





Article

Forecasting Financial Investment Firms' Insolvencies Empowered with Enhanced Predictive Modeling

Ahmed Amer Abdul-Kareem ^{1,*} , Zaki T. Fayed ¹, Sherine Rady ¹ , Salsabil Amin El-Regaily ¹ 
and Bashar M. Nema ² 

¹ Faculty of Computer and Information Sciences, Ain Shams University, Cairo 11566, Egypt

² Department of Computer Science, College of Science, Mustansiriyah University, Baghdad 10001, Iraq

* Correspondence: ahmed.amer.abdulkareem@gmail.com

Abstract: In the realm of financial decision-making, it is crucial to consider multiple factors, among which lies the pivotal concern of a firm's potential insolvency. Numerous insolvency prediction models utilize machine learning techniques try to solve this critical aspect. This paper aims to assess the financial performance of financial investment firms listed on the Iraq Stock Exchange (ISX) from 2012 to 2022. A Multi-Layer Perceptron predicting model with a parameter optimizer is proposed integrating an additional feature selection process. For this latter process, three methods are proposed and compared: Principal Component Analysis, correlation coefficient, and Particle Swarm Optimization. Through the fusion of financial ratios with machine learning, our model exhibits improved forecast accuracy and timeliness in predicting firms' insolvency. The highest accuracy model is the integrated MLP + PCA model, at 98.7%. The other models, MLP + PSO and MLP + CC, also exhibit strong performance, with 0.3% and 1.1% less accuracy, respectively, compared to the first model, indicating that the first model serves as a powerful predictive approach.

Keywords: insolvency; MLP; feature selection; financial ratio



Citation: Abdul-Kareem, Ahmed Amer, Zaki T. Fayed, Sherine Rady, Salsabil Amin El-Regaily, and Bashar M. Nema. 2024. Forecasting Financial Investment Firms' Insolvencies Empowered with Enhanced Predictive Modeling. *Journal of Risk and Financial Management* 17: 424. <https://doi.org/10.3390/jrfm17090424>

Academic Editor: Ștefan Cristian Gherghina

Received: 15 August 2024
Revised: 18 September 2024
Accepted: 20 September 2024
Published: 22 September 2024



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

1. Introduction

Insolvency forecasting (IF) seeks to determine the chances of a company becoming insolvent and therefore helps in evaluating its financial strength. IF includes models such as cash flow forecasts in which potential cash deficits can be detected and controlled, thus easing financial failure risks. Financial insolvency is associated with the terms failure, distress, bankruptcy, and default. This type of insolvency is regarded as an expensive event that affects a firm's ideal capital structure and leads to acts adverse to debt holders and stakeholders. However, financial insolvency can also improve corporate performance by encouraging changes and forcing managers to make riskier decisions (Almeida 2023).

Insolvency occurs when an entity's financial liabilities exceed its assets, resulting in an inability to pay debts as they become due. Therefore, insolvency is a critical financial condition that indicates a fundamental imbalance between what an entity owes and what it owns. Insolvency can manifest in different forms, such as cash-flow insolvency, where an entity lacks the necessary liquidity to meet its immediate payment obligations, or balance-sheet insolvency, where its total liabilities exceed the total value of its assets. Insolvency can have severe consequences, including legal actions by creditors, asset liquidation, debt restructuring negotiations, or the initiation of bankruptcy proceedings to address financial distress and facilitate a resolution of outstanding debts (Casey and Macey 2023).

Several insolvency risk statistical prediction schemes have been developed that have received significant attention from scholars and practitioners, such as the Z-score model (Pratiwi et al. 2023), the Zeta-score model (Vo 2023), and the logistic regression model (Gavurova et al. 2022). However, in the present complex business environment, managers have to determine the prospects of insolvency and act to prevent insolvency to ensure

solvency to avert probable business failure. This problem can also be solved by the new opportunities presented by machine learning and artificial intelligence for companies with large sets of data, as these approaches allow for analyzing the data and finding patterns and trends.

In a nutshell, IF mathematical models and AI-based approaches have some merits and demerits analogous to those of statistical models. These models are fed with fewer numbers of parameters, and they are capable of producing closer similarities in different datasets. Nevertheless, these models do not attempt to estimate the data structures, even though they are complicated. IF models developed on the use of AI are more complex in terms of learning capabilities; however, compared to the other models, they require more parameters, take longer to complete, and may lead to lower generality (Bassiouni et al. 2022).

The importance of this study is to address the gaps in several studies, including (Sharma et al. 2018), regarding some basic components, such as data pre-processing, parity of database distribution, and the effects of these in prediction, as depicted in Section 2. The disregard for these elements in prior research has resulted in the formulation of biased prediction models. To bridge these research gaps, we concentrate on feature selection methods to construct a precise model for early-stage financial insolvency prediction using genuine financial data from Iraqi enterprises. Our methodology also tackles data readiness, preprocessing, and data imbalance challenges.

This article introduces the enhanced forecasting model using Multi-Layer Perceptron (MLP). On the application and domain level, MLPs provide predictive power exhibiting strong forecasting skills in different financial forecasting tasks, such as stock price prediction, risk assessment, and portfolio optimization (Nayak et al. 2022). Nevertheless, understanding the constraints of these models is essential to ensure that predictions are not solely reliant on the most current techniques but are rooted in the most appropriate approach for the specific problem. With various types of ANNs showing promise in addressing financial forecasting challenges, this study concentrates on identifying the most effective models for forecasting insolvencies to attain optimal outcomes by using the hyperparameter tuning method. The findings indicate that feature selection proves significantly advantageous for such tasks, resulting in reduced complexity relative to alternative models. Principal Component Analysis, Particle Swarm Optimization, and correlation coefficients are utilized for the feature selection step. The Synthetic Minority Oversampling Technique (SMOTE) assesses various parameter configurations and addresses imbalanced datasets. A case study compares techniques using a unique real dataset. The study critically evaluates the procedures and variations in outcomes. The combination of these advanced machine learning algorithms aims to improve the precision and effectiveness of financial distress prediction, allowing for early interventions and strategic decision-making.

The subsequent parts of the paper are organized in the following manner: The related literature is examined in Section 2. Section 3 introduces the proposed enhanced forecast insolvency model. The experimental study is thoroughly explained in Section 4. Section 5 delves into the results and discussion employed. The conclusion is presented in Section 6. Limitations and directions for future research are detailed in Section 7.

2. Related Works

Financial distress prediction is crucial in risk management, investment decisions, and corporate governance. This review provides a comprehensive overview of approaches, strategies, and variables used in predicting financial distress, including the importance of financial ratios. The evaluation examines empirical data on the effectiveness of these predictors across different industries and economic conditions.

Karas and Režňáková (2020) revealed that operating cash-flow ratios with short-term debts significantly influence financial distress in SMEs and are more effective than traditional profit-based ratios. A hybrid modeling approach, including classification and regression trees, logistic regression, and PCA, was used to analyze the financial data of 4350 SMEs. The third model demonstrated the highest AUC of 0.93 in the learning sample.

Advanced solutions, including oversampling techniques, cost-sensitive approaches (CBoost algorithm), and ensemble-based models (XGBoost algorithm) for bankruptcy prediction, were illustrated in (Le 2021). The study in (Budianita et al. 2022) was concerned with the application of the C4.5 algorithm and Adaboost to predict financial distress companies. These algorithms were applied to five financial ratios from Altman's model to calculate the current ratio, total working capital to total assets, return on equity, return on assets, market value added, and asset turnover. This research examined 755 annual financial reports of firms listed on the Indonesia Stock Exchange during the period 2016–2019. The study's results showed that the integration of the C4 model with the Research and Technology University allowed the achievement of the objectives discussed in this paper. 5 and AdaBoost slightly outperformed this model in terms of accuracy in predicting financial distress at a threshold of $K = 10$ with an accuracy of 86. A total of 49% of participants accepted the completion of the survey questionnaires, the result achieved was that the C4.5 accuracy at $K = 8$ was 85.12%.

Performing financial distress prediction with a particular focus on imbalanced datasets was the main concern of (Safi et al. 2022); the authors proposed a metaheuristic optimization-based artificial neural network (MHOANN) with a cost-sensitivity fitness function. For fine-tuning the weights and biases of the artificial neural network (ANN) in the proposed work, PSO and CSO techniques were used. The evaluation of the method involved three datasets: all the variants of Spanish, Taiwanese and Polish. The level of accuracy obtained between the actual and prognosticated results differed across the datasets, thus the Spanish dataset yielded an accuracy ranging from 0.749 to 0.38 to 1.882, the UC dataset ranging from 0.789 to 0.888 and below, and the Polish dataset ranging from 0.790 to 0.848.

The authors (Yan et al. 2020) adopted Chinese manufacturing firms to model, analyze, and forecast earning management and financial distress. The authors examined the levels of association between financial ratios, macroeconomic factors, and the probability of financial failure. They employed financial performance indicators, such as net profit and net cash flow, in order to evaluate the effects of financial distress risk. It comprised a distributed Lagrange structure, logistic regression, and support vector machines (SVMs). In the process of feature selection, a Lasso penalty was applied. The lasso–logistic–distributed Lagrange (LLDL) model was the most effective, with an accuracy AUC of more than 95%, followed by the lasso–SVM–distributed Lagrange (LSVMDL) model, which had the greater G-mean and KS statistics. The two models also induced high levels of accuracy in predicting accounting ratios and macroeconomic variables.

The machine learning techniques used in (Shetty et al. 2022) for predicting bankruptcy were XGBoost, SVM, and deep learning neural networks. The following methods were used to predict bankruptcies with accounting data from the Belgian SMEs of the period 2002–2012 with accuracies of 82–83%.

The problems that arise from trying to forecast stock price movements using only historical data and from reiterating the need to account for factors such as news articles and social media activity were presented in (Maqbool et al. 2023). This work presented an intelligent system that involved Sentiment analysis programs (Valance Aware Dictionary and Sentiment Reasoner (VADER), TextBlob, Flair) and past data of stocks to predict the trends of the share market. The MLP Regressor model was used in forecasting the stock prices, where three kinds of sentiment scores were evaluated in their ability to predict the shifts in the tendencies of stock prices while testing the effectiveness of each sentiment analyzer on its own and along with other analyzers. Models with FLAIR (setting F), VADER joined with TextBlob (setting V + T), and F combined with T received the highest scores in trend predictions for the next day, which constituted 0.90. Moreover, the chosen single Flair sentiment analysis reached a high level of accuracy in the further prediction of the trend during the next 100 days (0.75).

The studies of Omari et al. (Al Omari et al. 2023) applied the ANN technique to evaluate insurance companies in the context of the Jordanian insurance industry. For this purpose, the researchers applied the multilayer perceptron (MLP) model with such input

predictors as subrogation, claims paid, market capitalization, and total shareholder equity. The refocusing on internal and external performance measurement involved analyses of the output variable, also known as total asset turnover (TAT). The findings obtained for this study revealed the fact that raising the number of iterations from 500 to 10,000 improved the ACC rate, precision rate, and F-score for the training dataset and the testing dataset. According to the 80% training sample, the ACC rate increased from 73%. From 7% at $n = 500$, the rate rose to 87%. From 8% at $n = 10,000$, the precision rate was raised from 56.8% to 79.63% and then decreased to 64%, while the F1-score increased from 57.9% to 70%. The same enhancements were also evidenced in the 20% test sample. Such findings asserted the flexibility improvement in the MLP model used in evaluating insurance companies' financial performance by exclusively using the total asset turnover (TAT) criterion.

The GALSTM-FDP model combined GA and LSTM, with LSTM standing for Financial Distress Prediction. This model made predictions with better accuracy compared to conventional models, such as Support Vector Machines (SVMs) and Decision Trees (DTs). The model was based on the data from 1953 firms of the MENA region and had a pureness of 50% to 91%. This work was a fusion of two methods, LSTM, which was exceptionally good at learning structures in financial time-series data, and GA, which was a form of optimization that mimicked the natural evolution process (Al Ali et al. 2023).

From the preceding information, it can be noted that little emphasis has been placed on cases addressing unbalanced datasets and other issues of preprocessing, which are arguably some of the most key factors that are capable of impacting the forecast performance model. Such a gap in the literature encourages more research and investigation in these directions. Our research contributes to filling this gap by investigating various feature selection approaches, addressing data preparation obstacles and preprocessing issues, and efficiently controlling data imbalance.

3. Materials and Methods

In this study, we integrate MLP and FS techniques to construct our Proposed Enhanced Model for Forecasting Insolvencies. The research methodology is shown in Figure 1.

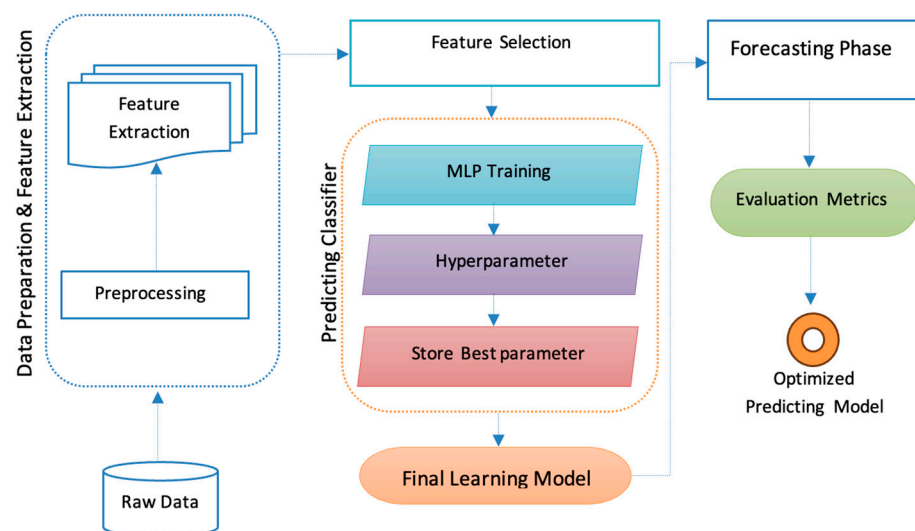


Figure 1. Proposed Enhanced Forecast Insolvency Model.

3.1. Data Preparation and Preprocessing

The datasets undergo cleaning and preprocessing to remove outliers and handle missing values to improve their quality using forward-filling by bringing forward the most recent observed value to fill in the gaps. The goal of this phase is to obtain purified data ready to be used in the next step.

To address the imbalanced issue in the dataset, the synthetic minority oversampling technique (SMOTE) is applied. The benefits of using the SMOTE include helping to balance the number of samples of different classes, which can, in turn, assist the classifier, especially in those that are affected by the class distribution. While random oversampling is somewhat similar to the duplication of records of the minority class, the SMOTE generates artificial records. This diversity is helpful in order to avoid overfitting the minority class during the model training process. The SMOTE's synthesis of new samples that have proximity to the existing minority class samples means that the classifier will be trained on a better and more discriminative decision surface, thereby enhancing generality. In the SMOTE, the parameter k , which controls the number of nearest neighbors used, is tunable, and it can be increased or decreased to make the generated synthetic samples closer or farther from other samples in the dataset, depending on the requirements. The SMOTE can be integrated with different types of ML models, such as decision trees, support vector machines, and neural networks; therefore, it is efficient in dealing with imbalanced datasets. Nevertheless, it is recommended to assess the accuracy of a model implemented after applying the SMOTE because, in some cases, it may not bring an improvement in the results according to the nature and characteristics of the data used. This technique works step by step as follows:

Step 1. Identify the Minority Class: SMOTE starts by identifying which class in the dataset is the minority class (the one with fewer samples).

Step 2. Choose a Sample: Randomly select an instance from the minority class.

Step 3. Find the Nearest Neighbors:

- For the selected instance, compute the distance to all other instances in the minority class. This is commonly done using a distance metric such as the Euclidean distance.
- Identify the k nearest neighbors of an instance from the minority class.

Step 4. Generate Synthetic Examples: For each of the k selected neighbors (X_n), create new synthetic instances:

- Randomly select one of the k nearest neighbors.
- Use the following formula to create a new synthetic instance X_{new} :

$$X_{new} = X_i + (I - X_i) \times \delta \quad (1)$$

where X_i is an element in K_{xi} : $X_i \in X_{min}$.

This finding suggests that the new instance X_{new} is created along the line segment between the original instance X_i and its selected neighbor X_n .

Step 5. Repeat: Repeat the process until the desired number of synthetic instances is created for the minority class, effectively balancing the dataset.

3.2. Feature Extraction

The subsequent phase involves feature extraction. Extracted data are used to calculate 24 financial ratios that make up the features of the predicting model. The calculations of such financial ratios are shown in Table 1.

Table 1. Financial Ratios Formula.

Group	Financial Ratios	Calculation Formula	Ref.
Liquidity	Current Ratio (CR)	Total Current Assets/Total Liabilities	(Radovanovic and Haas 2023)
	Instant Liquidity Ratio (ILR)	Cash/Total Current Liabilities	(Zhao et al. 2024)
	Net Working Capital (NWC)	Current Assets–Current Liabilities	(Zhu et al. 2021)
	Equity Ratio (ER)	Equity/Total Liabilities	(González and Keddad 2024)
	Debt Ratio (DR)	Total Liabilities/Equity	(Chakri et al. 2023)
Activity	Accounts Receivable Turnover (ART)	Net Sales/Accounts Receivable	(Wang et al. 2023)
	Fixed Asset Turnover (FAT)	Net Sales/Net Fixed Assets	(Lee 2023)
Profitability	Net Profit (NP)	Net Profit/Sales	(Zhang et al. 2022)
	Return on Equity (ROE)	Net Profit/Equity	(Salah 2020)
	Return on Capital (ROC)	Net Operating Profit/Working Capital	(Anton and Nucu 2020)
	Return on Assets (ROA)	Net Profit/Total Assets	(Mattos and Shasha 2024)
	Return on Investment (ROI)	Net Profit/Total Assets	(Xu et al. 2022)
	Proportion of a Corporation’s Assets (PA)	Cash/Total Assets	(Lohmann et al. 2022)
	Current Assets to Total Assets (CATA)	Current Assets/Total Assets	(Kim 2011)
	Total Current Liabilities (TCL)	Working Capital/Total Assets	(Wang et al. 2024)
	Working Capital Turnover (WCT)	Working Capital/Net Sales	(Kušter et al. 2023)
	Current Assets (NCA)	Net Profit/Current Assets	(Wu et al. 2022)
	Sales to Operating Profit (SOP)	Net Operating Profit/Net Sales	(Sang 2020)
	Asset Turnover (AT)	Net Sales/Total Assets	(Lin and Bai 2022)
	Sale to Current Assets (SCA)	Net Sales/Current Assets	(Kušter et al. 2023)
	Days’ Sale Uncollected (DSU)	$(365 \times \text{Accounts Receivable})/\text{Net Sales}$	(Jadiyappa and Shette 2024)
	Accounts Receivable (AR)	Accounts Receivable/Total Liabilities	(Tarjo et al. 2023)
Operating Return on Assets (OROA)	Net Operating Profit/Total Assets	(Lee 2023)	
Sale to Current Assets (SCA)	Current Assets/Net Sales	(Laitinen et al. 2023)	

These ratios, when applied, assist investors in analyzing the profitability of the investment firm and comparing it with other firms in the industry to identify potential risks or opportunities.

Liquidity ratios are important in finance since they point out the ability of the company to meet its current liabilities. Some key reasons why liquidity ratios are important are as follows. Liquidity shows the capacity to meet current obligations, which are important for the company’s operations. Liquidity ratios, therefore, help investors and creditors evaluate the risk of investing in or financing the business. A low liquidity ratio may cause problems in the company’s financial status; the firm may be unable to pay its debts. It may show how effectively a firm has been able to manage existing stocks of inventory, the amount of account receivables, and the accounts that are payable. A company that has a favorable liquidity status can participate in different opportunities that come up without having to look for financing. Lenders use liquidity ratios in an attempt to determine a company’s capability of repaying loans. Comparatively higher liquidity ratios may help in obtaining better terms in borrowings and hence lower interest rates. Liquidity ratios help an investor

and analyst of a particular firm evaluate how the firm is doing in comparison to other firms and its relevant industry (Zhao et al. 2024).

Turnover ratios, also referred to as activity ratios, are relevant in financial investment as they determine the efficiency of a company in utilizing its assets to generate revenues or cash flows. These ratios assist the investors in gaining insight into how efficient a company is in its operations. A higher turnover rate is often perceived to be an indicator of the good performance and efficient utilization of assets, improving the overall economic value of a business and hence the appeal to investors (Wang et al. 2023).

Profitability ratios are key financial indicators that are essential in everyday usage by investors and analysts while assessing the overall performance of a business in generating the net income relative to the total amount of sales, the total amounts of tangible and intangible assets, and the shareholder funds, respectively. These indicators can influence it in different ways. 1. one of these ways is performance appraisal. Therefore, profitability ratios, such as net profit margin, return on assets (ROA), and return on equity (ROE), help the investor to determine whether or not the firm is earning the best profit based on its revenues and investments. 2. From the profitability ratios, the investors who make investment decisions are in a position to compare firms that are in the same industries. Better forecasting is achieved through ratios, which may imply that higher ratios indicate that the firm is in a better financial position regarding operational efficiency, thereby facilitating investment decisions. 3. Based on investor attraction, such a firm may have very high levels of profitability, and there may be the possibility that such firms can easily expand in the future. Conversely, low ratios indicate inefficiency in operations, which may, in one way or the other, jeopardize the future growth of the company. 4. Regarding financial solvency, it is possible for investors to evaluate financial solvency by the use of ratios including ROE and ROA. This is especially because firms that have proved to be profitable over a certain period of time are usually less risky, particularly when it comes to investment during an economic crisis (Kim 2011).

3.3. Feature Selection Techniques

The selection of subset features plays a crucial role in optimizing the performance of a model. By avoiding the curse of dimensionality, this processing step simplifies the model, making it more interpretable, timely, and efficient for prediction, thereby mitigating the risk of overfitting. In this phase, we compare the impacts of three of the most distinct feature selection methods, namely PCA, the correlation coefficient, and PSO, on the performance of the proposed model in terms of the subset generated and the time of training.

3.3.1. Principal Component Analysis (PCA)

PCA is another method in data preprocessing and is a dimensionality reduction technique that can also be employed in carrying out feature selection by determining the best features within a dataset. The basic steps of this method for feature selection are as follows:

- Step 1: Standardize the Data.
- Step 2: Calculate the Covariance Matrix.
- Step 3: Calculate the Eigenvalues and Eigenvectors.
- Step 4: Sort Eigenvalues and Eigenvectors.
- Step 5: Select Principal Components.
- Step 6: Transform the Original Data.
- Step 7: Analyze the Principal Components.
- Step 8: Select the Features.

Therefore, by following the steps mentioned above, PCA can be applied to feature selection and dimensionality reduction for several problems, particularly for high-dimensional problems.

3.3.2. Correlation Coefficient

The correlation coefficient is one of the statistical measures used to quantify feature space, and it is normally used in feature selection. With the aid of CC, one is able to establish the magnitude and the directions of linear relationship between two variables. Its values fall between -1 and 1 .

$$CC = \begin{cases} 1 \Rightarrow a \text{ perfect positive linear relationship.} \\ -1 \Rightarrow a \text{ perfect negative linear relationship.} \\ 0 \Rightarrow no \text{ linear relationship.} \end{cases}$$

3.3.3. Practical Swarm Optimization-Based Feature Selection (PSO-FS)

PSO is a genetic algorithm that originated from the population and was developed by Kennedy and Eberhart in 1997 (Kennedy and Eberhart 1997). PSO, on the one hand, employs an iterative and iterative process that is modeled on a flock of birds for the attainment of the best solution. The basic steps of the PSO algorithm present are as follows.

3.4. Multi-Layer Perceptron (MLP)

This phase involves training and optimizing the predicting classifier. The MLP model is proposed, which is one of AI's most popular and simple forms for NN, and it has broken through in the area of financial forecasting and several other fields for multiple reasons, such as its nonlinearity property. This property suits our case since the variables of financial data usually show complex nonlinear relationships between them. MLPs, being nonlinear function learners, can successfully model these relationships. A second advantage of MLP is its ability to provide time series analysis. MLPs can be tailored to deal with time series data and to take into account the temporal dependencies and trends occurring in the financial market (Yi et al. 2023). However, it is important to note that MLPs also have limitations, such as overfitting, especially with a smaller amount of data. Therefore, the application of cross-validation is essential to prevent model overfitting.

The MLP model consists of one input layer with 24 input neurons, one or multiple hidden layers, and one output layer. The number of nodes in the output layer should match the number of classes in our multi-label problem, so the output layer for our problem has 2 nodes. With the help of backpropagation algorithms, these multilayer networks are trained using the 80% training set via a supervised method where the error in the output layer is computed first. Then, the weights in the preceding layers are adjusted through a backward iterative process.

3.5. Evaluation Metrics

In the improved methodology of the proposed model, the evaluation phase plays a crucial role since it helps in the assessment of the performance and the quality of the built machine learning model. During this phase, the objective is to see how well the models in question are able to and, more particularly, how accurate they are in terms of the forecasts produced. To this end, during the initial splitting of the dataset, 20 percent of the samples are allocated for test purposes. This test is carried out to test the models on data not included in the model development process. By utilizing this particular set of test data, it is possible to achieve a fair assessment of all of the models in terms of their predictive accuracy and performance when the new data extend beyond the current training dataset.

As part of the evaluation of the developed models, the current work focuses on a binary classification task and uses confusion matrices in the process of assessing all the developed models. Based on the confusion matrix, we can compute different measures of the performance of a given model. The selected measures are accuracy, precision, recall, F1-score, Jaccard, timing, and AUC.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (5)$$

$$Jaccard = \frac{TP}{TP + FP + FN} \quad (6)$$

4. Experimental Study

4.1. Dataset Description

In this section, we deal with the original dataset. The financial investment firms chosen for this study are listed on the ISX. The sampling period is from 2012 to 2022, which is equal to 10 years. These firms are analyzed and distinguished based on reports from the auditing, financial, and taxation databases and the ISX. All the variables used in the sample are obtained from standard financial reports, including income statements and balance sheets. Thus, the usefulness of this study is not limited by the fact that only Iraqi firm data are employed.

Our empirical results include 1440 instances for six firms, with 240 attributes allocated to each firm over 10 years, forming the basis for the experiments conducted. The dataset is split into 80% for training purposes and 20% for testing purposes. The dataset constitutes two categories: financial insolvency (coded as 1) and non-insolvency (coded as 0), depending on whether the firm is marked as insolvent (D or G) or not (S). This categorization is based on Altman's technique, as seen in Table 2.

Table 2. Description of Our Dataset Categories Based on Altman's Independent Variables.

Classify	Frequency	Percent
1 (D or G)	38	32.1
0 (S)	18	67.9
Total	56	100

The feature set comprises 24 distinct financial ratios, with the indicators (liquidity, activity, profitability, and efficiency), listed in Table 1. Feature selection is crucial in the process of machine learning since it enables us to remove features that cause overfitting, increases model understanding, increases accuracy, and decreases learning time. Reducing the dimension of the input data assists in making models more ordered and faster at making predictors, especially when time is important in training. The process then uses oversampling as a technique to help balance classes before the conventional model training process begins. As a result of such replication, oversampling offsets the dominating scale and enforces algorithms to consider otherwise marginalized yet essential classes.

4.2. MLP Configurations

Various configurations of MLPs for predicting insolvency are investigated, involving optimization techniques, activation functions, neuron quantities, and layer quantities, each with different levels of tolerance, which are necessary to create a successful neural network model. Nonetheless, since finding the best MLP structure is a difficult undertaking that is dependent on the quality of the data and the complexity of the issue, it is crucial to tune the model parameters. The parameter search space is listed in Table 2, which represents the minimum number of design decisions that an automated system must make for a given learning problem. Our experiments for model tuning address each of the parameter choices in Table 3.

Table 3. Proposed Insolvency Prediction Model Search Space.

Optimization Model	Adam	sgd							
Hidden Layer Sizes	10	20	30	33	50	55	70	80	100
Activation Function	ReLU	Logistic	Tanh	identity					
Alpha	0.07	0.1	0.5	0.7	0.0001	0.001	0.01	0.3	0.2
Learning Rate	Constant	Adaptive	0.001						

To assess the performance of our proposed methods, we conduct the following four experiments:

- (A) Evaluating the classical MLP classifier without any preprocessing or feature selection optimization processes.
- (B) Evaluating the optimized model version using data preprocessing (DP) that addresses curing missing values and imbalanced datasets (MLP + DP).
- (C) Evaluating the optimized simplified model version using additional feature selection for reducing the model complexity, data volume, and execution time (MLP + DP + FS). The experiment compares 3 different techniques for FS, which are PCA, CC, and PSO.
- (D) Evaluating a finally optimized fine-tuned version of the predicting model in (C), by tuning and optimizing the classifier hyperparameters.

5. Results and Discussion

This section presents the experiments illustrated in the previous section. Moreover, this section compares and evaluates results obtained by conducting a set of extensive experiments using the five evaluation metrics outlined at the end of Section 4.

Table 4 shows the performance results of the classical MLP classifier without any optimization processes (Experiment A). The table shows that the classical MLP without optimization is average, with a 0.857% precision, 0.667% recall, 0.748% F1-score, 0.667% accuracy, and 0.778% AUC. The training time is expected to take an excessively long duration due to the lack of efficient optimization algorithms, such as gradient descent and backpropagation, leading to slow learning, convergence, and possibly overfitting. The model achieves a processing time of 3:4 (Minute:Second) and uses all feature space.

Table 4. Results of Metrics of Classical MLP Classifiers Without Optimization Processes.

Model	Precision	Recall	F1-Score	Accuracy	Jaccard	AUC
MLP	85.7	66.7	74.8	66.7	69.21	77.8

Experiment B studies the effect of data preprocessing (DP) to address missing values and imbalanced datasets. The preprocessing steps in this experiment ensure data consistency and tackle the imbalance issue through the SMOTE, as illustrated by Figure 2.

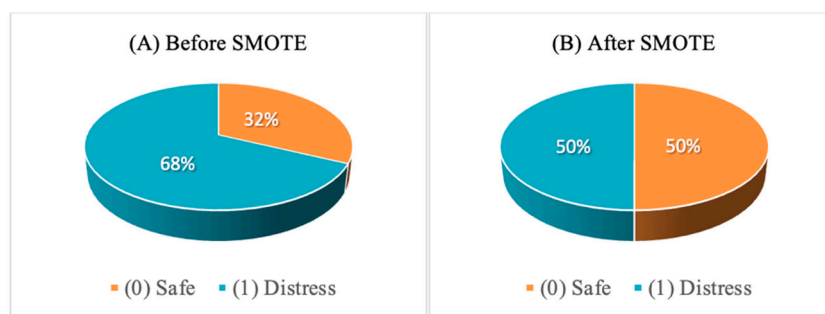


Figure 2. Application of SMOTE to an Imbalanced Dataset.

The Figure shows Pie charts comparing the numbers of instances labeled “Safe” and “Distress” before and after using a method called SMOTE. Before applying the SMOTE, 68% are “Safe” and 32% are “Distress”, while after applying the SMOTE, the percentages of “Safe” and “Distress” instances both become 50%. Table 5 shows the results of this experiment, in which the performance of MLPs is enhanced by 1.9% for precision, 9.8% for recall, 1.4% for F1-score, 0.9% for accuracy, and 2.9% for AUC.

Table 5. Results of Metrics of MLP + DP.

Model	Precision	Recall	F1-Score	Accuracy	Jaccard	AUC
MLP	87.6	67.5	76.2	67.6	70.39	80.7

Nonetheless, Experiment C carries out additional improvement experiments to achieve a more accurate and less complex model. In Experiment C, the effects of integrating FS processes are investigated. The 3 FS methods, Principal Component Analysis (PCA), Particle Swarm Optimization (PSO), and correlation coefficients are compared.

Integrating FS results in a decrease in the number of features by selecting relevant features, thereby reducing the model’s complexity and avoiding overfitting. Applying PCA has led to easier visualization and interpretation of data, reduced computational complexity and training time of machine learning models, and the removal of noise and irrelevant features (Figure 3).

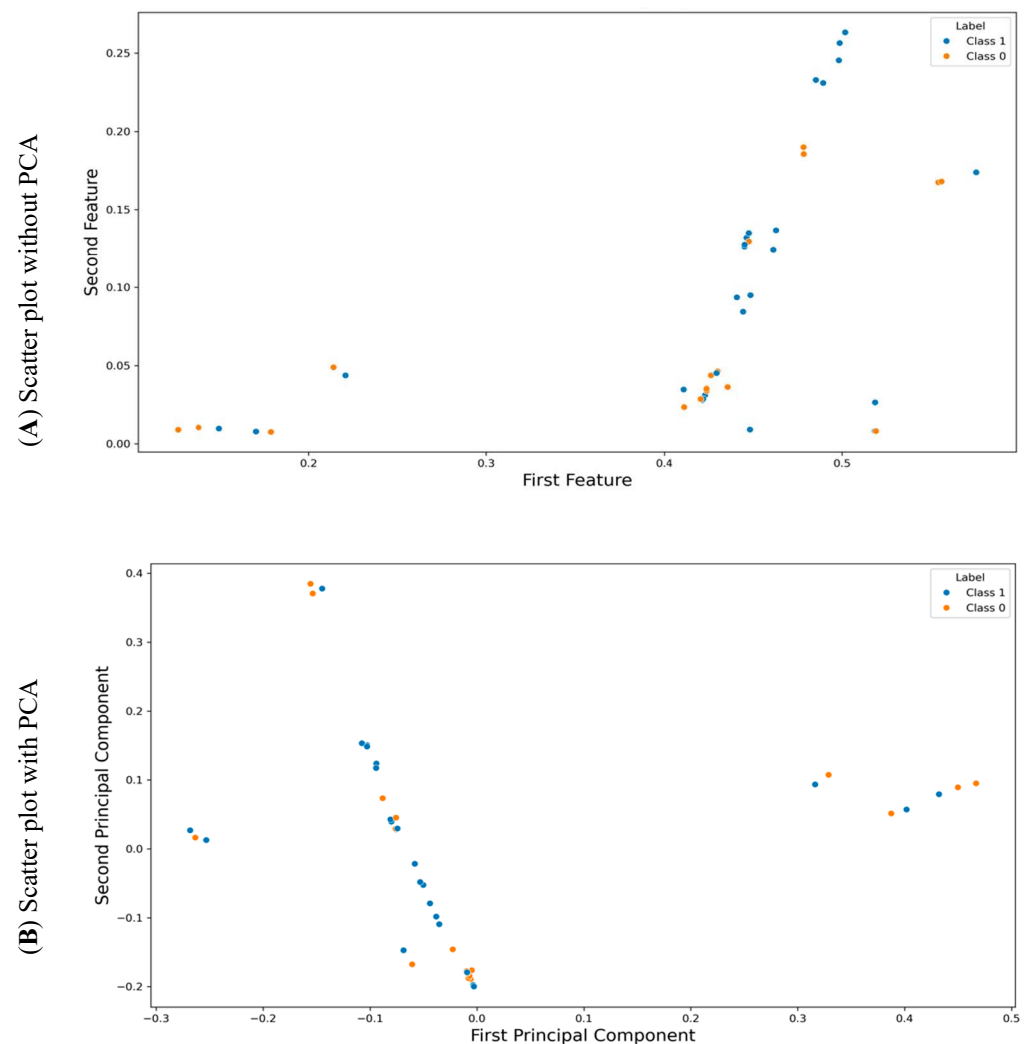


Figure 3. Scatter Plot Without/With PCA.

Determining the optimal number of PCs to retain dimensionality reduction is a crucial challenge in conducting PCA. In our data, the number of PCs is 6, as Figure 4 shows. To facilitate this decision-making process, it is helpful to evaluate an important parameter, which is the variance ratio. This ratio quantifies the proportion of variance explained by each PC. By visualizing this information on a scree plot, researchers can pinpoint the elbow point, as illustrated in Figure 5. From the most significant to the least significant principal component, the graphic shows the explained variance for each one. The plot’s “elbow” designates the beginning of the curve’s flattening. The number of major components that have to be kept for additional examination is indicated by this point. The “elbow” appears to consist of two or three parts in this figure. The cumulative explained variance, represented by the red line, indicates the percentage of the overall variance that can be attributed to the first n components.

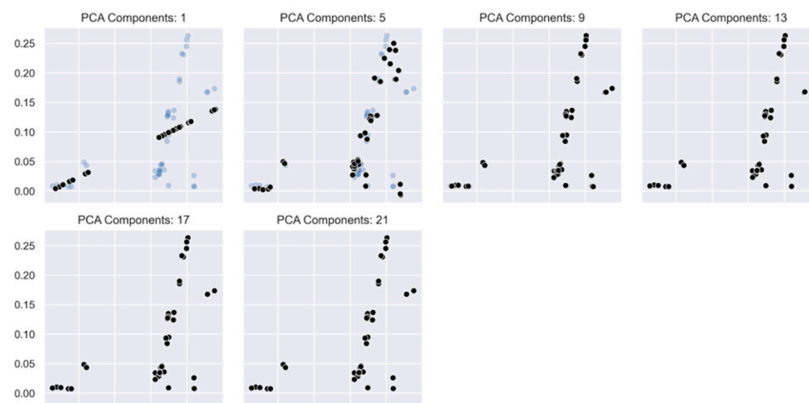


Figure 4. Optimal Number of PCs.

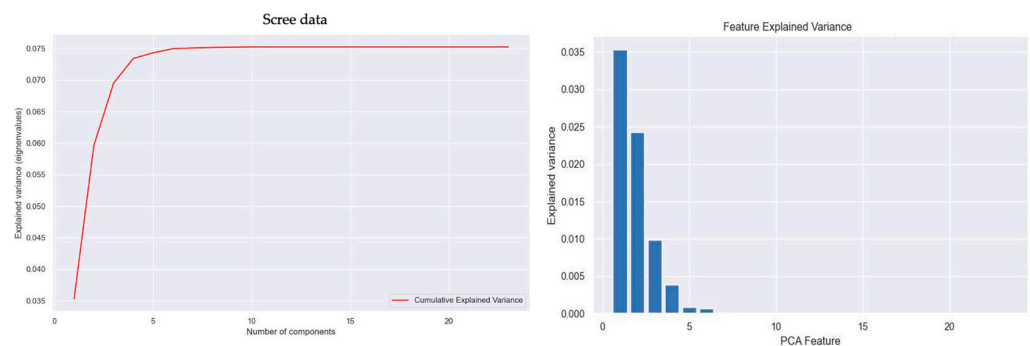


Figure 5. Scree Plot.

Figure 5 presents an explained variance of the principal components in PCA. X-Axis (Number of Components): On the horizontal axis, we have taken the number of the principal components from 1 to 20. Y-Axis (Explained Variance): The vertical axis illustrates the degree of explained variance or the eigenvalues of the principal components conveying how much of the variation in the dataset is retained by each of the principal components. Bar Chart: The blue bars are the explained variance for each of the principal components shown. The first few components (1 to 5) account for the majority of the data’s variance, unlike the remaining components. Cumulative Explained Variance: The first horizontal line in red indicates the Cumulative Explained Variance, which depicts the total variance that has been explained as more components are included. Initially, there is an upward trend, which demonstrates that the first few components contribute significantly towards the total variation. Elbow Point: The graph indicates an ‘elbow point’ around the first few components, where an increase in the number of components accounts for less and less variation in variance. Again, they indicate that beyond a certain number of components,

the united explained variance rises but only by a negligible margin thus implying that additional components are less influential in explaining variance.

This finding suggests that a small set, probably including only the first 5 components, accounts for most of the data variability. This result implies that it is possible to perform dimensionality reduction by only retaining such components, making the model simpler with the right amount of detail required.

An additional parameter to consider is the cumulative explained variance. This statistic sums the explained variance ratios to a certain PC. One can choose PCs that exceed a specific threshold, such as 80% or 90%.

Figure 6 is a line graph that shows the link between cumulative variance and the number of components. The goal of this graph is to determine the optimal number of components for capturing the majority of the data’s volatility. As shown in the figure, the curve at the first stage increases at a sharper angle, which means that the initial components account for a large extent of the variance. At some point, the last part of the curve starts making flat progress, signifying that subsequent components contribute to a less proportionate variation in the total variance. A dotted straight line drawn across the graph represents the 95% variability cumulative, which leads to extreme termination. The graph indicates that a certain amount of parts is needed to achieve this level, and the further addition of parts leads to a decrease in the contribution rate of variance. Finally, one’s experience in the field and the analytical aims impact the decision of which PCs to maintain. When considering the number of selected features and execution time, the MLP + PCA model achieves a processing time of 1.57:1569.80 and a feature space reduction of 10–20%.

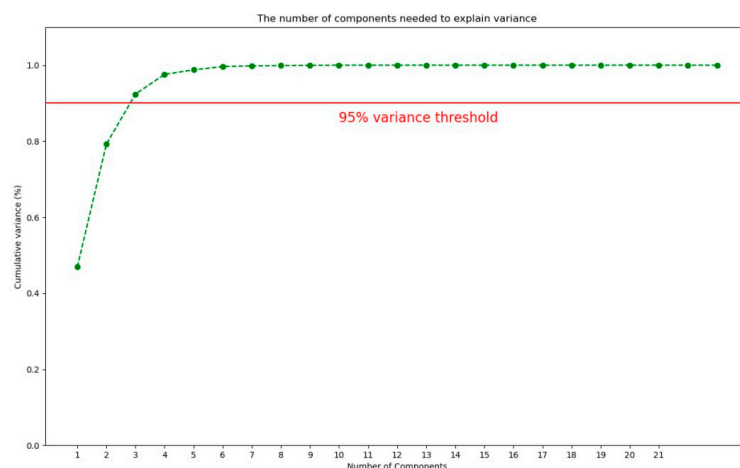


Figure 6. Number of Components Needed to Explain Variance.

The green line in the graph represents the cumulative explained variance as a function of the number of components. Each point on the line corresponds to the total variance explained by the components. As the number of components increased, the cumulative explained variance approached 1 (or 100%), indicating that more components accounted for more total variance in the dataset. The curve shows that the initial components contribute significantly to the explained variance, whereas additional components contribute less as the curve levels off. This visualization helps to understand how many components are needed to achieve the desired level of explained variance, such as the 95% threshold indicated by the red line.

The second FS employed is PSO. Global optimization contributes to the discovery of optimal solutions for intricate issues, refining parameter adjustments and hastening the convergence of machine learning algorithms. This process enables the tackling of nonlinear and multi-modal optimization challenges, thereby boosting model precision. Table 6 illustrates the 8 chosen features by PSO from the original set of 24:

Through this method, the features decrease from 24 to just 8, a reduction of 67%. When considering the number of selected features and execution time, the MLP + PSO model outperforms the MLP + PCA and MLP + CC models due to its shorter processing time of 1.05:1046.33 compared to 1.57:1569.80 and 1.16:1157.08, respectively. This difference results in a time discrepancy of 0.52:523.47 and 0.11:110.75, respectively, and a feature space reduction of 67%

The final FS employed is the correlation coefficient, which is a statistical metric indicating the strength of a linear relationship between two variables. The value ranges from -1 (inverse connection) to 1 (direct connection). A coefficient of 0 means no linear relationship, as shown in Figure 6.

Figure 7 depicts a correlation matrix obtained through the statistical analysis of financial data. The matrix depicts the correlation coefficients between several financial variables, which are represented by the rows and columns. Each cell in the matrix displays the correlation coefficient between the two variables assigned to that row and column. Positive values imply a positive correlation, whilst negative values suggest a negative correlation. The color intensity reflects the size of the association. The diagonal cells always carry the value 1 , which represents a variable's perfect correlation with itself. The color gradient illustrates the intensities of the relationships. Strong positive correlations, for example, are shown in deep blue, whereas strong negative correlations are depicted in dark red.

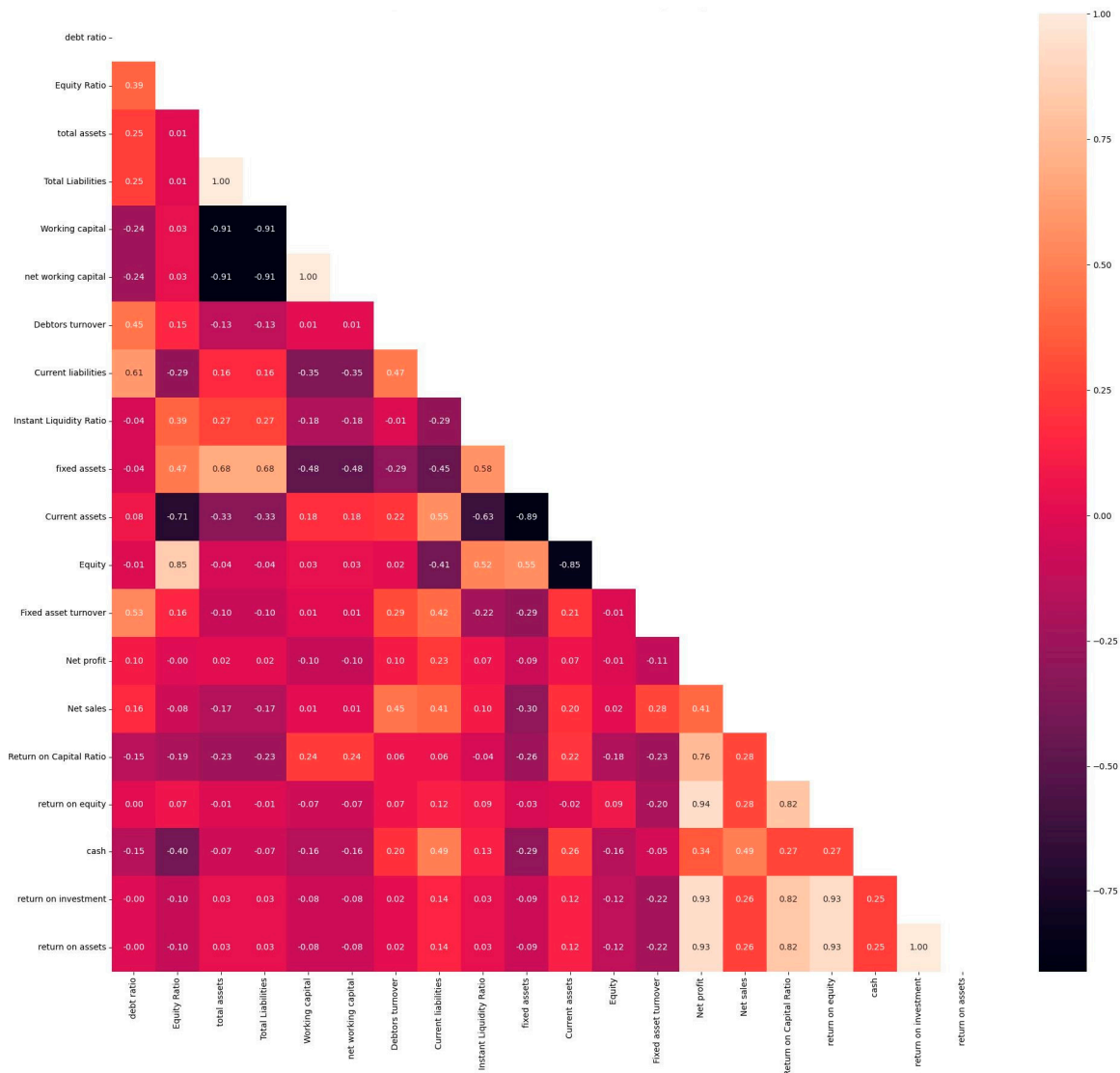


Figure 7. Correlation Heatmap of 10 Important Features With the Target.

Table 6. Selected Features Using PSO.

No. of Feature	Feature Index	Feature Name
1	0	Current Assets
2	8	Net Operating Profit
3	11	Working Capital
5	12	Current Ratio
6	13	Net Working Capital
7	14	Debtors’ Turnover
8	20	Instant Liquidity Ratio

Figure 8 illustrates the absolute values of the correlation coefficients between features and the predictive values of the top 10 features: current ratio, net working capital, working capital, cash, net profit, current liabilities, return on equity, return on assets, return on investment, and accounts receivable. When considering the number of selected features and execution time, the MLP + CC model achieves a processing time of 1.16:1157.08 and a feature space reduction of 58.3%.

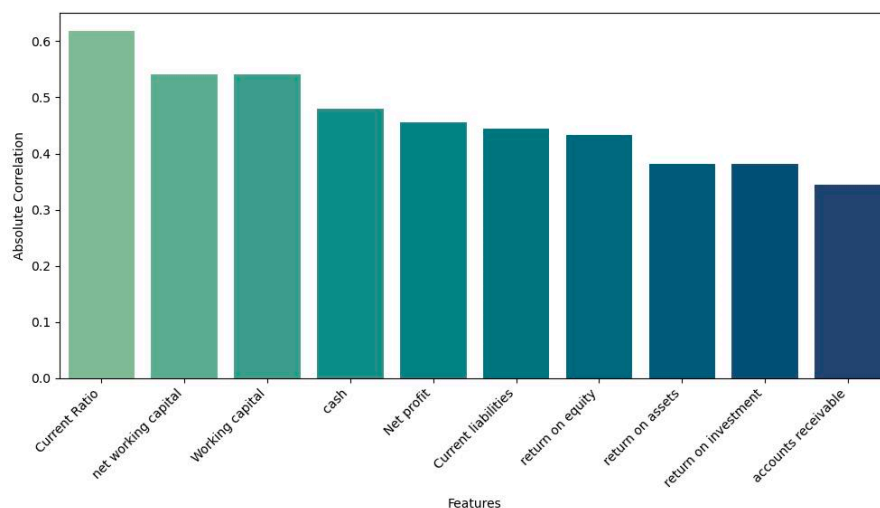


Figure 8. Highest Absolute Correlation Value of 10 Important Features.

In the first row of Table 7 with PCA-FS, the model shows enhancements of 1.4% for precision, 8.5% for recall, 5.8% for F1-score, 1.8% for accuracy, and 1.1% for AUC. In the second row of Table 7 using CC_FS, the model is boosted by 3.2% for precision, 2.6% for recall, 1.2% for F1-score, 1.8% for accuracy, and 1.9% for AUC. In the third row of Table 7 with PSO_FS, the model shows improvements of 0.006% for precision, 2.8% for recall, 1.5% for F1-score, 2.7% for accuracy, and 1.5% for AUC. When considering the number of selected features and execution time, the MLP + CC model achieves a processing time of 1.16:1157.08 and a feature space reduction of 58.3%.

Table 7. Results of Metrics for MLP + DP + FS.

Models + FS	Precision	Recall	F1	Accuracy	Jaccard	AUC
MLP + PCA	89	76	82	86	79.5	92.1
MLP + CC	84.4	93.1	88.5	86	87.7	99.3
MLP + PSO	87	95.6	91.6	95	91.9	95.5

The final Experiment D is the final fine-tuning step for the hyperparameters of the model achieved in Experiment C. Phase one of the popular techniques for hyperparameter tuning is using a random search algorithm. Random search is a simple yet effective method that involves randomly selecting hyperparameter values within a predefined range and evaluating the model performance with each set of values. In the beginning, the Hyperparameter Space is defined (Table 3 in Section 4). The optimal hyper-parameter discovered through the random search algorithm is demonstrated in Table 8.

Table 8. Optimal Hyperparameter Found by PCA, CC, and PSO.

Parameters	PCA	CC	PSO
Optimization Model	Adam	Adam	Adam
Hidden Layer Sizes	100	50	50
Activation Function	Identity	Identity	Identity
Alpha	0.7	0.001	0.001
Learning Rate	Adaptive	Constant	Adaptive
Elapsed Time [min:s:ms]	0:1.57:1569.80	0:1.16:1157.08	0:1.05:1046.33

In this study, many epochs are tested in the range of (3–100), and after many attempts, epoch 5 is selected, at which the best results are achieved in the shortest implementation time. Table 8 summarizes the estimators that achieved the best performance of these models. Essentially, the different combinations of activation functions, optimization algorithms, and architectures can cause the MLP to perform differently. The use of different metrics can provide insights into the strengths of each model. As mentioned in Table 3, different activation functions, 10 hidden layer sizes, 10 alpha values, and various learning rates and optimization algorithms are used to train and test the models. For all models, it is observed that the “Identity” function is the activation function that performs better when applied to our dataset. In our opinion, the model parameters depend on and are affected by the nature of the dataset. By the RandomizedSearchCV function in scikit-learn, the search is done over a predetermined range of hyperparameters. This method is efficient since it allows you to sample a predetermined number of hyperparameter combinations from specified distributions, especially in large search areas. The model, hyperparameter distribution, and search iteration count are all specified. You can assess the optimal parameters once the model has been fitted.

In the analysis of Table 9, it is observed that the diagnostic metrics of the models that achieve better performance are obtained using the Adam optimization algorithm. The second and third columns of Table 9 show the precision and recall rates of all models, respectively. The fourth, fifth, and sixth columns illustrate the F1, accuracy, and AUC rates, respectively. The model with the highest accuracy is the first MLP + PCA model with a value of 0.987, followed by the MLP + PSO and MLP + CC models with values of 0.983 and 0.976, respectively. On the other side, note that the second model (MLP + PSO) requires less time than other models at 1.05:1046.33, compared to the processing time of MLP + PCA and MLP + CC models at 1.57:1569.80 and 1.16:1157.08, respectively. Furthermore, only eight features, as a subset of the original set of 24 features, are employed in the MLP + PSO model.

Table 9. Results of Metrics for MLP + DP + FS (Fine-Tuning Version).

Model + FS	Precision	Recall	F1	Accuracy	Jaccard	AUC
MLP + PCA	99.9	99.9	99.8	98.7	96.5	99.9
MLP + CC	99.8	96.6	98.2	97.6	93.7	99.3
MLP + PSO	99.7	96.6	98.2	98.3	94.9	99.5

An essential tool for assessing and forecasting financial insolvency models is a confusion matrix. It offers a methodical approach for evaluating the prediction power of the model and comprehending the effects of various kinds of mistakes. For financial institutions and other stakeholders that depend on precise forecasts to make well-informed decisions, this assessment is essential. Figure 9 presents the confusion matrix of all the proposed models.

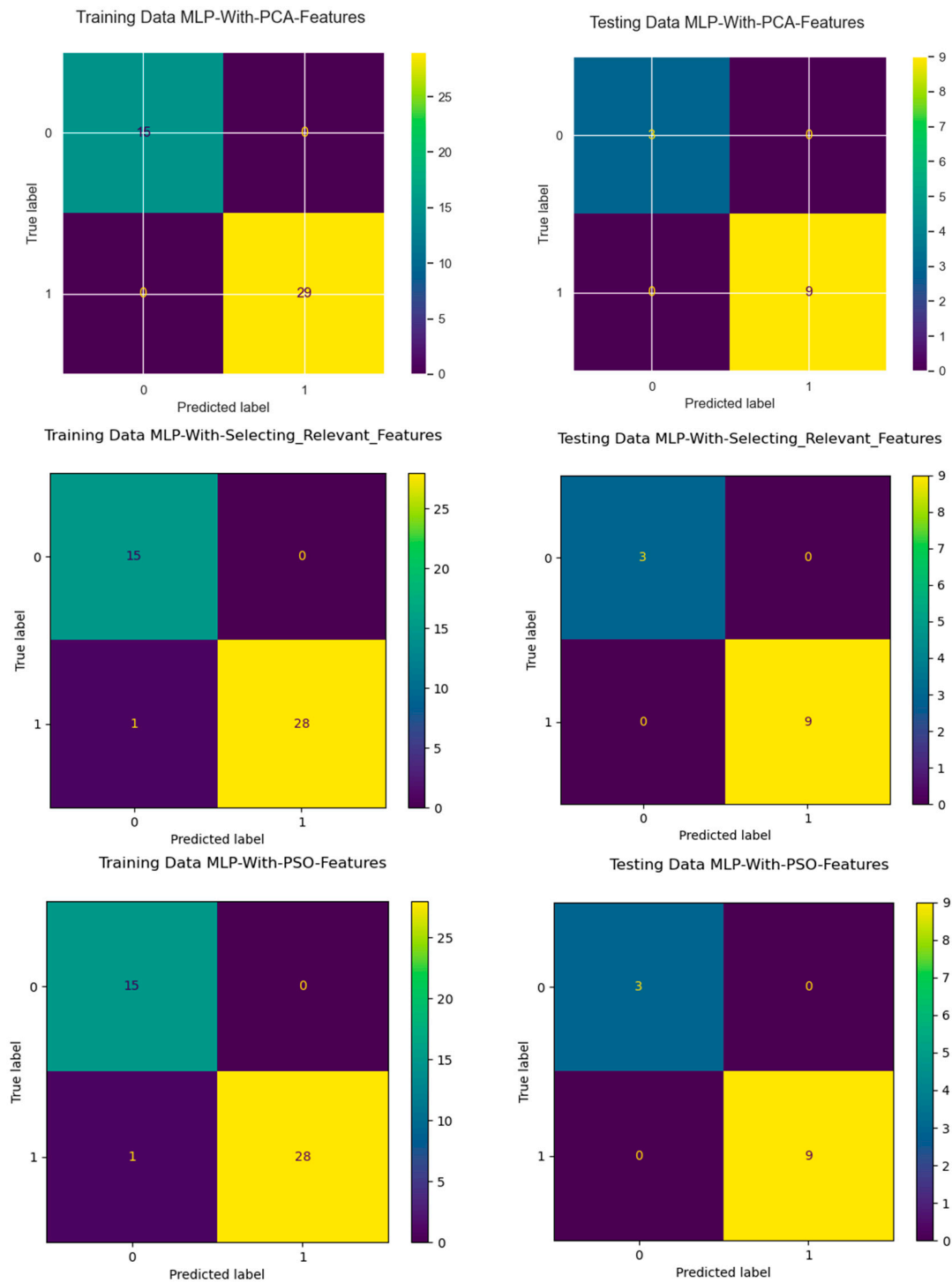


Figure 9. Confusion Matrix of Training and Testing Data Using PCA, CC, and PSO.

Figure 9 shows that the model is performing quite well, with a high number of true positives and true negatives, very few false negatives, and no false positives. This result

indicates a high sensitivity (or recall) for class 1, meaning the model is good at identifying class 1 instances. The absence of false positives suggests perfect precision for class 0, meaning that when the model predicts class 0, it is always correct. In general, the three proposed models achieve excellent performance. The difference between these models is approximately 1%. Our proposed models are compared to existing works from literature. Table 10 illustrates this comparison.

Table 10. Comparison Summary of Related Work and Proposed Models.

References & Years	Methodology	Dataset	FN	FS	HID	HPO	ACC
(Karas and Režňáková 2020)	RT & LR	Czech	15	×	×	×	93
(Le 2021)	XGBS	Korean	19	×	✓	×	82
(Budianita et al. 2022)	C4.5 + Adaboost	Indonesia	21	×	×	×	86
(Safi et al. 2022)	MHO-ANN	Spanish, Tawan, and Polish	36, 94, 64	×	×	✓	88
(Yan et al. 2020)	LLDL and LSVMDL	Chinese	27	✓	×	×	95
(Shetty et al. 2022)	XGBoost, SVM and NN	Belgian	4	×	×	×	83
(Maqbool et al. 2023)	MLP	Yahoo Finance	-	×	×	×	90
(Al Omari et al. 2023)	MLP	Jordanian	195	×	×	×	84
(Al Ali et al. 2023)	GALSTM	Osiris	31	×	×	✓	92
Proposed Models	MLP + PCA	Iraqi	20	✓	✓	✓	98.7
	MLP + CC		10	✓	✓	✓	97.6
	MLP + PSO		8	✓	✓	✓	98.3

Note: ACC (Accuracy), Number of Features (FN), Feature Selection (FS), Handling Imbalanced Dataset (HID), Hyperparameter Optimization (HPO), × (do not use), ✓ (use).

Table 10 lists a comparison of the proposed work concerning existing work. Though it is not for comparison because of the different datasets employed in different works, it displays the recent state of the art of the related works. The above table elaborates on the advantages of the three of our suggested models in terms of the Number of Features (FN), feature selection (FS), Handling Imbalanced Dataset (HID), Hyperparameter Optimization (HPO), and performance accuracy. The accuracy of our proposed model shows an increase of roughly 5–16% compared to related works. The suggested model also has another benefit of having fewer features due to the usage of FS, which utilizes a limited number (20, 10, 8 features) yet outperforms the others, thus reducing execution time. On the other side, we find it interesting to clarify the differences between investment firms and industrial/commercial firms in various aspects, particularly in terms of working capital, cash flows, and revenue sources:

- Working Capital: Working capital is, more often than not, a major requisite essential to facilitate business transactions in industrial and commercial firms. This factor includes the cost of stock, accounts receivable, and accounts payable. On the other hand, investment firms may work with lower working capital due to the fact that they mainly deal with financial securities, not blocked assets, in terms of inventory and receivables. Firms’ working capital, for the most part, relates primarily to cash flow requirements in trading operations.
- Cash Flows: The nature of the cash flows in these two types of firms is different. Industrial/commercial firms are involved in generating cash flows through revenue from their produced products and/or implementation of some services, so they can

be relatively more consistent but unpredictable depending on the market forces. Investment firms, on the other hand, obtain cash flows through fees, commissions, and investment income. These aspects can be a lot more unpredictable and sensitive to market forces, investment returns, and clients' workflow.

- Revenue Sources: Industrial and commercial organizations generate their revenues through manufacturing sales and service provision, whereas investment firms obtain their revenues through asset management fees, transaction fees, interest, and dividends on their investment portfolios. This has the effect of making their revenue functions more sensitive to the conditions of financial markets and, hence, vulnerable.

6. Conclusions

This research addresses a significant gap in previous studies that have not adequately emphasized pre-processing issues, the uneven distribution of databases, and the enhancement of predictive models in terms of financial insolvency prediction. Our proposed models tackle these gaps by processing the data and addressing the problem of class imbalance using SMOTE for improved model performance, reduced bias, increased generalization capabilities, and refined decision boundaries. For achieving a further optimized predictive model, feature selection is integrated.

Our proposed model is based on integrating an MLP with feature selection. Three techniques (PCA, CC, and PSO) are proposed for ultimately simplifying the model and reducing execution time while gaining even higher predictive accuracy. In this study, four models are examined: the classical MLP model, MLP + DP model, MLP + DP + FS model, and MLP + DP + FS + Fine tuning model.

The findings reveal that the MLP + PSO model fares better than the MLP + PCA and MLP + CC models while taking into account the number of selected features and execution time due to the MLP + PSO model's shorter processing time of 1.05:1046.33 in comparison to 1.57:1569.80 and 1.16:1157.08, respectively. The time difference that arises from this is 0.52:523.47 and 0.11:110.75, respectively. Furthermore, the MLP + PSO model maintains its position as the best predictor by using only eight features, which is a subset of the original 24 features. Furthermore, the Adam optimization algorithm performs better than the conventional SGD optimization technique. However, when accuracy is the only consideration, the results indicate that the MLP + PCA model performs better, outperforming other suggested models by 1%, with differences ranging from 5% to 16% compared to models in similar publications, as demonstrated in Table 10.

Our findings underscore the importance of feature selection from multiple sides, such as selecting the important features and reducing the time.

Based on the results of the experiments, it can be inferred that all methods effectively tackle the search problem. Configuring parameters for data uniformly can be challenging despite the MLP's strength. Adhering to MLP criteria necessitates considering factors such as sample variability, data intricacy, and suitable data preprocessing.

7. Limitations and Directions for Future Research

MLP is employed in this research in order to predict the financial performance of investment firms listed in ISX. In return, there are several limitations involved in the research. First, we only take 24 financial ratios to be the input variables of the MLP model. In the future, to address this limitation, we want to expand the number of different financial variables. Second, one machine learning technique is chosen. Thus, in the further works, it is necessary to take into account a combined approach.

In future research, we want to extend the possibilities to compare our models using more different financial variables, which will be helpful. In addition, we also expect to apply various neural network models and discuss the differences with our current work. Also, we intend to use XAI, also known as explainable artificial intelligence, which can effectively explain the reasoning behind outputs and decisions. It can be postulated that the findings of this study will have implications for enhancing decision-making and supporting

the policies in management. We also suggest that investment firms make use of the results from the above indicator to improve efficiency and performance.

Author Contributions: Conceptualization, A.A.A.-K. and Z.T.F.; formal analysis, A.A.A.-K.; funding acquisition, A.A.A.-K.; investigation, Z.T.F., S.R., S.A.E.-R. and B.M.N.; methodology, A.A.A.-K. and Z.T.F.; supervision, Z.T.F., S.R., S.A.E.-R. and B.M.N.; validation, Z.T.F. and S.R.; writing—original draft, A.A.A.-K.; writing—review and editing, A.A.A.-K., Z.T.F. and S.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data will be accessible upon request.

Acknowledgments: The author would like to extend their gratitude to the supervisors and reviewers.

Conflicts of Interest: The authors certify that there is no overlap between their personal and professional interests.

References

- Al Ali, Amal, Ahmed M. Khedr, Magdi El Bannany, and Sakeena Kanakkayil. 2023. GALSTM-FDP: A Time-Series modeling approach using hybrid GA and LSTM for financial distress prediction. *International Journal of Financial Studies* 11: 38. [CrossRef]
- Almeida, Luís. 2023. Risk and Bankruptcy Research: Mapping the State of the Art. *Journal of Risk and Financial Management* 16: 361. [CrossRef]
- Al Omari, Rania, Rami S. Alkhalwaldeh, and Jamil J. Jaber. 2023. Artificial neural network for classifying financial performance in Jordanian insurance sector. *Economies* 11: 106. [CrossRef]
- Anton, Sorin Gabriel, and Anca Elena Afloarei Nucu. 2020. The Impact of Working Capital Management on Firm Profitability: Empirical Evidence from the Polish Listed Firms. *Journal of Risk and Financial Management* 14: 9. [CrossRef]
- Bassiouni, Mahmoud M., Ripon K. Chakraborty, Omar K. Hussain, and Humyun Fuad Rahman. 2022. Advanced deep learning approaches to predict supply chain risks under COVID-19 restrictions. *Expert Systems with Applications* 211: 118604. [CrossRef]
- Budianita, Elvia, Rezi Yuliani, Ladda Suanmali, and Hidayati Rusnedy. 2022. Forecasting company financial distress: C4.5 and adaboost adoption. *DOAJ* 49: 300–7. Available online: <https://doaj.org/article/8ba779a4931b4d52a713eb65a12a8e39> (accessed on 18 September 2024).
- Casey, Anthony J., and Joshua C. Macey. 2023. Insolvency courts: General principles for systems design. *International Insolvency Review* 33: 23–39. [CrossRef]
- Chakri, Potta, Saurabh Pratap, Lakshay, and Sanjeeb Kumar Gouda. 2023. An exploratory data analysis approach for analyzing financial accounting data using machine learning. *Decision Analytics Journal* 7: 100212. [CrossRef]
- Gavurova, Beata, Sylvia Jencova, Radovan Bacik, Marta Miskufova, and Stanislav Letkovsky. 2022. Artificial intelligence in predicting the bankruptcy of non-financial corporations. *Oeconomia Copernicana* 13: 1215–51. [CrossRef]
- González, Oliver, and Benjamin Keddad. 2024. The piggy bank index: An intuitive risk measure to assess liquidity and capital adequacy in banks. *Finance Research Letters* 60: 104846. [CrossRef]
- Jadiyappa, Nemiraja, and Rachappa Shette. 2024. CSR regulation and the working capital management policy. *Global Finance Journal* 59: 100934. [CrossRef]
- Karas, Michal, and Mária Režňáková. 2020. Cash flows Indicators in the prediction of financial distress. *Engineering Economics* 31: 525–35. [CrossRef]
- Kennedy, J., and R. C. Eberhart. 1997. A discrete binary version of the particle swarm algorithm. Paper presented at 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, Orlando, FL, USA, October 12–15, vol. 5, pp. 4104–8. [CrossRef]
- Kim, Soo Y. 2011. Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis. *The Service Industries Journal* 31: 441–68. [CrossRef]
- Kušter, Denis, Bojana Vuković, Sunčica Milutinović, Kristina Peštović, Teodora Tica, and Dejan Jakšić. 2023. Early insolvency prediction as a key for sustainable business growth. *Sustainability* 15: 15304. [CrossRef]
- Laitinen, Erkki K., María-Del-Mar Camacho-Miñano, and Nora Muñoz-Izquierdo. 2023. A review of the limitations of financial failure prediction research. *Revista de Contabilidad* 26: 255–73. [CrossRef]
- Lee, Chia-Chi. 2023. Analyses of the operating performance of information service companies based on indicators of financial statements. *Asia Pacific Management Review* 28: 410–19. [CrossRef]
- Le, Tuong. 2021. A comprehensive survey of imbalanced learning methods for bankruptcy prediction. *IET Communications* 16: 433–41. [CrossRef]
- Lin, Boqiang, and Rui Bai. 2022. Machine learning approaches for explaining determinants of the debt financing in heavy-polluting enterprises. *Finance Research Letters* 44: 102094. [CrossRef]
- Lohmann, Christian, Steffen Möllenhoff, and Thorsten Ohliger. 2022. Nonlinear relationships in bankruptcy prediction and their effect on the profitability of bankruptcy prediction models. *Journal of Business Economics* 93: 1661–90. [CrossRef]

- Maqbool, Junaid, Preeti Aggarwal, Ravreet Kaur, Ajay Mittal, and Ishfaq Ali Ganaie. 2023. Stock prediction by integrating sentiment scores of financial news and MLP-regressor: A machine learning approach. *Procedia Computer Science* 218: 1067–78. [\[CrossRef\]](#)
- Mattos, Eduardo da Silva, and Dennis Shasha. 2024. Bankruptcy prediction with low-quality financial information. *Expert Systems with Applications* 237: 121418. [\[CrossRef\]](#)
- Nayak, Sarat Chandra, Satchidananda Dehuri, and Sung-Bae Cho. 2022. Intelligent financial forecasting with an improved chemical reaction optimization algorithm based dendritic neuron model. *IEEE Access* 10: 130921–43. [\[CrossRef\]](#)
- Pratiwi, Mutiara Restu, Yovita Vivianty Indriadewi Atmadjaja, and Indah Wahyu Ferawati. 2023. Prediction analysis of company Bankruptcy using comparison of the altman method (Z-Score) and grover method (G-Score) as an early warning system in pharmaceutical subsector companies. *Jurnal Maksipreneur: Manajemen, Koperasi, dan Entrepreneurship* 12: 486–98. [\[CrossRef\]](#)
- Radovanovic, Jelena, and Christian Haas. 2023. The evaluation of bankruptcy prediction models based on socio-economic costs. *Expert Systems with Applications* 227: 120275. [\[CrossRef\]](#)
- Safi, Salah Al-Deen, Pedro A. Castillo, and Hossam Faris. 2022. Cost-Sensitive Metaheuristic Optimization-Based Neural Network with Ensemble Learning for Financial Distress Prediction. *Applied Sciences* 12: 6918. [\[CrossRef\]](#)
- Salah, Wafaa. 2020. The International Financial Reporting Standards and Firm Performance: A Systematic Review. *Applied Finance and Accounting* 6: 1–10. [\[CrossRef\]](#)
- Sang, Bin. 2020. Application of genetic algorithm and BP neural network in supply chain finance under information sharing. *Journal of Computational and Applied Mathematics* 384: 113170. [\[CrossRef\]](#)
- Sharma, H. Kumar, K. Kumari, and S. Kar. 2018. Air passengers forecasting for Australian airline based on hybrid rough set approach. *Journal of Applied Mathematics, Statistics and Informatics* 14: 5–18. [\[CrossRef\]](#)
- Shetty, Shekar, Mohamed Musa, and Xavier Brédart. 2022. Bankruptcy prediction using machine learning techniques. *Journal of Risk and Financial Management* 15: 35. [\[CrossRef\]](#)
- Tarjo, Tarjo, Prasetyono Prasetyono, Eklamsia Sakti, Yusarina Mat-Isa, and Otniel Safkaur. 2023. Predicting fraudulent financial statement using cash flow shenanigans. *Verslas Teorija ir Praktika* 24: 33–46. [\[CrossRef\]](#)
- Vo, Duc Hong. 2023. Market risk, financial distress and firm performance in Vietnam. *PLoS ONE* 18: e0288621. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wang, Chengfu, Xiangfeng Chen, Xun Xu, and Wei Jin. 2023. Financing and operating strategies for blockchain technology-driven accounts receivable chains. *European Journal of Operational Research* 304: 1279–95. [\[CrossRef\]](#)
- Wang, Jun, Mao Li, Martin Skitmore, and Jianli Chen. 2024. Predicting construction company insolvent failure: A scientometric analysis and qualitative review of research trends. *Sustainability* 16: 2290. [\[CrossRef\]](#)
- Wu, Qun, Xinwang Liu, Jindong Qin, Ligang Zhou, Abbas Mardani, and Muhammet Deveci. 2022. An integrated generalized TODIM model for portfolio selection based on financial performance of firms. *Knowledge-Based Systems* 249: 108794. [\[CrossRef\]](#)
- Xu, Jian, Muhammad Akhtar, Muhammad Haris, Sulaman Muhammad, Olivier Joseph Abban, and Farhad Taghizadeh-Hesary. 2022. Energy crisis, firm profitability, and productivity: An emerging economy perspective. *Energy Strategy Reviews* 41: 100849. [\[CrossRef\]](#)
- Yan, Dawen, Guotai Chi, and Kin Keung Lai. 2020. Financial distress prediction and feature selection in multiple periods by lassoing unconstrained distributed lag non-linear models. *Mathematics* 8: 1275. [\[CrossRef\]](#)
- Yi, Kun, Qi Zhang, Wei Fan, Shoujin Wang, Pengyang Wang, Hui He, Ning An, Defu Lian, Longbing Cao, and Zhendong Niu. 2023. Frequency-domain MLPs are More Effective Learners in Time Series Forecasting. *arXiv arXiv:2311.06184*. [\[CrossRef\]](#)
- Zhang, Wen, Shaoshan Yan, Jian Li, Xin Tian, and Taketoshi Yoshida. 2022. Credit risk prediction of SMEs in supply chain finance by fusing demographic and behavioral data. *Transportation Research. Part E Logistics and Transportation Review* 158: 102611. [\[CrossRef\]](#)
- Zhao, Jinxian, Jamal Ouenniche, and Johannes De Smedt. 2024. Survey, classification and critical analysis of the literature on corporate bankruptcy and financial distress prediction. *Machine Learning with Applications* 15: 100527. [\[CrossRef\]](#)
- Zhu, Lei, Menghao Li, and Noura Metawa. 2021. Financial Risk Evaluation Z-Score model for intelligent IoT-based enterprises. *Information Processing & Management* 58: 102692. [\[CrossRef\]](#)

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