BANKRUPTCY PREDICTION USING MACHINE LEARNING – A META-ANALYSIS

https://doi.org/10.47743/jopafl-2022-26-06

Claudiu CLEMENT
Alexandru Ioan Cuza University of Iași, Faculty of Economics and Business Administration
Iași, România
office.claudiuclement@gmail.com

Mihaela DAVID
"Gheorghe Zane" Institute for Economic and Social Research - Romanian Academy, Iași Branch
Iaşi, Romania
mihaela_david88@yahoo.com

Danuț-Vasile JEMNA
Alexandru Ioan Cuza University of Iași, Faculty of Economics and Business Administration
Iași, România
danut.jemna@uaic.ro

Abstract: *This study is based on a meta-analysis of 64 studies in bankruptcy prediction using machine learning. The data on these studies was collected on six levels: algorithms, data balance, variable categories, variables types, industry, and region. The aim of this paper is to analyse the determinants of accuracy in bankruptcy prediction models. To achieve this aim, five Linear Mixed Effects models were developed. The results obtained show that while some factors are significant determinants for the accuracy of machine learning models in bankruptcy prediction (algorithm, data balance, industry, region), some factors as data type (continuous or continuous and categorical) and data category (financial or financial and non-financial) do not have an impact on accuracy prediction.*
Keywords: *Bankruptcy Prediction, Machine Learning, Meta-Analysis*

Introduction

Bankruptcy prediction is and always was of interest to investors, creditors, and governments. The timely identification of the imminent state of bankruptcy of a company is undoubtedly desirable. Initially discussed by FitzPatrick (1932), the problem of bankruptcy prediction is a classic one in the economics literature. At the same time, bankruptcy maintains equilibrium in an economy (Filipe Martins-da-Rocha et al., 2022). The most significant impact of accurate bankruptcy prediction is on the case of lending institutions. Banks are a good example because they must predict the possibility of default of a counterparty before deciding to grant or expand a loan. An accurate prediction will lead to better lending decisions and thus the avoidance of significant losses. In consequence, the extensive research into this field is unquestionably justified (Alaka et al., 2018). Initial studies on bankruptcy prediction models (BPM) are mostly focused on methods such as logistic regression or discriminant analysis. However, in the last 30 years, the landscape drastically changed, mainly because of the quick development of machine learning (ML) techniques.

There has been an increasing number of research papers addressing the topic of bankruptcy prediction through ML (Shi & Li, 2019). The main advantages of ML methods in BPM are the increased prediction accuracy over classical statistical methods (such as Z-Score) and the ability to handle large volumes of data (Hosaka, 2019). Although the vast majority of researchers and practitioners agree on the benefits of utilizing ML methods for predicting bankruptcy, there is little to no consensus on what method should be used for which use case. These models' performance is highly dependent on the algorithms chosen and the tweaking of their respective parameters. According to Alaka et al. (2018), the model choice is not objective and is often based on popularity or professional background because researchers do not have any evaluation material guiding them through which criteria a BPM should satisfy. In this context, there are significant differences in the reported results of studies without any apparent motivation for why this happens. Indeed, discrepancies between results can be generated by multiple factors such as different datasets, different tunning methods, or by particular characteristics of each method and their fit to a dataset. Methods such as Artificial Neural Networks (ANN), for example, are frequently misused due to the lack of guidelines and frameworks that would help researchers and practitioners identify which methods best suit their data conditions or research situation (Alaka et al., 2018). Moreover, the improper use of methods makes it very difficult for researchers to understand the strengths and limitations so they can adapt their choice of method to the research goals. Finally, the wrong choice of methods determines different studies to have a wide dispersion of performance results for the same method. Understanding these methods' performance drivers will help in choosing the right tool for the specific research data/purpose/situation.

In the context of scattered performance results (bankruptcy prediction is usually measured by accuracy, the area under the curve, and the ROC curve), this study aims to analyze the relationship between the bankruptcy prediction accuracy of different ML techniques and the factors that influence it through a meta-analysis. The meta-analysis on bankruptcy prediction accuracy enables to bring coherence in a field with inconsistent findings and explore what generates these differences. To the best of our knowledge, this study contributes by being the first meta-analytic study in the field of bankruptcy prediction models. As we detail in the literature review section, multiple studies address the topic through a review methodology but not through a meta-analysis. The fit of the methodology for the topic, the increasing number of studies, and the absence of agreement in the literature motivate the development of a meta-analysis of studies on machine learning methods applied to bankruptcy prediction. This will help see the problem through a quantitative lens and facilitate the enhancement of the model selection framework started by Alaka et al. (2018). The latter authors initiated the creation of a framework for model selection while briefly describing the advantages and disadvantages of the models. However, the decision on the type of model and its accuracy would depend upon factors such as sample size, data balance, variable category (whether the authors used financial or mixed variables), industry type (for example, construction or manufacturing), or data source region. None of the previous review studies analyse the prediction accuracy in relation to factors such as variables category, industry type, or data source region, which are included in our study.

The remainder of the paper is structured as follows. Section 2 provides an overview of the existing literature analysing the differences between the results of bankruptcy prediction studies using ML techniques. Section 3 describes the review process and the methodology adopted for the meta-analysis. Section 4 the results of the meta-analysis are shown. The paper ends with concluding remarks and references.

Previous research in bankruptcy prediction

Currently, both academics and professionals are applying ML methods to predict bankruptcy, and their use in BPM is rapidly increasing. According to Gissel et al. (2007), the main models utilized for BPM development are multiple discriminant analysis (MDA), logistic regression (LR), probit analysis, neural networks (NN), and, due to their frequent use in recent studies, we would add decision trees (DT), Ensemble Models (EM) and support vector machines (SVM) to the list. Among the classic methods, LR is frequently used in practice and academia because it is easy to implement and proves to provide users with satisfactory results (Hauser & Booth, 2011). One of the drawbacks of regression models is the incapability to manage the data imbalance that frequently occurs with bankruptcy databases due to the rarity of the phenomenon. Another class is "lazy algorithms" first mentioned by Aha (1997), here including k-Nearest Neighbor (kNN) and Case-Based Reasoning (CBR). Chen et al. (2011), Liang et al. (2016), and Le & Viviani (2018) use k-NN with promising results as a new candidate for powerful early warning systems for bankruptcy prediction. Furthermore, due to the large access to computing power, neural networks started to pick up at the beginning of the year 1990. Odom & Sharda (1990) and Tam & Kiang (1992) introduced artificial neural networks (ANN) for the prediction of corporate bankruptcies. Moreover, of all the methods used to predict bankruptcy, some of the simplest, in terms of mathematical complexity, are the methods based on decision trees. Frydman, Altman, & Kao (1985) were the first to use decision trees with positive results. Chen (2011) argues that decision trees are superior to logistic regression, especially for short term predictions.

The development of ML models is without a doubt a big leap forward for the bankruptcy prediction domain. However, the multitude of models is overwhelming, and the scattered results presented in the literature make model choice very difficult. Moreover, the improper use of these models, regularly due to not fully understanding their strengths and limits (Chung, Tan, & Holdsworth, 2008), leads to different or even biased results on the bankruptcy prediction accuracy. Therefore, BPM developers should choose a tool based on data characteristics/variable types/purposes/situations rather than just arbitrarily. Based on this lack of clarity, this study thus aims to evaluate the relationship between accuracy and a set of external factors (such as algorithm, variable's category, variable's type, industry, data balance, and region). In order to achieve this aim, this research adresses the results of recent studies which were not reviewed previously and examining the determinants of the differences in results.

As most studies conducted in this area apply a systematic review for analyzing the results, our approach aims at minimizing the subjectivity and have a base on a quantitative analysis with the ultimate goal of replicability and reliability of empirical results. In this respect, a meta-analysis of BPM studies is employed as it can integrate the results and provide a quantitative overview of the differences between results.

According to the Research Papers in Economics database, more than 2,000 meta-analyses are conducted, approaching very diverse topics (Havránek et al., 2020). Meta-analysis on finance, contrary to economics, is a relatively new research field that has been highly influenced by the meta-analytic research methods in management and economics (Geyer-Klingeberg et al., 2020b). Until December 2019, Geyer-Klingeberg et al. (2020) found 61 meta-analyses aggregating and comparing finance results. The authors' findings show that meta-analysis has been used as a tool that can facilitate progress in financial research, especially in areas with mixed evidence and controversial theoretical concepts.

To the best of our knowledge, on bankruptcy prediction, there was no meta-analysis developed. Although the literature on the subject is abundant, there are only systematic reviews on the topic. Clement (2020) presents a systematic review of the papers written on bankruptcy prediction between 2016 and 2020 and concluds that there is no clear path that a researcher should follow in addressing a bankruptcy prediction problem as the literature is very dispersed. Intending to create a tool selection framework, Alaka et al. (2018) study 49 journal articles on BPM published between 2010 and 2015 and concludes that there is no clear delimitation of which machine learning method is better than others, and there is still plenty of room for research. Shi & Li (2019) did a comprehensive systematic review on BPMs by analysing 496 academic articles published between 1968 and 2017. Their study has three main findings: (1) there is an abundance of academic papers, especially after the 2008 financial crisis; (2) there is little co-authorship in the research area as influential researchers were not working together for the development of the domain; (3) by far the most used and models for BPMs are LR and NN.

Methodology

Data sources and search strategy

A comprehensive search strategy was designed and completed within the Web of Science, Scopus, and ScienceDirect databases covering the full timeframe from 1968 through 2021. Google Scholar was also considered initially, but as in the case of Alaka et al. (2018) there was no filtering capacity and a significant number of papers loaded on various mixed topics hence this database was excluded. No search filters provided by the databases were used as well as no restriction by publication type (e.g. journal articles, conference proceedings). No studies other than those written in English are included in this study as we did not identify studies in other languages through the databases. We used the keywords "BANKRUPTCY PREDICTION" OR "FAILURE PREDICTION" OR "INSOLVENCY PREDICTION" AND "USING MACHINE LEARNING" OR "INTELLIGENT TECHNIQUES" to search for papers relevant to our study. We identified a list of 183 studies but selected only 64 for this study. No restriction was applied based on the type of machine learning algorithm used. A process flow of the methodology is presented in Figure 1 and the final sample of our selected literature is listed in Table 1.

Figure 2. Search Methodology

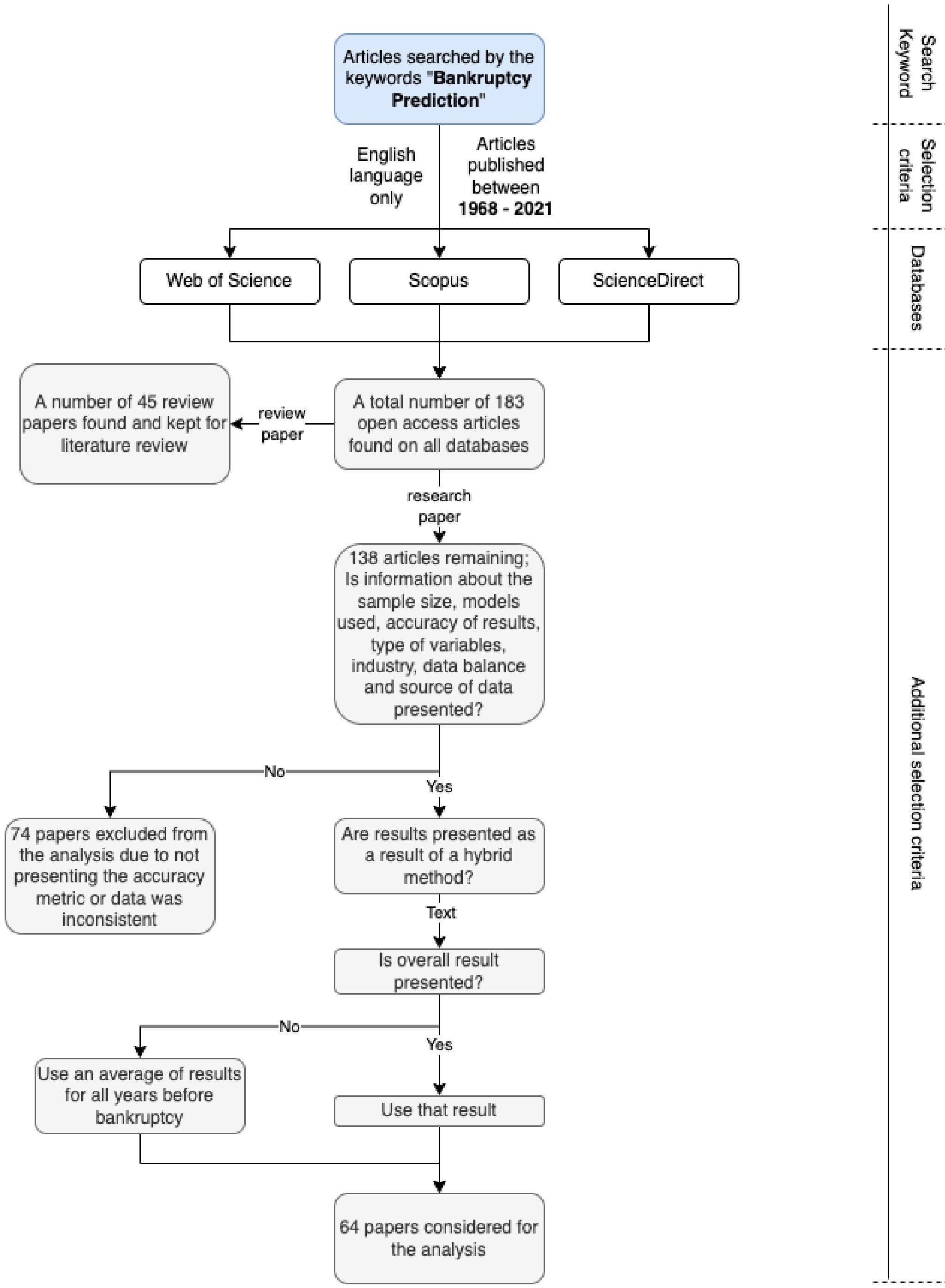

Table 1. List of selected studies

#	Authors	Year of study	#	Authors	Year of study
1	Altman	1968	33	Alifiah	2014
2	Fletcher, Goss	1993	34	Geng, Bose, Chen	2014
3	Canbas, Cabuk, Kilic	1997	35	Heo, Yang	2014
4	Dimitras et al.	1999	36	Kim, Kang, Kim	2014
5	Atyia	2001	37	Kim, Upneja,	2014
6	Lin, McClean	2001	38	Liao et al	2014
7	Min, Lee	2005	39	Lopez-Iturriaga, Pastor-Sanz	2014
8	Tsakonas et al.	2006	40	Tsai	2014
9	Alfaro, Garcia, Gamez, Elizondo	2007	41	Wang, Ma, Yang	2014
10	Hua et al.	2007	42	Gordini	2014
11	Guo	2008	43	Yu et al.	2014
12	Chen, Huang, Lin	2009	44	Alaminos, Del Castillo, Fernandez	2016
13	Cho, Kim, Bae	2009	45	Du Jardin	2016
14	Hung, Chen	2009	46	Sun et al	2016
15	Cho, Hong, Ha	2010	47	Liang et al.	2016
16	Kim, Kang	2010	48	Sartori, Mazzuchelli, Gregorio	2016
17	Tseng, Hu	2010	49	Antunes, Ribeiro, Pereira	2017
18	Van Gestel, Baesens, Martens	2010	50	Zelenkov, Fedorova, Chekrizov	2017
19	Yoon, Kwon	2010	51	Carmona, Climent, Momparler	2018
20	Chaudhuri	2011	52	Mai el al.	2018
21	Chen	2011	53	Gogas, Papadimitrou, Agrapetidou	2018
22	Li et al.	2011	54	Le, Viviano	2018
23	Sun, Jia, Li	2011	55	Obradovic et al.	2018
24	Hauser, Booth	2011	56	Affes, Hentati-Kaffel	2019
25	Yang, You, Li	2011	57	Agrawal, Maheshwari	2019
26	Huang et al.	2012	58	Chang	2019
27	Li, Sun	2012	59	Charalambakis, Garett	2019
28	Marques et al.	2012	60	Korol	2019
29	Xiao et al.	2012	61	Lukason, Andersson	2019
30	Olson, Delen, Meng	2012	62	Munoz-Izquierdo et al.	2019
31	Fedorova, Gilenko, Dovzhenko	2013	63	Aliaj, Anagnostopoulos, Piersanti	2020
32	Lee, Choi	2013	64	Filetti, Grech	2020

We selected studies based on the following inclusion criteria. Firstly, studies should use one statistical or intelligent technique for an empirical study predicting bankruptcy.

Secondly, studies should report an accuracy metric for the efficacy of the model tested in the empirical study. Third, studies were filtered for a specific time frame, following same approach as Alaka et al. (2018). From this perspective, this study is also similar to Balcaen & Ooghe (2006). Finally, studies reporting multiple accuracy rates for the same method or for different times before the bankruptcy event were included but the accuracy rate was calculated as an average.

We excluded papers reporting an efficacy metric other than accuracy (ROC Curve, AUC, etc.) as for keeping the consistency of the effect size in our study. Furthermore, we excluded papers predicting bankruptcy with other methods than the ones presented in the introduction due to some of them being only presented in one or two studies and thus making it difficult for comparative analysis. Finally, we did not include review studies presenting aggregated results of previous studies. Along the 64 papers studied there are 13 different algorithm types as shown in Table 2.

Table 2. Frequency of use of different types of algorithms in selected studies

	Algorithm Type	**Algorithm variations**
1	Bayesian Methods	3
2	Case-base reasoning	1
3	Clustering	1
4	Decision Trees	9
5	Discriminant Analysis	6
6	Ensemble Models	18
7	Genetic Analysis	1
8	KNN	1
9	Neural Networks	12
10	Regression Model	5
11	Revise	1
12	Rough Sets	2
13	SVM	4
	Total	**64**

Coding of data

Each paper was read thoroughly and a coded set of characteristics and statistical estimates was reported from each study. The coding was done by one coder with a second coder independently reconciling any inconsistencies. Research questions, research literature searching, compilation and coding were all done according to the guidelines from the Meta-analysis of Economics Research Reporting Guidelines of Stanley et al. (2013).

We collected the simple arithmetic mean-based accuracy ratio for the dependent variable, this metric being the most used across all papers. Hossin et al. (2011) argue that accuracy is mostly preferred by researchers because of its ease of calculation. Accuracy is calculated as below:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

The confusion matrix below (Table 3) highlights how each element from the accuracy equation is measured.

Table 3. Confusion Matrix

		PREDICTED CATEGORY	
		Positive	**Negative**
ACTUAL CLASS	**True**	True Positive (TP)	True Negative (TN)
	False	False Positive (FP)	False Negative (FN)

The included studies have been coded according to a number of factors listed out in Table 4. They included recording the (1) the prediction accuracy of the method, (2) the type of algorithm used for analysis, (3) the category of variables included in the selected studies, i.e. whether the data are financial or mixed, (4) the type of variables used for the analysis, i.e. whether the variables are continuous or mixed, (5) the industry the data belonged to, i.e. banking, mixed, or other (including construction, industrial, and restaurants), (6) the data balance, i.e. wheter the selected study used balanced or unbalanced data sets, and finally (7) the region where the study was performed or where the data belonged to. The information was collated for all 64 studies which resulted in the collection of 242 unique rows of information.

Table 4. Factors extracted from the literature review and the variables used for the meta-analysis

VARIABLES	DESCRIPTION
Dependent variable	
Accuracy	Simple arithmetic mean-based accuracy.
Independent variables	
Algorithm	Dummy variable denoting the use of Decision Trees; 1=yes; else 0; Dummy variable denoting the use of Discriminant Analysis; 1=yes; else 0; Dummy variable denoting the use of Ensemble Models; 1=yes; else 0; Dummy variable denoting the use of Neural Networks; 1=yes; else 0; Dummy variable denoting the use of Decision Trees; 1=yes; else 0; Dummy variable denoting the use of SVM; 1=yes; else 0; Dummy variable denoting the use of Other ML techniques; 1=yes; else 0; Where Bayesian Methods represents the reference category.
Variables' category	Dummy variable denoting the use of data from mixed (banking and non-banking) industries; 1=yes; else 0; Where financial data indicates the reference category.
Variables' type	Dummy variable denoting the use of mixed variables; 1=yes; else 0; Where continuous variables represents the reference category.
Industry	Dummy variable denoting the use of data from mixed industries; 1=yes; else 0; Dummy variable denoting the use of data from non-banking industries; 1=yes; else 0; Where banking industry represents the reference category.
Data balance	Dummy variable denoting the use of balanced data sets; 1=yes; else 0
Region of data	Dummy variable denoting the use of Asian datasets; 1=yes; else 0; Dummy variable denoting the use of Australia datasets; 1=yes; else 0;

	Dummy variable denoting the use of European datasets; 1=yes; else 0; Dummy variable denoting the use of Global datasets; 1=yes; else 0; Where American databases indicate the reference category.

Conducting the meta-analysis

Following the study of Varghese et al. (2020), the meta-analysis on the prediction accuracy was employed using linear mixed effects models. The advantage of using linear mixed effects models is that it allows for random effects to account for unobserved heterogeneity in bankruptcy prediction accuracy across studies. In addition to random effects, the fixed effects of the determinants of prediction accuracy listed in Table 4 were estimated.

In conducting the meta-analysis, five separate models were developed. In all models, all 64 studies were included (N = 220) and the main independent variable considered was the type of algorithm used in the analysis of the selected studies. In this respect, Model I tested the differences between methods and their impact on prediction accuracy. Model II examined the effect of data balance, along with that of algorithm type and their interaction, on the bankruptcy prediction accuracy. Model III investigated the impact of the type and the category of variables used in the selected studies on prediction accuracy. Model IV includes the industry's impact and the region where the data belong to in the studies addressed. Finally, Model V tested the effects of all fixed effects from the other models and the random effects generated by the studies. The results of the meta-analysis are discussed in Section 4.

Empirical results

Random effects: study

The use of mixed effects linear model allowed to account for intrinsic heterogeneities among studies. The results (provided in Table 5) indicate that there is a substantial degree of variance in the random effect of studies and that the variation between them had a major impact on prediction accuracy. Additionally, it was found that the variance of the random effect parameter was greater than the variance of the residual variance. The R^2 of Model V is higher than the R^2 of the other models including some of the fixed effects. This difference demonstrates that the accuracy is impacted significantly by the determinants studied.

Table 5. Results of the meta-analysis for all models

	Model I		Model II		Model III		Model IV		Model V	
Predictors	*Estimates*	*std. Error*	*Estimates*	*std. Error*	*Estimates*	*std. Error*	*Estimates*	*std. Error*	*Estimates*	*std. Error*
(Intercept)	0.70 ***	0.03	0.66 ***	0.04	0.70 ***	0.03	0.76 ***	0.05	0.73 ***	0.05
Fixed Effects										
Algorithms										

Decision Trees	0.13 ***	0.03	0.16 ***	0.04	0.13 ***	0.03	0.13 ***	0.03	0.16 ***	0.03
Discriminant Analysis	0.08 *	0.03	0.14 ***	0.04	0.08 *	0.03	0.08 **	0.03	0.14 ***	0.04
Ensemble Models	0.15 ***	0.03	0.22 ***	0.04	0.16 ***	0.03	0.15 ***	0.03	0.22 ***	0.03
Neural Networks	0.12 ***	0.03	0.17 ***	0.04	0.12 ***	0.03	0.13 ***	0.03	0.17 ***	0.03
Other	0.11 ***	0.03	0.17 ***	0.04	0.11 ***	0.03	0.11 ***	0.03	0.18 ***	0.04
Regression Model	0.09 **	0.03	0.12 ***	0.03	0.09 **	0.03	0.09 ***	0.03	0.12 ***	0.03
SVM	0.11 ***	0.03	0.16 ***	0.04	0.11 ***	0.03	0.10 ***	0.03	0.16 ***	0.03
Data Balance										
Unbalanced			0.12 *	0.06					0.13 *	0.06
*Algorithms*Balance*										
Decision Trees * Balance Unbalanced			-0.12 *	0.06					-0.12 *	0.05
Discriminant Analysis * Balance Unbalanced			-0.18 **	0.06					-0.19 **	0.06
Ensemble Models * Balance Unbalanced			-0.18 **	0.06					-0.17 **	0.05
Neural Networks * Balance Unbalanced			-0.15 *	0.06					-0.15 **	0.06
Other * Balance Unbalanced			-0.17 **	0.07					-0.18 **	0.06
Regression Model * Balance Unbalanced			-0.13 *	0.06					-0.12 *	0.05

SVM * Balance Unbalanced			-0.17 **	0.06					-0.16 **	0.05
Variables types										
Mixed					-0.14	0.08			-0.13	0.08
Variables category										
Mixed					-0.01	0.04			0.01	0.05
Industry										
Mixed							-0.11 *	0.04	-0.11 *	0.05
Other							-0.14 *	0.05	-0.14 **	0.06
Region										
Asia							0.05 *	0.02	0.04 *	0.02
Australia							0.02	0.02	0.02	0.02
Europe							-0.03	0.02	-0.04 *	0.02
Global							-0.00	0.05	-0.00	0.05
Random effects (variance)										
Study	0.008		0.009		0.008		0.009		0.009	
Residual	0.003		0.003		0.003		0.003		0.002	
Model performance parameters										
Sample size (cases)	242		242		242		242		242	
Sample size (study)	64		64		64		64		64	
AIC	-543.87		-545.35		-567.16		-544.58		-570.68	
Log likelihood	281.94		284.67		299.58		290.29		311.34	

R2 (only with fixed effects)	0.067	***	0.092	**	0.121	***	0.182	***	0.230	***
R2 (with fixed and random effects)	0.721	***	0.754	***	0.725	***	0.793	***	0.818	***

* $p<0.05$ ** $p<0.01$ *** $p<0.001$

Notes: The reference category for each variable is: Algorithm: *Bayesian Methods*; Balance: *Balanced*; Variables type: *Continuous*; Variables category: *Financial*; Industry: *Banking*; Region: *America*. The *Other* category for the variable Algorithm includes Case-base reasoning; Clustering; Genetic analysis; Revise; Rough sets. The *Other* category for the variable Industry includes: *Construction*; *Industrial*; *Restaurants*.

Model I: impact of algorithms on prediction accuracy

Prediction accuracy was found to have a significant relationship with algorithms. This confirms the hypothesis that the machine learning models used in bankruptcy prediction models have significantly different accuracy results. All algorithms display statistically significant differences from the reference method, the Bayesian Method, at the risk level of 0.05. The estimate's positive sign indicates that, when measured against the reference algorithm, the average accuracy of all algorithms is higher. The Ensemble Methods algorithm showed the most significant positive impact on prediction accuracy, followed by Decision Trees and Neural Networks.

Model II: impact of data balance on prediction accuracy

We tested the effect of data balance on algorithms' accuracy in this model. In light of the interaction between the data balance and the algorithm, the corresponding estimates are similar in sign and significance to Model I, indicating robust results. We found that the data balance has a substantial impact, and the estimate indicates that accuracy on unbalanced datasets is generally worse than accuracy on balanced datasets (0.12). We took into consideration the relationship between Bayesian Methods and Balanced data for the fixed effect of data balance. With the use of this, we were able to determine how various machine learning techniques and data balance ultimately affected accuracy. All machine learning models' prediction accuracy was revealed to be severely impacted by data unbalance when compared to the reference algorithm. All estimates are at least 0.05 level significant.

Model III: impact of both category and type of variables on prediction accuracy

Dummy variables for the category and type of variables were also utilized as predictor variables in Model III, in addition to the algorithms. We did not identify a statistically significant relationship between accuracy and the category or type of variables used in studies. This finding leads to two conclusions. First, the studies with mixed data (financial and non-financial data) failed to significantly outperform those only using financial data. Second, studies using mixed variables (continuous and categorical) failed to significantly outperform those only using continuous variables. These results seem contradictory in a way because non-financial data is frequently used in research today, and are mostly presented as a fruitful research direction. Note that these findings do not suggest that the use of mixed non-financial and financial data or a mix of continuous and noncontinuous data is to be avoided but that a quantitative analysis does not show the benefit of doing so.

Model IV: impact of industry and region on prediction accuracy

Data region and industry impact on accuracy were also investigated. It was observed that studies focusing on the banking industry have better accuracy than the ones focused on mixed industries. Also, datasets focused on other individual industries have a positive impact on accuracy. Moreover, with respect to the region of the datasets, only datasets from Asia and Europe exhibit significantly better accuracy performance than datasets from America. These findings show that models do not differ significantly in performance based on the region of data.

Model V: impact of all factors on prediction accuracy

Model V, which takes into account all factors, shows that the studies random effects had a large variation and that the differences between them have a considerable influence on prediction accuracy. Furthermore, the highest contribution of the random effect is shown by comparing the R^2 value associated with fixed effects and the R^2 value associated with both fixed effects and random effects (0.230 vs. 0.819). The AIC shows that Model V is better at explaining the variation of prediction accuracy. Moreover, after the inclusion of every variable in the final model, the estimated coefficients keep their sign and significance and are not significantly different in magnitude.

Conclusions

The use of ML techniques in bankruptcy prediction research was thoroughly surveyed in this meta-analysis. Sixty-four studies were reviewed and data were extracted on a number of six variables. With regard to data balance, variables' category, variables' type, industry, and region, this meta-analysis sought to explain the prediction accuracy of several machine learning algorithms. This study adds to the body of literature in several ways. First, our results demonstrate that there is a statistically significant difference in accuracy performance between machine learning models in the studies on the subject of bankruptcy prediction. These differences are mainly driven by the algorithm model, industry and region of data. Second, we were able to pinpoint some of the key factors that influence machine learning prediction accuracy by examining studies in relation to some variables that, to the best of our knowledge, have never been considered before in BPMs (data type, data category, industry, and region). Third, as far as it came to our attention, there are no studies that examine the factors that affect the accuracy of bankruptcy prediction models.

Disclaimer

This research is part of an expanded research presented in the PhD thesis of Claudiu Clement titled "Machine Learning Methods in Bankruptcy Prediction".

References

1. Aha, D. W. (1997). Editorial. In Lazy Learning (pp. 7–10). Springer Netherlands. https://doi.org/10.1007/978-94-017-2053-3_1
2. Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayi, S. O., Akinade, O. O., & 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
3. Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure : an overview of the classic statistical methodologies and their related problems. 38, 63–93. https://doi.org/10.1016/j.bar.2005.09.001
4. Chen, H. L., Yang, B., Wang, G., Liu, J., Xu, X., Wang, S. J., & Liu, D. Y. (2011). A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method. Knowledge-Based Systems, 24(8), 1348–1359. https://doi.org/10.1016/j.knosys.2011.06.008
5. Chen, M. Y. (2011). Predicting corporate financial distress based on integration of decision tree classification and logistic regression. Expert Systems with Applications, 38(9), 11261–11272. https://doi.org/10.1016/j.eswa.2011.02.173
6. Filipe Martins-da-Rocha, V., Mouallem Rosa, R., Martins-da-Rocha, V. F., & Rosa, R. M. (2022). Complete markets with bankruptcy risk and pecuniary default punishments. Economic Theory 2022, 1–16. https://doi.org/10.1007/S00199-022-01429-1
7. FitzPatrick, P. J. (1932). A comparison of the ratios of successful industrial enterprises with those of failed companies.
8. Frydman, H., Altman, E. I., & Kao, D.-L. (1985). Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress. Journal of Finance, 40(1), 269–291. https://econpapers.repec.org/RePEc:bla:jfinan:v:40:y:1985:i:1:p:269-91
9. Gissel, J. L., Giacomino, D., & Akers, M. D. (2007). A Review of Bankruptcy Prediction Studies: 1930-Present. In Journal of Financial Education (Vol. 33). Publisher Link.
10. Glass, G. V. (1976). Primary, Secondary, and Meta-Analysis of Research. Educational Researcher, 5(10), 3–8. https://doi.org/10.3102/0013189x005010003
11. Hauser, R. P., & Booth, D. (2011). Predicting Bankruptcy with Robust Logistic Regression. Journal of Data Science, 9, 565–584.
12. 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
13. Hossin M, Sulaiman M N, Mustapha A, Mustapha N, & Rahmat R W. (2011). A Hybrid Evaluation Metric for Optimizing Classifier. 28–29.
14. Le, H. H., & 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(February 2017), 16–25. https://doi.org/10.1016/j.ribaf.2017.07.104
15. Lee, Y. H. (2019). Strengths and Limitations of Meta-Analysis. The Korean Journal of Medicine, 94(5), 391–395. https://doi.org/10.3904/kjm.2019.94.5.391
16. Liang, D., Lu, C. C., Tsai, C. F., & 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
17. Odom, M. D., & Sharda, R. (1990). A neural network model for bankruptcy prediction. IJCNN. International Joint Conference on Neural Networks, 163–168. https://doi.org/10.1109/ijcnn.1990.137710
18. Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. Clinical Psychology Review, 17(8), 881–901. https://doi.org/10.1016/S0272-7358(97)00056-1
19. Shi, 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
20. Stanley, T. D., Doucouliagos, H., Giles, M., Heckemeyer, J. H., Johnston, R. J., Laroche, P., Nelson, J. P., Paldam, M., Poot, J., Pugh, G., Rosenberger, R. S., & Rost, K. (2013). Meta-analysis of economics research reporting guidelines. Journal of Economic Surveys, 27(2), 390–394. https://doi.org/10.1111/joes.12008
21. Tam, K. Y., & Kiang, M. Y. (1992). Managerial applications of neural networks: The case of bank failure predictions. Management Science, 38(7), 926–947. https://doi.org/10.1287/mnsc.38.7.926

22. Varghese, V., Chikaraishi, M., & Urata, J. (2020). Deep Learning in Transport Studies: A Meta-analysis on the Prediction Accuracy. Journal of Big Data Analytics in Transportation. https://doi.org/10.1007/s42421-020-00030-z