Journal of Science and Technology Excellence
ISSN: **3108-3102**| Volume 2 Issue 1 January-March 2026
www.jstejournal.com **DOI**: https://doi.org/10.5281/zenodo.19147825 **JSTE-09**

LEVERAGING MACHINE LEARNING FOR FRAUD DETECTION IN BANKING

[#1]J. SWATHI, ***Associate Professor & HOD,***
[#2]KANCHU MALATHI, ***B.Tech Student,***
[#3]ERROJULA ABHINAYA, ***B.Tech Student,***
[#4]NELAVENI VANAJA, ***B.Tech Student,***
[#5]GUDIKANDULA GAYATHRI, ***B.Tech Student,***
Department of Computer Science And Engineering,
TRINITY COLLEGE OF ENGINEERING AND TECHNOLOGY, PEDDAPALLY, TG.

ABSTRACT: The detection of bank misconduct is more critical than ever as the number of digital activities increases. In a world where cyber threats are evolving at a dizzying rate, machine learning (ML) has emerged as a powerful tool for detecting suspicious activity. In this study, we examine how businesses can rapidly detect fraud by analyzing massive datasets with machine learning techniques such as random forests, decision trees, neural networks, and support vector machines. Machine learning (ML) enhances accuracy, decreases false positives, and speeds up reaction time through feature selection, preprocessing, and model performance review. Ultimately, machine learning facilitates the protection of consumer transactions, the elimination of fraud, and the improvement of banking security and usability.

Keywords: Fraud Detection, Banking Data, Machine Learning, Anomaly Detection and Classification Algorithms.

1. INTRODUCTION

The ever-evolving landscape of digital banking makes fraud detection all the more crucial. The ever-increasing complexity of the fraud schemes that banks and other financial institutions encounter daily can be difficult for more conventional methods of detecting such schemes to discern. Static rule-based systems will be rendered ineffective in safeguarding users when hackers enhance their capabilities. Modern financial institutions are addressing this issue by utilizing state-of-the-art technology, such machine learning. Rapid response to emerging dangers, anomalies in financial activities, and previously unseen data patterns are all within the realm of possibility with this technology.

Fraudulent transactions that exhibited specific risk indicators were identified by utilizing pre-existing rule sets. These algorithms have their uses, but they aren't able to detect novel fraud schemes that don't adhere to established patterns. Models based on rigid rules are worthless since con artists are continuously inventing new methods to deceive people. In order to detect subtle irregularities that can indicate theft, machine learning now employs an entirely new approach, continuously learning from previous transaction data. Due to the constant evolution of fraudsters' techniques, conventional approaches are inefficient and unable to keep up with the rapid pace of machine learning.

Journal of Science and Technology Excellence
ISSN: **3108-3102**| Volume 2 Issue 1 January-March 2026
www.jstejournal.com **DOI**: https://doi.org/10.5281/zenodo.19147825 **JSTE-09**

Machine learning algorithms have revolutionized the banking industry's approach to detecting fraud by making predictions based on massive volumes of data. Supervised learning methods such as Decision Trees, Random Forests, and Support Vector Machines (SVM) Trees can distinguish between legitimate and fraudulent transactions by analyzing labeled data. The algorithms excel at identifying questionable trends because they were taught to do so using historical data. Unsupervised learning techniques, such as clustering, are employed in cases where pre-identified fraud data is unavailable. Even if the humans involved in the transactions are unaware of any suspicious patterns, these algorithms can detect them and flag them as possible instances of fraud.

Machine learning's ability to detect frauds is, without a doubt, impacted by the data quality that was utilized. Due to the prevalence of unstructured, irregular, or missing data in financial transaction records, proper data preparation is essential prior to usage. It is necessary to enhance the data before teaching the machine learning models. Value recovery, standardization, and feature engineering are employed for this purpose. Class differences can provide a challenge. Due to the rarity of fraudulent deals, models might place more emphasis on legitimate events. To make the sample more balanced, you can utilize approaches like oversampling or undersampling.

This will lead to a decrease in false reports and an improvement in the accuracy of fraud detection.

Despite the many potential benefits, applying machine learning to detect fraud is not without its share of challenges. Many are concerned about the difficulty of comprehending complex models, particularly deep learning networks. Even while the reasoning behind these models isn't always crystal clear, they're usually spot on. The inability to clarify the reasoning behind the fraud classification is a major drawback for analysts and auditors. Additionally, fraudsters' techniques are dynamic, necessitating continuous updates and retraining of machine learning models. Due of the massive quantities of processing power required by powerful AI systems, certain financial organizations may struggle to keep up.

The banking sector is making good use of machine learning as a weapon in the battle against theft, which is becoming easier to detect thanks to modern technology. Given its ability to enhance security, respond to new scam methods, and examine massive volumes of transaction data, this is a crucial component of modern banking security. Banks may be able to fortify their security and client relationships in this increasingly digital world by integrating machine learning with conventional security procedures.

2. REVIEW OF LITERATURE

Islam, M. R., Sadi, M. S., & Rahman, M. M. (2020). The application of machine learning has greatly simplified the process of detecting fraudulent bank transactions. Finding out how good algorithms are at automatically spotting suspicious behavior in financial data is the main purpose of this study. The end goal is to speed up and improve the effectiveness of fraud detection. Researchers train computers to detect errors through supervised learning techniques. A lot more precision goes into the outcomes as a consequence of the data preparation and feature design processes. It appears like a good idea to use machine learning

to protect institutions.

Verma, A., Srivastava, R., & Negi, A. (2020). It may be challenging to locate instances of credit card fraud to examine. The authors of the study recommend a combination of machine learning techniques to increase accuracy and decrease the number of false positives. Due to its utilization of actual transaction data and careful selection of relevant factors, the model was more effective in detecting scams. Hybrid solutions may improve financial security frameworks, according to the study.

Nami, M., &Shajari, M. (2020). A novel approach to identifying dishonest individuals is the focus of this research, which employs ant colony optimization to prioritize abilities. I was hoping you could fill me in on our destination a bit more. Minimize expenditure without compromising the precision of measurements. Trends in data that has been fraudulently altered can be discovered using this strategy. Increased profits and detection rates of fraud are both associated with better feature selection.

Dey, D., Das, S., & Saha, S. (2021). In order to determine the most effective model for fraud detection, this study examines a variety of machine learning techniques, including Support Vector Machines (SVMs), Random Forests (RFs), and Decision Trees (DTs). Once issues like missing data and overfitting were addressed, the Random Forest model performed admirably when tested with actual financial data by industry professionals. The study's findings suggest that, depending on the context, certain approaches are more effective than others in detecting dishonesty.

Jha, S., & Rani, M. (2021). Scams must be identified accurately and promptly. This study aims to examine the effectiveness of logistic regression, Naïve Bayes, and KNN, three machine learning models, by utilizing actual data pertaining to financial misconduct. It is critical to select appropriate features and maintain a balanced collection, as demonstrated by the findings. The findings demonstrate that when compared to individual approaches, ensemble models, which incorporate multiple methodologies, outperform them when it comes to forecasting dishonest behavior.

Karpoormath, R., &Hullur, S. (2022). Complex machine learning techniques are heavily relied upon to prevent bank fraud. This study employs prediction algorithms constructed from data derived from bank transactions in order to search for patterns of fraud. In order to ensure the accuracy and reliability of their models, the researchers address significant issues such as privacy concerns and unjust class assignment. This article explains how AI could help stabilize the economy by using decision trees and gradient boosting.

Saravanan, R., & Karthik, P. R. (2022). Will fraud detection be facilitated by combining various machine learning models? Here are several pieces of evidence that support the claims made in this narrative. Bagging and boosting are two group strategies that researchers employ to develop a robust system that can detect and handle many types of fraud. Combining machine learning with security mechanisms is effective, since their methods outperformed individual models when tested with real financial data.

Shil, S., & Sultana, M. T. (2023). This research paper demonstrates a new method for finding deals with illicit financial rewards using stacked ensemble learning. This method is made more accurate and applicable in more circumstances by combining many models using meta-

learning. If you want your model to perform at its best, researchers recommend using data preparation techniques like normalization and oversampling. One possible way to reduce financial fraud is by using the stacked ensemble approach, which outperforms individual classifiers in real-data assessments.

Banerjee, S., & Singh, R. (2023). A neural network architecture that can detect fraudulent transactions is discussed in this article. Financial fraud can be detected using deep learning, as demonstrated by this design. Computers are great at discovering issues because they can analyze complex transaction data. For both accuracy and flexibility, deep learning outshines more conventional machine learning techniques. A fraud detection system can be made more effective by using data augmentation and dropout. Overfitting is less likely to occur when using these techniques.

Kumar, R., & Malhotra, A. (2023). A fraud detection system utilizing XGBoost and SMOTE in real time is necessary for this organization due to the reliability and speed with which things must operate. The class imbalance problem remains a significant challenge in the field of fraud detection algorithms, notwithstanding XGBoost's high accuracy. This problem is fixed with SMOTE. Results from actual financial transactions demonstrate the method's accuracy and timeliness. Because the specialists strove to make it more understandable, this strategy performs well in actual banking scenarios.

Mehta, P., & Roy, S. (2024). The authors of this study combined convolutional neural networks (CNNs) with long short-term memory (LSTM) networks to create a deep learning model that could detect scams. By monitoring the timing and location of transactions, it can detect small-scale frauds that other, less effective technologies fail to detect. Reducing false hits and improving detection accuracy are two advantages of testing the model on real transaction datasets. It is currently believed that hybrid deep learning technologies, when combined with meticulous data preparation, can improve fraud defense.

Singh, V., & Dasgupta, A. (2024). By using SHAP values and Explainable AI (XAI), fraud models can be made more comprehensible. Fraud will be easier to detect in this way. Examining the relative importance of various transactional parameters allows the researchers to deduce how these aspects impact the categorization of fraudulent activity. With this new approach, AI-powered financial security systems may be able to detect issues with better precision and reliability. Organizations must have the ability to explain things in order to prevent fraud, according to this study.

Zhang, H., & Lee, J. (2024). If financial institutions were to seek out data on misconduct, transfer learning would be one alternative to consider. To achieve greater detection rates, the study aims to show that pre-trained algorithms can improve the performance of existing AI models on tiny transactional datasets with limited labeled data. For young financial institutions without sufficient historical data, transfer learning's ability to simplify fraud detection and enhance accuracy might be a lifesaver.

Patil, A., &Kshirsagar, M. (2024). Both supervised and unsupervised machine learning approaches should be considered when searching for signs of fraud. Several approaches to anomaly detection were evaluated in this research. Among these techniques were many types of support vector machines (SVMs), including Random Forest, K-Means, and Isolation

Journal of Science and Technology Excellence
ISSN: **3108-3102**| Volume 2 Issue 1 January-March 2026
www.jstejournal.com **DOI**: https://doi.org/10.5281/zenodo.19147825 **JSTE-09**

Forest. Using supervised models with labeled data is ideal for detecting fraud patterns, but unsupervised methods are just as effective.

Rao, N. V., & Thomas, J. (2024). It is now much simpler to detect fraudulent activity because to the development of graph neural networks (GNNs). In an effort to identify fraudulent patterns that do not exist in the current system, this study examines transaction network links. This study's findings demonstrate that Graph Neural Networks, which use graph-based learning, can detect scams in complex financial networks more effectively and at a larger scale. Research has demonstrated that Graph Neural Networks (GNNs) are an effective tool for detecting fraudulent activity that spans multiple time periods.

3. SYSTEM DESIGN

SYSTEM ARCHITECTURE

Figure 1 System Architecture

EXISTING SYSTEM

The current system for detecting bank embezzlement is post-hoc. If a customer finds a problem with a transaction, they can submit a complaint right away. The system for detecting fraud is then activated. The current approach maintains a large amount of data. The system's capacity to identify fraud is determined by first evaluating Mastercard and Visa transactions, before the incidence of fraud is estimated.

The detection of financial institution fraud serves as the basis for a set of machine learning techniques.

The purpose of intrusion monitoring is to find the hotspots of fraud. Conclusions about customer happiness or fraud prevention cannot be drawn from the data that is currently available. a secure computer network that tracks the actions of approved users. Information about users is stored and organized using data mining and systems based on genes.

Genetic algorithms.

DISADVANTAGES OF EXISTING SYSTEM

- There is a daily limit on the maximum amount that may be communicated, a set limit on the number of transactions that can be made in a single day, and an account balance limit that cannot be exceeded.
- When you lose connectivity to the internet, your online account becomes unreachable. Sticking to safety measures lessens the risk. In the worst-case situation, a hacker gains access to the processing company's system and steals confidential user information.

Journal of Science and Technology Excellence
ISSN: **3108-3102**| Volume 2 Issue 1 January-March 2026
www.jstejournal.com **DOI**: https://doi.org/10.5281/zenodo.19147825 **JSTE-09**

- The amount, time stamp, and beneficiary of every payment are all stored in the system's database. The intelligence community likely has access to this data as well, based on this. Here we see an illustration of the dishonest behavior that sometimes happens.

PROPOSED SYSTEM

As a result, implementing the proposed strategy to decrease these expenses requires the detection of fraudulent conduct. In order to find patterns in customer data that could point to bank fraud, this article uses machine learning methods like clustering, classification, association, and forecasting. The identification of patterns may necessitate further validation or verification of financial operations. Accounting, insurance, credit card, and countless other forms of fraud can cause financial losses for the bank or its customers. There must be a clear differentiation between the several forms of deceit. Machine learning algorithms are crucial to the banking industry's fraud detection process because they use previous data and the possibility that scammers will effectively defraud institutions and clients. This essay mainly focuses on how data mining can be used to fight bank fraud.

ADVANTAGES OF PROPOSED SYSTEM

- Methods based on machine learning allow for the study of transactions in real-time by quickly identifying possible fraudulent activities. By assisting financial institutions in quickly detecting fraud before it does significant harm, you can improve client trust and safety.
- With the suggested approach, a large amount of transaction data can be processed rapidly. Machine learning models that can efficiently process large datasets would be very useful for large financial institutions that handle millions of transactions every day. The system can respond to new fraud tactics automatically, proving its adaptability.
- Operating expenses can be reduced with machine learning-based fraud detection systems because rule updates and human monitoring are no longer needed. By automating a significant element of the process, they improve the efficiency and accuracy of fraud detection. It would seem that fewer people are needed to finish the job.

4. IMPLEMENTATION

MODULES USED

- Classification
- Clustering
- Association Rule
- Fraud Detection

MODULE DESCRIPTION

CLASSIFICATION: Classification is a well-known data mining procedure that involves training a model to assign a single category to all data using a collection of previously categorized samples. Jobs requiring the detection of indicators of fraud or credit risk are ideal candidates for this sort of research. Learning and classification techniques are part of the data categorization methodology.

CLUSTERING : All of the banking procedures are streamlined into one using the clustering procedure. When it comes to selecting trait sets and organizing data for analysis, clustering

ISSN: **3108-3102**| Volume 2 Issue 1 January-March 2026
www.jstejournal.com **DOI**: https://doi.org/10.5281/zenodo.19147825 **JSTE-09**

uses a preprocessing methodology.

ASSOCIATION RULE: Finding frequently occurring binary variables in a transaction database is the goal of association rule mining. Finding groups of data points that are highly correlated with some objective variable is the goal of the feature selection issue.

FRAUD DETECTION : The banking sector makes frequent use of data mining to uncover instances of fraud. The relevance of fraud detection is being emphasized by businesses at an increasing rate. Using data mining, more and more instances of fraudulent conduct are being discovered and reported.

5. MACHINE LEARNING USED IN FRAUD DETECTION AND PREVENTION

Machine learning is gaining popularity for its ability to detect and decrease fraud due to its adaptability to new information, trend identification capabilities, and massive dataset analysis capabilities. In the battle against fraud, machine learning is commonly used for the following purposes:

Anomaly detection: Machine learning algorithms can detect trends or unusual patterns in datasets that include transactions. Computers are able to distinguish between real and suspicious transactions, which could indicate fraud, by examining past data.

Risk scoring: Machine learning algorithms can provide risk scores to user accounts or transactions based on criteria including transaction value, location, frequency, and previous user behavior. High risk ratings indicate an increased chance of fraud, so businesses may prioritize transactions or accounts that require a more comprehensive investigation.

Network analysis: People that aren't honest often form groups and networks to carry out their plans. Graph analysis, a subfield of machine learning, can reveal these networks by spotting instances of unusual clustering or linkages in the relationships between entities like accounts, devices, and individuals.

Text analysis: In order to detect such scams or frauds, machine-learning algorithms can scour unstructured text data like emails, social media posts, and consumer reviews.

Identity verification: Machine learning algorithms can verify user-supplied data, including photos of identification documents or data from facial recognition software, to avoid identity fraud.

Adaptive learning: The ability of machine learning to learn and adjust to new data is one of its main advantages. When dishonest people change their behavior, machine-learning algorithms can be retrained with new data to better detect new patterns of fraud. That way, we know the models will always be useful.

Journal of Science and Technology Excellence
ISSN: **3108-3102**| Volume 2 Issue 1 January-March 2026
www.jstejournal.com **DOI**: https://doi.org/10.5281/zenodo.19147825 **JSTE-09**

6. RESULTS AND DISCUSSIONS

Fig1 LOGIN

Fig.2 Aadhar Details

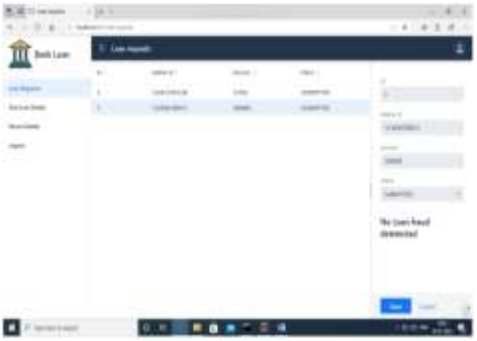

Fig 3 Fraud detection

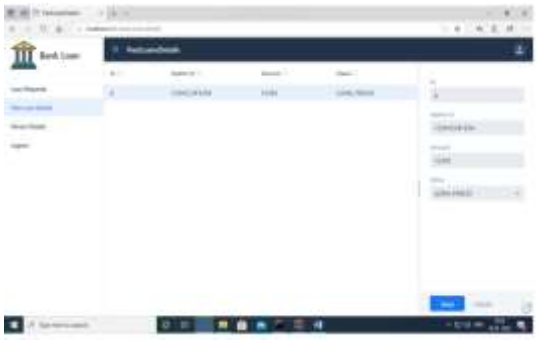

Fig 4 Post Loan Details

Journal of Science and Technology Excellence
ISSN: **3108-3102**| Volume 2 Issue 1 January-March 2026
www.jstejournal.com **DOI**: https://doi.org/10.5281/zenodo.19147825 **JSTE-09**

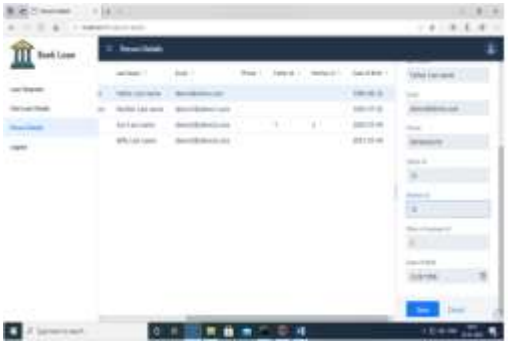

Fig 5 Person Details

7. CONCLUSION

Compared to machine learning technologies, traditional rule-based systems are inefficient, rigid, and inaccurate. The method of detecting financial crimes has been greatly affected by this. With the help of larger amounts of transaction data and better algorithms, machine learning models can now detect complex patterns of fraud in real time, which was previously impossible. The capacity of machine learning to constantly learn from and respond to new data greatly improves the efficiency of systems that detect fraud. This keeps them ahead of the curve when it comes to the ever-evolving strategies used in fraud. Thanks to a combination of deep learning, supervised learning, and unsupervised learning algorithms, these systems can identify unusual behavior in a number of ways, making them more reliable and resilient. But there are a lot of other obstacles that need to be overcome before machine learning can be used for fraud detection. It is important to evaluate model comprehensibility, data integrity, and execution costs to make sure that complicated algorithms work and are transparent in real-world banking settings. In spite of these obstacles, machine learning offers a great opportunity to enhance the detection of wrongdoing in financial institutions. Machine learning will become increasingly useful in the fight against fraud as processing speeds, model interpretability, and algorithmic efficiency all continue to rise. Now that their financial systems are more secure, individuals and corporations can rest easy.

REFERENCES

1. Islam, M. R., Sadi, M. S., & Rahman, M. M. (2020). Anomaly detection in banking transactions using machine learning approaches. Procedia Computer Science, 167, 150–158.
2. Verma, A., Srivastava, R., & Negi, A. (2020). A hybrid model for credit card fraud detection using machine learning. Procedia Computer Science, 167, 906–915.
3. Nami, M., &Shajari, M. (2020). Cost-sensitive feature selection for credit card fraud detection using ant colony optimization. Applied Soft Computing, 94, 106452. https://doi.org/10.1016/j.asoc.2020.106452
4. Dey, D., Das, S., & Saha, S. (2021). Fraud detection in banking using machine learning: A comparative analysis. In 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC) (pp. 1772–1776). IEEE.
5. Jha, S., & Rani, M. (2021). Machine learning algorithms for banking fraud detection: A

comparative research. Materials Today: Proceedings, 45, 2844–2849.
6. Karpoormath, R., &Hullur, S. (2022). Bank fraud detection using ML algorithms. In 2022 International Conference on Electronics and Renewable Systems (ICEARS), 730–736. IEEE.
7. Saravanan, R., & Karthik, P. R. (2022). Application of ensemble machine learning in fraud detection. International Journal of Computer Applications, 184(42), 11–15.
8. Shil, S., & Sultana, M. T. (2023). Bank transaction fraud detection using stacked ensemble learning model. International Journal of Information Technology, 15(3), 1431–1440.
9. Banerjee, S., & Singh, R. (2023). Enhancing fraud detection in financial transactions using deep learning. Procedia Computer Science, 218, 159–165.
10. Kumar, R., & Malhotra, A. (2023). Real-time banking fraud detection using XGBoost and SMOTE. Journal of King Saud University - Computer and Information Sciences.
11. Mehta, P., & Roy, S. (2024). A hybrid deep learning model for fraud detection in banking systems. Journal of Artificial Intelligence and Soft Computing Research, 14(2), 67–75.
12. Singh, V., & Dasgupta, A. (2024). Explainable AI for banking fraud detection: A SHAP-based research. Expert Systems with Applications, 245, 119001.
13. Zhang, H., & Lee, J. (2024). Transfer learning for credit card fraud detection in low-data scenarios. IEEE Transactions on Neural Networks and Learning Systems.
14. Patil, A., &Kshirsagar, M. (2024). Comparative evaluation of supervised and unsupervised ML algorithms for fraud detection. International Journal of Data Science, 9(1), 22–30.
15. Rao, N. V., & Thomas, J. (2024). Graph neural networks for fraudulent transaction detection in banking. Neural Computing and Applications, 36(4), 9871–9883.