

Un análisis bibliométrico de la predicción de quiebra empresarial con Machine Learning

**A Bibliometric Analysis of Business Bankruptcy
Prediction with Machine Learning**

Yuly Andrea Franco*

* Magíster en Finanzas, Universidad Externado de Colombia, Bogotá (Colombia). Docente-Investigadora, Observatorio de Economía y Operaciones Numéricas (ODEON), Universidad Externado de Colombia. Bogotá (Colombia). [yuly.franco1@uexternado.edu.co]; [ORCID ID: 0000-0003-2938-9331].

Artículo recibido: 10 de mayo de 2022

Aceptado: 27 de mayo de 2022

Para citar este artículo:

Franco, Y. A. (2022). Un análisis bibliométrico de la predicción de quiebra empresarial con Machine Learning. *Odeon*, 22, 87-126.

DOI: <https://doi.org/10.18601/17941113.n22.04>

Resumen

El objetivo de este artículo es presentar un análisis bibliométrico sobre el uso que han tenido las técnicas de Machine Learning (ML) en el proceso de predicción de quiebra empresarial a través de la revisión de la base de datos Web of Science. Este ejercicio brinda información sobre el inicio y el proceso de adaptación de dichas técnicas. Para ello, se identifican las diferentes técnicas de ML aplicadas en modelo de predicción de quiebras. Se obtiene como resultado 327 documentos, los cuales se clasifican por medida de evaluación del desempeño, área bajo la curva (AUC) y precisión (ACC), por ser las más utilizadas en el proceso de clasificación. Además, se identifica la relación entre investigadores, instituciones y países con mayor número de aplicaciones de este tipo. Los resultados evidencian que los algoritmos XGBoost, SVM, Smote, RF y DT presentan una capacidad predictiva mucho mayor que las metodologías tradicionales, enfocados en un horizonte de tiempo antes del suceso dada su mayor precisión. Así mismo, las variables financieras y no financieras contribuyen de manera favorable a dicha estimación.

Palabras clave: predicción; quiebra empresarial; Machine Learning; bibliometría.

Clasificación JEL: C53, G17, G33, C63.

Abstract

The aim of this article is to present a bibliometric analysis on the use that Machine Learning (ML) techniques have had in the process of predicting business bankruptcy through the review of the Web of Science database. This exercise provides information on the initiation and adaptation process of such techniques. For this, the different ML techniques applied in the bankruptcy prediction model are identified. As a result, 327 documents are obtained, of which they are classified by performance evaluation measure, the area under the curve (AUC) and precision (ACC), these being the most used in the classification process. In addition, the relationship between researchers, institutions and countries with the largest number of applications of this type is identified. The results show how the XGBoost, SVM, Smote, RF and D algorithms present a much greater predictive capacity than traditional methodologies; focused on a time horizon before the event given its greater precision. Similarly, financial and non-financial variables contribute favorably to said estimate.

Key words: Prediction; bankruptcy; Machine Learning; bibliometrics.

JEL classification: C53, G17, G33, C63.

Introducción

La evaluación y el análisis permanente de la sostenibilidad financiera es un factor de interés y constante evaluación para las sociedades. Su importancia reside en que permite la protección de los intereses económicos de un Estado y la preservación de la confianza, tanto de las empresas como de la economía. Su análisis y seguimiento permite minimizar los impactos causados con la liquidación o el deterioro de una empresa. Así mismo, la tendencia relativa a la insolvencia o quiebra.

La anterior perspectiva pone de manifiesto un interés particular por revisar y establecer técnicas que permitan a las empresas un mejor desempeño y un crecimiento por lo menos, en términos financieros, sostenible y que impacte en lo ambiental y en lo social, una triada que se encuentra directamente relacionada y cuyas variaciones se encuentran entrelazadas (Endrikat *et al.*, 2014).

En las últimas décadas, la predicción de quiebras ha sido un campo interés donde la mayoría de estudios de revisión se centran en modelos basados en métodos estadísticos, pero ejemplos recientes en la literatura muestran una tendencia al uso de múltiples herramientas de Machine Learning (ML). El objetivo principal ha sido identificar cómo el uso ML contribuye en la predicción de quiebras. El desempeño predictivo de los modelos de ML varía sustancialmente entre las muestras de países, lo que sugiere que la aplicabilidad de estos no se puede generalizar. La justificación de este artículo radica en su reconocimiento de que este campo de investigación se ha expandido dramáticamente en los últimos años y es importante evaluar su progreso y describir la tendencia de investigación a través del análisis y la visualización bibliométrica.

Por lo tanto, se hace relevante establecer una mirada sobre el inicio y el proceso de adaptación de dichas técnicas para su posterior aplicabilidad, uso y ventajas en el contexto de las empresas colombianas. Lo anterior permitirá, en general, establecer un punto de partida en términos de innovación empresarial y, sobre todo, de eficiencia en la sostenibilidad financiera de las empresas.

1. Quiebra empresarial

La quiebra empresarial se presenta cuando la organización mantiene en el tiempo la no capacidad de generar flujos de efectivo de su negocio, rentabilidad o cumplir con sus obligaciones a medida que vencen. Los estados financieros básicos de una empresa son los registros escritos de la situación financiera, e

incluyen informes estándar como son el Estado de situación financiera, el Estado de resultados integrales, el Estado de flujos de efectivo y el Estado de cambios en el patrimonio, todos ellos proporcionan una amplia gama de usuarios la información sobre los resultados de las operaciones de la empresa y su situación financiera, incluidos los flujos de caja del negocio (Huang y Yen, 2019).

El proceso de desarrollo de FL-Score sugiere que las variables financieras y económicas tienen un contenido informativo complementario para predecir la dificultad o la recuperación de las empresas, dichos modelos no contemplan gran cantidad de variables y clasifican de manera moderada; “tienen algunos supuestos restrictivos, tales como la normalidad y la independencia entre variables predictoras o de entrada. Teniendo en cuenta que la violación de estos supuestos [...] y las técnicas estadísticas pueden tener limitaciones para obtener la eficacia y validez” como resultado (Wang *et al.*, 2014); se propone, entonces, el uso del aprendizaje automático, como señalan Huang y Yen (2019, p. 2) “es deseable desarrollar modelos de predicción cuantitativos bastante precisos utilizando varios factores internos y externos”. Últimamente, las técnicas de aprendizaje automático están experimentando un aumento de interés y se utilizan ampliamente para construir modelos de predicción y clasificación, han demostrado superioridad sobre los métodos estadísticos tradicionales, esto gracias a que posibilitan una gran cantidad de variables, permitiendo que los métodos computacionales de última generación desarrollen modelos que evalúan la predicción de incumplimiento con gran precisión.

Autores como Heo y Yang afirman que las metodologías actuales con ML son mejores a la propuesta por Altman, así, “a partir de los resultados de la investigación [...] los algoritmos basados en el aprendizaje automático muestran una capacidad predictiva mucho mayor que la puntuación Z de Altman” (2014, p. 5) y, a diferencia de los modelos estadísticos, los métodos basados en inteligencia artificial no implican suposiciones sobre las especificidades de distribución de datos (Lahmiri y Bekiros, 2019, p. 8).

2. Aprendizaje automático

En 1955, John McCarthy propuso un proyecto de investigación en Dartmouth sobre inteligencia artificial (AI). Bajo esa propuesta, en 1956 se da paso a la comunidad estadística con el desarrollo de la modelación de algoritmos (Athey y Guido, 2019), lo que permite el paso al aprendizaje de máquinas, y nace así, como un subconjunto de la AI, el Machine Learning, acuñado por Arthur

Samuel en 1959. El aprendizaje automático o Machine Learning, como área de investigación, tiene por objetivo desarrollar métodos computacionales que sean capaces de “aprender” con la experiencia acumulada. En general, se busca establecer modelos o sistemas aptos para organizar el conocimiento existente o diseñar técnicas que permitan imitar el comportamiento humano de un experto de manera automática, y se pueda de igual manera operar gran cantidad de datos, esto debido a la creciente disponibilidad de datos a nivel de las organizaciones, en ocasiones no estructurados, pero de gran dimensión. Los algoritmos de aprendizaje automático se han empleado con éxito en el ámbito empresarial con el fin de abordar múltiples preguntas de investigación relacionadas con la dinámica de las empresas.

El Machine Learning cuenta con varias clases de algoritmos para su reconocimiento y toma de decisiones en el aprendizaje supervisado y no supervisado, con técnicas como máquinas de vectores de soporte (SVM) propuesta por Vapnik, que busca construir un hiperplano de separación en ese espacio, que maximice el margen entre los dos conjuntos de datos; algoritmos SMOTE (Synthetic Minority Oversampling Technique), enfoques conocidos en el área de preprocesamiento de datos desequilibrados, que crean puntos de datos sintéticos basados en la similitud de características del espacio considerando *k-nearest neighbors* a partir de ejemplos de minorías reales, mejorando el rendimiento de la clasificación para una clase minoritaria porque crea un mayor ámbito de decisión general; los árboles de decisión (DT), que toman la estructura de un árbol donde se dividen los datos en forma de hojas, el nodo interno corresponde a una prueba de atributo y cada rama es el resultado, las hojas representan las clases o distribuciones; Random forest (RF), creada por Leo Breiman basado en modelos de árboles de decisión, conocidos como árboles de clasificación y regresión generalizados (CART), esta genera repetidamente funciones de clasificación basadas en subconjuntos, seleccionando aleatoriamente un subconjunto de características de cada nodo del árbol; los modelos Extreme Gradient Boosting (XGBoost) apropiados para problemas de regresión, de clasificación binaria; al combinar árboles de regresión y aumento de gradientes, se resalta por evitar problemas de sobreajuste; las redes neuronales artificiales (ANN) son una estructura creada en capas con enlaces entre nodos, donde las variables de entrada determinan la primera capa del sistema y la última variable de salida (dependiente); entre otras.

Los algoritmos y las técnicas empleadas por el aprendizaje de máquinas han proporcionado avances promisorios y desempeños eficientes en términos del análisis de datos tanto para la predicción como para el reconocimiento de patrones

o los problemas de clasificación. Esto ha permitido el trabajo con conjuntos de datos de alta dimensionalidad y que en su mayoría no se encuentran sujetos a condiciones de normalidad o varianza específica, a una estructura particular o supuestos sobre los mismos.

Por ello, para encontrar técnicas que permitan prevenir una quiebra, durante las últimas décadas se han desarrollado sistemas provenientes de la estadística y de las ciencias de la computación que han permitido a las organizaciones alcanzar ciertos objetivos de manera más eficiente, como alerta temprana, que permitan identificar probabilidades de quiebra con diferentes horizontes de tiempo antes de dicho evento; de igual manera, identificar el modelo más preciso que pueda clasificar de manera correcta el mayor número de empresas para las que se activó un procedimiento de liquidación.

3. Medidas de desempeño

Las métricas de desempeño utilizadas con frecuencia para evaluar clasificadores son: precisión, exactitud, recuperación, F-measure, precisión general (ACC) y área bajo la curva (AUC); esta última es calculada a través del ROC (Receiver Operating Characteristic). Para problemas de clasificación binaria, como quiebra o no, las medidas de evaluación más utilizadas son AUC seguida por ACC.

El AUC es el área más pequeña bajo la curva (funciona mejor que otra que tiene un área más grande) y proporciona una medida agregada de rendimiento en todos los umbrales de clasificación posibles, generando una medida de rendimiento y solidez, es decir, la capacidad del modelo para clasificar correctamente los registros determinando cuál modelo predice mejor. Como recomendación de Hosmer Jr. *et al.* (2013) citado en Li *et al.* (2022), el AUC es aceptable cuando es superior a 0,7 o, como lo plantean Le *et al.* (2019), si un algoritmo tiene un AUC mayor que el de otro algoritmo, este algoritmo es mejor.

El AUC mide del equilibrio entre la tasa de verdaderos positivos¹ y la tasa de falsos positivos², calculando la capacidad de un clasificador para predecir correctamente las instancias positivas. El ACC se define como el porcentaje de instancias correctamente clasificadas.

1 Tasa de verdaderos positivos (TPR-true positive rate) es el porcentaje de instancias positivas clasificadas correctamente.

2 Tasa de falsos positivos (FRP-false positive rate) es el porcentaje de instancias negativas mal clasificadas.

4. Metodología

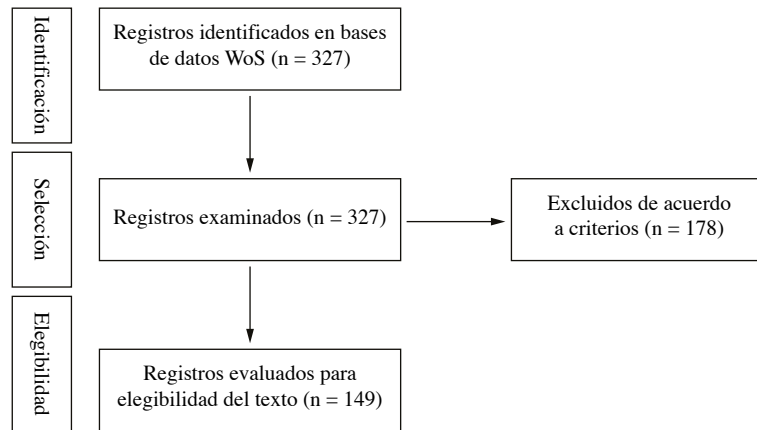
La metodología utilizada es a través del análisis bibliométrico de documentos que evidencian la relación de modelos ML y variables predictoras de quiebra, para ello se efectúa la revisión en un primer nivel de búsqueda a través de la base de datos Web of Science con las palabras clave “Machine Learning and Bankruptcy”, la cual arroja resultados desde el año 2004 hasta la fecha (2022). La muestra consta de 327 publicaciones académicas, sin la aplicación de parámetros booleanos a fin de tener una búsqueda más amplia, tomando un criterio de investigación inclusivo, es decir, no se limita a una disciplina, revistas específicas, tiempo o idioma, ni se ponen filtros en las diferentes disciplinas con el fin de obtener una visión integral de la investigación realizada en estas.

El estudio bibliométrico es un análisis cuantitativo de las variables que se pueden medir en la revisión de la literatura, como, por ejemplo, número de citas, relaciones entre autores, palabras clave, entre otros, lo que permite encontrar las tendencias en investigación de una temática en particular, y explorar y describir la literatura existente. Este se aplica en la herramienta VOSviewer con el uso de mapas bibliográficos determinando la correlación entre artículos, autores, temas, recursos e instituciones. Por medio de métodos matemáticos y estadísticos se presentan resultados cuantitativos de la comunicación escrita; es decir, proporciona información cuantitativa y objetiva sobre los resultados del proceso de investigación, a través de la búsqueda de mediciones estadísticas relacionadas con la producción y el comportamiento de las publicaciones.

En cada documento se identifica como están utilizando las técnicas de aprendizaje automático en la predicción de quiebras y su medida de desempeño. El valor agregado de esta revisión proviene de analizar sus aplicaciones sobre las técnicas de aprendizaje automático exploradas aplicadas en la predicción de quiebras que pueden ser cruciales en la gestión y la toma de decisiones para los gerentes de empresas corporativas y los responsables políticos, para posteriormente ser aplicadas en empresas de Colombia.

Después de la validación de los estudios mencionados, se realiza la selección de los trabajos que presentan sus medidas de desempeño, donde 149 son perfilados por medidas como ACC, AUC u otras. En la figura 1, se presenta el diagrama de flujo PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) para el proceso de revisión.

Figura 1: Diagrama de flujo PRISMA para la revisión sistemática de la literatura y clasificación de los artículos



Fuente: elaboración propia.

5. Resultados

Dado que en los últimos años se ha publicado un número cada vez mayor de artículos relacionados con la predicción de la quiebra empresarial, se evidencia también cómo su aumento se genera a través del uso de técnicas de aprendizaje automático, desde 2019 a la fecha las publicaciones continúan en crecimiento, como se puede ver en la tabla 1.

Tabla 1: Publicaciones de la aplicación de Machine Learning en la quiebra empresarial por año

Año	# artículos	Participación (%)
2004	2	0,61
2005	1	0,31
2006	4	1,22
2007	2	0,61
2008	7	2,14
2009	9	2,75
2010	7	2,14
2011	13	3,98

Año	# artículos	Participación (%)
2012	12	3,67
2013	5	1,53
2014	15	4,59
2015	14	4,28
2016	15	4,59
2017	12	3,67
2018	14	4,28
2019	41	12,54
2020	37	11,31
2021	64	19,57
2022	53	16,21
Total	327	100

Nota: cantidad de artículos publicados construida a partir de la revisión sistemática de literatura.

5.1 Análisis bibliométrico

Del total de documentos se obtienen 162 fuentes de publicaciones, de las cuales las 10 primeras posiciones se pueden observar en la tabla 2, donde la revista *Expert Systems with Applications* ocupa el primer lugar con un total de 47 artículos publicados en torno a la predicción de quiebras con tecnologías de ML (3,61 veces más que la segunda revista clasificada), estos cuentan con un total de 2.962 citas. El segundo lugar corresponde a la revista *Applied Soft Computing* con 13 artículos y 830 citas, seguido por la revista *European Journal of Operational Research* con 11 documentos y 1.235 citas, es decir, 405 más frente al segundo lugar. En la posición sexta y séptima se encuentran las revistas *Journal of Forecasting* y *Computational Economics* con 7 publicaciones, y citas de 110 y 104 respectivamente. En las últimas tres posiciones (8, 9 y 10) con 6 documentos, se encuentran las revistas *Knowledge-Based Systems*, *Expert Systems* y *Risks*, la primera con mayor número de citas (423), seguido por 127 y 18 respectivamente.

Las revistas con una fuerte participación de citas son *Expert Systems with Applications* y *European Journal of Operational Research* que, a pesar de no

estar en los 3 primeros lugares. Con gran influencia en la temática abordada están *Neurocomputing* y *Knowledge-Based Systems*.

Tabla 2: Top 10 de revistas por publicación y citación

	Nombre de la revista	Número de publicaciones	Citas totales
1.	Expert Systems with Applications	47	2.962
2.	Applied Soft Computing	13	830
3.	European Journal of Operational Research	11	1.235
4.	IEEE Access	9	123
5.	Neurocomputing	8	437
6.	Journal of Forecasting	7	110
7.	Computational Economics	7	104
8.	Knowledge-Based Systems	6	423
9.	Expert Systems	6	127
10.	Risks	6	18

Nota: top 10 de las revistas con mayor publicación y número de citas.

Con un filtro mínimo de 30 citaciones se construye la figura 2, cumpliendo un umbral de 94; los 10 documentos con mayor citación se muestran en la tabla 3, Min (2005) ocupa el primer lugar con un total de 508 citas, seguido por Lessmann (2015) con 443, Espejo (2010) con 352, Tsai (2008a) y Baboza (2017) con 303 y 270 citas. Los trabajos de los años 2010, 2015 y 2016 son los más frecuentes en la lista (tabla 3).

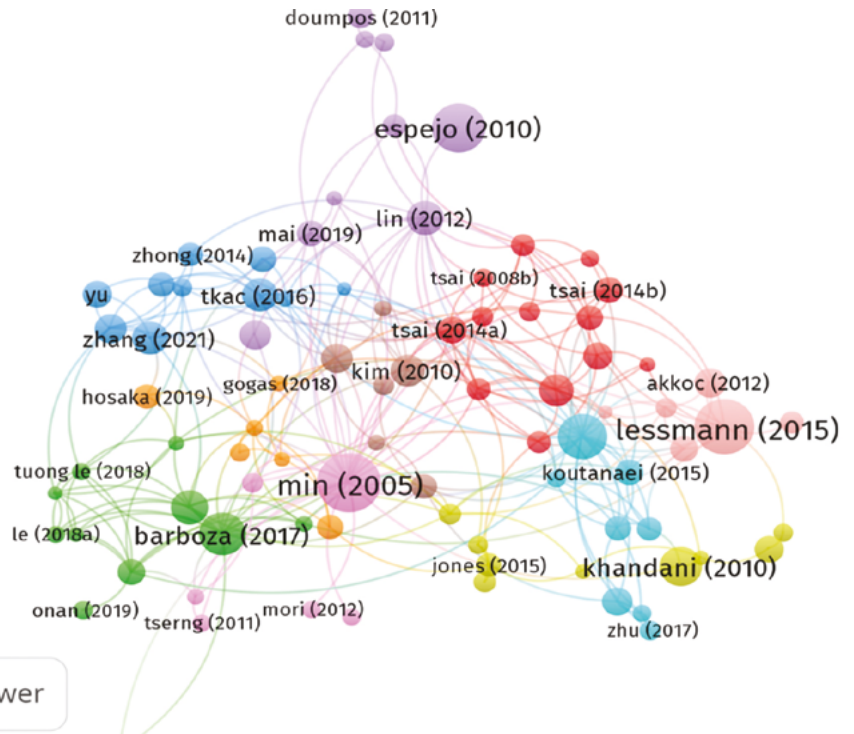
Tabla 3: Top 10 de documentos más citados

	Autor	Número de citas	Documento
1	Min (2015)	508	Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters
2	Lessmann (2015)	443	Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research
3	Espejo (2010)	352	A Survey on the Application of Genetic Programming to Classification

	Autor	Número de citas	Documento
4	Tsai (2008a)	303	Feature selection in bankruptcy prediction
5	Barboza (2017)	270	Machine learning models and bankruptcy prediction
6	Khandani (2010)	231	Consumer credit-risk models via machine-learning algorithms
7	Zieba (2016)	174	Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction
8	Lin (2012)	168	Machine Learning in Financial Crisis Prediction: A Survey
9	Zhang (2021)	165	A novel multi-stage ensemble model with enhanced outlier adaptation for credit scoring
10	Tkac (2016)	164	Artificial neural networks in business: Two decades of research

Nota: top 10 de los documentos más citados.

Figura 2. Citación de documentos



Fuente: elaboración propia con VOSviewer citación de documentos.

De igual manera cabe resaltar la cantidad de citas de los autores, donde para un total de 888 autores, con un mínimo de 5 documentos se forma una cadena de 7 autores principales, donde se resalta Tsai con 1.286 citas, seguido por Chen con 480 y Ribeiro con 209 citas. El primer autor se encuentra con una gran participación por citación en su documento “Feature selection in bankruptcy prediction”.

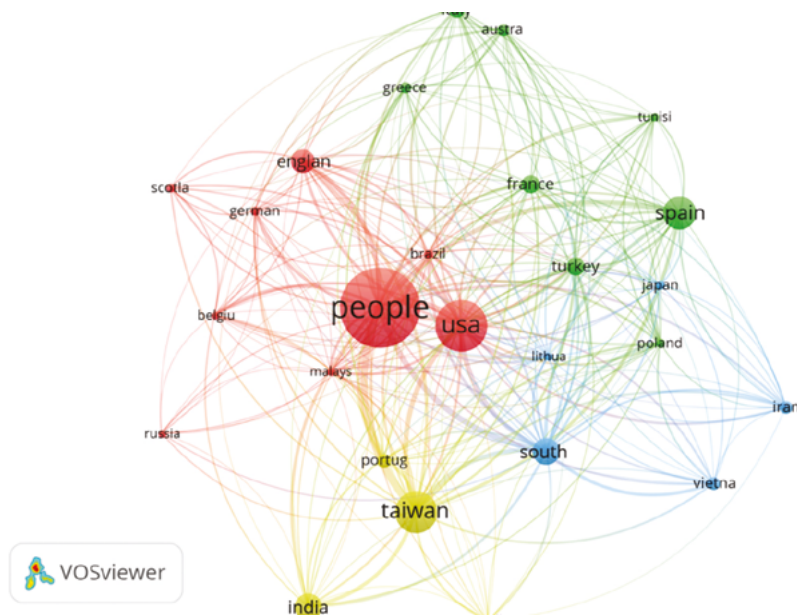
Del total de publicaciones en 493 organizaciones involucradas en la investigación científica, bajo un umbral de 5 documentos por organización, se obtienen 12 que cumplen dicha condición, sobresaliendo la National Central University y National Chung Cheng University de Taiwan, Hefei University of Technology, Chinese Academy of Sciences, Zhejiang University of Finance and Economics, Zhejiang Normal University, Wenzhou University y Southwestern University Finance and Economics ubicados en China; University Coimbra en Portugal, Universidad de Valencia en España; Korea Advanced Institute of Science & Technology en Korea del Sur y University of Sydney en Australia. Corroborando estas instituciones se encuentra la citación de documentos por países, donde China (people China) cuenta con 80 documentos y 2.054 citas, seguido por Taiwan con 36 documentos y 1.889, y Estados Unidos en tercer lugar con 47 documentos y 1.802 citas. En la tabla 4 se observan los 10 países con mayor número de citas y en la figura 3 el total de países por citación.

Tabla 4. Top 10 de documentos más citados por países

	Autor	Citaciones	Documentos
1	China	2.054	80
2	Taiwán	1.889	36
3	USA	1.802	47
4	Korea del Sur	1.171	21
5	España	863	27
6	Inglaterra	841	18
7	Turquía	478	13
8	India	357	21
9	Brasil	308	6
10	Francia	256	13

Nota: top 10 de los documentos más citados por países.

Figura 3. Total de países por citación



Fuente: elaboración propia con vosviewer citación por países.

Se construye la matriz de coocurrencia, la cual utiliza la relación conjunta de dos unidades léxicas, en una unidad superior, como una palabra, manteniendo la suposición de que los términos son mutuamente dependientes cuando su utilización es conjunta. Con el número mínimo de ocurrencias de una palabra clave establecido en 10, con 1.257 palabras clave, se alcanza un umbral de 55 grupos, donde Bankruptcy Prediction, Machine Learning, Neural-Networks, Classification, Support Vector Machines, Financial Ratios, Model, Financial Distress, Bankruptcy y Discriminant Analysis son las palabras con más coincidencia, como se ve por el tamaño de los círculos y la concentración en la figura 4.

Adicional a la anterior información se puede construir el índice de cocitación, el cual es la frecuencia con la que dos documentos son citados juntos por otros documentos, cuanto más se cita conjuntamente a dos autores o más se relacionan intelectualmente, entonces, centrados en la interrelación entre autores y coautores se puede identificar si están interconectados; de esta manera, se identifica un total de 8.543 autores y, con un filtro mínimo de 20 citas, se obtienen 118 grupos. Donde Altman con 322 citas, seguido por Tsai con 175 y Breiman con 164 cuentan con mayor cocitación.

Tabla 5. Revisión de literatura – Modelos evaluados a través de AUC

Año	Autores	Medición quiebra	AUC (%)	Mejor modelo
2012	Shian-Chang Huan, Yu-Cheng Tang, Chih-Wei Lee, Ming-Jen Chang	54 ratios financieros	97,89	KLFDA+MR-SVM
2014	Chih-Fong Tsaia, Yu-Feng Hsub, David C. Yenc	14 atributos sector australiano 20 atributo alemán 15 atributos sector japonés	87,23	DT
2015	Shakiba Khademolqorani, Ali Zeinal Hamadani, and FarimahMokhatab Rafiei	14 indicadores	84,00	DT
2015	Dong Zhao, Chunyu Huang, Yan Wei Fanhua Yu, Mingjing Wang, Huiling Chen	30 ratios financieros	92,31	KELM model
2016	Maciej Zieba, Sebastian K. Tomczak, Jakub M. Tomczak	64 variables	95,90	EXGB
2016	Aykut Ekinici, Halil Ibrahim Erdal	35 ratios financieros	91,50	RS-B-L
2016	David Johnstone	12 variables	93,12	Generalised Boosting
2017	Jun Huang, Haibo Wang, Gary Kochenberger	83 indicadores financieros	75,40	RF
2017	Zhiyong Li Jonathan Crook, Galina Andreeva	6 categorías de razones financieras	88,00	Eficiencia genérica
2017	Philippe du Jardin, David Veganzones, Eric Séverin	50 razones financieras	86,00	ELM-Boosting
2017	Wang Mingjing, Chen Huiling, Li Huai-zhong, Cai Zhennao, Zhao Xuehua, Tong Changfei, Li Jun, Xu Xin	30 variables	85,73	GWO-KELM
2017	Flavio Barboza, Herbert Kimura, Edward Altman	11 variables	92,97	Boosting
2017	Hong Hanh Le, Jean-Laurent Viviani	31 ratios financieros	81,90	ANNS_2
2018	Hafiz Alaka, Lukumon Oyedele, Hakeem Owolabi, Olugbenga Akinade, Muhammad Bilal, and Saheed Ajayi	29 indicadores financieros	92,68	Tuned ANN
2018	Tuong Le, Le Hoang Son, Minh Thanh Vo, Mi Young Lee and Sung Wook Baik	19 indicadores financieros	86,90	Cboost

Año	Autores	Medición quiebra	AUC (%)	Mejor modelo
2018	Tadaaki Hosaka	133 ratios financieros	92,00	CNN
2018	Feng Mai, Shaonan Tian, Chihoon Lee, Ling Ma	36 predictores	78,40	DL-Embedding
2018	Vicente García, Ana I. Marqués, J. Salvador Sánchez	14 conjuntos de datos financieros	98,20	RndF(C4.5)
2018	J. Uthayakumar, Noura Metawa, K. Shankar, S. K. Lakshmanaprabu.	64 variables	97,93	Propuesto
2018	Tuong Le, Mi Young Lee, Jun Ryeol Park, Sung Wook Baik	19 variables financieras	84,40	SMOTE + ENN Random Forest
2018	Pedro Carmona, Francisco Climent, Alexandre Momparker	30 variables	98,00	XGBoost
2018	Tamás Nyitrai, Virág Miklós	20 ratios financieros	83,10	Multilayer perceptron
2019	Tuong Le, Bay Vo, Hamido Fujita, Ngoc-Thanh Nguyen, Sung Wook Baik	19 índices para Corea. 10 indicadores para EE. UU. 10 indicadores para Japón	92,80	gXGBS_hist
2019	Tuong Le, Minh Thanh Vo, Bay Vo, Mi Young Lee, Sung Wook Baik	19 indicadores	87,10	Cboost-HAOC Oversampling method using SMOTE-ENN
2019	Yongming Song, Yi Peng	7 conjuntos de datos de clasificación financiera	96,19	SmoteBoost
2019	Birsen Eygi Erdogan, Süreyya Özoğür-Akyüz, Pınar Karadayı Atas.	No específica	77,98	Pruned linear kernel
2019	Georgios Manthoulis, Michalis Doumpos, Constantin Zopounidis, Emiliios Galariotis	8 indicadores	96,67	CAMELS + Diversification
2019	Martin Zoricák, Peter Gnip, Peter Drotár, Vladimír Gazda	20 atributos financieros	91,83	LSAD
2019	Makram Soui, Salima Smiti, Mohamed Wiem Mkaouer, Ridha Ejbali	24 ratios financieros Darden 26 ratios financieros Polonia	87,70	SAE
2019	Wei Xu, Yuchen Pan, Wenting Chen, Hongyong Fu	10 indicadores financieros	85,08	NIM

Año	Autores	Medición quiebra	AUC (%)	Mejor modelo
2019	Carlos Valencia, Sergio Cabrales, Laura García, Juan Ramírez, Diego Calderona	21 ratios financieros	85,76	RF
2019	Hossam Faris, Ruba Abukhurma, Waref Almanaseer, Mohammed Saadeh Antonio M. Mora, Pedro A. Castillo, Ibrahim Aljarah	15 variables financieras 21 variables no financieras	77,90	RF
2020	Yi Cao, Xiaoquan Liu, Jia Zhai, Shan Hua	31 variables	90,03	DNN
2020	Salima Smiti, Makram Soui	No especifica	98,00	BSM-SAES
2020	Said Marso, Mohamed EL Merouani	64 ratios financieros	96,25	ABCNN
2020	Gang Wanga, Jingling Ma, Gang Chen c, Ying Yang	39 variables financieras	96,09	RS2_ER
2020	Fayaz Hussain Tunion, Yi Ding, Amad Nabi Agha, Kinza Agha, Hafeez Ur Rehman Zubair Panhwar	230 razones financieras	94,30	SVM (bagging)
2020	Mark Clintworth, Dimitrios Lyridis, Evangelos Boulougouris	35 variables	88,40	XGB
2020	Xiaobo Tang, Shixuan Li, Mingliang Tan, Wenxuan Shi.	46 factores financieros	93,89	DNN
2020	Elena Gregova, Katarina Valaskova, Peter Adamko, Milos Tumpach, Jaroslav Jaros	14 ratios financieros	88,60	NN
2020	Chih-Fong Tsai	15 variables	97,10	SVM
2021	Guido Perboli, Ehsan Arabnezhad	24 variables	88,00	XGBoost
2021	Sami Ben Jabeur, Nicolae Stef, Pedro Carmona	57 ratios financieros	95,80	XGBoost
2021	Sami Ben Jabeur, Cheima Gharib, Salma Mefteh-Wali, Wissal Ben Arfi	18 variables	99,40	CatBoost
2021	Min Sue Park, Hwijae Son, Chongseok Hyun, Hyung Ju Hwang	25 categorías	98,10	LightGBM
2021	Flavio Barboza, Leonardo Fernando Cruz Basso, Herbert Kimura	20 variables	89,99	RF
2021	Nino Antulov-Fantulin, Raffaele Lagravinese, Giuliano Resce	26 indicadores financieros	98,30	Up-sampling
2021	Matthew Smith, Francisco Álvarez	17 variables	81,00	XGBoost

Año	Autores	Medición quiebra	AUC (%)	Mejor modelo
2021	Kuen-Liang Sue, Chih-Fong Tsai, Andy Chiu	101 variables China, 95 variables Taiwán	94,20	Bagging DT
2022	Umair Bin Yousaf, Khalil Jebran, Man Wang	23 indicadores financieros	100,00	RF
2022	Zhiyong Li, Chen Feng, Ying Tang	21 ratios financieros	98,30	Dynamic
2022	Yai Zou, Changchin Gao, Han Gao	47 ratios financieros	95,50	XGBoost-w
2022	Hongyi Qian, Baohui Wang, Minghe Yuan, Songfeng Gao, You Song	43 indicadores financieros	85,42	PIMP-XGBOOST
2022	Georgios Sermpinis, Serafeim Tsoukas Yiqun Zhang	No especifica	69,18	RF

Fuente: elaboración propia con la revisión de literatura, documentos clasificados con medida de evaluación AUC.

Tabla 6. Revisión de literatura – Modelos evaluados a través de AUC

Año	Autores	Medición Quiebra	ACC (%)	Mejor modelo
2004	Anja Cielen, Ludo Peeters, Koen Vanhoof	12 variables	86,40	DEA
2005	Jae H. Min, Young-Chan Lee	38 componentes principales	88,01	SVM
2006	Kin Keung Lai, Lean Yu, Wei Huang, and Shouyang Wang	No especifica	98,86	Meta modelo
2006	Po-Chang Ko, Ping-Chen Lin	39 variables	95,00	PSO-LRNN
2008	Chih-Fong Tsai	33 atributos	82,19	MLP
2009	Yi-Chung Hu	5 variables	91,54	ELECTRE-based SLP
2009	Rong-Ho Lin, Yao-Tien Wang, Chih-Hung Wu, Chun-Ling Chuang	21 atributos	83,30	HFP
2009	Hui Li, Jie Sun, Bo-Liang Sun	15 variables	91,50	OR-CBR
2010	Myoung-Jong Kim, Dae-Ki Kang	32 ratios financieros	76,47	Bagged NN
2010	Hui Li, Hai-Bin Huang, Jie Sun, Chuang Lin	30 ratios financieros	84,94	SVM
2011	Chih-Fong Tsai, Yu-Feng Hsu.	No especifica	88,41	SG-LR
2011	Ning Chena, Armando Vieira, Bernarde Ribeiro, João Duarte, João Neves	30 ratios financieros	90,70	LVQ

Año	Autores	Medición Quiebra	ACC (%)	Mejor modelo
2011	Arindam Chaudhuria, Kajal Deb.	45 variables	87,69	FSVM
2012	Bernardete Ribeiro, Catarina Silva, Ning Chen, Armando Vieira, João Carvalho das Neves.	30 ratios	89,40	SVM + MTL
2012	Wei-Yang Lin, Ya-Han Hu, and Chih-Fong Tsai	No especifica	90,14	Hybrido
2012	Mu-Yen Chen	21 variables financieros	98,02	SOM + SVR
2013	Fengyi Lin, Ching-Chiang Yeh, Meng-Yuan Lee.	16 variables	82,50	LLE+SVM
2013	Qi Yu, Yoan Miche, Eric Séverin, Amaury Lendasse	41 variables	96,87	LOO-IELM
2014	Junyoung Heo, Jin Yong Yang	11 variables	93,80	AdaBoost
2014	Gang Wang, Jian Mac, Shanlin Yang	30 variables	86,79	FS-Boosting
2014	Chih-Fong Tsai	5 conjuntos de datos	92,50	SOM + MLP
2015	Yang Lu, Nianyin Zeng, Xiaohui Liu, and Shujuan Yi	6 atributos	99,21	SPSO-SVM
2015	Xiangrong Zhang, Longying Hu	20 ratios financieros	94,13	NS-MKL
2015	Deron Liang, Chih-Fong Tsai, Hsin-Ting Wua	No especifica	92,98	CART
2016	Deron Liang, Chia-Chi Lu, Chih-Fong Tsai, Guan-An Shih	95 indicadores financieros y 95 indicadores de gobierno corporativo	79,10	SVM
2017	Deron Liang, Chih-Fong Tsai, An-Jie Dai, William Eberle	95 atributos empresas Taiwán 45 atributos empresas China 14 atributos empresas Australia 24 atributos empresas Alemania	93,00	CART
2017	D. Fernández-Arias, M. López-Martín, T. Montero-Romero, F. Martínez-Estudillo, F. Fernández-Navarro	26 variables	80,05	K-ELM
2017	Francisco Antunes Bernardete Ribeiro Francisco Pereira	30 variables	97,71	GP

Año	Autores	Medición Quiebra	ACC (%)	Mejor modelo
2018	Sinelnikova-Muryleva, Elena, Gorshkova, Taisija, Makeeva, Natalja	No especifica	96,62	RF
2018	Uthayakumar J, Noura etawa, K. Shankar, S. K. Lakshmanaprabu	No especifica	91,51	ACO-FCP
2019	Yu-Pei Huang, Meng-Feng Yen	16 variables	90,60	XGBoost
2019	Salim Lahmiri, Stelios Bekiros	6 atributos	99,96	GRNN
2019	H. Son, C. Hyun, D. Phan, H.J. Hwang	9 ratios financieros	88,00	XGBoost
2019	C. I. Johnpaul, V. N. K. Munaga, S. Prasad, G. R. Nickolas, Gangadharan	15 atributos	87,81	RF
2019	Wang Ping, Feng Wang, Aihua Wang, Yuncheng Huang	11 variables	86,25	ELM-Log
2020	Sen Zeng, L. Yaqin, Wanjun Yang, and Yanru Li	161 variables	81,94	GSPCA-SVM
2020	Abdollah Ansari, Ibrahim Said Ahmad, Azuraliza Abu Bakar, And Mohd Ridzwan Yaakub	66 atributos	99,73	PSO-MOA
2020	Santosh Kumar Shrivastav, P. Janaki Ramudu	16 variables	92,86	SVMLK
2020	Talha Mahboob Alam, Kamran Shaukat, Mubbashar Mushtaq	64 indicadores financieros	99,00	bosque de decisión-SMOTE
2020	P. K. Viswanathan, Suresh Srinivasan, N. Hariharan	No especifica	95,93	RF
2021	Anwer Mustafa Hilal, Hadeel Alsolai, Fahd N. Al-Wesabi, Mohammed Abdullah Al-Hagery Manar Ahmed Hamza, Mesfer Al Duhayyim	No especifica	98,83	CGOA
2021	H. Aljawazneh, M. Mora, P. García-Sánchez, P. A, Castillo-Valdivieso,	37 variables financieras y no financieras	94,14	Smote-ENN
2021	Hyeongjun Kim, Hoon Cho, Doojin Ryu	9 ratios	99,36	LSTM
2021	Nurul Alam, Junbin Gao, Stewart Jones	36 variables	93,71	Deep Learning
2021	Yong Ma, Hao Liu, Guangyu Zhai, Zongjie Huo	20 índices financieros	79,29	SAPSO-FSVM
2021	Nabeel Al-Milli, Amjad Hudaib, Nadim Obeid	65 variables	96,66	Enhanced GA

Año	Autores	Medición Quiebra	ACC (%)	Mejor modelo
2022	Ying Chen, Jifeng Guo, Junqin Huang, Bin Lin	13 variables distribuidas en 9 financieras y 4 no financieras	87,30	Sparse neural networks
2022	Shekar Shetty, Mohamed Musa, Xavier Brédart	No especifica	85,00	SVM
2022	Mohamed Elhoseny, Noura Metawa, Gabbor Sztano Ibrahim M. El-hasnony	14 atributos financieros Australia, 64 atributos financieros Polonia, 96 atributos financieros Taiwán	98,23	AWOA-DL
2022	Bambang Siswoyo, Zuraida Abal Abas, Ahmad Naim Che Pee, Rita Komalasari, Nano Suyatna	5 variables	97,00	BELM

Fuente: elaboración propia con la revisión de literatura, documentos clasificados con medida de evaluación ACC.

Las variables tanto financieras (liquidez, rentabilidad, gestión de efectivo, estructura de capital, entre otras) como no financieras (estructura de gobierno, divulgación de información, protección al inversor, entre otras) contribuyen de manera favorable a la predicción de los modelos. De los documentos analizados 92 detallan la cantidad y cuáles variables intervinieron en el proceso de predicción, donde se utilizaban como mínimo 5 y un máximo de 230; la cantidad de variables más utilizadas fueron 30, como se observa en la tabla 7.

Tabla 7: Análisis descriptivo de variables

Variables	
Media	33,87
Error típico	3,48
Mediana	24,50
Moda	30,00
Desviación estándar	33,42
Varianza de la muestra	1.116,66
Curtosis	14,69
Coefficiente de asimetría	3,31
Rango	225,00

Variables	
Mínimo	5,00
Máximo	230,00
Suma	3.115,83
Cuenta	92,00

Fuente: elaboración propia con análisis descriptivo de los documentos clasificados con medida de evaluación.

6. Conclusiones

El estudio de quiebras ha evidenciado un constante crecimiento. Dada su importancia y la necesidad del gran manejo de datos e información, se propone el uso de técnicas de Machine Learning; su incremento e influencia en dicho suceso se evidencia en las publicaciones por año, que eran de un dígito, iniciando con aumento en 2011, concentrando la mayor participación en los años de 2019 a 2022. Autores como Altman, Tsai y Breiman presentan una gran relación de apoyo en la investigación de los aportes de técnicas de ML en la predicción de quiebra, esto por la frecuencia con la que sus documentos son citados de manera conjunta en otros documentos.

De los documentos clasificados por medidas de desempeño 104 utilizan medidas como AUC y ACC, 54 y 50 respectivamente; se resalta que los modelos con mejores predicciones son los algoritmos de XGBoost, SVM, Smote, RF y DT, lo que confirma que los algoritmos basados en el aprendizaje automático muestran una capacidad predictiva mucho mayor a las metodologías tradicionales; se resalta que el horizonte de tiempo que genera una mayor precisión en la predicción de quiebra es un año antes del suceso, es decir, mientras más se aleja el evento su precisión disminuye, pero aun así la clasificación de quiebra o no quiebra es correcta.

Los autores confirman que el uso de las variables tanto financieras (liquidez, rentabilidad, gestión de efectivo, estructura de capital, entre otras) como no financieras (estructura de gobierno, divulgación de información, protección al inversor, entre otras) contribuye de manera favorable a la predicción. Adicionalmente, los estudios cuentan con una concentración en países como China, Taiwan y Estados Unidos, esto soportado por instituciones en torno a la investigación como: la National Central University y National Chung Cheng University de Taiwan y Hefei University of Technology, Chinese Academy of

Sciences, Zhejiang University of Finance and Economics, Zhejiang Normal University, Wenzhou University y Southwestern University Finance and Economics ubicados en China.

En coincidencia, los documentos utilizan palabras claves como “Bankruptcy prediction”, esta con mayor ocurrencia y fuerza de enlace total, y otras con una alta incidencia como son “machine learning”, “Bankruptcy”, “Financial Distress”, “Classification”, “Support Vector Machines” y “Financial Ratios”; elementos esenciales para la construcción de modelos predictores.

Actualmente, en Colombia no se cuenta con un estudio actualizado que permita identificar usos y ventajas de las técnicas de Machine Learning en el estudio de la sostenibilidad financiera, se propone para trabajos futuros poder determinar los efectos de emplear técnicas y algoritmos de ML en la determinación del desempeño empresarial en empresas colombianas.

Referencias

- Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayi, S. O., Akinade, O. O. y Bilal, M. (2018). Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Systems with Applications*, 94, 164-184. <https://doi.org/10.1016/j.eswa.2017.10.040>
- Alaka, H., Oyedele, L., Owolabi, H., Akinade, O., Bilal, M. y Ajayi, S. (2019). A Big Data Analytics Approach for Construction Firms Failure Prediction Models. *IEEE Transactions on Engineering Management*, 66(4), 689-698. <https://doi.org/10.1109/TEM.2018.2856376>
- Alam, N., Gao, J. y Jones, S. (2021). Corporate failure prediction: An evaluation of deep learning vs discrete hazard models. *Journal of International Financial Markets, Institutions and Money*, 75(266), 101455. <https://doi.org/10.1016/j.intfin.2021.101455>
- Alam, T. M., Shaukat, K., Mushtaq, M., Ali, Y., Khushi, M., Luo, S. y Wahab, A. (2021). Corporate Bankruptcy Prediction: An Approach towards Better Corporate World. *Computer Journal*, 64(11), 1731-1746. <https://doi.org/10.1093/comjnl/bxaa056>
- Aljawazneh, H., Mora, A. M., Garcia-Sanchez, P. y Castillo-Valdivieso, P. A. (2021). Comparing the performance of deep learning methods to predict companies' financial failure. *IEEE Access*, 9, 97010-97038. <https://doi.org/10.1109/ACCESS.2021.3093461>

- Al-Milli, N., Hudaib, A. y Obeid, N. (2021). Population diversity control of genetic algorithm using a novel injection method for bankruptcy prediction problem. *Mathematics*, 9(8), 1-18. <https://doi.org/10.3390/math9080823>
- Ansari, A., Ahmad, I. S., Bakar, A. A. y Yaakub, M. R. (2020). A hybrid metaheuristic method in training artificial neural network for bankruptcy prediction. *IEEE Access*, 8, 176640-176650. <https://doi.org/10.1109/ACCESS.2020.3026529>
- Antulov-Fantulin, N., Lagravinese, R. y Resce, G. (2021). Predicting bankruptcy of local government: A machine learning approach. *Journal of Economic Behavior and Organization*, 183, 681-699. <https://doi.org/10.1016/j.jebo.2021.01.014>
- Antunes, F., Ribeiro, B. y Pereira, F. (2017). Probabilistic modeling and visualization for bankruptcy prediction. *Applied Soft Computing Journal*, 60, 831-843. <https://doi.org/10.1016/j.asoc.2017.06.043>
- Appiahene, P., Missah, Y. M. y Najim, U. (2019). Evaluation of information technology impact on bank's performance: The Ghanaian experience. *International Journal of Engineering Business Management*, 11, 1-10. <https://doi.org/10.1177/1847979019835337>
- Barboza, F., Basso, L. F. C. y Kimura, H. (2021). New metrics and approaches for predicting bankruptcy. *Communications in Statistics: Simulation and Computation*, 0(0), 1-18. <https://doi.org/10.1080/03610918.2021.1910837>
- 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>
- Ben Jabeur, S., Stef, N. y Carmona, P. (2022). Bankruptcy Prediction using the XG-Boost Algorithm and Variable Importance Feature Engineering. *Computational Economics*. <https://doi.org/10.1007/s10614-021-10227-1>
- Bolaños Diaz , R. y Calderon Cahua, M. (2014). Introducción al meta-análisis tradicional. *Revista de Gastroenterología del Perú*, 34(1), 45-51.
- Botella, J. y Zamora, Á. (2017). El meta-análisis: una metodología para la investigación en educación. *Educación XXI*, 20(2), 17-38.

- Boyacioglu, M. A., Kara, Y. y Baykan, Ö. K. (2009). Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey. *Expert Systems with Applications*, 36(2 PART 2), 3355-3366. <https://doi.org/10.1016/j.eswa.2008.01.003>
- Bragoli, D., Ferretti, C., Ganugi, P., Marseguerra, G., Mezzogori, D. y Zammori, F. (2022). Machine-learning models for bankruptcy prediction: do industrial variables matter? *Spatial Economic Analysis*, 17, 156-177. <https://doi.org/10.1080/17421772.2021.1977377>
- Breiman, L. (2001). Machine Learning. *Random Forests*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>
- Cao, Y., Liu, X., Zhai, J. y Hua, S. (2022). A two-stage Bayesian network model for corporate bankruptcy prediction. *International Journal of Finance and Economics*, 27(1), 455-472. <https://doi.org/10.1002/ijfe.2162>
- Carmona, P., Climent, F. y Momparler, A. (2019). Predicting failure in the U.S. banking sector: An extreme gradient boosting approach. *International Review of Economics and Finance*, 61, 304-323. <https://doi.org/10.1016/j.iref.2018.03.008>
- Carmona, P., Dwekat, A. y Mardawi, Z. (2022). No more black boxes! Explaining the predictions of a machine learning XGBoost classifier algorithm in business failure. *Research in International Business and Finance*, 61, 101649. <https://doi.org/10.1016/j.ribaf.2022.101649>
- Chaudhuri, A. y De, K. (2011). Fuzzy Support Vector Machine for bankruptcy prediction. *Applied Soft Computing Journal*, 11(2), 2472-2486. <https://doi.org/10.1016/j.asoc.2010.10.003>
- Chen, M. Y. (2012). Visualization and dynamic evaluation model of corporate financial structure with self-organizing map and support vector regression. *Applied Soft Computing Journal*, 12(8), 2274-2288. <https://doi.org/10.1016/j.asoc.2012.03.046>
- Chen, N., Ribeiro, B., Vieira, A. S., Duarte, J. y Neves, J. C. (2011). A genetic algorithm-based approach to cost-sensitive bankruptcy prediction. *Expert Systems with Applications*, 38(10), 12939-12945. <https://doi.org/10.1016/j.eswa.2011.04.090>

- Chen, N., Vieira, A., Ribeiro, B., Duarte, J. y Neves, J. (2011). A stable credit rating model based on learning vector quantization. *Intelligent Data Analysis*, 15(2), 237-250. <https://doi.org/10.3233/IDA-2010-0465>
- Chen, S., Härdle, W. K. y Moro, R. A. (2011). Modeling default risk with support vector machines. *Quantitative Finance*, 11(1), 135-154. <https://doi.org/10.1080/14697680903410015>
- Chen, Y., Guo, J., Huang, J. y Lin, B. (2022). A novel method for financial distress prediction based on sparse neural networks with L1 / 2 regularization. *International Journal of Machine Learning and Cybernetics*, 13(7), 2089-2103. <https://doi.org/10.1007/s13042-022-01566-y>
- Chen, Z., Chen, W. y Shi, Y. (2020). Ensemble learning with label proportions for bankruptcy prediction. *Expert Systems with Applications*, 146, 113155. <https://doi.org/10.1016/j.eswa.2019.113155>
- Cheng, C., Jones, S. y Moser, W. J. (2018). Abnormal trading behavior of specific types of shareholders before US firm bankruptcy and its implications for firm bankruptcy prediction. *Journal of Business Finance and Accounting*, 45(9-10), 1100-1138. <https://doi.org/10.1111/jbfa.12338>
- Cho, S., Vasarhelyi, M. A., Sun, T. y Zhang, C. (2020). Learning from machine learning in accounting and assurance. *Journal of Emerging Technologies in Accounting*, 17(1), 1-10. <https://doi.org/10.2308/jeta-10718>
- Choi, H., Son, H. y Kim, C. (2018). Predicting financial distress of contractors in the construction industry using ensemble learning. *Expert Systems with Applications*, 110, 1-10. <https://doi.org/10.1016/j.eswa.2018.05.026>
- Cielen, A., Peeters, L. y Vanhoof, K. (2004). Bankruptcy prediction using a data envelopment analysis. *European Journal of Operational Research*, 154(2), 526-532. [https://doi.org/10.1016/S0377-2217\(03\)00186-3](https://doi.org/10.1016/S0377-2217(03)00186-3)
- Climent, F., Momparler, A. y Carmona, P. (2019). Anticipating bank distress in the Eurozone: An Extreme Gradient Boosting approach. *Journal of Business Research*, 101(June 2018), 885-896. <https://doi.org/10.1016/j.jbusres.2018.11.015>

- Clintworth, M., Lyridis, D. y Boulougouris, E. (2021). Financial risk assessment in shipping: A holistic machine learning based methodology. In *Maritime Economics and Logistics*. <https://doi.org/10.1057/s41278-020-00183-2>
- Cortés, E. A., Martínez, M. G. y Rubio, N. G. (2008). Linear discriminant analysis versus adaboost for failure forecasting. *Revista Española de Financiación y Contabilidad*, 37(137), 13-32. <https://doi.org/10.1080/02102412.2008.10779637>
- Danenas, P. y Garsva, G. (2009). Support vector machines and their application in credit risk evaluation process. *Transformations in Business and Economics*, 8(3 SUPPL. B), 46-58.
- Ding, K., Peng, X. y Wang, Y. (2019). A machine learning-based peer selection method with financial ratios. *Accounting Horizons*, 33(3), 75-87. <https://doi.org/10.2308/acch-52454>
- Ding, Y. y Simonoff, J. S. (2010). An investigation of missing data methods for classification trees applied to binary response data. *Journal of Machine Learning Research*, 11, 131-170.
- Drotár, P., Gnip, P., Zoričák, M. y Gazda, V. (2019). Small- and medium-enterprises bankruptcy dataset. *Data in Brief*, 25. <https://doi.org/10.1016/j.dib.2019.104360>
- du Jardin, P., Véganzones, D. y Séverin, E. (2019). Forecasting Corporate Bankruptcy Using Accrual-Based Models. *Computational Economics*, 54(1), 7-43. <https://doi.org/10.1007/s10614-017-9681-9>
- Ekinci, A. y Erdal, H. İ. (2017). Forecasting Bank Failure: Base Learners, Ensembles and Hybrid Ensembles. *Computational Economics*, 49(4), 677-686. <https://doi.org/10.1007/s10614-016-9623-y>
- Elhoseny, M., Metawa, N., Sztano, G. y El-hasnony, I. M. (2022). Deep Learning-Based Model for Financial Distress Prediction. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-022-04766-5>
- Endrikat, J., Guenther, E. y Hoppe, H. (2014). Making sense of conflicting empirical findings: A meta-analytic review of the relationship between corporate environmental and financial performance. *European Management Journal*, 735-751.

- Erdogan, B. E. (2013). Prediction of bankruptcy using support vector machines: An application to bank bankruptcy. *Journal of Statistical Computation and Simulation*, 83(8), 1543-1555. <https://doi.org/10.1080/00949655.2012.666550>
- Eygi Erdogan, B., Özögür-Akyüz, S. y Karadayı Ataş, P. (2021). A novel approach for panel data: An ensemble of weighted functional margin SVM models. *Information Sciences*, 557(xxxx), 373-381. <https://doi.org/10.1016/j.ins.2019.02.045>
- Faris, H., Abukhurma, R., Almanaseer, W., Saadeh, M., Mora, A. M., Castillo, P. A. y Aljarah, I. (2020). Improving financial bankruptcy prediction in a highly imbalanced class distribution using oversampling and ensemble learning: a case from the Spanish market. *Progress in Artificial Intelligence*, 9(1), 31-53. <https://doi.org/10.1007/s13748-019-00197-9>
- Farrokhi, A., Shirazi, F., Hajli, N. y Tajvidi, M. (2020). Using artificial intelligence to detect crisis related to events: Decision making in B2B by artificial intelligence. *Industrial Marketing Management*, 91(February), 257-273. <https://doi.org/10.1016/j.indmarman.2020.09.015>
- Fernández-Arias, D., López-Martín, M., Montero-Romero, T., Martínez-Estudillo, F. y Fernández-Navarro, F. (2018). Financial Soundness Prediction Using a Multi-classification Model: Evidence from Current Financial Crisis in OECD Banks. *Computational Economics*, 52(1), 275-297. <https://doi.org/10.1007/s10614-017-9676-6>
- Figlioli, B. y Lima, F. G. (2022). A proposed corporate distress and recovery prediction score based on financial and economic components. *Expert Systems with Applications*, 197, 116726. <https://doi.org/10.1016/j.eswa.2022.116726>
- García, V., Marqués, A. I. y Sánchez, J. S. (2019). Exploring the synergetic effects of sample types on the performance of ensembles for credit risk and corporate bankruptcy prediction. *Information Fusion*, 47, 88-101. <https://doi.org/10.1016/j.inffus.2018.07.004>
- Gogas, P., Papadimitriou, T. y Agrapetidou, A. (2018). Forecasting bank failures and stress testing: A machine learning approach. *International Journal of Forecasting*, 34(3), 440-455. <https://doi.org/10.1016/j.ijforecast.2018.01.009>
- Gregova, E., Valaskova, K., Adamko, P., Tumpach, M. y Jaros, J. (2020). Predicting financial distress of slovak enterprises: Comparison of selected traditional and

- learning algorithms methods. *Sustainability (Switzerland)*, 12(10). <https://doi.org/10.3390/SU12103954>
- Guerra, P. y Castelli, M. (2021). Machine learning applied to banking supervision a literature review. *Risks*, 9(7), 1-24. <https://doi.org/10.3390/risks9070136>
- Härdle, W., Lee, Y. J., Schäfer, D. y Yeh, Y. R. (2009). Variable selection and oversampling in the use of smooth support vector machines for predicting the default risk of companies. *Journal of Forecasting*, 28(6), 512-534. <https://doi.org/10.1002/for.1109>
- Heo, J. y Yang, J. Y. (2014). AdaBoost based bankruptcy forecasting of Korean construction companies. *Applied Soft Computing Journal*, 24, 494-499. <https://doi.org/10.1016/j.asoc.2014.08.009>
- Hilal, A. M., Alsolai, H., Al-Wesabi, F. N., Al-Hagery, M. A., Hamza, M. A. y Duha-yyim, M. Al. (2022). Artificial intelligence based optimal functional link neural network for financial data Science. *Computers, Materials and Continua*, 70(3), 6289-6304. <https://doi.org/10.32604/cmc.2022.021522>
- Hosaka, T. (2019). Bankruptcy prediction using imaged financial ratios and convolutional neural networks. *Expert Systems with Applications*, 117, 287-299. <https://doi.org/10.1016/j.eswa.2018.09.039>
- Hu, Y. C. (2009). Bankruptcy prediction using ELECTRE-based single-layer perceptron. *Neurocomputing*, 72(13-15), 3150-3157. <https://doi.org/10.1016/j.neucom.2009.03.002>
- Huang, J., Wang, H. y Kochenberger, G. (2017). Distressed Chinese firm prediction with discretized data. *Management Decision*, 55(5), 786-807. <https://doi.org/10.1108/MD-08-2016-0546>
- Huang, S. C., Tang, Y. C., Lee, C. W. y Chang, M. J. (2012). Kernel local Fisher discriminant analysis-based manifold-regularized svm model for financial distress predictions. *Expert Systems with Applications*, 39(3), 3855-3861. <https://doi.org/10.1016/j.eswa.2011.09.095>
- Huang, Y. P. y Yen, M. F. (2019). A new perspective of performance comparison among machine learning algorithms for financial distress prediction. *Applied Soft Computing Journal*, 83, 105663. <https://doi.org/10.1016/j.asoc.2019.105663>

- J, U., Metawa, N., Shankar, K. y Lakshmanaprabu, S. K. (2020). Financial crisis prediction model using ant colony optimization. *International Journal of Information Management*, 50(December), 538-556. <https://doi.org/10.1016/j.ijinfomgt.2018.12.001>
- Jabeur, S., Gharib, C., Mefteh-Wali, S. y Arfi, W. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, 166, 120658. <https://doi.org/10.1016/j.techfore.2021.120658>
- Jandaghi, G., Saranj, A., Rajaei, R., Ghasemi, A. y Tehrani, R. (2021). Identification of the Most Critical Factors in Bankruptcy Prediction and Credit Classification of Companies. *Iranian Journal of Management Studies*, 14(4), 817-834. <https://doi.org/10.22059/IJMS.2021.285398.673712>
- Jiang, C. Q., Liang, K., Chen, H. y Ding, Y. (2014). Analyzing market performance via social media: A case study of a banking industry crisis. *Science China Information Sciences*, 57(5), 1-18. <https://doi.org/10.1007/s11432-013-4860-3>
- Jindal, N. (2020). The Impact of Advertising and R&D on Bankruptcy Survival: A Double-Edged Sword. *Journal of Marketing*, 84(5), 22-40. <https://doi.org/10.1177/0022242920936205>
- 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>
- Karminsky, A. M. y Burekhin, R. N. (2019). Comparative analysis of methods for forecasting bankruptcies of Russian construction companies. *Business Informatics*, 13(3), 52-66. <https://doi.org/10.17323/1998-0663.2019.3.52.66>
- Khademolqorani, S., Zeinal Hamadani, A. y Mokhatab Rafiei, F. (2015). A hybrid analysis approach to improve financial distress forecasting: Empirical evidence from Iran. *Mathematical Problems in Engineering*, 2015. <https://doi.org/10.1155/2015/178197>
- Kim, H., Cho, H. y Ryu, D. (2020). Corporate default predictions using machine learning: Literature review. *Sustainability (Switzerland)*, 12(16), 1-11. <https://doi.org/10.3390/su12166325>

- Kim, H., Cho, H. y Ryu, D. (2022). Corporate Bankruptcy Prediction Using Machine Learning Methodologies with a Focus on Sequential Data. *Computational Economics*, 59(3), 1231-1249. <https://doi.org/10.1007/s10614-021-10126-5>
- Kim, M. J. y Kang, D. K. (2012). Classifiers selection in ensembles using genetic algorithms for bankruptcy prediction. *Expert Systems with Applications*, 39(10), 9308-9314. <https://doi.org/10.1016/j.eswa.2012.02.072>
- Kim, M. J. y Kang, D. K. (2010). Ensemble with neural networks for bankruptcy prediction. *Expert Systems with Applications*, 37(4), 3373-3379. <https://doi.org/10.1016/j.eswa.2009.10.012>
- Kim, M. J., Kang, D. K. y Kim, H. B. (2015). Geometric mean based boosting algorithm with over-sampling to resolve data imbalance problem for bankruptcy prediction. *Expert Systems with Applications*, 42(3), 1074-1082. <https://doi.org/10.1016/j.eswa.2014.08.025>
- Ko, P. C. y Lin, P. C. (2006). An evolution-based approach with modularized evaluations to forecast financial distress. *Knowledge-Based Systems*, 19(1), 84-91. <https://doi.org/10.1016/j.knsys.2005.11.006>
- Kristóf, T. y Virág, M. (2020). A Comprehensive Review of Corporate Bankruptcy Prediction in Hungary. *Journal of Risk and Financial Management*, 13(2), 35. <https://doi.org/10.3390/jrfm13020035>
- Lahmiri, S. y Bekiros, S. (2019). Can machine learning approaches predict corporate bankruptcy? Evidence from a qualitative experimental design. *Quantitative Finance*, 19(9), 1569-1577. <https://doi.org/10.1080/14697688.2019.1588468>
- Lai, K. K., Yu, L., Huang, W. y Wang, S. (2006). A novel support vector machine meta-model for business risk identification. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4099 LNAI(71433001), 980-984. https://doi.org/10.1007/11801603_118
- Le, H. H. y Viviani, J. L. (2018). Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios. *Research in International Business and Finance*, 44(June), 16-25. <https://doi.org/10.1016/j.ribaf.2017.07.104>

- Le, T., Lee, M. Y., Park, J. R. y Baik, S. W. (2018). Oversampling techniques for bankruptcy prediction: Novel features from a transaction dataset. *Symmetry*, *10*(4). <https://doi.org/10.3390/sym10040079>
- 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>
- Le, T., Vo, B., Fujita, H., Nguyen, N. T. y Baik, S. W. (2019). A fast and accurate approach for bankruptcy forecasting using squared logistics loss with GPU-based extreme gradient boosting. *Information Sciences*, *494*, 294-310. <https://doi.org/10.1016/j.ins.2019.04.060>
- Le, T., Vo, M. T., Vo, B., Lee, M. Y. y Baik, S. W. (2019). A Hybrid Approach Using Oversampling Technique and Cost-Sensitive Learning for Bankruptcy Prediction. *Complexity*. <https://doi.org/10.1155/2019/8460934>
- Li, H., Huang, H. Bin, Sun, J. y Lin, C. (2010). On sensitivity of case-based reasoning to optimal feature subsets in business failure prediction. *Expert Systems with Applications*, *37*(7), 4811-4821. <https://doi.org/10.1016/j.eswa.2009.12.034>
- Li, H., Sun, J. y Sun, B. L. (2009). Financial distress prediction based on OR-CBR in the principle of k-nearest neighbors. *Expert Systems with Applications*, *36*(1), 643-659. <https://doi.org/10.1016/j.eswa.2007.09.038>
- Li, X., Wang, F. y Chen, X. (2015). Support vector machine ensemble based on choquet integral for financial distress prediction. *International Journal of Pattern Recognition and Artificial Intelligence*, *29*(4). <https://doi.org/10.1142/S0218001415500160>
- Li, Z., Crook, J. y Andreeva, G. (2017). Dynamic prediction of financial distress using Malmquist DEA. *Expert Systems with Applications*, *80*, 94-106. <https://doi.org/10.1016/j.eswa.2017.03.017>
- Li, Z., Feng, C. y Tang, Y. (2022). Bank efficiency and failure prediction: a nonparametric and dynamic model based on data envelopment analysis. *Annals of Operations Research*, *315*(1), 279-315. <https://doi.org/10.1007/s10479-022-04597-4>
- Liang, D., Lu, C. C., Tsai, C. F. y Shih, G. A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European*

- Journal of Operational Research*, 252(2), 561-572. <https://doi.org/10.1016/j.ejor.2016.01.012>
- Liang, D., Tsai, C. F., Dai, A. J. y Eberle, W. (2018). A novel classifier ensemble approach for financial distress prediction. *Knowledge and Information Systems*, 54(2), 437-462. <https://doi.org/10.1007/s10115-017-1061-1>
- Liang, D., Tsai, C. F. y Wu, H. T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems*, 73(1), 289-297. <https://doi.org/10.1016/j.knosys.2014.10.010>
- Lin, F., Yeh, C. C. y Lee, M. Y. (2013). A hybrid business failure prediction model using locally linear embedding and support vector machines. *Romanian Journal of Economic Forecasting*, 16(1), 82-97.
- Lin, F., Yeh, C. C. y Lee, M. Y. (2011). The use of hybrid manifold learning and support vector machines in the prediction of business failure. *Knowledge-Based Systems*, 24(1), 95-101. <https://doi.org/10.1016/j.knosys.2010.07.009>
- Lin, R. H., Wang, Y. T., Wu, C. H. y Chuang, C. L. (2009). Developing a business failure prediction model via RST, GRA and CBR. *Expert Systems with Applications*, 36(2 Part 1), 1593-1600. <https://doi.org/10.1016/j.eswa.2007.11.068>
- Lin, W. C., Lu, Y. H. y Tsai, C. F. (2019). Feature selection in single and ensemble learning-based bankruptcy prediction models. *Expert Systems*, 36(1), 1-8. <https://doi.org/10.1111/exsy.12335>
- Lin, W. Y., Hu, Y. H. y Tsai, C. F. (2012). Machine learning in financial crisis prediction: A survey. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 42(4), 421-436. <https://doi.org/10.1109/TSMCC.2011.2170420>
- Liu, L. X., Liu, S. y Sathye, M. (2021). Predicting Bank Failures: A Synthesis of Literature and Directions for Future Research. *Journal of Risk and Financial Management*, 14(10), 474. <https://doi.org/10.3390/jrfm14100474>
- Lu, Y., Zeng, N., Liu, X. y Yi, S. (2015). A New Hybrid Algorithm for Bankruptcy Prediction Using Switching Particle Swarm Optimization and Support Vector Machines. *Discrete Dynamics in Nature and Society*, 2015. <https://doi.org/10.1155/2015/294930>

- Luo, B. (2022). A Method for Enterprise Network Innovation Performance Management Based on Deep Learning and Internet of Things. *Mathematical Problems in Engineering*, 2022. <https://doi.org/10.1155/2022/8277426>
- Ma, Y., Liu, H., Zhai, G. y Huo, Z. (2021). Financial Risk Early Warning Based on Wireless Network Communication and the Optimal Fuzzy SVM Artificial Intelligence Model. *Wireless Communications and Mobile Computing*, 2021. <https://doi.org/10.1155/2021/7819011>
- Mai, F., Tian, S., Lee, C. y Ma, L. (2019). Deep learning models for bankruptcy prediction using textual disclosures. *European Journal of Operational Research*, 274(2), 743-758. <https://doi.org/10.1016/j.ejor.2018.10.024>
- Manthoulis, G., Doumpos, M., Zopounidis, C. y Galariotis, E. (2020). An ordinal classification framework for bank failure prediction: Methodology and empirical evidence for US banks. *European Journal of Operational Research*, 282(2), 786-801. <https://doi.org/10.1016/j.ejor.2019.09.040>
- Marso, S. y Merouani, M. E. L. (2020). Bankruptcy prediction using hybrid neural networks with artificial bee colony. *Engineering Letters*, 28(4), 1191-1200.
- Min, J. H. y Lee, Y. C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28(4), 603-614. <https://doi.org/10.1016/j.eswa.2004.12.008>
- Mousavi, M. M. y Lin, J. (2020). The application of PROMETHEE multi-criteria decision aid in financial decision making: Case of distress prediction models evaluation. *Expert Systems with Applications*, 159, 113438. <https://doi.org/10.1016/j.eswa.2020.113438>
- Nyitrai, T. y Virág, M. (2019). The effects of handling outliers on the performance of bankruptcy prediction models. *Socio-Economic Planning Sciences*, 67(August 2017), 34-42. <https://doi.org/10.1016/j.seps.2018.08.004>
- Park, M. S., Son, H., Hyun, C. y Hwang, H. J. (2021). Explainability of machine learning models for bankruptcy prediction. *IEEE Access*, 9, 124887-124899. <https://doi.org/10.1109/ACCESS.2021.3110270>

- Perboli, G. y Arabnezhad, E. (2021). A Machine Learning-based DSS for mid and long-term company crisis prediction. *Expert Systems with Applications*, 174(July 2020), 114758. <https://doi.org/10.1016/j.eswa.2021.114758>
- Pérez-Pons, M. E., Parra-Dominguez, J., Hernández, G., Herrera-Viedma, E. y Corchado, J. M. (2022). Evaluation metrics and dimensional reduction for binary classification algorithms: A case study on bankruptcy prediction. *Knowledge Engineering Review*, 37(4), 8-10. <https://doi.org/10.1017/S026988892100014X>
- Ping, W., Wang, F., Wang, A. y Huang, Y. (2021). Risk Early Warning Research on China's Futures Company. *Emerging Markets Finance and Trade*, 57(8), 2259-2270. <https://doi.org/10.1080/1540496X.2019.1689355>
- Prasad, M. V. N. K., Nickolas, S. y Gangadharan, G. R. (2019). General representational automata using deep neural networks. *Data and Knowledge Engineering*, 122, 159-180. <https://doi.org/10.1016/j.datak.2019.06.004>
- Qian, H., Wang, B., Yuan, M., Gao, S. y Song, Y. (2022). Financial distress prediction using a corrected feature selection measure and gradient boosted decision tree. *Expert Systems with Applications*, 190(February 2021). <https://doi.org/10.1016/j.eswa.2021.116202>
- Ravi Kumar, P. y Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *European Journal of Operational Research*, 180(1), 1-28. <https://doi.org/10.1016/j.ejor.2006.08.043>
- Romero Martínez, M., Carmona Ibáñez, P. y Pozuelo Campillo, J. (2021). La predicción del fracaso empresarial de las cooperativas españolas. Aplicación del Algoritmo Extreme Gradient Boosting. *Revista de Economía Pública, Social y Cooperativa*, 255-288.
- Ribeiro, B., Silva, C., Chen, N., Vieira, A. y Carvalho Das Neves, J. (2012). Enhanced default risk models with SVM+. *Expert Systems with Applications*, 39(11), 10140-10152. <https://doi.org/10.1016/j.eswa.2012.02.142>
- Schalck, C. y Yankol-Schalck, M. (2021). Predicting French SME failures: new evidence from machine learning techniques. *Applied Economics*, 53(51), 5948-5963. <https://doi.org/10.1080/00036846.2021.1934389>

- Sermpinis, G., Tsoukas, S. y Zhang, Y. (2022). Modelling failure rates with machine-learning models: Evidence from a panel of UK firms. *European Financial Management*, (May). <https://doi.org/10.1111/eufm.12369>
- Shetty, S., Musa, M. y Brédart, X. (2022). Bankruptcy Prediction Using Machine Learning Techniques. *Journal of Risk and Financial Management*, 15(1). <https://doi.org/10.3390/jrfm15010035>
- Shi, Y. y Li, X. (2019). A bibliometric study on intelligent techniques of bankruptcy prediction for corporate firms. *Heliyon*, 5(12), e02997. <https://doi.org/10.1016/j.heliyon.2019.e02997>
- Shi, Y. y Li, X. (2019). An overview of bankruptcy prediction models for corporate firms: A systematic literature review. *Intangible Capital*, 15(2), 114-127. <https://doi.org/10.3926/ic.1354>
- Shrivastav, S. K. y Janaki Ramudu, P. (2020). Bankruptcy prediction and stress quantification using support vector machine: Evidence from Indian banks. *Risks*, 8(2). <https://doi.org/10.3390/risks8020052>
- Sinelnikova-Muryleva, E. V., Gorshkova, T. G. y Makeeva, N. V. (2018). Default forecasting in the Russian Banking sector. *Ekonomicheskaya Politika*, 13, 8-27. <https://doi.org/10.18288/1994-5124-2018-2-01>
- Sinelnikova-Muryleva, E. V., Gorshkova, T. G. y Makeeva, N. V. (2018). Default forecasting in the Russian Banking sector. *Ekonomicheskaya Politika*, 13(2), 8-27. <https://doi.org/10.18288/1994-5124-2018-2-01>
- Siswoyo, B., Abas, Z. A., Pee, A. N. C., Komalasari, R. y Suyatna, N. (2022). Ensemble machine learning algorithm optimization of bankruptcy prediction of bank. *IAES International Journal of Artificial Intelligence*, 11(2), 679-686. <https://doi.org/10.11591/ijai.v11.i2.pp679-686>
- Smith, M. y Alvarez, F. (2022). Predicting Firm-Level Bankruptcy in the Spanish Economy Using Extreme Gradient Boosting. In *Computational Economics* (Vol. 59). <https://doi.org/10.1007/s10614-020-10078-2>
- Smiti, S. y Soui, M. (2020). Bankruptcy Prediction Using Deep Learning Approach Based on Borderline SMOTE. *Information Systems Frontiers*, 22(5), 1067-1083. <https://doi.org/10.1007/s10796-020-10031-6>

- Son, H., Hyun, C., Phan, D. y Hwang, H. J. (2019). Data analytic approach for bankruptcy prediction. *Expert Systems with Applications*, 138, 112816. <https://doi.org/10.1016/j.eswa.2019.07.033>
- Song, Y. y Peng, Y. (2019). A MCDM-Based Evaluation Approach for Imbalanced Classification Methods in Financial Risk Prediction. *IEEE Access*, 7(Mcdm), 84897-84906. <https://doi.org/10.1109/ACCESS.2019.2924923>
- Soui, M., Smiti, S., Mkaouer, M. W. y Ejbali, R. (2020). Bankruptcy Prediction Using Stacked Auto-Encoders. *Applied Artificial Intelligence*, 34(1), 80-100. <https://doi.org/10.1080/08839514.2019.1691849>
- Sue, K. L., Tsai, C. F. y Chiu, A. (2021). The data sampling effect on financial distress prediction by single and ensemble learning techniques. *Communications in Statistics—Theory and Methods*, 0(0), 1-12. <https://doi.org/10.1080/03610926.2021.1992439>
- Sutiene, K., Luksys, K. y Kundeliene, K. (2021). Towards Automation of Short-Term Financial Distress Detection: A Real-World Case Study. *International Journal of Information Technology and Decision Making*, 20(4), 1299-1333. <https://doi.org/10.1142/S0219622021500334>
- Tang, X., Li, S., Tan, M. y Shi, W. (2020). Incorporating textual and management factors into financial distress prediction: A comparative study of machine learning methods. *Journal of Forecasting*, 39(5), 769-787. <https://doi.org/10.1002/for.2661>
- Tian, Y., Shi, Y. y Liu, X. (2012). Recent advances on support vector machines research. *Technological and Economic Development of Economy*, 18(1), 5-33. <https://doi.org/10.3846/20294913.2012.661205>
- Tkáč, M. y Verner, R. (2016). Artificial neural networks in business: Two decades of research. *Applied Soft Computing Journal*, 38, 788-804. <https://doi.org/10.1016/j.asoc.2015.09.040>
- Tsai, C. F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. *Information Fusion*, 16(1), 46-58. <https://doi.org/10.1016/j.inffus.2011.12.001>
- Tsai, C. F. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems*, 22(2), 120-127. <https://doi.org/10.1016/j.knosys.2008.08.002>

- Tsai, C. F. (2008). Financial decision support using neural networks and support vector machines. *Expert Systems*, 25(4), 380-393. <https://doi.org/10.1111/j.1468-0394.2008.00449.x>
- Tsai, C. F. (2020). Two-stage hybrid learning techniques for bankruptcy prediction*. *Statistical Analysis and Data Mining*, 13(6), 565-572. <https://doi.org/10.1002/sam.11482>
- Tsai, C. F., Hsu, Y. F. y Yen, D. C. (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing Journal*, 24, 977-984. <https://doi.org/10.1016/j.asoc.2014.08.047>
- Tsai, C. F. y Wu, J. W. (2008). Using neural network ensembles for bankruptcy prediction and credit scoring. *Expert Systems with Applications*, 34(4), 2639-2649. <https://doi.org/10.1016/j.eswa.2007.05.019>
- Tserng, H. P., Lin, G. F., Tsai, L. K. y Chen, P. C. (2011). An enforced support vector machine model for construction contractor default prediction. *Automation in Construction*, 20(8), 1242-1249. <https://doi.org/10.1016/j.autcon.2011.05.007>
- Tunio, F. H., Ding, Y., Agha, A. N., Agha, K. y Panhwar, H. U. R. Z. (2021). Financial Distress Prediction Using Adaboost and Bagging in Pakistan Stock Exchange. *Journal of Asian Finance, Economics and Business*, 8(1), 665-673. <https://doi.org/10.13106/jafeb.2021.vol8.no1.665>
- Uthayakumar, J., Metawa, N., Shankar, K. y Lakshmanaprabu, S. K. (2020). Intelligent hybrid model for financial crisis prediction using machine learning techniques. *Information Systems and E-Business Management*, 18(4), 617-645. <https://doi.org/10.1007/s10257-018-0388-9>
- Valencia, C., Cabrales, S., Garcia, L., Ramirez, J. y Calderona, D. (2019). Generalized additive model with embedded variable selection for bankruptcy prediction: Prediction versus interpretation. *Cogent Economics and Finance*, 7(1). <https://doi.org/10.1080/23322039.2019.1597956>
- Viswanathan, P. K., Srinivasan, S. y Hariharan, N. (2020). Predicting Financial Health of Banks for Investor Guidance Using Machine Learning Algorithms. *Journal of Emerging Market Finance*, 19(2), 226-261. <https://doi.org/10.1177/0972652720913478>

- 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>
- Wang, G., Ma, J., Chen, G. y Yang, Y. (2020). Financial distress prediction: Regularized sparse-based Random Subspace with ER aggregation rule incorporating textual disclosures. *Applied Soft Computing Journal*, 90, 106152. <https://doi.org/10.1016/j.asoc.2020.106152>
- Wang, H. y Liu, X. (2021). Undersampling bankruptcy prediction: Taiwan bankruptcy data. *PLoS ONE*, 16(7 July), 1-17. <https://doi.org/10.1371/journal.pone.0254030>
- Wang, M., Chen, H., Li, H., Cai, Z., Zhao, X., Tong, C., ... Xu, X. (2017). Grey wolf optimization evolving kernel extreme learning machine: Application to bankruptcy prediction. *Engineering Applications of Artificial Intelligence*, 63, 54-68. <https://doi.org/10.1016/j.engappai.2017.05.003>
- Whiting, D. G., Hansen, J. V., McDonald, J. B., Albrecht, C. y Albrecht, W. S. (2012). Machine learning methods for detecting patterns of management fraud. *Computational Intelligence*, 28(4), 505-527. <https://doi.org/10.1111/j.1467-8640.2012.00425.x>
- Xiaosi, X., Ying, C. y Haitao, Z. (2011). The comparison of enterprise bankruptcy forecasting method. *Journal of Applied Statistics*, 38(2), 301-308. <https://doi.org/10.1080/02664760903406470>
- Xu, W., Pan, Y., Chen, W. y Fu, H. (2019). Forecasting corporate failure in the Chinese energy sector: A novel integrated model of deep learning and support vector machine. *Energies*, 12(11). <https://doi.org/10.3390/en12122251>
- Yarahmadia, H., Shirib, M., Navidic, H. y Sharifid, A. (2021). A bankruptcy based approach to solving multi-agent credit assignment problem. *Ijnaa.Semnan.Ac.Ir*, 12(December), 1987-2018. https://ijnaa.semnan.ac.ir/index.php/themes/base/front/assets/plugins/journal/journal/article_5968_aeb4b80f492120d6f67d6f57c5e9c0e2.pdf
- Yousaf, U. Bin, Jebran, K. y Wang, M. (2022). A comparison of static, dynamic and machine learning models in predicting the financial distress of Chinese firms. *Journal for Economic Forecasting*, 1, 122-138.

- Yu, Q., Miche, Y., Séverin, E. y Lendasse, A. (2014). Bankruptcy prediction using Extreme Learning Machine and financial expertise. *Neurocomputing*, 128, 296-302. <https://doi.org/10.1016/j.neucom.2013.01.063>
- Zeng, S., Li, Y., Yang, W. y Li, Y. (2020). A financial distress prediction model based on sparse algorithm and support vector machine. *Mathematical Problems in Engineering*. <https://doi.org/10.1155/2020/5625271>
- Zhang, X. y Hu, L. (2016). A nonlinear subspace multiple kernel learning for financial distress prediction of Chinese listed companies. *Neurocomputing*, 177, 636-642. <https://doi.org/10.1016/j.neucom.2015.11.078>
- Zhang, Y., Liu, R., Heidari, A. A., Wang, X., Chen, Y., Wang, M. y Chen, H. (2021). Towards augmented kernel extreme learning models for bankruptcy prediction: Algorithmic behavior and comprehensive analysis. *Neurocomputing*, 430, 185-212. <https://doi.org/10.1016/j.neucom.2020.10.038>
- Zhao, D., Huang, C., Wei, Y., Yu, F., Wang, M. y Chen, H. (2017). An effective computational model for bankruptcy prediction using kernel extreme learning machine approach. *Computational Economics*, 49(2), 325-341. <https://doi.org/10.1007/s10614-016-9562-7>
- Zięba, M., Tomczak, S. K. y Tomczak, J. M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems with Applications*, 58, 93-101. <https://doi.org/10.1016/j.eswa.2016.04.001>
- Zoričák, M., Gnip, P., Drotár, P. y Gazda, V. (2020). Bankruptcy prediction for small- and medium-sized companies using severely imbalanced datasets. *Economic Modelling*, 84(February), 165-176. <https://doi.org/10.1016/j.econmod.2019.04.003>
- Zou, Y., Gao, C. y Gao, H. (2022). Business Failure Prediction Based on a Cost-Sensitive Extreme Gradient Boosting Machine. *IEEE Access*, 10, 42623-42639. <https://doi.org/10.1109/ACCESS.2022.3168857>