

CLASIFICACIÓN CREDITICIA MEDIANTE MODELOS DE AGREGACIÓN

María Bonilla

Universidad de Valencia

Ignacio Olmeda

Universidad de Alcalá

Rosa Puertas

Universidad de Valencia

RESUMEN:

La resolución de problemas de clasificación financieros mediante la utilización de modelos paramétricos y no paramétricos es cada vez más frecuente al ser de gran utilidad en los procesos de toma de decisiones. En concreto, tales modelos han sido aplicadas con gran éxito a problemas de clasificación de bonos, predicción de quiebras y evaluación de créditos. Sin embargo, la necesidad constante de mejorar su precisión ha originado que recientemente las investigaciones se dirijan hacia modelos de agregación que, en algunos casos, pueden superar a los modelos individuales. En el presente trabajo analizamos la capacidad predictiva de dos modelos híbridos basados en la agregación de modelos paramétricos y no paramétricos en un problema de concesión de tarjetas de crédito. Tras realizar un análisis comparativo, considerando costes asimétricos de clasificación errónea, concluimos que la agregación de modelos puede resultar ventajosa.

INTRODUCCIÓN

La gran mayoría de los préstamos que diariamente conceden las entidades financieras podrían ser otorgados mediante técnicas de clasificación, que en breves instantes, y en base a las características propias del cliente, determinen la probabilidad de que en un futuro éste pueda hacer frente a sus obligaciones de pago. Así pues, los modelos de *credit scoring* pueden definirse como sistemas de evaluación automática que permiten clasificar al individuo en dos grupos: el integrado por individuos con una elevada probabilidad de devolver el crédito y el de aquéllos individuos que, previsiblemente, no harán frente a sus compromisos crediticios. El objetivo principal de los modelos de clasificación crediticia será, por tanto, extraer una regla general de una muestra de observaciones que facilite el análisis de nuevas solicitudes de crédito, informando con alto grado de precisión, sobre la probabilidad de que la petición analizada resulte fallida.

Aunque recientemente están lográndose importantes mejoras en este ámbito, la preocupación por obtener modelos adecuados de concesión de créditos se remonta a principios de los años 50, cuando Durand (1951) elaboró un modelo discriminante de préstamos que permitía clasificarlos de acuerdo con su capacidad de devolución. A partir de este momento los estudios se han ido sucediendo, pasando de clásicos como Myers y Forgy (1963), Bierman y Hausman (1970), Orgler (1970) y Apilato et al. (1974) entre otros, en los que modelos paramétricos eran el centro de sus investigaciones, a estudios más sofisticados (Marose, 1990; Jenson, 1992; Piramuthu *et al.*, 1994; Dorrnsoro et al., 1997) en los que la utilización de modelos no paramétricos han permitido obtener importantes progresos en términos de exactitud.

La diferencia fundamental entre los modelos paramétricos y los no paramétricos reside en que los paramétricos suponen conocida la forma de la función discriminante, reduciéndose el problema a estimar los parámetros que la determinan, mientras que los no paramétricos suponen una forma funcional desconocida, y su objetivo consiste en aproximar aquélla mediante formas funcionales flexibles. Se trata, por tanto, de modelos mucho más flexibles que los anteriores, pero también más costosos en su construcción.

El objetivo de este trabajo consiste en analizar sistemas de hibridación evaluando las posibles mejoras predictivas frente a los modelos individualmente analizados. Para ello plantearemos dos formas de agregación, una paramétrica y otra no paramétrica, y compararemos, en términos de costes, los resultados obtenidos por dichos métodos con los alcanzados por otros modelos simples.

La estructura seguida en nuestro trabajo es la siguiente: en la sección segunda realizamos una breve revisión de los fundamentos teóricos que definen los modelos que hemos utilizado en nuestro estudio. En la sección tercera, planteamos la metodología formal seguida en la construcción del híbrido. En la sección cuarta, presentaremos los resultados obtenidos comparando la precisión, en un problema de clasi-

ficación de créditos, de los sistemas de agregación con otros modelos simples. Por último, en la sección quinta formulamos las principales conclusiones obtenidas en nuestro estudio.

FUNDAMENTOS DE LOS MODELOS DE CLASIFICACIÓN

Los modelos de clasificación tratan de encontrar la forma funcional, $f(x)$ que permita ajustar las observaciones de una muestra minimizando el error de clasificación. Un modelo de clasificación puede ser expresado formalmente:

$$P = f(x_1, x_2, \dots, x_n) + \varepsilon \quad (2.1)$$

donde x_i representan los atributos de un determinado individuo, P es el grupo al que pertenece el individuo (moroso o no moroso), $f(x)$ la función de clasificación que especifica la relación existente entre las variables utilizadas, y ε es la perturbación aleatoria que representa el término de error. El *credit scoring*, por tanto, puede ser definido como un problema de clasificación, consiste en encontrar la forma funcional $f(x)$ que permita discriminar a los individuos de acuerdo con sus características crediticias, sirviéndose para ello del comportamiento seguido en otros créditos concedidos.

Dependiendo de que $f(x)$ esté o no especificada previamente, los modelos utilizados se denominan paramétricos o no paramétricos, como indicamos anteriormente, los modelos paramétricos parten de una forma funcional conocida y el problema reside en estimar los parámetros que mejor ajusten las observaciones de muestra, mientras que los no paramétricos no asumen ninguna forma funcional y por tanto, además de estimar los parámetros, será necesario determinar la forma funcional $f(x)$ para que el error cometido también sea mínimo.

En nuestro trabajo hemos utilizado dos técnicas paramétricas, el Análisis Discriminante (AD) y el Logit, y tres no paramétricas: Árboles de Regresión y Clasificación (CART), Regresión Localmente Ponderada (RLP) y Redes Neuronales Artificiales (RNA). Por motivos de brevedad, realizaremos una revisión de los modelos no paramétricos empleados en el presente estudio, ya que los modelos paramétricos son bien conocidos al haber sido ampliamente analizados en numerosos estudios financieros.

El CART (Friedman, 1977) es una técnica no paramétrica de clasificación binaria que consiste en ir dividiendo sucesivamente la muestra original en submuestras procurando que éstas sean cada vez más homogéneas, utilizando para ello reglas univariantes que permitan determinar cual es la variable independiente que mejor discrimina la muestra. El modelo realizará divisiones sucesivas del conjunto de ejemplos analizando un punto de corte de un determinado atributo (por ejemplo, "¿son los ingresos superiores a x mill. pta/año?") de forma que la partición generada sea la que mayor homogeneidad induce en los subgrupos generados. Su utilización en problemas de clasificación es relativamente reciente si lo comparamos con los modelos; cabe citar los trabajos de Breiman et al. (1984), Frydman et al. (1985) o Marais et al. (1984).

CART presenta una estructura en forma de un árbol compuesto por una sucesión de nodos y ramas que constituyen respectivamente los grupos y divisiones que se van realizando de la muestra original. Cuando ya no sea posible realizar ninguna otra división que mejore la homogeneidad de los subgrupos se llega a un nodo terminal, que será la clase asignada por el modelo. El error total del árbol se calculará sumando los correspondientes a cada uno de estos nodos terminales.

El principal problema con el que se enfrenta este modelo es la complejidad de su estructura, problema común a todos los modelos no paramétricos, y que fácilmente puede desembocar en el sobreaprendizaje de las observaciones que componen la muestra. Por esta razón, en la construcción de los modelos CART no sólo se persigue crear conjuntos homogéneos, sino también se pretende obtener aquella estructura que presente una complejidad óptima. Bajo este doble objetivo resulta necesario penalizar la excesiva complejidad de la estructura.

La RLP consiste en ponderar no linealmente los ejemplos más parecidos a la observación cuya clasificación tratamos de establecer. El procedimiento para construir una función alisada $g(x)$ sigue los pasos: 1) tomar un ejemplo x y buscar los q vecinos más próximos a dicho punto, constituyendo así una vecindad $N(x)$ en términos de similitud de los atributos de ambos ejemplos, tomando para ello la distancia euclídea de los atributos, 2) calcular la distancia máxima entre x y cualquier punto del vecindario, 3) asignar pesos a cada uno de los puntos de la vecindad $N(x)$ a través de la función de ponderación tri-cúbica, estos puntos se ponderan en función de su distancia respecto de x , los puntos próximos a x poseen pesos elevados mientras que, por el contrario, a los puntos alejados de x se les asignan ponderaciones bajas, 4) sobre la base de estos pesos asignados se ajusta, mediante mínimos cuadrados ponderados, una función (lineal o cuadrática) $g(x)$ sobre el vecindario $N(x)$, con ello se obtiene el valor ajustado $g(x)$, 5) repetir

este procedimiento para cada valor de la variable predictora para el cuál se desea obtener una estimación $g(x)$.

Las RNA son modelos matemáticos que asemejan los sistemas de neuronas altamente interconectadas a modo de elementos de proceso, organizados jerárquicamente en capas y que permiten computar relaciones impulso-respuesta. La información procedente del exterior se introduce en el modelo por la capa de entrada y fluye por el resto de capas ocultas hasta la capa de salida que dará una determinada respuesta.

Las conexiones de los nodos llevan asociadas un peso que determina el efecto cuantitativo que cada nodo ejerce sobre el otro, de forma que la información procedente de otros nodos será ponderada antes de ser transformada y enviada a los nodos de la siguiente capa. Así pues, el valor de entrada de cada neurona se calculará sumando los valores de entrada ponderados por sus pesos correspondientes,

$$y_i = \sum_{i=1}^n x_i w_i \quad (2.3)$$

donde x_i representa los atributos del problema, w_i las ponderaciones asociadas, e y_i el valor de entrada. Posteriormente, a dicha entrada se le asocia una función de activación (que en nuestro caso ha sido la función logística) obteniendo de esta forma el valor de salida de cada nodo que será transmitido al resto de neuronas. Así pues, la solución de la red vendrá dada por la siguiente expresión,

$$y = \sum_{i=1}^n \beta_i F(x'w_i) \quad (2.4)$$

donde, y es la salida de la red, x' representa el vector transpuesto de atributos de entrada de la red, w_i se corresponde con los pesos asociados a las conexiones entre la capa de entrada y las capas ocultas, mientras que β_i son los pesos asociados a las conexiones entre la capa intermedia y la de salida, y F la función de activación. Recientemente las RNA han sido aplicadas con éxito a numerosos problemas de clasificación financiera (Serrano (1994), Brockett *et al.* (1994), Tam y Kiang (1992) y Dutta *et al.* (1994), entre otros).

Un problema común a todos los modelos no paramétricos, como ya hemos señalado anteriormente, es el sobreaprendizaje ocasionado por una sobreparametrización del modelo: el modelo es tan complejo que no es capaz de generalizar, limitándose a memorizar el conjunto de entrenamiento por lo que ante nuevas observaciones el error cometido aumenta considerablemente. Por tanto, resulta imprescindible obtener la estructura óptima de cada uno de los modelos no paramétricos, de manera que, tanto para el entrenamiento como para el test, el error sea mínimo. En el caso del CART el problema se resolverá podando aquellas ramas que presenten un alto grado de complejidad, en la RLP se tendrá que determinar el tamaño óptimo del vecindario, mientras que en las RNA el proceso es relativamente más costoso pues no sólo hay que determinar el número de capas ocultas, sino también el número de neuronas que las definen.

Con objeto de obtener el equilibrio idóneo entre la parametrización del modelo y la complejidad del problema hemos utilizado el método de validación cruzada propuesto por Stone (1974), que consiste en dividir aleatoriamente el conjunto de entrenamiento en diez partes iguales no solapadas, manteniendo en cada subconjunto la misma proporción de éxito o fracaso que existía en la muestra original. Mediante rotación, un conjunto formado por nueve particiones es elegido para realizar el entrenamiento del modelo, y la décima partición es utilizada para testearlo, repitiendo este proceso. El error de validación cruzada se calcula como una media de los resultados de estos diez pruebas, y la estructura óptima será aquella que minimice esta media. Elegida la estructura, la capacidad predictiva del modelo se valorará mediante el conjunto original de test, obteniendo así el error de predicción.

METODOLOGÍA PARA LA CONSTRUCCIÓN DE MODELOS HÍBRIDOS.

Aunque no son muchos los trabajos desarrollados en esta área, la idea de agregar modelos individuales con objeto de mejorar la potencia clasificadora de los mismos parte de los años 60, cuando Bates y Granger (1969) proponen un híbrido de modelos de predicción para aprovechar las ventajas comparativas de todos ellos. Posteriormente, Granger y Ramanathan (1984) plantean la posibilidad de combinar modelos individuales mediante una regresión lineal, tomando como variables independientes las predicciones de los diferentes modelos. Sin embargo, la utilización del modelo lineal no siempre resulta óptima debido a la posible existencia de relaciones cruzadas entre los clasificadores.

La literatura existente revela que los esfuerzos han ido dirigidos exclusivamente a la combinación de predicciones cuantitativas, mostrando poco interés hacia la agregación de predicciones cualitativas.

Clemen (1989) realiza una revisión de la literatura existente poniendo de relieve algunos de los posibles problemas inherentes a la agregación de predicciones cualitativas (Feather y Kaylen, 1989; Fan et al., 1996; Kamstra y Keneddy, 1998)

Un sistema híbrido es un comité de modelos construido a partir de una agregación eficiente de predicciones. La estructura es, por tanto, muy similar a la expresada en 2.1, la diferencia radica en que las variables independientes en lugar de ser características de las observaciones que componen la muestra, son predicciones de modelos individuales, es decir, tomaremos las predicciones de dichos modelos como variables independientes. La expresión del modelo híbrido la siguiente:

$$P_i = f(P_i^1, P_i^2, P_i^3, P_i^4, P_i^5) \quad (3.1)$$

donde $f(P)$, al igual que en 2.1, es la función de agregación que determina la relación existente entre las variables utilizadas, en nuestro problema dichas variables serán probabilidades de éxito o fracaso obtenidas en cada observación por cada uno de los modelos utilizados, P_i el resultado obtenido, y P_i^z las predicciones de los modelos, siendo $i=1,2,\dots,n$ las predicciones obtenidas sobre los conjuntos validados e $i=n+1, n+2, \dots, m$ las predicciones del conjunto de test, $z=1,2,3,4,5$ se corresponde con los distintos modelos, AD, logit, CART, RLP y RNA respectivamente.

El problema reside en que el híbrido debe ser construido utilizando sólo observaciones del conjunto de entrenamiento, y por el contrario su precisión tiene que ser evaluada sobre el conjunto de test. Kumar y Olmeda (1999), proponen aproximar el error de predicción mediante el empleo del error de validación cruzada, al ser éste un estimador insesgado de aquél (Stone, 1974).

Específicamente, dada una función de pérdida, d , que permita valorar las diferencias existentes entre el resultado obtenido, P_i , y la solución real de cada observación del conjunto de entrenamiento (c_1, c_2, \dots, c_n) y de test ($c_{n+1}, c_{n+2}, \dots, c_m$):

$$d_i = d(c_i, P_i) \quad (3.2)$$

el comité óptimo será aquel que permita agregar las predicciones, expresión 3.1., de forma que se minimice la función de pérdida sobre el conjunto de test, es decir,

$$D = \sum_{i=n+1}^m d_i \quad (3.3)$$

Por tanto, el híbrido que minimice la función de pérdida, D' , calculada sobre los conjuntos de test validados también minimizarán la función D .

$$D' = \sum_{i=1}^n d_i \quad (3.4)$$

Ahora bien, la condición necesaria para que esto sea cierto es que los conjuntos de entrenamiento y test presenten distribuciones similares, pues de otro modo el híbrido propuesto podría ser óptimo sobre el conjunto de entrenamiento y no serlo sobre el de test. Como veremos en la sección siguiente, hemos demostrado también que si esta condición es cierta los híbridos también resultará más precisos que los modelos simples, aunque los modelos utilizados en la agregación no hayan sido previamente optimizados.

Por tanto, la metodología empleada puede resumirse en los siguientes pasos:

1. Utilizar el método de validación cruzada con objeto de encontrar la estructura óptima de cada uno de los modelos simples.
2. Obtener las predicciones de los mejores modelos para cada uno de los conjuntos de test validados.
3. Elegir una función de pérdida y encontrar el híbrido que minimice dicha función a lo largo de estos conjuntos.
4. Utilizar el híbrido para calcular el error de predicción sobre el conjunto de test.

Para la construcción de la función de agregación $f(P)$ pueden utilizarse indistintamente técnicas paramétricas y no paramétricas, nosotros comparemos la precisión del híbrido construido a través del modelo logit y de un modelo CART.

RESULTADOS

La base de datos utilizada en nuestro estudio procede de un trabajo previo presentado por Quinlan (1987), formada por 690 observaciones de 14 características crediticias de individuos demandantes de una tarjeta de crédito, así como de los respectivos comportamientos crediticios posteriores a la concesión. El 55% de las observaciones correspondía a individuos que resultaron fallidos. El conjunto total se dividió aleatoriamente en dos subconjuntos: el "conjunto de entrenamiento", formado por 600 observaciones,

sobre el que se estiman los modelos individuales y se construyen los híbridos, y el “conjunto de test”, formado las 90 observaciones restantes.

Los resultados obtenidos con el empleo de modelos simples son objeto de un trabajo anterior (Bonilla et al., 1998), en dicho trabajo se demuestra que un modelo de RNA superaba significativamente la capacidad predictiva del resto de modelos. En la Tabla 1 resumimos los resultados que se obtuvieron:

TABLA I

MODELOS	ERRORES	
	ENTRENAMIENTO	TEST
AD	14.2%	12.2%
LOGIT	12.5%	12.2%
CART	14.6%	13.3%
RLP	12.3%	12.3%
RNA	12.3%	10.0%

En los problemas de clasificación, existen dos tipos de errores: el error de tipo I, cuando se aprueba un crédito que debería haberse rechazado pues el cliente ha resultado moroso, y el inverso, error tipo de II, cuando se rechaza un crédito que debería haberse aprobado, pues de haberse concedido el cliente hubiera cumplido con sus obligaciones crediticias. El coste asociado a cada uno de estos errores depende obviamente de la estrategia seguida por el prestamista en cada momento, por lo que, en este trabajo, hemos utilizado una amplia variedad de costes relativos con objeto de intentar cubrir cualquier situación que pudiera darse. A continuación, utilizamos la metodología de hibridación descrita, comparando, en términos de costes totales, los resultados de estos dos modelos con los obtenidos individualmente.

En la Tabla 2 presentamos los porcentajes de error y sus costes globales, entendiendo por costes globales la suma de los errores de clasificación, multiplicados por sus respectivos costes. Observamos que, para costes menores que 1/2 e iguales o superiores a 20 el híbrido de base logit supera significativamente al resto de modelos, incluso a las RNA que era el modelo simple que había resultado más preciso.

Como mencionamos, el procedimiento de hibridación debería resultar también adecuado con independencia de si los modelos individuales empleados son o no óptimos. Para comprobar esto, dividimos la muestra en dos subconjuntos del mismo tamaño, un conjunto de entrenamiento formado por 345 observaciones, y otro de test formado por el mismo número de observaciones, manteniendo distribuciones homogéneas e igual proporción de fallidos a la que existía en la muestra original. Este proceso lo realizamos dos veces para dar mayor consistencia a los resultados obtenidos. En las Tablas 3 y 4 mostramos los resultados obtenidos, y como podemos comprobar, el híbrido de base logit de nuevo domina al resto de modelos siempre que coste sea igual o menor que 1/5 e igual o mayor que 15. Como conclusión, si la distribución de los conjuntos de entrenamiento y test son similares, mediante la agregación de modelos también se obtienen mejores resultados que con los modelos individuales, incluso cuando para la construcción del híbrido los modelos utilizados no han sido previamente optimizados.

CONCLUSIONES

En el presente trabajo hemos realizado un análisis comparativo en términos de costes de cinco modelos diferentes, dos paramétricos (AD y logit) y tres no paramétricos (CART, RLP, RNA), frente a dos modelos híbridos consistentes en "comités" formados por las predicciones de dichos modelos.

Para poder analizar cualquier estrategia adoptada por el prestamista, se han utilizado una amplia variedad de costes, demostrando que cuando las asimetrías entre los errores tipo I y tipo II son significativas los sistemas de agregación propuestos proporcionan mayor capacidad predictiva que el resto de modelos. En concreto, cuando el error tipo I sea superior al tipo II, el híbrido de base logit bate a cualquier otro modelo incluso a las RNA, que era el modelo simple que presentaba mayor precisión.

Finalmente hemos comprobado que cuando se utilizan conjuntos de entrenamiento y test similares, también es posible la construcción de un híbrido que supere significativamente la capacidad predictiva de cualquier número de modelos simples, con independencia de si éstos son a o no óptimos.

REFERENCIAS

- APILADO, V.; WARNER, D.; DAUTEN, J. (1974): "EVALUATIVE TECHNIQUES IN CONSUMER FINANCE-EXPERIMENTAL RESULTS AND POLICY IMPLICATIONS FOR FINANCIAL INSTITUTIONS" *JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS*, 275-283.
- BIERMAN, H.; HAUSMAN, W. (1970): "THE CREDIT GRANTING DECISION", *MANAGEMENT SCIENCE* 16, B-519-532.
- BONILLA, M.; OLMEDA, I.; PUERTAS, R. (1998): "MODELOS ALTERNATIVOS DE CREDIT SCORING", EN A. PARTAL Y F. MORENO (EDS.), *VI FORO DE FINANZAS*, 733- 750.
- BREIMAN, L.; FRIEDMAN, J.; OLSHEN, R.; STONE, C. (1984): *CLASSIFICATION AND REGRESION TREES* WADSWORTH & BROOKS
- BROCKETT, P.; COOPER, W.; GOLDEN, L.; XIA, X. (1997): "A CASE STUDY IN APPLYING NEURAL NETWORKS TO PREDICTING INSOLVENCY FOR PROPERTY AND CASUALTY INSURERS" *JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY*, pp. 1153-1162
- DORRONSORO, J.; GINEL, F.; SÁNCHEZ, C.; SANTA CRUZ, C. (1997): "NEURAL FRAUD DETECTION IN CREDIT CARD OPERATIONS", *IEEE TRANSACTIONS ON NEURAL NETWORKS*, 8, 827-834.
- DURAND, D. (1951): "RISK ELEMENTS IN CONSUMER INSTALLMENT FINANCING" *STUDY N° 8, NATIONAL BUREAU OF ECONOMIC RESEARCH, NEW YORK*.
- DUTTA, S.; SHERKHAR, S. (1988): "BOND RATING: A NON-CONSERVATIVE APPLICATION ON NEURAL NETWORKS", IN R. TRIPPI AND E. TURBAN (EDS.), *NEURAL NETWORKS IN FINANCE AND INVESTING*, CHICAGO 1993, 257-273.
- FAN, D.K.; LAU, K.-N.; LEUNG, P.-L. (1996): "COMBINING ORDINAL FORECASTS WITH AN APPLICATION IN A FINANCIAL MARKET", *JOURNAL OF FORECASTING*, 15, 37-48.
- FEATHER, P.M.; KAYLEN, M.S. (1989): "CONDITIONAL QUALITATIVE FORECASTING", *AMERICAN JOURNAL OF AGRICULTURAL ECONOMICS*, 71, 195-201.
- FRIEDMAN, J.H. (1977): "A RECURSIVE PARTITIONING DECISION RULE FOR NONPARAMETRIC CLASSIFICATION" *IEEE TRANSACTIONS ON COMPUTERS*, 404-509
- FRYDMAN, H.; ALTMAN, E.; KAO, D. (1985): " INTRODUCING RECURSIVE PARTITIONING FOR FINANCIAL CLASSIFICATION: THE CASE OF FINANCIAL DISTRESS". *THE JOURNAL OF FINANCE*, 269-291.
- JENSON, H. (1992): "USING NEURAL NETWORKS FOR CREDIT SCORING", *MANAGERIAL FINANCE*, 18, 15-26.
- KAMSTRA, M.; KENNEDY, P. (1998): "COMBINING QUALITATIVE FORECASTS USING LOGIT" *INTERNATIONAL JOURNAL OF FORECASTING*, 14, 83-93.
- KUMAR, A.; OLMEDA, I (1999): "A STUDY OF COMPOSITE OR HYBRID CLASSIFIERS FOR KNOWLEDGE DISCOVERY", *INFORMS JOURNAL OF COMPUTING*, PRÓXIMA PUBLICACIÓN.
- MARAIS, M.L; PATELL, J.; WOLFSON, M. (1984): "THE EXPERIMENTAL DESIGN OF CLASSIFICATION MODELS: AN APPLICATION OF RECURSIVE PARTITIONING AND BOOTSTRAPPING TO COMMERCIAL BANK LOAN CLASSIFICATIONS". *JOURNAL OF ACCOUNTING RESEARCH*, 87-114.
- MAROSE, R. (1990): "A FINANCIAL NEURAL NETWORK APPLICATION", IN R. TRIPPI AND E. TURBAN (EDS.), *NEURAL NETWORKS IN FINANCE AND INVESTING*, CHICAGO 1993, 75-83.
- MYERS, J; FORGY, E. (1963): "THE DEVELOPMENT OF NUMERICAL CREDIT EVALUATION SYSTEMS" *JOURNAL OF AMERICAN STATISTICAL ASSOCIATION* 58, 799-806.
- ORGLER, Y. (1970): "A CREDIT SCORING MODEL FOR COMMERCIAL LOANS" *JOURNAL OF MONEY, CREDIT & BANKING* 2, 435-445.
- PIRAMUTHU, S.; SHAW, M.; GENTRY, J. (1994): "A CLASSIFICATION APPROACH USING MULTI-LAYERED NEURAL NETWORKS", *DECISION SUPPORT SYSTEMS*, 11, 509-522.
- QUINLAN, J.R. (1987): "SIMPLIFYING DECISION TREES". *INTL. JOURNAL MAN-MACHINE STUDIES* 27, 221-234.
- SERRANO, C. (1994): *LAS REDES NEURONALES ARTIFICIALES EN EL ANÁLISIS DE LA INFORMACIÓN CONTABLE*. TESIS DOCTORAL. DEPARTAMENTO DE CONTABILIDAD Y FINANZAS DE LA UNIVERSIDAD DE ZARAGOZA.
- STONE, M.(1974): "CROSS-VALIDATORY CHOICE AND ASSESSMENT OF STATISTICAL PREDICTIONS", *JOURNAL OF THE ROYAL STATISTICAL SOCIETY*, 36, 11-144.
- TAM, K.; KIANG, M. (1992): " MANAGERIAL APPLICATIONS OF NEURAL NETWORKS: THE CASE OF BANK FAILURE PREDICTIONS" *MANAGEMENT SCIENCE*, 38, 926-947.

TABLA 2

COSTE	HÍBRIDO CART		HÍBRIDO LOGIT		DA		LOGIT		CART		RLP		RNA	
	% ERROR	COSTE	%ERRO R	COSTE	%ERRO R	COSTE	%ERRO R	COSTE	%ERRO R	COSTE	%ERRO R	COSTE	%ERRO R	COSTE
1/100	24.44	4.2	43.33	0.4		9		6		11		7		4
1/50	24.44	4.4	41.11	0.7		9		6.1		11		7.1		4.1
1/10	21.11	4.6	22.22	2.9		9.2		6.5		11.1		7.5		4.5
1/5	23.33	8.2	17.78	4		9.4		7		11.2		8		5
1/4	21.11	8.5	15.55	5		9.5		7.3		11.3		8.3		5.3
1/3	20.00	9.3	14.44	5.6		9.7		7.7		11.3		6.7		5.7
1/2	16.66	10	13.33	7.5		10		8.5		11.5		9.5		6.5

COSTE	HÍBRIDO CART		HÍBRIDO LOGIT		DA		LOGIT		CART		RLP		RNA	
	% ERROR	COSTE	%ERROR R	COSTE	%ERROR R	COSTE	%ERROR R	COSTE	%ERROR R	COSTE	%ERROR R	COSTE	%ERROR R	COSTE
1	18.88	17	13.33	12	12.22	11	12.22	11	13.33	12	12.3	11	10	9
5	17.77	36	14.44	29		19		31		16		32		29
10	18.88	53	22.22	38		29		56		21		57		54
15	20.00	88	21.11	61		39		81		26		82		79
20	18.88	93	32.22	29		49		106		31		107		104
25	18.88	113	35.55	32		59		131		36		132		129
50	22.22	265	55.55	50		109		256		61		257		254
100	24.44	616	55.55	50		209		506		111		507		504

TABLA 3

COSTE	HÍBRIDO - CART		HÍBRIDO LOGIT		DA		LOGIT		CART		RLP	
	% ERROR	COSTE	% ERROR	COSTE	% ERROR	COSTE	% ERROR	COSTE	% ERROR	COSTE	% ERROR	COSTE
1/100	16.5	19.9	26.09	5.8		38.1		30.2		40.1		37.1
1/50	15.65	16.8	23.19	7.5		38.2		30.3		40.2		37.3
1/10	16.52	19.2	17.39	15.9		38.8		31.7		40.8		38.3
1/5	16.81	27.6	15.36	23.4		39.2		33.4		41.6		39.6
1/2	15.36	37	14.78	38		42		38.5		44		43.5
1	16.23	56	14.20	49	13.33	46	13.62	47	13.91	48	14.49	50
5	26.08	154	18.26	91		78		115		80		102
10	23.47	171	27.82	123		118		200		120		167
15	25.50	242	32.75	155		158		285		160		232
20	25.50	297	38.55	171		198		370		200		297
25	26.66	356	45.21	156		238		455		240		362
50	26.66	631	55.65	192		438		880		440		687
100	26.66	1181	55.65	192		838		1730		840		1337

TABLA 4

COSTE	HÍBRIDO - CART		HÍBRIDO LOGIT		DA		LOGIT		CART		RLP	
	% ERROR	COSTE	% ERROR	COSTE	% ERROR	COSTE	% ERROR	COSTE	% ERROR	COSTE	% ERROR	COSTE
1/100	17.39	21.4	27.83	5.9		37.1		30.2		39.1		37.1
1/50	17.1	19.8	25.50	6.7		37.2		30.3		39.2		37.3
1/10	16.52	21.9	17.10	15.8		37.8		31.7		39.8		38.4
1/5	15.07	26.4	16.23	25.6		38.6		33.4		40.6		39.6
1/2	13.91	37	14.49	37.5		41		38.5		43		43.5
1	13.33	46	14.78	51	13.04	45	13.62	47	13.62	47	14.49	50
5	23.19	132	17.97	94		77		115		79		102
10	25.8	215	27.54	131		117		200		119		167
15	25.50	284	31.88	166		257		275		262		279
20	24.34	331	40.00	195		332		360		337		364
25	24.92	422	48.40	167		407		445		412		449
50	25.50	774	55.65	192		437		880		439		687
100	31.88	1298	55.65	192		837		1730		935		1337