

DOS APLICACIONES EMPÍRICAS DE LAS REDES NEURONALES ARTIFICIALES A LA CLASIFICACIÓN Y LA PREDICCIÓN FINANCIERA EN EL MERCADO ESPAÑOL

Jorge de Andrés Sánchez*

Universidad Rovira i Virgili

En muchos problemas de carácter financiero, la relación entre las variables implicadas es más compleja de lo que hacen suponer los modelos teóricos habituales (por otra parte, de utilidad incuestionable); o bien la información viene dada de forma poco estructurada. Numerosos autores han mostrado la utilidad que presenta en estas circunstancias los instrumentos que proporciona la Inteligencia Artificial. El objeto del presente trabajo es arrojar evidencias de la utilidad de uno de estos instrumentos, las Redes de Neuronas Artificiales, en dos problemas clásicos de clasificación y predicción financiera: la predicción de la quiebra y la estimación de los rendimientos de carteras de valores, por lo cual comparamos su capacidad predictiva con la de los métodos multivariantes utilizados habitualmente en dichos contextos.

Palabras clave: inteligencia artificial, redes neuronales artificiales, predicción de la quiebra, modelos multifactoriales, CAPM, APT.

1. INTRODUCCIÓN

Ha sido ampliamente contrastada la efectividad que proporcionan los instrumentos de la denominada Inteligencia Artificial en diversas áreas del análisis financiero. Una de estas herramientas son las Redes Neuronales Artificiales (RNAs), basadas en la imitación de la estructura y activi-

(*) El autor agradece los útiles comentarios efectuados por un evaluador anónimo sobre una versión preliminar del trabajo, que han permitido mejorarlo notablemente. También agradece las constructivas aportaciones y la ayuda proporcionada por Cristian Vidal de Ventós.

dades del cerebro humano. Dos de sus características más relevantes son su capacidad para aproximar funciones complejas, tal como muestran Hornick *et. al.* (1989), y su utilidad para agrupar elementos correctamente a partir de determinadas variables. Por ello, las RNAs pueden considerarse un instrumento alternativo a numerosos métodos multivariantes clásicos, lo que las hace aptas, tal como exponen Hawley *et. al.* (1990) y Serrano y Gallizo (1997), para ser aplicadas en un amplio espectro de problemas financieros.

Entre dichas aplicaciones podemos remarcar la calificación de riesgos crediticios y rating de bonos a partir de las características de los emisores (Dutta y Shekhar, 1988; Richeson *et. al.*, 1994; Jensen, 1992 y Castromán y Porto, 2000); la valoración de derivados financieros (Trippi y DeSieno, 1992, en la predicción de precios de los futuros sobre el índice S&P 500; y Hutchinson *et. al.*, 1994 y Lechón *et. al.*, 1997 en la valoración de opciones sobre acciones); la predicción del tipo de interés futuro a partir de la información que contiene la estructura temporal de los tipos de interés (Swanson y White, 1995 y Pérez *et. al.*, 2000); o en el análisis de series temporales financieras, en contraposición a instrumentos econométricos más convencionales como la modelización ARIMA o ARCH (Refenes, 1996 y Bonilla *et. al.*, 2000). En nuestro trabajo son objeto de estudio dos de las aplicaciones que han tenido un mayor desarrollo en la literatura financiera: la predicción del fracaso empresarial y la predicción bursátil. En ambos casos, la tipología de RNA ensayada es la denominada como Perceptrón Multicapa (PMC).

En este artículo contrastamos, en el mercado financiero español, la utilidad de la aplicación de RNAs para un problema de clasificación y otro de predicción financiera. En la primera de ellas, aplicamos una RNA al problema clásico de la predicción de la quiebra, comparando los resultados que ofrece con los derivados de la utilización de la regresión logística (RL) y el análisis discriminante (AD), dos instrumentos usualmente utilizados en este contexto. Por tanto, en esta primera aplicación, el escenario de modelización es uniecuacional, con datos de corte transversal. Nuestro estudio está en la línea de los trabajos de Serrano y Martín (1993) y López y Flórez (2000), donde también se evalúa la utilidad de las RNAs en la predicción de la quiebra a una base de datos de los estados contables de 66 bancos españoles en el periodo 1977-1982, pero tiene algunos aspectos que la diferencian de estos trabajos. Nosotros, para aumentar los grados de libertad de los métodos que evaluamos (RNA, AD y RL), no tomamos un número menor de los 9 ratios contables inicialmente considerados en Serrano y Martín (1993), tal como hacen López y Flórez (2000), sino que son reducidos a un número menor de variables mediante análisis factorial, propuesta acorde con la de West (1985) para un modelo logit. Ello nos permitirá, *de facto*, trabajar con los 9 ratios inicialmente seleccionados.

Otro campo de estudio donde el empleo de las RNAs han encontrado una gran aplicación ha sido la predicción del precio de las acciones, siendo los enfoques normalmente empleados, el del análisis técnico – en esta línea encuentra el trabajo seminal de White (1988)- y el análisis fundamental –un trabajo representativo es Yoon y Swales (1991). El enfoque

que utilizaremos en nuestra aplicación es diferente a los dos anteriores. Nosotros nos basaremos en el marco conceptual que proporcionan los modelos factoriales de valoración de activos tipo Capital Asset Pricing Model (CAPM) de Sharpe (1964) o Arbitrage Pricing Theory (APT) de Ross (1976). En este sentido, Ahmadi (1996) sugiere (aunque no realiza ninguna aplicación empírica) que, dado que la relación rendimiento-factores sea posiblemente no lineal, al contrario de lo que proponen dichos modelos, una forma viable de implementar el APT es mediante el uso de RNAs; sugerencia que puede generalizarse a cualquier otro modelo multifactorial. Nosotros, partiendo de una especificación multifactorial, utilizamos un PMC para la estimación de los rendimientos de las carteras sectoriales en que se dividía la bolsa de Madrid durante el periodo 1990-2001 y comparamos los resultados obtenidos con los que se derivan de la aplicación de la regresión sistemas de ecuaciones lineales aparentemente no relacionadas (SUR)¹, metodología econométrica utilizada comúnmente en la estimación simultánea para varios activos o carteras de este tipo de modelos de valoración. Este segundo caso supone un escenario de modelización diferente, ya que es multiecuacional y se dispone de series temporales.

El resto del trabajo lo hemos estructurado de la siguiente forma. En el apartado que sigue, realizamos un repaso de los antecedentes más relevantes sobre la utilización de las RNAs en los dos problemas que abordamos. Posteriormente, exponemos el funcionamiento de una RNA del tipo PMC, que será el utilizado en nuestras aplicaciones empíricas y algunos aspectos relativos a la estimación y diseño de dicho tipo de RNAs. En los apartados cuarto y quinto analizamos los resultados que nos ofrece la utilización de las RNAs en la predicción de la quiebra y la estimación de rendimientos bursátiles, respectivamente, contrastándolos en ambos casos con los de instrumentos econométricos que habitualmente son utilizados. En el último apartado, señalamos las conclusiones del trabajo realizado que entendemos que son más relevantes.

2. APLICACIÓN DE LAS REDES NEURONALES A LA PREDICCIÓN DE LA QUIEBRA Y LA PREDICCIÓN BURSÁTIL

Desde la publicación de los trabajos de Beaver (1967) y Altman (1968), la predicción de la quiebra empresarial a partir de un conjunto de ratios que emanan de los estados contables ha devenido un tema clásico en la literatura contable. Diversos trabajos han mostrado la utilidad de las RNAs. En el mercado norteamericano, Tam y Kiang (1992) en el sector bancario y Wilson y Sharda (1994) y Rahimian *et. al.* (1996) en el industrial, arrojan evidencias de la superior capacidad predictiva de las RNA tipo PMC respecto a instrumentos más clásicos como el AD, la RL y los

(1) La abreviatura comúnmente utilizada, SUR, corresponde al nombre inglés *Seemingly Unrelated Regression*.

árboles de decisión. No obstante, Altman *et. al.* (1994), tomando una base de datos compuesta por 1000 PYMES italianas, si bien, vuelven a aportar una evidencia sobre la superioridad predictiva de las RNAs sobre el AD, ésta no es tan clara. Así, consideran que las RNAs deben ser consideradas como un instrumento complementario a los métodos más tradicionales, ya que éstos, al contrario que las RNAs, permiten no sólo realizar predicciones relativamente fiables, sino también determinar qué ratios son los más determinantes para predecir la quiebra en empresas de una determinada tipología.

En el ámbito español, es posiblemente la crisis bancaria del periodo 1977-1985, que afectó a un gran número de entidades de crédito, el suceso más estudiado desde la perspectiva de los modelos de predicción de la quiebra. Así, son remarcables los trabajos de Laffarga *et. al.* (1986) y Pina (1989), realizados con instrumentos multivariantes convencionales. Tomando como referencia este suceso y la base de datos utilizada en dichos trabajos y como variables explicativas, los mismos 9 ratios que Pina (1989), Serrano y Martín (1993) diseñan un modelo neuronal que combina el uso de PMCs y redes autoorganizadas. Si bien, los resultados que obtienen son satisfactorios, dado que logran clasificar correctamente a los bancos en un 94% de ocasiones, no ofrecen una comparación de los resultados que se obtendrían en el caso de utilizar otros métodos. No obstante, afirman que han constatado previamente que los resultados que ofrecen las RNAs son superiores a los que ofrece una RL. López y Flórez (2000), utilizando aproximadamente la misma base de datos (de la muestra original de 66 bancos, únicamente toman 52, la mitad de los cuales quebraron) observan mejores resultados en las predicciones efectuadas por una RNA del tipo PMC que las que realiza la RL y el AD. Asimismo, únicamente utilizan 6 ratios contables para clasificar las entidades bancarias en solventes y fracasadas.

Otro campo donde las redes neuronales han encontrado una intensiva aplicación es la predicción bursátil. Así, podemos encontrar dos tipos básicos de enfoques. En primer lugar, el basado en el análisis técnico. En este caso, se busca predecir los rendimientos y, por tanto, generar estrategias de negociación, a partir de los precios que en el pasado han tomado las acciones. La otra perspectiva utilizada por la literatura es el análisis fundamental, basado en la utilización de información contable.

Dentro del primer enfoque podemos mencionar a White (1988). Éste muestra la capacidad de las RNAs para predecir la rentabilidad de acciones (concretamente de IBM), a partir de los rendimientos pasados, comparando los resultados obtenidos con los de un modelo de series temporales autoregresivo. Mientras que el modelo autoregresivo no formula predicciones estadísticamente significativas, la RNA sí lo hace. Así, muestra que el análisis técnico es relevante, en contra de lo que indica la hipótesis del mercado eficiente en cualquiera de sus formas; y que la relación intertemporal de los rendimientos tiene un fuerte carácter no lineal que puede ser captado por las RNAs y no mediante un modelo de series temporales al uso. En esta línea, Robles y Naylor (1996) en el mercado bursátil de Chile y Yao y Poh (1996) en el de Malasia, muestran que a partir de PMCs (o en el caso de Robles y Naylor (1996), versiones ligeramente más

sofisticadas como las redes de Elman), es posible generar estrategias más eficientes que la simple de comprar y mantener, tomando como variables predictivas diversos osciladores técnicos.

En el mercado español, Fernández *et. al.* (2000) muestran la capacidad predictiva de los PMCs en el Índice General de la Bolsa de Madrid (IGBM). Las estrategias generadas por las RNAs (comprar si el índice sube y vender si éste baja), en ausencia de costes de transacción, se muestran más eficientes que las de comprar y mantener en escenarios estables o bajistas de la bolsa, si bien ello no se manifiesta en el episodio alcista analizado. Por otra parte, Pérez y Torras (2001) contrastan, en el IBEX-35, la capacidad predictiva de diversos modelos no lineales econométricos de series temporales (series del tipo ARCH, LSTAR, etc.) y de diversas arquitecturas de RNAs (PMCs, redes de Elman y redes de Jordan). Observan que es posible efectuar predicciones del IBEX-35 a partir de su comportamiento pasado, mostrándose las RNAs como un instrumento de predicción sensiblemente superior.

Utilizando como enfoque el análisis bursátil fundamental, Yoon y Swales (1991) parten de la opinión de analistas sobre la dirección de la empresa y sus estados contables para realizar predicciones sobre el rendimiento de las acciones en 10 sectores industriales diferentes. Observan una mayor capacidad predictiva de los PMCs que en el AD. Harvey *et. al.* (2000), utilizando como variables predictivas, los ratios PER y book-to-market y la rentabilidad por dividendo observan, en diferentes mercados emergentes, que los PMCs generan estrategias de compra y venta de acciones más eficientes que las efectuadas por una regresión lineal o una estrategia de comprar y mantener. Kanas (2001) compara, en el mercado inglés y alemán, el comportamiento de los PMCs y de la regresión lineal en la estimación de rendimientos a partir de la rentabilidad media por dividendo del mercado y del volumen negociado en anteriores periodos, con un resultado favorable a las RNAs.

Finalmente, debemos remarcar que la aportación de los instrumentos de inteligencia artificial a la predicción de la quiebra y de los rendimientos bursátiles, no se limita al uso de las RNAs. Respecto a la predicción de la quiebra, han sido ensayados diferentes métodos, los cuales presentan, en general, unos resultados prometedores. Así, Matsatsinis *et. al.* (1997) muestran la utilidad de los sistemas expertos y Kingdon y Feldman (1993) observan un buen comportamiento de los algoritmos genéticos. Por otra parte, Dimitras *et. al.* (1999), utilizando instrumentos de la teoría de conjuntos toscos (*rough sets*) y Matsatsinis *et. al.* (2003) con un sistema experto borroso, obtienen una mayor eficiencia predictiva que con métodos multivariantes convencionales.

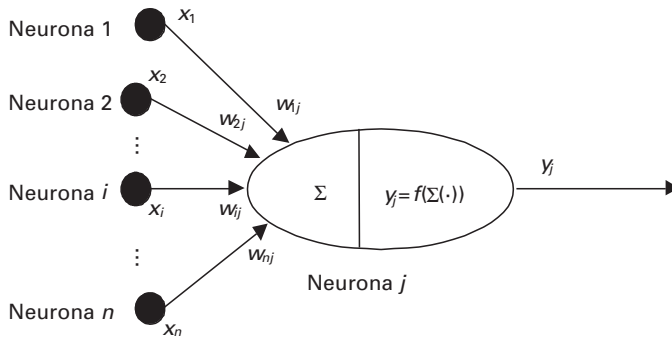
En un contexto de predicción bursátil, la mayor parte de trabajos que utilizan instrumentos de inteligencia artificial, suelen partir, de nuevo, del marco analítico que ofrecen el análisis fundamental y/o el análisis técnico. Así, John *et. al.* (1996) y Trippi y Lee (1996) muestran la utilidad de los sistemas expertos; mientras que Mahfoud y Mani (1996) obtienen buenos resultados con la utilización de algoritmos genéticos. Por otra parte, Spanos *et. al.* (2000) contrastan, con relativo éxito, la capacidad predictiva de

los sistemas expertos borrosos en el mercado de Atenas y Lam *et al.* (1996) diseñan un sistema borroso experto basado en indicadores técnicos. Algunos autores proponen combinar más de un instrumento de inteligencia artificial en este problema, como Wong *et al.* (1993), que diseñan un sistema experto borroso apoyado por una RNA y Kuo *et al.* (2001), que proponen un modelo que combina los algoritmos genéticos, los sistemas borrosos y las redes neuronales.

3. REDES NEURONALES ARTIFICIALES: LOS PERCEPTRONES MULTICAPA

Las RNAs son instrumentos de procesamiento de datos con una estructura y funciones que imitan al cerebro humano. De forma muy simplificada, se puede conceptualizar una RNA como la conexión de diversas unidades básicas de procesamiento de información, las *neuronas*, mediante tres tipos de capas: de entrada, ocultas y de salida. Una neurona recibe la información que le suministran las neuronas de la capa precedente, que vendrá dada por la combinación lineal de la información proveniente de cada una ponderada por unos coeficientes denominados *pesos sinápticos*. Asimismo, una neurona proporciona una salida hacia las neuronas de capas posteriores, que se hallan con una *función de transferencia o activación* –las más usuales pueden consultarse en Martín y Sanz (1997, p.18). Ello viene representado en el gráfico 1, donde x_i representa la salida de la i -ésima neurona precedente, w_{ij} el peso sináptico que liga a dicha neurona con la que es objeto de representación, la neurona j -ésima y $f(\cdot)$ la función de activación, que determina la salida de la j -ésima neurona (y_j) hacia el resto de neuronas.

Gráfico 1
ARQUITECTURA Y FUNCIONAMIENTO DE UNA NEURONA DE UNA RNA

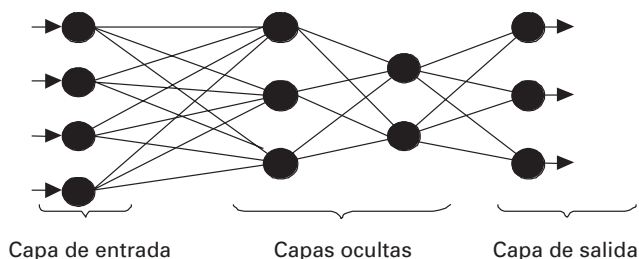


Debemos remarcar que existe una gran variedad de RNAs, dependiendo de su arquitectura, la forma en que se determinan los pesos sinápticos, denominada *entrenamiento*, y de la habilidad del cerebro humano

imitada. Una amplia panorámica sobre los diferentes tipos de RNAs puede encontrarse en Kuan y White (1994) y Martín y Sanz (1997). Los primeros centran su exposición en los PMCs o versiones ligeramente más sofisticadas como las redes de Elman y las de Jordan, poniendo énfasis en su similitud y las interacciones que pueden observarse con los modelos de regresión no lineal; y su capacidad aproximativa de procesos dinámicos no lineales. En cambio, Martín y Sanz (1997) realizan un análisis descriptivo, desde una perspectiva "convencional" en la literatura sobre RNAs, de una amplia variedad de tipologías (PMC's, mapas autoorganizados, redes de Hopfield, etc.).

El tipo de RNA que empleamos en este trabajo es el Perceptrón Multicapa (PMC), por otra parte, muy utilizado en aplicaciones financieras. Un PMC es una RNA que imita la capacidad humana de generalizar la información suministrada en una muestra compuesta por inputs y outputs y asociar, a un input no proporcionado anteriormente, un output. Este tipo de redes ajustan los pesos mediante entrenamiento² *supervisado*. A la red se le indica, para cada input suministrado, cuál es el output deseado, buscándose minimizar una función de error (normalmente, el error cuadrático) entre el output real y el output que ofrece la red.

Gráfico 2
ESTRUCTURA DE UN PMC



En el gráfico 2 puede observarse que un PMC consta de una capa de entrada, a la cual se suministra las observaciones de las variables independientes. Esta capa está conectada a un grupo de capas denominadas como capas ocultas (dos en el gráfico 2); las cuales, están unidas con la capa de salida, que proporcionará el valor estimado para las variables endógenas. La entrada a una neurona j , net_j , viene dada por:

(2) Por el contrario, la RNAs que ajustan los pesos con aprendizaje no supervisado no necesitan conocer los outputs deseados, de tal forma que organizan la información presentada a partir de las regularidades que se observan en los datos de entrada. Un ejemplo típico de estas redes son los mapas autoorganizados, muy útiles en la agrupación (*clustering*) de datos.

$$net_j = \sum_i w_{ij} x_i + \theta_j$$

donde θ_j es término conocido como umbral (*bias*) de la neurona j . Asimismo, para la salida de la neurona j , y_j , la función de transferencia más utilizada es la función sigmoideal:

$$y_j = \frac{1}{1 + \exp(-net_j)}$$

Los pesos y los umbrales, que deben minimizar el error cuadrático de las salidas que proporciona la red, se hallan mediante el algoritmo *back-propagation* descrito en Rummelhart y McClelland (1986), basado en la optimización no lineal mediante el descenso del gradiente, que puede considerarse, tal como indican Pérez *et. al.* (2000), como una versión recursiva de la técnica econométrica de mínimos cuadrados no lineales.

El número de neuronas que deben existir en cada capa oculta (que determinará el número de parámetros a estimar) y el número de capas ocultas, deberá ser suficiente para que la red aprenda los ejemplos de la muestra de entrenamiento. No obstante, su número no debe ser excesivo, ya que se puede producir un *sobreparametrizaje* que implique que, aunque la red ajusta muy bien la muestra de entrenamiento, no es capaz de generalizar el conocimiento adquirido, proporcionando muy malas estimaciones cuando se la proporciona inputs no suministrados durante el aprendizaje. Desgraciadamente, no existe un método claramente establecido para determinar la dimensión óptima de una red neuronal (número de capas, número de pesos, etc.) que proporcione el equilibrio óptimo entre el ajuste a la muestra de entrenamiento y la capacidad de generalización; aunque sí que existen criterios heurísticos y procedimientos de carácter empírico que pueden ser de ayuda.

Un criterio que sirve de guía para establecer la dimensión de la red, es la regla de Baum. Ésta indica que el error medio de generalización de un PMC es directamente proporcional al número de parámetros que deben ser ajustados e inversamente proporcional al número de ejemplos de entrenamiento. Respecto al número de capas ocultas, tal como apuntan Martín y Sanz (1997, p.216), no es común diseñar redes con más de dos. En cualquier caso, tal como muestran analíticamente Hornick *et. al.* (1989), un PMC con una única capa oculta constituye, en efecto, un aproximador universal de funciones. Por otra parte, algunos procedimientos muy utilizados en la determinación final de la red óptima consisten en establecer pesos comunes para grupos de neuronas, el "podado" de la red y el decaimiento de los pesos. El objetivo final de estas técnicas consiste en mejorar la capacidad de generalización de la red sin menoscabo del ajuste obtenido con la muestra de entrenamiento. Una explicación detallada de la implementación de estos procedimientos puede encontrarse en Martín y Sanz (1997, p. 75).

Por otra parte, Refenes (1995) aconseja realizar un análisis de sensibilidad, tanto en la fase de entrenamiento como en la de testeo para determinar cuál es la arquitectura final. También apunta la conveniencia de manipular el ratio de aprendizaje de la red (velocidad con que se produce

la convergencia de la red hacia el error mínimo), a medida que transcurre el proceso de ajuste de los parámetros. Se debe elegir aquella arquitectura que requiera ajustar un menor número de parámetros, si ofrece un error de ajuste en el entrenamiento similar a otras de mayor dimensión.

Finalmente, tal como indican Adya y Collopy (1998), a la hora de establecer si el ajuste proporcionado por una RNA es satisfactorio o no, es necesario comparar los resultados que ofrece con los que se obtienen con las técnicas que habitualmente se utilizan en el problema de predicción analizado. En este sentido, deben ser tenidas en cuenta, tanto la calidad del ajuste obtenido dentro de la muestra utilizada para el ajuste de los parámetros de los modelos como, sobre todo, la calidad de las predicciones efectuadas por los modelos en la submuestra utilizada para medir su capacidad de generalización; ya que la utilidad esencial de las RNAs es su potencia para realizar predicciones.

4. PREDICCIÓN DE LA QUIEBRA CON REDES NEURONALES ARTIFICIALES

En este epígrafe, contrastamos la capacidad predictiva de la RNA tipo PMC con dos de los instrumentos utilizados habitualmente en el problema de la predicción de la quiebra: la RL y el AD; siendo el hecho estudiado, la crisis bancaria española producida a principios de los años 80, problema ya abordado desde nuestra perspectiva en Serrano y Martín (1993) y López y Flórez (2000). No obstante, a diferencia de los trabajos anteriores, reducimos el conjunto de ratios seleccionados (los 9 utilizados en los trabajos de Pina, 1989 y Serrano y Martín, 2000) mediante el análisis factorial, lo cual, como ya anticipamos, permitirá obtener modelos con menos parámetros a ajustar, lo cual es especialmente interesante en esta aplicación, dado el pequeño tamaño de la muestra disponible.

4.1. Base de datos y metodología

La muestra analizada comprende los 66 bancos utilizados en los estudios de Serrano y Martín (1993) y Laffarga *et. al.* (1986), entre los que había 29 quebrados y 37 no quebrados durante el periodo 1977-1982. Los 9 ratios contables de los que partiremos para realizar las predicciones son los que se muestran en el cuadro 1, ya utilizados por Pina (1989) y Serrano y Martín (1993).

Cuadro 1
RATIOS EMPLEADOS POR PINA (1989)
Y SERRANO Y MARTÍN (1993)

Activo Circulante/Activo Total (R1)	Beneficio Neto/Neto (R6)
(Activo Circulante-Tesorería)/Activo Total (R2)	Beneficio Neto/Pasivo Exigible (R7)
Activo Circulante/Pasivo Exigible (R3)	Coste de Ventas/Ventas Brutas (R8)
Reservas/Pasivo Exigible (R4)	Cash-Flow/Pasivo Exigible (R9)
Beneficio Neto/Activo Total (R5)	

Dado que la muestra utilizada no es excesivamente amplia, optamos por reducir el número de las variables explicativas utilizando el análisis factorial. Los resultados obtenidos vienen dados en el cuadro 2.

Como variables explicativas hemos tomado finalmente los tres factores que presentan un valor propio mayor que 1 (los únicos que aparecen en el cuadro 2), ya que éstos explican el 93,39% de la varianza de los 9 ratios. Se puede comprobar que aunque el coeficiente de Kaiser-Meyer-Olkin no indica que el análisis factorial sea especialmente adecuado, tampoco lo rechaza; y, asimismo, el test de Barlett rechaza la hipótesis de que la matriz de correlaciones sea la unidad. También se muestran las cargas factoriales de los ratios con un nivel de significación de al menos el 95% en cada factor, por lo que se muestran aquellas que presentan un valor absoluto superior al 0,5. Asimismo, hemos empleado la rotación varimax en la determinación de los factores. Se puede comprobar que los tres factores tienen una fácil interpretación contable. El primero (F1) está relacionado con la rentabilidad y la productividad de la entidad, el segundo (F2) con su liquidez, mientras que el tercero (F3) refleja el grado de autofinanciación del banco.

Cuadro 2
RESULTADOS DEL ANÁLISIS FACTORIAL

Kaiser-Meyer-Olkin (KMO)=0,667

Test de esfericidad de Barlett: $\chi^2=1425,86$. Sign: 0,000%

Ratios	Factor 1 (F1)	Factor 2 (F2)	Factor 3 (F3)
R1		0,965	
R2		0,888	
R3		0,942	
R4			0,938
R5	0,977		
R6	0,962		
R7	0,975		
R8	-0,908		
R9	0,898		

El hecho de que los 3 factores seleccionados contengan un porcentaje realmente elevado de la varianza de las variables explicativas, así como el hecho de que éstos presenten una fácil interpretación contable, reafirman que es adecuado utilizar dichos factores en la realización de los ajustes de los modelos neuronales y estadísticos en lugar de los ratios inicialmente seleccionados. Además, debemos tener en cuenta el bajo tamaño de la muestra de que partimos. De esta forma, el número de variables explicativas (o inputs en terminología neuronal) queda reducido notablemente respecto a trabajos anteriores: en Serrano y Martín se utilizaban los 9 ratios del cuadro 1, mientras que López y Flórez (2000) utilizaban los ratios R1, R4 y R8 del cuadro 1, a los que añadían: Ventas/Activo total, Reservas/Neto y Tesorería/Activo Circulante.

Para contrastar la capacidad predictiva del PMC respecto a los métodos más convencionales, hemos dividido nuestra muestra en dos submuestras. La primera comprende 40 bancos, de los cuales, la mitad quebraron, y todos fueron seleccionados al azar. Con ésta entrenamos a la RNA y estimamos los parámetros de las funciones logística y discriminante. Posteriormente, con la segunda submuestra, que consta de 26 entidades (9 quebradas y 17 solventes), predecimos su grupo de pertenencia a partir de los parámetros estimados con anterioridad, contrastándose así la capacidad de extrapolación de cada metodología.

4.2. Resultados

Los resultados del ajuste de los modelos discriminante y logit vienen dados en el cuadro 3. Así, el modelo discriminante propuesto puede escribirse como:

$$Z_i = \alpha_0 + \alpha_1 F1_i + \alpha_2 F2_i + \alpha_3 F3_i \quad (1)$$

Donde Z_i es la puntuación o "Z-score" otorgada a la entidad de crédito i -ésima en función de los valores que toma su nivel de rentabilidad ($F1_i$), su liquidez ($F2_i$) y autofinanciación ($F3_i$); mientras que α_j , $j=0,1,2,3$, son los parámetros a estimar. Éstos han sido ajustados mediante el procedimiento de Fisher, que busca maximizar la variabilidad de la puntuación Z entre los grupos (2 en esta aplicación) respecto a la variabilidad del Z-score dentro de los grupos. Dado que el valor de Z_i -Z-score del i -ésimo banco - es una variable de carácter continuo, es preciso determinar el punto de corte, Z^* , a partir del cual la puntuación obtenida por una empresa lleva a clasificarla como solvente o quebrada. Para determinar dicho punto de corte partimos de una puntuación Z media correspondiente a cada grupo (centroide). Dado que se ha utilizado una muestra de estimación con el mismo número de empresas quebradas y solventes (20 en cada caso), Z^* se determina sin más que hallar la media aritmética de los centroides, que es 0. De esta forma, si $Z > 0$, clasificaríamos a la i -ésima entidad de crédito como solvente; mientras que en caso contrario, la empresa sería clasificada como quebrada.

Puede observarse en el cuadro 3 que el modelo discriminante especificado es aceptable, dado que tal como indica la lambda de Wilks³ de todo el modelo y el contraste χ^2 asociado, la diferencia del Z-score medio de los dos grupos es significativa a un nivel de inferior al 1%. Por otra parte, el signo de los coeficientes estimados para (1) indica la existencia de una relación positiva entre la salud de una entidad de crédito

(3) La lambda de Wilks es un estadístico cuyo rango de variación es el intervalo [0,1] y se utiliza para contrastar, en el marco del análisis discriminante, si la media de las puntuaciones Z , o del valor de las variables utilizadas para "discriminar" es significativamente diferente. Un valor cercano a 1 indicaría la inexistencia de diferencias significativas entre grupos mientras que su alejamiento de este valor indicaría una mayor diferencia entre éstos.

y su rentabilidad y autofinanciación; y una relación negativa de la liquidez de la entidad bancaria con su supervivencia. No obstante, respecto al factor liquidez, puede observarse que su lambda de Wilks es prácticamente 1, por lo que el valor de su estadístico χ^2 indica que la diferencia del mismo entre grupos no es significativa. Respecto a los otros dos factores, F1 y F3, su nivel de significación para explicar las diferencias entre grupos se sitúa por debajo del 10%, lo que nos permite no rechazar a dichos factores como significativos en la explicación de la quiebra.

Cuadro 3
RESULTADOS DE LA ESTIMACIÓN DISCRIMINANTE Y LOGIT

ANÁLISIS DISCRIMINANTE			REGRESIÓN LOGIT		
Función discriminante canónica no tipificada: $Z = 0,027 + 1,181 \cdot F1 - 0,039 \cdot F2 + 0,321 \cdot F3$			R ² de Cox y Snell: 0,646 R ² de Nagelkerke: 0,861		
Centroides canónicos no tipificados. Quebradas: -0'716. Solventes: 0'716			Significación global del modelo: Chi-Cuadrado: 41,519 (0,000)		
Signific. Del modelo: - Lambda de Wilkins: 0,649 - Chi-Cuadrado equivalente 15,752 (0,001)			Análisis de las variables explicativas		
Análisis ANOVA de los factores			Factor	Coficiente	Chi-Cuadrado
Factor	Lambda de Wilks	Estadístico F	Constante		
F1	0,681	17,783 (0,000)	F1	-2,0441	2,739 (0,098)
F2	1	0,009 (0,926)	F2	21,2012	3,670 (0,055)
F3	0,928	2,935 (0,095)	F3	-1,9609	1,435 (0,231)
				0,5939	0,152 (0,697)

Nota: Entre paréntesis, el nivel de significación

Por otra parte, el modelo logit especificado es:

$$P(Y_i=1) = \frac{e^{\alpha_0 + \alpha_1 F1_i + \alpha_2 F2_i + \alpha_3 F3_i}}{1 + e^{\alpha_0 + \alpha_1 F1_i + \alpha_2 F2_i + \alpha_3 F3_i}} \quad (2)$$

donde $P(Y_i=1)$ denota la probabilidad de que el banco i -ésimo pertenezca, dadas sus puntuaciones en los factores F1, F2 y F3, al grupo de bancos solventes. Por tanto, tal como es habitual en la práctica, tomaremos como probabilidad de corte para clasificar a un banco como solvente $P(Y_i=1) > 1/2$. Los coeficientes del modelo (2) han sido estimados con el método de máximo-verosimilitud, utilizando un algoritmo de introducción. Si el coeficiente asociado a un factor es positivo, ello será indicativo de que un incremento de dicho factor producirá un aumento en la probabilidad de supervivencia de la empresa; mientras que si el coeficiente fuera negativo, un aumento de dicho factor conllevaría un aumento de la probabilidad de fracaso.

En el cuadro 3 se ofrece el resultado obtenido con la estimación de la RL. El valor obtenido con los coeficientes de determinación de Cox y Snell

y de Nagelkerke⁴, que son los habitualmente utilizados para evaluar la capacidad de ajuste de una RL, indican que el ajuste global del modelo ha sido bastante satisfactorio. Ello es confirmado por el valor del estadístico χ^2 , que señala que el ajuste global es significativo para niveles inferiores al 1%. La interpretación del signo de los coeficientes estimados para las variables F1, F2 y F3 es análogo al que podemos deducir de la función discriminante ajustada. La relación entre la autofinanciación y la rentabilidad es positiva respecto a la probabilidad de supervivencia del banco, mientras que la probabilidad de supervivencia y la liquidez se relacionan de forma inversa. No obstante, el test de Wald sugiere que únicamente es un factor significativo para explicar la quiebra de un banco F1 (asociado con su rentabilidad y productividad).

A la hora de aplicar los modelos seleccionados, tanto en la muestra de entrenamiento como en la de testeo, hemos optado por mantener dentro de los modelos discriminante y logit las variables explicativas que no se mostraban estadísticamente significativas con el coeficiente que ha sido estimado para las mismas. Entendemos que la muestra utilizada en el ajuste de los parámetros es bastante pequeña, por lo que los tests de significación efectuados adolecen de potencia de contraste. En segundo lugar, nuestro principal interés se centra en contrastar la capacidad de ajuste a la muestra y la capacidad predictiva de las metodologías ensayadas. En este sentido, nos parece más adecuado partir de los coeficientes que proporcionan un mejor ajuste a la muestra, cuyos valores vienen dados en el cuadro 3.

La RNA que finalmente hemos empleado consta de una capa de entrada con 3 neuronas (una para cada factor) y una neurona en la capa de salida, que señala si la empresa quebrará, proporcionando una salida de 0; o tendrá éxito, tomando entonces como valor 1. La capa oculta finalmente utilizada consta de 4 neuronas, mientras que el número de épocas utilizadas en el entrenamiento de la RNA se situaba entorno a las 7.500, ya que a partir de entonces, la red disminuía muy lentamente su capacidad de aprendizaje. En la RNA, dado que la neurona de salida proporciona valores entre 0 y 1, la puntuación de corte elegida es, como en la RL, 1/2, de tal forma que si la salida fuera superior a 1/2, clasificaríamos al banco dentro del grupo de los que no quebraron. Respecto a las neuronas de la capa de entrada asociados a la liquidez y la autofinanciación, sobre los que había dudas en el AD y la RL de su significación estadística, hemos observado que varios de los pesos que las conectan con las neuronas de la capa oculta difieren notablemente de 0, lo que puede ser indicativo de la conveniencia de seguir utilizando también estas variables en la predicción de la quiebra. Asimismo, ello podría indicar que los métodos estadísticos ensayados han captado con una

(4) La interpretación de dichos coeficientes es análoga a la del coeficiente de determinación convencional. Su construcción, en cambio, es diferente. Ésta se basa en la razón de los valores que toma la función de verosimilitud cuando el valor de los coeficientes de las variables explicativas es 0 y su valor máximo-verosímil.

menor precisión que la RNA el carácter de la relación entre F2 y F3 y la contingencia de quiebra.

En el cuadro 4 se muestran los resultados de la clasificación que realizan las funciones discriminante y logit y el PMC, en la muestra de ajuste y en la de predicción. Tanto en la muestra de entrenamiento como en la muestra de comprobación puede observarse que el porcentaje de aciertos de la RNA es superior al de los dos métodos tradicionales. En la muestra de estimación, la RNA ofrece una fiabilidad del 100%, mientras que la regresión logit acierta el 95% de casos y el análisis discriminante, el 92,5%. En la muestra de comprobación, la RNA se continúa manteniendo con un porcentaje de aciertos más elevado que los otros métodos, aunque en todos los casos, la fiabilidad disminuye. No obstante, la RNA sigue proporcionando un error del tipo A del 0%, mientras que el porcentaje de errores del tipo B se sitúa en un nivel idéntico al de los métodos multivariantes clásicos.

Cuadro 4
RESULTADOS DE LA CLASIFICACIÓN DE LOS MÉTODOS ENSAYADOS

	AD			RL			RNA		
	Acierto	Error A	Error B	Acierto	Error A	Error B	Acierto	Error A	Error B
Muestra de ajuste (40 bancos)	92,5%	5%	10%	95%	5%	5%	100%	0%	0%
Muestra test (26 bancos)	76,92%	22,33%	23,53%	76,92%	22,33%	23,53%	84,61%	0%	23,53%

Nota: Error A corresponde a clasificar una empresa quebrada como sana, mientras que el error B corresponde a clasificar una empresa sana como quebrada. Aunque la literatura referente a la predicción de la quiebra normalmente los denomina como "error I" y "error II" respectivamente, hemos preferido no mantener esta nomenclatura para evitar la confusión que ello puede producir respecto al concepto de "error del tipo I" y "error del tipo II" en el contraste de hipótesis estadísticas.

Debemos mencionar que nuestros resultados son similares a los obtenidos en Serrano y Martín (1993) y López y Flórez (2000). Serrano y Martín (1993) obtienen un éxito con el PMC en las predicciones de la muestra test del 89% y un éxito en el global de la muestra del 94%, si bien, su muestra de ajuste es ligeramente inferior (toman 32 bancos). Por su parte, la red ajustada por López y Flórez (2000) muestra un éxito de un 89% en la muestra de testeo y un 96% para toda la muestra que utilizan (estimación y testeo), que se compone únicamente por 52 bancos (no por los 66 nuestros) y también encuentran evidencias de la superioridad predictiva de la RNA respecto al AD y la RL. La ligera discrepancia entre nuestros resultados y los obtenidos en trabajos anteriores puede ser imputable a que hayan sido utilizados diferentes bancos a los nuestros en las fases de entrenamiento/test (ya que éstos, en todos los casos, son seleccionados aleatoriamente) de tal forma que, en nuestra muestra de testeo, hayamos incluido los bancos con mayor dificultad para ser clasificados, mientras que éstos hayan sido incluidos en la muestra de entrenamiento en los

anteriores trabajos. Respecto al trabajo de López y Flórez (2000), también podría ocurrir que alguna de las entidades de crédito que nuestra red no ha sido capaz de clasificar correctamente no haya sido incluida en su muestra.

En cualquier caso, debemos remarcar el hecho de que nuestra RNA presenta una menor dimensión y obtiene unos resultados muy similares. De hecho, según la regla de Baum, el error de generalización medio de nuestra RNA debe ser ostensiblemente menor. Mientras que las neuronas de entrada tomadas en Serrano y Martín (1993) eran 9, considerándose una capa oculta con 10 neuronas, en López y Flórez (2000) la RNA utilizada constaba de 6 neuronas en la capa de entrada y 4 en la capa oculta.

5. ESTIMACIÓN MULTIFACTORIAL DEL RENDIMIENTO DE LAS CARTERAS DE VALORES CON REDES NEURONALES

Los modelos multifactoriales de valoración de activos tipo CAPM –Sharpe (1964)– o APT –Ross (1976) postulan la existencia de una relación lineal entre el rendimiento de los activos y la evolución de diferentes variables financieras y macroeconómicas. Entre las variables financieras podemos mencionar el rendimiento de una cartera de activos representativa del mercado bursátil (en el caso del CAPM es la única contemplada), determinados ratios fundamentales, etc. Respecto a las variables macroeconómicas, son especialmente utilizadas la variación no esperada de la inflación y el crecimiento de la actividad económica.

Es indudable que estos modelos presentan varios atractivos. En primer lugar, los parámetros implicados, las betas, que miden la sensibilidad del precio de un título ante variaciones de los factores, pueden ser obtenidos a partir de datos históricos mediante regresión lineal y es fácil establecer su capacidad explicativa. Por otra parte, tras determinarse la relación lineal rendimiento-factores, la proyección del rendimiento futuro de los títulos se reduce a la predicción de unas pocas variables que pueden ser estimadas con cierto grado de verosimilitud en horizontes temporales cortos.

Por otra parte, si aceptamos que el valor de una acción es el valor descontado de los dividendos que proporcionará en el futuro, deberemos concluir que los factores relevantes en la formación de su precio (y, por tanto, desde un punto de vista dinámico, de los rendimientos) son todos aquéllos que repercutan en los dividendos y la tasa de descuento. No obstante, la relación de estos factores con los precios es compleja y de difícil precisión analítica. También debe remarcarse que es un hecho bien conocido que las betas no se mantienen invariables en el tiempo, lo cual dificulta la proyección futura de los rendimientos a partir de los coeficientes obtenidos previamente. Tampoco debemos olvidar que entre dichos factores existen múltiples interacciones difíciles de cuantificar analíticamente. Por estas razones, Ahmadi (1996) sugiere que una forma viable de implementar el APT es el uso de RNAs, sugerencia que creemos que puede generalizarse a cualquier otra formulación factorial. En este apar-

tado exploramos las posibles ventajas de la estimación de un modelo multifactorial mediante RNAs, respecto a la utilización de los instrumentos de regresión lineal.

5.1. Base de datos y metodología

El objeto de análisis es el rendimiento mensual de las carteras sectoriales de la bolsa de Madrid durante el periodo 1990-2001 (144 observaciones). En este caso, entendemos que la muestra empleada, aunque no excesivamente elevada, tiene un tamaño al menos, bastante aceptable respecto a la utilizada en el caso de la predicción de la quiebra. En el cuadro 5 recogemos estos sectores⁵ y la abreviatura que se les asigna. Como indicadores del comportamiento de estas carteras, hemos tomado el rendimiento de los índices sectoriales de la bolsa de Madrid, medido como tanto efectivo mensual. Todos los datos han sido obtenidos de las series temporales del banco de España.

Cuadro 5
SECTORES DE LA BOLSA DE MADRID DE 1990-2001

Número (j)	Abreviatura	Sector
1	BAN	Bancos y Financieras
2	ELEC	Eléctricas
3	AL	Alimentación
4	CON	Construcción
5	CAR	Cartera e Inversión
6	MET	Metal-Mecánica
7	PETR	Petroquímico
8	COM	Comunicaciones
9	OTR	Otras Industrias y Servicios

Debemos remarcar que, en los modelos multifactoriales, la determinación de las variables conducentes a dar una explicación de las rentabilidades bursátiles pasa por elegir aquéllas que, de manera razonable o por investigaciones previas, puedan explicar dichas rentabilidades. Nosotros partimos de que el rendimiento de las carteras puede explicarse a partir de las siguientes variables:

- El rendimiento de la cartera de mercado (RM), que hemos supuesto que es la correspondiente al Índice General de la Bolsa de Madrid.

(5) En el año 2002 la bolsa de Madrid ha cambiado el número de sectores para los que elabora los índices, así como su composición y su nombre. Más detalles de dicho cambio pueden consultarse en la web de la bolsa de Madrid <<<http://www.BolsaMadrid.es>>>.

- El rendimiento de los títulos de Deuda Pública a más de dos años (DPMLP). Tal como apunta Iglesias (2000), cuando los tipos de interés de la deuda pública aumentan, la demanda de acciones disminuye, por lo que es una variable relevante en la formación de los precios bursátiles.

- La prima de riesgo que incorporan los títulos de renta fija con un vencimiento a más de dos años de las emisiones corporativas respecto al tipo de interés de la deuda pública con dicho vencimiento (PR). Reflejará, siguiendo a Ferson y Harvey (1991), la probabilidad de quiebra percibida por los inversores.

- La variación de la pendiente de la ETTI desde el mes anterior hasta el mes objeto de análisis (Δ PEND). Dicha pendiente, tal como proponen Ferson y Harvey (1991), la medimos por la diferencia relativa del rendimiento de la deuda pública con vencimiento en el medio y largo plazo (en nuestro caso, a más de dos años) y el rendimiento de los repos a 1 mes. Las variaciones en la pendiente de la ETTI anticipan los ciclos económicos, tal como muestra Harvey (1988, 1991) y Rico (2001) en España.

- La variación no esperada del IPC (Δ IPC), que medimos como la diferencia entre el IPC del mes anterior y el IPC del mes siguiente, tal como proponen Pérez y Aranda (2000), de forma que aceptamos, implícitamente, la Teoría de las Expectativas Adaptativas. Un incremento de la inflación supone un aumento del tipo de interés con el que se descuentan los dividendos. Por otra parte, un incremento de la inflación puede ser debido a un aumento de la demanda agregada, con el consiguiente incremento de los beneficios empresariales que ello supone. Así, se pone de manifiesto la relevancia, al menos a priori, de esta variable, lo cual justifica su inclusión habitual en los trabajos empíricos. No obstante, los comentarios anteriores nos llevan a remarcar que el signo que debe presentar la relación rendimiento-inflación no es fácil de precisar a priori.

- La variación en términos relativos del Índice de Producción Industrial (Δ IPI), que siguiendo a Pérez y Aranda (2000), hemos tomado como proxy para cuantificar el crecimiento económico. Dada la fuerte estacionalidad que presenta este índice en el mes de agosto, como valor del IPI en un determinado momento hemos tomado su media móvil de orden 12. El crecimiento económico estimula los beneficios empresariales y, por tanto, supone un estímulo para el aumento del precio de las acciones; por lo que en principio, debe ser también una variable relevante.

En la determinación de la relación rendimiento-factores hemos utilizado dos métodos:

- La utilización de la regresión de sistemas de ecuaciones aparentemente no relacionadas (SUR) estimada por mínimos cuadrados generalizados factibles, también conocido como estimador de Zellner, en lugar de estimar el modelo individualmente en cada cartera. En este caso, el modelo a estimar es:

$$R_{j,t} = \alpha_{0,j} + \alpha_{1,j} RM_t + \alpha_{2,j} DPMLP_t + \alpha_{3,j} PR_t + \alpha_{4,j} \Delta PEND_t + \alpha_{5,j} \Delta IPC_t + \alpha_{6,j} \Delta IPI_t + \varepsilon_{j,tr} \quad j=1,2,\dots,9 \quad (3)$$

El subíndice j denota el sector al que se refiere la ecuación (3) –ver cuadro 5–, mientras que t denota el número de observación al que hacemos referencia y $\varepsilon_{j,t}$ es el término de error asociado a la t -ésima observación del j -ésimo sector. La utilización de SUR estaría motivada porque, en la práctica, los términos de error de los modelos de estimación de rendimientos como el que presentamos, suelen estar correlacionados para los diferentes activos o carteras para los que son estimados dichos modelos. Por esta razón, la utilización de modelos SUR se ha mostrado como una práctica común cuando se estima un modelo multibeta para varias carteras o valores simultáneamente, siendo ejemplos Roll y Ross (1980), que utilizan mínimos cuadrados generalizados en la estimación de los parámetros del sistema o Thorbecke (1998). En esta línea, y en España, podemos mencionar los trabajos de Pérez y Aranda (2000) y Bruno *et. al.* (2002). Asimismo, acerca de la conveniencia de la utilización de la metodología SUR en este problema, también se manifiesta Greene (1999, p. 561) refiriéndose al CAPM. Obsérvese que, al utilizar un modelo SUR, se asume que el rendimiento del mercado es una variable exógena respecto al de cada cartera en particular, lo cual, por otra parte, es una hipótesis de partida dentro del marco de los modelos de valoración de activos en el que nos movemos.

- El método alternativo que ensayamos es la utilización de un PMC con seis neuronas en la capa de entrada (las 6 variables exógenas) y 9 neuronas de salida (los rendimientos de los 9 sectores).

Para comparar el poder de ajuste y predicción de ambos métodos hemos dividido la muestra total (que abarca desde 1990 a 2001) en dos submuestras. La primera abarca los años 1990-1999 (por tanto, contiene 120 observaciones) y con ella estimamos los parámetros del sistema mediante SUR y a la vez, entrenamos la RNA. Ello nos permitirá comparar la capacidad de ajuste a ambas metodologías. La segunda submuestra abarca el periodo 2000-2001. Con ella comparamos los rendimientos que predicen la regresión lineal y el PMC con los parámetros obtenidos en el primer subperiodo.

5.2. Resultados

En el cuadro 6 se ofrecen los resultados que ofrece la estimación SUR para el periodo 1990-1999, el coeficiente de determinación de cada regresión (R^2), el error cuadrático medio (ECM) y el índice de consistencia de Theil (Theil), que expresamos como la diferencia entre 1 y el índice de desigualdad de Theil.

A excepción del rendimiento de la cartera de mercado (RM), que es significativo siempre y proporciona, como era de esperar, betas positivas, el resto de variables explicativas no son en bastantes ocasiones estadísticamente significativas según los niveles de presunción estándar y, cuando lo son, pueden presentar signos diferentes entre carteras. No obstante, remarcamos que excepto en el caso de ΔIPC (que, no obstante, es significativa en la explicación del rendimiento del sector de empresas constructoras con un nivel del 11%), no hay ninguna variable que no sea

significativa en la explicación del rendimiento de al menos una de las carteras sectoriales. También debemos reconocer que Δ IPÍ sólo es significativa para explicar los rendimientos del sector de Alimentación.

Cuadro 6
RESULTADOS DE LA ESTIMACIÓN SUR
DURANTE EL PERIODO 1990-1999

	BAN	ELEC	AL	CONS	CAR	MET	PETR	COM	OTR
Constante	0,003 (0,805)	-0,002 (0,890)	0,006 (0,674)	0,003 (0,829)	0,031 (0,102)	-0,018 (0,336)	-0,005 (0,705)	0,016 (0,333)	-0,027 (0,096)
RM	1,082 (0,000)	0,836 (0,000)	0,966 (0,000)	1,187 (0,000)	0,691 (0,000)	1,105 (0,000)	0,888 (0,000)	0,901 (0,000)	1,051 (0,000)
DPMLP	0,046 (0,617)	0,136 (0,217)	0,187 (0,111)	(0,200) (0,131)	-0,198 (0,219)	-0,111 (0,498)	-0,083 (0,447)	-0,248 (0,073)	0,168 (0,231)
PR	-0,524 (0,251)	-0,266 (0,629)	-1,925 (0,001)	-1,958 (0,003)	-0,888 (0,270)	1,297 (0,112)	0,897 (0,102)	1,274 (0,066)	-0,074 (0,916)
Δ PEND	-0,050 (0,237)	0,055 (0,274)	-0,135 (0,012)	-0,285 (0,000)	-0,143 (0,054)	-0,063 (0,404)	-0,029 (0,561)	0,268 (0,000)	-0,232 (0,000)
Δ IPC	0,448 (0,695)	0,156 (0,910)	-1,254 (0,393)	2,666 (0,109)	-2,771 (0,169)	-2,873 (0,160)	-0,137 (0,920)	-1,107 (0,524)	1,391 (0,427)
Δ IPÍ	0,431 (0,485)	-0,821 (0,269)	1,526 (0,054)	1,212 (0,176)	0,543 (0,618)	-1,099 (0,319)	-1,132 (0,126)	0,250 (0,790)	0,512 (0,588)
R^2	84,16%	69,60%	72,60%	75,38%	44,70%	63,29%	71,42%	66,26%	66,98%
ECM	0,000824	0,001194	0,001359	0,001739	0,002561	0,002638	0,001184	0,001903	0,001934
Theil	79,56%	70,94%	71,87%	73,54%	55,92%	66,36%	71,58%	69,76%	68,41%

Nota: Entre paréntesis viene dado el nivel de significación

Una razón que podría explicar la poca significación de Δ IPC e Δ IPÍ es la existencia de colinealidad entre éstas variables y el resto, ya que entre las variables explicativas propuestas existen unas claras interacciones. Tal como ya apuntamos en el epígrafe dedicado al análisis de la predicción de la quiebra, nuestro fin primordial consiste en contrastar la capacidad de ajuste y predictiva de la RNA que proponemos respecto a la metodología econométrica ensayada; y la colinealidad no supone un problema en este contexto. Por esta razón, entendemos conveniente de nuevo mantener todas las variables explicativas inicialmente seleccionadas con los valores estimados para sus coeficientes pues, en cualquier caso, éstos pueden considerarse como los "mejores" a partir de la experiencia muestral disponible, ya que la existencia de colinealidad no supone la existencia de sesgo en las estimaciones de los coeficientes realizadas con SUR. También debemos recordar que, excepto la variable Δ IPC, todas las variables seleccionadas se han mostrado relevantes en la explicación de los rendimientos de al menos una de las carteras sectoriales. Asimismo, con indepen-

dencia de que pueda haber ciertos problemas de colinealidad, también puede existir un notable componente no lineal en la relación entre el rendimiento de las carteras y la variación del IPC y del IPI que no puede ser captado con el modelo de regresión. Este aspecto supone una razón adicional para seguir considerando estas variables en la predicción de rendimientos, ya que las RNAs sí son capaces de captar dicho componente.

En el cuadro 7 se ofrece el resultado de la calidad del ajuste del PMC de los rendimientos bursátiles para el periodo 1990-1999. En su especificación final, se han introducido 6 neuronas en una capa oculta, ya que se ha observado que añadir una capa adicional y/o más neuronas no disminuía sustancialmente el error de la RNA. Asimismo, en la fase de entrenamiento, el número de épocas ha sido 5.000, ya que a partir de ésta, el error que proporcionaba no disminuía. Hemos observado que varios pesos sinápticos asociados a las dos neuronas de la capa de entrada que recogen el valor de ΔIPC e ΔIPC están notablemente alejados de cero, lo que puede ser indicativo de que estas variables son relevantes para explicar los rendimientos bursátiles. El hecho de que su significación en el modelo de regresión fuera dudosa, puede ser debido a que la dependencia de los rendimientos de las carteras respecto a la inflación y el nivel de actividad industrial tiene un componente no lineal de cierta importancia, que con la regresión no ha sido captado.

Cuadro 7
RESULTADOS DE LA RNA PARA EL PERIODO 1990-1999

	BAN	ELEC	AL	CONS	CAR	MET	PETR	COM	OTR
R^2	85,47%	73,82%	71,19%	79,72%	55,38%	62,48%	72,71%	72,80%	73,96%
ECM	0,000852	0,001257	0,002139	0,00158	0,002085	0,00338	0,001145	0,001962	0,001801
Theil	79,43%	69,72%	66,00%	75,38%	62,26%	63,34%	72,00%	71,37%	70,53%

Un análisis comparativo de los cuadros 6 y 7 sugiere que hay cierta evidencia de la superioridad en el ajuste obtenido de la RNA respecto a SUR. Puede observarse que en 7 de las 9 carteras sectoriales, el coeficiente de determinación de la RNA es superior. No obstante, dicha afirmación queda en cierta forma diluida, ya que sólo en 5 carteras el índice de Theil muestra un mejor comportamiento de la RNA, mientras que hay 5 carteras en las que se produce un menor ECM cuando se utilizan métodos de regresión.

En el cuadro 8 se observan los resultados de las predicciones de los rendimientos de los sectores en los años 2000-2001 con las especificaciones SUR y neuronal. En este caso es donde se observa, de forma nítida, la superioridad de la capacidad predictiva de la RNA respecto a SUR. Mientras que el coeficiente de determinación sólo es mejor con la especificación SUR en una cartera (la de empresas de alimentación), podemos observar que el ECM cometido es, en todas las carteras, inferior en la RNA. Asimismo, el índice de Theil indica que la RNA produce predicciones más consistentes que la metodología SUR excepto en el caso de las

carteras de alimentación y la de otras industrias.

Cuadro 8
RESULTADOS DE LA PREDICCIÓN MEDIANTE SUR VERSUS RNA
PARA EL PERIODO 2000-2001

	SUR			RNA		
	R^2	ECM	Theil	R^2	ECM	Theil
BAN	54,34%	0,002386	62,83%	67,89%	0,001576	69,35%
ELEC	19,61%	0,005901	42,02%	48,14%	0,002176	49,22%
AL	7,14%	0,016552	34,48%	0,04%	0,007577	26,16%
CON	21,38%	0,010786	42,69%	38,27%	0,003866	49,64%
CAR	19,66%	0,005100	44,63%	29,89%	0,003877	47,40%
MET	4,97%	0,016017	34,03%	17,88%	0,006110	44,68%
PETR	0,05%	0,015137	30,00%	16,81%	0,004746	41,25%
COM	42,76%	0,006074	48,66%	48,63%	0,004621	53,59%
OTR	27,77%	0,004553	51,34%	27,99%	0,003009	50,36%

5. CONCLUSIONES

En este trabajo hemos realizado un análisis comparativo de la capacidad de ajuste y predicción de los métodos econométricos convencionales respecto al empleo de las RNAs del tipo PMC en un problema clásico de clasificación financiera, la predicción de la quiebra; y en la estimación de los rendimientos de carteras de valores a partir de una formulación multibeta. En la primera aplicación debemos realizar una modelización uniecuacional, trabajándose con datos de carácter transversal, mientras que en la segunda, los datos utilizados son series temporales, siendo la modelización de los rendimientos de las carteras, de carácter multicuacional. Aunque se trata de escenarios de modelización diferentes, en ambos casos, los resultados son favorables a las RNAs.

En el problema de la predicción de la quiebra, los resultados obtenidos evidencian la superioridad de las RNAs respecto a los métodos multivariantes habituales, tanto desde el punto de vista del ajuste obtenido como desde el punto de vista predictivo. Asimismo, los resultados que obtenemos con nuestra RNA son similares a los que obtienen las RNAs de Serrano y Martín (1993) y López y Flórez (2000) con la misma base de datos. No obstante, en nuestro trabajo, al reducir las variables explicativas inicialmente elegidas a un número limitado de factores, hemos utilizado un PMC de dimensión sustancialmente inferior a la de estos trabajos.

En la segunda aplicación hemos comparado, desde el punto de vista de una formulación multibeta de los rendimientos de las acciones, la capacidad predictiva de la metodología SUR y un PMC. Tanto con la

muestra de ajuste como con la de comprobación, la evidencia apuntaba a que era más eficiente el PMC. No obstante, debemos reconocer que ello no se mostraba de forma nítida en la fase de ajuste, pero sí que era mucho más evidente en la muestra de testeo.

Entendemos que los positivos resultados que obtenemos con las RNAs respecto al resto de métodos ensayados en las dos aplicaciones efectuadas, pueden ser imputados a la mayor capacidad de las RNAs para aproximar relaciones no lineales y, paralelamente, tener en cuenta las interrelaciones que existen entre las variables explicativas.

No obstante, tal como apuntan Altman *et. al.* (1994), no pensamos que las RNAs sean un instrumento sustitutivo de los métodos convencionales, sino más bien un complemento, ya que presentan ciertas carencias. En primer lugar, los coeficientes obtenidos con los métodos convencionales son más fáciles de interpretar, ya que son comprensibles intuitivamente y el contraste de su significación es sencillo. En cambio, la comprensión de los resultados que ofrece un PMC es más compleja, ya que la capa oculta actúa como una "caja negra", mientras que tampoco es sencillo determinar el significado de los pesos sinápticos.

Por otra parte, los métodos multivariantes convencionales son más fáciles de implementar y estimar que una RNA ya que, en esta última, el proceso de entrenamiento es relativamente lento. Además, no existe una fórmula generalmente aceptada sobre cómo determinar la dimensión óptima de una RNA, aunque tal como apuntamos en el tercer epígrafe, existen algunos criterios y técnicas que pueden ayudar a diseñar una RNA con una dimensión adecuada. No obstante, este aspecto dependerá, en última instancia, de la subjetividad y experiencia del investigador.

REFERENCIAS BIBLIOGRÁFICAS

- Adya, M. y Collopy, F. (1998): "How Effective are Neural Networks at Forecasting and Prediction? A Review and Evaluation", *Journal of Forecasting*, vol. 17, nº 5-6, pp. 89-94.
- Ahmadi, H. (1996): "Testability of Arbitrage Pricing Theory by Neural Networks", en Trippi; R.R y Turban, E. (coords.), *Neural Networks in Finance and Investing*, Irwing, Chicago, pp. 543-554.
- Altman, E.I. (1968): "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *Journal of Finance*, vol. 23, nº 4, pp. 589-609.
- Altman, E.I.; Marco, G. y Varetto, F. (1994): "Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks", *Journal of Banking and Finance*, vol. 18, nº 3, pp. 505-529.
- Beaver, W. (1967): "Financial Ratios as Predictors of Failure", *Selected Studies, 1966. The Journal of Accounting Research*, vol. 4, suplemento, pp. 71-111.

- Bonilla, M.; Marco, P.P. y Olmeda, I. (2000): "Forecasting Exchange Rates Volatilities Using Artificial Neural Networks", en Bonilla; M.; Casasús, T. y Sala, R. (coords.), *Financial Modelling*, Physica-Verlag, Heidelberg, pp. 57-68.
- Bruno, N.; Medina, U. y Morini, S. (2002): "Contraste factorial del arbitrage pricing theory en el mercado bursátil español", *Análisis Financiero*, nº 88, pp. 58-63.
- Castromán, J.L. y Porto, N. (2000): "Asistencia Neuronal en la toma de decisiones: una aplicación al credit scoring en un ambiente de incertidumbre", VII Congreso de la Sociedad Internacional de Gestión y Economía Fuzzy (SIGEF), septiembre, Chania (Grecia).
- Dimitras, A.I.; Slowinski, R.; Susmaga, R. y Zopounidis, C. (1999): "Business Failure Prediction Using Rough Sets", *European Journal of Operational Research*, vol. 114, nº 2, pp. 263-280.
- Dutta, S. y Shekhar, S. (1988): "Bond Rating: A Nonconservative Application of Neural Networks", *Proceedings of the IEEE International Conference of Neural Networks*, pp. 11443-11450.
- Fernández, F.; González, C. y Sosvilla, S. (2000): "On the Profitability of Technical Trading Rules Based on Artificial Neural Networks: Evidence from Madrid Stock Market", *Economic Letters*, vol. 69, nº 1, pp. 89-94.
- Ferson, W.E. y Harvey, C.R. (1991): "The Variation of Economic Risk Premiums", *Journal of Political Economy*, vol. 99, nº 2, pp. 385-415.
- Greene, W.H. (1998): *Análisis econométrico*, Prentice-Hall, Madrid, 1999.
- Harvey, C.R. (1988): "The Real Term Structure and Consumption Growth", *Journal of Financial Economics*, vol. 22, nº 2, pp. 305-333.
- Harvey, C.R. (1991): "The Term Structure and World Economic Growth", *The Journal of Fixed Income*, vol. 1, nº 1, pp. 7-19.
- Harvey, C.R; Travers, K.E. y Costa, M.J. (2000): "Forecasting Emerging Market Returns Using Neural Networks", *Emerging Markets Quarterly*, vol. 4, nº 2, pp. 1-12.
- Hawley, D.D; Johnson, J.D. y Raina, D. (1990): "Artificial Neural Systems: A New Tool for Financial Decision-Making", *Financial Analysts Journal*, vol. 46, nº 6, noviembre-diciembre, pp. 63-72.
- Hornick, K.; Stinchcombe, M. y White, H. (1989): "Multilayer Feedforward Networks Are Universal Approximators", *Neural Networks*, vol. 2, nº 5, pp. 359-366.
- Hutchinson, J.M.; Lo, A.W. y Poggio, T. (1994): "A non Parametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks", *The Journal of Finance*, vol. 39, nº 3, pp. 851-889.

- Iglesias, S. (2000): "La valoración de activos con riesgo: modelos de factor único vs. modelos multifactoriales. Un enfoque empírico", XIV Congreso Nacional y X Congreso Hispano-Francés de la Asociación Europea de Dirección y Economía de la Empresa (AEDEM), junio, Jaén.
- Jensen, H.L. (1992): "Using Neural Networks for Credit Scoring", *Managerial Finance*, vol. 18, nº 6, pp. 15-26.
- John, G.H.; Miller, P.P. y Kerber, R. (1996): "Stock Selection Using RECON", en Refenes, A.P; Mostafa, Y.A., Moody, J. y Weigend, A. (coords.), *Neural Networks in Financial Engineering*, World Scientific, Nueva Jersey, pp. 303-316.
- Kanas, A. (2001): "Neural Networks vs Linear Models of Stock Returns: An Application to the UK and German Stock Market Indexes", en Zopounidis; C.; Pardalos, P. y Baourakis, G. (coords.), *Fuzzy Sets in Management, Economics and Marketing*, World Scientific, New Jersey, pp. 181-193.
- Kingdon, J. y Feldman, K. (1995): "Genetic Algorithms for Bankruptcy Prediction", Research Report nº. 01-95, SearchSpace Ltd., Londres.
- Kuan, C. y White, H. (1994): "Artificial Neural Networks: An Econometric Perspective", *Econometric Reviews*, vol. 13, nº 6, pp. 1-91.
- Kuo, R.J; Chen, C.H. y Hwang, Y.C. (2001): "An Intelligent Stock Trading Decision Support System Through Integration of Genetic Algorithm Based Fuzzy Neural Network and Artificial Network", *Fuzzy Sets and Systems*, vol. 118, nº 1, pp. 47-52.
- Laffarga, J.; Martín, J.L. y Vázquez, J. (1986): "El pronóstico a largo plazo del fracaso en las instituciones bancarias: metodología y aplicaciones del caso español", *Esic-Market*, nº 54, pp. 113-123.
- Lam, K. P.P; Chiu, K.C. y Chan, W.K. (1996): "Stock Selection Using Recon", en Refenes, A.P; Mostafa, Y.A., Moody, J. y Weigend, A. (coords.), *Neural Networks in Financial Engineering*, World Scientific, New Jersey, pp. 317-329.
- Lechón, P.P; Ruiz, F.J. y Santamaría, R. (1997): "Valoración de opciones mediante inteligencia artificial: Contraste empírico con las opciones IBEX-35", *Actualidad Financiera*, nº 2, pp. 35-44.
- López, E. y Flórez, R. (2000): "Aplicación de dos modelos de redes neuronales artificiales para el análisis económico-financiero empresarial", *Revista Europea de Dirección y Economía de la Empresa*, vol. 9, nº 2, pp. 141-166.
- Mahfoud, S. y Mani, G. (1996): "Financial Forecasting Using Genetic Algorithms", *Applied Artificial Intelligence*, vol. 10, nº 6, pp. 543-565.
- Martín, B. y Sanz, A. (1997): *Redes neuronales y sistemas borrosos*, Rama, Madrid.

- Matsatsinis, N.; Doumpos, M. y Zopounidis, C. (1997): "Knowledge Acquisition and Representation for Expert Systems in the Field of Financial Analysis", *Expert Systems with Applications: An International Journal*, vol. 12, nº 2, pp. 247-262.
- Matsatsinis, N.; Kosmidou, K.; Doumpos, M. y Zopounidis, C. (2003): "A Fuzzy Decision Aiding Method for the Assessment of Corporate Bankruptcy", *Fuzzy Economic Review*, vol. 8, nº 1, pp. 13-24.
- Pérez, F. y Aranda, C. (2000): "Las intervenciones del Banco de España y la reacción del mercado bursátil", *Análisis Financiero*, nº 80, pp. 26-33.
- Pérez, J.; Torra, S. y Borrell, M. (2000): "Modelos para la predicción de los tipos de interés en el mercado interbancario: estructuras lineales, GARCH y redes neuronales artificiales", *Revista Asturiana de Economía*, nº 18, pp. 123-139.
- Pérez-Rodríguez, J. y Torras, S. (2001): "Diversas formas de predicción no lineal y contrastes de selección de modelos en la predicción de rendimientos del IBEX-35", Fedea, Documento de trabajo, serie de Economía Internacional, Monetaria y Financiera nº 94, Madrid.
- Pina, V. (1989): "La información contable en la predicción de la crisis bancaria", *Revista Española de Financiación y Contabilidad*, vol. 18, nº 58, pp. 309-338.
- Rahimian, E.; Singh, S.; Tammachote, T. y Rivmani, R. (1996): "Bankruptcy Prediction by Neural Network", en Trippi; R.R y Turban, E. (coords.), *Neural Networks in Finance and Investing*, Irwing, Chicago, pp. 243-259.
- Refenes, A.N. (1995): "Neural Network Design Considerations", en Refenes; A.N. (coord), *Neural networks in capital markets*, John Willey, Nueva York, pp. 10-25.
- Refenes, A.N. (1996): "Constructive Learning and its Application to Currency Exchange Rate Forecasting", en Trippi; R.R. y Turban, E. (coords.), *Neural Networks in Finance and Investing*, Irwing, Chicago, pp. 777-805.
- Richeson, L.; Zimmermann, R.A. y Barnett, K.G. (1994): "Predicting Consumer Credit Performance: Can Neural Networks Outperform Traditional Statistical Methods?", *International Journal of Applied Expert Systems*, vol. 2, nº 2, pp. 116-130.
- Rico, PP. (2001): "Estructura temporal de los tipos de interés y crecimiento económico en España", *Revista Española de Financiación y Contabilidad*, vol. 30, nº 108, pp. 393-407.
- Robles, J. y Naylor, C. (1996): "Applying Neural Networks in Cooper Trading: A Technical Analysis Simulation", en Refenes, A.P; Mostafa, Y.A.; Moody, J. y Weigend, A. (coords.), *Neural Networks in Financial Engineering*, World Scientific, New Jersey, pp. 234-240.

- Roll, R. y Ross, S. A. (1980): "An Empirical Investigation of the Arbitrage Pricing Theory", *The Journal of Finance*, vol. 35, nº 5, pp. 1073-1103.
- Ross, S.A. (1976): "The Arbitrage Theory of Capital Pricing", *Journal of Economic Theory*, vol. 13, nº 3, pp. 341-360.
- Rummelhart, D. y McClelland J. (1986): *Parallel distributed processing*, MIT Press, Cambridge.
- Serrano, C. y Gallizo, J.L. (1997): "Las redes neuronales artificiales en el tratamiento de la información financiera", <<<http://ciberconta.unizar.es>>>, Consulta en julio de 1997.
- Serrano, C. y Martín, B. (1993): "Predicción de la quiebra bancaria mediante el empleo de redes de neuronas artificiales", *Revista Española de Financiación y Contabilidad*, vol. 22, nº 74, pp. 153-171.
- Sharpe, W. (1964): "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk", *The Journal of Finance*, vol. 19, nº 3, pp. 425-442.
- Spanos, M.; Kosmidou, K. y Doumpos, M. (2000): "Portfolio Selection with Fuzzy Logic", VII Congreso de la Sociedad Internacional de Gestión y Economía Fuzzy (SIGEF), Chania (Grecia): pp. 91-101.
- Swanson, R. y White, H. (1995): "A Model-Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks", *Journal of Business and Economic Statistics*, vol. 13, nº 3, pp. 265-275.
- Tam, K.-Y. y Kiang, M.-Y. (1992): "Predicting Bank Failures: A Neural Network Approach", *Management Science*, vol. 38, nº 7, pp. 926-947.
- Thorbecke, W. (1997): "On Stock Market Returns and Monetary Policy", *The Journal of Finance*, vol. 52, nº 2, pp. 629-658.
- Trippi, R.R. y Desieno, D. (1992): "Trading Equity Index Futures with a Neural Network", *Journal of Portfolio Management*, vol. 18, nº 5 pp. 27-33.
- Trippi, R.R. y Lee, J. K. (1996): *Artificial intelligence in finance & investing*, Irwin, Chicago.
- West, R. G. (1985): "A Factor-Analytic Approach to Bank Condition", *Journal of Banking and Finance*, vol. 9, nº 2, pp. 253-266.
- White, H. (1988): "Economic Prediction Using Neural Networks: the Case of IBM Daily Stock Prices", *Proceedings of the IEEE International Conference of Neural Networks*, pp. II451-II458.
- Wilson, R.L. y Sharda, R. (1994): "Bankruptcy Prediction Using Neural Networks", *Decision Support Systems*, vol. 11, nº 5, pp. 545-557.

- Wong, F.S.; Wang, Pp.-Z.; Goh, T-H. y Quek, B. K. (1992): "Fuzzy Neural Systems for Stock Selection", *Financial Analysts Journal*, vol. 48, n° 1, enero-febrero, pp. 47-52.
- Yao, J.-T. y Poh, H.L. (1996): "Equity Forecasting: A Case Study in the KLSE Index", en Refenes, A.P; Mostafa, Y.A.; Moody, J. y Weigend, A. (coords.), *Neural Networks in Financial Engineering*, World Scientific, Nueva Jersey, pp. 341-253.
- Yoon, Y. y Sales, G. (1991): "Predicting Stock Price Performance: A Neural Network Approach", *Proceedings of the IEEE 24th annual International conference of Systems Sciences*, pp. 156-162.

ABSTRACT

In many financial problems the relation between relevant variables is more complex than theoretical models suppose (despite their clear usefulness). Moreover, in these problems the relevant information is usually ill-structured. In this respect, many authors have shown the suitability of Artificial Intelligence instruments. The aim of this paper is to provide evidence about the usefulness of Artificial Neural Networks in two classic topics of financial classification and financial prediction: failure business prediction and estimation of the stock portfolios return. In our applications, we compare the performance of the RNAs with the performance of classical multivariate methods that are used frequently in the analysis of those problems.

Key words: artificial intelligence, artificial neural networks, failure prediction, multiple factor pricing models, CAPM, APT.