International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Credit Card Fraud Detection

Vishal Chaudhari

Department of Information Technology, All India Shri Shivaji Memorial Society's, Institute Of Information Technology, Pune– 411001, Maharashtra, India

ABSTRACT

Online payment options have risen due to e-commerce and numerous other websites, which raises the possibility of online fraud. Both the owners of credit cards and financial institutions suffer large financial losses as a result of credit card theft. The primary goal is to create and implement a cutting-edge fraud detection method for real-time transaction data, which will analyse client transaction history. wherein cardholders are grouped according to the size of their transactions. The dataset was first subjected to a machine learning technique, which somewhat increased the accuracy of fraud detection. Later, three convolutional neural network-based designs are used to boost the effectiveness of fraud detection.

INTRODUCTION

Online commerce is continually expanding. Credit cards are used to make purchases of goods and services using both virtual and real cards, with the former being used for offline transactions and the latter for online ones. A form of identity theft known as credit card fraud (CCF) occurs when someone other than the account holder uses a credit card or account information for an unauthorized transaction. Fraud may come from a credit card that has been lost, stolen, or fraudulently fabricated. The growth of e-banking and various online payment environments has led to an increase in fraud, such as CCF, causing billions of dollars in losses annually. The purpose of supervised CCF detection is to build an existing transactional credit card payment model for machine learning (ML).

METHODOLOGY

Table 1 displays the credit card mainframe transaction table and identifies key features. Although the overall design of the transaction information table may vary significantly amongst card issuers, the key details captured would be under database control and available for fraud detection modelling.

Table 1:

No.	Name of Feature	Description
1	Account number	Related with account number
2	Open to buy	The availability of balance
3	Credit Limit	the associated account
4	Card number	Number of Credit card
5	Amount	by the merchant
6	Transaction Time	Time of the transaction
7	Transaction Date	Date of the transaction
8	Transaction Type	cash withdrawal and purchase
9	Currency Code	The currency code
10	Code	The Merchant business type code
11	Merchant Number	The merchant reference number
12	Country	takes place

13	Transaction City	place
14	Approval Code	request, it means approve or reject.

Table 2 presents the detail of the dataset containing 31 columns, including time, V1, V2, V3.....V28 as PCA applied features, amount, and class labels.

Table 2

No	Feature	Description
1.	Time	transaction
2.	attributes	These 28 columns show result of a PCA dimensionality reduction to protect user identities and sensitive features.
3.	Amount	Amount of transaction
4.	Class label	nonfraudulent and fraudulent

The difference between the model's prediction and the rock bottom dataset truth, where TP, TN, FP, and FN stand for true positive, true negative, false positive, and false negative, respectively, can be used as a confusion metric in traditional techniques of estimating ML classifiers.

i) ACCURACY

Accuracy= TP+TN / TP+TN+FP+FN

ii) PRECISION

Precision=TP / TP+FP

iii) F1-SCORE

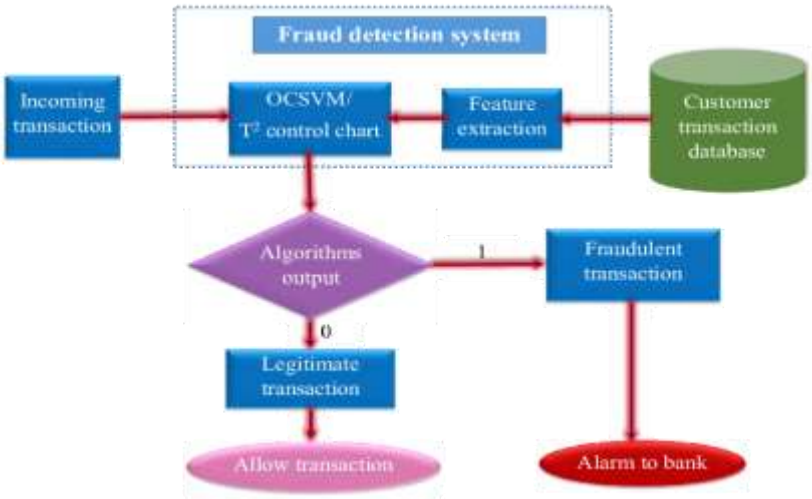
F=2 * Precision * Recall / Precision + Recall

iv) RECALL

Recall= TP / TP + FN

MODELING AND ANALYSIS

1) Diagrammatical design of Credit Card Fraud Detection



2) Difficulties of Credit Card Fraud Detection

- **Imbalanced data:** The credit card fraud detection data has imbalanced nature. It means that very small percentages of all credit card transactions are fraudulent. This cause the detection of fraud detection very difficult.
- **Overlapping data:** Many transactions may be considered fraudulent, while actually they are normal and reversely, a fraudulent transaction may also seem to be (false negative).
- **Fraud detection cost:** The system should take into account both the cost of fraudulent behavior that is detected and the cost of preventing it.
- **Lack of standard metrics:** There is no standard evaluation criterion for assessing and comparing the results of fraud detection systems.

RESULTS AND DISCUSSION

1) **DATA VISUALISATION:**

The dataset includes credit card transactions made by European cardholders in October 2018. Out of the 284,807 transactions in the dataset, 492 were scams that occurred within the last two days. It only includes numerical input variables that come from a PCA transformation. We are unable to provide more background on the original dataset's architecture and data due to the concealment issue. The seconds between the dataset's first transaction and each subsequent transaction are covered by the feature "Time." The distribution of the CCF dataset into fraudulent and nonfraudulent transactions is shown in Figure .

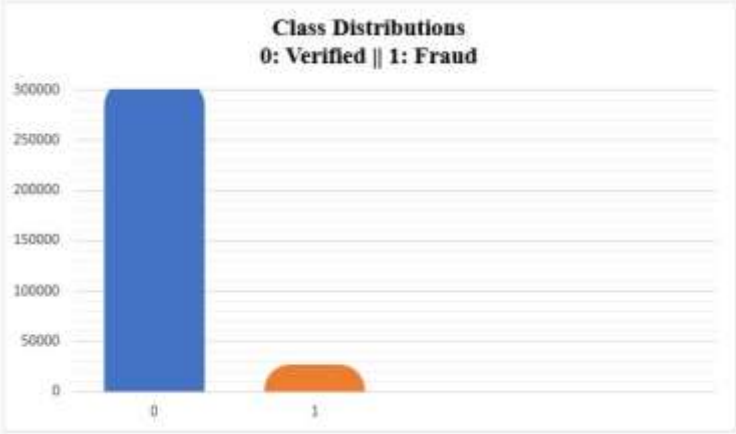


Fig: Class distribution of fraudulent and nonfraud transactions.

2) **TOP 10 ALGORITHMS IN MACHINE LEARNING FORFRAUD DETECTION:**

In the present study , the top ten ML algorithms are incorporated for the detection of credit card frauds. The list of these algorithms is given below:

1. Linear Regression
2. Logistic Regression
3. Decision Tree
4. SVM
5. Naïve Bayes
6. CNN
7. K-Means
8. Random Forest
9. Dimensionality Reduction Algorithms
10. Gradient Boosting Algorithms

These algorithms can also encompass association analysis, clustering, classification, statistical learning, and link mining.

CONCLUSION

The threat posed by CCF to financial institutions is growing. Fraudsters frequently develop novel fraud techniques. A strong classifier can handle the evolving fraud landscape. A fraud detection system's top aim is to accurately anticipate fraud situations while lowering the number of false-positive cases. ML approaches function differently depending on the specific business scenario. Different ML approaches are mostly driven by the type of incoming data. The efficacy of the model for detecting CCF is heavily influenced by the number of features, number of transactions, and correlation between the features.

ACKNOWLEDGEMENTS

With immense pleasure, We present the seminar report as per of the curriculum of the T.E. Information Technology Engineering. We wish to thank all the people who gave us unending support right from when the idea was conceived. I would like to express my deep and sincere gratitude to my seminar guide, **Mr. Pritesh Patil** for allowing me to do this work and providing invaluable guidance. I would like to express my deep gratitude to **Dr. P. B. Mane**, Principal, **Dr. M. A. Thalor**, Head of Department. They were always there with competent guidance and valuable suggestions throughout the pursuance of this presentation. I would like to appreciate all the respondents whose responses and coordination were most important for the presentation and who helped me a lot in collecting necessary information. Above all, no words can express my feeling to my parents, friends, and all those people who supported me during my seminar.

REFERENCES

[1] V. N. Dornadula and S. Geetha, ‘‘Credit card fraud detection using machine learningalgorithms,’’*Proc.Comput.Sci.*,vol.165,pp. 631–641,Jan.2019

[2] N. Kousika, G. Vishali, S. Sunandhana, and M. A. Vijay, ‘‘Machine learning based fraud analysis and detection system,’’ *J. Phys., Conf.*, vol. 1916, no. 1, May 2020

[3] P. Raghavan and N. E. Gayar, ‘‘Fraud detection using machine learning and deep learning,’’ in *Proc. Int. Conf. Comput. Intell Knowl.Economy (ICCIKE)*,Dec.2019

[4] A. O. Balogun, S. Basri, S. J. Abdulkadir, and A. S. Hashim, “Performance analysis of feature selection methods in software defect prediction: A search method approach,” Appl. Sci., vol. 9, no. 13, p. 2764, Jul. 2019.