ISSN: 2581-3404 (Online) IF: 5.68 (SJIF)
IJIRTM, Volume-9, Issue-1, January-2025

Credit Card Fraud Detection: Survey and Discussion

[1]Madhuree Sahu, [2]Ritu Prasad
[1]M.Tech. Research Scholar, [2]Assistant Professor
[1,2]Department of CSE, Technocrats Institute of Technology (Excellence), Bhopal (India)
[1]sahumadhuree070@gmail.com, [2]rituprasad_2023@gmail.com

Abstract. *Credit card fraud detection is a challenging problem that banks and credit card issuers are struggling with and are making huge efforts in implementing fraud detection systems. In today's banking system, fraud detection is often done by using rule based methods. However, the progress and development of machine learning techniques gives banks and financial institutions the possibility to detect an unusual situation faster for big financial data sets. Machine Learning is a branch of Artificial Intelligence that has become very popular, and useful, in the last 10-15 years. One definition of Machine Learning is that it is the semi-automated extraction of knowledge from data. In this research work discusses different machine learning based credit/debit card fraud detection and challenges.*

Keywords: Classification, On-line transactions, Machine learning, Automated teller machine.

Introduction

Credit and debit cards are some of the most common means of making transactions in the world. The ease with which they can be obtained and used is what makes them so popular. Global purchases of goods and services, cash advances, and withdrawals made with credit, debit, and prepaid cards reached more than 42 trillion USD in 2019, and are expected to increase to 56 trillion USD by 2025. Such a massive volume of transactions and the widespread use of payment cards encourage criminals to commit card fraud. Despite the low proportion of frauds among card transactions, total losses coming from fraudulent payments are still very high. In 2019, gross fraud losses to issuers, merchants, and acquirers of transactions reached 28.65 billion USD worldwide. Every year, these losses have been rising together with the total card transaction amount and are expected to keep rising in the future. In 14 selected EU countries, according to a study of the European Banking Authority (EBA), approximately 0.016% of card transactions were reported as fraudulent by issuers in the second half of 2020 (EBA 2022). In the case of transaction value, fraudulent payments make up 0.025% of the total value. Card fraud rates reported by acquirers are even higher. About 0.035% of transactions were identified as fraudulent, which translates to 0.046% of the total value. According to the EBA, the total card fraud value reported by issuers and acquirers in H2 2020 was more than EUR 440 million. Credit card frauds can be split into application and behavioral frauds. Application fraud happens when individuals obtain credit cards based on applications with false personal or financial information. Such fraudsters may try to spend as much as possible shortly after obtaining their card and then refuse to repay the debt. This type of fraud is quite common, but it is dominated by behavioral frauds. Behavioral frauds make up a large group that can be divided into four main categories.

These are Card-not-Present (CNP) fraud, counterfeit card fraud, lost/stolen card fraud, and card-never-arrived fraud. Credit card information can be stolen via malware attacks that traditionally target computers, but nowadays malicious code can also be included in smartphone apps. Users, therefore, need to be cautious and carefully choose the applications which they are installing. Social engineering is the other and possibly more dangerous method of obtaining card information. Fraudsters keep making more and more complex methods by which they persuade card owners to hand over all key information about their cards. The exact approaches differ but fraudsters usually act as trustworthy and legit institutions and try to manipulate their targets. They can use fake e-mails, sms messages, or even phone calls, in which fraudsters talk to their victims in order to receive credit card information. The final major card fraud type is skimming and counterfeit fraud. Those occur when card details are illegally taken and used to create counterfeit cards. Skimming is a technique used to scan card information from its magnetic stripe with a device called a skimmer. Skimmers can be installed in payment terminals but most often they are used in an Automated Teller Machine (ATM). They tend to be installed in place of a legit card reader in an ATM and most of the time they are nearly identical to real readers. Therefore for a common consumer, it is very difficult to tell if a skimmer is used in an ATM. On the other hand, since skimmers need to be this complex, they are also expensive and hard to obtain. Banks are also monitoring their ATM with cameras and thus it is difficult to install a skimmer without getting noticed. These factors combined mean that skimming and counterfeit card frauds are disappearing. Instead, fraudsters are shifting their interest to the internet and are starting to use various phishing methods. Banks often stay one step behind and are not able to prevent such attacks.

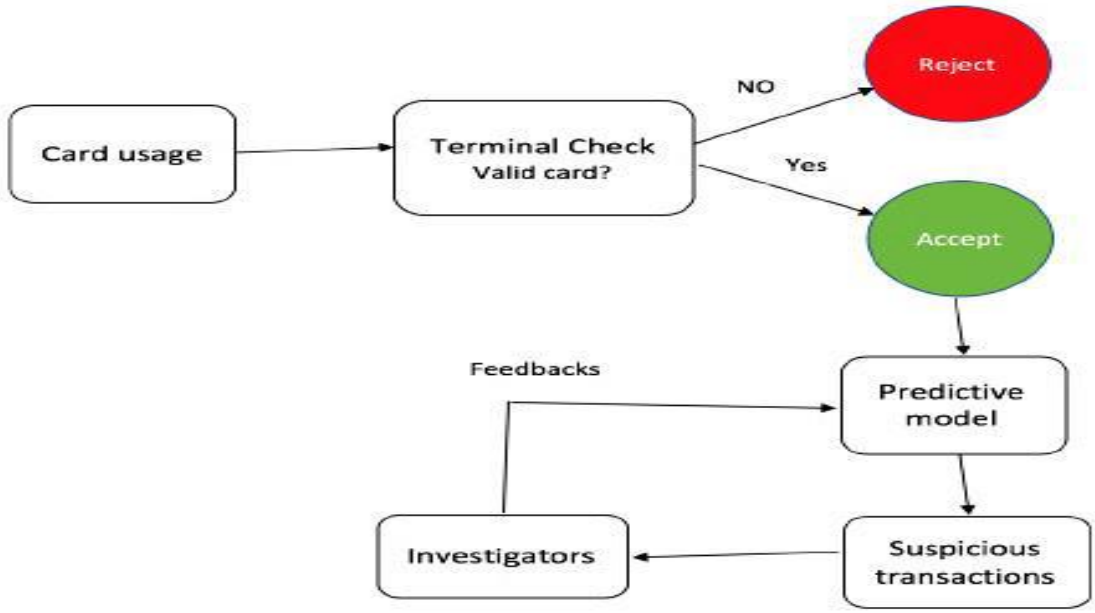

Figure 1: Fraud detection process.

II. Literature Review

This section provides a literature review of previous researches that used ML techniques for credit card fraud detection. Several machine learning methods have been proposed for credit card fraud detection. Specifically, supervised learning algorithms have shown to be highly effective in detecting credit card fraud, where labelled datasets containing previous transaction records are utilized to build machine learning models that can detect new fraudulent transactions.

[1] ML has provided methods such as Logistic Regression (LR), Support vector machines (SVM), Decision Trees (DT), Random Forest (RF), and K-Nearest Neighbors (KNN). However, these methods cannot meet the outstanding performance required to detect and predict unusual fraud patterns. In this regard, the contribution of this research work is to propose a framework for fraud detection (FFD). At first, to overcome the unbalanced data problem, the framework uses an under-sampling technique. Next, a feature selection (FS) mechanism is applied to select only relevant features. Then, a Support Vector Data Description (SVDD) is used to build the ML model. SVDD aims to create a tight boundary around regular data points to distinguish them from potential outliers or anomalies.

[2] The fraud detection system in banking organization relies on data-driven approach to identify the fraudulent transactions. In real time, detection of each and every fraudulent transaction becomes a challenging task as financial institutions need aggressive jobs running on the log data to perform a data mining task. This paper introduces a novel model for credit card fraud detection which combines ensemble learning techniques such as boosting and bagging. Their model incorporates the key characteristics of both the techniques by building a hybrid model of bagging and boosting ensemble classifiers. Experimentation on Brazilian bank data and UCSD-FICO data with our model shows sturdiness over the state-of-the-art ones in detecting the unseen fraudulent transactions because the problem of data imbalance was handled by a hybrid strategy.

[3] The advance in technologies such as e-commerce and financial technology (FinTech) applications has sparked an increase in the number of online card transactions that occur on a daily basis. As a result, there has been a spike in credit card fraud that affects card issuing companies, merchants, and banks. It is therefore essential to develop mechanisms that ensure the security and integrity of credit card transactions. In this research, we implement a machine learning (ML) based framework for credit card fraud detection using a real world imbalanced datasets that were generated from European credit cardholders. To solve the issue of class imbalance, they re-sampled the dataset using the Synthetic Minority over-sampling technique (SMOTE). This framework was evaluated using the following ML methods: Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), Extreme Gradient Boosting (XGBoost), Decision Tree (DT), and Extra Tree (ET). These ML algorithms were coupled with the Adaptive Boosting (AdaBoost) technique to increase their quality of classification. The models were evaluated using the accuracy, the recall, the precision, the Matthews Correlation Coefficient (MCC), and the Area under the Curve (AUC).

[4] The author described a decision tree as a tree-like graph made up of internal nodes that stand in for tests on attributes, branches that indicate the results of the tests, and leaf nodes that represent class labels.

The route chosen from the root node to the leaf determines the classification rules. The root node, which is the most obvious property to separate the data, is first selected to divide each input data set. Before the tree is formed, the attributes and values that will be used to analyze the input data at each intermediary node are identified. By moving from a root node to a leaf node and stopping at all internal nodes along the way, the tree can prefigure newly arriving data depending on the test conditions of the characteristics at each node. The primary difficulty lies in deciding which value to use to split a decision tree node.

[5] The fraudulent transactions may be detected via way of means of utilizing both this sort and integrating any of those methods. The version can study in a greater correct way via way of means of including new features. Several data mining strategies are being utilized by financial institutions and credit score card organizations for detecting fraud behaviors. The ordinary utilization sample of customers relying upon their beyond sports may be recognized via way of means of making use of any of those methods. Therefore, a comparative evaluation is made right here via way of means of analyzing exclusive fraud detection strategies proposed over the years.

[6] The most widely used method of payment for online transactions and fraud in everyday purchases is credit cards. Nowadays Fraudsters invent new techniques to commit fraudulent transactions which demand constant innovation for their detection techniques. The majority of methods based on artificial intelligence, fuzzy logic, NN, LR, NB, sequence alignment, DT, Bayesian networks, meta-learning, genetic programming, etc. have been created to identify different types of credit card fraud. Transaction strategies utilized in credit card fraud detection systems are surveyed in this study. Based on deep learning from a neural network, they employed twelve ML algorithms to detect credit card fraud. They benchmark and real-world performance. In addition to searching the dataset, Ada Boost and majority voting methods are applied to build the hybrid model. The accuracy and sensitivity achieved by the optimal random forest algorithm under the benchmark data are 95% and 91%, respectively.

[7] In the present time, credit card fraud detection has gained a lot of interest in the machine learning research community. This section presents the state-of-the-art techniques that have been applied for credit card fraud detection, which is categorized as probabilistic approach, individual learning approach and cost-sensitive learning-based fraud detection system. Risk induced Bayesian inference bagging is proposed for credit card fraud detection in which a novel bag creation strategy has been applied to re-balance the distribution of classes in terms of minority (fraudulent transaction) and majority (legitimate transaction) sample. Once the bag creation phase was over, the Bayesian classifier was trained and based on the probability score for a transaction being fraud afterward the class label was assigned using threshold value. The disadvantage of the RIBIB model was: it cannot be efficient for handling the concept drift problem.

III. Machine Learning Techniques

Progress in data science is not simply a matter of increased performance, speed, and storage. In addition to the type of information found in libraries, data generated in organizations, and established systems designed to gather and codify data, new forms of technology can use data that are both people-generated and machine-generated. These data are often chaotic and unstructured.

The connectedness allowed by data science is driving a new kind of discovery. People are using social networks to draw their own connections between friends, things, events, likes, dislikes, places, ideas, and emotions. Governments are analyzing social networks to stop terrorist acts. Businesses are mining social and transactional information for connections that will help them discover new opportunities. Scientists are building massive grids of connected data to tease out new findings, using AI and machine learning. As addressed in more detail below, these advances have allowed the emergence of computers that can help you perform tasks that previously had been tedious. Many of these machine-learning–aided tasks have been largely accepted and incorporated into the everyday practice of medicine. The performance of these machine tasks is not perfect and often requires a skilled person to oversee the process, but in many cases, it is good enough, given the need for relatively rapid interpretation of images and the lack of local expertise.

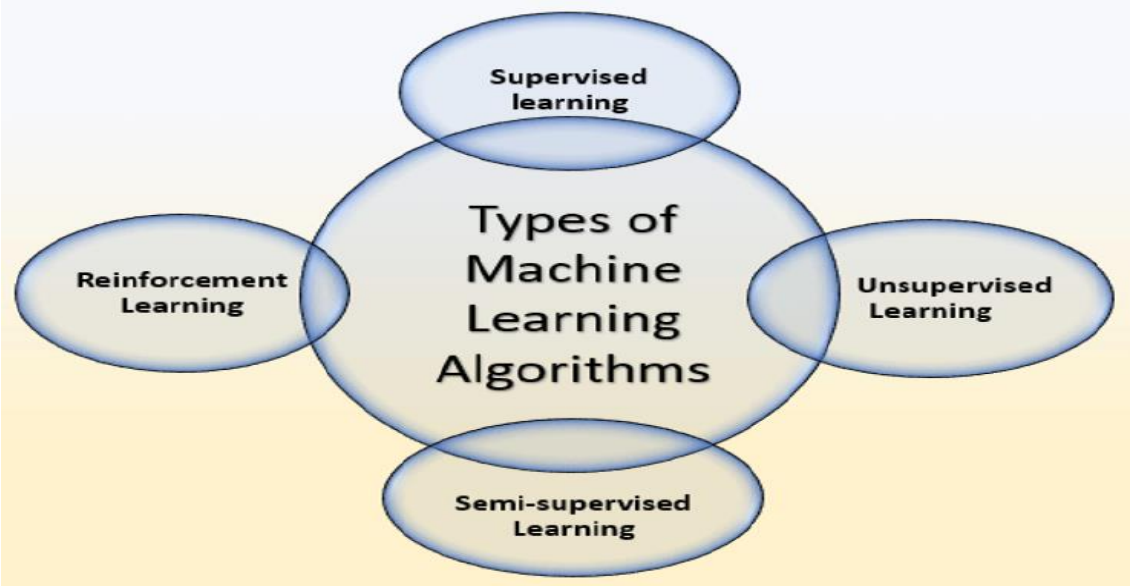

Figure 2: Types of machine learning.

IV. Conclusion

Financial frauds can have significant impacts on individuals, businesses, and society as a whole. Victims of fraud can suffer financial losses, damage to their credit scores, and emotional distress. In addition, fraud can decrease trust in financial institutions and systems, which can have broader implications for the economy and society. A large majority of losses from card fraud is borne by owners of the cards. Banks are only liable for a small proportion of losses. They may need to compensate a client when bank employees do not manage to block a lost card properly or in some cases when the law tells them to do so. In this research work discusses different techniques to detect the fraud in card based online transactions.

References:

[1] Ebenezer Esenogho, Ibomoiye Domor Mienye, "A Neural Network Ensemble with Feature Engineering for Improved Credit Card Fraud Detection", IEEE Access, 2022, pp. 16400-16408.

[2] Altyeb Altaher Taha, Sharaf Jameel Malebary, "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine", IEEE Access, 2020, pp. 25579-25588.

[3] Fawaz Khaled Alarfaj, Iqra Malik, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms", IEEE Access, 2022, pp. 39700-39715.

[4] Rashmi S. More, Chetan J. Awati, "Credit Card Fraud Detection Using Supervised Learning Approach", International Journal Of Scientific & Technology Research, 2020, pp. 216-220.

[5] V. S. S. Karthik, Abinash Mishra, "Credit Card Fraud Detection by Modelling Behaviour Pattern using Hybrid Ensemble Model", Arabian Journal for Science and Engineering, Springer 2021, pp. 1-12.

[6] Emmanuel Ileberi, Yanxia Sun, "Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost", IEEE Access, 2021, pp. 165286-165295.

[7] Altyeb Altaher Taha, Sharaf Jameel Malebary, "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine", IEEE Access, 2020, pp. 25579-25588.

[8] Fawaz Khaled Alarfaj, Iqra Malik, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms", IEEE Access, 2022, pp. 39700-39715.

[9] Rashmi S. More, Chetan J. Awati, "Credit Card Fraud Detection Using Supervised Learning Approach", International Journal Of Scientific & Technology Research, 2020, pp. 216-220.

[10] Jay Prakash Maurya, DK Rathore, S Joshi, M Manoria, V Richhariya, "Corporate Sector Fraud: Challenges and Safety", Machine Learning Applications for Accounting Disclosure and Fraud Detection, 2021, pp. 16-31.

[11] V. S. S. Karthik, Abinash Mishra, "Credit Card Fraud Detection by Modelling Behaviour Pattern using Hybrid Ensemble Model", Arabian Journal for Science and Engineering, Springer 2021, pp. 1-12.

[12] Emmanuel Ileberi, Yanxia Sun, "Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost", IEEE Access, 2021, pp. 165286-165295.

[13] Dileep M R, Navaneeth A V, "A Novel Approach for Credit Card Fraud Detection using Decision Tree and Random Forest Algorithms", IEEE, 2021, pp. 1025-1028.

[14] Deepak Rathore, Praveen Kumar, "Recent Trends in Machine Learning for Health Care Sector", International Journal of Innovative Research in Technology and Management, Vol-5, Issue-2, 2021.

[15] Mosa M. M. Megdad, Bassem S. Abu-Nasser, " Fraudulent Financial Transactions Detection Using Machine Learning", International Journal of Academic Information Systems Research, 2022, pp. 30-39.

[16] Pumsirirat, Apapan, and Liu Yan. "Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine." International Journal of advanced computer science and applications 9.1 (2018): 18-25.