ISSN: 2395-7852

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

Volume 10, Issue 3, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 6.551

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

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

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

Credit Card Fraud Detection Using Machine Learning

M. Rahini sudha [1] , CH . Inesh Reddy [2] , K . Bharath Reddy[3], Alzeemar Khan[4] , M. Mano[5]

[1] Assistant Professor, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

[2] UG Scholar, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

[3] UG Scholar, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

[4] UG Scholar, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

[5] UG Scholar, Department of Computer Science and Engineering, Velammal Institute of Technology, Panchetti, Chennai, India

ABSTRACT: Credit card fraud detection is presently the most frequently occurring problem in the present world. This is due to the rise in both online transactions and e-commerce platforms. Credit card fraud generally happens when the card was stolen for any of the unauthorized purposes or even when the fraudster usesthe credit card information for his use. In the present world, we are facing a lot of credit card problems Credit card fraud detection is an important study in the current era of mobile payment.Improving the performance of a fraud detection model and keeping its stability are very challengingbecause user's payment behaviors and criminals fraud behaviors are often changing.To detect the fraudulent activities the credit card fraud detection system was introduced. In this article we focus on obtaining Deep learning algorithm of legal and fraud transactions from the aspect of the lossfunction.

KEYWORDS: Credit Card Fraud Detection, Fraud Detection, Convolutional Neural Network

I. INTRODUCTION

Credit card fraud is a growing concern in the present world with the growing fraud in the government offices, corporate industries, finance industries, and many other organizations. In the presentworld, the high dependency on the internet is the reason for an increased rate of credit card fraud transactions but the fraud has increased not only online but also offline transactions. Though the data mining techniques are used the result is not much accurate to detect these credit card frauds. The only way to minimize these losses is the detection of the fraud using efficient algorithms which is a promising way to reduce the credit card frauds. As the use of the internet is increasing, a credit card is issued by the finance company. Having a credit card means that we can borrow the funds. The funds can be used for any of the purposes. When coming to the issuance of the card, the condition involved is that the cardholder will pay back the original amount they borrowed along with the additional charges they agreed to pay.

A credit card is said to be a fraud when some other person uses your credit card instead of you without your authorization. Fraudsters steal the credit card PIN or the account details to perform any of the unauthorized transactions without stealing the original physical card. Using the credit card fraud detection we could find out whether the new transactions are fraud one or a genuine one.

The fraud that is committed may involve the card such as a credit card or debit card. In this, the card itself acts as a fraudulent source in the transaction. The purpose of committing the crime may be to obtain the goods without paying money or to obtain the unauthorized fund. Credit cards are a nice target for fraud.

The reason is that in a very short time a lot of money can be earned without taking many risks and eventhe crime will take many weeks to be detected .In 2022, there were 1,579 data breaches and nearly 179 million records among which Credit card frauds were the most common form with 133,015 reports, then employment or tax-related frauds with 82,051 reports, phone frauds with 55,045 reports followed by bank frauds with 50,517 reports from the statics released by FTC [9].

II. RELATED WORKS

A research about a case study involving credit card fraud detection, where data normalization is applied before Cluster Analysis and with results obtained from the use of Cluster Analysis and Artificial Neural Networks on fraud detection has shown that by clustering attributes neuronal inputs can be minimized. And promising results can be obtained by using normalized data and data should be MLP trained. This research was based on unsupervised learning. Significance of this paper was to find new methods for fraud detection and to increase the accuracy of results. The data set for this paper is based on real life transactional data by a large company and personal details in data is kept confidential.[11] Accuracy of an algorithm is around 50%. Significance of this paper was to find an algorithm and to reduce the cost measure. The result obtained was by 23% and the algorithm they find was Bayes minimum risk

Multiple Supervised and Semi-Supervised machine learning techniques are used for fraud detection [8], but we aim is to overcome three main challenges with card frauds related dataset i.e., strong class imbalance, the inclusion of labelled and unlabeled samples, and to increase the ability to process a largenumber of transactions.

Different Supervised machine learning algorithms [3] like Decision Trees, Naive Bayes Classification, Least Squares Regression, Logistic Regression and SVM are used to detect fraudulent transactions in real-time datasets. Two methods under random forests [6] are used to train the behavioral features of normal and abnormal transactions. They are Random-tree-based random forest and CART-based. Even though random forest obtains good results on small set data, there are still some problems in case of imbalanced data. The future work will focus on solving the above-mentioned problem. The algorithm of the random forest itself should be improved.

Performance of Logistic Regression, K-Nearest Neighbour, and Naïve Bayes are analyzed on highly skewed credit card fraud data where Research is carried out on examining meta-classifiers and meta- learning approaches in handling highly imbalanced credit card fraud data.

Through supervised learning methods can be used there may fail at certain cases of detecting the fraud cases. A model of deep Auto-encoder and restricted Boltzmann machine (RBM) [2] that can construct normal transactions to find anomalies from normal patterns. Not only that a hybrid method is developed with a combination of Ad boost and Majority Voting methods [4].

III. PROPOSED METHOD

The proposed system is based on Convolutional Neural Network with deep learning for classification. This algorithm inspired by biological neural networks (where the brain is considered particularly important in the central nervous system) and are used in statistics and cognitive science. These are represented by the interconnection of neural systems from various input variables to the output, and it can be represented as mathematical functions that are configured to represent complexrelationships between inputs (independent variables) and outputs (dependent variables).

Card transactions are always unfamiliar when compared to previous transactions made the customer. This unfamiliarity is a very difficult problem in real-world when are called concept drift problems [1]. Concept drift can be said as a variable which changes over time and in unforeseen ways.These variables cause a high imbalance in data. The main aim of our research is to overcome the

problem of Concept drift to implement on real-world scenario. Table 1, [1] shows basic features that arecaptured when any transaction is made.

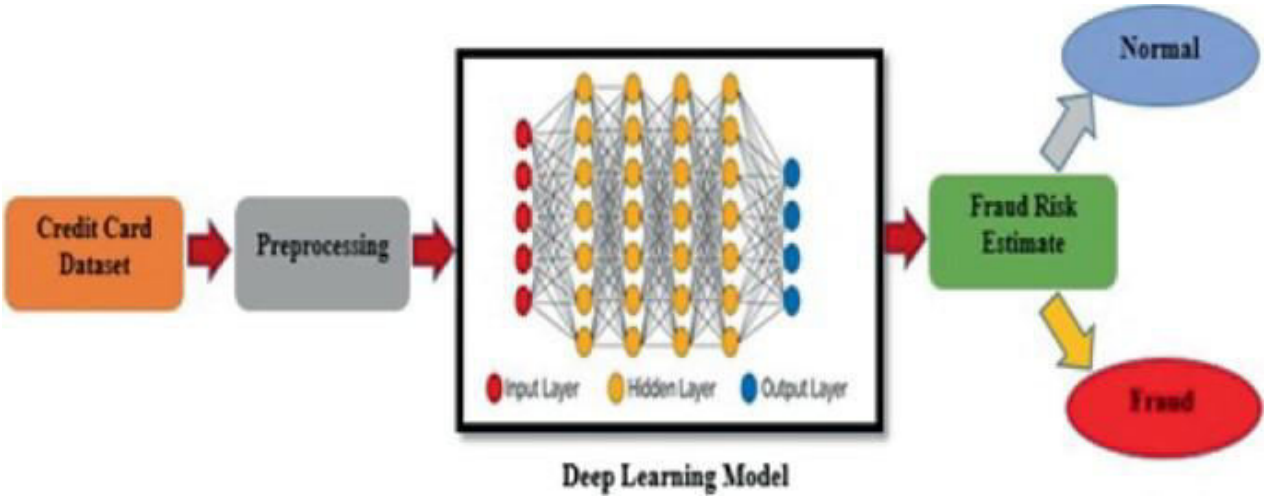

Fig 1. Architecture of proposed method

1. Data Collection: The first step is to collect credit card transaction data from various sources, suchas credit card companies and banks. This data includes transaction date, time, amount, location, cardholder information, and merchant information.

2. Data Pre-processing: Once the data is collected, it needs to be pre-processed to remove duplicates, missing values, and outliers. The pre-processed data is then normalized andconverted into a format that can be used for analysis.

3. Feature Engineering: Feature engineering involves selecting relevant features from the pre- processed data that are useful in detecting fraud. These features include transaction amount,location, time of the day, and cardholder information.

4. Model Selection: There are several deep learning algorithms that can be used for credit card fraud detection, such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) Networks, and Autoencoders. The appropriate algorithm is selected based on the size of the dataset, complexity of the problem, and computational resources available.

5. Training and Testing: The selected model is trained on a labeled dataset that contains bothfraudulent and non-fraudulent transactions. The trained model is then tested on a separatedataset to evaluate its performance.

Deployment: Once the model is trained and tested, it can be deployed in a real-time environment Credit card fraud is a major concern for both credit card companies and cardholders. One effective way toprevent credit card fraud is by developing a fraud detection system that can identify fraudulent transactions in real-time. Deep learning algorithms have shown great potential in detecting fraudulent activities in credit card transactions. Here is a proposed system for credit card fraud detection using deep learning algorithms: to detect fraudulent transactions in real-time. The system can send alerts to credit card companies and cardholders to take necessary actions.

7. Monitoring and Updating: The fraud detection system needs to be monitored and updated regularly to improve its performance and adapt to changing patterns of fraud.

IV. EXPERIMENTAL RESULTS

We have experimented few models on original as well as SMOTE dataset. The results are tabulated, which shows great differences in accuracy, precision and MCC as well. We even used one-class SVM which can be best used for binary class datasets. Since we have 2 classes in our dataset we can useone-class SVM as well.

Table 1, shows the results on the dataset before applying SMOTE and fig 2, shows the same resultsgraphically.

Methods	Accuracy	Precision	MCC
Local Outlier factor	0.8990	0.0038	0.0172
Isolation forest	0.9011	0.0147	0.1047
Support vector machine	0.9987	0.7681	0.5257
Logistic regression	0.9990	0.875	0.6766
Decision tree	0.9994	0.8854	0.8356
Random forest	0.9994	0.9310	0.8268

Table 1: Accuracy, Precision and MCC values before applying SMOTE,

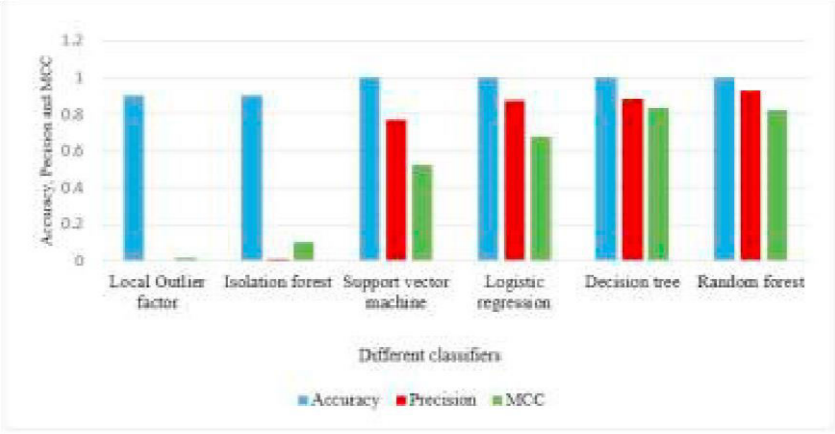

Fig 2: chart showing results on original dataset

One-Class SVM
Accuracy: 0.7009
Precision: 0.7015

Table 2, shows the results on the dataset after applying SMOTE and fig 3, shows the same resultsgraphically.

Methods	Accuracy	Precision	MCC
Local Outlier factor	0.4582	0.2941	0.1376
Isolation forest	0.5883	0.9447	0.2961
Logistic regression	0.9718	0.9831	0.9438
Decision tree	0.9708	0.9814	0.9420
Random forest	0.9998	0.9996	0.9996

Table 2: Accuracy, Precision and MCC values after applying SMOTE,

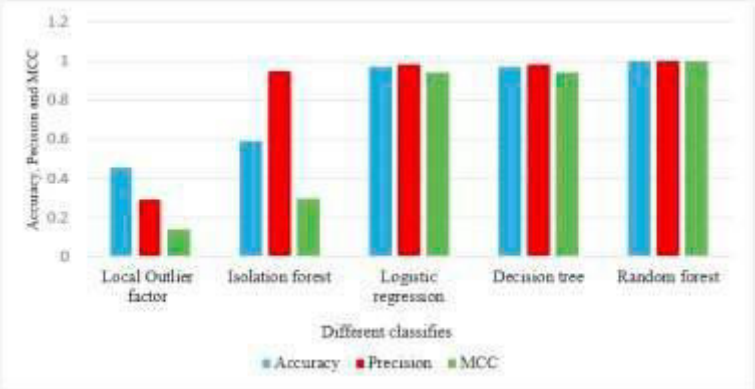

Fig 3: chart showing results on updated dataset

Fig 4, shows the comparison between the values of MCC on dataset before and after applying SMOTE.

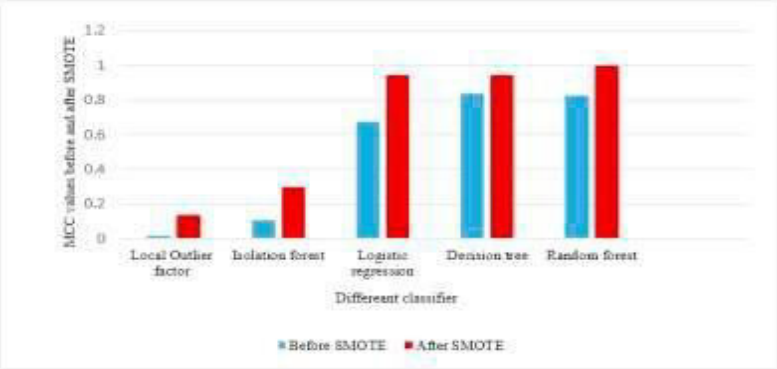

Fig 4: MCC parameter comparison between original and updated dataset

V. CONCLUSION

In this paper we developed a novel method for fraud detection, where customers are grouped based on their transaction and extract behavioral patterns to develop a profile for every cardholder. Then different classifiers are applied on three

different groups later rating scores are generated for every type of classifier. This dynamic changes in parameters lead the system to adapt to new cardholder's transaction behavior's timely. Followed by a feedback mechanism to solve the problem of concept drift. We observed that the Matthews Correlation Coefficient was the better parameter to deal with imbalance dataset. MCC was not the only solution. By applying the SMOTE, we tried balancing the dataset, where we found that the classifiers were performing better than before. The other way of handling imbalance dataset is to use one-class classifiers like one-class SVM. We finally observed that Logistic regression, decision tree and random forest are the algorithms that gave better results.

REFERENCES

1. Jiang, Changjun et al. "Credit Card Fraud Detection: A Novel Approach Using Aggregation Strategy and Feedback Mechanism." *IEEE Internet of Things Journal* 5 (2018): 3637-3647.
2. Pumsirirat, A. and Yan, L. (2018). 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).
3. Mohammed, Emad, and Behrouz Far. "Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study." IEEE Annals of the History of Computing, IEEE, 1 July 2018, doi.ieeecomputersociety.org/10.1109/IRI.2018.00025.
4. Randhawa, Kuldeep, et al. "Credit Card Fraud Detection Using AdaBoost and Majority Voting." IEEE Access, vol. 6, 2018, pp. 14277–14284., doi:10.1109/access.2018.2806420.
5. Roy, Abhimanyu, et al. "Deep Learning Detecting Fraud in Credit Card Transactions." 2018 Systems and Information Engineering Design Symposium (SIEDS), 2018, doi:10.1109/sieds.2018.8374722.
6. Xuan, Shiyang, et al. "Random Forest for Credit Card Fraud Detection." 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 2018, doi:10.1109/icnsc.2018.8361343.
7. Awoyemi, John O., et al. "Credit Card Fraud Detection Using Machine Learning Techniques: A Comparative Analysis." 2017 International Conference on Computing Networking and Informatics (ICCNI), 2017, doi:10.1109/iccni.2017.8123782.
8. Melo-Acosta, German E., et al. "Fraud Detection in Big Data Using Supervised and Semi-Supervised Learning Techniques." 2017 IEEE Colombian Conference on Communications and Computing (COLCOM), 2017, doi:10.1109/colcomcon.2017.8088206.
9. Sailusha R, Gnaneswar V, Ramesh R, Rao GR. Credit card fraud detection using machine learning. In2020 4th international conference on intelligent computing and control systems (ICICCS) 2020 May 13 (pp. 1264-1270). IEEE
10. Tanouz D, Subramanian RR, Eswar D, Reddy GP, Kumar AR, Praneeth CV. Credit card fraud detection using machine learning. In2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) 2021 May 6 (pp. 967-972). IEEE.

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

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

www.ijarasem.com