*Journal of Industrial and Systems Engineering*
Vol. 15, No. 1, pp. 306-327
Winter (January) 2023

# A review of credit rating models: A combined analysis and suggestions for future research

**Seyed Farid Mousavi[1], Arezoo Gazori-Nishabori[2]***

[1]*Department of Operation Management and Information Technology, Kharazmi University, Tehran, Iran*
[2]*Department of Industrial Engineering, South-Tehran Branch, Azad University, Tehran, Iran*
*mousavifarid@khu.ac.ir, arezoogazori.66@gmail.com*

**Abstract**

The key to solving the problem of obtaining complex facilities is to create a suitable credit rating model that can provide technical support for the approval of granting facilities provided by small and micro enterprises. Credit rating agencies perform assessment to support financial institutions in processing debts. Added literature in the field of credit rating from January 2015 to August 2023 was analyzed to discover opportunities for further research. Bibliometric analysis was used to understand the existing literature. Subsequently, through structured review theories, the methods used by researchers and credit rating agencies were examined. A hybrid literature review was developed by integrating bibliometric and structured review of research articles from widely recognized databases. A sample of 72 articles has been made and studied to identify the gaps in the field of credit rating and create a suitable solution to fill such gaps. The results showed that most studies appeared as post-financial crisis effects reported in 2016 and 2023. It contributes to the existing literature by encouraging researchers and credit rating agencies to develop a specific credit rating system by evaluating existing models and improvising them by adopting advanced techniques such as multiple regression, neural networks, aggregate learning, and machine learning.

**Keywords:** Credit rating, credit rating models, ensemble learning, machine learning, statistical method

## 1- Introduction

An approach to reducing the non-repayment of facilities and overdue claims of banks is to develop an accurate credit scoring system to measure customers' risk and determine their credit scores. The development of a credit scoring system decreases human errors, besides enhancing precision and accelerating credit risk determination. Thus, it shortens the facility granting time, reduces costs and facility provision risks, and increases efficiency and transparency in banks. Conventional statistical methods and artificial intelligence techniques play a crucial role in predicting credit risks. Many past studies are based on quantitative methods, while a few, such as Uthayakumar, Vengattaraman, and Dhavachelvan (2020), have adopted qualitative approaches to improving the performance of credit risk-predicting models. Quantitative methods have been developed for the recognition of economic models and types of customers and markets. They predict the default rate and determine the credit score of customers.

---

*Corresponding author
ISSN: 1735-8272, Copyright c 2023 JISE. All rights reserved

With respect to shifts in economic conditions, various social and political transformations, and their impacts on the economy and economic enterprises, it is imperative to develop an inclusive model that takes quantitative and qualitative conditions and experts' judgments into account. Abrupt shifts in the economies of the US, Greece, Venezuela, and Iran reflect variations in economic models, e.g., the 2001 terroristic attack and the 2008 financial crisis in the US, the debt crisis after the financial crisis in 2007-2008 in Greece, the extreme inflation rate in 2018 in Venezuela, sanction-resultant economic shocks in 2018 in Iran, and the global economic shock rooted in the outbreak of Covid-19. Concerning the variation in economic models, it is necessary to benefit from both qualitative and quantitative approaches to identify and model changes.

Credit Rate Agencies (CRAs) are independent third parties who determine the credit rate of business units according to financial performance. CRAs significantly contribute to debt processing and direct lenders about customers' credits. Lenders often utilize the information presented by these agencies to decide on the debt cost, given the debt interest rate or repayment terms. For example, loanees with lower credit scores face higher interest rates, while those with acceptable credit rates can receive extra loans with lower interest rates. Hence, the views of CRAs are important for both loaners and loanees. The respective literature on the participation of CRAs in the financial market was reinforced by investigations on the effect of credit rates on investors in global financial crises, during which the trade of CRAs was criticized extensively. Since that time, regulators have attempted to withdraw their power in financial markets (Ubarhande & Chandani, 2021). However, Florian (2016) justified the positive contribution of CRAs to financial markets, especially during financial crises. Wang and Yang (2020) reasoned that CRAs could warn financial markets of such crises in advance. As noted by Sangiorgi and Spatt (2017), the accuracy of credit rates in predicting financial problems decline during financial crises. Likewise, rating institutions underperform in structured financing during these crises.

An examination of various papers on credit rating reveals a dearth of research inclusively reviewing credit rating models centered on legal clients, especially in the banking industry, and analyzing the respective literature in this domain. Thus, this paper reviewed and analyzed the present literature on credit rating models by specifically focusing on juridical individuals. It considered the banking industry to identify contributions to this domain from January 2015 to August 2023 and presented contextual suggestions for future studies. In this regard, it was compiled into four sections. The second section presents the research methodology, and the third section investigates the literature on credit rating meticulously. The analyses are provided in section four, and conclusions and suggestions for future studies are presented in section five.

## 2- Methodology

The present paper reviewed the respective literature to identify contributions to credit rating models for juridical individuals in the banking industry during an eight-year period (2015-2023) and presented contextual suggestions for future studies. For this purpose, it followed a systematic review approach. To ensure that the literature was deeply examined, we defined a paper review protocol that enumerated the literature searching and quality criteria and introduced databases and sampling approaches.

### 2-1- The literature searching criteria and introduced databases

We selected papers from known databases, such as Scopus, Web of Science, and Google Scholar. The selected papers were published by Elsevier, Springer, Sage Publication, Taylor and Francis, Emerald Insight, IEEE, and other valid publications. A total of 72 research papers were examined. The following section enumerates the phases of paper selection for investigation.

### 2-2- Sampling

Databases, including Scopus, Web of Science, and Google Scholar, were selected for the examination of the present literature on credit rating. The research papers were extracted from the databases and screened by applied keywords like credit rating, credit scoring, credit rating model, and credit scoring model. Scopus and Web of Science are the most favorite databases in the world and present scientific resources and research materials in all disciplines. The editorial boards of the journals indexed in these databases and their affiliated publishers guarantee the selection of the best research works for publication. Table (1) displays the phases of paper extraction.

- A total of 299 papers were extracted from the Scopus, Web of Science, and Google Scholar databases in the first phase.
- Out of this number, 176 papers were identified based on the applied keywords, like credit rating, credit scoring, credit rating model, and credit scoring model.
- The selected 176 papers were examined initially, and 72 related papers were identified as the sample.

During this investigation, papers on credit rating, especially of juridical individuals, were selected due to their significance in the banking industry.

**Table 1.** Extraction process of selected papers for investigation

| Subject | Paper's Number |
|---|---|
| English Credit Rating Papers for 2015-2023 | 299 |
| Papers Filtered by Keyword Credit Rating Model | 176 |
| Selecting Papers | 72 |

## 2-3- Quality criteria

This study followed a concept-based approach to screen and examine 176 papers obtained based on keywords. Several concepts were considered in this respect, e.g., models or CRAs, techniques, credit rating scopes, factors employed for determining credit rates, and model comparisons. Furthermore, the author-based approach was not used since it might be unsuitable for a successful literature review and give rise to a summary, as noted by Ubarhande & Chandani (2021). The following section examines the literature on credit rating models in 72 extracted sample papers according to the research methodology.

# 3- Examining the literature on credit rating models

CRAs' evaluations of loanees' credits reflect strong emphases during the recovery of any financial crises, like the ones reported in 2008 and 2016, or non-financial crises, such as epidemics. After such crises, precautions and vigilance toward every financial transaction rise all over the world. Financial sectors have witnessed the emergence and intrusion of third-party assessors in such periods. During the recovery after the breakdown resulting from the Covid-19 epidemic, financial institutions bear the big responsibility of economic development and support by offering pecuniary aid. Customer credit is now the main keyword that appears as an intermediate in successful financial transactions to determine the level of credit risks. In this section, the present literature on credit rating is examined, and researchers' inclinations are identified. For this purpose, the background in areas shown in figure (1) is examined.

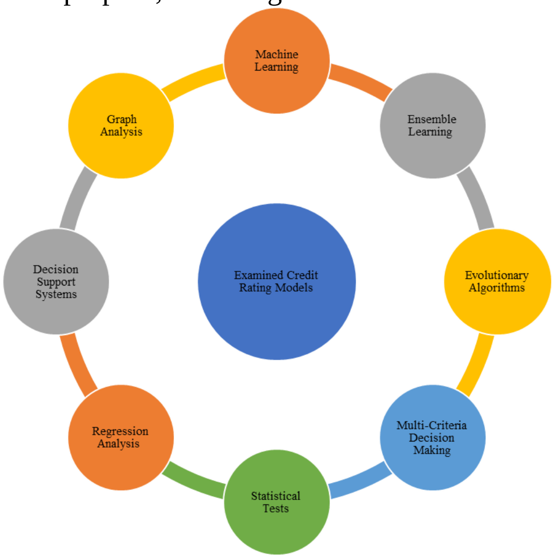


**Fig 1.** Literature on credit rating models

## 3-1- Machine learning

In the machine learning domain, Gao et al.(2023) proposed a new credit scoring model based on contrastive augmentation and tree-enhanced (CATE) setting mechanisms. This model automatically constructs explainable cross-features by using tree-based models to learn decision rules from the data and the significant of each local cross-feature is then derived through an attention mechanism. Qian et al. (2023) used from a novel end-to-end soft reordering one-dimensional convolution neural network (CNN) that could reorganize the original tabular data and make them more conducive to CNN learning. This method proposed for firm's credit scoring and obtained superior results than other models. Dong et al. (2023) introduced a loan-level dataset from loans micro finance company in China. In this paper investigated information credit scores on microloan risk management. Helder et al. (2022) proposed a feature selection technique based on variable neighborhood concept that called VNS. This method used to make default prediction in credit analysis problems. Javadpour et al. (2021) improved the efficiency of customer credit rating through machine learning algorithms in the cloud computing of macro data. In their study, they applied several forecasting algorithms to make predictions according to the integration of Ordered Weighting Averaging (OWA)-based results. Abdi (2021) presented a new multistage feature classification and selection approach for the credit risk scoring of customers in Iranian banks. The selected significant features were applied in the K-nearest neighbor and decision tree algorithms during a four-stage process. Ampountolas et al. (2021) presented a machine-learning approach to microcredit rating. In microlending markets, the lack of a credit record is a significant barrier to assessing loanees' credits and thus deciding on fair interest rates. The results showed that the Random Forest algorithm could well perform this task using the available data of customers, such as age, occupation, and location. Singh and Goel (2021) embarked on predicting the credit scores of bank customers using feature selection and data-mining algorithms. In this study, they compared the Random Forest and Logistic Regression algorithms to predict customers' credit scores by applying *k* best features to datasets.

Tran et al. (2021) ran an empirical analysis based on machine learning to score credits in the Vietnamese banking system. The experiments were performed by modern machine learning methods based on the LightGBM, CatBoost, and Random Forest group learning models. The experimental results showed improvements compared to the base algorithms, such as support vector machine or logistic regression. Agrawal, Ahirao and Dere (2021) assessed customer credit scoring by machine learning algorithms. The results displayed that logistic regression was the most accurate algorithm in the credit score-predicting model compared to the other machine learning algorithms. Djeundje et al. (2021) tried to improve credit scoring with alternative data, without which the type of available data was demographic. Models using emails and psychometric and demographic variables can render higher prediction accuracies.

Li, Xiao and Yang (2021) presented a new approach to scoring credits based on feature transformation and ensemble models. The feature transformation process, including boosting trees and auto-encoders, replaced manual feature engineering and solved the data imbalance problem. For the classification process, this paper designed a heterogeneous ensemble model by giving weights to the DNN and factorization machine. Tripathi et al. (2021) empirically analyzed machine learning algorithms to classify credit scores. Their study analyzed empirical results of mixing feature selection approaches with various classification methods. Kozeny (2015) focused on feature selection in a credit scoring model. They employed three different feature selection methods, including a filtering method (the K-squared test and correlation coefficients) and two wrapper methods (forward stepwise selection and backward stepwise selection), to reduce overfitting.

Nazari, Mehregan and Tehrani (2020) evaluated the effectiveness of data-mining techniques in the scoring of bank customers' credits using mathematical models. They considered the individual loanees of the Refah Bank of Zanjan province in Iran. Their findings revealed that the support vector machine and artificial neural networks were the most efficient examined techniques. Mukid et al. (2018) analyzed credit scoring by the weighted K-nearest neighbor algorithm and displayed the outperformance of the Gaussian and rectangular kernels. SULTANA (2018) evaluated an automated credit scoring system for financial services in developing countries. This study identified the primary deciding factors in developing an automated system for credit scoring purposes. The priority of employment, applicants' rights, previous facility reception records, the purpose for taking facilities, and number of requested facilities constituted optimal features.

Jadhav, He and Jenkins (2018) proposed a new approach, namely the *Information Gain Directed Feature Selection algorithm*, to selecting features in credit scoring programs. In this study, *m* advantageous features were released by the GA Wrapper mechanism, and three machine learning algorithms, i.e., KNN, Naïve Bayes, and SVM, were used for credit scoring. Shi and Xu (2016) used the fuzzy SVM algorithm with a new member function to present a credit scoring system, where the function displayed the varying contribution of every entry point to the SVM classifier hyperplane learning. Bunker, Zhang and Naeem (2016) attempted to improve their credit scoring model by mixing features derived from bank statements. In this study, Naïve Bayes was the best model in terms of its performance.

## 3-2- Ensemble learning

In the ensemble learning domain, Liu et al. (2022) proposed the new method based on the tree algorithm for credit scoring. This method considered a boosted tree as a base framework that focal-aware cost-sensitive light gradient boosting machine (LightGBM-focal).Ubarhande and Chandani (2021) employed the random forest algorithm to predict a customer-confirmed model for acquiring bank facilities. Singh et al. (2021) presented a multilevel classification and an ensemble approach based on the modified PSO clustering for credit scoring. They proposed a new approach under the title of multilevel classification and cluster-based ensemble, which embraced feature selection weaknesses and ensemble-based classification. Kang et al. (2021) presented a graph-based semi-supervised reject inference framework considering imbalanced data distribution for the credit scoring of consumers. Reject inference is a technique for deducing good-bad labels for rejected applicants and can overcome biases in credit scoring. This paper considered imbalanced data distribution to score consumers' credits.

Li and Chen (2020) compared the performance of ensemble learning in credit scoring models. They presented a comparative performance evaluation of ensemble algorithms, such as Random Forest, AdaBoost, XGBoost, LightGBM, and Stacking, in terms of accuracy, under-curve area, Kolmogorov-Smirnov statistic, Brier score, and model operation time during credit scoring. Qiu (2019) tried to predict credit risks in an imbalanced social lending environment based on the XGBoost algorithm. The credit data of small enterprises were imbalanced and made general classification models fail in predicting credit performance. The presented model integrated the benefits of XGBoost and KMeans.

Shema (2019) followed an effective credit scoring approach using limited mobile phone data. To screen loanees, digital loaners usually collect a large volume of data, such as relational models, data tied to social media activities, and accurate use of mobile phones, from their customers. These data show several potential privacy risks to loanees. This study revealed that accurate credit scoring models could be trained by airtime recharge data due to their less intrusion into loanees' privacy.

Aji and Dhini (2019) scored credits with a data-mining method and considered the mortgage loan as their case in Indonesia. The C4.5 decision tree and random forest were the applied classifiers. According to their findings, the high-accurate classifier was the random forest algorithm with AdaBoost (72.95%), while the worst prediction belonged to C4.5 with an accuracy of 68.7%. Chopra and Bhilare (2018) examined the uses of ensemble models in credit scoring systems. Their empirical analysis showed that the gradient-boosting model outperformed the basic decision tree learner. AlaÕraj and Abbod (2016) presented a classifiers consensus system approach for credit scoring. The empirical results, analyses, and statistical tests displayed that the proposed mixed method was more competent in improving prediction performance than all basic classifiers, logistic regression, MARS, and seven conventional mixed approaches in terms of the accuracy average and under-curve area.

Wickens (2016) examined ensemble classification based on supervised clustering for credit scoring. In the proposed approach, supervised clustering was used to divide the data samples of every class into several clusters. Then, the clusters of different classes were paired to develop some train subsets. AlaÕraj and Abbod (2016) presented a new hybrid ensemble credit scoring model according to the classifiers consensus system approach. To this end, they combined two data preprocessing methods based on Gabriel neighborhood graph editing and MARS, and the modeling phase of the new classifiers' combination was based on the consensus of various classification algorithms. Koutanaei, Sajedi and Khanbabaei (2015) presented a hybrid data-mining model of feature selection algorithms and ensemble learning classifiers for credit scoring. The results showed that PCA was the best feature selection algorithm. The classification findings also revealed that the AdaBoost method of the artificial neural network was a highly accurate classifier.

In the artificial neural networks' domain, Mercep et al. (2021) used deep neural networks for behavioral credit rating. Their proposed network outperformed several methods and was equally adapted to the XGBoost model. Li, Xiao and Yang (2021) ran a comparative analysis based on neural network models to reinforce credit scoring systems of micro and small businesses with soft information. Their study showed the advantageousness of the Back Propagation Neural networks (BPNN) in predicting loan classifications. Xiao, Xiao and Wang (2016) presented a new approach to credit scoring based on a cost-sensitive neural network ensemble. The proposed approach compared the weights of the multilevel classes with the main train data and enabled the multilevel basic neural networks to consider unbalanced classes. The prediction model could well compromise between default (bad credit applicants) and non-default (high-credited customers) classes.

Li and Sun (2021) examined the uses of the optimal segmentation RBF neural network algorithm in credit rating. The improved optimal segmentation algorithm was used to train the RBF neural network parameters to expand the width and center of the class and the RBF network model. Lastly, the differential objective function of the class was employed to realize the adaptive selection of the number of hidden nodes for adjusting the structure of the RBF network model dynamically. Pang, Wang and Xia (2021) examined a farmer's credit rating model and program based on a multilayer unified network with a linear classifier. Besides, they presented an index calculation formula per level of the three-level credit rating system and a formula for estimating the credit score of a one-way four-layer network.

Dadmohammadi and Ahmadi (2019) presented a combined learning approach to credit scoring using an adaptive hierarchical mix of experts in the Iranian banking industry. They developed a credit scoring model using a modular neural network based on hybrid ensemble learning. The proposed model comprised four robust neural networks that collectively built the Adaptive Hierarchical Mix of Experts (AHME). Ayouche, Aboulaich and Ellaia (2017) suggested a neural network approach for the partnership credit scoring classification problem. The new method presented for developing credit scoring models considered several partnership contract criteria in the microfinance institutions of Morocco using multilayer perceptron neural networks.

Soydaner and Kocadağlı (2015) employed artificial neural networks with gradient learning algorithms for credit scoring. They also considered some prerequisites, such as model convolution, overfitting, and optimal algorithm selection when training artificial neural networks. Kiruthika and Dilsha (2015) introduced a neural network approach to microfinance credit scoring. Their study compared logistic regression and artificial neural network models in their credit scoring of microfinance data. Tomczak and Zięba (2015) presented Classification Restricted Boltzmann Machine (ClassRBM) for comprehensible credit scoring models. In the first step, ClassRBM was trained as an independent classifier that was able to predict credit conditions but lacked any interpretable structure. To obtain a comprehensible model, they first evaluated the association of every binary feature through ClassRBM and then employed these values to create a scoring table (scorecard).

### 3-3- Evolutionary algorithms

In the evolutionary algorithms' domain, Xio et al. (2023) combined a new over-sampling method, the variational autoencoder (VAE), and a deep ensemble classifier, the deep forest (DF), and proposed a novel deep ensemble model for credit scoring in internet finance, VAE–DF. This method is an effective credit scoring tool, especially for the complex distributions of the datasets, highly class-imbalanced and non-linear datasets. Wu et al. (2022) proposed framework searches for a pool of Pareto-optimal credit scoring models with different features. The searching process did by binary multi-objective particle swarm optimization (BMOPSO) algorithm for three retail credit scoring datasets. Pławiak et al. (2020) presented a novel hierarchical network of learners to predict credit scores. This model employed deep learning, ensemble learning, supervised learning, layered learning, genetic feature selection, genetic optimization of learner algorithms, and new layered training (learner selection) besides cross-validation. Babaev et al. (2019) applied deep learning in credit facility-granting programs. This research suggested a new approach, namely the Embedding-Transactional Recurrent Neural Network, to estimate the credit scores of bank customers by examining the history of their debts and credit card transactions. Kozeny (2015) compared the performance of the alternative optimal function in genetic algorithms to score credits. He suggested an alternative optimal function according to a variable bitmask and compared its performance with optimal functions based on a polynomial equation and parameter range estimation. The results showed the advantageousness of bitmask over the other two approaches.

## 3-4- Multi-criteria decision making (MCDM)

In the multi-criteria decision-making domain, Katsimperis and Andrikopoulos (2021) developed a flexible commercial credit scoring model using multi criteria decision analysis. This method saved time and resources in addition to guaranteeing the quality of the outcomes. Fernando and Siagian (2021) suggested using MCDM to evaluate the credit applications of banks and employed the 5C index, including character, capacity, capital, economic conditions, and collaterals. Ehtesham Rasi, Karamipour and Arad (2020) rated actual customers of banks based on credit risks using MCDM and Artificial Intelligence Hyperbolic Regression (AIHR). The MCDM-based rating included customers' income, credits in the market, occupations, and contact duration with banks, besides collateral type, collateral value, and average account balance, for facilitating the credit risks of actual customers. The AIHR-based rating involved customers' credits in the market, incomes, occupations, and contact duration with banks, along with collateral value and type.

Nazari, Mehregan and Tehrani (2019) employed a hybrid model for credit scoring by focusing on the credit microloan customers of Refah Bank in Zanjan, Iran. Following an optimization linear programming approach, they employed the Utilities Additives DIS criminates (UTADIS) model to score the credits of bank customers. Among the merits of the proposed approach, we can refer to its high flexibility, mutual interaction with decision-makers, and updating ability under various macroeconomy conditions. The results indicated that customers' initial credit scores and ages had the maximum effect on their credit scores.

Shen, Sakai and Tzeng (2019) compared two new hybrid Multi-Relational Data Mining (MRDM) approaches to scoring consumers' credits in uncertain and vague judgment conditions. MRDM is a subset of hybrid MCDM that employs the benefits of machine learning, soft computations, and decision-making techniques. Their study employed DRSA, DANP, and the bipolar model for criteria selection and weighting and the fuzzy logic for application rating.

Chai et al. (2019) presented a multicriteria approach to modeling the credit rating of small enterprises in China. This paper used partial correlation and probit regression analyses and examined 687 small retail enterprises in a regional trade bank in China to develop a credit rating system consisting of 17 indices. Then, the TOPSIS software with the fuzzy C-means technique was used for the credit scoring of companies. Reference [31] presented a fuzzy decision support system for credit scoring. The proposed approach weighted the criteria by fuzzy MCDM and used TOPSIS with the risk distance function to rate alternatives based on the minimum risk.

Yotsawat, Wattuya and Srivihok (2021) presented a credit rating model based on a customer number bell-shaped distribution and introduced a multi-objective programming for this purpose. The first objective function minimized the absolute difference between the obligor number and perfect customer ratios by following a standard normal distribution. The second objective function minimized the total difference of the deviation between two adjacent credit scores' loss rates. These researchers mixed the two objective functions to ensure the obligor number distribution and the homogeneity of the loss rate and employed a genetic algorithm to solve the model.

## 3-5- Statistical tests

In the statistical tests' domain, Ashofteh and Bravo (2021) presented a conservative approach to online credit scoring. This study employed the non-parametric Kruskal-Wallis statistic to develop a conservative credit scoring model and probe the effect of modeling performance on the credit provider's interest. According to the findings, the new credit scoring approach revealed a reasonable coefficient of determination and an extremely low false negative rate. and Chen (2020) introduced a novel method for credit scoring by maximizing the Kolmogorov-Smirnov (K-S) statistic directly. The K-S statistic is one of the most significant performance evaluation criteria for scoring methods. However, neither of the present scoring methods directly deals with the K-S statistic in the modeling phase. This research proposed a new credit scoring approach that maximized the K-S statistic directly.

## 3-6- Regression analysis

In the regression analysis domain, Dumitrescu et al. (2022) proposed a high-efficient and interpretable credit scoring method, namely Penalized Logistic Tree Regression (PLTR), which employed decision tree information to improve the performance of the logistic regression. The laws extracted from various low-deep decision trees built by initial prediction variables were used as predictors in the PLTR model.

Verster et al. (2019) explored customer credit scoring models based on the credit cards of banks. This research discretized the credit card data of a bank, selected features by calculating the evidence weight and information value and divergence, and employed logistic regression for prediction. Hashemi Taba, Mahfoozi Mousavi and Khatavakhotan (2019) presented a newly developed algorithm with integrated criteria for the dynamic and intelligent rating of bank customers. The purpose of this research was to provide qualitative criteria for identifying the best customer credit rating model centered on financial transfer. Instead of concentrating on customer credit, this study employed the discredit derived from the concepts concerning system quality assurance.

Niu, Ren and Li (2019) embarked on credit scoring through machine learning by combining social network information with evidence from Peer to Peer (P2P) lending, which enabled individuals to raise loans from others directly and removed the intermediacy of financial institutions. The logistic regression results show a statistically significant correlation between social network information and loan default. The results of the random forest, AdaBoost, and LightGBM machine learning algorithms reveal that social network information can improve the prediction performance of loan default considerably. Chen and Xiang (2017) investigated credit scoring models based on the group Lasso regression. The models were developed based on the group Lasso logistic regression, where the tuning parameters of $\lambda$ were selected by the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and cross-validation.

Serrano-Cinca and Gutiérrez-Nieto (2016) used profit scoring as an alternative to credit scoring systems in P2P lending. The analyzed sample revealed that lenders who chose loans using a profit scoring system through multivariate regression outperformed those who employed a conventional credit scoring system based on logistic regression. Fernandes and Artes (2016) examined spatial dependence among credit risks and its improvement in credit scoring. From a dataset with the localization and default information of nine million Brazilian small and medium enterprises, the authors suggested a criterion of the local default risk according to the utilization of ordinary kriging. This variable was included in logistic credit scoring models as an explanatory variable.

### 3-7- Decision support systems

In the decision support systems domain, Zhou et al. (2023) designed an expert system for default prediction the credit scoring s of small firms in China. In this paper used SMOTE to deal with the imbalanced data and employ random forest to build predictive credit features. Tezerjan, Samghabadi and Memariani (2021) presented a hybrid model called ARF for credit scoring in complex systems. They used 5C criteria, including character, capacity, capital, collateral, and conditions, to score customers. The hybrid model could detect and predict the shocks of various sectors of the stock market based on ANFIS and RNN by employing historical data and indices. Then, the outcomes, together with other customer criteria, were inserted into a Fuzzy Rule Base (FRB) that finalized customers' scores. Abbasi Astamal and Rahimi (2019) designed an expert credit rating system for actual bank customers using fuzzy neural networks. The structured model developed by confirmatory analysis determined and classified input variables to the system. Then, a fuzzy expert system with six stages was modeled. For this purpose, the researchers designed a fuzzy system with financial capacity, support, reliability, and repayment records as its inputs and customer credit as its output.

### 3-8- Graph analysis

In the graph analysis domain, Jiang et al. (2023) compared the traditional models and generative adversarial networks (GANs) in solving the class imbalance problem of credit scoring. Paraiso et al. (2021) used network features for credit scoring in micro financing. Microfinance institutions' use of non-conventional data for credit scoring is highly beneficial when customers lack confirmable credit histories. In this respect, this paper relied on the data obtained from smart phones. The researchers created a network that connected a certain user to his/her phone contacts who were the users of a certain mobile phone application by observing ethical issues and protecting privacy. Furthermore, some feature extraction techniques, such as an introduction to centrality, were utilized. Óskarsdóttir et al. (2019) investigated the value of big data for credit scoring by improving financial inclusion through mobile phone data and social network analytics. To this end, they employed a unique mix of datasets, encompassing the records of contact details, information on credit accounts, and debts of customers, to create scorecards for credit card applicants. Table (A1) in the appendix A shows the most significant strengths and weaknesses of the credit rating models extracted from the topic literature.

## 4- Data analysis and discussion

The literature on credit rating is rich in terms of its availability and diversity. All examined papers in this review have been written systematically for the purpose of developing credit rating models. The present review ran a bibliographic analysis and investigated the collected literature with regard to the following headings.

**Table 2.** Analyzed headings

| Row | Heading |
|---|---|
| 1 | Annual Distribution of Published Studies |
| 2 | Spatial Distribution of Published Studies |
| 3 | Industry Considered for Published Studies |
| 4 | Publishers' Distribution for Examined Studies |
| 5 | Theories and Methods Followed by Studies |

### 4-1- Annual distribution of published studies

The annual distribution of the selected research papers published in three Scopus, Web of Science, and Google Scholars databases is shown in figure (2).

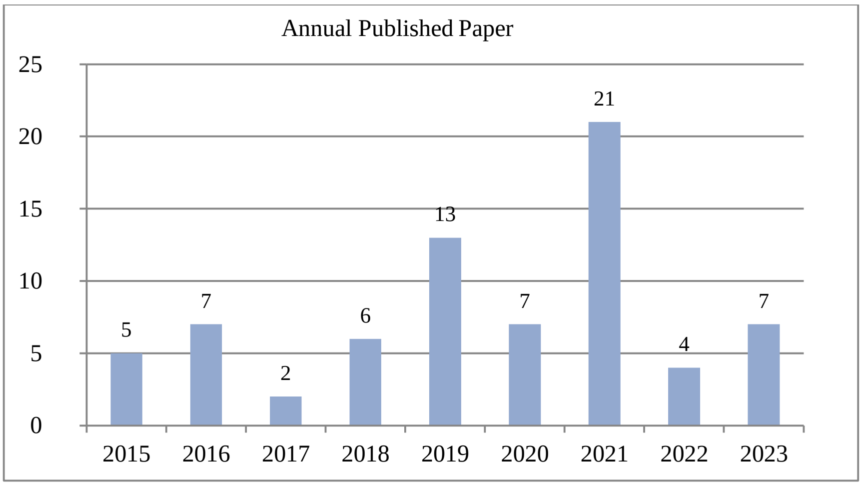


**Fig 2.** Annual Distribution of Published Studies

Almost 74% of the research papers have been published in the 2019-2023 period. An intense literature growth is observed for the year 2019. There is a 116% surge in the research numbers after the 2018 financial crisis. Likewise, the examined papers in 2021 include about 34% of the research literature by themselves.

### 4-2- Spatial distribution of published studies

Figure (3) illustrates the spatial distribution of 72 research papers examined in this section. For analysis, the collected research papers were categorized based on the country the research was implemented and were labeled under the titles of Asia, Europe, America (North and South), Africa, and Australia. Asia and Europe (Eurasia) involved about 81% of the entire examined research. In European and Asian countries, financial concerns and regulations are accompanied by unstable political environments (Ubarhande & Chandani, 2021). This issue develops different attitudes toward regulations. Hence, Eurasia has recorded the highest number of papers in the examined period (2015-2023). Bank regulations, like Basel norms, and financial crises are among the main reasons for meticulously evaluating and investigating credit rates in the mentioned domain. Lending institutions are concerned about credit rates for debt pricing (Tanja et al., 2019).

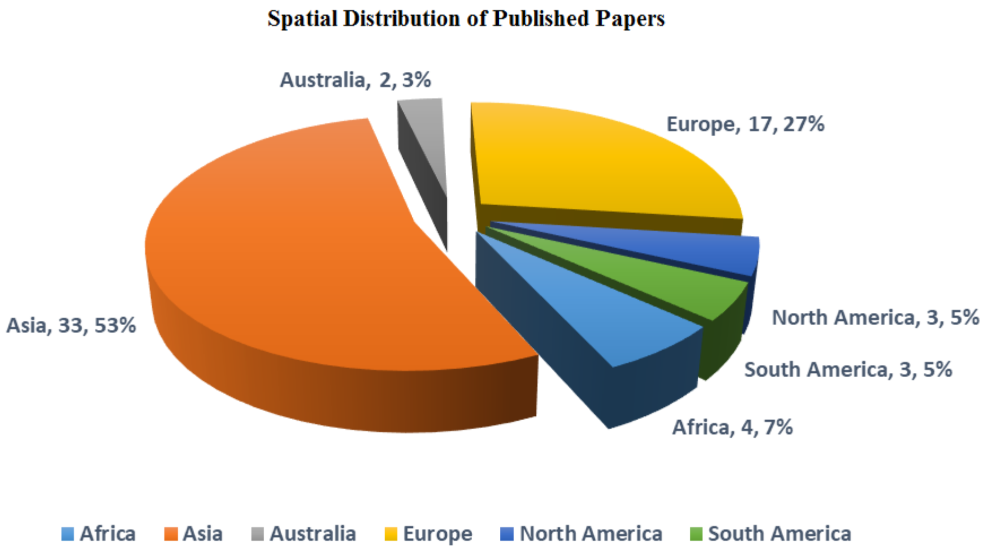


**Fig 3.** Demographic distribution of research papers

### 4-3- Considered industry for published studies

The credit rating mechanism, considered factors, and effect of credit rates differ according to the nature of the examined enterprise or tool. The close analysis of the industry is completely related and crucial since the credit rate is intensely influenced by the nature of the industry and information transfer inside the industry, as noted by Abad, Ferreras and Robles (2020). Figure (4) depicts different industries investigated during the seven-year examined period (2015-2023).

During their studies on credit rating in the past seven years, researchers have preferred the public sector (for banks) to collect and analyze data. It is because numerous studies tackle the databases evaluating credit rates or financial statements. Data collection and analysis become challenging if a sample is filtered for a certain industry, while the scope and application of results can be extended if we consider all industries for data collection and analysis. The second and third sectors privileged for analysis are micro-lending institutions and P2P loans since these sectors are major users of credit-related information.

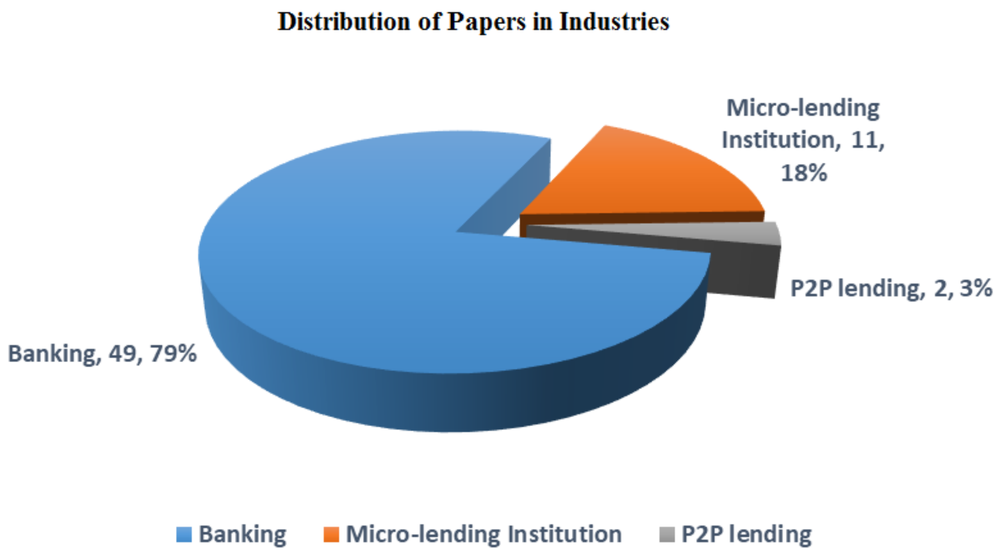


**Fig 4.** Considered industry for research on credit rating

### 4-4- Publishers' distribution for examined studies

Credible international publishers, such as Elsevier and Springer, have published about 48% of the papers. MDPI and the Islamic Azad University are ranked second by publishing 18% of the papers.

Other papers have also been published in popular foreign and domestic journals. Figure (5) compares the number of papers published in these seven years (2015-2023) for different publishers.

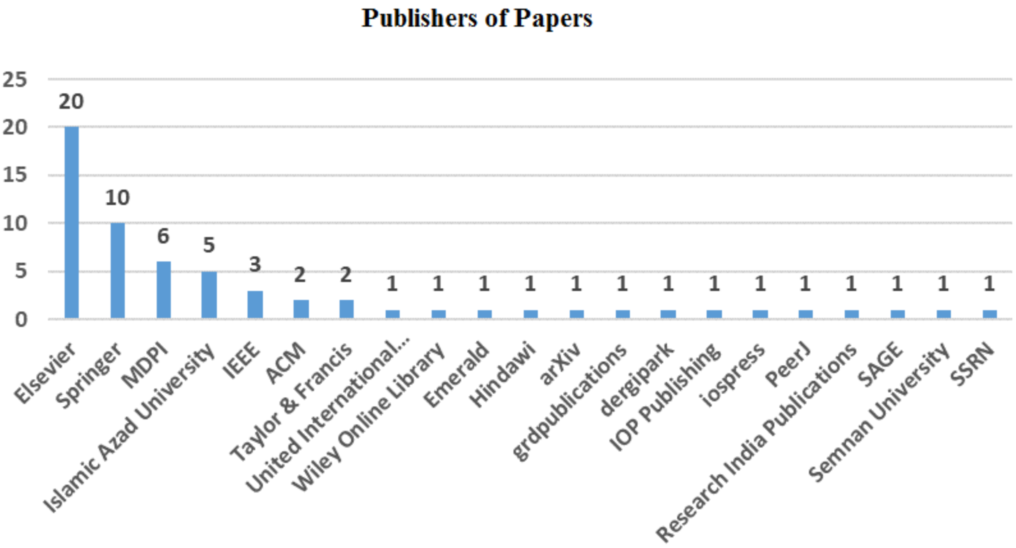


**Fig 5.** Publishers of examined papers on credit rating

## 4-5- Followed theories and methods

It is paramount for an inclusive literature review to analyze the goals and methods employed by researchers. Credit rating agencies mainly consider the financial data of organizations or tools for determining credit rates. During the past seven years, researchers have considered primary and secondary data to rate credits, develop validation models, and validate the scores allocated by credit rating agencies. In table (B1) in the appendix B, the reviewed papers are categorized into 12 main clusters in terms of their methodologies.

# 5- Conclusion and future suggestions

This paper reviewed the literature on credit rating during the 2015-2023 periods. Based on the meticulous analysis of the research papers on credit rating in terms of their purposes, methodologies, resources, collected data, research locations, etc., it was found that many studies tended to develop new frameworks and add them to the existing knowledge base. With technological developments, machine learning (neural networks and ensemble learning), artificial intelligence, and regression have become extensively popular and favored. We discovered that while working on credit rates in the past seven years, 79% of the researchers (59 out of 72 papers) preferred the public sector (for banks) to analyze data. Micro lending institutions and P2P loans allocated 18% (11 out of 72 papers) and 3% (2 out of 72 papers) to themselves, respectively. Since micro lending institutions and P2P lending expand day by day, researchers are expected to pay specific attention to these sectors. The present paper suggests particularly analyzing these classes in all credit rating elements and hybrid models by integrating advanced quantitative techniques with qualitative information and judgments to score credits.

Employing a voluminous dataset, especially for micro lending institutions and P2P loans, may enhance models' performance and provide more accurate estimations. Similarly, we may effectively control outliers while perceiving the constraints of machine learning algorithms. Including temporal aspects of credit risks is another promising direction for future studies. Non-interpretability as the prevalent deficit of ensemble classification is another potential research domain.

Trade-centered policies can be a future direction for evaluating models, provided that credit datasets should be informative in terms of profits and costs. Although numerical improvement is heeded as a valuable innovation, it is recommended that statistical tests be included for the validation of models' performance. The focus on scoring uses, especially with the general University of California, Irvine datasets, may be due to the confidentiality of the credit scoring scope. Recently, datasets of data-mining competitions and online P2P lending platforms have become available. Thus, other types of datasets should play roles in examining models' robustness in various circumstances. Besides, the datasets of online P2P lending platforms provide rich information that can be considered for the development of credit scoring models. A new research trend may also evolve in the future using private data to score profits. Furthermore, it is suggested that UCI datasets shape the bases of comparison due to their general uses in the literature.

## References


Abad, P., Ferreras, R. and Robles, M.D., (2020). Intra-industry transfer effects of credit risk news: Rated versus unrated rivals. *The British Accounting Review*, *52*(1), p.100815.

Abbasi Astamal, M. and Rahimi, R., (2019). Designing an Expert System for Credit Rating of Real Customers of Banks Using Fuzzy Neural Networks. *Advances in Mathematical Finance and Applications*, *4*(1), pp.89-102.

Abdi, F., (2021). A New Multi-Stage Feature Selection and Classification Approach: Bank Customer Credit Risk Scoring. *Journal of Industrial Engineering International*, *17*(1), pp.78-87.

Agrawal, S., Ahirao, P. and Dere, P., (2021), May. Credit Score Evaluation of Customer Using Machine Learning Algorithms. In *Proceedings of the 4th International Conference on Advances in Science & Technology (ICAST2021)*.

Aji, N.A. and Dhini, A., (2019), July. Credit scoring through data mining approach: A case study of mortgage loan in Indonesia. In *2019 16th International Conference on Service Systems and Service Management (ICSSSM)* (pp. 1-5). IEEE.

Ala'raj, M. and Abbod, M.F., (2016). A new hybrid ensemble credit scoring model based on classifiers consensus system approach. *Expert systems with applications*, *64*, pp.36-55.

Ala'raj, M. and Abbod, M.F., (2016). Classifiers consensus system approach for credit scoring. *Knowledge-Based Systems*, *104*, pp.89-105.

Ampountolas, A., Nyarko Nde, T., Date, P. and Constantinescu, C., (2021). A machine learning approach for micro-credit scoring. *Risks*, *9*(3), p.50.

Anderson, R., (2007). *The credit scoring toolkit: theory and practice for retail credit risk management and decision automation*. Oxford University Press.

Anderson, R.A., (2022). *Credit intelligence and modelling: Many paths through the forest of credit rating and scoring*. Oxford University Press.

Ashofteh, A. and Bravo, J.M., (2021). A conservative approach for online credit scoring. *Expert Systems with Applications*, *176*, p.114835.

Ayouche, S., Aboulaich, R. and Ellaia, R., (2017). Partnership credit scoring classification Probem: A neural network approach. *International Journal of Applied Engineering Research*, *12*(5), pp.693-704.

Babaev, D., Savchenko, M., Tuzhilin, A. and Umerenkov, D., (2019), July. Et-rnn: Applying deep learning to credit loan applications. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 2183-2190).

Bolton, C., (2010). *Logistic regression and its application in credit scoring* (Doctoral dissertation, University of Pretoria).

Bunker, R.P., Zhang, W. and Naeem, M.A., (2016). Improving a credit scoring model by incorporating bank statement derived features. *arXiv preprint arXiv:1611.00252*.

Chai, N., Wu, B., Yang, W. and Shi, B., (2019). A multicriteria approach for modeling small enterprise credit rating: evidence from China. *Emerging Markets Finance and Trade*, *55*(11), pp.2523-2543.

Chen, H. and Xiang, Y., (2017). The study of credit scoring model based on group lasso. *Procedia computer science*, *122*, pp.677-684.

Chopra, A. and Bhilare, P., (2018). Application of ensemble models in credit scoring models. *Business Perspectives and Research*, *6*(2), pp.129-141.

Dadmohammadi, D. and Ahmadi, A., (2019). A Combined Learning Approach for Credit Scoring Using Adaptive Hierarchical Mixture of Experts: Iranian Banking Industry. *AUT Journal of Modeling and Simulation*, *51*(1), pp.67-80.

Dastile, X., Celik, T. and Potsane, M., (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, *91*, p.106263.

Djeundje, V.B., Crook, J., Calabrese, R. and Hamid, M., (2021). Enhancing credit scoring with alternative data. *Expert Systems with Applications*, *163*, p.113766.

Dumitrescu, E., Hué, S., Hurlin, C. and Tokpavi, S., (2022). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research*, *297*(3), pp.1178-1192.

Dong, Y., Gou, Q., Qiu, H., (2023). Big tech credit score and default risk ——Evidence from loan-level data of a representative microfinance company in China. *China Economic Review, 81, pp. 102010.*

Ehtesham Rasi, R., Karamipour, M. and Arad, M., (2020). Rating the Actual Customers of Banks based on Credit Risk using Multiple Criteria Decision Making and Artificial Intelligence Hyperbolic Regression. *International Journal of Finance & Managerial Accounting*, *4*(16), pp.51-63.

Fang, F. and Chen, Y., (2019). A new approach for credit scoring by directly maximizing the Kolmogorov–Smirnov statistic. *Computational Statistics & Data Analysis*, *133*, pp.180-194.

Fernandes, G.B. and Artes, R., (2016). Spatial dependence in credit risk and its improvement in credit scoring. *European Journal of Operational Research*, *249*(2), pp.517-524.

Fernando, E. and Siagian, P., (2021). Proposal to use the analytic hierarchy process method evaluate bank credit submissions. *Procedia Computer Science*, *179*, pp.232-241.

Gao, Y., Xiao, H., Zhan, Zh., Cai, W., and Hu, X., (2023). CATE: Contrastive augmentation and tree-enhanced embedding for credit scoring. *Information Sciences, In press. (* https://doi.org/10.1016/j.ins.2023.119447)

Hashemi Taba, N., Mahfoozi Mousavi, S.K. and Khatavakhotan, A.S., (2019). A novel algorithm developed with integrated metrics for dynamic and smart credit rating of bank customers. In *Fundamental Research in Electrical Engineering: The Selected Papers of The First International Conference on Fundamental Research in Electrical Engineering* (pp. 787-799). Springer Singapore.

Harris, T., (2015). Credit scoring using the clustered support vector machine. *Expert Systems with Applications*, *42*(2), pp.741-750.

Hastie, T., Tibshirani, R., Friedman, J.H. and Friedman, J.H., (2009). The elements of statistical learning: data mining, inference, and prediction. *New York: springer*, 2, pp. 1-758.

Heler, V.G., Filomena, T.P., Ferreria, L., and Kirch, G., (2022). Application of the VNS heuristic for feature selection in credit scoring problems, *Maching Learning with Applications*, 9, pp. 100349.

Hovdenakk, A.H., (2021). *Machine learning vs logistic regression in credit scoring: A trade-off between accuracy and interpretability?* (Master's thesis, The University of Bergen).

Ignatius, J., Hatami-Marbini, A., Rahman, A., Dhamotharan, L. and Khoshnevis, P., (2018). A fuzzy decision support system for credit scoring. *Neural Computing and Applications*, *29*, pp.921-937.

Jiang, C., Lu, W., Wang, Zh., and Ding, Y., (2023). Benchmarking state-of-the-art imbalanced data learning approaches for credit scoring. *Expert System with Applications*, 213, pp. 118878.

Jadhav, S., He, H. and Jenkins, K., (2018). Information gain directed genetic algorithm wrapper feature selection for credit rating. *Applied Soft Computing*, *69*, pp.541-553.

Javadpour, A., Saedifar, K., Wang, G., Li, K.C. and Saghafi, F., (2021). Improving the efficiency of customer's credit rating with machine learning in big data cloud computing. *Wireless Personal Communications*, *121*(4), pp.2699-2718.

Kang, Y., Jia, N., Cui, R. and Deng, J., (2021). A graph-based semi-supervised reject inference framework considering imbalanced data distribution for consumer credit scoring. *Applied Soft Computing*, *105*, p.107259.

Karamizadeh, S., Abdullah, S.M., Zamani, M. and Kherikhah, A., (2015). Pattern recognition techniques: studies on appropriate classifications. In *Advanced Computer and Communication Engineering Technology: Proceedings of the 1st International Conference on Communication and Computer Engineering* (pp. 791-799). Springer International Publishing.

Katsimperis, V. and Andrikopoulos, A., (2021). Creating a flexible business credit rating model using multicriteria decision analysis. *Journal of Multi‐ Criteria Decision Analysis*, *28*(1-2), pp.45-67.

Kiesel, F., (2016). Do investors still rely on credit rating agencies? Evidence from the financial crisis. *The Journal of Fixed Income*, *25*(4), pp.20-31.

Kiruthika and Dilsha, M., (2015). A neural network approach for microfinance credit scoring. *Journal of Statistics and Management Systems*, *18*(1-2), pp.121-138.

Koutanaei, F.N., Sajedi, H. and 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*, pp.11-23.

Kozeny, V., (2015). Genetic algorithms for credit scoring: Alternative fitness function performance comparison. *Expert Systems with applications*, *42*(6), pp.2998-3004.

Laborda, J. and Ryoo, S., (2021). Feature selection in a credit scoring model. *Mathematics*, *9*(7), p.746.

Li, B., Xiao, B. and Yang, Y., (2021). Strengthen credit scoring system of small and micro businesses with soft information: Analysis and comparison based on neural network models. *Journal of Intelligent & Fuzzy Systems*, *40*(3), pp.4257-4274.

Li, X., and Sun, Y., (2021). Application of RBF neural network optimal segmentation algorithm in credit rating. *Neural Computing and Applications*, *33*, pp.8227-8235.

Li, Y., and Chen, W., (2020). A comparative performance assessment of ensemble learning for credit scoring. Mathematics 8: 1756.

Liu, W., Fan, H., Xia, M., and Xia, M., (2022). A focal-aware cost-sensitive boosted tree for imbalanced credit scoring. *Expert Systems with Applications*, 208, pp.118158.

Merćep, A., Mrčela, L., Birov, M. and Kostanjčar, Z.,(2020). Deep neural networks for behavioral credit rating. *Entropy*, *23*(1), p.27.

Mukid, M.A., Widiharih, T., Rusgiyono, A. and Prahutama, A., (2018), May. Credit scoring analysis using weighted k nearest neighbor. In *Journal of Physics: Conference Series* (Vol. 1025, No. 1, p. 012114). IOP Publishing.

Müller, A.C. and Guido, S., (2016). *Introduction to machine learning with Python: a guide for data scientists*. " O'Reilly Media, Inc.".

Nazari, A., Mehregan, M. and Tehrani, R., (2020). Evaluating the Effectiveness of Data Mining Techniques in Credit Scoring of Bank Customers Using Mathematical Models: A Case Study of Individual Borrowers of Refah Kargaran Bank in Zanjan Province, Iran. *International Journal of Nonlinear Analysis and Applications*, *11*(Special Issue), pp.299-309.

Nazari, A., Mehregan, M. and Tehrani, R., (2019). Using the Hybrid Model for Credit Scoring (Case Study: Credit Clients of microloans, Bank Refah-Kargeran of Zanjan, Iran). *Journal of Optimization in Industrial Engineering*, *12*(2), pp.65-78.

Niu, B., Ren, J. and Li, X., (2019). Credit scoring using machine learning by combing social network information: Evidence from peer-to-peer lending. *Information*, *10*(12), p.397.

Óskarsdóttir, M., Bravo, C., Sarraute, C., Vanthienen, J. and Baesens, B., (2019). The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics. *Applied Soft Computing*, *74*, pp.26-39.

Pang, S., Wang, S. and Xia, L., (2020). Farmer's Credit Rating Model and Application Based on Multilayer Unified Network with Linear Classifier. *Complexity*, *2020*, pp.1-13.

Pramod, C.P. and Pillai, G.N., (2021). K-Means clustering based Extreme Learning ANFIS with improved interpretability for regression problems. *Knowledge-Based Systems*, *215*, p.106750.

Paraíso, P., Ruiz, S., Gomes, P., Rodrigues, L. and Gama, J., (2021). Using network features for credit scoring in microfinance. *International Journal of Data Science and Analytics*, *12*, pp.121-134.

Pławiak, P., Abdar, M., Pławiak, J., Makarenkov, V. and Acharya, U.R., (2020). DGHNL: A new deep genetic hierarchical network of learners for prediction of credit scoring. *Information Sciences*, *516*, pp.401-418.

Qian, H., Ma, P., Gao, S., and Song, Y., (2023). Soft reordering one-dimensional convolutional neural network for credit scoring. *Knowldage –Based Systems,* 266, pp.110414.

Qiu, W., (2019), July. Credit risk prediction in an imbalanced social lending environment based on XGBoost. In *2019 5th International Conference on Big Data and Information Analytics (BigDIA)* (pp. 150-156). IEEE.

Sangiorgi, F. and Spatt, C., (2017). The economics of credit rating agencies. *Foundations and Trends® in Finance*, *12*(1), pp.1-116.

Serrano-Cinca, C. and Gutiérrez-Nieto, B., (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. *Decision Support Systems*, *89*, pp.113-122.

Shema, A., (2019), January. Effective credit scoring using limited mobile phone data. In *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development* (pp. 1-11).

Shen, K.Y., Sakai, H. and Tzeng, G.H., (2019). Comparing two novel hybrid MRDM approaches to consumer credit scoring under uncertainty and fuzzy judgments. *International Journal of Fuzzy Systems*, *21*, pp.194-212.

Shi, J. and Xu, B., (2016). Credit scoring by fuzzy support vector machines with a novel membership function. *Journal of Risk and Financial Management*, *9*(4), p.13.

Singh, D.K. and Goel, N., (2021). Bank Customer's Credit Score Prediction Using Feature Selection and Data Mining Algorithm. In *Progress in Advanced Computing and Intelligent Engineering: Proceedings of ICACIE 2020* (pp. 889-897). Springer Singapore.

Singh, I., Kumar, N., Srinivasa, K.G., Maini, S., Ahuja, U. and Jain, S., (2021). A multi-level classification and modified PSO clustering based ensemble approach for credit scoring. *Applied Soft Computing*, *111*, p.107687.

Soydaner, D. and Kocadağlı, O., (2015). Artificial neural networks with gradient learning algorithm for credit scoring. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, *44*(2), pp.3-12.

SULTANA, R., (2018). *An Evaluation of Automated Credit Scoring System for Financial Services in Developing Countries* (Doctoral dissertation, United International University).

Tezerjan, M.Y., Samghabadi, A.S. and Memariani, A., (2021). ARF: A hybrid model for credit scoring in complex systems. *Expert Systems with Applications*, *185*, p.115634.

Thomas, L., Crook, J. and Edelman, D., (2017). *Credit scoring and its applications*. Society for industrial and Applied Mathematics.

Tomczak, J.M. and Zięba, M., (2015). Classification restricted Boltzmann machine for comprehensible credit scoring model. *Expert Systems with Applications*, *42*(4), pp.1789-1796.

Tran, K.Q., Duong, B.V., Tran, L.Q., Tran, A.L.H., Nguyen, A.T. and Nguyen, K.V., (2021). Machine learning-based empirical investigation for credit scoring in Vietnam's banking. In *Advances and Trends in Artificial Intelligence. From Theory to Practice: 34th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2021, Kuala Lumpur, Malaysia, July 26–29, 2021, Proceedings, Part II 34* (pp. 564-574). Springer International Publishing.

Tripathi, D., Edla, D.R., Bablani, A., Shukla, A.K. and Reddy, B.R., (2021). Experimental analysis of machine learning methods for credit score classification. *Progress in Artificial Intelligence*, *10*, pp.217-243.

Uthayakumar, J., Vengattaraman, T. and Dhavachelvan, P., (2020). Swarm intelligence based classification rule induction (CRI) framework for qualitative and quantitative approach: An application of bankruptcy prediction and credit risk analysis. *Journal of King Saud University-Computer and Information Sciences*, *32*(6), pp.647-657.

Ubarhande, P. and Chandani, A., (2021). Elements of credit rating: a hybrid review and future research Agenda. *Cogent Business & Management*, *8*(1), p.1878977.

Vanara, R., Wani, P., Pawar, S., More, P. and Patil, P., (2021). Predication approval for bank loan using random forest algorithm. *International Journal of Progressive Research in Science and Engineering*, *2*(7), pp.137-142.

Verster, T., De Jongh, R., Greenberg, S., Fourie, E. and de Wet, D., (2019). A motivation for banks in emerging economies to adapt agency ratings when assessing corporate credit. *South African Journal of Economic and Management Sciences*, *22*(1), pp.1-11.

Wang, M. and Yang, H., (2020). Research on Customer Credit Scoring Model Based on Bank Credit Card. In *Intelligent Information Processing X: 11th IFIP TC 12 International Conference, IIP 2020, Hangzhou, China, July 3–6, 2020, Proceedings 11* (pp. 232-243). Springer International Publishing.

Wickens, M., (2016). The eurozone financial crisis: debt, credit ratings and monetary and fiscal policy. *Empirica*, *43*, pp.219-233.

Wu, Y., Huang, W., Tian, Y., Zhu, Q., Yu, L., (2022). An uncertainty-oriented cost-sensitive credit scoring framework with multi-objective feature selection, *Electronic Commerce Research and Applicatopns*, 53, pp.101155.

Xiao, H., Xiao, Z. and Wang, Y., (2016). Ensemble classification based on supervised clustering for credit scoring. *Applied Soft Computing*, *43*, pp.73-86.

Xiao, J., Zhong, *Y*., Jia, Y., Wang, Y., Li, R., Jiang, X., and Wang, Sh., (2023). A novel deep ensemble model for imbalanced credit scoring in internet finance. *International Journal of Forcasting*, *In press, https://doi.org/10.1016/j.ijforecast.2023.03.004*.

Yotsawat, W., Wattuya, P. and Srivihok, A., (2021). A novel method for credit scoring based on cost-sensitive neural network ensemble. *IEEE Access*, *9*, pp.78521-78537.

Zhang, Y. and Chi, G., (2018). A credit rating model based on a customer number bell-shaped distribution. *Management Decision*, *56*(5), pp.987-1007.

Zhang, Z., Li, Y., Li, Y., and Liu, S., (2023). A Local binary social spider algorithm for feature selection in credit scoring model. *Applied Soft Computing* , 144, pp. 110549.

Zhou, X., Cheng, S., Zhu, M., Guo, C., Zhou, S., Xu, P., Xue, Z. and Zhang, W., (2018). A state of the art survey of data mining-based fraud detection and credit scoring. In *MATEC Web of Conferences* (Vol. 189, p. 03002). EDP Sciences.

Zhou, Y., Shen, L., and Ballester, L., (2023). A two-stage credit scoring model based on random forest: Evidence from Chinese small firms. *International Review of Financial Analysis*, *89*, pp.102775.

# Appendix A:

**Table A1.** Most significant strengths and weaknesses of selected rating models

| Model | Cluster | Strengths | Weaknesses | Reference(s) |
|---|---|---|---|---|
| Statistical | Logistic regression | -It is relatively accurate in validation.<br>-Sample classification<br>-It is possible to convert coefficients to scorecard formats. | - Complex and noisy data create deficiency. They may not appear as complex patterns and relations like other machine learning models.<br>-Logistic regression needs data preprocessing.<br>-It cannot manage missed data, and we should convert some variables to obtain their nonlinear effects. | (Anderson, 2007)<br>(Hovdenakk, 2021) |
| Machine learning | K-nearest neighbor | -It is a simple and easy machine-learning model.<br>-It is explainable to non-specialists.<br>-Dynamically updating train samples by adding new cases to them<br>-It is relatively accurate.<br>-There are no assumptions on data structures.<br>-It is influenced by outliers. | -It may not discover convoluted patterns recognized by other complex models.<br>-It needs data processing.<br>-With big datasets, the training process can be time-consuming.<br>-If the train data are big, it will take a long time for the algorithm to search the whole train data.<br>-It cannot score the features of every special applicant. A large volume of memory and CPU is considered since the standard method preserves and reuses (almost) the entire train data.<br>-It breaks down in the presence of irrelevant r big data. | (Müller & Guido, 2017), (Bolton, 2010), and Thomas et al., 2017) |
| | Decision tree | -It is easy to comprehend and interpret the results of decision trees.<br>-Exerting control through variables that separate the proper predicted sample from the improper sample<br>-The algorithm is fast during training.<br>-It does not need data preprocessing. | -Overfitting arises by building deep and complex trees.<br>-Determining maximal depths for trees reduces the accuracy of train data.<br>-If the data are structurally complex, the decision tree should build deep trees to classify the target variable.<br>-A extremely deep and complex tree can overfit the train data.<br>-The tree is exposed to severe changes if a small shift occurs in train data.<br>-A small change can lead to intense variations in branches. Thus, decision trees experience high variance. | (Müller & Guido, 2017), (Hastie et al., 2009), and (Dastile et al., 2020) |
| | Support Vector Machine (SVM) | -It does not need data structure assumptions in non-parametric conditions.<br>- It is a novel data-mining method that can be used by optimization approaches for machine learning solutions.<br>-It focuses on the statistical learning principle. | -It is difficult to interpret features and standard formulas that lack commercial constraints and characteristics. | (Oreski, 2014)<br>(Harris, 2015) |
| | Artificial Neural Networks (ANNs) | -They are convincing tools for uncertain data<br>-Prediction accuracy privilege of ANNSs compared to statistical methods<br>-Generalizability, robustness, and the lack of any explicit problem explanation<br>-Capacity to manage a large volume of data<br>-Needing fewer statistical assumptions<br>-A nonlinear and non-parametric method | -Problems applied to small samples may possess irrelevant features.<br>-Long training and selection time is required.<br>-Overfitting arises when they are applied to voluminous datasets.<br>- The interpretation and trial and error processes are difficult in some studies.<br>-They are computationally dense and need numerous repetitions in pursuit of the ultimate solution.<br>-It is difficult and costly to maintain them, especially during continuous training for their up-to-dateness.<br>-They are ambiguous since interpreting relations is extremely difficult. | (Karamizadeh et al., 2015) |

**Table A1.** Continued

| Model | Cluster | Strengths | Weaknesses | Reference(s) |
|---|---|---|---|---|
| Ensemble learning | Random forest | -The high variance problem of decision trees is reduced by the Bagging technique.<br>-It can record complex structures, such as nonlinear relations, in data and prevent overfitting.<br>-It is stronger against variations in train data.<br>-The overfitting problem is decreased since models are developed based on countless random trees.<br>-It can manage outliers. | -Random forests create many convoluted trees.<br>-It is difficult to interpret why the random forest classifies loanees as proper or improper samples.<br>-They are the best only when prediction is the primary objective, and there is no need to comprehend basic relations. | (Hastie et al., 2009) and (Müller & Guido, 2017) |
| | Gradient boosting trees | -They do not create convoluted trees like the random forest<br>-Their performance is improved by the combination of several trees.<br>-They do not need to preprocess data.<br>-They decrease the overfitting probability of the train data by creating various trees. | -The algorithm needs parameter tuning.<br>-It may be slow at train time.<br>Similar to the random forest, it is not easy to interpret why this model classifies loanees as proper or improper samples.<br>When estimating the significance of a variable for this model, we can observe that some variables are ignored. | (Müller & Guido, 2017) |
| Evolutionary algorithm | Deep Genetic Hierarchical Network (DGHN) | -They manifest the best prediction performance for known credit scoring datasets. | -Similar to any other deep learning models, it requires long-term training and optimization for creating a convoluted system structure. | (Plawiak et al., 2020) |
| Graph analysis | Graph analysis | -It is easily implemented.<br>-Analysts simply comprehend the visual nature of results. | It needs high computational power for training and operation. | (Zhou et al., 2018) |
| Decision support systems | Fuzzy neural networks | -They are suitable methods for solving nonlinear problems.<br>-They are accurate measurement tools for uncertain and nonlinear concepts.<br>-They benefit from the training power of neural networks and the linguistic advantage of fuzzy systems.<br>-Neural networks are used to arrange data and recognize patterns. | -With an increase in the number of features, the network entrance partitioning exponentially increases the number of rules in the fuzzy inference systems.<br>-Partitioning the network entrance reduces its interpretability and raises the computational load. | (Pillai & Pramod, 2021) |

# Appendix B:

**Table B1.** Methodologies employed for developing credit rating models

| Main cluster | Methodology | No. of papers | Key authors |
|---|---|---|---|
| Machine Learning | SVM<br>Fuzzy SVM<br>KNN<br>K-means<br>Hierarchical Clustering<br>PCA<br>Factorization Machines<br>Fuzzy C-means<br>CNN | 30 | (Qian et al., 2023),(Javadpout et al., 2021), (Merćep et al., 2020), (Singh et al., 2021), (Nazai et al., 2020), (Shen et al., 2019), (Sultana, 2018), (Jadhav et al., 2018), (Shi & Xu, 2016), (Bunker et al., 2016), (Agrawal et al., 2021), (Paraíso et al., 2021), (Tripathi et al., 2021), (Laborda & Ryoo, 2021), (Li & Chen, 2020), (Pławiak et al., 2020), (Koutanaei et al., 2015), (Ala'raj & Abbod (a), 2016), (Xiao et al., 2016), (Ala'raj & Abbod (b), 2016), (Abdi, 2021), (Ampountolas et al., 2021), (Nazari et al., 2019), (Mukid et al., 2018), (Qiu, 2019), (Djeundj et al., 2021), (Li et al., (a) 2021), (Pławiak et al., 2020), (Li et al., (b), 2021), (Chai et al., 2019) |
| Decision Tree | CART<br>ID3 algorithm<br>ID5 algorithm<br>CATE<br>LightGBM-focal | 18 | (Gao et al., 2023),(Liu et al., 2022),(Javadpour et al., 2021), (Nazari et al., 2019), (Koutanaei et al., 2015), (Abdi, 2021), (Ampountolas et al., 2021), (Singh et al., 2021), (Nazai et al., 2020), (Aji & Dhini, 2019), (Sultana, 2018), (Bunker et al., 2016), (Agrawal et al., 2021), (Paraíso et al., 2021), (Tripathi et al., 2021), (Li & Chen, 2020), (Chopra & Bhilar, 2018), (Ala'raj & Abbod (a), 2016), (Ala'raj & Abbod (b), 2016) |
| Evolutionary Algorithms | Genetic Algorithm<br>PSO<br>Genetic Programming<br>GA wrapper | 9 | (Abdi, 2021), (Kozeny, 2015), (Pławiak et al., 2020), (Zhang & Chi, 2018), (Koutanaei et al., 2015), (Singh et al., 2021), (Dadmohammadi & Ahmadi, 2019), (Sultana, 2018), (Jadhav et al., 2018) |
| Artificial Neural Network | MLP<br>ANFIS<br>RNN<br>DNN<br>SOM<br>RBF<br>Multilayer Unified Network<br>Laterally Connected Neural Network<br>Learning Vector Quantizer Neural Network<br>Restricted Boltzmann Machine<br>Expert System(SMOTE) | 26 | (Zhou et al., 2023),(Javadpout et al., 2021), (Ampountolas et al., 2021), (Ashofteh & Bravo, 2021), (Singh et al., 2021), (Nazari et al., 2020), (Nazari et al., 2019), (Dadmohammadi & Ahmadi, 2019), (Sultana, 2018), (Ayouche et al., 2017), (Soydaner & Kocadağlı, 2015), (Kiruthika, & Dilsha, 2015), ((Djeundj et al., 2021), (Li et al., (a), 2021), (Tripathi et al., 2021), (Li & Chen, 2020), (Pławiak et al., 2020), (Koutanaei et al., 2015), (Ala'raj & Abbod (a), 2016), (Ala'raj & Abbod (b), 2016), (Tezerjan et al., 2021), (Babaev et al., 2019), (Merćep et al., 2020), (Li et al., (b), 2021), (Li & Sun, 2021), (Pang et al., 2020), (Tomczak & Zięba, 2015) |
| Multi Criteria Decision Making | OWA<br>Fuzzy AHP<br>AHP<br>UTADIS<br>DANP<br>DRSA<br>TOPSIS<br>Multi-objective Programming | 9 | (Javadpout et al., 2021), (Ignatius et al., 2019), (Katsimperis & Andrikopoulos, 2021), (Fernando & Siagian, 2021), (Ehtesham Rasi et al., 2020), (Nazari et al., 2019), (Shen et al., 2019), (Chai et al., 2019), (Zahng & Chi, 2018) |
| Rule-Based Systems | Fuzzy Rule Base | 4 | (Tezerjan et al., 2021), (Abbasi Astamal & Rahimi, 2019), (Sultana, 2018), (Pławiak et al., 2020), |
| Graph Analysis | Personalized PageRank<br>Spreading Activation<br>Node embedding algorithm<br>GAN | 2 | (Jiang, 2023),(Óskarsdóttir et al., 2019), (Paraíso et al., 2019) |
| Probabilistic model | Naive Bayes<br>Bayesian Networks | 12 | (Singh et al., 2021), (Jadhav et al., 2018), (Bunker et al., 2016), (Agrawal et al., 2021), (Paraíso et al., 2019), (Tripathi et al., 2021), (Li & Chen, 2020), (Koutanaei et al., 2015), (Ala'raj & Abbod (a), 2016), (Ala'raj & Abbod (b), 2016), (Nazari et al., 2019) |

**Table B2.** Continued

| Main cluster | Methodology | No. of papers | Key authors |
|---|---|---|---|
| Regression Analysis | Probit Regression<br>Kriging<br>MARS<br>Logistic Regression<br>Linear Discriminant Analysis<br>Hyperbolic Regression<br>Penalised-type Regression<br>Ridge regression<br>Penalized Logistic Tree Regression | 33 | (Chai et al., 2019), (Fernandes & Artes, 2016), (Ala'raj & Abbod (a), 2016), (Ala'raj & Abbod (b), 2016), (Merćep et al., 2020), (Singh & Goel, 2021), (Ashofteh & Bravo, 2021), (Singh et al., 2021), (Nazari et al., 2020), (Babaev et al., 2019), (Wang & Yang, 2020), (Óskarsdóttir et al., 2019), (Niu et al., 2019), (Sultana, 2018), (Chen & Xiang, 2017), (Kiruthika, & Dilsha, 2015), (Serrano-Cinca & Gutiérrez-Nieto, 2016), (Bunker et al., 2016), (Agrawal et al., 2021), (Paraíso et al., 2021), (Djeundj et al., 2021), (Li et al., (a) 2021), (Tripathi et al., 2021), (Laborda & Ryoo, 2021), (Li & Chen, 2020), (Fang & Chen, 2019), (Fernandes & Artes, 2016), (Ala'raj & Abbod (a), 2016), (Xiao et al., 2016), (Ehtesham Rasi et al., 2020), (Hashemi Taba et al., 2019), (Damirchi et al., 2022) |
| Ensemble learning | Random Forest<br>XGBoost<br>AdaBoost<br>Extra Trees Classifier<br>LightGBM<br>CatBoost<br>Soft Majority Voting<br>Stacking<br>GBDT<br>Random Subspace Method<br>Boosting<br>Bagging<br>VAE-DF | 26 | (Xiao et al., 2023), (Merćep et al., 2020), (Ampountolas et al., 2021), (Singh & Goel, 2021), (Tran et al., 2021), (Vanara et al., 2021), (Singh et al., 2021), (Kang et al., 2021), (Shema, 2019), (Aji & Dhini, 2019), (Niu et al., 2019), (Bunker et al., 2016), ((Agrawal et al., 2021), (Paraíso et al., 2021), (Tripathi et al., 2021), (Laborda & Ryoo, 2021), (Chopra & Bhilare, 2018), (Koutanaei et al., 2015), (Ala'raj & Abbod (a), 2016), (Ala'raj & Abbod (b), 2016), (Qiu, 2019), (Djeundj et al., 2021), (Li & Chen, 2020), (Babaev et al., 2019), (Singh et al., 2021), (Li et al., (b), 2021), (Xiao et al., 2016) |
| Statistical Tests | Kruskal-Wallis statistic<br>Kolmogorov-Smirnov Statistic | 2 | (Ashofteh & Bravo, 2021), (Fang & Chen, 2019) |
| Other Quantitative Methods | WLC<br>DEA<br>WOE<br>Information Value<br>Information Entropy<br>Expected Maximum Profit<br>HME<br>Forward Stepwise Selection<br>Backward Stepwise Selection<br>LBSA | 8 | (Zhang et al., 2023),(Nazari et al., 2020), (Wang & Yang, 2020), (Paraíso et al., 2021), (Jadhav et al., 2018), (Koutanaei et al., 2015), (Óskarsdóttir et al., 2019), (Dadmohammadi & Ahmadi, 2019), (Laborda & Ryoo, 2021) |