Int. J. Nonlinear Anal. Appl. In Press, 1–9
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2023.31084.4561

A model for validating bank customers using multilayer perceptron neural network and imperialist competitive algorithm (ICA)

Bahman Masomian[a], Seyed Mohsen Mirhosseini[b,*]

[a]*Department of Computer Engineering, Hi. C, Islamic Azad University, Hidaj, Iran*

[b]*Department of Computer Engineering, Ka. C, Islamic Azad University, Karaj, Iran*

(Communicated by Seyed Hossein Siadati)

Abstract

Given the highly competitive nature of the banking industry, financial and credit institutions continually seek to identify the most reliable and profitable customers. They are particularly concerned about loan defaults or delays in repayment, which can negatively impact economic growth. Credit scoring models are among the most effective tools in modern banking for evaluating customer creditworthiness. These models enable banks to assess credit requests with greater accuracy and lower cost. In recent years, machine learning techniques—especially predictive classifiers—have been extensively applied to credit scoring and customer classification. This study introduces a novel hybrid model that combines a Multilayer Perceptron (MLP) neural network with the Imperialist Competitive Algorithm (ICA). In the proposed approach, ICA is employed to optimise the hyperparameters of the MLP network. The model is tested on a dataset comprising 2,571 real customers from Saderat Bank, categorised into default and non-default classes based on 11 identified features. The results demonstrate that the proposed model achieves higher accuracy and lower prediction error in assessing customers' credit behaviour.

Keywords: Credit Risk, Credit Scoring, Probability of Default, Multilayer Perceptron Neural Network, Imperialist Competitive Algorithm
2010 MSC: 68T07

1 Introduction

In the financial system, credit is considered more valuable than gold. In addition, the high probability of default (PD) shakes the market confidence, which ultimately leads to its structural collapse [10]. Therefore, default control defends the financial system against financial risks by creating an appropriate credit scoring model through Predictive Classifiers. Moreover, lenders' losses can be minimised in a highly competitive market if the PD of borrowers is correctly predicted through a credit scoring model. Credit risk evaluation is usually based on credit scoring models, which are widely used in evaluating the default probability of loan applicants. The key issue in credit risk evaluation is how to classify loan applicants into two main groups: default and non-default. The appraiser may decide to reject

*Corresponding author
Email addresses: b.masomian@iau.ac.ir (Bahman Masomian), mohsen.mirhosseini@iau.ac.ir (Seyed Mohsen Mirhosseini)

Received: June 2023 *Accepted:* August 2023

or accept the loan application after completing the credit evaluation process. Credit scoring has attracted a lot of attention in the financial industry due to its role in credit risk management. Decades of research in credit scoring have led to some classic predictive models that have been developed to measure financial conditions and the probability of default. According to the efficient market hypothesis [4] in complex economic environments, classic credit scoring models fail to accurately predict the probability of default in the real world. Rapid advances in artificial intelligence and especially machine learning have led to the emergence of a solution to the above problem. Credit score prediction is a key aspect of risk management in financial institutions. The purpose of credit scoring is basically to classify applicants for credit facilities into two categories: good payers and bad payers. The challenging aspect of credit scoring is that the classes are highly imbalanced since a small percentage of bank customers show bad behaviour in repaying their received facilities. It is acknowledged that training highly unbalanced data models is a challenging task for machine learning algorithms. A wide variety of classification techniques have already been proposed in the credit scoring literature to deal with the specific aspect of unbalanced credit risk data, including statistical techniques such as Linear Discriminant Analysis (LDA), logistic regression, and nonparametric models such as k-nearest neighbour and decision trees.

Recently, neural network techniques have been increasingly used in a wide range of applications that were traditionally performed using statistical methods. These networks are trained by using data changes over time and modifying their weights in such a way that the output obtained from these methods is very close to the real output with the least difference. Research conducted by [5] demonstrates that artificial neural networks are a viable alternative to the discriminant analysis (LDA) and logistic regression methods for determining the credit ranking of customers, especially in situations where dependent and discrete variables have complex and nonlinear relationships. It has been reported that neural network models have higher accuracy compared to LDA and logistic regression models. Hyperparameters are parameters in a deep neural network that are set before the training process begins and cannot be learned from the data. They affect the behaviour and performance of the network during training and can significantly impact its effectiveness. Hyperparameter tuning in deep neural networks is a crucial step to optimise the performance of the model. Deep neural networks often have a large number of hyperparameters that can significantly impact the model's training process and final performance. Hyperparameter optimisation finds the optimal value of several hyperparameters, resulting in an optimised model that minimises the value of a predefined objective function on experimental data [14].

The Imperialist Competitive Algorithm (ICA) [2] is a method in evolutionary computing that addresses finding the optimal answer to various optimisation problems. This algorithm provides an algorithm for solving mathematical optimisation problems by mathematical modelling of the socio-political evolution process. In terms of application, this algorithm is included in the category of evolutionary optimisation algorithms. Like all the algorithms in this category, the ICA also forms an initial set of possible answers. Initial answers in the ICA are known as countries. The ICA gradually improves these initial answers (countries) and finally provides the appropriate answer to the optimisation problem (desirable country). By imitating the process of the social, economic, and political evolution of countries and by mathematically modelling parts of this process, this algorithm presents operators in a regular format as an algorithm that can help solve complex optimisation problems.

It is difficult to design a model for predicting credit risk due to a large number of missing values, high-dimensional data, and an unbalanced class. This article presents a hybrid model to predict the probability of repayment of the facilities provided to the customers of Saderat Bank, so that, according to the mentioned limitations and also the goals of the said bank, a suitable method can be used to measure the customers' credit and the repayment of the granted facilities. In the hybrid method, a Multilayer Perceptron (MLP) neural network is used to classify customers into two categories, good and bad, and the ICA is used to determine the optimal value of hyperparameters. The reason for using the Imperialist Competitive meta-heuristic algorithm is to find the optimal values of the hyperparameters of the proposed model; in effect, each of these hyperparameters can take different values, and therefore, there will be a large number of states for different values of each hyperparameter. Deterministic optimisation will find it difficult to find optimal values. Accordingly, the main questions of this article are as follows:

- Is a deep multi-layer perceptron neural network suitable for customer validation?
- Can determining the optimal values of hyper-parameters of a deep perceptron neural network by the ICA algorithm increase the efficiency of this network in validating bank customers?

2 Theoretical foundations of research

2.1 Credit risk

The possibility of non-repayment or late payment of principal and sub-facilities granted by banks is called credit risk. Designing a model for measuring and grading credit risk was done for the first time in 1909 by John Murray on bonds [8]. The great similarity of bank credit facilities to bonds caused the credit risk scoring of bank facilities, which means measuring the risk of not repaying the principal and interest of loans, to be taken into consideration by some researchers. In general, the risks that affect financial institutions can be classified into three levels as follows [13]:

- The first level is the risks that the financial institution has no control over and is only affected by, including government risk, policies, and economic, social, and natural cycles.
- The second level is the risks that the financial institution influences, but this influence is small and is more influenced by, as legal risk, reputation, and competition.
- The third level is the risks that affect the financial institution, but the financial institution can control and manage them by applying methods and tools, including credit risk, market risk, and operational liquidity.

The financial institution can only overcome and control third-level risks by employing risk management methods and tools. Meanwhile, credit risk is considered the most important risk due to its centrality, volume of operations, and especially its sensitivity. The risk of default or the possibility of not repaying the debt by the borrower is a loss that threatens the bank in the event of default. Therefore, credit risk is rooted in the possibility of default or non-repayment of the facility by the facility recipient, and the probability of its occurrence fluctuates between zero and one. Payment default is declared by a banking institution when the scheduled instalments are not made within a certain period after the due date. The default can be completely economic, that is, when the economic value of assets, which is the current value of expected future cash flows, becomes less than the value of unpaid debts [3]. The loss caused by default depends on the definition of default, and the definition of default also depends on the estimate of the probability of default (derived from past data). Credit scoring agencies consider the event of default when three months pass from the due date of a scheduled payment and no payment is made during this period, but the theoretical models of credit risk, which are after Merton's model [7] use the definition of economic default to measuring the number of losses. It is worth mentioning that various default events do not necessarily cause immediate losses, but they increase the probability of permanent default or bankruptcy. Default risk is measured by the probability that default will occur during a certain period. Of course, the default probability cannot be measured directly, but the past default statistics collected from within the credit system should be used. The financial institution can only overcome and control third-level risks by employing risk management methods and tools. Meanwhile, credit risk is considered the most important risk due to its centrality, volume of operations, and especially its sensitivity.

2.2 Imperialist Competitive Algorithm (ICA)

The Imperialist Competitive Algorithm (ICA) was presented by Atashpaz and Lucas in 2007. The ICA is a new algorithm in evolutionary calculations, which is based on the socio-political evolution of humans. Like other evolutionary algorithms, the ICA starts with a number of random initial populations, each of which is called a country. Some of the best countries are selected as Imperialist and the rest of the population are considered as colonies. In optimisation problems, considering the function f(x), it is tried to find the argument x in such a way that the cost corresponding to it is optimal. In an optimisation problem with Nvar variables, a country is represented by an Nvar×1 array, which is defined as equation 2.1.

$$country = [p_1, p_2, p_3, \dots, p_{Nvar}] \tag{2.1}$$

The cost of a country with the evaluation function f is obtained by the variables $(p_1, p_2, p_3, \dots, p_{Nvar})$ and through equation 2.2:

$$cost = f(country) = f(p_1, p_2, p_3, \dots, p_{Nvar}) \tag{2.2}$$

In the ICA, the number of N-countries of the initial country is created, and Nimp is chosen as the Imperialist from the best members of this population (countries with the best value of the cost function). The rest of the N-col countries form colonies, each of which belongs to an empire. The colonising countries, by applying the policy of assimilation

(assimilation) in line with the different axes of optimisation, pull the colonised countries towards them. According to their power, the Imperialists pull these colonies towards them. The total power of each empire is determined by calculating the power of both parts of it, i.e. the power of the Imperialist country, plus a percentage of the average power of its colonies. The imperialistic policy of attraction and competition forms the core of this algorithm. According to the policy of attraction, the imperialistic countries tried to alienate the colonised country by using methods such as destroying the colonised country's language, culture, and customs. In presenting this algorithm, this policy is done by moving the colonies of an empire towards it, according to a specific relationship. If, during the movement, a colony reaches a better position than the imperialist, the place of the two will be exchanged. Imperialistic competition forms another important part of this algorithm. During the imperialistic competition, the weak empires gradually lost their power and disappeared. The imperialistic competition leads to a state where there is only one empire in the world. This is the state when the ICA algorithm stops. Figure 1 shows the pseudocode of the imperialistic competitive algorithm.

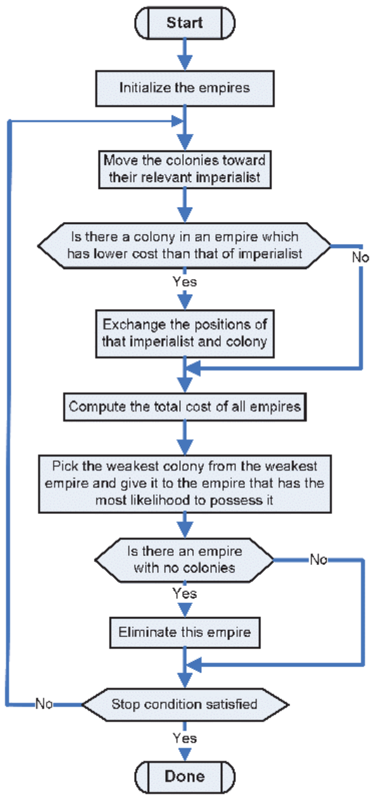

Figure 1: Figure 1. Flowchart of the ICA algorithm

2.3 Multilayer Perceptron Neural Network

Artificial neural networks are composed of simple processing elements called neurons. Neurons are simple computing units that receive several weighted inputs (can include a bias input) and can return a single output through an activation function. The idea of a single neuron can be extended to a Multilayer Perceptron (MLP) neural network by adding multiple layers containing multiple neurons to this network, where each neuron processes its inputs and produces an output value that is then shared with all Neurons in the next layer. The basic structure of a Multilayer Perceptron (MLP) neural network has a hidden layer and an output layer. To calculate the output of the hidden neuron i, in a network with one hidden layer, we process the weighted inputs and the bias term $b_i^{(1)}$ as follows:

$$h_i = f^{(1)}(b_i^{(1)} + \sum_{j=1}^{m} w_{ij} x_i) \tag{2.3}$$

where, W is the weight matrix and w_{ij} represents the connection weight of input j to hidden unit i. Similarly, the last layer output is calculated as follows:

$$y = f^{(2)}(b^{(2)} + \sum_{j=1}^{m_h} v_j h_j) \tag{2.4}$$

where m_h is the number of hidden neurons and v is the weight vector, and v_j is the weight that connects the hidden unit j to the output neuron. Finally, the network can model non-linear relationships in the data using activation functions $f^{(1)}$ and $f^{(2)}$. The sigmoid and RELU are two frequently used activation functions in deep neural networks.

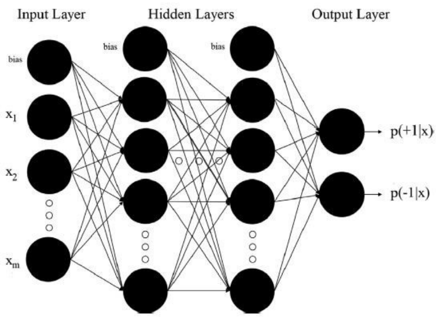

Figure 2: Figure 2. Multilayer Perceptron (MLP) Neural Network

Advances in research on neural networks and increased computing performance led to the creation of large networks with multiple hidden layers that make deep learning possible. The effect of this is that the network can propagate the weights through the network and thus can learn complexities in large data sets using multiple processing layers with complex structures. Figure 2 presents an example of a Multilayer Perceptron (MLP) Neural Network. This network is made using several layers of neurons connected with simple activation functions. In classification problems, a softmax function can be used as an activation function on the neurons located in the output layer. To classify (predict), the network uses the class of the output layer, which returns the neuron with the highest probability as the predicted class. The difference between the probability vector returned by the output layer and the actual label vector can be quantified as an error. The error rate determines how the weights are adjusted during network training.

3 Literature review

Much research has been done on the construction of credit scoring models. Various classification models have been used in this field. Logistic regression and decision trees are among the most used models in credit scoring, and it has been proven that both models are simple and efficient. More sophisticated machine learning techniques, such as support vector machines and neural networks, have also been widely used for credit scoring. In addition, recently, the use of hybrid methods that combine the advantages of different classifiers is also expanding rapidly. For example, [1] developed a multiple-classifier system that uses neural networks, support vector machines, decision trees, and simple Bayesian as base models. The proposed hybrid approach improved the prediction performance over other benchmarked classifiers on more than five real credit-scoring datasets. In recent years, due to the remarkable ability of deep learning, some studies have started to use deep learning algorithms for credit risk management. [6] designed a Deep Belief Network (DBN) for credit scoring. They compared it with a support vector machine, logistic regression, and multilayer perceptron on a credit default exchange dataset. The results showed that the DBN has the best performance among the compared methods. [12] proposed a boosted deep network that combines genetic programming and deep learning. The proposed model had the best accuracy against other traditional machine learning methods. [15] used a hybrid method by combining a Relief Algorithm (which performs feature selection) with a convolutional neural network for credit scoring. The combined relief-CNN model showed better performance in comparison to logistic regression and random forest. Multilayer perceptron networks and deep convolutional neural networks were used by [9] to evaluate the credit value of applicants. The deep convolutional network outperformed the deep perceptron network in the experiments. [11] conducted a comparative study of four multi-objective evolutionary algorithms and compared their performance. The experiment results confirmed the effectiveness of the evolutionary algorithm proposed by the authors in creating classification rules for credit risk analysis.

4 Methodology

This is an applied paper. The statistical population of the present article includes the real customers of Saderat Bank in Tehran, who have been examined with access to the data of 2571 subjects from 2020 to 2022, of which 2000 subjects are good customers and 571 subjects are bad customers. Multi-layer perceptron neural networks can perform complex classifications by using a sufficient number of layers and an appropriate number of neurons in each layer. Therefore, according to the problem investigated in this research, which is a type of classification problem, multi-layer perceptron neural networks were adopted to classify bank customers and, as a result, determine credit risk.

4.1 Independent and dependent variables of the research

The first step in evaluating and solving a problem is having a correct definition of the problem and identifying the effective factors. Thus, domestic and foreign valid articles were reviewed to select the variables affecting the research. Furthermore, the research independent variables were selected from among the variables that accounted for the credit scoring of the real customers of the banks in the literature. The articles on the real customers of banks' credit scoring have adopted various variables, and in some models, up to 50 variables have been identified with their significance. However, almost most of the research concluded that 10 to 15 variables significantly affect the probability of customer default. In this paper, we adopted 11 variables, including age, gender, education, loan amount, loan type, customer income, collateral type, the average balance of bank accounts, number of loans taken, repayment period, and type of installment as independent variables and the customer's credit status, i.e. whether it is good or bad, is considered as a dependent variable. In this research, independent and dependent variables were used as the input and the output of the neural network.

4.2 Details of the proposed method

Figure 3 shows the overall flowchart of the proposed method. The proposed method consists of two main steps. In the first step, the network configuration process is done by determining the optimal value of the model hyperparameters, and the model is prepared and trained in the second stage. Unlike model parameters that are learned during training by machine learning algorithms, hyperparameters are external variables in the network configuration whose value is determined before the start of learning. The values of hyperparameters have a great impact on the speed and performance of the network. The number of hidden layers, learning rate, activation function, batch size, number of epochs, initial weights, and optimiser algorithm are among these hyperparameters in deep neural networks. The process of setting hyperparameters is a serious challenge. For this purpose, we have used the ICA algorithm. There are 11 neurons in the first layer of the used neural network, equal to the number of features selected from customers. This network has four hidden layers. The first and second layers have 11 neurons. The third layer has five neurons, and the fourth layer has two neurons. In addition, there is a neuron in the output layer, in which the sigmoid activation function is used to classify customers. The training rate is equal to 0.0005, and the number of repetitions of network training is considered equal to 10. After optimising the hyperparameters of the perceptron neural network using by ICA algorithm, the model is trained using through training data.

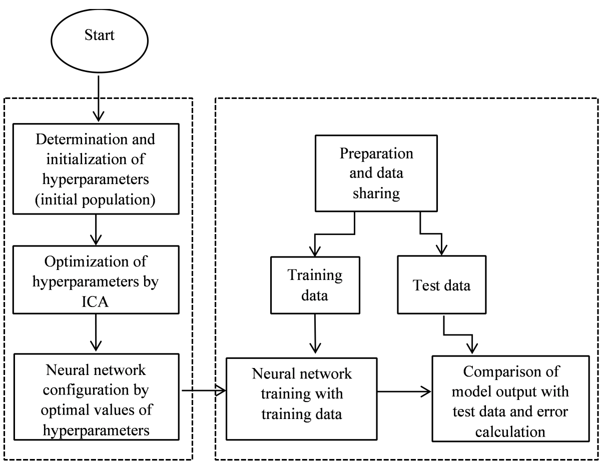

Figure 3: Figure 3. General flowchart of the proposed method

The the proposed method algorithmic steps are shown below:

1: Initialization of hyperparameter values
2: Optimizing the values of hyperparameters by imperialist competitive algorithm
3: Building a model of multi-layer perceptron neural network using optimal values of hyperparameters
4: Data normalization
5: Classifying the data into 70% training data and 30% test data
6: Model training using training data
7: Evaluating the model using evaluation criteria

5 Evaluation and results

5.1 Evaluation criteria

A model should be evaluated to ensure its efficiency. The purpose of evaluating the model is to estimate the selected model error, that is, to what extent the selected model works on unseen data. A good model not only works well on the trained data, but can also work well on unseen data. Therefore, we must make sure that the efficiency of the model does not decrease in the face of new input data. In this research, accuracy, precision and false detection rate criteria were used to evaluate the proposed method. How to calculate these criteria is given in relations (5) to (7).

$$Accuracy = (TP + TN)/(TP + FP + FN + TN) \tag{5.1}$$

$$Precision = TP/(TP + FP) \tag{5.2}$$

$$FalseAlarmRate(FAR) = FP/(FP + TN) \tag{5.3}$$

where, TP, FP, TN and FN are called true positive (TP), false positive (FP), true negative (TN) and false negative (FN), respectively. A true positive is when a customer is correctly identified as a good customer. A false positive is when a bad customer is mistakenly predicted to be a good customer. In a true negative, a customer is correctly predicted as bad customer, and in a false negative, a customer is mistakenly considered as a bad customer.

5.2 Experiments results

The proposed method was implemented using Python. For the implementation, the Keras package with TensorFlow Backend was used on a GeForce GTX 1070 graphics card with 8 GB of RAM. The data set was divided into two parts, training and testing, respectively, with a ratio of 70 to 30, and the model was trained ten times using them. The initial value of the number of countries in the ICA algorithm is considered equal to 100. The number of empires is equal to 5, and the number of iterations of the algorithm is considered equal to 80. In the experiments conducted to evaluate the proposed method (DNN + ICA), this method was compared with C5.0 decision tree, support vector machine (SVM), and the perceptron neural network with random values of hyperparameters (DNN). Table 1 shows that the proposed method has better results than the other compared methods in all three evaluation criteria.

Table 1: Table 1. The results of the experiment

Method	Accuracy	Precision	False Alarm Rate
C5.0	88.42	73.59	34.85
SVM	89.76	75.34	30.82
DNN	91.69	79.15	25.57
DNN+ICA	96.74	89.11	10.51

For a better comparison of the results, we used from bar chart in Figure 4.

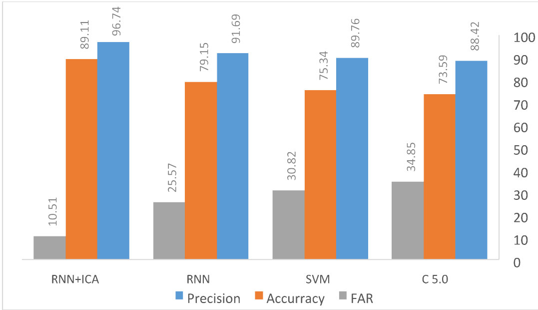

Figure 4: Bar chart diagram of experiment results

Answer to the first research question: The Support Vector Machine is a widely used method in bank customer classification with proven efficiency in previous research. Comparing the results of the perceptron neural network with this method and the decision tree method, which is one of the most widely used techniques in classification, shows that the hypothesis of the efficiency of the multilayer perceptron neural network is confirmed.

Answer to the second research question: The proposed method adopts the ICA algorithm to select the optimal value of hyperparameters, so it has better results than the perceptron neural network with random values of hyperparameters. Therefore, the hypothesis is confirmed that using the imperialistic competitive algorithm to determine the hyperparameters' optimal values can be effective.

6 Conclusion

Increasing profitability is one of the main goals of banks and any other financial organisations. Achieving this goal comes with increased risk. Credit risk management is critical to the bank. One of the most important tools for assessing risk is a system that identifies customers who are likely to default by identifying important and influential factors. Credit scoring models, based on the quantitative and qualitative criteria received from customers and using the current and past information of the applicant, evaluate the probability of non-repayment of the loan by him and give him a score. In this article, an acceptable model was presented for the credit scoring of bank customers using the information available in the credit files of the real customers of Saderat Bank. For this purpose, the effective features in the credit scoring of the customers of the banks were identified following a literature review. After identifying the influencing features, the appropriate neural network model was selected for customer credit scoring. For this purpose, the multilayer perceptron neural network was chosen, and the imperialist competitive algorithm was also used to determine the optimal values of the hyperparameters of this network. The experiments to evaluate the proposed method showed that the proposed model has superior performance compared to other methods and proved its effectiveness.

References

[1] M. Ala'raj and M. Abbod, *Classifiers consensus system approach for credit scoring*, Knowledge-Based Syst. **104** (2016), 89–105.

[2] E. Atashpaz-Gargari and C. Lucas, *Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition*, IEEE Cong. Evolut. Comput., 2007, pp. 4661–4667.

[3] J. Bessis, *Risk Management in Banking*, John Wiley & Sons, 2011.

[4] M. Faheem, *AI-Driven risk assessment models: Revolutionizing credit scoring and default prediction*, Iconic Res. Engin. J. **5** (2021), 177–186.

[5] T. Lee, C. Chiu, C. Lu, and I. Chen, *Credit scoring using the hybrid neural discriminant technique*, Expert Syst. Appl. **23** (2002), 245–254.

[6] C. Luo, D. Wu, and D. Wu, *A deep learning approach for credit scoring using credit default swaps*, Engin. Appl. Artific. Intell. **65** (2017), 465–470.

[7] R. Merton, *On the pricing of corporate debt: The risk structure of interest rates*, J. Finance **29** (1974), 449–470.

[8] G. Morton, *Managing Bank Risk: An Introduction to Broad Base Credit Engineering*, Academic Press, Amsterdam, 2003.

[9] V. Neagoe, A. Ciotec, and G. Cucu, *Deep convolutional neural networks versus multilayer perceptron for financial prediction*, Int. Conf. Commun., 2018, pp. 201–206.

[10] M. Peihani, *Basel committee on banking supervision: A primer on governance, history, and legitimacy—Part i*, Brill Rese. Perspect. Int. Bank. Secur. Law **1** (2016), no. 1, 1–87.

[11] M. Soui, I. Gasmi, S. Smiti, and K. Ghédira, *Rule-based credit risk assessment model using multi-objective evolutionary algorithms*, Expert Syst. Appl. **126** (2019), 144–157.

[12] K. Tran, T. Duong, and Q. Ho, *Credit scoring model: A combination of genetic programming and deep learning*, Future Technol. Conf., 2016, pp. 145–149.

[13] H. Van Greuning and S. Bratanovic, *Analyzing Banking Risk: A Framework for Assessing Corporate Governance and Risk Management*, World Bank Publications, 2020.

[14] M. Vazan, *Machine Learning and Data Science: Foundations, Concepts, Algorithms, and Tools*, Miaad Andisheh Pub., 2021. [In Persian]

[15] B. Zhu, W. Yang, H. Wang, and Y. Yuan, *A hybrid deep learning model for consumer credit scoring*, Int. Conf. Artific. Intell. Big Data, 2018, pp. 205–208.