

Original

# Optimización de la predicción de problemas financieros en empresas sanitarias privadas españolas aplicando algoritmos genéticos



Jesús María González-Martín<sup>a,\*</sup>, Agustín J. Sánchez-Medina<sup>b</sup> y Jesús B. Alonso<sup>c</sup>

<sup>a</sup> Unidad de Investigación, Hospital Universitario Gran Canaria Doctor Negrín, Las Palmas de Gran Canaria, España

<sup>b</sup> Instituto Universitario de Ciencias y Tecnologías Cibernéticas (IUCTC), Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, España

<sup>c</sup> Instituto para el Desarrollo Tecnológico y la Innovación en Comunicaciones (IDeTIC), Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, España

## INFORMACIÓN DEL ARTÍCULO

### Historia del artículo:

Recibido el 30 de enero de 2017

Aceptado el 9 de enero de 2018

On-line el 22 de agosto de 2018

### Palabras clave:

Empresas sanitarias  
Concurso de acreedores  
Z-Altman  
Predicción  
Algoritmos genéticos

## R E S U M E N

**Objetivo:** Presentar una metodología para optimizar, a través de la Z'-Score de Altman para empresas privadas, la predicción de entrada en situación de concurso de acreedores (bancarrota) en empresas privadas del sector sanitario español.

**Método:** El método propuesto consiste en la aplicación de los algoritmos genéticos (AG) para encontrar los coeficientes de la fórmula de la cadena de ratios propuestos por Altman en su versión para empresas privadas que optimicen la predicción en empresas privadas sanitarias españolas, maximizando la sensibilidad y la especificidad, y con ello reduciendo los errores de tipo I y tipo II. Con este propósito se ha utilizado una muestra de 5903 empresas del sector sanitario privado español obtenidas de las bases de datos de Sistema de Análisis de Balances Ibéricos (SABI) entre los años 2007 y 2015.

**Resultados:** El modelo predictivo obtenido con los AG presenta mayor exactitud, sensibilidad y especificidad que el propuesto por Altman para empresas privadas, tanto con los datos de test como con todos los datos de la muestra.

**Conclusiones:** El hallazgo más importante del presente estudio es establecer una metodología que logra identificar unos coeficientes optimizados para la Z de Altman, lo cual permite realizar una predicción más precisa de la bancarrota en las empresas sanitarias privadas españolas.

© 2018 SESPAS. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Optimization of the prediction of financial problems in Spanish private health companies using genetic algorithms

### A B S T R A C T

**Objective:** This paper presents a methodology to optimize, using Altman's Z-Score for private companies, the prediction of private companies of the Spanish health sector entering a situation of bankruptcy.

**Method:** The proposed method consists of the application of genetic algorithms (GA) to find the coefficients of the formula of the chain of ratios proposed by Altman in the version of the score for private companies which optimize the prediction for Spanish private health companies, maximizing sensitivity and specificity, and thereby reducing type I and type II errors. For this purpose, a sample of 5,903 companies from the Spanish private health sector obtained from the database of the Iberian Balance Analysis System (SABI) between 2007 and 2015 was used.

**Results:** The results show that the predictive model obtained with the AG presents greater accuracy, sensitivity and specificity than that proposed by Altman for private companies with both test data and all sample data.

**Conclusions:** The most important finding of this study was to establish a methodology that can identify the optimized coefficients for the Altman Z-Score, which allows a more accurate prediction of bankruptcy in Spanish private healthcare companies.

© 2018 SESPAS. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### Keywords:

Health companies  
Bankruptcy  
Z-Altman  
Prediction  
Genetic algorithms

## Introducción

La bancarrota es un evento crítico que puede causar grandes pérdidas a los prestamistas, accionistas, acreedores y naciones; por ello, el desarrollo de modelos de predicción de bancarrota ha sido uno de los principales temas de investigación en el área financiera<sup>1</sup>.

\* Autor para correspondencia.

Correo electrónico: [jesus.gonzalez119@alu.ulpgc.es](mailto:jesus.gonzalez119@alu.ulpgc.es) (J.M. González-Martín).

La literatura de la predicción de la bancarrota comienza en la década de 1930. FitzPatrick<sup>2</sup>, en 1932, comparó 13 ratios de empresas en bancarrota y no en bancarrota (activas) y descubrió que estas últimas poseían ratios más favorables. Hasta mediados de los años 1960, estas predicciones se basaban en un análisis univariable. El trabajo más reconocido fue el de Beaver<sup>3</sup>, quien encontró que un número determinado de ratios podían discriminar entre bancarrota y no bancarrota con hasta 5 años de antelación. En 1968, Edward Altman<sup>4</sup> publicó el primer análisis multivariable. Este autor utilizó el análisis discriminante múltiple para desarrollar un modelo predictivo (denominado Z-Score) usando cinco ratios financieras para predecir la bancarrota de empresas manufactureras:

- 1) (Activo circulante – Pasivo líquido) / Total activo.
- 2) Reservas / Total activo.
- 3) EBIT (*Earnings Before Interest and Taxes*) / Total activo.
- 4) Capitalización bursátil / Deudas totales.
- 5) Ingresos de explotación / Total activo.

Ohlson<sup>5</sup> presentó una alternativa utilizando modelos *logit*, y Zmijewski<sup>6</sup> usando modelos *probit*. En el año 2000, Altman<sup>7</sup> adaptó su fórmula propuesta en 1968 para el sector de las empresas privadas, cambiando la variable *Market value of equity* (capitalización bursátil) por *Book values of equity* (patrimonio neto).

Por otro lado, hasta 1990, las técnicas estadísticas multivariantes más usadas para la predicción de la bancarrota eran el análisis discriminante múltiple<sup>4,8</sup>, el *logit*<sup>5</sup> y el *probit*<sup>6</sup>. Sin embargo, estas técnicas deben cumplir unos supuestos, como la linealidad, la normalidad y la independencia de las observaciones, que no siempre se cumplen en el ámbito financiero<sup>9</sup>. En consecuencia, dichos métodos pueden tener limitaciones de validez<sup>10</sup>.

Desde finales de los años 1980, las técnicas de inteligencia artificial, como las redes neuronales y los algoritmos genéticos (AG), han demostrado ser menos vulnerables a estas suposiciones<sup>11,12</sup> y se han aplicado con éxito en la predicción de la bancarrota<sup>13,14</sup>. El primer artículo del uso de redes neuronales en la predicción de la bancarrota fue publicado en 1990 por Odom y Sharda<sup>15</sup>. En 2002, Shin y Lee<sup>10</sup> utilizaron los AG para predecir la bancarrota. En la [tabla I](#) del Apéndice *online* se muestra un resumen de los principales autores en este campo.

Cabe mencionar que la predicción de la bancarrota ha sido utilizada en multitud de sectores y países (p. ej., en empresas manufactureras<sup>4</sup> y no manufactureras<sup>7</sup>, en empresas de tamaño medio y pequeñas italianas<sup>16</sup>, y en empresas de Corea<sup>10</sup>). Sin embargo, no existe un modelo específico para el sector sanitario privado español. Por ello, este estudio propone la creación de un modelo *ad hoc* para predecir la salud financiera de estas empresas con 2 años de antelación<sup>17,18</sup> utilizando la técnica de los AG y las cinco variables propuestas por Altman<sup>7</sup> para las empresas privadas. De este modo, a través de los AG se calcularon unos nuevos coeficientes para las ratios que componen la Z'-Score de Altman adaptados a este sector (Z'-Score modificada), así como el nuevo punto de corte asociado a estos nuevos coeficientes para discriminar entre empresas en riesgo de problemas financieros y empresas sin riesgos.

La importancia de este sector en España es clara. Así, en el año 2015, el gasto de la sanidad española supuso el 9% del producto interior bruto<sup>19</sup>, del cual el 2,7% correspondió a la sanidad privada, siendo este dato superior respecto a otros países de nuestro entorno (1,7% en Alemania, 2,4% en Francia, 2,2% en Italia) (véase la [tabla II](#) del Apéndice *online*). También se puede observar cómo la colaboración entre el sector público y el privado va en aumento, y parece ser el camino que se pretende tomar en un futuro cercano<sup>20</sup>.

## Método

### Datos

Los datos se obtuvieron de la base de datos SABI (Sistema de Análisis de Balances Ibéricos)<sup>21</sup>. La muestra estuvo formada por aquellas empresas cuya última información disponible se encontraba entre los años 2007 y 2015. Debido a que se decidió realizar predicciones a 2 años, de todas las empresas se tomaron datos en dos cursos económicos diferentes. Del último año disponible se obtenía si la empresa estaba en concurso de acreedores o activa, y del año correspondiente a dos periodos anteriores al comentado se obtuvieron los datos contables necesarios para calcular las ratios propuestas por Altman<sup>8</sup>.

Los códigos CNAE (Clasificación Nacional de Actividades Económicas) utilizados para seleccionar las empresas del sector privado sanitario español fueron los siguientes: 2110, 2120, 3250, 4646, 4773, 4774, 7211, 7219, 7490, 8610, 8621, 8622, 8623 y 8690. De todas ellas, se seleccionaron las que tenían correctamente informadas las cinco ratios propuestas por Altman<sup>7</sup> (véase fórmula 1). Finalmente, la muestra quedó conformada por 5824 empresas activas y 79 en concurso de acreedores. Estas cifras reflejan que la proporción de una categoría es extremadamente pequeña respecto a la otra. En estas situaciones, los modelos de clasificación tienden a predecir a la mayor parte de las empresas en la clase mayoritaria, de tal forma que se aumenta la exactitud a costa de ignorar a la clase minoritaria. Para resolver este problema de datos no balanceados se ha utilizado la técnica de inframuestrear la categoría mayoritaria, de modo que ambas categorías estén igualmente representadas<sup>22</sup>. En la [tabla III](#) del Apéndice *online* se muestra la estadística descriptiva de la muestra.

### Técnica de los algoritmos genéticos

Los AG simulan la teoría de la evolución propuesta por Darwin<sup>16</sup>. Fueron desarrollados por Holland<sup>23</sup> y son procesos estocásticos robustos<sup>1,10</sup> que pueden usarse para resolver problemas de búsqueda y optimización a través de una función *fitness* o función objetivo. Los AG realizan el proceso de búsqueda en cuatro pasos: inicialización, selección, cruce y mutación<sup>10</sup>.

En el proceso de inicialización, una población de individuos, llamados cromosomas, se distribuyen de forma aleatoria por el rango de búsqueda definido para cada coeficiente y son evaluados a través de la función *fitness*<sup>16</sup>.

En el proceso de selección, aquellos cromosomas que presenten valores más óptimos en la función *fitness* serán seleccionados para crear la siguiente generación<sup>16</sup>. En el proceso de cruce se seleccionan dos individuos con valores óptimos en la función *fitness* y se unen de manera que el descendiente contiene información parcial de cada uno de sus progenitores; de esta forma, cabe esperar que el descendiente mejore el resultado de la función *fitness*<sup>10</sup>.

La mutación es un mecanismo por el que de forma aleatoria se seleccionan miembros de la población y se cambia también aleatoriamente alguno de sus genes. De esta forma se desplazan valores numéricos hacia zonas del espacio de búsqueda que no pueden ser alcanzadas por los otros operadores genéticos<sup>10</sup>.

### Análisis estadístico

Para solucionar el problema de los datos no balanceados se recurrió al procedimiento de validación cruzada dejando uno fuera, o *leave-one-out cross-validation* (LOOCV<sup>24</sup>). Esta técnica implica inframuestrear la categoría mayoritaria, seleccionando al azar el mismo número de empresas activas que de empresas en concurso de acreedores. Por ello, se parte de un conjunto de datos con 158 registros (79 empresas activas elegidas de forma aleatoria y 79 en

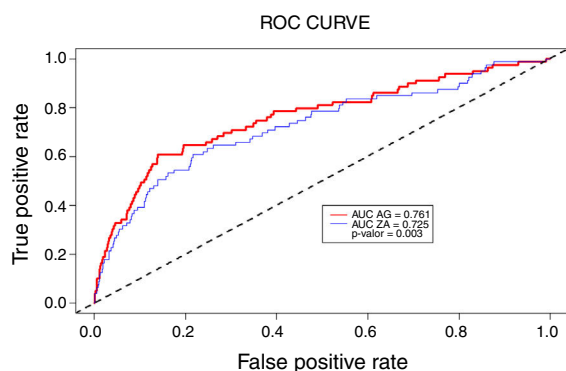


Figura 1. Procedimiento de validación cruzada dejando uno fuera.

Tabla 1  
Matriz de confusión

Clase actual	Predicción	
	Concurso de acreedores	Activa
Concurso de acreedores	Verdaderos positivos (VP)	Falsos negativos (FN)
Activa	Falsos positivos (FP)	Verdaderos negativos (VN)

concurso). Una vez definido el set de datos de entrenamiento con 157 registros y el de validación con un solo registro, al set de entrenamiento se le aplica la técnica de los AG, la cual devuelve diferentes soluciones, cada una de ellas formada con cinco coeficientes y un punto de corte (fig. 1). Por una parte, estas soluciones son acumuladas para obtener los coeficientes definitivos al final del proceso, y por otra parte se calculan las medias de cada coeficiente de estas soluciones, con lo que se obtienen cinco coeficientes y un punto de corte. Con este único resultado se procesa el set de validación y se construye la matriz de confusión. El mismo set es procesado utilizando la fórmula Z'-Altman para poder comparar los resultados obtenidos y determinar si el algoritmo propuesto los mejora (véase la tabla IV del Apéndice online, Fichero Test). El proceso se repite 158 veces, de tal modo que todos los registros son utilizados una vez como set de validación.

Con el objetivo de obtener unos resultados robustos, este proceso se repite 100 veces, es decir, se crean 100 sets de datos, con 158 registros cada uno. Una vez terminado el proceso, se calculan las medias de los coeficientes de todas las soluciones acumuladas durante todo el proceso para obtener los coeficientes definitivos.

El programa estadístico utilizado fue R Core Team (2016)<sup>25</sup>, y se usaron los paquetes GA<sup>26</sup>, fmsb<sup>27</sup> y pROC<sup>28</sup>.

#### Medidas de rendimiento

Las medidas de rendimiento utilizadas se extraen a partir de una matriz de confusión<sup>22</sup> (tabla 1), en la que los verdaderos positivos (VP) indican el número de empresas en concurso clasificadas correctamente. Los falsos negativos (FN) indican el número de empresas en concurso clasificadas erróneamente. Los verdaderos negativos (VN) indican el número de empresas activas clasificadas correctamente. Los falsos positivos (FP) indican el número de empresas activas clasificadas erróneamente.

La función *fitness* utilizada en los AG se ha definido como  $(0,30 * \text{Sensibilidad} + 0,70 * \text{Especificidad})$ , donde la sensibilidad<sup>29</sup> representa la proporción de VP,  $VP / (VP + FN)$ , y la especificidad es la proporción de VN,  $VN / (VN + FP)$ . Así pues, se ha realizado una media ponderada entre la sensibilidad y la especificidad, de tal forma que el modelo tenga en cuenta los dos tipos de empresa, aportando más peso a la especificidad debido a la existencia de datos no balanceados.

También se han utilizado otras medidas, como: a) la exactitud (*accuracy*),  $VP + VN / (VP + FP + FN + VN)$ , aunque presenta problemas cuando las clases no están equilibradas; b) la media geométrica (*G-Mean*), que se define como  $\sqrt{\text{Sensibilidad} * \text{Especificidad}}$ , dando la misma importancia a la sensibilidad y a la especificidad<sup>30</sup>; c) el área bajo la curva ROC, que representa de forma gráfica el porcentaje de empresas activas clasificadas incorrectamente  $(1 - \text{Especificidad})$  en el eje de abscisas y el porcentaje de empresas en concurso clasificadas correctamente (*Sensibilidad*)<sup>30</sup>, y cuando el área bajo esta curva tiene un valor 1, su máximo, implica que el modelo tiene un 100% de sensibilidad y de especificidad, con lo que la capacidad de discriminación del modelo es total, y así, cuanto mayor es su valor, mejor rendimiento tiene el modelo; d) el F1-Score, que representa la media armónica entre la sensibilidad y el valor predictivo positivo,  $VP / (VP + FP)$ ; y e) el coeficiente de correlación de Matthews (MCC) o coeficiente phi, que es una medida que puede utilizarse tanto en datos balanceados como no balanceados:

$$(VP * VN) - (FP * FN) / \sqrt{(VP + FP * (VP + FN) * (VN + FN))}$$

Una vez determinados los nuevos coeficientes y el nuevo punto de corte, se procedió a comparar el poder predictivo de ambos modelos. Para ello, se comparan los resultados obtenidos realizando predicciones con la Z'-Score modificada con los coeficientes calculados en el presente trabajo, con los resultados que se logran a través de la Z'-Score propuesta por Altman para empresas privadas<sup>7</sup>, la cual se detalla a continuación (la elección del trabajo de este autor se debió a que es uno de los más citados en relación a la predicción de la bancarrota<sup>31</sup>):

$$Z' - score \text{ Altman} = 0,717 * X1 + 0,847 * X2 + 3,107 * X3 + 0,42 * X4 + 0,998 * X5 \quad (1)$$

donde:

X1 = (Activo circulante - Pasivo líquido) / Total activo

X2 = Reservas / Total activo

X3 = EBIT / Total activo

X4 = Patrimonio neto / Deudas totales

X5 = Ingresos de explotación / Total activo

Para interpretar esta ratio se siguió el siguiente razonamiento. Si la puntuación Z' obtenida resultaba inferior a 1,23, se encuadraba a la empresa en la categoría de «riesgo de bancarrota», y en caso contrario, en la de «no riesgo de bancarrota».

#### Resultados

En primer lugar, se muestra la estadística descriptiva de los datos de la muestra desglosada por situación: «concurso de acreedores» o «activas». Se observa que las medias y las medianas son mayores en todas las ratios en las empresas activas (tabla 2).

Como ya se ha mencionado, los AG se utilizaron para obtener unos coeficientes mejorados y su respectivo punto de corte para la Z'-Score de Altman<sup>7</sup> (la tabla 3 muestra los parámetros utilizados para controlar el proceso de los AG<sup>22</sup>). Dichos coeficientes, que son calculados como la media de las distintas soluciones óptimas que proporcionan los AG, se muestran en la tabla 4 junto con su mediana y los valores mínimos y máximos. También se muestran los valores originales propuestos por Altman. De este modo, la fórmula 2 presenta el modelo de Altman con los coeficientes mejorados:

$$Z' - score \text{ modificada} = 0,436 * X1 + 0,510 * X2 + 2,79 * X3 + 7,407 * X4 + 2,908 * X5 \quad (2)$$

Para poder interpretar el resultado de la fórmula debe tenerse en cuenta lo siguiente:

**Tabla 2**  
Estadística descriptiva de los registros de la muestra

Variable	Empresas activas (n = 5824)		Empresas en concurso de acreedores (n = 79)	
	Media (DT)	Mediana (P25-P75)	Media (DT)	Mediana (P25-P75)
X1	0,24 (0,32)	0,23 (0,04-0,45)	0,12 (0,36)	0,1 (-0,01-0,36)
X2	0,31 (0,27)	0,26 (0,08-0,5)	0,21 (0,24)	0,11 (0,04-0,31)
X3	0,08 (0,18)	0,05 (0,01-0,13)	-0,01 (0,19)	0,02 (-0,03-0,04)
X4	0,47 (0,33)	0,48 (0,26-0,71)	0,22 (0,31)	0,19 (0,07-0,39)
X5	1,67 (1,7)	1,31 (0,77-2,1)	1,12 (1,24)	0,74 (0,27-1,6)

DT: desviación típica; P: percentil.

X1: (Activo circulante - Pasivo líquido) / Total activo.

X2: Reservas / Total activo.

X3: EBIT / Total activo.

X4: Patrimonio neto / Deudas totales.

X5: Ingresos de explotación / Total activo.

**Tabla 3**  
Parámetros utilizados en los AG

Descripción	Mínimo	Máximo
Rango de búsqueda X1	0,01	1,00
Rango de búsqueda X2	0,01	1,00
Rango de búsqueda X3	2,00	3,50
Rango de búsqueda X4	3,50	9,00
Rango de búsqueda X5	2,00	4,50
Punto de corte	4,00	5,50
Población inicial	1.000	
Tasa de cruce	0,6	
Tasa de mutación	0,1	
Número máximo de iteraciones	80	

X1: (Activo circulante - Pasivo líquido) / Total activo.

X2: Reservas / Total activo.

X3: EBIT / Total activo.

X4: Patrimonio neto / Deudas totales.

X5: Ingresos de explotación / Total activo.

- Si el resultado es mayor que 4,715 (punto de corte obtenido a través de los AG), la previsión es que la empresa no tendrá riesgo.
- Si el resultado es menor o igual que 4,715, la previsión es que la empresa está en riesgo de entrar en concurso de acreedores.

Las principales diferencias entre el modelo original y el propuesto en este trabajo se encuentran en los coeficientes de las variables X4 y X5 (véanse las fórmulas 1 y 2). Una vez calculados los coeficientes mejorados se procedió a comprobar si realmente proporcionan una mejor predicción. La [tabla 5](#) muestra los resultados de las medidas de rendimiento obtenidos durante el proceso de ejecución del conjunto de datos utilizados para el test. Se puede apreciar cierta mejora en el modelo propuesto respecto del modelo original de Altman.

Por otra parte, también se presenta la comparación de ambos modelos aplicados a la totalidad de los registros. Se observa cómo la sensibilidad mejora 7,6 puntos, la especificidad 2,04 puntos y el área bajo la curva ROC 3,6 puntos (estos datos se han obtenido a través de

la matriz de confusión mostrada en la [tabla IV](#) del Apéndice online). En la [figura 1](#) del Apéndice online se muestran las curvas ROC, en las que se aprecia el mayor poder predictivo del nuevo modelo.

## Discusión

La importancia del sector sanitario privado en España es una realidad de la que prácticamente nadie duda. Si a ello se une que el servicio que presta es crítico para la salud de las personas, resulta evidente que contar con empresas saneadas es una cuestión relevante. Así, poder conocer con antelación que una empresa tiene una alta probabilidad de tener problemas financieros es de gran interés para la propia empresa, ya que podría tomar medidas para paliar dicha situación. También lo sería para proveedores y clientes, pues tendrían una mayor información y, por tanto, más seguridad para plantear relaciones a largo plazo.

Por otro lado, cada sector económico posee unas características y peculiaridades financieras que las diferencian de otros. La rotación de activos, la rentabilidad, la liquidez, los ingresos, el endeudamiento, los tiempos de cobro a clientes o pago a proveedores, etc., son solo algunos de los parámetros que los diferencian<sup>32,33</sup>. Consecuentemente, tratar a todas las empresas de la misma forma para predecir si están saneadas o no puede inducir a errores en dicha predicción.

Por todo ello, en el presente trabajo se ha propuesto la mejora de uno de los modelos predictivos más utilizados, la Z'-Score de Altman. Esto se ha hecho estableciendo unos coeficientes para las ratios que propone el autor y un nuevo punto de corte, calculados *ad hoc* para el sector sanitario privado español. En el caso del sector sanitario, la diferenciación parece encontrarse en las variables patrimonio neto, deudas totales e ingresos de explotación. A la luz de los resultados, se observa que los nuevos coeficientes presentan una pequeña mejora respecto a las predicciones obtenidas con el modelo original de Altman<sup>7</sup>, reduciéndose los errores de tipo I y tipo II, es decir, se reducen las predicciones erróneas tanto de empresas en concurso de acreedores como de empresas sanas. No debe

**Tabla 4**  
Coeficientes Z-Score de Altman y estadística descriptiva de las soluciones obtenidas en el proceso AG

Coefficiente de la variable	Z-Score Altman	AG Media (DT)	AG Mediana (P25-P75)	AG Min-Max
X1	0,717	0,44 (0,24)	0,42 (0,24-0,62)	0,01-1,00
X2	0,847	0,51 (0,24)	0,52 (0,33-0,69)	0,01-1,00
X3	3,107	2,79 (0,34)	2,79 (2,54-3,04)	2,00-3,50
X4	0,42	7,41 (1,13)	7,77 (6,84-8,21)	3,50-9,00
X5	0,998	2,91 (0,47)	2,87 (2,57-3,15)	2,00-4,50
Punto de corte	1,23	4,72 (0,29)	4,72 (4,53-4,91)	4,00-5,50

DT: desviación típica; P: percentil.

X1: (Activo circulante - Pasivo líquido) / Total activo.

X2: Reservas / Total activo.

X3: EBIT / Total activo.

X4: Patrimonio neto / Deudas totales.

X5: Ingresos de explotación / Total activo.

El número de soluciones encontradas a lo largo del proceso es 551.346.

**Tabla 5**  
Comparación de las medidas de rendimiento con todos los datos del test y todos los registros

Medidas de rendimiento	Datos test		Todos los registros	
	AG	Z-Score Altman	AG	Z-Score Altman
Prevalencia	50	50	1,34	1,34
Sensibilidad ( <i>recall</i> )	56,37	53,16	60,76	53,16
Especificidad	85,7	84,66	85,49	83,45
Exactitud ( <i>accuracy</i> )	71,03	68,91	85,16	83,04
Valor predictivo positivo (precisión)	79,76	77,61	5,38	4,17
Valor predictivo negativo	66,26	64,38	99,38	99,24
G-Mean	0,70	0,67	0,72	0,67
F1-Score	0,66	0,63	0,10	0,08
MCC (coeficiente phi)	0,44	0,4	0,15	0,11
Índice Kappa	0,42	0,38	0,08	0,05
Área bajo la curva ROC	-	-	0,761	0,725 <sup>a</sup>

MCC: coeficiente de correlación de Matthews.

<sup>a</sup> La diferencia de AUC resulta significativa,  $p = 0,003$ .

pasarse por alto que, en el mundo empresarial, ambos problemas son graves. Predecir que una empresa puede entrar en concurso de acreedores cuando realmente está sana puede estigmatizarla, complicándole el acceso a créditos o provocando que sus clientes y proveedores la abandonen por falta de credibilidad<sup>34</sup>. Por otra parte, si el error es el contrario, es decir, considerarla sana cuando realmente no lo está, pone en riesgo a sus clientes y proveedores.

Por otro lado, puede afirmarse que los resultados obtenidos (exactitud del 85,16%) van en la misma dirección que varios modelos predictivos comentados previamente: para Odom y Sharda<sup>15</sup> del 81,81%, y para Shin y Lee<sup>10</sup> del 80,8%. Así, se obtiene un resultado similar en la evolución de la fórmula Z'-Score de Altman propuesta por Almamy et al.<sup>31</sup>, que obtienen una exactitud del 82,9%. En relación con el área bajo la curva ROC, la mejoría observada en este artículo es similar a la observada por Altman et al.<sup>35</sup> en su comparación entre modelos con información financiera y un segundo modelo en el que se añaden variables con información no financiera.

Finalmente, cabe señalar que las limitaciones de este estudio radican en que el uso de AG no garantiza que se encuentre la solución óptima del problema. Estas soluciones dependerán del rango de búsqueda, del tamaño de la población, y de las ratios de cruce y de mutación aplicadas. En cuanto a futuras líneas de investigación, podrían utilizarse otras técnicas de inframuestreo, como la validación cruzada de k-iteraciones. Por otro lado, sería interesante aplicar la metodología seguida en este trabajo a subsectores más específicos dentro del sector sanitario. Evidentemente, esto solo podrá hacerse cuando se tenga un número suficiente de empresas en concurso de acreedores.

### ¿Qué se sabe sobre el tema?

Se han propuesto diversas fórmulas para predecir la bancarrota en distintos sectores industriales y países. La Z-Score de Altman es una de las más utilizadas en el mundo financiero. Dicha fórmula general para empresas privadas ha sido adaptada para empresas con determinadas características, pero nunca para el sector sanitario.

### ¿Qué añade el estudio realizado a la literatura?

Se aporta una metodología que, para las empresas sanitarias privadas españolas, permite optimizar los coeficientes que multiplican a las ratios propuestos en la Z-Score de Altman, logrando así una fórmula que permite realizar predicciones más eficientes.

## Editor responsable del artículo

David Cantarero.

## Declaración de transparencia

El autor principal (garante responsable del manuscrito) afirma que este manuscrito es un reporte honesto, preciso y transparente del estudio que se remite a GACETA SANITARIA, que no se han omitido aspectos importantes del estudio, y que las discrepancias del estudio según lo previsto (y, si son relevantes, registradas) se han explicado.

## Contribuciones de autoría

A. Sánchez Medina es el responsable del artículo, contribuyó con la concepción y dirección del estudio, la recogida de datos, el análisis, el diseño del estudio, la interpretación de los datos y la elaboración final del manuscrito, y con importantes contribuciones intelectuales. José M<sup>a</sup> González-Martín contribuyó con la concepción del estudio, la recogida de datos, el análisis, el diseño del estudio, la interpretación de los datos y la elaboración final del manuscrito, y con importantes contribuciones intelectuales. Jesús B. Alonso contribuyó con la concepción del estudio, la recogida de datos, el análisis, el diseño del estudio, la interpretación de los datos y la elaboración final del manuscrito, y con importantes contribuciones intelectuales. Todos los autores aprobaron la versión final del trabajo enviado.

## Financiación

Ninguna.

## Conflictos de intereses

Ninguno.

## Anexo. Material adicional

Se puede consultar material adicional a este artículo en su versión electrónica disponible en [doi:10.1016/j.gaceta.2018.01.001](https://doi.org/10.1016/j.gaceta.2018.01.001).

## Bibliografía

- Kim M-JJ, Kang D-KK. Classifiers selection in ensembles using genetic algorithms for bankruptcy prediction. *Expert Syst Appl*. 2012;39:9308-14.
- FitzPatrick P. A comparison of the ratios of successful industrial enterprises with those of failed companies. Washington; 1932.
- Beaver WH. Financial ratios as predictors of failure. *Source J Account Res Empir Res Account Sel Stud*. 1966;4:71-111.
- Altman EI. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *J Finance*. 1968;23:589-609.
- Ohlson JA. Financial ratios and the probabilistic prediction of bankruptcy. *J Account Res*. 1980;18:109-31.
- Zmijewski ME. Methodological issues related to the estimation of financial distress prediction models. *J Account Res*. 1984;22:59-82.
- Altman EI. Predicting financial distress of companies: revisiting the Z-Score and ZETA<sup>®</sup> models background. New York Univ: Stern Sch Business; 2000. p. 1-54.
- Altman EI. Corporate financial distress: a complete guide to predicting, avoiding and dealing with bankruptcy. 3rd ed. New York: Wiley; 1983. p. 368.
- Deakin EB. A discriminant analysis of predictors of business failure. *J Account Res*. 1972;10:167.
- Shin K-S, Lee Y-J. A genetic algorithm application in bankruptcy prediction modeling. *Expert Syst Appl*. 2002;23:321-8.
- Messier WF, Hansen JV. Inducing rules for expert system development: an example using default and bankruptcy data. *Manage Sci*. 1988;34:1403-15.
- Chung H-MM, Tam KY. A comparative analysis of inductive-learning algorithms. *Intell Syst Accounting Financ Manag*. 1993;2:3-18.
- Angelini E, di Tollo G, Rolli A. A neural network approach for credit risk evaluation. *Q Rev Econ Financ*. 2008;48:733-55.

14. Atiya AF. Bankruptcy prediction for credit risk using neural networks: a survey and new results. *IEEE Trans Neural Networks*. 2001;12:929–35.
15. Odom MD, Sharda R. A neural network model for bankruptcy prediction. *Int Jt Conf Neural Networks*. 1990;2:163–8.
16. Gordini N. A genetic algorithm approach for SMEs bankruptcy prediction: empirical evidence from Italy. *Expert Systems with Applications Elsevier*. 2014;64:33–45.
17. Du Jardin p. Bankruptcy prediction using terminal failure processes. *Eur J Oper Res*. 2015;242:286–303.
18. Hennawy RHA, Morris RC. The significance of base year in developing failure prediction models. *J Bus Financ Account*. 1983;10:209–23.
19. OECD/EU. Health at a glance: Europe 2016 state of health in the EU cycle. OECD/EU; 2016.
20. Deloitte. El sector sanitario en España, situación actual y perspectivas de futuro. 2014. p. 26.
21. Bureau Van Dijk. SABI. Sistemas de Análisis de Balances Ibéricos. Vol. null. 2017. p. null.
22. Kim H-JJ, Jo N-OO, Shin K-SS. Optimization of cluster-based evolutionary under-sampling for the artificial neural networks in corporate bankruptcy prediction. *Expert Syst Appl*. 2016;59:226–34.
23. Holland JH. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control and artificial intelligence. University of Michigan Press; 1975. p. 183.
24. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Int Jt Conf Artif Intell*. 1995;2:1137–43.
25. R Core T. R Core Team, 2017. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (Consultado el 29/1/2017.) Disponible en: <http://www.R-project.org/>
26. Scrucca L. GA: a package for genetic algorithms in R. *J Stat Softw*. 2013;53:1–37.
27. Nakazawa M. fmsb: functions for medical statistics book with some demographic data. R Packag version 052. 2015. p. 1-40.
28. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:1–77.
29. Altman DG, Bland JM. Diagnostic tests 1: sensitivity and specificity. *BMJ*. 1994;308:1552.
30. Barandela R, Sánchez JS, García V, et al. *Stra8nit*. 2003;36:849–51.
31. Almamy J, Aston J, Ngwa LN. An evaluation of Altman's Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: evidence from the UK. *J Corp Financ*. 2016;36:278–85.
32. Etiennot H, Preve L, Allende VS. Working capital management: an exploratory study. *J Appl Financ*. 2012;1:162–76.
33. Rodríguez Ferrer T, Hernández García MC, González Gómez JI. Contribución del tamaño y el sector en la explicación de la rentabilidad empresarial. *Rev Española Financ y Contab*. 2000:903–30.
34. Etheridge HL, Sriram RS, Hsu HYK. A comparison of selected artificial neural networks that help auditors evaluate client financial viability. *Decis Sci*. 2000;31:531–50.
35. Altman EI, Sabato G, Wilson N. The value of non-financial information in small-sized enterprise risk management. *J Credit Risk*. 2010;6:1–33.