University of Tehran Press

Industrial Management Journal

Online ISSN: 3115 - 7386

Homepage: https://imj.ut.ac.ir

Designing an Optimization-Simulation Model for Credit Scoring and Loan Structuring Using a Memetic Algorithm: A Case Study of Corporate Banking Clients

Amir Khorrami [1], **Mahmoud Dehghan Nayeri** [2] ✉, **and Ali Rajabzadeh** [3]

1. Ph.D. Candidate, Department of Industrial Management, Faculty of Management and Economic, Tarbiat Modares University, Tehran, Iran. E-mail: amir.khorrami@modares.ac.ir
2. Corresponding author, Associate Prof., Department of Industrial Management, Faculty of Management and Economic, Tarbiat Modares University, Tehran, Iran. E-mail: mdnayeri@modares.ac.ir
3. Prof., Department of Industrial Management, Faculty of Management and Economic, Tarbiat Modares University, Tehran, Iran. E-mail: alirajabzadeh@modares.ac.ir

Article Info

Article type:
Research Article

Article history:
Received June 06, 2024
Received in revised form December 15, 2024
Accepted January 29, 2025
Available online June 10, 2025

Keywords:
credit risk, credit scoring, classification, memetic algorithm, optimization-simulation model

ABSTRACT

Objective: This paper introduces a groundbreaking optimization-simulation model, a novel approach that promises to revolutionize credit scoring and loan optimization for banks.

Methods: The proposed approach follows a three-stage framework: data preparation, credit scoring, and optimization simulation. In the data preparation stage, corporate client data, including bank loan information and financial statements, has been collected and processed to define and calculate relevant features. The credit scoring stage involved meticulous feature selection using the correlation method, followed by the rigorous training and testing of five classification methods: logistic regression (LR), K-nearest neighbors (KNN), artificial neural network (ANN), adaptive boosting (AdaBoost), and random forest (RF). Model performance has been evaluated using accuracy, F1-score, and area under the curve (AUC) to identify the most effective classifier. In the optimization-simulation stage, the Memetic Algorithm (MA) has been utilized to optimize loan characteristics, including loan size, interest rate, and repayment period, while minimizing the rate of loan defaults. Additionally, this stage incorporated the pre-trained credit scoring model to estimate the impact of loan characteristics on default probabilities.

Results: A case study was conducted using data from 1,000 corporate clients of Bank Tejarat. The optimization-simulation approach has successfully reduced the loan default rate from 33% to below 5%, a significant achievement that underscores its potential to mitigate banks' credit risk. This shows the effectiveness of the proposed method in reducing credit risk for banks. Additionally, the AdaBoost technique achieved the best performance among the credit assessment models.

Conclusion: The optimization-simulation approach combines determining the optimal loan specifications with the credit assessment process. This approach considers the impact of loan characteristics on the likelihood of customer default and utilizes this information to reduce banks' credit risk.

Cite this article: Khorrami, A., Dehghan Nayeri, M., & Rajabzadeh, A. (2025). Designing an optimization-simulation model for credit scoring and loan structuring using a memetic algorithm: A case study of corporate banking clients. ***Industrial Management Journal***, *17*(2), 27-57. https://doi.org/10.22059/imj.2025.370552.1008119

Publisher: University of Tehran Press.

DOI: https://doi.org/10.22059/imj.2025.370552.1008119

Introduction

An efficient and effective banking system is a necessary tool for the economic growth of any country (Soui et al., 2019). Among banking activities, lending is particularly important (Cheng & Qu, 2020). While providing loans offers benefits such as contributing to economic growth, it also introduces credit risk. Credit risk refers to the possibility that a borrower will fail to repay loan installments due to financial inability or unwillingness (Malik & Thomas, 2010; Ahmed & Rajaleximi, 2019). The risk of loan default consistently poses a significant challenge for banks and financial institutions, negatively impacting their performance and the financial system of countries (Bluhm et al., 2016). Although credit risk cannot be eliminated, it should be managed as effectively as possible (Dehghan Nayeri et al., 2021). Banks employ various methods to account for risk in their lending processes. The most appropriate approach involves measuring credit risk, and credit scoring—also known as credit rating, evaluation, or assessment—is the most widely adopted method (Bluhm et al., 2016; Ahmed & Rajaleximi, 2019). According to Thomas et al. (2017), credit scoring is the process of evaluating bank clients based on predefined criteria to determine their creditworthiness. Credit scoring has become especially crucial for managing and mitigating credit risk in contemporary banking.

The purpose of credit scoring methods is to predict the probability of default, that is, the failure of borrowers to repay loan installments on time (Ahmed & Rajaleximi, 2019). The general concept of credit scoring methods is to compare the characteristics and properties of a loan application with those of previous borrowers. If a client's characteristics are sufficiently similar to those who have defaulted on their loans, the loan application is rejected. Conversely, if the characteristics are acceptably similar to clients who have not defaulted, the loan application is accepted (Witzany & Witzany, 2017).

Historically, credit scoring has been based on two primary approaches: the judgmental method and the credit scoring model approach (Abdou & Pointon, 2011). In the judgmental approach, a professional credit analyst reviews and decides on loan applications, using their experience and subjective judgment and incorporating traditional methods such as the 5C criteria (Koutanaei et al., 2015). Despite the strengths of the judgmental approach, particularly in considering qualitative aspects and incorporating the credit analyst's experience, this approach has significant drawbacks. These include the inconsistency of decisions due to individual attitudes and preferences of decision-makers, the potential for inadvertent errors, and the time-consuming nature of the evaluation process. To address these shortcomings, credit scoring models have been proposed (Abdou & Pointon, 2011; Dastile et al., 2020).

Credit scoring models provide an automated and efficient assessment of loan applications, offering a sharp contrast to subjective judgmental approaches. These models leverage extensive

client databases, collecting necessary data upon receiving a new application. The applicant's creditworthiness is then quantified using a specific scoring model and compared against a predefined threshold, categorizing them as either a "good" or "bad" credit risk. Loan approvals are typically restricted to those classified as "good." A key advantage of credit scoring models is their inherent consistency—by eliminating human judgment, they ensure uniform decisions across comparable cases (Abdou & Pointon, 2011; Ahmed & Rajaleximi, 2019).

From a computational perspective, credit scoring models utilize classification methods to predict client labels—categorizing them as either "good" or "bad" credit risks. Over the past three decades, these models have evolved significantly.

Initially, statistical learning techniques such as linear discriminant analysis (LDA) and logistic regression (LR) gained prominence due to their efficiency, simplicity, and high interpretability (Dastile et al., 2020; Markov et al., 2022). The rapid advancement of computing tools in the 1990s, combined with pivotal events like the Basel Accord and the global financial crisis, fueled research interest in credit scoring models (Markov et al., 2022). This surge in attention facilitated the development of machine learning approaches, including decision trees (DT), artificial neural networks (ANN), and support vector machines (SVM) (Koutanaei et al., 2015; Dastile et al., 2020).

More recently, ensemble learning techniques—such as random forest (RF), extreme gradient boosting (XGBoost), and adaptive boosting (AdaBoost)—have attracted significant interest due to their enhanced predictive performance (Dastile et al., 2020; Markov et al., 2022).

Literature Background

The rapid advancements in credit scoring models over recent decades have led to a substantial increase in published research on the topic, both in Iran and globally. Tables 1 and 2 present key credit scoring methods identified in recent Persian and English scholarly articles, respectively.

Table 1. List of recent studies on credit scoring (in Persian)

Credit Scoring Method	Client Type	Previous Research
Logistic Regression (LR)	Individual	(Farbad & Mohammadi, 2014; Baharlou et al., 2016; Ghasemnia Arabi & Safaei Ghadikolaei, 2019; Mir et al., 2022; Moradi et al., 2022)
	Corporate	(Taghvi et al., 2008; Nili & Sabzevari, 2008; Parvizian et el., 2009; Mehrara et al., 2009)
Other Regression Methods	Individual	(Mehrara et al., 2009; Mir et al., 2022; Moradi et al., 2022; Koohi & Gholami, 2012; Rastegar & Eidi Goosh, 2020)
Data Envelopment Analysis (DEA)	Corporate	(Koohi & Gholami, 2012; Izadikhah & Shamsi, 2020; Izadikhah et al., 2021)
Artificial Neural Network (ANN)	Individual	(Farbad & Mohammadi, 2014; Ghasemnia Arabi & Safaei Ghadikolaei, 2019; Moradi et al., 2022; Hosseini & Zibaei, 2015; Maleki et al., 2020)
	Corporate	(Mehrara et al., 2009; Dadmohammadi & Ahmadi, 2017; Jandaghi et al., 2020)

Credit Scoring Method	Client Type	Previous Research
Decision Tree (DT)	Individual	(Farbad & Mohammadi, 2014; Khorrami et al., 2020; Kazemi et al., 2022)
K-nearest Neighbors (KNN)	Individual	(Kazemi et al., 2022)
Clustering	Individual	(Najibi & Mokhatab Rafiei, 2017; Moradi et al., 2022)
Random Forest (RF)	Individual	(Kazemi et al., 2022)
Genetic Algorithm (GA)	Individual	(Eghbali et al., 2017; Daneshvar et al., 2021)

Table 2. List of recent studies on credit scoring (in English)

Credit Scoring Method	Client Type	Previous Research
Logistic Regression (LR)	Individual	(Dželihodžić et al., 2018; Munkhdalai et al., 2019; Nalic & Martinovic, 2020; Boughaci et al., 2020; Biecek et al., 2021; Laborda & Ryoo, 2021; Hussin Adam Khatir & Bee, 2022; Lenka et al., 2022; Sum et al., 2022; Runchi et al., 2023; Aljadani et al., 2023; Gambacorta et al., 2024)
	Corporate	(Nehrebecka, 2018; Boughaci et al., 2020; Qadi et el., 2023)
Naïve Bayes (NB)	Individual	(Nalic & Martinovic, 2020; Boughaci et al., 2020; Tripathi et al., 2021; Hussin Adam Khatir & Bee, 2022; Lenka et al., 2022; Runchi et al., 2023)
	Corporate	(Boughaci et al., 2020)
K-nearest Neighbors (KNN)	Individual	(Boughaci et al., 2020; Laborda & Ryoo, 2021; Tripathi et al., 2021; Ampountolas et al., 2021; Hussin Adam Khatir & Bee, 2022; Lenka et al., 2022; Runchi et al., 2023; Aljadani et al., 2023)
	Corporate	(Boughaci et al., 2020)
Decision Tree (DT)	Individual	(Dželihodžić et al., 2018; Nalic & Martinovic, 2020; Tripathi et al., 2021; Ampountolas et al., 2021; Hussin Adam Khatir & Bee, 2022; Lenka et al., 2022; Runchi et al., 2023; Aljadani et al., 2023; Mushava & Murray, 2024)
Artificial Neural Network (ANN)	Individual	(Dželihodžić et al., 2018; Munkhdalai et al., 2019; Boughaci et al., 2020; Tripathi et al., 2021; Ampountolas et al., 2021; Hussin Adam Khatir & Bee, 2022; Sum et al., 2022; Aljadani et al., 2023)
	Corporate	(Boughaci et al., 2020)
Support Vector Machine (SVM)	Individual	(Munkhdalai et al., 2019; Nalic & Martinovic, 2020; Boughaci et al., 2020; Laborda & Ryoo, 2021; Tripathi et al., 2021; Lenka et al., 2022; Sum et al., 2022; Runchi et al., 2023; Rofik et al., 2024)
	Corporate	(Nehrebecka, 2018; Boughaci et al., 2020; Zhou, 2022)
Random Forest (RF)	Individual	(Munkhdalai et al., 2019; Biecek et al., 2021; Tripathi et al., 2021; Laborda & Ryoo, 2021; Ampountolas et al., 2021; Hussin Adam Khatir & Bee, 2022; Lenka et al., 2022; Runchi et al., 2023; Aljadani et al., 2023; Rofik et al., 2024; Xiao et al., 2024)
	Corporate	(Qadi et el., 2023)
Adaptive Boosting (AdaBoost)	Individual	(Boughaci et al., 2020; Ampountolas et al., 2021; Runchi et al., 2023; Aljadani et al., 2023)
	Corporate	(Boughaci et al., 2020; Qadi et el., 2023)
Extreme Gradient Boosting (XGBoost)	Individual	(Munkhdalai et al., 2019; Biecek et al., 2021; Ampountolas et al., 2021; Lenka et al., 2022; Runchi et al., 2023; Aljadani et al., 2023; Rofik et al., 2024; Krishna et al., 2024)
	Corporate	(Qadi et el., 2023)
Bagging	Individual	(Dželihodžić et al., 2018; Boughaci et al., 2020)
	Corporate	(Boughaci et al., 2020)
Deep Learning (DL)	Individual	(Munkhdalai et al., 2019; Xiao et al., 2024; Talaat et al., 2024)

Table 1 and Table 2 illustrate the diverse range of credit scoring methods employed in prior research to evaluate the creditworthiness of individual and corporate clients. In these studies, a set of features represented client attributes to predict their credit status. For individual clients, credit scoring methods utilized personal characteristics (e.g., age, gender, marital status) alongside financial characteristics (e.g., housing status, average account balance, income). Similarly, for corporate clients, company characteristics (e.g., history, field of activity, ownership status), and data extracted from financial statements, particularly financial ratios, were key factors in the credit assessment process. Furthermore, characteristics related to previously granted loans are likewise relevant to both individual and corporate clients. Key examples, widely used in numerous studies (including Rastegar & Eidi Goosh, 2020; Kazemi et al., 2022; Nalic & Martinovic, 2021; Hussin Adam Khatir & Bee, 2022), include loan size (amount), interest rate, and loan age (repayment period). Integrating loan-related characteristics with other individual/corporate client data can enhance the effectiveness of credit scoring methods in distinguishing between good and bad clients.

The majority of features employed in credit scoring models (e.g., age, gender, occupation for individual clients field of activity, and financial statements for corporate clients) are predetermined and beyond the bank's control. Consequently, the bank's primary action is to utilize these features within the scoring model to assess a client's creditworthiness and determine loan eligibility. However, a limited set of features within these models are variable from the bank's perspective and can be modified, specifically the characteristics of the loans granted. These adjustable features, relevant to both individual and corporate clients, include loan size, interest rate, and loan age.

Given that banks can adjust loan characteristics within certain limits, treating these parameters as fixed in the credit scoring process is logically inadequate. Since loan characteristics influence the probability of client default, a more effective approach would be to consider them as variables and explore their modification to minimize default risk. Our review of prior research indicates that this perspective has not yet been widely adopted in the field of credit scoring.

Notably, in optimal lending (loan portfolio optimization) studies (e.g., Metawa et al., 2016, 2017; Wang, 2021; Basu et al., 2023; Rong et al., 2023; Soltani et al., 2023; Yan et al., 2024), loan characteristics are treated as adjustable variables. However, these studies typically assume fixed and predetermined client credit statuses, thereby overlooking the potential impact of loan characteristics on creditworthiness.

Against this backdrop, the substantial impact of varying loan characteristics on client default probabilities has largely been overlooked in the research literature. The widespread adoption of credit scoring models emerged in response to escalating credit risk faced by banks and the financial losses associated with loan defaults. Given the critical role of credit scoring in financial

institutions' risk management, advancing the modeling framework with innovative solutions is imperative. This study seeks to mitigate client default probability—and, consequently, reduce banks' credit risk—by uniquely integrating loan characteristics as variables within the credit scoring process.

In terms of credit scoring models, our study closely aligns with the work of Ampountolas et al. (2021) and Adam Khatir & Bee (2022), as both studies employed all five models utilized in this research: LR, KNN, ANN, AdaBoost, and RF. Additionally, our approach exhibits significant similarities with the studies of Laborda & Ryoo (2021) and Runchi et al. (2023), as four of our five selected models overlap with their research. However, the primary distinction of the present study lies in the type of bank clients examined. While previous studies have focused exclusively on individual borrowers, this research uniquely investigates corporate clients.

The primary distinctions among the aforementioned studies (Ampountolas et al., 2021; Laborda & Ryoo, 2021; Adam Khatir & Bee, 2022; Runchi et al., 2023) largely stem from variations in their datasets, leading to differences in the characteristics and features considered within the credit scoring process. Additionally, these studies vary in their feature selection methods and the specific credit scoring models employed. Regarding client type, the research conducted by Boughaci et al. (2020), and Qadi et al. (2023) are the most closely related to the present study, as both examined corporate clients and employed largely similar credit scoring models. However, a critical distinction remains: all reviewed studies treated the characteristics of granted loans as fixed constants and did not explore the impact of adjusting these characteristics on client default probability. To the best of our knowledge, this study is the first to investigate how modifications in loan characteristics influence default rates among corporate clients.

This paper has presented an optimization-simulation model designed to simultaneously evaluate bank clients using credit scoring models and determine the optimal characteristics of their loans. The approach has begun with applying classification methods for credit scoring, followed by an analysis of how variations in loan characteristics influence client default probability through simulation. A memetic optimization algorithm has then been employed to identify the optimal loan parameters that minimize the bank's default rate.

To validate the proposed model, a case study has been conducted using data collected from corporate clients of a commercial bank in Iran. The subsequent section outlines the materials and methods used in this research, followed by a presentation and analysis of the case study results. Finally, key findings are discussed, and conclusions are drawn.

Materials and Methods

The flowchart in Figure 1 visually represents the framework of the proposed optimization-simulation model for credit scoring, which incorporates a memetic algorithm. As depicted in this figure, the model comprises three main stages: data preparation, credit scoring models, and the optimization-simulation model. First, during the data preparation stage, data has been collected, features have subsequently been defined and calculated based on this data, and data preprocessing has been performed. Next, in the credit scoring stage, relevant features have been selected, and the data has been divided into training and test subsets. Each credit scoring model (LR, KNN, ANN, AdaBoost, and RF) has then been trained and evaluated on the test data. The best-performing model has been selected for the subsequent stage based on predefined classification criteria. During the optimization-simulation stage, loan characteristics have been refined through iterative optimization using the memetic algorithm, while the selected credit scoring model has simulated client default probabilities. The following subsections will introduce and explain each stage in detail.

Data preparation

To evaluate the proposed method, data from thousands of corporate clients of Tejarat Bank in Iran who successfully obtained loans between 2017 and 2021 has been collected. This information has been provided anonymously by the bank for research purposes. In addition to data on the loans received, financial statement information for these clients has additionally been collected. The bank loan data included the type of loan, loan size (amount), loan interest rate, loan age (repayment period), and loan default status. The financial statement information comprised the companies' balance sheets and profit and loss statements for the year 2021.

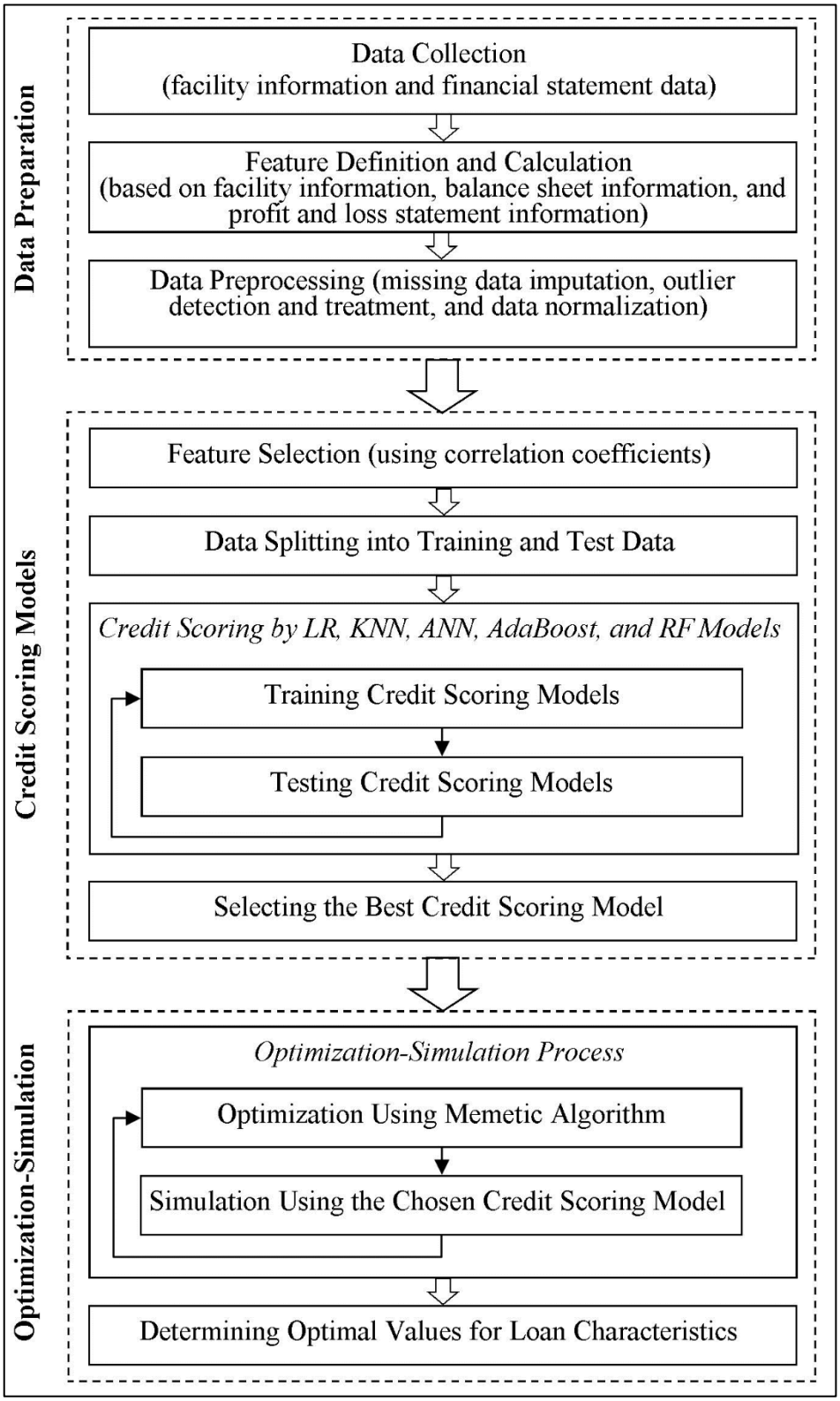

Figure 1. Overall framework of the proposed method

Feature definition

Based on the collected data, features have been defined and calculated for the credit scoring models. The data label for each client's default status has been defined as binary (0/1). Following Nehrebecka (2018), a loan has been considered in default if a client failed to pay their installments for three months, for any reason. Analysis of the data revealed that 415 out of the 1000 corporate clients (41.5%) had defaulted. Consistent with prior research (e.g., (Ampountolas et al., 2021)), the logarithm of the loan size has been calculated and used instead of the original value. In addition to loan size, the features included loan type (integer-coded), annual interest rate (%), and loan age (repayment period in years). Consequently, a total of four features have been derived from the bank loan information.

Financial statement information is valuable for assessing a company's overall condition, strengths, and weaknesses. To facilitate inter-company comparisons, financial ratios (Groppelli & Nikbakht, 2000) have commonly been employed. Based on the available data, 34 financial ratios have been calculated, encompassing liquidity, activity (efficiency), leverage (debt), and profitability ratios, along with other common metrics. Furthermore, the logarithm values of six key balance sheet items included for the clients: total current assets, total assets, total current liabilities, total liabilities, total equity, and capital. Consequently, a total of 40 characteristics have been derived from the financial statement information.

Data preprocessing

- To guarantee data quality before incorporating features into the credit scoring models, the following preprocessing steps were carried out:Missing Data Replacement: The collected raw data contained no missing values. However, instances of division by zero occurred during the calculation of financial ratios. These values, representing approximately 2% of the total data, have been replaced using the mean imputation method (Yang et al., 2021; Lenka et al., 2022).
- Outlier Detection and Treatment: Outliers have been detected and treated using the interquartile range (IQR) method (Acuna & Rodriguez, 2004; Dawson, 2011). Extreme outliers have been defined as data points below the lower limit (Q1 - 3 * IQR) or above the upper limit (Q3 + 3 * IQR), where Q1 and Q3 represent the first and third quartiles, respectively. These outliers, comprising about 3% of the total data, have been replaced with the values of the corresponding lower or upper limits.
- Data Normalization: To mitigate the adverse effects of varying feature ranges on the performance of credit scoring models, data normalization (scaling) has been performed.

The min-max normalization method has been employed to map the feature values to the range [0-1] (Yang et al., 2021; Deng et al., 2022).

Credit scoring models

Feature selection

As previously described, 44 features have been defined and calculated for use in the credit scoring models, comprising 4 loan characteristic features and 40 financial statement features. These features serve as the independent variables in the classification problem. The dependent variable, which the credit scoring models aim to predict, is the data label (default status). Employing a large number of features in classification models can reduce efficiency through overfitting and increase computational time. Therefore, it is crucial to select a subset of features with a greater capacity to estimate the probability of client default. In this study, the correlation coefficient method (Laborda & Ryoo, 2021) has been applied for feature selection.

The point-biserial correlation coefficient has been computed individually for each of the 44 features against the data label (default status). For each feature, if the probability (p-value) has been below the significance level ($\alpha = 0.05$), the relationship between the feature and default status has been deemed statistically significant. Consequently, these significant features have been selected for use in the credit scoring models. However, recognizing the crucial role of the primary loan characteristics—loan size, interest rate, and loan age—within the optimization-simulation model, these features have been included in the credit scoring models irrespective of the correlation coefficient test outcomes.

Classification methods

This study has employed five distinct classification methods for credit scoring, which are briefly introduced below.

Logistic Regression (LR): Logistic regression, a specific case of the generalized linear model (GLM), operates on principles similar to linear regression, with the key difference being its binary dependent variable. If a client defaults, the dependent variable is assigned a value of one; otherwise, it is zero. Logistic regression utilizes the logit function to model the relationship between the independent variables (features, denoted as X) and the dependent variable (data labels, denoted as Y). The probability of a data point belonging to class one (default) has been calculated using the following equation (Laborda & Ryoo, 2021):

$$\Pr(Y = 1|X) = p(X) = \frac{e^{\beta_0+\beta_1 X_1+\cdots+\beta_m X_m}}{1 + e^{\beta_0+\beta_1 X_1+\cdots+\beta_m X_m}} \quad (1)$$

The variable $p(X)$ represents the probability of default for a given client, ranging from zero to one. Consequently, a default threshold of 0.5 is commonly used to categorize clients into good and bad credit groups (Louzada et al., 2016; Laborda & Ryoo, 2021).

K-Nearest Neighbor (KNN): The fundamental principle of the KNN method posits that data points with similar input features tend to have similar output values. In this method, for a given data point, its distance to all other points in the dataset is calculated based on their feature vectors. Subsequently, the K data points with the smallest distances are identified as the nearest neighbors. The probability of the data point belonging to a specific class (good or bad credit) is then determined by calculating the majority class among these K nearest neighbors (Laborda & Ryoo, 2021; Adam Khatir & Bee, 2022).

Artificial Neural Network (ANN): Inspired by the human brain's functioning, artificial neural networks exhibit significant flexibility in solving classification problems. A widely utilized type is the Multilayer Perceptron Network (MLP), which comprises at least three layers: an input layer, one or more hidden layers, and an output layer. In the input layer, each feature corresponds to a neuron, while the output layer typically consists of a single neuron for binary classification. The hidden layers enable the network to model complicated relationships within the data. The mathematical connections between neurons in different layers are established through weights. Training the neural network involves determining these weights to minimize classification error and create an effective mapping between the input feature vector and the output data label (Koutanaei et al., 2015; Adam Khatir & Bee, 2022).

Adaptive Boosting (AdaBoost): AdaBoost is an ensemble learning method that trains a set of weak learners and combines their outputs through weighted averaging to produce the final prediction. These weak learners are typically simple models, such as linear discriminants or decision trees. A key advantage of AdaBoost over other boosting techniques is its ability to incorporate the errors of preceding learners when training subsequent ones. To achieve this, higher weights are assigned to the data points that were misclassified by the previous learner. Consequently, through an iterative process, the ensemble of learners is trained to minimize the overall model error (Koutanaei et al., 2015; Boughaci & Alkhawaldeh, 2020).

Random Forest (RF): Random Forest, another ensemble learning method, comprises a collection of decision trees that are aggregated to form the final prediction. These individual trees are intentionally diverse. To prevent over-reliance on specific features, each tree is trained on a random subset of the feature set. Furthermore, the data used to train each tree also varies; this is commonly achieved through resampling techniques such as bootstrapping. As a result of these differences in feature and data sampling, each tree is trained uniquely and generates distinct predictions. The final prediction of the random forest is then determined by aggregating these

individual tree predictions, often through majority voting (Laborda & Ryoo, 2021; Adam Khatir & Bee, 2022).

Evaluation criteria

To assess the performance of the credit scoring models, the dataset has been randomly partitioned into two subsets: a training set (70% of the data) and a test set (30% of the data). The models have been trained on the training data, and their performance has been subsequently evaluated on the test set. The evaluation process for each model has been initiated with the calculation of the confusion matrix. This matrix considers all four possible predicted outcomes for the data label (class): true positive (TP), false positive (FP), true negative (TN), and false negative (FN). Based on the values within the confusion matrix, the following criteria have been computed and used to evaluate the credit scoring models (Dastile et al., 2020; Adam Khatir & Bee, 2022):

- Precision: Indicates the proportion of instances the model identified as positive (defaulted) that were positive (defaulted).
- Recall: Indicates the proportion of actual positive instances (defaulted) that were correctly identified as positive by the model.
- F1-score: A harmonic mean of precision and recall, representing the model's overall ability to distinguish between actual positive and actual negative instances.
- Accuracy: Indicates the proportion of instances the model correctly identified as either positive (positive) or negative (negative).
- Area Under the Curve (AUC): Represents the classification model's performance based on the Receiver Operating Characteristic (ROC) curve, illustrating the extent to which the true positive rate (TPR) exceeds the false positive rate (FPR) in identifying positive instances.

Optimization-simulation model

Following the credit scoring stage, the best-performing model would be selected based on the aforementioned criteria. This selected model could then be integrated into the optimization-simulation process. This section would firstly provide a brief introduction to the memetic algorithm, followed by an explanation of how both the memetic algorithm and the chosen credit scoring model have been utilized within the optimization-simulation framework.

Memetic algorithm

The Genetic Algorithm (GA), first introduced by Holland (1973), is a well-established metaheuristic optimization method. Its fundamental principles are inspired by natural selection in biological evolution. Despite its innovative nature and significant impact on optimization

research, the original GA often exhibits suboptimal performance in numerous real-world optimization problems. To address these limitations, various enhanced versions of the genetic algorithm have been developed, broadly categorized as evolutionary algorithms (Sörensen & Sevaux, 2006). Among these improved algorithms, the Memetic Algorithm (MA) stands out as particularly successful.

The concept of the memetic algorithm stems from the idea of memes, which, in contrast to genes, can adapt based on environmental conditions. Moscato (1989) first employed the meme concept and introduced the memetic algorithm as a variant of the genetic algorithm, enhanced by the capability of individual learning for local enhancement. In the memetic algorithm, the local search operator is integrated with conventional genetic operators, thereby increasing the algorithm's efficiency and mitigating the risk of premature convergence (Neri & Cotta, 2012; Kumar & Memoria, 2020).

Model procedure

- The optimization-simulation model has been designed to identify the optimal loan characteristics for clients, aiming to minimize default rates and, in turn, reduce the bank's credit risk. The key components of this model are as follows: Optimization: Recognizing that loan characteristics influence client default rates, the optimal values for these characteristics can be identified to minimize the probability of client default. To this end, the primary loan characteristics—size, interest rate, and age—have been considered as decision variables in the optimization problem. The objective function has been to minimize the expected number of loan defaults, and the memetic algorithm has been employed to solve this problem.
- Simulation: Directly observing the impact of altering loan characteristics on a client's future default status is infeasible in the real world. To overcome this logical and practical limitation, the model has utilized a simulation approach to estimate the probability of client default. This simulation has been performed by invoking a pre-trained credit scoring model. Specifically, after modifying the loan characteristics (i.e., loan size, age, and interest rate) for a client, these updated characteristics have been input into the pre-trained credit scoring model to obtain a new estimate of that client's default status.

The variables of the optimization problem within the optimization-simulation model have been defined as follows: The first variable has been loan size (amount), a continuous variable represented logarithmically. The second and third variables have been loan interest rate and loan age, which are discrete variables. Their values have been selected from the predefined sets offered by the bank under study.

Given that the credit scoring models had been trained on the training data, the optimization-simulation model was applied exclusively to the test data, comprising 300 clients, to ensure a fair evaluation. The problem variables have been represented in chromosomes as follows: each chromosome consisted of 300 segments, with each segment corresponding to a single client (where i denotes the client index and n represents the total number of clients.). Within each segment, three genes have encoded the three loan characteristic variables for that client. Consequently, each chromosome has contained a total of 900 genes, organized as follows:

$$[\mathrm{LS}_1, \mathrm{IR}_1, \mathrm{LA}_1, \dots, \mathrm{LS}_i, \mathrm{IR}_i, \mathrm{LA}_i, \dots, \mathrm{LS}_n, \mathrm{IR}_n, \mathrm{LA}_n] \quad (2)$$

Where $\mathrm{LS_i}$, $\mathrm{IR_i}$, and $\mathrm{LA_i}$ represent the loan size, interest rate, and loan age of client i, respectively. For instance, a segment [10,4,5] for a client indicates a loan with a logarithmic size of 10 (equivalent to 10 billion Rials), an interest rate of 4%, and a repayment period of 5 years. When generating the initial population of chromosomes, the gene values have been randomly assigned according to these rules: for the first variable (loan size), the lower and upper bounds have been 9 and 13, respectively. For the second variable (loan interest rate), the feasible values have been {4, 18} percent. For the third variable (loan age), the feasible values were {1, 2, 3, 5, 10} years, based on the guidelines of the bank under study.

In the memetic algorithm, a single-point crossover has been employed, where the values of a randomly determined single gene from two selected chromosomes have been substituted to generate two new chromosomes. The roulette wheel method has been used for chromosome selection during crossover. The mutation operator has additionally been implemented as a single-point operation, where a randomly selected gene in the chromosome has its value randomly changed to form a new chromosome. Furthermore, the local search operator implemented rules similar to the mutation operator, with the distinction that after each new chromosome generation, its fitness has been compared to the previous one. If no improvement is achieved, the operation continued until a better solution is found or the maximum allowed repetitions are reached. The fitness function (objective function) aimed to minimize the number of client loan defaults. The fitness evaluation has been performed by invoking the pre-trained credit scoring model to estimate the default status of clients based on their updated loan characteristics.

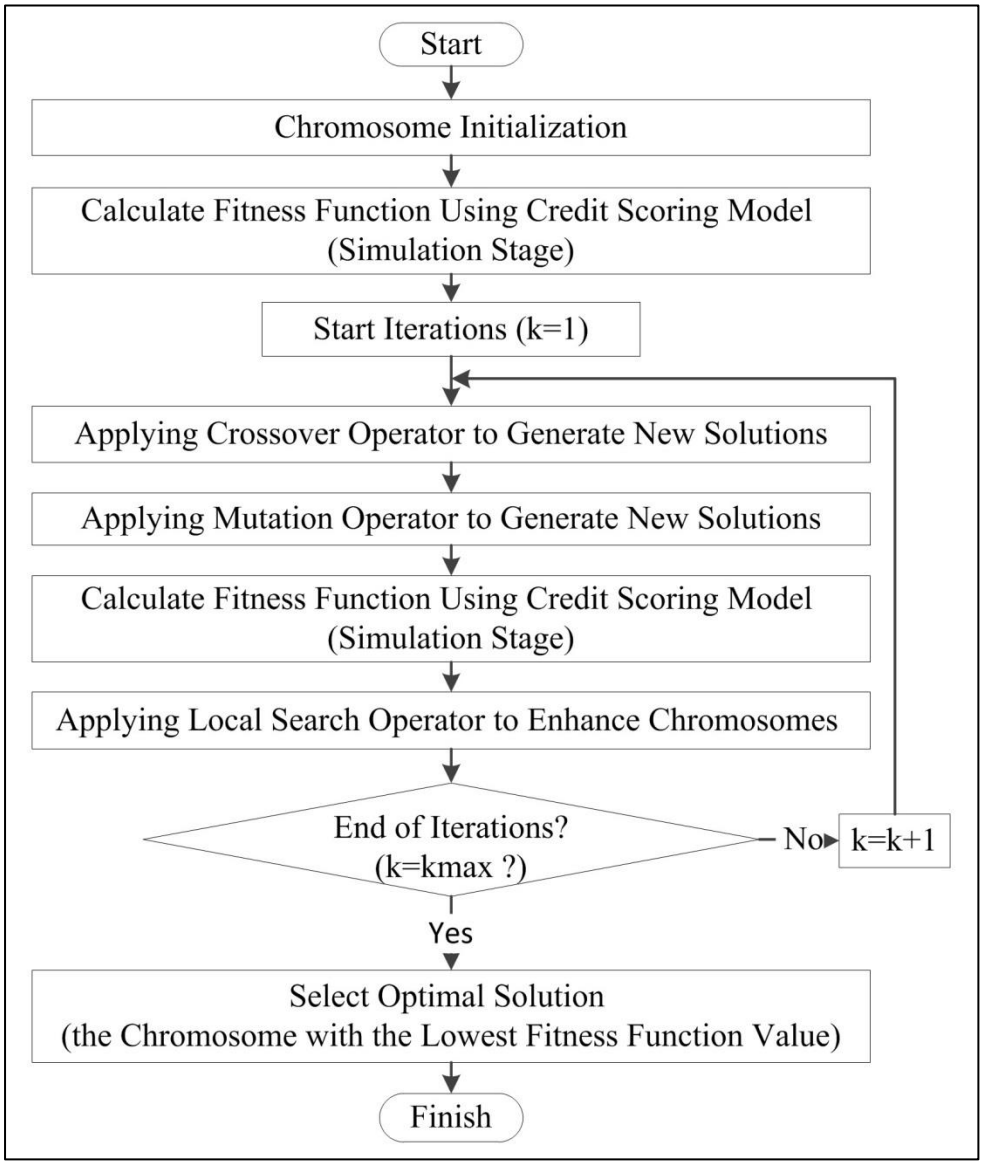

Figure 2. Implementation stages of the optimization-simulation model

To summarize the explanations provided in this section, the flowchart of the optimization-simulation model is shown in Figure 2. The algorithm begins by initializing the chromosomes based on their permissible values. Then, for each chromosome, the fitness function has been calculated using the pre-trained credit scoring model—a process referred to as the simulation stage. Next, the main loop of the algorithm commenced. In each iteration, new chromosomes (solutions) have been generated through crossover and mutation operators, followed by the calculation of their fitness functions. The local search operator has then been applied to refine the solutions. After completing the algorithm's iterations, the chromosome with the lowest objective function value has been selected as the optimal solution.

To examine the impact of each loan characteristic on client default status and the bank's credit risk, the optimization-simulation model has been implemented across four distinct cases, and the results were subsequently compared:

- Case 1: Loan age and interest rate were fixed; loan size was variable.
- Case 2: Loan size and age were fixed; interest rate was variable.
- Case 3: Loan size and interest rate were fixed; loan age was variable.
- Case 4: Loan size, interest rate, and age were all variable.

Results

A case study has been conducted using data collected from 1000 corporate clients of Tejarat Bank in Iran. Banking loan information and financial statement data, including balance sheets and profit and loss statements, have been gathered and analyzed. Accordingly, 44 features have been defined, calculated, and preprocessed.

For credit scoring, five classification models have been implemented using MATLAB's Statistical and Machine Learning Toolbox. The dataset has been randomly split into a training set (70%) and a test set (30%). In the KNN method, the number of neighbors (k) has been set to 7, and the distance metric between points has been Euclidean distance. In the ANN method, a multilayer perceptron network with two hidden layers containing 20 and 10 neurons, respectively, has been employed. The network has been trained using the Levenberg-Marquardt (LM) error propagation algorithm. In the AdaBoost method, linear discriminant learners have been applied with the number of training cycles set to 30 and the learning rate set to 0.1. In the RF method, the maximum number of decision splits has been set to 15, the minimum observations per leaf has been set to 2, and the number of training cycles has been set to 20.

In the memetic algorithm, the population size has been set to 30 chromosomes, the maximum number of iterations to 300, the crossover rate to 80%, the mutation rate to 20%, and the maximum number of local search attempts to 30. All data preparation and analysis have been performed using MATLAB software (version 2019a).Selected features

To identify the features exhibiting a significant relationship with the data label (default status), a correlation test has been conducted. The results of this analysis are presented in Table 3.

Table 3. Results of the correlation test for feature selection.

Feature	Correlation Coefficient	p-value	Status	Feature	Correlation Coefficient	p-value	Status
Total Current Assets	-0.0726	0.0217	✔	Shareholders' Equity to Total Assets Ratio	-0.1125	0.0004	✔
Total Assets	-0.0658	0.0376	✔	Shareholders' Equity to Total Liabilities Ratio	-0.1173	0.0002	✔
Total Current Liabilities	-0.0570	0.0718	-	Shareholders' Equity to Fixed Assets Ratio	-0.1356	0.0000	✔
Total Liabilities	-0.0483	0.1267	-	Debt Coverage Ratio	-0.0510	0.1067	-
Equity	-0.0831	0.0086	✔	Net Profit Margin	-0.0111	0.7254	-
Capital	-0.1001	0.0015	✔	Operating Profit Margin	-0.0169	0.5927	-
Current Ratio	-0.0370	0.2425	-	Gross Profit Margin	-0.0561	0.0763	-
Quick Ratio	-0.0088	0.7816	-	Return on Equity	0.0342	0.2804	-
Cash Ratio	-0.0749	0.0178	✔	Return on Assets	-0.0140	0.6594	-
Current Assets to Total Assets Ratio	-0.0146	0.6444	-	Return on Working Capital	-0.0368	0.2446	-
Inventory Turnover Period	-0.0383	0.2258	-	Return on Fixed Assets	-0.0747	0.0181	✔
Inventory to Working Capital Ratio	-0.0750	0.0177	✔	Equity to Debt Ratio	-0.0998	0.0016	✔
Working Capital Turnover	-0.0488	0.1231	-	Working Capital to Total Assets Ratio	-0.0384	0.2245	-
Fixed Asset Turnover	-0.0651	0.0395	✔	Retained Earnings to Total Assets Ratio	-0.0587	0.0634	-
Total Asset Turnover	0.0003	0.9934	-	Altman Z-Score	-0.0681	0.0314	✔
Debt Ratio	0.0849	0.0073	✔	Cost of Goods Sold to Revenue Ratio	0.0048	0.8797	-
Equity Ratio	-0.0805	0.0109	✔	Logarithm of Working Capital	-0.0534	0.0914	-
Debt-to-Equity Ratio	0.0878	0.0055	✔	Return on Investment	-0.0122	0.7005	-
Total Debt to Net Worth Ratio	0.0391	0.2162	-	Loan Amount/Size	0.0039	0.9030	✔
Current Debt to Net Worth Ratio	0.0160	0.6125	-	Loan Type	-0.0572	0.0708	✔
Long-Term Debt to Net Worth Ratio	0.0202	0.5240	-	Loan Interest Rate	0.0332	0.2940	✔
Fixed Assets to Net Worth Ratio	0.0303	0.3378	-	Loan Repayment Period	-0.0537	0.0894	✔

Each feature's correlation was considered significant if its p-value fell below the significance threshold ($\alpha = 0.05$). As shown in Table 3, 16 financial statement features met this criterion and were selected for inclusion in the credit scoring models. Among the four loan features, none exhibited statistical significance. However, since this study's methodology is centered on loan characteristics, these features were also incorporated. Consequently, a total of 20 features were selected for training and testing the credit scoring models.

Results of credit scoring models

Training results

For each credit scoring model, the Receiver Operating Characteristic (ROC) curve for the training data is presented in Figure 3, with the corresponding Area Under the Curve (AUC) value specified for each model. In these curves, the vertical axis represents the True Positive Rate (TPR), while the horizontal axis denotes the False Positive Rate (FPR). AUC values serve as key performance indicators, quantifying the extent to which the TPR exceeds the FPR. As shown in Figure 3, all five credit scoring models demonstrated strong performance, with AUC values exceeding 0.70.

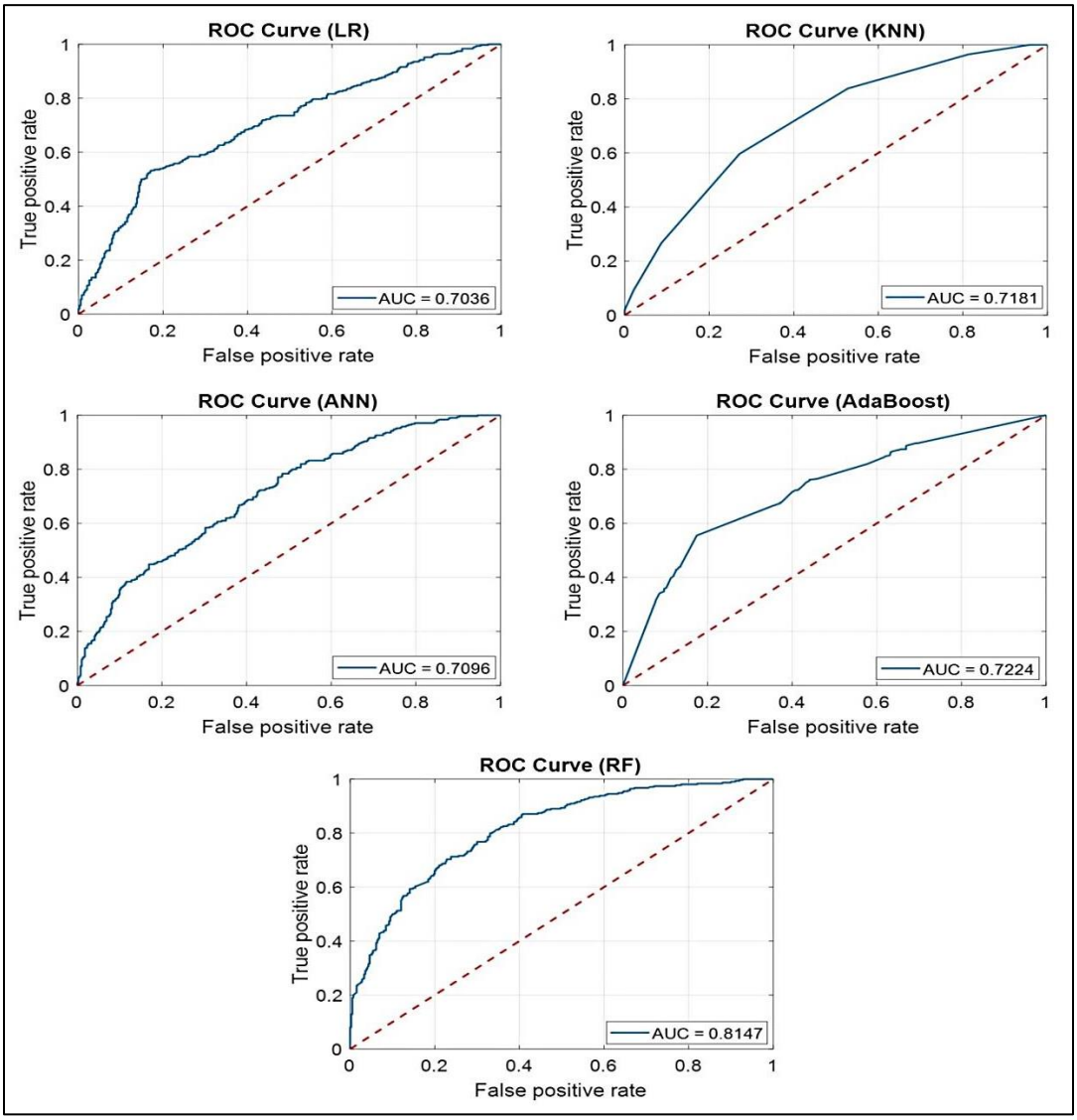

Figure 3. Resulting ROC curves of the credit scoring models (training data)

The evaluation criteria for the credit scoring models applied to the training data are summarized in Table 4, with the best performance for each criterion highlighted in bold. As shown in the table, precision and recall values exceeded 0.50 for almost all models. Given the variability in precision and recall across models, the F1-score— which integrates both metrics— is also presented. All models achieved an F1-score above 0.60, with the ANN method demonstrating the best performance at 0.6337.

Similarly, all models exhibited accuracy above 0.60, with the RF method delivering the highest accuracy at 0.7271. Additionally, AUC values across all models surpassed 0.70, with the RF method achieving the highest value at 0.8147. These results confirm that all models performed acceptably according to the defined evaluation criteria, indicating that underfitting (poor model fit) did not occur.

Table 4. Results of credit scoring models on training data

Model	Precision	Recall	F1-Score	Accuracy	AUC
Logistic Regression (LR)	0.6917	0.5355	0.6036	0.6886	0.7036
k-Nearest Neighbors (KNN)	0.6357	0.5968	0.6156	0.6700	0.7181
Artificial Neural Network (ANN)	0.5642	**0.7226**	**0.6337**	0.6300	0.7096
Adaptive Boosting (AdaBoost)	0.7167	0.5548	0.6255	0.7057	0.7224
Random Forest (RF)	**0.8182**	0.4935	0.6157	**0.7271**	**0.8147**

Test results

For each credit scoring model, the ROC curve for the test data is presented in Figure 4. As shown, all models achieved AUC values of approximately 0.60 or higher, indicating acceptable performance. Among the models, LR and AdaBoost demonstrated the strongest results.

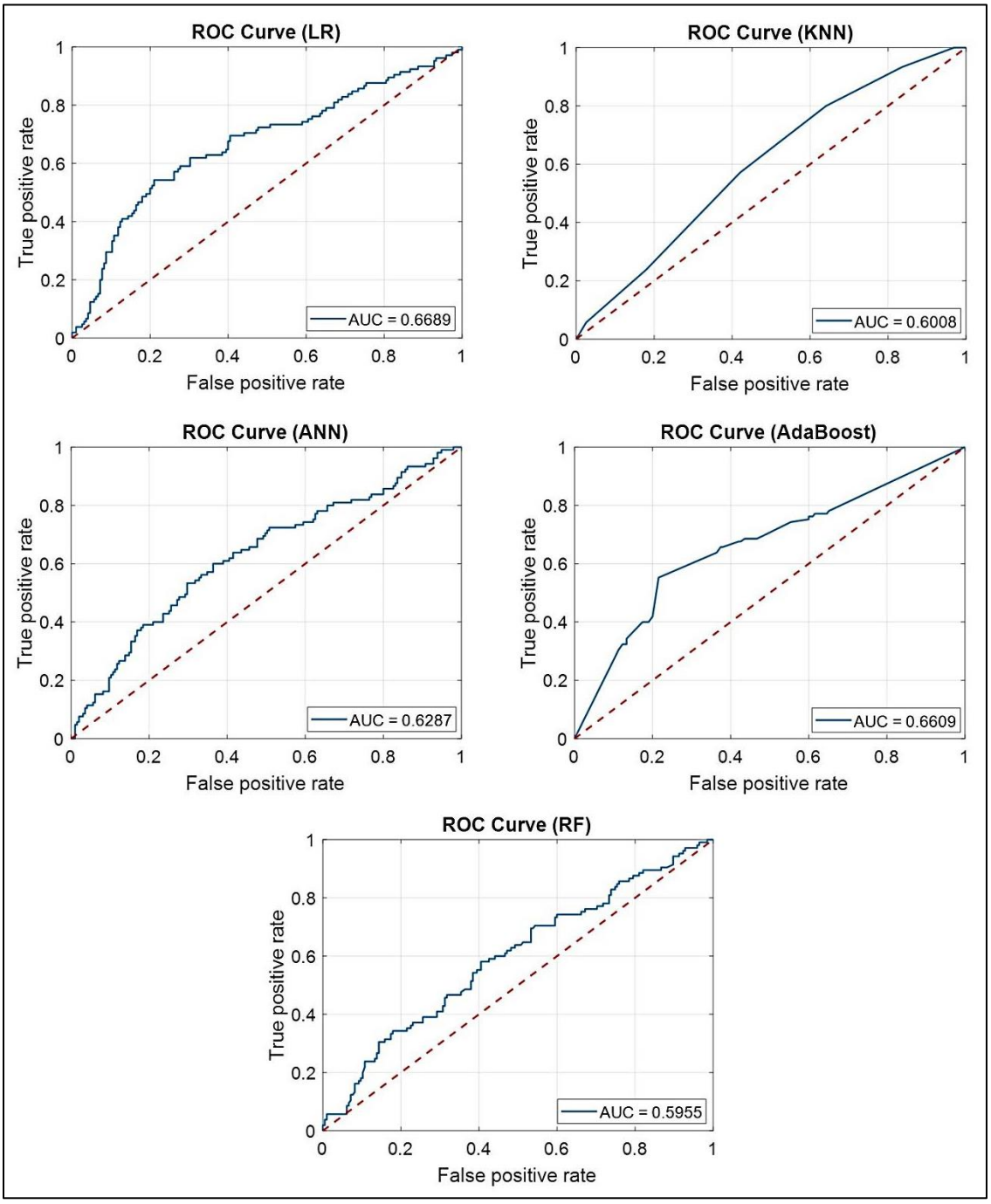

Figure 4. Resulting ROC curves of the credit scoring models (test data)

The evaluation criteria for the credit scoring models on the test data are summarized in Table 5. In terms of precision and recall, the AdaBoost and ANN methods demonstrated the best performance, respectively. The F1-score, which integrates both metrics, indicates that AdaBoost outperformed the other models overall. Additionally, AdaBoost achieved the highest accuracy

among the tested models. Regarding the AUC criterion, while the LR method recorded the highest value, the performance of AdaBoost was comparable.

Table 5. Results of credit scoring models on test data

Model	Precision	Recall	F1-Score	Accuracy	AUC
Logistic Regression (LR)	0.5481	0.5429	0.5455	0.6833	**0.6689**
k-Nearest Neighbors (KNN)	0.4225	0.5714	0.4858	0.5767	0.6008
Artificial Neural Network (ANN)	0.4302	**0.7048**	0.5343	0.5700	0.6287
Adaptive Boosting (AdaBoost)	**0.5800**	0.5524	**0.5659**	**0.7033**	0.6609
Random Forest (RF)	0.4429	0.2952	0.3543	0.6233	0.5955

The results presented in Table 5 indicate that the credit scoring models used in this study exhibited acceptable overall performance, with no signs of underfitting or overfitting. The models achieved an average F1-score of approximately 0.50, while the average accuracy and AUC values were around 0.63, suggesting satisfactory predictive capability. Considering the F1-score, accuracy, and AUC, the Adaptive Boosting (AdaBoost) method demonstrated the strongest performance among the tested models on the test data. Consequently, the trained AdaBoost model was selected as the simulation component in the optimization-simulation framework.

Results of optimization-simulation model

The optimization-simulation model was designed to minimize client default probability by adjusting the characteristics of granted loans. To simulate the impact of these modifications on default risk, the AdaBoost model—identified as the best-performing model on the test data—was utilized. The test dataset comprised 300 clients, of whom 33% (100 clients) had defaulted. The optimization-simulation model was executed across four distinct scenarios, and the results were systematically compared.

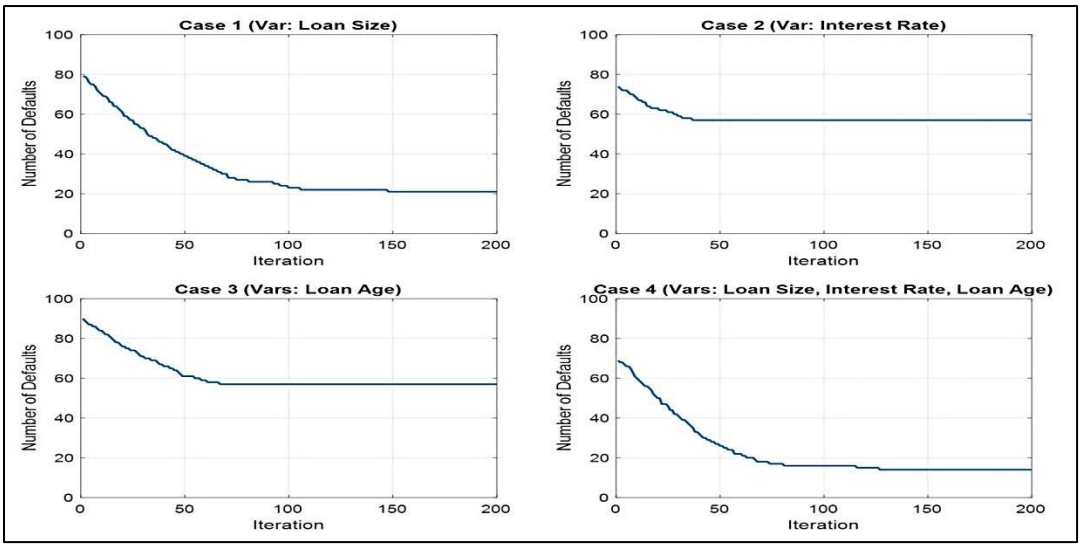

Figure 5. Convergence curve of memetic algorithm in different optimization cases

The convergence curves of the memetic algorithm for various optimization scenarios are illustrated in Figure 5. In these graphs, the vertical axis represents the objective function (the number of defaulted loans), while the horizontal axis indicates the algorithm's iterations.

As shown, the objective function value was initially high across all cases but gradually declined, ultimately converging to an optimal value. Since the initial population generation included solutions with fewer than 100 defaults, the starting point varied for each case. As the algorithm progressed, the application of crossover, mutation, and local search operators facilitated the discovery of improved solutions, leading to a steady reduction in the objective function value.

Convergence was achieved at approximately iterations 150, 40, 70, and 130 for cases 1 through 4, respectively. Each optimization process was completed within 30 minutes. The best performance was observed in case 4, where all three loan characteristics were considered as variables.

The results of the four optimization cases, along with the pre-optimization outcomes, are presented in Table 6. For each case, the percentage of defaulted clients is shown alongside the average values of the key variables: loan size, interest rate, and age.

Table 6. Results of optimization-simulation model across different cases.

Case	Default Rate (%)	Mean Value of Loan Size (Log)	Mean Value of Interest Rate (%)	Mean Value of Loan Age (Year)
Before Optimization	33.33%	11.21	15.90	3.51
Case 1 (Loan Size Optimization)	7.00%	**10.09**	15.90	3.51
Case 2 (Interest Rate Optimization)	19.00%	11.21	**9.74**	3.51
Case 3 (Loan Age Optimization)	19.00%	11.21	15.90	**5.22**
Case 4 (All Variables Optimization)	4.67%	**9.98**	**10.49**	**4.39**

In Case 1, loan size was the optimization variable. As shown in Table 6, adjusting the loan size reduced defaults from 100 clients (33.33%) to 21 clients (7%), marking a 79% decrease compared to the pre-optimization scenario. The average loan size decreased from 11.21 (representing over 100 billion Rials) to 10.09 (representing over 10 billion Rials), reflecting a reduction in the amount of loans granted.

In Case 2, the interest rate was the variable. Modifying the loan interest rate lowered defaults from 100 clients (33.33%) to 57 clients (19%), a 43% decrease. Here, the average interest rate declined from 15.90% to 9.74%.

In Case 3, loan age (repayment period) was the variable. Adjusting loan age reduced defaults from 100 clients (33.33%) to 57 clients (19%), also a 43% decrease. In this case, the average loan age increased significantly from 3.51 years to 22.5 years.

In Case 4, all three loan characteristics—loan size, interest rate, and age—were optimization variables. Adjustments to these factors reduced defaults from 100 cases (33.33%) to 14 cases (4.67%), marking an 86% decrease compared to the pre-optimization scenario. In this case, the average loan size and interest rate declined, while the average loan age increased relative to the pre-optimization conditions. As illustrated in Figure 6, the client default rate decreased significantly across all four optimization cases compared to the pre-optimization scenario. Among these cases, Case 4—where all three loan characteristics were variable—produced the largest reduction in default rate.

Additionally, among the first three optimization cases, Case 1, which optimized loan size, resulted in a greater decrease in defaults than the other two cases, while Cases 2 and 3 yielded similar default rates.

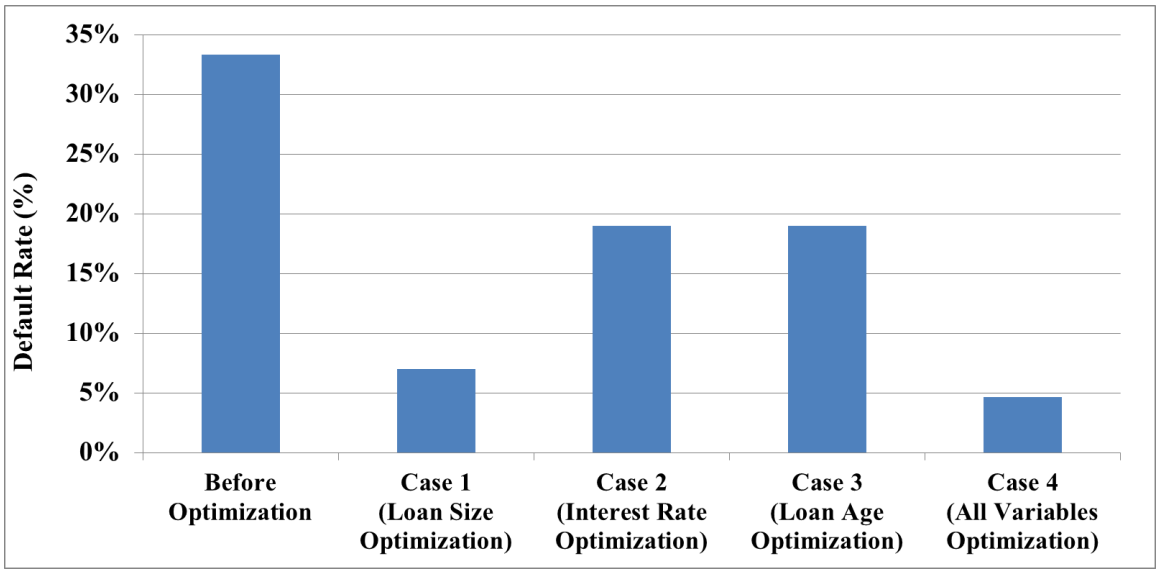

Figure 6. Client's default rate in different optimization cases

Discussion and Conclusion

In this study, an optimization-simulation model was developed to simultaneously perform credit scoring and determine the optimal characteristics of client loans. A case study was conducted using data from corporate clients of an Iranian commercial bank. The proposed method consisted of three stages: (I) Data Preparation: Forty-four features were defined and calculated based on financial statements and bank loans, followed by data preprocessing. (II) Feature Selection & Credit Scoring: Twenty features were selected for credit scoring models using the correlation

coefficient method. Five different models were trained, and the best-performing model was chosen based on its test data performance. (III) Optimization-Simulation: The selected credit scoring model was integrated into the optimization process to determine optimal loan characteristics using a memetic algorithm across four different cases.

Five classification models were employed for credit scoring: LR and KNN as classic baseline models; AdaBoost and RF as ensemble learning techniques, and ANN, which falls between traditional and advanced models. Performance evaluations confirmed that all models met the required minimum accuracy. Among them, AdaBoost demonstrated the highest F1-score and accuracy and ranked second in AUC performance. Thus, AdaBoost, an ensemble learning method, was selected as the best-performing model—aligning with recent research highlighting its effectiveness in credit scoring (Dastile et al., 2020; Markov et al., 2022).

The primary contribution of this study is the introduction of an optimization-simulation model. In this approach, after training the credit scoring models, loan characteristics were treated as adjustable variables within an optimization framework. Recognizing banks' ability to modify loan parameters within a defined range, the model incorporated simulation to assess how loan characteristics influence client default probability. During the simulation phase, all client-specific features—except loan characteristics—remained constant, allowing for precise analysis of default probability variations. The optimization problem was then formulated to determine the optimal loan characteristics, which were solved using a memetic algorithm. Results confirmed the effectiveness of this approach, demonstrating that the algorithm efficiently improved the objective function and converged to an optimal solution within a relatively short timeframe (under 30 minutes).

The optimization problem was examined in four cases:

- Case 1: Loan size was the variable.
- Case 2: Interest rate was the variable.
- Case 3: Loan age (repayment period) was the variable.
- Case 4: All three characteristics were variables simultaneously.

Optimization results revealed reductions in loan defaults of 79%, 43%, 43%, and 86% across these respective cases, confirming that the optimization-simulation model effectively reduces default probability by adjusting loan characteristics. Additionally, loan size exhibited the most significant impact on reducing defaults, followed by interest rate and loan age. These findings suggest that modifying loan terms to reduce installment amounts—either by lowering loan size and interest rate or extending the repayment period—effectively decreases default probability.

The optimization-simulation model presents promising opportunities for improving lending decisions. Traditional credit assessment methods first evaluate a client's creditworthiness before determining loan characteristics, often overlooking how loan parameters influence default probability. By integrating optimal loan determination within the credit scoring process, this model accounts for the impact of loan characteristics on default probability, thereby offering financial institutions a tool to mitigate credit risk more effectively.

Like all research, this study faced certain limitations: (I) Data Accessibility: Obtaining bank loan data for corporate clients—despite anonymity—was challenging. Public institutions such as the Central Bank could facilitate financial research by publishing anonymized loan datasets. This study utilized a sample of 1,000 corporate clients; access to larger datasets would enhance evaluation reliability. (II) Data Availability: Some financial ratios, particularly market value ratios, could not be calculated due to limited data. Additional client details, such as company establishment year, employee count, location, education level, and industry sector, were also unavailable. Expanding the dataset would further refine the proposed method. (III) Technical Constraints: The study utilized MATLAB, which limited the selection of credit scoring models (e.g., XGBoost was unavailable, so AdaBoost was used as an alternative). Future research can explore more diverse models using Python or R. (IV) Alternative Approaches: While this study employed optimization-simulation, future research could examine loan dynamics through other methods, such as game theory, dynamic systems theory, or behavioral economics. (V) Finally, the findings suggest that reducing loan size and interest rates, or extending repayment periods, significantly lowers default rates. Future research may leverage these insights from a banking policy perspective to refine lending procedures and guidelines, thereby supporting more sustainable credit risk management.

Data Availability Statement

Data available on request from the authors.

Ethical considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare no conflict of interest.

References

Abdou, H. A., & Pointon, J. (2011). Credit scoring, statistical techniques and evaluation criteria: a review of the literature. *Intelligent systems in accounting, finance and management*, *18*(2-3), 59-88. https://doi.org/10.1002/isaf.325

Acuna, E., & Rodriguez, C. (2004). A meta-analysis study of outlier detection methods in classification. *Technical paper, Department of Mathematics, University of Puerto Rico at Mayaguez*, *1*, 25.

Ahmed, M. I., & Rajaleximi, P. R. (2019). An empirical study on credit scoring and credit scorecard for financial institutions. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, *8*(7), 2278-1323.

Aljadani, A., Alharthi, B., Farsi, M. A., Balaha, H. M., Badawy, M., & Elhosseini, M. A. (2023). Mathematical Modeling and Analysis of Credit Scoring Using the LIME Explainer: A Comprehensive Approach. *Mathematics*, *11*(19), 4055. https://doi.org/10.3390/math11194055

Ampountolas, A., Nyarko Nde, T., Date, P., & Constantinescu, C. (2021). A machine learning approach for micro-credit scoring. *Risks*, *9*(3), 50. https://doi.org/10.3390/risks9030050

Baharlou, N., Aminbeydokhti, A. A., & Mohagheghnia, M. J. (2016). Comparison Between the Optimized Binary and Multinomial Logistic Regression for Credit Scoring of Real Costumers in Bank Refah Kargaran. *Economics Research*, *16*(63), 147-166. https://doi.org/10.22054/joer.2017.7587 *(in Persian)*

Basu, D., Mitra, S., & Verma, N. K. (2023). Mitigating credit risk: modelling and optimizing co-insurance in loan pricing. *Applied Economics*, *55*(29), 3422-3441. https://doi.org/10.1080/00036846.2022.2115000

Biecek, P., Chlebus, M., Gajda, J., Gosiewska, A., Kozak, A., Ogonowski, D., ... & Wojewnik, P. (2021). Enabling machine learning algorithms for credit scoring--explainable artificial intelligence (XAI) methods for clear understanding complex predictive models. *arXiv preprint*. https://doi.org/10.48550/arXiv.2104.06735

Bluhm, C., Overbeck, L., & Wagner, C. (2016). *Introduction to credit risk modeling*. CRC Press. https://doi.org/10.1201/9781584889939

Boughaci, D., & Alkhawaldeh, A. A. (2020). Appropriate machine learning techniques for credit scoring and bankruptcy prediction in banking and finance: A comparative study. *Risk and Decision Analysis*, *8*(1-2), 15-24. https://doi.org/10.3233/RDA-180051

Cheng, M., & Qu, Y. (2020). Does bank FinTech reduce credit risk? Evidence from China. *Pacific-Basin Finance Journal*, *63*, 101398. https://doi.org/10.1016/j.pacfin.2020.101398

Dadmohammadi, D., & Ahmadi, A. (2017). A novel hybrid approach for training a hierarchical ensemble of neural networks for credit scoring in banking industry. *Intenational Journal of Industrial Engineering and Production Management*, *28*(1), 161-174. *(in Persian)*

Daneshvar, A., Homayounfar, M., Nahavandi, B., & Salahi, F. (2021). A multi-objective approach to the problem of subset feature selection using meta-heuristic methods. *Industrial Management Journal*, *13*(2), 278-299. https://doi.org/10.22059/imj.2021.315625.1007809 *(in Persian)*

Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, *91*, 106263. https://doi.org/10.1016/j.asoc.2020.106263

Dawson, R. (2011). How significant is a boxplot outlier?. *Journal of Statistics Education*, *19*(2). https://doi.org/10.1080/10691898.2011.11889610

Dehghan Nayeri, M., Khazaei, M., & Alinasab Imani, F. (2021). Boundary judgments analysis of the beneficiaries in banking loan allocation. *Industrial Management Journal*, *13*(1), 27-52. https://doi.org/10.22059/imj.2021.311264.1007786 *(in Persian)*

Deng, X., Li, M., Deng, S., & Wang, L. (2022). Hybrid gene selection approach using XGBoost and multi-objective genetic algorithm for cancer classification. *Medical & Biological Engineering & Computing*, *60*(3), 663-681. https://doi.org/10.1007/s11517-021-02476-x

Dželihodžić, A., Đonko, D., & Kevrić, J. (2018). Improved credit scoring model based on bagging neural network. *International Journal of Information Technology & Decision Making*, *17*(06), 1725-1741. https://doi.org/10.1142/S0219622018500293

Eghbali, A., Razavi Hajiagha, S. H., & Amoozad, H. (2017). Performance comparison of genetic algorithm fitness function in customer credit scoring. *Industrial Management Journal*, *9*(2), 264-245. https://doi.org/10.22059/imj.2017.226860.1007191 *(in Persian)*

Farbad Bayazidi, E. & Mohammadi, M. (2014). Credit rating of real customers of Bank Mellat using data mining methods. *The first national congress for the development of monetary and banking management*, Tehran, Iran. *(in Persian)*

Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. (2024). How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. *Journal of Financial Stability*, 101284. https://doi.org/10.1016/j.jfs.2024.101284

Ghasemnia Arabi, N., & Safaei Ghadikolaei, A. (2019). Comparison of the performances of classical models and artificial intelligence in predicting bank customers' credit status. *Journal of Business Administration Researches*, *10*(20), 51-69. https://doi.org/10.29252/bar.2019.1320 *(in Persian)*

Groppelli, A. A., & Nikbakht, E. (2000). Finance, Barron's Educational Series. *Inc., ISBN*, *764112757*.

Holland, J. H. (1973). Genetic algorithms and the optimal allocation of trials. *SIAM journal on computing*, *2*(2), 88-105. https://doi.org/10.1137/0202009

Hosseini, S. A., & Zibaei, M. (2015). Credit risk management in agricultural bank of Mamasani using neural network model. *Agricultural Economics*, *9*(2), 103-119. *(in Persian)*

Hussin Adam Khatir, A. A., & Bee, M. (2022). Machine learning models and data-balancing techniques for credit scoring: What is the best combination? *Risks*, *10*(9), 169. https://doi.org/10.3390/risks10090169

Izadikhah, M., & Shamsi, M. (2020). Credit rating of the bank legal customers by using the improved modified Russell model (Case study: the legal customers of Arak Melli Bank). *Journal of New Researches in Mathematics, 5*(22), 111-126. *(in Persian)*

Izadikhah, M., Shamsi, M., Sheikhan, A., & Ghafouri, F. (2021). A New Hybrid Approach Based on Fitch Ranking and Data Envelopment Analysis to Evaluate the Credit Performance of the Bank's Legal Clients. *Journal of Decisions & Operations Research, 6*(4), 553-569. https://doi.org/10.22105/dmor.2021.250078.1225 *(in Persian)*

Jandaghi, G., Saranj, A., Rajaei, R., Qhasemi, A., & Tehrani, R. (2020). Evaluating credit risk based on combined model of neural network of pattern recognition and ants' colony algorithm. *Management and Development Process*, *33*(1), 133-170. https://doi.org/10.29252/jmdp.33.1.133 *(in Persian)*

Kazemi, A. Azimzadeh Tehrani, K. Abdali, A & Aryai, S (2022). Predicting the credit score of individuals using data mining algorithms (case study: customers of an Iranian state bank). *Monetary and Banking Research*, *48*, 327-360. *(in Persian)*

Khorrami, A., Taghavifard, M. T., & Khatami Firouzabadi, S. M. A. (2020). Credit status assessment of bank loan applicants using CBR method. *Industrial Management Studies*, *18*(59), 79-116. https://doi.org/10.22054/jims.2018.18574.1660 *(in Persian)*

Koohi, H & Gholami, R (2012). Credit rating of legal clients of the industry sector using data envelopment analysis (DEA) model. *Quantitative Studies in Management*, *3*(3), 115-138. *(in Persian)*

Koutanaei, F. N., Sajedi, H., & Khanbabaei, M. (2015). A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit scoring. *Journal of Retailing and Consumer Services*, *27*, 11-23. https://doi.org/10.1016/j.jretconser.2015.07.003

Krishna, S. J. S., Aarif, M., Bhasin, N. K., Kadyan, S., & Bala, B. K. (2024, July). Predictive Analytics in Credit Scoring: Integrating XG Boost and Neural Networks for Enhanced Financial Decision Making. In *2024 International Conference on Data Science and Network Security (ICDSNS)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICDSNS62112.2024.10691123

Kumar, R., & Memoria, M. (2020). A review of memetic algorithm and its application in traveling salesman problem. *Int. J. Emerg. Technol.*, *11*(2), 1110-1115.

Laborda, J., & Ryoo, S. (2021). Feature selection in a credit scoring model. *Mathematics*, *9*(7), 746. https://doi.org/10.3390/math9070746

Lenka, S. R., Bisoy, S. K., Priyadarshini, R., & Sain, M. (2022). Empirical analysis of ensemble learning for imbalanced credit scoring datasets: A systematic review. *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/6584352

Louzada, F., Ara, A., & Fernandes, G. B. (2016). Classification methods applied to credit scoring: Systematic review and overall comparison. *Surveys in Operations Research and Management Science*, *21*(2), 117-134. https://doi.org/10.1016/j.sorms.2016.10.001

Maleki, A., Zare, A., NiKoumaram, H., & Shahverdiani, S. (2020). Presenting an optimal credit risk model for crowdfunding process using multilayer perceptron (MLP) neural network. *Financial Engineering and Portfolio Management*, *11*(43), 271-290. *(in Persian)*

Malik, M., & Thomas, L. C. (2010). Modelling credit risk of portfolio of consumer loans. *Journal of the Operational Research Society*, *61*(3), 411-420. https://doi.org/10.1057/jors.2009.123

Markov, A., Seleznyova, Z., & Lapshin, V. (2022). Credit scoring methods: Latest trends and points to consider. *The Journal of Finance and Data Science*. https://doi.org/10.1016/j.jfds.2022.07.002

Mehrara, M., Mousaei, M., Tasavori, M. & Hassanzadeh, A. (2009). Credit rating of corporate clients of Parsian bank. *Economic Modeling*, *3*(10), 121-150. *(in Persian)*

Metawa, N., Elhoseny, M., Hassan, M. K., & Hassanien, A. E. (2016, December). Loan portfolio optimization using genetic algorithm: a case of credit constraints. In *2016 12th international computer engineering conference (ICENCO)* (pp. 59-64). IEEE. https://doi.org/10.1109/ICENCO.2016.7856446

Metawa, N., Hassan, M. K., & Elhoseny, M. (2017). Genetic algorithm based model for optimizing bank lending decisions. *Expert Systems with Applications*, *80*, 75-82. https://doi.org/10.1016/j.eswa.2017.03.021

Mir, M., Rafiee, H., Ensan, E., Shakeri Bostanabad, R., & Yazdanpanah, R. (2022). Ranking of factors affecting credit risk of agricultural facilities: A case study of fattening units in Ardabil province of Iran. *Agricultural Economics and Development*, *30*(3), 35-70. https://doi.org/10.30490/aead.2023.353565.1306 *(in Persian)*

Moradi, S., Mokhatab Rafiei, F. & Saghayi, A. (2022). Identification of dynamic patterns of credit risk of individual customers of banks and dinancial institutions (case study: three Iranian banks). *Monetary and Banking Research*, *15*(51), 121-154. *(in Persian)*

Moscato, P. (1989). On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms. *Caltech concurrent computation program, C3P Report*, *826*(1989), 37.

Munkhdalai, L., Munkhdalai, T., Namsrai, O. E., Lee, J. Y., & Ryu, K. H. (2019). An empirical comparison of machine-learning methods on bank client credit assessments. *Sustainability*, *11*(3), 699. https://doi.org/10.3390/su11030699

Mushava, J., & Murray, M. (2024). Flexible loss functions for binary classification in gradient-boosted decision trees: An application to credit scoring. *Expert Systems with Applications*, *238*, 121876. https://doi.org/10.1016/j.eswa.2023.121876

Najibi, M & Mokhatab Rafiei, F. (2017). Credit rating of bank customers using clustering with a performance comparison approach of three algorithms: Kohonen, K-Means, and Two-Step. *The 7th International Conference on Management and Accounting*, Tehran, Iran. *(in Persian)*

Nalić, J., & Martinovic, G. (2020). Building a credit scoring model based on data mining approaches. *International Journal of Software Engineering and Knowledge Engineering*, *30*(02), 147-169. https://doi.org/10.1142/S0218194020500072

Nehrebecka, N. (2018). Predicting the default risk of companies. Comparison of credit scoring models: LOGIT vs Support Vector Machines. *Econometrics. Ekonometria. Advances in Applied Data Analytics*, *22*(2), 54-73. https://doi.org/10.15611/eada.2018.2.05

Neri, F., & Cotta, C. (2012). Memetic algorithms and memetic computing optimization: A literature review. *Swarm and Evolutionary Computation*, *2*, 1-14. https://doi.org/10.1016/j.swevo.2011.11.003

Nili, M. & Sabzevari, H. (2008). Estimation and comparison of logit credit rating model with analytic hierarchy process (AHP) method. *Industrial Engineering & Management*, *24*(43), 105-117. *(in Persian)*

Parvizian, K., Zakawt, S. M & Mohammadian, M (2009). Internal ranking of bank customers using logit regression models. *Economic Research*, *6*, 61-89. *(in Persian)*

Pławiak, P., Abdar, M., Pławiak, J., Makarenkov, V., & Acharya, U. R. (2020). DGHNL: A new deep genetic hierarchical network of learners for prediction of credit scoring. *Information Sciences*, *516*, 401-418. https://doi.org/10.1016/j.ins.2019.12.045

Qadi, A., Trocan, M., Diaz-Rodriguez, N., & Frossard, T. (2023). Feature contribution alignment with expert knowledge for artificial intelligence credit scoring. *Signal, Image and Video Processing*, *17*(2), 427-434. https://doi.org/10.1007/s11760-022-02239-7

Rastegar, M. A., & Eidi Goosh, M. (2020). Credit risk modeling: spline based logistic regression survival approach. *Journal of Investment Knowledge*, *9*(34), 167-184. *(in Persian)*

Rofik, R., Aulia, R., Musaadah, K., Ardyani, S. S. F., & Hakim, A. A. (2024). The Optimization of Credit Scoring Model Using Stacking Ensemble Learning and Oversampling Techniques. J*ournal of Information System Exploration and Research*, *2*(1), 11-20. https://doi.org/10.52465/joiser.v2i1.203

Rong, Y., Liu, S., Yan, S., Huang, W. W., & Chen, Y. (2023). Proposing a new loan recommendation framework for loan allocation strategies in online P2P lending. *Industrial Management & Data Systems*, *123*(3), 910-930. https://doi.org/10.1108/IMDS-07-2022-0399

Runchi, Z., Liguo, X., & Qin, W. (2023). An ensemble credit scoring model based on logistic regression with heterogeneous balancing and weighting effects. *Expert Systems with Applications*, *212*, 118732. https://doi.org/10.1016/j.eswa.2022.118732

Soltani, A., Ehtsham Rathi, R., & Abedi, S. (2023). Providing a multi-objective mathematical model for optimizing the equipment and allocation of financial resources of the banking system. *Journal of Industrial Management*, *15*(2), 272-298. https://doi.org/10.22059/IMJ.2023.352863.1008011 *(in Persian)*

Sörensen, K., & Sevaux, M. (2006). MA| PM: memetic algorithms with population management. *Computers & Operations Research*, *33*(5), 1214-1225. https://doi.org/10.1016/j.cor.2004.09.011

Soui, M., Gasmi, I., Smiti, S., & Ghédira, K. (2019). Rule-based credit risk assessment model using multi-objective evolutionary algorithms. *Expert systems with applications*, *126*, 144-157. https://doi.org/10.1016/j.eswa.2019.01.078

Sum, R. M., Ismail, W., Abdullah, Z. H., & Shah, N. F. M. N. (2022). A New Efficient Credit Scoring Model for Personal Loan Using Data Mining Technique for Sustainability Management. *Journal of Sustainability Science and Management*. http://doi.org/10.46754/jssm.2022.05.005

Taghvi, M., Lotfi, A. & Sohrabi, A (2008). Credit risk model and rating of legal customers of Keshavarzi Bank. *Economic Research*, *4*, 99-128. *(in Persian)*

Talaat, F. M., Aljadani, A., Badawy, M., & Elhosseini, M. (2024). Toward interpretable credit scoring: integrating explainable artificial intelligence with deep learning for credit card default prediction. *Neural Computing and Applications*, *36*(9), 4847-4865. https://doi.org/10.1007/s00521-023-09232-2

Thomas, L., Crook, J., & Edelman, D. (2017). *Credit scoring and its applications*. Society for industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611974560

Tripathi, D., Edla, D. R., Bablani, A., Shukla, A. K., & Reddy, B. R. (2021). Experimental analysis of machine learning methods for credit score classification. *Progress in Artificial Intelligence*, *10*, 217-243. https://doi.org/10.1007/s13748-021-00238-2

Wang, S. (2021). Research on optimization model of bank credit strategy. *Financial Engineering and Risk Management*, *4*(3), 58-62. https://doi.org/10.23977/ferm.2021.040310

Witzany, J., & Witzany, J. (2017). *Credit risk management*. Springer International Publishing. https://doi.org/10.1007/978-3-319-49800-3

Xiao, J., Zhong, Y., Jia, Y., Wang, Y., Li, R., Jiang, X., & Wang, S. (2024). A novel deep ensemble model for imbalanced credit scoring in internet finance. *International Journal of Forecasting*, *40*(1), 348-372. https://doi.org/10.1016/j.ijforecast.2023.03.004

Yan, B., Liang, M., & Liu, L. (2024). Optimal decisions of retailer's loans from bank. *International Journal of Finance & Economics*, 1-8. https://doi.org/10.1002/ijfe.2979

Yang, F., Qiao, Y., Huang, C., Wang, S., & Wang, X. (2021). An automatic credit scoring strategy (ACSS) using memetic evolutionary algorithm and neural architecture search. *Applied Soft Computing*, *113*, 107871. https://doi.org/10.1016/j.asoc.2021.107871

Zhou, M. (2022). Credit Risk Assessment Modeling Method Based on Fuzzy Integral and SVM. *Mobile Information Systems*, *2022*. https://doi.org/10.1155/2022/3950210