International Research in Economics and Finance; Vol. 6, No. 2; May, 2022
ISSN 2529-8038 E-ISSN 2591-734X
Published by July Press

A Logistic Regression Model for Credit Risk of Companies in the Service Sector

Lobna Abid[1]

[1] Economic Department, University of Sfax, ISAAS, Tunisia

Correspondence: Lobna Abid, Assistant Professor, Economic Department, University of Sfax, Aeroport Roat km 4.5, ISAAS, Tunisia.

Received: March 23, 2022 Accepted: April 21, 2022 Online Published: May 28, 2022

doi:10.20849/iref.v6i2.1179 URL: https://doi.org/10.20849/iref.v6i2.1179

Abstract

Credit risk prediction is a vital issue in empirical studies as it has attracted the interests of many researchers. In the current study, a logistic regression model is used to evaluate determinants of payment default risks of companies in the service sector.

Data, which consist of six financial variables and two macro-economic variables, have been collected from the Tunisian Central Bank and World Development Indicators.

The obtained results show that debt, solvency and profitability ratios and a loan amount are the key firm-specific factor determining credit risk. Moreover, we further find that high level of inflation and the decrease of GDP growth rate are able to increase corporate credit risk.

Keywords: credit risk, Logistic Regression, corporate credit

1. Introduction

Different disciplines, such as economics and econometrics, have shown interest in the study of firm failures, which shows the extent to which the firm plays a prominent role in economy. In this sphere, the relevance of firms' default of payment, as a field of study, stems from the fact that a firm's cash flow cannot meet the contractually required payment.

Firms in financial trouble, and obviously having payment difficulties, face many circumstances affecting their value and the well-being of creditors. This gives an explanation to the presence of literature that takes into account the issue of companies facing financial. The excessive allocation of loans results in a high level of unpaid loans, which exposes banks to many risks revolving mainly around their issue of credit risk.

Tunisia is among the countries that are mostly affected by the credit risk problem compared to the countries in the MENA region (IMF, 2002, 2010). This situation has prompted the supervisory authorities to find an adequate solution in order to alleviate the effect of this phenomenon which paralyzes both the banking system and notably the repayment capacity of economic agents that has deteriorated further in recent years.

In fact, the study of bad credits has drawn our interest and encouraged us to consider in-depth the phenomenon of credit risk allocation mainly during this transitional period.

Therefore, we attempt to study the determinants of corporate credit risk that contribute in increasing bank liabilities. To deal in depth with this problem, we have recourse to a theoretical approach. This study empirically raises questions about the determinants of corporate credit risk because, by referring to cash flows of a company, we find that those determinants are not sufficient to meet the contractually required payment.

In this study, we adopt a Logistic Regression model, as a predictive technique capable of identifying credit risk determinants of corporate credit service sector. According to Malley et al. (2012), this method has remained subject to various specific assumptions. For example, at the time when the most adopted variables and also the anticipated interactions are not correctly included in the model, then certain problems related to the model mis-specification may occur. In this context, these authors argued that the standard regression model cannot deal with multicollinearity between independent variables.

Knowing that there are many techniques for all types of data sets, the main goal of the current study is to evaluate determinants of credit risks in the corporate service sector in Tunisia. This initiative is realized through a case study, and the findings are obtained by using the parametric method of logistic regression. Hence, we empirically attempt to detect the corporates' payment defaults. In the present study, ours sample consists of 1461 companies in the service sector. In fact, literature revealed that the focus on these techniques, mainly in developed countries, has yielded a paramount insist on a variety of models consisting of certain methodological benefits as well as difficulties. However, in developing countries, the evidence on credit scoring achievements is rather limited (Altman et al., 1979 and Dinh and Kleimeier, 2007).

In this context and to our knowledge, our study is most likely the first to offer such a considerable analysis of firms' default payment. Hence, our objective is to identify the determinants of corporate credit risk that largely contribute in increasing bank liabilities.

To reach this end, we first selected the indicators that are collected from the Tunisian Central Bank. Moreover, we explained how to combine these indicators in a credit scoring model. Conducting a study on a sample of 1461 firms, we found that debt, solvency and profitability ratios, loan amount, and macroeconomic factors are the most statistically significant variables allowing for the prediction of default payment in the corporate credit service.

Indeed, though this study is restricted to identify and distinguish between good and bad firms of a sample taken from the Tunisian Central Bank, the current paper is an initiative aiming at reducing impaid credits. Thus, the identification of firms' types can be generalized to other samples through using more variables that can increase the accuracies of the used models.

Our study is structured as follows: the goal and background of this paper are presented in the first section. The second section presents the review of literature on methods used for the prediction of default risk. The third section displays the data sources, the research methodology and the analysis of the obtained results. Section four is devoted to the conclusion and recommendations.

2. Theoretical Background

A number of statistical and machine learning models have been adopted in order to predict credit risks and bankruptcies. One of the statistical techniques that can be used to detect firm's failure is the discriminant analysis (DA). According to Altman (1968) and Beaver (1966), this model is considered as the most frequently used method just before 1980. In addition, multiple discriminant analysis (MDA) was applied to develop the Z-Score model in order to predict bankruptcy (Altman (1968). Since this method has certain limitations, researchers have substituted it for logit or probit regression analysis and for linear conditional probability models.

Dimitras et al. (1996) have carried out a thorough review of statistical methods which predict firms default payment. They showed that logistic regression is amply used to predict the probability of firms' financial difficulties.

In the same vein, and in order to distinguish between good and bad borrowers, Abid et al. (2016) used logistic regression and discriminant analyses. These authors proved that the logistic regression model yields a better classification rate in predicting customer types compared to discriminant analysis.

Still within the framework of predicting default payment, the adopted methods changed from univariate to multivariate models by means of machine learning methods. In this context, Ravi Kumar and Ravi (2007) employed many types of intelligent techniques. For instance, the neural network (NN) method is considered as the most widely-used technique. besides, there are many other data mining techniques, such as decision trees (DT) (Frydman, et al., 1985), the genetic algorithms (Shin & Lee, 2002), the Kohonen map (du Jardin & Séverin, 2012), the simulation analysis(Cohen et al., 2012), and the **s**upport vector machines (SVM) (Gestel et al.,2006).According to Chen and Du (2009), the Neural Network (NN) is a nonlinear mathematical approach outperforming single classifiers when researchers attempt to test complex data.

Danenas & Garsva (2015) used the linear Support Vector Machine (SVM) to address the credit risk issue. This method solves the imbalanced classes problem and can be used for larger data sets. Authors proved that this method offers equivalent results with regard to the logistic regression and RBF network. In the same sphere, other authors, such as Bellotti and Crook (2009), compared the SVM technique with many well-known algorithms using credit card dataset. These authors found that the SVM technique is more successful in terms of classification of default customers. Furthermore, Lee et al. (2006) assumed that SVM outperform MLP in assessing credit scoring when using consumer credit data. This result is similar to the findings of Schebesch and Stecking (2005) who compare the SVM technique with the Logistic Regression model relying on a database of applicants for building and loan credits.

In order to predict financial distress in Chinese companies, Geng et al. (2015) compared different data mining techniques, such as neural network, Decision trees and Support Vector Machines. Authors showed that neural network technique is the most accurate technique. Bemš et al. (2015) used a modified magic square approach to predict a company's default. The favor of this technique is the graphical interpretation of a company score which is offered as a diagram consisting of an illustration of individual factor that can serve to the total score value.

Another stream of research relating to credit scoring and bankruptcy prediction problems has applied a set of neural networks. For instance, Duman et al. (2012) and Kruppa et al. (2013) advocated that decision trees, with their ensemble variations, notably random forests, are proven to be among other significant models in the field of credit risk and bankruptcy prediction. This new trend of soft computing that indicates a certain variation within a dataset proves its potential in having better classification performance in the context of credit risk (Hsieh & Hung, 2010).Similarly, West et al. (2005) adopted bagging and boosting using the Multi-Layer Perceptron (MLP) as the base classifier and concluded that, in most cases, the ensemble techniques were higher than the single best model. In the context of classifying bad loans, Desai et al. (1996) maintained that the MLP and the modular neural network reveal accurate results compared to logistic regression models.

With reliance on basic concepts of fuzzy logic and MLP neural networks, Khashei et al. (2013) implemented a model of a hybrid binary credit risk prediction. In this model, authors employed fuzzy numbers in an attempt to better modeling uncertainties and complexities in financial database.

On the other hand, Du Jardin and Séverin (2011) used a new method named Kohonen map in order to ameliorate the financial failure prediction accuracy of companies. In the context of creditscoring, Leong (2016) proved that the Bayesian Network (BN) technique presents more efficient results than logistic regression, SVM, decision trees, and multilayer Perceptron (MLP) for a large dataset. In Australian Banks, Sanford and Moosa (2015) proved that the BN could predict many risks, such as operational risk events, aggregate operational loss distributions, and Operational Value-at-Risk. According to Kao et al. (2012), a Bayesian latent variable model with classification and regression tree approach is adopted for banks to accurately predict applicants' performance and their repayment behavior. Moreover, Masmoudi et al. (2019) proposed a discrete Bayesian network with latent variable in order to evaluate the payment default probability of Tunisian households. This model makes it possible to evaluate the probability of default by dealing with a multi-class situation. In order to evaluate credit risk in Brazil, Sousa et al. (2016) employed a new dynamic modeling structure. This model is more efficient compared to the static modeling methods.

The hybrid associative memory, which is considered as another technique used in evaluating credit risk, has yielded more accurate results compared to artificial neural network techniques (ANN) as well as the SVM technique. These findings indicate that this technique, the hybrid associative memory, be appropriate for predicting financial distress, mainly for highly imbalanced data (Sánchez *et al.*, 2016). In the same vein, Mselmi et al. (2017) compared Logit model, Artificial Neural Networks, Support Vector Machine techniques, Partial Least Squares, and a hybrid model integrating Support Vector Machine with Partial Least Squares to expect the financial distress of French companies. Research in this field indicates that SVM is the best adopted technique that has yielded a rate of 85.7% accuracy for one year prior to financial distress. Conversely, and two years before the financial distress, the hybrid model outperforms the SVM, the Logit model, the Partial Least Squares, and the ANN obtain an accuracy rate of 94.28%.

Luo et al. (2017) compared the deep learning technique with traditional techniques, such as the Multinomial logistic regression (MLR), the MLP and the SVM to scan the performances of credit scoring models. Authors found out that deep belief networks give the best performance and overtake the other algorithms and the MLR has the worst performance.

To evaluate financial problems, Tavana et al. (2018) have currently used two methods among the most recent machine learning techniques: the ANNs and the BNs.This two-intelligent technique consists of several algorithms and tests that are used to analyse data and validate the proposed model. In fact, the ANN approach is used to detect the general trend of risk through the study of two main influential factors of liquidity risk whereas the BN technique is applied to identify the most influential factor reflecting that the liquidity risk occurs regardless of measuring all indicators. In the same context, Junhui Xu, Zekai Lu & Ying Xie (2021) used the machine learning methods (random forest (RF), the extreme gradient reinforcement tree, the gradient reinforcement model (GBM), and the neural network (NN)) to predict factors significantly affecting payment. The accuracy of all four methods is greater than 90%, and the RF is superior to the other classification models.

3. Data and Research Methodology

3.1 Data

The Tunisian Central Bank (TCB) and the World Development Indicators constituted the database of the current paper that investigates 1461 companies in the service sector over the period 2015–2018.

The choice of the variable to be explained and the explanatory variables is an essential step that precedes the use of the econometric model dealing with the determinants of credit risk.

The dependent variable

Credit risk: we measure the credit risk by a binary variable, which takes the value 1 in the event of non-repayment or default of the credit at maturity and with a value 0 if this was not the case. In the sample provided, 1081 applications (representing 64 %) were credit worthy, whereas 380 applications (26 %) were not.

Independent variables

Table 1 groups the main variables that we have adopted in the current study as well as the data sources. Their definitions and measurements are also detailed in the same Table.

The explanatory variables are ranked among the most proven and cited risk factors in the literature, including profitability, solvency, coverage, growth, activity, debt, size, liquidity and monetary position.

Table 1. Proposed variables

Variables	Measures	Source
Dependent variable		
Credit risk	1 if there is a payment default) 0 (otherwise)	(TCB)
Explanatory variables		
Debt ratio	Total debt / Total Assets	(TCB)
Solvency ratio	Equity/ Total liabilities	(TCB)
Liquidity ratio	The Working capital/ Total assets	(TCB)
Profitability ratio	Earnings Before Interest and Taxes / Total Assets	(TCB)
Loan amount	Log loan size	(TCB)
Age	Age of firm	(TCB)
Macroeconomic explanatory variables:		
Inflation	Inflation rate	World development indicators (WDI)
GDP	GDP growth rate	World development indicators (WDI)

Source: literature review

Table 2 presents the descriptive statistical analysis, including all variables. This Table displays descriptive statistics relating to our data.

Table 2. Dataset descriptive statistics

Variables	Mean	(std)	Min	Max
Credit risk	0.077	0.26	0	1
Loan amount	5.78	0.98	0.301	8.75
Age	17.507	13.11	0.4	96.5
Debt ratio	0.74	0.62	0.0031	17.96
Solvency ratio	1.29	9.24	-0.944	316.0443
Liquidity ratio	0.04	0.62	-45.588	1.006
Profitability ratio	-0.033	0.688	-17.06	0.98
Inflation	5.17	1.36	3.62	7.307
GDP growth rate	1.69	0.518	1.195	2.483

3.2 Research Methodology

Logistic Regression Model:

According to Yap et al. (2011), the main interest of the Logistic Regression credit scoring model is to specify the conditional probability of each application that belongs to no particular class.

The standard model is used to model the probability of default in the banking sector (Lessmann et al. (2015). This model is used in order to explain a dichotomous dependent variable by explanatory variables. The evaluation of “good" and "bad" firms depends on the values of the explanatory variables of the applicant’s credit.

The logistic regression model, a common modelling technique capable of classifying the applicants into two groups by using a set of predictive variables, aims at specifying the generation of the predicted values of the dependent variable that lies in the interval between zero and one. The logistic regression model is represented in Equation (1):

$$\ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots.. + \beta_n X_n + \varepsilon_t$$

Where:

$\ln \frac{pi}{(1-pi)}$: represents the default event yi

Pi: the probability of default of firm i.

X: independent variables

β: the coefficient of the independent variable

ε: error term

The estimation of the logit model follows the maximum likelihood method. Indeed, the Ordinary Least Squares (OLS) method is not adapted in this case for several reasons. Moreover, this approach is the most frequently used in literature.

The likelihood is given by the following equation:

$$L = \prod_{y_{i=1}} F(\beta_i X_i^*) \prod_{yi=0} (1 - F(\beta_i X_i^*))$$

The Table below shows the results obtained from the logistic regression model:

Results’ Analysis by using the LR Analysis

Table 4. Estimation results of the Logit model

	Coeff	Std.Err	P-Value
Loan size	0.387	0.075	0.000***
Age	0.046	0.005	0.000***
Debt ratio	1.33	0.248	0.000***
Solvency ratio	0.013	0.0045	0.004***
liquidityratio	-0.024	0.1	0.807
Profitability ratio	-1.39	0.232	0.000***
Inflation	0.33	0.155	0.033**
GDP growth rate	-0.93	0.407	0.021**
cons	-6.14	0.585	0.000

Note: ***, ** and * indicate statistical significance at 1%, 5% and 10% level respectively.

Source: Author's calculations

The sample used in the estimate includes 1461 companies in the service sector. The analysis of the econometric results reported in the Table above leads to the following conclusions. In general, the results of the estimations show that a large number of selected variables are significant.

The loan size is the amount of credit the applicant is granted. Many studies have employed the loan size as a predictor variable. The results are ambiguous, and thus no clear expectations can be deduced. In our study, we prove that the coefficient on the loan size variable is positive and has a significant sign. This is expected because a high amount of credit decreases the probability for the creditor to honour their commitment. These findings are consistent with the results of Jin et al. (2019) who predict that the higher the amount of credit, the higher the risk of default apprehended by borrowers.

The results, indicating that the variable of age is statistically significant, show that the company’s age has an effect on the probability of default payment. This finding corroborates with Han et al.’s (2018) results.We can also explain this result by the fact that the degraded economic situation in Tunisia has a negative effect on the financial situation of companies whatever their age. In this context, the economic and financial environment suffered four salient facts, namely the resurgence of terrorism and its immediate corollary, the shock on tourism, the deficiency of external demand from the main European partners, the stagnation of the crisis, and the Libyan and internal social tensions. These unfavourable factors have weighed heavily on economic growth, which reflects the sluggishness of activity in most sectors and mainly the service sector such as tourism and other services.

Debt ratio positively and significantly affects the probability of credit default in our model. This finding indicates that the higher a firm's debt; the less cash it has to cover its liabilities. The indebtedness of the firm seems, therefore, a determining factor of the credit risk of Tunisian firms in the service sector. This finding is in accordance with Buchdadi et al. (2020), Tulcanaza et al. (2019) and Woo et al (2020) who examined the effect of debt ratio on credit risk in the logistics and shipping industry.

As expected, the solvency ratio, which measures the creditworthiness of the company, positively affects the credit risk of companies. This result confirms the studies carried out by Altman (1968) who shows that the solvency of a company is among the most significant indicators for predicting the risk of default. Furthermore, we note that there is a negative but not significant relationship between the liquidity ratio and the credit risk. These results contradict the studies carried out by Altman (1986), Gathecha (2016) and Kristanti et al. (2016), indicating that liquidity have a positive and significant effect on the credit risk of companies.

Profitability is a primary factor in determining solvency and liquidity. In our case, we show that the coefficient associated with the profitability indicator is negatively and statistically significant. This result is inconsistent with those found in previous studies of Su-Han Wo et al, (2020) who proved that the debt ratio has no significant effect on payment default.

Our empirical findings confirm that macroeconomic variables play an important role in the dynamics of impaired credit in Tunisia. The results can be summarized as follows. First, the empirical evidence illustrates a negative and significant impact of GDP on firm default payment. The GDP growth rate explains the economic situation of a country for a specific period. In fact, the higher the GDP rate, the lower the rate of unpaid loans, which shows that in times of economic growth, the rate of unpaid loans decreases. These results are consistent with various studies carried out by Louzis et al. (2012) and Abid et al. (2014) to determine the relationship between unpaid credits and the economic situation. This negative relationship implies that an economic slowdown generates an increase in the level of unpaid loans, and vice versa since in the event of economic growth (high GDP growth rate) the level of income and business activity reverses this trend, which promotes the companies' ability to honour their commitments. Second, inflation rate can also explain the economic situation of country. A high inflation rate causes a decline in the turnover of companies and subsequently evokes the problem of repaying their debts. The results of our study show that the inflation rate has a positive and significant impact on the level of unpaid credit. This result is consistent with the study of Fofack (2005) and Abid et al (2014)that confirm the positive relationship between the rate of inflation and the level of bad debts of different economic agents.

4. Conclusion and Perspectives

Credit scoring systems are fundamental tools adopted to prevent bad debt loss. Literature on credit risk maintains that companies that are ascribed to default payment are generally treated using information about their previous payment behaviour.

In this study, to assess factors affecting company credit risk, we have recourse to the logistic regression model. Thus, we have used a sample of 1461 companies that have been granted loans. Among this sample, 380 companies had a default up to 90 days, whereas 1081 had a clear history without any default. The results obtained from the estimation of a logit model allow us to conclude that the probability of firms' default is dependent not only on economic conditions (an increase in the inflation rate and a decrease in the GDP growth rate), but also on microeconomic variables (profitability, indebtedness, solvency) that increase this probability. This second piece of information is important in terms of the credit policy that banks should put at work for firms.

The ultimate goal of this part is to propose a set of recommendations to the detriment of many results that are the fruit of an analysis process, and whose goal is to reduce companies' default. The results of our study lead to interesting recommendations for the management and supervisory bodies of banks, which, in turn, can help to take the necessary measures in order to better control the default rate.

To better control the evolution of credits, Tunisian Central Bank (TCB) undertook certain measures to counteract the "alarming" allocation of credits. In this regard, it has tightened its prudential regulations in an unprecedented way by revising three fundamental regulations that create several direct restrictions on the ability of banks to finance the economy, such as the introduction of new requirements for the provision of classified claims, a higher minimum regulatory capital ratio and a more restrictive liquidity regulation whether in the short or long terms.

Despite the theoretical and empirical contributions and findings, this study has some limitations that may offer extensions and perspectives for future research. Since the results show that microeconomic factors affect the probability of firm failure, it is important to further analyse the other determinants of this risk. Specifically, certain firm characteristics reflecting a typical company profile, such as those relating to the firm, innovation capacity, strategic objectives, etc. which may influence the rate of firm failure could be further investigated.

However, one of the limitations of the empirical analysis conducted in our study is that the sample size is relatively small. It would undoubtedly be interesting to take into account other sectors of activity (industry, agriculture, and fishing) in order to better study various factors that have an impact on the credit risk of firms in Tunisia. We can also fructify our work by making an evaluation of the coronavirus effect, which has amply impacted the economic activity of Tunisian firms.

References

Abid, L., Masmoudi, A., & Zouari-Ghorbel, S. (2016). The consumer loans payment default predictive model: An application of the logistic regression and the discriminant analysis in a Tunisian Commercial Bank. *Journal of the Knowledge Economy*, 1-15. https://doi.org/10.1007/s13132-016-0382-8

Abid, L., Ouertani, M. N., & Zouari-Ghorbel, S. (2014). Macroeconomic and Bank-Specific Determinants of Household's Non-Performing Loans in Tunisia: a Dynamic Panel Data. *Procedia Economics and Finance, 13*, 58-68. https://doi.org/10.1016/S2212-5671(14)00430-4

Akkoc, S. (2012). An empirical comparison of conventional techniques, neural networks and the three stage hybrid adaptive neuro fuzzy inference system (ANFIS) model for credit scoring analysis: The case of Turkish credit card data. *European Journal of Operational Research, 222*(1), 168-178. https://doi.org/10.1016/j.ejor.2012.04.009

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The*

Journal of Finance, 23(4), 589-609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Beaver, W. H. (1966). Fnancial ratios as predictors of failure, Empirical Research in Accounting: Selected Studies, *Journal of Accounting Research, Supplement to Vol. 5*, 179-199. https://doi.org/10.2307/2490171

Beaver, W. H. (1968). Market prices, financial ratios and the prediction of failure. *Journal of Accounting Research*, Autumn, 179-192. https://doi.org/10.2307/2490233

Bellotti, T., & Crook, J. (2009). Support vector machines for credit scoring and discovery of significant features. *Expert Systems with Applications, 36*(2), 3302-3308. https://doi.org/10.1016/j.eswa.2008.01.005

Bemš, J., Star`y, O., Macaš, M., Žegklitz, J., & Pošík, P. (2015). Innovative default pre- diction approach. *Expert Systems with Applications, 42*(17-18), 6277-6285. https://doi.org/10.1016/j.eswa.2015.04.053

Buchdadi, A., Nguyen, X., Putra, F., & Dalimunthe, S. (2020). The effect of credit risk and capital adequacy on financial distress in rural banks. *Accounting, 6*(6), 967-974. https://doi.org/10.5267/j.ac.2020.7.023

Carlos, S.-C. (1996). Self-organizing neural networks for financial diagnosis. *Decision Support Systems, 17*(3), 227-238. https://doi.org/10.1016/0167-9236(95)00033-X

Cleofas-Sánchez, L., García, V., Marqués, A., & Sánchez, J. S. (2016). Financial distress prediction using the hybrid associative memory with translation. *Applied Soft Computing, 44*, 144-152. https://doi.org/10.1016/j.asoc.2016.04.005

Cohen, S., Doumpos, M., Neofytou, E., & Zopounidis, C. (2012). Assessing financial distress where bankruptcy is not an option: An alternative approach for local municipalities. *European Journal of Operational Research, 218*(1), 270-279. https://doi.org/10.1016/j.ejor.2011.10.021

Danenas, P., & Garsva, G. (2015). Selection of support vector machines based classi- fiers for credit risk domain. *Expert Systems with Applications, 42*(6), 3194-3204. https://doi.org/10.1016/j.eswa.2014.12.001

Desai, V., Crook, J., & Overstreet, G. (1996). A Comparison of Neural Networks and Linear Scoring Models in the Credit Union Environment. *European Computer Journal of Operational Research, 95*(6), 24-37. https://doi.org/10.1016/0377-2217(95)00246-4

Dimitras, A. I., Zanakis, S. H., & Zopounidis, C. (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research, 90,* 487-513. https://doi.org/10.1016/0377-2217(95)00070-4

Du Jardin, P., & Séverin, E. (2011). Predicting corporate bankruptcy using a self-orga- nizing map: An empirical study to improve the forecasting horizon of a financial failure model. *Decision Support Systems, 51*(3), 701-711. https://doi.org/10.1016/j.dss.2011.04.001

Du Jardin, P., & Séverin, E. (2012). Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time. *European Journal of Operational Research, 221*(2), 378-396. https://doi.org/10.1016/j.ejor.2012.04.006

Duman, E., Ekinci, Y., & Tanrıverdi, A. (2012). Comparing alternative classifiers for database marketing: The case of imbalanced datasets. *Expert Systems with Applications, 39*(1), 48-53. https://doi.org/10.1016/j.eswa.2011.06.048

Fletcher, D., & Goss, E. (1993). Forecasting with neural networks: An application using bankruptcy data. *Information & Management, 24*(3), 159-167. https://doi.org/10.1016/0378-7206(93)90064-Z

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://doi.org/10.1111/j.1540-6261.1985.tb04949.x

Gathecha, J. W. (2016). Effect of firm characteristics on financial distress of non-financial listed firms at Nairobi Securities Exchange. Doctoral dissertation, *Kenyatta University.*

Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research, 241*(1), 236-247. https://doi.org/10.1016/j.ejor.2014.08.016

Gestel, T. V., Baesens, B., Suykens, J. A. K., Van den Poel, D., Baestaens, D.-E., & Willekens, M. (2006). Bayesian kernel based classification for financial distress detection. *European Journal of Operational Research, 172*(3), 979-1003. https://doi.org/10.1016/j.ejor.2004.11.009

Han, J. T., Chen, Q., Liu, J. G., Luo, X. L., & Fan, W. (2018). The persuasion of borrowers' voluntary information in peer to peer lending: An empirical study based on elaboration likelihood model. *Computers in Human Behavior, 78*, 200-14. https://doi.org/10.1016/j.chb.2017.09.004

Hsieh, N.-C., & Hung, L.-P. (2010). A data driven ensemble classifier for credit scoring analysis. *Expert Systems with Applications, 37*(1), 534-545. https://doi.org/10.1016/j.eswa.2009.05.059

Jin, J., Shang, Q., & Ma, Q. (2019). The role of appearance attractiveness and loan amount in peer-to-peer lending: Evidence from event-related potentials. *Neuroscience Letters, 692*, 10-15. https://doi.org/10.1016/j.neulet.2018.10.052

Kao, L. J., Chiu, C. C., & Chiu, F. Y. (2012). A Bayesian latent variable model with classification and regression tree approach for behaviour and credit scoring. *Knowledge-Based Systems, 36,* 245-252. https://doi.org/10.1016/j.knosys.2012.07.004

Khashei, M., Rezvan, M. T., Hamadani, A. Z., & Bijari, M. (2013). A bi-level neural-based fuzzy classification approach for credit scoring problems. *Complexity, 18,* 46-57. https://doi.org/10.1002/cplx.21458

Kristanti, F. T., Rahayu, S., & Huda, A. N. (2016). The determinant of financial distress on Indonesian family firm. *Procedia-Social and Behavioral Sciences, 219*, 440-447. https://doi.org/10.1016/j.sbspro.2016.05.018

Kruppa, J., Schwarz, A., Arminger, G., & Ziegler, A. (2013). Consumer credit risk: Individual probability estimates using machine learning. *Expert Systems with Applications, 40*(13), 5125-5131. https://doi.org/10.1016/j.eswa.2013.03.019

Lee, T.-S., Chiu, C.-C., Chou, Y.-C., & Lu, C.-J. (2006). Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics & Data Analysis, 50*, 1113-1130. https://doi.org/10.1016/j.csda.2004.11.006

Leong, C. K. (2016). Credit risk scoring with Bayesian network models. *Computational Economics, 47*(3), 423-446. https://doi.org/10.1007/s10614-015-9505-8

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking stateof-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research, 247*, 124-136. https://doi.org/10.1016/j.ejor.2015.05.030

Li, Z., Crook, J., & Andreeva, G. (2013). Chinese companies distress prediction: An application of data envelopment analysis. *Journal of the Operational Research Society, 65*(3), 466-479. https://doi.org/10.1057/jors.2013.67

Lu, Z.-K., & Xie, Y. (2021). Loan default prediction of Chinese P2P market: a machine learning methodology *Scientific Report.*

Luo, C., Wu, D., & Wu, D. (2017). A deep learning approach for credit scoring using credit default swaps. *Engineering Applications of Artificial Intelligence, 65,* 465-470. https://doi.org/10.1016/j.engappai.2016.12.002

Masmoudi, K. *et al.*. (2019). Credit risk modeling using Bayesian network with a latent variable. *Expert Systems With Applications, 127,* 157-166. https://doi.org/10.1016/j.eswa.2019.03.014

Mselmi, N., Lahiani, A., & Hamza, T. (2017). Financial distress prediction: The case of French small and medium-sized firms. *International Review of Financial Analysis, 50*, 67-80. https://doi.org/10.1016/j.irfa.2017.02.004

Pope, D. G., & Sydnor, J. R. (2011). What's in a picture? Evidence of discrimination from prosper.com. *The Journal of Human Resources, 46*(1), 53-92. https://doi.org/10.1353/jhr.2011.0025

Ravi Kumar, P., & 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

Sanford, A., & Moosa, I. (2015). Operational risk modelling and organizational learning in structured finance operations: A bayesian network approach. *Journal of the Operational Research Society, 66*(1), 86-115. https://doi.org/10.1057/jors.2013.49

Schebesch, K. B., & Stecking R. (2005). Support vector machines for classifying and describing credit applicants: detecting typical and critical regions. *Journal of the Operational Research Society, 56*(9), 1082-1088. https://doi.org/10.1057/palgrave.jors.2602023

Shin, K. S., & Lee, Y. J. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert*

Systems with Applications, 23(3), 321-328. https://doi.org/10.1016/S0957-4174(02)00051-9

Sousa, M. R., Gama, J., & Brandão, E. (2016). A new dynamic modeling framework for credit risk assessment. *Expert Systems with Applications, 45*, 341-351. https://doi.org/10.1016/j.eswa.2015.09.055

Tavana, M., Abtahi, A.-R., Di Caprio, D., & Poortarigh, M. (2018). An artificial neural network and Bayesian network model for liquidity risk assessment in banking. *Neurocomputing, 275*, 2525-2554. https://doi.org/10.1016/j.neucom.2017.11.034

Tulcanaza Prieto, A. B., & Lee, Y. H. (2019). Internal and external determinants of capital structure in large Korean firms. *Global Business & Finance Review, 24*(3), 79-96. https://doi.org/10.17549/gbfr.2019.24.3.79

West, D., Dellana, S., & Qian, J. (2005). Neural networks ensemble strategies for financial decision applications. *Comp. Operat. Res., 32*, 2543-2559. https://doi.org/10.1016/j.cor.2004.03.017

Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. Decision Support Systems, 11(5), 545-557. https://doi.org/10.1016/0167-9236(94)90024-8

Woo, S. H., Kwon, M. S., & Yuen, K. F. (2020). Financial determinants of credit risk in the logistics and shipping industries. *Maritime Economics & Logistics, 3*(1), 1-23. https://doi.org/10.1057/s41278-020-00157-4

Woo, S-H., Kwon, M.-S., & Yuen, K. F. (2020). Financial determinants of credit risk in the logistics and shipping industries. *Maritime Economics & Logistics, 23*, 268-290. https://doi.org/10.1057/s41278-020-00157-4

Yap, P., Ong, S., & Husain, N. (2011). Using data mining to improve assessment of credit worthiness via credit scoring models. *Exp. Syst.*, *38*(10), 1374-1383. https://doi.org/10.1016/j.eswa.2011.04.147