



Propuesta de un modelo predictivo
de la variación del precio de acciones basado
en redes neuronales y análisis
de estados financieros¹
*Proposal of a stock prices variation
prediction model based on neural networks
and financial statements analysis*

MARTÍN PÉREZ-GUEVARA¹

mperezguevara@gmail.com

MultusConsulting C.A.

ERNESTO MORENO²

ernesto.moreno@multusconsulting.com

Universidad Metropolitana

Universidad Católica Andrés Bello

Recibido: 30/09/2011

Aceptado: 01/02/2012

-
- ¹ Artículo basado en el trabajo de grado presentado a la Universidad Metropolitana, titulado "Diseño de un modelo predictivo empleando redes neuronales supervisadas para pronosticar el comportamiento del precio de acciones de empresas en sectores industriales de EE.UU."
 - ² Graduado en la Universidad Metropolitana de Economía Empresarial con distinción Summa Cum Laude, mediante la presentación de trabajo de grado calificado sobresaliente y recomendado para publicación. Ha trabajado como asistente de Investigación para el desarrollo de metodologías y modelos de análisis de mercado y de Portafolio de inversión para MultusConsulting C.A.
 - ³ Economista con MSc en Economía Aplicada y especialización en Análisis Industrial de NYU, asesor de inversión registrado en Estados Unidos, profesor de economía en la UNIMET y de finanzas en la UCAB, con más de 13 años de experiencia en estrategia económica y análisis de mercado.



Resumen

Se propone el diseño de un modelo predictivo de los rangos de variación futura del precio de acciones, empleando redes neuronales y análisis de estados financieros. Para ello se forman las diversas variables de entrada del modelo a partir de una muestra de los reportes trimestrales y precios de acciones históricos de empresas de sectores industriales de EE.UU.

Para esto se implementa el “perceptrón multicapa” entrenado con el algoritmo de retropropagación con resiliencia empleando diversas variables, para finalmente evaluar el desempeño del modelo considerando la correcta clasificación de la muestra y una estimación de la certidumbre de la mencionada clasificación para cada uno de los rangos de variación futura del precio de las acciones. Además, se plantean diversas interrogantes adicionales surgidas de la evaluación de resultados que trascienden esta investigación y cuyas respuestas en un posterior desarrollo contribuirían posiblemente al perfeccionamiento del modelo propuesto.

Palabras clave: redes neuronales, precio de acciones, modelo predictivo.

Summary

A prediction model is proposed to predict future stock prices variations intervals based on neural networks and financial statements analysis. For which input variables are constructed with quarter financial statements and historical stock prices of a sample of EEUU domestic companies from the industrial sector.

Furthermore a multilayer perceptron trained with the resilient back-propagation algorithm using different variables is employed, to finally evaluate the performance of the model considering the ratio of correctly classified samples and the estimation of certainty of that classification regarding the different future stock prices variations intervals. Also new questions arise from the evaluation of results, that go beyond this research and which answers seems promising to possibly improve the proposed model.

Key words: neural networks, stock prices, prediction model.



Las redes neuronales

La era moderna de las redes neuronales comienza con el trabajo de McCulloch y Pitts (1943), citados por Haykin (2005). McCulloch fue un psiquiatra y neuroanatomista por entrenamiento que invirtió veinte años pensando sobre la representación de un evento en el sistema nervioso. Pitts fue un matemático prodigio, que se unió a McCulloch en 1942.

En su informe de investigación clásico, McCulloch y Pitts describen un cálculo lógico de las redes neuronales que unía los estudios de neurofisiología y lógica matemática. Su modelo formal de una neurona asumía una ley del “todo o nada”. Con un número suficiente de tales unidades simples, y conexiones sinápticas acomodadas adecuadamente y operando de forma sincronizada, McCulloch y Pitts mostraron que una red constituida de la mencionada manera pudiera, en principio, computar cualquier función computable. Esto fue un resultado sumamente significativo, y con ello es generalmente aceptado que nacieron las disciplinas de redes neuronales e inteligencia artificial.

En su forma más general, una red neuronal puede ser definida como una máquina diseñada para modelar la forma en que el cerebro realiza una cierta tarea o función de interés (como una máquina adaptativa). Siendo más formales en la definición, se diría que una red neuronal es un procesador distribuido en forma masivamente paralela hecho de unidades de procesamiento simples, que tiene una propensión natural a guardar conocimiento debido a la experiencia y hacerlo accesible para su uso.

El uso de redes neuronales posee diversas capacidades y propiedades sumamente, útiles como son: 1) la no linealidad, 2) el aprendizaje supervisado, 3) la adaptabilidad, 4) la respuesta evidenciada, 5) la información contextual, 6) la tolerancia a los fallos, 7) la implementación VLSI, 8) la uniformidad en el análisis y diseño, y 9) la analogía neurobiológica.

Todas las mencionadas características y propiedades de las redes neuronales, hacen de éstas una excelente herramienta para la resolución



de gran cantidad de problemas y las vuelven propicias para ser empleadas en la elaboración de modelos predictivos.

Una común implementación de red neuronal, empleada para la realización de tareas de clasificación de patrones y de predicción, es el “perceptrón multicapa”. Esta estructura será empleada en la investigación por el gran poder computacional que posee, debido a sus principales características.

Un perceptrón multicapa tiene tres características distintivas. La primera es que el modelo de cada neurona en la red emplea una función de activación no lineal. Es común observar entonces el empleo de la función logística sigmoide⁴, la cual está inspirada en la fase refractaria de las neuronas reales. La segunda es que la red tiene una o más capas de neuronas ocultas que no son parte de la entrada ni salida de datos de la red. Estas neuronas ocultas son las que le permitan a la red aprender funciones complejas. Y finalmente, la tercera característica es el alto grado de conectividad con el cual se construye este tipo de red neuronal.

Predicción de los mercados financieros

Predecir es el proceso mediante el cual se hacen proyecciones sobre el desempeño posterior basadas en datos históricos existentes. Una predicción precisa ayuda a la toma de decisiones y planeación para el futuro. Las predicciones otorgan a las personas el poder de modificar las variables presentes para obtener como resultado escenarios favorables.

La selección e implementación de una metodología adecuada de predicción ha sido siempre un problema importante de planificación y control para la mayoría de las empresas y agencias. La estabilidad organizacional y financiera de una organización depende de la precisión de

⁴ En general, una función sigmoide es una función real de variable real diferenciable, con una primera derivada no-negativa o no-positiva y con, exactamente, un punto de inflexión.

Representada por la fórmula:
$$P(t) = \frac{1}{1 + e^{-t}}$$



una predicción, en cuanto la información ganada sería usada para tomar decisiones clave en las áreas de recursos humanos, compras, mercadeo, publicidad y financiamiento de capital.

Cuando una variable ha de ser predicha, la dificultad para hacerlo depende de varios factores. Como principal tenemos el caso del patrón histórico de una variable y los elementos de entrada que afectan su comportamiento, lo cual puede incrementar la complejidad de la tarea de predicción. Un patrón de comportamiento histórico volátil puede sugerir que el factor a ser pronosticado es influenciado a su vez por múltiples otros factores, los cuales tienen efectos dinámicos que cambian con el tiempo. Algunos de los mencionados factores podrían no ser identificables y por ello se requeriría de una experticia ganada a través del tiempo para elaborar un modelo de predicción (Lakshminarayanan, 2005).

Hay diversas motivaciones para intentar predecir los precios en el mercado de acciones. La más básica se refiere a la búsqueda de beneficio económico. Cualquier sistema que pudiera identificar variaciones de precio específicas consistentemente dentro de un mercado dinámico podría hacer al poseedor de tal sistema bastante adinerado. De modo que muchos individuos, incluidos investigadores, profesionales de inversión e inversionistas promedio, continuamente buscan maneras de producir ese sistema superior que les permitirá obtener altos beneficios.

Otra motivación importante está relacionada con las áreas de investigación en finanzas. Ya que ha sido propuesta la Hipótesis del Mercado Eficiente (HME), que explica que al instante de descubrirse una oportunidad para tomar ventaja del mercado ésta deja de existir. La HME declara que no puede existir ningún sistema que consistentemente pueda vencer al mercado y que por lo tanto éste no puede ser predicho de ninguna forma (Lawrence, 1997).

Diversas investigaciones se han llevado a cabo para comprobar que el mercado puede ser predicho. Hsieh (1991), citado por Kalyvas (2001), probó para el S&P500 que sus retornos semanales de 1962 a 1989, sus retornos diarios de 1983 a 1989 y los retornos cada 15 minutos de 1988 no cumplían con las condiciones de la teoría del *paseo aleatorio*. Tsibouris

y Zeidenberg (1996), citados por Kalyvas (2001), encontraron evidencias en el mercado de acciones americano contra la forma débil de la HME. Mientras que White (1993), citado por Kalyvas (2001), no logró encontrar suficiente evidencia contra la HME cuando intentó predecir los retornos de la acción IBM diariamente (Kalyvas, 2001).

En general no ha habido consenso sobre la validez de la HME, pero muchos observadores del mercado tienden a creer en sus formas más débiles y por lo tanto no están dispuestos a compartir sistemas propietarios de inversión. Sin embargo, contrario a la HME, muchos investigadores alegan que el mercado de capitales y otros sistemas complejos muestran un comportamiento caótico. Entendiendo que el caos se refiere a sistemas determinísticos no lineales que aparentan ser aleatorios por la dificultad que existe para expresarlos teóricamente.

Debido a la habilidad que tienen las redes neuronales para modelar funciones no lineales, pareciera posible emplearlas para sobrepasar el desempeño de otros métodos de análisis clásicos y computacionales. Además de ser aplicadas para la predicción del mercado de acciones, las redes neuronales también han sido empleadas para realizar diversas tareas financieras. Hay diversos sistemas experimentales y comerciales diseñados para monitorear los mercados de mercancías y futuros, de intercambio de divisas, de planificación financiera, de estabilidad corporativa y para predecir posibles quiebras.

Los bancos sobre todo han empleado modelos de redes neuronales para revisar peticiones de crédito de sus clientes y evaluar las probabilidades de bancarrota. Mientras que los administradores de dinero o *money managers* los han empleado para construir portafolios de inversión rentables en tiempo real.

Es interesante destacar que, con el tiempo, se han desarrollado notables avances en la teoría detrás del diseño e implementación de las redes neuronales como herramienta de propósito general. Lo cual aumenta las posibilidades de emplearlas como técnica de análisis y predicción de los mercados de capitales. (Lawrence, 1997).



Sobre la metodología del modelo predictivo propuesto

El modelo parte de la implementación del perceptrón multicapa⁵ entrenado con el algoritmo de retropropagación con resiliencia. Para lo cual es necesario formar un conjunto de datos de entrada y de salida provenientes de la información financiera recabada para la investigación, que serán pre-procesados con base en los requerimientos de la construcción y del entrenamiento de la mencionada red neuronal.

Entonces, múltiples redes son entrenadas con diversas combinaciones de datos de entrada para hallar el conjunto de redes óptimas para la predicción de los establecidos intervalos de variación futura de precio. En la figura 1 es posible apreciar un resumen de la metodología desarrollada para el modelo.

Formación de la muestra

Para este estudio se tomó la información histórica de las acciones de las empresas domésticas en EE.UU., que son parte, según Edgar (el sistema de información de la SEC), de las 10000 empresas más grandes que cotizan en el mercado de capitales y que forman parte del sector de industriales, según la clasificación de empresas expuesta por la institución financiera Charles Schwab.

Así que se contó con una muestra de 687 empresas, para las cuales se recabó el ejemplar 27 contenido en los reportes trimestrales históricos de las mencionadas empresas, obtenido directamente de Edgar. Permitiendo ello, finalmente, con base en las consideraciones y necesidades de la investigación, formar una muestra de 5.215 conjuntos de información que contenían el valor de las variables financieras empleadas en el modelo y la relacionada variación de precios de la correspondiente acción para cuatro períodos diferentes, a 3 meses, a 6 meses, a 9 meses y a 12 meses.

⁵ Red neuronal empleada por excelencia para resolver problemas que involucran un algoritmo de aprendizaje por corrección de error, particularmente el de retropropagación, según Haykin (2005).

FIGURA 1

DIAGRAMA REPRESENTATIVO DE LA METODOLOGÍA DE LA INVESTIGACIÓN

	PASOS METODOLÓGICOS	DETALLE DE LOS PASOS METODOLÓGICOS	FUENTES Y ANTECEDENTES
INFORMACIÓN FINANCIERA	Formación de la muestra	<ul style="list-style-type: none"> Filtro de empresas de Charles Schwab Sistema de información de la SEC (EDGAR) Ejemplar 27 de los reporteros trimestrales 	<ul style="list-style-type: none"> Charles Schwab EDGAR
	Definición de variables	<ul style="list-style-type: none"> Variables de análisis de estados financieros con base porcentual Variables de análisis de razones financieras. 	<ul style="list-style-type: none"> Wild <i>et. al</i> (2007)
	Selección de variables (datos de entrada)	<ul style="list-style-type: none"> Análisis factorial con extracción de componentes principales y rotación ortogonal Varimax. 	<ul style="list-style-type: none"> Kim <i>et. al</i> (1988 y 1991)
	Determinación de rangos de variación precio (Datos de salida ideales)	<ul style="list-style-type: none"> Rangos inferior y superior. Rangos inferior, medio y superior. 	<ul style="list-style-type: none"> Haykin (2005)
IMPLEMENTACIÓN DE LAS REDES	Preprocesamiento de datos	<ul style="list-style-type: none"> Normalización. Decorrelación (análisis de componentes principales estandarizado). 	<ul style="list-style-type: none"> Haykin (2005) LeCun <i>et. al</i> (1998) Ramírez <i>et. al</i> (2007)
	Arquitectura de las redes neuronales	<ul style="list-style-type: none"> Función de activación sigmoide tangente hiperbólica. Capa de entrada con 1-6 neuronas Dos capas ocultas con diez neuronas Capa de salida con 2-3 neuronas. 	<ul style="list-style-type: none"> Haykin (2005) Le Cun <i>et. al</i> (1998)
	Entrenamiento de las redes neuronales	<ul style="list-style-type: none"> Retropropagación con resiliencia (Algoritmo de aprendizaje). 10% de la muestra para evaluación. Detención prematura de entrenamiento. 	<ul style="list-style-type: none"> Haykin (2005) Le Cun <i>et. al</i> (1998) Kalyvas (2001)
	Resultados del modelo	<ul style="list-style-type: none"> Presentación de variables asociadas, probabilidad y clasificación de cada red. Mejor red según la probabilidad de cada rango de variación de precio. 	

Fuente: Elaboración propia



Definición y selección de variables (datos de entrada)

Con base en la información disponible en la muestra formada y la teoría de análisis de estados financieros expuesta por Wild *et. al.* (2007), se consideraron 21 indicadores sustentados en el análisis de estados financieros porcentuales y 21 indicadores sustentados en el análisis de razones. A partir de ello, basado en los planteamientos teóricos de Haykin (2005) respecto al eficiente entrenamiento de un perceptrón multicapa, cuatro indicadores fueron desechados por presentar valores nulos repetidamente en un alto porcentaje de la muestra, semejando constantes, y se realizó un análisis factorial que permitiera representar grupos de indicadores altamente correlacionados con uno solo de ellos, disminuyendo así la dimensión de posibles datos de entrada de la red neuronal. En la Tabla 1 es posible apreciar las trece variables seleccionadas finalmente.

TABLA 1
VARIABLES SELECCIONADAS MEDIANTE EL ANÁLISIS FACTORIAL

NOMBRES DE LAS VARIABLES	REFERENCIA
Cuentas por cobrar netas %	A3
Inventario %	A4
Activos fijos %	A6
Depreciación %	A11
Acciones comunes %	A12
Otros valores de capital %	A13
Razón de la prueba de ácido	C2
Días de venta en inventarios	C5
Efectivo y valores a activo circulante	C7
Cuentas por cobrar neto a activo circulante	D5
Deuda total a capital	E1
Deuda de largo plazo a capital	E2
Margen de utilidad bruta	F1

Fuente: elaboración propia

Determinación de rangos de variación de precio (datos de salida ideales)

Ya que la variación de precio de las acciones es el fenómeno que se desea predecir, a partir de éste se deben formar los datos de salida ideales con los cuales se entrenará a la red neuronal. Así que, a partir de 90% de la muestra más antigua, se categorizó la variación de precio de las acciones en dos y tres rangos para cada uno de los cuatro marcos temporales propuestos en los conjuntos de información.

Definiendo los límites de los intervalos con base en una distribución equitativa de la muestra, ya que según Haykin (2005), es importante que la red se enfrente a una muestra de tamaño similar respecto a cada dato de salida que se desea la red aprenda a identificar. De modo que correspondieron aproximadamente 2.346 conjuntos de información a cada rango, al considerarse dos intervalos, y 1.564 conjuntos al considerarse tres.

Además, fue peculiar apreciar que la variación de precio en todos los marcos temporales considerados presentó una distribución prácticamente normal, con un pico ascendente resaltante centrado en el punto de variación nula de precio y más achatado en la medida que el marco temporal era mayor.

Finalmente es importante destacar que los datos de salida de la red neuronal, basados en los rangos de variación de precio determinados, fueron codificados como es propuesto en la teoría tradicional para el caso de dos rangos, asignando valores simétricos opuestos con un valor absoluto de 0,9. Mientras para el caso de tres rangos fueron codificados bajo el enfoque de la distancia equilátera euclidiana, de modo que la incorrecta clasificación de un rango es sopesada más eficientemente en los diversos pesos sinápticos de la red que en el caso del enfoque tradicional.

Preprocesamiento de datos

Luego partiendo de lo recomendado por Haykin (2005) y LeCun *et al.* (1998), se deben aplicar tres pasos a los datos de entrada de la red,

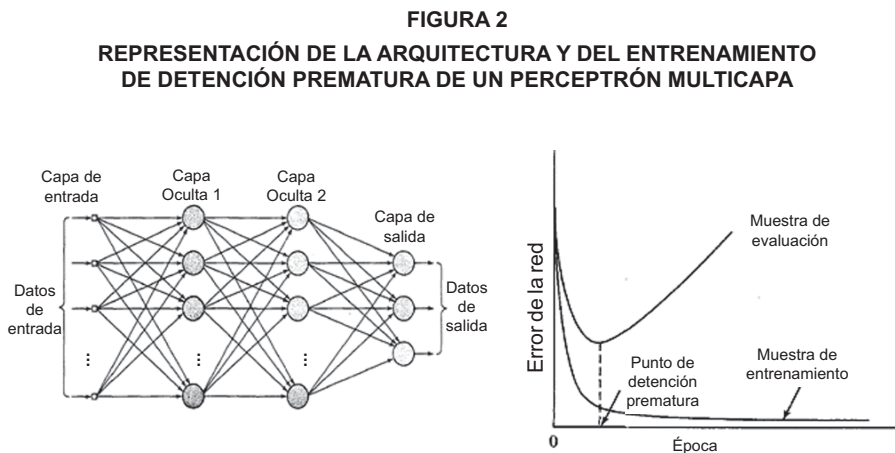


que son: a) sustraer la media a los valores; b) dividir los valores entre la desviación estándar; y c) decorrelacionar las variables.

Los primeros dos pasos son comúnmente llamados, de forma conjunta, normalización de las variables, y además de las ventajas que esta transformación aporta al proceso de entrenamiento de la red según LeCun (1998), permite también posteriormente realizar el análisis de componentes principales (ACP) estandarizado o centrado (Ramírez *et al.*, 2007), que conlleva a decorrelacionar las variables originales, obteniendo así nuevas variables decorrelacionadas que mantienen toda la información del conjunto de variables originales.

Arquitectura de las redes neuronales

Considerando el caso del perceptrón multicapa, que es el tipo de red implementado en el modelo y representado en la Figura 2, según Haykin (2005), la relación entre el tamaño de la muestra de datos debe ser mayor o igual a la cantidad de pesos sinápticos (tamaño de la red) entre el margen del error de clasificación permisible.



Fuente: Haykin (2005), pp. 181 y 238.

De modo que frente a la disposición de una muestra de entrenamiento de 4.698 conjuntos de datos y la determinación del máximo margen de error deseable de clasificación en un diez por ciento, se calculó un límite teórico para el tamaño de la estructura de la red de 469 parámetros libres, que definiendo sólo pesos sinápticos en la red, sin sesgos, conllevó a emplear combinaciones de hasta seis variables máximo como datos de entrada de las redes entrenadas.

Entonces, considerando las limitaciones estructurales impuestas por el tamaño de la muestra de datos disponible y recomendaciones realizadas por Haykin (2005), se determinó que la estructura de las redes neuronales estaría conformada por una primera capa de una a seis neuronas de entrada considerando las combinaciones de variables, por dos capas escondidas con diez neuronas ocultas cada una, y por una capa de salida de entre dos a tres neuronas de salida, considerando las dos posibles opciones de clasificación de rangos de variación de precio propuestas. Además se empleó como función de activación la sigmoide tangente hiperbólica, por su condición de simetría, lo cual es recomendado por LeCun (1998).

La estructura expuesta, conllevó entonces al entrenamiento de 33.176 redes distintas derivadas de las posibles combinaciones de las 13 variables en grupos de seis, cinco, cuatro, tres, dos y una variable para cada uno de los cuatro marcos temporales establecidos para cada una de las dos posibles clasificaciones de rangos de variación de precio determinadas.

Entrenamiento de las redes neuronales

En cuanto al algoritmo de aprendizaje, se determinó emplear la retropropagación con resiliencia, tomando como precedente el trabajo de Kalyvas (2001), el cual lo recomienda como el algoritmo más eficiente, considerando una comparación experimental que realiza frente a otros algoritmos alternativos. Así que por la implementación del mencionado algoritmo se hace necesario entrenar a la red empleando todos los campos de datos como un solo lote, de modo que en cada época



de entrenamiento todos los campos son introducidos a la red y evaluados conjuntamente.

En cuanto al proceso de entrenamiento de la red, se determinó que se emplearía diez por ciento de los 5.215 campos de la muestra de datos como conjunto de datos de evaluación. Por lo cual la muestra de entrenamiento quedó conformada por 4.693 campos de datos y la muestra de evaluación por 522.

También se decidió, con base en las recomendaciones de Haykin (2005), detener el proceso de entrenamiento buscando el mínimo del error de evaluación durante el proceso de entrenamiento, proceso representado en la Figura 2. Esta tarea no es sencilla, y dado que no necesariamente el error de entrenamiento y de evaluación fluctúan suavemente durante el proceso de entrenamiento y que es posible hallar mínimos locales en el error de evaluación, se procedió a considerar 50 épocas consecutivas de aumento del error de evaluación respecto al último mínimo alcanzado, como punto indicativo para detener el entrenamiento de la red.

Respecto a la presentación de resultados y evaluación de la red

Finalmente, basadas en el propósito de las redes neuronales, de predecir correctamente la variación de precio futuro, se determinó que la presentación de resultados de la red consistiera en: a) reportar la proporción de datos clasificados correctamente en sus respectivos rangos de variación de precio; y b) la probabilidad de que un dato clasificado en un rango particular, efectivamente pertenezca a ese rango y no sea una errada clasificación de otros rangos.

De modo que las diversas redes generadas por la metodología propuesta serán además evaluadas y comparadas en concordancia con las planteadas medidas de clasificación y probabilidad. Considerando a la “probabilidad de pertenencia al rango”, como el parámetro más importante y decisivo en cuanto a la determinación de las mejores redes, y al

‘grado de correcta clasificación del conjunto perteneciente al rango’ como el segundo factor más importante de evaluación.

Adicionalmente será reportado el conjunto de variables que forman parte de los mejores modelos, de forma que se pueda evaluar si algunos datos de entrada (variables) en particular contribuyen a mejorar la capacidad de predicción de la red para los diversos rangos de variación de precio estudiados.

Síntesis de resultados

Los resultados fueron analizados caracterizando y contrastando las redes neuronales que presentaron mejor desempeño predictivo respecto a cada uno de los intervalos de variación de precio establecidos en la investigación.

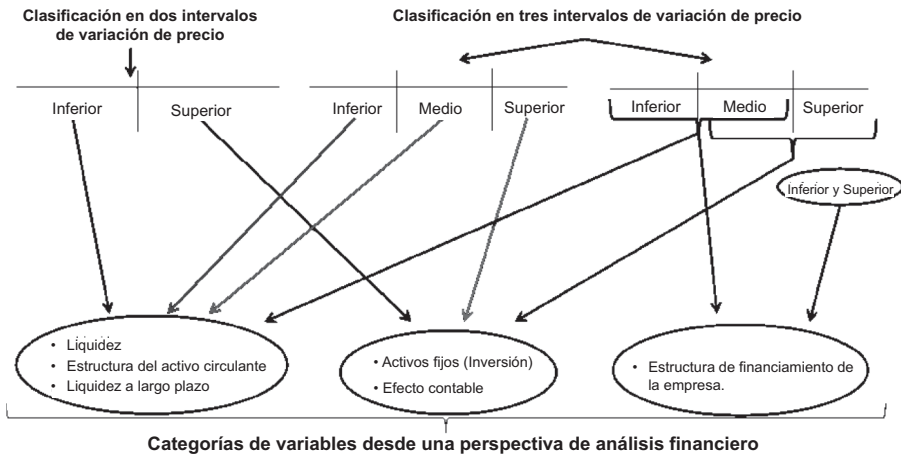
Apreciándose en primer lugar que ninguna red se destacó en la predicción general de todos los intervalos considerados, sino que por el contrario, cada red mostró destacarse en la predicción de variación de precio para intervalos específicos. Además, con base en la muestra de variables y empresas empleadas, se pudo apreciar un mejor desempeño del modelo para predecir ciertos intervalos respecto a otros.

El tipo de variables empleadas por las mejores redes respecto a cada uno de los intervalos de predicción puede ser apreciado en la Figura 3, en la cual se observa que existe una relación recurrente entre ciertos grupos de variables e intervalos de variación de precio.

Así que el modelo expone que variaciones negativas del precio de las acciones (intervalos inferiores) tienen relación con las variables representativas de la capacidad de la empresa para responder a requerimientos financieros y operativos a corto plazo (días de venta en inventarios y efectivo y valores a activo circulante), así como de su solvencia a largo plazo (deuda a largo plazo a capital) y su estructura de financiamiento (acciones comunes %). Mientras que las variaciones positivas del precio de las acciones están relacionadas con las decisiones de inversión de la



FIGURA 3
DIAGRAMA SÍNTESIS DE LAS CATEGORÍAS DE VARIABLES SEGÚN EL ANÁLISIS FINANCIERO, RELACIONADAS CON LAS MEJORES REDES SEGÚN EL INTERVALO DE VARIACIÓN DE PRECIO CONSIDERADO



Fuente: Elaboración propia.

Nota: El rango medio y el rango combinado inferior y superior están relacionados adicionalmente con diversas variables que no presentan consistencia en los diversos períodos de estimación considerados. Lo cual disminuye la aparente fuerza de relación entre estos grupos de intervalos y los grupos de variables identificados.

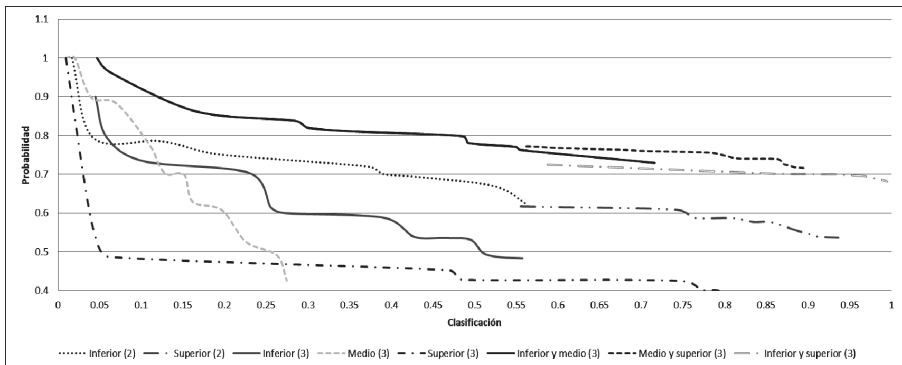
empresa (activos fijos %) y la manera en que se gestionan los activos a nivel contable (depreciación %).

La diferencia de desempeño del modelo para predecir los diversos intervalos de variación de precio puede ser apreciada en la Figura 4, la cual es un gráfico de curvas que representa los resultados de las mejores redes con base en probabilidad (certidumbre de que la acción clasificada en un rango efectivamente pertenece a éste) y clasificación (porcentaje de acciones pertenecientes al intervalo de variación de precio correctamente clasificado) para cada uno de los intervalos de variación de precio considerados, incluidas las combinaciones de intervalos en el caso de la clasificación en tres intervalos.

De modo que cada curva en la Figura 4 representa el conjunto de redes más eficientes respecto a los resultados de todas las redes entrenadas para cada intervalo. Donde la eficiencia se observa por lo alejado de la curva respecto a los ejes, dándose prioridad en la interpretación a la altura de la curva (valor de probabilidad) respecto a su distancia del eje vertical (valor de clasificación), y cada punto de la curva representa el desempeño de una red específica. Se destaca que no se distinguen los períodos de estimación considerados, que representarían partes de las curvas representadas en la Figura 4, o curvas inferiores a éstas.

FIGURA 4

GRÁFICO DE LA RELACIÓN DE PROBABILIDAD Y CLASIFICACIÓN DE LAS MEJORES REDES SEGÚN EL INTERVALO DE VARIACIÓN DE PRECIO CONSIDERADO



Fuente: Elaboración propia.

(2): Se refiere a dos intervalos de clasificación.

(3): Se refiere a tres intervalos de clasificación.

Nota: Se considera superior el desempeño de las redes neuronales en caso de presentar en primer lugar 'mejor probabilidad' y en segundo "mejor clasificación", de modo que las líneas superiores representan los intervalos mejor interpretados por las redes neuronales y, por lo tanto, las redes más eficientes en cada caso.



A partir de la Figura 4 es posible denotar que los intervalos inferiores de variación de precio son predichos por el modelo con mayor eficiencia que los intervalos superiores. Además se aprecia que la segmentación de las variaciones de precio en más intervalos benefició en gran medida al desempeño del modelo al considerar las combinaciones de intervalos. Pues la eficiencia de predicción de los intervalos combinados es superior a la de todos los otros intervalos interpretados por separado y su curva desciende con mayor suavidad a la vez que alcanza mejores niveles de clasificación.

Se observa además en la Figura 4 que existe una relación negativa entre los valores de probabilidad y clasificación, de modo que las redes capaces de clasificar un mayor número de variaciones de precio de acciones en un intervalo específico pierden necesariamente confianza en la mencionada clasificación. Esto se puede interpretar como que algunas redes identifican patrones de variación de precio relacionados específicamente con subgrupos de acciones de la muestra, donde en la medida en que los subgrupos son más pequeños, los patrones son más peculiares y precisos.

Finalmente fue interesante apreciar en los resultados, que se obtenía un mejor valor de probabilidad en las redes neuronales basadas en los mayores períodos de tiempo. Lo cual es contra-intuitivo, en cuanto se esperaba que tanto el valor de probabilidad como el de clasificación se vieran perjudicados por la incertidumbre, beneficiando así el más corto plazo. De modo que es posible que exista un período óptimo de predicción a ser descubierto e implementado en el modelo, que pudiera ser mayor al período de predicción de un año.

Conclusión, alcances y recomendaciones

Es importante destacar que se verificó la capacidad de un perceptrón multicapa con dos capas escondidas y apenas diez neuronas en cada capa, entrenado mediante la retropropagación con resiliencia, para identificar relaciones no lineales entre las consideradas variables basadas



en análisis de estados financieros y la variación futura del precio de acciones.

El énfasis de haber empleado sólo diez neuronas en cada capa se debe a la normal expectativa, al trabajar con redes neuronales, de necesitar un gran número de neuronas en cada capa para poder aproximar adecuadamente la función objetivo, y por lo tanto encontrar resultados significativos. Sin embargo, en esta investigación con sólo diez neuronas ya se obtuvieron resultados alentadores, mostrando la eficacia de la red neuronal para aproximar la función objetivo, que en nuestro caso es la predicción.

La estructura determinada para la red permitió un exitoso desarrollo del modelo predictivo, a pesar de haberse enfrentado esta investigación con limitaciones en la recolección de datos, en contraste con la cantidad ideal de datos para entrenar redes neuronales. Pues varias de las redes entrenadas mostraron valores de probabilidad superiores a 70%, con valores de clasificación también superiores a 70%, e incluso mejores, en el contexto del sector de industriales de los mercados de capitales norteamericanos.

Debido a que las diversas redes entrenadas mostraron especializarse en la predicción de intervalos de variación de precio específicos, resulta interesante plantearse la posibilidad de emplear conjuntamente redes distintas para formar una red neuronal compuesta, o sencillamente como múltiples filtros o advertencias en el proceso de formación y selección de estrategias de inversión.

Es interesante destacar que frente a los resultados del modelo, es posible emplearlo en diversas aplicaciones de selección y gestión de inversiones en los mercados de acciones. Ya que la capacidad de la red de distinguir rangos particulares de variación futura de precio permitiría elaborar complejas estrategias de inversión, empleando instrumentos financieros derivados, contribuir al desarrollo de filtros para la selección de acciones en la formación de portafolios de inversión, perfeccionar las estrategias implementadas para la posterior gestión de los mencionados portafolios de inversión, y asistir modelos de valoración de empresas.



Así que el modelo permitiría generar diversas estrategias de inversión, tanto activas como pasivas, aprovechando el trabajo conjunto de diversas redes entrenadas. Además es posible perfeccionar el modelo complementándolo con la elaboración e implementación de reglas de toma de decisiones enmarcadas en la programación de sistemas expertos.

Aparte del posterior desarrollo e implementación del modelo propuesto, sería importante plantearse una revisión más exhaustiva de los mecanismos de clasificación y agrupación de las empresas que fuesen empleados por el modelo, de modo que no sólo se utilicen los mecanismos tradicionales de clasificación industrial. Ya que nuevos criterios que permitieran identificar características cuantitativas y cualitativas comunes a grupos de empresas, pudieran contribuir enormemente a que fuesen correcta o incorrectamente predichas las variaciones de precio de sus acciones, siendo tal conocimiento crucial para refinar los alcances del modelo planteado y su utilidad.

Finalmente se recomienda para posteriores investigaciones explorar diversas cuestiones que escaparon del alcance de esta investigación, como es el caso de la definición de un mayor número de variables, posiblemente desde perspectivas distintas al análisis de estados financieros, el empleo de muestras de empresas más amplias pertenecientes a diversos contextos geográficos, la revisión de modificaciones a la estructura de la red neuronal –de disponerse de muestras más numerosas–, la búsqueda de períodos óptimos de predicción, y el desarrollo de modelos de evaluación más avanzados para medir y comprobar el desempeño de las redes neuronales entrenadas.



Referencias

- DASE, R.; PAWAR, D. (2010). "Application of Artificial Neural Network for stock market predictions: A review of literature". *International Journal of Machine Intelligence*, Volume 2, Issue 2, pp. 14-17, India.
- FERNANDO, P. y ANGUIANO-CARRASCO, C. (2010, enero-abril). El análisis factorial como técnica de investigación en Psicología. *Papeles del psicólogo*, Redalyc [versión electrónica], Vol. 31, pp. 18-33. Obtenido el 20 de noviembre de 2010, de <http://redalyc.uaemex.mx/src/inicio/ArtPdfRed.jsp?iCve=77812441003>
- HAYKIN, S. (2005). *Neural Networks a Comprehensive Foundation* (2nded.). India, Pearson Prentice Hall.
- HEATON, J. (2010). *Programming neural networks with encog 2 in java*. St. Louis, MO USA: WordsRU.com
- HURTADO, J. (2000). *Metodología de la investigación holística*. Venezuela, Caracas: Sypal.
- KIM, J.; MUELLER, C. (1988). *Factor Analysis, Statistical Methods and Practical Issues*. University of Iowa, USA: Sage publications, Inc.
- KIM, J.; MUELLER, C. (1991). *Introduction to Factor Analysis, what it is and how to do it*. University of Iowa, USA: Sage publications, Inc.
- KALYVAS, E. (2001). *Using neural networks and genetic algorithms to predict stock market returns*. Trabajo de grado, Master of Science in Advanced Computer Science, University of Manchester, UK: University of Manchester.
- LAKSHMINARAYANAN, S. (2005). *An integrated stock market forecasting model using neural networks*. Trabajo de grado, Master of Science, College of Engineering and Technology, Ohio University, USA: University of Ohio.
- LAWRENCE, R. (1997). *Using Neural Networks to Forecast Stock Market Prices*. Department of Computer Science, University of Manitoba, UK: University of Manitoba.
- LECUN, Y.; BOTTOU, L.; ORR, G. y MÜLLER, K. (1998). *Efficient BackProp*. USA: Springer.
- MURPHY, J. (2003). *Análisis técnico de los mercados financieros*. Barcelona: Ediciones Gestión 2000, S.A.



- RAMÍREZ, G. y VÁSQUEZ, M. (2007). *Análisis General* de su Curso de Análisis de Datos I. Universidad Central de Venezuela, Área de Postgrado en Estadística y Actuariado, Caracas: Manuscrito no publicado.
- REUTERS, (2002). *Curso de Bolsa*. Barcelona: Ediciones Gestión 2000.
- RIEDMILLER, M.; BRAUN, H. (s.f). *A Direct Adaptive Method for Faster Back-propagation Learning: The RPROP Algorithm*. University of Karlsruhe, Alemania: Autor.
- SMITH, L. (ed.) (2002). *A tutorial on Principal Componentes Analysis*. s.l: Autor.
- U.S. Census Bureau (2010). *ECPC Report 1*. Obtenido el 19 de noviembre de 2010, de http://www.census.gov/eos/www/naics/history/docs/report_1.pdf
- U.S. Census Bureau (2010). *Issues Paper N°1*. Obtenido el 19 de noviembre de 2010, de http://www.census.gov/eos/www/naics/history/docs/issue_paper_1.pdf
- U.S. Securities and Exchange Commission (2010). Visitado el 18 de noviembre de 2010, de <http://www.sec.gov>
- WILD, J.; SUBRAMANYAM, K. y HALSEY, R. (2007). *Análisis de estados financieros*. (9ª ed.). México: McGraw-Hill Interamericana.
- World Federation of Exchanges (2010). *Ts2 Market cap.XLS*. Obtenido el 18 de noviembre de 2010, de <http://www.world-exchanges.org/statistics/time-series/market-capitalization>
- WORLD FEDERATION OF EXCHANGES (2010). *Ts3 Share Trading.XLS*. Obtenido el 18 de noviembre de 2010, de <http://www.world-exchanges.org/statistics/time-series/value-share-trading>
- ZAPATA, J. (2006). *Efecto de la limpieza química en la permeabilidad de membranas cerámicas de ultrafiltración*. Tesis doctoral, Instituto de Biotecnología, Universidad de Granada, España: Editorial de la Universidad de Granada.
- ZEKIC, M. (s.f). *Neural Network Applications in Stock Market Predictions – A Methodology Analysis*. University of Josip Juraj Strossmayer in Osijek, Croatia: University of JosipJurajStrossmayer in Osijek.