fases anteriores, el algoritmo Backpropagation busca minimizar el error cometido por la red comparando con el patrón de datos real. Ahora bien, este mínimo puede ser local o global. Como el algoritmo presenta un gradiente descendente se asegura que con el cambio de los pesos se reduzca el error hasta encontrar el mínimo más cercano, por lo que en algún caso el proceso podría detenerse en algún mínimo local y no global. Así pues, según el problema planteado no tiene por qué ser necesaria la reducción al mínimo del error global, podría bastar y ser satisfactoria la solución obtenida con un valor de error de mínimo local.

El control del gradiente decreciente viene determinado por el parámetro  $\alpha$  que es la tasa de aprendizaje. En función del número escogido se incrementa o se reduce la velocidad del aprendizaje. A mayor importe, menor tiempo de aprendizaje, pero los saltos que se producen en los pesos son grandes y, por tanto, podría saltarse el mínimo y no poder alcanzarlo. En general suelen aplicarse unos valores entre 0.05 a 0.25.

Cuando la red llega a un mínimo, sea local o global, detiene el aprendizaje. Si se trata de un mínimo del error local y la solución que ofrece la red no es satisfactoria, pueden plantearse algunas soluciones para alcanzar el mínimo global o alcanzar un nuevo mínimo local menor que el anterior. Estas posibles soluciones pueden ser bien añadiendo más capas ocultas, introduciendo cambios en los parámetros de aprendizaje, o comenzar de nuevo el proceso de aprendizaje desde el principio, pero con pesos iniciales diferentes. En algunos paquetes informáticos es posible forzar a la red para que continúe con el aprendizaje, a pesar de alcanzar un mínimo local, y detener el proceso de aprendizaje bien cuando la red lleve un tiempo determinado bien cuando alcance un error de aprendizaje determinado (Oliver, 2013).

### **III. ESTUDIO COMPARATIVO DE LA APLICACIÓN DE LA RED NEURONAL.**

## CAPÍTULO III: ESTUDIO COMPARATIVO DE LA APLICACIÓN DE LA RED NEURONAL

En el presente capítulo se va a analizar dos redes neuronales Backpropagation para la predicción del IBEX 35, una red Backpropagation incluyendo todos los inputs y otra red, también de tipo Backpropagation, en la que se ha hecho una preselección de los inputs. Lo que se pretende demostrar es que a pesar de la capacidad de generalización que posee la red neuronal, la incorporación de los inputs en la red pueden provocar ruido y distorsiones en los resultados por lo que puede ser preferible realizar una preselección de inputs en función de su contribución en la red. La base de datos utilizada parte del 04/01/2006 al 24/04/2015. No obstante, finalmente se han cogido la muestra de las fechas 11/01/2006 al 31/12/2013. Todo dato comprendido en esas fechas servirá para realizar la predicción y poder analizar y comparar con una red en la que se incorporen todos los inputs. Los datos seleccionados son diarios. El análisis se ha realizado con las rentabilidades de los inputs.

El objeto de estudio no trata de construir el mejor modelo de red para la predicción del IBEX sino verificar que la preselección de los inputs mejora la calidad del entrenamiento y de las predicciones, tanto para las diferentes redes utilizadas para la preselección de los inputs como para la red con todos ellos.

Los diferentes índices sobre los que se va a trabajar son:

- Futuro IBEX
- IBEX
- DAX
- Futuro de Petróleo
- STANDARD AND POORS
- FTSE
- DOW JONES Industrial
- DOW JONES Transportes
- EURUSD.

A continuación, se muestra la tabla de correlaciones para verificar como de interrelacionados están y conocer la relación de estos activos.

	<i>FUTURO IBEX</i>	<i>IBEX</i>	<i>DAX</i>	<i>FUTURO DE PETRÓLEO</i>	<i>STANDARD AND POORS</i>	<i>FTSE</i>	<i>DOW JOHN INDUSTRIAL</i>	<i>DOW JOHN TRANSPORTES</i>	<i>EURUSD</i>
<b>FUTURO IBEX</b>	1	0,978677	0,796333	0,275682	0,633343	0,513334	0,610068	0,563553	0,262844
<b>IBEX</b>	0,978677	1	0,818895	0,281469	0,644348	0,528398	0,619488	0,575362	0,266019
<b>DAX</b>	0,796333	0,818895	1	0,285551	0,680241	0,583986	0,658012	0,616245	0,217513
<b>FUTURO DE PETRÓLEO</b>	0,275682	0,281469	0,285551	1	0,371718	0,371433	0,347209	0,278851	0,27049
<b>STANDARD AND POORS</b>	0,633343	0,644348	0,680241	0,371718	1	0,681782	0,974788	0,855199	0,312377
<b>FTSE</b>	0,513334	0,528398	0,583986	0,371433	0,681782	1	0,680544	0,603277	0,309531
<b>DOW JOHN INDUSTRIAL</b>	0,610068	0,619488	0,658012	0,347209	0,974788	0,680544	1	0,832193	0,308667
<b>DOW JOHN TRANSPORTES</b>	0,563553	0,575362	0,616245	0,278851	0,855199	0,603277	0,832193	1	0,260795
<b>EURUSD</b>	0,262844	0,266019	0,217513	0,27049	0,312377	0,309531	0,308667	0,260795	1

**Tabla 1** Tabla de correlaciones de todos los inputs.

Fuente: Elaboración propia.

En la Tabla 1 de correlaciones, se ve claramente la alta correlación que presentan el IBEX y el Futuro del IBEX, al estar relacionados directamente. También destaca la alta correlación que tienen el DAX y el IBEX así como los dos índices americanos Dow JONES Industrial y Standard and Poors. Si un input tiene una alta correlación con otro input, podría provocar multicolinealidad y un fuerte ruido en el proceso de aprendizaje de la red neuronal. Además, se observa la poca correlación entre EURUSD y el resto de inputs, en este caso estos inputs están poco correlacionados entre sí, con lo cual provocarán poco ruido en el proceso de aprendizaje de la red Backpropagation.

Para hacer el análisis debemos seleccionar los inputs que se van a utilizar para las redes neuronales y son los siguientes:

- Las medias móviles ponderadas de los 5 periodos anteriores de cada input.
- Los 5 días de retardo o sesiones de cada input (R1, R2, R3, R4, R5).

Se podría analizar las correlaciones cruzadas para cada uno de los inputs, pero en este caso se va a obviar.

A continuación, se presenta la tabla de correlaciones de las medias móviles de cada uno de los inputs para estudiar su comparativa con las correlaciones anteriores de los inputs y escoger las que más contribución tengan para la red neuronal.

	IBEX	MAIBEX	MAFIBEX	MADAX	MAFOIL	MAS.P	MAFTSE	MADJI	MADJT	MAEURUSD
IBEX	1	-0,04160821	-0,03360278	-0,02652967	-0,01762319	0,00422294	-0,04293082	0,00038525	0,00585722	0,0241782
MAIBEX	-0,04160821	1	0,99134682	0,79582772	0,30808111	0,72787093	0,70660754	0,70991114	0,64093085	0,35545537
MAFIBEX	-0,03360278	0,99134682	1	0,78717241	0,30446775	0,72106798	0,70010057	0,70344895	0,63367017	0,34942837
MADAX	-0,02652967	0,79582772	0,78717241	1	0,33594523	0,82149822	0,79576112	0,80682866	0,73060101	0,26631363
MAFOIL	-0,01762319	0,30808111	0,30446775	0,33594523	1	0,39885526	0,40172974	0,36487508	0,29111776	0,33876238
MAS.P	0,00422294	0,72787093	0,72106798	0,82149822	0,39885526	1	0,85212301	0,97803748	0,87404522	0,34889444
MAFTSE	-0,04293082	0,70660754	0,70010057	0,79576112	0,40172974	0,85212301	1	0,84020314	0,73869778	0,29539977
MADJI	0,00038525	0,70991114	0,70344895	0,80682866	0,36487508	0,97803748	0,84020314	1	0,84835854	0,33002505
MADJT	0,00585722	0,64093085	0,63367017	0,73060101	0,29111776	0,87404522	0,73869778	0,84835854	1	0,27857298
MAEURUSD	0,0241782	0,35545537	0,34942837	0,26631363	0,33876238	0,34889444	0,29539977	0,33002505	0,27857298	1

**Tabla 2 Tabla de correlaciones medias móviles.**

Fuente: Elaboración propia.

Se analiza si existe alta correlación entre las medias móviles con respecto al IBEX, en este caso, según los datos obtenidos, se mantienen las correlaciones con respecto a la tabla de activos que se ha comentado anteriormente. Según la tabla de correlaciones de las medias móviles, el input con mayor correlación es el Futuro del IBEX con respecto al IBEX y el menos correlacionado es el Futuro del petróleo. Observamos, por ejemplo, que la media móvil del Futuro del IBEX y el IBEX ha aumentado en dos décimas, con lo cual se confirma que son dos inputs altamente correlacionados y que podrían provocar ruido en el aprendizaje de la red neuronal; otro ejemplo que lo demuestra es el input Dow Jones Transportes y el DAX, que aumenta doce décimas con respecto a la anterior tabla de correlaciones. Anteriormente hemos destacado lo poco correlacionados que están el EURUSD y el resto de inputs, y sigue manteniéndose la baja correlación entre sus medias móviles.

Se puede afirmar, por tanto, la coincidencia que existe de las correlaciones tanto de las medias móviles de los inputs como de los propios inputs entre sí.

A continuación, se incluyen las tablas de correlaciones los retardos R1, R2, R3, R4 y R5 de cada input. En todos los datos que se presentan, no se aprecian correlaciones significativas y, por tanto, solo con los datos de correlación no se podría tomar una decisión clara de que inputs podría ser la mejor selección para que el proceso de aprendizaje de la red neuronal tenga el menor ruido posible.

	<i>IBEX</i>	<i>R1IBEX</i>	<i>R2IBEX</i>	<i>R3IBEX</i>	<i>R4IBEX</i>	<i>R5IBEX</i>
IBEX	1	0,01610809	-0,05182435	-0,04176525	0,00365641	-0,05660642
R1IBEX	0,01610809	1	0,0160492	-0,05183121	-0,04182598	0,00352867
R2IBEX	-0,05182435	0,0160492	1	0,0160654	-0,05183836	-0,04180634
R3IBEX	-0,04176525	-0,05183121	0,0160654	1	0,0160828	-0,05180216
R4IBEX	0,00365641	-0,04182598	-0,05183836	0,0160828	1	0,01608951
R5IBEX	-0,05660642	0,00352867	-0,04180634	-0,05180216	0,01608951	1

**Tabla 3 Tabla de correlaciones retardos del IBEX.**

Fuente: Elaboración propia.

En esta tabla se representan las correlaciones de los cinco últimos retardos o sesiones del activo IBEX, y se observa que en general todos los retardos del IBEX están poco correlacionados, se observa el Retardo 4 como el menos correlacionado de todos con respecto al IBEX; y el Retardo 5 como el más correlacionado.

	<i>IBEX</i>	<i>R1FIBEX</i>	<i>R2FIBEX</i>	<i>R3FIBEX</i>	<i>R4FIBEX</i>	<i>R5FIBEX</i>
IBEX	1	0,03198042	-0,05417678	-0,04611183	0,00113114	-0,05028438
R1FIBEX	0,03198042	1	0,01549194	-0,0449489	-0,05332215	-0,00431184
R2FIBEX	-0,05417678	0,01549194	1	0,0154949	-0,04497726	-0,05331204
R3FIBEX	-0,04611183	-0,0449489	0,0154949	1	0,01548355	-0,04494675
R4FIBEX	0,00113114	-0,05332215	-0,04497726	0,01548355	1	0,01544633
R5FIBEX	-0,05028438	-0,00431184	-0,05331204	-0,04494675	0,01544633	1

**Tabla 4 Tabla de correlaciones retardos del Futuro del IBEX.**

Fuente: Elaboración propia.

En esta tabla se representan las correlaciones de los cinco últimos retardos o sesiones del activo Futuro del IBEX. No se observa gran correlación entre los inputs, no obstante, coincide con los datos de la anterior tabla con que el menos correlacionado es el R4 del Futuro del IBEX, ya que está relacionado directamente con el IBEX; y el más correlacionado sería el R2, en este caso.

	<i>IBEX</i>	<i>R1DAX</i>	<i>R2DAX</i>	<i>R3DAX</i>	<i>R4DAX</i>	<i>R5DAX</i>
<i>IBEX</i>	1	0,01534433	-0,04566855	-0,0389661	0,03078916	-0,03184646
<i>R1DAX</i>	0,01534433	1	0,00635123	-0,03936053	-0,02979314	0,03437165
<i>R2DAX</i>	-0,04566855	0,00635123	1	0,00618557	-0,03927129	-0,02936787
<i>R3DAX</i>	-0,0389661	-0,03936053	0,00618557	1	0,00622374	-0,03897946
<i>R4DAX</i>	0,03078916	-0,02979314	-0,03927129	0,00622374	1	0,00605852
<i>R5DAX</i>	-0,03184646	0,03437165	-0,02936787	-0,03897946	0,00605852	1

**Tabla 5 Tabla de correlaciones retardos del DAX.**

Fuente: Elaboración propia.

En esta tabla se representan las correlaciones de los cinco últimos retardos o sesiones del activo DAX. Los datos obtenidos no muestran, al igual que en el resto de tablas una correlación significativa de los inputs, se destaca el R1 DAX como el menos correlacionado con respecto al IBEX; y el que mayor correlación presenta es R2 DAX.

	<i>IBEX</i>	<i>R1FOIL</i>	<i>R2FOIL</i>	<i>R3FOIL</i>	<i>R4FOIL</i>	<i>R5FOIL</i>
<i>IBEX</i>	1	0,00020133	-0,04908055	0,00865005	0,0225135	-0,01777903
<i>R1FOIL</i>	0,00020133	1	0,15622938	-0,04413382	-0,03008653	0,06700169
<i>R2FOIL</i>	-0,04908055	0,15622938	1	0,15624112	-0,0441077	-0,03002095
<i>R3FOIL</i>	0,00865005	-0,04413382	0,15624112	1	0,1562137	-0,04393957
<i>R4FOIL</i>	0,0225135	-0,03008653	-0,0441077	0,1562137	1	0,15644013
<i>R5FOIL</i>	-0,01777903	0,06700169	-0,03002095	-0,04393957	0,15644013	1

**Tabla 6 Tabla de correlaciones retardos del Futuro del Petróleo.**

Fuente: Elaboración propia.

En esta tabla se representan las correlaciones de los cinco últimos retardos o sesiones del activo Futuro del Petróleo. No se observa mucha correlación entre los inputs, pero se ve claramente que el Retardo 1 es el menos correlacionado; y el Retardo 2 sería el que más correlación presenta con respecto al IBEX.

	<i>IBEX</i>	<i>R1S.P</i>	<i>R2S.P</i>	<i>R3S.P</i>	<i>R4S.P</i>	<i>R5S.P</i>
IBEX	1	0,06841651	-0,07541478	-0,00529174	0,02977003	-0,05153462
R1S.P	0,06841651	1	0,12059001	-0,10080458	-0,00950551	0,01449827
R2S.P	-0,07541478	0,12059001	1	0,12068967	-0,10084037	-0,00918294
R3S.P	-0,00529174	-0,10080458	0,12068967	1	0,12073539	-0,10059966
R4S.P	0,02977003	-0,00950551	-0,10084037	0,12073539	1	0,12070653
R5S.P	-0,05153462	0,01449827	-0,00918294	-0,10059966	0,12070653	1

**Tabla 7** Tabla de correlaciones retardos de Standard and Poors.

Fuente: Elaboración propia.

En esta tabla se representan las correlaciones de los cinco últimos retardos o sesiones del activo Standard and Poors. Al igual que no se presenta alta correlación en el resto de tablas revisadas, en este caso se observa lo mismo. El input que menos correlación tiene es R3 Standard and Poors; en el que mayor correlación se observa es R1 Standard and Poors.

	<i>IBEX</i>	<i>R1FTSE</i>	<i>R2FTSE</i>	<i>R3FTSE</i>	<i>R4FTSE</i>	<i>R5FTSE</i>
IBEX	1	-0,02943436	-0,01591858	-0,03158303	0,01186467	-0,05314856
R1FTSE	-0,02943436	1	0,12606945	-0,04764146	-0,05583577	-0,01486534
R2FTSE	-0,01591858	0,12606945	1	0,12607923	-0,04765177	-0,05557946
R3FTSE	-0,03158303	-0,04764146	0,12607923	1	0,12616208	-0,04757386
R4FTSE	0,01186467	-0,05583577	-0,04765177	0,12616208	1	0,12606879
R5FTSE	-0,05314856	-0,01486534	-0,05557946	-0,04757386	0,12606879	1

**Tabla 8** Tabla de correlaciones retardos del FTSE.

Fuente: Elaboración propia.

En esta tabla se representan las correlaciones de los cinco últimos retardos o sesiones del activo FTSE. Tampoco en estos resultados de correlaciones se observa gran correlación entre los inputs, no obstante, se ve que el input con menor correlación es R4 FTSE; y el de mayor es R5 FTSE.

	<i>IBEX</i>	<i>R1DJI</i>	<i>R2DJI</i>	<i>R3DJI</i>	<i>R4DJI</i>	<i>R5DJI</i>
IBEX	1	0,06502347	-0,08297198	-0,00297207	0,03775941	-0,0568343
R1DJI	0,06502347	1	0,10873593	-0,08281481	-0,01241276	0,0341297
R2DJI	-0,08297198	0,10873593	1	0,10882994	-0,08287136	-0,01218389
R3DJI	-0,00297207	-0,08281481	0,10882994	1	0,10888407	-0,08266802
R4DJI	0,03775941	-0,01241276	-0,08287136	0,10888407	1	0,10880569
R5DJI	-0,0568343	0,0341297	-0,01218389	-0,08266802	0,10880569	1

**Tabla 9** Tabla de correlaciones retardos del Dow Jones Industrial.

Fuente: Elaboración propia.

En esta tabla se representan las correlaciones de los cinco últimos retardos o sesiones del activo Dow Jones Industrial. Todos los inputs están muy poco correlacionados, pero se ve claramente que el input con menos correlación es el R3 Dow Jones Industrial; y el input con más correlación es R2 Dow Jones Industrial.

	<i>IBEX</i>	<i>R1DJT</i>	<i>R2DJT</i>	<i>R3DJT</i>	<i>R4DJT</i>	<i>R5DJT</i>
IBEX	1	0,05783297	-0,05139174	-0,01619909	0,02151406	-0,03147239
R1DJT	0,05783297	1	0,14328152	-0,04281659	-0,03766457	-0,00693991
R2DJT	-0,05139174	0,14328152	1	0,14319441	-0,04281673	-0,03739865
R3DJT	-0,01619909	-0,04281659	0,14319441	1	0,14311561	-0,04312169
R4DJT	0,02151406	-0,03766457	-0,04281673	0,14311561	1	0,14304385
R5DJT	-0,03147239	-0,00693991	-0,03739865	-0,04312169	0,14304385	1

**Tabla 10** Tabla de correlaciones retardos del Dow Jones Transportes.

Fuente: Elaboración propia.

En esta tabla se representan las correlaciones de los cinco últimos retardos o sesiones del activo Dow Jones Transportes. No se aprecia correlación significativa entre los inputs, pero se observa que el menos correlacionado con el IBEX es R3; y el que mayor correlación presenta es R1.

	<i>IBEX</i>	<i>R1EURUSD</i>	<i>R2EURUSD</i>	<i>R3EURUSD</i>	<i>R4EURUSD</i>	<i>R5EURUSD</i>
<i>IBEX</i>	1	0,0364879	-0,01502289	0,00569514	0,03565561	-0,008563
<i>R1EURUSD</i>	0,0364879	1	0,1961369	-0,00154183	-0,03316483	0,02195673
<i>R2EURUSD</i>	-0,01502289	0,1961369	1	0,19638538	-0,00193518	-0,03370337
<i>R3EURUSD</i>	0,00569514	-0,00154183	0,19638538	1	0,19679617	-0,00155931
<i>R4EURUSD</i>	0,03565561	-0,03316483	-0,00193518	0,19679617	1	0,19734744
<i>R5EURUSD</i>	-0,008563	0,02195673	-0,03370337	-0,00155931	0,19734744	1

**Tabla 11 Tabla de correlaciones retardos del EURUSD.**

Fuente: Elaboración propia.

En esta tabla se representan las correlaciones de los cinco últimos retardos o sesiones del activo EURUSD. Tampoco se observa una gran correlación entre sus inputs, no obstante, se ve como el que menor correlación presenta es el R3 EURUSD; y el que mayor es el R4 EURUSD.

Como se ha indicado anteriormente, el objeto de estudio no es en sí mismo construir el mejor modelo de red para la predicción del IBEX 35, sino verificar que la preselección de los inputs mejora la calidad del entrenamiento y de las predicciones.

Por tanto, para la configuración de la red neuronal Backpropagation con previa preselección de inputs, dispondrá de la siguiente configuración:

- Una capa oculta
- 10.000 iteraciones
- Pesos iniciales configurados de manera aleatoria
- Los parámetros alfa y beta se establecen de forma equilibrada en un valor de 0,5 fijo para cada uno de ellos.
- No se ha partido la muestra para realizar testeo.

El estudio podría ampliarse utilizando otras configuraciones de red, otro tipo de redes, incluso otros inputs, también podrían añadirse más indicadores técnicos, pero está fuera del objeto de este análisis y se deja para futuras investigaciones.

Todas las redes neuronales se van a construir mediante el lenguaje de programación de R.

R es un lenguaje y entorno de computación y gráficos estadísticos desarrollado en los Laboratorios Bell (antes de AT & T, ahora Lucent Technologies) por John Chambers y su equipo. R ofrece una amplia variedad de técnicas gráficas estadística

(lineal y no lineal de modelado, pruebas estadísticas clásicas, análisis de series temporales, clasificación, agrupación,...) y proporciona una ruta de código abierto.

R es un conjunto integrado de servicios de software para la manipulación de datos, cálculo y representación gráfica que incluye:

- Un manejo eficaz de los datos y la instalación de almacenamiento
- Un conjunto de operadores para los cálculos de matrices
- Integración de herramientas intermedias para el análisis de datos
- Instalaciones gráficas para el análisis y visualización de datos
- Un simple y eficaz lenguaje de programación

Se van a realizar varios entrenamientos de redes para hacer la preselección de inputs, todos ellos con redes de tipo Backpropagation, con lo cual se construye:

- Una red del IBEX con todas las medias móviles de cada input (una red por cada input y sus medias móviles).
- Una red por cada retardo (R1, R2, R3, R4 y R5). Se ha cogido todos los retardos de cada input.

Todos estos inputs se entrenan en una misma red, todos los R1, R2, R3, R4 y R5. De cada uno de los Retardos con todos los inputs habrá que analizar que inputs están más correlacionados entre sí, aunque bien se podrían estudiar otro tipo de entrenamientos que no se van a dar en este análisis.

Todos estos entrenamientos se hacen para poder hacer una preselección de los inputs indicados y se va a ver aquellos que, en todas estas redes, tienen una mayor contribución en cada red. Para determinar la contribución de estos inputs se utiliza el algoritmo de Garson.

Según Garson, los pesos que conectan las variables en una red neural son parcialmente análogos a los coeficientes de parámetros en un modelo de regresión estándar, y puede ser utilizado para describir las relaciones entre las variables. Los pesos dictan la influencia relativa de la información que se procesa en la red de tal manera que variables de entrada que no son relevantes en su correlación con una variable de respuesta son suprimidas por los pesos.

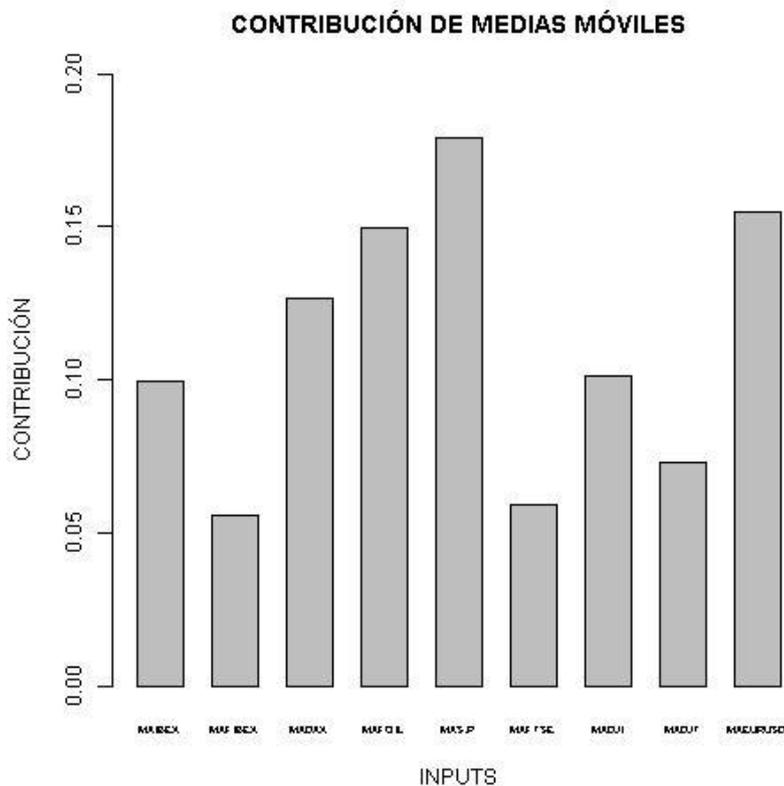
El efecto contrario se observa para los pesos asignados a las variables explicativas que tienen fuertes asociaciones positivas o negativas con una variable de respuesta. Una diferencia obvia entre una red neuronal y un modelo de regresión es que el número de pesos es excesivo en la antigua caso. Esta característica es ventajosa ya que hace que las redes neuronales sean muy flexibles para el modelado de

funciones no lineales con múltiples interacciones, aunque la interpretación de los efectos de las específicas variables es por supuesto un reto.

Un método descrito por Garson en 1991, identifica la importancia relativa de las variables en una red neuronal supervisada teniendo en cuenta los pesos del modelo. La importancia relativa (de la fuerza de asociación) de una variable específica, se puede determinar mediante la identificación de todas las conexiones ponderadas entre los nodos de interés, es decir, todos los pesos de conexión al nodo de entrada específico (inputs) que pasen a través la capa oculta a la variable de respuesta, se identifican como resultado de la red. Esto se repite para todas las demás variables explicativas hasta tener una lista con todos los pesos para cada input. Las conexiones son efectuadas por cada nodo de entrada y reducido en relación con todas las demás entradas. Se obtiene un único valor para cada variable explicativa que describe la relación con la variable de respuesta en el modelo. El algoritmo original indica la importancia relativa como la magnitud absoluta de 0 a 1, por tanto, la dirección de respuesta no puede ser determinada. El algoritmo actualmente sólo funciona para las redes neuronales con una capa oculta y una única respuesta variable. (Garson, 1991); (Goh, 1995); (Olden *et al* Jackson, 2002); (Olden *et al* Joy *et al* Death, 2004)

Para hacer la preselección de inputs que necesitamos, se han entrenado las redes que contienen los Retardos (R1, R2, R3, R4 y R5) para predecir el IBEX, por una parte y por otra las medias móviles de cada input. En las tablas de cotización anteriores pueden verse las medidas de contribución de cada uno con respecto al IBEX.

A continuación, se muestran las gráficas de contribuciones de los cinco últimos retardos o sesiones de cada input y de las medias móviles de los inputs, para escoger los que mayor contribución tienen en la red y hacer la preselección de los mismos. Para cada una de las redes, los resultados han sido los siguientes:



**Gráfica 1 Contribución de las medias móviles de los inputs.**

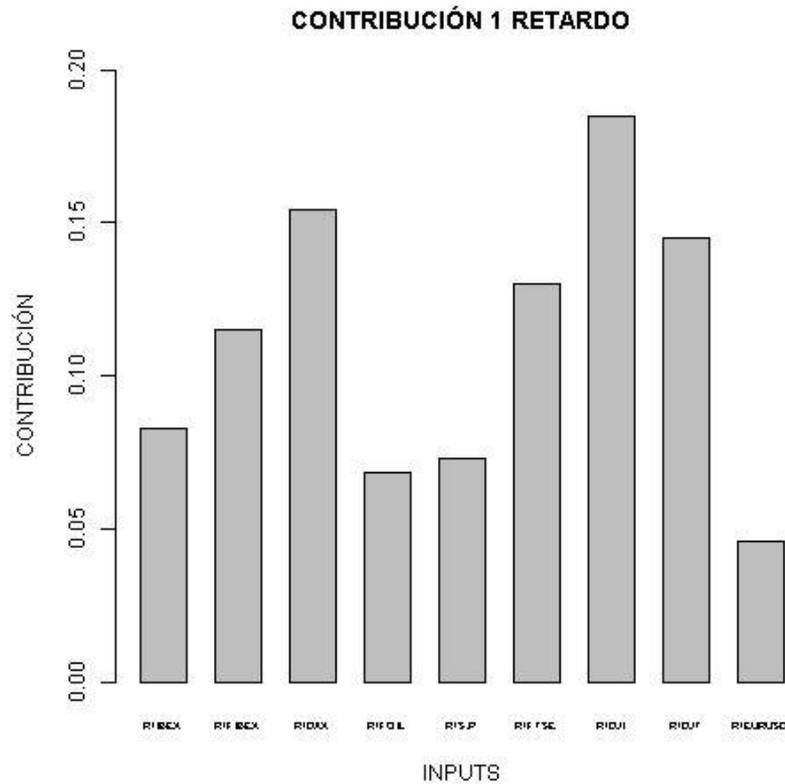
Fuente: Elaboración propia.

En esta gráfica se muestra la contribución que tienen las medias móviles de los inputs, dando como resultado que los tres inputs con mayor contribución son: Media Móvil del Standard and Poors; la Media Móvil del Futuro del Petróleo y Media Móvil de EURUSD. Comparando las gráficas de contribución de las medias móviles, con las tablas de correlaciones de las medias móviles elaboradas anteriormente, podemos deducir que los tres inputs con mayor correlación, coinciden, en este caso, estando entre los cinco primeros inputs con menor correlación con respecto al IBEX.

Se muestra a continuación, la tabla resumen de los datos de correlación de las medias móviles de los inputs:

	IBEX	MAIBEX	MAFIBEX	MADAX	MAFOIL	MAS.P	MAFTSE	MADJI	MADJT	MAEURUSD
IBEX	1	-0,04160821	-0,03360278	-0,02652967	-0,01762319	0,00422294	-0,04293082	0,00038525	0,00585722	0,0241782

**Tabla 12 Tabla de correlaciones de medias móviles**



**Gráfica 2 Contribución de los retardos 1 de los inputs.**

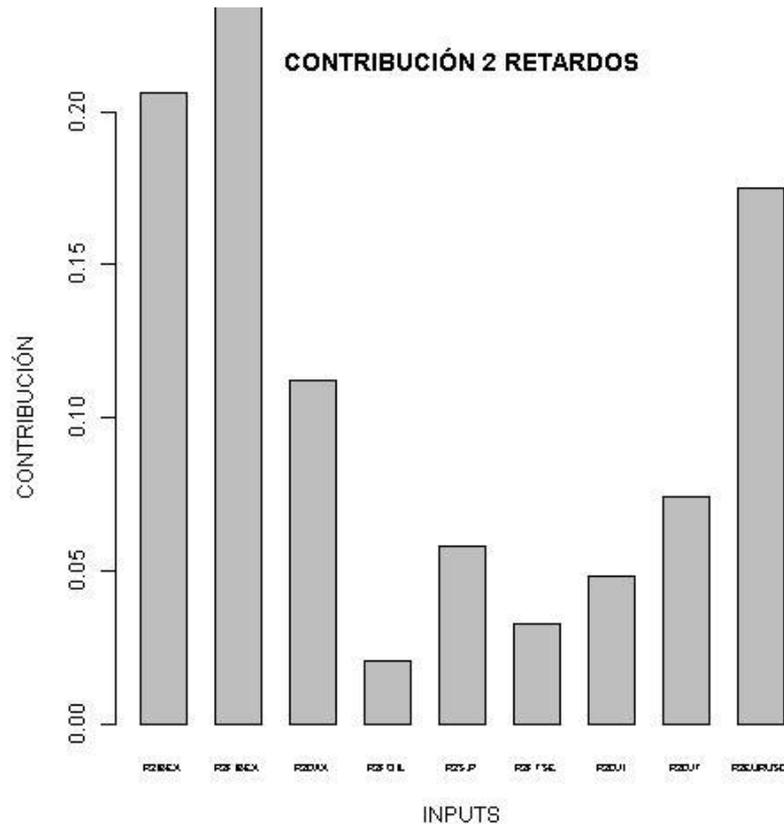
Fuente: Elaboración propia.

En esta gráfica se muestra la contribución que tienen los inputs con Retardo 1, dando como resultado que los tres inputs con mayor contribución son: R1 Dow Jones Industrial; R1 del DAX y R1 Dow Jones Transportes. En este caso, si observamos la gráfica presentada en relación a los datos de correlación existente entre los inputs con respecto al IBEX, explicados anteriormente, se observa la coincidencia del R1 DJI y R1DJT como los inputs con mayor correlación y a su vez con mayor contribución, no obstante, queda el DAX en segunda posición como el input con menor correlación y a su vez como uno de los tres con mayor contribución en la red.

Se muestra a continuación, la tabla resumen de los datos de correlación de los Retardos 1 de los inputs:

	IBEX	R1 IBEX	R1 FIBEX	R1 DAX	R1 FOIL	R1 S.P.	R1FTSE	R1 DJI	R1 DJT	R1 EURUSD
IBEX	1	0,01610809	0,03198042	0,01534433	0,00020133	0,06841651	-0,02943436	0,06502347	0,05783297	0,0364879

**Tabla 13 Tabla de correlaciones de los retardos R1**



**Gráfica 3 Contribución de los retardos 2 de los inputs.**

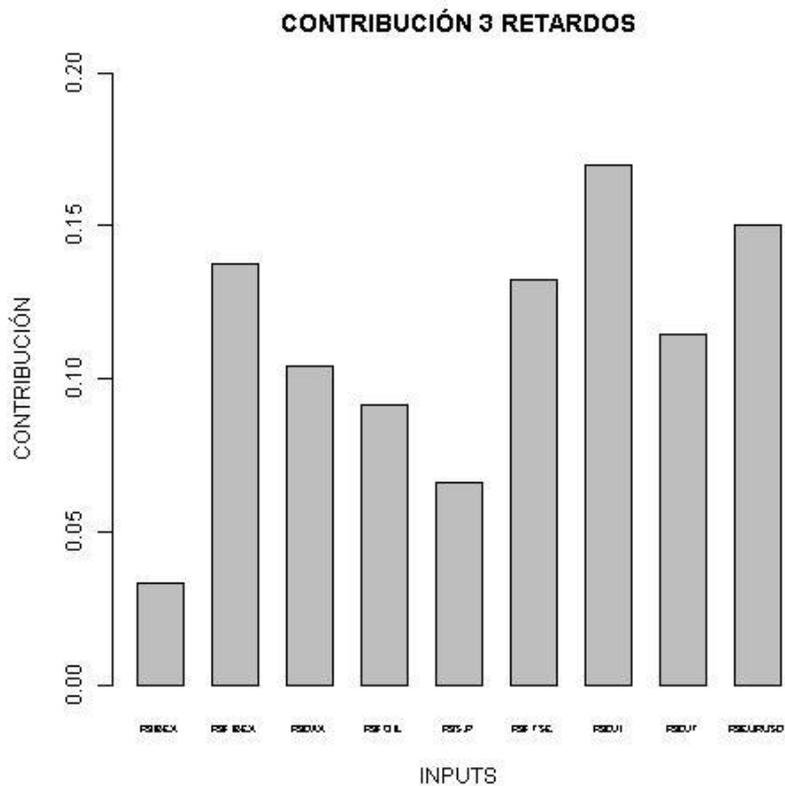
Fuente: Elaboración propia.

En esta gráfica se muestra la contribución que tienen los inputs con Retardo 2, dando como resultado que los tres inputs con mayor contribución son: R2 Futuro del IBEX; R2 IBEX y R2 EURUSD. Se comprueba que los resultados de la gráfica de contribución en contraposición a la tabla de correlación son similares, puesto que coinciden tanto el R2 FIBEX como el R2 IBEX como dos de los inputs con mayor contribución y con mayor correlación, y por el contrario, se observa que R2 EURUSD es uno de los inputs con mayor contribución y a su vez se posiciona como el de que menos correlación tiene con respecto al IBEX.

Se muestra a continuación, la tabla resumen de los datos de correlación de los Retardos 2 de los inputs:

	IBEX	R2 IBEX	R2 FIBEX	R2 DAX	R2 FOIL	R2 S.P.	R2 FTSE	R2 DJI	R2 DJT	R2 EURUSD
IBEX	1	-0,05182435	-0,05417678	-0,04566855	-0,04908055	-0,07541478	-0,01591858	-0,08297198	-0,05139174	-0,01502289

**Tabla 14 Tabla de correlaciones de los retardos R2**



**Gráfica 4 Contribución de los retardos 3 de los inputs.**

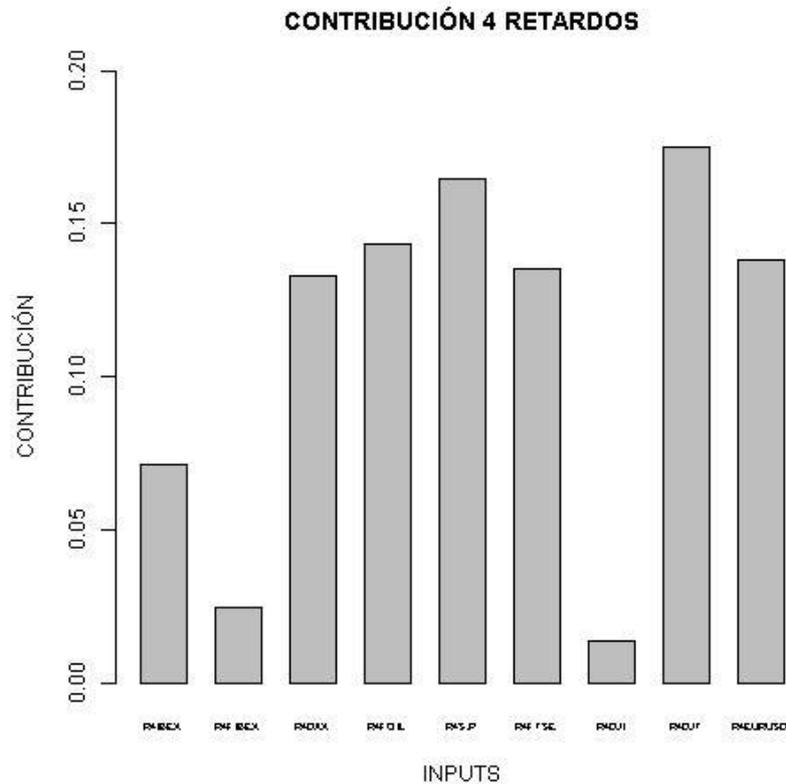
Fuente: Elaboración propia.

En esta gráfica se muestra la contribución que tienen los inputs con Retardo 3, dando como resultado que los tres inputs con mayor peso son: R3 Dow Jones Industrial; R3 EURUSD y R3 Futuro del IBEX. Se visualiza que tanto R3 DJI como R3 EURUSD son dos de los inputs con mayor contribución, según la gráfica mostrada, y a su vez, concuerdan como los dos inputs de menor correlación. Por otra parte, el input R3 FIBEX, se coloca como uno de los inputs con mayor correlación, así como de mayor contribución.

Se muestra a continuación, la tabla resumen de los datos de correlación de los Retardos 3 de los inputs:

	IBEX	R3 IBEX	R3 FIBEX	R3 DAX	R3 FOIL	R3 S.P.	R3FTSE	R3 DJI	R3 DJT	R3 EURUSD
IBEX	1	-0,04176525	-0,04611183	-0,0389661	0,00865005	-0,00529174	-0,03158303	-0,00297207	-0,01619909	0,00569514

**Tabla 15 Tabla de correlaciones de los retardos R3**



**Gráfica 5 Contribución de los retardos 4 de los inputs.**

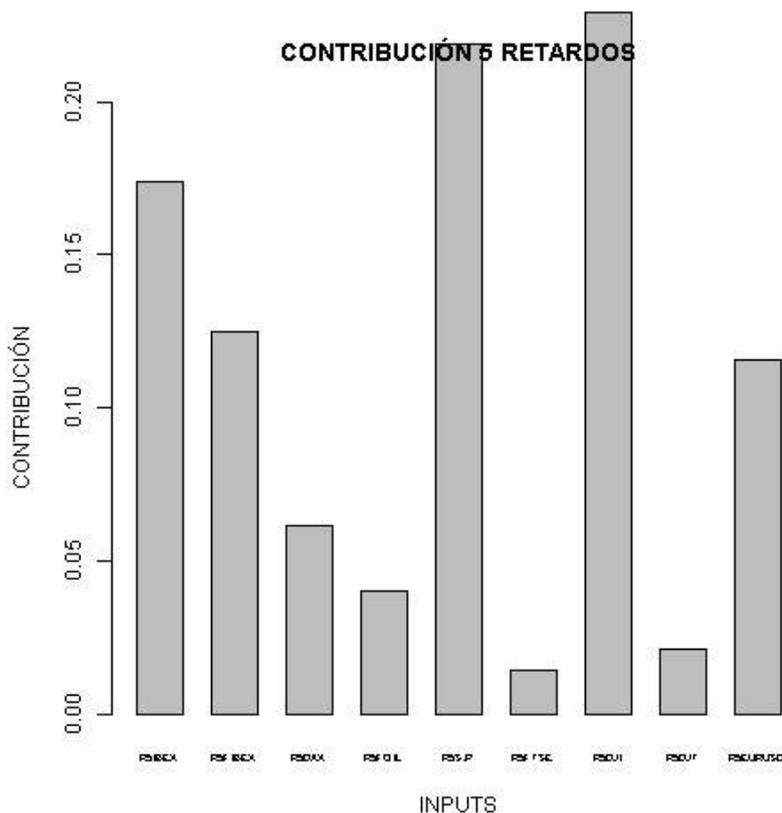
Fuente: Elaboración propia.

En esta gráfica se muestra la contribución que tienen los inputs con Retardo 4, dando como resultado que los tres inputs con mayor peso son: R4 Dow Jones Transportes; R4 Standard and Poors y R4 Futuro del petróleo. En este caso, se encuentran los tres inputs en una situación intermedia de valores de correlación, los cuáles no permiten confirmar la afirmación de que puedan ser los valores con mayor contribución y, a su vez, con mayor o menor correlación.

Se muestra a continuación, la tabla resumen de los datos de correlación de los Retardos 4 de los inputs:

	IBEX	R4 IBEX	R4 FIBEX	R4 DAX	R4 FOIL	R4 S.P.	R4FTSE	R4 DJI	R4 DJT	R4 EURUSD
IBEX	1	0,00365641	0,00113114	0,03078916	0,0225135	0,02977003	0,01186467	0,03775941	0,02151406	0,03565561

**Tabla 16 Tabla de correlaciones de los retardos R4**



**Gráfica 6 Contribución de los retardos 5 de los inputs.**

Fuente: Elaboración propia.

En esta gráfica se muestra la contribución que tienen los inputs con Retardo 5, dando como resultado que los tres inputs con mayor peso son: R5 Dow Jones Industrial; R5 Futuro del Petróleo y R5 IBEX. Como se puede observar, como en casos anteriores, se pone de manifiesto que dos de los inputs con mayor contribución son a su vez los que mayor correlación tienen, como son R5 DJI y R5 IBEX, siendo por contrapartida, el R5 FOIL.

Se muestra a continuación, la tabla resumen de los datos de correlación de los Retardos 5 de los inputs:

	IBEX	R5 IBEX	R5 FIBEX	R5 DAX	R5 FOIL	R5 S.P.	R5FTSE	R5 DJI	R5 DJT	R5 EURUSD
IBEX	1	-0,05660642	-0,05028438	-0,03184646	-0,01777903	-0,05153462	-0,05314856	-0,0568343	-0,03147239	-0,008563

**Tabla 17 Tabla de correlaciones de los retardos R5**

Así pues, se han escogido los tres inputs indicados con mayor contribución como inputs preseleccionados. Este ha sido un criterio arbitrario, podría haberse cogido otro número de inputs, pero se pretende reducir el número total de inputs.

Finalmente, los inputs seleccionados para construir la red han sido:

“R2FIBEX”; “R3FIBEX”; “R5FIBEX”; “R1DAX”; “R2DAX”; “R4FOIL”; “R4S.P”; “R5S.P”; “R1DJI”; “R3DJI”; “R5DJI”; “R1DJT”; “R4DJT”; “R2EURUSD”; “R3EURUSD”; “MAFOIL”; “MAS.P”; “MAEURUSD”.

Según el análisis efectuado con respecto a las gráficas de contribución y correlaciones entre los inputs, no podemos decir a ciencia cierta, en este caso, que los inputs con mayor contribución son también los que menor o mayor correlación presentan con respecto al IBEX. Para evaluar dicho estudio, se han comprobado los 6 casos de inputs (media móvil de cada input y R1, R2, R3, R4 y R5 de cada input), en total disponemos de 18 inputs, que, siguiendo un criterio de división de los valores con menor correlación, correlación neutral y mayor correlación, según sus posiciones, se ha obtenido que un 33,8% de valores de menor correlación son también los de mayor contribución en la red; un 38,8% de valores se encuentran en una correlación neutral; y, un 27,7% de valores de mayor correlación son también los de mayor contribución en la red. No podemos concluir, a la vista de estos resultados, que en esta preselección coincida, que los inputs de mayor contribución son los de mayor correlación y viceversa.

En este punto del análisis, se procede a realizar el entrenamiento de la red neuronal con todos los inputs preseleccionados y se compara con una red neuronal en la que se han incorporado todos los inputs. En ambos casos la estructura de la red neuronal es la ya comentada y utilizada anteriormente, excepto que en este caso se ha aumentado las iteraciones a 100.000.

Para evaluar la eficiencia o no de la preselección de inputs frente a la inclusión de todos los inputs en la red neuronal, se va a proceder a medir los diferentes tipos de red de la predicción del índice del IBEX 35. Existen muchos tipos de medidas del error. En este caso se van a calcular los siguientes tipos de error: Root Mean Square Error (RMSE), Mean Average Error (MAE), Mean Percentage Error (MPE), Mean Average Percentage Error (MAPE) que se describen como:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

$$MAE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|$$

$$MAE = \frac{1}{n} \sum_{i=1}^n \frac{\hat{y}_i - y_i}{y_i}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$

Siendo  $y_i$  el valor observado (rentabilidad del índice),  $\hat{y}_i$  el valor de predicción y  $n$  el número de predicciones realizadas.

Como se ha explicado anteriormente, es necesario calcular el error cometido por las predicciones para determinar qué red es más eficiente. Los resultados obtenidos se presentan en un cuadro resumen a continuación:

- **Error de predicción con todos los inputs**

	RMSE	MAE	MPE	MAPE
Test set	0,39200	0,39172	-7448,37700	17341,93995

**Tabla 18 Error de predicción con todos los inputs**

- **Error de predicción con los inputs seleccionados**

	RMSE	MAE	MPE	MAPE
Test set	0,39251	0,39219	-7356,89032	17256,40476

**Tabla 19 Error de predicción con los inputs seleccionados**

Se debe evaluar si hay alguna diferencia significativa en los resultados obtenidos en los errores de predicción de los inputs y hacer una comparativa entre los dos errores obtenidos. A simple vista, no hay ninguna diferencia significativa, quiere decir que la primera conclusión que podemos afirmar es que la hipótesis de hacer una preselección de inputs para la mejora del proceso de aprendizaje, para esta muestra concreta, para los inputs preseleccionados y para el proceso de los inputs seguidos, no obtiene ninguna ventaja respecto a no hacer esta preselección.

No obstante, existe una ligera diferencia a favor de la red que incluye todos los inputs en RSME y en el MAE, y sin embargo al revés en MPE y MAPE. Por tanto, no se puede sacar una conclusión clara con estos resultados. Incluso teniendo en cuenta esta diferencia, se utiliza mucho más tiempo para hacer la preselección de inputs y se

tendría que haber obtenido una diferencia notoria para que sea viable hacer la preselección.

Por tanto, podemos concluir, al menos con este estudio, que la capacidad de generalización que tiene la red Backpropagation, entendemos que supera al ruido que se produce en la red, porque al incorporar más variables, puede producirse multicolinealidad entre las mismas por la correlación cruzada entre las diferentes variables, por ejemplo, R1DAX con R1DJI no se ha llegado a analizar la correlación entre ellas y todo esto puede provocar ruido en el proceso de aprendizaje de la red neuronal.

## **IV. CONCLUSIONES.**

## CAPÍTULO IV: CONCLUSIONES

El planteamiento y objetivo de este estudio consiste en la modelización y predicción del precio del índice Ibex-35 mediante una red neuronal de tipo Backpropagation, en el que se ha realizado una preselección de los inputs. Dicha preselección se estimaba necesaria dado que, a pesar de la capacidad de generalización de un problema de la red Backpropagation, la incorporación de inputs altamente correlacionados entre sí podría provocar ruido en el proceso de aprendizaje de la red. Para evitar este hecho, se ha realizado una preselección de aquellos inputs que contribuyan con mayor peso en la predicción del índice Ibex 35, también mediante una red neuronal Backpropagation, pretendiendo conseguir una reducción del ruido de posibles interacciones entre dichos inputs.

En este proyecto se justifica si la metodología aplicada en las redes neuronales Backpropagation demuestra su teoría y mejora los resultados obtenidos haciendo una comparativa con la aplicación del método de redes neuronales, incluyendo todos los inputs, por una parte, y haciendo una preselección de los mismos, por otra. Por tanto, se construyen dos redes neuronales, ambas de tipo Backpropagation, una con todos los inputs, y otra con preselección, según el método indicado y que recordamos a continuación.

Respecto a la base de datos utilizada, parte de las fechas comprendidas entre 04/01/2006 al 24/04/2015. No obstante, finalmente se han cogido la muestra de las fechas 11/01/2006 al 31/12/2013. Los datos seleccionados son diarios y el análisis se ha realizado con las rentabilidades de los inputs.

El objeto de estudio no trataba de construir el mejor modelo de red para la predicción del IBEX sino verificar que la preselección de los inputs mejora la calidad del entrenamiento y de las predicciones, tanto para las diferentes redes utilizadas para la preselección de los inputs como para la red con todos ellos.

Los diferentes índices utilizados son:

- Futuro IBEX
- IBEX
- DAX
- Futuro de Petróleo
- STANDARD AND POORS
- FTSE
- DOW JONES Industrial
- DOW JONES Transportes
- EURUSD

La estructura de la red neuronal Backpropagation con preselección de inputs es la siguiente:

- Una capa oculta
- 10.000 iteraciones
- Pesos iniciales configurados de manera aleatoria
- Los parámetros alfa y beta se establecen de forma equilibrada en un valor de 0,5 fijo para cada uno de ellos.
- No se ha partido la muestra para realizar testeo.

Para la construcción de la red neuronal Backpropagation, con preselección de inputs, se han realizado varios entrenamientos de redes, con lo cual se construye:

- Una red del IBEX con todas las medias móviles de cada input (una red por cada input y sus medias móviles).
- Una red por cada retardo (R1, R2, R3, R4 y R5). Se ha cogido todos los retardos de cada input.

Finalmente, según los datos obtenidos en los entrenamientos, los inputs preseleccionados han sido:

“R2FIBEX”; “R3FIBEX”; “R5FIBEX”; “R1DAX”; “R2DAX”; “R4FOIL”; “R4S.P”; “R5S.P”; “R1DJI”; “R3DJI”; “R5DJI”; “R1DJT”; “R4DJT”; “R2EURUSD”; “R3EURUSD”; “MAFOIL”; “MAS.P”; “MAEURUSD”.

Por otra parte, la estructura de la red neuronal Backpropagation compuesta con todos los inputs, sin preselección, tiene la misma estructura, pero en lugar de 10.000 iteraciones, se han realizado 100.000 iteraciones.

Los resultados obtenidos en el estudio indican que, la preselección de inputs puede ser una alternativa para la mejora de la calidad de las predicciones de la red, dado que evita incorporar mucho ruido en la misma, no ha quedado claro que, para esta muestra y para estos inputs seleccionados, así como para la configuración de la red empleada, la preselección de los inputs tenga una ventaja clara y definida. Con la muestra utilizada, se ha llegado a la conclusión que la hipótesis de partida no es viable porque se pierde mucho tiempo en computación, para no obtener una mejora sustancial en los errores de predicción.

No obstante, se debería ampliar el estudio a otra muestra, otros inputs, o realizar el proceso de preselección no por tipo de retardo sino por bloque de activos, es decir, con todos los retardos de cada input de manera conjunta. De esta forma, tener una conclusión más definitiva al respecto.

## **V. BIBLIOGRAFÍA.**

## CAPÍTULO V: BIBLIOGRAFÍA

AYODELE ARIYO ADEBIYI; ADEREMI OLUYINKA ADEWUNI; CHARLES KOREDE AYO. (2014) Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction. Artículo publicado en Journal of Applied Mathematics. Volume 2014 (2014), Article ID 614342, pp 7

DEL CARPIO GALLEGOS, M. SC. (2005) Las Redes Neuronales Artificiales en las Finanzas. Industrial Data revista de investigación. ISSN: 1560-9146. Vol. 8, núm. 2, 2005 Universidad Nacional Mayor de San Marcos (Perú).

FREITAS, FABIO D.; DE SOUZA, ALBERTO F.; DE ALMEIDA, AILSON R. (2009) Prediction-based portfolio optimization model using neural networks. Artículo publicado en Journal Neurocomputing. Volume 72 Issue 10-12, Junio 2009, pp 2155-2170.

GARCÍA, MARTHA CECILIA; JALAL, AURA MARÍA; GARZÓN, LUIS ALFONSO; LÓPEZ, JORGE MARIO. (2013) Métodos para predecir índices bursátiles, pp 51-82. Artículo publicado en ECOS de Economía. ISSN 1657-4206 – No. 37 julio-diciembre 2013, pp 51-82.

GARSON, G.D. (1991). Interpreting neural network connection weights. Artículo publicado en Journal Artificial Intelligence Expert. Vol. 9, no. 1-2, pp. 119-165.

GOH, A.T.C. (1995). Back-propagation neural networks for modeling complex systems. Artículo publicado en Journal Artificial Intelligence in Engineering- ARTIF INTELL ENG. Vol. 9 no 3, pp 143-151.

GÓMEZ QUESADA , FRANCISCO JAVIER; FERNÁNDEZ GRACIANI, MIGUEL ÁNGEL; LÓPEZ BONAL, MARÍA TERESA; DÍAZ-MARTA, MARÍA ALONSO. (2010) Aprendizaje con redes neuronales artificiales. Revista Ensayos, nº 9- UCLM. Pp 169-179. Diciembre 1994. ISSN 0214-4824.

HUMBERTO CORTES CORREALES, DANIEL. (2002). Uso de Redes Neuronales Para Analizar Modelos Mecánicos de Tejidos Humanos. VI Congreso Colombiano De Elementos Finitos Y Modelamiento Numérico Uso De Redes Neuronales Para Analizar Modelos Mecánicos de Tejidos Humanos. Colombia (Mayo 2002)  
[https://www.researchgate.net/publication/267855840\\_Uso\\_de\\_Red\\_Neuronales\\_Para\\_Analizar\\_Modelos\\_Mecanicos\\_de\\_Tejidos\\_Biologicos](https://www.researchgate.net/publication/267855840_Uso_de_Red_Neuronales_Para_Analizar_Modelos_Mecanicos_de_Tejidos_Biologicos)

MINSKY, MARVIN; A.PAPERT SEYMOUR. (1987). Perceptrons: An introduction to Computational Geometry, Expanded Edition. Editorial The MIT Pres. Libro publicado 28 diciembre 1987.

MOLINA CABRERA, ALEXANDER; GARCÉS RUIZ, ALEJANDRO; TORO OCAMPO, ELIANA MIRLEDY. (2006) Pronóstico de bolsa de valores empleando técnicas inteligentes Tecnura. Vol. 9, núm. 18, pp. 57-66. Universidad Distrital Francisco José de Caldas, Bogotá (Colombia)

NAZEMI, A., ABBASI, B., OMIDI, F. (2015). Solving portfolio selection models with uncertain returns using an artificial neural network scheme. Applied Intelligence, Vol. 42, June 2015, pages 609-621. Doi: 10.1007/s10489-014-0616-z

OLDEN, J.D., JACKSON, D.A. (2002). Illuminating the 'black-box': a randomization approach for understanding variable contributions in artificial neural networks. Ecological Modelling. Pp 135-150.

OLDEN, J.D., JOY, M.K., DEATH, R.G. (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. Ecological Modelling 154. Pp 389-397.

OLIVER MUNCHARAZ, JAVIER. (2013) Modelización de la Volatilidad Condicional en Índices Bursátiles: Comparativa Modelo Egarch versus Red Neuronal Backpropagation. Editorial Universidad Politécnica de Valencia. Facultad de Administración y Dirección de Empresas.

ROSILLO CAMBLOS, RAFAEL; DUNIS CL; DE LA FUENTE, D.; PINO, R. (2012) Predicción del Índice IBEX-35 aplicando Máquinas de Soporte Vectorial y Redes Neuronales. Artículo 6th International Conference on Industrial Engineering and Industrial Management. XVI Congreso de Ingeniería de Organización.

VALENCIA REYES, MARCO ANTONIO; YÁÑEZ MÁRQUEZ, CORNELIO; SÁNCHEZ FERNÁNDEZ, LUIS PASTOR. (2006) Algoritmo Backpropagation para Redes Neuronales: conceptos y aplicaciones. Instituto Politécnico Nacional Centro de Investigación en Computación. No 125. Serie Verde.

VANSTONE, BRUCE; FINNIE, GAVIN. (2008) An empirical methodology for developing stockmarket trading systems using artificial neural networks. Journal ELSEVIER.

YONGHUI DAI; DONGMEI HAN; WEIHUI DAI. (2013) Modeling and Computin of Stock Index Forecasting Based on Neural Network and Markov Chain.