

El Algoritmo *See5* *versus* la metodología Rough Set. Una aplicación a la predicción de la insolvencia en empresas Españolas de seguros no-vida

Zuleyka DÍAZ MARTÍNEZ

Profesora Ayudante. Departamento de Economía Financiera y Contabilidad I.
Universidad Complutense de Madrid.
zuleyka@ccee.ucm.es

José FERNÁNDEZ MENÉNDEZ

Profesor Asociado. Departamento de Organización de Empresas.
Universidad Complutense de Madrid.
jfernan@ccee.ucm.es

M.^a Jesús SEGOVIA VARGAS

Profesora Asociada. Departamento de Economía Financiera y Contabilidad I.
Universidad Complutense de Madrid.
mjsegovia@ccee.ucm.es

Fecha de recepción: 12/01/2005

Fecha de aceptación: 26/07/2005

RESUMEN

La predicción de la insolvencia de las empresas aseguradoras se configura como una importante cuestión en el campo de la investigación financiera, debido a la necesidad de proteger al público en general a la vez que minimizar los costes asociados a este problema. En el presente trabajo se pretende examinar la aplicabilidad a la predicción de la insolvencia de empresas españolas de seguros no-vida de un método procedente del campo de la Inteligencia Artificial, el algoritmo *See5*, tomando como información de partida un conjunto de ratios de carácter financiero. También se comparan los resultados alcanzados con los que se obtienen aplicando la metodología *Rough Set*. Los métodos que se examinarán caen dentro del área conocida como Aprendizaje Automático y presentan las ventajas de ser fácilmente implementables, proporcionar resultados de interpretación sencilla y salvar algunos inconvenientes de las técnicas estadísticas convencionales.

Palabras clave: Insolvencia, Empresas Aseguradoras, *See5*, *Rough Set*.

See5 Algorithm *versus* *Rough Set* methodology. An application to the prediction of insolvency in Spanish non-life insurance companies

ABSTRACT

Prediction of insurance companies insolvency has arisen as an important problem in the field of financial research, in order to protect both society and customers and minimize the costs associated with

this issue. In this paper we intend to examine the suitability of a method coming from the field of the Artificial Intelligence to the prediction of insolvency among Spanish non-life insurance companies, the *See5* algorithm, taking a group of financial ratios as a starting point. Furthermore a comparative study of the performance of this technique and the *Rough Set* methodology is carried out. The methods that will be examined fall into the well-known area of Machine Learning and they present the advantage of being easy to implement as well as providing results of simple interpretation whilst avoiding some of the drawbacks of the conventional statistical techniques.

Keywords: Insolvency, Insurance Companies, See5, Rough Set.

SUMARIO: 1. Introducción. 2. El algoritmo de inducción de árboles de decisión *See5*. 3. Selección de datos y variables. 4. Resultados. 5. Conclusiones.

1. INTRODUCCIÓN

La predicción de la insolvencia es uno de los temas centrales del análisis financiero que ha suscitado el interés no sólo del ámbito académico sino también de un amplio abanico de usuarios relacionados con el mundo empresarial. Las entidades aseguradoras han sido objeto desde hace tiempo de regulación y control con el propósito de comprobar que mantienen una situación de solvencia suficiente para cumplir su objeto social. Ahora bien, siendo la protección de asegurados y beneficiarios amparados por un contrato de seguro la principal finalidad de la ordenación y supervisión de la actividad aseguradora, no puede el legislador dejar de considerar la enorme importancia del sector del seguro para el desarrollo de la economía en general, importancia que hace aconsejable, incluso imprescindible, una regulación y control estrictos que permitan evitar eventuales insolvencias que de producirse podrían ocasionar una pérdida de confianza en el funcionamiento del sector, de consecuencias nefastas para el conjunto de la actividad económica. En definitiva, en este sector, la detección precoz de insolvencias o de las condiciones que pueden llevar a que ésta acaezca es una cuestión de suma importancia e interés general.

Las investigaciones de tipo empírico acerca del fenómeno del fracaso empresarial se han llevado a cabo tradicionalmente a través de la aplicación de técnicas estadísticas de carácter paramétrico, como el Análisis Discriminante o la Regresión Logística, que utilizan ratios financieros como variables explicativas. Dadas las peculiaridades del sector asegurador, la mayoría de dichas investigaciones se centran en otros sectores de la economía. No obstante, en el sector de seguros español, cabe destacar los trabajos realizados por López Herrera, Moreno Rojas y Rodríguez Rodríguez (1994), Mora Enguñados (1994), Martín Peña, Leguey Galán y Sánchez López (1999) y Sanchís Arellano, Gil y Heras Martínez (2003), donde se pone de manifiesto la utilidad de estos métodos para valorar la situación financiera de este tipo de empresas.

Sin embargo, aunque los resultados obtenidos en los trabajos mencionados han sido satisfactorios, todas estas técnicas presentan el inconveniente de que parten de hipótesis más o menos restrictivas acerca de las propiedades distribucionales de las variables explicativas que, especialmente en el caso de la información contable, no se suelen cumplir. Además, dada su complejidad, puede resultar difícil extraer conclusiones de sus resultados para un usuario poco familiarizado con la técnica.

Por contra, las técnicas procedentes del campo de la Inteligencia Artificial, de carácter no paramétrico, no precisan de hipótesis preestablecidas sobre las variables de partida. Dentro de este tipo de técnicas, para el problema que nos ocupa son de gran utilidad las que se encuadran en el área del *Machine Learning* (Aprendizaje Automático). Un representante típico de esta categoría son las redes neuronales, de las que se han desarrollado un buen número de aplicaciones en los más variados campos. En cuanto a su aplicación a la predicción de crisis en el sector de seguros español, Martínez de Lejarza Esparducer (1999) obtiene en su trabajo resultados muy superiores a los del Análisis Discriminante. También el enfoque de los Conjuntos Aproximados —«Rough Sets» en terminología anglosajona— pertenece a esta categoría. Una aplicación al sector de seguros español de la *Teoría Rough Set* (Pawlak, 1991) la podemos encontrar en el trabajo de Segovia Vargas (2003), donde se alcanzan resultados muy satisfactorios y similares a los del Análisis Discriminante, pero evitando con la aplicación del Rough Set los inconvenientes derivados del incumplimiento de las hipótesis requeridas por el Análisis Discriminante sobre las variables de partida.

Un importante conjunto de procedimientos de Aprendizaje Automático es el formado por los distintos sistemas de inducción de árboles de decisión. Dentro de los estudios realizados siguiendo este enfoque, destaca el trabajo, con una orientación marcadamente estadística, de Friedman (1977), que sirvió como base para la construcción del sistema CART (*Classification and Regression Trees*), descrito en Breiman, Friedman, Olshen y Stone (1984).

Más inspirados directamente en el campo de la Inteligencia Artificial son los algoritmos de inducción de árboles de decisión ID3, C4.5 y See5 desarrollados por Quinlan (1979, 1983, 1986, 1988, 1993 y 1997), que han alcanzado una notable repercusión. Algunos estudios que implementan este tipo de metodología para la predicción de crisis empresariales son los de, entre otros, McKee (1995) y González Pérez, Correa Rodríguez y Blázquez Múrez (1999), realizados a partir de una muestra de sociedades cotizadas en bolsa y una muestra de PYMES, respectivamente.

El presente trabajo se inscribe dentro de esta tendencia, cada vez más acusada, a utilizar para el análisis de problemas de naturaleza económica y empresarial técnicas procedentes de las áreas del Aprendizaje Automático y la Inteligencia Artificial. Lo que se pretende en concreto será aplicar al sector de empresas españolas de seguros no-vida el algoritmo See5 a fin de obtener un modelo de predicción de insolvencias en dicho sector basándonos en ratios financieros, comparando además los resultados alcanzados con los que se obtienen aplicando la metodología Rough Set.

El resto del trabajo se estructura de la siguiente forma: en la sección 2 se exponen brevemente los fundamentos teóricos del algoritmo See5. La sección 3 se refiere a la

selección de datos y variables que intervienen en el estudio empírico. En la sección 4 se presentan los resultados obtenidos con el algoritmo See5 y su comparación con el enfoque Rough Set. Finalmente, en la sección 5 exponemos nuestras conclusiones.

2. EL ALGORITMO DE INDUCCIÓN DE ÁRBOLES DE DECISIÓN SEE5

Quizás de entre todos los métodos de aprendizaje automático sean los sistemas de aprendizaje basados en árboles de decisión los más fáciles de utilizar y de entender. Además, la estructura de condición y ramificación de un árbol de decisión es idónea para el problema que nos ocupa, el de clasificación de empresas como sanas o fracasadas.

Las sucesivas ramas de un árbol de decisión van realizando una serie de divisiones excluyentes y exhaustivas en el conjunto de elementos que se quieren clasificar. La diferencia básica entre los distintos algoritmos empleados radica en el criterio utilizado para realizar las particiones mencionadas.

La construcción automática de árboles de decisión se origina con los trabajos en el campo de las ciencias sociales de Morgan y Sonquist (1963) y Morgan y Messenger (1973). En estadística, resulta fundamental la aportación de Breiman *et al.* (1984), que proponen el algoritmo para construir árboles denominado CART (*Classification and Regression Trees*). Más o menos al mismo tiempo los sistemas de inducción de árboles de decisión comienzan a usarse en el campo del aprendizaje automático, donde alcanzan una notable repercusión los algoritmos desarrollados por Quinlan (1979, 1983, 1986, 1988, 1993 y 1997), y en ingeniería (Henrichon y Fu, 1969; Sethi y Sarvarayudu, 1982).

El algoritmo See5 (Quinlan, 1997) es la versión más moderna del ID3 y C4.5 desarrollados por este autor a lo largo de las dos últimas décadas. En él, el criterio utilizado para hacer las particiones se apoya en una serie de conceptos procedentes de la Teoría de la Información y ha experimentado a lo largo del tiempo una serie de notables mejoras. La idea central que comparte con otros algoritmos similares es la de tomar en cada rama del árbol para hacer la correspondiente partición aquella variable que proporciona más información de cara a clasificar los elementos que constituyen el *conjunto de entrenamiento* (el conjunto de datos usados para construir el árbol).

La cantidad de información que proporciona un mensaje o la realización de una variable aleatoria x es inversamente proporcional a la probabilidad de ocurrencia de ese mensaje o de esa realización de la variable aleatoria. Resulta habitual en Ingeniería de Comunicaciones o en Estadística medir esta cantidad en bits (Krippendorff,

1986; Reza, 1994; Ziemer y Tranter, 2002), que se obtienen como $\log_2 \frac{1}{p_x}$. El pro-

medio de esta magnitud para todas las posibles ocurrencias de la variable aleatoria x recibe el nombre de *entropía* de x :

$$H(x) = \sum_x p(x) \log_2 \frac{1}{p(x)}.$$

La entropía es una medida de la aleatoriedad o incertidumbre de x o de la cantidad de información que, en promedio, nos proporciona el conocimiento de x .

Análogamente, se puede definir la *entropía conjunta* de dos variables aleatorias x e y :
$$H(x, y) = \sum_{x,y} p(x, y) \log_2 \frac{1}{p(x, y)},$$

que es la cantidad de información que, en promedio, nos proporciona el conocimiento de x e y . La *entropía condicional* de x dada y , $H(x/y)$, se define como

$$H(x/y) = \sum_{x,y} p(x, y) \log_2 \frac{1}{p(x/y)},$$

y es una medida de la incertidumbre respecto a x cuando se conoce y . Es decir, es la cantidad de información que necesitamos para conocer plenamente x cuando ya tenemos la información suministrada por y . Se cumplirá siempre, lógicamente, que $H(x/y) \leq H(x)$, pues al conocer y tenemos más información que nos puede ayudar a reducir la incertidumbre sobre x . A esa reducción de incertidumbre se la denomina *información mutua* entre x e y : $I(x; y) = H(x) - H(x/y)$ que es la información que una de las variables nos transmite sobre la otra. Se cumple además que $I(x; y) = I(y; x)$. La información mutua es una magnitud similar a la covarianza pero tiene algunas propiedades que la hacen superior a ella.

Podemos considerar que x es una variable aleatoria que nos muestra la clase a la que pertenece un elemento, mientras que y_i , $i=1, 2, \dots, n$, son los atributos o variables que caracterizan a los elementos que queremos clasificar. Así, en nuestro caso, x indicará si la empresa pertenece al grupo de las sanas o al grupo de las fracasadas, mientras que las variables y_i serán los ratios financieros utilizados para la clasificación.

Inicialmente Quinlan seleccionaba para hacer cada partición aquella variable y_i que proporcionaba la máxima información sobre x , es decir, maximizaba $I(x; y_i)$ (magnitud que denominaba *gain*). Aunque esto proporciona buenos resultados, introduce un sesgo en favor de las y_i con muchos valores distintos. Para corregirlo, las sucesivas versiones del algoritmo seleccionan aquella y_i que maximiza la magnitud

$$\frac{I(x; y_i)}{H(y_i)},$$

a la que Quinlan denomina *gain ratio*. Ésta será el porcentaje de la información proporcionada por y_i que es útil para conocer x . Adicionalmente, para evitar seleccionar un atributo simplemente porque su entropía $H(y_i)$ sea pequeña, lo que aumentaría el valor del cociente anterior, se exige además que $I(x; y_i)$ sea razonablemente grande.

El proceso de construcción del árbol de decisión mediante la aplicación reiterada del procedimiento descrito finaliza cuando se alcanza la pureza del nodo, entendiendo por «nodo puro» aquél al que sólo corresponden casos pertenecientes a una de las clases del problema, o cuando la ramificación del árbol ya no suponga ninguna mejora.

Generalmente, este método recursivo de construcción de árboles de decisión conducirá a la generación de árboles muy complejos y excesivamente ajustados a

los datos del conjunto utilizado para dicha construcción. En consecuencia, harán una clasificación cuasi-perfecta. Esto, que en principio puede parecer óptimo, en realidad no lo es, ya que ajustarse demasiado a los datos de entrenamiento suele tener como consecuencia que el modelo sea muy específico y se comporte mal para nuevos elementos, especialmente si tenemos en cuenta que el conjunto de entrenamiento puede contener ruido, lo que hará que el modelo intente ajustarse a los errores, perjudicando su comportamiento global. Éste es un problema que en general presentan todas las técnicas de aprendizaje de un modelo a partir de un conjunto de datos de entrenamiento (Aprendizaje Automático), al que se conoce como «sobreajuste» (*overfitting*).

El modo más frecuente de limitar dicho problema en el contexto de los árboles de decisión consiste en eliminar condiciones de las ramas del árbol, consiguiendo con estas modificaciones la obtención de modelos más generales. Este procedimiento puede verse como un proceso de «poda» del árbol. Esto aumentará el error de clasificación sobre el conjunto de casos de entrenamiento, pero cabe esperar que lo disminuya sobre nuevos casos no usados en la construcción del árbol.

El algoritmo See5 permite combinar dos tipos de poda: prepoda y postpoda. El proceso de prepoda se realiza durante la construcción del árbol impidiendo la ramificación de nodos que contengan un número de ejemplos inferior a una cierta constante. Además, se implementa también un método de postpoda del árbol ajustado inicialmente que consiste en sustituir una rama del árbol por una hoja, en función de una tasa de error prevista o estimada. Consideremos que existe una hoja que cubre N casos clasificando incorrectamente E de ellos, lo que se puede interpretar suponiendo que existe una variable aleatoria que sigue una distribución binomial en la que el experimento se repite N veces obteniendo E errores. A partir de esto se estima la probabilidad de error p_e , que será la tasa de error prevista o estimada. Para ello se realiza una estimación de un intervalo de confianza para la probabilidad de error de la variable binomial y se toma como p_e el límite superior de ese intervalo (será una estimación pesimista). Entonces, para una hoja que cubra N casos, el número de errores previstos será $N \cdot p_e$. Si en lugar de una hoja tenemos una rama el número de errores previstos será la suma de los de cada una de sus hojas. De este modo, una rama será sustituida por una hoja (es decir, la rama será podada) cuando el número de errores previstos de ésta sea menor que el de aquélla.

El algoritmo See5 incorpora además funcionalidades adicionales como, entre otras, un método para obtener a partir del árbol un conjunto de reglas de interpretación más sencilla. Para una descripción detallada acerca de la implementación y características del algoritmo véase Quinlan (1993 y 1997).

Este tipo de sistemas clasificadores presentan indudables ventajas para su aplicación al ámbito financiero, tales como su carácter estrictamente no paramétrico —lo cual es importante teniendo en cuenta que la información que manejaremos en nuestro análisis, ratios contables, no suele cumplir las hipótesis requeridas por las técnicas estadísticas—, que no necesitan de la intervención del experto humano para la inferencia de los modelos clasificadores —ya que éstos se elaboran auto-

máticamente— y que, además, frente a las redes neuronales y métodos similares, de indiscutible poder predictivo, sus resultados son mucho más fácilmente interpretables que los modelos de «caja negra» suministrados por aquéllos.

3. SELECCIÓN DE DATOS Y VARIABLES

En cuanto a la muestra de empresas seleccionada para el análisis, ésta es la utilizada para la aplicación de la metodología Rough Set a la predicción de crisis empresariales en seguros no-vida en el trabajo de Segovia Vargas (2003). Consta de 36 empresas no fracasadas y 36 empresas fracasadas (en adelante, «buenas» y «malas»), emparejadas por tamaño —medido a través del volumen de primas— y tipo de negocio, eliminando así el efecto de estas variables en el estudio, y escogiendo como criterio de selección de las empresas fracasadas el hecho de haber sido intervenidas por la Comisión Liquidadora de Entidades Aseguradoras (CLEA)¹, por entender que se trata de una medida objetivamente determinable de las empresas que fracasan. La muestra abarca datos del periodo comprendido entre 1983 y 1993, extraídos de la publicación anual *Balances y cuentas. Seguros privados* de la Dirección General de Seguros y Fondos de Pensiones.

Una vez tomada la muestra, nos situamos en el periodo anterior al de la insolvencia para tratar de determinar qué indicios de este suceso nos proporcionan los datos de las cuentas anuales en forma de ratios. El éxito o fracaso de una empresa será entendido como una variable dependiente que deberá ser explicada por un conjunto de ratios financieros que actuarán como variables independientes. Así que de cada una de las empresas se han obtenido las cuentas anuales del año previo a la quiebra y, a partir de dicha información, se han calculado una serie de ratios, unos populares en la literatura contable para medir la solvencia empresarial y otros específicos del sector asegurador. En la tabla 1 se presentan los 19 ratios seleccionados.

De las 72 empresas de que consta la muestra original, hemos utilizado únicamente 27 empresas de cada una de las dos submuestras (una formada por las 36 empresas «buenas» y la otra por las 36 empresas «malas») para la elaboración de nuestro modelo, reservando las 9 restantes para poder comprobar la validez del mismo aplicándolo a empresas cuyos datos no hubieran sido utilizados en dicha elaboración. En consecuencia, tendremos una muestra de *entrenamiento* para obtener el árbol y reglas de decisión formada por 54 empresas y una muestra de *validación* para verificar su capacidad predictiva formada por 18 empresas. La selección de la muestra de validación se ha efectuado aleatoriamente.

Para la aplicación del algoritmo se utilizó el software *See5* de *Rulequest Research*.

¹ La CLEA tenía por objeto asumir la liquidación de aquellas entidades aseguradoras que por encontrarse en una situación patrimonial irregular, de previsible insolvencia, les encomendase el Ministerio de Economía y Hacienda o el órgano competente de la respectiva Comunidad Autónoma. En la actualidad, la actividad de la CLEA ha pasado a ser desempeñada por el Consorcio de Compensación de Seguros.

Tabla 1: Ratios empleados.

| Ratio | Definición |
|--------------|--|
| R1 | Fondo de Maniobra / Activo Total |
| R2 | Beneficio antes de Impuestos (BAI) / Capitales propios |
| R3 | Ingresos Financieros / Total Inversiones |
| R4 | BAI* / Pasivo Total BAI* = BAI + Amortizaciones + Provisiones + Resultados Extraordinarios |
| R5 | Total Primas de seguro directo / Capitales propios |
| R6 | Total Primas de negocio neto / Capitales propios |
| R7 | Total Primas de seguro directo / (Capitales propios + Provisiones Técnicas) |
| R8 | Total Primas de negocio neto / (Capitales propios + Provisiones Técnicas netas) |
| R9 | Capitales Propios / Pasivo Total |
| R10 | Provisiones Técnicas / Capitales Propios |
| R11 | Gastos Técnicos de seguro directo / Capitales propios |
| R12 | Gastos Técnicos de negocio neto / Capitales propios |
| R13 | Gastos Técnicos de seguro directo / (Capitales propios + Provisiones Técnicas) |
| R14 | Gastos Técnicos de negocio neto / (Capitales propios + Provisiones Técnicas netas) |
| R15 | (Gastos Técnicos de seguro directo + Comisiones y otros gastos de Explotación) / Total Primas adquiridas de seguro directo |
| R16 | (Gastos Técnicos de negocio neto + Comisiones y otros gastos de Explotación) / Total Primas adquiridas de negocio neto |
| R17 | Provisiones Técnicas de reaseguro cedido / Provisiones Técnicas |
| R18 | Ratio de Siniestralidad = Gastos Técnicos de seguro directo / Primas adquiridas de seguro directo |
| R19 | Ratio de Siniestralidad Neta = Gastos Técnicos de negocio neto / Primas adquiridas de negocio neto |

Fuente: Elaboración propia.

4. RESULTADOS

4.1. ALGORITMO SEE5

A continuación se presenta el árbol de decisión obtenido aplicando el algoritmo See5 a nuestra muestra, así como la evaluación del mismo tanto para la muestra utilizada en su elaboración como para la muestra de validación.

Como se puede observar, en el árbol aparecen únicamente 6 de los 19 ratios iniciales, lo que indica que 13 de los ratios empleados no aportan información relevante para clasificar a las empresas como «buenas» o «malas». Es decir, el árbol construido consta del mínimo número de atributos (variables) que se requieren para la clasificación eficiente de los ejemplos dados, con lo que es claro el gran poder explicativo de esta técnica.

```

R13 > 0.68:
...R9 <= 0.59: mala (14)
:   R9 > 0.59:
:     ...R15 > 0.99: mala (3)
:       R15 <= 0.99: buena (3)
R13 <= 0.68:
...R1 > 0.29: buena (20/2)
R1 <= 0.29:
...R2 > 0.04: mala (3)
R2 <= 0.04:
...R6 > 0.64: buena (3)
R6 <= 0.64:
...R9 <= 0.85: mala (4)
R9 > 0.85: buena (4/1)
    
```

Evaluation on training data (54 cases):

```

          Decision Tree
          -----
          Size          Errors

           8           3 (5.6%)

          (a)   (b)   <-classified as
          ----  ----
           27           (a): class buena
           3           24   (b): class mala
    
```

Evaluation on test data (18 cases):

```

          Decision Tree
          -----
          Size          Errors

           8           5 (27.8%)

          (a)   (b)   <-classified as
          ----  ----
           7           2   (a): class buena
           3           6   (b): class mala
    
```

Nuestro árbol se leería del modo siguiente:

- Si el ratio R13 es mayor de 0,68 y además el ratio R9 es menor o igual de 0,59, la empresa será «mala», siendo 14 el número de empresas de la muestra que verifican este hecho.
- Si el ratio R13 es mayor de 0,68 y el ratio R9 es mayor de 0,59 y el ratio R15 mayor de 0,99, la empresa será «mala», cumpliendo estas condiciones 3 empresas.

Y así continuaríamos descendiendo por el árbol, hasta completar un total de 8 hojas. Obsérvese que al final de cada hoja aparece un valor (n) o (n/m): n representa el número de empresas en la muestra que se clasifican de acuerdo a las condiciones que nos llevan hasta esa hoja y m el número de empresas mal clasificadas.

La evaluación de este árbol de decisión construido con la muestra de entrenamiento (54 empresas) indica que el árbol consta de 8 hojas y comete un total de 3 errores (5,6%), lo que supone un porcentaje de aciertos del 94,4%. También se muestra una *matriz de confusión* que señala el tipo de errores cometidos.

Por último, para comprobar la capacidad predictiva del árbol, se clasifican de acuerdo con éste las 18 empresas de la muestra de validación, obteniendo un porcentaje de clasificaciones correctas del 72,2%.

Aunque no en este caso, en ocasiones puede resultar difícil interpretar un árbol de decisión. See5 permite solventar este problema derivando, a partir del árbol, un conjunto de reglas más simples de la forma *si* (condiciones) - *entonces* (decisión). Las reglas que se obtienen a partir del árbol anterior son las siguientes:

Rules:

```

Rule 1: (20/2, lift 1.7)
  R1 > 0.29
  R13 <= 0.68
  -> class buena [0.864]

Rule 2: (12/1, lift 1.7)
  R2 <= 0.04
  R6 > 0.64
  R13 <= 0.68
  -> class buena [0.857]

Rule 3: (7/1, lift 1.6)
  R9 > 0.85
  -> class buena [0.778]

Rule 4: (14, lift 1.9)
  R9 <= 0.59
  R13 > 0.68
  -> class mala [0.938]

```

```

Rule 5: (7, lift 1.8)
  R13 > 0.68
  R15 > 0.99
  -> class mala [0.889]

Rule 6: (26/6, lift 1.5)
  R1 <= 0.29
  -> class mala [0.750]

Default class: buena

```

Cada regla consiste en:

- Una serie de estadísticas: un número n o n/m , seguido de otro número x ; n y m representan lo mismo que en el árbol y x es el resultado de dividir la precisión estimada de la regla entre la frecuencia relativa de la clase predicha en la muestra de entrenamiento. La precisión de la regla se estima mediante el denominador de Laplace $(n-m+1)/(n+2)$ (Niblett, 1987; Clark y Boswell, 1991).
- Una o más condiciones que deben ser satisfechas para que la regla sea aplicable.
- La clase predicha por la regla.
- Un valor entre 0 y 1 que indica la fiabilidad de la predicción estimada a través del ratio de Laplace.

También existe una clase por defecto (en este caso «buena») para cuando ninguna de las reglas sea aplicable.

El número de errores cometidos al clasificar mediante estas reglas y el tipo de los mismos coinciden, tanto con la muestra de entrenamiento como con la de validación, con los de las clasificaciones hechas con el árbol (aunque no siempre tiene por qué ser así).

A la luz de estos resultados, aquellas partes interesadas en evaluar la solvencia de una firma de seguros no-vida deberían tener en cuenta principalmente las siguientes cuestiones:

- La liquidez, medida a través del ratio R1. Si bien una de las cuestiones más importantes para asegurar el buen funcionamiento de cualquier tipo de empresa es la necesidad de un colchón de liquidez que le permita hacer frente a sus deudas a corto plazo sin tener que recurrir a la realización de sus activos fijos, en el caso de la empresa de seguros dicha necesidad reviste una mayor importancia debido a la inversión del proceso productivo que se da en este tipo de entidades.
- La rentabilidad de los recursos propios, medida a través del ratio R2. Este ratio informa acerca de la eficiencia de la empresa en la utilización de los fondos aportados por los accionistas.

- La solvencia en sentido estricto, medida a través de los ratios R6 y R13. Estos ratios recogen en sus numeradores la medida del riesgo anual de la compañía, ya sea basándose en la cuantía de las primas destinadas a cubrir los riesgos asegurados en el año (R6) o basándose en la valoración de los riesgos que han ocurrido en el ejercicio (siniestros) y que se recogen como Gastos Técnicos en la cuenta de Pérdidas y Ganancias (R13). Los denominadores muestran el soporte financiero de las empresas. En R6 se contemplan los capitales propios, lo que sugiere un soporte global del riesgo independiente de las características anuales, mientras que en R13 se suman a los capitales propios las provisiones técnicas, logrando así el soporte financiero real para el periodo analizado.
- La *autonomía financiera*, medida a través del ratio R9. Este ratio muestra qué proporción de las obligaciones totales de la empresa representan los fondos propios (capital y reservas). Esto confirma la importancia, desde el punto de vista de la solvencia, de la adecuación de estos fondos porque puedan ser necesarios para hacer frente en última instancia a una elevada siniestralidad no prevista.
- La política de tarificación, medida a través del ratio R15. Éste es uno de los ratios utilizados en Estados Unidos por la NAIC (*National Association of Insurance Commissioners*) en su sistema de alarma basado en ratios financieros IRIS (*Insurance Regulatory Information System*) para detectar qué compañías aseguradoras pueden necesitar atención reguladora (este ratio es denominado por la NAIC *Expense Ratio* —ratio de gastos—). Informa de la capacidad de afrontar los gastos anuales con las primas anuales correspondientes, indicando si la firma está llevando a cabo un proceso adecuado de tarificación que le permita hacer frente con los ingresos por primas a los gastos correspondientes.

A pesar de que los resultados que hemos obtenidos son satisfactorios, es posible mejorarlos recurriendo a la opción que incorpora See5 de *adaptive boosting*, basado en el trabajo de Freund y Schapire (1997). Muy brevemente, la idea consiste en generar varios clasificadores (árboles o conjuntos de reglas) en vez de sólo uno. Como primer paso, se construye un único árbol (o conjunto de reglas) del mismo modo que acabamos de ver, que cometerá algunos errores en la clasificación (3 en nuestro caso). Estos errores serán el foco de atención al construir el segundo clasificador en aras de corregirlos. En consecuencia, el segundo clasificador generalmente será diferente al primero y también cometerá errores que serán el foco de atención durante la construcción del tercer clasificador. Este proceso continúa para un número predeterminado de iteraciones o *trials*. Finalmente, cada caso será clasificado ponderando las categorías a las que haya sido asignado por los distintos clasificadores construidos. Los resultados obtenidos con este método son con frecuencia excepcionalmente buenos. Así, en nuestro caso, partiendo del primer árbol de decisión visto anteriormente los resultados que alcanzamos realizando 18 iteraciones son:

- Con la muestra de entrenamiento, el 100% de clasificaciones correctas, como podemos observar en la *matriz de confusión*:

| (a) | (b) | <-classified as |
|-------|-------|------------------|
| ----- | ----- | |
| 27 | | (a): class buena |
| | 27 | (b): class mala |

- Con la muestra de validación, el 83,3% de clasificaciones correctas:

| (a) | (b) | <-classified as |
|-------|-------|------------------|
| ----- | ----- | |
| 7 | 2 | (a): class buena |
| 1 | 8 | (b): class mala |

4.2. COMPARACIÓN CON EL ENFOQUE ROUGH SET

La teoría Rough Set fue originalmente desarrollada en los años ochenta por Z. Pawlak (1991) como herramienta formal para tratar con la incertidumbre o la vaguedad inherente a un proceso de decisión. Es un enfoque que se encuadra también dentro de las aplicaciones de la Inteligencia Artificial.

Podría decirse que en el análisis científico nos encontramos con tres categorías generales de imprecisión. La primera categoría ocurre cuando un acontecimiento es aleatorio en naturaleza, en cuyo caso, la imprecisión asociada con este tipo de acontecimientos puede describirse por la teoría estadística de la probabilidad. La segunda surge del hecho de que los objetos pueden no pertenecer exclusivamente a una única categoría sino que pueden pertenecer a varias categorías aunque con diferentes grados, en este caso, la imprecisión toma forma de pertenencia difusa a un conjunto. Esta segunda categoría es el objeto de la lógica *fuzzy*. La tercera categoría es la teoría *rough set* que es útil cuando las clases en las que han de clasificarse los objetos son imprecisas, pero, sin embargo, pueden aproximarse mediante conjuntos precisos (Nurmi, Kacprzyk y Fedrizzi, 1996). Estas diferencias muestran una de las principales ventajas de la teoría *rough set* y es que no se necesita ninguna información adicional acerca de los datos como puede ser una distribución de probabilidad en estadística o el grado o probabilidad de pertenencia en la teoría *fuzzy set*.

La filosofía del método se basa en la suposición de que con cada objeto del universo que estamos considerando se puede asociar alguna información (datos, conocimiento). Los objetos caracterizados por la misma información no son discernibles en vista de la información disponible. La relación de no diferenciación generó de este modo la base matemática para esta teoría.

La información imprecisa provoca la no diferenciación de los objetos en términos de datos disponibles y evita, en consecuencia, su asignación precisa a un conjunto. «Rough» podría traducirse por «vago, impreciso». Por tanto, y de manera intuitiva, un rough set es un conjunto de objetos que, en general, no pueden ser caracterizados de manera precisa en términos de la información disponible. Si esta información consiste en los valores adoptados por una serie de atributos, diremos que un rough set es un conjunto de objetos (por ejemplo, empresas) que, en general, no pueden ser caracterizados de manera precisa en términos de los valores de dichos atributos (por ejemplo, ratios financieros).

Brevemente, el enfoque *rough set* consiste en descubrir dependencias entre atributos en una tabla de información y reducir el conjunto de los mismos eliminando aquéllos que no son esenciales para caracterizar el conocimiento acerca de los objetos analizados. Un *reducto* se define como el mínimo subconjunto de atributos que asegura la misma calidad de clasificación que el conjunto de todos ellos. De la tabla de información reducida pueden derivarse reglas de decisión en forma de sentencias lógicas (*si* <condiciones> *entonces* <decisión>). Para un mayor detalle acerca de esta teoría remitimos al lector al trabajo de Segovia Vargas (2003), donde se exponen sus principales conceptos y se lleva a cabo la aplicación de esta metodología a la misma muestra de empresas que hemos empleado con el algoritmo See5. El análisis *rough set* de la información se desarrolló utilizando el programa ROSE (Predki, Slowinski, Stefanowski, Susmaga y Wilk, 1998; Predki y Wilk, 1999) creado en la Universidad Tecnológica de Poznań (Polonia).

De la aplicación de la metodología *rough set* a nuestra muestra de entrenamiento se obtuvieron 229 reductos (conjunto de ratios que clasifican a las 54 empresas en «buenas» y «malas» igual que si tomásemos en consideración los 19 ratios originales), de los cuales se seleccionó el formado por los ratios R3, R4, R9, R14 y R15, teniendo en cuenta los siguientes criterios: a) el reducto debía contener el menor número de atributos que fuera posible; b) debía contener los atributos considerados más significativos para la evaluación de la solvencia de las empresas; c) tras seleccionar algunos reductos que cumpliesen los dos primeros requisitos, se exigió una tercera condición consistente en que el reducto que se fuera a seleccionar no contuviera ratios que estuviesen correlacionados.

Hemos de señalar que para la aplicación de este método previamente se hicieron discretos los valores continuos de los ratios. En las aplicaciones prácticas, los sistemas de información a menudo contienen datos de carácter continuo y discreto. El análisis *rough set* de los sistemas de información da buenos resultados cuando los dominios de los atributos son conjuntos finitos cuyo número cardinal es relativamente bajo. Este requisito a menudo se satisface cuando los atributos tienen un carácter cualitativo (discreto). Sin embargo, los atributos que toman valores arbitrarios de unos intervalos dados, es decir, tienen un carácter cuantitativo, deben ser tratados para traducir sus valores a términos cualitativos. Esta manipulación implica una división del dominio original en algunos subintervalos y la asignación de códigos cualitativos a dichos subintervalos (véanse tabla 2 y tabla 3).

Tabla 2: Lista de subintervalos (enfoque Rough Set)

| Ratio | 1° | 2° | 3° | 4° |
|-------|--------------------|------------------|------------------|--------------------|
| R1 | $(-\infty, 0.115]$ | $(0.115, 0.295]$ | $(0.295, 0.475]$ | $(0.475, +\infty)$ |
| R2 | $(-\infty, 0]$ | $(0, 0.1]$ | $(0.1, 0.07]$ | $(0.07, +\infty)$ |
| R3 | $(-\infty, 0.03]$ | $(0.03, 0.06]$ | $(0.06, 0.11]$ | $(0.11, +\infty)$ |
| R4 | $(-\infty, 0.03]$ | $(0.03, 0.08]$ | $(0.08, 0.26]$ | $(0.26, +\infty)$ |
| R5 | $(-\infty, 0.565]$ | $(0.565, 1.565]$ | $(1.565, 3.29]$ | $(3.29, +\infty)$ |
| R6 | $(-\infty, 0.525]$ | $(0.525, 1.38]$ | $(1.38, 2.715]$ | $(2.715, +\infty)$ |
| R7 | $(-\infty, 0.455]$ | $(0.455, 0.725]$ | $(0.725, 1.22]$ | $(1.22, +\infty)$ |
| R8 | $(-\infty, 0.46]$ | $(0.46, 0.7]$ | $(0.7, 1.18]$ | $(1.18, +\infty)$ |
| R9 | $(-\infty, 0.14]$ | $(0.14, 0.35]$ | $(0.35, 0.68]$ | $(0.68, +\infty)$ |
| R10 | $(-\infty, 0.04]$ | $(0.04, 0.545]$ | $(0.545, 2.97]$ | $(2.97, +\infty)$ |
| R11 | $(-\infty, 0.27]$ | $(0.27, 1.095]$ | $(1.095, 2.43]$ | $(2.43, +\infty)$ |
| R12 | $(-\infty, 0.27]$ | $(0.27, 0.845]$ | $(0.845, 1.815]$ | $(1.815, +\infty)$ |
| R13 | $(-\infty, 0.27]$ | $(0.27, 0.49]$ | $(0.49, 0.82]$ | $(0.82, +\infty)$ |
| R14 | $(-\infty, 0.225]$ | $(0.225, 0.435]$ | $(0.435, 0.765]$ | $(0.765, +\infty)$ |
| R15 | $(-\infty, 0.98]$ | $(0.98, 1.055]$ | $(1.055, 1.27]$ | $(1.27, +\infty)$ |
| R16 | $(-\infty, 1]$ | $(1, 1.09]$ | $(1.09, 1.29]$ | $(1.29, +\infty)$ |
| R17 | $(-\infty, 0]$ | $(0, 0.065]$ | $(0.065, 0.19]$ | $(0.19, +\infty)$ |
| R18 | $(-\infty, 0.515]$ | $(0.515, 0.68]$ | $(0.68, 0.785]$ | $(0.785, +\infty)$ |
| R19 | $(-\infty, 0.515]$ | $(0.515, 0.655]$ | $(0.655, 0.75]$ | $(0.75, +\infty)$ |

Fuente: Elaboración propia.

Tabla 3: Asignación de códigos a los subintervalos (enfoque Rough Set)

| Ratio | 1° | 2° | 3° | 4° |
|-------|----|----|----|----|
| R1 | 1 | 2 | 3 | 4 |
| R2 | 1 | 2 | 3 | 4 |
| R3 | 1 | 2 | 3 | 4 |
| R4 | 1 | 2 | 3 | 4 |
| R5 | 1 | 3 | 4 | 2 |
| R6 | 1 | 3 | 4 | 2 |
| R7 | 1 | 3 | 4 | 2 |
| R8 | 1 | 3 | 4 | 2 |
| R9 | 1 | 3 | 4 | 2 |
| R10 | 1 | 3 | 4 | 2 |
| R11 | 1 | 4 | 3 | 2 |
| R12 | 1 | 4 | 3 | 2 |
| R13 | 1 | 4 | 3 | 2 |
| R14 | 1 | 4 | 3 | 2 |
| R15 | 1 | 4 | 3 | 2 |
| R16 | 1 | 4 | 3 | 2 |
| R17 | 1 | 3 | 3 | 2 |
| R18 | 4 | 3 | 2 | 1 |
| R19 | 4 | 3 | 2 | 1 |

Fuente: Elaboración propia.

Una vez seleccionado el reducto, se derivaron 27 reglas de decisión (véase tabla 4), que constituyen un algoritmo de clasificación que puede ser utilizado para evaluar cualquier otra empresa. De las posibles estrategias que incorpora ROSE para la generación de reglas de decisión, se escogió la estrategia consistente en la obtención del mínimo conjunto de reglas que cubriesen todos los objetos de la tabla de decisión, lo que implica la clasificación correcta del 100% de las empresas de la muestra de entrenamiento. Otro tipo de estrategias podrían conducir a sistemas más reducidos y por ende más fáciles de manejar e interpretar en la práctica.

Tabla 4: El algoritmo de 27 reglas (enfoque Rough Set)

| Número de regla | R3 | R4 | R9 | R14 | R15 | Decisión 0 = "mala" 1 = "buena" | Fuerza | Empresas |
|-----------------|----|----|----|-----|-----|---------------------------------------|--------|------------------------------|
| 1 | 2 | | 4 | | | 0 | 4 | E2, E14, E18, E32 |
| 2 | | 3 | | 2 | | 0 | 4 | E17, E30, E43, E35 |
| 3 | 1 | 4 | | | | 0 | 3 | E7, E10, E31 |
| 4 | | 1 | 4 | | | 0 | 3 | E13, E28, E33 |
| 5 | | | 3 | 3 | | 0 | 2 | E1, E4 |
| 6 | 3 | | 1 | | | 0 | 2 | E6, E8 |
| 7 | 2 | | 3 | 4 | | 0 | 2 | E11, E15 |
| 8 | 4 | | | 1 | | 0 | 2 | E12, E16 |
| 9 | 4 | 1 | | | | 0 | 2 | E16, E29 |
| 10 | 2 | | 1 | 3 | | 0 | 1 | E9 |
| 11 | | | 3 | 2 | | 0 | 1 | E3 |
| 12 | 3 | | 2 | | 2 | 0 | 1 | E5 |
| 13 | 2 | 2 | | 2 | | 0 | 3 | E2, E18, E36 |
| 14 | | | | 1 | 3 | 1 | 3 | E102, E113, E132 |
| 15 | | | 2 | | 4 | 1 | 5 | E114, E117, E131, E133, E134 |
| 16 | | 1 | 3 | | | 1 | 4 | E101, E109, E111, E115 |
| 17 | 4 | 2 | | | | 1 | 2 | E106, E110 |
| 18 | 3 | 4 | | | 1 | 1 | 2 | E103, E105 |
| 19 | 1 | 2 | 4 | | | 1 | 2 | E112, E129 |
| 20 | | 2 | | | 2 | 1 | 1 | E135 |
| 21 | | 3 | 1 | | | 1 | 1 | E104 |
| 22 | 2 | | | 1 | | 1 | 2 | E116, E130 |
| 23 | 3 | | | 4 | | 1 | 3 | E107, E111, E115 |
| 24 | 4 | | | 3 | | 1 | 2 | E106, E108 |
| 25 | 1 | 3 | 2 | | | 1 | 1 | E128 |
| 26 | 1 | 1 | | | 2 | 1 | 1 | E136 |
| 27 | | 4 | | 2 | 4 | 1 | 1 | E118 |

Fuente: Elaboración propia.

Finalmente, el poder predictivo de este algoritmo se contrastó con la muestra de validación formada por 18 empresas, obteniendo un 77,78% de clasificaciones correctas.

En la tabla 5 se presentan los porcentajes de clasificaciones correctas obtenidas con los dos métodos (See5 y Rough Set), desagregadas por clase. Tal y como puede observarse en dicha tabla, los resultados de ambas metodologías muestran su capacidad para responder de manera eficiente al problema de la predicción del fracaso empresarial, siendo alternativas muy fiables a las técnicas estadísticas convencionales, y más aún en el caso del algoritmo See5, que obtiene un porcentaje de aciertos superior con la muestra de validación clasificando correctamente el 88,89% de las empresas «malas» frente al 77,78% de la metodología Rough Set, lo cual es importante teniendo en cuenta que precisamente lo que nos interesa captar es la insolvencia.

Tabla 5: Clasificaciones correctas

| See5 | | |
|----------------------------------|---------------------------------|------------------------------|
| Clasificaciones correctas | Muestra de entrenamiento | Muestra de validación |
| Empresas “buenas” | 100% | 77,78% |
| Empresas “malas” | 100% | 88,89% |
| Total | 100% | 83,33% |

| Rough Set | | |
|----------------------------------|---------------------------------|------------------------------|
| Clasificaciones correctas | Muestra de entrenamiento | Muestra de validación |
| Empresas “buenas” | 100% | 77,78% |
| Empresas “malas” | 100% | 77,78% |
| Total | 100% | 77,78% |

Fuente: Elaboración propia.

5. CONCLUSIONES

Los dos métodos que hemos aplicado son estrictamente no paramétricos, lo que les convierte en superiores a las técnicas estadísticas en el sentido de que se adecúan mejor a la información contable, que suele presentar datos interrelacionados, incompletos, adulterados o erróneos; proporcionan modelos muy sencillos entendibles fácilmente por el analista humano, ya sea en forma de árboles o reglas de decisión. Además, dan buenos resultados incluso cuando se trabaja con escaso número de datos, aspecto éste importante en algunas aplicaciones al ámbito financiero.

Sin embargo, cabe señalar algunas ventajas que el algoritmo See5 presenta frente a la metodología Rough Set: ésta trabaja mejor con datos discretos, bien sean variables cualitativas o variables cuantitativas previamente convertidas en discretas, mientras que el algoritmo See5 acepta atributos de tipo discreto o continuo, sin ninguna clase de limitación; el subconjunto de variables relevantes se derivan automáticamente con See5, así como la importancia de cada una de ellas en el proceso de asignación, por contra, la metodología Rough Set proporciona muchos subconjuntos, entre los cuales habrá de seleccionar el analista para derivar las reglas de decisión; además, aunque no se ha visto aquí, See5 posee una ventaja muy importante desde el punto de vista económico, ya que permite considerar distintos costes de clasificación errónea, mientras que Rough Set sólo computa el número global de errores sin distinguir si se trata de clasificar una empresa sana como fracasada o clasificar una fracasada como sana, error este último que resultaría mucho más grave.

Por otro lado, aunque los resultados de la validación externa indican que ambos modelos pueden considerarse adecuados, no debemos olvidar algunas de sus limitaciones a la hora de utilizarlos con fines predictivos, como son el hecho de que la muestra utilizada, además de ser relativamente pequeña, abarca datos del periodo comprendido entre 1983 y 1993, por lo que teniendo en cuenta las notables transformaciones que desde entonces ha sufrido el sector en su regulación, estructura y funcionamiento, extrapolar los resultados a día de hoy no sería aconsejable en modo alguno. Asimismo, debemos tener presente que nuestro análisis se ha realizado al margen de algunos factores, como el factor tamaño. No obstante, si hubiésemos optado por no realizar el emparejamiento de las empresas de acuerdo a su tamaño, el peso tan determinante de dicha variable en la probabilidad de quiebra, especialmente en el sector del seguro, podría haber oscurecido el papel de las variables de carácter financiero en las que estamos especialmente interesados.

También hemos de considerar que, dado el reducido tamaño muestral, habría sido más conveniente utilizar la totalidad de empresas para la elaboración de los modelos y llevar a cabo una validación cruzada «jackknife». Sin embargo, esto carece de sentido para el enfoque Rough Set debido al papel que juega el decisor humano en la selección de reductos. Al eliminar un elemento de la muestra el decisor deberá seleccionar un nuevo reducto y lo razonable es que se trate del mismo original, salvo que sea posible uno con menor número de ratios. Esto ocurrirá en un número muy reducido de casos, que pueden ser considerados como atípicos. Por ello, para mantener la comparabilidad entre las dos técnicas, se ha optado por un método de validación apropiado para ambas.

Además, sería interesante encontrar la manera de examinar la situación de una empresa con información no sólo del periodo anterior al que pretendemos evaluar, sino con varios periodos de desfase. Para comprobar el poder explicativo de los ratios con distintos desfases respecto al periodo que pretendemos examinar, se podrían elaborar diferentes modelos para cada uno de los años previos a la crisis. Asimismo, si nuestro objetivo es diseñar un modelo que pueda predecir la crisis *a priori*, sería conveniente elaborar un modelo teniendo en cuenta los valores de los ratios de varios años consecutivos antes de la crisis tomados conjuntamente. En este sentido, nuestro trabajo podría ser considerado como un primer paso en ese camino.

BIBLIOGRAFÍA

- BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R.A. y STONE, C. J.: Classification and Regression Trees. Belmont, California: Wadsworth International Group, 1984. 358 p. ISBN: 0-412-04841-8.
- CLARK, P. y BOSWELL, R. Rule Induction with CN2: Some Recent Improvements. En: KODRATOFF, Y. *Machine Learning - Proceedings of the Fifth European Conference (EWSL-91)*. Berlin: Springer-Verlag, 1991. pp. 151-163. ISBN: 3-540-53816-X.
- FREUND, Y. y SCHAPIRE, R. E.: A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, Vol. 55, no1, 1997. pp. 119-139.
- FRIEDMAN, J. H.: A recursive partitioning decision rule for non-parametric classification. *IEEE Transactions on Computers*, Vol. 26, n.º 4, 1977. pp. 404-408.
- GONZÁLEZ PÉREZ, A. L., CORREA RODRÍGUEZ, A. y BLÁZQUEZ MÚREZ, J. A.: Perfil del fracaso empresarial para una muestra de pequeñas y medianas empresas. Zaragoza: Comunicación presentada al X Congreso AECA, 1999.
- HENRICHON, Jr., E. G. y FU, K. S.: A nonparametric partitioning procedure for pattern classification. *IEEE Transactions on Computers*, 18, 1969. pp. 614-624.
- KRIPPENDORFF, K.: Information theory: structural models for qualitative data. Beverly Hills: Sage, 1986. 96 p. ISBN: 0-8039-2132-2.
- LÓPEZ HERRERA, D., MORENO ROJAS, J. y RODRÍGUEZ RODRÍGUEZ, P.: Modelos de previsión del fracaso empresarial: Aplicación a entidades de seguros en España. *Esic Market*, 84, abril-junio, 1994. pp. 83-125.
- MARTÍN PEÑA, M. L., LEGUEY GALÁN, S. y SÁNCHEZ LÓPEZ, J. M.: Solvencia y estabilidad financiera en la empresa de seguros: Metodología y evaluación empírica mediante análisis multivariante. Madrid: Cuadernos de la Fundación Mapfre Estudios, no 49, 1999. 254 p. ISBN: 84-894-2932-4.
- MARTÍNEZ DE LEJARZA ESPARDUCER, I.: Previsión del fracaso empresarial mediante redes neuronales: un estudio comparativo con el análisis discriminante. En: BONSON PONTE, E.: *Tecnologías Inteligentes para la Gestión Empresarial*. Madrid: RA-MA Editorial, 1999. pp. 53-70. ISBN: 84-789-7338-9.
- McKEE, T.E. Predicting bankruptcy via induction. *Journal of Information Technology*, 10, 1995. pp. 26-36.
- MORA ENGUÍDANOS, A.: Los modelos de predicción del fracaso empresarial: una aplicación empírica del logit. *Revista Española de Financiación y Contabilidad*, Vol. 23, no 78, 1994. pp. 203-233.
- MORGAN, J. N. y MESSENGER, R. C.: THAID: a Sequential Search Program for the Analysis of Nominal Scale Dependent Variables. Ann Arbor, Michigan: Technical report, Institute for Social Research, University of Michigan, 1973.
- MORGAN, J. N. y SONQUIST, J. A.: Problems in the analysis of survey data, and a proposal, *Journal of the American Statistical Association*, 58, 1963. pp. 415-434.
- NIBLETT, T.: Constructing decision trees in noisy domains. En: BRATKO, I. y LAVRAČ, N. *Progress in Machine Learning (proceedings of the 2nd European Working Session on Learning)*. Wilmslow, UK: Sigma, 1987. pp. 67-78. ISBN: 1-85058-088-X.
- NURMI, H., KACPRZYK, J. y FEDRIZZI, M.: Probabilistic, fuzzy and rough concepts in social choice. *European Journal of Operational Research*, Vol. 95, n.º 2, 1996. pp. 264-277.
- PAWLAK, Z.: Rough Sets. Theoretical Aspects of Reasoning about Data. Dordrecht, Boston, London: Kluwer Academic Publishers, 1991. 252 p. ISBN: 0-792-31472-7.

- PREDKI, B., SLOWINSKI, R., STEFANOWSKI, J., SUSMAGA, R. y WILK, S. ROSE: Software Implementation of the Rough Set Theory. En: POLKOWSKI, L. y SKOWRON, A.: *Rough Sets and Current Trends in Computing, Lecture Notes in Artificial Intelligence*. Berlin: Springer-Verlag, 1998. pp. 605-608. ISBN: 3-540-64655-8.
- PREDKI, B. y WILK, S.: Rough Set Based Data Exploration Using ROSE System. En: RAS, Z. W. y SKOWRON, A.: *Foundations of Intelligent Systems, Lecture Notes in Artificial Intelligence*. Berlin: Springer-Verlag, 1999. pp. 172-180. ISBN: 3-540-65965-X.
- QUINLAN, J. R.: Discovering rules by induction from large collections of examples. En: MICHIE, D.: *Expert systems in the microelectronic age*. Edimburgh: Edimburgh University Press, 1979. 212 p. ISBN: 0-852-24381-2.
- QUINLAN, J. R.: Learning efficient classification procedures. En: MICHALSKI, R. S., CARBONELL, J. G. y MITCHELL, T. M.: *Machine learning: An Artificial Intelligence approach*. Palo Alto, California: Tioga Press, 1983. p. 463-482. ISBN: 0-934613-09-5.
- QUINLAN, J. R. Induction of decision trees. *Machine Learning*, Vol. 1, n.º 1, 1986. pp. 81-106.
- QUINLAN, J. R.: Decision trees and multi-valued attributes. En: HAYES, J. E., MICHIE, D. y RICHARDS, J.: *Machine Intelligence 11 - Towards an Automated Logic of Human Thought*. Oxford, UK: Oxford University Press, 1988. pp. 305-318. ISBN: 0-198-53718-2.
- QUINLAN, J. R.: C4.5: Programs for machine learning. San Mateo, California: Morgan Kaufmann, 1993. 302 p. ISBN: 1-55860-238-0.
- QUINLAN, J. R. See5.URL: <http://www.rulequest.com/see5-info.html>, 1997.
- REZA, F. M.: An introduction to Information Theory. New York: Dover Publications, 1994. 496 p. ISBN: 0-486-68210-2.
- SANCHÍS ARELLANO, A., GIL, J. A. y HERAS MARTÍNEZ, A.: El Análisis Discriminante en la previsión de la insolvencia en las empresas de seguros de no vida. *Revista Española de Financiación y Contabilidad*, Vol. 32, n.º 116, 2003. pp. 183-233.
- SEGOVIA VARGAS, M. J.: Predicción de crisis empresariales en seguros no vida mediante la metodología Rough Set. Tesis Doctoral, Madrid: Universidad Complutense de Madrid, 2003. 275 p.
- SETHI, I. K. y SARVARAYUDU, G. P. R.: Hierarchical classifier design using mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4, 1982. pp. 441-445.
- ZIEMER, R. E. y TRANTER, W. H.: Principles of Communications: Systems, Modulation, and Noise. New York: John Wiley & Sons, 2002. 648 p. ISBN: 0-471-39253-7.