# REAL-TIME FRAUD PREVENTION IN DIGITAL BANKING A CLOUD AND AI PERSPECTIVE

**Article** *in* Journal of Emerging Technologies and Innovative Research · May 2023

1 author:

Jeyasri Sekar
**16** PUBLICATIONS **407** CITATIONS

# REAL-TIME FRAUD PREVENTION IN DIGITAL BANKING A CLOUD AND AI PERSPECTIVE

**Jeyasri Sekar**

Senior Software Engineer

USA

*Abstract:* Real-time fraud prevention in digital banking involves employing advanced technologies such as machine learning and cloud infrastructure to detect and mitigate fraudulent transactions instantly, safeguarding financial institutions and customers from potential financial losses and security breaches. Some challenges in real-time fraud prevention in digital banking include handling large volumes of data in real-time, ensuring the accuracy and reliability of machine learning models, addressing evolving fraud tactics, balancing between false positives and false negatives, and maintaining compliance with regulatory requirements regarding data privacy and security. This paper developed a comprehensive approach to detect and mitigate fraudulent transactions in real-time using cloud-based infrastructure and advanced artificial intelligence techniques. The process begins with data collection from various sources, followed by preprocessing to ensure data integrity and completeness. Feature extraction techniques, including dimensionality reduction, are applied to identify key attributes indicative of fraudulent behavior. Feature selection is optimized using Improved Red Piranha Optimization (IRPO) to enhance model performance. Subsequently, machine learning models, such as Support Vector Machines (SVM), and Naïve Bayes (NB) are developed and deployed to classify transactions as fraudulent or legitimate. By leveraging cloud computing and AI, this framework enables timely detection and prevention of fraudulent activities, contributing to the security and trustworthiness of digital banking systems.

*Keywords:* Real-Time Fraud Prevention; Digital Banking; IRPO; SVM; Cloud; Machine Learning.

**INTRODUCTION:**

The current financial environment places a high priority on fraud prevention in digital banking due to the proliferation of online banking services and the quick advancement of technology, which have opened up new avenues for fraudulent activity. Digital transactions are convenient, but they also increase the danger of cybercrime, which includes identity theft, financial fraud, and illegal access [1, 2]. Financial institutions are thus under tremendous pressure to put strong security measures in place in order to protect the assets of their clients and preserve consumer confidence in the digital banking environment. The breadth and sophistication of fraud schemes have increased dramatically as a result of the growth of digital banking platforms and the growing reliance on electronic payment methods [3]. Traditional fraud detection measures have considerable hurdles as criminals consistently develop advanced tactics to exploit flaws in digital systems. As a result, there is an increasing need for creative solutions that make use of cutting-edge technology to improve digital banking's capacity to prevent fraud [4, 5]. In the past, rule-based systems and manual procedures dominated the financial industry's fraud detection landscape. These methods had limitations when it came to monitoring changing risks and identifying minute trends that would point to fraudulent activity [6]. But the emergence of artificial

intelligence (AI) and machine learning has transformed fraud prevention by allowing automated, data-driven methods that can detect aberrant activity with previously unheard-of precision and analyze massive volumes of transactional data in real-time [7].

Anomaly detection models and supervised learning classifiers are two examples of machine learning algorithms that have become extremely effective tools for identifying fraudulent transactions in digital banking. These algorithms use past transaction data to discover patterns of acceptable conduct and spot anomalies that could be indicative of fraud [8, 9]. Machine learning systems are capable of adapting to new threats and improving their efficacy over time by continually studying transactional trends and changing their models based on fresh data. Fraud protection in digital banking is still a complicated and multidimensional problem, especially with the major gains made possible by machine learning [10]. The intrinsic imbalance in transaction data, where fraudulent transactions account for a very tiny percentage of total activity, is one of the main challenges. Traditional machine learning algorithms may find it difficult to reliably identify fraudulent activity in highly skewed datasets due to this imbalance [11]. In order to overcome this difficulty, scholars and industry professionals have investigated a range of strategies, including feature engineering, data preprocessing, and ensemble learning approaches, to enhance the efficacy of fraud detection models [12, 13]. Furthermore, integrating cutting-edge technology like cloud and deep learning has the potential to improve digital banking systems' security and resistance against fraud [14, 15].

The contributions of this paper are manifested below,

- The paper introduces novel technique IRPO for feature selection, enhancing model performance and efficacy in fraud detection.
- It demonstrates the application of various machine learning models, including support vector machines (SVM), random forest, and neural networks, in classifying transactions as fraudulent or legitimate.
- The framework leverages cloud computing and artificial intelligence to enable timely detection and prevention of fraudulent activities, contributing to the security and trustworthiness of digital banking systems.

**LITERATURE REVIEW:**

To overcome the imbalance in credit card transaction datasets, Ding *et al.* [16] devised a method in 2023 that improves the generator component of the Variational Autoencoder Generative Adversarial Network (VAEGAN). Our goal is to increase the training set of ensembles learning classification models by providing diversified and reliable minority class data through the proposal of a unique oversampling approach.

A machine learning approach based on real-world unbalanced datasets from European cards was suggested by Ileberi *et al.* [17] in 2021. Utilizing the Synthetic Minority over-sampling Technique (SMOTE), the class imbalance was addressed. Accuracy, recall, precision, MCC, and AUC were assessed for this framework, which includes Support Vector Machine, Logistic Regression, Random Forest, XGBoost, Decision Tree, and Extra Tree algorithms with Adaptive Boosting. It was also verified on a heavily biased artificial fraud dataset.

A detection framework utilizing quantum machine learning (QML) and Support Vector Machine (SVM) supplemented with quantum annealing solvers was employed by Wang *et al.* in 2022 [18]. QML is shown to be faster and more accurate than twelve other machine learning techniques when tested on two datasets, especially when dealing with severely skewed bank loan data. Although feature selection only slightly improves accuracy, it significantly speeds up detection.

A Spatial-Temporal Attention-based Graph Network (STAGN) for fraud detection was introduced in 2022 by Cheng *et al.* [19]. Using spatial-temporal attention and graph neural networks to learn transaction graph characteristics, STAGN achieves superior performance in precision-recall curves and AUC compared to state-of-the-art techniques. STAGN's adaptability in user behavior-based tasks is demonstrated by empirical research involving domain experts, who prove its effectiveness in detecting suspicious transactions, identifying fraud hotspots, and revealing trends. Wang *et al.* [20] employed learning automata in 2021 as part of an adaptive learning strategy. We improve flexibility and optimize window settings with the introduction of the Learning Automatic Window (LAW). In response to fluctuations in fraudulent transaction patterns, LAW dynamically modifies the window settings.

Hashemi *et al.* [21] created an approach in 2023 that makes use of class weight-tuning hyperparameters. We handle imbalances in data while maintaining practical concerns by employing Bayesian optimization. To improve LightGBM performance, we incorporate CatBoost and XGBoost and recommend weight-tuning as a preprocessing step for imbalanced data. Moreover, we use deep learning to fine-tune hyperparameters, with an emphasis on the suggested weight-tuning method.

Wang *et al.* [22] proposed CAeSaR in 2023, a system that satisfies these requirements through the use of two novel techniques: TELSI, which effectively integrates decision strategies from multiple modules, ensuring decision explainability, and TRTPT, which divides functions based on temporal positions relative to a reference fraudulent transaction.

**PROPOSED METHODOLOGY:**

Digital banking fraud prevention makes use of tools for the prompt identification and blocking of fraudulent transactions. But problems still exist, including managing massive data sets, guaranteeing model correctness, and upholding regulatory compliance. This research offers a comprehensive approach to the real-time identification and prevention of fraudulent transactions, utilizing cutting-edge AI algorithms and cloud-based infrastructure to improve accuracy and efficiency.

Overall Proposed Architecture:
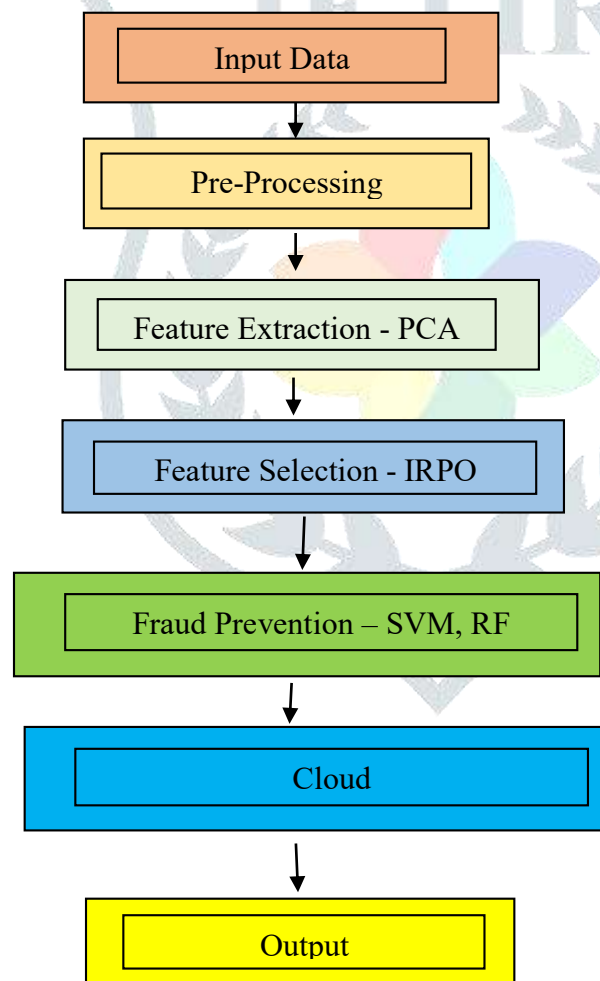


Pre-Processing:

In order to ensure dataset completeness and perform preprocessing, this work uses approaches for data normalization, standardization, and imputation. Preprocessing techniques such as data normalization and standardization are employed to guarantee that numerical characteristics within a dataset are consistently scaled, hence facilitating equitable comparisons across varying ranges. Standardization changes data to have a mean of 0 and a standard deviation of 1, whereas normalization usually adjusts data

to a range of 0 to 1 or -1 to 1. These methods are essential to machine learning algorithms because they keep characteristics that have bigger scales from predominating over smaller ones when the model is being trained. To keep a dataset complete, data imputation requires filling in missing values. There are several reasons why data may be missing, including incomplete surveys, equipment failures, and mistakes made during data entry. Imputation techniques like mean, median, or predictive imputation estimate and substitute reasonable values for missing values, preserving the dataset's suitability for analysis and modelling.

Feature Extraction:

In this paper, PCA is employed as a feature extraction technique. Principal Component Analysis (PCA) is a potent statistical method for exploring data and reducing dimensionality. High-dimensional data is transformed into a lower-dimensional space while retaining as much of the original variance as feasible. This is how it functions. By determining the major components—the paths along which the data fluctuates most—PCA is able to do this. Because these major components are uncorrelated and orthogonal to one another, they are perfect for streamlining data representation and cutting down on redundancy. PCA identifies the most relevant characteristics from the data using mathematical procedures involving eigenvalues and eigenvectors, enabling a more compact representation while maintaining the necessary structure and patterns. In addition to reducing dimensionality, PCA is useful for feature extraction, noise reduction, and data visualization in a variety of fields, supporting the study and understanding of data. It is important to use caution while interpreting PCA findings, bearing in mind the trade-offs between information loss and dimensionality reduction, in order to guarantee that the transformed data yields significant insights.

Feature Selection:

In this study, IRPO is employed to enhance classification accuracy by selecting relevant features and refining the collected data, ultimately improving model performance. Native to southern China and the eastern Himalayas, the red panda is a small mammal recognized for its reddish-brown fur and unique markings. Flourishing in temperate forests with dense bamboo cover, it excels at climbing trees. Its diet primarily consists of bamboo leaves and shoots, relying on sharp senses and climbing prowess. The design of the RPO approach is inspired by these natural attributes.

**Mathematical Modelling**

Initialization:

Functioning as a population-based metaheuristic algorithm, the RPO technique employs red pandas as symbolic representations of individual members. Each red panda represents a potential solution to a problem variable, positioned within the search space. Mathematically, each red panda is depicted as a vector, forming a matrix where rows represent potential solutions and columns hold values for associated problem variables. Initially, red panda coordinates within the search space are randomly initialized using Eq. (3) and Eq. (4). This approach facilitates exploration and exploitation of the solution space to find optimal solutions.

$$Y = \begin{bmatrix} Y_1 \\ \cdot \\ \cdot \\ \cdot \\ Y_i \\ \cdot \\ \cdot \\ Y_M \end{bmatrix}_{M \times n} = \begin{bmatrix} Y_{1,1} \dots Y_{1,j} \dots Y_{1,n} \\ \cdot \\ \cdot \\ \cdot \\ Y_{i,1} \dots Y_{i,j} \dots Y_{i,n} \\ \cdot \\ \cdot \\ Y_{M,1} \dots Y_{M,j} \dots Y_{M,n} \end{bmatrix}_{M \times n} \qquad (3)$$

$$y_{i,j} = lob_j + r_{i,j} \cdot (upb_j - lob_j), i = 1,2, \dots, M, j = 1,2, \dots, n \qquad (4)$$

The population matrix holding the red panda locations is represented by $Y$ in the RPO technique, where $Y_i$ stands for the $ith$ red panda (possible solution) and $Y_{i,j}$ for its $jth$ dimension (problem variable). $M$ is the total number of red pandas, and $n$ is the number of variables that are the problem. The $jth$ problem variable's lower and upper limits are indicated by the variables $lob_j$ and $upb_j$, respectively, and random integers $r_{i,j}$ inside the interval $[0,1]$ are used. The positions of each red panda act as potential solutions,

making it possible to assess the objective function associated with each one. A matrix of the form provided by Eq. (5) can be used to represent the final set of evaluated objective function values.

$$f = \begin{bmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_i \\ \cdot \\ \cdot \\ f_M \end{bmatrix}_{M \times 1} = \begin{bmatrix} f(Y_1) \\ \cdot \\ \cdot \\ \cdot \\ f(Y_i) \\ \cdot \\ \cdot \\ f(Y_M) \end{bmatrix}_{M \times 1} \tag{5}$$

The value obtained by the $ith$ red panda is indicated by $f_i$, and $f$ represents the vector of values of the objective function. These values of the objective function are essential for evaluating the caliber of potential solutions. The greatest and lowest values of the objective function are used to identify the best and worst potential solutions, respectively. These potential solutions are modified appropriately during every iteration. Iterative upgrades to potential solutions for the best possible problem-solving are part of the RPO's exploration and exploitation phases.

Phase 1: Exploration Strategy – Foraging:

During the initial phase of RPO, red pandas' positions mimic their foraging behavior in the wild. Leveraging their adeptness in detecting food sources, each red panda evaluates the locations of others with better objective function values as potential feeding grounds. These prospective food positions are identified through comparisons of objective function values, with each red panda randomly selecting one position using Eq. (6). This process simulates the exploration for optimal solutions in the search space.

$$pfs_i = \{Y_k | k \in \{1,2,..,M\} and \ f_k < f_i\} \cup \{Y_{best}\} \tag{6}$$

Based on a comparison with the location of the best candidate solution $Y_{best}$, the suggested food sources for each red panda $pfs_i$ are identified. Approaching these sources causes large positional shifts that improve ability of algorithm to globally search and explore. By determining new locations in relation to the food source (best candidate solution), red pandas' foraging behavior can be replicated. Eq. (7) and Eq. (8) are used to update the red panda's location if the objective function value improves at the new location.

$$Y_i^{p1}: y_{i,j}^{p1} = y_{i,j} + r.\left(sfs_{i,j} - Is.y_{i,j}\right) \tag{7}$$

$$Y_i = \begin{cases} Y_i^{p1}, f_i^{p1} < f_i \\ Y_i, else \end{cases} \tag{8}$$

The new location of the $ith$ red panda as ascertained from the RPO's first phase is represented by $Y_i^{p1}$. Objective function is denoted by $f_i^{p1}$, and its position in the $jth$ dimension is indicated by $y_{i,j}^{p1}$. For the $ith$ red panda, $sfs_i$ denotes the preferred food source, and $sfs_{i,j}$ denotes its location in the $jth$ dimension. $Is$ is a randomly chosen number from the set $\{1, 2\}$, and the variable $r$ is a random value between 0 and 1.

Phase 2: Proficiency in ascending and perching on trees (exploitation):

During the second phase of RPO, red pandas' tree-climbing behavior guides their positioning. These animals typically rest on trees for extended periods and move to nearby trees for food after ground foraging. This behavior induces slight positional adjustments, enhancing the RPO algorithm's exploitation and local search capabilities in promising areas. Mathematically, this behavior entails computing new positions for each red panda and updating previous positions if the objective function improves, as described in Eq. (9) and Eq. (10). This process mimics the iterative refinement of solutions as red pandas navigate the search space in pursuit of optimal solutions.

$$Y_{i,j}^{p2} = y_{i,j} + \frac{lob_j + r_{i,j} \cdot (upb_j - lob_j)}{t}, i = 1,2,\dots,M, j = 1,2,\dots,n, t = 1,2,\dots,T \quad (9) \qquad Y_i = \begin{cases} Y_i^{p2}, f_i^{p2} < f_i \\ Y_i, else \end{cases}$$

$$(10)$$

The $ith$ red panda's modified position, obtained from the second phase of RPO, is represented by $Y_i^{p2}$. Objective function is shown by $f_i^{p2}$, and its position in the jth dimension is indicated by $Y_{i,j}^{p2}$. A random number between 0 and 1 represents the variable $r$. The symbol $t$ denotes the algorithm's iteration counter, whereas $T$ stands for the maximum iterations.

**Fraud Prevention**

Develop machine learning models such as supervised learning classifiers namely SVM, and NB to detect fraudulent transactions based on selected features.

SVM:

SVM is a powerful supervised learning technique that may be used for both classification and regression issues. By identifying the optimal hyperplane in a high-dimensional space that divides data points of different classes by the largest amount, SVM aims to maximise the margin of separation in classification. This hyperplane is chosen to decrease classification error and maximise the separation between each class's support vectors, or nearest data points. SVM works well with both linearly and non-linearly separable data by using a variety of kernel functions to transform the data into a higher-dimensional space where a linear decision boundary may be applied. SVM, which were originally created for classification applications, have been adjusted to handle nonlinear regression challenges using an ε-insensitive loss function. The input data, x, in support vector regression is transformed into a higher-dimensional feature space via a kernel function. Next, a linear model f(x,ω) is built in this modified space.

The term $g_j(x)$ represents a series of nonlinear transformations, while $b$ denotes the bias term. Estimation quality is evaluated using the ε-insensitive loss function. In regression scenarios, errors exceeding the threshold $\varepsilon$ are penalized by the loss function, typically resulting in a sparse decision rule representation with notable algorithmic and representational benefits. One drawback of SVM is their inability to provide direct probability estimates, necessitating computationally intensive fivefold cross-validation for calculation.

Naive Bayes (NB):

Based on Bayes' theorem, the Naive Bayes (NB) algorithm is a probabilistic classification method that determines the likelihood of a hypothesis given the available data. Because of its efficacy and efficiency, NB is utilised extensively in machine learning applications despite its simplicity, especially in text categorisation and spam filtering. The "naive" name comes from the algorithm's assumption of feature independence, which means that each feature is treated independently of all others given the class label. Because of this, calculating probabilities is much easier, and even with big datasets, NB remains computationally efficient. When performing classification tasks, NB multiplies the prior probability of each class by the conditional likelihood of witnessing the characteristics given each class to get the probability of each class label given a set of features. For the input data, the class with the highest probability is projected to be the class label. Even while NB's assumption of feature independence may oversimplify real-world data, it frequently works well in practice and is especially helpful with high-dimensional datasets.

**Result and Discussion**

Dataset Description:

The Fraud Detection on Bank Payments dataset [23] is a synthetic dataset generated to simulate transactions from a financial payment system. It contains information about various transactions, including features such as transaction amount, merchant

category, transaction time, and whether the transaction is fraudulent or not. This dataset is designed to resemble real-world banking transaction data, allowing researchers and data scientists to develop and test fraud detection algorithms and models. The goal of analyzing this dataset is to identify patterns and characteristics associated with fraudulent transactions, enabling financial institutions to implement effective fraud prevention measures and protect their customers from financial losses.

Overall Performance Analysis of Proposed and Existing Model:

The table provides a comprehensive comparison of the proposed SVM-NB model with existing methods, including SVM, RF, and Naive Bayes. The metrics evaluated include Accuracy, Precision, Recall, and F1-score. The proposed SVM-NB model demonstrates superior performance across all metrics, with an impressive Accuracy of 98.46%. It exhibits high Precision (96.74%), Recall (95.46%), and F1-score (97.74%), indicating its proficiency in correctly classifying instances while minimizing false positives and false negatives. In contrast, SVM and RF achieve lower Accuracy (92.32% and 92.07%, respectively) and slightly inferior Precision, Recall, and F1-scores compared to the proposed model. Naive Bayes, although competitive in Precision and Recall, lags behind in overall performance with the lowest Accuracy (91.54%) and F1-score (90.82%) among the methods evaluated. Overall, the results underscore the effectiveness of the proposed SVM-RF model in classification tasks, offering superior performance compared to existing methods.

Table: Performance Analysis of Existing and Proposed Model

| Methods | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| Proposed SVM-NB | 98.46 | 96.74 | 95.46 | 97.74 |
| SVM | 92.32 | 91.86 | 92.07 | 92.3 |
| RF | 92.07 | 91.13 | 91.54 | 91.94 |
| NB | 91.54 | 91.07 | 91.86 | 90.82 |

Overall Graphical Representation: Existing and Proposed Graphical Representation
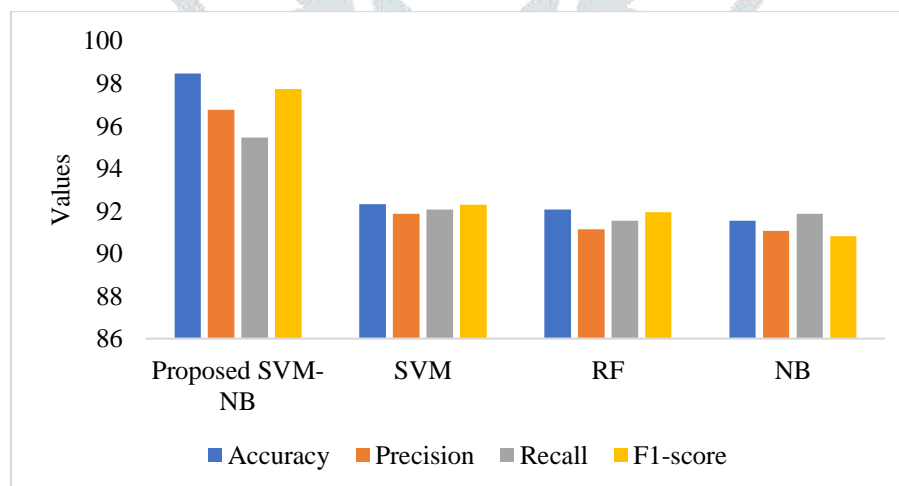


The graphical representation comparing the existing approach with the proposed method. It illustrates the performance metrics of both methods, showcasing the superiority of the proposed framework. The graph highlights key evaluation criteria such as accuracy, precision, recall, and F1-score. Through this visual comparison, it becomes evident how the proposed approach outperforms the existing method in terms of fraud detection effectiveness, providing a clear overview of the advancements achieved by the new model.

**CONCLUSION:**

This study created a thorough method that makes use of cutting-edge artificial intelligence algorithms and cloud-based infrastructure to identify and stop fraudulent transactions in real time. To ensure data integrity and completeness, preprocessing was done after data was collected from many sources. Key indicators indicative of fraudulent conduct was identified by applying feature extraction techniques, such as dimensionality reduction. To improve model performance, feature selection was adjusted by IRPO. Later, to distinguish transactions as fraudulent or valid, machine learning models were created and used, such as NB and SVM. This framework increased the security and reliability of digital financial systems by utilizing cloud computing and artificial intelligence to allow prompt identification and prevention of fraudulent activity.

**REFERENCES:**

[1] Găbudeanu, L., Brici, I., Mare, C., Mihai, I. C., & Șcheau, M. C. (2021). Privacy intrusiveness in financial-banking fraud detection. Risks, 9(6), 104.

[2] Hashemi, S. K., Mirtaheri, S. L., & Greco, S. (2022). Fraud detection in banking data by machine learning techniques. IEEE Access, 11, 3034-3043.

[3] Vanini, P., Rossi, S., Zvizdic, E., & Domenig, T. (2023). Online payment fraud: from anomaly detection to risk management. Financial Innovation, 9(1), 66.

[4] Sarma, D., Alam, W., Saha, I., Alam, M. N., Alam, M. J., & Hossain, S. (2020, July). Bank fraud detection using community detection algorithm. In 2020 second international conference on inventive research in computing applications (ICIRCA) (pp. 642-646). IEEE.

[5] Eshghi, A., & Kargari, M. (2019). Introducing a new method for the fusion of fraud evidence in banking transactions with regards to uncertainty. Expert Systems with Applications, 121, 382-392.

[6] Daliri, S. (2020). Using harmony search algorithm in neural networks to improve fraud detection in banking system. Computational Intelligence and Neuroscience, 2020.

[7] Darwish, S. M. (2020). A bio-inspired credit card fraud detection model based on user behavior analysis suitable for business management in electronic banking. Journal of Ambient Intelligence and Humanized Computing, 11(11), 4873-4887.

[8] Singh, A., Jain, A., & Biable, S. E. (2022). Financial fraud detection approach based on firefly optimization algorithm and support vector machine. Applied Computational Intelligence and Soft Computing, 2022.

[9] Darwish, S. M. (2020). A bio-inspired credit card fraud detection model based on user behavior analysis suitable for business management in electronic banking. Journal of Ambient Intelligence and Humanized Computing, 11(11), 4873-4887.

[10] Daliri, S. (2020). Using harmony search algorithm in neural networks to improve fraud detection in banking system. Computational Intelligence and Neuroscience, 2020.

[11] Cui, J., Yan, C., & Wang, C. (2021). ReMEMBeR: Ranking metric embedding-based multicontextual behavior profiling for online banking fraud detection. IEEE Transactions on Computational Social Systems, 8(3), 643-654.

[12] Eshghi, A., & Kargari, M. (2019). Introducing a new method for the fusion of fraud evidence in banking transactions with regards to uncertainty. Expert Systems with Applications, 121, 382-392.

[13] Karthik, V. S. S., Mishra, A., & Reddy, U. S. (2022). Credit card fraud detection by modelling behaviour pattern using hybrid ensemble model. Arabian Journal for Science and Engineering, 47(2), 1987-1997.

[14] C. Iscan, O. Kumas, F. P. Akbulut and A. Akbulut, "Wallet-Based Transaction Fraud Prevention Through LightGBM With the Focus on Minimizing False Alarms," in IEEE Access, vol. 11, pp. 131465-131474, 2023, doi: 10.1109/ACCESS.2023.3321666.

[15] W. Ning, S. Chen, S. Lei and X. Liao, "AMWSPLAdaboost Credit Card Fraud Detection Method Based on Enhanced Base Classifier Diversity," in IEEE Access, vol. 11, pp. 66488-66496, 2023, doi: 10.1109/ACCESS.2023.3290957.

[16] Y. Ding, W. Kang, J. Feng, B. Peng and A. Yang, "Credit Card Fraud Detection Based on Improved Variational Autoencoder Generative Adversarial Network," in IEEE Access, vol. 11, pp. 83680-83691, 2023, doi: 10.1109/ACCESS.2023.3302339.

[17] E. Ileberi, Y. Sun and Z. Wang, "Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost," in IEEE Access, vol. 9, pp. 165286-165294, 2021, doi: 10.1109/ACCESS.2021.3134330.

[18] H. Wang, W. Wang, Y. Liu and B. Alidaee, "Integrating Machine Learning Algorithms With Quantum Annealing Solvers for Online Fraud Detection," in IEEE Access, vol. 10, pp. 75908-75917, 2022, doi: 10.1109/ACCESS.2022.3190897.

[19] D. Cheng, X. Wang, Y. Zhang and L. Zhang, "Graph Neural Network for Fraud Detection via Spatial-Temporal Attention," in IEEE Transactions on Knowledge and Data Engineering, vol. 34, no. 8, pp. 3800-3813, 1 Aug. 2022, doi: 10.1109/TKDE.2020.3025588.

[20] C. Wang, C. Wang, H. Zhu and J. Cui, "LAW: Learning Automatic Windows for Online Payment Fraud Detection," in IEEE Transactions on Dependable and Secure Computing, vol. 18, no. 5, pp. 2122-2135, 1 Sept.-Oct. 2021, doi: 10.1109/TDSC.2020.3037784.

[21] S. K. Hashemi, S. L. Mirtaheri and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," in IEEE Access, vol. 11, pp. 3034-3043, 2023, doi: 10.1109/ACCESS.2022.3232287.

[22] C. Wang, S. Chai, H. Zhu and C. Jiang, "CAeSaR: An Online Payment Anti-Fraud Integration System With Decision Explainability," in IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 3, pp. 2565-2577, 1 May-June 2023, doi: 10.1109/TDSC.2022.3186733.

[23] Dataset taken from: "https://www.kaggle.com/code/turkayavci/fraud-detection-on-bank-payments", dated 26-04-2024.