Bankruptcy Prediction Using First-Order Autonomous Learning Multi-Model Classifier

Amine Sabek[1] | *Univresity of Tamanghasset, Tamanghasset, Algeria*
Jakub Horák[2] | *Institute of Technology and Business, České Budějovice, Czechia*
Hussam Musa[3] | *Matej Bel University, Banska Bystrica, Slovakia*
Amélia Ferreira da Silva[4] | *Porto Accounting and Business School, Porto, Portugal*

Received 3.6.2024, **Accepted** (reviewed) 15.7.2024, **Published** 13.12.2024

Abstract

Research background: Bankruptcy and financial distress prediction has always been an integral part of any financial management system. It gives an indication to stakeholders to take precautionary measures in order to avoid losses. The traditional approaches for prediction, including logistic regression and discriminant analysis, are constrained by their inability to deal with complex and high-dimensional data (Odom and Sharda, 1990; Min and Lee, 2005). Recent developments in the field of machine learning, and particularly autonomous learning classifiers, present a potential proposed alternative.
Purpose: The purpose of this paper is to propose a first-order autonomous learning classifier (F-O $ALMM_0$) for predicting bankruptcy of business entities and individuals.
Design/methodology/approach: The data file contained a total of 352 companies obtained from the Kaggle database and incorporating 83 financial ratios. Initially, the model's performance was assessed as a preliminary step, but the results were average, followed by the application of Principal Component Analysis (PCA) to enhance the quality of the input's variables. Afterwards, the number of independent variables was reduced to 26. Thus, the results were improved.

Keywords	***DOI***	***JEL code***
Bankruptcy prediction, first-order, autonomous learning, Multi-Model Classifier, Principal Components Analysis	*https://doi.org/10.54694/stat.2024.30*	*C45, C53, C67, G17, G33*

1 Univresity of Tamanghasset, 10034 Airport road, Tamanghasset, Algeria. E-mail: sabek.amine@univ-tam.dz. ORCID: <https://orcid.org/0000-0002-6970-4183>.
2 Institute of Technology and Business in České Budějovice, School of Expertness and Valuation, Okružní 517/10, 37001 České Budějovice, Czechia. E-mail: horak@mail.vstecb.cz. ORCID: <https://orcid.org/0000-0001-6364-9745>.
3 Department of Finance and Accounting, Faculty of Economics, Matej Bel University, Tajoveského 10, 975 90 Banska Bystrica, Slovakia. E-mail: hussam.musa@umb.sk. ORCID: <https://orcid.org/0000-0002-4492-8770>.
4 Porto Accounting and Business School, CEOS PP ISCAP, Polytechnic of Porto, Rua Dr. Roberta Frias 712, 4200-465 Porto, Portugal. E-mail: acfs@iscap.ipp.pt. ORCID: <https://orcid.org/0000-0002-8366-9863>.

INTRODUCTION

The need for an accurate bankruptcy prediction model has become more urgent in recent times. This is because financial markets have become so complex that even the failure of large corporate firms causes a huge impact on the economy. The conventional statistical models are satisfactory up to some extent but generally fail due to the dynamic and nonlinear nature of the financial data. The paper introduces a new first-order autonomous learning classifier, F-O ALMM$_0$, to predict bankruptcy. This model reduces the inefficiencies inherent in traditional models by employing the latest machine learning techniques.

Bankruptcy is one major issue that many companies in different industries and sectors face. It occasions a crumbling of the financial structure of a company, halt in its operation, and dislocates employees into unemployment (Štefko et al., 2020). Basically, it means the company cannot meet its obligations in terms of employees, distributors, suppliers, shareholders, and lenders (Horváthová and Mokrišová, 2018). Only the assets that are in the hands of the judicial authorities remain, and these will be sold to pay off the company's debts. Many things lead to bankruptcy, but the end result is the same. Of these, however, the most significant is neglect; that is, not heeding the signs of bankruptcy before it takes hold (Kanapickienė et al., 2023). Ignoring these signs and doing nothing to arrest them will definitely bring a company to bankruptcy (Zhang et al., 2021).

Bankruptcy prediction is considered to be one of the important solutions in developed countries, being one of the major dependencies of risk management in large corporations (Safi et al., 2022). At the outset, statistical models were devised by focusing on specific financial ratios which actually turned out to possess excellent predictive abilities pertaining to bankruptcy (Srebro et al., 2021).

Artificial intelligence is nowadays considered one of the most powerful contemporary tools in bankruptcy prediction. Previous studies have shown that classification capabilities of intelligent models were more accurate and better than earlier statistical models (Odom and Sharda, 1990; Wilson and Sharda, 1994; Cooper, 1999; Jo et al., 1997; Min and Lee, 2005; Zieba, Tomczak and Tomczak, 2016; Salehi and Pour, 2016; Karas and Reznakova, 2017; Belas et al., 2017). For this reason, research in this area has become very attractive since quality of the results can always be improved. The more advanced AI becomes, the more accurate bankruptcy prediction is.

The aim of this paper is to utilize the F-O ALMM$_0$ model for bankruptcy prediction, which is considered a very important activity not only for companies but also for individuals, as it will help them discover their financial health. This model is a hybrid intelligent multi-model classifier. To test this predictive ability, 83 financial ratios were depended on, and a sample consisting of 352 companies, totaling 2 661 financial instances. This sample will be divided into a training sample and a testing sample. Moreover, the process of financial input purification will be examined to determine its role in improving the quality of this model. The Principal Component Analysis (PCA) technique will be relied upon for this verification. Finally, the results of the model will be compared before and after the input processing. The main question addressed in this paper is: How accurate is the F-O ALMM$_0$ model in bankruptcy prediction, and how effective is the PCA technique in improving the quality of financial inputs and the model's accuracy?

Major contribution of this research is to develop and validate the F-O ALMM$_0$ model for bankruptcy prediction, enhanced by applying PCA in order to improve the quality of the input variables. This contribution comes in various ways:

In this paper, a novel integration of self-directed learning classifiers and dimensionality-reduction techniques are proposed to create an improved method of prediction that is more accurate and reliable.

The empirical results obtained in this study underline the effectiveness of the F-O ALMM$_0$ model in developing the body of knowledge with practical and theoretical implications. The tool works for the early warning of potential bankruptcies, thus helping financial analysts, policymakers, and investors in decision-making processes.

This research fills the gap in literature through different approaches to applying machine learning techniques to financial prediction and provides new insights and empirical evidence that can be used to support the effectiveness of these methods.

The various limitations of traditional models – revealing how F-O $ALMM_0$ overcomes such limitations – will give further depth to the research in the current discussion on improving the models of bankruptcy prediction.

In the following part of the contribution, an overview of the literature is given, focused primarily on specific studies dealing with the issue of predicting bankruptcy of business entities. This is followed by the Methods chapter, in which the used data, methods and selection of variables are described. In the chapter Case Study, the research outputs are presented and evaluated, respectively the ability of the F-O $ALMM_0$ model to predict bankruptcy is tested. The final chapter summarizes the achieved results and describes their application in practice as well as research limitations.

1 LITERATURE REVIEW

The inability of macroeconomic policies to sufficiently address systemic risks, as evidenced in the financial crisis of 2008, is regarded by many as part of the cause of the crisis (Aliu et al., 2022). Indeed, a significant number of analysts argue that the low interest rate policy adopted by the Federal Reserve was the main cause of the housing bubble in the US (Dufour and Orhangazi, 2014). This imperative of increasing lending led to a speculative attitude in lending, and risk management practices were sidetracked (Smaranda, 2014). In such endeavors, various methodologies starting from classical statistical approaches to some of the most promising emerging artificial intelligence techniques have been implemented (Zhou et al., 2012).

The realm of failure prediction literature has been predominantly influenced by the dominance of multiple discriminant analysis for an extended period (Bauer, 2012). Beaver (1966) pioneered the application of univariate discriminant analysis. Blum (1974) formulates the model for failing companies and demonstrates that incorporating periodic revisions into the model fails to enhance the precision of forecasting. Ohlson (1980) introduced conditional probability models as an alternative to the multiple discriminant analysis.

Machine Learning, a sub-domain of Artificial Intelligence, focuses on the development of methodologies and approaches that enable computers to learn (Krulický et al., 2020). Machine Learning can be formulated in various ways (Yu, 2013). The most popular computational intelligence methodologies are found to be very effective in solving nonlinear problems (Zvaríková et al., 2022). Furthermore, these methodologies have the high capability to extract meaningful information from imprecise data and discover complex patterns that cannot be perceived by humans or traditional systems (Cleofas-Sanchez et al., 2016). Notably, Wu and Wang (2000) pioneered the application of neural networks for the assessment of credit risk specifically pertaining to small and medium-sized enterprises (SMEs). Serrano-Cinca (1996) employed self-organizing maps as a methodological approach. Desai et al. (1996) determined that the modular neural network and Multi-Layer-Perceptron (MLP) exhibited particular efficacy in accurately forecasting non-performing loans. Final and Fatih Oglu (2002) developed a hybrid classifier incorporating associative memories and self-organizing maps (SOM) in their approach to speaker recognition. Min and Lee (2005) utilized support vector machines (SVMs) to address the issue of bankruptcy prediction. Hsieh (2005) devised a credit scoring model that used the K-means clustering algorithms and SOM in order to ascertain the optimal inputs for a feed-forward Multilayer Perceptron (MLP). Glezakos et al. (2010) propose an alternative assessment and assert that logistic regression models exhibit high efficacy. Chen et al. (2011) introduced an evolutionary method to concurrently optimize the complexity and weights of a learning vector quantization network, with a focus on symmetric cost preference. Lin and Yang (2012) developed a rolling-logit model that allows the forecasting of corporate bankruptcy in the Taiwan Security Exchange, using current information as well as past information. Cao (2012) introduced a new multiple classifier ensemble model called MCELCCh-FDP that combines different classifiers using firm life cycle and Choquet integral in addressing financial distress. Serrano-Cinca and Gutiérrez-Nieto (2013) used Partial Least Square Discriminant Analysis as a predictive tool for the 2008 banking crisis in the United States. Khashei et al. (2013) applied essential principles of the MLP neural networks and fuzzy logic to construct a hybrid binary credit risk prediction model. Tsai et al. (2014) undertook an extensive study that aimed to compare classifier ensembles using three commonly employed classification techniques, namely decision trees (DT), SVM, and MLP neural networks. Giordani et al. (2014) discussed how adding spline functions to a logistic bankruptcy model improves prediction accuracy by 70% to 90%. This approach identifies complex nonlinear relationships between firm distress and financial metrics of leverage, earnings, and liquidity. Kou et al. (2014) proposed a multi-criteria decision-making (MCDM) framework for prioritizing various clustering algorithms. Kim et al. (2015) proposed the GMBoost, which is a geometric mean-based boosting algorithm and is one of the potential remedies for the class imbalance problem. In the research by Barboza et al. (2017), with a view to predicting bankruptcy a year ahead, a number of machine learning models were fitted, including boosting, support vector machines, bagging, and random forest. Li et al. (2017) used a linear programming algorithm to calculate the efficiency of company stability. Traczynski (2017) introduced a Bayesian model averaging approach to predict bankruptcy, addressing uncertainty in identifying the correct model. Key findings are that only the ratio of total liabilities to total assets and the volatility of market returns consistently predict default across various industries. This new method, which combines information from multiple models or includes industry-specific factors, performs better than traditional single-model approaches.

Angelov and Gu (2017) introduced an innovative 0-order multi-model classifier named Autonomous Learning Multiple-Model ($ALMM_0$-0). Tang et al. (2019) introduced an evolutionary pruning neural network (EPNN) model to predict bankruptcy. Soares et al. (2020) used the zero-order Autonomous Learning Multiple-Model ($ALMM_0$-0*) neuro-fuzzy methodologies, with the primary aim of categorizing diverse cardiac ailments based on auditory signals. Santos et al. (2022) presented the First-Order Autonomous Learning Multi-Model ($ALMM_0$) system as a regressor, which demonstrated the potential for seamless adaptation into a binary classifier. Sabek and Saihi (2023) made a comparison of the results between logistic regression and artificial neural networks in the forecast of financial distress for Saudi Arabia and Algeria. Rainarli and Sabek (2023) applied many machine learning methods to train the prediction model and process missing values and imbalanced data. Sabek (2023) compared two varieties of Artificial Neural Networks (ANNs) to Logistic Regression (LR) in the prediction of financial distress. His conclusion was that the superiority of the networks over LR depends on factors such as the specific network's type and its suitability for the given issue. Sabek and Horak (2023) used Gaussian Process Regression (GPR) to predict financial distress, optimized its hyperparameters to extract the optimal model, and then compared it with other machine learning models. They found that GPR achieved very suitable results. Altman et al. (2023) presented and examined the Omega Score, a new metric designed to improve the prediction of defaults in small and medium-sized enterprises (SMEs). They reconsider the traditional models of default prediction and estimate the effectiveness of the Omega Score in identifying SMEs that are at risk. Valaskova et al. (2023) explored the issue of bankruptcy forecasting in the Visegrad Group countries after the outbreak of COVID-19. They showed how economic disturbances caused by the pandemic had changed bankruptcy risk factors and suggested implications for financial management and policy development in the post-pandemic period.

In summary, therefore, the 2008 financial crisis can be qualified as partial inability of macroeconomic policies to address systemic risks from the housing market. For instance, many analysts have pointed out the low-interest-rate policy by the Federal Reserve as key in the formation of the housing bubble. The speculative lending that followed took no heed of risk management and set base for the wide-scale financial instability.

In the field of failure prediction, traditional methodologies, mostly multiple discriminant analysis, have occupied center stage in this domain for quite a long time. Of late, there have been innovations such as the introduction of univariate discriminant analysis, conditional probability models, and a range of machine learning techniques. More specifically, machine learning methodologies, including neural networks, self-organized maps, and SVMs, have displayed prowess for handling nonlinear phenomena and extracting relevant information from imprecise data.

Key contributions in bankruptcy prediction and credit risk assessment include the use of neural networks for SMEs, modular neural networks and MLPs for non-performing loans, and hybrid classifiers based both on associative memories and on self-organizing maps, for speaker recognition. Other innovative approaches entail logistic regression models, evolutionary methods for the optimization of learning vector quantization networks, and rolling-logit models to predict corporate bankruptcy.

Other methodologies proposed for predicting financial distress and bankruptcy have been the multiple classifier ensemble models, boosting algorithms with geometric mean – based variants, and Bayesian model averaging approaches. They encompass information from various models or involve industry-specific factors, even if they generally outperform the single-model solutions.

Machine learning has been in focus of late which has been experimented upon boosting, bagging, and random forests to predict bankruptcy. Methods like Omega Score for SMEs, Gaussian Process Regression (GPR), and analysis of bankruptcy risk factors post-COVID-19 have also been highlighted for their significant contributions to the field.

Furthermore, Autonomous Learning has become more significant. Researchers have developed new age models that are programmed to learn and improve autonomously with the passage of time without human intervention. Most important ones include the zero-order Autonomous Learning Multiple-Model ($ALMM_0$) and the First-Order Autonomous Learning Multi-Model (ALMM1). The $ALMM_0$ model represents how complex data patterns like cardiac ailments based on auditory signals can be effectively categorized using neuro-fuzzy methodologies. As binary classifiers, it easily adapts. Similarly, the First-Order Autonomous Learning Multi-Model ($ALMM_0$) system points out significant potential as both a regressor and a classifier in relation to autonomous learning in financial risk management and more areas of business beyond this scope.

The field of bankruptcy prediction has undergone a transformation from traditional statistical methods to more advanced machine learning methods and autonomous learning models, providing more precise and dependable tools for managing financial risk.

The first author is a leading expert in the sphere of financial distress and bankruptcy prediction, having an impressive record of scholarly publications that unequivocally prove his in-depth knowledge of and further innovativeness in this domain. He definitely turns out to be outstanding in his collaborative work on how to cope with challenges of missing and imbalanced data in bankruptcy prediction using machine learning. Moreover, he has compared artificial neural networks to logistic regression, proving their models on differentiating financial distress. His research in the optimization of hyperparameters in Gaussian Process Regression further proves skill in predictive accuracy. This proves the adaptability and efficiency of the techniques within different economic contexts. Further, his comparative evaluation of CA Score, Kida, and Springate models for financial distress prediction in Algeria serves as evidence of his comprehensive analytical studies and dedication to advancing the field.

2 METHODS

In this paper, as a first step, the ability of F-O $ALMM_0$ to predict bankruptcy will be tested by training the model using a training sample consisting of 2 001 financial instances, and then testing it using a testing sample consisting of 660 financial instances. In the second phase, the PCA technique will be used to extract only those principal components which have the biggest influence on the dependent variable and hence reduce the input size and improve its quality. The model will then be tested again. The results of the model before and after using the PCA technique will be compared.

F-O $ALMM_0$ is a Multi-model developed by Gu and Angelov (2018), for binary classification purposes and this is basically consistent with the purpose of the current study.[5]

The initial version of the model, Zero-Order $ALMM_0$, was created by Angelov and Gu (2017), and it has been employed in numerous prior research investigations for the purpose of classification. The model has consistently demonstrated a remarkable proficiency in its classification capabilities (Angelov and Gu, 2017; Soares et al., 2020).

Angelov and Gu (2017) examine the $ALMM_0$ general applicability with data drawn from diverse sources without any geographic and temporal limitation. In contrast, Soares et al. (2020) are interested in heart sound classification, and their recorded data usually comes from publicly available medical databases. Their time ranges are not precisely stated but were generally of data up until about 2020.

This motivated us to test the second version of the model First-Order $ALMM_0$ for classification, mainly because, to the best of our knowledge, this version of the model has not been investigated and tested before. According to Angelov et al. (2018), $ALMM_0$ has been realized to form a generic system which can be easily applied for the purpose of multi-model systems connected to probabilistic or other local models. The system is completely data-driven; therefore, it lets its structure be defined by non-parametric data clouds generated from empirical observations and makes no assumption about the distribution or properties of data in any form. This makes the new system capable of acquiring meta-parameters directly from the data and recursively updated, making the efficiency of memory usage and computational calculations within the algorithm more enhanced.

According to Angelov and Gu (2017) and Soares et al. (2020), adopting the self-learning method has several advantages:

- Adaptability: The model learns and updates itself in real-time based on new data, thus being relevant and accurate.
- Simplicity and Interpretability: Being simple and easy to understand makes this method useful in applications.
- Real-Time Processing: Data is processed fast; hence, it is best for scenarios necessiting immediate feedback.
- Robustness: Handles noise and outliers very well, boosting its performance on imperfect data. On the other side, this method is not without its deficiencies:
- Limited to Zero-Order: The zero-order model cannot deal with complex relationships in the data, thus crippling its overall efficiency.
- Scalability Issues: The model could be difficult to manage and scale with increasing data.
- Dependence on Initial Data Quality: It depends a lot on the quality of the first training data set.
- Complexity in Real-World Implementation: This can turn out to be complex in real-world scenarios due to integration and resource management problems in the implementation of the methodology.

2.1 Autonomous learning of Multi-Model systems

For several decades, multi-model systems have been in use in a rather wide spectrum of applications within adaptive control, observers, predictors, and classifiers, and have proven to be an effective tool in dealing with difficulties stemming from uncertainties related to measurements and motion. Actually, their operation is based on the ancient principle of «divide and rule,» where complex problems are broken down into a series of more feasible ones and then integrated together (Angelov et al., 2018).

---

[5] The model code and demo are publicly published at: <*https://www.researchgate.net/publication/322446053_FirstOrder_Autonomous_Learning_Multi-Model_System_source_code_Matlab_version*>.

Autonomous Learning Systems (ALSs) can be perceived as the physical embodiments of artificial intelligence. ALSs can be conceptualized as a convergence of sensor-equipped computational platforms (machines/devices) equipped with software algorithms, enabling these systems to acquire evolving intelligence through interaction with the self-monitoring and external environment. Some of the very basic properties of ALSs include the ability of self-adaptation and self-monitoring; therefore, self-learning or autonomous learning, learning of new knowledge, and update of existing knowledge are very crucial (Angelov, 2012).

The next section describes the learning process for the $ALMM_0$ system in some detail, structured around two major steps: structure identification and parameter identification: (Angelov and Gu, 2018)

For every recently acquired data sample, denoted as x_{K+1}, the global mean μ_K and the average scalar products X_K are updatedto μ_{K+1} and X_{K+1}.

The unimodal discrete density at the x_{K+1} and the central points of the existing data clouds $\mu_{K,i}$ (i = 1, 2 ,…, N_K) are computed using the following equation:

$$D\kappa(x) = \frac{1}{1+\frac{\|x-uk\|^2}{\sigma^2 k}} \quad . \tag{1}$$

Denoted by $D_{K+1}(x_{K+1})$ and $D_{K+1}(\mu_{K,i})$ (i = 1, 2 ,…, N_K). The following principle is examined to determine if x_{K+1} has the capability to generate a novel rule:

$$\text{Cond.1} \qquad \text{IF } (D_{k+1}(x_{K+1}) > \max_{I = 1,2,\dots,N_k} (D_{k+1}(\mu_{K,i}))) \ . \tag{2}$$

$$\text{Or } (D_{k+1}(x_{K+1}) < \min_{I = 1,2,\dots,N_k} (D_{k+1}(\mu_{K,i}))) \ .$$

Then (X_{K+1} is a new focal point).

In case condition 1 is satisfied, a new rule is generated, depending on the value of X_{K+1}. A very important step would then be to check for a possible overlapping between the newly acquired data cloud and the previously existing data clusters. A principle of preventing overlap is utilized in view of the following:

$$\text{Cond.2} \qquad \text{IF } (D_{k+1,i}(x_{K+1}) \geq \frac{1}{1+n^2} \quad .$$

Then the i^{th} facal point and the respective data could needs to be replaced by a new one. (3)

Where: $D_{k+1,i}(x_{K+1})$ is the unimodal discrete density computed per rule (data cluster) using the following equation:

$$D_{k+1,i}(x_{K+1}) = 1 + \frac{\frac{1}{s_{ki}^2\left\|x_{k+1}-u_{k,i}\right\|^2}}{(s_{k,i}+1)(s_{k,i}X_{k,i}+\|x_{k+1}\|^2)-\|x_{k+1}+s_{k,i}u_{k,i}\|^2} \quad . \tag{4}$$

The logical basis for considering $D_{k+1,i}(x_{K+1}) \geq 1/(1 + n^2)$ arises from the well-known Chebyshev inequality, which elucidates the probability of a specific data sample, denoted as x to be n time standard deviation, away from the mean, μ:

$$P(\|x-u\|^2 \leq n^2\,\sigma^2) \geq 1-\tfrac{1}{n^2}\,. \tag{5a}$$

By employing the unimodal discrete density, the Chebyshev inequality can be restated in an elegant manner:

$$P(D_{k+1,i}(x_{K+1}) \geq -\frac{1}{1+n^2}) \geq 1-\tfrac{1}{n^2}\,. \tag{5b}$$

Here, n = 0.5 is used. That is, $(D_{k+1,i}(x_{K+1}) \geq 0.8$ for x_{K+1} is less than $\sigma/2$ away from the central point of the i^{th} data cloud. Put differently, x_{K+1} demonstrates a close proximity to all the points of the i^{th} data cloud. Consequently, x_{K+1} will be able to replace the focal point of the i^{th} data cloud.

In the event that Condition 1 is satisfied and Condition 2 is unfulfilled, a new rule termed «data cloud» with the focal point x_{K+1} is inserted.

$$N_{k+1} \leftarrow N_k + 1, \tag{6a}$$

$$S_{k+1,}N_{k+1} \leftarrow 1, \tag{6b}$$

$$u_{k+1,}N_{k+1} \leftarrow x_{k+1}, \tag{6c}$$

$$X_{k+1,}N_{k+1} \leftarrow \|x_{k+1}\|^2. \tag{6d}$$

In contrast, when Conditions 1 and 2 are concurrently met, the current overlapping data cluster (assuming the i^{th} data cloud) is substituted by a novel one with the central point x_{K+1}, denoted as $(N_{k+1} \leftarrow N_k)$.

$$S_{k+1,i} \leftarrow \left\lceil \frac{1-s_{K,i}}{2} \right\rceil, \tag{7a}$$

$$u_{k+1,i} \leftarrow \frac{x_{k+1}+u_{k,i}}{2}, \tag{7b}$$

$$X_{k+1,i} \leftarrow \frac{\|x_{k+1}\|^2+X_{k,i}}{2}. \tag{7c}$$

The aforementioned principle aims to prevent the discarding of previously gathered information within the ALMM0 system, because the novel data cloud may be initialized by an abnormal data sample.

In the event that Condition 1 fails to meet the required criteria, the value of x_{K+1} is allocated to the closest existing data cloud based on the utilization of the following equation:

$$\text{IF } (j^*) = \underset{i=1,2,\ldots,N}{\text{Arg min}} \left(\|x-u_i\|\right) \text{ Then } (\mathrm{G}_{j*} \leftarrow x). \tag{8}$$

The corresponding quantities are updated as follows $(N_{k+1} \leftarrow N_k)$:

$$S_{k+1,i} \leftarrow S_{k,i} + 1, \tag{9a}$$

$$u_{k+1,i} \leftarrow \frac{S_{k,i}}{S_{k+1,i}} u_{k,i} + \frac{1}{S_{k+1,i}} x_{k+i}, \tag{9b}$$

$$X_{k+1,i} \leftarrow \frac{S_{k,i}}{S_{k+1,i}} X_{k,i} + \frac{1}{S_{k+1,i}} \| x_{k+i} \|^2 . \tag{9c}$$

The descriptors (sample count, dot product, and average) of the remaining data clusters remain unchanged during the subsequent processing iteration. In $ALMM_0$, each data cloud serves as a foundation for constructing the antecedent (IF) part of the fuzzy rules.

2.2 The training algorithm of the First-Order Autonomous Multi-Model

The training algorithm starts by initializing the system and the first cloud as follows (Santos et al., 2022):

$$\begin{cases} K \leftarrow 1 \\ u \leftarrow x_1 \\ E(\|x\|^2) \leftarrow \|x_1\|^2 \\ N \leftarrow 1 \\ f_1 \leftarrow x_1 \\ X_1 \leftarrow \|x_1\|^2 \\ M_1 \leftarrow 1 \\ C_1 \leftarrow \Omega_0 I_{[(n+1)\times(n+1)]} \\ a_1 \leftarrow 0_{[(n+1)\times 1]} \\ B_1 \leftarrow 1 \\ P_1 \leftarrow 0 \end{cases} \tag{10}$$

In the aforementioned context, μ represents the overall mean value of the data points that have been analyzed. N denotes the total count of samples that have undergone analysis. F_k signifies the central point from cloud k. X_k represents the average scalar product of the data points scrutinized by the same cloud. Represents the total count of samples that utilized in generating the aforementioned cloud. Corresponds to the iteration number at which the cloud was formed, and denotes the sum of all previously normalized densities associated with the cloud. M_k denotes the total number of samples used in generating the above-mentioned cloud. B_k is the iteration number where this cloud was created and P_k refers to the summation of all previous normalized densities λ linked to this cloud.

After the initialization, the algorithm proceeds to analyze the following sample.

For the remaining samples, a tri-phase process is followed. The system undergoes each stage consecutively for every sample during the training process. The three stages encompass the ensuing procedures, such as cloud creation/ antecedents update, stale rule removal, the consequents update.

- Cloud creation/Antecedents update:

The initial phase commences with the incrementing of K while simultaneously updating the system's global parameters, namely μ and $E(X^2)$. Subsequently, the unimodal global density is computed for each focal point and the sample under examination:

$$D(x_j) < \min(D(f_i)) \vee D(x_j) < \max(D(f_i)) . \tag{11}$$

If false, the nearest cloud is found, using Formula (12):

$$l = \underset{i = 1,2,\ldots,N}{\operatorname{Arg\,min}} (\|x_j - f_i\|) . \tag{12}$$

Then, the found cloud antecedents are updated, using Formula (13):

$$\begin{cases} M_l \leftarrow M_l + 1 \\ f_l \leftarrow \dfrac{M_l - 1}{M_l} f_l + \dfrac{1}{M_l} x_j \\ X_l \leftarrow \dfrac{M_l - 1}{M_l} X_l + \dfrac{1}{M_l} \| x_j \| \end{cases} . \tag{13}$$

Then, the algorithm proceeds to the next phase.

In the event that condition 1 is true, it becomes necessary to generate the antecedents for a novel rule. Nevertheless, there exists a possibility of sample x_j overlapping with the existing cloud antecedents. Therefore, it becomes imperative to check the position the sample compared to all clouds. In order to accomplish this objective, an update is applied to each cloud based on Formula (13). Subsequently, the unimodal local density of sample x_j is computed for each cloud utilizing their updated antecedents. The logical value of Formula (14) is then cheched:

$$\max(Di(xj) > 0.8. \tag{14}$$

If false, the absence of any identified overlap indicates the necessity for generating a novel cloud using Formula (15):

$$\begin{cases} N \leftarrow N + 1 \\ f_N \leftarrow x_j \\ M_N \leftarrow 1 \\ X_N \leftarrow \| x_j \| \\ B_N \leftarrow K \\ C_N \leftarrow C_1 \leftarrow \Omega_0 I_{[(n+1)\times(n+1)]} \\ A_N \leftarrow \sum_{i=1}^{N-1} \dfrac{A_t}{N-1} \end{cases} . \tag{15}$$

Then, the algorithm proceeds to the subsequent phase. If true, and an overlap is detected, the current existing cloud is determined based on the Formula (16):

$$\mathrm{L} = \underset{i = 1,\dots,\mathrm{N}}{\arg\min} \; (\mathrm{D_i(x_j)}) . \tag{16}$$

Afterwards, a novel cloud is generated over the existing on, using Formula (17):

$$\begin{cases} f_l \leftarrow \dfrac{x_j - f_l}{2} \\ M_l \leftarrow ceil\left(\dfrac{1 - M_l}{2}\right) \\ X_l \leftarrow \dfrac{\|x_j\|^2 + \mathrm{E}(\|x\|^2)_l}{2} \\ B_l \leftarrow K \\ P_l \leftarrow 0 \cdot \end{cases} \tag{17}$$

Then, the algorithm proceeds to the second phase.

- Removal of stale rules:

In the second step, the algorithm initiates updating the local density of x_j for all updated clouds. In the second step, the algorithm starts by updating the local density of x_j for all updated cloud.

They are subsequently normalised according to Formula (18):

$$\lambda_i = \frac{D_i\left(x_j\right)}{\sum_{i=1}^{N} D_i\left(x_j\right)} . \tag{18}$$

Upon acquiring the normalized densities, the utility of each rule, denoted as ηi, is computed based on its antecedent utilizing Formula (19):

$$P_i \leftarrow P_i + \lambda_i . \tag{19}$$

If B_i = K, Then:

$$\eta_i \leftarrow 1 . \tag{20}$$

Otherwise,

$$\eta_i \leftarrow \frac{1}{\mathrm{K} - B_i} P_i . \tag{21}$$

Subsequently, the utility of each rule is subjected to a comparison with a minimum admissible value, η_0, through Formula (22) (condition 3):

$$\eta_i \leftarrow \eta_0 . \tag{22}$$

For every logical value that holds true, the corresponding rule is eliminated, leading to a decrement in the number of clouds. Irrespective of the logical value of condition 3 for each cloud, the algorithm proceeds to the next phase.

- Consequent parameters update:

In conclusion, the algorithm proceeds to update the consequent parameters utilizing the Formulas (23) and (24):

$$C_i \leftarrow C_i - \frac{\lambda_i C_i u_j u_j^T C_i}{1 + \lambda_i u_j^T C_i u_j} , \tag{23}$$

$$a_i \leftarrow a_i - \lambda_i C_i u_j (y_i - u_j^T a_i) . \tag{24}$$

3 DATA AND VARIABLES

In this section, a comprehensive explanation of the data used in the study is provided. Ready-made data extracted from Kaggle were relied upon for this purpose. It should be noted that not all the data was used; some of it was excluded in order to organize the training and testing samples appropriately, as depicted in Table 1.

Table 1 Data set description

Data set characteristics	Multivariate	Area	Business	Number of instances	2 661
Attribute characteristics	Real	Number of attributes	83	Bankrupt instances	1 247
Associated tasks	Classification	Number of companies	352	Normal instances	1 414

Note: This data is multivariate, as it includes 83 financial ratios. It is extracted from the financial statements of real, non-fictitious companies, prepared for classification purposes, and in the field of business. The data is extracted from the Kaggle database. It includes 3 672 financial instances for 422 companies, but this data was filtered and only 2 661 financial instances for 352 companies were used. Divided into 1 247 bankruptcy instances and 1 414 healthy instances.
Source: Own construction

Table 1 illustrates the characteristics of the data used, which are financial data associated with real companies, not fictional ones, specifically designed for the purpose of studying bankruptcy prediction. It is noteworthy that the number of predictors is 83 (financial ratios).

The selection process of the independent variables to be used for bankruptcy detection was beyond the authors. As explained above, these data were already existing and downloaded from the Kaggle website. The large number of 83 variables probably indicates the intention of the data creators to include all financial ratios that could be useful for predicting bankruptcy and which have been mainly used in previous relevant studies.

After data refinement, the analysis relied on data from 352 companies, with a total of 2 661 financial instances distributed as follows: 1 247 bankruptcy instances and 1 414 healthy instances. In Table 2, the total sample is divided into a training sample and a testing sample.

The training sample consisted of 187 companies and included a total of 2 001 financial instances, which were divided into 896 bankruptcy and 1 105 healthy instances. Regarding the test sample, there were data from 165 companies that provided a total of 660 financial instances: 351 cases of bankruptcy and 309 healthy instances.

Table 2 Data set divisions

Train	187 Companies	2 001 Instances	Bankrupt	896
			Normal	1 105
Test	165 Companies	660 Instances	Bankrupt	351
			Normal	309

Note: After examining and sorting the original data, the study sample was extracted, which pertains to the data of 352 companies, with a total of 2 661 financial instances. Since the model must undergo training and testing processes, the data was divided into a training sample that includes 2 001 instances, divided into 896 bankruptcy instances, 1 105 health instances. Secondly, a test sample, which includes 660 instances, divided into 351 bankruptcy instances, 309 health instances.
Source: Own construction

4 RESULTS

In this section, the ability of the F-O ALMM_0 model to predict bankruptcy will be tested. As explained in the previous section, the study sample is divided into a training sample and a testing sample. Using MATLAB, the samples were included. On the other hand, the model code was incorporated, and then the model instructions were applied. As a first step, the command to input the samples into the model was given, followed by training and verification. Lastly, the model's ability to predict bankruptcy was tested. Tabel 1 shows results of the classification accuracy.

Table 3 Confusion matrix

Observed		Predicted	
		Y	
		Bankrupt	Normal
Actual Y	Bankrupt	254	97
	Normal	119	190
Accuracy		67.27	

Note: The confusion matrix aids to identify several elements, firstly, the model's overall classification accuracy, in this case, was 67.27%, this rate is the result of dividing the total number of correctly classified instances by the total number of instances. The intersection of observed bankruptcy and predicted bankruptcy expresses instances that are correctly classified, and the intersection of observed bankruptcy with predicted normal expresses instances that are incorrectly classified. The intersection of observed normal and predicted bankruptcy expresses instances that are incorrectly classified, and the intersection of observed normal with predicted normal expresses instances that are correctly classified.
Source: Own construction

Table 3 shows the classification accuracy for the F-O ALMM$_0$ model after training and subsequent testing. From this table, one can see that the overall accuracy reached 67.27%, which is fairly reasonable for an intelligent model. It correctly classified 190 healthy instances and misclassified 119 instances. It successfully classified 254 bankruptcy instances and misclassified 97 instances, correspondingly. As such, it can be said that the model is considerably challenged during classification at the side of the classification of the healthy instances. Table 4: Accuracy prediction measures, Mean Absolute Error (MAE), Mean Square Error (MSE), and Root Mean Square Error (RMSE), evaluating the model on the ability to predict bankruptcy. Moreover, Recall – a sensitivity measure – was used to test the model for its ability to detect bankruptcy instances. Further on, True Negative Rate (TNR) was applied as a measure of the model's ability to recognize healthy instances.

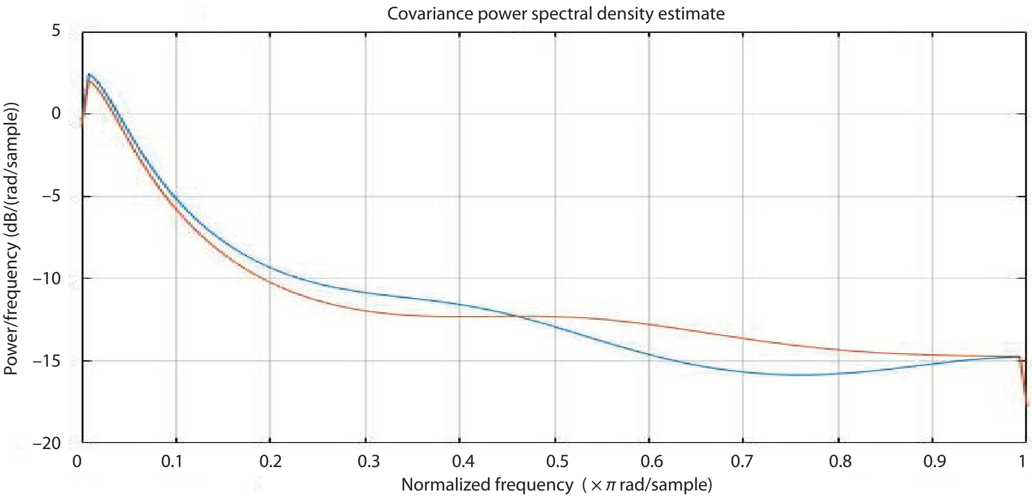
Table 4 Error values

Model	Type I	Type II	Type III	Recall	TNR
F-O ALMM$_0$	0.327	0.327	0.572	68.10	66.20

Note: To evaluate the classification accuracy more specifically, the most important measures that help in determining the model error were used, type I, Mean Absolute Error, type II, Mean Square Error, type III, Root Mean Square Error. Although Recall and TNR are elements of the confusion matrix, they are presented in this table for clarification purposes. Recall to assess the model's capacity in detecting instances of bankruptcy. TNR to evaluate the model's proficiency in identifying healthy instances.
Source: Own construction

Table 4 illustrates the values of the measures used to assess the model's error. These values exhibit relatively high levels, particularly RMSE. This indicates that the model encounters difficulties in accurately classifying bankruptcy. This is reflected in the overall accuracy of 67.27 %. Besides the Recall and TNR measures, moderate values are presented. Note that the TNR rate is lower than the Recall rate, indicating that the model has more difficulty classifying healthy instances compared to classifying instances of bankruptcy. This is also in line with our earlier observation from Table 3. Figure 1 further demonstrates a disparity between actual vs. predicted values.

Figure 1 Predicted vs actual plot



Note: A figure expressing the relationship between the actual values and the predicted values in the form of two curves, the first curve in blue represents the actual values, while the red curve represents the predicted values. The figure highlights the observed convergence between the two curves at some point, and the observed dissonance at some point.
Source: Own construction

The blue curve indicates the actual values, while the red shows the expected values. It can be noted that the two curves, at a point within the range of 0 to 0.1, came very close but this did not persist to result in divergence in other ranges. In order to improve the accuracy of the model in predicting bankruptcy, it was necessary to first improve the quality of the inputs. As mentioned earlier, the results of the model were not satisfactory, and this can be attributed to the presence of impurities that needs to be eliminated. That is in case it removes unneeded variables that confuse the learning of the model and retains only the variables which have a high impact on the dependent variable. It is important to note that 83 predictors are a very large number and should be reduced to a number that allows the model to be well trained. However, the number of predictors should not be reduced arbitrarily; rather a systematic method called principal component analysis, or PCA, ought to be followed. This is one such technique for extracting those independent variables which have most influence on the dependent variable. A prerequisite to the use of PCA is that there must be some multicollinearity between the predictors; if there is no multicollinearity, the inputs are all independent, and hence PCA would never be needed.

Table 5 illustrates the results of the test for detecting multicollinearity among the independent variables. It should be noted that the variables presented in the Table 5 are a random sample of the overall test results used for illustrative purposes only. Based on the results shown in Table 5, the presence of linear multicollinearity among the variables is inferred. This is attributed to the shrinkage of Tolerance values and the inflation of VIF values. Regarding the Tolerance measure, the more its values are inflated and approached 1, the more this indicates the fading of linear multicollinearity, and vice versa. As observed in Table 5, all Tolerance values are very small and close to zero. With respect to the VIF measure, the more its values shrink and do not exceed the threshold of 3, the more this indicates the fading of linear multicollinearity, and vice versa. Furthermore, it is noteworthy that all VIF values in the Table 5 confirm the presence of multicollinearity since their values exceeded 10 in all instances. Table 6 presents the initial and extracted values of the independent variables. As mentioned earlier, the variables presented in the Table 6 are a random sample of the overall test results used for illustrative purposes only.

Table 5 Multicollinearity test before PCA

Predictors	Collinearity statistics	
	Tolerance	VIF
X25	0.036	27.455
X33	0.039	25.357
X34	0.021	48.174
X38	0.011	93.841
X48	0.045	22.404
X51	0.023	43.290
X63	0.015	67.559
X64	0.012	84.181
X70	0.010	95.799
X73	0.007	148.365
X77	0.009	110.877
X81	0.033	30.697

Note: A statistical test using linear regression for the purpose of examining the selected data, and ascertaining whether there is an overlapping relationship between the independent variables or not. The test depends on two basic indicators, Tolerance and VIF, if the values of tolerance are small and do not approach 1, and the values of VIF are very inflated and exceed 3. This indicates the existence of multicollinearity between the variables.
Source: Own construction

Table 6 Communalities

Predictors	Initial	Extraction
X25	1	0.960
X33	1	0.917
X34	1	0.992
X38	1	0.970
X48	1	0.969
X51	1	0.851
X63	1	0.848
X64	1	0.931
X70	1	0.925
X73	1	0.847
X77	1	0.959
X81	1	0.976

Note: A statistical sub-test of the outputs of the Principal Components Analysis. This test is based on two basic indicators, the initial values, the extracted values. The second indicator is the most important, as it indicates the extent to which the data is well represented in the appropriate manner that aids in extracting suitable principal components. If the extracted values exceed 0.75, this indicates the success of the statistical test in extracting the suitable components.
Source: Own construction

Through Table 6, it is observed that the initial value remains constant at 1, which is favorable. However, our primary concern lies in the extracted value, as it indicates the extent to which the data is well represented in the appropriate manner that aids in extracting suitable principal components.

It is evident from Table 6 that the extracted value exceeds 0.75 in all instances, which is highly suitable and indicates the effectiveness of PCA in extracting the principal components. Table 7 illustrates the correlation between the financial ratios and the extracted principal components. It is worth noting once again that the variables and components presented in the Table 7 are a random sample from the overall test results, used for illustrative purposes only.

Table 7 Rotated components matrix

Financial ratios	Components							
	1	2	3	4	5	6	7	8
X25	0.031	**0.340**	0.032	–0.025	0.018	–0.002	0.018	0.907
X33	0.045	–0.003	0.200	0.002	0.165	–0.006	**0.916**	0.004
X34	0.022	0.002	–0.021	–0.026	–0.003	–0.002	0.006	–0.003
X38	0.022	–0.044	–0.030	–0.018	–0.007	**0.959**	–0.007	–0.002
X48	0.076	0.093	–0.038	–0.041	0.021	–0.005	0.004	**0.970**
X51	**0.381**	0.164	0.135	–0.061	–0.085	–0.285	0.102	0.134
X63	–0.178	–0.038	0.008	**0.850**	0.007	–0.035	0.017	–0.011
X64	–0.794	0.016	0.056	**0.343**	–0.006	–0.011	–0.021	–0.037
X70	**0.347**	0.048	–0.036	–0.840	0.032	0.012	–0.002	0.069
X73	–0.307	0.067	0.111	0.274	0.005	0.017	–0.019	–0.013
X77	**0.927**	–0.005	–0.018	–0.065	0.028	0.023	0.026	0.033
X81	0.105	0.121	–0.041	–0.060	0.034	–0.003	0.008	**0.966**

Note: A statistical sub-test of the outputs of the Principal Components Analysis. The test expresses the relationship between the extracted principal components and the financial ratios. Principal components that have relationships of values greater than 0.3 with three or more financial ratios are considered strong components. Principal components that have relationships of values greater than 0.3 with fewer than three components are considered weak components.
Source: Own construction

As a final result of the PCA test, 26 principal components were extracted. These components have the highest influence on the dependent variable. Table 7 illustrates the relationship between the extracted principal components and the financial ratios. It is worth noting that a component that does not have a high correlation with a value ≥ 0.3 with three or more financial ratios is considered a weak component and preferable to be excluded. 18 strong components were observed, alongside 8 components showing weak correlation. To verify the disappearance of multicollinearity among predictors, the new data will be subjected to the test of multicollinearity.

Table 8 demonstrates that multicollinearity has definitely disappeared, and the results presented above are contrary to the results of the multicollinearity test prior to PCA testing, as indicated in Table 5. It is noteworthy that the Tolerance has inflated and become constant at a value of 1, while the VIF value has decreased and does not exceed 3 in all instances. Now that the principal components have been extracted, the model can be tested again. But experimenting with 26 components relying was found to be more accurate than relying on 18 components only. That means the weak components also play a significant role in improving accuracy. Table 9 shows the accuracy of the model classification after applying the PCA and extracting the suitable principal components.

Table 8 Multicollinearity test after PCA

Independent variables	Collinearity statistics	
	Tolerance	VIF
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	1	1
8	1	1
9	1	1
10	1	1
11	1	1
12	1	1

Note: A statistical test using linear regression for the purpose of examining the selected data, and ascertaining whether there is an overlapping relationship between the independent variables or not. The test depends on two basic indicators, Tolerance and VIF, if the values of tolerance are small and do not approach 1, and the values of VIF are very inflated and exceed 3. This indicates the existence of multicollinearity between the variables.
Source: Own construction

Table 9 Confusion matrix after PCA

Observed		Predicted	
		Y	
		Bankrupt	Normal
Actual Y	Bankrupt	269	82
	Normal	67	242
Accuracy		77.42	

Note: The confusion matrix aids to identify several elements, firstly, the model's overall classification accuracy, in this case, was 77.42%, this rate is the result of dividing the total number of correctly classified instances by the total number of instances. The intersection of observed bankruptcy and predicted bankruptcy expresses instances that are correctly classified, and the intersection of observed bankruptcy with predicted normal expresses instances that are incorrectly classified. The intersection of observed normal and predicted bankruptcy expresses instances that are incorrectly classified, and the intersection of observed normal with predicted normal expresses instances that are correctly classified.
Source: Own construction

Table 9 presents the accuracy of classification from F-O $ALMM_0$ model after training and testing. As noted, the overall accuracy was 77.42%, which is satisfactory in its totality. Improvement in the model is highly significant because it managed to classify correctly 269 bankruptcy instances; whereas, it misclassified 82 instances. It also classified 242 healthy instances correctly and misclassified 67. Notice that the misclassification rate of the model concerning the healthy instances has decreased, and the classification ability has increased after applying the PCA technique. Table 10 focuses on the model's ability to predict bankruptcy by measures of prediction accuracy.

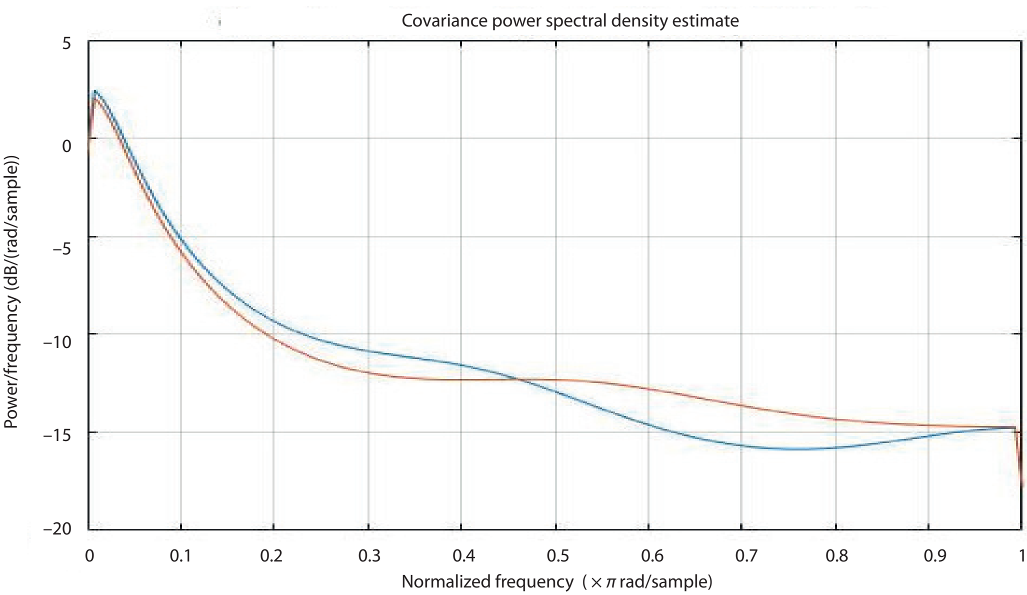
Table 10 Error values after PCA

Model	Type I	Type II	Type III	Recall	TNR
F-O ALMM$_0$	0.227	0.227	0.476	79.82	74.61

Note: To evaluate the classification accuracy more specifically, the most important measures that help in determining the model error were used, type I, Mean Absolute Error, type II, Mean Square Error, type III, Root Mean Square Error. Although Recall and TNR are elements of the confusion matrix, these metrices are shown in this table for clarification purposes. Recall to assess the model's capacity in detecting instances of bankruptcy. TNR to evaluate the model's proficiency in identifying healthy instances.
Source: Own construction

Table 10 illustrates the values of the model's error. It is noteworthy that these values have significantly decreased, indicating an improvement in the model's quality in predicting bankruptcy. This is reflected in the overall accuracy of 77.42. Additionally, the Recall and TNR indicate suitable values as well. It is interesting to note that the TNR rate still stands below the Recall rate, indicating that the model still tends to misclassify healthy instances in comparison with bankruptcy instances. It can be noted that there is a difference between the actual and expected values from Figure 2.

Figure 2 Predicted values vs actual plot



Note: A figure expressing the relationship between the actual values and the predicted values in the form of two curves, the first curve in blue represents the actual values, while the red curve represents the predicted values. The figure highlights the observed convergence between the two curves at some point, and the observed dissonance at some point.
Source: Own construction

The impact of the PCA technique is clearly evident through Figure 2, where the level of compatibility between the actual values and the expected values is noticeable. In Figure 1, the compatibility was limited to the horizontal range (0–0.1) only, but in this case, the compatibility between the two curves takes a longer range (0–0.3). Then, a slight divergence in the range (0.3–0.7) is observed, followed by renewed compatibility in the range (0.7–1). Table 11 illustrates the comparison between the results of the model before and after using PCA.

Table 11 Parameters for evaluation

Measure	Before PCA	After PCA
Error 1	0.327	0.227
Error 2	0.327	0.227
Error 3	0.572	0.476
Sig.	0.00	0.00
R^2	0.116	0.297
Cov.	0.084	0.136
Recall	68.10	79.82
TNR	66.20	74.61
Precision	66.20	74.69
F_1	0.672	0.771
Acc.	67.27	77.42

Note: To compare the different results of the model before and after using PCA, the most important mathematical and statistical measures for performance evaluation were combined. Firstly, the error measures, type I, Mean Absolute Error, type II, Mean Square Error, type III, Root Mean Square Error. Secondly, the elements of the confusion matrix, Recall, TNR, Precision, F1, and Accuracy. Thirdly, the statistical measures, Significance, Coefficient of determination, and Covariance.
Source: Own construction

Through Table 11, the model can be evaluated mathematically using error measures, and statistically using certain statistical measures. The statistical significance of the model can be assessed through (Sig), the correlation between the actual and predicted values can be evaluated using (R^2), and the degree of correlation in the variation between actual and predicted values can be assessed using (Cov). F1-score is the harmonic of Recall and precision, and the higher the F1, the better the predictive accuracy of the classification procedure. All in all, it can be noted that all the values given in Table 11 reflect an improvement in the quality of the model after applying the PCA.

5 DISCUSSION

The results presented in Tables 3 to 11 provide valuable insights into the performance and improvements made to the F-O $ALMM_0$ model for bankruptcy prediction. In this discussion, the analysis and interpretation of these results aim to assess the model's accuracy, identify challenges faced, and evaluate the impact of Principal Component Analysis (PCA) on model enhancement.

5.1 Interpreting findings in the context of existing research

Table 3 illustrates the initial classification accuracy of the F-O $ALMM_0$ model. The overall attained accuracy of 67.27% is at a moderate level for an intelligent model. Even though it classifies 254 instances of bankruptcy correctly, it struggles with the classification of healthy instances, as it misclassified 119 of them. This obviously proves that there are major challenges in separating the two instances, which is very important in financial risk assessment. This agrees with previous research indicating that most financial models often misclassify healthy firms because their financial characteristics overlap with those of distressed firms (Altman, 1968; Ohlson, 1980).

5.2 Challenges in model performance

Table 4 presents an overview of error measures used to evaluate model performance. A relatively high RMSE with a medium range for Recall and TNR, it indicates that the model suffers from classifying bankruptcy correctly. Besides, one can mention that the rate of TNR is lower than Recall indicating thus a greater challenge in classifying healthy instances. This challenge is also consistent with the findings from related studies where models create a bias towards the bankruptcy instances due to its relatively lower occurrence in datasets (Beaver, McNichols and Rhie, 2005).

5.3 Improving input data quality with PCA

The discussion then focuses on the necessity of optimizing the quality of input data. It will be shown that the number of 83 predictors will cause problems with the training of the model. A solution will be the systematic use of PCA in extracting influential independent variables and eliminating multicollinearity among predictors, as presented in Table 5. In this table, the results indicate that there is a significant problem of multicollinearity since Tolerance values approaching zero and VIF valuesexceeding 10, a fact that means the necessity of using PCA.

Table 6 reveals that the extraction of principal components in all instances is above 0.75, which proves that it is appropriate to apply PCA for reducing dimensions. From the PCA test shown in Table 7, it can be seen that 26 principal components are obtained. Notably, 18 are strongly correlated and 8 are weakly correlated to financial ratios. The selection procedure reduced the multicollinearity, as confirmed in Table 8 with Tolerance reaching the constant value of 1 and VIF values not exceeding 3. Remarkably, it is relying on all 26 components, which seems more efficient to improve the accuracy of classification.

The extraction of 26 principal components, of which 18 correlated highly with financial ratios, supports the literature that suggests dimensionality reduction can be a method to enhance the performance of models by reducing multicollinearity (Jolliffe, 2002).

5.4 Enhanced model performance post-PCA

Table 9 presents the improved performance of F-O ALMM$_0$ model after the use of PCA. The accuracy increased to 77.42%, hence there is a significant improvement. It can be noted that this model misclassifies fewer healthy instances, demonstrating the efficacy of principal component analysis in addressing the challenges of the model. Table 10 further highlights the improved quality of the model in the significantly decreased values of error. The Recall and TNR values improve considerably; however, the rate of TNR remains low, indicating the continued challenge in classifying healthy instances. Table 11 presents a comprehensive assessment of model performance, including statistical measures. It is shown that the improvement in the quality of the model is very high. The results obtained from F1-Score also prove the predictive ability of the model. The results align with previous studies advocating PCA's utility in optimizing predictive models by improving feature relevance andreducing dimensionality (Wold, Esbensen and Geladi, 1987).

5.5 Explanation for results

The improved performance of the model can be attributed to many factors. Dimensionality reduction in the input data through PCA certainly helped lessen the impact of both multicollinearity and overfitting. By selecting the most influential principal components, the model focused on a lesser number of relevant features, hence improving the predictive accuracy. Further, from the result showing strong correlations between the selected components and financial ratios indicate that PCA effectively retained significant information for accurate classification.

The initial moderate performance and improvement afterwards underline quite significantly that data preprocessing is very essential for a machine learning model. In this case, probably the large number of predictors used initially might have introduced noise and redundancy, which PCA effectively reduced. This aligns with prior literature highlighting feature selection and dimensionality reduction as a significant approach in improving model performance (Guyon and Elisseeff, 2003).

5.6 Discussion of similarities and differences

Compared to existing studies, we notice both similarities and differences. The medium accuracy initially, then its improved performance after PCA are align with prior studies that highlight the advantages of dimensionality reduction (Wold, Esbensen and Geladi, 1987; Jolliffe, 2002; Guyon and Elisseeff, 2003). In contrast, Chen and Du's (2009) study experimental results show that factor analysis exacerbates the misclassification errorleading to failure companies being incorrectly identified as healthy companies. This may be due to the peculiarity of sample, such as the distribution of financial health or due to economic environment during the collection of data.

5.7 Summarizing the discussion

In summary, the application of PCA played a significant role in the enhancement of the performance of the F-O ALMM$_0$ model for bankruptcy prediction. Accuracy of the model, statistical significance, and error measures have all improved, with particular benefits in reducing misclassification of healthy instances. These results suggesting proper preprocessing and dimensionality reduction are important steps when developing effective predictive models within the financial domain.

Future research may explore more dimensionality reduction techniques and integration of alternative machine learning algorithms to increase further robustness in the model. Further, examination into how the diverse financial environments and different sample sizes impact model performance may provide deeper insights into the generalizability of the findings.

By contextualizing these results within existing research, providing explanations for observed outcomes, and summarizing key points, this discussion section aligns with standard expectations and offers a comprehensive analysis of the study's findings and implications.

CONCLUSION

This paper tests the predictive ability of the intelligent Multi-model, F-O ALMM$_0$ for bankruptcy prediction. A large study sample is used in this respect, consisting of a training sample with 2 001 financial instances and a testing sample with 660 financial instances. After reviewing the literature and explaining the model's structure, the practical part was divided into four main stages. In the first stage, the model's performance was tested after training it using 83 predictors. However, the model yielded only average or modest results, achieving a classification accuracy of only 67.27 %. This raised the question: What is wrong? And how can the model's results be improved?

The first thing that caught attention was the sheer number of predictors, which is indeed beneficial for gaining a comprehensive understanding of a company's financial status in a given fiscal year. However, as much as it is advantageous, it can also become a drawback, because the data may contain impurities and conflicting information that impede the model's learning ability. Therefore, as a second stage, the data was processed using Principal Component Analysis (PCA) technique, extracting 26 principal components, including 8 weak components. Nevertheless, experimental results demonstrated that relying on all components without excluding the weak ones yielded better outcomes.

In the third stage, the performance of the model was tested once more with the extracted predictors. Obvious improvement in the quality of the model was realized as it attained a classification accuracy of 77.42%. In the fourth stage, a comparison was made between the results before and after applying PCA on the model with relying on mathematical and statistical measures. It has been concluded that this intelligent model achieves highly appropriate results in bankruptcy prediction, especially when the input features are pre-processed using the PCA.

The research would develop and validate the F-O $ALMM_0$ model for the advancement of financial prediction field, wherein, through the application of principal component analysis to improve the quality of variables provided as input, it would not be merely a theoretical advancement about autonomous learning method but providing a tangible solution to enhance financial stability and decision-making. The findings are of practical relevance to investors, financial analysts, and policymakers, providing a robust tool for the early warning system of potential bankruptcies.

Besides these significant findings, the research presents a number of limitations. The first major limitation pertains to conduct a comparison analysis between the initial and subsequent versions of the model. Further, this study may involve testing the model using a larger and diverse sample that includes diverse temporal and spatial scope.

This approach can be adapted for application in other markets using more realistic data from established sources. Besides, the same approach can be adapted to categorize challenges across various fields beyond finance. The future research is expected to develop hybrid intelligent models further to address classification issues in both the finance and marketing fields. Also, it is anticipated to test other data processing tools, with a focus on Lasso Regression particularly, be used to enhance the processing methodology of raw data.

ACKNOWLEDGMENT

This research was supported by Institute of Technology and Business in České Budějovice, the project: IVSUZO2301 – the impact of the circular economy on the share prices of companies listed on the stock exchange.

This research was supported by the Scientific Grant Agency of Slovak Republic under progect VEGA No. 1/0479/23 – Research of circular consumer behavior in the context of STP marketing model. The authors would like to express their gratitude to the Scientific Grant Agency of the Ministry of Education, Research, Development, and Youth of the Slovak Republic for financial support of this research and publication. However, the responsibility for the facts stated, opinions expressed, and the conclusions drawn are entirely of the authors.

This work is financed by Portuguese national funds through FCT – Fundação para a Ciência e Tecnologia, under the project UIDB/05422/2020.

References

ALIU, F., HASKOVA, A., SULER, P. (2022). Sustainability of electricity prices and the consequences for the Prague Stock Exchange [online]. *Entrepreneurship and Sustainability Issues*, 10(2): 473–494. <https://doi.org/10.9770/jesi.2022.10.2(30)>.

ALTMAN, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 23(4): 589–609.

ALTMAN, E. I., BALZANO, M., GIANNOZZI, A., SRHOJ, S. (2023). The Omega Score: An improved tool for SME default predictions [online]. *Journal of the International Council for Small Business*, 4(4): 362–373. <https://doi.org/10.1080/26437015.2023.2186284>.

ANGELOV, P. (2012). *Autonomous learning systems: from data streams to knowledge in real time*. John Wiley & Sons, Ltd. ISBN 978-1-118-48176-9

ANGELOV, P., GU, X. (2017). Autonomous learning multi-model classifier of 0-Order (ALMMo-0). *Evolving and Adaptive Intelligent Systems,* Ljubljana, Slovenia, pp. 1–7.

ANGELOV, P. P., GU, X. J. C., PRÍNCIPE, J. C. (2018). Autonomous Learning Multimodel Systems From Data Streams [online]. *IEEE Transactions on Fuzzy Systems*, 26(4): 2213–2224. <https://doi.org/10.1109/TFUZZ.2017.2769039>.

BARBOZA, F., KIMURA, H., ALTMAN, E. (2017). Machine learning models and bankruptcy prediction [online]. *Expert Systems With Applications,* 83: 405–417. <https://doi.org/10.1016/j.eswa.2017.04.006>.

BAUER, J. (2012). *Bankruptcy Risk Prediction and Pricing: Unravelling the Negative Distress Risk Premium*. PhD Thesis, Cranfield School of Management, Cranfield University, England, 255 p.
BEAVER, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 4: 71–111.
BEAVER, W. H., MCNICHOLS, M. F., RHIE, J. W. (2005). Have Financial Statements Become Less Informative? Evidence from the Ability of Financial Ratios to Predict Bankruptcy [online]. *Review of Accounting Studies*, 10(1): 93–122. <https://doi.org/10.1007/s11142-004-6341-9>.
BELAS, J., MISANKOVA, M., SCHONFELD, J., GAVUROVA, B. (2017). Credit risk management: financial safety and sustainability aspects [online]. *Journal of Security and Sustainability Issues*, 7: 79–93. <https://doi.org/10.9770/jssi.2017.7.1(7)>.
BLUM, M. (1974). Failing Company Discriminant Analysis [online]. *Journal of Accounting Research*, 12(1): 1–25. <https://doi.org/10.2307/2490525>.
CAO, Y. (2012). MCELCCh-FDP: Financial distress prediction with classifier ensembles based on firm life cycle and Choquet integral [online]. *Expert Systems with Application,* 39(8): 7041–7049. <https://doi.org/10.3390/math8081275>.
CHEN, N., RIBEIRO, B., VIEIRA, A. S., DUARTE, J., NEVES, J. C. (2011). A genetic algorithm-based approach to cost-sensitive bankruptcy prediction [online]. *Expert Systems with Applications,* 38(10): 12939–12945. <https://doi.org/10.1016/j.eswa.2011.04.090>.
CHEN, W., DU, Y. (2009). Using neural networks and data mining techniques for the financial distress prediction model [online]. *Expert Systems with Applications*, 36(2): 4075–4086. <https://doi.org/10.1016/j.eswa.2008.03.020>.
CLEOFAS-SANCHEZ, L., GARCIA, V., MARQUES, A. I., SANCHEZ, J. S. (2016). Financial distress prediction using the hybrid associative memory with translation [online]. *Applied Soft Computing,* 44: 144–152. <https://doi.org/10.1016/j.asoc.2016.04.005>.
COOPER, J. C. B. (1999). Artificial neural networks versus multivariate statistics: an application from economics. *Journal of Applied Statistics,* 26(8): 909–921.
DESAI, V. S., CROOK, J. N., OVERSTREET JR, G. A. (1996). A comparison of neural networks and linear scoring models in the credit union environment. *Eur. J. Oper. Res*, 95(1): 24–37.
DUFOUR, M., ORHANGAZI, O. (2014). Capitalism, Crisis, and Class: The United States Economy after the 2008 Financial Crisis [online]. *Review of radical political economics,* 46(4): 461–472. <https://doi.org/10.1177/0486613414537981>.
FINAL, M. Y., FATIH OGLU, Y. (2002). Self organizing map and associative memory model hybrid classifier for speaker recognition. In: *Proc. 6th Seminar on Neural Network Applications in Electrical Engineering*, Belgrade, pp. 71–74.
GIORDANI, P., JACOBSON, T., SCHEDVIN, E. VON, VILLANI, M. (2014). Taking the Twists into Account: Predicting Firm Bankruptcy Risk with Splines of Financial Ratios [online]. *Journal of Financial and Quantitative Analysis*, 49(4): 1071–1099. <https://doi:10.1017/S0022109014000623>.
GLEZAKOS, M., MYLONAKIS, J., OIKONOMOU, K. (2010). An empirical research on early bankruptcy forecasting models: Does logit analysis enhance business failure predictability? *European Journal of Finance & Banking Research*, 3: 1–15.
GU, X., ANGELOV, P. P. (2018). First-order autonomous learning multi-model system (Source code, Matlab version) [online]. [cit. 9.8.2023] <https://www.researchgate.net/publication/322446053_First-Order_Autonomous_Learning_Multi-Model_System_source_code_Matlab_version>.
HORVÁTHOVÁ J., MOKRIŠOVÁ, M. (2018). Risk of Bankruptcy, Its Determinants and Models [online]. *Risks,* 64(117). <https://doi.org/10.3390/risks6040117>.
HSIEH, N. C. (2005). Hybrid mining approach in the design of credit scoring models [online]. *Expert Syst. Appl*, 28(4): 655–665. <https://doi.org/10.1016/j.eswa.2004.12.022>.
JO, H., HAN, I., LEE, H. (1997). Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis [online]. *Expert Systems with Applications*, 13(2): 97–108. <https://doi.org/10.1016/S0957-4174(97)00011-0>.
JOLLIFFE, I. T. (2002). *Principal Component Analysis.* 2nd Ed. Springer Series in Statistics, Springer.
KANAPICKIENĖ, R., KANAPICKAS, T., NEČIŪNAS, A. (2023). Bankruptcy Prediction for Micro and Small Enterprises Using Financial, Non-Financial, Business Sector and Macroeconomic Variables: The Case of the Lithuanian Construction Sector [online]. *Risks*. 11(5): 1–33. <https://doi.org/10.3390/risks11050097>.
KARAS, M., REZNAKOVA, M. (2017). Predicting the bankruptcy of construction companies: A CART-based model [online]. *Engineering Economics*, 28: 145–154. <https://doi.org/10.5755/j01.ee.28.2.16353>.
KHASHEI, M., REZVAN, M. T., HAMADANI, A. Z., BIJARI, M. (2013). A bi-level neural-based fuzzy classification approach for credit scoring problemss [online]. *Complexity*, 18: 46–57. <https://doi.org/10.1002/cplx.21458>.
KIM, M. J., KANG, D. K., KIM, H. B. (2015). Geometric mean based boosting algorithm with over-sampling to resolve data imbalance problem for bankruptcy prediction [online]. *Expert Systems with Applications,* 42(3): 1074–1082. <https://doi.org/10.1016/j.eswa.2014.08.025>.
KOU, G., PENG, Y., WANG, G. (2014). Evaluation of clustering algorithms for financial risk analysis using MCDM methods [online]. *Information Sciences,* 275: 1–12. <https://doi.org/10.1016/j.ins.2014.02.137>.
KRULICKÝ, T., KALINOVÁ, E., KUČERA, J. (2020). Machine learning prediction of USA export to PRC in context of mutual sanction [online]. *Littera Scripta*, 13(1): 83–101. <https://doi.org/10.36708/Littera_Scripta2020/1/6>.

LI, Z., CROOK, J., ANDREEVA, G. (2017). Dynamic prediction of financial distress using Malmquist DEA [online]. *Expert Systems with Applications*, 80: 94–106. <https://doi.org/10.1016/j.eswa.2017.03.017>.

LIN, C., YANG, T. (2012). Validation the Role of Previous Information in Predicting TSE Corporation Bankruptcy. *Procedia – Social and Behavioral Sciences*, 57: 560–565.

MIN, J., LEE, Y. (2005). Bankruptcy Prediction Using Support Vector Machine with Optimal Choice of Kernel Function Parameters [online]. *Expert Systems with Applications,* 28(4): 603–614. <https://doi.org/10.1016/j.eswa.2004.12.008>.

ODOM, M. D., SHARDA, R. (1990). A neural network model for bankruptcy prediction. *IJCNN International Joint Conference on Neural Networks,* 2: 163–168.

OHLSON, J. A. (1980), Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research,* 18(1): 109–131.

RAINARLI, E., SABEK, A. (2023). Mitigating challenges: Handling mis sing values and imbalanced data in bankruptcy prediction using machine learning [online]. *Littera Scripta, Economics, Corporate finance, Finance and Valuation*, 16(2): 79–96. <https://doi.org/10.36708/Littera_Scripta2023/2/6>.

SABEK, A. (2023). Unveiling the diverse efficacy of artificial neural networks and logistic regression: A comparative analysis in predicting financial distress [online]. *Croatian Review of Economic, Business and Social Statistics (CREBSS)*, 9(1): 16–32. <http://doi.org/10.2478/crebss-2023-0002>.

SABEK, A. (2024). A Comparative Analytical Evaluation of CA Score, Kida, Springate Models for Financial Distress Prediction: The Case of Algeria. *Revue Finance & marchés,* 11(1): 38–50.

SABEK, A., HORAK, J. (2023). Gaussian Process Regression´s Hyperparameters Optimization to Predict Financial Distress [online]. *Retos, Revista de Ciencias Administracion y Economia*, 13(26): 273–289. <http://doi.org/10.17163/ret.n26.2023.06>.

SABEK, A., SAIHI, Y. (2023). Crunching Numbers, Making Decisions: Artificial Intelligence and Statistics for Financial Distress Forecasting in Algeria and Saudi Arabia [online]. *CAFI, Comptabilité, Actuariat, Finance & Information*, 6(2): 183–201. <https://doi.org/10.23925/cafi.62.60718>.

SAFI, SA-D, CASTILLO, P.A., FARIS, H. (2022). Cost-Sensitive Metaheuristic Optimization-Based Neural Network with Ensemble Learning for Financial Distress Prediction [online]. *Applied Sciences*, 12(14): 1–32. <https://doi.org/10.3390/app12146918>.

SALEHI, M., POUR, M. D. (2016). Bankruptcy prediction of listed companies on the Tehran Stock exchange [online]. *International Journal of Law and Management*, 58: 545–561. <https://doi.org/10.1108/IJLMA-05-2015-0023>.

SANTOS, F., VENTURA, R., SOUSA, J. M. C., VIEIRA, M. S. (2022). First-Order Autonomous Learning Multi-Model Systems for Multiclass Classification tasks. *2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE),* Padua, Italy, 1–6.

SERRANO-CINCA, C. (1996). Self organizing neural networks for financial diagnosis [online]. *Decision Suport Systems,* 17(3): 227–238. <https://doi.org/10.1016/0167-9236(95)00033-X>.

SERRANO-CINCA, C., GUTIÉRREZ-NIETO, B. (2013). Partial Least Square Discriminant Analysis for bankruptcy prediction [online]. *Decision Support Systems*, 54(3): 1245–1255. <https://doi.org/10.1016/j.dss.2012.11.015>.

SMARANDA, C. (2014). Scoring functions and bankruptcy prediction models – case study for Romanian companies [online]. *Procedia Economics and Finance*, 10: 217–226. <https://doi.org/10.1016/S2212-5671(14)00296-2>.

SOARES, E., ANGELOV, P., GU, X. (2020). Autonomous Learning Multiple-Model zero-order classifier for heart sound classification [online]. *Applied Soft Computing,* Vol 94. <https://doi.org/10.1016/j.asoc.2020.106449>.

SREBRO, B, MAVRENSKI, B, BOGOJEVIĆ, V., KNEŽEVIĆ, S., MILAŠINOVIĆ, M., TRAVICA, J. (2021). Bankruptcy Risk Prediction in Ensuring the Sustainable Operation of Agriculture Companies [online]. *Sustainability,* 13(14): 1–17. <https://doi.org/10.3390/su13147712>.

ŠTEFKO R, HORVÁTHOVÁ J., MOKRIŠOVÁ M. (2020). Bankruptcy Prediction with the Use of Data Envelopment Analysis: An Empirical Study of Slovak Businesses [online]. *Journal of Risk and Financial Management*, 13(9): 1–15. <https://doi.org/10.3390/jrfm13090212>.

TANG, Y., J. JI, J., ZHU, Y., GAO, S., TANG, Z., TODO, Y. (2019). A Differential Evolution-Oriented Pruning Neural Network Model for Bankruptcy Prediction [online]. *Complexity,* pp. 1–21. <https://doi.org/10.1155/2019/8682124>.

TRACZYNSKI, J. (2017). Firm Default Prediction: A Bayesian Model-Averaging Approach [online]. *Journal of Financial and Quantitative Analysis*, 52(3): 1211–1245. <https://doi:10.1017/S002210901700031X>.

TSAI, C. F., HSU, Y. F., YEN, D. C. (2014). A comparative study of classifier ensembles for bankruptcy prediction [online]. *Applied Soft Computing,* 24: 977–984. <https://doi.org/10.1016/j.asoc.2014.08.047>.

VALASKOVA, K., GAJDOSIKOVA, D., BELAS, J. (2023). Bankruptcy prediction in the postpandemic period: A case study of Visegrad Group countries [online]. *Oeconomia Copernicana*, 14(1): 253–293. <https://doi.org/10.24136/oc.2023.007>.

WILSON, R., SHARDA, R. (1994). Bankruptcy prediction using neural networks [online]. *Decision Support Systems*, 11(5): 545–557. <https://doi.org/10.1016/0167-9236(94)90024-8>.

WOLD, S., ESBENSEN, K., GELADI, P. (1987). Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1–3): 37–52.

WU, C., WANG, X.M. (2000). A neural network approach for analyzing small business lending decisions [online]. *Review of Quantitative Finance and Accounting,* 15: 259–276. <https://doi.org/10.1023/A:1008324023422>.

YU, Q. (2013). *Machine Learning for Corporate Bankruptcy Prediction*. PhD Thesis, School of Science, Aalto University, Department of Information and Computer Science, Finland.

ZHANG, Y., LIU, R., HEIDARI, A. A., WANG, X., CHEN, Y., WANG, M., CHEN, H. (2021). Towards augmented kernel extreme learning models for bankruptcy prediction: Algorithmic behavior and comprehensive analysis [online]. *Neurocomputing*, 430: 185–212. <https://doi.org/10.1016/j.neucom.2020.10.038>.

ZHOU, L., LAI, K. K., YEND, J. (2012). Empirical models based on features ranking techniques for corporate financial distress prediction [online]. *Computers and Mathematics with Applications*, 64(8): 2484–2496. <https://doi.org/10.1016/j.camwa.2012.06.003>.

ZIEBA, M., TOMCZAK, S. K., TOMCZAK, J. M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction [online]. *Expert Systems with Applications*, 58: 93–101. <https://doi.org/10.1016/j.eswa.2016.04.001>.

ZVARÍKOVÁ, K., MACHOVÁ, V., NICA, E. (2022). Cognitive Artificial Intelligence Algorithms, Movement and Behavior Tracking Tools, and Customer Identification Technology in the Metaverse Commerce. *Review of Contemporary Philosophy,* 21: 171–187.