Journal of Business Sectors
Volume 03 Issue 01 June 2025
ISSN 2989-3445 www.jobsjournal.eu

Surviving the Pandemic: Financial Distress Prediction for Slovak SME Manufacturers

Frederik Rech
School of Economics, Beijing Institute of Technology, China
frederikrech@gmail.com
0000-0002-6392-4951

Cyrus Isaboke
School of International Education Hunan University of Finance and Economics, China;
isabokecyrus@gmail.com
0000-0002-9263-5235

Hanying Xu
College of Arts, Humanities and Social Sciences, The University of Edinburgh; United Kingdom
s2302815@ed.ac.uk

ABSTRACT

Research background: In this era of economic uncertainty, small and medium-sized manufacturing enterprises (SMEs) face an increasing risk of financial distress, often with devastating consequences for employment, supply chains, and overall economic stability. The ability to predict corporate financial distress is crucial for financial stability, risk management, and economic resilience, particularly for small and medium-sized enterprises, which are highly vulnerable to financial shocks.

Purpose of the article: The purpose of this study is to develop a financial distress prediction model for Slovak SME manufacturing firms by identifying key financial indicators that distinguish bankrupt from non-bankrupt companies over a three-year period before failure. The study also examines how these indicators evolve as financial distress approaches and evaluates the impact of external shocks, particularly the COVID-19 pandemic, on their predictive power. The goal is to enhance early detection of financial instability, supporting more effective risk management in the manufacturing sector.

Methods: To develop a robust financial distress prediction model, we employ a three-step methodology. First, we address data imbalance using the SMOTE, which generates synthetic samples to balance the minority class and improve model performance. Next, we implement a Genetic Algorithm for variable selection, optimizing the choice of predictors by minimizing the Schwarz Information Criterion. Finally, we use Logistic Regression to model financial distress, ensuring interpretability and statistical rigor.

Findings & Value added: While EBIT/Total Assets, Cash Flow, and Cash Flow/Total Assets consistently reduce financial distress risk, liquidity measures such as Cash and Cash Equivalents/Current Liabilities shift in relevance—acting as protective factors closer to failure but signaling inefficiencies further from financial distress. The model demonstrates strong predictive performance, maintaining high accuracy, recall, and AUC even when tested on real-world unbalanced data, confirming its practical applicability. These findings emphasize the need for adaptive financial distress prediction models that reflect the dynamic nature of financial distress and provide more effective early warning systems in an era of heightened economic uncertainty.

RECEIVED: February 6 ⊙ **ACCEPTED:** March 25 ⊙ **PUBLISHED ONLINE:** June 30

KEYWORDS: financial distress; SMEs; Slovakia

JEL CLASSIFICATION: O14, M21, G33

CITATION: Rech, F., Isaboke, C., & Xu, H. (2025). Surviving the Pandemic: Financial Distress Prediction for Slovak SME Manufacturers. *Journal of Business Sectors,* 3(1), 41–51. https://doi.org/10.62222/SNRN2189

INTRODUCTION

As economic units, digital platforms create the basis for In an era of economic uncertainty, small and medium-sized manufacturing enterprises (SMEs) face an increasing risk of financial distress, often with devastating consequences for employment, supply chains, and overall economic stability. Unlike large corporations, SMEs operate with limited financial buffers, making them particularly susceptible to market shocks, supply chain disruptions, and shifting credit conditions. The COVID-19 pandemic further exposed these vulnerabilities, leading to an urgent need for accurate and adaptable financial distress prediction (FDP) models. However, predicting financial distress remains a complex challenge, as variations in methodologies, feature selection, and economic conditions often raise concerns about model reliability and applicability (Zvarikova et al., 2017). As global financial instability persists, refining FDP approaches tailored to SMEs is crucial for ensuring early intervention, financial stability, and long-term resilience in the manufacturing sector.

The globalization of economies has intensified the frequency and consequences of corporate financial distresses, where financial instability in one nation can quickly escalate into widespread economic disruptions (Klieštik et al., 2018). The COVID-19 pandemic, unlike the gradual unfolding of the 2008 financial crisis, caused an immediate and worldwide economic shock, leading to severe disruptions in business operations and permanent closures for many firms (Musa et al., 2022). Companies faced major operational challenges, including halted production, disrupted supply chains, and financial instability, with varying impacts across industries and regions (Christine et al., 2020). The shift to remote work further underscored economic vulnerabilities, as not all sectors could adapt, resulting in widespread layoffs, particularly among younger workers (Lambovska et al., 2021). Businesses also grappled with capital shortages, rising costs, and restricted credit access, further reducing their ability to generate value (Mitan et al., 2021). Emerging markets bore the brunt of these financial pressures (Christine et al., 2020). However, unlike the 2008 crisis, large-scale fiscal measures and interventions by central banks helped stabilize economies and accelerate recovery from the pandemic's economic consequences (Harjoto & Rossi, 2023).

In Slovakia, the groundwork for financial distress prediction has been established through conventional research methods by scholars like Gulka (2016), Kovacova & Kliestik (2017), Svabova et al. (2018), Boda & Uradnicek (2019), and Gajdosikova & Valaskova (2023). Recent methodological innovations have enhanced the accuracy and relevance of these prediction models. Notable advancements include Mihalovič's (2018) hybrid model combining genetic algorithms with neural networks; the CatBoost algorithm by Papík et al. (2023); neural network ensembles analyzed by Durica et al. (2023); automated machine learning with H2O AutoML; deep learning methods like Gradient-boosting and AdaBoost by Horváthová et al. (2024); and dynamic graph theory models by Horváthová et al. (2023).

Although significant advancements have been made in financial distress prediction modeling, previous studies primarily analyzed Slovak companies either at an aggregate level or broadly across entire sectors. Such approaches overlook the specific vulnerabilities of SMEs within economically critical sectors, such as manufacturing. Furthermore, there is limited research examining how the COVID-19 pandemic has specifically influenced the predictive relevance and dynamics of key financial indicators for SME manufacturers. This study addresses these gaps by narrowing the analytical focus explicitly to Slovak SME manufacturers and investigating how the pandemic altered the stability and predictive power of traditional financial indicators over time. Theoretically, it contributes by enhancing the understanding of how external economic shocks affect the temporal evolution of financial distress predictors, highlighting the changing roles of liquidity, profitability, and leverage measures in unprecedented economic conditions.

This study examines financial indicators that distinguish bankrupt from solvent SMEs in Slovakia's manufacturing sector up to three years before financial distress. As a vital part of the national economy, the sector requires specialized predictive models due to its unique characteristics. The research also investigates whether the COVID-19 crisis introduced new challenges affecting these indicators. The methodology includes SMOTE to address class imbalance, Genetic Algorithm (GA) optimization using Schwarz Information Criterion (SIC) for variable selection, and Logistic Regression (LR) to develop interpretable financial distress prediction models.

The results identified profitability, cash flow, and liquidity as the key predictors of financial distress, with their importance shifting as firms approached financial distress. EBIT/Total Assets, Cash Flow, and Cash Flow/Total Assets consistently showed a negative association with financial distress risk, highlighting the role of profitability and operational efficiency. Liquidity measures, such as Cash and Cash Equivalents/Current Liabilities, acted as protective factors closer to financial distress but became positively associated with risk three years prior, suggesting inefficiencies from excessive liquidity. The model performed well on real-world unbalanced data, demonstrating its robustness and ability to provide reliable financial risk assessments. Despite being trained on a balanced dataset, it achieved strong accuracy, recall, and AUC, confirming its practical applicability in predicting financial distress.

THEORETICAL BACKGROUND

Predicting financial distress in Slovakia has progressed through the development of different analytical approaches. Early research incorporated use of multiple discriminant analysis on agricultural companies, setting the stage for subsequent improvements. Building on this, researchers have increasingly recognized the time-varying nature of FDP models and the need to adapt them to local and sector-specific conditions. Refinements in logit, probit, and discriminant analysis (Gulka, 2016; Mihalovič, 2016; Kovacova & Kliestik, 2017) have improved predictive accuracy, while studies on model selection and adaptability (Boda & Uradnicek, 2019) underscore the importance of industry-specific approaches. Applications in agriculture, small businesses, and the heating sector (Gajdosikova & Valaskova, 2023; Svabova et al., 2018; Stefko et al., 2020) further highlight the necessity of tailored models. Additionally, validation of established techniques, such as Altman's Z-score (Boďa & Úradníček, 2016; Adamko & Svabova, 2016, Vavrek et al., 2021), continues to ensure model relevance in an evolving economic landscape.

Further contributions to refining FDP models under Slovak conditions have come from various studies employing diverse analytical approaches. Kliestik et al. (2020) emphasized the importance of financial ratios in maintaining corporate financial soundness within transitional economies, highlighting liquidity and leverage ratios as key predictors. Their analysis of over 400 models across Eastern Europe revealed variations in preferred ratios influenced by political and economic contexts. Meanwhile, Musa (2019) stressed the interdisciplinary relevance of FDP models, advocating for their use beyond financial management to support broader economic decision-making. Despite these advancements, challenges remain regarding practical adoption, as Lesáková et al. (2020) identified a lack of awareness and utilization of formal prediction methods among Slovak companies, particularly SMEs. Addressing this gap, Sponerová (2021) demonstrated that combining traditional financial indicators with modified variables significantly enhances accuracy for both Czech and Slovak SMEs, while Svabova et al. (2020) developed hybrid models using discriminant analysis and logistic regression, achieving prediction accuracies exceeding 90%, thereby showcasing the potential of tailored approaches for SMEs.

In recent years, traditional statistical methods in FDP have become less popular, as advanced machine learning techniques have demonstrated superior predictive accuracy. Horváthová & Mokrišová (2020) proposed Data Envelopment Analysis as a viable non-parametric alternative to logistic regression, highlighting its adaptability to various financial health assessment scenarios. Musa et al. (2024) compared logistic regression and neural networks, concluding that neural networks offer higher predictive accuracy. Similarly, Horváthová et al. (2024) evaluated gradient boosting models alongside traditional financial indicators, further confirming the advantages of machine learning. Other studies have focused entirely on a machine learning techniques, such as CatBoost and H2O AutoML, both of which achieved high classification performance (Papík et al., 2023; 2024). Additionally, deep learning models like AdaBoost and Gradient Boosting have shown strong results in sector-specific applications (Horváthová et al., 2023). Beyond refining existing models, some research has sought to develop entirely new approaches. One such innovation is dynamic graph theoretical models, which provide a novel framework for financial distress prediction (Horváthová et al., 2023). These advancements highlight the ongoing shift in Slovak FDP research toward more sophisticated, data-driven approaches, ensuring greater adaptability to the complexities of modern financial environments.

Genetic algorithms (GAs) are search techniques modeled after natural selection, first introduced by Holland (1975). They are commonly used in finance and started being applied to bankruptcy prediction in the early 2000s (Atiya, 2001; Shin & Lee, 2002) thereby enhancing precision by optimizing feature selection. What makes GAs particularly unique in financial distress prediction (FDP) is their ability to customize fitness functions based on specific objectives and models. This flexibility allows GAs to prioritize relevant features and fine-tune parameters, making them highly adaptable across different predictive frameworks. Wu et al. (2024) incorporated GAs into an ensemble feature selection method, improving bankruptcy prediction for Chinese A-share companies. Nayak & Rout (2023) demonstrated that Random Forest models perform best when GA is used to refine feature selection. Khedr et al. (2022) highlighted the enhanced accuracy of a hybrid MLP_GA model in the MENA region, while Safi et al. (2022) applied GAs to optimize neural networks, improving performance, particularly in imbalanced datasets. In the conditions of Slovakia, Kanász et al. (2023) utilized GA to optimize a shallow auto encoder ensemble for bankruptcy prediction in SMEs, refining the auto encoder threshold for improved accuracy.

A notable contribution to the field comes from Acosta & Fernandez-Rodriguez (2014) and Acosta et al. (2019), who demonstrated that combining GA with a traditional statistical model not only enhances predictive accuracy but also preserves interpretability, addressing the black-box issue. Their work showed that GA-driven feature selection using SIC (Bayesian Information Criteria) optimizes logistic regression by refining variable selection. This highlights the broader potential of hybrid approaches, where GAs can be leveraged to improve both traditional and machine learning models, ensuring a balance between performance and transparency in financial distress prediction.

Based on the reviewed literature, we propose the following research questions to further explore financial distress prediction for Slovak SMEs in the manufacturing sector:

- **RQ1**: How do key financial predictors identified in previous studies (e.g., liquidity, leverage, profitability) perform in predicting financial distress specifically among Slovak SME manufacturers?
- **RQ2**: Can advanced analytical methods, such as hybrid approaches combining Genetic Algorithms and Logistic Regression, improve predictive accuracy and adaptability over traditional FDP models for SMEs within Slovakia's manufacturing sector?

RESEARCH OBJECTIVE, METHODOLOGY AND DATA

This study seeks to determine the key financial variables that differentiate companies that went bankrupt from those that did not, examining a three-year period before financial distress. It also investigates whether the COVID-19 pandemic introduced significant disruptions to these predictive factors in large manufacturing firms in Slovakia. Analyzing the dataset from 2017 to 2020 is crucial, as it captures financial conditions both before and during the pandemic, allowing us to assess how financial indicators evolved under unprecedented economic stress. By examining this period in depth, we can generate more relevant recommendations for companies to enhance their financial resilience and preparedness for future crises. The methodology incorporates SMOTE to address data imbalance, a Genetic Algorithm with the Schwarz Information Criterion for selecting variables, and Logistic Regression for building predictive models. The dataset includes records from 2,032 Slovak manufacturing firms. The data is sourced from the Orbis database, a global repository of corporate and financial information. To maintain consistency, the study focuses on firms categorized by Orbis as large or very large. Covering the period from 2017 to 2019, the dataset features 43 financial indicators along with a measure of company age. The dependent variable, financial distress status, is classified based on the firms' financial situation as of 2020.

We lacked access to specific data regarding actual financial distress occurrences, as this information was unavailable. To ensure the credibility and robustness of our analysis, we relied on methodologies established by previous studies, including those of Kliestik et al. (2018), Gajdosikova & Valaskova (2023), and Valaskova et al. (2023). According to these authors, a company is classified as insolvent based on three key conditions: (1) its liabilities exceed its assets, resulting in negative equity; (2) it consistently reports negative post-tax profits; and (3) its equity-to-debt ratio falls below 0.08. Under this framework, a company is categorized as bankrupt if it either meets condition (1) alone or satisfies both conditions (2) and (3) together. Using this method, we identified 1,859 companies as healthy (non-bankrupt) and 173 as bankrupt, corresponding to a financial distress rate of 9.31%.

To address the issue of predicting corporate failure, there is no universal agreement on which financial ratios are the most effective, leading to considerable differences in how these ratios are classified and utilized. Generally, previous studies group financial ratios based on shared attributes. Following the methodology of Acosta-González et al. (2019), financial ratios in this study are categorized into seven distinct groups, each reflecting a particular dimension of a company's performance: Profitability, Cash Flow, Liquidity, Leverage, Balance Sheet Structure, Efficiency, and Other Indicators. A detailed list of these ratios, along with other internal variables used in this study, is provided in Table 1.

Table 1: Initial set of variables used

No.	Description	No.	Description
	Profitability		
1	EBIT/total assets	5	Net income/equity
2	EBIT/operating revenue	6	Net income/total assets
3	EBIT/(shareholders' funds + non-current liabilities)	7	Sales/total assets
4	(EBT/total assets)	8	Operating income/ interest paid
9	Operating revenue/(shareholders' funds + non-current liabilities)		
	Cash Flow		
10	Cash flow	13	Cash flow/total liabilities
11	Cash flow/total sales	14	Cash flow/current liabilities
12	Cash flow/total assets	15	Cash flow/shareholders' funds
	Liquidity		
16	Current assets/current liabilities	20	Working capital/total assets
17	Current liabilities/ total assets	21	(Current assets - stocks)/current liabilities
18	Cash and cash equivalents/total liabilities	22	(Shareholders' funds/(shareholders' funds + non-current liabilities))
19	Cash and cash equivalents/current liabilities	23	Shareholders' funds/ non-current liabilities
	Leverage		
24	Total liabilities/shareholders' funds	29	Fixed assets/shareholders' funds
25	Total liabilities/total assets	30	Shareholders' funds/ total liabilities
26	Long term debt/shareholders' funds	31	(Shareholders' funds + long term debt)/ total liabilities
27	Long term debt/current liabilities	32	(Shareholders' funds + long term debt)/ current liabilities
28	Long term debt/total assets	33	(Shareholders' funds + long term debt)/ total assets
	Balance Sheet Structure		
34	Cash and cash equivalents/total assets	37	(Shareholders' funds + capital)/total assets
35	Current assets/total assets	38	Shareholders' funds – capital/total assets
36	Fixed assets/total assets		
	Efficiency		
39	Total liabilities/EBITDA	41	EBITDA/sales
40	Current assets/total sales	42	Stock/sales
	Age and Size		
43	Age	44	Size(LN(total assets))

Source: own processing

To ensure accessibility, we present the methodology in a clear, step-by-step format, enhancing transparency and enabling both academic and non-academic audiences to understand and replicate our study.

Step 1: To handle missing data, we applied the MICE method using the <mice> package (van Buuren & Groothuis-Oudshoorn, 2011). Companies with over 50% missing values were excluded to maintain the reliability of the imputations. After completing the imputation process, we derived all required financial variables for the years 2017 to 2019. Firms were then categorized as either bankrupt or non-bankrupt based on their 2020 status, which served as the reference year for classification. During the analysis, many variables contained zero values, particularly in metrics related to debt. These zeros posed computational issues, such as division by zero, when included in financial ratios. To address this, we replaced zero values with a small positive constant. Furthermore, the dataset included a substantial number of outliers. To manage these effectively, we applied capping to all numeric variables and utilized z-score standardization to achieve consistent scaling across the data.

Step 2: To ensure a robust evaluation, the dataset was first split into training and testing subsets using an 80:20 ratio, with balancing techniques applied exclusively to the training data. Our dataset, consisting of 1,859 non-bankrupt and 173 bankrupt companies, exhibits significant class imbalance, which can bias classification models toward the majority class and reduce accuracy for predicting financial distress (Galar et al., 2012). Challenges include increased error rates for the minority class, difficulty in defining decision boundaries, and the unsuitability of traditional accuracy metrics for imbalanced data (Chawla, 2009). To address this, we applied SMOTE using the <smotefamily> package. SMOTE generates synthetic samples for the minority class by interpolating between existing data points, effectively balancing the dataset and improving model performance.

Step 3: Before selecting features, we first addressed multicollinearity by eliminating highly correlated variables within each year's dataset. This step ensures that redundant or collinear predictors do not distort the model's estimations, leading to more stable and interpretable results. To select features, we implemented a Genetic Algorithm (GA) with a fitness function based on the Schwarz Information Criterion (SIC), as proposed by Acosta-González et al. (2019). This fitness function evaluates subsets of features and prioritizes those that achieve a balance between simplicity and predictive accuracy by minimizing the SIC. GAs, inspired by the principles of natural selection, optimize solutions by iteratively evolving a population of candidates (chromosomes). Each chromosome represents a potential feature subset. The GA process involves the following steps:

1. Initialization: Create an initial population of feature combinations.
2. Selection: Identify the best-performing chromosomes based on the fitness function.
3. Crossover: Combine selected chromosomes to generate new solutions.
4. Mutation: Introduce random alterations to maintain variation in the population.
5. Evaluation: Measure the performance of the updated population.
6. Iteration: Repeat these steps until a predefined stopping condition is met, such as a fixed number of generations.

In this study, the GA used a population size of 1,000 and was run for a maximum of 200 iterations. We implemented the process using the <GA> package in R (Scrucca, 2013). For more details on the methodology, see Acosta-González & Fernández-Rodríguez (2014).

Step 4: After identifying the most relevant variables through GA-SIC optimization, we employed Logistic Regression (LR) to construct a bankruptcy prediction model using the training dataset. LR is a statistical approach designed to estimate the likelihood of a binary outcome-such as whether a company will go bankrupt-based on multiple explanatory variables. The model predicts the probability that an instance belongs to a particular category (e.g., bankrupt vs. non-bankrupt) by applying a logistic function, which constrains output values within the range of 0 to 1.

The logistic function is mathematically defined as:

$$P\left(Y=1 \mid X\right)=\frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_n X_n}}$$

In this equation: P (Y = 1 | X) denotes the probability of financial distress, β0 is the intercept term, β_0, β_1,… β_n represent the coefficients corresponding to the predictor variables X_1, X_2,…X_n , and e is Euler's number (the base of the natural logarithm).

To implement the LR model, we utilized the <glm> function from the <stats> package in R (R Core Team, 2022). This method is widely used due to its simplicity and interpretability, allowing for a clear understanding of how each predictor variable contributes to financial distress probability. Given its effectiveness in binary classification problems, LR remains a valuable tool for financial risk assessment. For further insights into Logistic Regression, refer to Hosmer et al. (2013).

Step 5: To evaluate model performance, we applied standard binary classification metrics, including accuracy, precision, recall, and specificity. The assessment was conducted on the original, unbalanced testing dataset to ensure real-world applicability. After validation, we analyzed the logistic regression coefficients to examine how each predictor influenced bankruptcy risk over the 1–3 years preceding financial distress, identifying trends in variable significance and effect strength over time.

RESULTS AND DISCUSSION

The initial set of 44 financial ratios in table 1 was carefully selected to represent five key categories: profitability, cash flow, liquidity, leverage, and balance sheet structure, along with measures of efficiency, age, and size. To enhance the stability and reliability of the logistic regression models, highly correlated variables were first removed through a pre-processing step. This reduction minimized multicollinearity issues and ensured that the remaining variables provided unique and independent contributions to the models. Following this, a genetic algorithm was applied to identify the most significant predictors of financial distress across three logistic regression models, each representing firms at different distances from financial distress. Specifically, Model 1 corresponds to firms one year prior to financial distress, Model 2 to firms two years prior, and Model 3 to firms three years prior.

Table 2: Logistic Regression Models with Key Predictors Selected via Genetic Algorithm

Selected variables	M1	M2	M3	FRQ
(Intercept)	-18.4768 (0.0000)	-8.8970 (0.0000)	-1.9600 (0.0000)	
EBIT/Total Assets	-1.3318 (0.0000)	-0.8031 (0.0000)	-0.7680 (0.0000)	3
Sales/total assets		-0.2381 (0.0017)	-0.7079 (0.0000)	2
Operating income/interest paid		0.2517 (0.0000)	0.3643 (0.0000)	2
Cash flow	-0.4021 (0.0067)	-0.5191 (0.0000)	-0.3996 (0.0000)	3
Cash flow/Total assets	-0.8390 (0.0011)	0.8002 (0.0000)	0.4797 (0.0000)	3
Cash flow/total liabilities	6.4530 (0.0000)			1
Cash flow/Shareholders' funds	-0.4337 (0.0000)			1
Cash and cash equivalents/Total liabilities	-29.9348 (0.0000)	-19.1027 (0.0000)	1.3335 (0.0000)	3
Cash and cash equivalents/Current liabilities	-14.6442 (0.0000)	-1.7171 (0.0002)	-0.8254 (0.0000)	3
Working capital/total assets			0.3461 (0.0002)	1
(Current assets - stocks)/current liabilities	1.2989 (0.0043)			1
Shareholders' funds/non-current liabilities	-2.4721 (0.0000)			1
Total liabilities/shareholders' funds	-0.2701 (0.0003)			1
Total liabilities/total assets	16.6832 (0.0000)		2.0993 (0.0000)	2
Long term debt/total assets			-0.3247 (0.0000)	1
Cash and cash equivalents/total assets	4.4850 (0.0000)	2.6062 (0.0001)	-0.7696 (0.0000)	3
(Shareholders' funds + capital)/total assets	6.4099 (0.0000)	-1.4042 (0.0000)		2
Shareholders' funds – capital/total assets	5.7108 (0.0000)	-1.9584 (0.0000)	-0.4743 (0.0000)	3
Stock/sales			-0.1778 (0.0013)	1

Note: p-values are presented in parentheses. M1–M3: model 1–3. FRQ – frequency. Source: own calculation

The intercept in each model represents the predicted log-odds of default when all financial ratios are at their average (i.e., z-score = 0). Thus, the highly negative intercept in Model 1 (−18.48) implies an extremely low baseline probability of default for a "typical" firm one year before failure. By contrast, the less negative intercepts in Models 2 (−8.90) and 3 (−1.96) indicate successively higher baseline risks when forecasting two or three years prior to financial distress. This progression underscores that the further out one attempts to predict distress, the greater the model's baseline likelihood of default becomes, even before accounting for firm-specific differences in financial ratios.

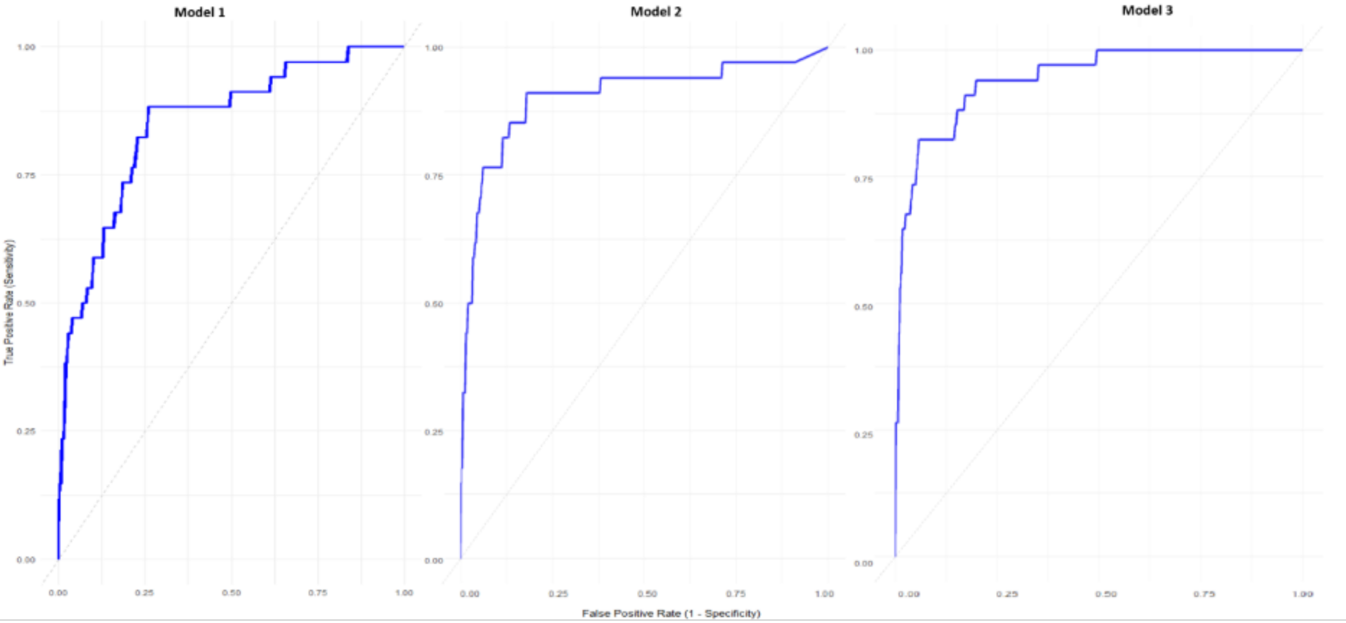
In Model 1, the top variables-Cash and Cash Equivalents/Total Liabilities, Cash and Cash Equivalents/Current Liabilities, Total Liabilities/Total Assets, EBIT/Total Assets, and Cash Flow/Total Assets-demonstrate the overriding importance of liquidity, manageable leverage, and operating profitability. The two cash-based ratios appear with large negative coefficients, implying that more liquid assets relative to liabilities significantly reduce short-term insolvency risks. Meanwhile, the positively signed Total Liabilities/Total Assets reinforces how high leverage becomes particularly dangerous in the final stages before distress, while EBIT/Total Assets and Cash Flow/Total Assets indicate that strong earnings and internal financing capacities are critical safeguards. In Model 2, three of the previous top variables (Cash and Cash Equivalents/Total Liabilities, Total Liabilities/Total Assets, and EBIT/Total Assets) remain highly influential, but two newcomers-Sales/Total Assets (negative) and Cash Flow (negative)-replace the narrowly targeted liquidity ratios from Model 1. This shift suggests that, as the horizon expands, a firm's ability to generate revenues and maintain a healthy overall cash flow gains prominence, even though leverage and liquidity still matter. By Model 3, Total Liabilities/Total Assets and EBIT/Total Assets continue to be powerful predictors, highlighting the persistent effects of leverage and profitability over the long term. However, Cash and Cash Equivalents/Total Liabilities

Table 3: Classification Metrics for the Logistic Regression Models

Model	TP	TN	FP	FN	AUC	Accuracy	Recall	Specificity	F1 Score
Model 1	333	28	6	38	0.9437	0.8914	0.8976	0.8235	0.9381
Model 2	316	29	5	55	0.9038	0.8519	0.8518	0.8529	0.9132
Model 3	305	23	11	66	0.85	0.8099	0.8221	0.6765	0.8875

Source: own research

Figure 1: ROC curves for the three logistic regression models



Source: own research

reenters with a positive coefficient, indicating that simply stockpiling cash too far in advance may not guarantee resilience if accompanied by other structural weaknesses. Completing the top five, Cash and Cash Equivalents/Current Liabilities and Sales/Total Assets both hold negative coefficients, underscoring that immediate liquidity coverage and active revenue generation remain significant protective factors-even in the longer run-against the likelihood of eventual default.

Table 3 presents the results of the models in a real-world setting using the original, unbalanced dataset, illustrating how effectively they distinguish financially distressed firms from healthy ones at different time horizons. Although the training data was balanced with SMOTE, the test set remained true to its real-world proportions to ensure practical relevance. Model 1 achieves the highest AUC (0.9437) and leads in accuracy, recall, and F1 Score, suggesting it excels at recognizing firms on the brink of financial distress one year in advance without excessively misclassifying healthy firms. Model 2 shows a moderate decline in AUC (0.9038) but achieves the highest specificity (0.8529), indicating that it generates fewer false positives when forecasting two years ahead. Meanwhile, Model 3, which attempts a three-year prediction window, experiences a more pronounced drop in AUC (0.8500) and specificity (0.6765), highlighting the growing difficulty of correctly identifying healthy companies at longer horizons. Nevertheless, its recall (0.8221) and F1 Score (0.8875) remain respectable, indicating that it continues to detect a substantial portion of at-risk firms. Taken together, these metrics underscore the trade-off between extending the forecast horizon and maintaining reliable discrimination: further ahead the model attempts to predict financial distress, the more challenging it becomes to distinguish truly healthy companies from those quietly progressing toward default.

Plots in figure 1 depict the ROC curves for the three logistic regression models, with the x-axis representing the False Positive Rate (1 – Specificity) and the y-axis representing the True Positive Rate (Sensitivity). Visually, Model 1 exhibits the highest curve, confirming its superior AUC and thus stronger ability to separate distressed from healthy firms at various probability thresholds. Model 2 still maintains a relatively steep rise in true positive rate at low false positive rates-reinforcing its strong specificity-but it levels out below Model 1's maximum performance. Model 3 maintains an initially steep ascent, indicating that it can achieve a decent sensitivity at relatively low false positives, but the curve flattens sooner than the other two, reflecting its lower overall AUC and diminished specificity. In practical terms, these curves confirm the trade-off between capturing more true positives (i.e., identifying distressed firms) and avoiding false positives (i.e., misclassifying healthy firms), with Model 1 achieving the most favorable balance of the three.

Our findings align in part with those of Gajdosikova & Valaskova (2023), who flagged Current Liabilities/Total Assets as a vital determinant in their one-year model for manufacturing and construction companies during the COVID-19 pandemic-echoing our results. However, our study shows that this ratio remains significant not just one year prior but also three, four, and five years before financial distress, indicating a more prolonged influence on financial health. Conversely, Shareholders' Funds/Total Liabilities and Total Liabilities/Total Assets, which they identified for near-term predictions, become significant in our models primarily at longer horizons. Their work also underscored the role of firm size (using dummy variables) in predicting survival for medium and large companies, whereas our model tested size as a continuous measure (via LN (Total Assets)) and found no significant predictive effect among small and medium-sized (SME) manufacturing firms. These differences suggest that size may carry less weight for financial distress prediction when focusing specifically on SMEs, where financial distress is more likely driven by liquidity constraints, cash flow volatility, and operational inefficiencies rather than absolute firm size.

Our approach-leveraging a Genetic Algorithm for variable selection-provides insights that go beyond prior studies using more constrained methodological frameworks. In line with Valaskova et al. (2018) and Kovacova & Kliestik (2017), we confirm the predictive power of Current Liabilities/Total Assets for the near term (one-year models). Yet, our results demonstrate that for large manufacturers, Total Liabilities/Total Assets appears more relevant when predicting distress over multiple years, diverging from the single-year emphasis in earlier work. Likewise, while Valaskova et al. (2023) identified Total Liabilities/Total Assets and Net Income/Total Assets in one-year predictions post-pandemic, those ratios surfaced in our models only for longer-term forecasts. Such divergences highlight the importance of continuously updating risk models to account for evolving economic conditions, especially amidst seismic events like the COVID-19 crisis.

Our results reinforce these observations by demonstrating that some variables-particularly those related to liquidity and leverage-remain consistently important throughout multiple time horizons, whereas others, such as firm size and certain profitability indicators, exert more nuanced or transient influence. The overarching implication is that models benefiting from flexible, data-driven selection processes are better equipped to detect shifts in a firm's financial vulnerability, particularly in times of prolonged economic turbulence. By refining predictive frameworks to capture these dynamics, stakeholders-ranging from investors to policymakers-can make more informed decisions aimed at sustaining firms through current and future crises.

CONCLUSION

This study aimed to identify key financial variables distinguishing bankrupt firms from stable ones over a three-year period before failure, with a focus on the impact of the COVID-19 pandemic on financial distress predictors in Slovak SMEs. We used SMOTE to address data imbalance, a Genetic Algorithm with the Schwarz Information Criterion for feature selection, and Logistic Regression for predictive modeling. The dataset included 2,032 Slovak SMEs, classified as large or very large, with financial data from 2017 to 2019 and financial distress status determined as of 2020. Firms were categorized based on financial distress criteria, resulting in 173 bankrupt and 1,859 non-bankrupt firms.

The results highlight that profitability, cash flow, and liquidity are the key predictors of financial distress, with their significance changing as firms approach financial distress. EBIT/Total Assets, Cash Flow, and Cash Flow/Total Assets consistently showed a negative association with financial distress risk, emphasizing the importance of operational efficiency and profitability. Liquidity measures, like Cash and Cash Equivalents/Current Liabilities, acted as protective factors closer to financial distress but displayed a positive association with financial distress risk three years prior, suggesting that excessive liquidity in earlier stages may signal inefficiencies. As the prediction horizon extended, Sales/Total Assets and Cash Flow became more important, highlighting the growing role of revenue generation and overall cash flow health. Leverage ratios, such as Total Liabilities/Total Assets, remained significant throughout, reinforcing the importance of financial structure. The model performed well on real-world unbalanced data, despite being trained on a SMOTE-balanced dataset, demonstrating its robustness and practical applicability. The model maintained high accuracy, recall, and AUC, confirming its potential for reliable real-time financial risk assessment. These findings underscore the need for adaptive financial distress prediction models that can evolve with changing financial conditions.

Theoretical contributions of this study include advancing the understanding of how financial distress predictors shift in relevance over different time horizons and how external shocks like the COVID-19 pandemic impact their predictive power. By integrating advanced machine learning techniques such as Genetic Algorithms for feature selection, the study enhances the robustness and accuracy of financial distress prediction models. From a practical perspective, the study provides valuable insights for decision-makers, including investors, financial institutions, and policymakers, by identifying critical financial indicators that can be monitored to detect early signs of financial instability. Furthermore, the results offer actionable recommendations for enhancing risk management practices, particularly in manufacturing SMEs, helping firms build resilience against potential future economic disruptions.

This study has several limitations. First, the dataset was limited to Slovak manufacturing SMEs, which may restrict the generalizability of the results to firms in other industries or regions. Additionally, the lack of direct financial distress records required the use of financial distress proxies, which, while based on established criteria, may introduce classification biases. Furthermore, although logistic regression provides interpretable results, alternative machine learning approaches, such as random forests or neural networks, could be explored to enhance predictive accuracy. Another limitation arises from the independent z-score standardization applied to each model, which prevents direct comparison of variable importance over time. Implementing global standardization techniques could allow for a more precise longitudinal analysis of how financial indicators evolve as financial distress approaches.

Future research could build on these findings by expanding the scope of analysis to different industries, incorporating macroeconomic variables, or applying alternative machine learning techniques to refine predictive accuracy. Additionally, examining the impact of external shocks, such as financial crises or geopolitical disruptions, on financial distress prediction models could provide valuable insights into the adaptability of financial distress indicators. By continuously refining predictive frameworks, future studies can contribute to the development of more effective early warning systems, enabling investors, policymakers, and financial institutions to make more informed decisions in assessing corporate financial health.

REFERENCES

1. Acosta-González, E., & Fernández-Rodríguez, F. (2014). Forecasting Financial Failure of Firms via Genetic Algorithms. *Computational Economics*, 43 (2), 133-157. https://doi.org/10.1007/s10614-013-9392-9
2. Acosta-González, E., Fernández-Rodríguez, F., & Ganga, H. (2019). Predicting Corporate Financial Failure Using Macroeconomic Variables and Accounting Data. *Computational Economics*, 53 (1), 227-257. https://doi.org/10.1007/s10614-017-9737-x
3. Adamko, P., & Svabova, L. (2016). Prediction of the risk of financial distress of Slovak companies. *8th International Scientific Conference Managing and Modelling of Financial Risks*.
4. Atiya, A. F. (2001). Financial distress prediction for credit risk using neural networks: A survey and new results. *IEEE Transactions on Neural Networks*, 12(4). https://doi.org/10.1109/72.935101
5. Boda, M., & Uradnicek, V. (2019). Predicting Financial Distress of Slovak Agricultural Enterprises. *Ekonomický Časopis*, 67 (4), 426-452.
6. Boďa, M., & Úradníček, V. (2016). The portability of altman's Z-score model to predicting corporate financial distress of Slovak companies. *Technological and Economic Development of Economy*, 22 (4), 532-553. https://doi.org/10.3846/20294913.2016.1197165
7. Chawla, N. V. (2009). Data Mining for Imbalanced Datasets: An Overview. In Data Mining and Knowledge Discovery Handbook. https://doi.org/10.1007/978-0-387-09823-4_45
8. Christine, M., Besart, A.-A., Cirera, A. X., Cruz, M., Davies, E., Grover, A., Iacovone, L., Kilinc, U., Medvedev, D., Maduko, F. O., Poupakis, S., Torres, J., & Tran, T. T. (2020). Unmasking the Impact of COVID-19 on Businesses Firm Level Evidence from Across the World. Retrieved: January 9, 2025, from: http://www.worldbank.org/prwp
9. Durica, M., Mazanec, J., & Frnda, J. (2023). Neural network modelling of non-prosperity of Slovak companies. *Engineering Management in Production and Services*, 15 (3). https://doi.org/10.2478/emj-2023-0016
10. Gajdosikova, D., & Valaskova, K. (2023). Financial distress Prediction Model Development and its Implications on Financial Performance in Slovakia. *Economics and Culture*, 20 (1). https://doi.org/10.2478/jec-2023-0003
11. Galar, M., Fernandez, A., Barrenechea, E., Bustince, H., & Herrera, F. (2012). A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches. I*n IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 42 (4), 463-484. https://doi.org/10.1109/TSMCC.2011.2161285
12. Gulka, M. (2016). The prediction model of financial distress of enterprises operating in conditions of SR. *Biatec*, 24 (6), 5-9.
13. Harjoto, M. A., & Rossi, F. (2023), Market reaction to the COVID-19 pandemic: evidence from emerging markets. *International Journal of Emerging Markets*, 18 (1), 173-199. https://doi.org/10.1108/IJOEM-05-2020-0545
14. Holland, J. H. (1975). Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. *University of Michigan Press*.

15. Horváthová, J., Mokrišová, M., & Schneider, A. (2024). The Application of Machine Learning in Diagnosing the Financial Health and Performance of Companies in the Construction Industry. *Information*, 15 (6), 355. https://doi.org/10.3390/info15060355
16. Horváthová, J., Mokrišová, M., & Bača, M. (2023). Financial distress Prediction for Sustainability of Businesses: The Application of Graph Theoretical Modeling. *Mathematics*, 11 (24), 4966. https://doi.org/10.3390/math11244966
17. Horváthová, J., & Mokrišová, M. (2020). Comparison of the results of a data envelopment analysis model and logit model in assessing business financial health. *Information (Switzerland)*, 11 (3), 1-20. https://doi.org/10.3390/info11030160
18. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression: Third Edition. *In Applied Logistic Regression*: Third Edition. https://doi.org/10.1002/9781118548387
19. Kanász, R., Gnip, P., Zoričák, M., & Drotár, P. (2023). Financial distress prediction using ensemble of autoencoders optimized by genetic algorithm. *PeerJ Computer Science*, 9. https://doi.org/10.7717/peerj-cs.1257
20. Khedr, A. M., El Bannany, M., Kanakkayil, S., & Yuldashev, M. (2022). Application of boruta feature selection in enhancing financial distress prediction performance of hybrid MLP-GA. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3584202.3584298
21. Kliestik, T., Valaskova, K., Lazaroiu, G., Kovacova, M., & Vrbka, J. (2020). Remaining financially healthy and competitive: The role of financial predictors. *Journal of Competitiveness*, 12 (1), 74–92. https://doi.org/10.7441/joc.2020.01.05
22. Kliestik, T., Vrbka, J., & Rowland, Z. (2018). Financial distress prediction in visegrad group countries using multiple discriminant analysis. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 13 (3). https://doi.org/10.24136/eq.2018.028
23. Kovacova, M., & Kliestik, T. (2017). Logit and probit application for the prediction of financial distress in slovak companies. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12 (4), 775-791. https://doi.org/10.24136/eq.v12i4.40
24. Lambovska, M., Sardinha, B., & Belas, J. (2021). Impact of the Covid-19 pandemic on youth unemployment in the European Union. *Ekonomicko-manazerske spektrum*, 15 (1), 55-63. https://doi.org/10.26552/ems.2021.1.55-63
25. Lesáková, Ľ., Gundová, P., & Vinczeová, I. (2020). The practice of use of models predicting financial distress in Slovak companies. *Journal of Eastern European and Central Asian Research*, 7 (1), 123-136. https://doi.org/10.15549/jeecar.v7i1.369
26. Mihalovič, M. (2016). Performance comparison of multiple discriminant analysis and logit models in financial distress prediction. *Economics and Sociology*, 9 (4). 101-118. https://doi.org/10.14254/2071-789X.2016/9-4/6
27. Mihalovič, M. (2018). The use of scoring models in predicting the default of economic entities in the Slovak Republic. (Využitie skóringových modelov pri predikcii defaultu ekonomických subjektov v Slovenskej republike). *Politicka Ekonomie*, 66 (6), 689-708. https://doi.org/10.18267/J.POLEK.1226
28. Mitan, A., Siekelova, A., Rusu, M., & Rovnak, M. (2021). Value-based management: A case study of Visegrad Four countries. *Ekonomicko-manazerske spektrum*. 15 (2), 87-98. https://doi.org/10.26552/ems.2021.2.87-98
29. Musa, H. (2019). Default Prediction Modelling for Small Enterprises: Case Slovakia. *34th IBIMA Conference*. Retrieved: January 13, 2025, from: https://ibima.org/accepted-paper/default-prediction-modelling-for-small-enterprises-case-slovakia/
30. Musa, H., Rech, F., Musova, Z., Yan, C., & Pinter, L. (2024). Financial distress prediction using machine learning: The case of Slovakia. *In A. Tsounis, N., Vlachvei (Ed.), Applied Economic Research and Trends*. ICOAE 2023. Springer Proceedings in Business and Economics. Springer. Cham. https://doi.org/10.1007/978-3-031-49105-4_34
31. Musa, H., Rech, F., Yan, C., & Musova, Z. (2022). The Deterioration of Financial Ratios During the Covid-19 Pandemic: Does Corporate Governance Matter? *Folia Oeconomica Stetinensia*, 22 (1), 219-242. https://doi.org/10.2478/FOLI-2022-0011
32. Nayak, S. M., & Rout, M. (2023). Genetic Algorithm and GAN based Hybrid Model for Financial distress Prediction. *OCIT 2023 - 21st International Conference on Information Technology, Proceedings*. https://doi.org/10.1109/OCIT59427.2023.10430554
33. Papík, M., Papíková, L., Kajanová, J., & Bečka, M. (2023). CatBoost: The Case of Financial distress Prediction. *Lecture Notes in Networks and Systems*, 487. https://doi.org/10.1007/978-3-031-08084-5_3
34. R Core Team. (2022). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. https://www.r-project.org/
35. Safi, S. A. D., Castillo, P. A., & Faris, H. (2022). Cost-Sensitive Metaheuristic Optimization-Based Neural Network with Ensemble Learning for Financial Distress Prediction. *Applied Sciences (Switzerland)*, 12 (14). https://doi.org/10.3390/app12146918
36. Scrucca, L. (2013). GA: A package for genetic algorithms in R. *Journal of Statistical Software*, 53 (4). https://doi.org/10.18637/jss.v053.i04
37. Shin, K. S., & Lee, Y. J. (2002). A genetic algorithm application in financial distress prediction modeling. *Expert Systems with Applications*, 23 (3). https://doi.org/10.1016/S0957-4174(02)00051-9

38. Sponerová, M. (2021). Financial distress Prediction: The Case of the Czech Republic and Slovakia. *Selected Papers (Part of ITEMA Conference Collection)*, 65-71. https://doi.org/10.31410/itema.s.p.2021.65
39. Stefko, R., Jencova, S., & Vasanicova, P. (2020). The Slovak Spa Industry and Spa Companies: Financial and Economic Situation. *Journal of Tourism and Services*, 11 (20), 28-43. https://doi.org/10.29036/JOTS.V11I20.137
40. Svabova, L., Durica, M., & Podhorska, I. (2018). Prediction of Default of Small Companies in the Slovak Republic. *Economics and Culture*, 15 (1), 88-95. https://doi.org/10.2478/JEC-2018-0010
41. Svabova, L., Michalkova, L., Durica, M., & Nica, E. (2020). Business failure prediction for Slovak small and medium-sized companies. *Sustainability (Switzerland)*, 12 (11). https://doi.org/10.3390/su12114572
42. Valaskova, K., Gajdosikova, D., & Belas, J. (2023). Financial distress prediction in the postpandemic period: A case study of Visegrad Group countries. *Oeconomia Copernicana*, 14 (1), 253-293. https://doi.org/doi:10.24136/oc.2023.007
43. Valaskova, K., Kliestik, T., Kovacova, M., Radisic, M., & Mirica, C. O. (2018). Financial distress prediction in specific economic conditions of Slovakia: Multiple discriminant analysis. *Proceedings of the 32nd International Business Information Management Association Conference,* IBIMA 2018 - Vision 2020: Sustainable Economic Development and Application of Innovation Management from Regional Expansion to Global Growth.
44. van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45 (3). https://doi.org/10.18637/jss.v045.i03
45. Vavrek, R., Gundová, P., Vozárová, I. K., & Kotulič, R. (2021). Altman model verification using a multi-criteria approach for slovakian agricultural enterprises. *E + M: Ekonomie a Management*, 24 (1), 146-164. https://doi.org/10.15240/TUL/001/2021-1-010
46. Wu, C., Chen, X., & Jiang, Y. (2024). Financial distress prediction based on ensemble feature selection and improved stacking algorithm. *Kybernetes*. https://doi.org/10.1108/K-08-2023-1428
47. Zvarikova, K., Spuchlakova, E., & Sopkova, G. (2017). International comparison of the relevant variables in the chosen financial distress models used in the risk management. *Oeconomia Copernicana*, 8 (1), 145-157.