Artificial Intelligence Predicting Bankruptcy

Subjects: Business, Finance

Contributor: Stanislav Letkovský , Sylvia Jenčová , Petra Vašaničová

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.

bankruptcy prediction logistic regression artificial intelligence

1. Introduction

In financial decision making, the ability to predict bankruptcy or anticipate challenges in meeting financial obligations holds paramount importance (Brygała 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 research 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 Brygała (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).

2. Artificial Intelligence Predicting Bankruptcy

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:

- Result criterion (model accuracy and interpretation of results);
- Data criterion (sample size, dispersion, and variable selection);
- 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.

References

1. Brygała, Magdalena. 2022. Consumer Bankruptcy Prediction Using Balanced and Imbalanced Data. Risks 10: 24.
2. Korol, Tomasz. 2019. Dynamic Bankruptcy Prediction Models for European Enterprises. Journal of Risk and Financial Management 12: 185.
3. Garcia, John. 2022. Bankruptcy prediction using synthetic sampling. Machine Learning with Applications 9: 100343.
4. Altman, Edward I. 1968. Financial Ratios, Discriminant Analysis and The Prediction of Corporate Bankruptcy. The Journal of Finance 23: 589–609.
5. Ohlson, James A. 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research 18: 109.
6. 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.
7. du Jardin, Philippe. 2018. Failure pattern-based ensembles applied to bankruptcy forecasting. Decision Support Systems 107: 64–77.
8. 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.
9. 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.
10. 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.
11. 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.

12. 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.
13. Fitzpatrick, Paul Joseph. 1932. A Comparison of Ratios of Successful Industrial Enterprises with Those of Failed Firm. Certified Public Accountant 6: 727–31.
14. Fisher, Ronald Aylmer. 1936. Linear Discriminant Analysis. Statistics & Discrete Methods of Data Sciences 392: 1–5.
15. Zadeh, Lotfi Aliasker. 1965. Fuzzy sets. Information and Control 8: 338–53.
16. Beaver, William H. 1966. Financial Ratios as Predictors of Failure. Journal of Accounting Research 4: 71.
17. 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.
18. Zmijewski, Mark. 1984. Methodological Issues Related to the Estimation of Financial Distress Prediction Models. Journal of Accounting Research 22: 59–82.
19. 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.
20. Mućko, Przemysław, and Adam Adamczyk. 2023. Does the bankrupt cheat? Impact of accounting manipulations on the effectiveness of a bankruptcy prediction. PLoS ONE 18: e0280384.
21. 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.
22. 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.
23. Wang, Haoming, and Xiangdong Liu. 2021. Undersampling bankruptcy prediction: Taiwan bankruptcy data. PLoS ONE 16: e0254030.
24. 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.
25. 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.

26. Ptak-Chmielewska, Aneta. 2019. Predicting Micro-Enterprise Failures Using Data Mining Techniques. Journal of Risk and Financial Management 12: 30.
27. Korol, Tomasz. 2021. Examining Statistical Methods in Forecasting Financial Energy of Households in Poland and Taiwan. Energies 14: 1821.
28. Korol, Tomazs, and Anestis Fotiadis. 2022. Implementing artificial intelligence in forecasting the risk of personal bankruptcies in Poland and Taiwan. Oeconomia Copernicana 13: 407–38.