Quantitative Finance and Economics

QFE, 6(3): 405–432.
DOI: 10.3934/QFE.2022018
Received: 05 April 2022
Revised: 22 June 2022
Accepted: 29 June 2022
Published: 11 July 2022

http://www.aimspress.com/journal/QFE

Research article

Impact of macroeconomic indicators on bankruptcy prediction models: Case of the Portuguese construction sector

Ana Sousa[1], Ana Braga[2] and Jorge Cunha[2,*]

[1] Department of Production and Systems Engineering, School of Engineering, University of Minho, Braga, Portugal
[2] ALGORITMI Research Centre, University of Minho, Braga, Portugal

* **Correspondence:** Email: jscunha@dps.uminho.pt.

Abstract: The importance of macroeconomic indicators on the performance of bankruptcy prediction models has been a contentious issue, due in part to a lack of empirical evidence. Most indicators are primarily centered around a company's internal environment, overlooking the impact of the economic cycle on the status of the company. This research brings awareness about the combination of microeconomic and macroeconomic factors. To do this, a new model based on logistic regression was combined with principal component analysis to determine the indicators that best explained the variations in the dataset studied. The sample used comprised data from 1,832 Portuguese construction companies from 2009 to 2019. The empirical results demonstrated an average accuracy rate of 90% up until three years before the bankruptcy. The microeconomic indicators with statistical significance fell within the category of liquidity ratios, solvency and financial autonomy ratios. Regarding the macroeconomic indicators, the gross domestic product and birth rate of enterprises proved to increase the accuracy of bankruptcy prediction more than using only microeconomic factors. A practical implication of the results obtained is that construction companies, as well as investors, government agencies and banks, can use the suggested model as a decision-support system. Furthermore, consistent use can lead to an effective method of preventing bankruptcy by spotting early warning indicators.

Keywords: bankruptcy; construction sector; logistic regression; macroeconomic indicators; microeconomic indicators; principal component analysis; ROC curve

JEL Codes: C51, C52, C81, C82, G33

Abbreviations: ACC: accuracy; AUC: area under the curve; CAE: statistical classification of economic activities; CLT: central limit theorem; FPR: false positive rate; GA: genetic algorithm; GDP: gross domestic product; MDA: multivariate discriminant analysis; n.a.: non-available; n.s.: non-significant; NN: neural networks; PCA: principal component analysis; ROC: receiver operating characteristic; stepAIC: stepwise Akaike information criterion; SVM: support vector machine; TPR: true positive rate; VIF: variance inflation factor.

1. Introduction

Since the global financial crisis of 2008, the research on bankruptcy prediction has received considerable attention because of the negative economic impact on most sectors and the survival of many companies (Shi & Li, 2019). Therefore, and considering the recent COVID-19 pandemic as well, the ability to identify early warning signals is critical for companies, investors, government agencies and banks (Kapliński, 2008; Perboli & Arabnezhad, 2021).

The construction sector remains one of the most difficult to study, owing to its variations depending on economic cycle fluctuations (European Commission, 2021, October; dos Santos & Silva, 2019; Tserng et al., 2012) and the uncertainty that each project entails due to its uniqueness and duration (Choi et al., 2018; Horta & Camanho, 2013). In Portugal, the construction sector had one of the highest bankruptcy rates among all sectors between 2008 and 2011 (PORDATA, 2022). Therefore, the search for tools capable of predicting future circumstances has become undeniably important in helping managers to prevent further decline or eventual bankruptcy by implementing the necessary measures in advance (Koksal & Arditi, 2004; Tinoco & Wilson, 2013).

The precise concept of bankruptcy is still unknown (Shi & Li, 2019). Thus, comparing the various studies that have been developed is exceedingly challenging (Bellovary et al., 2007). Altman (1968) and Ohlson (1980), two of the most well-known authors on this particular topic, used an exclusively legal definition, characterizing all companies that are legally in this situation as bankrupt. More recently, Altman & Hotchkiss (2006) highlighted four commonly found terminologies in the literature: failure, insolvency, default and bankruptcy. In the opinion of the authors, these terms are commonly used interchangeably, but their formal meaning is unique. In addition to these terms, there are studies focused on the analysis of companies in financial distress, which allow early warning signals to be detected. However, this field is still underdeveloped, as objectively defining the point at which a company enters financial distress is difficult (Tinoco et al., 2018; Mselmi et al., 2017; Platt & Platt, 2002).

According to Bellovary et al. (2007), several authors have attempted to study and have developed various models to solve this problem, which is the prediction of bankruptcies. Early studies used statistical models based on techniques such as univariate analysis (Beaver, 1966), multivariate discriminant analysis (Altman, 1968; Taffler, 1984) and logistic regression (Ohlson, 1980; Zavgren, 1985). To overcome the limitations of the traditional models, recent studies have brought new artificial intelligence alternatives, including neural networks (Boritz & Kennedy, 1995; Ben Jabeur et al., 2022; Lee & Choi, 2013; Pompe & Bilderbeek, 2005), genetic algorithms (Etemadi et al., 2009; Uthayakumar et al., 2020; Wu et al., 2007) and support vector machines (Barboza et al., 2017; Lessmann et al., 2015; Min et al., 2006).

Financial ratios have been used as key indicators of business failure in bankruptcy prediction models for a long time (Altman, 1968; Beaver et al., 2005). Overall, prior research indicates that financial ratios reflect a high ability to predict bankruptcy (e.g., Altman, 1968; Beaver, 1966; Ohlson, 1980).

However, as reported by Beaver et al. (2005) and Tinoco & Wilson (2013), combining financial ratios with other types of variables can act as a complement to the models and increase their predictive power. In this regard, several studies, such as those conducted by Acosta-González et al. (2019), Altman (1983), ben Jabeur et al. (2021), Liu (2004), Platt & Platt (1994) and Tinoco & Wilson (2013), have proved the significance of macroeconomic variables in predicting the risk of company bankruptcy. Companies have faced various crises in recent decades that historical models have been unable to predict. The unanticipated COVID-19 pandemic, which has been impacting the whole world since 2019 and prompting some businesses to shut down or activate contingency measures, is a recent example. According to Boratyńska (2016) and Habib et al. (2020), corporate bankruptcy occurs numerous times during an economic crisis because of macroeconomic issues, mostly owing to a reduction in sales.

Considering these recent developments in the literature, the main purpose of this study was to shed additional light on this debate by proposing a model that predicts the bankruptcy of Portuguese companies in the construction sector. While previous researchers have primarily focused on microeconomic indicators, the studies examining the relationship between microeconomic and macroeconomic indicators are limited. An innovative approach with principal component analysis (PCA) and logistic regression models were used to begin filling this information gap (Mbaluka et al., 2022; Sulaiman et al., 2021). This technique is feasible to prevent the consequences of multicollinearity and high variance inflation factor (VIF) values in a modeling process by reducing the influence of highly correlated independent variables a priori in a regression problem (Daoud, 2017; Liu et al., 2003). Furthermore, its practical application can improve the decision-making process for credit risks. For this reason, the following research question was formulated: Can macroeconomic indicators improve the performance of bankruptcy prediction models, besides the inclusion of microeconomic indicators?

The remainder of the paper is organized as follows. Section 2 describes the methodology used in the work, including the sample and data collection techniques, definition of the dependent variable, independent variables chosen for the model, and statistical analyses required for the development of the proposed model. Section 3 presents the empirical results, which include a description of the dataset, the variables chosen for the models and their performance. Section 4 discusses the importance of microeconomic and macroeconomic variables in bankruptcy prediction models, as well as the development of the selected variables over time. Finally, the conclusions and suggestions for future research are described in Section 5.

2. Materials and methods

2.1. Sample and data collection

The microeconomic data were collected from Orbis[1], a powerful database specializing in financial information from the Bureau van Dijk, and the macroeconomic data were collected from PORDATA[2], a database that gathers data on multiple aspects of society in Portugal.

During the microeconomic data preprocessing, the sample data were split into two types of companies: active ones (those that are operating normally) and bankrupt ones. For the bankrupt

[1]https://www.bvdinfo.com/en-gb/our-products/data/international/orbis

[2]https://www.pordata.pt/

companies, the period covered by the observations in the dataset ranged from 2009 to 2019, while the year 2019 was chosen for the active companies. All Portuguese companies with a statistical classification of economic activities (CAE) related to the construction sector, that is, CAE 41, 42 and 43, were considered. In addition to the previously mentioned characteristics, the following criteria were applied: (i) the microeconomic information was available for at least three consecutive years, (ii) the companies' assets were greater than zero and (iii) the companies with blank indicators or classed as non-available or non-significant are excluded. Therefore, the sample data contained 1,832 companies, 274 of which were bankrupt and 1,558 of which were active.

2.2. Variable selection

2.2.1. Dependent variable

A central issue in this study was how to classify a company as bankrupt since there is no clear-cut definition of bankruptcy that allows the researcher to easily classify them (Tinoco & Wilson, 2013). The dependent variable is binary, corresponding to two states: the company is bankrupt (coded as 1) or is not (coded as 0). Therefore, as stated in Section 1, while the term bankruptcy is commonly used to classify companies, it should be noted that this term is used in a broad sense in this paper. That is, this word encompasses not only the companies that have blatantly gone bankrupt, but also all of those that may have been facing severe financial difficulties and, as a result, a high risk of bankruptcy (or entering insolvency proceedings). To identify whether a company should be included in the group of bankrupted companies, we used information from the ORBIS database regarding a company's status. In this sense, we considered the companies to be bankrupt if they changed their status from active to one of the following statuses: active (default of payment), active (insolvency proceedings), bankruptcy or dissolved (bankruptcy).

As in the study performed by Tinoco & Wilson (2013), the classification of a company as being bankrupt adopted in this study has the advantage of not being contingent upon the legal definition of bankruptcy. That is, the group of companies classified as bankrupt includes all those companies in a difficult financial situation (financially distressed companies), regardless of whether they were considered bankrupt from a legal point of view.

2.2.2. Independent variables

A total of 26 independent variables that should be perceived as possible predictors of bankruptcy were considered in this study based on the following criteria: (i) popularity in the literature, (ii) statistical significance, (iii) the number of times it is cited in the literature, (iv) relevance to the issue under research and (v) the availability in the previously mentioned datasets. All of the variables analyzed were quantitative, referring to various microeconomic and macroeconomic indicators. According to Neves (2012), the microeconomic indicators are often expressed in categories since their employability is dependent on factors such as their aim or the information they give. As a result, the ratios were categorized as follows: liquidity ratios, profitability ratios, activity ratios, financial risk and leverage ratios, solvency and financial autonomy ratios and other internal ratios (Table 1). In addition to the 20 microeconomic indicators, six macroeconomic variables were chosen (Table 2).

Table 1. Potential microeconomic variables for prediction of bankruptcy.

Category	Variable	Description	Source	Sign
Liquidity ratios	LIQG	Current assets/current liabilities	Acosta-González et al. (2019); Beaver (1966); Chen et al. (2016); Cheng & Hoang (2015); Choi et al. (2018); Correia (2012); Costa (2014); Giriūniene et al. (2019); Heo & Yang (2014); Horta & Camanho (2013); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Režňáková (2017c); Karas & Srbová (2019); Lagesh et al. (2018); Obradović et al. (2018); da Pimenta (2015); da Rosa (2017); Shumway (2001); Silva (2014); Ng et al. (2011); Tserng et al. (2014); Tserng et al. (2012); Vieira et al. (2013); Yan et al. (2020); Zoričák et al. (2020)	-
	LIQR	Current assets-inventory/ current liabilities	Acosta-González et al. (2019); Chen et al. (2016); Cheng & Hoang (2015); Choi et al. (2018); Obradović et al. (2018); da Rosa (2017); Ng et al. (2011); Tserng et al. (2014); Tserng et al. (2012); Yan et al. (2020); Zoričák et al. (2020)	-
	FMAT	Working capital/total assets	Acosta-González et al. (2019); Altman (1968); Beaver (1966); Carvalho et al. (2020); Cheng & Hoang (2015); Choi et al. (2018); Correia (2012); Costa (2014); Giriūniene et al. (2019); Heo & Yang (2014); Horta & Camanho (2013); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Režňáková (2017c); Karas & Srbová (2019); Nouri & Soltani (2016); Obradović et al. (2018); Ohlson (1980); Oliveira (2014); Pacheco et al. (2019); Pham Vo Ninh et al. (2018); da Pimenta (2015); da Rosa (2017); Sánchez-Lasheras et al. (2012); Shumway (2001); Ng et al. (2011); Tserng et al. (2014); Vieira et al. (2013); Vo et al. (2019)	-
	FMV	Working capital/sales	Beaver (1966); Heo & Yang (2014); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Režňáková (2017c); Karas & Srbová (2019)	-
	FCAT	Cash flow/total assets	Acosta-González et al. (2019); Beaver (1966); Correia (2012); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Srbová (2019); Vieira et al. (2013)	-

Continued on next page

Category	Variable	Description	Source	Sign
Profitability ratios	ROE	Net income/shareholder funds	Choi et al. (2018); Correia (2012); Horta & Camanho (2013); Karas & Režňáková (2017c); Karas & Srbová (2019); Nouri & Soltani (2016); Pacheco et al. (2019); da Pimenta (2015); da Rosa (2017); Silva (2014); Ng et al. (2011); Tserng et al. (2014); Tserng et al. (2012); Vieira et al. (2013); Zoričák et al. (2020)	-
	EBITAT	Earnings before interest and taxes/ total assets	Acosta-González et al. (2019); Altman (1968); Carvalho et al. (2020); Heo & Yang (2014); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Režňáková (2017c); Karas & Srbová (2019); Lucanera et al. (2020); Oliveira (2014); Pacheco et al. (2019); Pham Vo Ninh et al. (2018); da Pimenta (2015); da Rosa (2017); Sánchez-Lasheras et al. (2012); Shumway (2001); Vo et al. (2019); Yan et al. (2020)	-
	RLV	Net income/sales	Beaver (1966); Cheng & Hoang (2015); Choi et al. (2018); Correia (2012); Costa (2014); Horta & Camanho (2013); da Rosa (2017); Silva (2014); Tserng et al. (2014); Tserng et al. (2012); Vieira et al. (2013); Yan et al. (2020); Zoričák et al. (2020)	-
	ROA	Net income/total assets	Acosta-González et al. (2019); Beaver (1966); Carvalho et al. (2020); Chen et al. (2016); Cheng & Hoang (2015); Choi et al. (2018); Correia (2012); Costa (2014); Horta & Camanho (2013); Jones & Wang (2019); Karas & Srbová (2019); Ohlson (1980); Oliveira (2014); Pacheco et al. (2019); da Pimenta (2015); da Rosa (2017); Shumway (2001); Silva (2014); Ng et al. (2011); Tserng et al. (2014); Tserng et al. (2012); Vieira et al. (2013); Yan et al. (2020); Zoričák et al. (2020)	-
	EBITDAV	Earnings before interest, taxes, depreciation and amortization/sales	Karas & Režňáková (2017a); Karas & Režňáková (2017b); da Rosa (2017)	-
Activity ratios	VAT	Sales/total assets	Acosta-González et al. (2019); Altman (1968); Carvalho et al. (2020); Cheng & Hoang (2015); Choi et al. (2018); Correia (2012); Costa (2014); Giriūniene et al. (2019); Heo & Yang (2014); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Režňáková (2017c); Karas & Srbová (2019); Nouri & Soltani (2016); Oliveira (2014); da Pimenta (2015); da Rosa (2017); Sánchez-Lasheras et al. (2012); Shumway (2001); Silva (2014); Tserng et al. (2014); Tserng et al. (2012); Vieira et al. (2013); Yan et al. (2020); Zoričák et al. (2020)	-
	ACV	Current assets/sales	Acosta-González et al. (2019); Beaver (1966); Karas & Srbová (2019); da Pimenta (2015); da Rosa (2017)	+

Continued on next page

Category	Variable	Description	Source	Sign
Financial risk and leverage ratios	PTAT	Total liabilities/total assets	Acosta-González et al. (2019); Carvalho et al. (2020); Chen et al. (2016); Choi et al. (2018); Correia (2012); Costa (2014); Karas & Srbová (2019); Nouri & Soltani (2016); Ohlson (1980); Oliveira (2014); Pacheco et al. (2019); da Pimenta (2015); da Rosa (2017); Shumway (2001); Tinoco & Wilson (2013); Ng et al. (2011); Tserng et al. (2014); Tserng et al. (2012); Vieira et al. (2013); Yan et al. (2020); Zoričák et al. (2020)	+
	PTCP	Total liabilities/shareholder funds	Acosta-González et al. (2019); Cheng & Hoang (2015); Choi et al. (2018); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Režňáková (2017c); Obradović et al. (2018); da Pimenta (2015); da Rosa (2017); Silva (2014); Tserng et al. (2014); Tserng et al. (2012); Yan et al. (2020); Zoričák et al. (2020)	+
	EBITDAJS	EBITDA/interest expenses	Costa (2014); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Režňáková (2017c); Karas & Srbová (2019); da Pimenta (2015); da Rosa (2017); Silva (2014); Tinoco & Wilson (2013)	-
Solvency and financial autonomy ratios	SOLV	Shareholder funds/total assets	Jones & Wang (2019); Karminsky & Burekhin (2019); Obradović et al. (2018); Oliveira (2014); Pacheco et al. (2019); da Pimenta (2015); da Rosa (2017)	-
	PCAT	Current liabilities/total assets	Acosta-González et al. (2019); Beaver (1966); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Režňáková (2017c); da Rosa (2017); Yan et al. (2020)	+
	AFAT	Fixed assets/total assets	Acosta-González et al. (2019); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karas & Režňáková (2017c); Lucanera et al. (2020); Ng et al. (2011); Yan et al. (2020)	-
Other internal ratios	DIM	$log_{(Totalassets)}$	Acosta-González et al. (2019); Carvalho et al. (2020); Heo & Yang (2014); Horta & Camanho (2013); Karas & Režňáková (2017a); Karas & Režňáková (2017b); Karminsky & Burekhin (2019); Lucanera et al. (2020); Oliveira (2014); da Rosa (2017); Ng et al. (2011)	-
	NUMC	Number of employees	Oliveira (2014); Pacheco et al. (2019); da Rosa (2017)	-

Note: The sign "-" indicates that the greater the value of this variable, the lower the likelihood of bankruptcy. The "+" sign indicates that the greater the value of this variable, the greater the likelihood of bankruptcy. The acronyms of the variables were kept in Portuguese.

Table 2. Potential macroeconomic variables for prediction of bankruptcy.

Variable	Description	Source	Sign
TJ	Interest rate (annual average)	Acosta-González et al. (2019); Carvalho et al. (2020); Hudson (1986); Liu (2004); Nouri & Soltani (2016); Oliveira (2014); Young (1995); Živković (2016)	+
PIB	GDP annual growth	Carvalho et al. (2020); Giriūniene et al. (2019); Jones & Wang (2019); Oliveira (2014); Yan et al. (2020)	-
TI	Inflation rate (annual average)	Acosta-González et al. (2019); Carvalho et al. (2020); Giriūniene et al. (2019); Jones & Wang (2019); Nouri & Soltani (2016); Oliveira (2014); Pham Vo Ninh et al. (2018); Vo et al. (2019); Yan et al. (2020)	+
TD	Unemployment rate (annual average)	Acosta-González et al. (2019); Carvalho et al. (2020); Giriūniene et al. (2019); Jones & Wang (2019); Yan et al. (2020); Živković (2016)	-
TNE	Birth rate of enterprises (annual average)[3]	Altman (1983); Cuthbertson & Hudson (1996); Hudson (1986); Liu (2004); Platt & Platt (1994); Young (1995)	+
TJCH	Interest rate on mortgages (annual average)	Pacheco et al. (2019); da Rosa (2017)	+

Note: The sign "-" indicates that the greater the value of this variable, the lower the likelihood of bankruptcy. The "+" sign indicates that the greater the value of this variable, the greater the likelihood of bankruptcy. The acronyms of the variables were kept in Portuguese.

2.3. Data analysis

All statistical analyses implemented to develop the models were carried out using R statistical software (4.0.3) and RStudio version 4.0.3 (Figure 1). It is crucial to highlight that the methodology utilized in this study involves estimating three models by utilizing the same set of independent variables across three years before the bankruptcy.

[3]According to the definition of the OECD, the birth rate of enterprises corresponds to the number of enterprise births in the reference period (t) divided by the number of enterprises active in t (OECD Statistics, 2022).

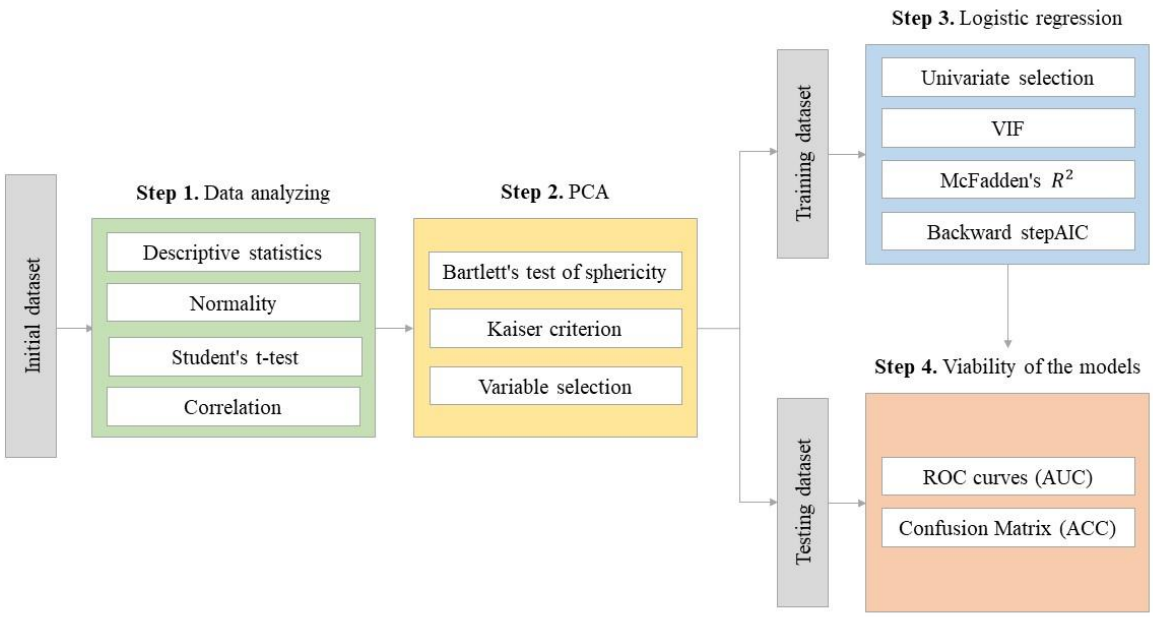

Figure 1. Flowchart of the proposed statistical methodology.

Step 1. The dataset was analyzed using basic descriptive statistics such as means, medians, standard deviations, minimums and maximums (Murphy, 2021). Furthermore, the normality assumption was checked by performing a Kolmogorov-Smirnov test (Kwak & Park, 2019), and the homogeneity of variance was tested using the *var.test* function (Kim & Cribbie, 2018). Independent sample t-tests were applied with a significance level of 0.05. The purpose of this test was to determine if there was a difference between the active companies and bankrupt companies (Wellek & Blettner, 2012). Even though the Kolmogorov-Smirnov test results showed that none of the variables follow the normal distribution, we decided to use a parametric test for the following reasons: (i) according to the central limit theorem, if the sample size is greater than 30, no nonparametric test is required (Kwak & Kim, 2017) and (ii) non-parametric tests have less statistical power than parametric tests (Kwak & Kim, 2017; Lydersen, 2015). To determine the correlation between the independent variables, the *cor* function was employed. It was determined that the variables were highly correlated if they had a value greater than $+0.70$ or less than -0.70 (Asuero et al., 2006).

Step 2. The PCA method was used to reduce the dimensions of the dataset while retaining as much information as possible (Lafi & Kaneene, 1992). The sample data adequacy were determined using Bartlett's sphericity test (Wu, 2021). To perform the PCA on the R software platform, the *prcomp* function was used. Furthermore, the principal components required for the model were investigated by using the three most common approaches: the proportion of variance explained criterion, the eigenvalue criterion and the scree plot criterion (Gotts et al., 2020; Wood et al., 2021). The *fviz contrib* function was employed to choose the variables that contribute the most to each of the components. As a result, all variables above the red line were selected.

Step 3. The data were randomly divided into training (70%) and testing (30%) data. The logistic regression models were constructed by using the *glm* function in the training dataset. Because several independent variables in the models were correlated, a univariate selection technique was adopted, which meant that the variables were chosen one by one. For each model, the VIF was used to determine the level of multicollinearity. To avoid this problem, the VIF value should be less than 10 (Hair et al., 2019; Serrano-Cinca et al., 2019). McFadden's R^2 was then calculated to assess the quality of the models, with values ranging from 0 to 1 (Abdallah, 2018; Ma & Li, 2021). Finally, the optimal model was constructed using the backward stepwise Akaike information criterion, which allows for the reduction and selection of the most essential variables (Liu et al., 2022; Zhang, 2016).

Step 4. The performance of the logistic regression models was evaluated by analyzing the receiver operating characteristic (ROC) curves for the training and test datasets. The ROC curves are plotted with sensitivity or true positive rate versus 1—specificity or false positive rate (Hajian-Tilaki, 2013; Phillips et al., 2006). On the other hand, the area under the curve (AUC) reflects the area beneath the ROC curve (Carneiro et al., 2017). Thus, the *auc* function computes one measure of—accuracy to discriminate between the active and bankrupt companies (Hajian-Tilaki, 2013). Overall, a model with an AUC value greater than 0.90 indicates high predictive power (Manel et al., 2001; Tserng et al., 2012, 2014). There are other metrics for evaluating model performance, such as the accuracy (ACC) calculated from the confusion matrix. According to some authors, this metric is less reliable than the AUC because it requires the establishment of a single cut-off value (Bowers & Zhou, 2019; Ling et al., 2003). These two metrics were computed to evaluate the model's performance.

3. Results

3.1. Descriptive analysis

Table 3 presents the descriptive statistics for the active and bankrupt companies one-year prior to bankruptcy, including the minimums, medians, averages, standard deviations and maximums. The descriptive statistics revealed, as predicted, that bankrupt companies had significantly different characteristics from active companies. These had higher levels of liquidity, profitability and solvency ratios. In contrast, the bankrupt companies had a greater amount of indebtedness. Furthermore, when the averages of the variables were examined in greater detail, it was discovered that the average values of the ratios LIQG, LIQR, FMAT, FMV, ROE, ROA, RLV, EBITAT, FCAT, AFAT, EBITDAV, VAT, EBITDAJS and SOLV were higher for active companies; thus, it is expected that the higher the value of these ratios, the lower the probability of bankruptcy. The average values for the ACV, PTAT, PTCP and PCAT indicators, on the other hand, were higher for the bankrupt companies. As a result, the higher the value of these indicators, the greater the probability of a company approaching bankruptcy, necessitating specific attention from managers to its increase.

The correlations between the independent variables are presented in Table 4 and Table 5. The *cor* function shows that the following microeconomic variables were relatively highly correlated: EBITAT with ROA (0.981), FCAT with ROA (0.925), LIQG with LIQR (0.858), EBITDAV with RLV (0.976), FCAT with EBITAT (0.919) and ACV with FMV (0.972). The SOLV variable and PTAT variable had a perfect negative correlation of −1[4]. In terms of the macroeconomic variables, almost all correlations were relevant, except those between PIB and TI, TI and TD, TI and TNE and TNE and TJCH.

[4]This perfect negative correlation is because these two financial ratios are complementary (the sum of both is equal to 1).

Table 3. Summary statistics for Model 1.

Variable	State	Minimum	Median	Mean	Standard deviation	Maximum
LIQG	Active	0.095	2.132	3.894	6.486	98.067
	Bankruptcy	0.510	1.474	2.429	2.474	17.609
LIQR	Active	0.029	1.758	3.062	5.400	98.067
	Bankruptcy	0.001	1.038	1.290	1.449	12.026
FMAT	Active	−2.084	0.290	0.310	0.262	0.979
	Bankruptcy	−0.264	0.426	0.439	0.271	0.982
FMV	Active	−6.480	0.227	1.189	16.720	640.196
	Bankruptcy	−0.615	0.603	4.869	31.611	504.515
FCAT	Active	−0.674	0.063	0.081	0.103	0.665
	Bankruptcy	−0.429	0.010	0.006	0.070	0.293
ROE	Active	−8.087	0.079	0.060	0.551	2.117
	Bankruptcy	−8.559	0.006	−0.197	0.867	0.597
EBITAT	Active	−0.620	0.046	0.068	0.106	0.597
	Bankruptcy	−0.472	0.021	0.014	0.070	0.252
RLV	Active	−37.213	0.024	0.051	2.188	76.945
	Bankruptcy	−8.563	0.003	−0.124	0.625	0.432
ROA	Active	0.093	−0.659	0.027	0.044	0.565
	Bankruptcy	−0.502	0.001	−0.014	0.066	0.204
EBITDAV	Active	−34.862	0.072	0.114	2.121	75.091
	Bankruptcy	−2.706	0.064	0.090	0.654	9.169
VAT	Active	0.001	1.161	1.291	0.811	6.365
	Bankruptcy	0.002	0.621	0.730	0.615	4.290
ACV	Active	0.083	0.634	2.327	27.393	1041.760
	Bankruptcy	0.196	1.413	6.522	35.786	564.382
PTAT	Active	0.008	0.598	0.573	0.221	0.995
	Bankruptcy	0.297	0.805	0.785	0.125	0.994
PTCP	Active	0.008	1.488	3.058	7.907	199.488
	Bankruptcy	0.422	4.130	6.991	12.497	168.456
EBITDAJS	Active	−1337933	17.652	5856	147875	5451071
	Bankruptcy	−273.705	1.462	3.324	46.218	552.836
SOLV	Active	0.005	0.402	0.427	0.221	0.992
	Bankruptcy	0.006	0.195	0.215	0.125	0.703
PCAT	Active	0.006	0.359	0.379	0.221	0.962
	Bankruptcy	0.056	0.529	0.528	0.245	0.978
AFAT	Active	0	0.154	0.213	0.197	0.982
	Bankruptcy	0	0.099	0.156	0.158	0.948
DIM	Active	1.198	2.598	2.708	0.603	6.186
	Bankruptcy	1.866	3.483	3.493	0.635	5.595
NUMC	Active	1	8	19.349	99.787	3418
	Bankruptcy	1	16	30.011	55.662	719

Continued on next page

Variable	State	Minimum	Median	Mean	Standard deviation	Maximum
TJ	Active	0.023	0.023	0.023	−	0.023
	Bankruptcy	0.024	0.055	0.054	0.010	0.067
PIB	Active	0.022	0.022	0.022	−	0.022
	Bankruptcy	−0.041	−0.009	−0.099	0.022	0.035
TI	Active	0.003	0.003	0.003	−	0.003
	Bankruptcy	−0.008	0.014	0.016	0.016	0.037
TD	Active	0.065	0.065	0.065	−	0.065
	Bankruptcy	0.070	0.108	0.113	0.027	0.162
TNE	Active	0.120	0.120	0.120	−	0.120
	Bankruptcy	0.077	0.088	0.094	0.017	0.133
TJCH	Active	0.012	0.012	0.012	−	0.012
	Bankruptcy	0.014	0.032	0.034	0.011	0.054

Note: Rates are constant throughout the year and do not fluctuate.

In general, the high correlation observed between practically all of the microeconomic variables is related to a common numerator or denominator. Nonetheless, the correlation between macroeconomic variables is largely explained by their behavior as a function of the economic cycle. For example, when the gross domestic product grows, interest rates and inflation rise, while unemployment falls. People will also be more likely to start businesses because they are more confident that their investments will be rewarded. After all, the economy is experiencing higher levels of income and demand.

Table 4. Correlation coefficients for the microeconomic variables.

	LIQG	LIQR	FMAT	FMV	FCAT	ROE	EBITAT	RLV	ROA	EBITDAV	VAT	ACV	PTAT	PTCP	EBITDAJS	SOLV	PCAT	AFAT	DIM	NUMC
LIQG	1.000																			
LIQR	0.858	1.000																		
FMAT	0.175	0.014	1.000																	
FMV	0.035	−0.016	0.115	1.000																
FCAT	−0.004	0.067	−0.191	−0.062	1.000															
ROE	0.026	0.038	−0.017	−0.013	0.483	1.000														
EBITAT	0.012	0.064	−0.095	−0.047	0.919	0.513	1.000													
RLV	0.060	0.094	−0.045	−0.393	0.127	0.048	0.125	1.000												
ROA	0.027	0.076	−0.092	−0.042	0.925	0.537	0.981	0.140	1.000											
EBITDAV	0.059	0.090	−0.030	−0.249	0.123	0.045	0.121	0.976	0.135	1.000										
VAT	−0.146	−0.063	−0.241	−0.111	0.289	0.120	0.230	0.003	0.188	−0.024	1.000									
ACV	0.040	0.001	0.080	0.972	−0.058	−0.013	−0.044	−0.323	−0.036	−0.208	−0.119	1.000								
PTAT	−0.341	−0.381	0.018	0.009	−0.286	−0.193	−0.287	−0.057	−0.303	−0.049	0.005	−0.004	1.000							
PTCP	−0.078	−0.102	0.050	0.027	−0.178	−0.482	−0.171	−0.017	−0.172	−0.011	−0.102	0.022	0.425	1.000						
EBITDAJS	−0.010	−0.001	−0.024	−0.177	0.073	0.016	0.069	0.108	0.064	0.106	0.007	−0.196	−0.048	−0.013	1.000					
SOLV	0.341	0.381	−0.018	−0.009	0.286	0.193	0.287	0.057	0.303	0.049	−0.005	0.004	−1.000	−0.425	0.048	1.000				
PCAT	−0.489	−0.432	−0.119	−0.014	−0.186	−0.112	−0.164	−0.030	−0.173	−0.031	0.122	−0.022	0.619	0.236	−0.015	−0.619	1.000			
AFAT	−0.122	−0.080	−0.443	−0.054	0.130	−0.007	−0.010	0.033	−0.010	0.040	−0.032	−0.044	−0.028	−0.042	0.040	0.028	−0.149	1.000		
DIM	−0.093	−0.184	0.114	0.109	−0.170	−0.020	−0.082	0.018	−0.075	0.043	−0.387	0.121	0.167	0.075	0.003	−0.167	0.153	0.039	1.000	
NUMC	−0.050	−0.041	−0.037	−0.010	−0.014	0.000	−0.006	−0.001	−0.009	−0.001	0.004	−0.009	0.028	−0.006	−0.003	−0.028	0.058	0.032	0.302	1.000

Table 5. Correlation coefficients for the macroeconomic variables.

	TJ	PIB	TI	TD	TNE	TJCH
TJ	1.000					
PIB	−0.832	1.000				
TI	0.738	−0.459	1.000			
TD	0.822	−0.790	0.602	1.000		
TNE	−0.776	0.821	−0.535	−0.836	1.000	
TJCH	0.973	−0.745	0.745	0.710	−0.623	1.000

3.2. Principal component analysis

PCA was applied to choose the variables that explain as much as possible of the variance in the dataset and allow them to be characterized and reduced after having previously confirmed whether these variables were associated (see Section 3. 1). In this context, the number of components to be preserved must be fixed. Thus, only the top five components were kept since the variables that contributed the most to these components produced the greatest outcomes in the estimated models. It should be noted that their summed contribution to the total variation was 61.87%, which is less than 70%, a common value found in the literature (the subjective nature of this criterion is emphasized). Figure 2 depicts the *fviz contrib* function results for the first five components.

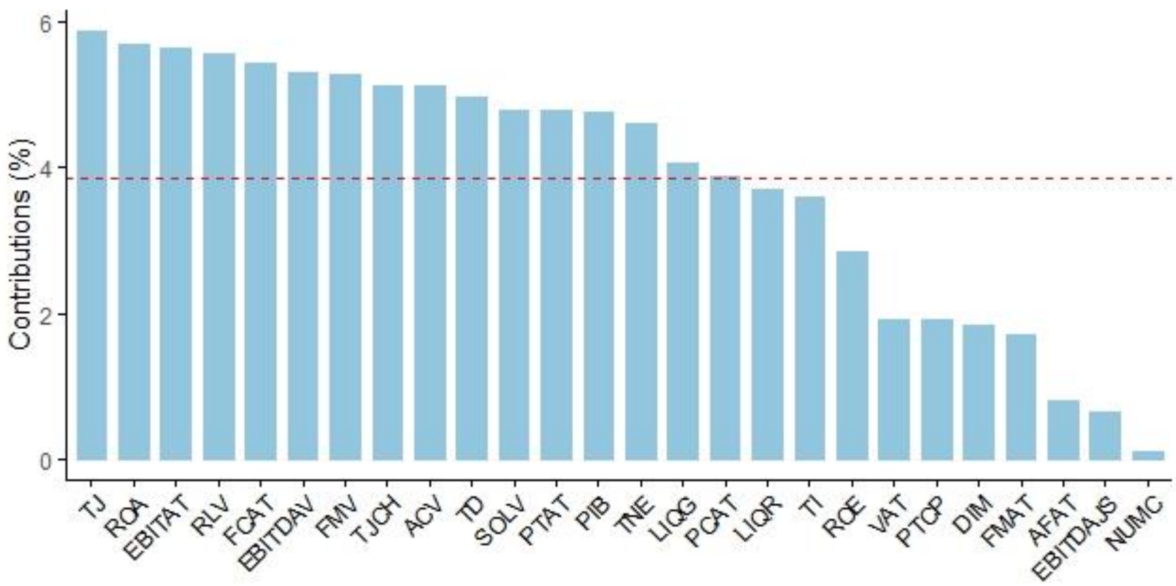

Figure 2. Variables that have the greatest impact on the first five principal components.

It is important to note that the variables that contribute the most to the first five components are above the red line in Figure 2. As a result, the following independent variables were chosen for the model: TJ, ROA, EBITAT, RLV, FCAT, EBITDAV, FMV, TJCH, ACV, TD, SOLV, PTAT, PIB, TNE, LIQG and PCAT.

3.3. Empirical prediction models

The bankruptcy model proposed in this study entails the estimation of three models for three different temporal moments: Model 1 predicts bankruptcy from one year prior to the observation of bankruptcy, Model 2 predicts bankruptcy from two years prior and Model 3 predicts bankruptcy from three years prior. Table 6 provides an overview of the findings. For all three models, the microeconomic variables that appear to be important predictors are cash flow to total assets and shareholder funds to total assets. Thus, companies with a higher cash flow-to-total assets ratio are less likely to fail. The negative sign of the shareholder funds-to-total assets ratio indicates that a high value of this variable reduces the probability of failure. On the other hand, the effects of current liabilities on total assets differ between the models. As a result, this variable is only statistically significant in Model 3, and its sign becomes positive over time. The macroeconomic variables are remarkably important across models. As expected, companies are less likely to fail when PIB increases. The TNE variable is significantly related to bankruptcy in Model 2 and Model 3. However, the coefficient for this variable has an unexpected sign, indicating that the higher the value of this macroeconomic indicator, the less likely is a company to be in a difficult financial situation. The VIF values for the five variables considered in this study were all less than 2, indicating that multicollinearity is not a problem in the estimations, and that the coefficient levels obtained are reliable. In Model 1, McFadden's R^2 was 87%, indicating that this model explains 87% of the total variability of the dependent variable. The McFadden's R^2 values for Model 2 and Model 3 were 68% and 65%, respectively.

Table 6. Outcomes of logistic regression models.

	Model 1		Model 2		Model 3	
FCAT	−5.601	**	−4.5346	***	−3.0898	*
SOLV	−3.991	***	−4.7503	***	−4.3135	***
PCAT	−1.871		−0.2309		1.3061	*
PIB	−799.103	***	−285.9563	***	−263.4298	***
TNE	−357.092		−128.3724	***	−184.3570	***
Constant	58.664	*	20.4592	***	25.8439	***
AIC	149.93		362.73		393.09	
McFadden's R^2	0.873		0.676		0.648	

Note: *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

3.4. Performance evaluation for logistic regression models

The area under the ROC curve (AUC) measures the real performance of a statistical model, and a value near 1.0 indicates excellent performance. All of the models demonstrated high performance on the training and test datasets. The first model achieved the best performance with an AUC value of 0.963 for the test dataset. Thus, this statistical model can correctly predict the bankruptcy of a company with a probability of 96.3%. Figure 3 presents the ROC curve of each model based on microeconomic and macroeconomic indicators for the test dataset.

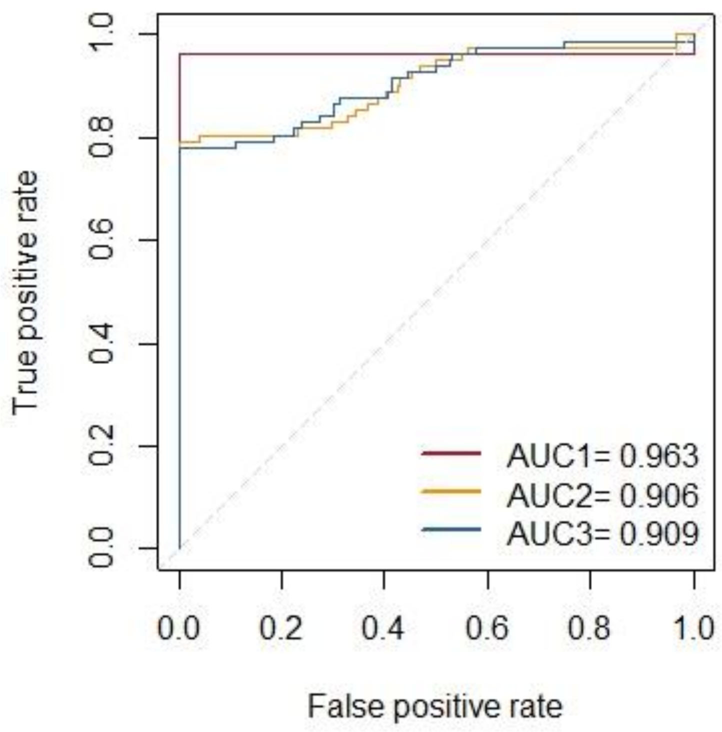

Figure 3. ROC curves indicating model performance.

A confusion matrix, on the other hand, represents the number of elements correctly predicted and those incorrectly classified. (Note that the optimal cut-off was computed using the *InformationValue* package.) Tables 7, 8 and 9 depict the confusion matrices derived for Models 1, 2 and 3, respectively. The results indicate that all active companies were correctly classified. However, as the time horizon increased, so did the number of bankruptcy companies classify as false positives. The accuracy of the models was calculated using the matrices mentioned below.

Table 7. Confusion matrix for Model 1 (optimal cut-off = 0.20).

		Actual condition	
		Active	Bankrupt
Predicted condition	Active	467	4
	Bankrupt	0	78
Overall accuracy: ACC = 99.5%; Sensitivity = 96.3%; Specificity = 100%.			

Table 8. Confusion matrix for Model 2 (optimal cut-off = 0.35).

		Actual condition	
		Active	Bankrupt
Predicted condition	Active	467	17
	Bankrupt	0	65
Overall accuracy: ACC = 96.9%; Sensitivity = 79.3%; Specificity = 100%.			

Table 9. Confusion matrix for Model 3 (optimal cut-off = 0.31).

		Actual condition	
		Active	Bankrupt
Predicted condition	Active	467	21
	Bankrupt	0	61

Overall accuracy: ACC = 96.7%; Sensitivity = 78.0%; Specificity = 100%.

4. Discussion

4.1. Models with microeconomic and macroeconomic variables

Macroeconomic factors are now recognized as important predictors of bankruptcy, affecting all companies (Acosta-González et al., 2019). By linking the microeconomic and macroeconomic variables in a new model, it was possible to explore the question of whether macroeconomic indicators improve the performance of bankruptcy prediction models, besides the inclusion of microeconomic indicators. As a result, this study offers a consistent and highly accurate model with the two types of variables, suggesting that companies in the construction sector are sensitive to changes in the macroeconomic conditions of the country. This finding corroborates the results obtained by Acosta-González et al. (2019), Carvalho et al. (2020), Giriūniene et al. (2019), Oliveira (2014), Tinoco & Wilson (2013) and Yan et al. (2020). The predictor variables included in the model exhibited different characteristics over time, including statistical significance, as well as the expected[5] and obtained signs (Table 10 and Table 11).

Table 10. Variables included in the models.

Category	Variable	Model 1	Model 2	Model 3
Liquidity ratios	FCAT	X	X	X
Solvency and financial autonomy ratios	SOLV	X	X	X
	PCAT			X
Macroeconomic indicators	PIB	X	X	X
	TNE		X	X

Note: The X means that the variable has statistical significance.

Table 11. Expected signs and output results according to the variables in each model.

Category	Variable	Expected sign	Model 1	Model 2	Model 3
Liquidity ratios	FCAT	-	-	-	-
Solvency and financial autonomy ratios	SOLV	-	-	-	-
	PCAT	+	-	-	+
Macroeconomic indicators	PIB	-	-	-	-
	TNE	+	-	-	-

Note: The sign "-" indicates that the greater the value of this variable, the lower the likelihood of bankruptcy. The "+" sign indicates that the greater the value of this variable, the greater the likelihood of bankruptcy.

[5]Theoretical signs established by the authors studied.

The models were formed based on three microeconomic indicators, one of which is related to the liquidity ratio, and the others to solvency and the financial autonomy ratios. These categories are critical to understanding the bankruptcy event. According to Altman (1968) "*a firm with a poor profitability and/or solvency record may be regarded as a potential bankrupt"*. Furthermore, Dimitras et al. (1996) examined 47 articles on bankruptcy prediction models and discovered that "*the most important financial ratios came from the solvency category*". According to Lee et al. (1996), "*in case of failure prediction liquidity seems to play an important role*". Recently, Altman & Hotchkiss (2006) stated that "*in general, ratios measuring profitability, liquidity, leverage, and solvency, and multidimensional measures, like earnings and cash flow coverage, prevailed as the most significant indicators*". The authors also stated that "*the order of their importance is not clear since almost every study cited a different ratio as being the most effective indication of impending problems*". As a result, it can be deduced that the categories of indicators chosen for the model developed in this study are, according to several authors, the most significant.

The cash flow/total assets (FCAT) ratio was statistically significant in all models in terms of its effect on liquidity. As a result, it was found that this ratio is crucial in explaining bankruptcy in the short and medium term. The expected negative coefficient indicates that the higher the liquidity of a company, the lower the likelihood of bankruptcy. In other words, businesses generate positive net financial resources from their operations, which strengthens their financial position. Out of all of the authors studied, only Correia (2012) and Vieira et al. (2013) demonstrated the predictive power of this indicator in their models. Furthermore, Vieira et al. (2013) stated that "*the ratio that has more capacity to predict bankruptcy over the different periods in the Portuguese construction sector is the ratio of cash-flow to total assets*", a finding similar to Correia (2012). This result, however, differs from those presented by Acosta-González et al. (2019), Beaver (1966), Karas & Režňáková (2017a), Karas & Režňáková (2017b) and Karas & Srbová (2019), because none of these authors considered it important or statistically significant for their models.

The shareholder funds/total assets (SOLV) ratio was always statistically significant and negatively related to bankruptcy in the category of solvency and financial autonomy ratios, demonstrating a result consistent with the initial expectation. Therefore, the lower the value of this ratio, the greater the likelihood of companies in the Portuguese construction sector will go bankrupt in the short or medium term. Entities with greater financial strength and a larger capacity to meet non-current obligations are thus less reliant on external funds. Jones & Wang (2019), Karminsky & Burekhin (2019), Obradović et al. (2018), Pacheco et al. (2019) and da Rosa (2017) all mentioned that this ratio is important for their models, though the signs obtained by Obradović et al. (2018) were opposite of what was expected. Oliveira (2014) and da Pimenta (2015), on the other hand, did not find evidence that solvency was important in explaining bankruptcy. The current liabilities/total assets (PCAT) ratio was found to be statistically significant only in Model 3. As expected, this ratio had a positive coefficient, suggesting that increasing the proportion of current liabilities to total assets increases the likelihood of bankruptcy in the medium term. In general, these businesses finance a larger portion of their total assets with short-term liabilities (either short-term bank loans or supplier credit, for example). This will put the company under a lot of financial strain because it needs to be able to generate money to meet those responsibilities in the short term. The sign of this indicator in Models 1 and 2 was the opposite of what was expected, but it was not statistically significant. These findings appear to indicate that this ratio's predictive ability improves over time, as it was only statistically significant in Model 3. Another intriguing aspect of the PCAT ratio has been demonstrated by Karas

& Režňáková (2017c), who investigated the impact of this ratio in the construction and manufacturing sectors and discovered that it appears to be specific to the construction sector, as it was one of the least significant variables in the manufacturing sector and one of the most significant in the construction sector. The studies by Acosta-González et al. (2019), Karas & Režňáková (2017a), Karas & Režňáková (2017b), Karas & Režňáková (2017c) and Yan et al. (2020) corroborate the results obtained in this study, as these authors reported a positive sign for this indicator, as well as highlighted its merits in predicting the bankruptcy of construction companies.

Two macroeconomic indicators were used in the models: the gross domestic product (PIB) and birth rate of enterprises (TNE). The PIB indicator was statistically significant in all models tested, corroborating the results obtained by Carvalho et al. (2020), Jones & Wang (2019), Oliveira (2014) and Yan et al. (2020). The results obtained differ from those obtained by Giriūniene et al. (2019). This variable's coefficient estimate had a negative sign, implying that the lower the PIB growth rate, the higher the risk of bankruptcy. This conclusion was predictable given that PIB represents all goods and services produced by a given country; thus, a high PIB indicates a healthy and balanced economy. This result is even more relevant when one considers that, by definition, the construction sector is highly sensitive to the economic cycle. The TNE indicator demonstrated statistical relevance in Models 2 and 3, and, despite the PIB indicator's strong capacity for prediction when combined with the microeconomic indicators, the addition of this indicator improved the quality of the models. This result is comparable to those presented by Altman (1983), Cuthbertson & Hudson (1996), Hudson (1986), Liu (2004), Platt & Platt (1994) and Young (1995). However, the estimated coefficient had a negative sign, implying that a decrease in TNE increases the likelihood of failure. At the outset of the study, it was expected that a higher enterprise birth rate might predict bankruptcy since fiercer competition would exist between companies in the construction sector. One possible explanation for the obtained result is that the time horizon chosen for the failed companies is correlated with the number of bankruptcies because, surprisingly, the years with the highest number of bankruptcies were also the years with the lowest TNE indicator. It is also worth noting the TNE indicator's lack of available data for 2019. As a result, the average for the previous five years was calculated because it remained constant during this period.

4.2. Comparing performance measures of the models

In general, these models appear to be highly accurate up to three years before the bankruptcy (around 90%), as based on the measures AUC and ACC, as well as be of excellent quality (Table 12). Model 1 performed better in terms of AUC and ACC, primarily for the training dataset (96% and 99%, respectively). Nonetheless, Models 2 and 3 yielded curves that were reasonably close to the ideal model (i.e., equal to 1). Acosta-González et al. (2019), Choi et al. (2018) and Ohlson (1980) all observed similar behavior in terms of the model's predictive ability. However, it is also possible to confirm that several of the works mentioned throughout this paper, namely Altman (1968), Costa (2014), Horta & Camanho (2013), Karminsky & Burekhin (2019), Kuběnka & Myšková (2019), Pacheco et al. (2019), Tinoco & Wilson (2013), Tserng et al. (2014) and Zoričák et al. (2020), presented models with much lower explanatory power.

On the other hand, as the time reference period before bankruptcy lengthened, the accuracy rate (Beaver, 1966) and quality decreased (Succurro et al., 2019). According to Beaver (1966), one of the possible explanations for these findings is the deterioration of bankrupt companies over time, because the difference in the averages of the indicators between active and bankrupt companies grows as bankruptcy approaches. Another possible explanation is that the microeconomic indicators used in the models are typically based on data from a year before companies are identified as bankrupt, and thus may not be the most appropriate measures to explain bankruptcy in the two and three years preceding it, i.e., an indicator that is important in the year preceding bankruptcy may not be important in the earlier years preceding bankruptcy (Karas & Režňáková, 2017c).

Even though the PCA results showed that some macroeconomic indicators significantly contributed to the variability of the data, the entire process of building the model with only microeconomic indicators was repeated to understand the impact of macroeconomic indicators on the prediction of bankruptcies in the Portuguese construction sector. The model developed with only microeconomic indicators performed worse, with AUC, ACC and quality values of around 80% and less than 22% (quality) (Table 12).

This confirmed that the model proposed with microeconomic and macroeconomic indicators is more efficient and effective. Furthermore, the disparity in the values obtained by the two models suggests that the Portuguese construction sector is sensitive to macroeconomic changes. Acosta-González et al. (2019) reached a similar conclusion: "*companies in the Spanish construction sector are more sensitive to changes in macroeconomic factors than to their accounting ratios*".

Table 12. Measures of performance for the models.

	Micro and macroeconomic variables			Microeconomic variables		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
test data (AUC)	96.3%	90.6%	90.9%	82.6%	80.9%	77.9%
accuracy (ACC)	99.5%	96.9%	96.7%	84.9%[a]	84.9%[b]	84.9%[c]
McFadden's R^2	87.3%	67.6%	64.8%	22.0%	16.8%	15.8%

Note: ACC values computed for optimal cut-off: [a] Optimal cut-off = 0.8346; [b] Optimal cut-off = 0.6516; [c] Optimal cut-off = 0.6523.

To summarize, this research emphasizes the importance of ratios that measure a company's liquidity, solvency and financial autonomy, as well as the importance of macroeconomic indicators (that condition the context in which all companies operate) in predicting bankruptcy. In addition, the proposed model's high accuracy demonstrates its applicability in the Portuguese construction sector. Given the findings, it would be beneficial for businesses in this sector to use the model developed and, as a result, devise solutions that reduce and prevent their financial situation from worsening. This model is equally important for banks, public entities and investors, primarily in terms of risk analysis, because it allows for the assessment of a company's viability and stability.

5. Conclusions

The purpose of this study was to develop a model that is able to predict the probability of bankruptcy in the Portuguese construction sector while also including company-specific variables and macroeconomic indicators. Therefore, the study's major contribution to the literature is the identification of a parsimonious model that predicts the probability of a company's bankruptcy based on a limited number of company-specific variables and macroeconomic indicators. The combination appears to improve the model's performance significantly, corroborating and extending the findings of previous empirical studies. Indeed, research on whether indicators from the external environment or indicators within the company influence the probability of bankruptcy has attracted the attention of several researchers. On the one hand, a company's specific indicators (here, proxied by financial ratios) drive its profitability and market success. On the other hand, it is acknowledged that external indicators (here, proxied by macroeconomic indicators) that are beyond the control of company managers affect a company's performance. And, these external indicators are important because company managers are unable to accurately predict and manage the impact of changes in macroeconomic indicators on the company's performance. However, due to the high correlations between the variables, PCA was required to determine which ones best explained the variation in the dataset.

Regarding the performance measures for the models, the results showed an overall predictive accuracy of over 90% up to three years before the bankruptcy, with quality ranging from 65% to 87%. The decrease in predictive ability and quality over time is primarily due to the time elapsed before bankruptcy occurred. Solvency, as well as cash flow/total assets and current liabilities/total assets, are microeconomic indicators that help explain the bankruptcy event. Furthermore, it has been demonstrated that solvency and the cash flow/total assets ratios are critical for analyzing the situations of companies in the construction sector in the short and medium term, and that the current liabilities/total assets ratio is relevant in the medium term. In terms of macroeconomic indicators, it was found that the gross domestic product and birth rate of enterprises in the construction sector are critical in predicting bankruptcy.

A major implication of the results obtained is that as the economic cycle has become increasingly important owing to recent economic crises; it is critical for companies in the construction sector to employ the created model and, consequently, build solutions that minimize and prevent their financial condition from worsening. Therefore, these findings shed new light on bankruptcy prediction models aimed at the construction sector, emphasizing its sensitivity to macroeconomic conditions (Acosta-González et al., 2019; Carvalho et al., 2020; Giriūniene et al., 2019). This model is also beneficial for banks by way of their credit risk assessment activities and government agencies by way of public policy design; it also helps investors to have a better understanding of a company's reality when deciding about their investments.

This study faced several limitations, including a lack of key indicator information in Orbis, which resulted in a significant drop in the amount of usable data, primarily in the group of bankrupt enterprises, as well as in the number of independent variables. Because PORDATA did not have information on the TNE indicator for 2019, it was necessary to compute the average for the previous five years (similar values in this period). Another limitation of the study is that we were unable to identify a significant sample of Orbis companies in bankruptcy. It should also be noted that not all companies use the same accounting standards, which may be misleading in terms of their true economic and financial status and affect the models' predictive power. Nevertheless, this difficulty was mitigated to some extent by using the same database (Orbis) for all companies and the existence of a procedure for standardizing the information obtained from the enterprises, allowing for better comparison of this information. Finally, the high correlation between several of the selected indicators is evident, which is to be expected given that the majority were obtained in nearly identical ways. To address this issue, a more in-depth examination of the ratios used in bankruptcy prediction models is recommended, as previous studies have primarily focused on those with the greatest degree of popularity in the literature, with no rational theory guiding their selection (Acosta-González & Fernández-Rodríguez, 2014; Karels & Prakash, 1987).

Some avenues for further research are as follows. First, it is envisaged to include another category of indicators, such as market indicators (e.g., the market value of equity, share price and market capitalization/total debt ratio). Some authors (e.g., Pham Vo Ninh et al., 2018; Tinoco & Wilson, 2013) have already confirmed that the combination of these three types of indicators is beneficial for predicting bankruptcy. Second, replicating this research project for a larger sample as a more representative sample of the population, specifically in the group of failed companies, would improve the model performance and then determine whether the results obtained would remain. Third, a study on the effects of the COVID-19 pandemic on the economic and financial situation of Portuguese and foreign companies should be conducted to gain a more comprehensive understanding of the response of organizations, as well as the sectors most harmed. Finally, it would be interesting to develop an interface that construction companies, as well as investors, government agencies and banks, could use to obtain all possible information about the state of their company and relevant predictions in a simple and intuitive manner. This interface can be created by using the shiny package, which is available in RStudio.

Although we believe that our study could be replicated using a sample of companies in other sectors and/or countries, and that similar insights are likely to be obtained, it is necessary to be cautious regarding the generalization of the conclusions drawn in this study. The complexity of the process associated with the bankruptcy of a company should be recognized. It is possible, for example, that institutional contexts, regulatory frameworks or cultural beliefs might influence that process, which makes it difficult to fully capture all of the factors driving the failure of a company.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

Abdallah FDM (2018) Statistical Modelling of Categorical Outcome with More than Two Nominal Categories. *Am J Appl Math Stat* 6: 262–265. https://doi.org/10.12691/ajams-6-6-7

Acosta-González E, Fernández-Rodríguez F (2014) Forecasting Financial Failure of Firms via Genetic Algorithms. *Comput Econ* 43: 133–157. https://doi.org/10.1007/s10614-013-9392-9

Acosta-González E, Fernández-Rodríguez F, Ganga H (2019) Predicting Corporate Financial Failure Using Macroeconomic Variables and Accounting Data. *Comput Econ* 53: 227–257. https://doi.org/10.1007/s10614-017-9737-x

Altman EI (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *J Financ* 23: 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman EI (1983) Why businesses fail. *J Bus Strat* 3: 15–21. https://doi.org/10.1108/eb038985

Altman EI, Hotchkiss E (2006) *Corporate Financial Distress and Bankruptcy* (3rd ed.), John Wiley & Sons, Inc.

Asuero AG, Sayago A, González AG (2006) The correlation coefficient: An overview. *Crit Rev Anal Chem* 36: 41–59. https://doi.org/10.1080/10408340500526766

Barboza F, Kimura H, Altman E (2017) Machine learning models and bankruptcy prediction. *Expert Syst Appl* 83: 405–417. https://doi.org/10.1016/j.eswa.2017.04.006

Beaver WH (1966) Financial Ratios As Predictors of Failure. *J Account Res* 4: 71–111. https://doi.org/10.2307/2490171

Beaver W, McNichols M, Rhie JW (2005) Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy. *Rev Account Stud* 10: 93–122. https://doi.org/10.1007/s11142-004-6341-9

Bellovary J, Giacomino D, Akers MD (2007) A Review of Bankruptcy Prediction Studies: 1930–Present. *J Financ Educ* 33: 1–42. https://www.jstor.org/stable/41948574

Boratyńska K (2016) Corporate bankruptcy and survival on the market: Lessons from evolutionary economics. *Oecon Copernic* 7: 107–129. https://doi.org/10.12775/OeC.2016.008

Boritz JE, Kennedy DB (1995) Effectiveness of neural network types for prediction of business failure. *Expert Syst Appl* 9: 503–512. https://doi.org/10.1016/0957-4174(95)00020-8

Bowers AJ, Zhou X (2019) Receiver Operating Characteristic (ROC) Area Under the Curve (AUC): A Diagnostic Measure for Evaluating the Accuracy of Predictors of Education Outcomes. *J Educ Stud Placed Risk* 24: 20–46. https://doi.org/10.1080/10824669.2018.1523734

Carneiro P, Braga AC, Barroso M (2017) Work-related musculoskeletal disorders in home care nurses: Study of the main risk factors. *Int J Ind Ergonom* 61: 22–28. https://doi.org/10.1016/j.ergon.2017.05.002

Carvalho PV, Curto JD, Primor R (2020) Macroeconomic determinants of credit risk: Evidence from the Eurozone. *Int J Financ Econ,* 1–19. https://doi.org/10.1002/ijfe.2259

Chen JH, Su MC, Annuerine B (2016) Exploring and weighting features for financially distressed construction companies using Swarm Inspired Projection algorithm. *Adv Eng Inform* 30: 376–389. https://doi.org/10.1016/j.aei.2016.05.003

Cheng MY, Hoang ND (2015) Evaluating contractor financial status using a hybrid fuzzy instance based classifier: Case study in the construction industry. *IEEE T Eng Manage* 62: 184–192. https://doi.org/10.1109/TEM.2014.2384513

Choi H, Son H, Kim C (2018) Predicting financial distress of contractors in the construction industry using ensemble learning. *Expert Syst Appl* 110: 1–10. https://doi.org/10.1016/j.eswa.2018.05.026

Correia C (2012) *Previsão da insolvência: evidência no setor da construção* [Dissertação de Mestrado, Universidade de Aveiro]. Repositório Institucional da Universidade de Aveiro. http://hdl.handle.net/10773/9573

Costa HA (2014) *Modelo de previsão de falência: o caso da construção civil em Portugal* [Dissertação de Mestrado, Universidade do Algarve, Repositório da Universidade do Algarve]. http://hdl.handle.net/10400.1/8321

Cuthbertson K, Hudson J (1996) The determinants of compulsory liquidations in the U.K. *Manch Sch* 64: 298–308. https://doi.org/10.1111/j.1467-9957.1996.tb00487.x

Daoud JI (2017) Multicollinearity and Regression Analysis. *J Phys (Conference Series)* 949: 1–6. https://doi.org/10.1088/1742-6596/949/1/012009

Dimitras AI, Zanakis SH, Zopounidis C (1996) A survey of business failures with an emphasis on prediction methods and industrial applications. *Eur J Oper Res* 90: 487–513. https://doi.org/10.1016/0377-2217(95)00070-4

Etemadi H, Rostamy AAA, Dehkordi HF (2009) A genetic programming model for bankruptcy prediction: Empirical evidence from Iran. *Expert Sys Appl* 36: 3199–3207. https://doi.org/10.1016/j.eswa.2008.01.012

European Commission (2021 October) *European Construction Sector Observatory*. Available from: https://ec.europa.eu/docsroom/documents/47918/attachments/1/translations/en/renditions/native

Giriūniene G, Giriūnas L, Morkunas M, et al. (2019) A comparison on leading methodologies for bankruptcy prediction: The case of the construction sector in Lithuania. *Economies* 7: 1–20. https://doi.org/10.3390/economies7030082

Gotts SJ, Gilmore AW, Martin A (2020) Brain networks, dimensionality, and global signal averaging in resting-state fMRI: Hierarchical network structure results in low-dimensional spatiotemporal dynamics. *NeuroImage* 205: 1–17. https://doi.org/10.1016/j.neuroimage.2019.116289

Habib A, Costa MD, Huang HJ, et al. (2020) Determinants and consequences of financial distress: review of the empirical literature. *Account Financ* 60: 1023–1075. https://doi.org/10.1111/acfi.12400

Hair JF, Black WC, Babin BJ, et al. (2019) *Multivariate Data Analysis* (8th ed.), Cengage Learning.

Hajian-Tilaki K (2013) Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp J Int Med* 4: 627–635. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3755824/

Heo J, Yang JY (2014) AdaBoost based bankruptcy forecasting of Korean construction companies. *Appl Soft Comput* 24: 494–499. https://doi.org/10.1016/j.asoc.2014.08.009

Horta IM, Camanho AS (2013) Company failure prediction in the construction industry. *Expert Syst Appl* 40: 6253–6257. https://doi.org/10.1016/j.eswa.2013.05.045

Hudson J (1986) An analysis of company liquidations. *Appl Econ* 18: 219–235. https://doi.org/10.1080/00036848600000025

ben Jabeur S, Mefteh-Wali S, Carmona P (2021) The impact of institutional and macroeconomic conditions on aggregate business bankruptcy. *Struct Change Econ D* 59: 108–119. https://doi.org/10.1016/j.strueco.2021.08.010

Ben Jabeur S, Stef N, Carmona P (2022) Bankruptcy Prediction using the XGBoost Algorithm and Variable Importance Feature Engineering. *Comput Econ,* 1–27. https://doi.org/10.1007/s10614-021-10227-1

Jones S, Wang T (2019) Predicting private company failure: A multi-class analysis. *J Int Financ Mark Inst Money* 61: 161–188. https://doi.org/10.1016/j.intfin.2019.03.004

Kapliński O (2008) Usefulness and credibility of scoring methods in construction industry. *J Civil Eng Manage* 14: 21–28. https://doi.org/10.3846/1392-3730.2008.14.21-28

Karas M, Režňáková M (2017a) Predicting the bankruptcy of construction companies: A CART-based model. *Eng Econ* 28: 145–154. https://doi.org/10.5755/j01.ee.28.2.16353

Karas M, Režňáková M (2017b) The potential of dynamic indicator in development of the bankruptcy prediction models: The case of construction companies. *Acta Univ Agr Silviculturae Mendelianae Brunensis* 65: 641–652. https://doi.org/10.11118/actaun201765020641

Karas M, Režňáková M (2017c) The stability of bankruptcy predictors in the construction and manufacturing industries at various times before bankruptcy. *Ekonomika Manage* 20: 116–133. https://doi.org/10.15240/tul/001/2017-2-009

Karas M, Srbová P (2019) Predicting bankruptcy in construction business: Traditional model validation and formulation of a new model. *J Int Stud* 12: 283–296. https://doi.org/10.14254/2071-8330.2019/12-1/19

Karels GV, Prakash AJ (1987) Multivariate Normality and Forecasting of Business Bankruptcy. *J Bus Financ Account* 14: 573–593. https://doi.org/10.1111/j.1468-5957.1987.tb00113.x

Karminsky A, Burekhin R (2019) Comparative analysis of methods for forecasting bankruptcies of Russian construction companies. *Bus Inf* 13: 52–66. https://doi.org/10.17323/1998-0663.2019.3.52.66

Kim YJ, Cribbie RA (2018) ANOVA and the variance homogeneity assumption: Exploring a better gatekeeper. *British J Math Stat Psychol* 71: 1–12. https://doi.org/10.1111/bmsp.12103

Koksal A, Arditi D (2004) Predicting Construction Company Decline. *J Constr Eng Manage* 130: 799–807. https://doi.org/10.1061/(asce)0733-9364(2004)130:6(799)

Kuběnka M, Myšková R (2019) Obvious and hidden features of corporate default in bankruptcy models. *J Bus Econ Manage* 20: 368–383. https://doi.org/10.3846/jbem.2019.9612

Kwak SG, Kim JH (2017) Central limit theorem: The cornerstone of modern statistics. *Korean J Anesthesiology* 70: 144–156. https://doi.org/10.4097/kjae.2017.70.2.144

Kwak SG, Park SH (2019) Normality Test in Clinical Research. *J Rheumatic Dis* 26: 5–11. https://doi.org/10.4078/jrd.2019.26.1.5

Lafi SQ, Kaneene JB (1992) An explanation of the use of principal-components analysis to detect and correct for multicollinearity. *Prev Vet Med* 13: 261–275. https://doi.org/10.1016/0167-5877(92)90041-D

Lagesh MA, Srikanth M, Acharya D (2018) Corporate Performance during Business Cycles: Evidence from Indian Manufacturing Firms. *Global Bus Rev* 19: 1–14. https://doi.org/10.1177/0972150918788740

Lee KC, Han I, Kwon Y (1996) Hybrid neural network models for bankruptcy predictions. *Decis Support Syst* 18: 63–72. https://doi.org/10.1016/0167-9236(96)00018-8

Lee S, Choi WS (2013) A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis. *Expert Syst Appl* 40: 2941–2946. https://doi.org/10.1016/j.eswa.2012.12.009

Lessmann S, Baesens B, Seow HV, et al. (2015) Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *Eur J Oper Res* 247: 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Ling CX, Huang J, Zhang H (2003) *AUC: A better measure than accuracy in comparing learning algorithms* [Paper presentation]. Conference of the Canadian Society for Computational Studies of Intelligence, Berlin, Heidelberg. Available from: https://doi.org/10.1007/3-540-44886-1_25

Liu J (2004) Macroeconomic determinants of corporate failures: Evidence from the UK. *Appl Econ* 36: 939–945. https://doi.org/10.1080/0003684042000233168

Liu RX, Kuang J, Gong Q, et al. (2003) Principal component regression analysis with SPSS. *Comput Meth Prog Biomed* 71: 141–147. https://doi.org/10.1016/S0169-2607(02)00058-5

Liu W, Jiang Q, Sun C, Liu S, et al. (2022) Developing a 5-gene prognostic signature for cervical cancer by integrating mRNA and copy number variations. *BMC Cancer* 22: 1–16. https://doi.org/10.1186/s12885-022-09291-z

Lucanera JP, Fabregat-Aibar L, Scherger V, et al. (2020) Can the SOM analysis predict business failure using capital structure theory? Evidence from the subprime crisis in Spain. *Axioms* 9: 1–13. https://doi.org/10.3390/AXIOMS9020046

Lydersen S (2015) Statistical review: Frequently given comments. *Ann Rheumat Dis* 74: 323–325. https://doi.org/10.1136/annrheumdis-2014-206186

Ma J, Li C (2021) *A comparison of Logit and Probit models using Monte Carlo simulation* [Paper presentation]. 2021 40th Chinese Control Conference (CCC), Shanghai, China. Available from: https://doi.org/10.23919/CCC52363.2021.9550250

Manel S, Ceri Williams H, Ormerod SJ (2001) Evaluating presence-absence models in ecology: The need to account for prevalence. *J Appl Ecology* 38: 921–931. https://doi.org/10.1046/j.1365-2664.2001.00647.x

Mbaluka MK, Muriithi DK, Njoroge GG (2022) Application of Principal Component Analysis and Hierarchical Regression Model on Kenya Macroeconomic Indicators. *Eur J Math Stat* 3: 26–38. https://doi.org/10.24018/ejmath.2022.3.1.74

Min SH, Lee J, Han I (2006) Hybrid genetic algorithms and support vector machines for bankruptcy prediction. *Expert Syst Appl* 31: 652–660. https://doi.org/10.1016/j.eswa.2005.09.070

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

Murphy KR (2021) In praise of Table 1: The importance of making better use of descriptive statistics. *Ind Organ Psychol* 14: 461–477. https://doi.org/10.1017/IOP.2021.90

Neves JCD (2012) *Análise e Relato Financeiro—Uma visão integrada de gestão* (5th ed.), Texto Editores, Lda.

Ng ST, Wong JM, Zhang J (2011) Applying Z-score model to distinguish insolvent construction companies in China. *Habitat Int* 35: 599–607. https://doi.org/10.1016/j.habitatint.2011.03.008

Nouri BA, Soltani M (2016) Designing a bankruptcy prediction model based on account, market and macroeconomic variables (Case Study: Cyprus Stock Exchange). *Iranian J Manage Stud* 9: 125–147. https://doi.org/10.22059/ijms.2016.55038

Obradović DB, Jakaić D, Rupić IB, et al. (2018) Insolvency prediction model of the company: The case of the republic of serbia. *Econ Res-Ekon Istraz* 31: 138–157. https://doi.org/10.1080/1331677X.2017.1421990

OECD Statistics (2022) *SDBS Business Demography Indicators (ISIC Rev. 4) : Birth rate of enterprises*. Available from: https://stats.oecd.org/index.aspx?queryid=81074

Ohlson JA (1980) Financial Ratios and the Probabilistic Prediction of Bankruptcy. *J Account Res* 18: 109–131. https://doi.org/10.2307/2490395

Oliveira MPG (2014) *A insolvência empresarial na indústria transformadora portuguesa: as determinantes financeiras e macroeconómicas* [Dissertação de Mestrado, Universidade do Porto]. Repositório Aberto da Universidade do Porto. Available from: https://repositorio-aberto.up.pt/handle/10216/77110

Pacheco L, Rosa R, Oliveria Tavares F (2019) Risco de Falência de PME: Evidência no setor da construção em Portugal. *Innovar* 29: 143–157. https://doi.org/10.15446/innovar.v29n71.76401

Perboli G, Arabnezhad E (2021) A Machine Learning-based DSS for mid and long-term company crisis prediction. *Expert Syst Appl* 174: 1–12. https://doi.org/10.1016/j.eswa.2021.114758

Pham Vo Ninh B, Do Thanh T, Vo Hong D (2018) Financial distress and bankruptcy prediction: An appropriate model for listed firms in Vietnam. *Econ Syst* 42: 616–624. https://doi.org/10.1016/j.ecosys.2018.05.002

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Model* 190: 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

da Pimenta IC (2015) *Modelos de previsão de falência - análise econométrica do setor da construção civil na UE* [Dissertação de Mestrado, Universidade do Porto]. Repositório Aberto da Universidade do Porto. Available from: https://repositorio-aberto.up.pt/handle/10216/81446

Platt HD, Platt MB (1994) Business cycle effects on state corporate failure rates. *J Econ Bus* 46: 113–127. https://doi.org/10.1016/0148-6195(94)90005-1

Platt HD, Platt MB (2002) Predicting corporate financial distress: Reflections on choice-based sample bias. *J Econ Financ* 26: 184–199. https://doi.org/10.1007/bf02755985

Pompe PPM, Bilderbeek J (2005) The prediction of bankruptcy of small- and medium-sized industrial firms. *J Bus Venturing* 20: 847–868. https://doi.org/10.1016/j.jbusvent.2004.07.003

PORDATA (2022) *Taxa de mortalidade das empresas: total e por sector de actividade económica*. Available from: https://www.pordata.pt/Portugal/Taxa+de+mortalidade+das+empresas+total+e+por+sector+de+actividade+económica-2888

da Rosa RFC (2017) *Risco de falência de PME: evidência no setor da construção em Portugal* [Dissertação de Mestrado, Universidade de Aveiro]. Repositório Institucional da Universidade de Aveiro. Available from: http://hdl.handle.net/10773/23050

Sánchez-Lasheras F, De Andrés J, Lorca P, et al. (2012) A hybrid device for the solution of sampling bias problems in the forecasting of firms' bankruptcy. *Expert Syst Appl* 39: 7512–7523. https://doi.org/10.1016/j.eswa.2012.01.135

dos Santos AR, Silva N (2019) Sectoral concentration risk in Portuguese banks' loan exposures to non-financial firms. *Banco Portugal Econ Stud,* 1–17. https://www.bportugal.pt/en/paper/sectoral-concentration-risk-portuguese-banks-loan-exposures-non-financial-firms

Serrano-Cinca C, Gutiérrez-Nieto B, Bernate-Valbuena M (2019) The use of accounting anomalies indicators to predict business failure. *Eur Manage J* 37: 353–375. https://doi.org/10.1016/j.emj.2018.10.006

Shi Y, Li X (2019) An overview of bankruptcy prediction models for corporate firms: A systematic literature review. *Intang Cap* 15: 114–127. https://doi.org/10.3926/ic.1354

Shumway T (2001) Forecasting bankruptcy more accurately: A simple hazard model. *J Bus* 74: 101–124. https://doi.org/10.1086/209665

Silva AFR (2014) *Bankruptcy forecasting models civil construction* [Dissertação de Mestrado, Instituto Universitário de Lisboa]. Repositório do Iscte—Instituto Universitário de Lisboa. Available from: http://hdl.handle.net/10071/10978

Succurro M, Arcuri G, Costanzo GD (2019) A combined approach based on robust PCA to improve bankruptcy forecasting. *Rev Account Financ* 18: 296–320. https://doi.org/10.1108/RAF-04-2018-0077

Sulaiman MS, Abood MM, Sinnakaudan SK, et al. (2021) Assessing and solving multicollinearity in sediment transport prediction models using principal component analysis. *ISH J Hydraul Eng* 27: 343–353. https://doi.org/10.1080/09715010.2019.1653799

Taffler RJ (1984) Empirical models for the monitoring of UK corporations. *J Bank Financ* 8: 199–227. https://doi.org/10.1016/0378-4266(84)90004-9

Tinoco MH, Holmes P, Wilson N (2018) Polytomous response financial distress models: The role of accounting, market and macroeconomic variables. *International Review of Financial Analysis*, *59*, 276–289. https://doi.org/10.1016/j.irfa.2018.03.017

Tinoco MH, Wilson N (2013) Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *Int Rev Financ Anal* 30: 394–419. https://doi.org/10.1016/j.irfa.2013.02.013

Tserng HP, Chen PC, Huang WH, et al. (2014) Prediction of default probability for construction firms using the logit model. *J Civ Eng Manag* 20: 247–255. https://doi.org/10.3846/13923730.2013.801886

Tserng HP, Liao HH, Jaselskis EJ, et al. (2012) Predicting Construction Contractor Default with Barrier Option Model. *J Constr Eng M* 138: 621–630. https://doi.org/10.1061/(asce)co.1943-7862.0000465

Uthayakumar J, Metawa N, Shankar K, et al. (2020) Financial crisis prediction model using ant colony optimization. *Int J Inf Manage* 50: 538–556. https://doi.org/10.1016/j.ijinfomgt.2018.12.001

Vieira ES, Pinho C, Correia C (2013) Insolvency prediction in the Portuguese construction industry. *Marmara J Eur Stud* 21: 143–164. Available from: https://www.researchgate.net/publication/263037318_Insolvency_prediction_in_the_Portuguese_construction_industry

Vo DH, Pham BNV, Ho CM, et al. (2019) Corporate Financial Distress of Industry Level Listings in Vietnam. *J Risk Financ Manage* 12: 1–17. https://doi.org/10.3390/jrfm12040155

Wellek S, Blettner M (2012) On the Proper Use of the Crossover Design in Clinical Trials. *Dtsch Arztebl Int* 109: 276–281. https://doi.org/10.3238/arztebl.2012.0276

Wood MD, Simmatis LER, Jacobson JA, et al. (2021) Principal Components Analysis Using Data Collected From Healthy Individuals on Two Robotic Assessment Platforms Yields Similar Behavioral Patterns. *Front Hum Neurosci* 15: 1–12. https://doi.org/10.3389/fnhum.2021.652201

Wu CH, Tzeng GH, Goo YJ, et al. (2007) A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. *Expert Syst Appl* 32: 397–408. https://doi.org/10.1016/j.eswa.2005.12.008

Wu T (2021) Quantifying coastal flood vulnerability for climate adaptation policy using principal component analysis. *Ecol Indic* 129: 1–12. https://doi.org/10.1016/j.ecolind.2021.108006

Yan D, Chi G, Lai KK (2020) Financial Distress Prediction and Feature Selection in Multiple Periods by Lassoing Unconstrained Distributed Lag Non-linear Models. *Mathematics* 8: 1–29. https://doi.org/10.3390/math8081275

Young G (1995) Company liquidations, interest rates and debt. *Manch Sch Econ Soc Stud* 63: 57–69. https://doi.org/10.1111/j.1467-9957.1995.tb01448.x

Zavgren CV (1985) Assessing the Vulnerability to failure of American Industrial Firms: a Logistic Analysis. *J Bus Financ Account* 12: 19–45. https://doi.org/10.1111/j.1468-5957.1985.tb00077.x

Zhang Z (2016) Variable selection with stepwise and best subset approaches. *Ann Transl Med* 4: 1–6. https://doi.org/10.21037/atm.2016.03.35

Žiković IT (2016) Modelling the impact of macroeconomic variables on aggregate corporate insolvency: Case of Croatia. *Econ Res-Ekon Istraz* 29: 515–528. https://doi.org/10.1080/1331677X.2016.1175727

Zoričák M, Gnip P, Drotár P, et al. (2020) Bankruptcy prediction for small- and medium-sized companies using severely imbalanced datasets. *Econ Model* 84: 165–176. https://doi.org/10.1016/j.econmod.2019.04.003