ISSN Online: 2821-1936

Transactions on Data Analysis in Social Science

Journal Homepage: https://transoscience.ir

Evaluation and Comparison of Classification Model Performance in Predicting Corporate Credit Ratings Using Artificial Intelligence: A Case Study of the Tehran Stock Exchange

Sh. Shaghaghi Shahri[1], O. Rahmani Seryasat[2,*]

[1] Department of Financial Management, Faculty of Management, Tehran East Branch, Islamic Azad University, Tehran, Iran
[2] Assistant Professor, Department of Electrical Engineering, Shams Higher Education Institute, Gorgan, Iran

ARTICLE INFO

Article History:
Received 15 February 2024
Received in revised form 22 March 2024
Accepted 12 June 2024
Available online 20 June 2024

Keywords:
Corporate Credit Ratings, Support Vector Machine, Neural Network, k-Nearest Neighbors, Decision Tree, Performance Evaluation, Tehran Stock Exchange

ABSTRACT

This article examines and evaluates the performance of four classification models in predicting corporate credit ratings. The models under study include Support Vector Machine (SVM), Artificial Neural Network (Neural Network), k-Nearest Neighbors (KNN), and Decision Tree. The data used includes features such as exports, company age, production volume, external auditing, foreign ownership, ownership type, and company size, all extracted from the financial statements of companies listed on the Tehran Stock Exchange. The data was divided into training and testing sets, standardized, and then used for training and evaluating the models. The performance of the models was assessed based on accuracy, precision, recall, ROC AUC score, and confusion matrix. The results indicate that the Decision Tree model, with an accuracy of 1.000 and an ROC AUC score of 1.000, exhibited the best performance in predicting corporate credit ratings. The SVM and Neural Network models demonstrated very good performance with an accuracy of 0.995 and an ROC AUC score of 0.999. The KNN model showed acceptable performance with an accuracy of 0.990 and an ROC AUC score of 0.993. This study demonstrates that classification models can effectively aid in predicting corporate credit ratings, with the Decision Tree model being identified as the best option in this context.

1. INTRODUCTION

One of the most important sources of short-term financing is trade credit. Trade credit is influenced by a variety of factors that affect both the availability of financial resources and the quality of corporate management. This study aims to examine the factors influencing trade credit among firms listed on the Tehran Stock Exchange. Given the vital role of trade credit in corporate financing and its impact on economic performance, identifying the key determinants of trade credit is essential. These factors may include export activities, company age, production volume, external auditing, foreign ownership, ownership structure, and firm size. The primary objective of this research is to identify and analyze these factors and provide strategic recommendations for enhancing trade credit

* Corresponding Author: omid.seryasat@gmail.com
Assistant Professor, Department of Electrical Engineering, Shams Higher Education Institute, Gorgan, Iran

http://dx.doi.org/10.47176/TDASS.2024.109

among companies. The significance of this study lies in the fact that trade credit plays a crucial role in improving firms' financial flexibility and liquidity.

1.1. Problem Statement

Improving trade credit can enable firms to operate under better financial conditions and enhance their competitiveness in the market. This research can also serve as a guide for corporate managers and economic policymakers aiming to improve firms' financial performance. Trade credit is a critical short-term financing resource that allows firms to postpone payments for goods and services. It depends on factors such as financial strength, sales volume, efficient management, and the level of trust in the company. Among various financing methods, trade credit stands out as one of the most significant, enabling companies to procure raw materials or essential goods without the need for immediate cash payments. This approach is vital for both large and small businesses and can significantly influence their success or failure.

Traditionally, credit rating assessments of firms have relied on subjective expert opinions, often influenced by personal judgment and emotion. There is an evident need for a more objective and data-driven approach to estimating firms' credit ratings using economic indicators. In this study, we propose an intelligent model that utilizes artificial intelligence to estimate firms' credit ratings. Expert knowledge was employed to identify relevant features influencing credit scores, which were subsequently used in AI-based models for estimation. Python programming language was used for implementing the model.

1.1.1. Research Questions

1. What are the key parameters influencing the trade credit of companies?
2. What is the most effective artificial intelligence method for predicting firms' trade credit?
3. What are the evaluation criteria for AI models in predicting corporate trade credit?

This study aims to develop an artificial intelligence-based model for predicting corporate trade credit. Four machine learning methods were employed, and the best-performing approach was selected based on established evaluation metrics. The statistical population comprises 115 companies listed on the Tehran Stock Exchange during the years 2014 to 2021.

1.2. Factors Affecting Corporate Financing

Export Activity: Exporting firms generally enjoy better trade credit due to their access to international markets and foreign currency earnings.

Firm Age: Older firms tend to have a more established financial history, increasing lender confidence.

Production Volume: Firms with higher production volumes are typically better equipped to meet market demands, which can enhance their trade credit.

External Auditing: Independent auditing enhances the reliability of financial statements, thereby improving creditworthiness.

Foreign Ownership: Firms with foreign investment often have greater access to financial resources and support.

Ownership Structure: Whether a company is privately or publicly owned can influence the level of trade credit it receives.

Firm Size: Larger firms generally have more financial and human capital, enabling them to secure greater trade credit.

Trade credits can be categorized into different types depending on the company's activities and financial needs. Common types include:

- **Consumer Credit**: Provided to end users to purchase goods and services on an installment basis.
- **Commercial (Trade) Credit**: Offered to businesses to acquire raw materials or products needed for operations.
- **Bank Credit**: Short-term financial support provided by banks to meet immediate funding needs.

Credit rating is a vital tool for assessing a company's ability to repay its debts. It is typically conducted by financial and credit institutions and helps companies understand and improve their credit risk profiles.

2. RESEARCH BACKGROUND

Corporate credit ratings are critical for assessing the financial risk of companies, enabling financial institutions to make informed lending and investment decisions. Traditionally, credit rating prediction relied on manual evaluations and statistical models such as logistic regression and discriminant analysis, which often struggled to capture complex, non-linear relationships in financial data [1]. The advent of artificial intelligence (AI) and machine learning (ML) has transformed this landscape, offering advanced tools to improve predictive accuracy and adaptability in dynamic economic conditions [2].

AI-based classification models, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Random Forests (RF), and Gradient Boosting (GB), have shown superior performance over traditional statistical methods in credit risk assessment [3]. For instance, a study utilizing a seven-year dataset from S&P Capital IQ Pro demonstrated that ANNs and GB outperformed logistic regression in capturing non-linear relationships among financial and business risk variables [4]. Similarly, ensemble methods like XGBoost have achieved remarkable accuracy (up to 99.4%) in predicting credit card customer defaults, highlighting their potential for corporate credit rating applications [5].

The integration of alternative data sources, such as digital footprints and social network analytics, has further enhanced the predictive power of AI models. Research has shown that incorporating call-detail records into credit scoring models significantly improves performance, as measured by the Area Under the Curve (AUC) [6]. Additionally, feature selection techniques, including filter, wrapper, and embedded methods, play a crucial role in identifying relevant predictors, thereby enhancing model efficiency and interpretability [7]. However, challenges such as data imbalance and model transparency remain significant hurdles in deploying AI-driven credit rating models [8].

Comparative studies of classification models are essential to identify the most effective algorithms for specific datasets and contexts. For example, a study comparing logistic regression, SVMs, k-Nearest Neighbors (kNN), and ANNs found that factorization machines achieved higher accuracy in credit risk assessment due to their ability to model feature interactions [9]. Another analysis revealed that Random Forests and ANNs, combined with oversampling techniques, consistently outperformed traditional classifiers across diverse datasets [10]. These findings underscore the importance of evaluating model performance using robust metrics such as accuracy, F1 score, AUC, and the Kolmogorov–Smirnov (KS) statistic [11].

Recent advancements in explainable AI (XAI) have addressed concerns about the interpretability of complex ML models. Techniques like Shapley values have been used to elucidate the contribution of alternative data features, such as loan structure and external scores, in credit scoring [12]. Moreover, hybrid models combining traditional statistical methods with ML, such as clustering-based logistic regression with multilayer perceptrons, have shown improved accuracy in handling imbalanced datasets [13]. Despite these advancements, the literature highlights the need for standardized benchmarks to compare AI models against traditional methods like the FICO credit scoring system, particularly using real-world datasets [14].

The use of AI in corporate credit rating prediction also raises ethical and regulatory considerations. AI-based models can inadvertently introduce biases, necessitating careful data preprocessing and performance evaluation to ensure fairness and compliance [15]. Furthermore, the computational complexity of certain models, such as factorization machines, requires optimization to handle sparse datasets effectively [16]. These challenges highlight the need for comprehensive studies that not only evaluate model performance but also address practical implementation issues in corporate credit rating systems.

Other studies conducted within the country also indicate that various factors influence the commercial creditworthiness of companies. For example, the study by Aflatouni and Nemati (2019) examined the impact of production volume and exports on corporate credit ratings. Their findings revealed that higher production levels and increased export activity can lead to improved commercial creditworthiness [17].

In another study, Hoseini et al. (2018) investigated the effect of company age and ownership structure on corporate credit. The results suggested that older companies with private ownership generally enjoy higher credit ratings [18].

Rezaei et al. (2016) explored the role of external auditing in shaping corporate creditworthiness. Their research demonstrated that audits conducted by independent institutions can enhance stakeholders' trust in financial statements, thereby improving a company's credit standing [19].

Gopalan et al. (2022) found that firms with external audits and foreign ownership typically benefit from stronger credit profiles. Furthermore, companies with higher production capacities and active participation in export markets tend to maintain better credit ratings [20].

Fan et al. (2013) studied the effects of asset volatility and profitability on corporate creditworthiness. The study concluded that firms with lower asset volatility and higher profitability are more likely to have superior credit standings [21].

A study by Alipour and Parsinejad (2018) in Transactions on Machine Intelligence investigated various neural architectures feedforward MLP and feedback NARX networks alongside an ANFIS model to forecast stock portfolio prices. The neural networks outperformed ANFIS in next-day closing price prediction in their simulation, while ANFIS showed robustness with limited data. This highlights the potential of custom neural network designs for accurate financial forecasting [22].

Finally, Demirgok et al. (2010) analyzed the influence of ownership structure and banking systems on corporate credit. Their findings indicated that companies with concentrated ownership and access to robust banking systems typically enjoy higher levels of commercial creditworthiness [23].

This study aims to fill the gap in the literature by systematically evaluating and comparing the performance of various AI-based classification models for predicting corporate credit ratings. By leveraging a diverse set of financial and non-financial variables and employing rigorous evaluation metrics, the research seeks to identify the most effective models and provide insights into their practical applicability for financial institutions.

3. CONCEPTUAL MODEL OF THE STUDY

Based on the theoretical foundations and literature review, a conceptual model has been designed to investigate the factors influencing corporate creditworthiness. This model includes the following independent variables: exports, company age, production volume, external auditing, foreign ownership, ownership type, and company size, with corporate creditworthiness as the dependent variable.

This section describes the methodology used to evaluate and compare the performance of four classification models in predicting corporate credit ratings. The models examined are Support Vector Machine (SVM), Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), and Decision Tree.

4. DATA PREPARATION

The dataset includes the following features:

- Company exports

- Company age (ranging from 1 to 40 years)
- Production volume
- External auditing status (presence or absence)
- Foreign ownership status (presence or absence)
- Ownership type (private or public)
- Company size (small, medium, or large)
- Label (corporate creditworthiness)

These data were extracted from the financial statements of companies listed on the Tehran Stock Exchange. The data are divided into independent variables (features) and a dependent variable (label). Independent variables include exports, company age, production volume, external auditing, foreign ownership, ownership type, and company size, while the dependent variable is the corporate credit rating.

For data preparation, the dataset was split into training and testing subsets with an 80:20 ratio. Additionally, feature scaling was applied to standardize the variables to a uniform scale.

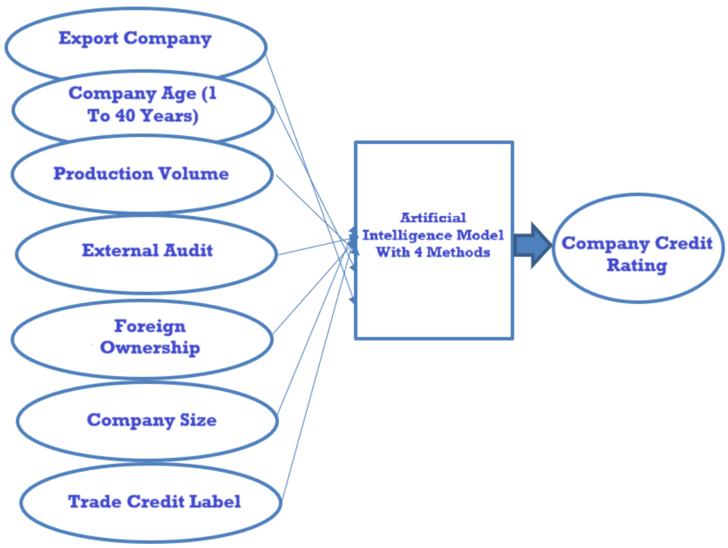

Fig. 1. Conceptual Model of the Study

1.3. Classification

The Support Vector Machine (SVM) model is a powerful method for data classification that separates data into two distinct classes by finding an optimal hyperplane. This model was implemented using the sklearn library. The Artificial Neural Network (ANN) model consists of multiple layers of neurons, where each layer is connected to the next through weights and biases. This model was implemented using the sklearn library and the MLPClassifier function. The K-Nearest Neighbors (KNN) model classifies new samples based on the distance to the nearest training samples. It was implemented using the sklearn library and the KNeighborsClassifier function. The Decision Tree model uses a tree structure for decision-making, where each node evaluates one feature. This model was implemented using the sklearn library and the DecisionTreeClassifier function.

1.4. Model Evaluation

This section describes the criteria and metrics used to assess and compare the performance of the classification models.

Accuracy: The overall accuracy of the models is calculated as the ratio of correctly predicted instances to the total predictions.

Precision and Recall: These metrics evaluate the model's performance in identifying positive and negative samples accurately.

ROC Curve and AUC: The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) metric are used to assess the models' ability to discriminate between classes.

Confusion Matrix: This matrix displays the counts of correct and incorrect predictions for each class.

5. IMPLEMENTATION

The implementation was done in Python, utilizing the Pandas, Numpy, Sklearn, and Matplotlib libraries for data preparation, model training, and performance evaluation.

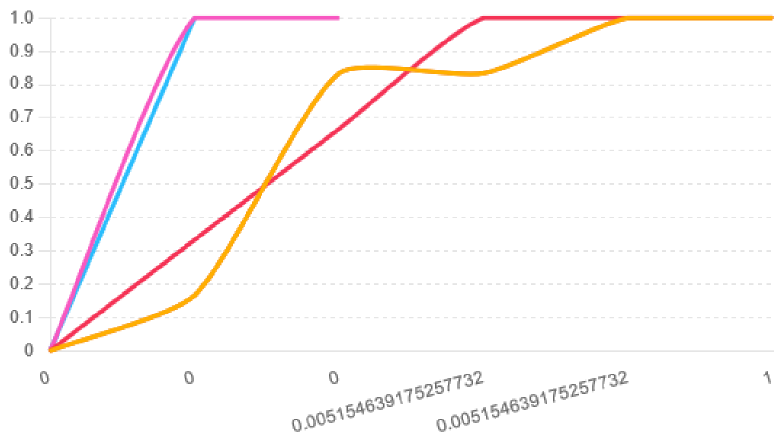

Fig. 2. ROC Curves of Classification Methods

The Receiver Operating Characteristic (ROC) curve is a crucial tool for evaluating the performance of classification models. This curve illustrates a model's ability to distinguish between positive and negative classes. The horizontal axis represents the False Positive Rate (FPR), while the vertical axis represents the True Positive Rate (TPR), also known as Sensitivity.

ROC curves were plotted for the four models under study: SVM, Neural Network, KNN, and Decision Tree. The interpretation of each ROC curve is as follows:

- The SVM model achieved an accuracy of 0.995 and an ROC AUC score of 0.999, demonstrating excellent performance. Its ROC curve lies close to the top-left corner, indicating a high ability to discriminate between classes.
- The Neural Network model also showed excellent performance with an accuracy of 0.995 and an ROC AUC score of 0.999. Its ROC curve closely resembles that of the SVM model, signifying strong class separation capability.
- The KNN model attained an accuracy of 0.990 and an ROC AUC of 0.993, indicating good performance. Its ROC curve is slightly lower than those of the SVM and Neural Network models but still reflects good discrimination ability.

- The Decision Tree model achieved perfect accuracy (1.000) and an ROC AUC score of 1.000, indicating outstanding performance. Its ROC curve forms a perfect square at the top-left corner, representing unparalleled ability to distinguish between classes.

These ROC curves reflect the predictive performance of the models in forecasting corporate credit ratings. The Decision Tree model, with an ROC AUC of 1.000, demonstrates the best performance and exceptional capability in differentiating positive and negative classes. The SVM and Neural Network models also perform very well, while the KNN model performs slightly worse than the other three but still provides good predictive ability.

In summary, the closer the ROC curve is to the top-left corner and the higher the AUC score, the better the model's performance in class separation. Therefore, the Decision Tree model is the most suitable choice for predicting corporate credit ratings on this dataset.

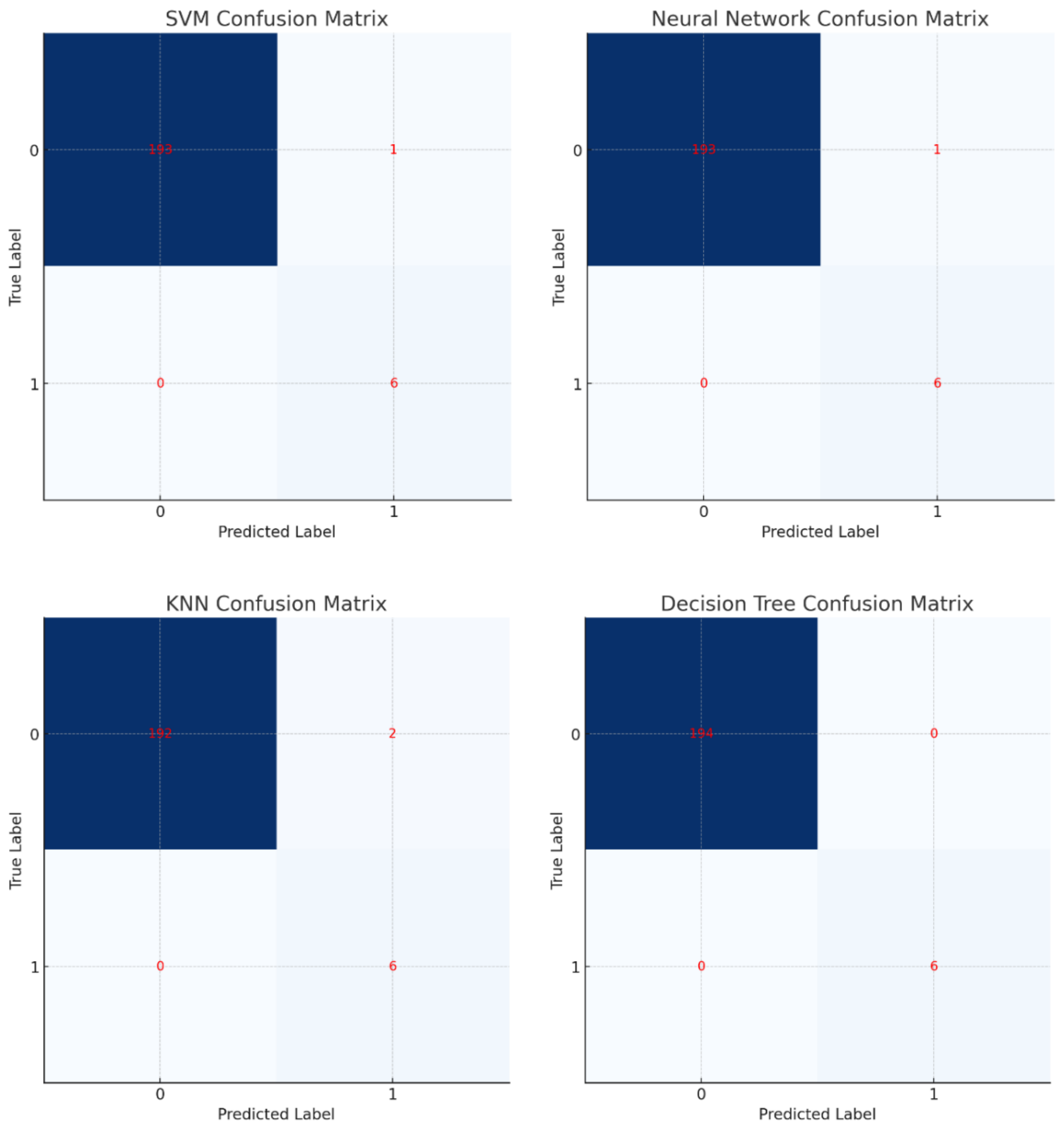

Fig. 3. Confusion Matrices for Classification Methods

The confusion matrix is an essential tool for evaluating the performance of classification models. It helps us understand how accurately the model predicts different classes. Each confusion matrix consists of four key values:

- True Negatives (TN): The number of negative samples correctly classified as negative.
- True Positives (TP): The number of positive samples correctly classified as positive.
- False Positives (FP): The number of negative samples incorrectly classified as positive.
- False Negatives (FN): The number of positive samples incorrectly classified as negative.

The SVM model correctly classified 193 negative samples as negative (TN), misclassified 1 negative sample as positive (FP), correctly classified all 6 positive samples as positive (TP), and did not misclassify any positive samples as negative (FN).

The Neural Network model showed the same performance as the SVM model, with 193 true negatives, 1 false positive, 6 true positives, and 0 false negatives.

The KNN model correctly classified 192 negative samples as negative (TN), misclassified 2 negative samples as positive (FP), correctly classified all 6 positive samples as positive (TP), and did not misclassify any positive samples as negative (FN).

The Decision Tree model correctly classified all 194 negative samples as negative (TN), with no false positives (FP). It correctly classified all 6 positive samples as positive (TP) and had no false negatives (FN).

Table. 1. Comparison of Results for Corporate Credit Rating Prediction

Model	Accuracy	Precision	Recall	ROC AUC Score
SVM	0.995	0.857	1.000	0.999
Neural Network	0.995	0.857	1.000	0.999
KNN	0.990	0.750	1.000	0.993
Decision Tree	1.000	1.000	1.000	1.000

6. CONCLUSION

This study evaluated and compared the performance of four classification models—Support Vector Machine (SVM), Neural Network, K-Nearest Neighbors (KNN), and Decision Tree—in predicting corporate credit ratings using data from companies listed on the Tehran Stock Exchange. The results demonstrate varying capabilities among these models in distinguishing between companies with high and low creditworthiness.

The Decision Tree model achieved the best performance, correctly classifying all samples with 100% accuracy and an ROC AUC score of 1.000. Due to its simple and interpretable structure, this model can serve as an effective tool for financial and credit managers in assessing corporate credit ratings.

Both the SVM and Neural Network models also performed very well, each making only one misclassification while maintaining high accuracy and recall. These models are particularly suitable for handling complex and nonlinear data and can be effectively used alongside the Decision Tree model.

The KNN model showed comparatively weaker performance but still delivered acceptable results. It can be especially useful in scenarios where there is insufficient data to train more complex models.

Overall, the findings indicate that classification models can effectively assist in predicting corporate credit ratings. Given the superior performance of the Decision Tree model, it is recommended that financial and credit managers adopt this model for evaluation and prediction tasks. Additionally, employing a combination of different models may enhance prediction accuracy and reliability, leading to better and more informed decision-making.

Declaration

We acknowledge that we used ChatGPT to enhance the academic writing of our manuscript while ensuring the originality and integrity of our work.

Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

Declaration of Interest

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

REFERENCES

[1] Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589–609. https://doi.org/10.2307/2978933

[2] Chen, N., Ribeiro, B., & Chen, A. (2016). Financial credit risk assessment: A recent review. *Artificial Intelligence Review, 45*(1), 1–23. https://doi.org/10.1007/s10462-015-9434-x

[3] Addo, P. M., Guegan, D., & Hassani, B. (2018). Credit risk analysis using machine and deep learning models. *Risks, 6*(2), 38. https://doi.org/10.3390/risks6020038

[4] Jang, J., Choi, H., Bae, H., Lee, S., Kwon, M., & Jung, M. (2023). CXL-ANNS: Software-Hardware Collaborative Memory Disaggregation and Computation for Billion-Scale Approximate Nearest Neighbor Search. In *USENIX Annual Technical Conference*.

[5] Chang, V., Sivakulasingam, S., Wang, H., Wong, S. T., Ganatra, M. A., & Luo, J. (2024). Credit risk prediction using machine learning and deep learning: A study on credit card customers. *Risks, 12*(11), 174. https://doi.org/10.3390/risks12110174

[6] Ridzuan, N. N., Masri, M., Anshari, M., Fitriyani, N. L., & Syafrudin, M. (2024). AI in the financial sector: The line between innovation, regulation and ethical responsibility. *Information, 15*(8), 432. https://doi.org/10.3390/info15080432

[7] Jemai, J., & Zarrad, A. (2023). Feature selection engineering for credit risk assessment in retail banking. *Information, 14*(3), 200. https://doi.org/10.3390/info14030200

[8] Huang, X., Liu, X., & Ren, Y. (2018). Enterprise credit risk evaluation based on neural network algorithm. *Cognitive Systems Research, 52*, 317–324. https://doi.org/10.1016/j.cogsys.2018.07.023

[9] Quan, J., & Sun, X. (2024). Credit risk assessment using the factorization machine model with feature interactions. *Humanities and Social Sciences Communications, 11*(1), 1–10. https://doi.org/10.1057/s41599-024-02700-7

[10] Antoniadis, A., Georgiou, K., Charmanas, K., Mittas, N., & Angelis, L. (2023, October). Benchmarking classifiers for loan default prediction using archetypal analysis. In *Proceedings of the 7th International Conference on Algorithms, Computing and Systems* (pp. 13–18). https://doi.org/10.1145/3631908.3631911

[11] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357. https://doi.org/10.1613/jair.953

[12] Hlongwane, R., Ramaboa, K. K. M., & Mongwe, W. (2024). Enhancing credit scoring accuracy with a comprehensive evaluation of alternative data. *PLOS ONE, 19*(5), e0303566. https://doi.org/10.1371/journal.pone.0303566

[13] Xu, T. (2024). Comparative analysis of machine learning algorithms for consumer credit risk assessment. *Transactions on Computer Science and Intelligent Systems Research, 4*, 60–67. https://doi.org/10.62051/r1m3pg16

[14] 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

[15] Sadok, H., Sakka, F., & El Maknouzi, M. E. H. (2022). Artificial intelligence and bank credit analysis: A review. *Cogent Economics & Finance, 10*(1), 2023262. https://doi.org/10.1080/23322039.2021.2023262

[16] Zhang, L., Yu, R., Lin, Y., & Yang, Y. (2024, August). A more precise credit score computation method based on big data and machine learning. In *Proceedings of the 2024 7th International Conference on Machine Learning and Machine Intelligence (MLMI)* (pp. 34–38). https://doi.org/10.1145/3696271.3696277

[17] Aflatouni, M., & Nemati, H. (2019). Investigating the impact of production volume and exports on corporate creditworthiness. *Journal of Business Management*, 12(3), 45–63.

[18] Hosseini, A., et al. (2018). The effect of company age and ownership type on commercial creditworthiness. *Journal of Financial Research*, 10(2), 34–51.

[19] Rezaei, M., et al. (2016). The role of external auditing in corporate creditworthiness. *Auditing Journal*, 9(1), 22–39.

[20] Gopalan, R., et al. (2022). The impact of foreign ownership and external auditing on trade credit. Journal of Financial Studies, 45(4), 123-142.

[21] Fan, J. P., et al. (2013). Asset volatility and profitability impacts on trade credit. Financial Research Letters, 12(2), 98-112.

[22] Alipour, M. A. and Parsinejad, S. (2018). Designing an Appropriate Neural Network to Assist in More Accurate Stock Portfolio Predictions. *Transactions on Machine Intelligence*, *1*(4), 228-242. doi: 10.47176/TMI.2018.228

[23] Demirgok-Kunt, A., et al. (2010). Ownership structure and banking systems' influence on trade credit. Journal of Corporate Finance, 16(3), 304-317