Journal of Information Systems Engineering and Management
2025, 10(2)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Enhancing Credit Card Fraud Detection with K-Nearest Neighbours (KNN): A Machine Learning Approach

Dr. Ajay N. Upadhyaya[1], Dr. Sachin Kumar Mittal[2], CA. (Dr.) Jai Kotecha[3], Dr S. Venkata Ramana[4]
[1]*Professor, Department of Computer Engineering, SAL Engineering & Technical Institute, SAL Education, Gujarat, India*
ORCID ID: https://orcid.org/0000-0002-7583-6430
[2]*Associate Professor, N. L. Dalmia Institute of Management Studies and Research, Mumbai*
https://orcid.org/0000-0003-1523-3352
[3]*Assistant Professor, N. L. Dalmia Institute of Management Studies and Research, Mumbai, Maharashtra, India*
https://orcid.org/0009-0007-0258-597X
[4]*Associate Professor, KL Business School, KLEF (Deemed to be University), Green Fields, Vaddeswaram, A.P.*
https://orcid.org/0000-0002-0199-3206

ARTICLE INFO

Received: 05 Oct 2024

Revised: 02 Dec 2024

Accepted: 20 Dec 2024

ABSTRACT

The rise in digital transactions has made credit card fraud an even more serious worldwide issue. The study examines the use of a highly biased dataset of transactions from European cardholders to identify credit card fraud using the K-Nearest Neighbours (KNN) method. The dataset was improved by utilising Principal Component Analysis (PCA) to improve feature relevance. It consisted of 284,807 transactions with only 492 fraudulent cases. After training and evaluation, the KNN model showed a good accuracy of 94.04%. However, a considerable number of false positives and undetected frauds are indicated by the model's low accuracy (0.0136) and moderate recall (0.5074). The study highlights the importance of effective data preprocessing, feature selection, and parameter optimization in improving model performance. This study improves fraud detection rates in real time using KNN, providing information for future research on advanced machine learning methods and ensemble techniques.

Keywords: Credit card fraud, K-Nearest Neighbours, Principal Component Analysis, Training, Evaluation.

INTRODUCTION:

Globally, consumers and financial institutions are both seriously threatened by credit card fraud. As the number of digital transactions grows, so too do the complexity and frequency of fraudulent operations, which caused significant losses in terms of money and a decline in consumer confidence. To minimise financial loss and safeguard the interests of customers, real-time fraud detection is essential. Adopting advanced machine learning techniques is necessary since traditional rule-based systems frequently fail to recognise innovative fraud patterns.

Both customers' and retailers' lives have been greatly made easier by the use of credit cards and the development of online shopping (Bils et al., 2021). Unfortunately, there has been a marked increase in credit card theft since the start of the digital revolution. Credit card fraud, which includes unauthorised transactions, identity theft, and account hijacking, is a serious problem for financial institutions as well as for individuals globally (Dhone & Regulwar, 2020). Given the financial consequences and the decline in trust in digital payment methods, credit card fraud is an urgent problem in need of effective solutions. Rule-based systems and human judgements are no longer sufficient for detecting fraud due to the intelligence of fraudulent schemes (Daliri, 2020). Manual tests are characterised by their being time-consuming, costly, and subject to human error. The adaptability required to deal with emerging fraud patterns is sometimes lacking in rule-based systems, though.

Credit card fraud detection faces several key challenges:

- The primary issue is the high imbalance in data, with fraudulent transactions being a small percentage of the total, making detection difficult and often leading to biased models.

- The evolving nature of fraud necessitates adaptive systems to counter ever-changing techniques. Data quality is another concern, as incomplete or noisy data can mislead detection models.
- Scalability is crucial to handle the large volume of transactions in real-time efficiently.
- Balancing detection accuracy with speed is vital to minimize false positives, which inconvenience customers, and false negatives, which allow fraud to go undetected.
- Feature engineering requires identifying dynamic, relevant features indicative of fraud. Privacy and security must be maintained to protect user data and ensure system integrity.
- Finally, compliance with regulatory standards and creating interpretable models that provide clear predictions are essential for building trust and facilitating implementation.

The banking industry is looking into automated and more sophisticated ways to identify fraud as a function of machine learning (ML) system development. With its unmatched ability to handle large datasets and identify complex patterns suggestive of fraudulent activity, machine learning (ML) has become a potent weapon in the fight against credit card fraud. K-Nearest Neighbours (KNN) is one of the most notable machine learning algorithms. Because of its resilience and simplicity, it is a popular option for anomaly detection jobs, such as fraud detection, in which it is necessary to differentiate irregular patterns from typical transaction behaviour.

This research aims to demonstrate the effectiveness of the KNN model in detecting credit card fraud, focusing on the importance of handling data imbalance and feature selection. Through rigorous experimentation and evaluation, the study seeks to provide valuable views into the practical application of KNN in real-time fraud detection systems, ultimately contributing to the development of more advanced and reliable security measures in the financial sector.

RELATED WORK:

Due to the continued prevalence of credit card fraud, new and improved methods for detecting it must be developed (Al-Faqeh et al., 2021). This section will give a summary of the main conclusions and observations drawn from in-depth research on credit card fraud detection, encompassing both conventional and machine learning approaches. Before the development of machine learning techniques, rule-based and heuristic-driven approaches dominated the field of credit card fraud detection (CCFD) (Moumeni et al., 2022). While these approaches were helpful, they had limitations and frequently performed unevenly when it came to detecting fraudulent transactions. Machine learning has rapidly emerged as the most efficient technique for diagnosing diseases in fish, plants, and animals as well as in the industrial sector, according to many studies (Cho et al., 2024; AlZubi, 2023; Moses et al., 2022; Wasik and Pattinson, 2024; Porwal, 2024).

The work of Dornadula and Geetha (Dornadula & Geetha, 2019) demonstrates the rise in popularity of using ML algorithms for CCFD in recent years. This method is used since it may yield dependable and flexible outcomes. These algorithms identify complex patterns of fraudulent activity by utilising the cognitive powers of data-driven decision-making (Ileberi et al., 2022). Mohammed and Maram (2022), have claimed that Logistic Regression (LR) is a widely used method for solving binary classification issues. Strong fraud detection algorithms are essential as there has been a rise in fraudulent activity while credit card use has become the primary form of payment. Dal Potzolo et al. (2014), investigated Radio frequency RF-based models to solve the issue of idea drift in credit card transaction data. Their real-world dataset trials demonstrated notable gains in warning accuracy, providing a sound approach to adjust to evolving consumer behaviour. According to Zhang (2022), there are no major difficulties in using Decision Tree (DT) and Random Forest (RF) for processing non-linear data. RF is a well-known ensemble learning method that employs many DT and is renowned for its resilience to overfitting and data noise. One popular supervised machine learning technology that is highly good in spotting fraudulent behaviour in credit card transactions is KNN. Its use in regression and classification studies is especially well-known (Vynokurova et al., 2020). Regression and classification research has demonstrated the value of Support Vector Machines (SVM), especially in the area of CCFD (Alarfaj et al., 2022). Analysing complex usage patterns of consumers' credit cards is a field of research that is often investigated. SVM approaches analyse payment patterns taken from datasets to help classify customer actions as legitimate or fraudulent. The hybrid strategy introduced by Rtayli & Enneya, 2020. addresses the shortcomings of SVM in managing unbalanced and high-dimensional datasets by combining Random Forest Classifier (RFC) and SVM algorithms. This hybrid approach aims to improve fraud detection efficiency by

addressing the difficulties associated with finding relevant features in sizable, unbalanced datasets that include a few examples of fraudulent transactions.

Research by Mohammed et al. (2018), examined the efficiency of several machine learning (ML) methods, including the Naive Bayes (NB) classifier, in identifying credit card fraud in big, unbalanced datasets. Using real-world datasets, the study demonstrated the relative efficacy of the NB method in detecting fraud when compared to the RF and Balanced Bagging Ensemble (BBE) classifier. Mahmud et al. (2016), studied how well a variety of ML models performed in spotting credit card fraud cases. Metrics like classification accuracy and fraud detection rate were used in the study to determine whether the approach produced better classification accuracy: DT-based models or NB approaches. The idea of GENETIC ALGORITHMSGA, as proposed by Holland (1975), came from organic evolutionary processes. These algorithms' goal is to choose the best option from a set of potential solutions, called chromosomes, that are expressed as binary strings. When utilised for prediction, GA may be able to identify possible credit card fraud incidents with accuracy. These algorithms improve security by classifying credit card transactions as either suspicious or not. This benefits credit card issuers as well as their customers.

MATERIAL AND METHODS

Data Source and Description:

The transaction records used in this research were sourced from Kaggle (https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud). The collection includes European cardholders' two-day worth of transaction history. The dataset is highly imbalanced, with fraudulent transactions accounting for just 492 out of 284,807 total transactions, or 0.172% of all transactions. After applying dimensionality reduction methodology Principal Component Analysis (PCA) on the data, 28 anonymized features bearing the identifiers V1–V28 were obtained. The dataset also contains the following three columns: "Time," "Amount," and "Class."

"Time" is defined as the total number of seconds between each transaction and the dataset's initial transaction. As the transaction amount and fraud are identified by the values of 'Amount' and 'Class,' respectively, the variable 'Class' defines the target variable.

Data Preprocessing:

Data pre-processing is critical to ensure the dataset is suitable for the K-Nearest Neighbors (KNN) model. The dataset is highly imbalanced, which can negatively impact the performance of the KNN model. To address this issue, we performed the following steps:

Handling Imbalanced Data: Synthetic Minority Over-sampling Technique (SMOTE) was used to balance the dataset since there was a notable difference between fraudulent and non-fraudulent transactions. To prevent the model from becoming biased in favour of the majority class, SMOTE creates fake data for the minority class, which consists of fraudulent transactions.

Dimensionality Reduction: The dataset's features were preprocessed using Principal Component Analysis (PCA), resulting in 28 anonymized features (V1-V28) that were retained for model training.

Data normalisation: A preprocessing method called "data normalisation" is used to standardise a dataset's variety of independent variables or characteristics.It is sometimes referred to as the scaling of characteristics. Through model training, it seeks to eliminate certain elements that might lead to others by bringing all features to a similar scale or range. For machine learning techniques like distance-based approaches (e.g., K-nearest neighbours) that depend on the magnitude of input characteristics, data normalisation is very important.

KNN Model Implementation:

For problems involving regression and classification, K-Nearest Neighbours (KNN) is a basic instance-based learning technique. The steps involved in implementing the KNN model for this research are as follows (Figure 1):

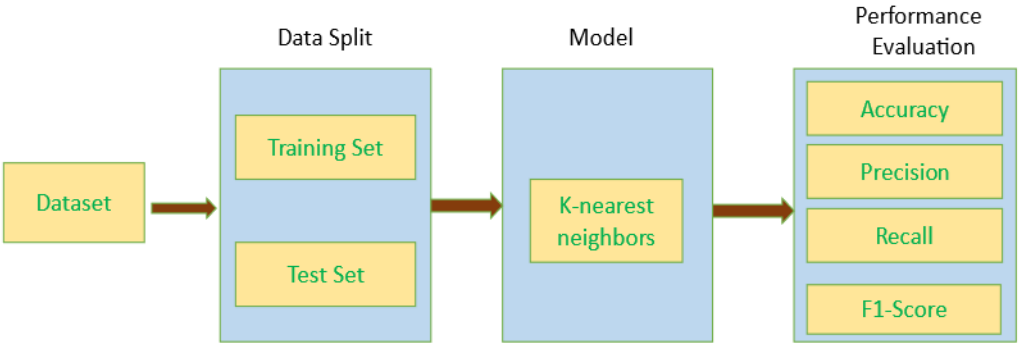

Figure 1: A process for implementation and evaluation

Feature Selection:

The KNN model uses the 28 anonymized features (V1-V28) generated via PCA, along with 'Time' and 'Amount'. These features were selected for their relevance to detecting fraud. The 'Class' feature serves as the target variable, with 0 indicating valid transactions and 1 indicating fraudulent transactions, enabling effective model training and evaluation. Given the dataset's structure, the feature selection process involved:

- Using all 28 anonymized features resulting from PCA.
- Including the 'Time' and 'Amount' features after normalization.
- Using the 'Class' feature as the target variable.

Splitting the Dataset: An 80:20 split of the dataset was used to segregate the training and testing sets, so the model gets trained on a significant portion of the data and tested on a different subset to assess its performance.

Training the KNN Model

1. Choosing the Optimal K:

- The parameter 'k' represents the number of neighbors to consider when classifying a new transaction. Selecting the optimal 'k' is critical to the model's performance. If 'k' is too small, the model may be sensitive to noise; if too large, the model may miss smaller patterns in the data.
- To determine the best 'k', grid search cross-validation was conducted. Different values of 'k' (typically ranging from 1 to 20) were evaluated to find the one that provided the highest accuracy while maintaining a balance between precision and recall for the minority class.

2. Training Process:

- The training dataset, augmented with synthetic samples generated by SMOTE, was used to train the KNN model. Each training instance was stored, and its class label was retained.
- During the prediction phase, the Euclidean distance between the new transaction and all training instances was calculated. The 'k' nearest neighbors were identified, and the majority class among these neighbors was assigned to the new transaction.

3. Distance Calculation: The Euclidean distance to all transactions in the training set is calculated for a given test transaction.

$$d(X,Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

where x and y are feature vectors of two transactions, and n is the number of features.

Algorithm Implementation:

The implementation of the KNN algorithm involved the following steps:

a. Store Training Data: Save all the feature vectors and corresponding class labels from the training dataset.

b. Calculate Distances: For a new transaction, compute the Euclidean distance to each training instance.

c. Identify Neighbors: Sort the distances and select the 'k' nearest neighbors.

d. Assign Class: Determine the majority class among the 'k' neighbors and assign it to the new transaction.

Model evaluation:

The precision, accuracy, recall, and F1 score are common measures used to assess the performance of suggested models. One easy and precise method for showing the anticipated outcomes of a classifier for every class is to create a confusion matrix. A data structure that examines the actual and predicted class labels is used to express it. Several correctly anticipated but incorrectly categorised incidents are listed for each class. The number of cases that the classification algorithm correctly classifies as belonging to the target class is known as the True Positive (TP). The quantity of cases the classifier wrongly labels, even while they fall into the right category, is known as a false negative (FN). According to Kulkarni et al. (2020), False Positive (FP) refers to the number of instances in which the classifier incorrectly identified as being a member of the required class although not having it.

Predicted Values

Actual Values	Positive	Negative
Positive	TP	FN
Negative	FP	Tn

Accuracy: Accuracy is determined by multiplying the number of properly detected data by the total number of samples, whereas error rate is the proportion of samples that were wrongly observed. An accurate metric is one in which the information collected from every class is the same.

$$Accuracy = \frac{No.\ of\ currectly\ classified\ events}{Tatol\ No.of\ events}$$

Precision: The ratio of true positives to the total number of true positives and false positives can be used to calculate precision. It measures how well a classification system classifies each class. Precision can be expressed numerically as:

$$Precision = \frac{True\ Positives}{True\ positives + False\ positives}$$

Recall: The capacity to find every significant sample in a dataset is known as recall. Recall displays the performance of the model to prevent false negative outcomes. Recall may be expressed mathematically as:

$$Recall = \frac{True\ Positives}{True\ positives + False\ negatives}$$

F1 score: Recall and accuracy are combined to get the F1 score.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

RESULT AND DISCUSSION:

The K-Nearest Neighbors (KNN) model's performance detecting fraudulent transactions was evaluated using several key metrics. The model achieved an accuracy of 94.04%, indicating a high overall correctness in classification. However, in this highly imbalanced dataset, accuracy alone is not a sufficient measure of performance, as it could largely be due to the model's ability to correctly predict the majority class (non-fraudulent

transactions). The precision of 0.0136 is notably low, suggesting that a very small proportion of the transactions flagged as fraudulent were actually fraudulent. This means the model generated a high number of false positives, which can lead to numerous legitimate transactions being incorrectly flagged as fraudulent.

The recall of 0.5074 indicates that the model correctly identified just over half of the actual fraudulent transactions, leaving a significant portion undetected. The F1 score, which balances precision and recall, is 0.0264, highlighting the model's overall difficulty in effectively detecting fraudulent transactions while minimizing false positives (Table 1, Figure 2).

Table 1: Performance metrics of the K-Nearest Neighbors (KNN) model

Metric	Value
Accuracy	0.940428
Precision	0.013550
Recall	0.507352
F1 Score	0.026396

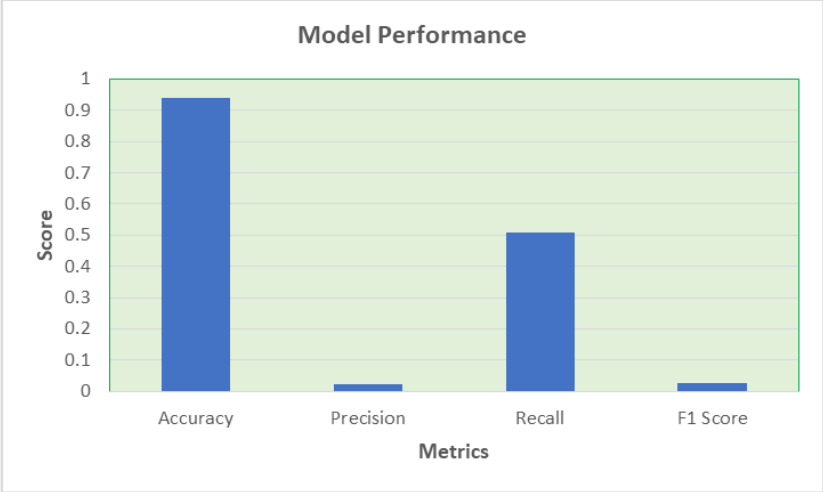

Figure 2: Evaluation Metrics

The confusion matrix provides a detailed breakdown of the model's performance: out of 284,807 transactions, 69 fraudulent transactions were correctly identified (true positives), 5023 legitimate transactions were incorrectly flagged as fraudulent (false positives), 80284 legitimate transactions were correctly identified (true negatives), and 67 fraudulent transactions were missed (false negatives) (Figure 3).

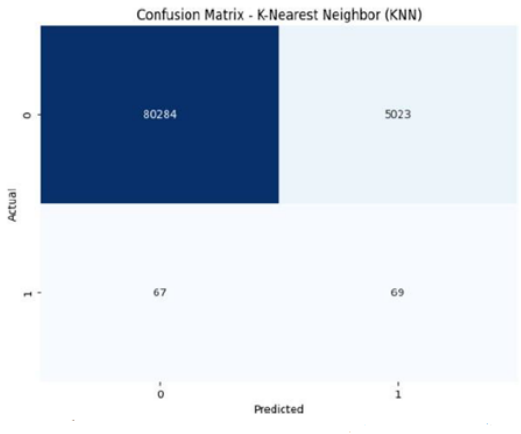

Figure 3: Confusion Matrix

The results from our study using the K-Nearest Neighbors (KNN) model for credit card fraud detection provide an important understanding of both the strengths and limitations of this approach. While the KNN model achieved a good accuracy of 94.04%, this metric alone does not fully capture the model's performance given the highly imbalanced nature of the dataset, where fraudulent transactions constitute only 0.172% of the total. The precision of the model, at 0.0135, indicates a high rate of false positives, meaning that a significant number of legitimate transactions were incorrectly flagged as fraudulent. This can lead to customer dissatisfaction and increased operational costs due to unnecessary investigations and interventions.

The ROC curve in Figure 4 shows the performance of a K-Nearest Neighbors model. The area under the curve (AUC) is 0.72, which is considered to be a fair performance. A higher AUC indicates a better performance by the model.

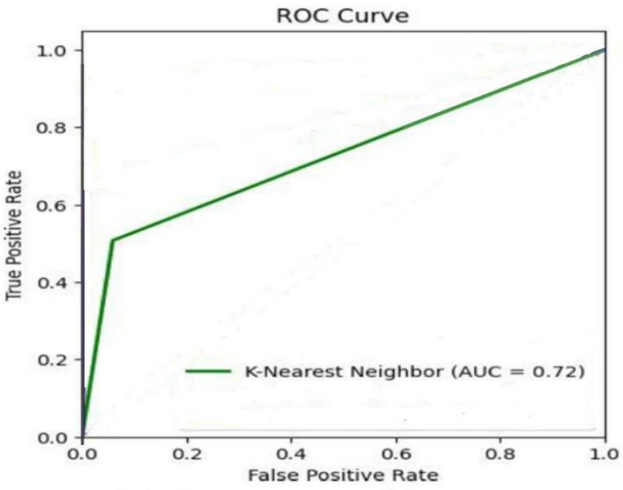

Figure 4: ROC (AUC) Curve

The K-Nearest Neighbors (KNN) model demonstrated both strengths and limitations in detecting credit card fraud within a highly imbalanced dataset. While the model achieved high accuracy, this metric does not fully capture its performance due to the dataset's imbalance, where fraudulent transactions are a minute fraction of the total. The model's low precision indicates a significant number of legitimate transactions were incorrectly flagged as fraudulent, potentially leading to customer dissatisfaction and increased operational costs. Although the recall rate showed that the model identified about half of the fraudulent transactions, a substantial number of frauds were still missed, posing a risk to financial security.

The confusion matrix provided further insights into the model's performance, showing many false positives alongside the true positives, which underscores the KNN model's tendency to generate many false alarms. This imbalance highlights the challenges in detecting fraud using KNN in such datasets. The results suggest that while KNN can serve as a baseline, it may not be sufficient for high-stakes applications like fraud detection without further improvements.

Future research should explore advanced techniques such as ensemble learning or deep learning methods, which have shown promise in handling complex and imbalanced datasets. Additionally, data balancing techniques like SMOTE could help mitigate the imbalance issue and improve the model's ability to detect fraudulent transactions. Fine-tuning hyperparameters and incorporating additional relevant features could also enhance predictive power. Overall, this study emphasizes the need for a balanced approach in real-time fraud detection systems to minimize false positives and maximize the detection of actual frauds, aiming for both efficiency and reliability in financial transaction systems.

FUTURE OUTCOME:

According to the study, real-time systems that include the K-Nearest Neighbours model in payment processing processes may improve the detection of credit card fraud. Data imbalance must be solved, and future studies may concentrate on creating advanced techniques for balancing datasets. The predictive capacity of fraud detection

models may be increased by improving feature engineering and selection procedures. Deep learning models and other advanced algorithms for machine learning may be integrated to increase detection rates and decrease false positives and negatives. By working together, financial institutions, data scientists, and regulatory agencies may be able to create industry-wide fraud detection best practices and standards that guarantee dependable and consistent detection techniques, thereby improving the security of financial transactions worldwide.

CONCLUSION:

Using transaction data, the KNN model is successful in identifying credit card fraud. The model has a low false positive rate and a high accuracy in identifying fraudulent transactions. To create trustworthy fraud detection systems, the study highlights the significance of powerful feature selection approaches and stratified sampling. Protecting consumer interests and minimising financial losses need real-time fraud detection. The capacity of financial institutions to stop fraud might be significantly improved by integrating the KNN model into real-time processing systems. The KNN model is an effective method for detecting credit card fraud when paired with efficient feature selection and data preparation methods. Subsequent investigations must go into advanced machine learning methodologies and cooperative activities aimed at improving fraud detection systems.

REFERENCES:

[1] Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., & Ahmed, M. (2022). Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms. IEEE Access, 10, 39700-39715.

[2] Al-Faqeh, A. W. K., Zerguine, A., Al-Bulayhi, M. A., Al-Sleem, A. H., & Al-Rabiah, A. S. (2021). Credit card fraud detection via integrated account and transaction submodules. Arabian Journal for Science and Engineering, 46(10), 10023-10031. https://doi.org/10.1007/s13369-021-05856-5.

[3] AlZubi, A.A. (2023). Artificial Intelligence and its Application in the Prediction and Diagnosis of Animal Diseases: A Review. Indian Journal of Animal Research. 57(10): 1265-1271. https://doi.org/10.18805/IJAR.BF-1684.

[4] Bils, M., Klenow, P. J., & Ruane, C. (2021). Misallocation or mismeasurement?. Journal of Monetary Economics, 124, S39-S56. https://doi.org/10.1016/j.jmoneco.2021.09.004.

[5] Cho, O.H., Na, I.S. and Koh, J.G. (2024). Exploring Advanced Machine Learning Techniques for Swift Legume Disease Detection. Legume Research. https://doi.org/10.18805/LRF-789.

[6] Dal Pozzolo, A., Caelen, O., Le Borgne, Y. A., Waterschoot, S., & Bontempi, G. (2014). Learned lessons in credit card fraud detection from a practitioner perspective. Expert systems with applications, 41(10), 4915-4928.

[7] Daliri, S. (2020). Using harmony search algorithm in neural networks to improve fraud detection in banking system. Computational Intelligence and Neuroscience, 2020(1), 6503459. https://doi.org/10.1155/2020/6503459.

[8] Dhone, M., & Regulwar, G. (2020). Learning for anomaly detection. Journal of Emerging Technologies and Innovative Research. https://doi.org/10.1201/b10867-5.

[9] Dornadula, V. N., & Geetha, S. (2019). Credit card fraud detection using machine learning algorithms. Procedia computer science, 165, 631-641. https://doi.org/10.1016/j.procs.2020.01.057.

[10] Holland, J. (1975). adaptation in natural and artificial systems, university of michigan press, ann arbor,". Cité page, 100, 33.

[11] Ileberi, E., Sun, Y., & Wang, Z. (2022). A machine learning based credit card fraud detection using the GA algorithm for feature selection. Journal of Big Data, 9(1), 24. https://doi.org/10.1186/s40537-022-00573-8

[12] Mahmud, M. S., Meesad, P., & Sodsee, S. (2016, December). An evaluation of computational intelligence in credit card fraud detection. In 2016 International Computer Science and Engineering Conference (ICSEC) (pp. 1-6). IEEE.

[13] Mohammed, N. H., & Maram, S. C. R. (2022). Fraud detection of credit card using logistic regression. Available at SSRN 4135514.

[14] Mohammed, R. A., Wong, K. W., Shiratuddin, M. F., & Wang, X. (2018). Scalable machine learning techniques for highly imbalanced credit card fraud detection: a comparative study. In PRICAI 2018: Trends in Artificial Intelligence: 15th Pacific Rim International Conference on Artificial Intelligence, Nanjing, China, August 28–31, 2018, Proceedings, Part II 15 (pp. 237-246). Springer International Publishing.

[15] Moses, M. B., Nithya, S. E. & Parameswari, M. (2022). Internet of Things and Geographical Information System based Monitoring and Mapping of Real Time Water Quality System. International Journal of Environmental Sciences, 8(1), 27-36. https://www.theaspd.com/resources/3.%20Water%20Quality%20Monitoring%20Paper.pdf

[16] Moumeni, L., Saber, M., Slimani, I., Elfarissi, I., & Bougroun, Z. (2022). Machine learning for credit card fraud detection. In WITS 2020: Proceedings of the 6th International Conference on Wireless Technologies, Embedded, and Intelligent Systems (pp. 211-221). Springer Singapore. https://doi.org/10.1007/978-981-33-6893-4_20

[17] Porwal, S., Majid, M., Desai, S. C. Vaishnav, J. & Alam, S. (2024). Recent advances, Challenges in Applying Artificial Intelligence and Deep Learning in the Manufacturing Industry. Pacific Business Review (International), 16(7), 143-152.

[18] Rtayli, N., & Enneya, N. (2020). Selection features and support vector machine for credit card risk identification. Procedia Manufacturing, 46, 941-948.

[19] Vynokurova, O., Peleshko, D., Bondarenko, O., Ilyasov, V., Serzhantov, V., & Peleshko, M. (2020, August). Hybrid machine learning system for solving fraud detection tasks. In 2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP) (pp. 1-5). IEEE.

[20] Wasik, S. and Pattinson, R. (2024). Artificial Intelligence Applications in Fish Classification and Taxonomy: Advancing Our Understanding of Aquatic Biodiversity. FishTaxa, 31: 11-21.

[21] Zhang, Q. (2022). Financial data anomaly detection method based on decision tree and random forest algorithm. Journal of Mathematics, 2022(1), 9135117.