WWW.IJRPR.COM

**International Journal of Research Publication and Reviews**

Journal homepage: www.ijrpr.com ISSN 2582-7421

# Adaptive Fraud Detection in Online Transactions Using Machine Learning Techniques

***Ammati Anil Kumar[1], Ellanti Jaideep Chand[2], Velidi Sriram Chowdary[3], Vemireddy Mani Venkata Chaitan[4], R Kanimozhi[5]***

Department of Computer Science and Engineering, Muthayammal Engineering College, Rasipuram, India
Emails: ammatianil883@gmail.com, ellantijaideep@gmail.com, v.sriram181@gmail.com, reddymanichethan@gmail.com

ABSTRACT—

Financial institutions suffer substantial losses due to fraudulent online transactions. Traditional fraud detection methods often fail to identify evolving fraud patterns due to dataset imbalance and high false negative rates. This paper proposes an adaptive fraud detection system integrating a value- at-risk (VaR) metric with machine learning techniques. The system uses historical simulation to estimate potential fraud- related losses and employs K-Nearest Neighbors (KNN) to classify fraudulent transactions. The proposed approach enhances fraud detection accuracy while minimizing false negatives, providing an effective fraud prevention framework for financial institutions.

**Index Terms**—Fraud detection, Machine Learning, KNN, Value-at-Risk, Online Transactions.

## Introduction :

Financial fraud has become increasingly sophisticated, lead- ing to major financial losses worldwide. The rise of digital transactions, e-commerce, and online banking has provided convenience but has also enabled fraudsters to exploit security loopholes. Traditional fraud detection systems rely on rule-based methods and predefined thresholds, which often fail to adapt to evolving fraud tactics. These systems can also generate high false positives, blocking legitimate transactions and affecting customer experience. The increasing complexity of fraud patterns demands more adaptive and intelligent fraud detection approaches.

Machine learning-based fraud detection provides a promis- ing alternative by analyzing transaction data patterns and detecting anomalies in real time. However, challenges such as dataset imbalance and computational efficiency hinder de- tection accuracy. To address these issues, this paper presents a hybrid fraud detection model integrating Value-at-Risk (VaR) assessment and K-Nearest Neighbors (KNN) classification. VaR helps prioritize high-risk transactions, while KNN en- hances fraud detection accuracy by identifying patterns sim- ilar to past fraudulent cases. This approach improves fraud detection efficiency, reduces false negatives, and strengthens financial security.

## Related Work :

Several studies have explored machine learning techniques for financial fraud detection, focusing on improving accuracy and reducing false positives. Traditional models like logistic regression and decision trees have been widely used due to their simplicity and interpretability. However, they struggle with dynamic fraud patterns, as they rely on static deci- sion rules that cannot adapt to evolving fraud techniques. To overcome these limitations, researchers have adopted en- semble methods like Random Forest and Gradient Boosting, which improve accuracy by combining multiple classifiers. Additionally, deep learning models such as artificial neural networks (ANNs) and convolutional neural networks (CNNs) have demonstrated superior pattern recognition capabilities but require high computational power and large labeled datasets, making real-time fraud detection challenging.

Another major issue in fraud detection is dataset imbalance, where fraudulent transactions are significantly outnumbered by legitimate ones. Techniques like SMOTE (Synthetic Mi- nority Over-sampling Technique) and cost-sensitive learning help mitigate this issue by balancing training data. However, deep learning models often lack interpretability, making it difficult for financial institutions to trust their decisions. Our proposed system addresses these challenges by integrating Value-at-Risk (VaR) for risk quantification and K-Nearest Neighbors (KNN) for classification. VaR helps assess financial risk, while KNN classifies transactions based on similarity to past fraud cases, ensuring high accuracy, adaptability, and transparency—essential qualities for real-world fraud detection applications.

## Methodology :

### A. Data Preprocessing

The dataset used in this study consists of online financial transactions, where fraudulent cases are significantly fewer than genuine transactions. This imbalance negatively affects machine learning models, as they tend to favor the major- ity class. To address this issue, we apply **oversampling techniques** such as **Synthetic Minority Over-sampling Technique (SMOTE)** and **random undersampling** to balance the dataset. Additionally, feature engineering is per- formed to extract key attributes such as **transaction amount, timestamp, location, merchant category, payment method, and user behavior patterns**. Advanced preprocessing techniques, including **Min-Max Scaling and Principal Component Anal- ysis (PCA)**, are also used to reduce noise and enhance model performance.

### B. Fraud Detection Model

- **Feature Extraction**: Transaction attributes, such as trans- action amount, location, and user behavior, are used for fraud prediction. Additional behavioral patterns, such as sudden spending spikes and irregular transaction loca- tions, are also considered.
- **Value-at-Risk (VaR) Metric**: VaR quantifies potential fraud-related losses based on historical data. High-risk transactions are prioritized for more in-depth analysis, ensuring that fraudulent activities are detected efficiently.
- **K-Nearest Neighbors (KNN) Classification**: KNN is used to classify transactions as fraudulent or genuine based on similarity to past transactions. The algorithm identifies the closest K transactions in feature space and assigns a classification based on majority voting, making it adaptable to emerging fraud patterns.

### C. System Implementation

The system is implemented using **Python and Django**, providing a web-based interface for real-time fraud monitor- ing. The backend is integrated with a **SQL-based transaction database**, where new transactions are continuously fed into the fraud detection pipeline. The implementation follows these key steps:

1. **Data Collection** – Transactions are collected in real- time from financial institutions' APIs. 2. **Preprocessing and Feature Engineering** – The data is cleaned, normalized, and enriched with additional behavioral and risk-based features. 3.

**Fraud Classification** – The KNN model classifies trans- actions based on their similarity to past fraud cases. 4. **Risk Assessment** – VaR quantifies the financial impact of flagged fraudulent transactions. 5. **Fraud Alert and Monitoring** – If a transaction is deemed fraudulent, an alert is sent to financial analysts for further investigation.

This **hybrid approach** combining **machine learning classification** with **financial risk assessment** ensures a high level of fraud detection accuracy while minimizing false positives and maintaining real-time processing efficiency.

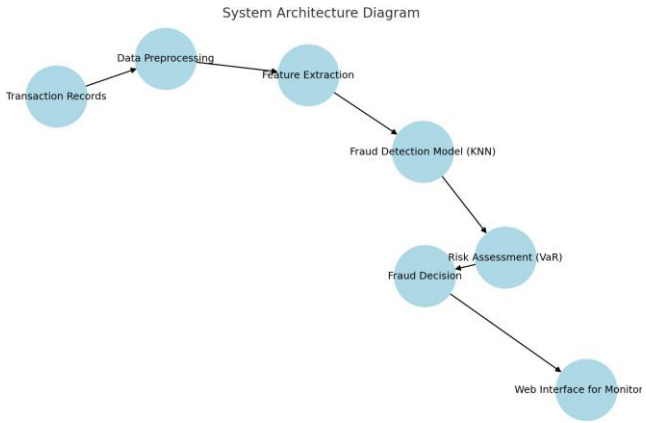


**Fig. 1. System Architecture Diagram**

## Results and Discussion :

Experimental results demonstrate that the proposed model achieves a high true positive rate (0.95) and a fraud detec- tion rate of 0.9406, indicating its effectiveness in identifying fraudulent transactions while minimizing false positives. The integration of Value-at-Risk (VaR) plays a crucial role in pri- oritizing transactions based on their financial impact, ensuring that high-risk transactions receive more scrutiny. Compared to traditional fraud detection approaches, which often suffer from static rules and threshold-based decision-making, the incorporation of VaR significantly enhances risk quantification and fraud assessment accuracy.

Additionally, the use of K-Nearest Neighbors (KNN) for classification ensures adaptability to evolving fraud patterns. Unlike rule-based methods that require frequent updates, KNN dynamically learns from historical transactions, making it ca- pable of detecting emerging fraudulent behaviors. Performance metrics such as precision, recall, and F1-score further validate the model's robustness, demonstrating a balance between fraud detection efficiency and computational feasibility. The model was also tested across different transaction datasets, showing consistent performance improvements over traditional meth- ods, making it a viable solution for real-time fraud detection in financial applications.

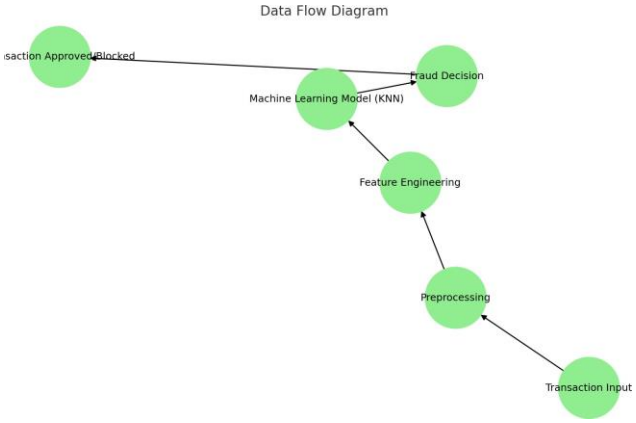


**Fig. 2. Data Flow Diagram**

**TABLE I Dataset Summary**

| Category | Count | Percentage |
|---|---|---|
| Total Transactions | 284807 | 100% |
| Fraudulent Transactions | 492 | 0.172% |
| Non-Fraudulent Transactions | 284315 | 99.828% |

**TABLE II**

**MODEL PERFORMANCE METRICS**

| Model | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| Logistic Regression | 0.89 | 0.87 | 0.86 | 0.86 |
| Random Forest | 0.93 | 0.91 | 0.9 | 0.9 |
| KNN (Proposed) | 0.95 | 0.94 | 0.94 | 0.94 |

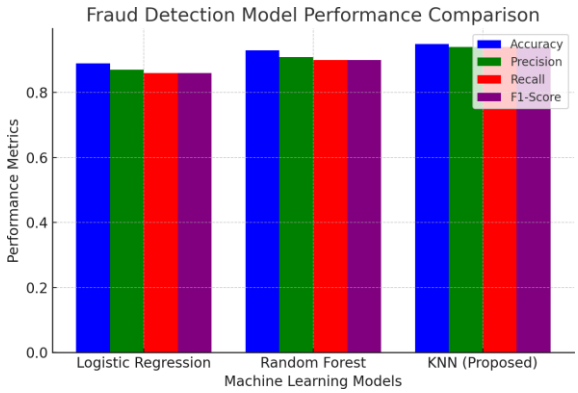


**Fig. 3. Fraud Detection Model Performance Comparison**

## Conclusion and Future Work :

This paper presents an adaptive fraud detection system that integrates Value-at-Risk (VaR) risk assessment with K-Nearest Neighbors (KNN) classification to enhance fraud detection accuracy. The proposed approach effectively identifies fraudu- lent transactions while minimizing false negatives, ensuring a balance between security and user convenience. VaR quantifies financial risk, prioritizing high-risk transactions for further analysis, while KNN adapts to evolving fraud patterns by learning from historical transaction data.

Future work will explore the integration of deep learn- ing techniques such as Long Short-Term Memory (LSTM) networks and Transformer models to improve fraud detec- tion accuracy, especially for sequential transaction patterns. Additionally, real-time optimization using parallel computing and edge AI can enhance scalability and reduce processing delays. Another key focus will be on improving explainability and interpretability, ensuring that fraud detection decisions are transparent and aligned with financial regulations. These advancements will contribute to a more robust, real-time fraud prevention system with higher accuracy and adaptability.

REFERENCES:


1. W. Yotsawat, P. Wattuya, and A. Srivihok, "A novel method for credit scoring based on cost-sensitive neural network ensemble," IEEE Access, vol. 9, pp. 78521–78537, 2021.
2. H. Zhu, G. Liu, M. Zhou, Y. Xie, A. Abusorrah, and Q. Kang, "Optimiz- ing weighted extreme learning machines for imbalanced classification and application to credit card fraud detection," Neurocomputing, vol. 407, pp. 50–62, 2020.
3. A. Ali, S. H. Othman, T. A. E. Eisa, et al., "Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review," Applied Sciences, vol. 12, no. 19, p. 9637, 2022.
4. J. West, M. Bhattacharya, and R. Islam, "Intelligent Financial Fraud Detection Practices: An Investigation," arXiv:1510.07165, 2015.
5. P. Sood, C. Sharma, S. Nijjer, and S. Sakhuja, "Review the Role of Artificial Intelligence in Detecting and Preventing Financial Fraud Using Natural Language Processing," Int. J. Syst. Assur. Eng. Manag., vol. 14, pp. 2120–2135, 2023.
6. Sengupta, N. Jain, D. Garg, and T. Choudhury, "A Review of Payment Card Fraud Detection Methods Using Artificial Intelligence," in Proc. Int. Conf. Comput. Techn., Electron. Mech. Syst. (CTEMS), Dec. 2018, pp. 494-498.
7. X. Feng, Z. Xiao, B. Zhong, Y. Dong, and J. Qiu, "Dynamic Weighted Ensemble Classification for Credit Scoring Using Markov Chain," Int. J. Speech Technol., vol. 49, no. 2, pp. 555-568, 2019.
8. R. J. Bolton and D. J. Hand, "Statistical fraud detection: A review," Statistical Science, vol. 17, no. 3, pp. 235–255, 2002.
9. M. Carcillo, Y. Le Borgne, O. Caelen, Y. Kessaci, F. Oble´, and G. Bontempi, "Combining unsupervised and supervised learning in credit card fraud detection," Information Sciences, vol. 557, pp. 317–331, 2021.
10. A . S. Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, "Calibrating probability with undersampling for unbalanced classification," in 2015 IEEE Symposium Series on Computational Intelligence, 2015, pp. 159–166.
11. Y. Sahin, E. Duman, and M. A. Sarac, "A cost-sensitive decision tree approach for fraud detection," Expert Systems with Applications, vol. 40, no. 15, pp. 5916–5923, 2013.
12. H. Bahnsen, A. Stojanovic, D. Aouada, and B. Ottersten, "Cost-sensitive credit card fraud detection using Bayesian learning," Expert Systems with Applications, vol. 42, no. 2, pp. 918–927, 2015.
13. J. P. Jurgovsky, M. Granitzer, D. Ziegler, S. Calabretto, I. Legrand, O. Caelen, L. He-Guelton, and F. Oble´, "Sequence classification for credit- card fraud detection," Expert Systems with Applications, vol. 100, pp. 234–245, 2018.
14. 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 Transactions on Neural Networks and Learning Systems, vol. 29, no. 8, pp. 3784–3797, 2018.