e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.521

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707187|

A Smart Way to Identify Credit Card Fraud Using Machine Learning Techniques

Sagar G M, Prof. Sravanthi Kalal

Student, Department of MCA, AMC Engineering College, Bengaluru, India

Assistant Professor, Department of MCA, AMC Engineering College, Bengaluru, India

ABSTRACT: Credit cards are currently among the most popular payment mechanisms for both offline and online transactions, driven by advancements in electronic commerce systems and communication technologies. However, this popularity also increases the hazard of fraud. Fraudsters continually develop new methods to deceive people, leading to significant financial losses for individuals and businesses each year. The smart and crafty nature of credit card criminals makes it difficult to detect credit card fraud. Furthermore, the detection procedure is made more difficult by the notable data imbalance that fraud detection systems exploit. Thus, it is essential to have quick and accurate ways to spot fraudulent credit card transactions. This study introduces the Gradient Boosting Classifier as a novel machine learning approach for detecting credit card fraud. Trial outcomes prove that this method achieves 100% training accuracy and 91% test accuracy, outperforming previous machine learning techniques.

KEYWORDS: innovative technique, gradient boosting, machine learning.

I.INTRODUCTION

Machine learning mentions to the ability of a computer algorithmic system to learn from past experiences and improve without explicit programming. It is a branch of artificial intelligence that uses data and statistical techniques to predict outcomes and generate actionable insights. The principle is that a computer can produce accurate results by learning from data or examples. Machine learning is closely related to data mining and Bayesian predictive modeling, where the computer uses algorithms to respond to data inputs.

One common application of machine learning is generating recommendations. For instance, Netflix provides recommendations based on users' viewing histories. Tech companies use unsupervised learning to create personalized suggestions that enhance user experience. Machine learning is also used to automate tasks like fraud detection, predictive maintenance, and portfolio optimization.

Unlike traditional programming, where a programmer defines all rules based on expert consultation, machine learning allows the framework to learn and improve on its own. In traditional programming, the complexity of the system determines the amount of rules needed, making maintenance challenging. In machine learning, the system studies from examples, much like humans learn from experience. The more data it has, the better it can predict outcomes. Conversely, with less data, the system's accuracy decreases.

II. LITERATURE REVIEW

Attioui, A. [1] notes that significant advancements in machine learning have enabled the expansion of real-time, interactive, intelligent systems for various data processing and classification tasks. The accuracy and precision of these systems depend on the temporal and logical consistency of the data and the speed of feedback generation. In order to increase accuracy and precision, financial institutions are investing in data analysis technologies and algorithm enhancements for fraud detection systems, which is the subject of this study. Numerous machine learning-based approaches have been recognized, but few studies compare deep learning paradigms, and there is a need for real-time solutions. This paper proposes a deep neural networkbased real-time credit card deception discovery system using an auto-encoder for real-time transaction classification. Our model outperforms four other binary classification models in terms of accuracy and recall.

Kataria, A. [2] discusses the potential of artificial intelligence in automating financial risk assessment for businesses and credit bureaus through machine learning. This initiative aims to assess and analyze credit card delinquency risk, providing a predictive framework for credit bureaus. Machine learning aids risk assessment by differentiating between honest and fake transactions in diverse datasets. Financial institutions can halt fund disbursement for flagged transactions. The study contrasts several models of machine learning including random forests, RUSBoost, decision trees, logistic regression, multilayer perceptrons, and nearest neighbor algorithms.

Hashim, A. S. [3] states that Software Defect Prediction (SDP) models rely on software metrics. The quality of these models depends on the dataset's quality, with high dimensionality being a significant issue. Feature selection (FS) is a common solution, but empirical research on FS techniques for SDP often yields inconsistent results. This inconsistency arises from the varied computational underpinnings of FS approaches and the different search tactics used.

Bandaranayake, B. [4] describes a policy campaign by Australia's Victorian Department of Education and Early Childhood Development to prevent fraud and corruption. The policy framework reflects a broad, decentralized structure of governance and accountability. The case highlights the complexity of policy implementation and the practical approach adopted by the Department. Large-scale educational institutions can study appreciated lessons from this case, despite the lack of simple solutions for preventing fraud and corruption.

BOUAHIDI, E. [5] discusses the growing use of credit cards for electronic payments, increasing the risk of fraud for financial institutions and service providers. Effective fraud detection systems are needed to mitigate significant annual losses.

Nevertheless, false alerts might result from machine learning systems' frequent inability to take into account fraud sequences or behavioral shifts. This research presents a Long Short-Term Memory (LSTM) network-based credit card fraud detection system that integrates transaction sequences. The approach uses credit cardholders' historical purchase records to increase the accuracy of fraud detection. Studies show that the suggested methodology detects fraud with higher accuracy.

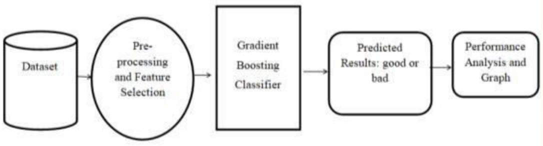
Dataset
Pre-processing and Feature Selection
Gradient Boosting Classifier
Predicted Results: good or bad
Performance Analysis and Graph

Fig 1 : System Architecture

III.PROPOSED MODEL

Machine learning is the capability of a computer algorithm to learn from past experiences and improve without explicit programming. It uses data and statistical techniques to predict outcomes and generate actionable insights, forming a crucial part of artificial intelligence.

The premise is that computers can generate precise results by learning from data instances. Machine learning is closely related to data mining and Bayesian predictive modeling, where an algorithm processes input data to produce responses.

A common machine learning task is recommendation generation. For instance, Netflix uses it to recommend shows based on users' viewing history. Tech companies employ unsupervised learning to create personalized suggestions that enhance user experience.

Machine learning also automates tasks like fraud detection, predictive maintenance, and portfolio optimization. Unlike traditional programming, which requires extensive rule-setting based on expert consultation, machine learning relies on the computer's learning process. Similar to human learning through experience, the machine improves its predictions with more data exposure.

IV.LITERATURE ANALYSIS

Attioui, A. et al. [1] noted significant advancements in machine learning, facilitating real-time, interactive, intelligent systems across various domains, with fraud detection. Banks and financial institutions are investing in enhancing algorithms and data analysis technologies for more accurate fraud detection systems. Despite numerous machine learning-based approaches, there is a lack of comparative studies on deep learning paradigms, and existing works often neglect the necessity of real-time solutions. This study suggests a system for detecting credit card fraud in real time that is based on a deep neural network using an auto-encoder for real-time classification of transactions as authorized or fraudulent. The system's performance is likened with four other binary classification models, showing promising results in terms of accuracy and recall.

Kataria, A. [2] highlighted the potential of artificial intelligence to automate the financial risk assessment process for businesses and credit bureaus. Machine learning distinguishes between genuine and fraudulent transactions, aiding in risk assessment and enabling financial institutions to halt funds disbursement in case of fraud alerts. Models like random forests, RUSBoost, decision trees, logistic regression, multilayer perceptrons, and nearest neighbor algorithms are used for this purpose.

Hashim, A. S. [3] discussed Software Defect Prediction (SDP) models built using software metrics, emphasizing the impact of data quality on model effectiveness. Feature selection (FS) addresses high dimensionality issues, though empirical research on FS techniques for SDP often yields inconsistent results, making FS strategy selection challenging.

Bandaranayake, B. [4] presented a case study on fraud and corruption control by Australia's Victorian Department of Education and Early Childhood

Development, illustrating the complexity of policy implementation in large, decentralized educational institutions.

Bouahidi, E. [5] addressed how the rising use of credit cards for electronic payments is making financial institutions more susceptible to theft. The study suggests utilizing Long Short-Term Memory (LSTM) networks for a fraud detection system., which incorporates transaction sequences to improve detection accuracy by leveraging previous purchase histories.

V. PRESENT MODEL

Raghavan et al. describe an auto-encoder as a real neural network capable of encoding and decoding data. This technique is used to train auto-encoders without anomalous locations, with reconstruction error indicating anomalies categorized as "fraud" or "no fraud."

Carcillo et al. used a mixed method combining unsupervised outlier scores with a classifier for fraud detection, implementing different granularity levels for outliers. Another method by Carta et al. used a modified discrete Fourier transform model to detect fraud based on frequency patterns, addressing imbalanced class distribution and cold-start issues.

Despite the promotion of techniques like CNN and LSTM for image classification and NLP, the application of deep learning (DL) approaches remains limited in credit card recognition due to data pre-processing challenges.

VI. RECOMMENDED METHOD

This study proposes using the Gradient Boosting Classifier to detect fraudulent credit card transactions. The method integrates the classifier's parameters into the system, aiming to distinguish between authentic and fraudulent transactions. The proposed method's effectiveness is evaluated using real-world datasets from Kaggle, achieving 100% training accuracy and 91% test accuracy.

Key components of the future technique include data collection, pre-processing, model application, prediction output, performance analysis, and graphical depiction. The experiment was conducted using an Intel Core i3 processor with 8GB of RAM, employing Flask for the web interface and Python for machine learning techniques.

The Gradient Boosting Classifier-based method offers potential improvements in accuracy over the current random forest approach, identifying intricate patterns by training predictors to correct each other's errors. The method is highly flexible, supporting optimization on various loss functions and hyper-parameter tuning to improve function fit. It handles categorical and numerical values without preprocessing and manages missing data without imputation.

VII. CONCLUSION

Preventing credit card fraud is crucial as it supports the growing use of credit cards. Financial institutions continually face substantial and persistent financial losses due to fraud, making the detection of fraudulent transactions increasingly challenging. Hence, developing more effective methods is imperative. This paper employs Gradient Boosting Classifier to propose an intelligent approach for fraud detection in credit card transactions. Through a run of experiments using real-world data, we evaluated We assessed the execution of our method using performance analysis metricsOur test findings demonstrate that our approach not only surpassed other algorithms for machine learning, but moreover achieved the highest accuracy. These results emphasize the effectiveness of our method and underscore the significance of employing robust parameter optimization strategies to improve predictive performance.

REFERENCES

1. Y. A. Abakarim, M. Lahby, and A. In Attioui, "An effective real-time model for credit card fraud detection based on deep learning," in Proceedings of the Twelfth International Conf. cerebral, Systems:
2. Theories Appl., Oct. 2018, pp. 17, doi: 10.1145/3289402.3289530.
3. H. L. Abdi and J. Williams, "Principal component analysis," Wiley
4. Interdisciplinary Rev. Statistics, July 2010, pp. 433-459, doi: 10.1002/wics.101.
5. S. Basri, S. J. Abdulkadir, A. O. Balogun, and A. Hashim, "Facilitating user authorization from imbalanced data logs of credit cards using artificial intelligence," Mobile Information Systems, Oct. 2020, doi: 10.1155/2020/8885269.

6. B. Bandaranayake, "A case study of the Victorian Department of Education and Early Childhood Development in Australia: Fraud and corruption control at the education system level," J. Cases Educate.,
7. Dec. 2014, pp. 34-53, doi:
8. 10.1177/15554589145496669.
9. M. Szelg, R. Baaszczyński, A. T. Matuszyk, and M. Filho de Almeida, "Expert system detects auto loan fraud using a dominance-based rough set strategy," Applied, Jan. 2021, doi: 10.1016/j.eswa.2020.113740.
10. F. Cartella, Y. Funabiki, T. Akishita, D. Yamaguchi, O. Anunciacao, and O.
11. Elshocht, "Interleaved sequence RNNs for fraud detection," in Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,
12. 2020, pp. 3101-3109, doi:
13. 10.1145/3394486.3403361.
14. V. N. Dornadula and S. Geetha, "Methods of credit card fraud detection using machine learning," in Proceedings of Computer Science, Jan. 2019, pp. 631-641, doi: 10.1016/j.procs.2020.01.057.

INTERNATIONAL JOURNAL OF

MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com