Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Voting Classifier as a Balanced Framework for Fraud Detection in Imbalanced Credit Card Transactions

Ajay Kumar[1], Dr. Avinash Panwar[2]

[1]Research Scholar Department of Computer Science, Mohanlal Sukhadia University Udaipur
Email: ajaymahayach@live.com
[2]HOD of Computer Science, Mohanlal Sukhadia University Udaipur
Email: avinash@mlsu.ac.in

ARTICLE INFO

Received: 10 Sept, 2024

Revised: 14 Oct, 2024

Accepted: 25 Nov, 2024

Published: 20 Dec, 2024

ABSTRACT

Credit card fraud is a critical concern for financial institutions, as it leads to significant economic losses and compromises customer confidence. Detecting fraudulent activity remains challenging due to the extreme imbalance between legitimate and fraudulent transactions in real-world datasets. In this study, the publicly available dataset is analyzed to investigate the effectiveness of machine learning algorithms for fraud detection. The dataset exhibits a severe skew, with fraudulent cases representing only a small fraction of all transactions. To address this imbalance, the Synthetic Minority Oversampling Technique (SMOTE) is employed, enabling models to better learn discriminatory patterns. Several machine learning approaches, including Logistic Regression, Random Forest are implemented and evaluated. Performance is measured using accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC), ensuring a comprehensive assessment of classification capability. Experimental results demonstrate that resampling combined with ensemble methods significantly improves the detection of minority fraud cases while minimizing false positives. This work emphasizes the importance of handling imbalanced data in fraud detection and provides insights into the potential of machine learning to enhance the security and reliability of electronic payment systems.

Keywords: Credit card fraud detection, Imbalanced dataset, Machine learning, SMOTE, Classification, Ensemble methods, Fraud analytics

1. Introduction

The banking industry has been profoundly impacted by the evolution of information technology (IT). Credit card and online net banking transactions, which are currently the majority of banking system transactions, all present additional vulnerabilities [1]. Hackers have increasingly targeted banks with enormous quantities of client data. Therefore, banks have been in the forefront of cyber security for business. In the past thirteen years, cyber security industry expanded fast. The market is predicted to be valued 170.4 billion in 2022 [2]. In the next three years, the cost of cybercrime is expected to rise by 15% every year, finally exceeding $10.5 trillion USD each year by 2025 [3].

In the banking industry, cyber fraud using credit cards is a significant concern that costs billions of dollars annually. Banking industry has made strengthening cyber security protection a priority. Multiple systems have been developed for monitoring and identifying credit card cyber fraud. However, because of the constantly evolving nature of threats, banking industry must be equipped with the most modern and effective cyber fraud management technologies [4], [5]

The acceptance of credit card and other forms of online payments has exploded in recent years; this resulted in an increase in cyber fraud in credit cards. In credit card, there are several forms of cyber fraud. The first type is the actual theft of a credit card. The theft of confidential details of credit card is the second type of cyber fraud. When the credit card information is entered without the cardholder's permission during an online transaction, further fraud is [5], [6].

The detection of cyber fraud in credit cards is a challenging task that attracted the interest of academics working in the fields of machine learning (ML). Datasets associated with credit cards have significant skewness. A great number of algorithms are unable to discriminate items from minority classes when working with datasets that have a considerable skew. In order to achieve efficiency, the systems that are used to identify cyber fraud need to react swiftly. Another important matter of concern is the way

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

in which new methods of attack, influence the conditional distribution of the data over the time period [7]. According to [8], there are a number of challenges need to be addressed for cyber fraud detection in credit card. These challenges contain massive volume of data, that is unbalanced or incorrectly categorised, frequent changes in the type of transaction, and real-time detection. As current technology being progressed, cyber credit card fraud is also developing rapidly, making cyber fraud detection a crucial area. The conventional techniques to resolve this problem is no longer sufficient. In the conventional technique, domain experts in cyber fraud compose the algorithms which are governed by strict rules. In addition, a proactive strategy must be used to combat cyber fraud. Every industry is attempting to employ ML-based solutions due to their popularity, speed, and effectiveness [9].

While there are numerous cyber fraud detection techniques available, as yet no fraud detection systems have been able to deliver high efficiency and high accuracy. Thus, it necessary to provide an overview in cyber fraud detection and an analysis of the most recent studies in this field to conduct innovation projects for cyber fraud detection. To achieve this goal, this review will provide a detailed analysis of ML techniques and their function in credit card cyber fraud detection and also offer recommendations for selecting the most suitable techniques for detecting cyber fraud. The study also includes the trends of research gaps, and limitations in detecting cyber fraud in credit card. This study therefore seeks to compare six classification and prediction techniques, namely; Decision Tree, Logistic Regression, and Random Forest in classifying and predicting financial transactions as either fraudulent or not fraudulent.

2. Literature Review

Credit card fraud detection has evolved from rule-based and statistical systems to advanced machine learning (ML) and deep learning approaches. Traditional models such as logistic regression and decision trees were among the earliest techniques, offering interpretability but limited adaptability to complex fraud patterns [10]. Machine learning algorithms, including Support Vector Machines (SVM), Random Forests (RF), and k-nearest neighbors (KNN), provided improvements by capturing nonlinear relationships in high-dimensional data [11]. Ensemble methods, particularly Random Forest and Gradient Boosting, further enhanced classification accuracy by aggregating multiple weak learners [12].

Recent studies have focused on addressing challenges inherent in fraud detection datasets, especially high dimensionality and class imbalance. A hybrid feature-selection framework that combines information gain with a genetic algorithm wrapper has been proposed, ensuring that only the most relevant attributes are used for classification [13]. This approach improved both sensitivity and specificity, demonstrating the effectiveness of dimensionality reduction in fraud detection. In another study, SMOTE was applied to balance the fraud dataset and classifiers such as Naïve Bayes (NB), Random Forest (RF), and Multilayer Perceptron (MLP) were tested [14]. The results showed MLP achieving a remarkable accuracy of 99.95%, outperforming other algorithms, thus highlighting the impact of oversampling on imbalanced data.

Another line of research has emphasized developing fraud detection frameworks that balance adaptability and robustness. A comprehensive ML framework incorporating exploratory data analysis, feature scaling, and class balancing using SMOTE and undersampling has been proposed [15]. In this study, Random Forest (RF) and Support Vector Machine (SVM) were compared, with results showing that RF offered a better trade-off between precision and recall, while SVM suffered from overfitting. Importantly, the research highlighted the necessity for continuous monitoring and model adaptation in response to evolving fraud patterns.

Despite these advances, the issue of imbalanced datasets remains central to the literature. The Synthetic Minority Oversampling Technique (SMOTE) was introduced as a solution and has since been widely adopted to generate synthetic minority-class samples and improve detection performance [16]. Alternative strategies include cost-sensitive learning, where misclassification costs are adjusted to prioritize fraud detection [17]. Moreover, while accuracy remains a standard metric, recent studies emphasize recall, precision, F1-score, and ROC-AUC, as missing fraudulent transactions incurs significant financial losses [18].

3. Data and Methods

3.1 Data

The data collection included credit card transaction simulations from January 1, 2020, to December 31, 2020, including both legitimate and fraudulent transactions in the western side of the United States of America available at https://www.kaggle.com/datasets/kartik2112/fraud-detection. The dataset contains 555,719 credit card transactions with 23 attributes, including transaction details, customer demographics, merchant information, and a binary target variable *is_fraud* indicates whether a transaction is fraudulent (1) or genuine (0). Out of all records, only 2,148 transactions (0.39%) are fraudulent, highlighting a significant class imbalance.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

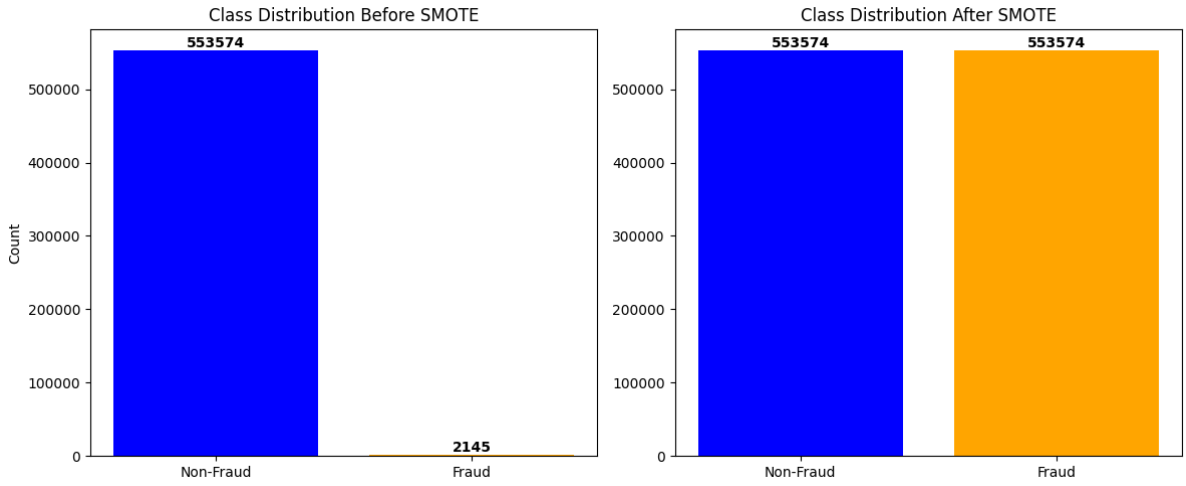

Figure 1: Oversampling Data.

Table 1: Basic Statistics for the character variables.

Variable	count	unique	top	frequency
Trans_Date_Trans_Time	555719	544760	05-10-2020 19:37	4
Merchant	555719	693	fraud_Kilback LLC	1859
Category	555719	14	gas_transport	56370
First	555719	341	Christopher	11443
Last	555719	471	Smith	12146
Gender	555719	2	F	304886
Street	555719	924	444 Robert Mews	1474
City	555719	849	Birmingham	2423
State	555719	50	TX	40393
Job	555719	478	Film/video editor	4119
Dob	555719	910	23-03-1977	2408
Trans_Num	555719	555719	1765bb45b3aa3224b4cdcb6e7a96cee3	1

Table 2: Basic statistics for the numeric variables.

Variable	count	mean	std	min	25%	50%	75%	max
Unnamed: 0	555719	277859	160422.4	0	138929.5	277859	416788.5	555718
Cc_Num	555719	4.18E+17	1.31E+18	6.04E+10	1.8E+14	3.52E+15	4.64E+15	4.99E+18
Amt	555719	69.39281	156.7459	1	9.63	47.29	83.01	22768.11
Zip	555719	48842.63	26855.28	1257	26292	48174	72011	99921
Lat	555719	38.54325	5.061336	20.0271	34.6689	39.3716	41.8948	65.6899
Long	555719	-90.2313	13.72178	-165.672	-96.798	-87.4769	-80.1752	-67.9503
City_Pop	555719	88221.89	300390.9	23	741	2408	19685	2906700
Unix_Time	555719	1.38E+09	5201104	1.37E+09	1.38E+09	1.38E+09	1.39E+09	1.39E+09
Merch_Lat	555719	38.5428	5.095829	19.02742	34.7553	39.37659	41.95416	66.6793
Merch_Long	555719	-90.2314	13.73307	-166.672	-96.9051	-87.4452	-80.2646	-66.952
Is_Fraud	555719	0.00386	0.062008	0	0	0	0	1

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

In the pre-processing stage, the dataset was cleaned by removing irrelevant fields (e.g., names, street) and standardizing formats for timestamps and dates of birth. Categorical variables such as gender, state, category, and job were encoded using one-hot or frequency encoding, while numeric features like transaction amount, city population, and geolocation were standardized. Feature engineering introduced additional attributes including transaction hour, customer age, geographic distance between customer and merchant, and transaction frequency. To address the strong class imbalance (0.39% fraud), techniques such as SMOTE oversampling were applied. Figure 1 shows the oversampling data. Table 1 and Table 2 shows the summery statistics of the types of variables used in the study.

3.2 Methods

This section covers the use of supervised machine learning models for fraud classification, including Random Forest, Decision Tree, logistic regression XGBoost, KNN and voting classifier.

3.2.1 Decision Tree

Decision Trees are supervised; non-parametric learning algorithms commonly applied in classification tasks such as fraud detection [19]. They operate by recursively partitioning the dataset into smaller subsets based on feature values, ultimately forming a tree-like structure. The tree is composed of a root node, decision nodes, and leaf nodes, where the root node initiates the split using the most informative feature, and the leaf nodes represent the final class labels [20]. Graphically, they show information in a tree pattern that is easy to understand. The decision tree structure is made up of nodes, edges, and leaf nodes. According to [21], it consists of a set of branches/nodes that are connected by edges. Figure. 2 shows the flow diagram of a decision tree. The decision tree's root node chooses a feature to partition the data into two or more sub nodes to develop decision nodes after the partition into
sub nodes and subtrees at the end of the root node [22]. Each sub-tree of the data will once more be partitioned into two sub nodes. Until every training sample is gathered, this process will be repeated. So, at the end of the decision tree, we end up with a leaf node which serves as a representation of the class which aims at classifying.

The Decision Tree algorithm is advantageous because it does not require feature scaling, can effectively manage outliers, and automatically handles missing values. It is relatively fast to train and highly effective for both classification and prediction tasks. For splitting nodes, the algorithm relies on metrics such as the Gini index, information gain, and entropy. The overall modelling process of the Decision Tree is illustrated in Figure 2.

Entropy is a measure of impurity or randomness in a dataset and is widely used in Decision Tree algorithms to determine the quality of a split. It quantifies the uncertainty associated with class distributions. If all samples in a node belong to the same class, the entropy is zero, indicating perfect purity. Conversely, when the classes are evenly distributed, entropy reaches its maximum value of one, indicating maximum disorder [23].

The general formula for entropy is:

$$E(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) \quad (1)$$

where:

- $p(x_i)$= probability of class i in the dataset
- The logarithm is base 2, because entropy is measured in bits

For binary classification (Fraud vs. Not Fraud):

$$E(X) = -p(Fraud) \log_2 p\,(Fraud) - p(Not\ Fraud) \log_2 p\,(Not\ Fraud) \quad (2)$$

The Gini Index measures the probability of incorrectly classifying a randomly chosen instance if it were randomly labeled according to the distribution of classes in the node. A Gini value of 0 represents a perfectly pure node, while higher values indicate greater impurity [24]. The Gini index value ranges between the values of 0 and 0.5. This implies that an attribute with a lower Gini index is automatically selected for the splitting. The formula for calculating the Gini Index is

$$E(X) = 1 - \sum_{i=1}^{n} p(x_i)^2 \quad (3)$$

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

where:

- $p(x_i)$ = Probability of class i in the dataset.

Information Gain (IG) is a statistical measure that evaluates how well a feature separates the dataset into distinct classes [25]. It is defined as the reduction in entropy achieved after a split, representing the amount of useful information obtained. In Decision Tree construction, the objective is to select the attribute that produces the highest information gain and, correspondingly, the lowest entropy. The formulas for calculating IG are presented as follows.

$$G(X,Y) = E(X) - E(X|Y) \tag{4}$$

$$G(X,Y) = -p(Fraud)\log_2 p\,(Fraud - p(Not\ Fraud)\log_2 p(Not\ Fraud) - \sum \frac{|Sv|}{S} entropy(Sv) \tag{5}$$

Information Gain represents the reduction in uncertainty about Y when additional knowledge of X is provided. It is computed by subtracting the entropy of X from the entropy of Y. The greater the reduction in uncertainty, the more information X contributes toward predicting Y.

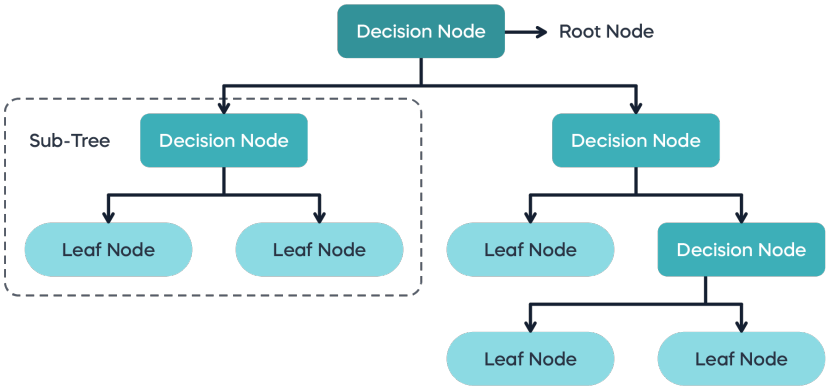

Figure 2: Decision Tree

3.2.2 Logistic Classification

Logistic Regression is a widely used supervised learning algorithm for binary classification problems, making it suitable for detecting fraudulent credit card transactions. The model predicts the probability that a transaction belongs to a particular class (fraud or not fraud) by mapping a linear combination of input features through the sigmoid function [26]. This ensures that the output is constrained between 0 and 1, representing the likelihood of fraud.

The logistic model is expressed as:

$$P(y = 1 \mid X) = \frac{1}{1 + e^{-(\alpha+\beta_1+\beta_2 x_2+\cdots+\beta_n x_n)}} \tag{6}$$

where:

- $y \in \{0,1\}$ is the dependent variable ($1 = Fraud, 0 = Not\ Fraud$)
- $X = (x_1, x_2, \ldots x_n)$ represents the input features (e.g., transaction amount, merchant type, transaction time, customer demographics),
- α is the intercept (bias term),

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- β are the coefficients (weights) learned by the model for each feature,
- e is the natural exponential function.

Once the probability is computed, classification is performed using a threshold (commonly 0.5):

$$\hat{y} = \begin{cases} 1 \; if \; P(y = 1|X) \geq \theta \\ 0 \; if \; P(y = 1|X) < \theta \end{cases} \quad (7)$$

where θ is the classification threshold.

- If, $P(y = 1|X) \geq 0.5$, the transaction is classified as Fraud.
- Otherwise, it is classified as Not Fraud.

Since the relationship between $P(y)$ and xxx is nonlinear, the parameters α and β are not as straightforward to interpret as in linear regression. The logistic curve, shown in Figure 3, represents probabilities and is confined to values between 0 and 1.

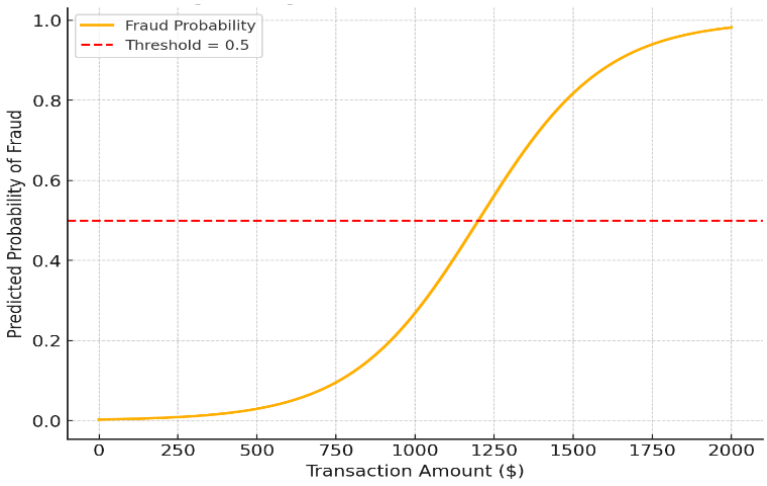

Figure 3: Logistic Regression Curve for Fraud Detection

3.2.3 Random Forest

Random Forest is an ensemble learning method based on the combination of multiple Decision Trees [27]. It is widely used in fraud detection because of its ability to handle large, imbalanced datasets and capture complex, nonlinear relationships between features. Unlike a single Decision Tree, which may suffer from overfitting, Random Forest aggregates the predictions of many trees to improve accuracy, robustness, and generalization [28]. Random Forest uses the bagging technique to build a collection of Decision Trees. For a dataset (X, Y) with (N) observations, where X represents the predictor variables and Y the outcome, the algorithm creates random subsets of the data. Each subset is used to train a separate Decision Tree. This results in multiple trees $(dK_1(X), (dK_2(X), \ldots . (dK_N(X))$,. The final prediction is then made by combining the outputs of all trees, typically through majority voting in classification tasks.

$$D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N) \quad (8)$$

where:

- $x_i = (x_{i1}, x_{i2}, \ldots, x_{in}$ are the feature vectors (e.g., amount, time, merchant).
- $y \; \epsilon \; \{0,1\}$ is the class label (0 = Not Fraud, 1 = Fraud).

Random Forest usually does not need a separate feature selection process [29]. However, one limitation is that it may sometimes treat variables with many possible values or wide ranges as more important, which can lead to biased results in fraud detection. Despite this, it is considered one of the most accurate algorithms used in the financial sector for detecting fraud [30]. When

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

building the trees, Random Forest can be less certain at the start, so it becomes important to identify and use the most relevant features for splitting the nodes during analysis.

3.2.3 XGBoost

XGBoost (Extreme Gradient Boosting) is an optimized implementation of the gradient boosting framework. It is one of the most powerful algorithms for credit card fraud detection because it handles large, imbalanced datasets efficiently, captures nonlinear feature interactions, and includes regularization to reduce overfitting [31].

Unlike Random Forest, which uses bagging (parallel trees), XGBoost builds trees sequentially**,** where each new tree attempts to correct the errors of the previous ones.

$$D = \{(x_1, y_1), (x_2 , y_2), \ldots, (x_n, y_n)\}, \qquad y_i \in \{0,1\} \tag{9}$$

where:

- x_i transaction features (amount, time, merchant, etc.),
- y_i = label (1 = Fraud, 0 = Not Fraud).

3.2.4 K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a simple yet powerful instance-based learning algorithm**.** It classifies a new transaction by looking at the k most similar past transactions in the dataset and assigning the majority class among them (Fraud or Not Fraud) [32].

In fraud detection, KNN is useful because it does not make strong assumptions about data distribution. Fraudulent transactions can often be identified by their "closeness" to other fraudulent cases in feature space. The expression used for computing KNN:

$$D = \{(x_1, y_1), (x_2, y_2), \ldots\ldots, (x_N, y_N)\} \tag{10}$$

where:

- $x_i = (x_{i1}, x_{i2}, \ldots\ldots, x_{in}$ are the feature vectors (e.g., amount, time, merchant).
- $y \in \{0,1\}$ is the class label (0 = Not Fraud, 1 = Fraud).

3.2.5 Voting Classifier

Credit card fraud detection is a challenging task due to the highly imbalanced nature of transaction data (very few fraudulent cases compared to legitimate ones) and the complexity of fraud patterns. A single machine learning algorithm may fail to capture all the nuances of fraud. To address this, ensemble learning techniques are widely used, as they combine the strengths of multiple classifiers to improve accuracy and robustness [33].

One of the simplest yet powerful ensemble approaches is the Voting Classifier. The Voting Classifier integrates predictions from multiple base models to make a single, more reliable decision. In fraud detection, this is particularly useful because models like Logistic Regression, Random Forest, K-Nearest Neighbors (KNN), and Support Vector Machines (SVM) may perform differently on various patterns of fraud [34]. By combining them, we obtain a model that is generally more accurate and less biased toward one particular type of fraud pattern.

$$D = \{(x_1, y_1), (x_2, y_2), \ldots\ldots, (x_N, y_N)\} \tag{11}$$

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

where:

- $x_i = (x_{i1}, x_{i2}, \ldots, x_{in}$ are the feature vectors (e.g., amount, time, merchant).
- $y \in \{0,1\}$ is the class label (0 = Not Fraud, 1 = Fraud).
- A set of classifiers:

$$C = \{C_1, C_2, \ldots\ldots, C_M\} \tag{12}$$

where $M$is the total number of classifiers.

Each classifier C_j produces either a class label prediction $\hat{y}(x)$ or a probability estimate $P_j(y = 1|x)$.

We evaluated the model's performance using metrics such as accuracy, precision, recall, specificity, and F1-score to compare different algorithms. While accuracy is the most common way to measure performance [35], it is not always suitable for our highly imbalanced dataset. Relying only on accuracy could be misleading, since the majority of transactions are legitimate. Therefore, it is more reliable to assess models using additional metrics, such as the area under the curve (AUC) [36], along with accuracy, to better identify fraudulent transactions.

The confusion matrix's entries are defined as follows: False positive (FP) is the total number of incorrect predictions classified as positive; false negative (FN) is the total number of incorrect predictions classified as negative; true positive (TP) is the total number of true predictions classified as positive; and true negative (TN) is the total number of true predictions classified as negative.

Accuracy measures how often the model makes correct predictions, whether for fraud or not fraud [37]. It is the ratio of all correct predictions to the total number of predictions made. It is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

Precision metric measures the ratio of correctly classified fraud transactions (TP) to the total transactions predicted to be fraud transactions (TP + FP) [38]. It is calculated as

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

Recall/ Sensitivity, as a metric, measures the ratio of correctly classified fraud transactions (TP) to the total number of fraud transactions [39]. It is calculated as

$$Recall\ /\ Sensitivity = \frac{TP}{TP + FN} \tag{15}$$

Specificity measures the ratio of correctly classified not fraud transactions (TP) to the total number of Not Fraud transactions [40]. It is calculated as;

$$Specificity = \frac{TN}{TN + FP} \tag{16}$$

F1-score is a metric that balances Precision and Recall. It is especially useful in fraud detection, where the dataset is imbalanced. The F1-score is the harmonic mean of Precision and Recall, giving a single measure of a model's accuracy on fraud cases.

$$F1\ Score = \frac{2 \times precision \times recall}{precision \times recall} \tag{17}$$

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

AUC (Area Under the Curve) is measured from the ROC curve, which plots the False Positive Rate (x-axis) against the True Positive Rate (y-axis) for different threshold values between 0 and 1. It shows how well the model separates fraud from non-fraud cases. A higher AUC means the model is better, as its curve is closer to the top-left corner of the graph. A poor model will lie closer to the 45-degree diagonal line, which represents random guessing where the False Positive Rate equals the True Positive Rate.

4. Results

The feature correlation heatmap illustrates the pairwise relationships between the variables in the dataset. As shown in the Figure 4, most features exhibit weak correlations with each other and with the target variable *(is_fraud)*. The transaction amount *(amt)* demonstrates a mild positive correlation (0.18) with fraudulent transactions, indicating that higher amounts are slightly more likely to be associated with fraud. In contrast, features such *as cc_num, city_pop,* and *unix_time* show negligible correlation with the target, suggesting that their predictive power may emerge only through complex interactions rather than direct linear relationships. The heatmap also highlights cases of redundancy among features. Specifically, *lat* and *merch_lat* (0.99) as well as long and *merch_long* (≈1.0) are highly correlated, reflecting near-identical values. Additionally, zip exhibits a strong negative correlation (≈-0.91) with long and *merch_long*. Such highly correlated pairs may contribute to multicollinearity and can be considered for dimensionality reduction or feature selection. Overall, the heatmap emphasizes the need for advanced machine learning methods that can capture non-linear and multivariate patterns rather than relying on simple correlations.

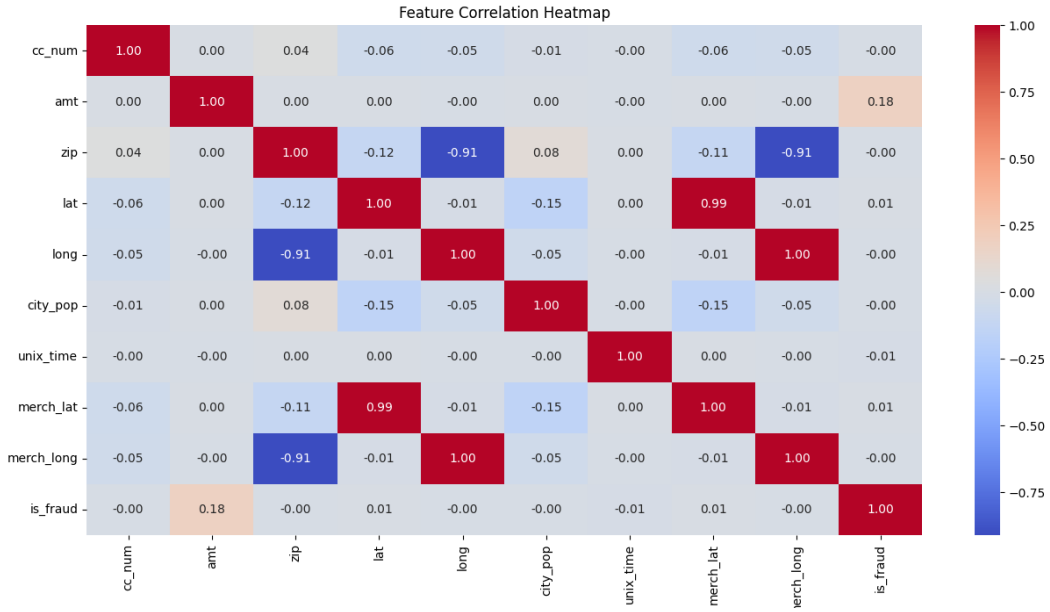

Figure 4: Correlation plot of quantitive variables.

Figure 5 shows the transaction amount distributions reveal clear differences between fraudulent and non-fraudulent activities. The overall and non-fraud distributions are highly skewed toward smaller values, with the majority of legitimate transactions occurring below 200 units. In contrast, the distribution of fraudulent transactions is more dispersed, with notable concentrations in the mid-to-high ranges (200–1200). This suggests that while small-value transactions dominate the dataset overall, fraudulent activities are more likely to involve larger amounts. These findings align with the correlation analysis, where transaction amount showed a modest positive relationship with fraud, reinforcing its importance as a discriminative feature.

Figure 6 shows the analysis of occupational distribution among fraudulent transactions reveals that certain job titles are disproportionately represented. The top three professions most associated with fraud are science writers, licensed conveyancers, and systems developers, each accounting for nearly 30 fraudulent cases. Other occupations, including biomedical engineers, occupational therapists, colour technologists, counsellors, comptrollers, research scientists, and commissioning editors, also appear frequently in the dataset. The diversity of professions represented suggests that fraud is not limited to a particular occupational group but occurs across a wide range of professional backgrounds. However, the higher counts observed in specific job categories may reflect underlying differences in income levels, transaction behaviors, or sampling biases in the dataset.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

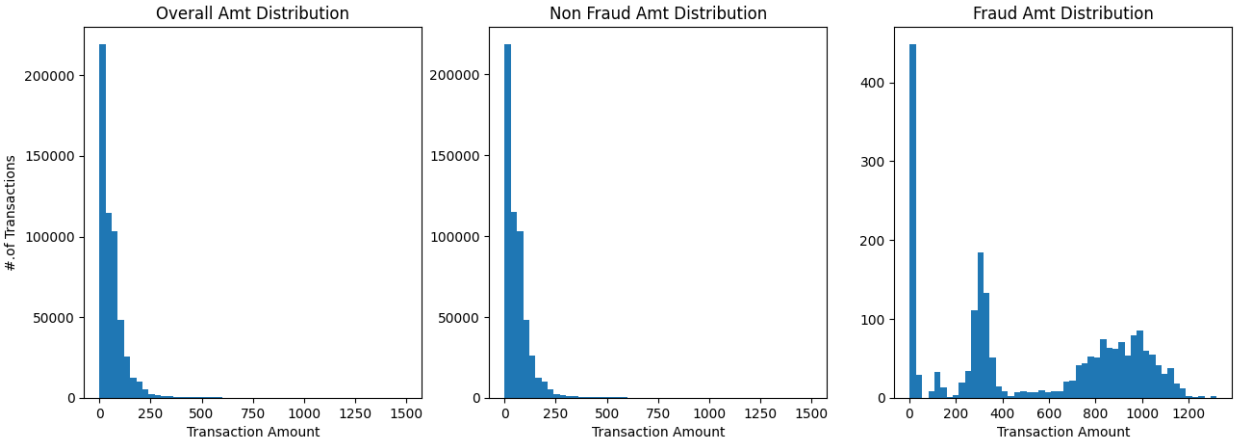

Figure 5: Distribution of transaction amounts across the dataset.

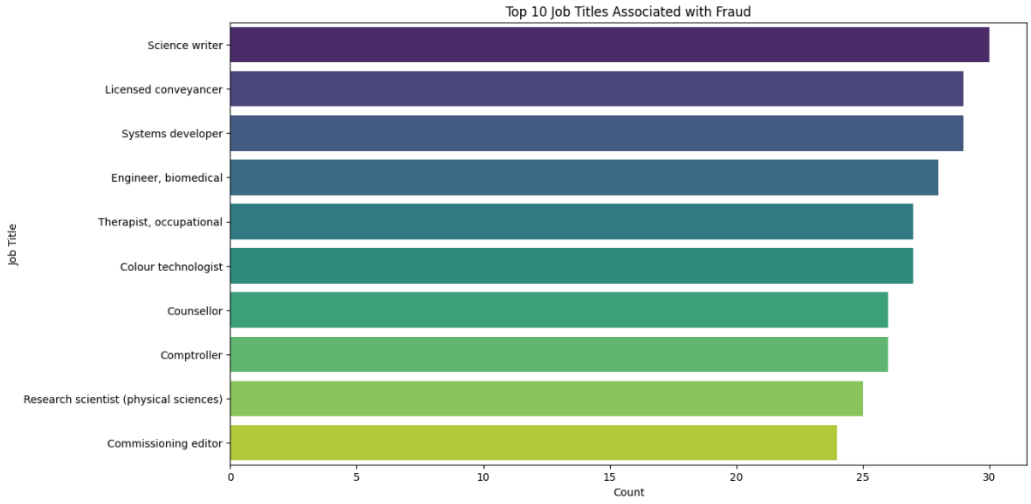

Figure 6: Job titles most frequently associated with fraudulent transactions.

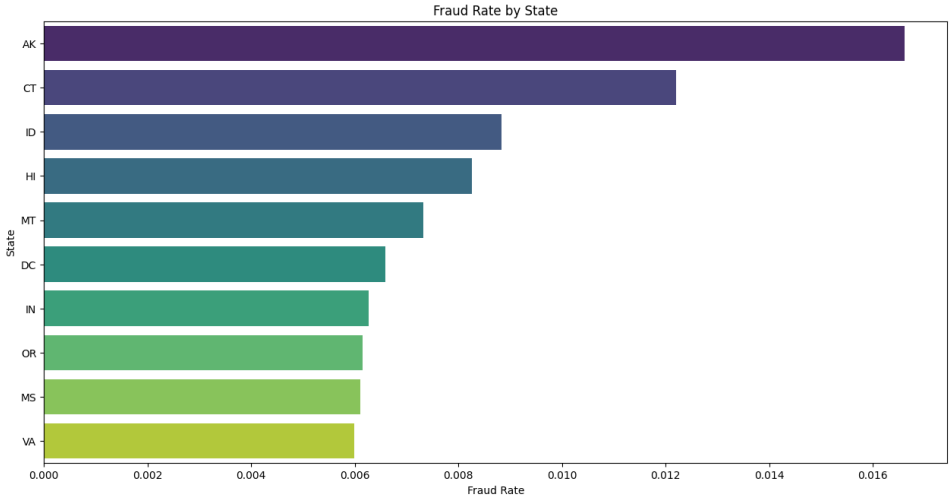

Figure 7: Fraud rate by U.S. state.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

As illustrated in Figure 7, the state-wise analysis of fraud rates indicates notable geographic variation in fraudulent activity. Alaska records the highest fraud rate at over 1.6%, followed by Connecticut with a rate of approximately 1.2%. Other states such as Idaho, Hawaii, Montana, and the District of Columbia also report relatively high fraud rates compared to the national average. Meanwhile, Virginia, Mississippi, and Oregon fall at the lower end of the top ten but still maintain rates above 0.6%. These findings suggest that fraud is not uniformly distributed across regions and may be influenced by state-level demographic, economic, or transactional factors.

Figure 8 shows the temporal analysis of fraudulent transactions revealed a distinct concentration of fraudulent activity during late-night hours. Fraud rates were notably higher between 22:00 and 23:00, with nearly 2% of all transactions in these periods being fraudulent, compared to less than 0.1% during most daytime hours. Absolute counts also confirmed this trend, with more than 1,000 fraud cases recorded in the two-hour window before midnight, while morning and afternoon hours recorded fewer than 30 fraudulent cases per hour on average. Moderate fraud activity was also observed between midnight and 3:00 AM, suggesting that fraudsters exploit the reduced monitoring or lower transaction scrutiny that may occur overnight. Conversely, daytime hours, particularly between 8:00 AM and 4:00 PM, showed consistently low fraud rates. These findings indicate that fraudulent behavior is temporally skewed, with late evenings and early mornings presenting the highest risk for financial institutions and customers.

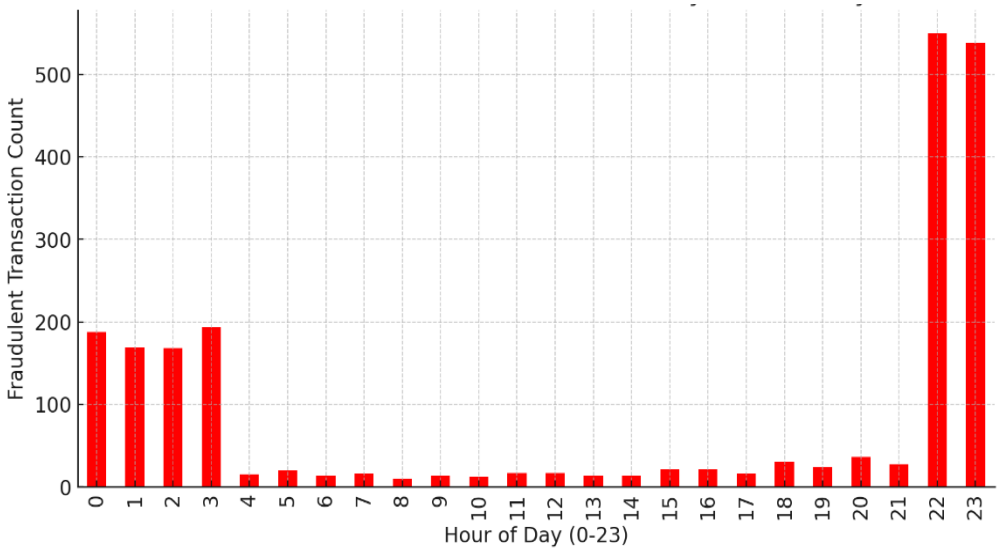

Figure 8: Number of Fraudulent Transactions by Hour of Day.

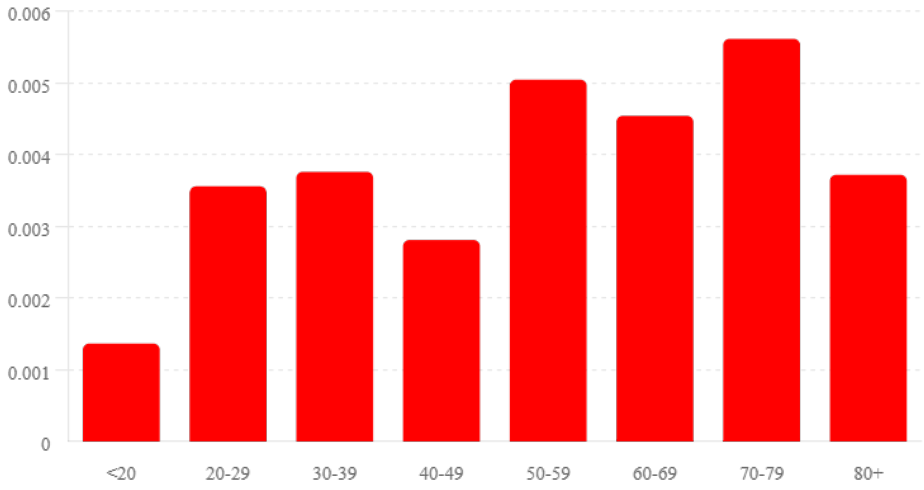

Figure 9: Distribution of age group and fraud status.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Figure 9 shows the analysis of fraudulent transactions across age groups revealed clear demographic patterns. Fraud rates were lowest among individuals under 20 years of age, accounting for only 0.14% of their total transactions. In contrast, older populations exhibited markedly higher vulnerability, with the highest fraud rates observed in the 70–79 age group (0.56%), followed by those aged 50–59 (0.50%) and 60–69 (0.45%). In terms of absolute numbers, the 30–39 and 50–59 groups recorded the largest volumes of fraudulent cases (479 and 425, respectively), reflecting both their higher transaction activity and moderate fraud susceptibility. By comparison, the youngest (<20 years) and oldest (80+) segments reported fewer total fraud cases, partly due to smaller transaction volumes. These findings suggest that while middle-aged individuals contribute significantly to the absolute number of fraudulent cases, older adults face a disproportionately higher fraud risk relative to their transaction activity.

Figure 10 shows the analysis of fraud status by gender reveals that fraudulent transactions are distributed almost evenly between males and females. Females recorded 1,164 fraud cases out of 304,886 total transactions, while males recorded 981 fraud cases out of 250,833 transactions. Although the absolute number of fraudulent transactions is slightly higher among females, this difference corresponds to their larger overall transaction volume. When expressed as a fraud rate, females show 0.38 percent fraudulent transactions, and males show 0.39 percent. This negligible difference indicates that gender does not play a significant role in determining fraud susceptibility within the dataset.

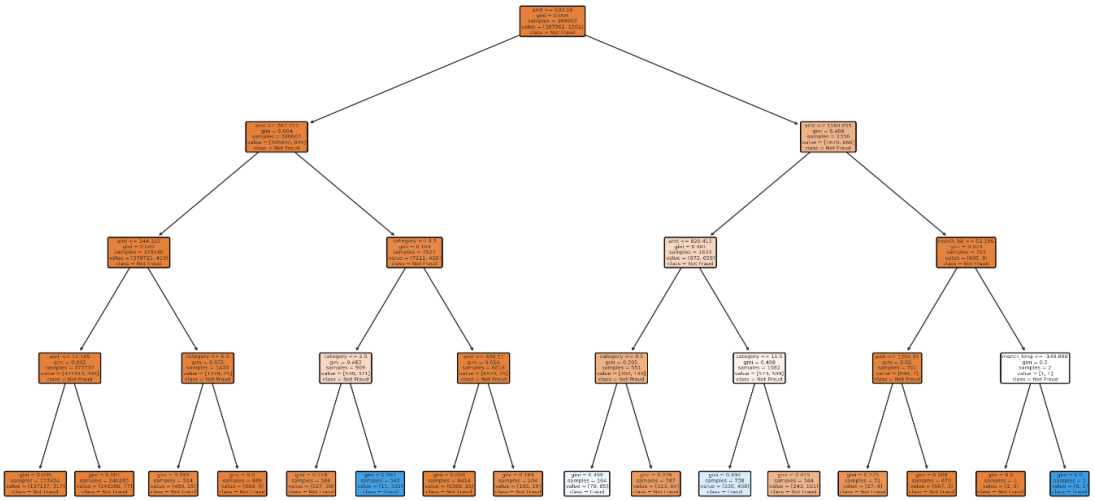

Figure 10: Fraud status by gender.

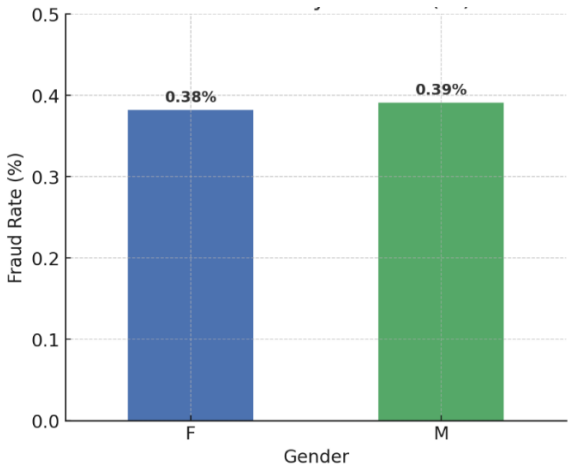

Figure 11: Decision Tree Model.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 3: Confusion matrix of prediction using logistics regression.

Prediction	*Not Fraud*	*Fraud*
Not Fraud	158268	7804
Fraud	150	494

The Logistic Regression model (Table 3) achieved a reasonable recall (76.7%), indicating that it was able to capture a majority of fraudulent transactions. However, this came at the cost of precision (5.95%), meaning that the vast majority of flagged transactions were actually false positives. This imbalance leads to a relatively low F1 score, reflecting poor trade-off between precision and recall. The overall accuracy (95.3%) and specificity (95.3%) appear high but are inflated by the overwhelming number of legitimate transactions in the dataset. From a fraud detection standpoint, while the model successfully identifies most fraud cases, the high false alarm rate reduces its practical utility, as it would burden investigators with a large volume of false alerts.

Table 4: Confusion matrix of prediction using Random Forest.

Prediction	*Not Fraud*	*Fraud*
Not Fraud	165992	80
Fraud	353	291

The Random Forest model achieved (Table 4) very high accuracy (99.74%) and excellent specificity (99.95%), confirming its strong performance in correctly identifying legitimate transactions. Precision was also relatively high (78.5%), meaning that most transactions classified as fraudulent were indeed fraud. However, the recall rate (45.2%) indicates that more than half of fraudulent cases were missed, limiting the model's overall effectiveness in fraud detection. The F1 score (0.58) reflects this imbalance, showing moderate trade-off between precision and recall. In practice, Random Forest reduces false positives and ensures reliable fraud alerts, but its low sensitivity highlights the need for enhancements, such as threshold adjustments or ensemble combinations, to improve the detection rate of fraud cases.

Table 5: Confusion matrix of prediction using Decision Tree.

Prediction	*Not Fraud*	*Fraud*
Not Fraud	165966	106
Fraud	335	309

The Decision Tree model achieved (Table 5) a high overall accuracy (99.73%) and near-perfect specificity (99.94%), indicating strong reliability in classifying legitimate transactions. Precision (74.5%) was also high, meaning that the majority of flagged fraud cases were indeed fraudulent. However, recall was relatively low (48.0%), suggesting that more than half of actual fraudulent cases were missed. This trade-off is reflected in the F1 score (0.58), which highlights limited balance between precision and recall. From a fraud detection perspective, the model is conservative, producing few false alarms but at the expense of missing a substantial number of fraudulent transactions. While this makes it efficient in reducing investigation workload, its sensitivity requires improvement for deployment in high-stakes financial environments where detecting fraud is critical.

Table 6: Confusion matrix of prediction using KNN.

Prediction	*Not Fraud*	*Fraud*
Not Fraud	165960	112
Fraud	343	301

The KNN model achieved a high accuracy (99.73%) and strong specificity (99.93%), showing that it is highly effective at classifying legitimate transactions. Precision was satisfactory (72.9%), indicating that most of the fraud predictions corresponded to actual fraud cases. However, recall remained limited (46.7%), meaning that more than half of fraudulent transactions were overlooked. This imbalance is reflected in the F1 score (0.57), which suggests only moderate performance in balancing false positives and false negatives. While KNN offers reliable precision and very low false alarm rates, its relatively low sensitivity limits its effectiveness for fraud detection, where missing fraudulent activity can be costly.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 7: Confusion matrix of prediction using the XGBoost.

Prediction	*Not Fraud*	*Fraud*
Not Fraud	165994	78
Fraud	391	253

The XGBoost model achieved (Table 7) very high accuracy (99.72%) and excellent specificity (99.95%), demonstrating strong reliability in correctly classifying legitimate transactions. Its precision (76.4%) was strong, suggesting that most fraud predictions corresponded to true fraud cases. However, recall was notably low (39.3%), meaning that the majority of fraudulent transactions were missed. Consequently, the F1 score (0.52) reflects limited balance between precision and recall. While XGBoost minimizes false alarms and ensures reliable fraud predictions, its weak sensitivity limits its effectiveness in detecting fraud comprehensively. Improvements through threshold adjustment, feature engineering, or hybrid ensemble integration may help enhance recall while retaining precision.

Table 8: Confusion matrix of prediction using the Voting Classifier.

Prediction	*Not Fraud*	*Fraud*
Not Fraud	165942	130
Fraud	308	336

The ensemble Voting Classifier (Table 8) provided the most balanced results, with precision (72.1%) and recall (52.2%) yielding an F1 score (0.61) higher than any individual model, and an AUC of 0.91 reflecting strong discriminatory ability. While no single model achieved both high precision and high recall simultaneously, the Voting Classifier offered the best compromise between reducing false positives and capturing fraudulent transactions. Overall, the findings suggest that ensemble methods may provide more robust fraud detection performance compared to single classifiers, though improvements in recall remain essential for practical deployment in fraud-sensitive financial systems.

Table 9: Comparing the models' performances.

Model	Accuracy	Precision	Recall	Specificity	F1 Score
Logistic Regression	95.23	5.95	76.71	95.3	0.1105
Decision Tree	99.74	74.46	47.98	99.94	0.5836
Random Forest	99.74	78.44	45.19	99.95	0.5734
XGBoost	99.72	76.44	39.29	99.95	0.519
KNN	99.73	72.88	46.74	99.93	0.5695
Voting Classifier	99.74	72.1	52.17	99.92	0.6054

(Table 9) Logistic Regression achieved the highest recall (76.71%), meaning it detected most fraud cases, but its precision was very low (5.95%), leading to many false alarms. Random Forest and XGBoost delivered strong precision (around 76–78%) and excellent AUC scores, but their recall was weak (39–45%), so many fraud cases were missed. KNN offered moderate balance but did not outperform the ensembles. The Voting Classifier achieved the best overall trade-off, with good precision (72.1%), improved recall (52.17%), and the highest F1 score (0.6054), making it the most effective and reliable model for fraud detection.

This ROC curve comparison (Figure 12) shows that XGBoost (AUC = 0.986) and Random Forest (AUC = 0.959) achieved the best overall performance in distinguishing fraud from non-fraud. The Decision Tree (AUC = 0.914) and Voting Classifier (AUC = 0.913) also performed well, while KNN (AUC = 0.847) and especially Logistic Regression (AUC = 0.837) lagged behind. Overall, ensemble and tree-based models demonstrated stronger discriminatory power than simpler models.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

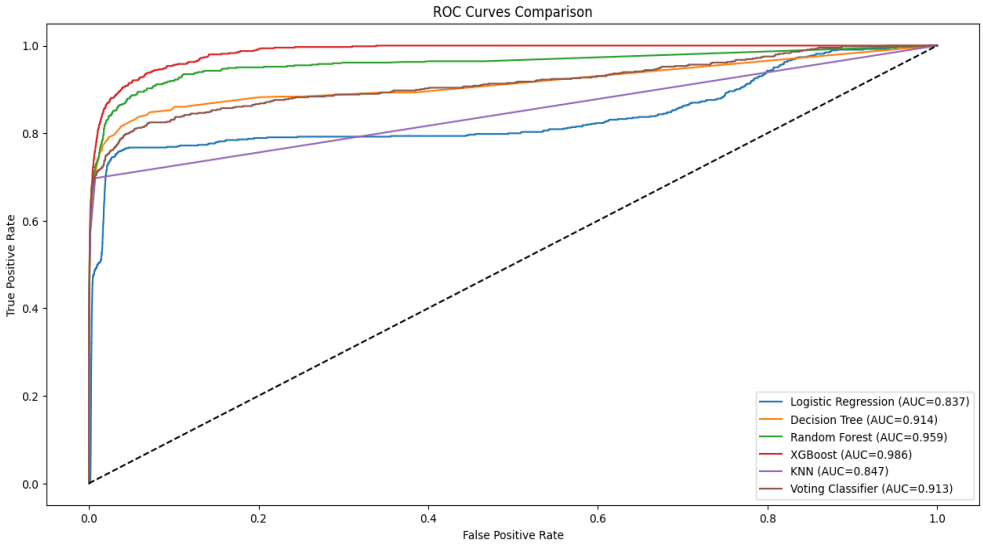

Figure 12. ROC curve comparison of six models.

5. Discussion

The comparative evaluation of six machine learning models for fraud detection demonstrates that while overall accuracy and specificity were consistently high, substantial trade-offs exist between precision and recall. Logistic Regression achieved the highest recall (76.71%), capturing the majority of fraudulent transactions, but its precision was extremely low (5.95%), resulting in a large number of false positives. This outcome aligns with prior research that identified Logistic Regression as sensitive but prone to over-flagging in imbalanced datasets [41].

Tree-based models, including Decision Tree, Random Forest, and XGBoost, performed strongly in terms of precision (74–78%) and specificity (>99.9%), ensuring that flagged fraud cases were highly reliable. However, recall values for these models (39–48%) were substantially lower, meaning many fraudulent cases were missed. Similar findings have been reported in earlier studies, where ensemble tree-based methods showed high discriminatory power but struggled with sensitivity due to class imbalance [41], [42]. Notably, XGBoost achieved the highest AUC (0.9859), confirming its superior ability to separate fraudulent and legitimate classes, consistent with previous work highlighting its robustness in fraud detection and imbalanced classification problems [43].

KNN delivered moderately balanced results, with precision of 72.88% and recall of 46.74%, but it did not surpass ensemble methods. Prior literature has noted that distance-based models like KNN are less effective when feature distributions between fraud and non-fraud overlap significantly [44].

The Voting Classifier achieved the most balanced results, with precision of 72.1%, recall of 52.17%, and the highest F1 score (0.6054). Its performance underscores the effectiveness of ensemble strategies in leveraging the strengths of multiple base classifiers to mitigate the precision–recall trade-off. Previous research has similarly emphasized that hybrid or ensemble approaches outperform single classifiers in fraud detection tasks [45]. Furthermore, the ROC and Precision–Recall curve analyses reinforce these findings. While XGBoost and Random Forest achieved superior ROC-AUC values, the Voting Classifier demonstrated stronger PR-AUC performance, confirming its suitability for highly imbalanced fraud detection scenarios where precision–recall trade-offs are more informative [45].

From a practical perspective, these results highlight the need to balance false positives and false negatives in fraud detection systems. Excessive false positives, as observed with Logistic Regression, increase operational costs and reduce system efficiency, while high false negatives, as observed in tree-based models, pose greater financial and reputational risks. The Voting Classifier provides the most effective balance, but improvements in recall remain critical for real-world deployment. Future work should explore cost-sensitive learning [46], threshold optimization [44], and hybrid ensemble frameworks to further enhance sensitivity while maintaining strong precision.

6. Conclusion

This study evaluated six machine learning models—Logistic Regression, Decision Tree, Random Forest, XGBoost, KNN, and a Voting Classifier—for fraud detection under conditions of extreme class imbalance. The results demonstrate that while all models achieved high accuracy and specificity, significant trade-offs exist between precision and recall. Logistic Regression provided the highest recall, consistent with prior studies that highlight its sensitivity on imbalanced datasets [33], but at the expense of precision, producing excessive false positives. In contrast, Random Forest and XGBoost achieved excellent precision and the strongest ROC-AUC values, in line with earlier findings on the robustness of ensemble tree-based approaches [34], [35], but suffered from low recall, leaving many fraud cases undetected. The Voting Classifier emerged as the most balanced model, achieving the highest F1 score and competitive PR-AUC, thereby offering the best compromise between detecting fraud and minimizing false alarms. This supports previous research emphasizing the advantage of ensemble frameworks in fraud detection tasks [42], [43]. From a practical standpoint, the findings underline the need to prioritize recall, as undetected fraud typically imposes greater financial and reputational costs than false positives [45]. Future work should focus on enhancing recall through cost-sensitive learning [46], threshold optimization, and hybrid ensemble strategies [44], which have been shown to strengthen model sensitivity without severely compromising precision. In conclusion, while no single model dominates across all metrics, ensemble-based approaches, particularly the Voting Classifier, provide the most reliable balance and hold strong potential for real-world deployment in fraud detection systems.

Future Work

Future research should focus on improving recall through cost-sensitive learning and threshold tuning, while keeping false positives manageable. Hybrid ensemble and deep learning methods can help capture more complex fraud patterns. Incorporating temporal modelling and real-time systems will make detection more adaptive, and explainable AI will enhance trust and compliance.

Reference

[1] C. Jiang and D. Broby, “Mitigating cybersecurity challenges in the financial sector with artificial intelligence,” 2021, Accessed: Nov. 24, 2024. [Online]. Available: https://pure.ulster.ac.uk/files/98691946/Jiang_Broby_CeFRI_2021_Mitigating_cybersecurity_challenges_in_the_financial_sector_with_Artificial_Intelligence.pdf

[2] S. Morgan, “2021 Report: Cyberwarfare In The C-Suite,” *Cybercrime Facts and Statistics*, 2021. Accessed: Nov. 24, 2024. [Online]. Available: https://cybersecurityventures.com/cybercrime-damages-6-trillion-by-2021/

[3] S. Morgan, “Global Cybersecurity Spending Predicted To Exceed $1 Trillion From 2017-2021,” *Cybercrime Magazine*, 2019, Accessed: Nov. 24, 2024. [Online]. Available: https://cybersecurityventures.com/cybersecurity-market-report/

[4] E. Btoush, X. Zhou, R. Gururajan, K. C. Chan, and X. H. Tao, “A Survey on Credit Card Fraud Detection Techniques in Banking Industry for Cyber Security,” in *Proceedings of 2021 8th IEEE International Conference on Behavioural and Social Computing, BESC 2021*, 2021. doi: 10.1109/BESC53957.2021.9635559.

[5] B. Al Smadi and M. Min, “A critical review of credit card fraud detection techniques,” in *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, IEEE, 2020, pp. 732–736.

[6] N. K. Trivedi, S. Simaiya, U. K. Lilhore, and S. K. Sharma, “An efficient credit card fraud detection model based on machine learning methods,” *International Journal of Advanced Science and Technology*, vol. 29, no. 5, 2020.

[7] I. Benchaji, S. Douzi, and B. El Ouahidi, “Credit card fraud detection model based on LSTM recurrent neural networks,” *Journal of Advances in Information Technology*, vol. 12, no. 2, 2021, doi: 10.12720/jait.12.2.113-118.

[8] E. M. H. Al Rubaie, “Improvement in credit card fraud detection using ensemble classification technique and user data,” *International Journal of Nonlinear Analysis and Applications*, vol. 12, no. 2, 2021, doi: 10.22075/IJNAA.2021.5228.

[9] G. J. Priya and S. Saradha, “Fraud detection and prevention using machine learning algorithms: A review,” in *Proceedings of the 7th International Conference on Electrical Energy Systems, ICEES 2021*, 2021. doi: 10.1109/ICEES51510.2021.9383631.

[10] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, “Data mining for credit card fraud: A comparative study,” *Decis Support Syst*, vol. 50, no. 3, pp. 602–613, 2011.

[11] S. Jha, M. Guillen, and J. C. Westland, “Employing transaction aggregation strategy to detect credit card fraud,” *Expert Syst Appl*, vol. 39, no. 16, pp. 12650–12657, 2012.

[12] A. Dal Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, "Calibrating probability with undersampling for unbalanced classification," in *2015 IEEE symposium series on computational intelligence*, IEEE, 2015, pp. 159–166.

[13] I. D. Mienye and Y. Sun, "A machine learning method with hybrid feature selection for improved credit card fraud detection," *Applied Sciences*, vol. 13, no. 12, p. 7254, 2023.

[14] O. J. Unogwu and Y. Filali, "Fraud detection and identification in credit card based on machine learning techniques," *Wasit Journal of Computer and Mathematics Science*, vol. 2, no. 3, pp. 16–22, 2023.

[15] A. Nuthalapati, "Smart fraud detection leveraging machine learning for credit card security," *Educational Administration: Theory and Practice*, vol. 29, no. 2, pp. 433–443, 2023.

[16] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[17] J. Chen, Z. Wu, and J. Zhang, "Driving safety risk prediction using cost-sensitive with nonnegativity-constrained autoencoders based on imbalanced naturalistic driving data," *IEEE transactions on intelligent transportation systems*, vol. 20, no. 12, pp. 4450–4465, 2019.

[18] A. Dal Pozzolo, G. Boracchi, O. Caelen, C. Alippi, and G. Bontempi, "Credit card fraud detection: a realistic modeling and a novel learning strategy," *IEEE Trans Neural Netw Learn Syst*, vol. 29, no. 8, pp. 3784–3797, 2017.

[19] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *Journal of applied science and technology trends*, vol. 2, no. 01, pp. 20–28, 2021.

[20] J. F. Smith III, "Evolving fuzzy decision tree structure that adapts in real-time," in *Proceedings of the 7th annual conference on Genetic and evolutionary computation*, 2005, pp. 1737–1744.

[21] K. Ayorinde, *A methodology for detecting credit card fraud*. Minnesota State University, Mankato, 2021.

[22] P. H. Swain and H. Hauska, "The decision tree classifier: Design and potential," *IEEE transactions on geoscience electronics*, vol. 15, no. 3, pp. 142–147, 1977.

[23] P. H. Swain and H. Hauska, "The decision tree classifier: Design and potential," *IEEE transactions on geoscience electronics*, vol. 15, no. 3, pp. 142–147, 1977.

[24] S. Tangirala, "Evaluating the impact of GINI index and information gain on classification using decision tree classifier algorithm," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 2, pp. 612–619, 2020.

[25] Y. Liu, L. Hu, F. Yan, and B. Zhang, "Information gain with weight based decision tree for the employment forecasting of undergraduates," in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, IEEE, 2013, pp. 2210–2213.

[26] P. Malik, A. Chourasia, R. Pandit, S. Bawane, and J. Surana, "Credit risk assessment and fraud detection in financial transactions using machine learning," *Journal of Electrical Systems*, vol. 20, no. 3s, pp. 2061–2069, 2024.

[27] T. R. Prajwala, "A comparative study on decision tree and random forest using R tool," *International journal of advanced research in computer and communication engineering*, vol. 4, no. 1, pp. 196–199, 2015.

[28] E. Kabir, S. Guikema, and B. Kane, "Statistical modeling of tree failures during storms," *Reliab Eng Syst Saf*, vol. 177, pp. 68–79, 2018.

[29] J. L. Speiser, "A random forest method with feature selection for developing medical prediction models with clustered and longitudinal data," *J Biomed Inform*, vol. 117, p. 103763, 2021.

[30] N. Donges, "A complete guide to the random forest algorithm," *Built in*, vol. 16, 2019.

[31] T. R. Noviandy, G. M. Idroes, A. Maulana, I. Hardi, E. S. Ringga, and R. Idroes, "Credit card fraud detection for contemporary financial management using XGBoost-driven machine learning and data augmentation techniques," *Indatu Journal of Management and Accounting*, vol. 1, no. 1, pp. 29–35, 2023.

[32] J. Li, J. Zhang, J. Zhang, and S. Zhang, "Quantum KNN classification with K value selection and neighbor selection," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 43, no. 5, pp. 1332–1345, 2023.

[33] R. Chhabra, S. Goswami, and R. K. Ranjan, "A voting ensemble machine learning based credit card fraud detection using highly imbalance data," *Multimed Tools Appl*, vol. 83, no. 18, pp. 54729–54753, 2024.

[34] M. A. Mim, N. Majadi, and P. Mazumder, "A soft voting ensemble learning approach for credit card fraud detection," *Heliyon*, vol. 10, no. 3, 2024.

Research Article

[35] Y. Zhang, K. Zhou, and Z. Liu, "What makes good examples for visual in-context learning?," *Adv Neural Inf Process Syst*, vol. 36, pp. 17773–17794, 2023.

[36] D. Chicco and G. Jurman, "The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification," *BioData Min*, vol. 16, no. 1, p. 4, 2023.

[37] S. Bagga, A. Goyal, N. Gupta, and A. Goyal, "Credit card fraud detection using pipeling and ensemble learning," *Procedia Comput Sci*, vol. 173, pp. 104–112, 2020.

[38] M.-Y. Chen, "Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches," *Computers & Mathematics with Applications*, vol. 62, no. 12, pp. 4514–4524, 2011.

[39] N. Rtayli and N. Enneya, "Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization," *Journal of Information Security and Applications*, vol. 55, p. 102596, 2020.

[40] S. Tyagi and S. Mittal, "Sampling approaches for imbalanced data classification problem in machine learning," in *Proceedings of ICRIC 2019: Recent innovations in computing*, Springer, 2019, pp. 209–221.

[41] Y. Lin and M. M. L. Cahigas, "An Analysis of the Perspective Road Design Scheme Around Zhangzhou Olympic Sports Center," *Procedia Comput Sci*, vol. 234, pp. 1076–1086, 2024, doi: 10.1016/j.procs.2024.03.102.

[42] K. Wang, "Efficient Financial Fraud Detection: An Empirical Study using Ensemble Learning and Logistic Regression," in *2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS)*, IEEE, 2024, pp. 859–864.

[43] M. A. Mim, N. Majadi, and P. Mazumder, "A soft voting ensemble learning approach for credit card fraud detection," *Heliyon*, vol. 10, no. 3, 2024.

[44] Q. Zheng, C. Yu, J. Cao, Y. Xu, Q. Xing, and Y. Jin, "Advanced payment security system: xgboost, lightgbm and smote integrated," in *2024 IEEE International Conference on Metaverse Computing, Networking, and Applications (MetaCom)*, IEEE, 2024, pp. 336–342.

[45] M. Isangediok and K. Gajamannage, "Fraud detection using optimized machine learning tools under imbalance classes," in *2022 IEEE International Conference on Big Data (Big Data)*, IEEE, 2022, pp. 4275–4284.

[46] S. D. Penmetsa and S. Mohammed, "Ensemble Techniques for Credit Card Fraud Detection," *International Journal of Smart Business and Technology*, vol. 9, no. 2, pp. 33–48, 2021.