

A logistic regression approach to long-term bankruptcy prediction: The role of financial and non-financial indicators

Denis Kuster¹, Aleksandar Majstorovic², Veljko Dmitrovic³

¹ Schneider Electric LLC, Financial Controlling Department, Serbia, ORCID: 0000-0001-6697-6792, kuschter@yahoo.com;

² University Business Academy, Faculty of Economics and Engineering Management, Department of Economics, Serbia, ORCID: 0000-0002-6884-9209, aleksandar.majstorovic@fimek.edu.rs;

³ University of Belgrade, Faculty of Organizational Sciences, Department of Financial Management and Accounting, Serbia, ORCID: 0000-0001-8924-9758, veljko.dmitrovic@fon.bg.ac.rs.

Abstract: The main aim of the research is to examine the possibility of developing logistic regression (LR) model that could reliably predict the bankruptcy of Serbian companies three years in advance based on financial and non-financial variables. This is important both for business owners and external stakeholders. Owners can predict failure on time and define remedial measures and action plans in accordance with that. External stakeholders, on the other hand, can use these models to identify financial risks before deciding to start cooperation with a specific company. The main motive for the research stems from the lack of bankruptcy prediction models in the scientific community of the Republic of Serbia, especially when it comes to long-term prediction. It is necessary to predict bankruptcy early enough to be able to take measures. A prediction one year in advance, which is a common case in the existing literature, may be too late to preserve the business's future. According to the authors' findings, no long-term prediction models have been developed for the Serbian market. Existing traditional models are developed for foreign countries, meaning they are not suitable for developing countries like Serbia. The research sample includes 94 companies of all sizes and is balanced: half of the companies are healthy, and the other half are bankrupt. A total of 36 financial and 7 non-financial independent variables are included in the modelling. Financial analysis is done in MS Excel, while statistical analysis (logistic regression) is done in IBM's SPSS program v. 26. The research results demonstrate that statistical and financial analyses are effective for bankruptcy prediction modelling, considering that the generated model has significant predictive (classification) power of 80%.

Keywords: Bankruptcy, insolvency, statistics, strategy, financial analysis, modelling.

JEL Classification: G33, M40, C44, C52.

APA Style Citation: Kuster, D., Majstorovic, A., & Dmitrovic, V. (2025). A logistic regression approach to long-term bankruptcy prediction: The role of financial and non-financial indicators. *E&M Economics and Management*, 28(2), 165–179. <https://doi.org/10.15240/tul/001/2025-5-010>

Early Access Publication Date: March 31, 2025.

Introduction

Insolvency is always an actual and important topic. Kovacova et al. (2019) claim that the issue of bankruptcy prediction is spread

worldwide. Various explanatory variables and techniques have been used in the construction of bankruptcy prediction models. A company is said to be insolvent or under financial

distress if it is unable to pay its debts as they become due, which is aggravated if the value of the firm's assets is lower than its liabilities (Becerra et al., 2005).

According to Maras and Kozar (2021), bankruptcy proceedings in the Republic of Serbia have been prescribed by the Law on bankruptcy (The Republic of Serbia) and are initiated by the petition of the creditor, debtor or liquidator as authorized petitioners. Adopting a positive decision on such petition, in case the court determines the existence of one of the prescribed legal bankruptcy conditions, represents the opening of the bankruptcy proceedings. The main aim of bankruptcy proceedings is to enable the most favourable collective settlement of creditors' claims by selling the property of a bankrupt legal entity (Kuster, 2022).

The main aim of the research is to develop logistic regression model for bankruptcy prediction three years in advance. Despite various models developed worldwide, the necessity of developing a specific model for the Republic of Serbia is present. The traditional performance rankings, such as return on equity and return on assets, are not sufficient indicators of business performance (Milosevic et al., 2021). Not many research papers in Serbia include non-financial variables, which is the added value of the research. In addition, to the best authors' knowledge, no research has been conducted for the Serbian economy to predict bankruptcy three years in advance.

The main motive for the research stems from the fact the existing bankruptcy prediction models are, in most cases, developed for other countries that have different legal, accounting, and other requirements. Bankruptcy predictions are not explored enough in Serbia, nor in the Balkans, which emphasizes the importance of this research. Alihodzic (2013) and Muminovic et al. (2011) confirmed that the existing models are not always suitable for the Serbian market. Nikolic et al. (2018) state that one of the biggest problems in the Serbian economy is a large number of bankruptcies in the SME sector. Furthermore, according to the data published by the Business Registers Agency (2023), the bankruptcy rate for 2022 is 4.3%, with 5,817 companies being in the bankruptcy proceedings (out of 135,490). This nominates Serbia as an interesting research field, considering the fact the current bankruptcy in the EU is 3.1% (Trading

Economics, 2024). Bankruptcy prediction models support the economic stability and growth of Serbia by identifying vulnerable businesses at an early stage. This enables stakeholders to define intervention measures and reduce risks. Besides the above-mentioned, bankruptcy prediction modelling is important as a support for better decision-making in developing countries such as Serbia.

Through a comprehensive review of existing research encompassing a wide range of bankruptcy prediction models (e.g., Bellovary et al., 2007; Ratajczak et al., 2022), it has been determined that the overall accuracy of bankruptcy prediction models generally falls within the following range: 65–85% (for prediction three years in advance). Many authors, e.g., Du Jardin, 2009; Sharma and Mittal (2024), stated that, besides financial ratios, the non-financial variables are important in bankruptcy prediction. That being said, the research hypothesis is defined as follows:

H1: Bankruptcy in the Republic of Serbia can be predicted three years in advance with an overall accuracy of at least 70% using logistic regression that incorporates both financial and non-financial indicators.

H1.1: The logistic regression algorithm can detect patterns in data that predict bankruptcy for Serbian companies three years in advance, achieving an overall accuracy of at least 70%.

H1.2: Financial indicators serve as significant predictors of bankruptcy in Serbia three years in advance, with at least one financial ratio included in the final model.

H1.3: Non-financial indicators serve as significant predictors of bankruptcy in Serbia three years in advance, with at least one indicator included in the final model.

The research is organized as follows: the first section includes an overview of the existing literature related to bankruptcy and financial distress predictions. Afterwards, the sample and research methodology are presented. The fourth section is split into two primary sections. The first one explains model development, while the second section shows model accuracy testing. The last section is related to concluding remarks.

1 Theoretical background

Shi and Li (2019) generated a review of bankruptcy prediction models based on 321 most

Tab. 1: Bankruptcy prediction models – review

Classical statistical methods		
Ranking	Method	Number of papers
1	Logistic regression (logit)	123
2	Discriminant analysis	52
3	Multivariate discriminant analysis and Z-score	33
4	Hazard	19
5	Logit and probit	7
6	Probit	6
Machine learning and artificial intelligence models		
Ranking	Method	Number of papers
1	Neural network	56
2	Support vector machine	32
3	Decision tree	21
4	Generic algorithm	20
5	Fuzzy	17
6	Rough set	13
7	Data mining	11

Source: Shi and Li (2019)

relevant research papers. As can be seen in Tab. 1, the most frequently used technique in bankruptcy prediction is logistic regression (123/321 research papers). It is important to point out that one research paper can use more than one analysis method, and therefore, the sum of all the results is not exactly 321 papers.

The majority of business prediction models belong to the quantitative ones and their inputs are the results of technical financial analysis (Klepac & Hampel, 2018). In the literature review, we focused on the bankruptcy prediction models that are based on logistic regression only, as that is the technique used in the paper.

Ohlson (1980) is one of the pioneers of logistic regression analysis for the prediction of corporate bankruptcy. Results of the mentioned research indicate that net result to total assets and total debt to total assets are the best factors for bankruptcy prediction. Luoma and Laitinen (1991) developed a logit model with seven indicators based on a sample of Finnish firms. The predictive power of the model is 73.5% and 71% for bankrupt and

non-bankrupt firms, respectively. Fletcher and Gross (1993) proposed logit model with 3 factors and general application. The classification power of the model is 71%.

Laitinen and Laitinen (2000) applied logit analysis and developed four models to predict bankruptcy even three years in advance. Step-wise logit model with ten variables included has an accuracy of 75, 65 and 67%, one, two and three years before bankruptcy occurs. Gaeremynck and Willekens (2003) proposed a logit analysis with 8 factors for Belgian private companies, and generated a model of 72% accuracy. Altman and Sebato (2007) generated a logistic regression model using unlogged and logged variables with total classification power of 75% and 87%, respectively. The focus was on small and medium entities (SMEs), and the sample consisted of 2,000 US firms. Hu and Ansell (2009) proposed logistic regression among other models they developed in the research. Variable selection was done using stepwise method and the final scope included 10 both internal and external variables. The developed LR model based on internal factors has

predictive power of 90, 87, and 82% for prediction of bankruptcy, one, two and three years in advance.

Tseng and Hu (2010) introduced four models for bankruptcy prediction, and one of them is logit model. Holdout approach was used, and a total of 77 companies was divided into training (61 firms) and testing sets (16 firms) on a random basis. Three variables were selected as important using the stepwise method: working capital to operational expenses, after-tax result to total assets and Δ cash to total liabilities. Developed LR model has predictive power of 77% for the training dataset and 86% for the testing set. Cho et al. (2010) suggested four LR models based on different variables selection (decision trees and stepwise methods). The sample included 1,000 companies, and ratio of failed and existing companies is 1:1. The highest performing model has predictive power of 72% and includes seven profitability, liquidity, activity, stability, labour productivity but also growth ratios. Yoon and Kwon (2010) generated LR model based on a sample of 10,000 companies. The number of solvent and insolvent companies is equal. The final model has a classification power of 70%.

Divsalar et al. (2012) based their research on a 136-firm sample, which was not balanced. Developed LR model includes four variables (profitability, solvency and liquidity scope) after reduction and has 76% classification power for testing dataset. Shie et al. (2012) included 54 companies in the research and developed model that has 73% predictive power, based on profitability, leverage, cash, liquidity and efficiency ratios. Stanisic et al. (2013) proposed LR modelling for Serbian companies. Developed model has precision of 75%. The model included raw variables (EBITDA and number of employees), but also turnover and debt ratios.

Tseng et al. (2014) developed LR model by combining five groups of ratio indicators (liquidity, leverage, turnover, profitability and market group) and generated three models for predicting bankruptcy up to three years in advance. The highest performing model has the following classification power for one, two and three years in advance (respectively): 79, 79 and 75%. Gordini (2014) applied LR modelling on a sample of 3,100 companies. The variables included are mostly in the scope of profitability, turnover, liquidity, debt and cash ratios. Those were selected based on VIF and

stepwise analysis. Three models were developed. Accuracy of the models for bankruptcy prediction one, two and three years in advance follows (respectively): 69, 68 and 67%. Du Jardin (2015) developed LR multi-sector models (retail, construction and services) based on several test samples. The best performing model for all sectors has following classification power: 81% one year before bankruptcy, 78.5% two years before bankruptcy and 75.5% three years before bankruptcy. Variables included in modelling are: accounts payable turnover ratio, quick liquidity ratio, receivables turnover ratio, current ratio and net-income-to-shareholder-funds ratio.

Besic Obradovic et al. (2018) generated log-regression model for insolvency prediction of Serbian companies one year in advance. Overall model accuracy 88% for the training sample and 82.5% for the testing dataset. The model includes the following ratio variables: working capital, self-financing and business effectiveness. Cheong and Ramasamy (2019) examined the bankruptcy prediction of 536 bankrupted and non-bankrupted banks in the USA. Logit approach was in place, and the conclusion is that high return on equity, as well as the adequacy of capital, have a negative impact on bankruptcy. Logit model for Greece firms, developed by Papana and Spyridou (2020), performs with the following classification power: 66, 62.5, and 68%, one, two and three years prior to bankruptcy, respectively. After variables reduction using *t*-test, total of 9 ratios (mostly profitability, liquidity, asset structure and solvency) was included in the modelling. Sricharoenchit and Hensawang (2021) developed a bankruptcy prediction model in the automotive industry based on a sample of companies from Thailand. Financial ratios and corporate governance variables were included. The final model has predictive power of 75%. Papik and Papiková (2024) suggested a comparison of log regression and machine learning methods based on a total of 6 techniques. The logistic regression model reached total accuracy of 68.2%.

2 Research methodology

This study presents a new bankruptcy prediction model specifically designed for Serbian companies, aiming to predict bankruptcy three years in advance. Our model stands apart from existing ones by using a unique set

of predictors. We independently calculated both financial and non-financial indicators in MS Excel, carefully screened the data, and carried out the model's training and testing in SPSS. This custom-built approach allowed us to tailor each step of the development process, ensuring the model's alignment with Serbia's regulatory and business environment.

The research sample includes 94 entities of various sizes from the Republic of Serbia. Half of the companies are solvent, and the other half went bankrupt. The sample is balanced based on methodology of many other authors that examined this subject, meaning there

are 47 solvent and 47 bankrupt companies in the sample. Companies are not only balanced in terms of solvency/bankruptcy, but also in terms of age and turnover, as this sort of modelling requires an adequate sample. One part of the sample (80%) is used for model development, and the other part (20%) is used for model testing.

It is important to isolate the external impact of crises whenever possible and choose only stable business years in the sample. Therefore, we are focusing on the period before any time-lagged COVID-19 pandemic impacts. Companies that started bankruptcy proceedings

Tab. 2: Independent variables overview

Category	#	Symbol	Variable calculation method	#	Symbol	Variable calculation method
Financial variables	1	ROA	Net income/total assets	19	Inv / S	Inventory/sales
	2	CR	Current assets/current liabilities	20	OI / TA	Operating income/total assets
	3	WC / TA	Working capital/total assets	21	OCF / S	Operations cash flow/sales
	4	RER	Retained earnings/total assets	22	NI / S	Net income/sales
	5	EBIT / TA	EBIT/total assets	23	LTD / TA	Long-term debt/total assets
	6	S / TA	Sales/total assets	24	NW / TA	Net worth/total assets
	7	QR	(Current assets-inventories)/current liabilities	25	C / CL	Cash/current liabilities
	8	TD / TA	Total debt/total assets	26	OCF / CL	Operations cash flow/current liabilities
	9	CA / TA	Current assets/total assets	27	WC / S	Working capital/sales
	10	DR	Total liabilities/total assets	28	Cap / A	Capital/assets
	11	C / TA	Cash/total assets	29	NS / TA	Net sales/total assets
	12	OCF / TA	Operations cash flow/total assets	30	NW / TL	Net worth/total liabilities
	13	OCF / TL	Operations cash flow/total liabilities	31	NCI	No-credit interval
	14	LIQ	Current liabilities/total assets	32	$CFNI / D$	Cash flow (net income)/debt
	15	OCF / TD	Operations cash flow/total debt	33	OCF	Operations cash flow
	16	QA / TA	Quick assets/total assets	34	OE / OI	Operating expenses/operating income
	17	CA / S	Current assets/sales	35	QA/S	Quick assets/sales
	18	EBIT / Int	EBIT/interest	36	S / Inv	Sales/inventory
Non-financial variables	1	LOG_TA	Log(total assets)	5	$\ln(S / Inv)$	Log(sales/inventory)
	2	Headcount	Total number of employees	6	Log(S)	Log(sales)
	3	L / C	Labour/capital employed	7	$\ln(\text{Age})$	$\ln(\text{operating years before bankruptcy})$
	4	OY	Operating years before bankruptcy	–	–	–

Source: own

in 2020 are considered as bankrupted. The others still operating in the mentioned year are considered healthy. Data from the three business years preceding the bankruptcy declaration are utilized for calculating variables, as the primary goal of the research is to develop a model capable of accurately predicting bankruptcy three years ahead.

As presented in Tab. 2, total of 36 financial variables and 7 non-financial variables is included in the modelling. Financial variables are selected based on research of Bellovary et al. (2007), while non-financial variables were proposed by the following authors: Frydman et al. (1985), Tian and Yu (2017), Leshno and Spector (1996), Barniv et al. (1997), Park and Han (2002). Data for independent variables calculation is available on the Serbian Business Registers Agency website.

Status of solvency or bankruptcy is a dependent variable. In order to perform statistical modelling, this variable has been coded as follows: 0 = bankrupted entity and 1 = solvent entity. IBM's SPSS v.26 program and binary LR (logistic regression – LR), with “Stepwise – Forward: LR” algorithm is used for statistical modelling. Field (2009) stated that in case of forward LR algorithm, computer begins with a model that includes only a constant and then adds single predictors to the model based on specific

criteria. The criteria is the value of score statistic: the variable with the most significant score statistic is added to the model. Therefore, this methodology is the right choice when no previous research is available or when a lot of independent variables are included in the analysis initially. LR uses the log-ratio to assign a company to either bankrupt or non-bankrupt class (Veganzones & Séverin, 2018). Logistic regression is a widely used statistical technique. It is used to predict the probability of an event's occurrence by fitting data onto a logistic curve. Logistic regression estimates the value of two class labels or sequence variables (Tsai & Hsu, 2011). Logistic function can be expressed as follows:

$$P(\text{bankruptcy}) = \frac{1}{1 + e^{-z}} \quad (1)$$

where: P – probability of default; z – explanatory variables linear regression. If z is given input, its output is $f(z)$.

The definition of variables of input z follows:

$$z = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k \quad (2)$$

where: β_0 – intercept; $\beta_1, \beta_2, \beta_k$ – coefficients of x_1, x_2, x_k . When all the independent variables are equal to zero, the intercept is the z value.

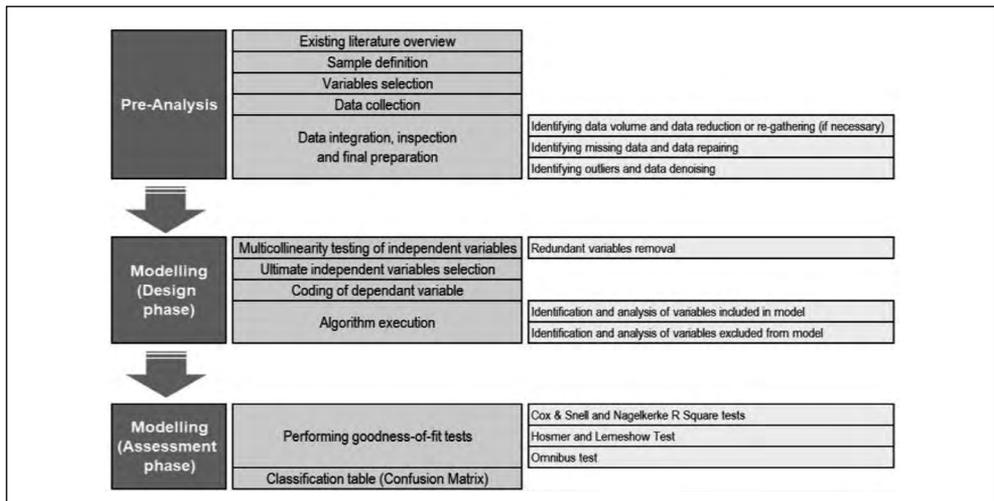


Fig. 1: Research framework

Source: Kuster (2023)

The quality of the developed models will be examined using the following tests: Cox and Snell R^2 , Nagelkerke R^2 , Hosmer and Lemeshow, Omnibus and Wald. Also, confusion matrix will be in place with the following indicators: accuracy, precision, recall/sensitivity and specificity:

$$\text{Accuracy} = \frac{\text{True positive} + \text{True negative}}{\text{True positive} + \text{True negative} + \text{False positive} + \text{False negative}} \quad (3)$$

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \quad (4)$$

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \quad (5)$$

$$\text{Specificity} = \frac{\text{True negative}}{\text{True negative} + \text{False positive}} \quad (6)$$

The whole research framework indicating a step-by-step program is presented in Fig. 1. The research consists of three phases: pre-analysis, model design phase and model testing phase.

3 Results

3.1 Bankruptcy prediction model – Design phase

According to Salmeron et al. (2018), the presence of high collinearity in a regression model suggests that the conclusions drawn from the analysis may be unreliable. Multicollinearity often occurs in analyses due to high intercorrelation of variables (Gasic et al., 2023). Collinearity detection has to be the first step in every statistical analysis. Analysis is done via variance inflation factor (VIF). VIF analysis was conducted for all the independent variables.

Tab. 3: Variance inflation factor (VIF)

Variables	Collinearity statistics	
	Tolerance	VIF
<i>RER</i>	0.455	2.196
<i>OCF/TL</i>	0.698	1.434
<i>LIQ</i>	0.712	1.405
<i>QA/TA</i>	0.539	1.856
<i>EBIT/Int</i>	0.808	1.238
<i>OI/TA</i>	0.628	1.591
<i>OCF/S</i>	0.410	2.440
<i>C/CL</i>	0.620	1.614
<i>OCF/CL</i>	0.561	1.783
<i>NW/TL</i>	0.396	2.523
<i>NCI</i>	0.765	1.308
<i>LOG_TA</i>	0.448	2.230
<i>OCF</i>	0.525	1.906
<i>OE/OI</i>	0.458	2.181
<i>S/Inv</i>	0.465	2.151
<i>L/C</i>	0.496	2.017
<i>Ln(Age)</i>	0.650	1.538

Note: Coefficients – dependent variable: status 0/1.

Source: own

Cohen et al. (2003) and O'Brien (2007) stated that VIF values higher than 10 indicate a problem. Menard (2001) claims that VIF value higher than 5 is cause for concern. After the removal of redundant variables, the problem of multicollinearity is not present since no value exceeds 3.0. Only variables that passed the test are listed in Tab. 3.

As previously mentioned, stepwise forward logistic regression is used. The main advantage of this method is the fact that this algorithm selects only those variables that have a huge impact in outcome prediction, which further leads to stable models without redundant variables and overfitting. The bankruptcy prediction model

was generated in three steps, where the third is the final with the best accuracy (Tab. 4).

It is worth mentioning that the parameters for all three steps will be presented as a part of the research and modelling process, but the analysis will always be focused on the final model parameters – Step 3. Three variables (*RER*, *OCF/CL* and *LOG_TA*) are included in the final equation for bankruptcy prediction. In Tab. 5, we tested the importance of the selected variables, i.e., what happens if some of them are removed. Since the "Sig. of change" is lower than the confidence level 0.05, it can be concluded that all three variables contribute significantly to the model and, therefore, should remain in the equation.

Tab. 4: Variables included in the bankruptcy prediction model (variables in the equation)

Step	Variable	B	S. E.	Wald	df	Sig.	Exp (B)
Step 1 ^a	<i>RER</i>	5.395	1.636	10.878	1	0.001	220.403
	Constant	-1.004	0.364	7.596	1	0.006	0.366
Step 2 ^b	<i>RER</i>	7.126	2.035	12.256	1	0.000	1,243.314
	<i>OCF/CL</i>	0.459	0.369	1.542	1	0.214	1.582
Step 3 ^c	Constant	-1.220	0.399	9.370	1	0.002	0.295
	<i>RER</i>	7.397	2.092	12.504	1	0.000	1,630.993
	<i>OCF/CL</i>	0.625	0.281	4.969	1	0.026	1.869
	<i>LOG_TA</i>	-1.388	0.496	7.827	1	0.005	0.250
	Constant	6.224	2.638	5.566	1	0.018	504.920

Note: a – variable(s) entered on step 1: *RER*; b – variable(s) entered on step 2: *OCF/CL*; c – variable(s) entered on step 3: *LOG_TA*.

Source: own

Tab. 5: Selected variables removal testing (model if term removed)

Variable		Model log likelihood	Change in -2 log likelihood	df	Sig. of the change
Step 1	<i>RER</i>	-52.679	16.778	1.000	0.000
Step 2	<i>RER</i>	-50.857	20.700	1.000	0.000
	<i>OCF/CL</i>	-44.290	7.567	1.000	0.006
Step 3	<i>RER</i>	-46.334	21.275	1.000	0.000
	<i>OCF/CL</i>	-42.402	13.412	1.000	0.000
	<i>LOG_TA</i>	-40.507	9.622	1.000	0.002

Source: own

Developed logistic regression models for predicting bankruptcy three years in advance can be written mathematically as follows:

$$P(x) = \frac{1}{1 + e^{-(6.224 - 1.388 \text{ LOG}(TA) + 0.625 \text{ OCFCL} + 7.397 \text{ RER})}} \quad (7)$$

3.2 Bankruptcy prediction model – Assessment phase

To begin with, Wald test results are presented in Tab. 5. This test indicates if an independent variable has a statistically significant impact on the observed dependent variable. In other words, it demonstrates if a specific financial or non-financial variable is important for

bankruptcy prediction. Considering the fact that p -value for this test (Sig.) is lower than 0.05 in case of all the variables, it can be concluded that the selected indicators have a significant impact on prediction power of the developed model.

The omnibus test can be used to analyse whether or not the model is reliable. It demonstrates how well the LR model predicts results, and how well this model predicts the risk (Pallant, 2009). When significance for Omnibus test (Tab. 6) is less than 0.05, model is considered to be well-fitted. The developed model is well-fitted taking into account the significance level that is lower than 0.05.

Tab. 6: Omnibus tests (tests of model coefficients)

Step	Item	Chi-square	df	Sig.
Step 1	Step	16.778	1.000	0.000
	Block	16.778	1.000	0.000
	Model	16.778	1.000	0.000
Step 2	Step	7.567	1.000	0.006
	Block	24.345	2.000	0.000
	Model	24.345	2.000	0.000
Step 3	Step	9.622	1.000	0.002
	Block	33.967	3.000	0.000
	Model	33.967	3.000	0.000

Source: own

Tab. 7: Cox & Snell and Nagelkerke R tests (model summary)

Step	-2 log likelihood	Cox & Snell R square	Nagelkerke R square
1	88.580	0.198	0.264
2	81.013	0.274	0.365
3	71.392	0.360	0.481

Source: own

According to Beslic Obradovic et al. (2018), the measure of R^2 ranges from 0 to 1, and values higher than 0.4 indicate that LR model is well-fitted. If the value of R^2 is 1, that means

the independent variables (predictors) can completely explain the dependent variable, which is rarely the case in practice. Braun et al. (2013) stated that values in the range 0.24–0.32 are

considered as acceptable for these indicators. Given the values 0.36 and 0.48 (Tab. 7), it can be said that the bankruptcy prediction model is well-fitted.

The Hosmer-Lemeshow test (Tab. 8) is based on the calculation of λ^2 observed (actual) and predicted (expected) value of the dependent variable. It compares the original variables

and prediction, in other words, it compares if there is a statistically significant difference between them (Pallant, 2009). Zenzerovic and Perusko (2009) stated that a model is well-fitted when Hosmer and Lemeshow test exceeds the level of 0.05. Considering the significance levels for the developed model, it can be concluded that it is well-fitted.

Tab. 8: Hosmer and Lemeshow test

Step	Chi-square	df	Sig.
1	9.171	7	0.241
2	11.527	8	0.174
3	7.948	8	0.439

Source: own

Tab. 9: Confusion matrix

Observed			Predicted					
			Selected cases (status 0/1)		Percentage correct (%)	Unselected cases (status 0/1)		Percentage correct (%)
			0	1		0	1	
Step 1	Status 0/1	0	29	9	76.3	5	4	55.6
		1	16	22	57.9	5	4	44.4
	Overall percentage (%)				67.1			50.0
Step 2	Status 0/1	0	30	8	78.9	5	4	55.6
		1	15	23	60.5	4	5	55.6
	Overall percentage (%)				69.7			55.6
Step 3	Status 0/1	0	30	8	78.9	6	3	66.7
		1	7	31	81.6	2	7	77.8
	Overall percentage (%)				80.3			72.2

Note: Classification table – the cut value is 0.500.

Source: own

The overall accuracy of the final developed model is 80.3% (72.2% for testing sample). The precision of the model is 78.9% (66.7% for testing sample). Recall (sensitivity) has value of 81.1% (75.0% for testing sample). Specificity of the model is 79.5% (70.0% for the testing sample). All the results are presented in Tab. 9.

4 Discussion

The developed models are useful for business owners, suppliers, customers, and banks. On the one hand, business owners can calculate the LR equation using their ratios and non-financial data and, therefore, identify whether there is a risk of bankruptcy. On the other hand, the same process can be executed by all

the other external stakeholders. Detailed statistical analysis with explanations can be useful for other authors as a starting point for further research that is based on logistic regression. Also, careful selection of variables can be useful for future research to develop new bankruptcy prediction models using different samples. An important contribution of this research is the inclusion of both financial and non-financial variables, but also development of the model that predicts bankruptcy three years in advance. In most cases, predicting bankruptcy for only one year before its occurrence does not leave enough time for the application of any healing measures. Significant advantage of logistic regression models compared other machine learning models (e.g., decision trees, neural networks) is the fact that everyone can use the regression equation easily by inputting specific values for the variables (ratios); there is no black box dilemma.

The research has several limitations. To begin with, model is developed as general, meaning it can further experience difficulties trying to classify a company as part of some specific economic activity. The idea for further research is to collect a sample of companies that belong only to one economic activity and develop activity-based model. Also, the accuracy of prediction is mainly driven by the correctness of financial statements. Furthermore, the research sample is relatively small, but that limitation is compensated by using advanced statistical program and careful data analysis. Another idea for the next research is to involve more observations in the sample, or even to develop a model that could be applicable for all Balkan countries. To conclude with, this model does not include any macroeconomic variables, and those are important for predicting bankruptcy in the long run. Future research should include macroeconomic and some other non-financial variables in order to raise the classification power of the model.

The developed model that predicts insolvency three years in advance is a significant advancement in Serbian economic theory and practice. Three parameters are identified as important for the prediction: retained earnings ratio, operations cashflow to current liabilities and logarithm of total assets. Retained earnings are an important indicator, demonstrating cumulative profits of a company. Higher levels of retained earnings indicate that the company

is capable of generating profits. Companies with high retained earnings are less dependent of external financing. Low ratio of retained earnings to total assets may be an indicator of higher bankruptcy risk, due to reliance on external debts and insufficient profits for re-investment. Strategies like cost controlling and revenue boosting, as well as rising operational efficiency are convenient way to boost retained earnings. When it comes to log value of total assets, it is important that larger assets base means higher collateral for secured financing. Companies should work on increasing of their assets base through acquiring new technologies, as well as entering new markets. Operational cashflow compared to current liabilities demonstrates if a company is able to cover its short-term obligations using cash generated only from operations related to the core business. Higher values of the indicator are a sign of better liquidity. Companies should strategically work on optimizing their cash flows and building a sustainable cash pool from their core business activities. The first step in the mentioned process is optimization of working capital, but also reduction of unnecessary expenses.

Conclusions

Bankruptcy will never be outdated; it is always current, since there is no economy that is not affected by companies' insolvency. This question is especially important for developing countries like Serbia. It is important to develop prediction models specifically for one country due to differences in economic situation, business climate, legal and other requirements. Bankruptcy prediction models are treated as tools of higher financial analysis (Kubenka et al., 2021). There are many differences and disagreements about the idea and how to apply the methods of modern accounting analysis, and there is a certain number of differences regarding the range and limits of modern accounting analysis (Majstorovic, 2016). A bankruptcy prediction model based on logistic regression allows stakeholders to reliably assess a company's financial health using just a basic understanding of ratio indicators, without the need for in-depth accounting knowledge or expertise in financial analysis.

As previously mentioned, not many research papers globally were oriented towards developing models that could predict bankruptcy three years in advance. Best to the author's knowledge, no research in Serbia has covered

this wide prediction timeline yet. The developed model for prediction of bankruptcy three years in advance has overall classification accuracy in the range of 70–75% (for the test sample), together with a model of Tseng et al. (2014). The developed model has lower performance than the following prediction models: Hu and Ansell (2009) and Du Jardin (2015). The developed Y-3 LR model performs better than the following prediction models: Laitinen and Laitinen (2000), Gordini (2014) and Papan and Spyridou (2020).

The research followed a structured process, starting with a relevant literature review in order to establish solid base knowledge. Afterwards, a sample of 94 entities was defined (divided as 80%:20% = training:test sample), followed by selection of explanatory variables. The research initially included a combination of more than 40 financial and non-financial variables and the final developed model includes three of them. In the first step, 17 variables out of the initial 43 were selected using VIF analysis, as this test indicates the existing of multicollinearity that is not acceptable in LR modelling. Further reduction is done via stepwise LR algorithm where only 3 variables were selected in the third step as important for the model (*RER*, *OCFCL* & *LOG_TA*). The model's performance was evaluated through goodness-of-fit tests (Omnibus test, Cox & Snell test, Nagelkerke test, and Hosmer & Lemeshow test) and assessed with a confusion matrix for both training and testing datasets, providing a robust validation of the model's predictive accuracy and relevance within Serbia's economic environment. Final accuracy of the developed model is 80% (72% for the test sample). In other words, this study resulted in a new, unique prediction model aiming to predict bankruptcy in Serbian three years before it occurs.

Practical findings of the research highlight several important aspects: i) warning signs identification – the model is easy to use, and it enables the stakeholders to take proactive measures; ii) financial institutions support – the developed model is useful for investors, creditors and other institutions in the Republic of Serbia as it helps in managing risks and supports decision making; iii) promotion of sustainable business – bankruptcy prediction models like this motivate companies to take steps to improve financial health by addressing the problems that are identified by the model, and

that further leads to overall business environment stability; and iv) regulatory enhancements – the developed model showed that operations cashflow compared to current liabilities is one of the important predictors of bankruptcy. Poor cash management leads to financial problems, and therefore regulatory authorities may use this insight to launch new initiatives, like offering favourable loans to support businesses and help them maintain stable cashflows.

The idea for future research is to include even macroeconomic variables in the modelling to rise prediction accuracy. Also, the suggestion is to provide greater research sample or to develop a bankruptcy prediction model for specific economic activity (manufacturing, trade).

Serbia, as a developing economy, faces distinct economic and legal challenges that influence the financial health of its businesses. The country is still navigating a period of economic transition, using a Serbian sample for the model ensures that it accurately reflects the country's specific economic and regulatory environment. This is crucial, as there is currently no model that predicts bankruptcy in Serbia with a three-year lead time. Neighbouring Balkan countries such as Croatia, Bosnia and Herzegovina, North Macedonia, and Montenegro share similar economic structures, legal frameworks, and industrial profiles. These nations also face typical challenges encountered by developing economies. As a result, the model developed for Serbia could be a useful foundation for these countries, though it would need to be adjusted for local differences, including variations in business practices. In essence, while the model is tailored to Serbia, it has the potential to be adapted for other Balkan countries with some refinement. Pilot testing the model in these countries would offer valuable insights into its cross-border applicability and help determine the necessary modifications for improved predictive accuracy.

To conclude with, auxiliary hypotheses are accepted:

- i) Logistic regression algorithm is capable of identifying patterns in data related to predicting bankruptcy three years in advance as model of total accuracy of 80% was generated (72.2% for the test sample);
- ii) Financial variables serve as significant predictors of bankruptcy in Serbia three years in advance as two of them are included in the final LR equation (*RER* and *OCFCL*);

iii) Non-financial variables indicators serve as significant predictors of bankruptcy in Serbia three years in advance as one of them is included in the final LR equation (Log_TA).

Therefore, the main hypothesis is accepted. Bankruptcy in the Republic of Serbia can be predicted three years in advance with an overall accuracy of at least 70% using logistic regression that incorporates both financial and non-financial indicators.

Acknowledgements: *This work was supported by the University of Rijeka, under Grant: ZIP-UNIRI-2023-12.*

References

- Alihodzic, A. (2013). Testing the Kralicek DF indicator application on the Belgrade Stock Exchange. *Bankarstvo*, 42(3), 70–95.
- Altman, E. I., & Sabato, G. (2007). Modeling credit risk for SMEs: Evidence from the U.S. market. *Abacus*, 43(3), 332–357. <https://doi.org/10.1111/j.1467-6281.2007.00234.x>
- Barniv, R., Agarwal, A., & Leach, R. (1997). Predicting the outcome following bankruptcy filing: A three-state classification using neural networks. *International Journal of Intelligent Systems in Accounting, Finance & Management*, 6(3), 177–194. [https://doi.org/10.1002/\(sici\)1099-1174\(199709\)6:3<177::aid-isaf134>3.0.co;2-d](https://doi.org/10.1002/(sici)1099-1174(199709)6:3<177::aid-isaf134>3.0.co;2-d)
- Becerra, V. M., Galvão, R. K. H., & Abou-Seada, M. (2005). Neural and wavelet network models for financial distress classification. *Data Mining and Knowledge Discovery*, 11(1), 35–55. <https://doi.org/10.1007/s10618-005-1360-0>
- Bellovary, J. L., Giacomino, D. E., & Akers, M. D. (2007). A review of bankruptcy prediction studies: 1930 to present. *Journal of Financial Education*, 33, 1–42.
- Beslic Obradovic, D., Jaksic, D., Beslic Rupic, I., & Andric, M. (2018). Insolvency prediction model of the company: The case of the Republic of Serbia. *Economic Research-Ekonomska Istrazivanja*, 31(1), 139–157. <https://doi.org/10.1080/1331677x.2017.1421990>
- Braun, A., Müller, K., & Schmeiser, H. (2013). What drives insurers' demand for cat bond. *The Geneva Papers on Risk and Insurance – Issues and Practice*, 38(3), 580–611. <https://doi.org/10.1057/gpp.2012.51>
- Business Registers Agency. (2023). *Annual financial statement*. https://apr.gov.rs/upload/Portals/0/GFI_2023/Godisnji_izvestaj_o_poslovanju_privrede_u_2022.pdf
- Cheong, C. W. H., & Ramasamy, S. (2019). Bank failure: A new approach to prediction and supervision. *Asian Journal of Finance & Accounting*, 11(1), 111–140. <https://doi.org/10.5296/ajfa.v11i1.14455>
- Cho, S., Hong, H., & Ha, B.-C. (2010). A hybrid approach based on the combination of variable selection using decision trees and case-based reasoning using the Mahalanobis distance: For bankruptcy prediction. *Expert Systems with Applications*, 37(4), 3482–3488. <https://doi.org/10.1016/j.eswa.2009.10.040>
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates.
- Divsalar, M., Roodsaz, H., Vahdatinia, F., Norouzzadeh, G., & Behrooz, A. H. (2011). A robust data-mining approach to bankruptcy prediction. *Journal of Forecasting*, 31(6), 504–523. <https://doi.org/10.1002/for.1232>
- Du Jardin, P. (2009). Bankruptcy prediction models: How to choose the most relevant variables? *Bankers. Markets & Investors*, 98(1–2), 39–46.
- Du Jardin, P. (2015). Bankruptcy prediction using terminal failure processes. *European Journal of Operational Research*, 242(1), 286–303. <https://doi.org/10.1016/j.ejor.2014.09.059>
- Field, A. (2009). *Discovering statistics using SPSS*. SAGE.
- Fletcher, D., & Goss, E. (1993). Forecasting with neural networks: An application using bankruptcy data. *Information & Management*, 24, 159–167.
- Frydman, H., Altman, E. I., & Kao, D. L. (1985). Introducing recursive partitioning for financial classification: The case of financial distress. *The Journal of Finance*, 40(1), 269–291. <https://doi.org/10.1111/j.1540-6261.1985.tb04949.x>
- Gaeremynck, A., & Willekens, M. (2003). The endogenous relationship between audit-report type and business termination: Evidence on private firms in a non-litigious environment. *Accounting and Business Research*, 33(1), 65–79. <https://doi.org/10.1080/00014788.2003.9729632>
- Gasic, D., Berber, N., & Slavić, A. (2023). Effects of the compensation system and socio-demographic factors on the attitudes of

employees in the Republic of Serbia. *Management: Journal of Sustainable Business and Management Solutions in Emerging Economies*, 29(2). <https://doi.org/10.7595/management.fon.2022.0016>

Gordini, N. (2014). A genetic algorithm approach for SMEs bankruptcy prediction: Empirical evidence from Italy. *Expert Systems with Applications*, 41(14), 6433–6445. <https://doi.org/10.1016/j.eswa.2014.04.026>

Hu, Y.-C., & Ansell, J. (2007). Measuring retail company performance using credit scoring techniques. *European Journal of Operational Research*, 183(3), 1595–1606. <https://doi.org/10.1016/j.ejor.2006.09.101>

Klepac, V., & Hampel, D. (2018). Predicting bankruptcy of manufacturing companies in EU. *E&M Economics and Management*, 21(1), 159–174. <https://doi.org/10.15240/tul/001/2018-1-011>

Kovacova, M., Kliestik, T., Valaskova, K., Durana, P., & Juhaszova, Z. (2019). Systematic review of variables applied in bankruptcy prediction models of Visegrad group countries. *Oeconomia Copernicana*, 10(4), 743–772. <https://doi.org/10.24136/oc.2019.034>

Kubenka, M., Capek, J., & Sejkora, F. (2021). A new look at bankruptcy models. *E&M Economics and Management*, 24(3), 167–185. <https://doi.org/10.15240/tul/001/2021-3-010>

Kuster, D. (2022). Financial ratio indicators as early predictors of business failure: Evidence from Serbia. *The Annals of the Faculty of Economics Subotica*, 49, 67–83. <https://doi.org/10.5937/AnEkSub2200005K>

Kuster, D. (2023). Efficiency of working capital & assets management in the function of SMEs bankruptcy prediction. *Management: Journal of Sustainable Business and Management Solutions in Emerging Economies*, 28(3), 57–70. <https://doi.org/10.7595/management.fon.2023.0007>

Laitinen, E. K., & Laitinen, T. (2000). Bankruptcy prediction. *International Review of Financial Analysis*, 9(4), 327–349. [https://doi.org/10.1016/s1057-5219\(00\)00039-9](https://doi.org/10.1016/s1057-5219(00)00039-9)

Law on bankruptcy. Law No. 104/2009, 99/2011 – other Law, 71/2012 – decision of Constitutional Court, 83/2014, 113/2017, 44/2018 and 95/2018. In Official Gazette of the Republic of Serbia. The Republic of Serbia.

Leshno, M., & Spector, Y. (1996). Neural network prediction analysis: The bankruptcy case. *Neurocomputing*, 10(2), 125–147. [https://doi.org/10.1016/0925-2312\(94\)00060-3](https://doi.org/10.1016/0925-2312(94)00060-3)

Luoma, M., & Laitinen, E. K. (1991). Survival analysis as a tool for company failure prediction. *Omega*, 19(6), 673–678. [https://doi.org/10.1016/0305-0483\(91\)90015-1](https://doi.org/10.1016/0305-0483(91)90015-1)

Majstorovic, A., Miloradic, J., Andzic, S., Andzic, R., Divac, G., Anicic, A., & Mirkovic, Z. (2016). Legal aspects of financial analysis in agribusiness companies in Serbia. *Economics of Agriculture*, 63(4), 1379–1393. <https://doi.org/10.5937/ekopolj1604379m>

Maras, I., & Kozar, V. (2021). Protection of rights of secured creditors in the bankruptcy debtor assets sales procedure. *Ekonomika Preduzeca*, 69(6–7), 369–383. <https://doi.org/10.5937/ekopre2106369m>

Menard, S. (2001). *Applied logistic regression analysis* (2nd ed.). SAGE.

Milosevic, N., Dobrota, M., Dmitrovic, V., & Barjaktarovic Rakocevic, S. (2021). Managerial perception of human capital, innovations, and performance: Evidence from banking industry. *Engineering Economics*, 32(5), 446–458. <https://doi.org/10.5755/j01.ee.32.5.26032>

Muminović, S., Pavlović, V., & Cvijanović, J. M. (2011). Predictive ability of various bankruptcy prediction z-score models for Serbian publicly listed companies. *Industrija*, 39(3), 1–12.

Nikolić, N., Jovanović, I., Nikolić, Đ., Mihajlović, I., & Schulte, P. (2018). Investigation of the factors influencing SME failure as a function of its prevention and fast recovery after failure. *Entrepreneurship Research Journal*, 9(3), 20170030. <https://doi.org/10.1515/erj-2017-0030>

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 688–690. <https://doi.org/10.1007/s11135-006-9018-6>

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109–131. <https://doi.org/10.2307/2490395>

Pallant, J. (2009). *SPSS: Guide for survival*. Mikro knjiga.

Papana, A., & Spyridou, A. (2020). Bankruptcy prediction: The case of the Greek market. *Forecasting*, 2(4), 505–525. <https://doi.org/10.3390/forecast2040027>

Papik, M., & Papíková, L. (2024). Automated machine learning in bankruptcy prediction of manufacturing companies. *Procedia Computer Science*, 232, 1428–1436. <https://doi.org/10.1016/j.procs.2024.01.141>

- Park, C.-S., & Han, I. (2002). A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction. *Expert Systems with Applications*, 23(3), 255–264. [https://doi.org/10.1016/S0957-4174\(02\)00045-3](https://doi.org/10.1016/S0957-4174(02)00045-3)
- Ratajczak, P., Szutowski, D., & Szulcowska-Remi, A. (2022). Long-term bankruptcy prediction. Systematic literature review. *Research in Corporate Finance. SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4054665>
- Salmerón, R., García, C. B., & García, J. (2018). Variance inflation factor and condition number in multiple linear regression. *Journal of Statistical Computation and Simulation*, 88(12), 2365–2384. <https://doi.org/10.1080/00949655.2018.1463376>
- Sharma, S., & Mittal, A. K. (2024). Evolution of financial, non-financial, and macroeconomic predictors in corporate bankruptcy: A comprehensive review. *Economic Sciences*, 20(2), 70–83. <https://doi.org/10.69889/0vxzzm07>
- Shi, Y., & Li, X. (2019). An overview of bankruptcy prediction models for corporate firms: A systematic literature review. *Intangible Capital*, 15(2), 114–127. <https://doi.org/10.3926/ic.1354>
- Shie, F. S., Chen, M.-Y., & Liu, Y.-S. (2012). Prediction of corporate financial distress: An application of the America banking industry. *Neural Computing and Applications*, 21(7), 1687–1696. <https://doi.org/10.1007/s00521-011-0765-5>
- Sricharoenchit, N., & Hensawang, S. (2021). Bankruptcy prediction model using logit regression in the automotive sector. In *Proceedings of the 1st International Conference on Sustainable Management and Innovation, ICoSMI 2020* (pp. 250–270). EAI. <https://doi.org/10.4108/eai.14-9-2020.2304465>
- Stanisic, N., Mizdrakovic, V., & Knezevic, G. (2013). Corporate bankruptcy prediction in the Republic of Serbia. *Industrija*, 41(4), 145–159. <https://doi.org/10.5937/industrija41-4024>
- Tian, S., & Yu, Y. (2017). Financial ratios and bankruptcy predictions: An international evidence. *International Review of Economics & Finance*, 51, 510–526. <https://doi.org/10.1016/j.iref.2017.07.025>
- Trading Economics. (2024). *Bankruptcies*. <https://tradingeconomics.com/country-list/bankruptcies>
- Tsai, C., & Hsu, Y. (2011). A meta-learning framework for bankruptcy prediction. *Journal of Forecasting*, 32(2), 167–179. <https://doi.org/10.1002/for.1264>
- Tseng, F.-M., & Hu, Y.-C. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications*, 37(3), 1846–1853. <https://doi.org/10.1016/j.eswa.2009.07.081>
- Tseng, H. P., Chen, P.-C., Huang, W.-H., Lei, M. C., & Tran, Q. H. (2014). Prediction of default probability for construction firms using the logit model. *Journal of Civil Engineering and Management*, 20(2), 247–255. <https://doi.org/10.3846/13923730.2013.801886>
- Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems*, 112, 111–124. <https://doi.org/10.1016/j.dss.2018.06.011>
- Yoon, J. S., & Kwon, Y. S. (2010). A practical approach to bankruptcy prediction for small businesses: Substituting the unavailable financial data for credit card sales information. *Expert Systems with Applications*, 37(5), 3624–3629. <https://doi.org/10.1016/j.eswa.2009.10.029>
- Zenzerovic, R., & Perusko, T. (2009). Evaluating the assumption of going concern in a transitional environment – Empirical findings in Croatia. *Ekonomski Pregled*, 60(7–8), 348–368.