International Journal of Innovative Research and Advanced Studies (IJIRAS)
Volume 12 Issue 4, April 2025

ISSN: 2394-4404

Credit Card Fraud Detection Using Machine Learning

Prathibha.U

Assitant Professor Department of Software System, BSc AIML Sri Krishna Arts and Science College, Coimbatore, Tamilnadu, India

Sudhakar.S

Sai Aravind Prakash.J

Ankush Dey.K

BSc AIML Sri Krishna Arts and Science College, Coimbatore, Tamilnadu, India

Abstract: ***It is vital that credit card companies are able to give consumers data on all transactions that make up their credit card statements and account information. Verified fraudulent credit card transactions to identify such transactions so only customers can make informed choices about the transaction. For items that they have not yet purchased they are not charged for those purchases. Such Problems can be tackled with data science and its importance. Machine Learning and Machine Learning can't be overstated. This Project aims to illustrate the modelling of a data set using Matlab 3D in-house. The project will develop a dataset of data collected from the Machine learning for credit card fraud detection.***

This project aims to enhance the current research on credit card fraud detection by applying machine learning techniques, providing valuable insights into the performance of various modeling methods. The findings of this research will not only facilitate more efficient identification of fraudulent activities but will also aid financial institutions in establishing strong security protocols to safeguard consumers against potential fraud.

Keywords: Credit Card Fraud Detection, Machine Learning, Data Science, Transaction Classification, Financial Security.

I. INTRODUCTION

The fraud in credit card transactions is unauthorized and non-existent. To use a customer account without their consent by someone else than the user you're supposed to be using it or those who have registered and signed up for the account you want to use is a minor offense. owner of that account. Necessary prevention measures can be To stop these abuses and the behaviour of such fraudulent schemes are taken measures that could have contributed to the development of these fraudulent programs. Practices can be studied to minimize it and protect against it. In other words credit card payments will continue to be charged and processed in the future. Fraud can be defined as a case in which a person uses someone else's name to cheat on another person' If another person's credit card has been used for personal reasons when the owner and the user have not notified the bank about them but the card was cancelled or issued Card issuing authorities do not know that the card is valid until the card is issued. being used. Fraud detection involves monitoring the activities of the victim - the enterprise - and the activities of other victims. User populations of users to estimate, perceive or avoid uses in order to prevent or increase the use of new technologies.

II. HISTORY OF CREDIT CARD FRAUD DETECTION

Since the 1980s credit card fraud detection has evolved significantly in response to increasing complexity of fraudulent activities. In the early days of credit cards introduced by Diners Club and later by American Express and BankAmerica (now Visa) in the 1950s fraud checking relied on manual verification as merchant banks checked lists of blacklisted card By the 1980s and 1990 banks implemented rule-based fraud detection systems that flagged a suspicious transaction based on predefined conditions such as unusually large purchases or transactions from different locations in a

short time. The 2010s saw further advancements with AI and behavioural analytics, including geolocation device fingerprinting and neural networks to improve accuracy.

Fraud detection methods are continuously developed to defend against fraud in the field. Criminals are adapting their fraudulent strategies in the face of threats from fraud. These frauds are classified as:

- ✓ Credit Card fraud: online and offline
- ✓ Card Theft
- ✓ Account Bankruptcy
- ✓ Device Intrusion
- ✓ Application Fraud
- ✓ Counterfeit Card
- ✓ Telecommunication Fraud

III. APPLICATIONS OF CREDIT CARD FRAUD DETECTION

Credit card fraud detection plays a crucial role in various industries and financial ecosystems. Its applications span multiple areas in order to enhance security, prevent losses reduce financial loss and protect consumers ensuring financial security across industries:

A. BANKING AND FINANCIAL INSTITUTIONS

Real-Time Transaction Monitoring: Real time transaction monitoring is a crucial component of credit card fraud detection. It allows banks and financial institutions to identify and prevent fraudulent activities as they occur.

Instant Transaction Alerts: Instant transaction alerts are a crucial security feature in credit card fraud detection, helping customers and financial institutions quickly identify unauthorized transactions. Suspicious Activity Warnings: Suspicious activity warnings are an essential feature in credit card fraud detection, helping banks and financial institutions identify and prevent fraudulent transactions in real time.

B. AI-POWERED ANOMALY DETECTION

Machine Learning Algorithms for Fraud Detection: Machine learning algorithms are of great importance for the detection of fraud of credit cards. The increase in the amount of automated fraudulent transactions forms the basis of this technique.

Unusual Spending Pattern Alerts: A warning is triggered by transactions that substantially beyond the cardholder's usual spending patterns. Several transactions in a short period of time or high-value purchases could be deemed suspicious.

High-frequency trading: Develop advanced trading algorithms for faster and more accurate financial decisions.

IV. CHALLENGES IN CREDIT CARD FRAUD DETECTION

Despite significant advancements in fraud detection, financial institutions still face numerous challenges in effectively identifying and preventing fraudulent transactions. One of the most pressing issues is the continuous evolution of fraudulent techniques. Fraudsters frequently develop new methods to bypass security measures, making it difficult for traditional detection systems to keep up. They employ tactics such as synthetic identity fraud, where fake identities are created using real and fake information, and account takeovers, in which stolen credentials are used to gain unauthorized access to accounts. The use of advanced technologies, such as deepfake AI and automated bots, further complicates fraud detection efforts.

Another major challenge is the sheer volume of financial transactions processed daily. Millions of transactions occur globally, and distinguishing between legitimate and fraudulent ones requires high computational efficiency and real-time processing. Implementing fraud detection systems that can analyze massive datasets without causing delays is a demanding task. Furthermore, fraud detection models must be regularly updated to adapt to new fraud patterns, requiring continuous learning and retraining.

False positives pose another critical challenge in fraud detection. While fraud detection algorithms aim to minimize fraudulent transactions, they sometimes misclassify legitimate transactions as fraudulent. When genuine customers experience declined transactions due to false positives, it can lead to frustration and loss of trust in financial institutions. Balancing fraud detection sensitivity while reducing false positives is a crucial aspect that requires optimization of machine learning models.

Data privacy and regulatory compliance are also significant concerns. Financial institutions handle large volumes of sensitive customer data, making data security a top priority. Implementing fraud detection solutions that comply with global regulations such as the General Data Protection Regulation (GDPR) and the Payment Card Industry Data Security Standard (PCI DSS) is essential. Additionally, sharing fraud-related data across financial institutions could enhance fraud detection but raises concerns about data privacy and confidentiality. Addressing these challenges requires a collaborative approach involving advanced AI-driven solutions, industry-wide cooperation, and strict adherence to data protection laws.

V. FUTURE TRENDS IN CREDIT CARD FRAUD DETECTION

As fraud tactics evolve, the future of credit card fraud detection will rely on advanced technologies to strengthen security measures. One of the most promising developments is the integration of artificial intelligence (AI) and deep learning in fraud detection systems. AI-powered models can analyze vast amounts of transaction data, identify unusual spending patterns, and detect fraudulent activities in real time. Deep learning techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), can enhance the accuracy of fraud detection by learning complex patterns in transaction behaviors.

Blockchain technology is another transformative trend in fraud prevention. Blockchain's decentralized nature ensures that transaction records are immutable and transparent,

making it difficult for fraudsters to manipulate transaction data. Smart contracts can also be utilized to enhance security, as they automate financial transactions with predefined conditions, reducing human intervention and potential security risks.

The rise of biometric authentication methods is expected to add an additional layer of security to credit card transactions. Fingerprint scanning, facial recognition, and voice authentication are becoming increasingly popular in banking and financial services. These biometric security measures make it harder for fraudsters to gain unauthorized access to accounts, thereby reducing fraud cases related to identity theft and account takeovers.

Multi-factor authentication (MFA) and tokenization will also play crucial roles in the future of fraud prevention. MFA requires users to provide multiple forms of verification before accessing their accounts, such as a combination of passwords, OTPs, and biometrics. Tokenization replaces sensitive card information with encrypted tokens, ensuring that transaction details remain secure even if intercepted by cybercriminals. Additionally, real-time fraud intelligence sharing among financial institutions and regulatory bodies will improve fraud detection capabilities, allowing for proactive threat mitigation.

With the increasing adoption of machine learning and AI-driven fraud detection, continuous model improvement and adaptive learning will be necessary to stay ahead of emerging fraud techniques. The financial industry will need to invest in cutting-edge research and collaborate with cybersecurity experts to ensure that fraud prevention systems remain robust and effective in an ever-changing digital landscape.

VI. SYSTEM MODEL AND METHODOLOGY

Our system utilizes various machine learning algorithms such as Logistic Regression (LG), Support Vector Machine (SVM), XGBoost, Random Forest, Decision Tree, and K-Nearest Neighbors (KNN) for credit card fraud detection. Logistic Regression is used for predicting categorical dependent variables by providing probabilistic values between 0 and 1. The SVM algorithm aims to create the best decision boundary, or hyperplane, to segregate data points into distinct classes. KNN works by identifying the K closest neighbors to a data point and making predictions based on their classifications. Random Forest enhances accuracy by combining multiple decision trees and selecting the majority decision, while Decision Trees make hierarchical classifications. Our system employs an oversampling method for data balancing, and after testing, SVM has shown the highest accuracy, which is measured using the ROC (Receiver Operating Characteristic) curve.The dataset used in this system contains transactions made by European credit cardholders in September 2013 over two days. Out of 284,807 transactions, only 492 are fraudulent, resulting in a highly imbalanced dataset where fraud accounts for just 0.172% of transactions. Due to confidentiality reasons, the original features are not disclosed, and numerical input variables have been transformed using Principal Component Analysis (PCA). The dataset includes features such as "Time" (elapsed time between transactions) and "Amount" (transaction value), while the "Class" variable indicates whether a transaction is fraudulent (1) or legitimate (0). Given the dataset's class imbalance, the recommended accuracy measure is the Area Under the Precision-Recall Curve (AUPRC), as a confusion matrix alone is insufficient.

Principal Component Analysis (PCA) is employed as a dimensionality reduction technique that transforms high-dimensional data into a smaller-dimensional subspace while retaining the highest variance. It mathematically defines an orthogonal linear transformation that projects data onto new coordinate systems, where the first principal component captures the maximum variance, followed by the second component, and so on. To address data imbalance, we apply SMOTE (Synthetic Minority Oversampling Technique), a statistical method that increases the number of samples in the minority class by generating synthetic data points based on existing nearest neighbors.

Our proposed system uses an open-source dataset (credit_card.csv), read using the Pandas library. Preprocessing steps include checking for null values and scaling the data. The dataset is split into training and testing sets using the Sklearn library. Model training involves passing data through different machine learning algorithms, and the ROC curve is plotted for evaluation. During model testing, test values are passed through the trained model to generate predictions, which are compared with actual values using a confusion matrix. The final model is selected based on overall accuracy and generalization performance.

For software requirements, Jupyter Notebook serves as an open-source web application for coding, visualization, and documentation. Anaconda Navigator provides a GUI for managing Python packages, environments, and applications without using the command-line interface (CLI). The system relies on several Python libraries, including NumPy for numerical computing and linear algebra operations, Pandas for data manipulation and analysis, Keras for implementing deep learning neural networks, and Sklearn (Scikit-learn) for machine learning and statistical modeling. These tools collectively enable efficient fraud detection by leveraging data science and machine learning techniques.

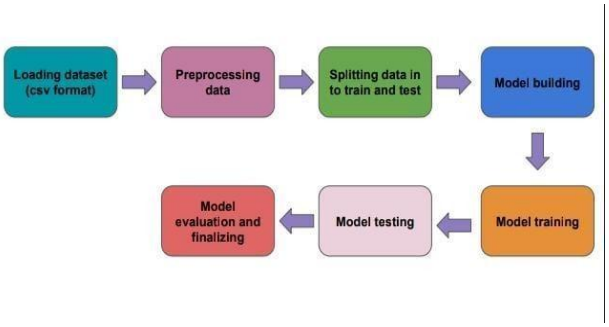
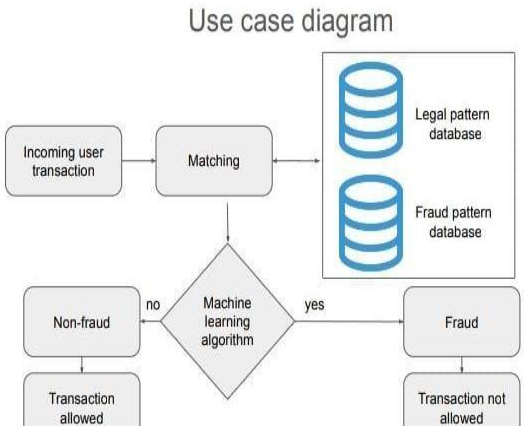


Figure 1: Proposed system

ISSN: 2394-4404

VII. USE CASE DIAGRAM



VIII. CONCLUSION

Credit card fraud detection is a crucial aspect of financial security that has evolved significantly over the years. From manual verification methods in the early days to AI-driven real-time monitoring, fraud detection systems have continuously improved to combat sophisticated fraud techniques. However, the financial industry still faces numerous challenges, including the evolving nature of fraud, the need for real-time data processing, the issue of false positives, and concerns over data privacy and regulatory compliance. Addressing these challenges requires ongoing research, innovation, and industry-wide collaboration.

The future of credit card fraud detection lies in the adoption of AI, deep learning, blockchain, and biometric authentication technologies. These advancements will enable financial institutions to detect fraud with greater accuracy and efficiency while enhancing security for consumers. Multi-factor authentication, tokenization, and real-time fraud intelligence sharing will further strengthen fraud prevention strategies. However, to stay ahead of fraudsters, financial institutions must remain vigilant, continuously update their fraud detection models, and invest in robust security measures.

As digital transactions continue to grow, the role of machine learning and AI in fraud detection will become even more critical. By leveraging cutting-edge technologies and adopting a proactive approach to fraud prevention, the financial industry can create a safer and more secure environment for consumers. Ultimately, a combination of advanced fraud detection techniques, global cooperation, and regulatory compliance will help mitigate credit card fraud risks and protect financial systems from emerging threats.

REFERENCES

[1] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed., Morgan Kaufmann, 2011.
[2] A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153-1176, 2016.
[3] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.
[4] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, 2000.
[5] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785-794, 2016.
[6] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine Learning: A Review of Classification and Combining Techniques," Artificial Intelligence Review, vol. 26, pp. 159-190, 2006.
[7] 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.