ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

Volume 11, Issue 2, March 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

IMPACT FACTOR: 7.583

| www.ijarasem.com | ijarasem@gmail.com | +91-9940572462 |

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

Credit Card Fraud Detection Based on Ensemble Machine Learning Classifiers

Dr.S.LAVANYA, KOTAKONDA MALLIKARJUNA, NARA KONDAIAH,

AKKILI NAVEENKUMAR REDDY

Professor, Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

ABSTRACT: Credit card fraud has existed ever since credit cards were introduced, resulting in financial losses, identity theft, severe security threatas, and misuse of personal information. Such a situation already dire at an individual level only worsens when an organization gets involved. One of the complicated approaches in machine learning is credit card fraud detection and it offers essential financial suggestions for both banks and customers. Absolute privacy is maintained for all ML research work under credit card fraud detection we dive deep into research process to propose best topics as per your needs. Our most qualified writers will communicate with you so that all research encounters are solved. Frequent updating of technology and massive resources with professional researchers is our key step to success for credit card fraud detection project.

KEYWORDS: Classification, Credit Card Fraud, Ensemble Learning, Machine Learning (ML), Synthetic Minority Over Sampling Technique (SMOTE), Random Under Sampling (RUS).

I. INTRODUCTION

Financial transactions attribute to an integral and crucial part of our day to day lives. As a matter of fact, in today's digital world financial transactions have gone electronic or online, and the countries of the world are already moving towards a cashless society. The convergence of three industrial giants' telecommunication, banking and retail sectors have instigated the rapid growth as well as widespread use of digital platforms one being, ecommerce, which has led to a surge in online transactions. As a consequence of increased digital transactions, this has led to escalation in the number of fraudulent and illegitimate transactions(1). The ease in use and global accessibility of Internet and comprehensive increase in ecommerce as well as other digital platforms in recent years, brings about the need for increased safety and secureness in financial transactions. Thus, security poses a major concern for all digital transactions. Fraudulent activities can severely damage the financial sector(2) Among the various digital payment systems prevalent, card payment systems are the most generally accepted, convenient and widely employed means of payment. The swift expansion of communication technology has facilitated significant expansion of digital payment systems, especially in the bank card systems.

Over the years, governments have pitched in steps that curtail typical banking and ATM transactions, leading to significant increase in transactions involving bank cards. The pandemic and post pandemic scenario too have added and aided in a significant amount of card burgeoning digital transactions. The year 2022 alone shows that, over 4.5 billion individuals around the world rely on online platform purchases, with sales expenses totaling around 5.54 trillion US dollars. The worldwide E-commerce sales report shows an almost 5-fold increase in online payments from 2014 to 2022. The most predominant digital payment modes identified are Card payments, Mobile payments, Unified Payment Interface (UPI), Internet banking, Prepaid cards, Digital Wallets, and Micro ATM. This study focuses on the classification of fraudulent transactions made with credit cards from regular valid transactions.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

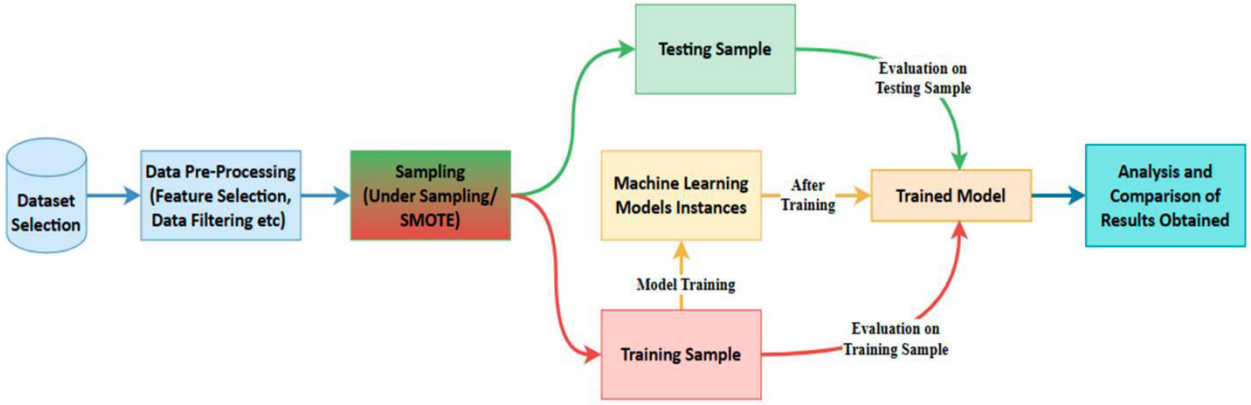


Fig 1: Enhancing Credit Card

Digital transactions have been advocated on various platforms such as e-commerce platforms because it makes handling and dealing with cash easier to both the users and the retailers. However, the tremendous volume of internet money transfers has led to increased attempts of fraud. The global financial industry is experiencing a significant threat from credit card fraud, making it a challenging subject for all businesses. As the technology involved in digital transactions are upgraded the fraudsters too evolve with newer techniques making it a challenging task.(4) Frauds can occur in many ways from Application Frauds, loss of cards, usage of imprints, ROC pumping, Magstrips to Phishing. Money Fraud is an illegal activity in which a person or a swindler directly or indirectly uses the money of the victim for and by fake transactions with information from the victim's credit card. As the usage of credit/debit card or net banking is experiencing significant growth, the possibility of numerous fraudulent activities is also increasing. These may include sharing of card details, personal details, and One Time Password (OTP) to anonymous and most probably fake calls (5). Credit card frauds are of mainly two categories - internal fraud and external fraud, which are based on the instigator of frauds, either a firsthand party or third party. Developing countries like India are a major marketplace for ebusiness based on the population and buying trends of customers.. It is predicted that by 2025, the country's digital transactions will grow by 71%. Thus, usage of digital payments such as credit cards is on a rise in the country This has led to a substantial increase in exploitation, misuse, and frauds of digital payments including credit cards. Lack of efficient Financial fraud detection systems has added to the woes in this field(6). Many countries like UK are expected to become cashless societies by 2026. Hence fraud detection and monitoring systems are the need of the hour.

II. BACKGROUND AND RELATED WORK

There are seven major stakeholders involved in credit card transactions. The: Card holder, Merchant, Payment Gate, Payment Processor, Card Payment System, Issuing Bank, and Acquiring Bank. A Cardholder is the customer or owner of the credit card used to make purchases. The customer presents the card either using a device connected with merchant system - Point-of-sale (POS) or via ecommerce site. A Merchant is the business owner who can sell products to the cardholder. The Payment gateway denotes the software that communicates the transaction details from merchant to payment processor. A Payment processor communicates with merchants, merchant banks, card networks and other entities to make card payments. The Credit card network signifies card payment networks deployed with payment processors to facilitate communication between merchant and issuing bank. Few popular credit card network providers are Visa, Master card, American Express, RuPay, and so on. The bank that provides credit card to the customers is designated as the Issuing bank. An Acquiring bank is the Merchant's bank which accepts credit card transactions of a card holder. When the customer presents the card to the merchant for payment through online or offline mode, the merchant sends the transaction details to the payment gateway, which passes the transaction to the payment processor using a secure channel. The payment processor sends the transaction authorization request to card payment network, where the details are verified by communicating with Customer or Issuing bank. Finally, the Issuing bank sends the acknowledgement to the customer.

Several recent studies have utilized different machine learning and deep learning methods such as Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Naive Bayes (NB), Multilayer Perceptron (MLP), AdaBoost, Support Vector Machines (SVM), K-Nearest Neighbor (KNN), Linear Regression, Ensemble learning, besides using deep learning techniques such as Autoencoders(9), Convolutional Neural Networks (CNN), LSTM, Recurrent Neural Networks (RNN),Generative Opponent networks (10) and more for credit card fraud detection. ML techniques on a whole, are seen useful for spam filtering, weather prediction, classification, prediction and diagnosis of diseases as proposed by (11), (12), forecasting problems (13)and many more. In the case of financial transactions ML based classifiers help in classifying fraudulent transactions from normal transactions, identifying abnormal transactions from

patterns, differentiating fraudsters and regular customers through credit card profiling and so on. Fraudsters, on the other hand, employ different means to circumvent the detection mechanisms. (14) compared Logistic Regression and Random Forest for credit card fraud detection and their study illustrated the superior performance of Random Forest over Logistic regression.

III. METHODOLOGIES USED

The objective of this paper is to examine the evaluation performance of several advanced data mining and machine learning techniques on credit card data and propose a suitable model for credit card fraud classification. According the literature survey done various models are seen to be imperative in credit card classification(27) .The most suitable machine learning models are selected for base learners of credit card fraud detections from literature study These classifiers are applied individually and the performance of each is evaluated using two sampling techniques RUS, SMOTE. Ensemble techniques are also seen to help in improving performance. Hence Bagging and Boosting ensembles are used. The sections below discuss some of the relevant classifiers used in this study. Each classifier has its own pros and cons (28)

The K nearest neighbor method is popularly used non-parametric supervised learning technique in many classification tasks such as disease classification(29) It is an instance-based method which classifies objects based on the closest feature set in the given data (30). Studies imply that this supervised learning algorithm is one of the best classifier algorithms in credit card fraud detection. KNN achieves consistently high performance. The K-NN algorithm compares the new data with existing predictions and places the new data in the most relevant output group.

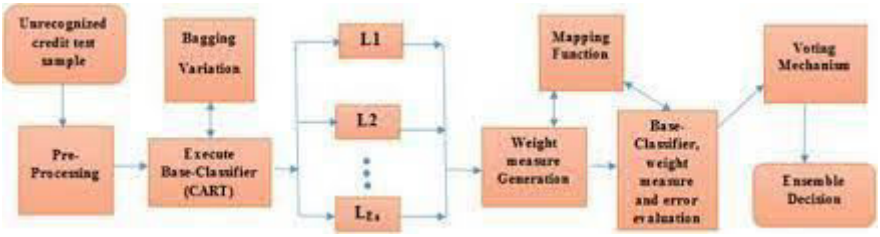


Fig 2: MPROVING ACCURACY IN CREDIT

Logistic regression is a statistical technique applied for classification and regression. Logistic regression models have illustrated good performance in classification problems(31), (32). To forecast the likelihood of counterfeit credit cards, a logistic regression approach is applied. Maximum likelihood estimate is used in logistic regression analysis to determine group membership (33). However, in order to evaluate the results of the group membership prediction with precision and accuracy, a preliminary analysis of the cleaned dataset is performed to see if the logistic regression assumptions were met.

Two assumptions underpin the Naive Bayes classifier. To begin, all features in an entry must be categorized so that they all contribute equally to the decision-making process, secondly, all characteristics should be statistically unrelated (8). The Bayes rule is used to classify an instance by applying it to each of the classes it belongs to. The model classifies the 2 categories of transactions fraudulent and valid transactions based on this rule. Naive Bayes models have seen to give good performance in classification tasks(34). (35) illustrated that Naïve Bayes was good for removing noise in credit card dataset.

The SVM algorithm is a supervised machine learning technique used in credit card fraud detection. The essential concept behind the SVM classification algorithm is to create a hyper plane as the decision plane, with the largest distance between the positive and negative modes [16]. SVM uses a kernel function to transfer the data to a specified very high-dimensional space and identifies the hyper plane that optimizes the margin between the two classes. The usage of appropriate kernel functions in SVM enhances classification process(36). SVM usually illustrates good performance for large datasets and hence it is considered suitable for dealing with the credit card fraud detection problem.

IV. RESULTS & DISCUSSIONS

The different models are evaluated by applying feature selection and without feature selection on the normal data, and resampled data. The proposed model is a binary classification problem and various metrics are used to assess the

performance of the binary classification models. The classification accuracy, precision, Recall, F1 score and MCC score are measured for actual data, data sampled using Random under sampling and SMOTE are measured with and without feature selection. The measures are defined as follows. It is a supervised learning technique applied in classification, regression and prediction jobs. It is usually represented as a treelike structure in which the features of the dataset are represented as internal nodes and decision rules as branches and the outcome as leaf nodes. Instances are classified using decision trees by sorting them down the tree from the root to a leaf node, which provides the classification.

Feature selection is the process of choosing a subset of features from a larger dataset for accurate predictions (39). It removes irrelevant and repetitive features thereby reducing the computational cost and improving storage space (25)In the experiment, the highly relevant features were selected using WEKA. Among 30 distinctive features, the Positively corelated features to the target variable - V11, V4, V2, V21, V19, V20, V8, V27, V28, Amount, V26, V25, V22 were selected using the Pearson Correlation coefficient feature selection method. Several feature selection techniques have been proposed in literature and have illustrated performance enhancement of classification(20). A correlation based feature selection method helps in identifying relation between features(22).

Fig 3: Result analysis

Random Forest classifier is considered as an ensemble of decision trees with bagging. This classifier has been seen to aid in classification tasks such as disease classification(11). Compared to individual decision trees it is a much sought after method owing to the accuracy of its predictions. It is based on a random selection of features; Random Forest models decide where to partition the data. Random Forest models incorporate differentiation because each tree splits based on different features, rather than splitting at comparable features at each node throughout the model. Because of the higher level of differentiation, there is a larger ensemble to aggregate over, resulting in a more accurate.

V. CONCLUSION

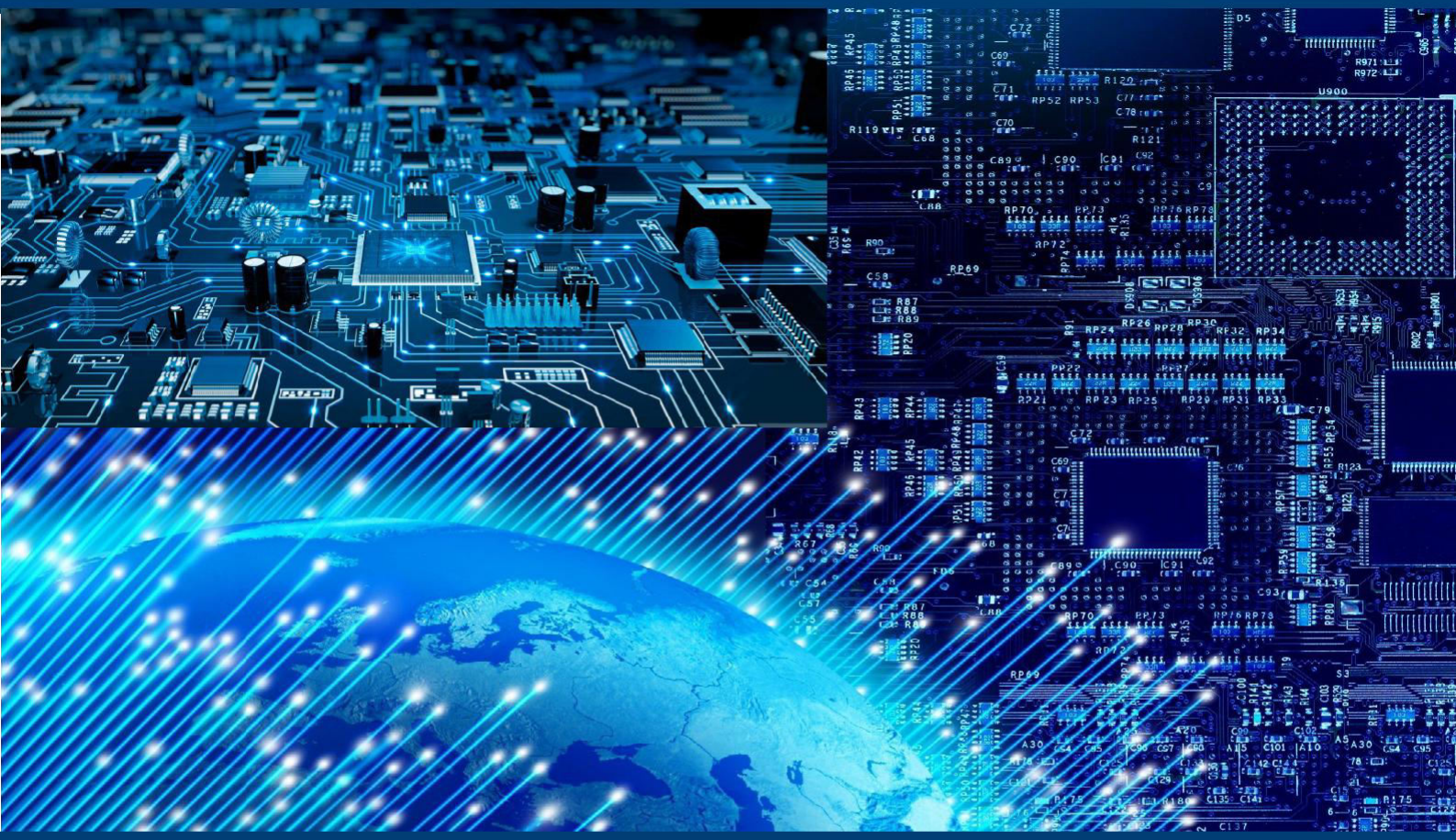
The study depicted the effectiveness of Machine learning algorithms in classifying credit card fraudulent transactions and regular valid transactions amongst the huge volume of credit card transactions. A comprehensive study of ten powerful machine learning methods was done and a model for classification of fraudulent transactions and valid transactions was developed. The proposed model was an ensemble of six machine learning classifiers. The six classifiers were selected based on performance and constituted as an ensemble which were combined using majority voting – both hard and soft voting. Ensemble learning methods were perceived to outperform individual learning methods. Both hard and soft voting illustrated stable results and soft voting classifiers were seen to perform better with normal data having no feature selection. A future work will to be see the performance provided by various techniques such as federated learning and deep learning in classification of genuine and fraudulent cases.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

REFERENCES

[1]. Bin Sulaiman R, Schetinin V, Sant P. Review of Machine Learning Approach on Credit Card Fraud Detection. Hum-Centric Intell Syst. 2022;2(1–2):55–68.
[2]. Makki S, Assaghir Z, Taher Y, Haque R, Hacid MS, Zeineddine H. An experimental study with imbalanced classification approaches for credit card fraud detection. IEEE Access. 2019;7:93010–22.
[3]. Visa credit cards in circulation Statista 2020 [Internet]. Available from: https://www.statista.com/statistics/379046/worl dwide-retail-e-commerce-sales/
[4]. Itoo F, Singh S. Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. Int J Inf Technol. 2021;13:1503–11.
[5]. Asha RB, KR SK. Credit card fraud detection using artificial neural network. Glob Transit Proc. 2021;2(1):35–41.
[6]. Dal Pozzolo A, Boracchi G, Caelen O, Alippi C, Bontempi G. Credit card fraud detection: a realistic modeling and a novel learning strategy. IEEE Trans Neural Netw Learn Syst. 2017;29(8):3784–97.
[7]. Randhawa K, Loo CK, Seera M, Lim CP, Nandi AK. Credit card fraud detection using AdaBoost and majority voting. IEEE Access. 2018;6:14277–84.
[8]. Awoyemi JO, Adetunmbi AO, Oluwadare SA. Credit card fraud detection using machine learning techniques: A comparative analysis. In: 2017 international conference on computing networking and informatics (ICCNI). IEEE; 2017. p. 1–9.
[9]. Sanober S, Alam I, Pande S, Arslan F, Rane KP, Singh BK, et al. An enhanced secure deep learning algorithm for fraud detection in wireless communication. Wirel Commun Mob Comput. 2021;2021:1–14.
[10].Singh KD, Singh P, Kang SS. Ensembled-based credit card fraud detection in online transactions. In: AIP Conference Proceedings. AIP Publishing LLC; 2022. p. 050009.



International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |