Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
A new metaheuristic optimization model for financial crisis prediction: Towards sustainable development

Mohamed Elhoseny a,b,* , Noura Metawa a,c, Ibrahim M. El-hasnony a

a Mansoura University, Egypt
b College of Computing and Informatics, University of Sharjah, the United Arab Emirates
c College of Business Administration, University of Sharjah, the United Arab Emirates

ARTICLE INFO

Keywords:
Financial crisis prediction
Sustainable development
Deep learning
Political optimizer
Machine learning
Hyperparameter tuning

ABSTRACT

Global crises such as the COVID-19 pandemic and other recent environmental, financial, and economic disasters have weakened economies around the world and marginalized efforts to build a sustainable economy and society. Financial crisis prediction (FCP) has a significant impact on the economy. The growth and strength of a country’s economy can be gauged by accurately predicting how many companies will fail and how many will succeed. Traditionally, there have been a number of approaches to achieving a successful FCP. Despite this, there is a problem with the accuracy of classification and prediction and with the legality of the data that is being used. Earlier studies have focused on statistical, machine learning (ML), and deep learning (DL) models to predict the financial status of a company. One of the biggest limitations of most machine learning models is model training with hyper-parameter fine-tuning. With this motivation, this paper presents an outlier detection model for FCP using a political optimizer-based deep neural network (OD-PODNN). The OD-PODNN aims to determine the financial status of a firm or company by involving several processes, namely preprocessing, outlier detection, classification, and hyperparameter optimization. The OD-PODNN makes use of the isolation forest (iForest) based outlier detection approach. Moreover, the PODNN-based classification model is derived, and the DNN hyperparameters are fine-tuned to boost the overall classification accuracy. To evaluate the OD-PODNN model, three different datasets are used, and the outcomes are inspected under varying performance measures. The results confirmed the superiority of the proposed OD-PODNN methodology over recent approaches.

1. Introduction

New metrics-based planning approaches are becoming more prevalent to address the current sustainability challenges faced by cities. A densely populated area has serious social, economic, and environmental consequences. Because of this, new demands, climate change, and economic downturns necessitate a flexible infrastructure and an adaptable environment. As a result, planners can better allocate available resources based on the performance of the studied processes and the social, ecological, and economic uncertainties in the system. Using model-based planning approaches can significantly enhance performance-based planning (PBP). We used artificial intelligence to deal with these disasters and achieve sustainable development after the COVID-19 pandemic.

Many issues have arisen across many fields, not the least of which is finance, resulting from the problems caused by COVID-19 and the global upheavals. The emergence of financial issues necessitated the development of a model for predicting the crisis outcomes from COVID-19. You have financial distress when you cannot meet your obligations because of the inability of your company to make enough money to do so. Financial distress can be caused by a high percentage of illiquid assets, high fixed costs, or revenue vulnerable to economic downturns. When warning signs of financial distress go unnoticed, it can be dangerous. As a result, the company’s liabilities rise to an unmanageable level, and it may be impossible to recover from a state of dire financial straits. Academics continue to focus on predicting financial distress because it is critical for firms and stakeholders such as capital market players, investors, and lenders in general.

Due to the significant implications of business crises on the sustainable development of society and the economy (international debt), they have been stimulating research to precisely explore the impact of bankruptcy and figure out ways to prevent it. To decrease the influence
of this crisis, companies could utilize the economy for help/funding from financial organizations. In contrast, decision-makers in the financial organization are trying to figure out those corporations, i.e., apparently announced bankruptcy in the future [1]. For that purpose, bankruptcy prediction or company crisis focuses on assessing future performance and financial health. The current study has primarily concerned the financial aspect, gaining outcomes with better predictive rates in the short-term; generally, twelve months [2]. But, because of diverse principles, this method tends to be more precise for more significant and medium–larger businesses.

Financial crisis prediction (FCP) is a significant field that assists financial institutions in making decisions at the appropriate moment [3] for sustainable development purposes. It is because of the intention of incorrect decision-making in corporations, which might result in bankruptcy/financial crises and affect clients, investors, vendors, etc. The current development in information technology allows accomplishing different kinds of data associated with the threat levels of companies from different manners. Most people depend on the predictor decision in assessing a massive amount of information [4]. However, several factors might contain an impact on analyses of the performance. AI and statistical approaches are utilized to identify FCP. In FCP, AI is used in several ways. It is used to build models that predict if a financial institution may suffer a crisis.

The primary concern is that financial variable filters from public financial declarations like financial ratios hold vast information regarding firm financial conditions convenient for FCP [5]. Using financial data and other data from organizations’ tactical competitiveness to design an effective technique is arduous. Aside from databases and AI, data mining technologies are widely used in other ways. FCP uses data mining in two ways: decision-making and early warning. Acting to avoid financial collapse is seen as useful. It then aids financial decision-makers in selecting and evaluating businesses for investment/collaboration. Deeper data analysis requires more resources and computational time. Filtering data with erroneous information is difficult. Hence, the procedure of extracting an enormous quantity of information is a critical part of recognizing financial failure, mainly for FCP [6]. The feature selection (FS) is an essential pre-processing phase in data mining. It proposes to filter the repetitive and unwanted features from the original information.

Filters, wrappers, and embedded-based processes are classified as FS systems. The wrapper system uses a learning mechanism to validate the results of the specified feature subsets. Although extensively used, wrapper techniques have several flaws, such as learner limits and processing complexity. Embedded techniques are easier than wrapper methods because they employ a different learning mechanism [7]. They used the filter approach due to restrictions. The filter approach selects features and learners before determining feature subsets. The FS issues are procedures. Detecting optimal features from accessible features is an NP-hard issue. Diverse strategies are offered to discover inferior solutions quickly. A genetic algorithm (GA) and ant colony optimization (ACO) were used to select critical traits. They aren’t used commercially, mostly for FCP.

Earlier research has focused on statistical, machine learning (ML), and deep learning (DL) models to forecast the financial health of a corporation. One of the main constraints is model training with hyperparameters fine-tuning. This paper presents an outlier detection model for FCP that is based on a political optimizer-based deep neural network (OD-PODNN). The objective of the OD-PODNN technique is to determine the financial status of a firm or company to ensure its long-term viability. Multiple processes are involved in the proposed OD-PODNN technique; these include data preprocessing, outlier detection, classification, and hyperparameter optimization. In addition, the OD-PODNN technique makes use of an outlier detection approach based on the isolation forest (iForest). Mainly, the iForest approach is applied to eradicate the occurrence of outliers in the financial data. The iForest technique has a tree-based outlier detection method with linear time difficulty and superior precision and is appropriate to maximum dimension and a considerable number of records. The DNN methodology can be utilized to classify the financial status of a company.

Furthermore, the PODNN-based classification model is developed, and the hyperparameter tuning of the DNN is carried out to improve the model’s overall classification accuracy. The political optimizer (PO) selects the DNN hyperparameters. This approach is a human social behavior stimulated metaheuristic which maps the concepts of multiple-party-political systems algorithmically and mathematically. It includes five stages: inter-party election, party switching, election campaign and government formation, parliamentary affairs, constituency allocation, and party formation. All the political members represent candidate solutions, and the population is separated into subclasses named political parties. It is necessary to perform performance validation on the OD-PODNN process with the help of benchmark datasets, and the results are then examined from various perspectives.

The main contributions of this paper can be summarized as follows:

- Data preprocessing using methods for missing values imputation, data cleaning, feature selection, and outlier detection and correction.
- Outlier detection by the effective iForest technique.
- Data classification by the DNN.
- Fine-tuning the DNN hyperparameters using the political optimizer.
- Evaluating the proposed model using three datasets.
- Performance evaluation using the accuracy, precision, recall, kappa, and F-score.

This paper’s reset is laid out as follows. The related work is discussed in Section 2. Section 3 delves into the model’s core concepts. It’s all laid out in detail in Section 4. Section 5 provides a detailed description of the experimental analysis and subsequent discussion. Finally, Section 6 provides the conclusion for the paper.

2. Related works

The theoretical and empirical literature has reached a growing consensus regarding the significant impact that the development of the financial system and financial crisis has on economic growth. In Song et al. [8], a genetic algorithm (GA) based model and statistical filter methods are employed for identifying the optimal feature for the support vector machine (SVM) method. The presented model is developed for having the ability to concurrently improve the parameters and features of the SVM. The simulation result shows on the information from Chinese companies shows that the GA-based method might extract smaller number of features with high precision than statistical filter methods. Uthayakumar et al. [9] proposed a cluster-based classification method, includes 2 phases: a fitness-scaling chaotic genetic ant colony algorithm (FSCGACA) and enhanced K-means clustering based classification method. Initially, an enhanced K-means method is invented for eliminating the incorrectly clustered information. Next, a rule-based method is chosen for designing the provided database. Finally, FSCGACA is used to seek the ideal parameter of the rule-based technique.

In [10], the communication features of financial crises in stock markets are examined according to the complicated network. The influence factors and propagation method of financial crises in global stock markets are analyzed qualitatively by wide-ranging applications of the quantitative and qualitative analyses, by taking transmission theory and complicated network of financial crises as theoretic bases. Sankhwar et al. [11] present a new prediction architecture for FCP by incorporating IGWO and FNC methods. An IGWO model is derived through incorporating tumbling effects and GWO model. The proposed IGWO-based F model is applied for discovering the optimum feature from the financial information. For classification method, FNC is applied.

Samitas et al. [12] research on “Earlier Warning System” (EWS) by
exploring feasible contagion risk, based financial network. EWS indicator improves crisis predictive method performances. With machine learning algorithms and network analysis, they detect indications of contagion risks on the dates in which the researchers witness substantial growth in centralities and correlations. Perboli and Arabnejad [13] emphasize long- and mid-term bankruptcy predictions (around sixty months) aiming at medium/small institutions. The important role of this work is a considerable development of the predictive performance in the short-term (12 months) with ML models, then an innovative method, when making correct long- and midterm predictions.

An SVM-based ensemble model for FCP was introduced in [2], and a comparison was made with the classical SVM classifier. The results show that the SVM ensemble method outperforms the SVM classifier in terms of performance. The accuracy and precision of Taiwan stock exchange data affect the model’s ranking, as Lin et al. [14] found when combining SVM with traditional models. Data mining techniques are used to conduct a thorough investigation of FCP in 107 Chinese companies [15]. Decision trees and SVMs were found to be less effective than neural networks.

It is currently useful to solve the classification task using bio-inspired algorithms. Quantitative as well as qualitative classification of financial data was achieved using ant miner by Uthayakumar et al. [16]. The FS method is used to develop an effective enhanced boosting technique known as FS-Boosting [17]. Improved accuracy and diversity can be achieved with the use of the FS process in boosting. Lin et al. [18] presents an innovative approach to the concept of FS, which incorporates expert knowledge and wrapper design. To begin, the financial data features are divided into seven sets and used a method of wrapper to select features from these sets. Traditional FS methods are shown to be less accurate in the simulations, whereas the new method is superior. A wide range of GWO algorithm applications are also documented in academic literature [19]. There are a variety of FCP models, but their performance can be further studied. The tumbling effect strategy has not been used in any of the aforementioned studies to address the issues with the GWO algorithm’s limitations.

Failure, on the other hand, referred to a situation in which the company was unable to pay its creditors, preferred stockholders and suppliers or had a bill that was overdue. All of these events led to a halt in the company’s operations [20]. As a result, the "failure" can be defined as a situation in which "the realized rate of return on invested capital, with allowances for risk consideration, is significantly and continually lower than prevailing rates of similar investments," a definition that did not indicate the discontinuity of the firm. On the other hand, the Federal Deposit Insurance Corporation (FDIC) was able to broker the sale of Washington Mutual’s assets to JP Morgan Chase in 2008 for US$ 1.9 billion [21]. "failure ‘Insolvency’ was used to describe a company’s financial woes and was used to denote a company’s negative net worth [22]. Taking on that much debt is the most common cause of insolvency. As stated by Ross et al. [23], financial distress occurs when operating cash flows are insufficient to meet current obligations (such as trade credit or interest expenses), and the firm is forced to take corrective action. For the purposes of this definition, a debtor defaulted meant that they had not paid a debt they were supposed to pay. When someone couldn’t pay their debts, they were called “insolvent” in the legal sense of the word. As a legal term, "bankruptcy" was used to describe a situation where a company was either insolvent or had fallen behind on payments. Several studies on Chinese firms [24] had fallacies in the ST label as a symbol of financial distress. There are four stages that ST companies go through, including the omission or reduction of the annual dividend payment due to a lack of cash, default on loan payments leading to legal action, reorganization or takeover and transfer to Asset Management Companies, delisting from a stock exchange, and for the disposal of assets. Taken as examples of insolvent firms, ST companies are used for two primary reasons. When a company’s finances begin to go downhill, it usually takes a long time. ST is a good ex ante example of financial distress because it occurs before bankruptcy even occurs. Due to the complicated bankruptcy process in China, researchers are unable to obtain data on bankrupt companies. Because they lacked a database of bankruptcy-related information on Chinese listed companies, they defined financially troubled companies as PT and ST companies in [25].

Lin et al. [26] examine a wider coverage of financial features, such as the suggested financial ratios from TEJ (Taiwan Economic Journal) databases besides those financial ratios examined in previous studies. The goal is to determine possibly valuable but formerly aware financial features for good predictive performance. In this work, they employed data mining systems to recognize 5 advantageous financial ratios, which two of them, tax rates and constant 4 quarterly EPS are formerly unaware of the study field. Uthayakumar et al. [27] presented an ACO-FCP method that integrates 2 stages: ACO-IFS and ACO-DC algorithms.

As discussed in the previous literature, financial crisis prediction (FCP) remains an exciting research area among research communities and businesspeople in business intelligence for sustainable development. Predicting a company’s financial health has previously been studied using statistical, deep-learning models and machine-learning. With this motivation, this paper presents an outlier detection with a political optimizer-based deep neural network (OD-PODNN) model for FCP. The current study uses feature selection, outlier detection, correction, and finally, the DNN hyperparameters are fine-tuned.

3. Methods and overviews

This section provides the main methods and algorithms that the proposed model based on. These methods include the deep neural network, isolation forest, and the political optimizer algorithm.

3.1. Deep neural network

DNNs, which are artificial neural networks with multiple hidden layers, are used to simulate neurotransmissions in the human brain in order to learn new information. In the DNN, data is input into the input layer, processed by neurons in the hidden layer, and then input into the next hidden layer until the result is input into the output layer to obtain the predictive value, then back propagation is conducted to correct the result. Such a structure is more efficient in building prediction models and abstraction, building more complex models by simulation, and is good at processing big data. The calculating method of the hidden layer performed where the variable is multiplied by the weight, the bias is added, then the activation function (AF) is used for nonlinear transformation, and the obtained result is input into the next neuron. The above steps are repeated until the predicted value is finally calculated. Hence, Eq. (1) is formed. The DNN referred to in this study is all composed of dense layers (all fully connected layers).

\[ Y_i = AF \cdot \text{Sum} \]  

(1)

3.2. Isolation forest

The isolation forest (IF) technique is applied for outlier detection to remove the unwanted instances from the dataset. IF provides maximum prediction quality with supreme efficiency while resolving high-dimensional data [28]. In IF, the existence of abnormalities is irrelevant to the prediction function. Finally, the IF is defined as an unsupervised and non-parametric model which is operated well if the outliers have not existed in the training sample.

Measurements have been made of abnormalities with maximum vulnerabilities in the context of an isolation-related model [29]. Here, isolation is a process used for isolating the sample from the remaining instances. Path length (depth) from the root node to the termination node in a tree depends on the number of splitting tasks. To measure the outliers, this path length may be used because the random dividing process is a standard short operation for anomaly points [30]. A forest’s
average path length is used to calculate the observation’s abnormal score. In order to improve processing efficiency, outliers are predicted using IF in two phases due to the large volume of data in the operating state. Isolation trees were first developed by using a subset of the training dataset to build them. Each test sample, secondly, is sent to the isolated trees in a forest to measure the distance traveled by the test sample. To develop 1 trees for training, a sub-sample is divided into recursive sample divisions until each sample is isolated. A split task is performed by a random selection of a parameter, and deciding a split measure among the higher and lower values of decided parameter. Isolation trees have been developed by using split operations, which are described by a split variable and its corresponding value. The isolation tree is constructed using a randomly selected subsample X from the input data X. There are 2 variables have to be elected for isolation forest for training phase, the size of sub-samples ψ and count of trees t.

As the count of sub-samples ψ are minimum when compared with data gathered from N and t Trees are projected in as training forest significantly. In estimation phase, a forest of isolation trees is created by the applied dataset. Outliers are instances with short average path lengths in a i Trees. A path length h(x) has been accomplished by counting edges from root node to termination node in which instance x has been placed by i Tree.

An abnormal score is essential for outlier prediction model for estimating the degree of outliers quantitatively. The maximum height of i Tree is computed by the size of sub-samples ψ, the average height is dependent to the order of log ψ. In order to eliminate the comparison of path lengths by using iTrees created with distinct sub-sample sizes, a normalized outlier value has been developed in this study. Here, the analysis model along with a BST for estimating the average path length of i Tree. In case of sample set with ψ instances, Preiss determined the average path length of ineffective searches in BST by:

\[
c(ψ) = \begin{cases} 
2H(ψ - 1) - 2(ψ - 1)/ψ & \text{for } ψ > 2 \\
1 & \text{for } ψ = 2 \\
0 & \text{otherwise}
\end{cases}
\] (2)

where \(H(i)\) implies the harmonic value evaluated by \(\ln(i + 0.5772156649)\) (Euler’s constant). \(c(ψ)\) defines the average path length for ψ instances. The anomaly score of instance x is illustrated as:

\[
s(x, ψ) = 2^{-c(x)}
\] (3)

where \(E(h(x))\) denotes the h(x) from blend of iTrees. Based on Eq. (2), the given conclusions are retrieved as follows:

- If \(E(h(x)) \to 0, \to 1\). It shows that when average path length of an instance is nearby 0, it is highly simple for isolating the sample. This instance is assumed as an anomaly.
- If \(E(h(x)) \to ψ - 1, s \to 0\). It implies that when average path length of an instance is nearby massive heights of i Trees, it is extremely complex for isolating the instance. It is assumed as a normal and simple instance in a data set.
- If \(E(h(x)) \to c(ψ), s \to 0.5\). It implies that it is not evident whether instance is good or not.

Afterward, outlier values of the instances ought to be measured; these samples are organized according to the reducing order for identifying the required outliers. The predicted abnormalities showcase that the structure of a product is various from healthy baseline. Such outlier’s field data has been employed for consecutive step for identifying the parameters which contributes in outlier’s condition of a product.

3.3. Political optimizer

Metaheuristics based on sociology are referred to as “political optimizer” (PO). Multi-party political systems [31] are mapped out by this method. In a tabular format, the population is arranged in a logical way. From the perspective of the column, solutions represent the candidates running for office in the same district, while from the perspective of the row, solutions represent the members of the same political party. Fig. 3 depicts such a logical division. As a result of this logical separation, each solution is both a member of the party and a candidate for election. The following are the five phases of PO [32]:

3.3.1. Party formation and constituency allocation

There are n political parties, and there are n constituencies in which the people are divided logically such that \(n = \sqrt{\text{size_of_population}}\) as seen in Fig. 1. A candidate solution is denoted by \(p^{k}_{b}\), where b th constituency is denoted by b and a th party denoted by a denotes. The winner of constituency b is identified by the notation \(c^{k}_{b}\), which stands for the best solution in constituency b. Also, \(p^{*}_{a}\) is the best solution and party leader of party a and is denoted by \(p^{*}_{a}\). The next phase begins after the parties have been formed and constituencies assigned.

3.3.2. Election campaign

During this phase, the algorithm’s primary position-updating mechanism is implemented. PO introduces a position-updating strategy known as RPPUS (recent past-based position updating strategy) in the original paper. First, the party leader’s position is updated, and then that of the constituency winner’s position is updated, using two equations each consisting of three cases. For each case, we consider the search agent’s current location, its previous location, and its current position relative to a reference solution (constituency winner or party leader).

3.3.3. Party switching

After the election campaign, each candidate (search agent) is randomly selected and swapped with the candidate from the opposing party who is the least fit, according to a probability determined by an adaptive parameter \(λ\). After a number of iterations, a party’s party-switching rate \(λ\) decreases monotonically until it reaches 0. It is during this stage that the delicate balance between exploration and exploitation can be achieved.

3.3.4. Election

These results are used to determine which solutions are most likely to win in each constituency and which party leaders are most likely to be reelected at this point.
3.3.5. Parliamentary affairs

Constituency winners are once again interacting with one another in order to improve their positions.

4. The proposed model

The OD-PODNN model is introduced in this section. OD-PODNN model is composed of three stages: detection of outliers using iForest; classification using DNN; and tuning of hyperparameters using the PO model. Fig. 2 illustrates the complete work processes of the presented OD-PODNN technique.

The proposed model starts by dataset preparation through the preprocessing stage. In the preprocessing stage handle the dataset missing value by replacing the missing value with mean or mode that reasonable

Fig. 2. Block diagram of OD-PODNN model.

Fig. 3. Block diagram of iForest based outlier detection.
features are assigned a decimal value between 0 and 1, with 1 denoting the largest possible value. The process of feature selection is used to eliminate the irrelevant features that is considered redundant features and misleading the learning model. outlier detection and correction is performed using the isolation forest algorithm (IFA). After data preprocessing the cleaned data is passed to deep learning with the political optimization (PO) algorithm for hyper parameters tuning. The details for the main process of the proposed model discussed in the following subsections.

4.1. Stage 1: iForest based outlier detection

Primarily, the iForest approach is applied to eradicate the occurrence of outliers in the financial data. IForest technique has been presented as Liu et al. [33]; this technique has been tree based outlier detections method with linear time difficulty and superior precision and appropriate to maximum dimension and huge number of records as seen in Fig. 3.

As anomalies are “less and distinct,” they are further susceptible that isolated. During the record oriented RT, data are recursively cut still every record is isolated. Arbitrary partitions generate outlier data as shorter path length as record with different attribute value has been highly possible that divided in early partition. The IForest has of few iTrees (Isolation Tree). All the iTrees are binary trees [34]. The execution phases are as follows:

1) Arbitrarily choose a set amount of instance points in the trained record as subsamples and fit it during the root node of trees.
2) Arbitrarily identify an attribute as well as arbitrarily make the cut point during the present node record, cutting point has been created amongst the maximal as well as minimal value of the detailed attribute from the present node record.
3) The hyperplane was created in this cutting point, and the recording space of present node has been separated as to 2 subspaces: the record lesser than p from the stated attribute is put as to the left child of present node, and the record superior to or equivalent to p is set the right child of present node.
4) Recursively apply steps 2 & 3, still, the child nodes are only one data or iTree is attained the determined as height.

Afterward receiving this iTree, the trained of iForests have been terminated, and next, it can estimate the testing record utilizing the created iForests. To the testing data, assume it traverses all iTrees and afterward computes height of data which finally fall on every tree. Next, it can obtain the average height of record from all the trees. When the average height has been smaller when compared to the provided threshold, next the data has been regarded as outlier. Fig. 4 illustrates the iForest based outlier detection process steps.

4.2. Stage 2: DNN based classification

In the second phase, the DNN methodology can be utilized to classify the financial status of a company. A DNN is an ANN method which consists of hidden, output, and input layers. The hidden layer executes a group of nonlinear functions which is expressed by:

\[ Z = \text{sig}(W \cdot x + \text{bias}) \]

In the equation, \( x \) denotes the input of each node, \( \text{sig} \) and \( \text{bias} \) represents the weight and bias vector correspondingly and \( \text{sig} \) characterizes the sigmoid activation function, i.e. \( \frac{1}{1 + e^{-x}} \). In this presented method, 2 hidden layers are taken into account and increase the mean absolute error (MAE) of DNN, optimum election of the weight metrics is needed hence, SAR method has been employed. SAR is inspired by the research conducted by humans at the time of search and rescue process. The search and rescue operations consist of individual stage and social stage. In the search method, group member collects the clue. The clue left at the search through the group member is saved in the memory matrix (O) where the human position is saved in the location matrix (W). The slopes matrix B using size \( N \cdot D \), has the left clue and the human position is denoted by:

\[ B = \begin{bmatrix} W \ O \end{bmatrix} \]

(5)

The 2 stages of human search are modeled in the following [35]. i) Social stage: The search direction is represented as \( SD_i = (W_i - B_i) \) in which \( k \neq i \). The novel solution is made by:

\[ W_i' = \begin{cases} B_i + r_i (W_i - B_i), & \text{if } f(B_i) > f(W_i) \\ W_i + r_i (W_i - B_i), & \text{if } r_i > SE \\ W_i, & \text{otherwise} \end{cases} \]

(6)

Now \( f(B_i) \) & \( f(W_i) \) represent the FF value for \( B_i \& W_i \). \( r_i \) & \( r_2 \) indicates arbitrary values in the interval of \([-1, 1]\) and \([0, 1]\), SE is method variable that ranges from zero and one. Fig. 4 illustrates the framework of DNN model. A candidate solution for the DNN using the political optimizer is shown in Fig. 5.

ii) Individual Stage: According to the existing location human recognize their novel location and \( i \& \text{human} \) is represented as,

\[ W_i = W_i + r_i (B_i - B_i), \quad i \neq k \neq m \]

(7)

Each solution must find in the solution space, when the novel location is outer the solution space is enhanced by

\[ W_i' = \begin{cases} \frac{w_{ij} + w_{ij}^\text{max}}{2} & \text{if } w_{ij}^\text{max} > w_{ij}^\text{max} \\ \frac{w_{ij} + w_{ij}^\text{min}}{2} & \text{if } w_{ij}^\text{max} < w_{ij}^\text{min} \end{cases} \]

(8)

In which \( W_i^\text{max} \) & \( W_i^\text{min} \) indicates the maximal and minimal of the threshold. The efficacy of detecting the global optimum solution can be improved as follows

\[ ME_i = \begin{cases} w_i & \text{if } f(W_i) > f(W_i) \\ ME_i & \text{otherwise} \end{cases} \]

(9)
Let \( M_{E_n} \) be the \( n \)th saved clues location in the memory matrix and \( n \) means an arbitrary integer value ranges from 1 and \( N \). In clue search method when optimal clue isn’t found near the present location afterward a specific amount of searches, humans go to a novel location. Initially, an unsuccessful search number (USN) is fixed to 0.

\[
\text{USN}_i = \begin{cases} 
\text{USN}_i + 1 & \text{if } f(W'_i) > f(W_i) \\
0 & \text{otherwise}
\end{cases}
\]

Once the USN values are larger when compared to the highest USN, the human got an arbitrary position in the searching space using Eq. (12), and the values of \( \text{USN}_i \) is fixed to 0.

\[
w_{ij} = w_{j\min} + r_4(w_{j\max} - w_{j\min}); i = 1, \ldots, D
\]

In which \( r_4 \) in the range of zero and one.

4.3. Stage 3: PO based hyperparameter tuning process

Finally, the DNN hyperparameters are selected by the application of PO. This approach [36] is a human social behavior stimulated meta-heuristic which maps the concepts of multiple-party-political systems algorithmically and mathematically. It includes 5 stages: inter-party election, party switching, election campaign, and government formation, parliamentary affairs, constituency allocation, and party formation. All the political members represent candidate solutions and the population is separated into subclasses named political parties. PO contains 2 exclusive characteristics:

- Logical division of population amongst constituencies and parties to allocate double roles to all the candidate solutions. Use of the current location of a candidate solution in the location update system. This method is called a Recent Past-based Position Updating Strategy (RPPUS).

All the phases play a different role in meeting distinct optimization problems that are discussed in these subsections.

4.3.1. Party formation and constituency allocation

The population is separated into \( n \) groups (party) all having n
solution (party member). Besides performing the roles of party members, a solution as well plays as an election candidate which rational ly splits the population into n constituency. In the novel PO, the mapping can be saved since there is single parameter n which determines the amount of members in all parties, the amount of constituencies, and number of parties. Furthermore, a candidate contests election from accurately single constituency.

4.3.2. Election campaign

In actual politics, the candidate visits their constituency and convinces the voter to elect them beforehand election. Mathematically PO expresses this stage to upgrade the locations of the solution in the searching space beforehand fitness computations. All the members (candidate solutions) initially upgrade their location regarding the winner of its constituency and party leader. Additionally, the learning from the mistake in prior election is mapped by the use of prior location of the candidate solutions in location-update method. The idea is arithmetically expressed in Eqs. (13) and (14). Furthermore, the step-wise demonstration can be shown in Algorithm 1. Where, $p_j(t)$ embodies $k$th parameter of $j$th member of $i$th political party in iteration $t$, $p_i^*$ means the leader of $i$th party, $c_j$ represents winner of $j$th constituency, $m^*$ symbolizes the $k$th parameter of party leader next constituency winners, then $f$ indicates the objective function.

$$p_j(t+1) = \begin{cases} 
m^* + r(m^* - p_j(t)), & \text{if } p_j(t-1) \leq p_j(t) \leq m^* \text{ and } m_j(t) \geq m^* \\
m^* + (2r-1)m^* - p_j(t), & \text{if } p_j(t-1) \leq m^* \text{ and } m_j(t) \geq p_j(t) \\
m^* + (2r-1)m^* - p_j(t) - (t), & \text{if } m^* \leq p_j(t-1) \leq p_j(t) \text{ and } m_j(t) \geq p_j(t) \\
m^* + (2r-1)m^* - p_j(t) - (t), & \text{if } m^* \leq p_j(t-1) \leq p_j(t) \text{ and } m_j(t) \leq p_j(t) \\
m^* + (2r-1)m^* - p_j(t) - (t), & \text{if } m^* \leq p_j(t-1) \leq p_j(t) \text{ and } m_j(t) \leq p_j(t) 
\end{cases}$$

$$f(p_j) = \begin{cases} 
\lambda + (1 - \frac{t}{T}) \times \lambda_{\text{max}} & \text{if } t \leq T \text{ and } \lambda_{\text{max}} \text{ denotes a generally initiated with 1. The index } q \text{ of minimum fit members of arbitrarily chosen party } r \text{ is defined by:} \\
\arg\max_{p_j} \{ f(p_j) \} 
\end{cases}$$

It is noteworthy that, in the party swapping stage formulated, $p_j^*$ and $p_i^*$ are switched; but, to accurately recreate the outcomes, a slight variation is needed in party replaced method.

4.3.4. Election

The objective function is used to map the inter-party election stage. An election candidate’s popularity is measured by the number of votes he or she receives. The more votes you get, the better. The best-fitting candidate is declared the winner in each and every district. E.g., $c_j$ represents the winner of $j$th constituency. In the election stage, the constituency winners and party leaders are upgraded. It is worth noting that the party leader of $i$th party is estimated by Eq. (17).

$$q = \arg\min_{p_j \in \mathbb{R}^n} f(p_j), c_j = p_i$$

Whereas $n$ represents overall amount of constituencies, $p_j$ signifies the leader of $j$th party and $f(p_j)$ calculates the fitness of $p_j$. The constituency winner of $j$th constituency is estimated by Eq. (18).

$$q = \arg\max_{p_j \in \mathbb{R}^n} f(p_j), c_j = p_i$$

In which $n$ indicates overall amount of parties, $c_j$ symbolizes the winner of $j$th constituency and $f(p_j)$ calculates the fitness of $p_j$.

### Table 1

| Dataset                | Source | instances | attributes | $\theta$ of class | Bankrupt/Non-Bankrupt | Attributes characteristics |
|------------------------|--------|-----------|------------|-------------------|-----------------------|---------------------------|
| Australian             | UCI    | 690       | 14         | 2                 | 383/207               | Categorical, Integer, Real|
| Analcat Dataset        | stern  | 50        | 5          | 2                 | 25/25                 | Categorical, Integer, Real|
| Polish company’s dataset | UCI    | 10,503    | 64         | 2                 | 495/10008             | Real                      |
4.3.5. Parliamentary affairs

Afterward an inter-party election, the winner becomes parliamentarian and runs the government. All the constituency winners update its location regarding an arbitrarily chosen constituency winner based on:

\[
\begin{align*}
\text{c}^j(t+1) &= \text{c}^j(t) + (2\epsilon - 1)\left|\text{c}^j(t) - \text{c}^j(t)\right|
\end{align*}
\]

(19)

Let \(\text{c}^j\) be the winner of \(j\)th constituency i.e., being upgraded, \(c\) denotes a winner of an arbitrarily elected constituency, and \(\epsilon\) represent arbitrarily made in the interval of zero and one. The pseudocode for the PO algorithm is shown in Algorithm 1.

Algorithm# 1. Political Optimizer Algorithm.

```
Input ← n (number of constituencies, political parties and party members), \lambda_{\text{max}} \text{(upper limit of the party switching rate)}, T_{\text{max}} \text{(total number of iterations)}
Output ← final population \(P(T_{\text{max}})\)

Initialize \((n \times n)\) candidate members \(P\)
compute the fitness of each member \(p^j\)
compute the set of the party leaders \(P^j\) and the set of the constituency winners \(C^j\)

\(t = 1; \quad p^j(t-1) = P^j; \quad f(P(t-1)) = f(P); \quad \lambda = \lambda_{\text{max}}\)
while \(t \leq T_{\text{max}}\) do
    \(p_{\text{reg}} = p^j;\)
    \(f(p_{\text{reg}}) = f(P)\)
    foreach \(p^j \in P\) do
        foreach \(p^j \in P\) do
            \(p^j = \text{ElectoinCampaign}(p^j, p^j(t-1), p^j, c^j)\)
        end
    end
end

PartySwitching \((P, C^j)\) /* Election phase */
compute the fitness of each member \(p^j\)
compute the set of the party leaders \(P^j\) and the set of the constituency winners \(C^j\)
Parliamentary Affairs \((C^j, P^j);\)
\(p^j(t-1) = P_{\text{reg}};\)
\(f(P(t-1)) = f(P_{\text{reg}});\)
\(\lambda = (\lambda - \lambda_{\text{max}}/T_{\text{max}});\)
\(t = t + 1;\)
```

5. Experimental validation

The performance validation of the OD-PODNN technique takes place using three sets of data, Australian [37], Analects Dataset [38] and Polish companies [39]. The results are examined in terms of different evaluation parameters. The details associated with the data sets are shown in Table 1.

- **Australian dataset:** considers 690 samples with 14 input financial attributes. The dataset is associated for binary classification tasks. In this file, you’ll find information on credit card application processes. To ensure data secrecy, all attribute names and values have been replaced with meaningless symbols. Because of the good mix of features in this dataset—continuous, nominal with small values, and nominal with large values—it’s worth looking into. A few values are omitted as well. The Australian dataset contains 307 non-bankrupt samples and 383 bankrupt samples.

- **Analects dataset:** considers 50 samples with about 5 attributes. The dataset is distributed into two or binary classes for bankrupt and non-bankrupt samples. The Analecta dataset contains 25 samples for non-bankrupt and the same for bankrupt category.

- **Polish company’s dataset:** considers 10,503 samples with 64 input financial attributes. The dataset is associated for binary classification tasks. EMIS, a database of information about emerging markets around the world, was used to gather the information. During the years 2000–2012, the insolvent enterprises were examined, while the still operational firms were assessed between the years 2007 and 2013. The Polish dataset contains 10,008 non-bankrupt samples and 495 bankrupt samples.

The results discussion for the three datasets are explained in the following. Firstly, a detailed comparative results analysis of the OD-PODNN technique with existing techniques on the Australian Credit datasets is shown in Table 2.

By inspecting the results in terms of precision, the LR and RBFNetworks have attained least outcome with the lower precision of 65.79 % and 86.31 % respectively. In line with, the DeepNN model has gained slightly improved outcome with the precision of 92.43% whereas even increased precision of 96.20 % has been accomplished by the TLBO-Deep model. However, the OD-PODNN technique has resulted in a maximum precision of 98.98 %.

With examining the outcomes with respect to recall, the LR and RBF Networks have gained worse outcomes with the minimal recall of 85.23 % and 81.53 % correspondingly. Also, the DeepNN approach has reached somewhat enhanced outcomes with the recall of 87.62 % whereas even enhanced recall of 90.29 % has been accomplished by the TLBO-Deep manner. But, the OD-PODNN methodology has resulted in a maximal recall of 96.43 %.

By scrutinizing the results in terms of accuracy, the LR and RBF Networks have attained least outcome with the lower accuracy of 79.71 % and 85.71 % respectively. Besides, the DeepNN technique has gained somewhat higher outcomes with an accuracy of 91.34 % whereas even higher accuracy of 94.05 % has been accomplished by the TLBO-Deep model. Finally, the OD-PODNN approach has resulted in increased accuracy of 97.65 %.

By studying the outcomes with respect to F-score, the LR and RBF Networks have attained least outcome with the lower F-score of 74.26 % and 83.86 % respectively. In line with, the DeepNN model has gained slightly improved outcome with the F-score of 89.96 % whereas even increased F-score of 93.15 % has been accomplished by the TLBO-Deep model. Lastly, the OD-PODNN technique has resulted in a maximum F-score of 96.09 %.

By reviewing the results in terms of kappa, the LR and RBF Networks have reached least result with the minimal kappa of 70.24 % and 57.97 % respectively. In line with, the DeepNN approach has gained somewhat superior outcome with the kappa of 82.30 % whereas even superior kappa of 87.91 % has been accomplished by the TLBO-Deep manner. At last, the OD-PODNN approach has resulted in a superior kappa of 94.76 %.

Secondly, a brief comparative outcomes analysis of the OD-PODNN approach with existing methods on the Analcat dataset is shown in Table 3.

| Methods          | Precision (%) | Recall (%) | Accuracy (%) | F-score (%) | Kappa (%) |
|------------------|---------------|------------|--------------|-------------|-----------|
| OD-PODNN         | 98.98         | 96.43      | 97.65        | 96.09       | 94.76     |
| TLBO-Deep        | 96.20         | 90.29      | 94.05        | 93.15       | 87.91     |
| Deep NN          | 92.43         | 87.62      | 91.34        | 89.66       | 82.30     |
| Logistic Regression | 65.79      | 85.23      | 79.71        | 74.26       | 70.24     |
| RBF Network      | 86.31         | 81.53      | 85.21        | 83.86       | 57.97     |

The results discussion for the three datasets are explained in the following. Firstly, a detailed comparative results analysis of the OD-PODNN technique with existing techniques on the Australian Credit datasets is shown in Table 2.

By inspecting the results in terms of precision, the LR and RBFNetworks have attained least outcome with the lower precision of 65.79 % and 86.31 % respectively. In line with, the DeepNN model has gained slightly improved outcome with the precision of 92.43% whereas even increased precision of 96.20 % has been accomplished by the TLBO-Deep model. However, the OD-PODNN technique has resulted in a maximum precision of 98.98 %.

With examining the outcomes with respect to recall, the LR and RBF Networks have gained worse outcomes with the minimal recall of 85.23 % and 81.53 % correspondingly. Also, the DeepNN approach has reached somewhat enhanced outcomes with the recall of 87.62 % whereas even enhanced recall of 90.29 % has been accomplished by the TLBO-Deep manner. But, the OD-PODNN methodology has resulted in a maximal recall of 96.43 %.

By scrutinizing the results in terms of accuracy, the LR and RBF Networks have attained least outcome with the lower accuracy of 79.71 % and 85.71 % respectively. Besides, the DeepNN technique has gained somewhat higher outcomes with an accuracy of 91.34 % whereas even higher accuracy of 94.05 % has been accomplished by the TLBO-Deep model. Finally, the OD-PODNN approach has resulted in increased accuracy of 97.65 %.

By studying the outcomes with respect to F-score, the LR and RBF Networks have attained least outcome with the lower F-score of 74.26 % and 83.86 % respectively. In line with, the DeepNN model has gained slightly improved outcome with the F-score of 89.96 % whereas even increased F-score of 93.15 % has been accomplished by the TLBO-Deep model. Lastly, the OD-PODNN technique has resulted in a maximum F-score of 96.09 %.

By reviewing the results in terms of kappa, the LR and RBF Networks have reached least result with the minimal kappa of 70.24 % and 57.97 % respectively. In line with, the DeepNN approach has gained somewhat superior outcome with the kappa of 82.30 % whereas even superior kappa of 87.91 % has been accomplished by the TLBO-Deep manner. At last, the OD-PODNN approach has resulted in a superior kappa of 94.76 %.

Secondly, a brief comparative outcomes analysis of the OD-PODNN approach with existing methods on the Analcat dataset is shown in Table 3.

| Methods          | Precision (%) | Recall (%) | Accuracy (%) | F-score (%) | Kappa (%) |
|------------------|---------------|------------|--------------|-------------|-----------|
| OD-PODNN         | 100.00        | 98.95      | 98.12        | 98.61       | 96.45     |
| TLBO-Deep        | 100.00        | 92.30      | 96.00        | 96.00       | 92.01     |
| Deep NN          | 96.00         | 85.00      | 90.00        | 90.56       | 80.00     |
| Logistic Regression | 92.00      | 85.18      | 88.00        | 88.46       | 76.00     |
| RBF Network      | 80.00         | 71.42      | 74.00        | 75.47       | 48.00     |
By inspecting the results in terms of precision, the LR and RBFNet-works have gained least outcome with the lower precision of 92 % and 80 % respectively. In line with, the DeepNN system has gained somewhat enhanced outcome with the precision of 96 % whereas even increased precision of 100 % has been accomplished by the TLBO-Deep approach. But, the OD-PODNN methodology has resulted in a maximal precision of 100 %.

By inspecting the results in terms of recall, the LR and RBFNetworks have attained least outcome with the lower recall of 85.18 % and 71.42 % respectively. Similarly, the DeepNN model has attained slightly enhanced outcomes with the recall of 85 % whereas even increased recall of 92.30 % has been accomplished by the TLBO-Deep model. However, the OD-PODNN technique has resulted in a maximal recall of 98.95 %.

By reviewing the results in terms of accuracy, the LR and RBFNet-works have attained least outcome with the lower accuracy of 88 % and 74 % correspondingly. Likewise, the DeepNN model has gained slightly improved outcomes with an accuracy of 90 % whereas even increased accuracy of 96 % has been accomplished by the TLBO-Deep model. Finally, the OD-PODNN technique has resulted in a higher accuracy of 98.12 %.

By scrutinizing the results with respect to F-score, the LR and RBFNetworks have attained least outcome with the lower F-score of 88.46 % and 75.47 % respectively. Followed by, the DeepNN model has gained somewhat increased outcome with the F-score of 90.56 % whereas even increased F-score of 96 % has been accomplished by the TLBO-Deep model. But, the OD-PODNN technique has resulted in a maximum F-score of 98.61 %.

By studying the outcomes in terms of kappa, the LR and RBFNet-works have reached least outcome with the lower kappa of 76 % and 48 % correspondingly. In line with, the DeepNN model has gained somewhat improved results with the kappa of 80 % whereas even increased kappa of 92.01 % has been accomplished by the TLBO-Deep model. In addition, the OD-PODNN technique has resulted in a maximal kappa of 96.45 %.

Finally, a brief comparative outcomes analysis of the OD-PODNN approach with existing methods on the Polish company’s dataset is shown in Table 4.

By inspecting the results in terms of precision, the LR and RBFNet-works have gained least outcome with the lower precision of 92 % and 80 % respectively. In line with, the DeepNN system has gained somewhat enhanced outcome with the precision of 96 % whereas even increased precision of 100 % has been accomplished by the TLBO-Deep approach. But, the OD-PODNN methodology has resulted in a maximal precision of 100 %.

By inspecting the results in terms of recall, the LR and RBFNetworks have attained least outcome with the lower recall of 92.30 % and 85.18 % respectively. Similarly, the DeepNN model has attained slightly enhanced outcomes with the recall of 92.30 % whereas even increased recall of 95.35 % has been accomplished by the TLBO-Deep model. However, the OD-PODNN technique has resulted in a maximal recall of 97.979 %.

By reviewing the results in terms of accuracy, the LR and RBFNet-works have attained least outcome with the lower accuracy of 97.5 % and 97.12 % correspondingly. Likewise, the DeepNN model has gained slightly improved outcomes with an accuracy of 98.5 % whereas even increased accuracy of 99.3 % has been accomplished by the TLBO-Deep model. Finally, the OD-PODNN technique has resulted in a higher accuracy of 99.8 %.

By scrutinizing the results with respect to Kappa, the LR and RBFNetworks have attained least outcome with the lower kappa of 72.439 % and 68.23 % respectively. Likewise, the DeepNN model has gained slightly improved outcomes with an accuracy of 82.82 % whereas even increased accuracy of 93.23 % has been accomplished by the TLBO-Deep model. But, the OD-PODNN technique has resulted in a maximum F-score of 98.61 %.

By reviewing the results in terms of accuracy, the LR and RBFNet-works have attained least outcome with the lower accuracy of 97.5 % and 97.12 % correspondingly. Likewise, the DeepNN model has gained slightly improved outcomes with an accuracy of 98.5 % whereas even increased accuracy of 99.3 % has been accomplished by the TLBO-Deep model. Finally, the OD-PODNN technique has resulted in a higher accuracy of 99.8 %.

By scrutinizing the results with respect to Kappa, the LR and RBFNetworks have attained least outcome with the lower kappa of 72.439 % and 68.23 % respectively. Followed by, the DeepNN model has gained somewhat increased outcome with the kappa of 82.82 % whereas even increased kappa of 93.23 % has been accomplished by

| Methods     | Precision | Recall | Accuracy | F-score | Kappa  |
|-------------|-----------|--------|----------|---------|--------|
| OD-PODNN    | 98.377    | 97.979 | 99.828   | 98.178  | 97.088 |
| TLBO-Deep   | 91.828    | 95.353 | 99.381   | 93.557  | 93.233 |
| Deep NN     | 87.777    | 79.797 | 98.524   | 83.597  | 82.827 |
| Logistic Reg.| 74.947    | 72.525 | 97.562   | 73.716  | 72.439 |
| RBF Network | 69.184    | 70.303 | 97.124   | 60.799  | 68.230 |

Table 4

Performance evaluation of Polish company’s dataset dataset using various classifiers.

Fig. 7. Accuracy analysis of OD-PODNN model on Austrailan Credit, Analcat and Polish company datasets.
Fig. 8. Kappa analysis of OD-PODNN model on Australian Credit, Analcat and Polish company datasets.

Fig. 9. Accuracy graph of OD-PODNN model on Australian Credit and Analcat dataset.
the TLBO-Deep model. But, the OD-PODNN technique has resulted in a maximum kappa of 97.088 %.

By studying the outcomes in terms of F-score, the LR and RBFNet-works have reached least outcome with the lower F-score of 73.7 % and 69.7 % correspondingly. In line with, the DeepNN model has gained somewhat improved results with the F-score of 83.59 % whereas even increased F-score of 93.557 % has been accomplished by the TLBO-Deep model. In addition, the OD-PODNN technique has resulted in a maximal F-score of 98.178 %.

Fig. 7(a–c) shows the accuracy analysis of the proposed model OD-PODNN for the three different datasets Australian Credit, Analcat and Polish company. The different accuracies for different datasets prove that the proposed model outperforms other techniques.

Fig. 9(a–c) shows the different kappa for the proposed model compared with other different related models. The results show that the proposed model outperforms other different techniques.

The accuracy graph of the LSTM manner under the Australian Credit dataset is showcased in Fig. 9(a). With more epochs, the LSTM method has achieved higher validation and training accuracy. Another interesting finding is that the LSTM model has the highest validation accuracy compared to training accuracy of any model tested to date. On the Analcat Credit dataset, the LSTM model’s accuracy graph is shown in Fig. 9(b). The LSTM model’s validation and training accuracies improved as the number of epochs increased. In addition, it has been found that the LSTM model has a higher validation accuracy than training accuracy.

Loss analysis of the LSTM model on the Australian Credit dataset is shown in Fig. 10 (a). As can be seen in the graph, the LSTM approach has reduced loss while increasing epoch count. When training loss is taken into account, the LSTM technique results in a reduced amount of validation loss. Data from Analcat Credit is used to illustrate loss graph analysis using an LSTM model in Fig. 10 (b). The figure outperformed the LSTM technique, which has seen its loss decrease as the number of epochs has increased. According to the LSTM method, the validation loss is less than the training loss.

From all previous results, we can conclude the superiority of the proposed model (OD-PODNN) in the three datasets.

6. Conclusion

This study has developed an efficient OD-PODNN model to assist in determining the financial status of a firm or company. The proposed OD-PODNN technique involves several stages: data preprocessing, outlier detection, classification, and hyperparameter optimization. In addition, the OD-PODNN technique used the isolation forest (iForest) based outlier detection approach. Moreover, the PODNN-based classification model is derived with its hyperparameters to boost the overall classification accuracy. The performance validation of the OD-PODNN model was evaluated using a benchmark dataset, and the results were inspected under varying aspects. Therefore, the OD-PODNN system can be applied as a proper tool for FCP. In future work, the proposed model will be used for different datasets and try new metaheuristic algorithms for feature selection and parameter optimization.

CRediT authorship contribution statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] K. Bluwstein, M. Buckmann, A. Joseph, M. Kang, S. Kapadia, O. Simek, Credit growth, the yield curve and financial crisis prediction: evidence from a machine learning approach, 2020.
[2] T.J. Green, Big Data Analysis in Financial Markets, Colorado Technical University, 2019.
[3] S. Vadlamudi, The impacts of machine learning in financial crisis prediction, Asian Bus. Rev. 10 (3) (2020) 171–176.
[4] Y. Yang, C. Yang, Research on the application of GA improved neural network in the prediction of financial crisis, in: Proceedings of the 2020 12th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), 2020, pp. 625–9.
[5] H. Choi, H. Son, C. Kim, Predicting financial distress of contractors in the construction industry using ensemble learning, Expert Syst. Appl. 110 (2018) 1–10.
[6] S. Shahzad, A. Abid, A. Shahzad, U. Zahoor, Interactions between financial indicators and bank financial strength ratings: a case study on BRIC countries, Int. J. Manag. 11 (8) (2020).
[7] Y. Shi, X. Li, An overview of bankruptcy prediction models for corporate firms: a systematic literature review, Intang. Cap. 15 (2) (2019) 114–127.
[8] X. Song, Y. Ding, J. Huang, Y. Ge, Feature selection for support vector machine in financial crisis prediction: a case study in China, Expert Syst. 27 (4) (2010) 299–310.
[9] J. Uthayakumar, N. Metawa, K. Shankar, S.K. Lakshmanaprabu, Intelligent hybrid model for financial crisis prediction using machine learning techniques, Inf. Syst. E-Bus. Manag. 18 (4) (2020) 617–645.
[10] L. Song, Y. Zhang, Banking network structure and transnational systemic risk contagion—the case of the European Union, Financ. Res. Lett. 39 (2021), 101660.
[11] S. Sankhwar, D. Gupta, K.C. Ramya, S.S. Rani, K. Shankar, S.K. Lakshmanaprabu, Improved grey wolf optimization-based feature subset selection with fuzzy neural classifier for financial crisis prediction, Soft Comput. 24 (1) (2020) 101–110.
