International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/IJSRCSEIT

Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms

AVS Radhika[1], M Sai Chandana[2], Jeedipally Sneha[2]
Assistant Professor[1], Students[2]
Department of CSE, Bhoj Reddy Engineering College for Women, Hyderabad, India

ARTICLEINFO

Article History:

Accepted: 01 May 2023
Published: 20 May 2023

Publication Issue
Volume 9, Issue 3
May-June-2023

Page Number
163-166

ABSTRACT

The usage of credit cards for online and regular purchases is exponentially increasing and so is the fraud related with it. A large number of fraud transactions are made every day. Various modern techniques like artificial neural network Different machine learning algorithms are compared, including Logistic Regression, Decision Trees, Random Forest, Artificial Neural Networks, Logistic Regression, K-Nearest Neighbors, and K-means clustering etc. are used in detecting fraudulent transactions. This paper uses genetic algorithm, and neural network which comprises of techniques for finding optimal solution for the problem and implicitly generating the result of the fraudulent transaction. The main aim is to detect the fraudulent transaction and to develop a method of generating test data. This algorithm is a heuristic approach used to solve high complexity computational problems. The implementation of an efficient fraud detection system is imperative for all credit card issuing companies and their clients to minimize their losses.

Keywords: Machine learning, Credit card, Electronic commerce, Fraud detection.

I. INTRODUCTION

A credit card is a thin handy plastic card that contains identification information such as a signature or picture, and authorizes the person named on it to charge purchases or services to his account - charges for which he will be billed periodically. Today, the information on the card is read by automated teller machines (ATMs), store readers, bank and is also used in online internet banking system. They have a unique card number which is of utmost importance. Its security relies on the physical security of the plastic card as well as the privacy of the credit card number. There is a rapid growth in the number of credit card transactions which has led to a substantial rise in fraudulent activities. Credit card fraud is a wide-ranging term for theft and fraud committed using a credit card as a fraudulent source of funds in a given transaction. Generally, Most of the credit card fraud

detection systems are based on artificial intelligence, Meta learning and pattern matching.

II. RELATED WORK

1] Vimala Devi. J et al. To detect counterfeit transactions, three machine-learning algorithms were presented and implemented. There are many measures used to evaluate the performance of classifiers or predictors, such as the Vector Machine, Random Forest, and Decision Tree. These metrics are either prevalence dependent or prevalence-independent. Furthermore, these techniques are used in credit card fraud detection mechanisms, and the results of these algorithms have been compared.

2] Popat and Chaudhary. supervised algorithms were presented Deep learning, Logistic Regression, Nave Bayesian, Support Vector Machine (SVM), Neural Network, Artificial Immune System, K Nearest Neighbour, Data Mining, Decision Tree, Fuzzy logic based System, and Genetic Algorithm are some of the techniques used. We compared machine-learning algorithms to prediction, clustering, and outlier detection.

3] Deepa and Akila . For fraud detection, different algorithms like Anomaly Detection Algorithm, K-Nearest Neighbor, Random Forest, K-Means and Decision Tree were used. Based on a given scenario, presented several techniques and predicted the best algorithm to detect deceitful transactions. To predict the fraud result, the system used various rules and algorithms to generate the Fraud score for that certain transaction.

4] Kibria and Sevkli. Using the grid search technique, create a deep learning model. The built model's performance is compared to the performance of two other traditional machine-learning algorithms: logistic regression (LR) and support vector machine (SVM). The developed model is applied to the credit card data set and the results are compared to logistic regression and support vector machine models.

5] Borse Suhas and Dhotre Machine Learning's Naive Bayes classification was used to predict common or fraudulent transactions.

6] Asha R B et al. have proposed a deep learning-based method for detecting fraud in credit card transactions. Using machine-learning algorithms such as support vector machine, k-nearest neighbor, and artificial neural network to predict the occurrence of fraud.

III.PROPOSED SYSTEM

There are lots of issues that make this procedure tough to implement and one of the biggest problems associated with fraud detection is the lack of both the literature providing experimental results and of real-world data for academic researchers to perform experiments on. The reason behind this is the sensitive financial data associated with the fraud that has to be kept confidential for the purpose of customer's privacy. Now, here we enumerate different properties a fraud detection system should have in order to generate proper results. The system should be able to handle skewed distributions, since only a very small percentage of all credit card transactions is fraudulent. There should be a proper means to handle the noise. Noise is the errors that is present in the data, for example, incorrect dates. This noise in actual data limits the accuracy of generalization that can be achieved, irrespective of how extensive the training set is. Another problem related to this field is overlapping data. Many transactions may resemble fraudulent transactions when actually they are genuine transactions.

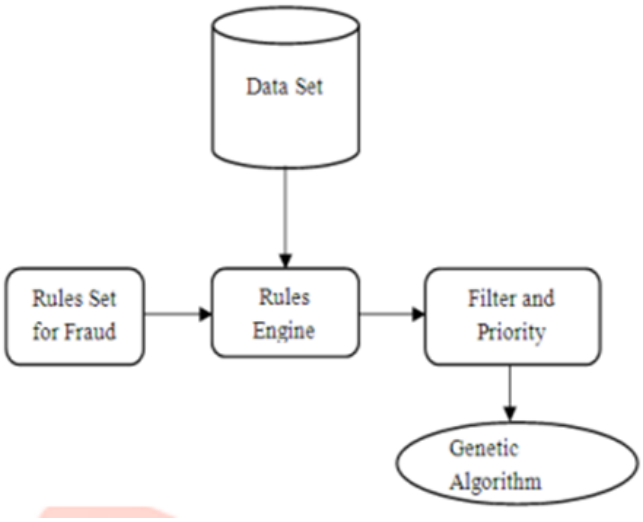

Fig 1. System Design

The above architectural design describes the work structure of the system. The data warehouse contains the customer data. This customer data is subjected to the rules engine and again ,the rules engine comprises of the rules set.

The filter and priority module sets the priority for the data and hence. Plays a very important role in the system. then the filter data is sent to the genetic algorithm module which performance its functions and generates the output.

IV.RESULTS

In this section, we report our experimental study that we performed with selected machine learning algorithms and imbalance classification approaches. First, we provide a detailed description of the design of experiments followed by the results and a discussion. Finally, we discuss some critical shortcomings we discovered in our experiments.

Design of Experiments

This section briefly presents the workflow of our experiments, the dataset used, the selection of target variables and performance measure.

A] Workflow of Experiments

Our experimental study is organized as follows. The experiments are presented and discussed in two phases. In the first phase, eight classification methods are compared. The comparison was carried out with respect to three parameters including the following: accuracy, sensitivity, and the Area under PrecisionRecall Curve (AUPRC). This comparison results in selecting the most suitable algorithms including the SVM and ANN.

In the second phase, the selected algorithms are used in comparing selected imbalance classification approaches such as Random Oversampling, One-Class Classification and Cost Sensitive. Then, the SVM is used as a binary classification tool, and compared to the One-Class Classification SVM and Cost Sensitive SVM. Also, the ANN is applied and compared to the Auto-Associative Neural Network.

B] Dataset and Variable Selection

The dataset used in our experiment contains credit card fraud labeled data. It contains ten million credit card transactions described by 8 variables listed here:

- Cust ID is an auto increasing integer value that represents the customer ID: This variable is removed later as it has no relevance for detecting fraud.
- Gender: represents the customer's gender.
- State: represents the state in which the customer lives in the United States.
- Card holder: is the number of cards that the customer holds (maximum 2).
- Balance: indicates the balance on the credit card in USD.
- Num Trans: is a discrete variable that represents the number of transactions made to date.
- NumInt Trans: is a discrete variable representing the number of international transactions made to date.
- Credit Line: denotes the customer's credit limit.

- Fraud Risk: the binary target variable, taking the values 0 denoting legitimate transaction, and 1 denoting fraudulent transaction.

V. CONCLUSION

Credit card fraud becomes a serious concern to the world. Fraud brings huge financial losses to the world. This urged Credit card companies have been invested money to create and develop techniques to reveal and reduce fraud. The prime goal of this study is to define algorithms that confer the appropriate, and can be adapted by credit card companies for identifying fraudulent transactions more accurately, in less time and cost. Different machine learning algorithms are compared, including Logistic Regression, Decision Trees, Random Forest, Artificial Neural Networks, Logistic Regression, K-Nearest Neighbors, and Kmeans clustering. Because not all scenarios are the same, a scenario-based algorithm can be used to determine which scenario is the best fit for that scenario.

VI. REFERENCES

[1]. K. Sim, V. Gopalkrishnan, A. Zimek, and G. Cong -"A survey on enhanced subspace clustering," Data Mining Knowledge Discovery, vol. 26, no. 2, pp. 332–397, 2019.

[2]. S. Mcskimming -"Trade-based money laundering: Responding to an emerging threat," Deakin Law Rev, vol. 15, no. 1, 2020.

[3]. Nitu Kumari, S. Kannan and A. Muthukumaravel - "Credit Card Fraud Detection Using Genetic-A Survey" published by MiddleEast Journal of Scientific Research , IDOSI Publications, 2014

[4]. Satvik Vats, Surya Kant Dubey, Naveen Kumar Pandey - "A Tool for Effective Detection of Fraud in Credit Card System", published in International Journal of Communication Network Security ISSN: 2231 – 1882, Volume-2

[5]. S.H. Projects and W. Lovo , —JMU Scholarly Commons Detecting credit card fraud : An analysis of fraud detection techniques,‖ 2020. [2] S. G and J. R. R, —A Study on Credit Card Fraud Detection using Data Mining Techniques,‖ Int. J. Data Min. Tech. Appl., vol. 7, no. 1, pp. 21–24, 2018, doi: 10.20894/ijdmta.102.007.001.004

[6]. V. N. Dornadula and S. Geetha —Credit Card Fraud Detection using Machine Learning Algorithms, Procedia Comput. Sci., vol. 165, pp. 631–641, 2019, doi: 10.1016/j.procs.2020.01.057.

[7]. A. H. Alhazmi and N. Aljehane - A Survey of Credit Card Fraud Detection Use Machine Learning, 2020 Int. Conf. Computes. Inf. Technol. ICCIT 2020, pp. 10–15, 2020, doi: 10.1109/ICCIT-144147971.2020.9213809

[8]. M. Kanchana, V. Chadda , and H. Jain - Credit card fraud detection,‖ Int. J. Adv. Sci. Technol., vol. 29, no. 6, pp. 2201–2215, 2020, doi: 10.17148/ijarcce.2016.5109. [14] A. RB and S. K. KR, —Credit Card Fraud Detection Using Artificial Neural Network,Glob. Transitions Proc, pp. 0–8,2021,doi: 10.1016/j.gltp.2021.01.006