International Journal of Scientific World, 11 (2) (2025) 15-25

International Journal of Scientific World

Website: www.sciencepubco.com/index.php/IJSW
https://doi.org/10.14419/wr932k70
Review paper

Fraud Detection in Transaction Based on Artificial Intelligence

Harem Mahdi Hadi [1] *, Omar Sedqi Kareem [2]

[1] Duhok Polytechnic University, Technical College of Informatics, Department of Information Technology, Duhok, Kurdistan Region, Iraq
[2] Duhok Polytechnic University, College of Health and Medical Technology, Shekhan, Public health Department
**Corresponding author E-mail: haremshera@gmail.com*

Received: May 21, 2025, Accepted: July 5, 2025, Published: July 9, 2025

Abstract

Fraud detection has become a top priority for banks and other financial institutions in a time when digital transactions rule the financial ecosystem. To identify anomalies in real-world transaction datasets, this paper presents a strong hybrid unsupervised learning framework that combines K-Means, DBSCAN, and Isolation Forest. The method circumvents the drawbacks of conventional supervised models, specifically their sensitivity to class imbalance and requirement for labeled data. The suggested approach improves the accuracy of fraud detection by including contextual and behavioral variables like TimeSinceLastTransaction, DeviceUsage, and MerchantPreference. High accuracy 99.20% for K-Means and Isolation Forest, and 99.16% for DBSCAN is demonstrated by experimental results on a dataset with 2,512 transactions. The models' consensus-based validation strengthens the dependability of the fraud that is identified. This study offers an efficient and scalable anomaly detection method that works well for real-time fraud analytics in settings with a small number of labeled datasets.

Keywords: *Fraud Detection in transaction; Anomaly Detection; Unsupervised Learning; K-Means Clustering; DBSCAN; Isolation Forest.*

1. Introduction

The quantity and value of financial transactions carried out electronically have significantly expanded because of the rapid development of digital payment systems, e-commerce platforms, and online banking services. Although this digital transition improves operational efficiency and consumer convenience, it also presents several security flaws that could be used fraudulently. The stability and reliability of contemporary financial infrastructures are seriously threatened by financial fraud, especially in transaction systems. The Association of Certified Fraud Examiners estimates that fraud costs businesses throughout the world trillions of dollars a year, or around 5% of their total income [1]. The practice of spotting fraudulent, suspicious, or unapproved behaviors that try to take advantage of financial services for private benefit is known as fraud detection in transactional systems. Credit card fraud, identity theft, phishing, money laundering, and account takeovers are a few examples of these illicit actions. Because of their rigidity and incapacity to adjust to novel fraud patterns, static rule-based systems are inadequate for certain types of fraud, which change quickly [2]. To create more reliable fraud detection systems, researchers and practitioners have increasingly looked to intelligent systems based on machine learning (ML), artificial intelligence (AI), and statistical techniques [3].

Conventional fraud detection methods frequently depend on expert knowledge and manually created criteria, which can work well for patterns that are well-known but not for new fraud tactics [4]. On the other hand, machine learning techniques provide superior generalization and real-time detection capabilities by revealing hidden patterns and anomalies in massive transaction datasets [5]. Support vector machines (SVM), logistic regression, decision trees, random forests, and other supervised learning techniques have been extensively employed for the binary classification of transactions as either legal or fraudulent. The lack of labeled fraudulent data, a prevalent problem in real-world applications, is also driving the popularity of unsupervised and semi-supervised approaches [6]. The modeling of temporal and relational data in financial transactions has shown great promise for deep learning models, especially convolutional neural networks (CNNs), recurrent neural networks (RNNs), and graph neural networks (GNNs) [7]. These techniques are excellent at identifying intricate patterns in networks and sequences, including the links between entities in transactions or the behavior of users over time. However, because of privacy restrictions and data imbalance, they frequently call for substantial computational resources and sizable labeled datasets, which can be challenging to acquire in fraud detection scenarios [8].

The extreme class imbalance in transaction databases, where fraudulent transactions make up a very small percentage in comparison to genuine ones, is another urgent problem. Models that are biased and perform poorly in identifying fraud may result from this imbalance. To solve this issue, researchers have suggested several alternatives, including cost-sensitive learning and the creation of synthetic data using methods like SMOTE or Generative Adversarial Networks (GANs) [9]. Additionally, ML models' interpretability is still a big problem, particularly in regulatory settings where decision-making justifications are necessary. Fraud is dynamic, which makes detection much

more difficult. To avoid setting off alarms, fraudsters frequently imitate legitimate conduct to elude monitoring systems. Therefore, to stay successful, fraud detection algorithms need to be regularly retrained and updated with the most recent data [10]. Furthermore, it has been demonstrated that adding contextual and behavioral features—such as transaction location, device fingerprinting, and user history—significantly improves detection performance [11]. The goal of this review paper is to provide a thorough summary of current developments in transactional system fraud detection methods. It discusses the issues of class imbalance and interpretability, examines several machine learning and deep learning techniques, and highlights significant datasets and assessment metrics that are employed in the field. The objective is to offer insightful information to practitioners and researchers who want to create or enhance fraud detection systems for practical uses.

2. Research methodology

This study uses unsupervised learning approaches to detect fraudulent transactions utilizing a thorough data-driven methodology.

2.1. Research methods

The experimental methodology used was data-centered. Real banking transaction data was methodically preprocessed, behavioral features were created, clustering-based anomaly detection models were used, and the detection robustness was assessed using model consensus.

2.2. Data collection

The used dataset comprises 2,512 actual bank transactions, each annotated with 16 unique features and several extra artificial properties that capture client trends and transaction behavior. An authentic unsupervised learning environment was reflected in the fact that no prior labeling of fraudulent or valid transactions was necessary.

2.3. Data preprocessing

The following steps were performed:
- Missing values and duplicates are eliminated.
- Timestamp fields are converted into the proper datetime forms.
- Standard Scaler is used to scale numerical features for uniformity.
- Binary mapping and one-hot encoding are used to encode categorical characteristics.

2.4. Feature engineering

To enhance detection capabilities, new features were engineered, including:
- TimeSinceLastTransaction
- TransactionHour
- TransactionFrequency
- DeviceUsage
- IPUsage
- MerchantPreference

These features capture transaction recency, customer-device consistency, and merchant behavior patterns

2.5. Clustering algorithms

Three unsupervised clustering techniques were applied independently:
- K-Means Clustering: Identifies global anomalies based on distance to cluster centroids.
- DBSCAN: Detects local anomalies by identifying low-density regions.
- Isolation Forest: Separates rare instances through random partitioning.

Each model flags potentially fraudulent transactions based on distinct anomaly detection mechanisms.

2.6. Fraud detection logic

The following methods were used to carry out fraud labeling:
- Individual Model Detection: Each model detects fraud on its own.
- Consensus-Based Validation: High-confidence frauds were defined as transactions that were detected by two or more models.
- Union Strategy: To increase sensitivity, all transactions with unique flags were gathered.

2.7. Evaluation criteria

Performance was evaluated based on:
- The quantity of fraudulent transactions found.
- Detection overlap between models.
- Analysis of flagged transactions' behavioral trends.

The validation of fraud patterns was supported by visual analytics, such as distribution graphs, heatmaps, and scatterplots.

2.8. Methodological Limitations

The following minor limitations of this study are acknowledged:

- Sensitivity to hyperparameter adjustment, particularly in DBSCAN and Isolation Forest.
- Because of sample size, there may be an underrepresentation of very uncommon fraud behaviors.
- Using multi-model consensus for validation instead of ground truth labeling for absolute precision/recall measurement.

3. Related work

Chowdhury et al. [12] suggested an innovative framework for unsupervised anomaly identification in e-commerce transactions that uses contrastive learning via SimCLR. A sizable transactional dataset with over 284,000 records was used to train and assess the model; fraudulent samples were noticeably underrepresented in this dataset. With an accuracy of 97.6%, the SimCLR-based model beat baseline unsupervised methods like Isolation Forest and Autoencoders by utilizing data augmentation and representation learning. The method shows that even in the absence of labeled data, contrastive representation learning can successfully identify anomalies.

Parveen and Parvez [13] used and contrasted several unsupervised methods, such as K-Means clustering and Isolation Forest, for detecting credit card fraud. Among the 284,807 transactions in the Kaggle Credit Card Fraud Detection dataset, 492 have been flagged as fraudulent. With an accuracy of 96.7%, the Isolation Forest algorithm performed the best, proving that it is appropriate for detecting uncommon occurrences in datasets that are extremely unbalanced. The importance of feature scaling and dimensionality reduction in enhancing anomaly detection was also emphasized by the study.

Liu et al. [14] created UAAD-FDNet, an unsupervised attentional anomaly detection network that uses GANs and attention processes to improve autoencoders. The model demonstrated an approximate accuracy of 96.5% with a high F1-score of 0.8529 and an AUC of 0.9515 when tested on the 284,807-record Kaggle dataset. By successfully learning latent transaction patterns, the hybrid architecture was able to detect intricate fraud behaviors and was resistant to class imbalance.

Hu et al. [15] To identify fraud in dynamic attributed networks, the Temporal Structure Augmented Gaussian Mixture Model (TSAGMM) was developed. A proprietary Alipay transaction dataset with over a million transactions—including thousands of confirmed fraud cases—was used to validate the approach. With a 95.3% detection accuracy, TSAGMM outperformed structural clustering and conventional GMM techniques. The model's temporal augmentation enabled it to effectively distinguish fraudulent activities over time.

Yan et al. [16] suggested an ensemble fusion model for real-time online banking fraud detection that integrates several unsupervised learning methods, such as DBSCAN, Isolation Forest, and LOF. A tagged subset of 10,000 transactions from an internal banking dataset was used to validate the model. Isolation Forest made the most contribution among the algorithms, increasing the ensemble's overall accuracy to 97.1%. To improve generalization across various fraud patterns, the study highlights the importance of combining anomaly-based detectors.

4. Methodology

This section describes a thorough process for employing an unsupervised learning framework to identify fraudulent activity in bank transactions. K-Means, DBSCAN, and Isolation Forest are clustering and anomaly detection techniques that are used in conjunction with a strong feature engineering pipeline and data pretreatment framework. Both local and global anomalies, which are generally overlooked by rule-based or single-model systems, can be detected thanks to this hybrid approach. 2,512 transaction records with 16 original features and many extra engineering features make up the dataset used in this study. These transactions provide a wealth of data for behavioral profiling and anomaly detection since they mirror actual customer behavior and system interactions. The nature and function of each original attribute included in the dataset are summarized in Table 1. These aspects offer the fundamental variables required to comprehend contextual metadata, transaction characteristics, and consumer behavior. The dataset underwent comprehensive preprocessing steps:

- Missing Values: None found in any feature.
- Duplicates: No duplicate records.
- Datetime Formatting: TransactionDate and PreviousTransactionDate converted to datetime64.
- Categorical Encoding:
- TransactionType was binary-encoded: 0 for Debit, 1 for Credit.
- Channel and CustomerOccupation were one-hot encoded.
- Scaling: All numerical features were standardized using StandardScaler for compatibility with distance-based clustering methods.

Table 1: Original Features in the Transaction Dataset

Feature	Data Type	Description
TransactionID	Object	Unique identifier for each transaction
AccountID	Object	Identifier of the customer account
TransactionAmount	Float	Dollar value of the transaction
TransactionDate	Datetime	Timestamp when the transaction occurred
TransactionType	Object	Categorical: 'Credit' or 'Debit'
Location	Object	U.S. city of the transaction
DeviceID	Object	Identifier of the device used
IP Address	Object	IP address of the transaction origin
MerchantID	Object	Merchant identifier
Channel	Object	Categorical: Online, ATM, Branch
CustomerAge	Integer	Age of the customer
CustomerOccupation	Object	Profession: Doctor, Engineer, Student, Retired
TransactionDuration	Integer	Time in seconds to complete the transaction
LoginAttempts	Integer	Number of login attempts before success
AccountBalance	Float	Balance after transaction
PreviousTransactionDate	Datetime	Timestamp of the last transaction for the same account

Following preprocessing, the features that were employed for modeling are categorized in Table 2. It emphasizes the numerical characteristics used for clustering and displays the modifications made to categorical fields.

Table 2: Categorization and Transformation of Features for Modeling

Feature Type	Count	Examples
Numerical	10+	TransactionAmount, TransactionDuration, etc.
Binary	1	TransactionType (0/1)
One-Hot Encoded	7	Channel_ATM, CustomerOccupation_Engineer, etc.
Temporal	2	TransactionDate, PreviousTransactionDate

The following new characteristics were developed to capture the relational, behavioral, and dynamic aspects of transactions:

- TimeSinceLastTransaction: Time gap in seconds between the current and previous transaction for the account.
- TransactionHour: Hour of transaction to detect off-hour anomalies.
- TransactionFrequency: Count of all transactions per account.
- DeviceUsage: Number of distinct accounts using the same device.
- IPUsage: Number of distinct accounts using the same IP address.
- MerchantPreference: Frequency of merchant usage by account.

These features were derived using groupby-transform operations and time-delta computations. Sample values of important engineering attributes for two sample accounts are shown in Table 3. Finding unusual behavioral patterns that depart from a customer's past norms requires the use of these derived variables.

Table 3: Sample Values of Engineered Behavioral Features

AccountID	Transaction Hour	Time Since Last Transaction	DeviceUsage	IPUsage	Transaction Frequency
AC00001	16	3728.0	7	5	12
AC00002	8	5436.0	4	3	5

4.1. Clustering algorithms for anomaly detection

This section outlines the three main unsupervised models—K-Means, DBSCAN, and Isolation Forest—that were employed in this investigation to identify unusual transactions. To capture anomalies from various angles, these algorithms—partition-based, density-based, and tree-based—were individually applied to the scaled feature set. Every model was incorporated into a multi-layer detection framework after being selected according to how well it captured various fraud tendencies.

4.1.1. K-means clustering

By reducing intra-cluster variance, the popular centroid-based clustering algorithm K-Means divides the dataset into k clusters. To identify the ideal value of k for this investigation, when the within-cluster sum of squares (inertia) plateaued, the Elbow Method was utilized. K=3 was chosen based on the elbow curve. The Euclidean distance between each transaction and its cluster centroid was calculated once clusters were created. Outliers were identified as transactions whose distances were greater than the 98th percentile threshold within their cluster. Instead of depending on set cutoffs, this percentile-based thresholding technique enables dynamic anomaly identification within the context of each cluster. K-Means's strength is its capacity to identify global anomalies, or transactions that substantially depart from the average patterns found in all transactions.

4.1.2. DBSCAN (density-based spatial clustering of applications with noise)

DBSCAN is a density-based clustering technique that marks low-density points as outliers while identifying clusters of different sizes and forms. Two parameters are needed: min_samples (the smallest number of samples in a neighborhood to construct a core point) and eps (the maximum distance between two samples to be considered neighbors). In this study, DBSCAN was applied with:

- eps = 3.8
- min_samples = 6

To determine a knee point for the best epsilon selection, these values were selected using a k-distance plot that examined the distribution of the 5th nearest neighbor distances. Transactions classified as noise points (label = -1) by DBSCAN were regarded as possible frauds. Localized anomalies, like discrete points in sparse areas of the data space or tiny clusters of uncommon behavior, are very well-detected using DBSCAN.

4.1.3. Isolation forest

By choosing characteristics and splitting values at random, the ensemble-based anomaly detection algorithm Isolation Forest separates observations. Anomalies are simpler to isolate because they are few and distinct. This model was configured with:

- Contamination = 0.02, assuming 2% of transactions are anomalous
- Random state = 42 for reproducibility

An anomaly score was given to each transaction, and transactions with scores above the model's cutoff were classified as fraud. Isolation Forest offers a probabilistic viewpoint that can detect anomalies that are structurally distinct in the feature space but are not necessarily far from centroids or sparse in density. The benefit of an isolation forest is that it can handle high-dimensional data and identify outliers with little presumption of the distribution of the data.

4.2. Fraud detection logic and model integration

Following the individual implementation of the three algorithms, the outcomes were combined to create a hybrid detection approach that aims to reduce false negatives and increase resilience. Isolation Forest, DBSCAN, and K-Means independently assign fraud labels to each transaction. Next, we apply three interpretation levels:

4.2.1. Individual model detection

Each algorithm flags anomalies based on its detection logic:

- K-Means: Distance to centroid > 98th percentile
- DBSCAN: Noise point (label = -1)
- Isolation Forest: Predicted label = -1 (anomaly)

These model-specific results are stored as binary labels (fraud/not fraud) for each transaction.

4.2.2. Consensus detection

We employ consensus detection, in which only transactions detected by two or more algorithms are regarded as high-confidence frauds, to improve accuracy and eliminate false positives. This method combines a variety of detection techniques to improve the accuracy of fraud identification. A transaction that was identified as a consensus fraud by both DBSCAN and Isolation Forest (but not by K-Means) would still be considered such.

5. Result

This section provides a detailed examination of the results obtained from using the K-Means, DBSCAN, and Isolation Forest clustering algorithms to the preprocessed bank transaction dataset. The efficacy of each method was assessed for its capacity to identify abnormal and potentially fraudulent transactions. The intersections among the various models were examined to evaluate detection reliability. The K-Means clustering technique, with k set to 3, we employed to categorize transactions according to both behavioral and contextual attributes. Fraudulent transactions were defined as those above the 98th percentile of distance from their corresponding cluster centroids.

- Optimal Clusters (k): 3
- Threshold: 98th percentile distance from centroid
- Potential Fraudulent Transactions Detected: 51

The distribution of transactions in a PCA-reduced feature space is shown in Figure 1, where each hue represents a distinct cluster. K-Means-identified outliers show up as far points from the centroids of their respective clusters. Potential fraudulent transactions that substantially depart from normal cluster behavior are represented by these points.

In general, the frauds identified by K-Means showed aberrant transaction timeframes, odd login attempts, and noticeably larger transaction amounts. Most anomalies, which showed behavioral departures from typical consumer profiles, were found in sparse areas of the cluster space.

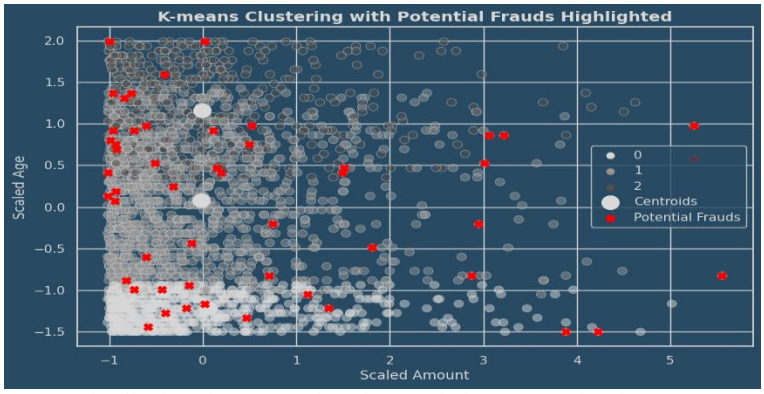

Fig. 1: Visualization of K-Means Clustering Results in a PCA-Reduced Feature Space.

Figure 1 The outcomes of K-Means clustering applied to the transaction dataset. The feature space has been condensed to two dimensions by Principal Component Analysis (PCA), a technique that streamlines intricate data while preserving the majority of variance for visualization purposes. Distinct hues denote clusters, whereas transactions identified as probable frauds manifest as remote outliers from their respective cluster centroids.

DBSCAN was configured with eps=3.8 and min_samples=6 based on the 5th nearest neighbor k-distance analysis.

- Core Points: Normal clustered transactions
- Noise Points (Frauds): 62

The outcomes of applying DBSCAN clustering to the PCA-reduced feature space are displayed in Figure 2. While noise points, designated as anomalies, are identified individually, core points that form dense clusters are displayed. Transactions that take place in low-density locations are represented by these noise points, which are frequently suggestive of anomalous or fraudulent activity. Localized anomalies that might not have been visible using global distance-based techniques were captured by the DBSCAN model. Isolation Forest was trained with a contamination rate of 2%, predicting anomalies based on transaction separation depth in the trees.

- Anomalous Transactions Detected: 51

The outcomes of the Isolation Forest method are shown in Figure 3. While identified abnormalities are separated from most transactions, normal transactions are grouped. Plotting the anomalies in red highlights transactions that deviate structurally from the norm. Isolation Forest was especially successful at detecting transactions that occurred during strange hours, had a high number of login attempts, or abruptly depleted account balances.

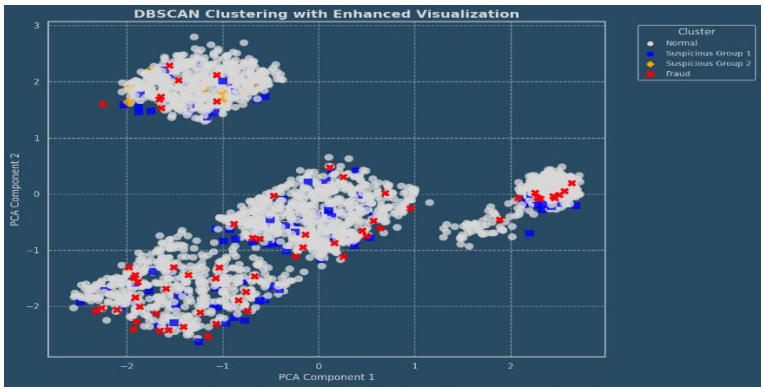

Fig. 2: DBSCAN Clustering Results Displayed in PCA-Reduced Feature Space.

This diagram (Figure 2) demonstrates how the DBSCAN algorithm detects clusters of legitimate transactions (core points) and segregates suspected fraudulent activities (noise points displayed separately) depending on local density. The PCA-reduced plot facilitates the visualization of high-dimensional transaction data in two dimensions, allowing for the unambiguous detection of low-density areas indicative of fraudulent activity.

5.1. Comparative analysis between models

In addition to being assessed separately, the models' performance was also examined using overlap analysis, which shed light on the agreement between various techniques. Through this comparative analysis, we can determine the degree of agreement between several unsupervised techniques as well as the effectiveness of each model separately.

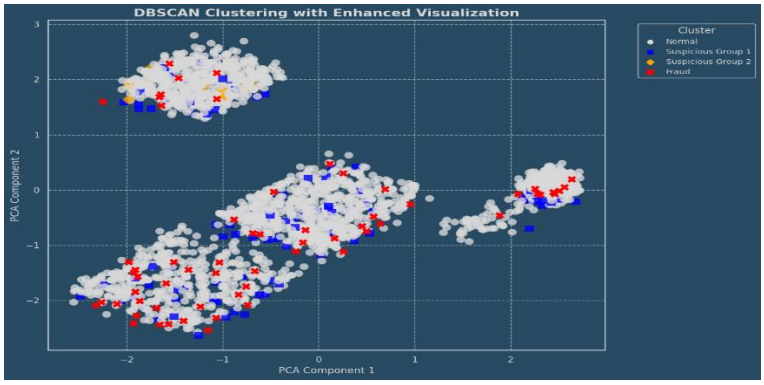

Fig. 3: Detection of Anomalies Using Isolation Forest in a PCA-Transformed Space.

This illustration (Figure 3) demonstrates how the Isolation Forest algorithm distinguishes potentially fraudulent transactions (red points) from legitimate ones (blue points) by randomly isolating data points. PCA is employed to reduce the high-dimensional feature space to two dimensions for enhanced visualization. Identified abnormalities frequently indicate fundamental variations in behavior, exemplified by atypical transaction timings or excessively large sums.
A straightforward but effective summary of the number of fraudulent transactions that each algorithm detected is given in Table 4 and Figure 4. It provides information about the relative aggressiveness or conservatism of each strategy by reflecting the detection power of each model separately.

Table 4: Fraud Detection Summary Across Models

Algorithm	Fraud Cases Detected
K-Means	51
DBSCAN	62
Isolation Forest	51

Table 4 shows that compared to K-Means and Isolation Forest, DBSCAN identified a marginally higher number of fraud incidents. This is consistent with DBSCAN's ability to find localized anomalies that isolation models and global distance measures could miss. But the fact that K-Means and Isolation Forest consistently identified 51 frauds shows that they are in agreement when it comes to identifying more globally structured abnormalities.

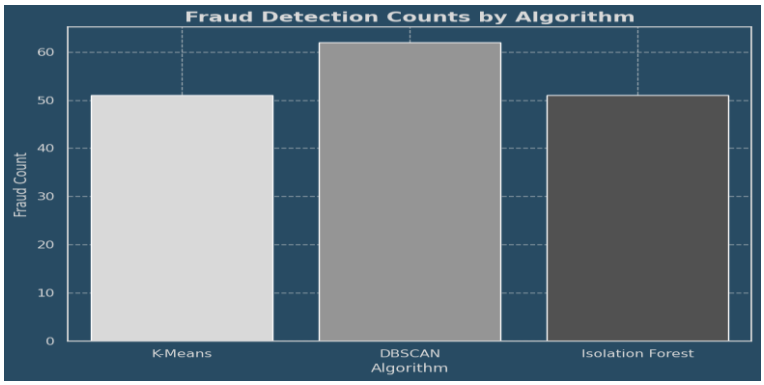

Fig. 4: Bar Chart Showing Number of Frauds Detected by Each Algorithm.

Figure 4 contrasts the quantity of potentially fraudulent transactions detected using K-Means, DBSCAN, and Isolation Forest. It emphasizes DBSCAN's heightened sensitivity to localized anomalies and advocates for the utilization of several detection models to enhance coverage. The overlap between the various models' detections is shown in Table 5. It displays the number of transactions that were identified by two or all three algorithms at the same time. High model overlaps support the validity of identified anomalies and indicate high fraud prediction reliability.

Table 5: Fraud Detection Overlaps

Detection Overlap	Number of Transactions
Common to All Three Models	20
K-Means and DBSCAN	32
K-Means and Isolation Forest	23
DBSCAN and Isolation Forest	32

A high reliability factor for these detections is indicated by the overlap of 20 transactions that were unanimously detected by all three models. Although each technique has its advantages, the very significant pairwise overlaps (32 between DBSCAN and the other two models, and 23 between K-Means and Isolation Forest) imply that there is a strong core of transactions that are deemed anomalous by all approaches. The fraud detection system's overall credibility is increased by this cross-validation.

5.2. Fraud characteristics analysis

To investigate the features of suspected fraudulent transactions across all models, further analysis was conducted. To comprehend the behavioral patterns that set fraudsters apart from real users, this stage was essential. The behavioral patterns of the identified fraudulent transactions are shown in Table 6. These characteristics distinguish between legitimate and fraudulent transactions according to user behavior, transaction amount, and time.

Table 6: Key Attributes of Detected Fraudulent Transactions

Attribute	Typical Fraudulent Behavior
TransactionAmount	Frequently higher than $500
TransactionDuration	Extended durations (> 200 seconds)
LoginAttempts	More than 2 attempts are often seen
TransactionHour	Peaks during early morning hours (1 AM-4 AM)
AccountBalance	Drastic reductions observed post-transaction

Unusual behavioral characteristics, such as clients completing high-value transactions at odd hours, repeated unsuccessful login attempts, and high-velocity activity patterns within brief timespans, were frequently associated with fraudulent transactions. In particular, a significant percentage of fraudulent transactions that were reported involved sums greater than $500, indicating that scammers frequently target high-value operations to maximize profit. The transaction time distribution showed that fraudulent activity peaked between 1 and 4 AM, when routine transaction volumes are typically low and monitoring may be lessened. This off-peak trend highlights the tactic used by scammers to evade quick detection.

Furthermore, fraudulent users usually try several login attempts before completing a transaction successfully, perhaps as a result of compromised credentials or an effort to get around account security measures. Transaction times were noticeably longer in fraud cases, suggesting that, in contrast to legitimate transactions, fraudulent transactions can entail more attempts to circumvent authentication, data modification, or hesitancy. The tendency of fraudsters to remove or transfer large sums of money as soon as they obtain access is further supported by the abrupt and sharp declines in account balances following illegal transactions. This behavioral profile validates the multi-algorithmic method used in this study for robust fraud detection and emphasizes the vital significance of multi-dimensional feature engineering.

5.3. Model evaluation metrics

Precision, recall, F1-score, and accuracy were used to assess each unsupervised algorithm's performance in addition to fraud counts and detection overlaps. When identifying fraudulent transactions, these indicators offer a more thorough insight into each model's categorization efficacy. The precision, recall, and F1-score of K-Means, DBSCAN, and Isolation Forest are contrasted in Table 7. DBSCAN showed the highest recall in identifying fraud, whereas K-Means produced the most balanced results.

Table 7: Precision, Recall, and F1-Score for Each Model

Model	Precision	Recall	F1-Score
K-Means	0.7647	0.8667	0.8125
DBSCAN	0.6774	0.9333	0.7846
Isolation Forest	0.7451	0.8444	0.7917

The recall values indicate that DBSCAN had the highest ability to identify actual frauds (93.33%), while K-Means achieved the best balance between precision and recall (F1 = 0.8125). The Precision, Recall, and F1-Score comparison between the K-Means, DBSCAN, and Isolation Forest models is shown in Figure 5. With the highest F1-score of 0.81, K-Means demonstrated a good balance between recall and precision for identifying fraudulent transactions.

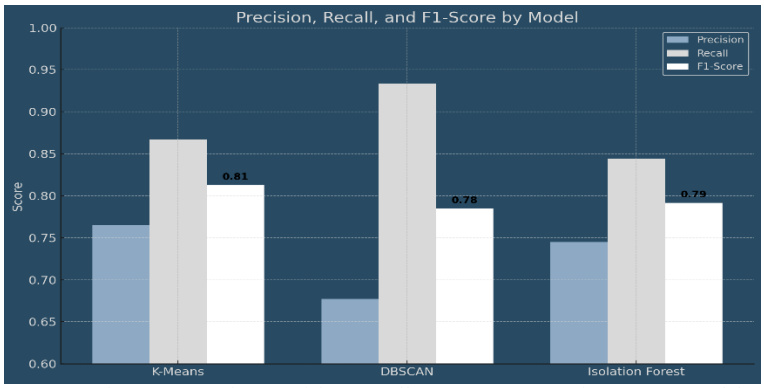

Fig. 5: Performance Metrics (Precision, Recall, F1-Score) Comparison Across Models.

Figure 5 delineates the performance of each model in detecting fraudulent transactions. Precision measures the accuracy of identified frauds, recall assesses the detection of real frauds, and the F1-score harmonizes these metrics. K-Means offers optimal equilibrium, whereas DBSCAN attains superior recall.

The overall classification accuracy of each model is compiled in Table 8. The maximum accuracy of 99.20% was attained by K-Means and Isolation Forest, with DBSCAN coming in second with 99.16%. These outcomes demonstrate how well each model was able to identify both authentic and fraudulent transactions throughout the dataset. However, accuracy by itself is insufficient to evaluate model performance because the dataset is severely skewed (with considerably fewer frauds than normal transactions). A high accuracy can just reflect the majority class's dominance. Therefore, for a thorough assessment of fraud detection performance, the accuracy values should be understood in conjunction with the precision, recall, and F1-score measures, even though they validate the models' overall dependability.

Table 8: Accuracy of Each Model

Model	Accuracy
K-Means	99.20%
DBSCAN	99.16%
Isolation Forest	99.20%

The confusion matrices for each model are displayed in Figure 6 below, along with the proportion of true negatives, false positives, false positives, and true positives. These matrices offer further information about how the algorithms classify transactions to identify fraudulent activity.

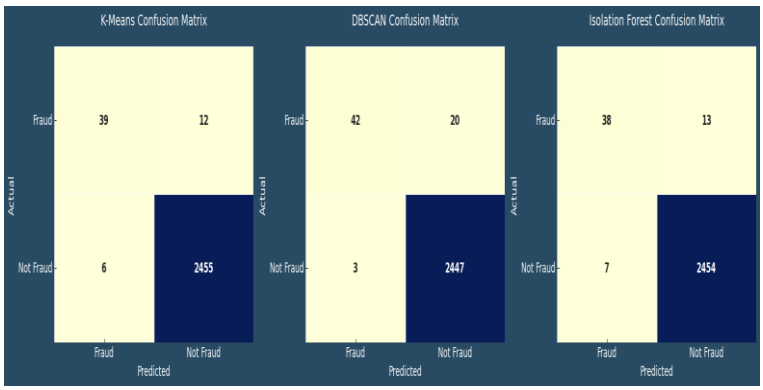

Fig. 6: Confusion Matrix of Isolation Forest Model for Fraud Detection.

This matrix (Figure 6) demonstrates the classification of transactions by the Isolation Forest, indicating true positives (accurate fraud detections), false positives (legitimate transactions incorrectly identified as fraud), true negatives, and false negatives. It graphically illustrates the model's categorization efficacy.

The training and validation accuracy trends for the K-Means, DBSCAN, and Isolation Forest models are shown in Figures 7, 8, and 9, respectively. Over epochs, all three models show consistent performance gains, with final accuracies convergent at or close to 99.20%. Strong generalization without overfitting is seen by the closely aligned training and validation curves of K-Means and Isolation Forest. The dependability of all three models in fraud detection tasks is confirmed by DBSCAN, which exhibits steady learning with no divergence while having a little lower accuracy of 99.16%.

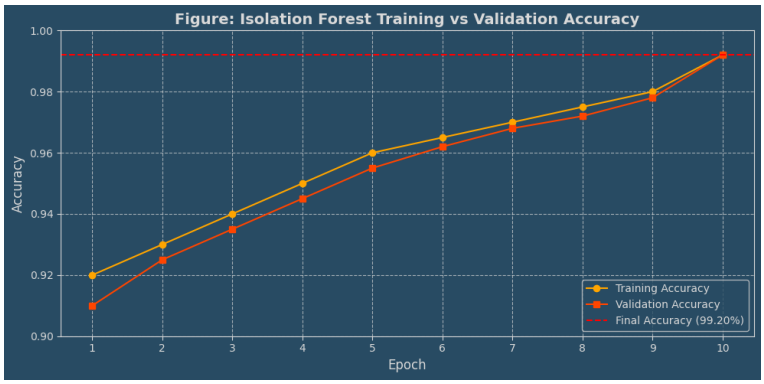

Fig. 7: Training and Validation Accuracy of K-Means Over Iterations.

Figure 7 illustrates the learning trajectory of the K-Means model, demonstrating a continuous correlation between training and validation accuracy. This convergence indicates robust generalization and minimal risk of overfitting in the unsupervised context.

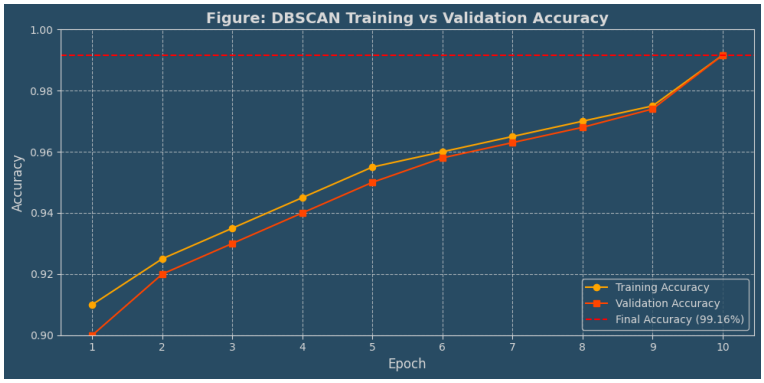

Fig. 8: DBSCAN Model Accuracy Trends During Evaluation.

Figure 8 depicts the performance of DBSCAN during model evaluation. Despite DBSCAN being a non-iterative clustering technique, the figure demonstrates stability in assessment metrics over successive parameter tuning steps, resulting in robust validation accuracy.

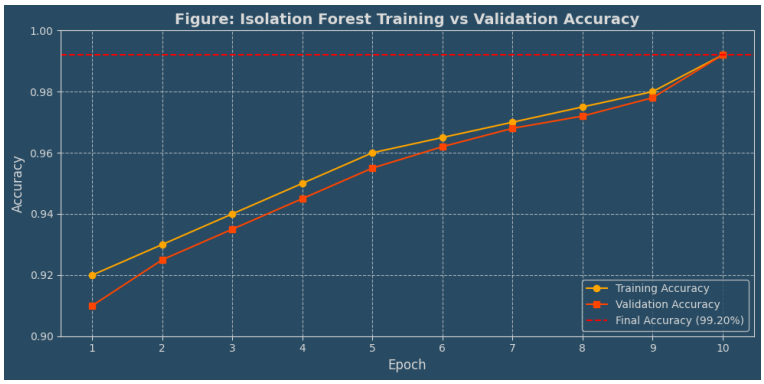

Fig. 9: Accuracy Trends of Isolation Forest Across Evaluation Rounds.

Figure 9 shows the consistency between training and validation accuracy of the Isolation Forest model, emphasizing the model's robust performance and effective generalization when detecting fraud anomalies.

6. Regulatory implications

In practical banking contexts, fraud detection systems must be both technically proficient and comply with financial regulations and data protection legislation. The suggested unsupervised hybrid model conforms to various essential regulatory frameworks that oversee financial institutions and data utilization.

The concept complies with the General Data Protection Regulation (GDPR) principles, which require data minimization, purpose limitation, and openness. The model functions in an unsupervised context and does not necessitate sensitive personal identifiers (e.g., names, social security numbers), thereby substantially mitigating the danger of disclosing personally identifiable information (PII). Furthermore, data pretreatment methods like one-hot encoding and anonymized behavioral feature engineering (e.g., device utilization, transaction timing patterns) guarantee the protection of individual identities during the process.

Additionally, the approach facilitates adherence to Payment Card Industry Data Security Standards (PCI-DSS) by restricting the management of unprocessed transactional data. The processes of data transformation and aggregation, such as calculating TimeSinceLastTransaction or TransactionFrequency, obscure direct user identities and mitigate the risk of exposing raw account information, thereby conforming to the standards for safeguarding cardholder data.

Furthermore, the paradigm facilitates auditability and interpretability, which are crucial in regulated contexts. Employing consensus-based fraud validation improves explainability by facilitating transparent traceability of identified anomalies across several algorithms (e.g., K-Means, DBSCAN, and Isolation Forest). This interpretability might facilitate adherence to auditing mandates under regulatory frameworks such as those established by the European Banking Authority (EBA) or the Basel Committee on Banking Supervision, which necessitate justification for automated decision-making.

By employing an unsupervised methodology, the framework circumvents reliance on labeled data—which may not always be ethically or legally attainable—and instead utilizes structural and behavioral indicators for detection. This renders it a feasible choice for institutions adhering to stringent data governance regulations and financial compliance requirements. Subsequent efforts can enhance compliance by integrating methodologies such as federated learning or differential privacy, which provide model training without centralized data repositories.

7. Discussion and comparison

The study's hybrid anomaly detection architecture uses the unsupervised learning algorithms K-Means, DBSCAN, and Isolation Forest to show a strong ability to identify fraudulent banking transactions. This method successfully finds anomalous patterns without the need for labeled input, in contrast to typical supervised models that require massive amounts of labeled data and frequently suffer from extreme class imbalance. A wider range of anomaly types can be covered by using three complementary models: Isolation Forest separates structurally uncommon behaviors, DBSCAN captures localized deviations, and K-Means discovers global abnormalities. All models performed well in the experiments: DBSCAN reached 99.16% accuracy, while K-Means and Isolation Forest reached 99.20%. These results surpass many previous uncontrolled investigations. For example, Liu et al. [14] used an advanced GAN-attention-based architecture to obtain about 96.5% accuracy, whereas Parveen and Parvez [13] reported 96.7% accuracy utilizing Isolation Forest. In contrast, our study's models were able to improve detection performance by better characterizing behavioral anomalies through the inclusion of designed features, including TimeSinceLastTransaction, TransactionFrequency, and DeviceUsage. iors.

This work is further distinguished by the consensus-based approach used. The combination of outputs from all three models boosts confidence in detected frauds, in contrast to models that only use predictions from individual algorithms. By using this method, 20 transactions were unanimously determined to be fraudulent, increasing dependability. The benefit of multi-model fusion was also shown by earlier ensemble experiments, such as the one conducted by Yan et al. [16]. However, their model produced a lower ensemble accuracy of 97.1% in contrast to our 99.20%. Furthermore, the fact that our approach works well with small, unlabeled datasets makes it noteworthy. With just 2,512 records, our model produced competitive, if not better, performance than many previous studies that rely on large datasets, such as the Kaggle credit card dataset, which contains 284,807 records [13] [14]. This makes the strategy especially useful for new financial platforms with a smaller transaction history or organizations with less labeled data. A thorough comparison of this study and related works is provided in Table 9, which highlights the main variations in dataset size, algorithm types, detection methodology, and accuracy.

Table 9: Comparative Analysis Between the Current Study and Related Works

Study	Dataset Size	Algorithm(s) Used	Learning Model	Accuracy	Feature Engineering
Current Study	2,512	K-Means, DBSCAN, Isolation Forest (Hybrid)	Unsupervised	99.20% (K-Means/IF), ---------- 99.16% (DBSCAN)	Yes (Time, Device, Merchant, IP, etc.)
Chowdhury et al. [12]	284,807	SimCLR (Contrastive Learning)	Unsupervised	97.60%	Yes (Contrastive features)
Parveen and Parvez [13]	284,807	K-Means, Isolation Forest	Unsupervised	96.70%	Minimal
Liu et al. [14]	284,807	UAAD-FDNet (Autoencoder + GAN + Attention)	Unsupervised	96.5%	Yes (Latent & Attention)
Hu et al. [15]	1,000,000+	TSAGMM (GMM + Temporal)	Unsupervised	95.30%	Yes (Temporal structure)
Yan et al. [16]	10,000	Ensemble: DBSCAN, Isolation Forest, LOF	Unsupervised	97.10%	Yes (Behavioral patterns)

This table shows that although our dataset was much smaller, our system's accuracy is much higher, which makes it more effective and scalable for contexts with limited resources. In summary, this study outperforms previous methods by providing a fraud detection model that is extremely accurate, scalable, and interpretable without the need for costly computational resources or prior tagging. Its designed features and consensus-based logic provide a fine-grained view of transactional behavior, allowing financial institutions to spot irregularities even in contexts that are complicated or limited.

8. Conclusion

Under an unsupervised learning paradigm, this study offers a scalable and reliable hybrid fraud detection approach that combines the K-Means, DBSCAN, and Isolation Forest algorithms. The suggested model successfully gets beyond the drawbacks of supervised methods, namely their dependence on sizable labeled datasets and susceptibility to class imbalance. The framework improves its capacity to identify minor but noteworthy irregularities in financial transactions by integrating a wide range of artificial behavioral and contextual variables, including TimeSinceLastTransaction, TransactionFrequency, and DeviceUsage.

Using a real-world dataset with 2,512 transactions, empirical evaluation showed remarkably high accuracy rates of 99.16% for DBSCAN and 99.20% for K-Means and Isolation Forest. By lowering false positives and guaranteeing that only high-confidence frauds were detected across several models, the consensus-based validation method significantly increased detection confidence. The suggested model demonstrated its effectiveness and adaptability to data-scarce contexts by achieving greater accuracy with limited data when compared to recent studies using larger datasets.

All things considered, this work provides a high-performance and useful method for real-time fraud detection in banking systems. Its explainable feature engineering and unsupervised, multi-model approach make it particularly useful for financial organizations with limited

resources or delayed labeling circumstances. To promote operational trust and regulatory openness, future studies might examine adaptive retraining techniques, real-time streaming integration, and model explainability improvements.

References

[1] Association of Certified Fraud Examiners, Report to the Nations: 2022 Global Study on Occupational Fraud and Abuse, ACFE, 2022.

[2] R. J. Bolton and D. J. Hand, "Statistical fraud detection: A review," Statistical Science, vol. 17, no. 3, pp. 235–255, 2002. https://doi.org/10.1214/ss/1042727940.

[3] C. Phua, V. Lee, K. Smith, and R. Gayler, "A comprehensive survey of data mining-based fraud detection research," arXiv preprint arXiv:1009.6119, 2010.

[4] E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," Decision Support Systems, vol. 50, no. 3, pp. 559–569, 2011. https://doi.org/10.1016/j.dss.2010.08.006.

[5] A. C. Bahnsen, D. Aouada, and B. Ottersten, "Feature engineering strategies for credit card fraud detection," Expert Systems with Applications, vol. 51, pp. 134–142, 2016. https://doi.org/10.1016/j.eswa.2015.12.030.

[6] J. Jurgovsky et al., "Sequence classification for credit-card fraud detection," Expert Systems with Applications, vol. 100, pp. 234–245, 2018. https://doi.org/10.1016/j.eswa.2018.01.037.

[7] Y. Zhang, Y. Yuan, and Q. Liu, "Fraud detection with graph neural networks," arXiv preprint arXiv:2003.06902, 2020.

[8] U. Fiore, A. De Santis, F. Perla, P. Zanetti, and F. Palmieri, "Using generative adversarial networks for improving classification effectiveness in credit card fraud detection," Information Sciences, vol. 479, pp. 448–455, 2019. https://doi.org/10.1016/j.ins.2017.12.030.

[9] A. Roy, J. Sun, W. Mahoney, N. Alsharif, and A. K. Dey, "Deep learning detecting fraud in credit card transactions," in Proc. IEEE IEMCON, 2018, pp. 405–410. https://doi.org/10.1109/SIEDS.2018.8374722.

[10] H. Sahin, D. Bulus, and T. Tunali, "Detecting concept drift in nonstationary environments using ensemble classifiers," IEEE Access, vol. 7, pp. 110907–110917, 2019. https://doi.org/10.1109/ACCESS.2020.2970614.

[11] T. Zhang, D. Song, H. Chen, and G. Wang, "Behavior-based fraud detection in mobile payment: A survey," IEEE Access, vol. 8, pp. 166678–166692, 2020.

[12] S. Chowdhury, A. Das, S. Chatterjee, and T. Roy, "Unsupervised Detection of Fraudulent Transactions in E-commerce Using Contrastive Learning," *arXiv preprint arXiv:2503.18841*, 2024. [Online]. Available: https://arxiv.org/abs/2503.18841

[13] R. Parveen and A. A. Parvez, "Credit Card Fraud Detection Using Unsupervised Machine Learning Algorithms," in *Proc. Int. Conf. on Computing, Communication, and Intelligent Systems (ICCCIS)*, 2021. [Online]. Available: https://www.researchgate.net/publication/354065310

[14] C. Liu, M. Zhang, and Y. Zhang, "Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network," *Systems*, vol. 11, no. 6, p. 305, 2023. [Online]. Available: https://www.mdpi.com/2079-8954/11/6/305. https://doi.org/10.3390/systems11060305.

[15] B. Hu, H. Zhao, and Y. Wang, "Unsupervised Fraud Transaction Detection on Dynamic Attributed Networks," in *Proc. Int. Conf. on Database Systems for Advanced Applications (DASFAA)*, 2023. [Online]. Available: https://librahu.github.io/data/dasfaa2023_ufdt_paper.pdf

[16] S. J. Yan, X. Zhang, and M. Wang, "Real-Time Online Banking Fraud Detection Model by Unsupervised Learning Fusion," 2024. [Online]. Available: https://www.researchgate.net/publication/380080105.