



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 3, March 2023

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)

# Fraud Detection in Online Payment Services Credit Card Fraud Detection using XGboost Algorithms

SUGANPRIYAN .S, MUHAMMED MITHILAJ, MUHAMMED NAVAL, MUHAMMED NAHAS  
NISHANTH.N

Department of Computer Science and Engineering, Mahendra Institute of Technology, Thiruchengode, Namakkal, Tamil Nadu, India

## ABSTRACT:

Banking industry has the major activity of lending money to those who are in need of money. In order to payback the principle borrowed from the depositor bank collects the interest made by the principle borrowers. Credit risk analysis is becoming an important field in financial risk management. Many credit risk analysis techniques are used for the evaluation of credit risk of the customer dataset. The evaluation of the credit risk datasets leads to the decision to issue the loan of the customer or reject the application of the customer is the difficult task which involves the deep analysis of the customer credit dataset or the data provided by the customer. In this paper we are surveying different techniques for the credit risk analysis which are used for the evaluation for the credit risk datasets. Credit card fraud is a serious problem in financial services. Machine learning algorithm based fraud detection scheme is implemented for detect the fraud card. Hybrid methods which use AdaBoost and majority voting methods are applied. To evaluate the model efficacy, a publicly available credit card data set is used. Then, a real-world credit card data set from a financial institution is analyzed.

## I. INTRODUCTION

Credit card fraud is a wide-ranging term for theft and fraud committed using or involving a payment card, such as a credit card or debit card, as a fraudulent source of funds in a transaction. The purpose may be to obtain goods without paying, or to obtain unauthorized funds from an account. Credit card fraud is also an adjunct to identity theft. According to the United States Federal Trade Commission, while the rate of identity theft had been holding steady during the mid-2000s, it increased by 21 percent in 2008. However, credit card fraud, that crime which most people associate with ID theft, decreased as a percentage of all ID theft complaints for the sixth year in a row.

Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on models and inference instead. It is seen as a subset of artificial intelligence. Machine learning algorithms build a mathematical model of sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task. Machine learning algorithms are used in the applications of email filtering, detection of network intruders, and computer vision, where it is infeasible to develop an algorithm of specific instructions for performing the task. Machine learning is closely related to computational statistics, which focuses on making predictions using computers. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a field of study within machine learning, and focuses on exploratory data analysis through unsupervised learning. In its application across business problems, machine learning is also referred to as predictive analytics.

## II. LITRATURE SURVEY

**2.1 Title:** Use of optimized Fuzzy C-Means clustering and supervised classifiers for automobile insurance fraud detection

**Authors:** Sharmila Subudhi, SuvasiniPanigrahi

**Year:** 2020



**Description:**

This paper presents a novel hybrid approach for detecting frauds in automobile insurance claims by applying Genetic Algorithm (GA) based Fuzzy C-Means (FCM) clustering and various supervised classifier models. Initially, a test set is extracted from the original insurance dataset.

**2.2. Title:** credit card fraud detection using self- Organizing maps

**Authors:** Vladimir ZASLAVSKY and Anna STRIZHAK

**Year:** 2019

**Description:**

Nowadays, credit card fraud detection is of great importance to financial institutions. This article presents an automated credit card fraud detection system based on the neural network technology. The authors apply the Self-Organizing Map algorithm to create a model of typical cardholder's behavior and to analyze the deviation of transactions. Thus finding suspicious transactions.

**2.3. Title:** Evaluating Credit Card Transactions in the Frequency Domain for a Proactive Fraud Detection Approach

**Authors:** Roberto Saia and Salvatore Carta

**Year:** 2019

**Description:**

The massive increase in financial transactions made in the e-commerce field has led to an equally massive increase in the risks related to fraudulent activities. It is a problem directly correlated with the use of credit cards, considering that almost all the operators that offer goods or services in the e-commerce space allow their customers to use them for making payments.

**2.4. Title:** Detecting credit card fraud by genetic algorithm and scatter search

**Authors:** Ekrem Duman, M. Hamdi Ozcelik

**Year:** 2019

**Description:**

In this study we develop a method which improves a credit card fraud detection solution currently being used in a bank. With this solution each transaction is scored and based on these scores the transactions are classified as fraudulent or legitimate. In fraud detection solutions the typical objective is to minimize the wrongly classified number of transactions. However, in reality, wrong classification of each transaction do not have the same effect in that if a card is in the hand of fraudsters its whole available limit is used up. Thus, the misclassification cost should be taken as the available limit of the card.

### III. EXISTING SYSTEM

Our existing has made a detail study on fraud detection using the method of natural observation of the events happened from the customer side. The existing has worked on the methods of collecting the data from the social media and framing them in terms of big data models and working on the challenges existed the field. Implemented a system which supports in the detection of the scams or frauds in the field of the business by recording the transactions and there by building a model using data mining models. In existing system AdaBoost algorithm has been implemented. The Random Forest (RF) algorithm has been implemented. Loss from credit card fraud affects the merchants, where they bear all costs, including card issuer fees, charges, and administrative charges.

#### DRAWBACKS

- Difficult to identify
- Insecure

### IV. PROPOSED SYSTEM

The proposed system first step is data collection in this step data collection from Kaggle. After collecting data to processing using machine learning algorithm. The processing data is converted into principle component analysis. In this step using 10 different component. Data splitting process to split a data from principle component analysis and k-fold cross validation used. Model development process using machine learning algorithm to train a data to send next step of the process. Performance evaluation implemented this project using formulas. Finally we deployed the project using model selection. Majority voting is frequently used in data classification, which involves a combined model with at least two algorithms.

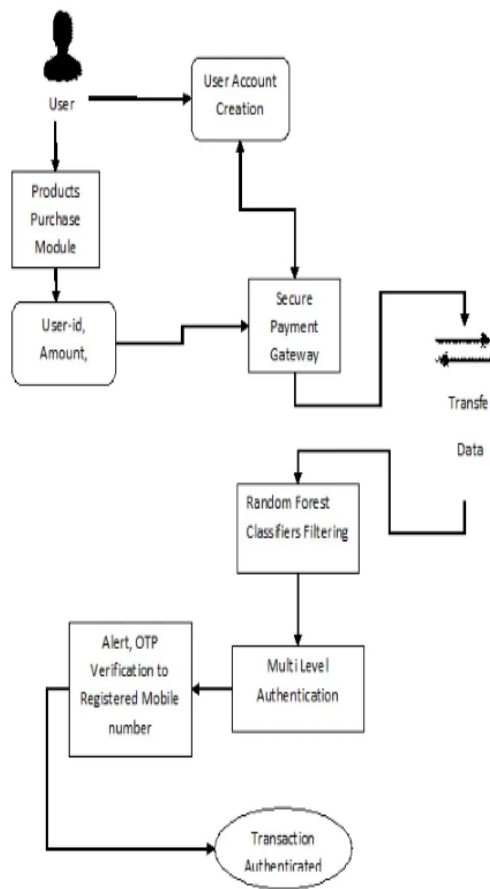


Machine learning based algorithm is implemented for fraud detection. Each algorithm makes its own prediction for every test sample. It will be extended to online learning models. In addition, other online learning models will be investigated. The use of online learning will enable rapid detection of fraud cases, potentially in real-time. In the proposed system the KNN algorithm is used. The k-nearest neighbor algorithm is a non-parametric method used for classification and regression. By using KNN the processing time is reduced and also process the larger datasets.

**ADVANTAGES:**

- Stable system
- More accuracy
- Less time to predict the fraud

**BLOCK DIAGRAM**



**V.RESULTS**



CREDIT CARD FRAUD DETECTION

#	Title	Value
1	Card No	544394704899
2	Location	Thani
3	Last Transaction Amount	4800
4	Status	Non Fraud





CREDIT CARD FRAUD DETECTION

#	Title	Value
1	Card No	411091937853
2	Location	Erode
3	Last Transaction Amount	2300
4	Status	Fraud



CREDIT CARD FRAUD DETECTION

#	Title	Value
1	Card No	175623709144
2	Location	Ranipet
3	Last Transaction Amount	2200
4	Status	Non Fraud



CREDIT CARD FRAUD DETECTION

#	Title	Value
1	Card No	175623709144
2	Location	Ranipet
3	Last Transaction Amount	2200
4	Status	Non Fraud

### V.CONCLUSIONS

It is worth keep in mind that objective of the paper is to surveying on the different classifier which are used in the credit risk evaluation. In this paper different types of classifiers are discussed and also different types of ensemble classifiers are briefed. The dataset which are used in the classifier is discussed in the paper. We have analyzed and compare their accuracies using different types classifiers and from comparison table we found that the ELM classifier gives better accuracies compare to other classifiers that is ELM gives 96.33(%) in German dataset and 96.32(%) in Australian dataset.

### VI.FUTURE WORK

The methods studied in this paper will be extended to online learning models. In addition, other online learning models will be investigated. The use of online learning will enable rapid detection of fraud cases, potentially in real-time. This in turn will help detect and prevent fraudulent transactions before they take place, which will reduce the number of losses incurred every day in the financial sector.

### REFERENCES

[1] Y. Sahin, S. Bulkan, and E. Duman, "A cost-sensitive decision tree approach for fraud detection," Expert Systems with Applications, vol. 40, no. 15, pp. 5916–5923, 2013.

[2] A. O. Adewumi and A. A. Akinyelu, "A survey of machine-learning and nature-inspired based credit card fraud detection techniques," International Journal of System Assurance Engineering and Management, vol. 8, pp. 937–953, 2017.



- [3] The Nilson Report (October 2016) [Online]. Available: [https://www.nilsonreport.com/upload/content\\_promo/The\\_Nilson\\_Report\\_10-17-2016.pdf](https://www.nilsonreport.com/upload/content_promo/The_Nilson_Report_10-17-2016.pdf)
- [4] N. S. Halvaiee and M. K. Akbari, "A novel model for credit card fraud detection using Artificial Immune Systems," *Applied Soft Computing*, vol. 24, pp. 40–49, 2014.
- [5] N. Mahmoudi and E. Duman, "Detecting credit card fraud by modified Fisher discriminant analysis," *Expert Systems with Applications*, vol. 42, no. 5, pp. 2510–2516, 2015.
- [6] R. Saia and S. Carta, "Evaluating Credit Card Transactions in the Frequency Domain for a Proactive Fraud Detection Approach," In *Proceedings of the 14th International Joint Conference on e-Business and Telecommunications*, vol. 4, pp. 335–342, 2017.
- [7] E. Duman, A. Buyukkaya, and I. Elikucuk, "A novel and successful credit card fraud detection system Implemented in a Turkish Bank," In *IEEE 13th International Conference on Data Mining Workshops (ICDMW)*, pp. 162–171, 2013.
- [8] M. Seera, C. P. Lim, K. S. Tan, and W. S. Liew, "Classification of transcranial Doppler signals using individual and ensemble recurrent neural networks," *Neurocomputing*, vol. 249, pp. 337-344, 2017.
- [9] Y. Li, C. Yan, W. Liu, and M. Li, "A principle component analysis- based random forest with the potential nearest neighbor method for automobile insurance fraud identification," *Applied Soft Computing*, to be published. DOI: 10.1016/j.asoc.2017.07.027.
- [10] F. H. Chen, D. J. Chi, and J. Y. Zhu, "Application of Random Forest, Rough Set Theory, Decision Tree and Neural Network to Detect Financial Statement Fraud–Taking Corporate Governance into Consideration," In *International Conference on Intelligent Computing*, pp. 221–234, Springer, 2014.





**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.118**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details