Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 8, 2025

SECURING DIGITAL TRANSACTIONS: MACHINE LEARNING FRAMEWORKS FOR FRAUD DETECTION IN PAYMENT SYSTEMS

Muhammad Inam ul Haq[*1], Muhammad Zarar[2], Shehryar Qamar Paracha[3], Ahmad Nawaz Shah[4], Muhammad Hamza[5], Warda Hussaini[6], Muhammad Sohail[7], Muhammad Awais[8]

[*1,3,7,8]*University name: The Islamia University of Bahawalpur, Pakistan Department: Department of Electronic Engineering*
[2]*University name: Tianjin University, China Department: College of Intelligent & Computing*
[4]*Pir Mehar Ali Shah Arid Agriculture University, Rawalpindi Department of Computer Science*
[5]*PhD scholar HITEC University Taxila*
[6]*University name: University of Karachi Department: Department of Public Administration*

[*1]engr.m.inamulhaq753@gmail.com , [2]zararkhan@tju.edu.cn , [3]shehryar.paracha@iub.edu.pk, [4]ahmadnawazshah511@gmail.com, [6]hussainiwarda57@gmail.com, [7]muhammad.sohail.ee@gmail.com, [8]choudhryawais268@gmail.com

DOI: https://doi.org/10.5281/zenodo.16976607

Keywords

Article History
Received: 26 May, 2025
Accepted: 07 August, 2025
Published: 26 August, 2025

Corresponding Author: *
Muhammad Inam ul Haq

Abstract

The widespread adoption of digital payment systems has revolutionized financial transactions, offering seamless and efficient services globally. However, this advancement has heightened cybersecurity risks, with fraudulent transactions posing a significant threat. These fraudulent activities result in financial losses and undermine trust in digital platforms, necessitating robust detection mechanisms. This paper proposes a machine learning (ML)-based framework for detecting fraudulent transactions in online payment systems. Seven models, including Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Trees (DT), Random Forest (RF), Extreme Gradient Boosting (XGB), Naive Bayes (NB), and Support Vector Classifier (SVC), were evaluated. To address the challenge of imbalanced datasets, over- sampling and under sampling techniques were applied, enhancing model sensitivity to minority-class samples. Experimental results revealed that ensemble models, particularly RF and XGB, achieved the highest accuracy of 99.3% and 99.4%, respectively. These models outperformed simpler classifiers such as LR and NB in key metrics, including precision, recall, and F1-score. This research highlights the potential of ML to strengthen cybersecurity in digital payment systems. By addressing challenges such as data imbalance and scalability, the proposed framework provides actionable insights for developing next-generation fraud detection systems, enhancing trust and security in the digital economy.

Index Terms–Fraud Detection, ML, Cybersecurity, Over- sampling, Under sampling, Ensemble Models, Digital Payment Systems..

INTRODUCTION

The proliferation of digital payment systems has transformed the financial landscape, enabling seamless and efficient transactions globally. However, this digital revolution has also ushered in significant cybersecurity challenges, particularly in the realm of fraud detection [1]. Cybersecurity threats, such as phishing, account takeovers, and unauthorized access, have become increasingly sophisticated, exploiting vulnerabilities in financial systems to cause substantial financial and reputational damage to individuals and organizations alike [2], [3].

Fraud detection in online payment systems represents a critical pillar of cybersecurity, aiming to identify and mitigate unauthorized transactions in real time. Traditional rule-based systems, which rely on static thresholds and predefined pat- terns, have historically been employed for fraud prevention. While effective in detecting known fraud scenarios, these systems lack the adaptability to identify emerging, dynamic threats [2]. Moreover, such systems often generate a high number of false positives, leading to inefficiencies and increased operational costs for fraud analysts [4].

ML has emerged as a powerful alternative, leveraging data- driven approaches to improve fraud detection accuracy and scalability. Unlike rule-based systems, ML models can learn complex patterns from historical transaction data, enabling them to generalize and adapt to novel fraud tactics [5]. Recent studies have demonstrated the efficacy of supervised and unsupervised ML models, as well as deep learning frameworks, in identifying fraudulent behavior with high precision and recall [6]. These advancements have opened new avenues for integrating artificial intelligence (AI) into cybersecurity measures, addressing the limitations of traditional approaches and providing robust, scalable solutions.

However, a significant challenge in fraud detection is the inherent imbalance in transaction datasets, where fraudulent transactions constitute a small fraction of the overall data. This imbalance can lead to biased model predictions, prioritizing the majority class (legitimate transactions) at the expense of minority class detection (fraudulent transactions). Address- ing this issue requires the implementation of data-balancing techniques, such as oversampling and under sampling, which enhance model sensitivity to fraudulent activities by ensuring a more equitable distribution of classes during training [7].

The primary contributions of this paper are as follows:

- We propose a novel framework for fraud detection in on- line payment systems, leveraging ML models to enhance cybersecurity measures.
- We address the issue of class imbalance in fraud datasets by implementing and evaluating data-balancing techniques, including oversampling and under sampling.
- We conduct a comprehensive evaluation of multiple ML algorithms, comparing their performance in terms of accuracy, precision, recall, F1-score, and ROC-AUC.
- We analyze the implications of integrating these models into real-world cybersecurity systems, highlighting their potential to reduce false positives and strengthen trust in digital payment platforms.

The remainder of this paper is organized as follows. Section 2 discusses the current cybersecurity landscape and fraud detection frameworks, providing an overview of existing challenges and advancements in the field. Section 3 presents the methodology, including data preprocessing, exploration data analysis, and the development of ML models. Section 4 evaluates the experimental results, comparing the performance of various algorithms and discussing their implications for cybersecurity. Finally, Section 5 concludes with a summary of findings and outlines potential directions for future research.

II. CYBERSECURITY LANDSCAPE AND FRAUD DETECTION FRAMEWORKS

The exponential growth of digital transactions has been accompanied by a proportional rise in cybersecurity threats, particularly in the financial sector. Fraudulent activities such as unauthorized access, phishing, and account takeovers are becoming increasingly sophisticated, necessitating the deployment of robust fraud detection mechanisms. Over the years, researchers have explored various methodologies to combat these challenges, ranging from rule-based systems to advanced AI models.

Khurana [8] analyzed the application of predictive AI in e- commerce payment systems, focusing on real-

time transaction security. The study introduced predictive analytics to monitor transaction behavior and flag suspicious activities. However, the reliance on predefined rules and static thresholds rendered the system ineffective against emerging and dynamic fraud schemes. Similarly, Chang et al. [9] proposed a hybrid approach combining DT, LR, and anomaly detection techniques. While their model achieved an accuracy of 92%, it struggled with datasets exhibiting high variability, leading to a reduction in generalizability.

Deep learning techniques have also been extensively explored for fraud detection. Udayakumar et al. [6] introduced the "Deep Fraud Net" framework, utilizing deep neural net- works for the detection and classification of financial fraud. The framework demonstrated high precision and sensitivity but was constrained by overfitting issues in datasets with unbalanced class distributions. Similarly, Obeng et al. [10] explored ensemble learning techniques, including RF and gradient boosting, for fraud detection. While ensemble mod- els improved detection rates, their computational complexity limited real-time applicability in large-scale systems.

ML techniques such as clustering and outlier detection have shown promise in detecting fraudulent activities. For instance, Aschi et al. [2] investigated the use of unsupervised models like k-means clustering to detect anomalous patterns in financial transactions. Despite their success in identifying novel fraud schemes, unsupervised methods often suffer from high false positive rates, which can overwhelm security analysts and increase operational costs. Kumar et al. [7] addressed these limitations by integrating cost-sensitive learning to handle imbalanced datasets, though the technique exhibited reduced performance on smaller datasets.

Anomaly detection frameworks, such as those proposed by Agrawal [3], have also been applied to payment security systems. The study examined density-based methods like DBSCAN and isolation forests for identifying outliers. While effective for static datasets, these methods lacked adaptability to evolving fraud patterns in real-time systems. Additionally, Priya and Saradha [11] conducted a comprehensive review of ML algorithms for fraud detection, highlighting the limitations of traditional algorithms in addressing the dynamic nature of cybersecurity threats.

Despite these advancements, several research gaps remain unaddressed. First, most studies focus on improving detection accuracy but neglect the operational impact of high false positive rates, which can erode trust in automated systems. Second, the reliance on static models limits the adaptability of fraud detection systems to emerging threats. Finally, the issue of dataset imbalance continues to hinder the effectiveness of ML models, as demonstrated by numerous studies [4], [12]. Addressing these gaps requires a holistic approach that integrates advanced ML techniques, robust data-balancing mechanisms, and scalable architectures to enhance real-time fraud detection capabilities.

This paper builds on existing work by addressing these challenges through the implementation of advanced ML models that combine supervised and unsupervised learning approaches. By evaluating a variety of algorithms and integrating balancing techniques, we aim to bridge the gap between detection accuracy and operational efficiency in real-world cybersecurity applications.

III. METHODOLOGY

This section elaborates on the proposed methodology, lever- aging advanced ML models and balancing techniques for on- line payment fraud detection. Figures derived from the dataset

[13] provide critical insights into the nature and distribution of the data. We employ complex mathematical notations and equations to formalize the methodology, ensuring a robust and systematic approach to fraud detection.

A. Dataset Analysis and Feature Distribution

The dataset comprises transaction records across various payment types, as shown in Figure 1. It is evident that CASH_OUT and PAYMENT are the most common transaction types, followed by CASH_IN. Minor classes such as TRANSFER and DEBIT exhibit significantly fewer samples, leading to a class imbalance that poses challenges for ML algorithms.

The fraudulent transactions, constituting a mere 0.1% of the total dataset, are visualized using a donut chart in Figure 2.

B. Data Balancing Techniques

Figure 3 illustrates under sampling and oversampling strategies. Under sampling reduces the majority class size by sampling, while oversampling creates synthetic data for the minority class:

$$\mathbf{x}_{\text{synthetic}} = \mathbf{x}_{\text{minority}} + \lambda(\mathbf{x}_{\text{neighbor}} - \mathbf{x}_{\text{minority}}), \quad \lambda \sim U(0, 1).$$

Here, $\mathbf{x}_{\text{minority}}$ is a randomly chosen minority class sample, and $\mathbf{x}_{\text{neighbor}}$ is one of its KNN. These methods address the imbalance ratio:

$$\text{IR} = \frac{|\{i : y_i = 0\}|}{|\{i : y_i = 1\}|}.$$

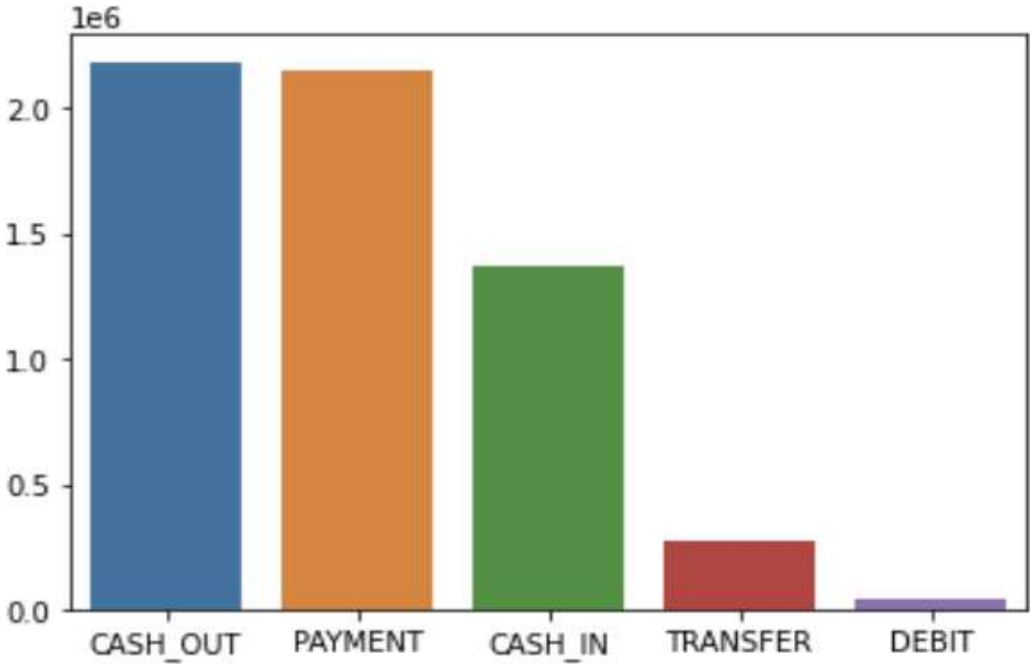

Fig. 1: Distribution of transaction types in the dataset.

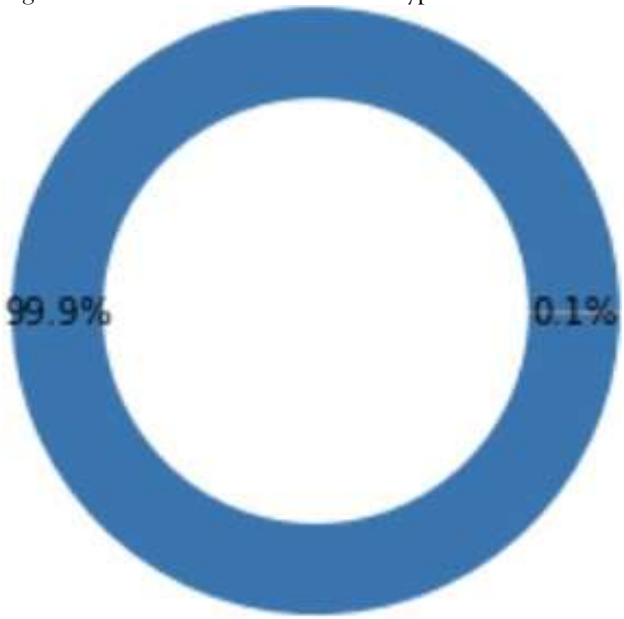

Fig. 2: Fraud vs. Legitimate transaction ratio.

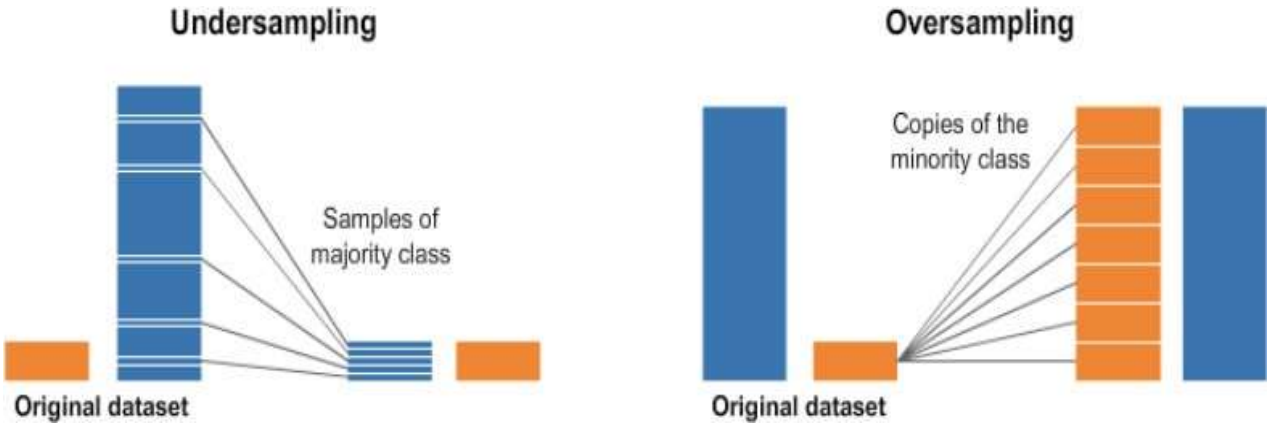

Fig. 3: Under sampling vs. Oversampling techniques.

C. Correlation Analysis

Figure 4 shows the correlation matrix, highlighting relationships among features. Strong correlations between oldbalanceOrg, newbalanceOrg, and amount suggest potential feature engineering opportunities. The correlation matrix is defined as:

$$R = \frac{X^T X}{n}$$

where R is the correlation matrix, X is the feature matrix, and n is the number of samples.

D. Algorithm for Fraud Detection

The steps of the proposed fraud detection framework are summarized in Algorithm 1. Mathematical rigor ensures re- producibility and scalability.

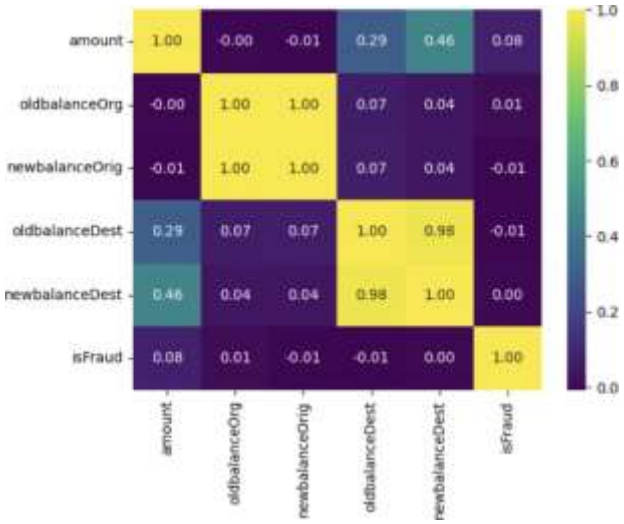

Fig. 4: Correlation matrix of transaction features.

Algorithm 1: Proposed Fraud Detection Framework

Input: Transaction dataset $D = \{(x_i, y_i)\}_{i=1}^{n}$ threshold τ model f.

Step 1: Data Preprocessing

Normalize features as follows: $x_i \leftarrow \frac{x_i - \mu}{\sigma}$ where μ and σ represent the feature means and standard deviations, respectively.

Apply sampling techniques to balance the dataset:

- **Oversampling:** Replicate or generate synthetic samples of the minority class.
- **Under sampling:** Randomly remove samples from the majority class.

Step 2: Data Balancing

Step 3: Model Training

Train f(x) on the balanced dataset using the binary cross-entropy loss function:

$$L = \frac{1}{n}\sum_{i=1}^{n}[-y_i \, log\,(f(x_i)) - (1-y_i)\, log\,(1-f(x_i))]$$

Step 4: Evaluation

Compute performance metrics:

$$Precision = \frac{TP}{TP+FP}, \qquad Recall = \frac{TP}{TP+FN}$$

where:

- TPTPTP = True Positives
- FPFPFP = False Positives
- FNFNFN = False Negatives

Step 5: Deployment

Deploy the optimized model f^* for real-time fraud detection.

E. Balanced Dataset Distribution

Figure 5 shows the target distribution after applying balancing techniques, ensuring an equal representation of classes for improved model training.

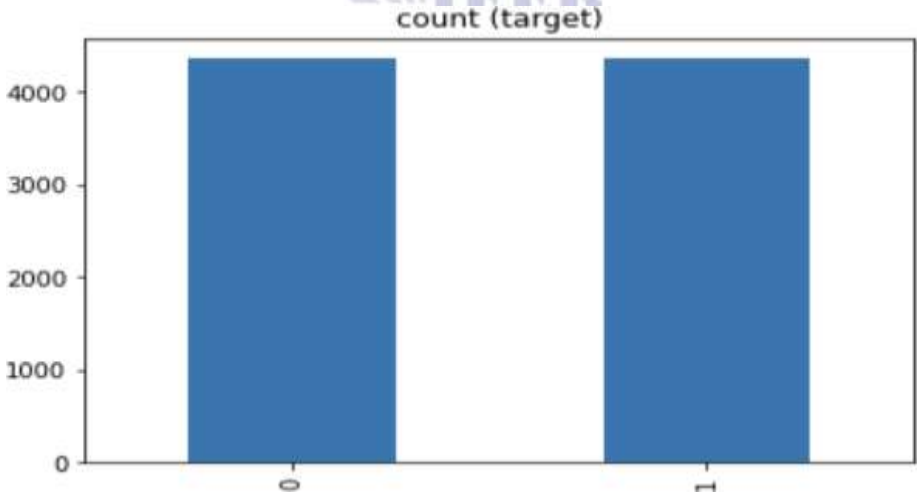

Fig. 5: Balanced target distribution post-data preprocessing.

F. ML Models and Training Approach

To develop a robust fraud detection system, we employ a diverse set of ML models. These models are selected based on their capability to handle binary classification problems and their complementary strengths. The following models are considered:

- **LR:** A linear model that predicts probabilities by applying a logistic function. It is effective for baseline comparisons and interpretable results.

- **KNN:** A non-parametric model that classifies data points based on their proximity to k nearest neighbors in feature space.
- **DT:** A tree-based model that partitions the feature space into regions by recursively splitting on feature values to minimize impurity.
- **RF:** An ensemble of DT trained on bootstrap samples, aggregating their predictions to improve accuracy and reduce overfitting.
- **XGB:** A gradient-boosting framework that optimizes an objective function iteratively, achieving high accuracy and computational efficiency.
- **NB:** A probabilistic classifier based on Bayes' theorem, assuming feature independence, which is computationally efficient for high-dimensional data.
- **SVC:** A discriminative classifier that finds the optimal hyperplane to separate classes in a high-dimensional space, using kernel functions to handle non-linearity.

1) **Training and Evaluation Approach:** To ensure a fair evaluation of the models, we adopt the following systematic training approach:

2)

a) Data Splitting: The dataset D is split into three subsets: training (80%), validation (10%), and test (10%) sets. Let Dtrain, Dval, and Dtest represent these subsets. Stratified sampling is applied to maintain the original class distribution across all subsets:

$D = D_{train} \cup D_{val} \cup D_{test}, \quad D_{train} \cap D_{val} = \emptyset.$

*b) **Hyperparameter Optimization:*** Each model is tuned to optimize its performance using grid search or randomized search on the validation set. Hyperparameters considered include:

- **LR**: Regularization strength $\lambda \in \{0.1, 1, 10\}$.
- **KNN:** Number of neighbors $k \in \{3, 5, 10\}$, distance metric $\in$ {Euclidean, Manhattan}.
- **DT:** Maximum depth $d \in \{5, 10, 20\}$, minimum samples per split $\in \{2, 5, 10\}$.
- **RF:** Number of estimators $n \in \{50, 100, 200\}$, maximum features $\in$ {sqrt, log2}.
- **XGB:** Learning rate $\eta \in \{0.01, 0.1, 0.3\}$, maximum depth $d \in \{3, 6, 9\}$.
- **NB:** No hyperparameters are tuned due to its simplicity.
- **SVC:** Regularization parameter $C \in \{0.1, 1, 10\}$, kernel type $\in$ {linear, rbf}.

c) Training: The models are trained on Dtrain using the balanced dataset. The loss function L is minimized during training:

$$L = \frac{1}{n}\sum_{i=1}^{n}[-y_i \log(f(x_i)) - (1-y_i)\log(1-f(x_i))]$$

where f (xi) is the predicted probability for the i-th instance.

d) Evaluation: Model performance is evaluated on Dtest using metrics such as precision, recall, F1-score, and ROC-AUC. The evaluation focuses on minimizing false positives while maintaining high recall to enhance cybersecurity measures.

$$F1 = \frac{2 . Precision. Recall}{Precision + Recall}$$

$$ROC\_AUC = \int_0^1 TPR(FPR)\, d(FPR)$$

where:

- TPR = True Positive Rate
- FPR = False Positive Rate

This systematic approach ensures a fair and comprehensive evaluation of models, facilitating the selection of the most effective classifier for real-time fraud detection.

IV. RESULTS AND ANALYSIS

This section presents the evaluation of the ML models using various performance metrics. The results are analyzed through tables, figures, and comparative insights to highlight the strengths and limitations of each

model in detecting fraudulent transactions.

A. Model Performance Overview

Table I summarizes the key performance metrics of eachML model, including training accuracy, testing accuracy, precision, recall, and F1-score. These metrics provide a comprehensive evaluation of the models' effectiveness in addressing the binary classification problem.

TABLE I: Performance Metrics of ML Models (Acc: Accuracy, Prec: Precision, Rec: Recall, F1: F1-Score)

Model	Train Acc	Test Acc	Prec	Rec	F1
LR	0.918	0.921	0.921	0.921	0.921
KNN	0.981	0.985	0.985	0.985	0.985
DT	1.000	0.992	0.992	0.992	0.992
RF	1.000	0.993	0.993	0.993	0.993
XGB	0.997	0.994	0.994	0.994	0.994
NB	0.578	0.566	0.566	0.566	0.566
SVC	0.956	0.966	0.966	0.966	0.966

B. Comparison of Metrics

Figure 6 presents a side-by-side comparison of precision, recall, and F1-scores across all models. The ensemble models, such as RF and XGB, consistently achieve higher scores, reflecting their ability to handle complex decision boundaries and imbalanced datasets.

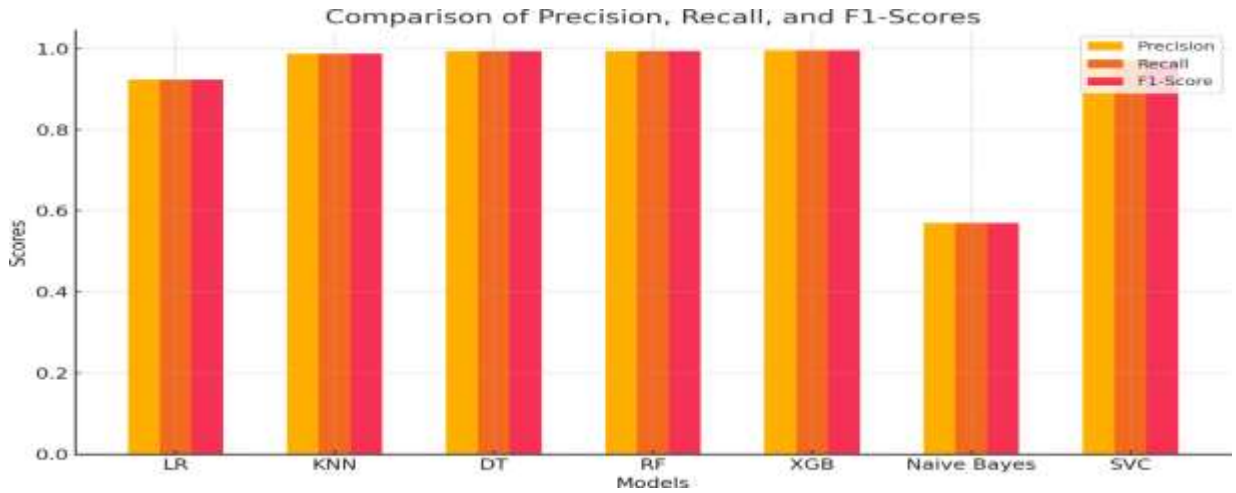

Fig. 6: Comparison of precision, recall, and F1-scores across models.

C. Training vs Testing Accuracy

The training and testing accuracies of the models are visualized in Figure 7. While most models demonstrate minimal overfitting, NB shows a significant performance drop from training to testing, indicating its limited capacity to generalize well on unseen data.

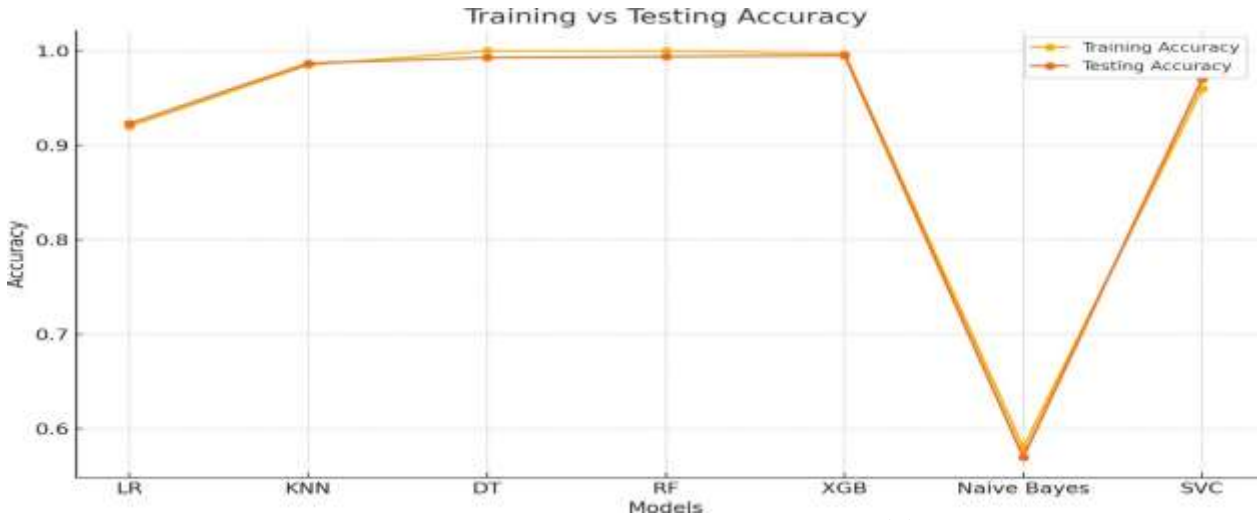

Fig. 7: Training vs Testing Accuracy Across Models.

D. Comparison of Correct and Incorrect Predictions

Figure 8 compares the number of correct and incorrect predictions (TP+TN vs. FP+FN) for each model. This visualization emphasizes the high accuracy of ensemble models such as RF and XGB.

E. F1-Score Trends

The distribution of F1-scores across models is highlighted in Figure 9. RF and XGB maintain the highest F1-scores, making them ideal candidates for cybersecurity applications where both precision and recall are critical.

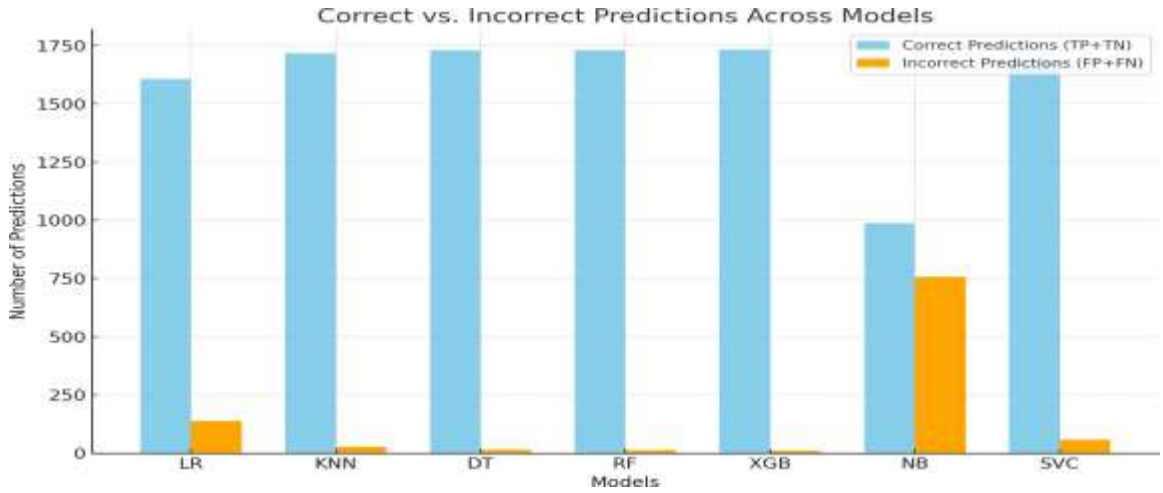

Fig. 8: Comparison of Correct vs. Incorrect Predictions Across Models.

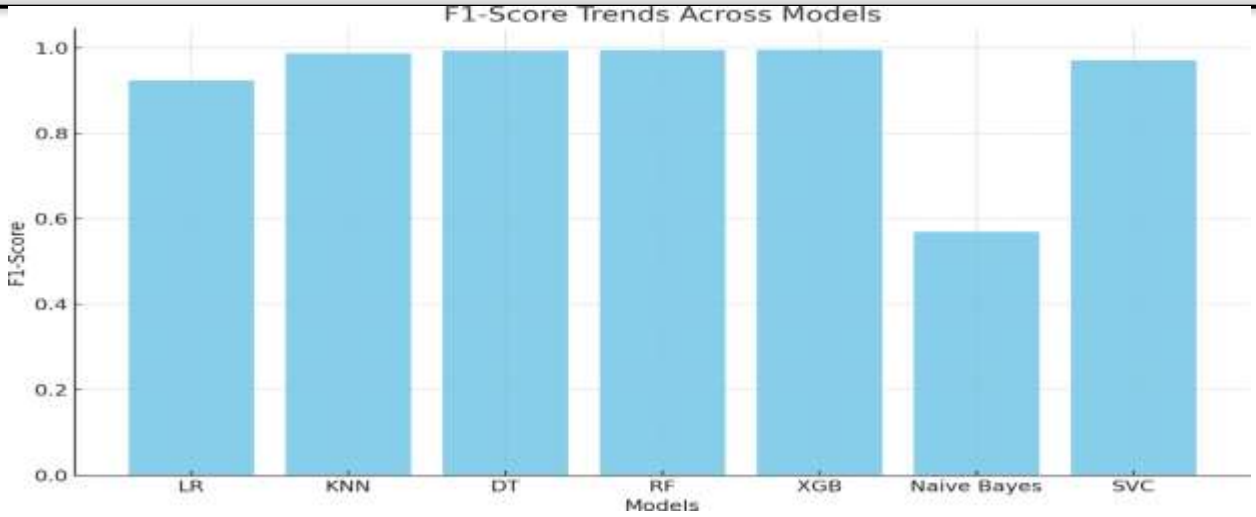

Fig. 9: F1-Score Trends Across Models.

F. Insights and Analysis

The analysis reveals the following key insights:

- **Ensemble models excel:** RF and XGB consistently out- perform other classifiers, demonstrating superior accuracy and generalization.
- **Simple models struggle:** LR and NB exhibit lower performance, particularly in recall and F1-score, due to their inability to model complex patterns effectively.
- **Balanced performance of SVC:** The SVC achieves high precision and recall, making it a suitable choice for applications requiring a balance between false positives and false negatives.
- **Generalization issues in NB:** The significant drop in NB' performance from training to testing indicates its sensitivity to data distributions and limited robustness.

The results underline the importance of selecting appropriate models based on the specific requirements of the cybersecurity framework, such as the need to minimize false alarms or maximize fraud detection rates.

V. CONCLUSION

This study explored the application of ML models for detecting fraudulent transactions in online payment systems, addressing critical challenges in cybersecurity. The experimental results demonstrated that ensemble models, such as RF and XGB, consistently outperform other classifiers, achieving superior accuracy, precision, recall, and F1-scores. These models effectively capture complex decision boundaries, making them particularly suitable for real-world fraud detection scenarios where minimizing false positives and false negatives is crucial. The study also emphasized the importance of addressing the inherent data imbalance in fraud detection datasets. Techniques such as oversampling and under sampling were employed to ensure an equitable representation of both majority and minority classes. These approaches significantly improved model sensitivity to fraudulent transactions while maintaining overall accuracy. However, the trade-off between computational cost and performance, particularly in ensemble methods, remains a critical consideration for real-time deployment. Despite these promising findings, several limitations were identified. The evaluation was conducted on a single dataset, which may not fully capture the diversity of fraud patterns across various domains. Additionally, the focus on standard performance metrics, such as accuracy, precision, recall, and F1-score, leaves other aspects—such as interpretability, fairness, and robustness to adversarial attacks—open for further exploration. Future research could investigate the integration of

hybrid models that combine supervised learning for classification with unsupervised anomaly detection to uncover previously unseen fraud patterns. Additionally, implementing explainable AI techniques would improve transparency and trust in these systems, particularly for stakeholders in financial institutions. Testing these models on more diverse and larger datasets would further validate their applicability in real-world scenarios. In conclusion, this research highlights the transformative potential of ML in enhancing cybersecurity frameworks for digital payment systems. By leveraging advanced models and addressing key challenges such as data imbalance and scalability, organizations can significantly improve their fraud detection capabilities. These findings provide a foundation for the development of next-generation fraud detection systems, fostering trust and security in the rapidly evolving digital economy.

References

Sonam Rani and Aiit Mittal. Securing digital payments a comprehensive analysis of ai driven fraud detection with real time transaction mon- itoring and anomaly detection. In *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, volume 6, pages 2345–2349. IEEE, 2023.

Massimiliano Aschi, Susanna Bonura, Nicola Masi, Domenico Messina, and Davide Profeta. Cybersecurity and fraud detection in financial transactions. In *Big Data and Artificial Intelligence in Digital Finance: Increasing Personalization and Trust in Digital Finance Using Big Data and AI*, pages 269–278. Springer, 2022.

Shobhit Agrawal. Enhancing payment security through ai-driven anomaly detection and predictive analytics. *International Journal of Sustainable Infrastructure for Cities and Societies*, 7(2):1–14, 2022.

Oluwabusayo Adijat Bello, Adebola Folorunso, Jane Onwuchekwa, Oluomachi Eunice Ejiofor, Folake Zainab Budale, Maryann Nwanneka Egwuonwu, et al. Analysing the impact of advanced analytics on fraud detection: A machine learning perspective. *European Journal of Computer Science and Information Technology*, 11(6):103–126, 2023.

Halima Oluwabunmi Bello, Adebimpe Bolatito Ige, and Maxwell Nana Ameyaw. Adaptive machine learning models: concepts for real-time financial fraud prevention in dynamic environments. *World Journal of Advanced Engineering Technology and Sciences*, 12(02):021–034, 2024.

R Udayakumar, A Joshi, SS Boomiga, and R Sugumar. Deep fraud net: A deep learning approach for cyber security and financial fraud detection and classification. *Journal of Internet Services and Information Security*, 13(4):138–157, 2023.

Prasanna Kumar, Darshini Y Gowda, and Ashwini M Prakash. Machine learning in cybersecurity: A comprehensive survey of data breach detection, cyber-attack prevention, and fraud detection. *Pioneering Smart Healthcare 5.0 with IoT, Federated Learning, and Cloud Security*, pages 175–197, 2024.

Rahul Khurana. Fraud detection in ecommerce payment systems: The role of predictive ai in real-time transaction security and risk management. *International Journal of Applied Machine Learning and Computational Intelligence*, 10(6):1–32, 2020.

Victor Chang, Alessandro Di Stefano, Zhili Sun, Giancarlo Fortino, et al. Digital payment fraud detection methods in digital ages and industry 4.0. *Computers and Electrical Engineering*, 100:107734, 2022.

Shadrack Obeng, Toluwalase Vanessa Iyelolu, Adetola Adewale Akin- sulire, and Courage Idemudia. Utilizing machine learning algorithms to prevent financial fraud and ensure transaction security. *World Journal of Advanced Research and Reviews*, 23(1):1972–1980, 2024.

G Jaculine Priya and S Saradha. Fraud detection and prevention using machine learning algorithms: a review. In *2021 7th International Conference on Electrical Energy Systems (ICEES)*, pages 564–568. IEEE, 2021.

Ananta Prasad Nanda, Krishna Kumar Veluri, and Dhananjay Beura. Role of ai in enhancing digital payment security. 2024.

Jainil Coder. Online payment fraud detection dataset. https://www.kaggle.com/datasets/jainilcoder/online-payment-fraud-detection/data, 2023