

PRONÓSTICO DE INSOLVENCIA
EMPRESARIAL EN COLOMBIA
A TRAVÉS DE INDICADORES
FINANCIEROS

Diego Andrés Correa-Mejía
Universidad de Antioquia
Colombia

Mauricio Lopera-Castaño
Universidad de Antioquia
Colombia

Diego Andrés Correa-Mejía
Mauricio Lopera-Castaño
JEL: G10, G11, G12

Pronóstico de insolvencia empresarial en Colombia a través de indicadores financieros

Resumen

La insolvencia empresarial afecta tanto a las empresas que entran en este proceso como a sus proveedores de bienes y servicios. Esta investigación hace uso de indicadores financieros para pronosticar la insolvencia empresarial con un año de anticipación. El estudio fue aplicado a 2.988 empresas que reportaron información financiera a la Superintendencia de Sociedades (Colombia) durante el año 2017, de las cuales 127 entraron en proceso de insolvencia en 2018. El pronóstico considera indicadores financieros de liquidez, rentabilidad y endeudamiento, y contrasta los resultados de la regresión logística con el algoritmo boosting. Se concluye que los indicadores financieros permiten pronosticar la insolvencia empresarial, sin embargo se debe recurrir a metodologías no tradicionales como el algoritmo boosting que consideren la asimetría de la información.

Palabras clave: Insolvencia empresarial, indicadores financieros, análisis financiero, algoritmo boosting, regresión logística.

Prévisions de l'insolvabilité des entreprises en Colombie au moyen d'indicateurs financiers

Résumé

L'insolvabilité commerciale affecte à la fois les entreprises qui entrent dans ce processus et leurs fournisseurs de biens et services. Cette recherche utilise des indicateurs financiers pour prévoir l'insolvabilité des entreprises un an à l'avance. L'étude a été appliquée à 2988 entreprises qui ont communiqué des informations financières à la Surintendance des entreprises (Colombie) en 2017, dont 127 sont entrées en insolvabilité en 2018. La prévision prend en compte les indicateurs financiers de liquidité, de rentabilité et d'endettement et oppose les résultats de la régression logistique à l'algorithme de boosting. Il est conclu que les indicateurs financiers permettent de prévoir l'insolvabilité des entreprises. Cependant, des méthodologies non traditionnelles telles que l'algorithme de boosting qui considèrent l'asymétrie d'information devraient être utilisées.

Mots clés: Insolvabilité des entreprises, indicateurs financiers, analyse financière, algorithme de boosting, régression logistique.

Forecast of business insolvency in Colombia through financial indicators

Abstract

Business insolvency affects both companies that enter this process and their suppliers of goods and services. This research uses financial indicators to forecast business insolvency one year in advance. The study was applied to 2,988 companies that reported financial information to the Superintendency of Companies (Colombia) during 2017, of which 127 went into insolvency in 2018. The forecast considers financial indicators of liquidity, profitability and indebtedness, and contrasts the results of the logistic regression with the boosting algorithm. It is concluded that financial indicators allow predicting business insolvency. However, non-traditional methodologies such as the boosting algorithm that consider the information asymmetry should be used.

Keywords: Business insolvency, financial indicators, financial analysis, boosting algorithm, logistic regression.

Artículo de Investigación

Pronóstico de insolvencia empresarial en Colombia a través de indicadores financieros¹

INFORMACIÓN DEL ARTÍCULO

Recepción del artículo: 17/09/2018
Concepto de evaluación: 29/10/2018
Aceptación del artículo: 15/12/2018

Diego Andrés Correa-Mejía*
Universidad de Antioquia, Colombia

Mauricio Lopera-Castaño
Universidad de Antioquia, Colombia

INTRODUCCIÓN

La insolvencia empresarial es regulada en Colombia mediante la ley 1116 (2006) y se considera un estado especial dentro del sistema de la bancarrota toda vez que allí entran las empresas con problemas financieros que buscan opciones para no llegar directamente a la desaparición empresarial (Upegui y Londoño, 2011). Según lo planteado por Wilches (2008), en el momento en el que una empresa entra en proceso de insolvencia empresarial, tanto el Estado como sus proveedores de bienes y servicios comienzan a servirle de apoyo con el fin de evitar su extinción y conservar los empleos que la empresa genera. En este sentido, la empresa insolvente tiene la posibilidad de renegociar sus deudas bajo nuevos parámetros de acuerdo con su capacidad de pago (Rodríguez, 2007).

A pesar de que el propósito de la ley 1116 (2006) es conservar el dinamismo de la economía previniendo la rápida desaparición de las empresas y preservando el empleo (Vélez, 2011), es de resaltar que la situación financiera de los acreedores se afecta ya que no recibirán los flujos de caja por los pagos del cliente insolvente en el tiempo que tenían previsto (Jovanovik, Todorovic, y Grbic, 2017). De esta manera, los proveedores en general deberían hacer un estudio minucioso sobre la situación financiera de sus potenciales clientes con el fin de medir el riesgo que tienen de incurrir en un estado de insolvencia en el futuro (Jayasekera, 2018). Esta investigación se desarrolla con el objetivo de pronosticar la insolvencia empresarial con un año de anticipación haciendo uso de indicadores financieros.

Estudios previos han concentrado sus esfuerzos en predecir la quiebra empresarial por medio del uso de medidas financieras. Los primeros estudios desarrollados en el tema fueron los de Beaver (1966) y Altman (1968) quienes combinaron estrategias de análisis financiero con modelos estadísticos con el fin de medir la probabilidad que determinada empresa tenía de entrar en procesos de bancarrota. A partir de

1. Resultados del proyecto de investigación "Pronóstico de insolvencia en Colombia", financiado por la Universidad de Antioquia (Acta número: 2017-18709).

*Autor de correspondencia: diegoa.correa@udea.edu.co

estos estudios, el tema de la quiebra empresarial ha sido desarrollado tanto desde el enfoque estadístico como financiero.

No obstante, según lo planteado por Amendola, Giordano, Parrella, y Restaino (2017) existe una brecha en este tipo de estudios ya que desde el enfoque estadístico se han desarrollado modelos predictivos cada vez más precisos, pero no se ha hecho una adecuada delimitación de la bancarrota empresarial, ni se ha hecho un uso adecuado de los indicadores financieros para realizar la predicción. Por otro lado, desde el enfoque financiero se han establecido los indicadores financieros más apropiados en la evaluación de continuidad empresarial, sin embargo, las metodologías utilizadas para la predicción han sido tradicionalistas y sin considerar la asimetría de información que se presenta en los procesos de bancarrota (Calabrese y Osmetti, 2015).

Este estudio se realizó considerando la información financiera de 2.988 empresas que reportaron información financiera a la Superintendencia de Sociedades durante el año 2017, de las cuales 127 entraron en proceso de insolvencia en 2018. Se eligieron indicadores financieros de liquidez, rentabilidad y endeudamiento considerando lo expuesto por Correa-garcía (2005), para realizar la predicción se aplicó el modelo de regresión logística y el algoritmo boosting con el fin de contrastar los resultados entre la metodología tradicional y no tradicional respectivamente.

El desarrollo del artículo se encuentra compuesto por la revisión de la literatura, luego se presenta la metodología donde se describen las características de la muestra estudiada para realizar el pronóstico y se describen los modelos utilizados. Posteriormente, se presentan los resultados obtenidos donde se realiza el contraste entre los resultados del modelo predictivo tradicional y del no tradicional. Por último, se presentan las conclusiones, recomendaciones y líneas de investigación futuras que pueden ser estudiadas a partir del estudio realizado.

REVISIÓN DE LA LITERATURA

La insolvencia empresarial en Colombia es regulada por la ley 1116 (2006) con el fin de preservar la empresa como unidad de explotación económica y fuente generadora de empleo (Correa, Pulgarín, Muñoz, y Álvarez, 2010). De acuerdo con lo establecido por Wilches (2008), la ley diferencia dos estados especiales: la insolvencia o reorganización empresarial y la liquidación judicial. Las empresas que se declaran en estado de insolvencia, deben cumplir con al menos uno de los dos requisitos expuestos por la ley 1116 (2006): cuando se encuentre en cesación de pagos o cuando exista incapacidad de pago inminente.

La cesación de pagos se presenta cuando una empresa incumple el pago de dos o más obligaciones con sus proveedores por más de noventa días. De acuerdo con lo planteado por Wilches (2008) cuando se incurre en cesación de pagos, los acreedores tienen el derecho de cobrar la obligación, sin embargo, el deudor también tiene el derecho de declararse insolvente con el fin de pagar sus obligaciones de manera ordenada y evitar un deterioro mayor de su situación financiera. El estado de insolvencia que se presenta por causa de cesación de pagos puede ser aclamado tanto por la empresa deudora como por sus proveedores de bienes y servicios demostrando los tiempos de morosidad que tiene el deudor (Romero, Melgarejo, y Vera, 2015).

Por otra parte, el estado de insolvencia también se puede presentar cuando exista incapacidad de pago inminente. Según lo establece la ley 1116 (2006), el estado de insolvencia empresarial argumentado en la existencia de incapacidad de pago inminente solo puede hacerse desde el interior de la empresa deudora y debe estar sustentado en condiciones de mercado o estructurales que puedan afectar razonablemente el cumplimiento de las obligaciones de una empresa (Wilches, 2009). En la identificación de las condiciones estructurales interviene la situación financiera de la empresa deudora y a partir de los indicadores financieros es posible medir la probabilidad de incurrir en estado de insolvencia. En este contexto, a partir de la evaluación financiera de la empresa, la administración puede argumentar el estado de insolvencia empresarial por incapacidad de pago inminente.

Los primeros estudios de predicción de la bancarrota fueron desarrollados por Beaver (1966) y Altman (1968) quienes combinaron la información financiera con modelos estadísticos con el fin de tener una mejor aproximación al momento de definir sobre los riesgos de continuidad que podría tener una empresa. A partir de ese momento, se ha producido un desarrollo teórico y empírico sobre la predicción de quiebra empresarial bajo el enfoque estadístico y financiero.

De acuerdo con Amendola et al. (2017) existe una brecha entre los dos enfoques ya que desde el enfoque estadístico los avances se han basado en el desarrollo de modelos predictivos que permitan obtener mayor precisión en la medición de la probabilidad de quiebra, descuidando el adecuado uso de los indicadores financieros y los estudios desarrollados bajo el enfoque financiero han centrado sus análisis en el correcto uso de indicadores financieros pero usando modelos tradicionales que no consideran la distribución de los datos y por ende no se consiguen resultados precisos.

Barboza, Kimura, y Altman (2017) exponen que el estudio de la quiebra empresarial ha pasado de las metodologías tradicionales como el análisis discriminante y la regresión logística a metodologías más modernas como redes neuronales, máquinas de soporte vectorial, algoritmo boosting, entre otras. El avance metodológico ha surgido principalmente porque las metodologías tradicionales no consideran la asimetría de información de los datos (Le, Son, Vo, Lee, y Baik, 2018). Según lo expuesto por Calabrese y Osmetti (2013) la quiebra es un evento raro en el que el número de empresas que no están en este estado superan en gran cantidad al número de empresas quebradas, lo cual afecta directamente el proceso predictivo.

Con el fin de superar la barrera de la asimetría de información que trae consigo la distribución de los datos, el desarrollo de los estudios que utilizan metodologías tradicionales como Ben (2017), Kovacova y Kliestik (2017) y Cultrera y Brédart (2016) han sesgado la muestra utilizando la misma cantidad de empresas quebradas y no quebradas en la predicción con el fin de mejorar la precisión del modelo.

No obstante, la falencia que tienen los estudios desarrollados desde el enfoque estadístico según lo plantea Amendola et al. (2017) consiste en que no le han brindado la importancia requerida a los indicadores financieros como variables predictivas, lo cual es visible en la cantidad de indicadores y de categorías financieras que se utilizan para la predicción. En el estudio realizado por Tian y Yu (2017) se utilizaron 29 indicadores financieros, Ben (2017), Jabeur y Fahmi (2018) utilizaron una batería de 33 indicadores financieros para desarrollar sus respectivos modelos.

El uso de grandes cantidades de indicadores afecta negativamente la practicidad de los modelos (Dinca, Baba, Dinca, Dauti, y Deari, 2017) ya que se requiere mayor información financiera sobre las empresas para realizar la evaluación de su riesgo crediticio (Eling y Jia, 2018). En este sentido, es necesario considerar las categorías y tipos indicadores que permiten evaluar la situación financiera de las empresas. De acuerdo con Correa-garcía (2005) el desempeño financiero de las empresas se encuentra dividido en tres categorías: liquidez, rentabilidad y endeudamiento.

La liquidez es un aspecto de corto plazo que mide la capacidad que tienen las empresas para hacer frente a sus obligaciones inmediatas (Gupta, Gregoriou, y Ebrahimi, 2018). A partir de este análisis se pueden evaluar riesgos de liquidez (default) que se presentan cuando las empresas no tienen los flujos de caja o los recursos fácilmente convertibles en efectivo para responder ante sus proveedores de bienes y servicios (Jovanovik et al., 2017). De acuerdo con Dinca et al. (2017) través del análisis de la liquidez, los acreedores pueden evaluar la capacidad de pago que tiene un potencial cliente y con base en este análisis es posible determinar el monto de crédito que puede ser aprobado.

Por otra parte, la rentabilidad es un aspecto organizacional que da cuenta de la sostenibilidad de la empresa en el largo plazo (Jabeur y Fahmi, 2018). A través de los análisis de rentabilidad se hacen relaciones entre la cantidad que han invertido tanto los accionistas como la empresa y los resultados que se obtienen en la explotación del objeto social (Jayasekera, 2018). Según lo expuesto por Bredart, Vella, y Bonello (2018), el análisis de la rentabilidad además de ser un aspecto fundamental en la evaluación del desempeño financiero empresarial, también genera la oportunidad de conocer la generación de flujos de efectivo futuros. De acuerdo con Correa-garcía (2005), en el largo plazo tanto la liquidez como la rentabilidad deben converger.

Por último, el endeudamiento es la tercer categoría que según Hebous y De Mooij (2018) siempre se debe considerar en el análisis financiero. Por medio del análisis del endeudamiento no solo se puede evaluar el nivel de deuda que tiene una empresa (Lyandres y Zhdanov, 2013) sino que también es posible establecer el grado de compromiso que tienen los flujos de caja futuros y de esta manera establecer la capacidad de pago (Yazdanfar y Öhman, 2015).

En este estudio se consideraron indicadores de las tres categorías con el fin de lograr una evaluación financiera holística de las empresas consideradas. Adicionalmente, se hizo uso de una metodología tradicional como lo es la regresión logística con el fin de contrastar los resultados predictivos con la metodología no tradicional como es el algoritmo boosting que según Jones, Johnstone, y Wilson (2017) considera la asimetría de los datos y se pueden lograr mejores predicciones que con las metodologías tradicionales.

METODOLOGÍA

En este apartado se describe la muestra de las empresas utilizadas en el desarrollo del estudio, se establecen los indicadores financieros utilizados y su forma de medición y por último, se exponen las características tanto del modelo de regresión logística como el del algoritmo boosting.

MUESTRA

Para el desarrollo de la investigación, la información financiera fue recolectada de los reportes que se encuentran en la Superintendencia de Sociedades con corte al 31 de diciembre del 2017. Adicionalmente, la información de las empresas declaradas en proceso de insolvencia empresarial en el año 2018 fue tomada de la misma Superintendencia.

Inicialmente, la cantidad de empresas que reportaron información financiera a la Superintendencia de Sociedades con corte al 2017 fue de 16.062, no obstante, debido a que algunas empresas no presentaban información necesaria para el cálculo de los indicadores financieros o presentaban inconsistencias en la información, se realizaron las depuraciones necesarias con el fin de tener la información precisa para desarrollar los modelos predictivos.

Luego de realizadas las depuraciones, la muestra final quedó representada en 2.861 empresas no insolventes y 127 empresas insolventes. En la tabla 1 se muestra la cantidad de empresas estudiadas y su distribución entre insolventes y no insolventes.

Tabla 1.
Empresas insolventes y no insolventes

Categoría	Empresas	Frecuencia
No insolventes	2.861	96%
Insolventes	127	4%
Total	2.988	100%

La composición de la muestra exhibe claramente la existencia de la asimetría en la distribución de los datos ya que el 4% son empresas que entraron en estado de insolvencia para el año 2018 y el 96% son empresas no insolventes, lo cual es coherente por lo expuesto por Calabrese y Osmetti (2015).

VARIABLES

La variable dependiente del estudio es la insolvencia empresarial, la cual es dicotómica y toma el valor de 0 cuando una empresa no está en estado de insolvencia y el valor de 1 cuando se encuentra en insolvencia. Dado que la variable de respuesta se encuentra entre 0 y 1, el resultado que se espera es la medición de la probabilidad que una empresa entre en proceso de insolvencia utilizando información financiera con un año de anticipación.

Los indicadores financieros utilizados para realizar la estimación de los modelos se encuentran dentro de las categorías de liquidez, rentabilidad y endeudamiento, siguiendo lo expuesto por Correa-garcía (2005). Para el desarrollo de la predicción se utilizaron en total 15 indicadores financieros de los cuales 4 corresponden a la categoría de liquidez, 7 a rentabilidad y 4 a endeudamiento.

En la tabla 2 se presenta el resumen de los indicadores financieros utilizados para la predicción, su abreviatura y forma de medición.

Tabla 2.
Resumen indicadores financieros

Categoría	Nombre	Abreviación	Medición
	Ciclo operacional	Cic_op	DI + DC [1]
Liquidez	Razón corriente	Raz_cte	$\frac{\text{Activo_corriente}}{\text{Pasivo_corriente}}$ [2]
	Relación CTNO/ ventas	Rel_CTNO_Ventas	$\frac{\text{CTNO}}{\text{Ventas}}$ [3]
	Relación flujo de caja libre operacional y ventas	FCL_Ventas	$\frac{\text{FCLO}}{\text{Ventas}}$ [4]
Rentabilidad	Rentabilidad del activo	ROA	$\frac{\text{Utilidad_Neta}}{\text{Activo_Total}}$ [5]
	Rentabilidad del patrimonio	ROE	$\frac{\text{Utilidad_Neta}}{\text{Patrimonio}}$ [6]
	Productividad del activo no corriente operacional	Ptividad_Anoco	$\frac{\text{Ventas}}{\text{ANOCO}}$ [7]
	Margen bruto	Mg_bruto	$\frac{\text{Utilidad_Bruta}}{\text{Ventas}}$ [8]
	Margen operacional	Mg_operacional	$\frac{\text{Utilidad_Operacional}}{\text{Ventas}}$ [9]
	Margen antes de impuestos	Mg_antesimpto	$\frac{\text{UAI}}{\text{Ventas}}$ [10]
	Margen neto	Mg_netto	$\frac{\text{Utilidad_Neta}}{\text{Ventas}}$ [11]
Endeudamiento	Nivel de endeudamiento	Niv_end	$\frac{\text{Pasivo}}{\text{Pasivo} + \text{Patrimonio}}$ [12]
	Apalancamiento financiero	Apa_fro	$\frac{\text{Pasivo_financiero}}{\text{Patrimonio}}$ [13]
	Apalancamiento total	Apa_total	$\frac{\text{Pasivo}}{\text{Patrimonio}}$ [14]
	Cobertura de intereses desde la utilidad operacional	Cob_int_UO	$\frac{\text{Utilidad_Operacional}}{\text{Gasto_financiero}}$ [15]

Fuente: Elaboración propia.

Donde:

DI: Días de inventarios.

DC: Días de cuentas por cobrar.

CTNO: Capital de trabajo neto operacional

FCLO: Flujo de caja operacional

ANOCO: Activo no corriente operacional

UAI: Utilidad antes de impuestos

En los estudios relacionados con la predicción de la insolvencia empresarial, la varianza de los parámetros se debe estabilizar con el fin de optimizar los resultados de los modelos predictivos (Jones et al., 2017). Esta situación se presenta ya que las empresas consideradas tienen resultados financieros dispersos y se presentan datos atípicos dentro de los indicadores financieros, lo cual se debe a que los indicadores financieros de las empresas no insolventes en general tienen comportamientos diferentes a los indicadores de empresas en insolvencia (Nyitrai y Virág, 2018). Para el desarrollo de este estudio, se hizo uso de la transformación propuesta por Yeo and Johnson (2000) ya que posibilita el trabajo con variables mayores, iguales o menores a cero.

MODELO

Para realizar el contraste de los resultados predictivos, a continuación se explican los modelos utilizados para el desarrollo de la investigación.

REGRESIÓN LOGÍSTICA

La regresión logística es útil cuando la variable de interés es dicotómica. Como lo aseguran Pérez, Lopera, y Vásquez (2017) en este tipo de modelo, la variable dependiente solo puede tomar dos valores que son complementarios. En este sentido, el resultado de la estimación de la regresión logística arroja un valor que se encuentra entre 0 y 1, y que sirve para predecir la probabilidad de que se produzca el evento definido en función de las variables independientes definidas.

El modelo de regresión logístico está explicado por la siguiente ecuación:

$$P_i = \frac{e^z}{1 + e^z} \quad [1]$$

Donde representa la probabilidad de que una empresa entre en proceso de insolvencia empresarial y x_i son las variables independientes que se relacionan en la tabla 2.

ALGORITMO BOOSTING

El algoritmo boosting como lo señalan Hastie, Tibshirani, y Friedman, (2017), es una poderosa herramienta de clasificación que combina la salida de muchos clasificadores débiles para producir un poderoso comité. El algoritmo más popular de este método fue desarrollado por Freund y Shapire (1997) y es conocido como el AdaBoost.M1. Para introducir este algoritmo los autores consideran un problema de clasificación binaria donde la variable de respuesta es codificada como $Y_i \in \{-1, 1\}$. Dado un vector de variables explicativas $x_i = (1, x_{2i}, x_{3i}, \dots, x_{ki})'$, un clasificador $G(x_i)$ produce una predicción tomando uno de los dos valores $\{-1, 1\}$. La tasa de error en el conjunto de entrenamiento es

$$\bar{e} = \frac{1}{N} \sum_{i=1}^N I(y_i \neq G(x_i)) \quad [2]$$

La media del error de las predicciones futuras es $E(I(y_i \neq G(x_i)))$. Un clasificador débil es uno cuya tasa de error es ligeramente mejor que la de un clasificador aleatorio. El propósito del Boosting es secuencialmente aplicar el clasificador débil a versiones modificadas de los datos, por lo tanto produciendo una secuencia de clasificadores débiles $G_m(x_i)$, $m = 1, 2, \dots, M$. Las predicciones finales de todos los clasificadores son combinadas a través de una mayoría de voto ponderada para producir las predicciones finales.

$$G(x_i) = \text{sign} \left(\sum_{m=1}^M \alpha_m G_m(x_i) \right) \quad [3]$$

En este caso los $\alpha_1, \alpha_2, \dots, \alpha_M$ son calculados por el algoritmo boosting, y dan el peso que tiene cada clasificador individual en el clasificador final. Su efecto es dar más alta influencia a los clasificadores más precisos en la secuencia. Las modificaciones de los datos en cada paso consisten en aplicar pesos w_1, w_2, \dots, w_N a cada una de las observaciones de entrenamiento (x_i, y_i) , $i = 1, 2, \dots, N$. Inicialmente todos los pesos son hechos $w_i = 1/N$, así que en el primer paso el clasificador entrena el modelo en los datos en la forma usual. Para cada una de las m iteraciones sucesivas los w_i son modificados y el algoritmo de clasificación es aplicado a estos datos modificados. En el paso m , las observaciones que fueron mal clasificadas por el clasificador $G_{m-1}(x_i)$ tienen su peso incrementado, mientras los pesos son disminuidos para aquellas observaciones que fueron correctamente clasificadas. Entonces a medida que el algoritmo progresa las observaciones que son más difíciles de clasificar correctamente reciben mayores pesos. Cada clasificador sucesivo es por lo tanto forzado a concentrarse en aquellas observaciones de entrenamiento que son mal clasificadas en el paso previo. Formalmente los pasos del algoritmo AdaBoost.M1 son los siguientes:

1. Haga $w_i = 1/N$ para $i = 1, 2, \dots, N$.
2. Para $m = 1, 2, \dots, M$
 - a. Ajuste el clasificador $G_m(x_i)$ débil a los datos de entrenamiento usando los pesos w_i .
 - b. Calcule la tasa de error del paso m ,

$$\bar{e}_m = \frac{\sum_{i=1}^N w_i I(y_i \neq G(x_i))}{\sum_{i=1}^N w_i} \quad [4]$$

c. Calcule $\alpha_m = \log((1 - \bar{e}_m) / \bar{e}_m)$ [5]

d. Haga $w_{i,m} = w_{i,m-1} e^{\alpha_m I(y_i \neq G(x_i))}$ [6]

3. Calcule la predicción dada por $G(x_i) = \text{sign} \left(\sum_{m=1}^M \alpha_m G_m(x_i) \right)$ [7]

RESULTADOS

Inicialmente se evaluó que los indicadores financieros escogidos en la realización de la predicción, tuvieran mejor desempeño en las empresas no insolventes que en las compañías insolventes. En la tabla 3 se muestran las diferencias de medias entre los dos grupos de empresas estudiados.

Tabla 3.
Diferencia de medias

Variable	Media No insolventes	Media insolventes	Estadístico t	P-valor
Cic_op	0.4899275	0.5475945	-5.2818	0.00000
Raz_cte	0.5596243	0.5668736	-0.40605	0.68540
Rel_CTNO_Ventas	0.5460833	0.5451744	0.52653	0.59940
FCL_Ventas	0.4683961	0.4733454	-1.18	0.24020
ROA	0.4265986	0.4148215	8.8599	0.00000
ROE	0.5180227	0.518242	-0.060164	0.95210
Ptividad_Anoco	0.4589981	0.3477476	5.8549	0.00000
Mg_bruto	0.6807604	0.6700276	1.2113	0.22790
Mg_operacional	0.4293685	0.4252707	2.9837	0.00336
Mg_antesimpto	0.3809004	0.3751153	3.2255	0.00158
Mg_netto	0.4007059	0.3948359	3.1237	0.00219
Niv_end	0.2977816	0.3880065	-10.278	0.00000
Apa_fro	0.4056256	0.4958455	-3.6644	0.00035
Apa_total	0.5023683	0.5030102	-0.1368	0.89140
Cob_int_UO	0.5510206	0.5505573	1.5697	0.11660

Fuente: Elaboración propia.

En términos generales, se observa que las empresas no insolventes cuentan con mejor ciclo operacional que las compañías insolventes ya que su media es menor (Correa, Martínez, Ruiz, y Yepes, 2018). Adicionalmente, las empresas que están en situación de insolvencia empresarial tienen menor desempeño financiero de acuerdo con sus indicadores de rentabilidad (Correa-garcía, Gómez, y Londoño, 2018), lo cual es de esperarse para este tipo de firmas ya que una de las causas para entrar en este estado son los resultados financieros adversos. Por último, como es de esperarse de acuerdo a lo planteado por Pérez et al. (2017), las empresas insolventes cuentan con un mayor grado de endeudamiento que representa un mayor compromiso de sus recursos frente a terceros y por ende mayor probabilidad de incurrir en impagos de sus obligaciones.

En este estudio, se plantea una comparación entre la predicción a través del modelo de regresión logística, el cual es considerado un modelo de predicción tradicional, y el algoritmo boosting que hace parte de las nuevas tendencias en la predicción de la insolvencia empresarial. En la tabla 4 se presenta la matriz de confusión, la cual muestra la precisión de cada modelo para clasificar a las compañías insolventes y no insolventes. La muestra fue separada aleatoriamente en dos grupos con el fin de entrenar cada modelo y luego evaluarlo. El 80% del total de los datos fueron utilizados para el entrenamiento, lo cual representó 2.289 empresas no insolventes y 102 compañías insolventes. El 20% de los datos restantes se utilizaron para evaluar el modelo, lo cual corresponde a 572 empresas no insolventes y 25 insolventes.

Tabla 4.
Matriz de confusión

Concepto	Algoritmo boosting		Regresión logística	
	Insolventes	No insolventes	Insolvente	No insolvente
Insolventes	80.0%	14.5%	0.0%	0.5%
No insolventes	20.0%	85.5%	100.0%	99.5%
Total	100.0%	100.0%	100.0%	100.0%

Fuente: Elaboración propia.

La matriz de confusión muestra la proporción de compañías que fueron clasificadas correcta e incorrectamente. En este caso, el 80% de las empresas insolventes fueron clasificadas de manera correcta usando el algoritmo boosting, mientras que usando la regresión logística no fue posible lograr un acierto en la clasificación de este tipo de empresas. Por otra parte, el 85,5% de las empresas no insolventes fueron clasificadas correctamente utilizando el algoritmo boosting, en tanto el 99,5% de las empresas no insolventes se clasificaron de manera acorde utilizando la regresión logística.

Con el fin de evaluar los resultados obtenidos en ambos modelos y determinar cuál es más conveniente utilizar para realizar pronósticos, se analizan los errores de tipo I y tipo II que existen en ambos resultados. En la tabla 5 se presentan las probabilidades de incurrir en este tipo de errores con cada modelo.

Tabla 5.
Probabilidad de incurrir en Error tipo I y Error tipo II

Modelo	Error tipo I	Error tipo II
Algoritmo boosting	14.5%	19.7%
Regresión logística	0.5%	100.0%

Fuente: Elaboración propia.

Los tipos de errores que se muestran en la tabla 5 muestran la probabilidad de clasificar a una empresa que se encuentra en estado de insolvencia o no insolvencia en un estado contrario. El error tipo I indica que bajo el algoritmo boosting existe una probabilidad del 14.5% de clasificar una empresa que no está en estado de insolvencia de manera incorrecta, mientras que esta probabilidad es de 0.5% haciendo uso de la regresión logística. Por otra parte, se muestra que bajo el uso del algoritmo boosting se presenta una probabilidad de 19.7% de clasificar una empresa insolvente como si no tuviera esta situación, en tanto que esta probabilidad es del 100% bajo la regresión logística.

Cometer el error tipo I y tipo II tiene diferentes consecuencias financieras para las empresas que otorgan los créditos. Según lo establece Jabeur y Fahmi (2018), cometer el error tipo I significa para las compañías una reducción de sus utilidades ya que se va a presentar un rechazo al otorgamiento del crédito a clientes que tienen la capacidad de pagar las respectivas obligaciones sin incurrir en incumplimientos. En este sentido, la regresión logística muestra una superioridad frente al algoritmo boosting en la precisión para evitar cometer el error de tipo I.

Sin embargo, el algoritmo boosting (probabilidad del 19.7%) muestra una mayor precisión frente a la regresión logística (probabilidad del 100%) en la incursión del error tipo II. El alto grado de incurrir en el error tipo II utilizando la regresión logística se debe a que este tipo de regresión asume normalidad en los datos observados (Pérez et al., 2017) y como se puede evidenciar, en los estudios de insolvencia empresarial la realidad muestra que son mayores las empresas que no se encuentran en este estado.

A diferencia del error tipo I, el error tipo II trae consecuencias directas frente a la posición financiera de la empresa acreedora. De acuerdo con lo establecido por Wang, Ma, y Yang (2014), cuando un acreedor incurre en error tipo II les está otorgando créditos a clientes que tienen alta probabilidad de entrar en un estado de insolvencia ya que no tienen capacidad de pago.

Las consecuencias financieras para los acreedores que incurren en error tipo II además de tener una reducción de sus utilidades derivado del gasto deterioro que se debe reconocer de sus cuentas por cobrar, también tiene una implicación directa sobre el efectivo de la empresa toda vez que los productos vendidos a los clientes riesgosos, tienen una alta probabilidad de que no sean pagados. De esta manera, se puede determinar que es más perjudicial para las empresas incurrir en error tipo II que en el error tipo I.

CONCLUSIONES

A través de este estudio se aborda la insolvencia empresarial considerando el ámbito jurídico, financiero y estadístico con el fin de lograr una mayor precisión en la predicción de este evento empresarial que afecta no solo a la empresa insolvente sino también a sus diferentes acreedores. A través de la identificación de las empresas que se declararon en insolvencia en el año 2018 en Colombia y de acuerdo con los indicadores financieros que suministran información relevante sobre la situación financiera de las empresas, se realizó la comparación entre los resultados predictivos obtenidos mediante la regresión logística (metodología tradicional) y el algoritmo

boosting (metodología no tradicional) con el propósito de evaluar la precisión del pronóstico.

Considerando los resultados mostrados en las matrices de confusión de ambas metodologías, se puede concluir que los indicadores financieros utilizados cuentan con poder para predecir la insolvencia empresarial. Adicionalmente, el algoritmo boosting cuenta con mayor precisión que la regresión logística para predecir la insolvencia empresarial. Lo anterior se sustenta en que a partir de la regresión logística no es posible omitir el error tipo II, el cual es altamente perjudicial para la posición financiera de los acreedores.

Los resultados de esta investigación generan la posibilidad a los clientes, gobiernos, inversionistas, bancos y proveedores de bienes y servicios, de contar con información relevante que puede incidir en las decisiones de compra, acuerdos de pagos, otorgamiento de créditos o inversiones en determinada empresa. En este contexto, los grupos de interés mencionados pueden tomar decisiones financieras más precisas que les permitan reducir sus pérdidas en el futuro.

Finalmente, dentro de las limitaciones en el desarrollo de esta investigación se encuentra principalmente la disponibilidad de la información financiera de las empresas. Debido al gran número de información faltante en la base de datos de la Superintendencia de Sociedades que se requería para calcular los indicadores financieros utilizados, las empresas consideradas en el estudio, a pesar de ser un volumen importante, pudieron haber representado a una población mayor si se contara con toda la información.

Como posible línea de investigación futura se sugiere realizar estudios predictivos a través de metodologías no tradicionales haciendo la comparación entre el poder predictivo que tienen los indicadores financieros tradicionales y los inductores de valor.

REFERENCIAS BIBLIOGRAFICAS

- Altman, E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4), 589-609.
- Amendola, A., Giordano, F., Parrella, M., y Restaino, M. (2017). Variable selection in high-dimensional regression: a nonparametric procedure for business failure prediction. *Applied Stochastic Models in Business and Industry*, 33(4), 355-368. <https://doi.org/10.1002/asmb.2240>
- Barboza, F., Kimura, H., y Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83, 405-417. <https://doi.org/10.1016/j.eswa.2017.04.006>
- Beaver, W. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 4(1966), 71-111. <https://doi.org/DOI: 10.2307/2490171>
- Ben, S. (2017). Bankruptcy prediction using Partial Least Squares Logistic Regression. *Journal of Retailing and Consumer Services*, 36(November 2016), 197-202. <https://doi.org/10.1016/j.jretconser.2017.02.005>

- Bredart, X., Vella, V., y Bonello, J. (2018). Machine Learning Models for Predicting Financial Distress. *Journal of Research in Economics*, 2(2), 174–185. <https://doi.org/10.24954/JORE.2018.22>
- Calabrese, R., y Osmetti, S. A. (2013). Modelling small and medium enterprise loan defaults as rare events: The generalized extreme value regression model. *Journal of Applied Statistics*, 40(6), 1172–1188. <https://doi.org/10.1080/02664763.2013.784894>
- Calabrese, R., y Osmetti, S. A. (2015). Improving forecast of binary rare events data: A gam-based approach. *Journal of Forecasting*, 34(3), 230–239. <https://doi.org/10.1002/for.2335>
- Correa-garcía, J. A. (2005). De la partida doble al análisis financiero. *Contaduría Universidad de Antioquia*, (46), 169–194.
- Correa-garcía, J. A., Gómez, S., y Londoño, F. (2018). Indicadores financieros y su eficiencia en la explicación de la generación de valor en el sector cooperativo. *Rev.fac.cienc.econ.*, XXVI(2), 129–144.
- Correa, D., Laura, M., Camila, R., y Alejandra, Y. (2018). Los indicadores de costos: una herramienta para gestionar la generación de valor en las empresas industriales colombianas. *Estudios Gerenciales*, 34(147), 190–199. <https://doi.org/10.18046/j.estger.2018.147.2643>
- Correa, J., Pulgarín, A., Muñoz, L., y Álvarez, M. (2010). Marco normativo y antecedentes de la revelación contable en Colombia. *Contaduría Universidad de Antioquia*, (56), 269–292.
- Cultrera, L., y Brédart, X. (2016). Bankruptcy prediction: The case of Belgian SMEs. *Review of Accounting and Finance*, 15(1), 101–119. <https://doi.org/10.1108/RAF-06-2014-0059>
- Dinca, G., Baba, M. C., Dinca, M. S., Dauti, B., y Deari, F. (2017). Insolvency risk prediction using the logit and logistic models: Some evidences from Romania. *Economic Computation and Economic Cybernetics Studies and Research*, 51(4), 139–157.
- Eling, M., y Jia, R. (2018). Business failure, efficiency, and volatility: Evidence from the European insurance industry. *International Review of Financial Analysis*, 59, 58–76. <https://doi.org/10.1016/j.irfa.2018.07.007>
- Gupta, J., Gregoriou, A., y Ebrahimi, T. (2018). Empirical comparison of hazard models in predicting SMEs failure. *Quantitative Finance*, 18(3), 437–466. <https://doi.org/10.1080/14697688.2017.1307514>
- Hastie, T., Tibshirani, R., Friedman, J. (2017). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, Stanford, California.
- Hebous, S., y De Mooij, R. (2018). Curbing Corporate Debt Bias: Do Limitations to Interest Deductibility Work? *Journal of Banking y Finance*, 1–11. <https://doi.org/10.1016/j.jbankfin.2018.07.013>

- Jabeur, S. Ben, y Fahmi, Y. (2018). Forecasting financial distress for French firms: a comparative study. *Empirical Economics*, 54(3), 1173–1186. <https://doi.org/10.1007/s00181-017-1246-1>
- Jayasekera, R. (2018). Prediction of company failure: Past, present and promising directions for the future. *International Review of Financial Analysis*, 55, 196–208. <https://doi.org/10.1016/j.irfa.2017.08.009>
- Jones, S., Johnstone, D., y Wilson, R. (2017). Predicting Corporate Bankruptcy: An Evaluation of Alternative Statistical Frameworks. *Journal of Business Finance and Accounting*, 44(1–2), 3–34. <https://doi.org/10.1111/jbfa.12218>
- Jovanovik, D., Todorovic, M., y Grbic, M. (2017). Financial Indicators as Predictors of Illiquidity. *Romanian Journal of Economic Forecasting*, 20(1), 128–149.
- Kovacova, M., y Kliestik, T. (2017). Logit and Probit application for the prediction of bankruptcy in Slovak companies. *Equilibrium-Quarterly Journal of Economics and Economic Policy*, 12(4), 775–791. <https://doi.org/10.24136/eq.v12i4.40>
- Le, T., Son, L. H., Vo, M. T., Lee, M. Y., y Baik, S. W. (2018). A cluster-based boosting algorithm for bankruptcy prediction in a highly imbalanced dataset. *Symmetry*, 10(7), 1–12. <https://doi.org/10.3390/sym10070250>
- Ley 1116. (2006). Diario Oficial No. 46.494 de 27 de diciembre de 2006, Colombia, diciembre 27 de 2006.
- Lyandres, E., y Zhdanov, A. (2013). Investment opportunities and bankruptcy prediction. *Journal of Financial Markets*, 16(3), 439–476. <https://doi.org/10.1016/j.finmar.2012.10.003>
- Nyitrai, T., y Virág, M. (2018). The effects of handling outliers on the performance of bankruptcy prediction models. *Socio-Economic Planning Sciences*, 1–9. [https://doi.org/10.1016/0304-3762\(82\)90059-1](https://doi.org/10.1016/0304-3762(82)90059-1)
- Pérez, J., Lopera, M., y Vásquez, F. (2017). Estimación de la probabilidad de riesgo de quiebra en las empresas colombianas a partir de un modelo para eventos raros. *Cuadernos de Administración*, 30(54), 7–38. <https://doi.org/10.11144/Javeriana.cao30-54.eprqe>
- Rodríguez, J. (2007). *Nuevo régimen de insolvencia*. Bogotá, Colombia: Universidad Externado de Colombia.
- Romero, F., Melgarejo, Z., y Vera, M. (2015). Fracaso empresarial de las pequeñas y medianas empresas (pymes) en Colombia. *Suma de Negocios*, 6(13), 29–41. <https://doi.org/10.1016/j.sumneg.2015.08.003>
- Tian, S., y Yu, Y. (2017). Financial ratios and bankruptcy predictions: An international evidence. *International Review of Economics and Finance*, 51, 510–526. <https://doi.org/10.1016/j.iref.2017.07.025>

- Upegui, A. y Londoño, Á. (2011). *Comentarios al régimen de insolvencia empresarial*. Bogotá, Colombia: Legis Editores.
- Vélez, L. (2011). ¿Qué tan bueno es el sistema de insolvencia empresarial colombiano? *Revista Supersociedades*, 2, 5-6.
- Wang, G., Ma, J., y Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications*, 41(5), 2353–2361. <https://doi.org/10.1016/j.eswa.2013.09.033>
- Wilches, R. (2008). Vacíos e inconsistencias estructurales del nuevo régimen de insolvencia empresarial colombiano. Identificación y propuestas de solución. *Vniversitas*, (117), 197–218.
- Wilches, R. (2009). La insolvencia transfronteriza en el derecho colombiano. *Revista de Derecho*, (32), 162–198.
- Yazdanfar, D., y Öhman, P. (2015). Debt financing and firm performance : an empirical study based on Swedish data. *The Journal of Risk Finance*, 16(1), 102–118. <https://doi.org/10.1108/JRF-06-2014-0085>