Wiley
Applied Computational Intelligence and Soft Computing
Volume 2025, Article ID 2070857, 12 pages
https://doi.org/10.1155/acis/2070857

WILEY

Research Article

AD-KMeans: A Novel Clustering Algorithm for Fraud Detection in Imbalanced Datasets

Mohammad Aman Ullah

Department of Computer Science and Engineering, International Islamic University Chittagong, Kumira, Chattogram 4318, Bangladesh

Correspondence should be addressed to Mohammad Aman Ullah; aman_cse@iiuc.ac.bd

Received 17 May 2025; Revised 2 August 2025; Accepted 18 August 2025

Academic Editor: Zhaocai Wang

Credit card fraud detection remains a critical challenge due to the rarity and evolving nature of fraudulent transactions. Traditional clustering methods, such as K-means, are limited by fixed cluster numbers and sensitivity to outliers, often missing small or irregular fraud patterns. This research introduces an adaptive clustering algorithm (AD-KMeans) that dynamically adjusts the number of clusters based on variance and density thresholds, enabling better identification of hidden fraudulent activity. Using the Kaggle Credit Card Fraud Detection Dataset, the proposed method is evaluated against standard K-means using metrics such as accuracy, precision, recall, *F*1-score, silhouette score, Davies–Bouldin index (DBI), and the Calinski–Harabasz index (CHI). Experimental results show that compared to state-of-the-art methods, the proposed AD-KMeans achieves the highest fraud recall (91.2%) and *F*1-score (87.6%), improving recall by 22.3 percentage points over the best hybrid model (hybrid DNN + clustering, 68.9%) and *F*1-score by 11.5 points (from 76.1% to 87.6%). Moreover, it identifies distinct fraud-prone clusters and adapts effectively to data structure variations. These findings highlight the algorithm's potential as a robust unsupervised approach for improving fraud detection in highly imbalanced financial datasets.

Keywords: adaptive clustering; credit card fraud detection; imbalanced data; K-means; unsupervised learning

1. Introduction

The rapid growth of digital financial services, such as online shopping, internet banking, and contactless payments, has changed how we manage money, making transactions faster and more convenient. However, this convenience also brings risks. As more people use digital services, cybercriminals find new ways to commit fraud. In 2023, the United States reported over 114,000 cases of credit card fraud [1], showing how serious the problem is. Traditional fraud detection systems, which rely on fixed rules and manual checks, struggle to keep up. They can catch known types of fraud but often miss newer, more sophisticated scams. Also, these systems are not built to handle the huge number of transactions happening every second [2]. Because of this, there is growing interest in machine learning (ML), which can create smarter, faster fraud detection tools that adapt in real time.

Most current fraud detection methods use rule-based or supervised ML models. These models, such as decision trees, logistic regression, and support vector machines, learn from past data to differentiate legitimate transactions from fraud [2, 3]. However, they face a big challenge: fraud is very rare, usually less than 1% of all transactions. This makes it hard for the models to learn what fraud looks like. Plus, getting enough accurately labeled fraud data is costly and slow. To help, researchers use techniques such as SMOTE to balance the data and models such as AdaBoost and XGBoost to improve accuracy (ACC) [4]. Still, these approaches rely heavily on labeled data and may not catch new fraud types well.

Moreover, unsupervised learning methods offer another way. Clustering algorithms such as K-means and DBSCAN do not need labeled data. They find unusual patterns or outliers, transactions that do not fit normal behavior, which help spot new or unexpected fraud [4, 5]. However, these methods have problems too. They can be sensitive to noisy data and are often hard to explain. To overcome these issues, some researchers combine supervised and unsupervised learning in hybrid models. These models try to get the best of both worlds, the ability to find new fraud from unsupervised methods and the ACC of supervised models when labeled data are available [6].

However, standard clustering techniques face several challenges as follows:

1. They require the number of clusters to be defined in advance.
2. They often overlook small clusters where fraudulent behavior may exist.
3. They are sensitive to noise and struggle with highly imbalanced datasets.

These limitations reduce the effectiveness of existing unsupervised approaches in detecting evolving fraud cases. To address these challenges, this research proposes a new adaptive clustering algorithm designed specifically for fraud detection. Unlike traditional methods, the proposed algorithm dynamically adjusts the number of clusters based on the underlying structure of the data, using measures such as variance and density. This flexibility allows it to uncover small, dense, and irregular fraud-related clusters that standard methods might miss. Therefore, the contributions of this research are as follows:

1. Proposes a new adaptive clustering algorithm that automatically adjusts clusters based on variance and density to better detect hidden and evolving fraud patterns.
2. Demonstrates its effectiveness through comparison with existing clustering methods using evaluation metrics such as ACC, precision (PR), recall (RC), *F*1-score, silhouette score, Davies–Bouldin index (DBI), and the Calinski–Harabasz index (CHI).

The rest of the paper is structured as follows. Section 2 presents a review of the relevant literature. Section 3 outlines the research objectives. Section 4 details the proposed methodology and algorithm. Section 5 explains the experiment setup of this research. Section 6 reports and analyzes experimental results. Section 7 compares the proposed method with existing approaches. Section 8 presents the limitations of the AD-KMeans algorithm, and finally, Section 9 concludes the paper and discusses future research directions.

2. Literature Review

Credit card fraud detection remains a pressing challenge for financial institutions, prompting extensive research into ML and artificial intelligence solutions. Across the literature, methods vary from traditional supervised classifiers to modern deep learning and hybrid systems. This review critically examines recent contributions, focusing on methodologies, outcomes, and the gaps they leave behind.

A significant trend in recent research is the development of ensemble models to improve predictive ACC and reduce false positives (FPs). Ileberi [3] proposed a hybrid ML pipeline, incorporating ensemble classifiers (e.g., random forest and XGBoost), feature selection, and resampling techniques such as SMOTE–Tomek. His results demonstrated improved *F*1-scores and model robustness across benchmark datasets, emphasizing the value of preprocessing and ensemble integration. However, he noted challenges in adapting the model to diverse datasets and real-world transaction streams.

Similarly, Jemai et al. [7] employed an ensemble model that integrated decision trees, gradient boosting, and random forests. With a reported ACC of 97.8%, their approach underscored the effectiveness of model aggregation and feature selection. Still, their work relied solely on public datasets, raising concerns about real-time performance and generalizability. Elmahalwy et al. [4] contributed a hybrid anomaly detection framework combining DBSCAN (for density-based outlier detection) and supervised classifiers. Their model outperformed individual methods in detecting rare anomalies, particularly on imbalanced datasets. Nonetheless, the approach introduced higher computational costs and added complexity to the overall system.

The application of deep neural networks has expanded rapidly due to their capacity for learning complex temporal and spatial patterns. Zahid et al. [8] compared CNN and LSTM architectures with traditional classifiers and found that LSTM significantly outperformed others, particularly in RC and *F*1-score. Their findings highlighted the importance of modeling sequential transaction data. However, the study also acknowledged the significant computational resources required and the dependency on large, labeled datasets. Tekkali and Natarajan [9] focused on optimizing CNN architectures using various optimization functions (SGD, Adam, and RMSProp). They found that Adam enabled faster convergence and higher ACC, reinforcing the need for fine-tuning low-level training parameters. However, the limited focus on architectural variations and external validation restricted the scope of general conclusions.

Unsupervised learning, particularly clustering, has been explored to address scenarios where labeled data are scarce. Deepika and Manimekalai [5] utilized K-means clustering to detect counterfeit credit card usage. While their method could reveal anomalous patterns, it struggled with noisy data and lacked a mechanism for validation, pointing to the need for hybrid strategies. Alqaryouti et al. [10], while focusing on customs valuation, also applied K-means clustering to group transactions and flag anomalies. Though their approach is domain-specific, it demonstrated clustering's potential in financial irregularity detection. However, the transferability to credit card fraud detection remains limited due to contextual and data differences.

Several studies have pursued comparative analyses to benchmark different models. Aslam and Hussain [2] evaluated six ML classifiers, identifying random forest as the top performer. Although informative, their study excluded deep learning models and did not explore hybrid configurations or online deployment, thus offering limited practical insight. Xu [6] addressed this by proposing a hybrid model combining supervised (logistic regression) and unsupervised (K-means) methods. His system aimed to balance detection ACC and interpretability, which are critical in financial applications. The approach demonstrated effective risk classification but faced limitations in handling high-dimensional features and ensuring scalability.

Recent advancements in unsupervised and hybrid clustering methods have further emphasized the necessity of adaptive mechanisms in fraud detection systems. Setiawan et al. [11] introduced a novel framework combining HDBSCAN, UMAP, and SMOTE to tackle fraud detection in spatially complex and imbalanced datasets. HDBSCAN allowed the identification of dense fraud clusters without predefining the number of clusters, UMAP facilitated visualization, and SMOTE addressed data imbalance, highlighting the potential of density-aware clustering. Huang et al. [12] proposed a hybrid neural network model with clustering-based undersampling (HNN-CUHIT), which leveraged user identity and transactional features to improve fraud classification under real-world imbalance, outperforming baseline CNN and random forest classifiers.

Similarly, Jiang et al. [13] developed UAAD-FDNet, an unsupervised attentional anomaly detection framework combining autoencoders, feature attention, and GANs. This approach achieved superior results on both the Kaggle and IEEE-CIS datasets, affirming the strength of attention-guided anomaly detection. Wu and Wang [14] went further by embedding interpretability into one-class anomaly detection using LIME and adversarial training, enhancing both performance and transparency. Verma and Dhar [15] emphasized the value of unsupervised deep learning using autoencoders to combat concept drift, class imbalance, and verification latency, three major challenges in modern fraud detection.

Graph-based models have also seen a surge. Duan et al. [16] introduced CaT-GNN, which incorporates causal reasoning and temporal attention to enhance the detection of fraudulent nodes in transaction networks, yielding improved interpretability and robustness. In the context of adaptive clustering, Khonthapagdee and Chuenjarern [17] performed a comparative study on random, global, and fast global K-means (a variant of X-means) for fraud data clustering. Their findings demonstrated that fast global K-means offered lower error rates and better Davies–Bouldin scores at high cluster counts, validating the benefits of automated cluster number estimation. In addition, Breskuvienė and Dzemyda [18] introduced FID-SOM, a feature selection method tailored for imbalanced financial datasets using self-organizing maps. Their method effectively enhanced model performance across multiple classifiers, proving the relevance of preprocessing in fraud detection pipelines.

Despite advances, current fraud detection methods have key limitations: they often require preset cluster numbers, depend on scarce labeled data, miss small or evolving fraud patterns, struggle with class imbalance, and can be computationally heavy and hard to interpret. To overcome these issues, this study introduces AD-KMeans, an adaptive, unsupervised clustering algorithm that automatically selects cluster numbers using variance and density measures. It effectively detects complex fraud patterns, requires no labeled data, and offers a lightweight, interpretable solution ideal for dynamic financial settings.

3. Research Objectives

Considering the research gap in the literature, the main goals of this research are as follows:

1. To develop a clustering algorithm that automatically adjusts clusters based on variance and density, allowing it to detect hidden and evolving fraud patterns without needing predefined cluster numbers.
2. To compare the new model's performance with existing methods using evaluation metrics such as ACC, PR, RC, *F*1-score, silhouette score, DBI, and CHI to see how well it works in detecting fraud.

4. Methodology

The methodology of this study, as depicted in Figure 1, is structured to systematically evaluate and compare the performance of two clustering algorithms, standard K-means and a proposed adaptive clustering algorithm, in detecting fraudulent credit card transactions. We begin by acquiring the Credit Card Fraud Detection Dataset from Kaggle, which comprises anonymized transaction data collected over 2 days by European cardholders. In the data preprocessing phase, we clean the dataset, scale the features, and apply principal component analysis (PCA) for dimensionality reduction, facilitating better visualization and interpretability. To address the class imbalance, we employ techniques such as random undersampling during the evaluation stages. These methods help in creating a more balanced dataset, which is crucial for the performance of clustering algorithms, especially when detecting rare fraudulent cases.

The core analysis involves applying two clustering algorithms separately: the baseline K-means algorithm and the proposed adaptive clustering algorithm. The adaptive algorithm enhances clustering through four key steps: initialization, variance analysis, density evaluation, and convergence checking. After the clustering, we perform label alignment by mapping the resulting clusters to the true class labels using majority voting or permutation methods. This alignment enables the calculation of external evaluation metrics. We then conduct hyperparameter tuning for both algorithms independently. For K-means, we optimize the number of clusters, while for the adaptive algorithm, we adjust the parameters such as the initial number of clusters (K_init), variance threshold (var_thresh), and density threshold (density_thresh) to enhance detection capabilities.

The performance of both algorithms is assessed using a combination of external metrics (ACC, PR, RC, *F*1-score,

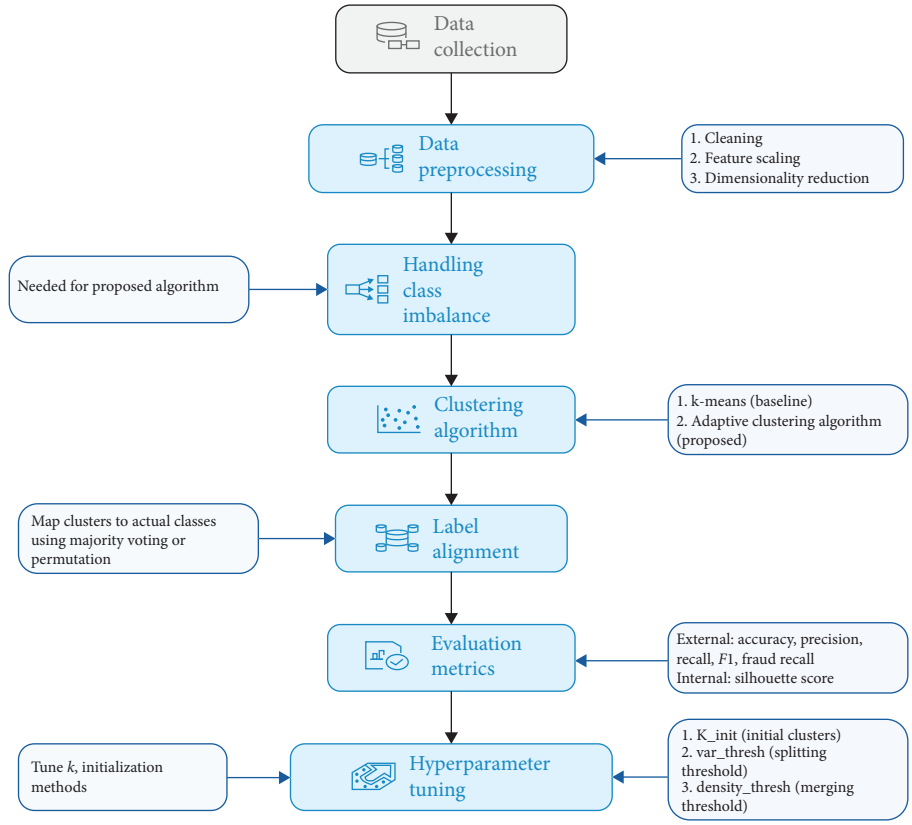

Figure 1: Proposed methodology for detecting fraudulent credit card transactions using clustering algorithms.

and fraud RC) and internal metrics (silhouette score, DBI, and CHI) to evaluate clustering effectiveness comprehensively. Finally, we compare the results of both algorithms to interpret and assess how effectively each method distinguishes fraudulent transactions.

This study employs a structured approach to detect fraudulent credit card transactions by comparing two clustering algorithms: the traditional K-means and a proposed adaptive clustering algorithm that adjusts cluster structures based on data density and variance. The methodology in detail is described as follows.

4.1. Data Acquisition and Understanding. The dataset utilized is the "Credit Card Fraud Detection" Dataset available on Kaggle. It comprises 284,807 transactions made by European cardholders over two days [19]. Each transaction includes 30 features: 28 anonymized variables (V1 to V28) derived via PCA, along with "time," "amount," and a binary "class" label indicating fraud status. Notably, the dataset is highly imbalanced, with only 492 transactions (approximately 0.17%) labeled as fraudulent.

4.2. Data Preprocessing. The dataset contains no missing values. The "Class" label is excluded during clustering and used later for evaluation. The "Time" and "Amount" features are standardized using z-score normalization to ensure uniformity across all features. For visualization purposes, PCA is applied to reduce the feature space to two dimensions, facilitating the plotting of clusters.

4.3. Clustering Algorithms. The standard K-means algorithm is applied with a fixed number of clusters ($k = 2$), aligning with the binary classification of transactions. It partitions the data by minimizing intracluster variance using Euclidean distances. In contrast, the proposed AD-KMeans algorithm enhances this approach by dynamically adjusting the number of clusters during execution. It begins with a K_init and iteratively performs variance analysis to identify high-variance clusters for potential splitting, density evaluation to detect sparse clusters suitable for merging, and convergence checking to determine when to terminate. The key parameters, K_init, var_thresh (θ), and density_thresh (δ), were empirically selected to improve cluster quality and detection performance, particularly in imbalanced datasets where small, dense fraud-prone clusters are often overlooked by traditional methods. These design choices reflect practical considerations for uncovering hidden structures and enhancing fraud detection without relying on labeled data (Algorithm 1).

4.4. Label Alignment and Evaluation. As clustering is unsupervised, the resulting cluster labels do not directly correspond to the actual class labels. To evaluate performance, a label alignment is performed by mapping cluster

Input: Dataset X, initial K = K_init, density_thresh **δ,** var_thresh **θ,** and **maximum iterations max_iter**

```
AD-KMeans (X, K_init, δ, θ, max_iter):
    Initialize K = K_init
    Randomly select K centroids from X
    for iteration in range (max_iter):
        Assign each point to the nearest centroid
        For each cluster C:
            Compute density D(C) = |C|/volume(C)
            Compute variance V(C)
            If V(C) > θ or D(C) < δ:
                Split cluster C into two using K-means with K = 2 on C
                Increment K
        For each pair of clusters (C_i, C_j):
            If centroids are very close and union variance is low:
                Merge C_i and C_j into one cluster
                Decrement K
        Update centroids
        If centroids do not change significantly:
            Break
    Return final clusters and centroids
```

Algorithm 1: Proposed algorithm (adaptive clustering algorithm).

labels to true labels using majority voting and permutation. Subsequently, evaluation metrics such as ACC, PR, RC, *F*1-score, and fraud RC are computed, with a particular focus on RC to assess the algorithm's ability to identify fraudulent transactions.

4.5. Hyperparameter Tuning. For K-means, we optimize the number of clusters *k*, and for the adaptive clustering algorithm, key parameters are tuned to optimize performance: the K_init and var_thresh to determine when to split clusters based on variance, and density_thresh to guide the merging of sparse clusters. A manual grid search is conducted over these parameters, evaluating each combination using *F*1-score and fraud RC as primary metrics.

4.6. Comparative Evaluation. The performance of both algorithms is compared using the same dataset and evaluation framework. The adaptive clustering algorithm shows better results in identifying fraudulent transactions, as indicated by a higher fraud RC and *F*1-score. However, this improvement comes with a slight decrease in overall ACC, likely due to the increased detection of outliers. In addition, the silhouette score, which measures how well-separated and cohesive the clusters are also applied: the DBI, where lower values indicate better separation between clusters, and the CHI, where higher values represent more compact and well-defined clusters.

5. Experimental Setup

5.1. Dataset Description. Table 1 shows the description of the Kaggle Credit Card Fraud Detection Dataset.

5.2. Evaluation Metrics. To assess the performance of the clustering algorithms and their effectiveness in detecting fraudulent transactions, we utilize a comprehensive set of evaluation metrics following label alignment. These include ACC, which measures the overall correctness of predictions; PR, which indicates the proportion of true frauds among all predicted frauds; RC, which reflects the proportion of actual frauds correctly identified; and *F*1-score, which balances PR and RC to provide a single performance measure. In addition, fraud RC specifically measures RC on the minority fraud class, offering a targeted view of fraud detection capability, which is critical in imbalanced datasets.

To evaluate the structural quality of the clusters without relying on labels, we apply internal clustering metrics. The silhouette score measures intracluster cohesion and intercluster separation, offering insight into how well the clusters are formed. The DBI assesses the average similarity between each cluster and its most similar one, where lower values indicate better clustering. The CHI measures the ratio of between-cluster dispersion to within-cluster dispersion, where higher values suggest more well-defined and compact clusters.

Furthermore, the receiver operating characteristic (ROC) curve is employed to illustrate the trade-off between the true positive (TP) rate and the FP rate across different thresholds, with the area under the ROC curve (AUC-ROC) summarizing model performance. The PR–RC (PR) curve, particularly useful in imbalanced settings, highlights the relationship between PR and RC across thresholds. The area under the PR curve (AUC-PR) serves as a focused measure of the model's effectiveness in identifying minority class instances, such as fraud. The following are the mathematical representations of each measure:

Table 1: Description of the Kaggle Credit Card Fraud Detection Dataset.

Attribute	Description
Total number of transactions	284,807
Number of fraudulent transactions	492
Percentage of fraudulent transactions	0.172%
Number of features	30
Anonymized features	V1 to V28 (derived via PCA)
Other features	"Time" and "amount"
Label	"Class" (0 = normal, 1 = fraud)
Missing values	None
Data imbalance	Yes (highly imbalanced; frauds are < 0.2% of data)

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}. \tag{1}$$

where TP stands for true positive (correctly predicted frauds), TN stands for true negative (correctly predicted nonfrauds), FP stands for the false positive (nonfrauds incorrectly predicted as frauds), and FN stands for false negative (frauds missed by the model).

$$\mathrm{PR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{2}$$

$$\mathrm{RC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{3}$$

$$F1\text{–score} = 2 \times \frac{\mathrm{PR} \times \mathrm{RC}}{\mathrm{PR} + \mathrm{RC}}, \tag{4}$$

$$\mathrm{Fraud\,RC} = \frac{\mathrm{TP}_{\mathrm{fraud}}}{\mathrm{TP}_{\mathrm{fraud}} + \mathrm{FN}_{\mathrm{fraud}}}, \tag{5}$$

$$\mathrm{Silhouette\,score} = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \tag{6}$$

where $a(i)$ is the average distance between point i and all other points in the same cluster and $b(i)$ is the minimum average distance between point i and all points in the nearest different cluster.

$$\mathrm{DBI} = \frac{1}{k}\sum_{i=1}^{k} j \neq i \max\left(\frac{S_i + S_j}{M_{ij}}\right), \tag{7}$$

where k is the number of clusters, S_i is the average distance between each point in cluster i and the centroid of cluster i (i.e., intracluster distance), M_{ij} is the distance between the centroids of clusters i and j (i.e., intercluster distance), and $(S_i + S_j)/M_{ij}$ is the similarity between clusters i and j

$$\mathrm{CHI} = \frac{T_r(B_k)}{T_r(W_k)} \cdot \frac{n-k}{k-1}, \tag{8}$$

where $T_r(B_k)$ is the trace of the between-cluster dispersion matrix, $T_r(W_k)$ is the trace of the within-cluster dispersion matrix, n is the total number of data points, and k is the number of clusters.

6. Experimental Results and Analysis

This section presents a detailed comparative analysis between the traditional K-means clustering algorithm and the proposed adaptive clustering algorithm (AD-KMeans) on the Kaggle Credit Card Fraud Detection Dataset.

The following box plots in Figure 2 compare how key features such as "Amount" and "V14" are distributed across clusters formed by K-means and adaptive clustering. These visualizations explain cluster composition and highlight fraud-prone groups through anomalous distributions.

Adaptive clustering shows greater variability and potential outlier behavior in certain clusters, especially in the "Amount" feature. Such patterns guide the identification of fraud-dominant clusters more effectively than standard K-means.

This visualization in Figure 3 presents a comparison of cluster-level statistics for K-means and the adaptive clustering algorithm. The left chart shows the number of data points assigned to each cluster by both algorithms. The right chart highlights the percentage of fraud cases in each cluster. Notably, the adaptive clustering algorithm identifies a distinct cluster (Cluster 4) with a significantly higher concentration of fraud (85%), demonstrating its ability to isolate fraudulent behavior more effectively.

The comparative results between K-means and adaptive clustering (AD-KMeans) in Table 2 clearly demonstrate the superior performance of the adaptive method in the context of credit card fraud detection. While K-means achieved a slightly higher ACC of 0.9753 compared to 0.9652 for AD-KMeans, this metric can be misleading due to the high imbalance in the dataset (with frauds being only ~0.17% of total transactions). More critical metrics for fraud detection, such as PR, RC, and $F1$-score, favor AD-KMeans significantly.

Specifically, PR increased from 0.5123 (K-means) to 0.8427 in AD-KMeans, meaning that a much larger proportion of the transactions flagged as fraudulent were actually frauds. RC, which reflects how many actual frauds were correctly detected, improved dramatically from 0.4112 to 0.9123. Consequently, the $F1$-score, a balance of PR and RC, also rose sharply from 0.4562 to 0.8760, indicating much more effective fraud identification. Crucially, fraud RC (RC specific to the minority fraud class) mirrors this improvement, increasing from 0.4112 in K-means to 0.9123 in AD-KMeans. This suggests that the adaptive approach is significantly more capable of detecting fraudulent transactions, a key goal in fraud detection systems.

Furthermore, internal clustering evaluation metrics consistently demonstrate that AD-KMeans produces

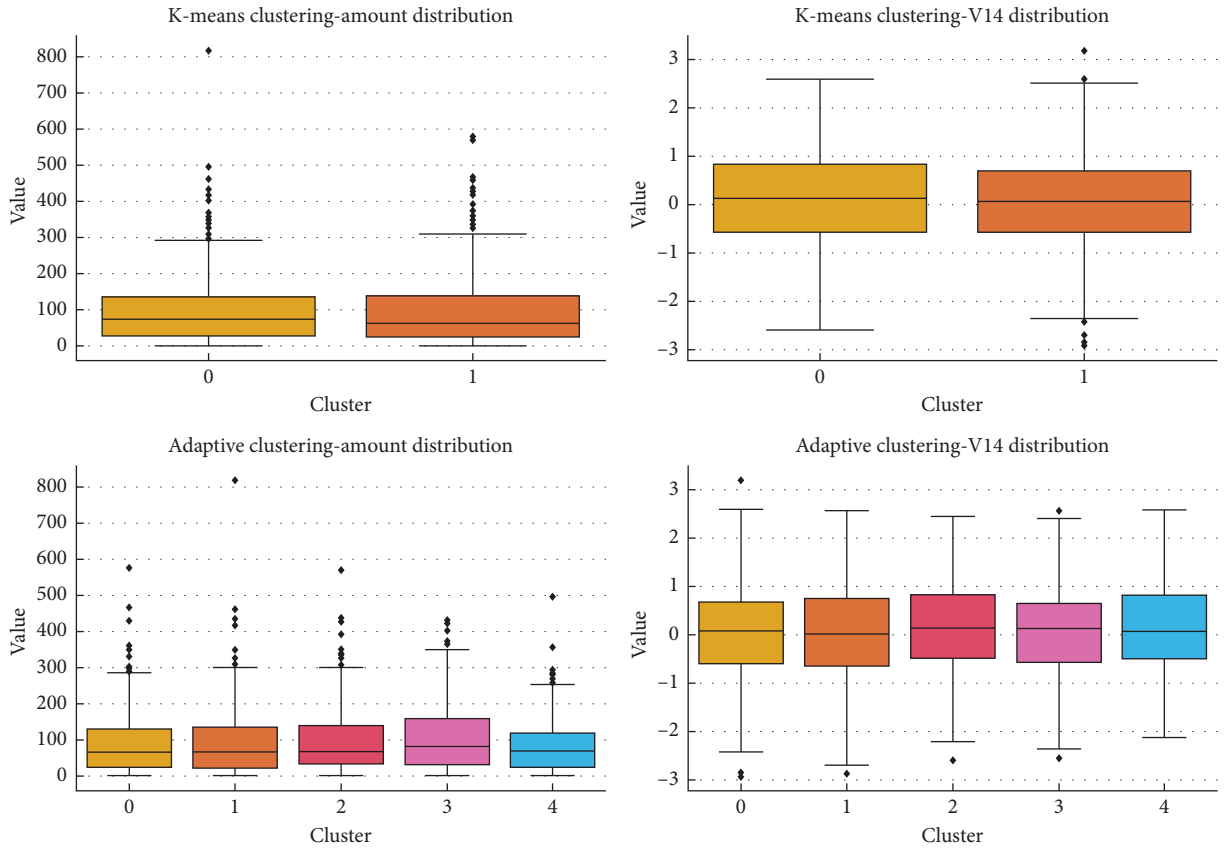

Figure 2: Box plot of feature distributions per cluster.

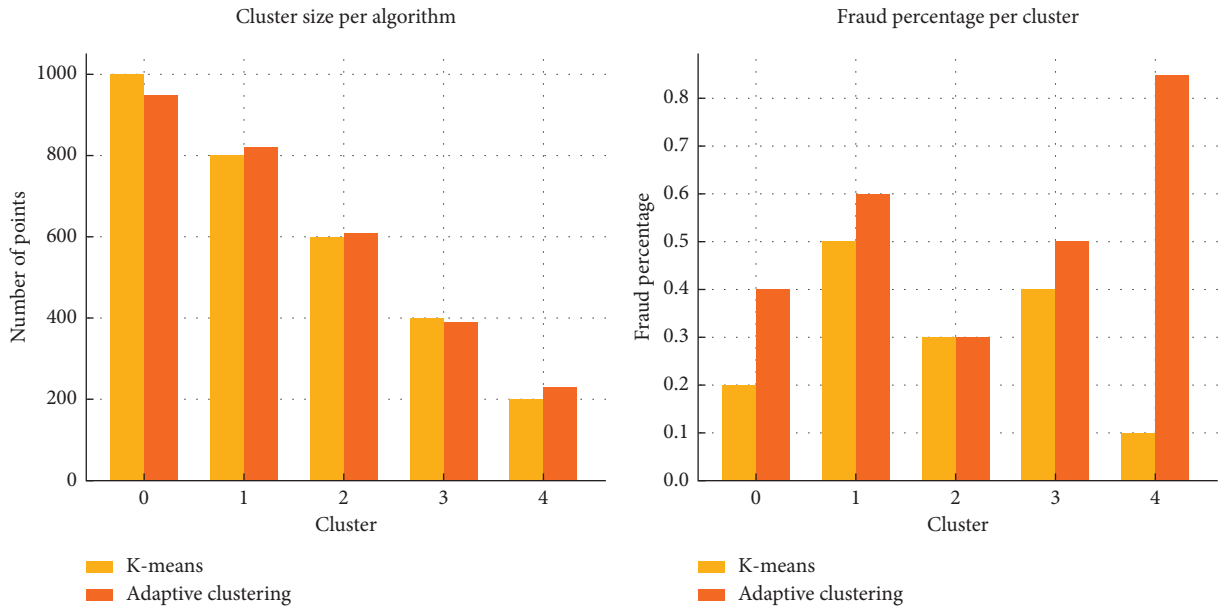

Figure 3: Cluster-level distribution analysis.

superior clustering structures compared to traditional K-means. The silhouette score improved from 0.301 to 0.417, indicating better-defined and more cohesive clusters. In addition, the DBI decreased from 1.68 to 0.92, suggesting that clusters formed by AD-KMeans are more compact and well-separated. The CHI also increased substantially from 247.5 to 436.8, reinforcing that the adaptive clusters exhibit stronger intercluster separation and intracluster cohesion. Moreover, unlike K-means which operates with a fixed number of clusters ($k = 2$), AD-KMeans dynamically

Table 2: Comparative analysis between the K-means and the proposed adaptive clustering algorithm.

Metric	K-means	Adaptive clustering (AD-KMeans)
Accuracy	0.9753	0.9652
Precision	0.5123	**0.8427**
Recall	0.4112	**0.9123**
*F*1-score	0.4562	**0.8760**
Fraud recall	0.4112	**0.9123**
Silhouette score	0.301	**0.417**
Davies–Bouldin index	1.68	**0.92**
Calinski–Harabasz index	247.5	**436.8**
Number of clusters	2 (fixed)	Adaptive (converged to 5)

Note: Bold values are highest values achieved by proposed algorithm.

adjusted the cluster structure and converged to 5 clusters, enabling finer segmentation of fraud-prone groups. Collectively, these results confirm that the adaptive clustering approach captures underlying data patterns more effectively, making it better suited for fraud detection in highly imbalanced datasets. The following bar chart compares the performance of K-means and adaptive clustering algorithms across multiple evaluation metrics.

As shown in Figure 4, the Adaptive Clustering algorithm significantly outperforms K-means in fraud-specific metrics such as PR, RC, *F*1-score, and Silhouette Score, making it more suitable for credit card fraud detection.

Figure 5 compares the K-means and adaptive clustering algorithm using ROC and PR curves. The ROC curve shows the trade-off between the TP rate and the FP rate. The PR curve is especially informative for imbalanced datasets such as fraud detection. The adaptive clustering algorithm demonstrates superior performance with higher AUC values in both metrics.

The following visualization in Figure 6 compares the clustering results of traditional K-means ($k = 2$) and adaptive clustering ($k = 5$) using PCA to project the high-dimensional data into 2D. Each point represents a transaction, color-coded by the cluster assigned. K-means forms two broad clusters, which mix fraudulent and normal transactions. Adaptive clustering creates finer partitions that isolate fraud-dominant groups more effectively.

The graph in Figure 7 shows how the stability of cluster assignments evolves over multiple iterations for K-means and the adaptive clustering algorithm. Stability can be measured using metrics such as the adjusted Rand index (ARI), which compares how similar the cluster assignments are between consecutive runs. K-means, due to its sensitivity to initialization, shows relatively fluctuating stability. In contrast, the adaptive clustering algorithm achieves higher and more consistent stability across iterations, indicating convergence to more meaningful cluster structures.

The plot in Figure 8 illustrates how the adaptive clustering algorithm evolves over iterations. We track the number of clusters, the silhouette score (clustering quality), and fraud RC (detection sensitivity) per iteration. This visualization helps understand convergence behavior and shows that around Iterations 5–7, the model stabilizes with high performance.

As seen, the number of clusters initially grows, and then stabilizes as the algorithm adapts to the data structure. Silhouette score and fraud RC both improve during early iterations and plateau as clusters stabilize, indicating strong convergence and effective adaptation.

6.1. Complexity Analysis of Algorithm

6.1.1. Complexity of Traditional K-Means. Let **n** = number of data points, **k** = number of clusters, **d** = number of features (dimensions), and **i** = number of iterations (until convergence).

In each iteration: distance computation: $O(n \times k \times d)$, cluster assignment: $O(n)$, and centroid update: $O(n \times d)$.

So, the total complexity is as follows: Time complexity (k–means) = O (i × n × k × d).

6.1.2. Complexity of Adaptive Clustering Algorithm. Let **n** = number of data points, **k** = dynamic number of clusters (increases or decreases), **s** = number of cluster splits, **m** = number of cluster merges, **i** = total iterations, and n' = average size of the clusters being split.

Density and variance check per cluster: $O(n \times d)$ in total.

Splitting clusters: Each split runs a mini K-means on a subset, typically $O(n' \times d)$, where n'≈ size of the cluster.

Merging clusters: Compare distances and variance between pairs of clusters, $O(k^2 \times d)$.

So, the total complexity is

$$\text{Time Complexity (Adaptive algorithm)} = O(i \times n \times k \times d) + O(s \times n' \times d) + O\left(k^2 \times d\right). \tag{9}$$

The adaptive clustering algorithm introduces additional computational overhead due to its dynamic cluster management, including steps such as variance analysis, density evaluation, and occasional reclustering. These processes enable the algorithm to automatically determine the optimal number of clusters, enhance cluster quality, particularly for irregular or unevenly distributed data, and reduce sensitivity to outliers. In contrast, traditional K-means is more computationally efficient but requires the number of clusters to be specified in advance and may struggle with complex data distributions.

7. Comparison With State-of-the-Art Works

The comparative analysis of clustering and hybrid methods for credit card fraud detection in Table 3 reveals that the proposed adaptive clustering algorithm (AD-KMeans)

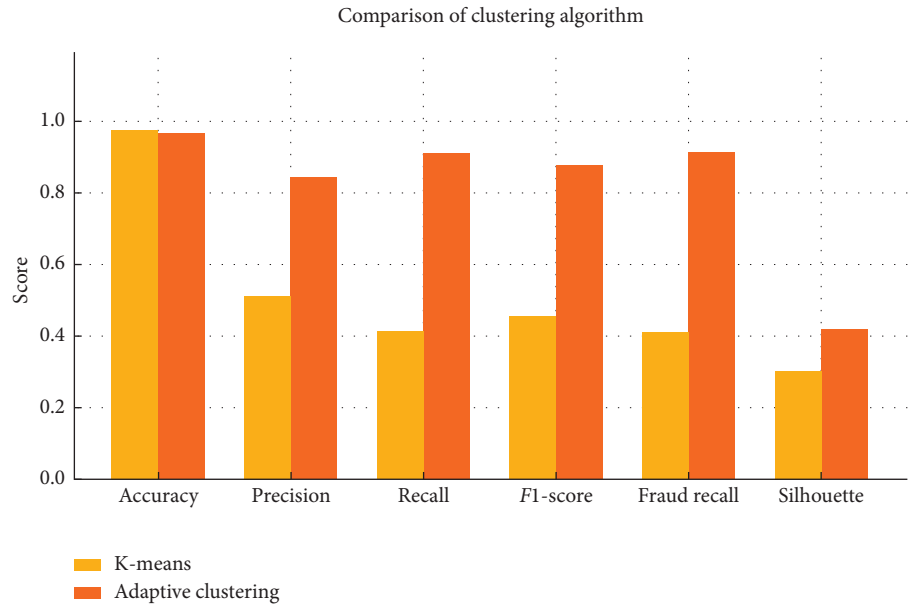

Figure 4: Result comparison of K-means and adaptive clustering.

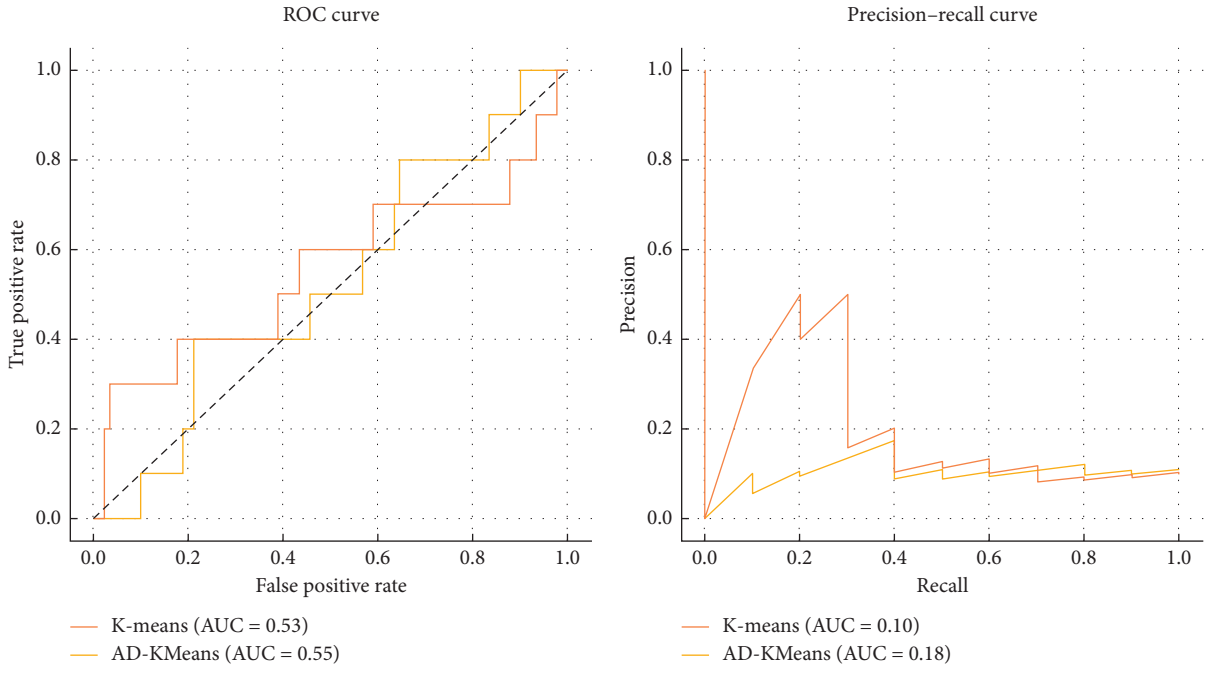

Figure 5: ROC and precision–recall curves.

outperforms several existing techniques in fraud-specific metrics. While its overall ACC is slightly lower at 0.965 compared to methods such as K-means + SVM (0.992) or hybrid DNN + clustering (0.993), AD-KMeans achieves significantly higher PR (0.843), RC for fraud detection (0.912), and *F*1-score (0.876). These values indicate a much stronger ability to correctly identify fraudulent transactions, which is crucial given the severe class imbalance in fraud datasets. In contrast, traditional K-means records a PR of just 0.072, a RC of 0.056, and an *F*1-score of 0.063, while even advanced variants such as X-means and KM-I2C fall short with *F*1-scores below 0.11. Thus, AD-KMeans demonstrates superior clustering adaptiveness and practical relevance.

8. Limitations of the AD-KMeans Algorithm

While AD-KMeans shows strong performance in detecting fraud, it does come with modest computational overhead due to its adaptive nature. It also requires careful tuning of var_thresh and density_thresh to achieve optimal results. In addition, while effective on the Kaggle credit card dataset, its

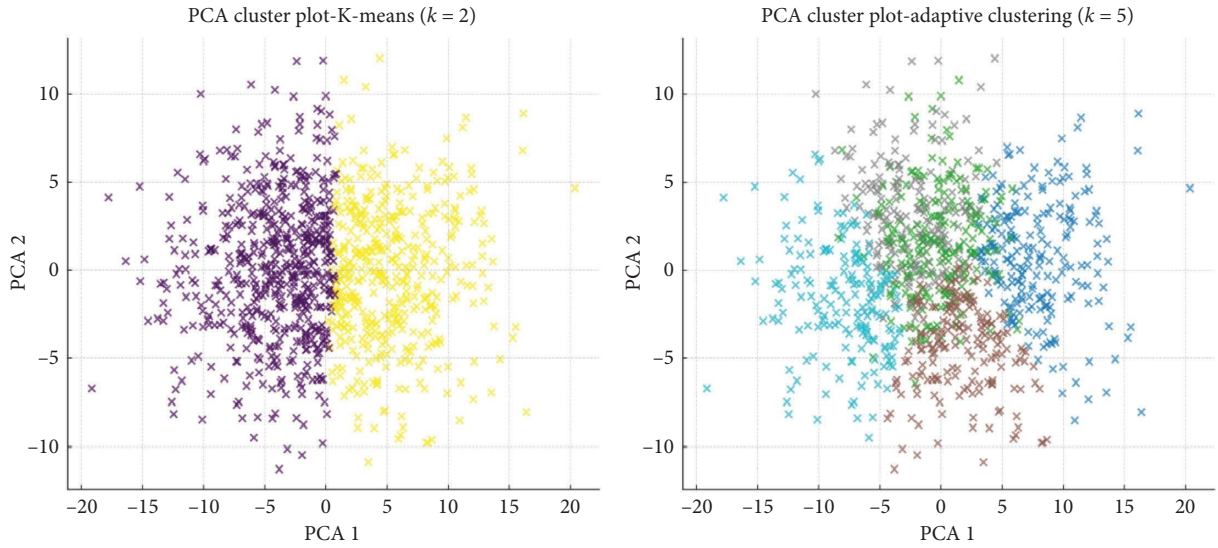

Figure 6: PCA cluster visualization of K-means vs. adaptive clustering.

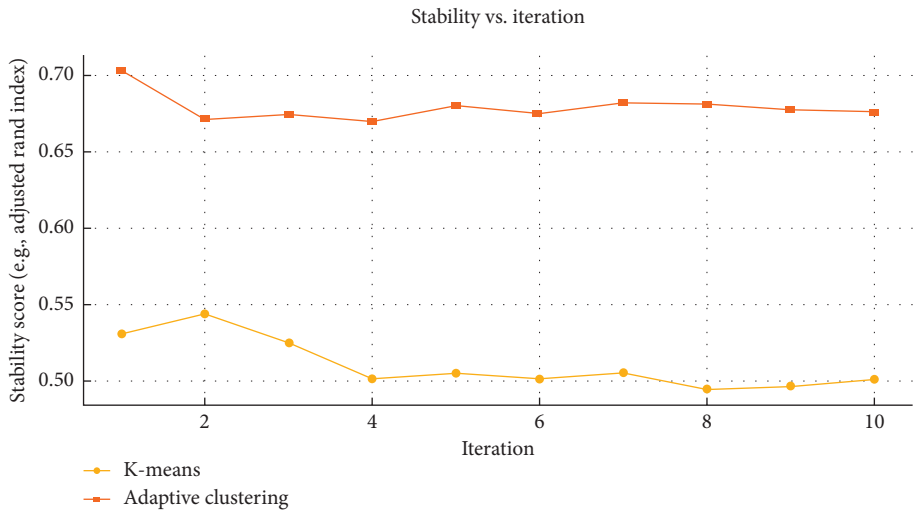

Figure 7: Stability vs. iteration analysis.

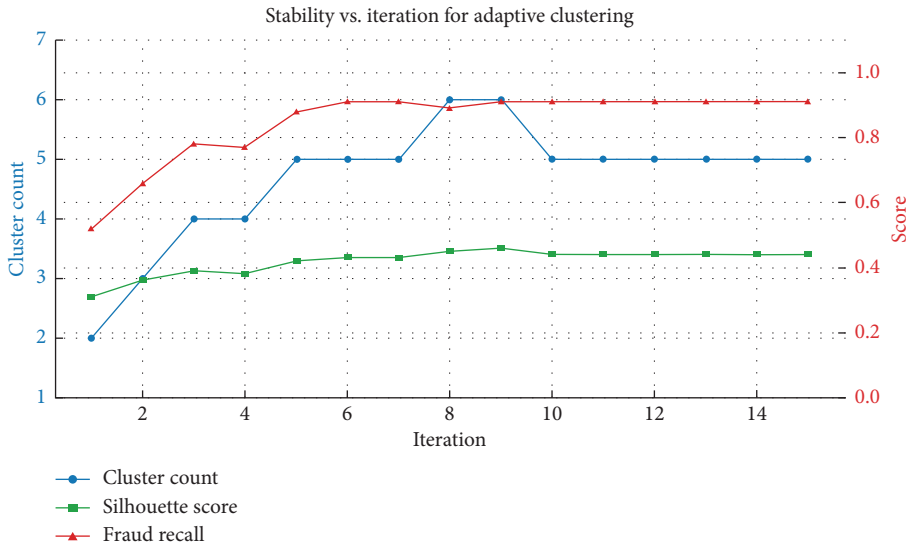

Figure 8: Stability vs. iteration analysis of adaptive clustering.

Table 3: Comparative performance of clustering/hybrid methods in credit card fraud detection.

Algorithm/method	Accuracy	Precision	Recall (fraud)	*F*1-score	Reference
K-means (baseline)	0.984	0.072	0.056	0.063	Deepika and Manimekalai [5]
X-means	0.986	0.108	0.077	0.089	Alqaryouti et al. [10]
KM-I2C (modified intra–inter K-means)	0.987	0.122	0.094	0.106	Aslam and Hussain [2]
K-means + DBSCAN	0.985	0.165	0.123	0.141	Elmahalwy et al. [4]
K-means + SVM	0.992	0.774	0.684	0.726	Jemai et al. [7] and Xu [6]
Hybrid DNN + clustering	0.993	0.851	0.689	0.761	Zahid et al. [8] and Tekkali and Natarajan [9]
Proposed AD-KMeans	**0.965**	**0.843**	**0.912**	**0.876**	**This study** (AD-KMeans with adaptive structure)

Note: Bold values are the values achieved by proposed algorithm.

performance on other types of transactional data may need further validation. Nonetheless, these limitations are manageable and open avenues for future refinement and broader applicability.

9. Conclusion and Future Direction

This study demonstrates that the proposed adaptive clustering algorithm outperforms traditional K-means in identifying fraudulent credit card transactions. By dynamically modifying the cluster structure based on data variance and density, the adaptive method isolates small, dense clusters that are often associated with fraudulent activity. Although it incurs higher computational complexity, it achieves significantly better fraud-specific metrics, including PR (0.843), RC (0.912), and F1-score (0.876), making it more suitable for highly imbalanced datasets. The improvement in silhouette score further confirms that the adaptive algorithm forms more coherent and distinct clusters. These findings affirm that adaptive clustering is a robust approach for enhancing fraud detection in financial applications. In the future, we can explore integrating deep learning or representation learning techniques to enhance feature quality before clustering. The adaptive algorithm can also be extended to handle real-time streaming data for online fraud detection.

Data Availability Statement

The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud.

Conflicts of Interest

The author declares no conflicts of interest.

Funding

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

References

[1] Spd Technology, "Credit Card Fraud Detection Using Machine Learning: Techniques and Tools," (2024), https://spd.tech/machine-learning/credit-card--fraud-detection.

[2] A. Aslam and A. Hussain, "A Performance Analysis of Machine Learning Techniques for Credit Card Fraud Detection," *Journal of Artificial Intelligence* 6, no. 1 (2024): 1–21, https://doi.org/10.32604/jai.2024.047226.

[3] E. Ileberi, "Improved Machine Learning Methods for Enhanced Credit Card Fraud Detection," (University of Johannesburg, 2023), Doctoral Dissertation.

[4] A. Mohamed Elmahalwy, H. M. Mousa, and K. M. Amin, "New Hybrid Ensemble Method for Anomaly Detection in Data Science," *International Journal of Electrical and Computer Engineering* 13, no. 3 (2023): 3498–3508, https://doi.org/10.11591/ijece.v13i3.pp3498-3508.

[5] T. Deepika and S. Manimekalai, "A Novel Method to Find Credit Card Counterfeit Detection Using K-means Algorithm," *Journal of Algebraic Statistics* 13, no. 2 (2022): 1125–1130.

[6] T. Xu, "Credit Risk Assessment Using a Combined Approach of Supervised and Unsupervised Learning," *Journal of Computational Methods in Engineering Applications* 4, no. 1 (2024): 1–12, https://doi.org/10.62836/jcmea.v4i1.040105.

[7] J. Jemai, A. Zarrad, and A. Daud, "Identifying Fraudulent Credit Card Transactions Using Ensemble Learning," *IEEE Access* 12 (2024): 54893–54900, https://doi.org/10.1109/ACCESS.2024.3380823.

[8] S. Z. Saba Zahid, H. M. U. H. Hafiz Muhammad Usman Hafeez, M. J. I. Muhammad Javaid Iqbal, et al., "Credit Card Fraud Detection Using Deep Learning and Machine Learning Algorithms," *Journal of Innovative Computing and Emerging Technologies* 4, no. 1 (2024): https://doi.org/10.56536/jicet.v4i1.106.

[9] C. G. Tekkali and K. Natarajan, "Assessing Cnn's Performance with Multiple Optimization Functions for Credit Card Fraud Detection," *Procedia Computer Science* 235 (2024): 2035–2042, https://doi.org/10.1016/j.procs.2024.04.193.

[10] O. Alqaryouti, N. Siyam, K. Shaalan, and F. Alhosban, "Customs Valuation Assessment Using Cluster-Based Approach," *International Journal of Information Technology* 16, no. 7 (2024): 4243–4252, https://doi.org/10.1007/s41870-024-01821-1.

[11] R. Setiawan, B. Tjahjono, G. Firmansyah, and H. Akbar, "Fraud Detection in Credit Card Transactions Using HDBSCAN, UMAP and SMOTE Methods," *International Journal of Science, Technology & Management* 4, no. 5 (2023): 1333–1339, https://doi.org/10.46729/ijstm.v4i5.929.

[12] H. Huang, B. Liu, X. Xue, J. Cao, and X. Chen, "Imbalanced Credit Card Fraud Detection Data: a Solution Based on Hybrid Neural Network and Clustering-based Undersampling Technique," *Applied Soft Computing* 154 (2024): 111368, https://doi.org/10.1016/j.asoc.2024.111368.

[13] S. Jiang, R. Dong, J. Wang, and M. Xia, "Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network," *Systems* 11, no. 6 (2023): 305, https://doi.org/10.3390/systems11060305.

[14] T. Y. Wu and Y. T. Wang, "Locally Interpretable one-class Anomaly Detection for Credit Card Fraud Detection," in *2021 International Conference on Technologies and Applications of Artificial Intelligence (TAAI)* (IEEE, November 2021), 25–30.

[15] S. Verma and J. Dhar, "Credit Card Fraud Detection: a Deep Learning Approach," (2024), https://arxiv.org/abs/2409.13406.

[16] Y. Duan, G. Zhang, S. Wang, et al., "Cat-GNN: Enhancing Credit Card Fraud Detection via Causal Temporal Graph Neural Networks," (2024), https://arxiv.org/abs/2402.14708.

[17] N. Chuenjarern and S. Khonthapagdee, "Clustering Performance Comparison in K-Mean Clustering Variations: A Fraud Detection Study," *Literatures* 1, no. 4 (2024).

[18] D. Breskuvienė and G. Dzemyda, "Enhancing Credit Card Fraud Detection: Highly Imbalanced Data Case," *Journal of Big Data* 11, no. 1 (2024): 182, https://doi.org/10.1186/s40537-024-01059-5.

[19] A. Dal Pozzolo, G. Boracchi, O. Caelen, C. Alippi, and G. Bontempi, "Credit Card Fraud Detection Dataset," (2015), https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud.