Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 46 (2025), No. 1, pp. 75-79
DOI : 10.47974/JIOS-1853

Improved security in credit cards via duplicitous contract detection

Suman Yadav [†]
Ruchi Sharma *
Annu Dabas [§]
Department of Electronics and Communication Engineering
Bharati Vidyapeeth's College of Engineering
Paschim Vihar 110063
New Delhi
India

Charu Arora [‡]
Department of Applied Sciences
Bharati Vidyapeeth's College of Engineering
Paschim Vihar 110063
New Delhi
India

Gaurav Mitra [@]
Apoorva Aggarwal [#]
Department of Electronics and Communication Engineering
Bharati Vidyapeeth's College of Engineering
Paschim Vihar 110063
New Delhi
India

Arvind Rehalia [$]
Department of Information and Technology Engineering
Bharati Vidyapeeth's College of Engineering
Paschim Vihar 110063
New Delhi
India

[†] *E-mail:* suman.yadav@bharatividyapeeth.edu
* *E-mail:* ruchi.sharma@bharatividyapeeth.edu (Corresponding Author)
[§] *E-mail:* annu.dabas@bharatividyapeeth.edu
[‡] *E-mail:* charu.arora@bharatividyapeeth.edu
[@] *E-mail:* gaurav.mitra@bharatividyapeeth.edu
[#] *E-mail:* apoorva.aggarwal@bharatividyapeeth.edu
[$] *E-mail:* arvind.rehalia@bharatividyapeeth.edu

Sanjeev Sanyal [^]
Department of Computer Science
IMS Engineering College
Ghaziabad
Uttar Pradesh
India

Monika Arora [%]
Department of Mathematics
Miranda House
University of Delhi
Delhi 110007
India

Abstract

Credit card fraud is among the most prominent financial frauds in the ever-growing industry. Credit card firms must detect fraudulent transactions to ensure clients are not billed for products they did not purchase. With technological advancements, fraudulent transactions have increased, driven by the proliferation of online payment options. Machine learning algorithms play a pivotal role in detecting fraud by analyzing transaction behavior. This study presents a model achieving 99.92% accuracy using techniques like Random Forest Classifier, SVC, and SGD Classifier. The model employs dataset preprocessing and sampling techniques such as SMOTE and SMOTEENN. Our findings highlight effective fraud detection mechanisms and their relevance to the financial industry.

Subject Classification: *94A15.*

Keywords: *Machine learning, Random forest classifier, Credit card fraud detection, Pipelines, SMOTE, SVC, SGD classifier.*

1. Introduction

Credit card fraud refers to unauthorized transactions performed without the user's consent. E-commerce growth has increased online fraud likelihood. Researchers have explored various machine learning (ML) methods to identify such fraud, emphasizing the need for efficient detection mechanisms. This paper investigates different ML algorithms to develop a near-perfect fraud detection model by analyzing transaction patterns. The fraud taxonomy spans financial fraud, phishing, telecom fraud, and more.

[^] *E-mail:* **sanjeev.sanyal@imsec.ac.in**
[%] *E-mail:* **monika.arora@mirandahouse.ac.in**

Fraudulent transactions pose challenges that ML algorithms can address by analyzing large datasets. Techniques like Bayesian methods, Artificial Neural Networks, and Random Forests have been explored. This study proposes an effective model for detecting fraudulent credit card transactions, achieving high accuracy.

2.Literature Review

Several researchers have proposed models for credit card fraud detection:

- **Awoyemi et al.** [1] utilized Naive Bayes and logistic regression, achieving 97.92% accuracy.
- **Ogwueleka** [2] achieved 100% accuracy using Artificial Neural Networks with a purely fraudulent dataset.
- **Xuan et al.** [3] implemented Random Forest, achieving 98.67% accuracy.
- **Singh et al.** [4] employed SVM for fraud detection, comparing models like logistic regression, Naive Bayes, and ensemble learning.

These studies highlight the importance of robust ML techniques and reliable datasets in fraud detection.

3. Proposed Methodology

The proposed model incorporates multiple ML pipelines to address the imbalanced dataset challenge, where legitimate transactions vastly outnumber fraudulent ones. Key steps include:

3.1 *Dataset*

The dataset, "creditcard.csv," contains 284,807 transactions, with 492 fraudulent. It includes 31 parameters, anonym zed to protect user privacy.

3.2 *Sampling Techniques*

- **SMOTE (Synthetic Minority Oversampling Technique)**: Oversamples minority classes to balance the dataset.
- **SMOTEENN**: Combines over- and under-sampling techniques to refine dataset balance.

3.3 *Classifiers*

- **Random Forest Classifier**: A versatile algorithm excelling in classification tasks.
- **SVC**: Effective for multi-class and binary classification.
- **SGD Classifier**: Employs stochastic gradient descent, suitable for large-scale data.

4. Implementation

The implementation process involves:

- **Data Preprocessing**: Removing null values and discrepancies.
- **Pipeline Creation**: Employing three pipelines with Random Forest, SVC, and SGD classifiers.
- **Training and Testing**: Combining classifier outputs using a dictionary to enhance efficiency.

Accuracy metrics for the classifiers are summarized below:

Classifier	Scaling	Sampling Technique	Accuracy
SGD Classifier	Standard Scalar	SMOTE	97.61%
Random Forest	Standard Scalar	SMOTE	99.92%
SVC	Standard Scalar	SMOTE	98.80%

5. Conclusion and Future Scope

This study successfully implemented ML pipelines for fraud detection, achieving 99.92% accuracy with Random Forest Classifier. Future work could focus on integrating advanced ensemble techniques and exploring larger, more diverse datasets to achieve even higher accuracy and reliability

References

[1] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," in *Proc. IEEE Int. Conf. Comput. Netw. Informatics*, Lagos, Nigeria (2017).

[2] F. N. Ogwueleka, "Data mining application in credit card fraud detection system," *J. Eng. Sci. Technol.*, vol. 6, no. 3, pp. 311–322 (2011).

[3] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, "Random forest for credit card fraud detection," in *Proc. IEEE Int. Conf. Netw., Sens. Control*, Zhuhai, China (2018).

[4] G. Singh, R. Gupta, and M. D. S. Chandel, "A machine learning approach for detection of fraud based on SVM," *Int. J. Sci. Eng. Technol.*, vol. 2, no. 5, pp. 455–459 (2012).

Received August, 2024