

Análisis de los determinantes del riesgo de crédito en presencia de carteras de bajo incumplimiento. Una nueva propuesta de aplicación¹

Raquel Flórez López
Universidad de León

RECIBIDO: 28 de noviembre de 2006

ACEPTADO: 7 de febrero de 2007

Resumen: Si bien Basilea II plantea la conveniencia de utilizar modelos internos o IRB para la gestión del riesgo de crédito, su desarrollo resulta complejo para las denominadas carteras de bajo incumplimiento (LDP), caracterizadas por un número reducido de datos de eventos de pérdida. La presencia de este tipo de carteras exige una rigurosa selección de los determinantes del riesgo crediticio y una metodología de análisis consistente, estructurada y bien documentada.

Este trabajo analiza las peculiaridades de las carteras LDP, proponiendo una metodología que incluye la selección de los determinantes más relevantes, el análisis de sus relaciones y la validación de los resultados obtenidos. El modelo se aplica a una muestra de bancos españoles, comparando los resultados con estudios previos y comprobando su potencial utilidad.

Palabras clave: Basilea II / Riesgo de crédito / Quiebra bancaria / Selección de variables / Modelos estadísticos / Análisis bayesiano / Aprendizaje Máquina.

Exploring for the Determinants of Credit Risk in Presence of Low Default Portfolios. A New Approach

Abstract: Although Basel II agreement establishes the convenience of internal-rating based (IRB) models, the use of this approach becomes very complicated for the so-called low-default portfolios (LDP) where banks may have limited loss data. In presence of such LDP, a rigorous selection of the determinants of credit risk must be established, together to a consistent, well-structured and adequately documented methodology of analysis.

This article analyses the typical characteristics of LDP, and proposes a methodology that includes the selection of main credit risk drivers, the analysis of internal relationships between them and the validation of final results. This model is applied to a sample of Spanish banks; results are compared to previous literature and the potential utility is observed.

Key Words: Basel II / Credit risk / Bank failure / Feature selection / Statistical models / Bayesian analysis / Machine learning.

INTRODUCCIÓN

La gestión del riesgo de crédito constituye una decisión crítica para las entidades financieras, que afecta a los resultados y a la posición de solvencia a corto y medio plazo.

Tradicionalmente los reguladores financieros han tratado de controlar este riesgo mediante el establecimiento de sistemas estandarizados de control, basados en evaluaciones externas. No obstante, el nuevo entorno económico requiere modelos de gestión más sofisticados y flexibles, cuyo desarrollo constituye uno de los principales retos del sector financiero para los próximos años (Carey, 2001).

Recientemente, el nuevo Acuerdo de Capital de Basilea II apuntó la conveniencia de gestionar el riesgo de crédito mediante sistemas internos (propuesta IRB), empleando los datos propios de cada banco para la estimación de los diversos componentes del riesgo de crédito y, en particular, para el análisis de la probabilidad de incum-

plimiento (PD) de las distintas carteras (CSBB, 2004 y 2005a).

Ahora bien, la aplicación del enfoque IRB exige que los bancos proporcionen definiciones consistentes y bien estructuradas de sus sistemas internos y de gestión, al tiempo que deben demostrar la robustez y eficiencia de los modelos aplicados. El cumplimiento de estos requisitos resulta particularmente complejo en el caso de carteras de bajo incumplimiento ("low default portfolios" o LDP), caracterizadas por el reducido tamaño de las muestras, la compleja cuantificación de su riesgo de crédito y la difícil validación estadística de los sistemas de gestión a ellas asociados.

Si bien habitualmente se ha prestado poca atención al problema de la selección de las variables explicativas del riesgo de crédito, siendo habitual la consideración de los indicadores más utilizados en estudios previos o la definición *ad hoc* de los determinantes por parte del propio investigador, en presencia de carteras LDP la se-

lección de atributos adquiere una mayor relevancia, dado que el reducido tamaño de la muestra sólo permite un número limitado de indicadores, que han de ser fuertemente relevantes y suficientemente explicativos.

Este trabajo propone una nueva metodología para el análisis de los determinantes del riesgo de crédito, considerando tres fases diferenciadas: la selección de los atributos más explicativos de la PD, el análisis de las relaciones funcionales existentes entre los atributos mediante el desarrollo de diversos modelos de clasificación (estadísticos y de Aprendizaje Máquina) y la validación de los resultados obtenidos a través de técnicas de remuestreo (bootstrap).

Para contrastar el modelo se analiza una muestra de 52 bancos españoles, fracasados y sanos durante el periodo 1978-1982, que ya ha sido objeto de numerosos estudios previos. La comparación de los resultados obtenidos con la literatura en la materia permitirá comprobar si la metodología propuesta resulta o no adecuada para la caracterización de las carteras LDP, en general, y de la insolvencia bancaria, en particular.

El trabajo se estructura como sigue. En la Sección 2 se realiza un breve análisis del nuevo entorno regulatorio, destacando las peculiaridades de las carteras LDP ante el Acuerdo de Basilea II. La Sección 3 presenta la nueva metodología para el análisis de los determinantes del riesgo de crédito, cuyo contraste empírico se realiza en la Sección 4 respecto a una muestra de bancos españoles. Por último, la Sección 5 incluye las principales conclusiones obtenidas, así como las líneas de investigación futuras a desarrollar.

PROBLEMÁTICA DEL RIESGO DE CRÉDITO PARA CARTERAS DE BAJO INCUMPLIMIENTO ANTE BASILEA II

En Junio de 2004, el Comité de Supervisión Bancaria del Banco Internacional de Pagos de Basilea aprobó un nuevo acuerdo, conocido como Basilea II, que define un nuevo marco riesgo-sensible destinado a mejorar la gestión del conjunto de riesgos financieros a los que se enfrentan los bancos internacionales (crédito, mercado y operacional, principalmente).

En el ámbito del riesgo de crédito, Basilea II introduce una nueva propuesta de gestión, cono-

cida como “propuesta basada en calificaciones internas o IRB” (“*Internal rating-based approach*”), que considera que las medidas internas de riesgo de crédito desarrolladas por los propios bancos pueden servir como punto de partida para el establecimiento de distintas categorías de riesgos y para el cálculo de los requisitos mínimos de capital.

La estimación de la probabilidad de incumplimiento (PD) resulta el componente más crítico para el desarrollo de modelos internos IRB (CSBB, 2004, pfo. 391). En el caso particular de exposiciones frente a empresas, soberanos y bancos, el Nuevo Acuerdo propone el empleo de alguna o varias de las siguientes técnicas: 1) experiencia interna de incumplimiento, relativa al empleo de datos históricos sobre fallidos; 2) asociación a datos externos, referente a la asociación de los grados internos de riesgo de crédito con alguna escala externa, como las desarrolladas por agencias de rating especializadas; y 3) modelos estadísticos de incumplimiento, relacionados con el empleo de técnicas estadísticas de predicción de insolvencia, suficientemente documentadas.

En particular, la aplicación de modelos estadísticos de incumplimiento, como los modelos de credit scoring, representa una alternativa muy atractiva en la práctica, aunque precisa demostrar un alto nivel de capacidad predicativa y justificar suficientemente la selección de los determinantes utilizados (Martín Marín y Trujillo Ponce, 2004).

Si bien el documento marco del Acuerdo de Basilea II recoge las reglas generales para la gestión del riesgo de crédito de la generalidad de las posiciones, la importancia y peculiares características de las denominadas “carteras de bajo incumplimiento” ha motivado la reciente publicación de un documento especializado en este ámbito (CSBB, 2005b).

La definición de las “carteras de bajo incumplimiento” (“*low-default portfolios*” o LDP) es compleja, no pudiéndose establecer un límite que clasifique de forma inequívoca las posiciones de una entidad en “de bajo incumplimiento” y “de no bajo incumplimiento”.

No obstante, entre las carteras que podrían estar próximas a las LDP destacan particularmente las siguientes: 1) carteras que históricamente han experimentado un número reducido

de incumplimientos, y que suelen considerarse (aunque no siempre) de bajo riesgo; *por ejemplo*, exposiciones frente a soberanos, bancos, compañías aseguradoras o empresas con ratings muy elevados; 2) carteras de tamaño reducido, bien a nivel global o respecto a un segmento particular del banco; *por ejemplo*, empresas de transporte, financiación de proyectos, etc.

Una de las carteras LDP más significativas para las entidades financieras es la referente a las exposiciones frente a otros bancos, como consecuencia del volumen de las operaciones interbancarias y de la emisión de instrumentos financieros por parte de estas entidades.

Las carteras tipo LDP se caracterizan por la difícil cuantificación de su riesgo de crédito y la complicada validación de los sistemas IRB a ellos asociados, dado que:

- La estimación de la PD a partir de datos históricos podría no verificar el requisito de confiabilidad estadística establecido por el Nuevo Acuerdo, debido al escaso número de observaciones utilizadas.
- La validación del modelo mediante la comparación temporal de la PD estimada y la PD realmente acontecida (“*backtesting*”) podría proporcionar evidencias relacionadas con una escasa precisión del sistema IRB.

Estas limitaciones obligarían a emplear el enfoque estándar basado en evaluaciones externas para la gestión de los LDP, lo que simplificaría enormemente los resultados obtenidos, máxime si se considera que estas carteras constituyen una porción muy sustancial de los activos bancarios. Ante esta situación, CSBB (2005b) establece un conjunto de recomendaciones que eviten la exclusión de las carteras LDP del enfoque IRB, con el objetivo de incrementar la riqueza de los datos disponibles, facilitar la cuantificación del riesgo de crédito y mejorar la validación de los sistemas en ausencia de datos históricos suficientes.

Entre ellas, destacan las propuestas para la mejor cuantificación del riesgo de crédito, tales como:

- Incremento del número de datos internos, mediante la combinación de segmentos con similares características de riesgo.

- Combinación de distintas categorías de rating y estimación de la probabilidad conjunta de incumplimiento de todas ellas. En particular, para los bancos que asocien sus grados internos con escalas externas de calificación podría ser conveniente combinar las categorías AAA y AA, ó AAA, AA y A, para ciertas posiciones LDP, o incluso distinguir únicamente entre posiciones en situación de cumplimiento o incumplimiento (problema de clasificación binaria).
- Estimación de la PD a partir de datos con horizontes temporales superiores a un año.

Respecto a la validación de los sistemas IRB desarrollados para carteras LDP, se plantea, entre otras medidas, el uso de herramientas de benchmarking, el empleo de juicios expertos y la combinación de distintas subcarteras LDP en una cartera global, que permita comparar la probabilidad media predicha de incumplimiento con la tasa de fallidos realmente observada.

Por último, en presencia de carteras LDP, los supervisores analizarán de forma preferente si el banco ha prestado una atención razonable a mejorar la consistencia de sus datos internos, para lo que debe haber realizado un conjunto de etapas bien documentadas. Si el supervisor considera que las fases ejecutadas no han sido suficientes para el tratamiento de los datos escasos, podrá realizar distintas actuaciones que dependerán de las circunstancias de la entidad, la jurisdicción en la que opere y las características de la cartera.

PROPUESTA PARA EL ANÁLISIS DE LOS DETERMINANTES DEL RIESGO DE CRÉDITO

En este trabajo se presenta una nueva metodología para el análisis y modelado de los determinantes del riesgo de crédito que, si bien podría utilizarse respecto a la generalidad de problemas de clasificación, resulta particularmente relevante en el caso de carteras de bajo incumplimiento.

La metodología que se propone presenta una arquitectura en tres etapas, como sigue: 1) *selección de variables explicativas*; 2) *modelos de análisis multivariante*, a partir de las variables

previamente seleccionadas; y 3) *evaluación y validación de los modelos*, mediante el empleo de diversos índices y técnicas de remuestreo (bootstrap). Por último, los resultados obtenidos son analizados y comparados entre sí al objeto de conocer las estrategias que mejor caracterizan a la muestra analizada.

SELECCIÓN DE VARIABLES EXPLICATIVAS

La ausencia de una teoría inequívoca sobre las causas del fracaso empresarial ha llevado a desarrollar un proceso de selección de variables explicativas con la finalidad de obtener un subconjunto de potenciales determinantes que sea necesario y suficiente para la caracterización del riesgo de crédito.

En las últimas décadas, el problema de la selección de atributos ha recibido una considerable atención por parte de los investigadores vinculados al ámbito del Aprendizaje Máquina, habiéndose propuesto diversos algoritmos en este ámbito². John *et al.* (1994) clasifican las distintas propuestas en dos grandes grupos, atendiendo al tipo de función utilizada para analizar la eficacia del modelo: algoritmos tipo “filter” y algoritmos tipo “wrapper”; asimismo, podría identificarse una tercera categoría de técnicas, conocida como algoritmos tipo “embedded” (Blum y Langley, 1997).

Los algoritmos “filter” tratan de evaluar la relevancia de un conjunto de variables explicativas a partir exclusivamente de la muestra analizada, sin considerar los modelos que se desarrollarán para su análisis (Blum y Langley, 1997). Entre sus principales ventajas destaca el reducido tiempo de computación que necesitan y su independencia respecto a los modelos multivariantes que utilice el usuario. Por el contrario, entre sus inconvenientes destaca la no consideración de la interdependencia que pudiera existir entre los datos y el algoritmo de inducción a utilizar. En este trabajo, se propone la aplicación de dos algoritmos “filter” diferentes:

- La variante Relief-F del algoritmo Relief (Kira y Rendell, 1992; Kononenko, 1997). Esta propuesta emplea un esquema de votación basado en el algoritmo del k -vecino más cercano para

la evaluación de las variables explicativas, generando un ranking final de atributos relevantes. Su principal limitación radica en la selección del punto de corte discriminatorio entre atributos explicativos y no explicativos, que habitualmente es definido por el propio usuario; como regla práctica, suelen seleccionarse aquellos atributos cuyos pesos superan el valor 0 o 0,1. Asimismo, puede presentar problemas en presencia de atributos altamente correlacionados entre sí.

- El algoritmo WACSA (Florez, 2002). Esta propuesta parte de los fundamentos de los algoritmos Relief-F y EUBAFES (Scherf y Brauer, 1997), si bien trata de solventar sus limitaciones mediante el análisis de las correlaciones parciales entre los atributos potenciales. Asimismo, se emplea un algoritmo de búsqueda diferente, que combina el Enfriamiento Simulado, el algoritmo del k -vecino más cercano y las cadenas de Markov. Finalmente se obtiene una lista de atributos ordenados según su relevancia, existiendo de nuevo el problema de la selección del punto de corte discriminatorio; como regla práctica puede seleccionarse aquel punto de corte tal que $w_i \geq k \cdot w_{i+1}$ ($k=2$ o $k=3$ generalmente), siendo w_i el nivel de relevancia asociado al k -ésimo atributo (Sestito y Dillon, 1994).

Las técnicas “wrapper” evalúan el desempeño de distintos subconjuntos de atributos, definidos mediante algún algoritmo de inducción, a partir de su eficacia respecto a la clasificación de la muestra analizada para un modelo multivariante determinado (análisis discriminante, árboles de decisión, etc.) Si bien existe evidencia de que estos algoritmos obtienen un desempeño superior a los modelos “filter”, presentan un elevado coste computacional, al tiempo que pueden verse afectados por sobreajuste al emplear el mismo algoritmo para la selección de atributos y la clasificación de patrones (John y Kohavi, 1995).

- En este trabajo se propone una técnica “wrapper” que lleva a cabo la selección de variables mediante la combinación de técnicas bayesianas y cadenas de Markov a través del modelo conocido como “Gibbs Sampler” (Dellaportas

et al., 1997; Kuo y Mallick, 1998). Se han considerado diversas variantes del algoritmo, empleando distribuciones normales, logísticas y de Bernoulli para la definición de la distribución a priori de la variable endógena y las relaciones entre las variables exógenas y endógena. Los resultados obtenidos proporcionan estimaciones a posteriori de la probabilidad de inclusión de cada atributo en los modelos de predicción de la insolvencia empresarial, si bien el punto de separación entre los atributos relevantes y no debe definirse por el usuario; como regla práctica suelen seleccionarse aquellos atributos cuya probabilidad a posteriori excede el 50%.

Los algoritmos “*embedded*” se caracterizan por integrar el proceso de selección de atributos dentro del propio algoritmo de clasificación. Este es el caso, por ejemplo, de los árboles de decisión o de algunos modelos de poda de Redes Neuronales Artificiales (RNAs) como los aplicados en este trabajo:

- El método de las unidades no contributivas (NCU), que emplea medidas estadísticas para la detección de las unidades de la red que no contribuyen al aprendizaje global del sistema. Para cada atributo se obtiene un índice de relevancia final, calculado mediante la aplicación del método “HVS” (White, 1989).
- El algoritmo “Optimal Brain Surgeon” (OBS), que aproxima el cambio en la función de error de la red cuando se eliminan determinados pesos del sistema (Stork y Hassibi, 1993), empleando de nuevo el método “HVS” para la obtención de una medida de relevancia de cada atributo.
- La aproximación a las derivadas parciales de redes neuronales hacia adelante (DP), que genera una estimación no lineal de las derivadas parciales de la red respecto a cada atributo potencial, considerando el valor medio de los datos analizados. Los valores absolutos de tales derivadas permiten obtener una medida significativa de la relevancia de cada indicador (McNelis, 1997).

En este trabajo, los modelos NCU, OBS y DP se han combinado en un índice global medio pa-

ra las redes neuronales, al objeto de reducir el sesgo del proceso final de selección de atributos. En este punto cabe destacar que los modelos que incluyen algoritmos de selección de variables tipo “*embeded*” suelen presentar problemas en presencia de un elevado número de atributos irrelevantes, reduciéndose sustancialmente el desempeño conseguido (Kira y Rendell, 1992). Asimismo, la presencia de correlaciones entre los atributos puede afectar significativamente a los resultados (Blum y Langley, 1997), al tiempo que pueden producirse fenómenos de sobreajuste. Como consecuencia, estos modelos suelen combinarse con algoritmos “*filter*” (aplicándose en una segunda fase a partir de los atributos seleccionados por estos últimos) y/o algoritmos “*wrapper*” (utilizándose para definir la función de evaluación de los distintos subconjunto de atributos).

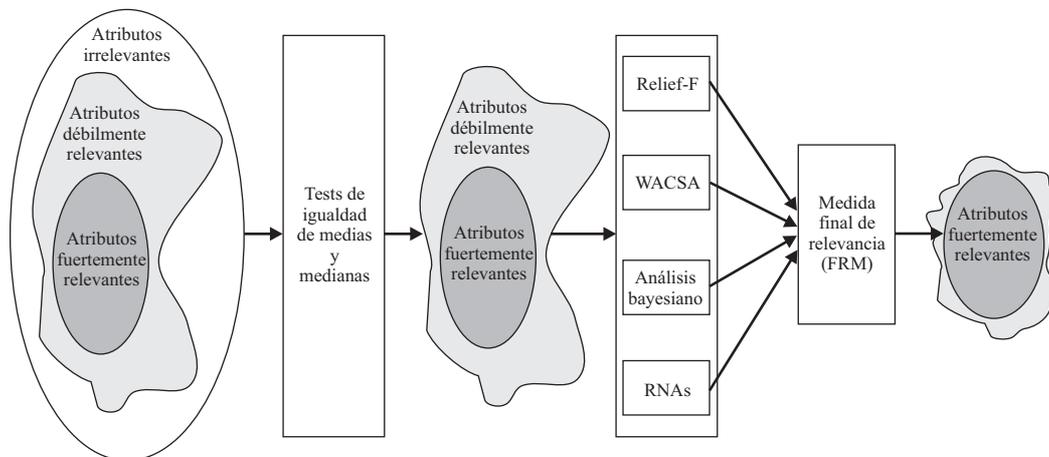
A partir de los modelos previos, en este trabajo se plantea un proceso híbrido de selección de variables, como sigue (figura 1)³:

- Paso 1. Aplicación de tests estadísticos de igualdad de medias y medianas para dos grupos (empresas fracasadas y no fracasadas), respecto a cada uno de los potenciales atributos explicativos. Aquellas variables que no rechacen las hipótesis nulas (H_0) de igualdad de medias y de igualdad de medianas se eliminan del estudio.
- Paso 2. Aplicación de los algoritmos Relief-F (R), WACSA (W), análisis bayesiano (B), y Redes Neuronales Artificiales (A) sobre los atributos no eliminados en el paso 1, obteniendo un índice de relevancia para cada modelo y variable.
- Paso 3. Normalización de cada índice de relevancia (R , W , B , A) respecto al intervalo $[0,1]$. De esta forma, se obtiene un indicador normalizado para cada i -ésimo atributo y método (R_{iN} , W_{iN} , B_{iN} , A_{iN}).
- Paso 4. Desarrollo de una medida final de relevancia (FRM) para cada i -ésimo atributo, como sigue:

$$FRM_i = \frac{1}{4} \cdot (R_{iN} + W_{iN} + B_{iN} + A_{iN})$$

- Paso 5. Obtención de un listado de medidas FRM_i , ordenadas de forma decreciente. El sub-

Figura 1.- Esquema del proceso de selección de variables



FUENTE: Elaboración propia.

conjunto final de atributos se elige de acuerdo con el punto de corte tal que $FRM_i \geq k.FRM_{i+1}$, siendo recomendable el empleo de valores próximos a $k=2$ ó $k=3$ (Sestito y Dillon, 1994).

MODELOS DE ANÁLISIS MULTIVARIANTE

Al objeto de analizar los determinantes del riesgo de crédito y, en particular, las variables explicativas del proceso de insolvencia vinculado con las entidades bancarias españolas, se han considerado distintos modelos de análisis multivariante tanto en el ámbito estadístico como de Aprendizaje Máquina: análisis discriminante, análisis logit y árbol de decisión C4.5. La selección de estos modelos se debe a su popularidad en la literatura en la materia, lo que facilitará el análisis comparativo de la eficacia de la metodología propuesta.

El *análisis discriminante* (MDA) constituye un modelo estadístico de clasificación empleado para el análisis de una variable dependiente categórica definida por dos o más clases, utilizando como predictores un conjunto de K variables independientes medidas, al menos, en escala de intervalo. En su versión lineal, el MDA obtiene $M-1$ funciones discriminantes, siendo M el número de clases y N el conjunto de datos analizados:

$$Z_{im} = \alpha_m + \sum_{k=1}^K \beta_{mk} X_{ik}, \quad i = 1, \dots, N \quad m = 1, \dots, M-1$$

Entre las asunciones del modelo destaca la normalidad multivariante de las variables inde-

pendientes y la igualdad de las matrices de covarianzas entre las clases.

El *análisis logit* constituye una de las técnicas estadísticas multivariantes más utilizadas para el estudio de variables dependientes no métricas, siendo su forma funcional la siguiente:

$$\rho_i = \text{Prob}(Y_i = 1 / X_{1i}, \dots, X_{Ki}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + \dots + \beta_K X_{Ki})}}$$

donde Y es la variable dependiente (no métrica) y K es el conjunto de variables independientes, métricas o no, (X_1, \dots, X_K) . En este punto suele ser habitual la denominación:

$$Z_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_K X_{Ki}$$

tal que:

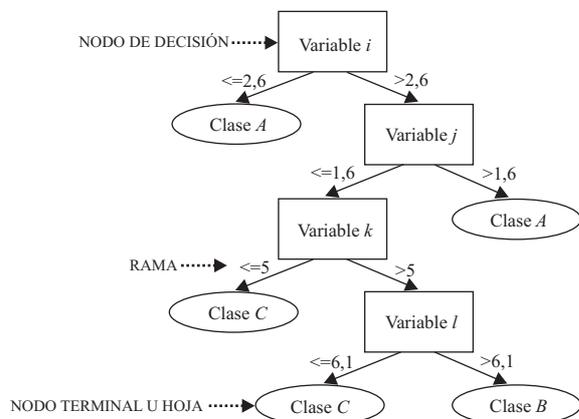
$$\rho_i = \text{Prob}(Y_i = 1 / X_{1i}, \dots, X_{Ki}) = \frac{1}{1 + e^{-Z_i}}$$

Mediante el análisis logit se consigue que la relación entre la variable independiente y los atributos explicativos no sea lineal, adoptando una forma de S entre 0 y 1. Este hecho resulta fundamental para comprender la validez del análisis logit, que parte de la hipótesis de que cambios de idéntica cuantía en las variables exógenas no afectan por igual a la variable endógena, lo que resulta particularmente oportuno en el caso de problemas económico-financieros.

La regresión logística se encuentra restringida en su forma básica a dos grupos, si bien en formulaciones alternativas puede considerar más categorías (análisis multilogit).

Los árboles de decisión son modelos de clasificación jerárquica y secuencial que dividen recursivamente un conjunto de datos (N). Cada árbol integra tres componentes básicos: los nodos de decisión, las ramas, y los nodos terminales (hojas) (figura 2).

Figura 2.- Modelo de árbol de decisión



FUENTE: Elaboración propia.

El árbol se construye mediante la partición determinista de la muestra inicial N (incluida en el nodo inicial) en varios subconjuntos descendientes (definidos en los nodos-hijo) a través de algún test lógico, conectándose los distintos nodos mediante ramas. El proceso de partición continúa hasta que se cumple una determinada condición de parada, dando lugar a un conjunto de nodos terminales (hojas) designados por una etiqueta de clase. Finalmente, cada rama desde el nodo raíz hasta cada nodo terminal puede interpretarse como una regla de decisión (Breiman *et al.*, 1984).

Existen distintos algoritmos de inducción de árboles de decisión (IDT) que pueden clasificarse atendiendo a diversos criterios, tales como el tipo de hiperplano utilizado para dividir los datos, los tests lógicos empleados en los nodos de decisión o los esquemas de simplificación (poda) aplicados al modelo.

En este trabajo, se considera uno de los modelos de árboles de decisión más empleados en el ámbito económico-financiero, el algoritmo C4.5

(Quinlan, 1993), que propone dos criterios diferenciados para la partición de los datos (“criterio de ganancia” y “criterio del ratio de ganancia”), emplea un mecanismo de detención del algoritmo basado en el test Chi-cuadrado de significación estadística y utiliza un proceso de poda pesimista.

EVALUACIÓN Y VALIDACIÓN DE LOS MODELOS

La evaluación de la capacidad predictiva de los distintos modelos de clasificación puede realizarse utilizando distintas medidas de desempeño, tales como la matriz de clasificación, el porcentaje de observaciones correctamente caracterizadas, el criterio de información de Akaike (medida AIC) o estadísticos basados en la medida R -cuadrado.

En este trabajo se propone el empleo de las matrices de clasificación derivadas de cada método si bien, dado que se analizarán muestras de tamaño reducido procedentes de una cartera de bajo incumplimiento, los resultados obtenidos podrían verse afectados por sobreajuste (“overfitting”). Al objeto de evitar esta limitación, se considerarán dos estrategias adicionales basadas en técnicas de bootstrap (alternativas 0.632E y 0.632E+ para problemas de clasificación) (Efron, 1986; Efron y Tibshirani, 1993 y 1995).

De acuerdo con la alternativa 0.632E, a partir de la muestra inicial de N observaciones se definirán B submuestras de bootstrap con reemplazamiento de tamaño N , desarrollándose el modelo de clasificación para cada una de ellas y estimándose el error cometido respecto a los casos no incluidos en cada submuestra. Finalmente, el error esperado real se obtiene como sigue:

$$\hat{Err}_{.632E} = 0.368 \times \overline{err} + 0.632 \times E_0$$

siendo \overline{err} el error generado por el clasificador para la muestra original (error aparente o de sustitución) y E_0 el error medio derivado de las distintas submuestras de bootstrap.

Por lo que respecta al número de submuestras de bootstrap a generar, a efectos estadísticos se recomienda un número no inferior a 25 y a efectos prácticos un número no superior a 200, sien-

do 50 un número adecuado para la mayoría de los casos (Efron y Tibshirani, 1995).

Posteriormente, Efron y Tibshirani (1995) desarrollaron una nueva variante, conocida como 0.632+E, que trata de evitar aún más los problemas de sobreajuste mediante el establecimiento de ponderaciones variables para el error de sustitución y el error medio de bootstrap, basadas en las probabilidades a priori y a posteriori de las distintas clases analizadas.

APLICACIÓN EMPÍRICA

Al objeto de evaluar la validez de la metodología propuesta para el análisis de las carteras LDP, a continuación se estudia el conocido problema de la crisis bancaria española para el periodo 1978-1982, caracterizado por una muestra de datos de tamaño muy reducido.

El problema de la insolvencia bancaria ha sido ampliamente analizado por la literatura relacionada con el fracaso empresarial, tanto en el ámbito nacional como internacional (Altman *et al.*, 1981). Si bien se han propuesto diversos modelos para su análisis, no existe una teoría inequívoca sobre las causas y síntomas del fracaso bancario (Myers, 1998), siendo habitual el empleo de diversos ratios económico-financieros como variables explicativas de la situación de insolvencia.

En el *ámbito nacional*, las investigaciones se han centrado en la crisis bancaria que tuvo lugar en España entre los años 1977 y 1985, que constituyó probablemente la crisis financiera más importante acontecida hasta el momento (Laffarga Briones *et al.*, 1985; Pina Martínez, 1989). Los anexos 1 y 2 recogen algunas de las aportaciones más relevantes en este ámbito.

Si bien todos los estudios se han centrado en el estudio de las relaciones entre las variables explicativas del fracaso bancario (definido como la intervención de la entidad por parte del Fondo de Garantía de Depósitos) para distintos periodos de tiempo, muchos de ellos no han desarrollado procesos específicos de selección de variables, limitándose a considerar los ratios más habituales en la materia o los indicadores seleccionados previamente por otros autores [en particular, los 9 ratios definidos por Pina Martínez

(1989) han sido ampliamente utilizados]. Por su parte, las investigaciones que han incluido procesos de selección de atributos utilizan modelos sencillos basados en tests estadísticos procedentes de los modelos clasificatorios utilizados (análisis discriminante, análisis logit, regresión lineal, árboles de decisión). Este enfoque, conocido como “embedded”, puede presentar una serie de limitaciones, tales como el sobreajuste del modelo y la reducción de su capacidad de caracterización, lo que resulta particularmente relevante en el caso de carteras LDP.

MUESTRA DE DATOS Y DEFINICIÓN DE VARIABLES

En el presente trabajo se ha considerado una muestra de 52 bancos españoles para el periodo 1978-1992, dividida en dos grupos diferenciados: 1) submuestra de bancos fracasados, integrada por 26 bancos intervenidos por el Fondo de Garantía de Depósitos en el periodo analizado; 2) submuestra de bancos sanos, formada por 26 entidades, emparejadas a las anteriores, que no fueron sometidas a intervención en el periodo.

Si bien los individuos analizados no coinciden exactamente con las muestras empleadas en los principales estudios previos, las diferencias resultan poco significativas, por lo que el análisis comparativo de resultados puede servir como referencia adicional sobre la validez y utilidad de la metodología propuesta.

La tabla 1 resume las características más destacadas de la muestra analizada, cuyo detalle se recoge en el anexo 3.

Por lo que respecta a la definición de la variable dependiente, se ha considerado una variable cualitativa medida en escala nominal de acuerdo con la siguiente relación:

- Banco fracasado: 1
- Banco no fracasado (sano): 0

En cuanto a las variables independientes, la ausencia de una teoría formal sobre el fracaso bancario ha llevado a considerar un número muy amplio de ratios potencialmente explicativos (65 indicadores), definidos a partir de la teoría clásica

Tabla 1.- Características de la muestra analizada

Definición de fracaso	Intervención de la entidad por el Fondo de Garantía de Depósitos (FGD).
Horizonte temporal	Periodo 1978-1982. Se excluyen las 25 entidades del holding RUMASA, expropiadas en 1983, debido a sus peculiares características.
Muestra seleccionada	<i>Submuestra de bancos fracasados:</i> 26 bancos fracasados durante 1977-1982. Para el mismo periodo, Laffarga Briones <i>et al.</i> (1984) consideran 25 entidades (excluyen Banco de Madrid), Pina Martínez (1989) analiza 20 individuos (excluye Banco Industrial de los Pirineos, Banco de Navarra, Banco Cantábrico, Banco Meridional, Banco de Valladolid y Banco de Préstamo y Ahorro) y Serrano Cinca (1994) considera 29 bancos, si bien tres de ellos suelen considerarse fracasados en 1983 (Banco de Expansión Industrial, Banco Simeón y Banco Garriga Nogués). <i>Submuestra de bancos sanos:</i> 26 bancos no intervenidos durante 1977-1982. Estas entidades se seleccionaron mediante emparejamiento a través del tamaño (volumen medio de depósitos para los cinco años previos a la intervención del banco al que se encuentran asociadas), utilizando una relación aproximada de 1:1.
Origen de los datos	Anuario Estadístico de la Banca Privada 1976-1981. Se ha considerado la información contable pública de cada entidad para uno y dos años antes del ejercicio en el que tuvo lugar su intervención (o la intervención del banco al que se encuentra emparejada), de forma similar a Laffarga Briones <i>et al.</i> (1985) y Pina Martínez (1989). Por su parte, autores como Serrano Cinca (1994), Martín del Brío y Serrano Cinca (1993) y Bonsón Ponte <i>et al.</i> (1997a) utilizan para las empresas sanas los estados contables publicados en el último año de la muestra (1982).

FUENTE: Elaboración propia.

ca del análisis financiero y de las particulares características del negocio bancario, incluyendo asimismo las variables utilizadas por estudios previos en la materia (anexo 2). Dichos ratios se han calculado a partir de las Cuentas Anuales de las entidades analizadas, clasificándose como sigue:

- Ratios de endeudamiento y cobertura de intereses (11 ratios): Cocientes indicativos del nivel de endeudamiento de la empresa, el coste derivado de éste y la composición del pasivo de la entidad.
- Ratios de liquidez y estructura de activo (17 ratios): Indicadores relacionados con la capacidad de la entidad para devolver sus deudas y la estructura de sus inversiones.
- Ratios de rotación y actividad (11 ratios): Cocientes relativos a la capacidad de la entidad para obtener ingresos a partir de un nivel específico de activos e indicadores relacionados con el nivel de utilización de las inversiones.
- Ratios de rentabilidad, generación de recursos y costes (21 ratios): Indicadores relacionados con los resultados obtenidos por la entidad, los recursos finales generados y los costes soportados para la obtención de éstos.
- Otros ratios de productividad (5 ratios): Indicadores relacionados con la productividad relativa de los factores productivos (fundamentalmente trabajo y capital, aproximados por el número de empleados y el número de oficinas operativas).

SELECCIÓN DE ATRIBUTOS

A partir de los 65 ratios inicialmente definidos para uno y dos años antes del fracaso, se procedió a seleccionar los atributos más explicativos para cada ejercicio mediante la aplicación del modelo previamente presentado.

De esta forma, en primer lugar, se contrastó la hipótesis de igualdad de medias y medianas respecto a cada indicador⁴ (paso 1), rechazándose 23 ratios para el primer año y 22 indicadores para el segundo ejercicio previo al fracaso.

Al objeto de conseguir un número más reducido de atributos relevantes, se aplicaron los algoritmos Relief-F, WACSA, análisis bayesiano y Redes Neuronales Artificiales (paso 2), normalizándose los resultados de cada técnica (R, W, B, A) en el intervalo $[0,1]$ (paso 3). Por último, se calculó la medida final de relevancia FRM previamente definida (paso 4), recogiendo en la tabla 2 los indicadores más significativos para cada ejercicio, ordenados de mayor a menor según su relevancia (paso 5).

A partir de los resultados previos, se seleccionaron como puntos de corte los valores $FRM_i/FRM_{i-1}=1,485$ (año 1) y $FRM_i/FRM_{i-1}=1,130$ (año 2) que, si bien resultan ligeramente inferiores a los recomendados por Sestito y Dillon (1994), son los más adecuados para las muestras analizadas, permitiendo obtener un conjunto reducido de atributos relevantes:

Tabla 2.- Resultados del proceso de selección de variables

AÑO 1			AÑO 2		
Ratio	FRM	FRM _t /FRM _{t-1}	Ratio	FRM	FRM _t /FRM _{t-1}
<i>RENT15</i>	0,578	1,175	<i>END8</i>	0,647	1,467
<i>END8</i>	0,492	1,160	<i>LIQ17</i>	0,441	1,007
<i>RENT9</i>	0,424	1,181	<i>LIQ15</i>	0,438	1,007
<i>LIQ11</i>	0,359	1,029	<i>LIQ6</i>	0,435	1,030
<i>LIQ13</i>	0,349	1,485	<i>RENT4</i>	0,422	1,052
<i>END3</i>	0,235	1,017	<i>LIQ7</i>	0,412	1,051
<i>RENT1</i>	0,231	1,022	<i>RENT12</i>	0,392	1,048
<i>RENT4</i>	0,226	1,009	<i>RENT9</i>	0,374	1,130
<i>RENT6</i>	0,224	1,004	<i>LIQ5</i>	0,331	1,047
<i>RENT16</i>	0,223	1,005	<i>LIQ3</i>	0,316	1,003
<i>RENT17</i>	0,222	1,047	<i>ROT8</i>	0,315	1,019
<i>END10</i>	0,212	1,005	<i>RENT2</i>	0,309	1,055
<i>LIQ3</i>	0,211	1,024	<i>ROT11</i>	0,293	1,021
<i>LIQ1</i>	0,206	1,005	<i>ROT7</i>	0,287	1,055
<i>RENT13</i>	0,205		<i>LIQ8</i>	0,272	

NOTA: Sólo se incluyen los 15 indicadores más relevantes. En negrita se indica el punto de corte.

FUENTE: Elaboración propia.

• *Primer año antes del fracaso bancario (5 ratios):*

– Ratios de endeudamiento:

$$END8 = \frac{\text{Reservas}}{\text{Neto}}$$

– Ratios de liquidez:

$$LIQ11 = \frac{\text{Activos líquidos}}{\text{Pasivo exigible}}$$

$$LIQ13 = \frac{\text{Créditos} + \text{Cartera de valores}}{\text{Activo total}}$$

– Ratios de rentabilidad:

$$RENT9 = \frac{\text{Resultado neto}}{\text{Capital} + \text{Reservas}}$$

$$RENT15 = \frac{\text{"Cash flow"}}{\text{Neto}}$$

La definición de activos líquidos procede de Serrano (1994), quien considera las partidas del Activo Circulante más líquidas de la empresa. El término “cash-flow” se aproxima a los recursos procedentes de las operaciones, al calcularse como la suma del beneficio neto antes de intereses e impuestos, las amortizaciones y las provisiones practicadas.

Como puede observarse, en el subconjunto final de variables relevantes no intervienen ratios relacionados con la rotación ni con la productividad de la entidad, sino que los indicadores vinculados con la rentabilidad, la liquidez y el endeudamiento de la empresa parecen ser los más significativos respecto a la predicción a corto plazo del fracaso bancario analizado. Si bien los ratios *LIQ11* y *LIQ13* ya se habían incluido en estudios previos, el resto de ratios resultan novedosos respecto a investigaciones anteriores.

• *Segundo año antes del fracaso bancario (8 ratios):*

– Ratios de endeudamiento:

$$END8 = \frac{\text{Reservas}}{\text{Neto}}$$

– Ratios de liquidez:

$$LIQ6 = \frac{\text{Tesorería}}{\text{Activos líquidos}}$$

$$LIQ7 = \frac{\text{Tesorería}}{\text{Activo circulante total}}$$

$$LIQ15 = \frac{\text{Créditos}}{\text{Activo total}}$$

$$LIQ17 = \frac{\text{Pasivo exigible (no bancario)}}{\text{Derechos de cobro (no bancarios)}}$$

– Ratios de rentabilidad:

$$RENT4 = \frac{\text{Resultado bruto de explotación}}{\text{Activo total}}$$

$$RENT9 = \frac{\text{Resultado neto}}{\text{Capital} + \text{Reservas}}$$

$$RENT12 = \text{Aproximación al apalancamiento financiero} = \frac{\text{Resultado neto}}{\text{Capital} + \text{Reservas}} - \frac{\text{Intereses de deudas}}{\text{Pasivo exigible}}$$

La mayoría de los ratios corresponden a indicadores de liquidez, que ganan peso específico respecto al primer año; así, destaca la inclusión

de dos indicadores altamente relacionados entre sí, *LIQ6* y *LIQ7*, lo que parece reforzar a priori la importancia de los ratios de tesorería respecto a la predicción de la insolvencia empresarial a dos ejercicios vista. Asimismo, cabe destacar la inclusión del indicador de apalancamiento financiero *RENT12* y la gran relevancia que continúa presentando el ratio de autofinanciación *END8* y el de rentabilidad de los fondos propios *RENT9*. Como puede comprobarse, los indicadores *LIQ7* y *LIQ15* ya habían sido incluidos por algunos estudios previos, mientras que los cuatro ratios restantes resultan novedosos.

MODELOS ESTADÍSTICOS MULTIVARIANTES

A partir de los indicadores previos, se desarrollaron dos modelos multivariantes para cada uno de los años analizados: análisis discriminante y análisis logit, utilizando el proceso de inclusión de variables “paso a paso” al objeto de garantizar la significatividad global de los modelos obtenidos. Cabe destacar que, debido a la inexistencia de normalidad multivariante y homocedasticidad muestral, así como a la elevada correlación entre algunos pares de atributos, los resultados obtenidos por los modelos estadísticos deben interpretarse con precaución, especialmente respecto a los coeficientes atribuidos a cada variable explicativa.

Los resultados obtenidos para los *modelos discriminantes* se resumen en la tabla 3, pudiendo comprobarse la existencia de signos coherentes para los atributos y la destacada importancia de los ratios relacionados con el

“cash-flow” de la empresa, su liquidez a corto plazo y su tasa de autofinanciación (año 1), así como el apalancamiento financiero y la relación entre el volumen de depósitos (pasivo) y los préstamos concedidos (activo) de naturaleza no interbancaria (año 2).

El desempeño real de cada modelo se evaluó mediante técnicas de bootstrap, que proporcionaron una estimación del error real de generalización (tabla 4).

De esta forma, el análisis discriminante alcanza porcentajes de acierto del 87% y del 83% para la muestra total (años 1 y 2), situándose su capacidad de generalización en el 85% y 81% (años 1 y 2), porcentajes muy elevados si se considera el carácter tradicionalmente pesimista de los modelos de bootstrap.

Respecto a la muestra total, la comparación de los resultados obtenidos con los estudios previos en la materia permite observar que el modelo MDA presenta un desempeño ligeramente mejor al conseguido por Laffarga *et al.* (1986b; 1987) con idéntico número de ratios para el primer año analizado (76,6%), situándose en cifras similares respecto al segundo ejercicio. Por su parte, los resultados obtenidos se asemejan mucho a los obtenidos por Serrano Cinca (1994; 1997), si bien el número de ratios utilizados es significativamente inferior.

Ahora bien, si se considera la capacidad de generalización obtenida, el modelo propuesto resulta significativamente más robusto que las propuestas anteriores, superando ampliamente al porcentaje de acierto sobre test obtenido por Fernández y Olmeda (1995), con un 61,76% pa-

Tabla 3.-Modelos de análisis discriminante (años 1 y 2)

Año 1						
Iteración	Acción		Coeficiente estandarizado	Coeficiente no estandarizado	Lambda de Wilks	Nivel signif. crítico
	Entrada	Salida				
1	<i>RENT15</i>	-	-0,680	-4,491	0,687	0,000
2	<i>LIQ11</i>	-	-0,758	-5,792	0,498	0,000
3	<i>END8</i>	-	-0,548	-2,961	0,415	0,000
	Constante			+5,968		

Año 2						
Iteración	Acción		Coeficiente estandarizado	Coeficiente no estandarizado	Lambda de Wilks	Nivel signif. crítico
	Entrada	Salida				
1	<i>RENT12</i>	-	-0,718	-8,012	0,721	0,000
2	<i>LIQ17</i>	-	-0,665	-3,732	0,591	0,000
	Constante			+4,054		

FUENTE: Elaboración propia.

ra el primer año y por De Andrés Sánchez (2005) con un 76,92% para el mismo periodo. Por su parte, Serrano Cinca (1994) alcanza un porcentaje de aciertos del 86% para el primer ejercicio, ligeramente superior al conseguido en este trabajo, si bien utiliza la técnica jackknife que suele ser más optimista que los modelos de bootstrap (Efron, 1982).

En cuanto al *análisis logit*, los modelos obtenidos se resumen en la tabla 5.

Los resultados obtenidos por el modelo logit resultan congruentes con el análisis discriminante previo en términos de signos y variables, si

bien el error cometido respecto a las muestras de bootstrap decrece, tanto respecto al primer como al segundo año analizado (tabla 6).

Así, el análisis logit alcanza porcentajes de acierto del 95% y del 83% para la muestra total (años 1 y 2), siendo su capacidad de generalización estimada del 90% y 81% (años 1 y 2). De nuevo estos porcentajes resultan muy elevados teniendo en cuenta el carácter pesimista de los métodos de evaluación utilizados.

La comparación con los estudios previos permite observar, respecto a la muestra total, porcentajes de acierto ligeramente superiores a

Tabla 4.- Análisis discriminante - estimación del error real cometido (años 1 y 2)

AÑO 1	Error aparente (\overline{err})	E0 estimado (\hat{Err}_{E0})		Error real estimado	
		Media	Intervalo de confianza $\alpha=0.05$	Media	Intervalo de confianza $\alpha=0.05$
Bootstrap 0.632					
Error total	0,135	0,160	(0,015 , 0,305)	0,151	(0,059 , 0,242)
Error tipo I	0,154	0,147	(-0,063 , 0,367)≈[0 , 0,367)	0,150	(0,017 , 0,289)
Error tipo II	0,115	0,162	(-0,056 , 0,380)≈[0 , 0,380)	0,145	(0,007 , 0,282)
Bootstrap 0.632+					
Error total	0,135	0,160	(0,015 , 0,305)	0,151	(0,057 , 0,245)
Error tipo I	0,154	0,147	(-0,063 , 0,367)≈[0 , 0,367)	0,149	(0,013 , 0,292)
Error tipo II	0,115	0,162	(-0,056 , 0,380)≈[0 , 0,380)	0,145	(0,004 , 0,289)
Error tipo I: B. Mas Sardá, B. Industrial de Cataluña, B. Crédito e Inversiones, B. Valladolid, B. Comercial Occidental, B. Asturias.					
Error tipo II: B. Internacional de Comercio, B. Huesca, B. Intercontinental Español.					

AÑO 2	Error aparente (\overline{err})	E0 estimado (\hat{Err}_{E0})		Error real estimado	
		Media	Intervalo de confianza $\alpha=0.05$	Media	Intervalo de confianza $\alpha=0.05$
Bootstrap 0.632					
Error total	0,173	0,202	(0,079 , 0,326)	0,191	(0,114 , 0,267)
Error tipo I	0,231	0,236	(0,017 , 0,454)	0,234	(0,096 , 0,372)
Error tipo II	0,115	0,184	(-0,071 , 0,439)≈[0 , 0,439)	0,159	(-0,003 , 0,320)≈[0 , 0,320)
Bootstrap 0.632+					
Error total	0,173	0,202	(0,079 , 0,326)	0,192	(0,112 , 0,274)
Error tipo I	0,231	0,236	(0,017 , 0,454)	0,235	(0,095 , 0,373)
Error tipo II	0,115	0,184	(-0,071 , 0,439)≈[0 , 0,439)	0,162	(-0,011 , 0,334)≈[0 , 0,334)
Error tipo I: Banca Catalana, B. Industrial de los Pirineos, B. Madrid, B. Crédito Comercial.					
Error tipo II: B. Comercial Español, B. Mercantil de Tarragona, B. Huesca.					
NOTA: Error tipo I: Clasificar a una entidad fracasada como sana. Error tipo 2: Clasificar a una entidad sana como fracasada.					

FUENTE: Elaboración propia.

Tabla 5.- Modelos de análisis logit (años 1 y 2)

	Coeficientes logísticos		Significación estadística individual		Correlac. parcial	Significación estadística global	
	β_i	$SE(\beta_i)$	Wald	Sig.	R_i	Chi-cuadr.	Sig.
RENT15	-38,874	14,208	7,487	0,014	-0,220	49,049	0,000
LIQ11	-22,070	8,978	6,043	0,006	-0,168		
Constante	21,988	8,284	7,045	0,008	0,206		
	Coeficientes logísticos		Significación estadística individual		Correlac. parcial	Significación estadística global	
	β_i	$SE(\beta_i)$	Wald	Sig.	R_i	Chi-cuadr.	Sig.
LIQ17	-11,449	4,585	6,235	0,013	-0,245	34,934	0,000
RENT12	-25,290	9,348	7,318	0,007	-0,299		
Constante	12,745	4,880	6,820	0,009	0,276		

FUENTE: Elaboración propia.

Tabla 6.- Análisis logit - estimación del error real cometido (años 1 y 2)

AÑO 1	Error aparente (\overline{err})	E0 estimado (\hat{Err}_{E0})		Error real estimado	
		Media	Intervalo de confianza $\alpha=0.05$	Media	Intervalo de confianza $\alpha=0.05$
Bootstrap 0.632					
<i>Error total</i>	0,058	0,128	(0,001 , 0,245)	0,102	(0,022 , 0,176)
<i>Error tipo I</i>	0,077	0,112	(-0,128 , 0,353) \approx [0 , 0,353)	0,099	(-0,052 , 0,251) \approx [0 , 0,251)
<i>Error tipo II</i>	0,038	0,140	(-0,023 , 0,303) \approx [0 , 0,303)	0,102	(-0,001 , 0,205) \approx [0 , 0,205)
Bootstrap 0.632+					
<i>Error total</i>	0,058	0,128	(0,001 , 0,245)	0,105	(0,020 , 0,183)
<i>Error tipo I</i>	0,077	0,112	(-0,128 , 0,353) \approx [0 , 0,353)	0,100	(-0,052 , 0,256) \approx [0 , 0,256)
<i>Error tipo II</i>	0,038	0,140	(-0,023 , 0,303) \approx [0 , 0,303)	0,108	(-0,004 , 0,220) \approx [0 , 0,220)
<i>Error tipo I:</i> B. Catalana, B. Alicante.					
<i>Error tipo II:</i> B. de Comercio.					
AÑO 2	Error aparente (\overline{err})	E0 estimado (\hat{Err}_{E0})		Error real estimado	
		Media	Intervalo de confianza $\alpha=0.05$	Media	Intervalo de confianza $\alpha=0.05$
Bootstrap 0.632					
<i>Error total</i>	0,173	0,198	(0,071 , 0,325)	0,189	(0,109 , 0,269)
<i>Error tipo I</i>	0,231	0,203	(-0,004 , 0,411) \approx [0 , 0,411)	0,213	(0,082 , 0,345)
<i>Error tipo II</i>	0,115	0,185	(-0,030 , 0,399) \approx [0 , 0,399)	0,159	(0,023 , 0,294)
Bootstrap 0.632+					
<i>Error total</i>	0,173	0,198	(0,071 , 0,325)	0,189	(0,009 , 0,272)
<i>Error tipo I</i>	0,231	0,203	(-0,004 , 0,411) \approx [0 , 0,411)	0,213	(0,077 , 0,349)
<i>Error tipo II</i>	0,115	0,185	(-0,030 , 0,399) \approx [0 , 0,399)	0,163	(0,017 , 0,308)
<i>Error tipo I:</i> B. Mas Sardá, B. Industrial de Cataluña, B. Crédito e Inversiones, B. Valladolid, B. Comercial Occidental, B. Asturias.					
<i>Error tipo II:</i> B. Internacional de Comercio, B. Huesca, B. Intercontinental Español.					

FUENTE: Elaboración propia.

los conseguidos por Laffarga *et al.* (1987), Pina Martínez (1989) y Serrano Cinca (1994 y 1997) respecto al primer año, y similares a los obtenidos por De Andrés Sánchez (2005). En cuanto al segundo ejercicio, el porcentaje de acierto resulta ligeramente inferior al obtenido por Pina Martínez (1989), si bien este autor analiza un porcentaje de empresas inferior y considera un número más elevado de indicadores.

De nuevo, la capacidad de generalización del modelo constituye su ventaja más evidente, dado que los porcentajes de acierto superan a los resultados sobre muestras de test obtenidos por Pina Martínez (1989), Fernández y Olmeda (1995) y de De Andrés Sánchez (2005).

ÁRBOLES DE DECISIÓN

La aplicación del algoritmo C4.5 para la construcción de árboles de decisión ha permitido obtener los árboles presentados en la figura 3, empleando el criterio de partición basado en el "ratio de ganancia", el método de poda conocido como "pessimistic error rate" ($p=0,25$) y fijando un mínimo de 3 individuos por hoja (5% del total de la muestra).

Para cada árbol se recoge la estructura final obtenida, así como las reglas lógicas que se derivan de él⁵. Para cada regla se resumen, entre paréntesis, el porcentaje de error cometido respecto a la clase final identificada, el número de ejemplos que pertenecen a cada clase, una aproximación al estadístico j y una aproximación al estadístico $(1-p)$ para determinar la relevancia de la regla.

Como puede observarse, las variables más significativas para la predicción del fracaso bancario continúan siendo los ratios *RENT15* y *LIQ11* para el primer año (acompañado del indicador *RENT9* relativo a la rentabilidad financiera), así como las variables *LIQ17* y *RENT12* para el segundo año (acompañados de dos atributos relativos a la liquidez inmediata de las entidades), lo que resulta congruente con las técnicas estadísticas previas.

Por lo que respecta a la capacidad caracterizadora de los modelos, la tabla 7 recoge las estimaciones de los errores reales para cada año analizado.

Como puede observarse, los árboles de decisión C4.5 alcanzan porcentajes de acierto próximos al 95% y 98% para el primer y segundo año

Figura 3.- Árboles de decisión –C4.5 (años 1 y 2)

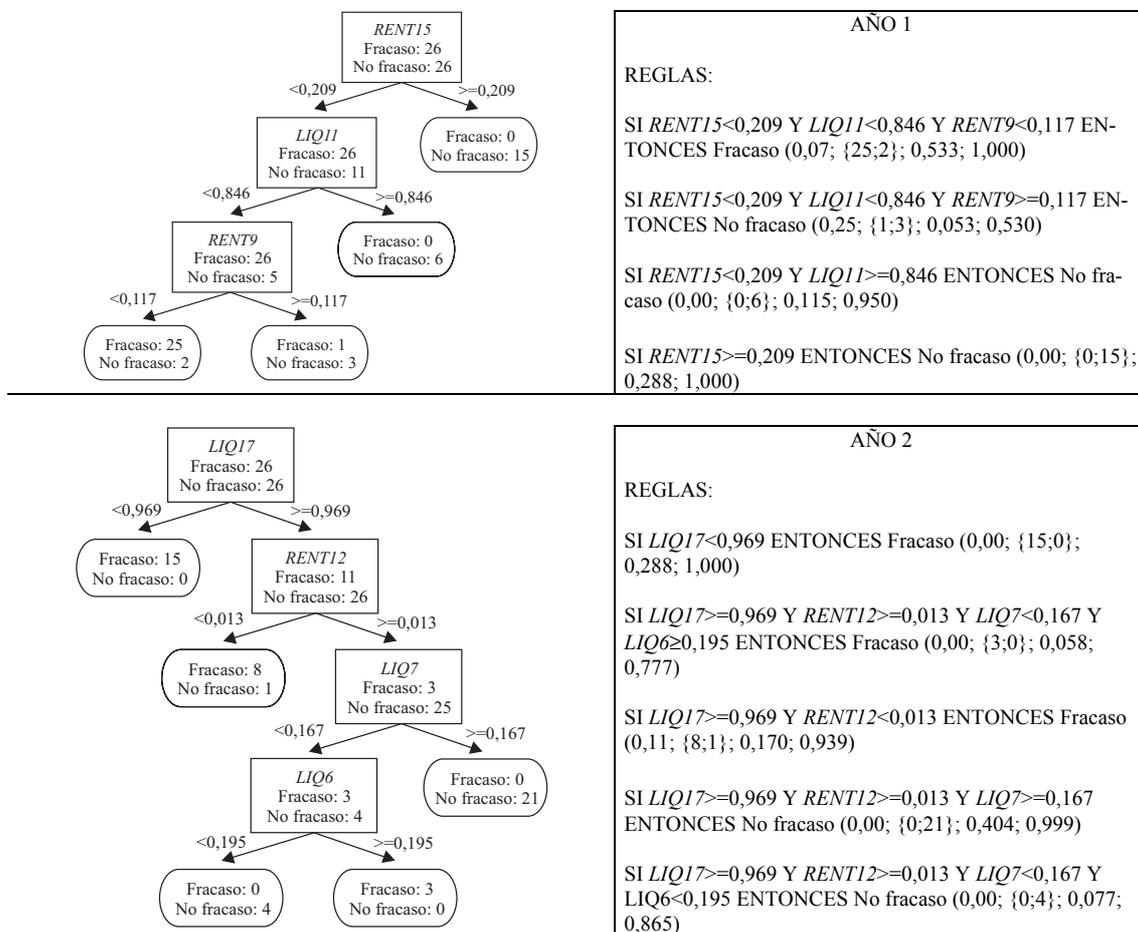


Tabla 7.- Árbol de decisión C4.5 - estimación del error real cometido (años 1 y 2)

AÑO 1	Error aparente (\overline{err})	E0 estimado (\hat{Err}_{E0})		Error real estimado	
		Media	Intervalo de confianza $\alpha=0.05$	Media	Intervalo de confianza $\alpha=0.05$
Bootstrap 0.632					
Error total	0,058	0,207	(0,041 , 0,372)	0,152	(0,047 , 0,256)
Error tipo I	0,038	0,165	(-0,089 , 0,419) \approx [0 , 0,419)	0,118	(-0,042 , 0,279) \approx [0 , 0,279)
Error tipo II	0,077	0,240	(-0,004 , 0,483)	0,180	(0,026 , 0,334)
Bootstrap 0.632+					
Error total	0,058	0,207	(0,041 , 0,372)	0,165	(0,046 , 0,285)
Error tipo I	0,038	0,165	(-0,089 , 0,419) \approx [0 , 0,419)	0,127	(-0,051 , 0,306) \approx [0 , 0,306)
Error tipo II	0,077	0,240	(-0,004 , 0,483)	0,197	(0,017 , 0,376)
Error tipo I: B. Valladolid.					
Error tipo II: B. de Comercio, B. Comercial Español.					
AÑO 2	Error aparente (\overline{err})	E0 estimado (\hat{Err}_{E0})		Error real estimado	
		Media	Intervalo de confianza $\alpha=0.05$	Media	Intervalo de confianza $\alpha=0.05$
Bootstrap 0.632					
Error total	0,019	0,230	(0,072 , 0,387)	0,152	(0,103 , 0,252)
Error tipo I	0,000	0,213	(-0,006 , 0,433) \approx [0 , 0,433)	0,135	(-0,004 , 0,274) \approx [0 , 0,274)
Error tipo II	0,038	0,251	(-0,015 , 0,516) \approx [0 , 0,251)	0,173	(0,005 , 0,341)
Bootstrap 0.632+					
Error total	0,019	0,230	(0,072 , 0,387)	0,178	(0,069 , 0,296)
Error tipo I	0,000	0,213	(-0,006 , 0,433) \approx [0 , 0,433)	0,160	(-0,005 , 0,325) \approx [0 , 0,325)
Error tipo II	0,038	0,251	(-0,015 , 0,516) \approx [0 , 0,251)	0,200	(-0,002 , 0,402) \approx [0 , 0,402)
Error tipo I: -					
Error tipo II: B. Huesca.					

FUENTE: Elaboración propia.

(muestra total), con una capacidad de generalización próxima al 85% para cada uno de los ejercicios analizados.

Si se comparan estos resultados con investigaciones previas puede observarse que, respecto a la muestra total, los resultados obtenidos superan a los alcanzados por Fernández y Olmeda (1995) y resultan ligeramente inferiores a los obtenidos por Bonsón Ponte *et al.* (1997a, 1997b) para el primer año analizado.

Si se considera el error real estimado, los árboles obtenidos superan ampliamente a la capacidad de caracterización de los modelos utilizados por los autores previos, cuyo porcentaje de acierto sobre test resulta inferior al 80%.

De esta forma, se ha comprobado la capacidad de generalización del modelo obtenido respecto a la muestra analizada, tanto por lo que respecta a los modelos estadísticos como a los árboles de decisión analizados. Esta característica resulta particularmente valiosa para el tratamiento de carteras de bajo incumplimiento, cuya validación resulta compleja dado que las técnicas tradicionales de análisis suelen generar porcentajes de acierto reducidos.

CONCLUSIONES Y DESARROLLOS FUTUROS

En el presente trabajo se ha analizado la problemática de la gestión del riesgo de crédito ante el nuevo Acuerdo de Basilea II y, en particular, las dificultades para desarrollar modelos internos de gestión respecto a las carteras de bajo incumplimiento (LDP), caracterizadas por muestras reducidas que dificultan la formulación de modelos robustos y estadísticamente válidos.

En particular, se ha considerado la problemática de una de las carteras LDP más habituales, la integrada por exposiciones frente a otros bancos, observando la necesidad de conseguir estimaciones fiables de la probabilidad de incumplimiento, lo que precisa utilizar técnicas de análisis avanzadas y especializadas en el tratamiento de muestras reducidas.

Con este objetivo, en este trabajo se ha planteado una nueva metodología para la gestión del riesgo de crédito, estructurada en tres fases: la selección de las variables explicativas más rele-

vantes del fracaso empresarial a través de un esquema híbrido, el análisis de las relaciones funcionales entre los determinantes mediante diversos modelos multivariantes (análisis discriminante, análisis logit, árboles de decisión) y la evaluación de los resultados obtenidos mediante técnicas de bootstrap.

El modelo ha sido aplicado a una conocida muestra de bancos españoles para el periodo 1978-1982, considerando uno y dos años antes del fracaso, al objeto de comparar adicionalmente los resultados obtenidos con los estudios previos en la materia.

De esta forma, ha podido observarse que los ratios más explicativos del fracaso bancario están integrados por indicadores de rentabilidad y liquidez y, en menor medida, por indicadores de endeudamiento. Los porcentajes de acierto obtenidos se reducen a lo largo del tiempo, siendo el modelo logit la propuesta más adecuada para el primer año analizado (95% para la muestra total y 90% para el modelo generalizado) y el árbol de decisión C4.5 el modelo más relevante para el segundo ejercicio (98% para la muestra total y 85% para el modelo generalizado).

Por su parte, la comparación de los resultados obtenidos con los estudios anteriores en la materia ha permitido observar que, si bien los porcentajes de acierto sobre la muestra total resultan ligeramente superiores a los conseguidos por la mayoría de las aportaciones previas, las principales mejoras se consiguen respecto a la capacidad de generalización del modelo, que resulta mucho más elevada, lo que facilita la implementación de esta metodología en el ámbito bancario, que precisa de modelos suficientemente validados.

No obstante, cabe destacar que Basilea II exige no sólo una validación fuera de muestra u "out-of-sample" (como la estimada en este trabajo) sino también una validación a lo largo del tiempo u "out-of-time" (CSBB, 2004; pfo. 420), que constituye una de las líneas de investigación a desarrollar en el futuro.

Asimismo, se plantea la conveniencia de ampliar el estudio a otros algoritmos de clasificación utilizados en el ámbito de la crisis bancaria, en particular las redes neuronales tipo Perceptrón Multicapa y Mapa SOFM de Kohonen, que por motivos de extensión no se han considerado en este trabajo, así como a carteras LDP alternativas (entidades aseguradoras, soberanos, etc).

ANEXO 1

PRINCIPALES ESTUDIOS EMPÍRICOS RESPECTO A LA CRISIS BANCARIA ESPAÑOLA

Autores	Muestra	Emparejamiento	Modelos	Crit. selección ratios	Nº ratios iniciales	Crit. reducción ratios	Nº ratios finales	Porcentaje de clasificación correcta
Laffarga <i>et al.</i> (1985)	22 frac./25 sanos (1978-1982)	Sí, según el valor medio de los depósitos en los últimos 5 años	ANOVA y análisis discriminante (univariante) 1 año antes	-Empleo frecuente en el análisis financiero. -Ratios computables según datos. -Criterio de investigadores y grado de significación estadística (test F).	15	Jerarquía de ratios (univariante).	15 (2 son más significativos)	Ratio "beneficio neto antes de impuestos/pasivo exigible": 91,5%. Ratio "beneficio neto antes de impuestos/activo total": 91,5%.
Laffarga <i>et al.</i> (1986a)	22 frac./25 sanos (1er año) 24 frac./24 sanos (2º año) 1978-1982	Sí, según el valor medio de los depósitos en los últimos 5 años	ANOVA, test de clasificación dicotómica, MDA 1 y 2 años antes	-Empleo frecuente en el análisis financiero. -Ratios computables según datos. -Criterio de investigadores y grado de significación estadística (test F).	15	Jerarquía de ratios (ANOVA y test dicotómico). Análisis MDA Proceso "Stepwise"	Univariante: 15 ratios (2 más significativos para cada año). MDA: 2 a 5 (año 1), 3 a 5 (año 2)	<u>Año 1 (univariante):</u> Ratio "beneficio neto antes de impuestos/pasivo exigible": 91,5%. Ratio "beneficio neto antes de impuestos/activo total": 91,5%. <u>Año 2 (univariante):</u> Ratio "beneficio neto antes de impuestos/capital+reservas": 77,1%. Ratio "gastos y amort. del ejercicio / productos del ejercicio": 79,2%. <u>MDA (año 1):</u> 85,1% (2 ratios), 76,6% (3 ratios), 83,0% (4 ratios), 85,1% (5 ratios). <u>MDA (año 2):</u> 83,3% (3 ratios), 83,3% (4 ratios), 83,3% (5 ratios).
Laffarga <i>et al.</i> (1986b)	22 frac./23 sanos (3er año) 21 frac./21 sanos (4º año) 20 frac./21 sanos (5º año) 1978-1982	Sí, según el valor medio de los depósitos en los últimos 5 años	ANOVA, test de clasificación dicotómica, MDA 3, 4 y 5 años antes	-Empleo frecuente en el análisis financiero. -Ratios computables según datos. -Criterio de investigadores y grado de significación estadística (test F).	15	Jerarquía de ratios (ANOVA y test dicotómico). Análisis MDA Proceso "Stepwise"	Univariante: 15 ratios (2-3 más significativos para cada año). MDA: 3 a 4 (años 3, 4 y 5).	<u>Año 3 (univariante):</u> Ratio "beneficio neto antes de impuestos/activo total": 77,8%. Ratio "beneficio neto antes de impuestos/pasivo exigible": 77,8%. <u>Año 4 (univariante):</u> Ratio "beneficio neto antes de impuestos/capital+reservas": 78,6%. Ratio "beneficio neto antes de impuestos/pasivo exigible": 66,7%. <u>Año 5 (univariante):</u> Ratio "reservas/pasivo total": 82,9%. Ratio "beneficio neto antes de impuestos/capital+reservas": 80,5%. <u>MDA (año 3):</u> 80,0% (3 ratios), 75,3% (4 ratios). <u>MDA (año 4):</u> 73,8% (3 ratios), 76,2% (4 ratios). <u>MDA (año 5):</u> 75,6% (3 ratios), 80,5% (4 ratios).
Laffarga <i>et al.</i> (1987)	22 frac./25 sanos (1978-1982)	Sí, según el valor medio de los depósitos en los últimos 5 años	MDA y logit 1 año antes	Juicio investigadores	4	No se utiliza	MDA: 4 ratios Logit: 4 ratios	<u>MDA (año 1):</u> 86,91% (entrenamiento). <u>Logit (año 1):</u> 93,40% (entrenamiento).
Pina Martínez (1989)	18 frac./20 sanos (1er año) 18 frac./20 sanos (2º año) 17 frac./20 sanos (3º año) 1978-1982 (+ test de 25 frac./25 sanos, periodo 1983-1985)	No	Logit 1 a 3 años antes y trianuales	Los más comúnmente utilizados. Emplea un modelo general de regresión y desecha los datos con un valor t poco significativo.	9	Análisis Logit Proceso "Stepwise"	Logit: 3 (año 1), 3 (año 2), 4 (año 3), 3 (ratio combinado trianual)	<u>Logit (año 1):</u> 92,11% (entrenamiento), 80% (test fracasados). <u>Logit (año 2):</u> 92,11% (entrenamiento); 88,00% (test fracasados). <u>Logit (año 3):</u> 91,89% (entrenamiento); 92,00% (test fracasados). <u>Logit (combinado trianual):</u> 100,00% (entrenamiento); 92,00% (test fracasados).
Laffarga <i>et al.</i> (1988; 1991)	22 frac./25 sanos (1er año) 24 frac./24 sanos (2º año) 22 frac./23 sanos (3er año) 21 frac./21 sanos (4º año) 20 frac./21 sanos (5º año) 1978-1982	Sí, según el valor medio de los depósitos en los últimos 5 años	ANOVA, test de clasificación dicotómica, MDA 1 a 5 años antes	-Empleo frecuente en el análisis financiero. -Ratios computables según datos. -Criterio de investigadores y grado de significación estadística (test F).	15	Jerarquía de ratios (ANOVA y test dicotómico). Análisis MDA Proceso "Stepwise"	Univariante: 3 ratios (para los 5 años de predicción). MDA: 3 a 4 (años 1 a 5).	<u>Univariante:</u> 90-95% (1er año), 75-80% (2º-5º año). <u>MDA (año 1):</u> 76,6% (3 ratios), 83,0% (4 ratios). <u>MDA (año 2):</u> 83,3% (3 ratios), 83,3% (4 ratios). <u>MDA (año 3):</u> 80,0% (3 ratios), 75,3% (4 ratios). <u>MDA (año 4):</u> 73,8% (3 ratios), 76,2% (4 ratios). <u>MDA (año 5):</u> 75,6% (3 ratios), 80,5% (4 ratios).
Rodríguez Fernández (1986, 1987)	31 frac./49 sanos (1981-1983)	No	Logit 1y 2 años antes	-Índices contables frecuentes.	15	Análisis logit Proceso "Stepwise"	2 ratios	91,25% (datos de 1980), 100% (datos de 1981).
Rodríguez Fernández (1989a; 1989b)	31 frac./49 sanos (1981-1983 (adicionalmente, la muestra se divide al 50% en entrenamiento y test)	No	Análisis univariante, MDA 1 y 2 años antes	-Índices contables frecuentes.	15	Análisis MDA Proceso "Stepwise"	3 ratios	<u>Univariante:</u> Ratio "fondos propios libres/inversiones crediticias": 80,00%. <u>MDA (datos de 1981):</u> 100,00% (muestra total); 97,44% (test). <u>MDA (datos de 1980):</u> 91,25% (muestra total); 89,74% (test).

Autores	Muestra	Emparejamiento	Modelos	Crit. selección ratios	Nº ratios iniciales	Crit. reducción ratios	Nº ratios finales	Porcentaje de clasificación correcta
Martínez Mongay <i>et al.</i> (1989)	37 frac./52 sanos 1979-1983 (+ test de 15 frac./15 sanos, 1984-1985)	No	Análisis probit 1 y 2 años antes	- Ratios contables diversos y año de fundación	13	Combinación de ratios (prueba y error)	9 (modelo final) y 5 (modelo control)	<u>Análisis probit:</u> 96,63% (modelo final y de control sobre muestral total). No se proporcionan resultados para la muestra de test.
Serrano Cinca (1994)	29 frac./37 sanos 1978-1983	Sí, al igual que Laffarga <i>et al.</i> (1986)	Análisis univariante Análisis factorial, escalamiento multidimensional (MDS) y análisis profit, Red SOFM de Kohonen, conglomerados, Perceptrón uni y Multicapa, MDA, logit 1 año antes.	- Ratios utilizados por Pina (1989), previa comprobación de su poder discriminatorio	9	Análisis factorial	Univariante: 6 ratios más significativos. Factorial: 3 factores MDS: 2 y 3 dimensiones Red SOFM: 9 Perceptrón: 9 MDA: 9 Logit: 9	<u>MDA. Perceptrón unicapa:</u> 89,39% (muestra total, 9 ratios). 86,36% (jackknife). <u>Logit:</u> 93,94% (muestra total, 9 ratios). 95,45% (muestra total, 8 partidas contables). <u>Red SOFM de Kohonen:</u> No se proporciona (14x14 neuronas). <u>Perceptrón multicapa (9x4x1 neuronas):</u> 93,94%-100% (muestra total). 94,00% (jackknife).
Serrano Cinca y Mar Molinero (1994); Mar Molinero y Serrano Cinca (2001)	29 frac./37 sanos 1978-1983	Sí, al igual que Laffarga <i>et al.</i> (1986)	Escalamiento multidimensional (MDS) y análisis profit, combinados con análisis de componentes principales. Análisis logit. Análisis univariante 1 año antes	- Ratios utilizados por Pina (1989), previa comprobación de su poder discriminatorio	9	No se aplica	9, en dos dimensiones principales (MDS). Análisis logit: 2 factores Univariante: 6 ratios más significativos	<u>Análisis logit sobre los valores obtenidos por el MDS y logit:</u> 94% (punto de corte: 0,5%). <u>Univariante:</u> No se proporciona.
Serrano Cinca (1997)	29 frac./37 sanos 1978-1983	Sí, al igual que Laffarga <i>et al.</i> (1986)	Análisis univariante (Lambda de Wilks), MDA, logit, Perceptrón uni y multicapa 1 año antes	- Ratios utilizados por Pina (1989), previa comprobación de su poder discriminatorio	9	Análisis MDA y logit: Proceso "Stepwise"	Univariante: 6 ratios más significativos MDA: 8 (excluye "beneficio neto/pasivo exigible"). Logit: 9 Perceptrón: 9	<u>MDA:</u> 87,88%. <u>Logit:</u> 93,94%. <u>Perceptrón:</u> 89,39% (unicapa lineal), 86,36% (unicapa lineal con jackknife), 92,42% (unicapa logístico), 93,94%-100% (multicapa).
Fernández y Olmeda (1995)	29 frac./37 sanos 1977-1985 (34 entrenamiento y 32 test y validación cruzada con v=6)	Sí, al igual que Laffarga <i>et al.</i> (1986)	MDA, logit, Perceptrón Multicapa, MARS, C4.5 (parámetros estimados por prueba y error). También combinaciones de métodos	- Ratios utilizados por Pina (1989).	9	No se aplica	9	<u>MDA:</u> 90% (entrenamiento); 61,76% (test). <u>Logit:</u> 95,29% (entrenamiento); 76,47% (test). <u>MARS:</u> 94,11% (entrenamiento); 79,41% (test). <u>Perceptrón multicapa (no se facilita la arquitectura):</u> 97,05% (muestra total); 82,35% (test). <u>C4.5:</u> 84,70% (entrenamiento); 79,41% (test). Las combinaciones de métodos NN+logit, NN+logit+C4.5+MDA, NN+logit+MARS+DA mejoran ligeramente los resultados individuales.
Serrano Cinca y Martín del Brío (1993)	29 frac./37 sanos 1978-1983	Sí, al igual que Laffarga <i>et al.</i> (1986)	Perceptrón multicapa, red SOFM de Kohonen 1 año antes	- Ratios utilizados por Pina (1989), previa comprobación de su poder discriminatorio	9	No se aplica	9	<u>Perceptrón Multicapa:</u> 94,00% (9x6x1 neurona). <u>Red SOFM de Kohonen:</u> 96,96-98,48% (14x14 neuronas).
Martín del Brío y Serrano Cinca (1993)	29 frac./37 sanos 1978-1983	Sí, al igual que Laffarga <i>et al.</i> (1986)	Red SOFM de Kohonen 1 año antes	- Ratios utilizados por Pina (1989), previa comprobación de su poder discriminatorio	9	No se aplica	9	No se proporciona. Se lleva a cabo un análisis temporal de la evolución de la situación de las entidades a lo largo del mapa de Kohonen y se compara con los resultados del Perceptrón Multicapa desarrollado en Serrano Cinca y Martín del Brío (1993).
Martín del Brío y Serrano Cinca (1995)	29 frac./37 sanos 1978-1983	Sí, al igual que Laffarga <i>et al.</i> (1986)	Red SOFM de Kohonen Análisis cluster 1 año antes	- Ratios utilizados por Pina (1989), previa comprobación de su poder discriminatorio	9	No se aplica	9	No se proporciona. Se lleva a cabo un análisis temporal de la evolución de la situación de las entidades a lo largo del mapa de Kohonen y se compara éste con los resultados del análisis cluster.
Bonsón Ponte <i>et al.</i> (1997a; 1997b)	29 frac./37 sanos 1978-1983	Sí, al igual que Laffarga <i>et al.</i> (1986)	Árbol de Decisión C4.5 1 año antes	- Ratios utilizados por Pina (1989), previa comprobación de su poder discriminatorio	9	Selección por el algoritmo C4.5	4-6	<u>C4.5:</u> 100,00% (muestra total, 6 nodos de profundidad). 77,30% (jackknife árboles clásicos). 83,33% (jackknife árbol de moda).

Autores	Muestra	Emparejamiento	Modelos	Crit. selección ratios	Nº ratios iniciales	Crit. reducción ratios	Nº ratios finales	Porcentaje de clasificación correcta
López González y Flórez López (2000)	26 frac./26 sanos 1978-1982 (34 entrenamiento y 18 test)	Sí, según tamaño	Perceptrón Multicapa y red SOFM de Kohonen 1 año antes	- Diversidad de ratios utilizados en la literatura y en estudios teóricos.	31	Análisis de Componentes Principales	6	<i>Perceptrón multicapa (6x4x1)</i> : 100% (muestra total); 88,89% (test). <i>Kohonen (10x10)</i> : 96,15% (muestra total).
De Andrés Sánchez (2005)	29 frac./37 sanos 1978-1983 (40 entrenamiento, 26 test)	Sí, al igual que Laffarga <i>et al.</i> (1986)	MDA, logit, probit, Perceptrón multicapa, Análisis Factorial 1 año antes	- Ratios utilizados por Pina (1989).	9	Análisis factorial	3 factores	<i>MDA</i> : 92,50% (entrenamiento), 76,92% (test). <i>Logit</i> : 95,00% (entrenamiento), 76,92% (test). <i>Probit</i> : 90,00% (entrenamiento), 80,97% (test). <i>Perceptrón Multicapa</i> : 100,00% (entrenamiento), 84,61% (test). (1 capa oculta con 4 neuronas).

ANEXO 2

VARIABLES EXPLICATIVAS UTILIZADAS EN ALGUNOS DE LOS PRINCIPALES ESTUDIOS SOBRE LA CRISIS BANCARIA ESPAÑOLA

Ratio	Tipo de ratio	Referencias									
		Laffarga <i>et al.</i> (1986a)	Laffarga <i>et al.</i> (1986b)	Laffarga <i>et al.</i> (1985)	Laffarga <i>et al.</i> (1991)	Pina Martínez (1989)	Serrano Cinca (1994; 1997). Fernández y Olmeda (1995). De Andrés Sánchez (2005)	Bonsón Ponte <i>et al.</i> (1997a; 1997b)	Rodríguez Fernández (1986; 1987)	Rodríguez Fernández (1989a; 1989b)	
Capital circulante / Pasivo exigible	Liquidez	(1)		(1)							
Beneficio neto antes de impuestos / Activo total	Rentabilidad	(1), (2)	(3), (5)	(1)	(1), (3), (5)						
Créditos más cartera de valores / Activo total	Liquidez	(1), (2)	(4)	(1)							
Activo fijo / Activo total	Estructura del activo	(2)	(5)	(1)							
Reservas / Pasivo total	Rentabilidad acumulada	(1), (2)	(3), (4), (5)	(1)							
Beneficio neto antes de impuestos / Pasivo exigible	Rentabilidad	(1), (2)	(4)	(1)							
Activo circulante / Total activo	Liquidez	(1), (2)	(3), (5)	(1)	(1), (3), (5)	I (2)	(1)	(1)			
Créditos / Capital y ganancias retenidas	Adecuación de capital	(1), (2)	(3), (4), (5)	(1)	(1), (3), (5)						
Créditos / Activo total	Liquidez	(2)		(1)	(1), (3), (5)						
Activo fijo / Activo total	Estructura del activo	(1), (2)	(3), (4)		(1), (2), (3), (4), (5)					(1), (2)	
Neto patrimonial / Pasivo total	Beneficio acumulado				(1), (2), (3), (4), (5)						
Beneficio antes de impuestos / Capital más reservas	Rentabilidad	(1), (2)	(3), (4), (5)	(1)	(1), (2), (3), (4), (5)						
Cartera de créditos y obligaciones / Activo total	Liquidez				(2), (4)						
Pasivo circulante / Pasivo total	Estructura del pasivo				(2), (4)						
Beneficio antes de impuestos / Pasivo total	Rentabilidad				(2), (4)						
Créditos / Activo total					(2)						
(Activo circulante – tesorería) / Activo total						I, (1), (2), (3), (c)	(1)	(1)			
Activo circulante / Pasivo exigible	Liquidez	(2)		(1)		I, (1), (3)	(1)	(1)			
Reservas / Pasivo exigible						I, (3)	(1)	(1)			
Beneficio neto / Activo total						I, (3)	(1)	(1)	(1), (2)	(1), (2)	
Beneficio neto / Neto						I, (1), (3)	(1)	(1)			
Beneficio neto / Pasivo exigible						I, (c)	(1)	(1)			
Coste de ventas / Ventas brutas						I (2)	(1)	(1)			
Cash-Flow(*) / Pasivo exigible						I, (c)	(1)	(1)			

Ratio	Tipo de ratio	Referencias								
		Laffarga <i>et al.</i> (1986a)	Laffarga <i>et al.</i> (1986b)	Laffarga <i>et al.</i> (1985)	Laffarga <i>et al.</i> (1991)	Pina Martínez (1989)	Serrano Cínc (1994; 1997). Fernández y Olmeda (1995). De Andrés Suárez (2005)	Bonsón Ponte <i>et al.</i> (1997a; 1997b)	Rodríguez Fernández (1986; 1987)	Rodríguez Fernández (1989a; 1989b)
Pasivo exigible / Pasivo total	Estructura del pasivo	(2)	(4)							
Tesorería / Activo circulante	Liquidez									
Capital circulante / Gastos del ejercicio	Liquidez			(1)						
Gastos y amortización del ejercicio / Productos del ejercicio	Eficiencia operativa	(2)		(1)						
Intereses y comisiones pagadas / Productos del ejercicio	Cobertura de los intereses y comisiones pagadas	(1)		(1)						
Fondos propios libres / Inversiones crediticias	Adecuación del capital								(1), (2)	(1), (2)

NOTAS:
 - (*): En este caso el cash-flow se entiende como beneficio neto más amortizaciones más provisiones.
 - Entre paréntesis se recoge, en su caso, el año anterior a la quiebra del que se han tomado los datos de las empresas analizadas.
 - Para Pina Martínez (1989) se consideran los ratios inicialmente incluidos en el análisis (1) y los seleccionados por el análisis logit para uno, dos y tres años antes del fracaso, y modelo combinado (c).

ANEXO 3

MUESTRA ANALIZADA: BANCOS FRACASADOS Y SANOS (1978-1982)

Año de fracaso	Banco fracasado	Volumen medio depós.	Banco sano emparejado	Volumen medio depós.
1982	(1) B. Unión	52.535	(27) B. Fomento	62.976
1982	(2) B. Mas Sardá	33.738	(28) B. Internacional de Comercio	33.076
1982	(3) B. Levante	13.262	(29) Banca Jover	14.803
1982	(4) Banca Catalana	89.210	(30) B. Pastor	112.531
1982	(5) B. Industrial de Cataluña	36.638	(31) B. Zaragozano	68.512
1982	(6) B. Barcelona	9.986	(32) B. de Vitoria	8.614
1982	(7) B. Girona	1.749	(33) Banca Pueyo	1.616
1982	(8) B. Alicante	16.189	(34) B. Crédito Balear	17.203
1982	(9) B. de Crédito e Inversiones	27.056	(35) B. Castilla	28.978
1981	(10) B. Industrial de los Pirineos	3.087	(36) B. Pequeña y Mediana Empresa	1.381
1980	(11) B. Madrid	60.480	(37) B. Guipuzcoano	41.208
1978	(12) B. Navarra	4.942	(38) B. Comercial Español	5.005
1978	(13) B. Cantábrico	2.077	(39) B. Mercantil de Tarragona	2.436
1978	(14) B. Meridional	1.953	(40) Banca Abel Matutes	2.113
1978	(15) B. Valladolid	12.583	(41) B. Galicia	6.085
1979	(16) B. de Crédito Comercial	2.522	(42) Sindicato de Banqueros de Barcelona	2.897
1982	(17) B. de Préstamo y Ahorro	12.838	(43) B. Huesca	9.683
1981	(18) B. de Descuento	9.922	(44) B. de Comercio	7.245
1981	(19) B. Comercial Occidental	12.247	(45) B. de Finanzas	10.306
1981	(20) B. Occidental	24.075	(46) B. Comercial Trasatlántico	28.751
1980	(21) B. Industrial del Mediterráneo	12.482	(47) B. Intercontinental Español	14.843
1980	(22) B. Catalán de Desarrollo	25.211	(48) B. Industrial de Bilbao	23.545
1980	(23) B. Promoción de Negocios	4.374	(49) B. del Progreso	3.232
1980	(24) B. López Quesada	12.797	(50) B. de Vasconia	12.103
1979	(25) B. de Asturias	5.052	(51) B. de Europa	6.438
1978	(26) B. de Granada	12.707	(52) B. Financiación Industrial	16.223

NOTAS

1. Trabajo presentado en las VI Jornadas de Predicción de la Insolvencia Empresarial, Carmona (Sevilla), 9 y 10 de noviembre de 2006.
2. Véase Dash y Liu (1997) para una revisión detallada.
3. La hibridación de distintos modelos de selección de atributos constituye una propuesta

- emergente que ha demostrado un óptimo comportamiento para la detección de atributos fuertemente relevantes (Cornuejols *et al.*, 2005; Olsson y Oard, 2006).
4. Para el contraste de igualdad de medias se utilizó el test paramétrico de la t de Student y las pruebas no paramétricas U de Mann-Whitney, Z de Kolmogorov-Smirnov, test de rachas de Wald-Wolfowitz y test de Moses. Para el con-

traste de la igualdad de medianas se utilizó el test de la mediana.

5. Se trata de las reglas lógicas inicialmente obtenidas, si bien podrían aplicarse algoritmos de simplificación que redujeran aún más la complejidad de estas sentencias.

BIBLIOGRAFÍA

- ALTMAN, E.I. (1981): *Financial Handbook*. New York: John Wiley & Sons.
- ANDRÉS SÁNCHEZ, J. DE (2005): "Comparativa de métodos de predicción de la quiebra: Redes neuronales artificiales vs. métodos estadísticos multivariantes", *Partida Doble*, núm. 168, (julio-agosto), pp. 105-113.
- BLUM, A.; LANGLEY, P. (1997): "Selection of Relevant Features and Examples in Machine Learning", *Artificial Intelligence*, núm. 97, pp. 245-271.
- BONSÓN PONTE, E.; ESCOBAR RODRÍGUEZ, T.; MARTÍN ZAMORA, M.P. (1997a): "Decision Tree Induction Systems. Applications in Accounting and Finance", en E. Bonsón Ponte y G. Sierra Molina [ed]: *Intelligent Technologies in Accounting and Business. Proceedings of the III International Meeting on Artificial Intelligence in Accounting, Finance and Tax*. Huelva.
- BONSÓN PONTE, E.; ESCOBAR RODRÍGUEZ, T.; MARTÍN ZAMORA, M.P. (1997b): "Sistemas de inducción de árboles de decisión: Utilidad en el análisis de crisis bancarias", *Ciberconta. Revista electrónica de Contabilidad*. Universidad de Zaragoza, Departamento de Contabilidad y Finanzas. (Disponible en <http://ciberconta.unizar.es/Biblioteca/0007/árboles.html>).
- BREIMAN, L.; FRIEDMAN, J.H.; OLSHEN, R.A.; STONE, C.J. (1984): *Classification and Regression Trees*. New York: Chapman and Hall.
- CAREY, M.; HRYCAY, M. (2001): "Parameterizing Credit Risk Models with Rating Data", *Journal of Banking and Finance*, núm. 25, pp. 197-270.
- COMITÉ DE SUPERVISIÓN BANCARIA DE BASILEA (CSBB) (2004): *International Convergence of Capital Measurements and Capital Standards. A Revised Framework*. Basilea: Bank for International Settlements.
- COMITÉ DE SUPERVISIÓN BANCARIA DE BASILEA (CSBB) (2005a): *Studies on the Validation of Internal Rating Systems*. (Working Paper, 14). Basilea: Bank for International Settlements.
- COMITÉ DE SUPERVISIÓN BANCARIA DE BASILEA (CSBB) (2005b): *Validation of low-default portfolios in the Basel II Framework*. (Newsletter, 6). Basilea: Bank for International Settlements.
- CORNUEJOLS, A.; FROIDEVAUX, C. Y MARY, J. (2005): "Comparing and Combining Feature Estimation Methods for the Analysis of Microarray Data", *JOBIM-05, Journées Ouvertes Biologie, Informatique, Mathématiques*. Lyon.
- DASH, M.; LIU, H. (1997): "Feature Selection for Classification", *Intelligent Data Analysis*, núm. 1, pp. 131-156.
- DELLAPORTAS, P.; FORSTER, J.J.; NTZOUFRAS, I. (1997): *On Bayesian Model and Variable Selection Using MCMC*. Atenas: Athens University of Economics and Business, Department of Statistics.
- EFRON, B. (1982): *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia: Society for Industrial and Applied Mathematics.
- EFRON, B. (1986): "How Biased is the Apparent Error Rate of the Prediction Rule?", *Journal of the American Statistical Association*, núm. 81, pp. 461-470.
- EFRON, B.; TIBSHIRANI, R. (1993): *An Introduction to the Bootstrap*. New York: Chapman and Hall.
- EFRON, B.; TIBSHIRANI, R. (1995): *Cross Validation and the Bootstrap: Estimating the Error Rate of a Prediction Rule*. (Technical Report, 176). Palo Alto: Stanford University.
- FERNÁNDEZ, E.; OLMEDA, I. (1995): "Bankruptcy Prediction with Artificial Neural Networks", *Lecture Notes on Computational Sciences*, vol. 930, pp. 1142-1146.
- FLÓREZ LÓPEZ, R. (2002): "Reviewing RELIEF and its Extensions: A New Approach for Estimating Attributes for Binary and Multi-Class Problems", *Proceedings of the 2002 IEEE Conference on Data Mining*. Maebashi City.
- JOHN, G.; KOHAVI, R. (1995): "Feature Subset Selection Using the Wrapper Model: Overfitting and Dynamic Search Space Topology", *Proceedings of the First International Conference on Knowledge Discovery and Data Mining*, pp. 192-197.
- JOHN, G.H.; KOHAVI, R.; PFLEGER, K. (1994): "Irrelevant Features and the Subset Selection Problem", *Proceedings of the Eleventh International Conference on Machine Learning*, pp. 121-129.
- KIRA, K.; RENDELL, K.A. (1992): "The Feature Selection Problem: Traditional Methods and a New Algorithm", *Proceedings Ninth International Conference on Artificial Intelligence*, pp. 129-134. MIT Press.
- KONONENKO, I. (1997): "Non-Myopic Attribute Estimation in Regression", *Applied Intelligence*, vol. 7, núm. 1, pp. 39-55.
- KUO, L.; MALLICK, B. (1998): "Variable Selection for Regression Models", *Sankhya B*, núm. 60, pp. 65-81.
- LAFFARGA BRIONES J.; MARTÍN, J.L.; VÁZQUEZ,

- M.J. (1985): "El análisis de la solvencia de las instituciones bancarias: Propuesta de una metodología y aplicaciones a la Banca española", *Esic-Market*, núm. 48, (2º trim.), pp. 51-73.
- LAFFARGA BRIONES J.; MARTÍN, J.L.; VÁZQUEZ, M.J. (1986a): "El pronóstico a corto plazo del fracaso en las instituciones bancarias: metodología y aplicaciones a la Banca española", *Esic-Market*, núm. 53, (3º trim.), pp. 59-116.
- LAFFARGA BRIONES, J., MARTÍN, J.L.; VÁZQUEZ, M.J. (1986b): "El pronóstico a largo plazo del fracaso en las instituciones bancarias: metodología y aplicaciones al caso español", *Esic-Market*, núm. 54, (4º trim.), pp. 113-167.
- LAFFARGA BRIONES, J.; MARTÍN, J.L.; VÁZQUEZ, M.J. (1987): "Predicción de la crisis bancaria española: La comparación entre el análisis logit y el análisis discriminante", *Cuadernos de Investigación Contable*, núm. 1, (otoño), pp. 103-110.
- LAFFARGA BRIONES J.; MARTÍN, J.L.; VÁZQUEZ, M.J. (1991): "La quiebra de la quiebra bancaria: el caso español". *Revista Española de Financiación y Contabilidad*, vol. 20, núm. 66, (enero-marzo), pp. 151-163.
- LÓPEZ GONZÁLEZ, E. Y FLÓREZ LÓPEZ, 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, núm. 2, pp. 139-164.
- MAR MOLINERO, C.; SERRANO CINCA, C. (2001): "Bank Failure: A Multidimensional Scaling Approach", *The European Journal of Finance*, núm. 7, pp. 165-183
- MARTÍN DEL BRÍO, B.; SERRANO CINCA, C. (1993): "Self-Organizing Neural Networks for the Analysis and Representation of Data: Some Financial Cases", *Neural Computing & Applications*, vol. 1, pp. 193-206.
- MARTÍN MARÍN, J.L.; TRUJILLO PONCE, A. (2004): "El nuevo acuerdo de Basilea y la gestión del riesgo de crédito", *Harvard Deusto. Finanzas y Contabilidad*, núm. 58, pp. 50-59.
- MARTÍNEZ MONGAY, C.; NAVARRO, M.V.; SANZ, F. (1989): "Selección y explotación de los sistemas de alarma y prevención de quiebra", *Investigaciones Económicas*, (supl.), pp. 135-141.
- MCNELIS, P.D. (1997): *Neural Networks and Genetic Algorithms: Tools for Forecasting and Risk Analysis in Financial Markets*. Washington D.C.: Georgetown University.
- MYERS, S.F. (1998): "El sistema ideal de gestión de riesgos. Un debate sobre las realidades", en Fundación BBV [ed]: *La gestión del Riesgo de Mercado y de Crédito. Nuevas Tendencias de Valoración*, pp. 173-191. Bilbao: Fundación BBV.
- OLSSON, J.S.; OARD, D.W. (2006): "Combining Feature Selectors for Text Classification", *Proceedings of the ACM Fifteenth Conference on Information and Knowledge Management*. Washington D.C.
- PINA MARTÍNEZ, V. (1989): "Estudio empírico de la crisis bancaria", *Revista Española de Financiación y Contabilidad*, vol. 28, núm. 58, (enero-marzo), pp. 309-338.
- QUINLAN, R.S. (1993): *C4.5. Programs for Machine Learning*. San Mateo: Morgan Kaufmann.
- RODRÍGUEZ FERNÁNDEZ, J.M. (1986): "Crisis en los bancos privados españoles: un modelo logit", *II Jornadas de Economía Industrial*. Madrid.
- RODRÍGUEZ FERNÁNDEZ, J.M. (1987): "Crisis en los bancos privados españoles: un modelo logit", *Investigaciones Económicas*, (supl.), pp. 59-64.
- RODRÍGUEZ FERNÁNDEZ, J.M. (1989a): "Análisis de las insolvencias bancarias en España: un modelo empírico". *Moneda y Crédito*, núm. 189, pp. 187-227.
- RODRÍGUEZ FERNÁNDEZ, J.M. (1989b): "The Crisis in Spanish Private Banks: A Logit Analysis", *Finance*, vol. 10, núm. 1, (junio), pp. 69-88.
- SCHERF, M.; BRAUER, W. (1997): *Feature Selection by Means of a Feature Weighting Approach*. (Technical Report, FKI-221-97). Munich: Technische Universität München.
- SERRANO CINCA, C. (1994): *Las redes neuronales artificiales en el análisis de la información contable*. (Tesis doctoral). Zaragoza: Universidad de Zaragoza.
- SERRANO CINCA, C. (1997): "Feedforward Neural Networks in the Classification of Financial Information", *European Journal of Finance*, vol. 3, núm. 3, (septiembre), pp. 183-202.
- SERRANO CINCA, C.; MAR MOLINERO, C. (1994): "Bank Failure: A Multidimensional Scaling Approach", *17th Annual Congress of the European Accounting Association*. Venice.
- SERRANO CINCA, C.; MAR MOLINERO, C. (1993): "Predicción de la crisis bancaria mediante el empleo de redes neuronales artificiales", *Revista Española de Financiación y Contabilidad*, vol. 22, núm. 74, pp. 153-176.
- SESTITO, S.; DILLON, T. (1994): *Automated Knowledge Acquisition*. New York: Prentice Hall.
- STORK, D.; HASSIBI, B. (1993): "Second Order Derivatives for Network Pruning: Optimal Brain Surgeon", en T.J. Sejnowski, G.E. Hinton y D.S. Touretzky, D.S. [ed.]: *Advances in Neural Information Processing Systems (NIPS 5)*, pp. 164-171. San Mateo: Morgan Kaufmann.
- WHITE, H. (1989): "Learning in Artificial Neural Networks: A Statistical Perspective", *Neural Computation*, núm. 1, pp. 425-464.