



Article

Is Artificial Intelligence Really More Accurate in Predicting Bankruptcy?

Stanislav Letkovský , Sylvia Jenčová and Petra Vašaničová *

Faculty of Management and Business, University of Presov, 080 01 Presov, Slovakia;
stanislav.letkovsky@smail.unipo.sk (S.L.); sylvia.jencova@unipo.sk (S.J.)

* Correspondence: petra.vasanicova@unipo.sk

Abstract: Predicting bankruptcy within selected industries is crucial because of the potential ripple effects and unique characteristics of those industries. It serves as a risk management tool, guiding various stakeholders in making decisions. While artificial intelligence (AI) has shown high success rates in classification tasks, it remains uncertain whether its use significantly enhances the potential for early warning of impending problems. The following question arises: will classical methods eventually replace the effectiveness of these advanced techniques? This paper sheds light on the fact that even classical methods continue to achieve results that are not far behind, highlighting their enduring importance in financial analysis. This paper aims to develop bankruptcy prediction models for the chemical industry in Slovakia and to compare their effectiveness. Predictions are generated using the classical logistic regression (LR) method as well as AI techniques, artificial neural networks (ANNs), support vector machines (SVMs), and decision trees (DTs). The analysis aims to determine which of the employed methods is the most efficient. The research sample consists of circa 600 enterprises operating in the Slovak chemical industry. The selection of eleven financial indicators used for bankruptcy prediction was grounded in prior research and existing literature. The results show that all of the explored methods yielded highly similar outcomes. Therefore, determining the clear superiority of any single method is a difficult task. This might be partially due to the potentially reduced quality of the input data. In addition to classical statistical methods employed in econometrics, there is an ongoing development of AI-based models and their hybrid forms. The following question arises: to what extent can these newer approaches enhance accuracy and effectiveness?

Keywords: bankruptcy; prediction; logistic regression; artificial intelligence; artificial neural networks; support vector machine; decision tree



Citation: Letkovský, Stanislav, Sylvia Jenčová, and Petra Vašaničová. 2024. Is Artificial Intelligence Really More Accurate in Predicting Bankruptcy? *International Journal of Financial Studies* 12: 8. <https://doi.org/10.3390/ijfs12010008>

Academic Editor: Vassilios Papavassiliou

Received: 16 October 2023
Revised: 11 January 2024
Accepted: 15 January 2024
Published: 18 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In financial decision making, the ability to predict bankruptcy or anticipate challenges in meeting financial obligations holds paramount importance (Brygala 2022). The consequences of financial failure have a significant impact on creditors. Therefore, the importance of prediction has enormously increased in recent decades. Accurate prediction offers various advantages, such as an increased debt collection rate and reduced costs in credit analysis, among others (Korol 2019). Bankruptcy often manifests as a longer term outcome with unclear indications. Artificial intelligence (AI) has the capability to detect hidden patterns signaling this condition, necessitating the analysis of a larger volume of data and studying their behavior under varied conditions. This study aims to compare the results of the most widely used AI methods and logistic regression in the specific context of the chemical industry in the Slovak Republic, a crucial component of the country's economy. Using identical input data, the objective is to provide an objective comparison of the outcomes generated by each method. The selection of predictors undergoes rigorous analysis, considering a wide array of frequently used indicators. This paper presents an original

solution tailored to a narrowly specialized environment, with potential applicability in other countries and time periods, encouraging broader comparisons and deeper investigations into the problem. As noted by [Brygala \(2022\)](#), the efficiency of logistic regression with unbalanced data is lower than with balanced data. Given the low proportion of bankrupt samples, developing accurate prediction models remains a challenge ([Garcia 2022](#)), highlighting the difficulty of obtaining high quality and sufficient data. Research on bankruptcy prediction models is undoubtedly crucial. A high number of failures could be devastating for the business sector. The model's performance strongly depends on the used tools. Most studies choose a model based on its popularity or professional background. Only a few works (e.g., [Altman 1968](#); [Ohlson 1980](#)) focus on expert analysis with the creation of their own model. The reason may be the lack of a comparison of the relative performance of the tools with respect to the required prediction criteria ([Alaka et al. 2018](#)). There are numerous statistical methods available for detecting the potential risk of bankruptcy, with logistic regression being the most widely used and yielding good results. In an effort to enhance accuracy, AI-based techniques are gaining prominence. While these methods hold the promise of improvement, some studies indicate minimal or no increase in accuracy. According to [du Jardin \(2018\)](#), the classical model reflects a rather elementary view of bankruptcy, treating it as the outcome of a historical process independent of time, reducible to a specific set of measures. However, in reality, businesses with similar financial profiles exhibit different failure rates. Some of them demonstrate greater adaptability and resilience to failure, often developing this capability at the onset of potential failure. Factors that can only be analyzed over time elude the grasp of traditional models.

Standard and previously proposed models are often unsuitable as they do not account for the specific nuances of a particular environment. A notable contribution of this study is the proposal of specific models tailored to the chemical industry of the Slovak Republic, enhancing the potential of financial management within this sector. These models can function as an early warning system for potential bankruptcy, benefiting both creditors and the company's management. Going beyond the standard prediction of one year before bankruptcy, this study also provides predictions two years in advance, allowing for early problem identification. Furthermore, this extended prediction horizon serves as a benchmark for comparing the effectiveness of various prediction methods. A similar comparative approach can be found in the work of [Aker and Karavardar \(2023\)](#), where they attempt to predict bankruptcy even three years in advance. Additionally, [Gavurova et al. \(2022\)](#) propose models for similar conditions, specifically within the engineering and automotive industry in Slovakia.

The entire principle of prediction models is based on finding a function that separates bankrupt from non-bankrupt samples with the highest possible reliability. Basic tasks in the field of bankruptcy prediction may include the following: (i) defining criteria for bankruptcy; (ii) selecting (searching for) predictive indicators; (iii) selecting (searching for) a method (model) capable of distinguishing a bankrupt company (prediction); (iv) evaluating (comparing) the success of models and the cost of misclassification.

The field of bankruptcy prediction remains relatively unexplored due to the absence of the exact application procedure for specific conditions. Numerous models have been developed for particular conditions, industries, countries, and diverse businesses ([Kliestik et al. 2023](#); [Nagy et al. 2023](#)). When applied, it is unclear which one is the most appropriate or the most effective to use for the specific data types, industries, or conditions. According to [Alaka et al. \(2018\)](#), the number of applications is inappropriate, highlighting the need for a systematic comparison of models. Effective models are lacking and create a research gap for specific industries ([Chen et al. 2021](#)).

The chemical industry is an important component of the Slovak economy. A bankruptcy model for the current conditions of this specific industry and country is missing. To fill this research gap is one of the motivations of this paper.

This paper aims to develop bankruptcy prediction models for the chemical industry in Slovakia and to compare their effectiveness. Predictions are generated using the classical

logistic regression (LR) method as well as AI techniques (artificial neural networks (ANNs), support vector machines (SVMs), and decision trees (DTs)). The analysis aims to determine which of the employed methods is the most efficient. A range of frequently used financial and economic indicators were analyzed to formulate the models. Out of these, 11 predictors were identified by the authors as pivotal for prediction. Notably, AV/S, E/TA, ROA, and TD/E exhibited the highest informative value in explaining the dependent variable. The dataset initially comprised 1221 companies in 2020 and 1206 in 2019, all operating in the chemical industry. Following meticulous preprocessing, the data were refined to 608 samples for 2020 and 605 for 2019. All models demonstrated an overall prediction accuracy exceeding 95%, positioning them as effective and relatively reliable. Minimal differences among the methods prevented a clear assessment of superiority for any specific model.

The remainder of this paper is organized as follows. In the next section, we provide a literature review. Next, we present the data and used methods. The presentation of the results and a discussion follows. The final section summarizes our findings, the limitations of our research, and possibilities for future research.

2. Literature Review

Prediction begins by comparing indicators of healthy and bankrupt enterprises (Fitzpatrick 1932), followed by discriminant analysis (Fisher 1936) and fuzzy set techniques (Zadeh 1965). One of the most well-known and frequently used models to date is the univariate analysis by Beaver (1966) and the z-score by Altman (1968). Furthermore, the popularity of discriminant analysis has increased due to work in the finance field by Taffler (1982). However, these conventional methods have limitations related to linearity, normality, and multicollinearity. The next stage of development involves the application of statistical methods such as logit (Ohlson 1980) and probit (Zmijewski 1984). With the advancement in technology, AI-based methods have emerged, and the work of Odom and Sharda (1990) is considered a pioneer in the prediction field using ANNs.

While bankruptcy can occur suddenly due to unexpected events, it is often possible to predict it by using the appropriate methods. Estimation errors can be caused by unreliable accounting statements, where data might be intentionally or unintentionally distorted (Mućko and Adamczyk 2023). Despite the extensive research, determining the superiority of any method remains unclear (Shin et al. 2005). Most models achieve high accuracy in the short term but experience significant declines over time (Korol 2019).

A common problem is sample imbalance which results in inaccurate predictions. Prediction errors have a negative impact on the company's financial health. Addressing sample imbalance in classification tasks can be approached through data-level techniques, algorithm-level adjustments, or hybrid methods. Preprocessing, which involves changing (reducing or increasing) the size of sets and equalizing their distribution, is also a simple and effective method. Oversampling is more commonly used (e.g., Chawla et al. 2002; Garcia 2022), while undersampling techniques receive less attention (Wang and Liu 2021; Brygała 2022). Zoričák et al. (2020) investigated sample imbalance in small and medium enterprises.

In their systematic study, Alaka et al. (2018) categorized the criteria for bankruptcy models into three fundamental categories:

1. Result criterion (model accuracy and interpretation of results);
2. Data criterion (sample size, dispersion, and variable selection);
3. Model properties criterion (design time, assumptions, variable relationship, etc.).

Finally, they compared the frequency of the usage of the individual methods. The highest frequency is achieved by ANNs (25%), followed by LR (20%) and SVMs (16%). Each of these methods has its own strengths and weaknesses. It can be concluded that no single model stands out as clearly superior when considering all of the identified bankruptcy criteria.

Shin et al. (2005) compared the results of SVMs and ANN-B, highlighting the higher accuracy of SVMs, especially when dealing with a small number of samples. When a large number of training sets are available, the results become comparable. Iturriaga and

[Sanz \(2015\)](#) proposed a hybrid model combining ANNs and a self-organizing map (SOM). This hybrid model predicts bankruptcy using ANNs one year prior to the event and applies the model to data from 2 and 3 years before the bankruptcy. They then created a SOM by combining the results, which provides a visual representation of the various risk profiles. They compared their results with discriminant analysis, LR and SVMs, showing the predominance of ANN accuracy.

[Korol \(2019\)](#) introduced a model for EU companies comparing the fuzzy sets, ANN, and DT methods. The evaluation included assessing efficiency drop up to 10 years before bankruptcy. [Ptak-Chmielewska \(2019\)](#) compared LR, SVMs, Boosting, ANNs, and DTs, finding that LR's performance matches that of the other methods like SVMs and DTs. [Wang and Liu \(2021\)](#) investigated the impact of the sample imbalance on the accuracy and proposed an undersampling method using SVMs, LR, neural networks, linear discriminant analysis, and random forest, among others. [Brygała \(2022\)](#) addressed sample imbalance and tested LR. [Chen et al. \(2021\)](#) developed a hybrid model that selects the most suitable prediction type (Naive Bayes, K-nearest neighbor, DTs, bagging, or LR) based on the data. They achieved the best results with DTs and LR, while bagging performed the worst. [Korol \(2021\)](#) and [Korol and Fotiadis \(2022\)](#) compared classical methods of multivariate discriminant analysis, LR and DTs, with LR emerging as the dominant performer.

3. Materials and Methods

3.1. Data

The dataset comprised enterprises from the Slovak chemical industry, sourced from the Register of Financial Statements of the Slovak Republic for the years 2019–2021. Slovakia belongs to the industrialized countries. Slovak industry employs more than 23% of the population, and its share in GDP has averaged 23% (21.4–24.3%) over the last ten years. The chemical industry in Slovakia, even if it is not the largest industry, forms a significant part of it because it is among the most profitable (in 2022, it was the most profitable; in 2021, it was the second).

Bankruptcy is the result of a longer term activity. As mentioned above, we used data for 2020 and 2021, which were influenced by the COVID-19 pandemic. We used data for 2019 because the chemical industry was partly affected by the pandemic, in exactly the opposite way, as this period was a period of prosperity for some of its business entities (e.g., pharmaceutical companies). It is also interesting to examine this period from the point of view of bankruptcy prediction. The bankruptcy samples encompassed companies with an equity-to-total-indebtedness ratio below 0.08 in 2021. The examined data were divided into two sets:

1. Set 1 for the year 2020 (one year before the bankruptcy assessment);
2. Set 2 for the year 2019 (two years before bankruptcy).

In preparation for the machine learning algorithms, the data were scaled through standardization. This process eradicates dependence on units and positional parameters, ensuring a mean of 0 and a standard deviation of 1 for the samples. To eliminate outliers, the interquartile range method was employed, with the outlier limit set at three times the interquartile range.

Each of the two sets was partitioned into three segments using a random number generator, roughly maintaining a distribution ratio of 50:20:30 for training, holdout, and testing. The combination of training and holdout segments was utilized for model creation and is referred to as the “training only” set. This larger portion comprised samples for model creation (70% of the dataset), while the smaller portion served as the test set (samples excluded from modeling). The literature does not specify the procedure or the exact conditions for dividing the samples. Also, the work of [Gavurova et al. \(2022\)](#) did not show a statistically significant difference between the ratios 60:40, 70:30, and 80:20. Based on this, and also through empirical testing, the determined ratio was the optimal choice (a sufficient number of samples for validation and a sufficient number of samples for modeling as well). The initial sample count totaled 1221 for the year 2020 and 1206 for the

year 2019. After the removal of incomplete data and outlier samples, the sets were reduced to 608 samples for 2020 and 605 samples for 2019. The training set consisted of 428 samples for 2020 and 425 samples for 2019, while the test set contained 180 samples for both years. Among these, the number of bankrupt samples were 45 enterprises, representing less than 8% of the dataset.

3.2. Methods

In our analysis, we have established that a company is classified as bankrupt if its equity-to-total-debt ratio exceeds 8%. ([Act No. 7/2005 Coll. on Bankruptcy and Restructuring \(2013\)](#) defines bankruptcy using this indicator.) Moreover, [Act No. 513/1991 Coll. Commercial Code \(2013\)](#) sets the minimum limit for the analyzed period at 0.08. Predictive models were constructed using classical LR as well as several AI methods (SVMs, DTs, and ANNs).

The objective of the modeling process is to enhance the accuracy and the efficiency of the prediction, making the selection of indicators a pivotal factor ([Abraham et al. 2022](#)). [du Jardin \(2010\)](#) illustrates that employing bulk selection with ANNs outperforms selection based on the existing literature. While the primary aim is not to identify the best indicators, this work's indicator selection is grounded in prior research and the existing literature. Specifically, we used the following indicators:

- Added-value-to-sales ratio (gross margin) = AV/S ;
- Return on sales (earnings before interests and taxes to sales) = ROS;
- Equity-to-total-assets ratio = E/TA ;
- Current-ratio-to-total-assets ratio = CuR/TA ;
- Current ratio (current-assets-to-current-liabilities ratio) = CuR ;
- Total assets = TA ;
- Return on equity (earnings after taxes to equity) = ROE;
- Return on assets (earnings before interests and taxes to total assets) = ROA;
- Assets turnover = AT ;
- Total-debt (liabilities)-to-equity ratio = TD/E ;
- Net-working-capital-to-total-assets ratio = NWC/TA .

Among the chosen indicators, the gross margin (AV/S) holds significant prominence as evidenced by [Ptak-Chmielewska \(2019\)](#), [Ben Jabeur \(2017\)](#), and [Hsieh et al. \(2006\)](#). In the profitability domain, the return on sales (ROS) indicator is widely acknowledged, according to [Zmeškal et al. \(2023\)](#), [Thanh-Long et al. \(2022\)](#), [Štefko et al. \(2021\)](#), [Ptak-Chmielewska \(2019\)](#), and [Ben Jabeur \(2017\)](#). The return on equity (ROE), defined by [Fitzpatrick \(1932\)](#) and emphasized by [Thanh-Long et al. \(2022\)](#) and [Yousaf and Bris \(2021\)](#), is deemed one of the most crucial predictors of bankruptcy. Similarly, return on assets (ROA), an indicator of profitability, has been employed in prediction models since [Altman \(1968\)](#), and subsequently by [Yousaf and Bris \(2021\)](#), [Ptak-Chmielewska \(2019\)](#), and [Zhang et al. \(1999\)](#).

The E/TA and CuR/TA indicators are considered significant according to [Ptak-Chmielewska \(2019\)](#), [Ben Jabeur \(2017\)](#), and [Shin et al. \(2005\)](#), within the context of liquidity. The current ratio (CuR) is widely used in liquidity assessment, as observed in [Thanh-Long et al. \(2022\)](#), [Yousaf and Bris \(2021\)](#), and [Zhang et al. \(1999\)](#), along with the total assets (TA) indicator. The assets turnover (AT) indicator, reflecting asset utilization efficiency, was identified as a pivotal predictor of bankruptcy by [Altman \(1968\)](#), [Zhang et al. \(1999\)](#), [Štefko et al. \(2021\)](#), [Ptak-Chmielewska \(2019\)](#), and [Zhang et al. \(1999\)](#). Other selected indicators, including TD/E and NWC/TA , were chosen in line with varying author recommendations and according to [Jenčová et al. \(2020\)](#) and [Gajdosikova et al. \(2023a, 2023b\)](#).

The indicators' selection was conditioned by the literature analysis, which informs us about the most important and frequently used bankruptcy predictors. Another factor is the quality of the input data, which partially limits their selection. Through empirical testing and analysis, we have selected eleven indicators that we consider the most important for the bankruptcy assessment.

The modeling was conducted using the JASP program, which permits the utilization of all analyzed methods. The configuration for each method was determined empirically to achieve optimal performance.

3.2.1. Logistic Regression (LR)

LR stands as a classic method and continues to be frequently employed in bankruptcy prediction. This method operates under the assumption of a logistic probability distribution. Crucial criteria for this method include the non-collinearity of independent variables and an adequate sample size (Ptak-Chmielewska 2019). The formula can be expressed in terms of an odds ratio:

$$\frac{\pi}{1 - \pi} = e^{(\alpha + b_1x_1 + \dots + b_nx_n)} \quad (1)$$

and then to logit:

$$\text{logit}(y) = \ln\left(\frac{\pi}{1 - \pi}\right) = \alpha + b_1x_1 + \dots + b_nx_n \quad (2)$$

where π is the probability of the event, α is the intercept, b is the regression coefficients, and x is the predictors. Coefficients are estimated using maximum likelihood.

Determining the optimal cut-off point ensures the efficacy of the model. While a commonly used value is 0.5, this may not guarantee the most optimal distribution (Brygała 2022).

3.2.2. Artificial Neural Network (ANN)

Models based on ANNs are often referred to as black boxes due to their inability to explain the rationale behind classifying a given sample as bankrupt or non-bankrupt, a critical aspect of such classification. Nonetheless, these models achieve high accuracy and they have gained increased usage with technological advancements. At the core of these models is the neuron which aggregates and transforms inputs into outputs through activation functions. When neurons of the same type are interconnected, they form layers. Neurons (layers) that accept input to the network constitute the input layer, while those providing network output make up the output layer. The remaining neurons form hidden nodes (hidden layers). A network comprises an input and an output layer along with no or several hidden layers. When information moves only from input to output without feedback, it is termed "Feed Forward" (FF), while networks with feedback constitute recurrent networks. In prediction, the backpropagation of error ("Backpropagation"-BP) learning method is most commonly used.

An advantage of ANNs is their ability to bypass assumptions about data distribution, enabling the representation of non-linear relationships between dependent and independent variables (Iturriaga and Sanz 2015). However, as per Shin et al. (2005), ANN-based models are limited by their network settings. Configuring them can prove challenging due to numerous control parameters (e.g., layer count, node count, activation function, learning method, and stopping criteria).

3.2.3. Support Vector Machine (SVM)

SVM is analogous to the quadratic optimization and involves searching for hyperplanes in space to maximize the distance from data points. It can deduce the optimal solution based on a limited amount of training sample data. The structure employs risk minimization, enabling it to achieve high generalization and excellent pattern recognition capabilities (Shin et al. 2005). It employs a linear model to establish non-linear class boundaries by mapping the inputs to a multidimensional space (Iturriaga and Sanz 2015). An SVM supports both classification and regression tasks, accommodating continuous as well as categorical variables (Ptak-Chmielewska 2019). The comprehensive elucidation of SVMs can be found in the book by Cortes and Vapnik (1995).

Based on support vectors, an SVM constructs an estimation function for non-linear class boundaries. Support vectors represent training points closest to an optimal separation. An SVM determines a hyperplane to minimize errors. Notably, an SVM offers advantages

compared to ANNs, including simpler parameter configuration (only 2 free parameters—upper bound and kernel function), superior potential for finding the optimal solution (ANNs can become trapped in local minima), and applicability even with a small sample size.

3.2.4. Decision Tree (DT)

The DT method employs entropy to quantify the discriminatory power of samples (Quinlan 1986). The underlying principle relies on decision-making rules, with decisions structured hierarchically based on their importance (if the indebtedness factor is deemed more crucial than liquidity, it is evaluated earlier, and therefore occupies a higher position within the decision-making tree). The process involves stepwise hierarchical segmentation of the samples. This starts with the entire dataset (root) and is guided by rules which divide the samples into smaller groups (nodes). The terminal group is termed a leaf and is not subjected to further division. Ultimately, each sample is assigned to one of these leaves.

This approach lacks coefficients or calculations; it relies solely on division rules. The primary challenge is the risk of overfitting. Notably, its strengths encompass results that are easily interpretable, model flexibility, insensitivity to missing data, and no requirement for a normal distribution (Ptak-Chmielewska 2019).

4. Results and Discussion

All models exhibited the capability to identify 50% of bankrupt samples in the 2020 test set (one year before bankruptcy), while achieving a perfect accuracy of 100% for non-bankrupt samples, leading to an overall accuracy exceeding 96%. The training set’s accuracy for this year was slightly lower. For the scenario two years before bankruptcy, the models displayed the ability to accurately identify 29–36% of bankrupt samples in the test set, resulting in an overall accuracy ranging from 94% to 97%. Remarkably, all models demonstrated almost identical outcomes for this scenario (see Table 1).

Table 1. Prediction results for all models.

Sample	Obs.	Predicted												
		LR			SVM			DT			ANN			
		0	1	Correct	0	1	Correct	0	1	Correct	0	1	Correct	
2 years	Training	0	393	1	99.75%	393	1	99.75%	393	1	99.75%	393	1	99.75%
		1	17	14	45.16%	16	15	48.39%	15	16	51.61%	17	14	45.16%
		Overall	96.47%	3.53%	95.76%	96.24%	3.76%	96.00%	96.00%	4.00%	96.24%	96.47%	3.53%	95.76%
	Testing	0	166	0	100.00%	166	0	100.00%	166	0	100.00%	166	0	100.00%
		1	10	4	28.57%	9	5	35.71%	9	5	35.71%	9	5	35.71%
		Overall	97.78%	2.22%	94.44%	97.22%	2.78%	95.00%	97.22%	2.78%	95.00%	97.22%	2.78%	95.00%
1 year	Training	0	396	1	99.75%	397	1	99.75%	397	1	99.75%	396	1	99.75%
		1	18	13	41.94%	17	13	43.33%	16	14	46.67%	19	12	38.71%
		Overall	96.73%	3.27%	95.56%	96.73%	3.27%	95.79%	96.50%	3.50%	96.03%	96.96%	3.04%	95.33%
	Testing	0	166	0	100.00%	166	0	100.00%	166	0	100.00%	166	0	100.00%
		1	7	7	50.00%	7	7	50.00%	7	7	50.00%	7	7	50.00%
		Overall	96.11%	3.89%	96.11%	96.11%	3.89%	96.11%	96.11%	3.89%	96.11%	96.11%	3.89%	96.11%

Note: Obs. denotes observation. LR is logistic regression, SVM is support vector machine, DT denotes decision tree, and ANN is artificial neural network.

The results show that all investigated methods achieved very similar results. This conclusion corresponds to that of Shin et al. (2005). It is not possible to decide on the superiority of any method. Many authors achieve higher accuracy when using AI rather than classical models, or their results are comparable. However, classical statistical models are not more accurate than AI. The AI-based models in this work similarly did not show statistically

significantly higher accuracy than LR. On the contrary, they all had a better ability to identify bankrupt samples two years before bankruptcy. The results of [Iturriaga and Sanz \(2015\)](#), [Ptak-Chmielewska \(2019\)](#), and [Gavurova et al. \(2022\)](#) showed higher accuracy of AI models.

The important factor is the research sample size. The used research set shows a typical feature, which is the imbalance between bankrupt and non-bankrupt samples. [Zoričák et al. \(2020\)](#) investigated this problem. Sometimes, this characteristic should be compensated for by using the methods of undersampling or oversampling. The undersampling method with a low number of data could only lead to higher distortion (e.g., [Wang and Liu 2021](#); [Brygała 2022](#)). Conversely, the oversampling method could improve the prediction (e.g., [Chawla et al. 2002](#); [Garcia 2022](#)).

4.1. LR

The LR method facilitates the modeling of a linear relationship between independent variables (financial indicators) and the dependent variable (bankruptcy). The stepwise method consistently eliminates the predictor with the lowest weight in each step. The initial null model encompasses all of the analyzed indicators. The resultant model comprises three of the most crucial indicators: AV/S, E/TA, and ROA. The multicollinearity diagnosis for the resulting model is shown in Table 2. The selected cut-off value adheres to the standard setting of 0.5.

Table 2. Multicollinearity diagnostics.

Variable	Tolerance	VIF
AV/S	0.925	1.082
E/TA	0.946	1.057
ROA	0.940	1.064

Note: VIF denotes variance inflation factor.

One year prior to bankruptcy, the overall prediction accuracy reached 96.1% for the test set and 95.6% for the training set. However, the capacity to distinguish bankrupt samples was merely 50% for the test set and 42% for the training set (see in Table 1).

For the 2-year interval preceding bankruptcy, the accuracy achieved 94.4%. The accuracy for bankrupt samples in the test set was 28.6%, and for non-bankrupt samples, it reached 45.2% (see Table 1). While this model exhibited the lowest accuracy within the comprehensive assessment, the difference in performance was marginal, making it difficult to definitively establish the superiority of any specific method. The results of the performance metrics are in Table 3. Summaries of LR models are in Tables A1 and A2 in Appendix A.

Table 3. Performance metrics.

Metrics	Value
Accuracy	0.958
AUC	0.873
Sensitivity	0.452
Specificity	0.997
Precision	0.933
F-measure	0.609

Note: AUC denotes area under the ROC curve.

LR has already been applied by [Zhang et al. \(1999\)](#) in the USA, who achieved an accuracy of only 78%. In contrast to this work, the authors applied balanced datasets (an equal number of bankrupt and non-bankrupt samples). The selected predictors did not fit this model, except for ROA, which remains representative even after examining

multicollinearity. Although the predictor NWC/TA was equally chosen, it exhibited collinearity with other predictors in this dataset and was discarded in the case of LR. Ptak-Chmielewska (2019) used AT and CuR, similarly to this case, but applied qualitative indicators in addition to proportional ones, which may lead to better predictions. The model was closest to that of Gavurova et al. (2022). The indicators consistently selected were ROS, ROA, ROE, TA, AT, NWC/TA, TD/E, and CuR. Iturriaga and Sanz (2015) showed the superiority of AI over LR, even in the case of 3 years before bankruptcy. Ben Jabeur (2017) applied the PLS-LR (partial least squares LR) method to resolve multicollinearity, resulting in a set of components representing the original predictors. Unlike the stepwise method, where correlated predictors are discarded, PLS transmits the information of the dependent variable and all input indicators, which could improve the accuracy in this case as well. Aker and Karavardar (2023) left the selection of predictors to random forest, and on their data, LR achieved the lowest performance compared to DTs and SVMs.

4.2. ANN

Based on the reviewed literature, the most suitable neural network type seems to be multilayer perceptron (MLP) with feed-forward (FF) connections (without recursion) and learning based on backpropagation (BP). The input layer comprised 12 nodes (11 indicators + 1 bias) and the output was either a value of 0 (healthy company) or 1 (bankrupt company). The model incorporated just one hidden layer with three nodes (+1 bias). Empirical verification demonstrated that increasing the number of hidden nodes and layers had minimal impact on the accuracy in this context, as the chosen values achieved optimal performance. This finding aligns with studies by Gavurova et al. (2022) and the comparative study by Perez (2006). The hyperbolic tangent activation function was selected for its applicability to handle negative values of certain variables.

The model achieved an overall accuracy of 96.1% on the test set and 95.3% on the training set for the 1-year interval before bankruptcy. The accuracy for the bankrupt samples in the test set matched LR at 50%, and for the training set, it reached 38.7%. The accuracy was comparatively lower for the 2-year interval before bankruptcy (as indicated in Table 1). Tables 4 and 5 give the evaluation metrics for ANN models. The network weights of the ANN models are in Tables A3 and A4 in Appendix A.

Table 4. ANN model—evaluation metrics for 1 year.

Metrics	0	1	Average/Total
Support	166	14	180
Accuracy	0.961	0.961	0.961
Precision (Positive Predictive Value)	0.960	1.000	0.963
Recall (True Positive Rate)	1.000	0.500	0.961
False Positive Rate	0.500	0.000	0.250
False Discovery Rate	0.040	0.000	0.020
F1 Score	0.979	0.667	0.955
Matthews Correlation Coefficient	0.693	0.693	0.693
Area Under Curve (AUC)	0.750	0.750	0.750
Negative Predictive Value	1.000	0.960	0.980
True Negative Rate	0.500	1.000	0.750
False Negative Rate	0.000	0.500	0.250
False Omission Rate	0.000	0.040	0.020
Threat Score	11.857	1.000	6.429
Statistical Parity	0.961	0.039	1.000

Note: all metrics are calculated for every class against all other classes.

Table 5. ANN model—evaluation metrics for 2 years.

Metrics	0	1	Average/Total
Support	166	14	180
Accuracy	0.950	0.950	0.950
Precision (Positive Predictive Value)	0.949	1.000	0.953
Recall (True Positive Rate)	1.000	0.357	0.950
False Positive Rate	0.643	0.000	0.321
False Discovery Rate	0.051	0.000	0.026
F1 Score	0.974	0.526	0.939
Matthews Correlation Coefficient	0.582	0.582	0.582
Area Under Curve (AUC)	0.676	0.536	0.606
Negative Predictive Value	1.000	0.949	0.974
True Negative Rate	0.357	1.000	0.679
False Negative Rate	0.000	0.643	0.321
False Omission Rate	0.000	0.051	0.026
Threat Score	9.222	0.556	4.889
Statistical Parity	0.972	0.028	1.000

Note: all metrics are calculated for every class against all other classes.

The use of the MLP network type with one hidden layer aligns with [Tsai and Cheng \(2012\)](#), [Iturriaga and Sanz \(2015\)](#), [Ptak-Chmielewska \(2019\)](#), and [Thanh-Long et al. \(2022\)](#). Conversely, [Korol \(2019\)](#) employed an MLP with two hidden layers. The learning rate range of 0.4–0.8 is supported by other studies ([Odom and Sharda 1990](#); [du Jardin 2010](#); [Thanh-Long et al. 2022](#)). The maximum number of iterations during learning was set to 500, differing from the approaches found in the literature. Another distinction is the learning method; [Iturriaga and Sanz \(2015\)](#) utilized a batch learning method, which they described as more stable compared to the online learning method employed in this work. The choice of predictors can have a significant impact. An unbalanced dataset may affect the accuracy and place more weight on non-bankrupt (minority) samples, a typical issue in bankruptcy data. [Thanh-Long et al. \(2022\)](#) used SMOTE to balance the dataset for use with ANN-BP, a method detailed by [Chawla et al. \(2002\)](#). [Sigrist and Leuenberger \(2023\)](#) enhanced prediction accuracy over ANNs using the boost technique.

4.3. SVM

The method constitutes a supervised learning algorithm that categorizes samples into two groups (bankrupt/non-bankrupt), with the maximum possible gap between these categories based on training samples. The overall achieved accuracy was 96.1% for the 1-year interval and 95% for the 2-year interval (see [Table 1](#)). Remarkably, the results closely mirror those of the ANN model, with the accuracy on the test set being identical. The algorithm setting involved linear weights of the 3rd degree, gamma parameter of 1, and cost of constraints violation of 3. [Tables 6 and 7](#) give the evaluation metrics for the SVM models.

[Tsai and Cheng \(2012\)](#) and [du Jardin \(2018\)](#) describe the use of RBF as the foundation of SVMs, which exhibit a strong predictive ability, a finding supported by [Ptak-Chmielewska \(2019\)](#) where this method achieved the best results. Conversely, [Aker and Karavardar \(2023\)](#) label this method as the weakest among AI methods for predicting 1 year before bankruptcy. The underpinnings of this model were established based on the work of [Zoričák et al. \(2020\)](#). [Shin et al. \(2005\)](#) concluded that an SVM has a higher accuracy rate and better generalization than ANN-BP, a conclusion not validated in this case, as the

results of this model were approximately the same for ANNs or DTs. For a more thorough investigation of this method, testing with several kernel functions would be necessary.

Table 6. SVM model—evaluation metrics for 1 year.

Metrics	0	1	Average/Total
Support	166	14	180
Accuracy	0.961	0.961	0.961
Precision (Positive Predictive Value)	0.960	1.000	0.963
Recall (True Positive Rate)	1.000	0.500	0.961
False Positive Rate	0.500	0.000	0.250
False Discovery Rate	0.040	0.000	0.020
F1 Score	0.979	0.667	0.955
Matthews Correlation Coefficient	0.693	0.693	0.693
Area Under Curve (AUC)	0.750	0.750	0.750
Negative Predictive Value	1.000	0.960	0.980
True Negative Rate	0.500	1.000	0.750
False Negative Rate	0.000	0.500	0.250
False Omission Rate	0.000	0.040	0.020
Threat Score	11.857	1.000	6.429
Statistical Parity	0.961	0.039	1.000

Note: all metrics are calculated for every class against all other classes.

Table 7. SVM model—evaluation metrics for 2 years.

Metrics	0	1	Average/Total
Support	166	14	180
Accuracy	0.950	0.950	0.950
Precision (Positive Predictive Value)	0.949	1.000	0.953
Recall (True Positive Rate)	1.000	0.357	0.950
False Positive Rate	0.643	0.000	0.321
False Discovery Rate	0.051	0.000	0.026
F1 Score	0.974	0.526	0.939
Matthews Correlation Coefficient	0.582	0.582	0.582
Area Under Curve (AUC)	0.679	0.679	0.679
Negative Predictive Value	1.000	0.949	0.974
True Negative Rate	0.357	1.000	0.679
False Negative Rate	0.000	0.643	0.321
False Omission Rate	0.000	0.051	0.026
Threat Score	9.222	0.556	4.889
Statistical Parity	0.972	0.028	1.000

Note: all metrics are calculated for every class against all other classes.

4.4. DT

In a similar vein, the DT method serves as a supervised learning algorithm that relies on observations at the tree's root to make bankruptcy decisions at its leaves. Interestingly, the prediction outcomes for the test set align with those of SVMs and ANNs, even though the training set's accuracy was the highest among all models (see Table 1). The algorithm set-

ting contained a minimum of 50 observations for split and 10 in terminal, with a maximum iteration depth of 30. Tables 8 and 9 give the evaluation metrics for the DT models.

Table 8. DT model—evaluation metrics for 1 year.

Metrics	0	1	Average/Total
Support	166	14	180
Accuracy	0.961	0.961	0.961
Precision (Positive Predictive Value)	0.960	1.000	0.963
Recall (True Positive Rate)	1.000	0.500	0.961
False Positive Rate	0.500	0.000	0.250
False Discovery Rate	0.040	0.000	0.020
F1 Score	0.979	0.667	0.955
Matthews Correlation Coefficient	0.693	0.693	0.693
Area Under Curve (AUC)	0.750	0.750	0.750
Negative Predictive Value	1.000	0.960	0.980
True Negative Rate	0.500	1.000	0.750
False Negative Rate	0.000	0.500	0.250
False Omission Rate	0.000	0.040	0.020
Threat Score	11.857	1.000	6.429
Statistical Parity	0.961	0.039	1.000

Note: all metrics are calculated for every class against all other classes.

Table 9. DT model—evaluation metrics for 2 years.

Metrics	0	1	Average/Total
Support	166	14	180
Accuracy	0.950	0.950	0.950
Precision (Positive Predictive Value)	0.949	1.000	0.953
Recall (True Positive Rate)	1.000	0.357	0.950
False Positive Rate	0.643	0.000	0.321
False Discovery Rate	0.051	0.000	0.026
F1 Score	0.974	0.526	0.939
Matthews Correlation Coefficient	0.582	0.582	0.582
Area Under Curve (AUC)	0.679	0.679	0.679
Negative Predictive Value	1.000	0.949	0.974
True Negative Rate	0.357	1.000	0.679
False Negative Rate	0.000	0.643	0.321
False Omission Rate	0.000	0.051	0.026
Threat Score	9.222	0.556	4.889
Statistical Parity	0.972	0.028	1.000

Note: all metrics are calculated for every class against all other classes.

The TD/E and E/TA indicators carry the most substantial weight in DT-based prediction. It is logical that TD/E holds significant predictive value given that its inverse forms the basis for bankruptcy definition in this study. Giving its inherent importance, this indicator's weight from prior years is naturally pronounced. Notably, this indicator's significance is highlighted in the research by Štefko et al. (2021), Yousaf and Bris (2021), and

Ben Jabeur (2017). The E/TA indicator finds application in the works of Ptak-Chmielewska (2019) and Ben Jabeur (2017).

Additionally, other pivotal indicators include CuR/TA, NWC/TA, ROA, and CuR. Among the analyzed indicators, the DT models have identified these as crucial for bankruptcy determination. On the other hand, the DT models have labeled other analyzed indicators as less relevant for predicting bankruptcy (see relative importance in Tables 10 and 11).

Table 10. DT model—feature importance for 2 years.

Variable	Relative Importance
TD/E	32.609
E/TA	30.435
CuR/TA	15.217
NWC/TA	10.870
ROA	6.522
CuR	4.348

Table 11. DT model—feature importance for 1 year.

Variable	Relative Importance
TD/E	31.481
E/TA	29.630
CuR/TA	16.667
NWC/TA	11.111
ROA	9.259
CuR	1.852

Aker and Karavardar (2023) and Chen et al. (2021) reported DTs as the best-performing model, a finding not validated by this study. A notable distinction is the type of data, as our dataset does not permit the application of market indicators. Chen et al. (2021) highlighted AT as one of the three effective predictive attributes in their DT models. In contrast, in this case, despite the inclusion of this indicator in the input parameters, it was excluded from the resulting DT model due to its low importance in explaining the dependent variable.

5. Conclusions

We conclude that all of the methods achieved good accuracy, and the differences between the individual methods were low.

One of the main limitations is the lower quality of the input data. The first problematic symptom is a low number of bankruptcy samples. With a total of only 45 bankruptcy samples, of which 14 were randomly chosen for testing, the representation remains below 8%. This limitation might be mitigated by incorporating samples from diverse industries or other countries. However, such a strategy introduces issues of comparability due to varying conditions and standards. Moreover, the data face the challenge of an imbalance between the bankrupt and non-bankrupt sets. It is worth noting that this study's primary focus did not involve exploring the impact of data size or imbalance. Another limitation lies in the specific dataset (one country, one industry). It is a challenge to apply this research in another industry or country. Future research could propose models for the next period and compare their accuracy. The recommendation is also to apply the oversampling method to balance the imbalance of the samples, which could increase the accuracy.

As the prediction time horizon extends, a noticeable decline in prediction accuracy is observed. The selection of indicators is limited due to missing data in the financial statements. Few companies meet the conditions for calculating all indicators. In Slovak

conditions, the selection of data necessary for accurate analysis is significantly limited by the size of the industry and its fragmentation.

This paper contributes to the existing literature by designing specific bankruptcy models of the chemical industry applicable in Slovak conditions. The results show that the use of AI-based techniques does not reduce the prediction accuracy. On the contrary, these techniques can increase the prediction accuracy, especially in a longer time horizon.

Author Contributions: Conceptualization, S.L., S.J. and P.V.; methodology, S.L. and S.J.; software, S.L.; validation, S.L., S.J. and P.V.; formal analysis, P.V.; investigation, S.L., S.J. and P.V.; resources, S.L., S.J. and P.V.; data curation, S.L. and S.J.; writing—original draft preparation, S.L., S.J. and P.V.; writing—review and editing, P.V.; visualization, S.L. and P.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Cultural and Educational Grant Agency of the Ministry of Education, Science, Research, and Sport of the Slovak Republic, grant No. 001PU-4/2022-KEGA. This research was funded by the Scientific Grant Agency of the Ministry of Education, Science, Research, and Sport of the Slovak Republic and the Slovak Academy of Sciences, grant No. 1/0741/20-VEGA.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: For requests concerning the data, please contact the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. LR model summary—bankruptcy for 1 year.

Model	Deviance	AIC	BIC	df	ΔX^2	p	McFadden R ²	Nagelkerke R ²	Tjur R ²	Cox & Snell R ²
1	146.302	170.302	219.012	416			0.000	0.000	0.379	0.000
2	146.401	168.401	213.051	417	0.099	0.753	0.000	0.000	0.376	0.000
3	146.445	166.445	207.037	418	0.044	0.833	0.000	0.000	0.375	0.000
4	146.570	164.570	201.102	419	0.125	0.724	0.000	0.000	0.374	0.000
5	146.789	162.789	195.262	420	0.218	0.640	0.000	0.000	0.372	0.000
6	147.609	161.609	190.023	421	0.820	0.365	0.000	0.000	0.369	0.000
7	149.054	161.054	185.409	422	1.445	0.229	0.000	0.000	0.363	0.000
8	150.479	160.479	180.775	423	1.426	0.232	0.000	0.000	0.362	0.000
9	151.871	159.871	176.107	424	1.391	0.238	0.000	0.000	0.362	0.000

Note: AIC denotes Akaike information criterion, BIC denotes Bayesian information criterion, df is degree of freedom, and p is p -value. Null model contains nuisance parameters: AV/S, ROS, E/TA, CuR/TA, CuR, TA, ROE, ROA, AT, TD/E, and NWC/TA.

Table A2. LR model summary—bankruptcy for 2 years.

Model	Deviance	AIC	BIC	df	ΔX^2	p	McFadden R ²	Nagelkerke R ²	Tjur R ²	Cox & Snell R ²
1	130.981	154.981	203.606	413			0.000	0.000	0.438	0.000
2	131.003	153.003	197.576	414	0.022	0.882	0.000	0.000	0.438	0.000
3	131.114	151.114	191.634	415	0.111	0.739	0.000	0.000	0.438	0.000
4	131.275	149.275	185.744	416	0.162	0.688	0.000	0.000	0.440	0.000
5	131.540	147.540	179.957	417	0.265	0.607	0.000	0.000	0.441	0.000
6	131.977	145.977	174.341	418	0.437	0.509	0.000	0.000	0.434	0.000

Table A2. *Cont.*

Model	Deviance	AIC	BIC	df	ΔX^2	p	McFadden R ²	Nagelkerke R ²	Tjur R ²	Cox & Snell R ²
7	133.052	145.052	169.364	419	1.075	0.300	0.000	0.000	0.430	0.000
8	134.742	144.742	165.003	420	1.690	0.194	0.000	0.000	0.417	0.000
9	136.329	144.329	160.537	421	1.587	0.208	0.000	0.000	0.411	0.000

Note: AIC denotes Akaike information criterion, BIC denotes Bayesian information criterion, df is degree of freedom, and p is p -value. Null model contains nuisance parameters: AV/S, ROS, E/TA, CuR/TA, CuR, TA, ROE, ROA, AT, TD/E, and NWC/TA.

Table A3. ANN model—network weights for 1 year.

Node	Layer		Node	Layer	Weight
Intercept		→	Hidden 1	output	−1.103
AV/S	input	→	Hidden 1	output	−0.664
ROS	input	→	Hidden 1	output	−0.257
E/TA	input	→	Hidden 1	output	0.779
CuR/TA	input	→	Hidden 1	output	1.302
CuR	input	→	Hidden 1	output	0.666
TA	input	→	Hidden 1	output	0.054
ROE	input	→	Hidden 1	output	0.653
ROA	input	→	Hidden 1	output	1.888
AT	input	→	Hidden 1	output	−0.700
TD/E	input	→	Hidden 1	output	0.384
NWC/TA	input	→	Hidden 1	output	−0.043
Intercept		→	Hidden 2	output	−0.494
AV/S		→	Hidden 2	output	1.491
ROS	input	→	Hidden 2	output	−0.897
E/TA	input	→	Hidden 2	output	0.008
CuR/TA	input	→	Hidden 2	output	−0.735
CuR	input	→	Hidden 2	output	−0.868
TA	input	→	Hidden 2	output	−0.474
ROE	input	→	Hidden 2	output	−0.697
ROA	input	→	Hidden 2	output	−1.580
AT	input	→	Hidden 2	output	0.981
TD/E	input	→	Hidden 2	output	−1.634
NWC/TA	input	→	Hidden 2	output	0.606
Intercept	input	→	Hidden 3	output	−0.435
AV/S		→	Hidden 3	output	1.201
ROS	input	→	Hidden 3	output	−0.993
E/TA	input	→	Hidden 3	output	1.316
CuR/TA	input	→	Hidden 3	output	−1.275
CuR	input	→	Hidden 3	output	−1.193
TA	input	→	Hidden 3	output	−0.429
ROE	input	→	Hidden 3	output	−0.271
ROA	input	→	Hidden 3	output	−0.600

Table A3. *Cont.*

Node	Layer		Node	Layer	Weight
AT	input	→	Hidden 3	output	0.763
TD/E	input	→	Hidden 3	output	−0.454
NWC/TA	input	→	Hidden 3	output	0.670
Intercept	input	→	1	output	1.809
Hidden 1		→	1	output	−1.258
Hidden 2	input	→	1	output	−2.146
Hidden 3	input	→	1	output	0.936
Intercept	input	→	0	0	0.068
Hidden 1		→	0	0	0.003
Hidden 2	−1	→	0	0	0.262
Hidden 3	−1	→	0	0	−0.272

Note: the weights are input for the hyperbolic tangent activation function.

Table A4. ANN model—network weights for 2 years.

Node	Layer		Node	Layer	Weight
Intercept		→	Hidden 1	output	−0.025
AV/S	input	→	Hidden 1	output	−0.695
ROS	input	→	Hidden 1	output	−0.328
E/TA	input	→	Hidden 1	output	0.471
CuR/TA	input	→	Hidden 1	output	1.227
CuR	input	→	Hidden 1	output	−0.681
TA	input	→	Hidden 1	output	0.982
ROE	input	→	Hidden 1	output	0.077
ROA	input	→	Hidden 1	output	−1.075
AT	input	→	Hidden 1	output	−1.351
TD/E	input	→	Hidden 1	output	0.360
NWC/TA	input	→	Hidden 1	output	0.180
Intercept		→	Hidden 2	output	−2.427
AV/S	input	→	Hidden 2	output	−0.135
ROS	input	→	Hidden 2	output	−0.097
E/TA	input	→	Hidden 2	output	−0.582
CuR/TA	input	→	Hidden 2	output	−0.249
CuR	input	→	Hidden 2	output	−0.928
TA	input	→	Hidden 2	output	0.307
ROE	input	→	Hidden 2	output	−0.255
ROA	input	→	Hidden 2	output	0.170
AT	input	→	Hidden 2	output	0.540
TD/E	input	→	Hidden 2	output	−1.035
NWC/TA	input	→	Hidden 2	output	−0.679
Intercept		→	Hidden 3	output	0.940
AV/S	input	→	Hidden 3	output	−1.224

Table A4. Cont.

Node	Layer		Node	Layer	Weight
ROS	input	→	Hidden 3	output	−0.160
E/TA	input	→	Hidden 3	output	1.316
CuR/TA	input	→	Hidden 3	output	0.175
CuR	input	→	Hidden 3	output	−0.663
TA	input	→	Hidden 3	output	−0.478
ROE	input	→	Hidden 3	output	−0.580
ROA	input	→	Hidden 3	output	1.358
AT	input	→	Hidden 3	output	0.376
TD/E	input	→	Hidden 3	output	−1.210
NWC/TA	input	→	Hidden 3	output	−0.795
Intercept		→	1	output	1.151
Hidden 1	input	→	1	output	−0.344
Hidden 2	input	→	1	output	−0.971
Hidden 3	input	→	1	output	0.229
Intercept		→	0	0	0.447
Hidden 1	−1	→	0	0	0.029
Hidden 2	−1	→	0	0	0.402
Hidden 3	−1	→	0	0	−0.069

Note: the weights are input for the hyperbolic tangent activation function.

References

- Abraham, Rebecca, Mahmoud El Samad, Amer M. Bakhach, Hani El-Chaarani, Ahmad Sardouk, Sam El Nemar, and Dalia Jaber. 2022. Forecasting a Stock Trend Using Genetic Algorithm and Random Forest. *Journal of Risk and Financial Management* 15: 188. [CrossRef]
- Act No. 513/1991 Coll. Commercial Code. 2013. Bratislava: Iura Edition. Available online: <https://www.wipo.int/wipolex/en/text/466905> (accessed on 20 September 2023).
- Act No. 7/2005 Coll. on Bankruptcy and Restructuring. 2013. Bratislava: Iura Edition. Available online: <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2005/7/> (accessed on 20 September 2023).
- Aker, Yusuf, and Alper Karavardar. 2023. Using Machine Learning Methods in Financial Distress Prediction: Sample of Small and Medium Sized Enterprises Operating in Turkey. *Ege Academic Review* 23: 145–62. [CrossRef]
- Alaka, Hafiz A., Lukumon O. Oyedele, Hakeem A. Owolabi, Vikas Kumar, Saheed O. Ajayi, Olugbenga O. Akinade, and Muhammad Bilal. 2018. Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Systems with Applications* 19: 164–84. [CrossRef]
- Altman, Edward I. 1968. Financial Ratios, Discriminant Analysis and The Prediction of Corporate Bankruptcy. *The Journal of Finance* 23: 589–609. [CrossRef]
- Beaver, William H. 1966. Financial Ratios as Predictors of Failure. *Journal of Accounting Research* 4: 71. [CrossRef]
- Ben Jabeur, Sami. 2017. Bankruptcy prediction using Partial Least Squares Logistic Regression. *Journal of Retailing and Consumer Services* 36: 197–202. [CrossRef]
- Brygała, Magdalena. 2022. Consumer Bankruptcy Prediction Using Balanced and Imbalanced Data. *Risks* 10: 24. [CrossRef]
- Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research* 16: 321–57. [CrossRef]
- Chen, You-Shyang, Chien-Ku Lin, Chih-Min Lo, Su-Fen Chen, and Qi-Jun Liao. 2021. Comparable Studies of Financial Bankruptcy Prediction Using Advanced Hybrid Intelligent Classification Models to Provide Early Warning in the Electronics Industry. *Mathematics* 9: 2622. [CrossRef]
- Cortes, Corinna, and Vladimir Vapnik. 1995. Support-vector networks. *Machine Learning* 20: 273–97. [CrossRef]
- du Jardin, Philippe. 2010. Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing* 73: 2047–60.
- du Jardin, Philippe. 2018. Failure pattern-based ensembles applied to bankruptcy forecasting. *Decision Support Systems* 107: 64–77. [CrossRef]
- Fisher, Ronald Aylmer. 1936. Linear Discriminant Analysis. *Statistics & Discrete Methods of Data Sciences* 392: 1–5.

- Fitzpatrick, Paul Joseph. 1932. A Comparison of Ratios of Successful Industrial Enterprises with Those of Failed Firm. *Certified Public Accountant* 6: 727–31.
- Gajdosikova, Dominika, George Lăzăroiu, and Katarina Valaskova. 2023a. How Particular Firm-Specific Features Influence Corporate Debt Level: A Case Study of Slovak Enterprises. *Axioms* 12: 183. [\[CrossRef\]](#)
- Gajdosikova, Dominika, Katarina Valaskova, Tomas Kliestik, and Maria Kovacova. 2023b. Research on Corporate Indebtedness Determinants: A Case Study of Visegrad Group Countries. *Mathematics* 11: 299. [\[CrossRef\]](#)
- Garcia, John. 2022. Bankruptcy prediction using synthetic sampling. *Machine Learning with Applications* 9: 100343. [\[CrossRef\]](#)
- Gavurova, Beata, Sylvia Jencova, Radovan Bacik, Marta Miskufova, and Stanislav Letkovsky. 2022. Artificial intelligence in predicting the bankruptcy of non-financial corporations. *Oeconomia Copernicana* 13: 1215–51. [\[CrossRef\]](#)
- Hsieh, Wen-Kuei, Shang-Ming Liu, and Sung-Yi Hsieh. 2006. Hybrid Neural Network Bankruptcy Prediction: An Integration of Financial Ratios, Intellectual Capital Ratios, MDA, and Neural Network Learning. Paper presented at the 9th Joint International Conference on Information Sciences (JCIS-06), Kaohsiung, Taiwan, October 8–11.
- Iturriaga, Félix J. López, and Iván Pastor Sanz. 2015. Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks. *Expert Systems with Applications* 42: 2857–69. [\[CrossRef\]](#)
- Jenčová, Sylvia, Róbert Štefko, and Petra Vašaničová. 2020. Scoring Model of the Financial Health of the Electrical Engineering Industry's Non-financial Corporations. *Energies* 13: 4364. [\[CrossRef\]](#)
- Kliestik, Tomas, Marek Nagy, and Katarina Valaskova. 2023. Global value chains and industry 4.0 in the context of lean workplaces for enhancing company performance and its comprehension via the digital readiness and expertise of workforce in the V4 nations. *Mathematics* 11: 601. [\[CrossRef\]](#)
- Korol, Tomasz. 2019. Dynamic Bankruptcy Prediction Models for European Enterprises. *Journal of Risk and Financial Management* 12: 185. [\[CrossRef\]](#)
- Korol, Tomasz. 2021. Examining Statistical Methods in Forecasting Financial Energy of Households in Poland and Taiwan. *Energies* 14: 1821. [\[CrossRef\]](#)
- Korol, Tomasz, and Anestis Fotiadis. 2022. Implementing artificial intelligence in forecasting the risk of personal bankruptcies in Poland and Taiwan. *Oeconomia Copernicana* 13: 407–38. [\[CrossRef\]](#)
- Mučko, Przemysław, and Adam Adamczyk. 2023. Does the bankrupt cheat? Impact of accounting manipulations on the effectiveness of a bankruptcy prediction. *PLoS ONE* 18: e0280384. [\[CrossRef\]](#) [\[PubMed\]](#)
- Nagy, Marek, George Lăzăroiu, and Katarina Valaskova. 2023. Machine Intelligence and Autonomous Robotic Technologies in the Corporate Context of SMEs: Deep Learning and Virtual Simulation Algorithms, Cyber-Physical Production Networks, and Industry 4.0-Based Manufacturing Systems. *Applied Sciences* 13: 1681. [\[CrossRef\]](#)
- Odom, Marcus D., and Ramesh Sharda. 1990. A neural network model for bankruptcy prediction. Paper presented at 1990 IJCNN International Joint Conference on Neural Networks, San Diego, CA, USA, June 17–21.
- Ohlson, James A. 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research* 18: 109. [\[CrossRef\]](#)
- Perez, Muriel. 2006. Artificial neural networks and bankruptcy forecasting: A state of the art. *Neural Computing and Applications* 15: 154–63. [\[CrossRef\]](#)
- Ptak-Chmielewska, Aneta. 2019. Predicting Micro-Enterprise Failures Using Data Mining Techniques. *Journal of Risk and Financial Management* 12: 30. [\[CrossRef\]](#)
- Quinlan, John Ross. 1986. Induction of decision trees. *Machine Learning* 1: 81–106. [\[CrossRef\]](#)
- Shin, Kyung-Shik, Taik Soo Lee, and Hyun-jung Kim. 2005. An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications* 28: 127–35. [\[CrossRef\]](#)
- Sigrist, Fabio, and Nicola Leuenberger. 2023. Machine learning for corporate default risk: Multi-period prediction, frailty correlation, loan portfolios, and tail probabilities. *European Journal of Operational Research* 305: 1390–406. [\[CrossRef\]](#)
- Štefko, Róbert, Petra Vašaničová, Sylvia Jenčová, and Aneta Pachura. 2021. Management and Economic Sustainability of the Slovak Industrial Companies with Medium Energy Intensity. *Energies* 14: 267. [\[CrossRef\]](#)
- Taffler, Richard J. 1982. Forecasting Company Failure in the UK Using Discriminant Analysis and Financial Ratio Data. *Journal of the Royal Statistical Society. Series A (General)* 145: 342. [\[CrossRef\]](#)
- Thanh-Long, Ngo, Tran Minh, and Le Hong-Chuong. 2022. A Back Propagation Neural Network Model with the Synthetic Minority Over-Sampling Technique for Construction Company Bankruptcy Prediction. *International Journal of Sustainable Construction Engineering and Technology* 13: 68–79. [\[CrossRef\]](#)
- Tsai, Chih-Fong, and Kai-Chun Cheng. 2012. Simple instance selection for bankruptcy prediction. *Knowledge-Based Systems* 27: 333–42. [\[CrossRef\]](#)
- Wang, Haoming, and Xiangdong Liu. 2021. Undersampling bankruptcy prediction: Taiwan bankruptcy data. *PLoS ONE* 16: e0254030. [\[CrossRef\]](#)
- Yousaf, Muhammad, and Petr Bris. 2021. Assessment of bankruptcy risks in Czech companies using regression analysis. *Problems and Perspectives in Management* 19: 46–55. [\[CrossRef\]](#)
- Zadeh, Lotfi Aliasker. 1965. Fuzzy sets. *Information and Control* 8: 338–53. [\[CrossRef\]](#)
- Zhang, Guogiang, Michael Y. Hu, B. Eddy Patuwo, and Daniel C. Indro. 1999. Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research* 116: 16–32. [\[CrossRef\]](#)

- Zmeškal, Zdeněk, Dana Dluhošová, Karolina Lisztwanová, Antonín Pončík, and Iveta Ratmanová. 2023. Distribution Prediction of Decomposed Relative EVA Measure with Levy-Driven Mean-Reversion Processes: The Case of an Automotive Sector of a Small Open Economy. *Forecasting* 5: 453–71. [[CrossRef](#)]
- Zmijewski, Mark. 1984. Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research* 22: 59–82. [[CrossRef](#)]
- Zoričák, Martin, Peter Gnip, Peter Drotár, and Vladimír Gazda. 2020. Bankruptcy prediction for small- and medium-sized companies using severely imbalanced datasets. *Economic Modelling* 84: 165–76. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.