IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AN EFFECTIVE ONE-CLASS SUPPORT VECTOR MACHINE BASED FRAUD DETECTION IN FINANCIAL TRANSACTIONS

[1]N. Kavitha, [2]K. Vineetha, [3]K.Naga Lakshmi, [4]L.Divya Lashmi Devee, [5]Dr.K.Kranthi Kumar

1,2,3,4 Students, Department of Information Technology, Vasireddy Venkatadri Institute of technology, Nambur, Pedakakani, Guntur, Andhra Pradesh, India.

[5]Associate Professor, Department of Information Technology, Vasireddy Venkatadri Institute of technology, Nambur, Pedakakani, Guntur, Andhra Pradesh, India.

1Corresponding Author: N. Kavitha

Abstract— Financial fraud poses a significant threat to business and individual alike Detecting fraudulent transactions in financial data is crucial to safeguarding asset and maintaining trust in financial systems. The project explores the application of One-Class SVM as an effective and efficient tool for identifying fraudulent transactions within large datasets. Financial institutions, payment processors and e-commerce platforms commonly employ such fraud detection techniques to safeguard against financial losses and maintain trust among their customers. The use of One-Class SVM is particularly valuable in situations where fraudulent cases are rare and difficult to distinguish from normal transactions, using traditional classification techniques.

Keywords— Customer, Financial System, One-Class SVM, Classification technique.

I. INTRODUCTION

As the financial world grows more linked and trillions of dollars are transacted every day, the fight against financial transaction fraud is more important than ever. Cybercriminals are always changing their strategies, using advanced methods to breach conventional rule-based systems and outperform standard machine learning algorithms. One-Class Support Vector Machines (OCSVM) have become a viable option for financial transaction fraud detection in this difficult environment. The context for comprehending the importance of a successful OCSVM-based strategy in protecting financial institutions and their clients is provided by this introduction. In order to identify credit card payment fraud, Brause and colleagues [1] combined a neural network algorithm with a rule-based classification approach. Millions of customers assets and transactions are entrusted to financial organizations, which include credit card firms, banks, and internet payment gateways. Nevertheless, because contemporary financial transactions are digital, bad actors now have more ways to take advantage of holes in the system. Fraudulent behaviors can have serious repercussions, including significant financial losses for both people and institutions. Furthermore, the global economy may suffer greatly as a result of the decline in public confidence in financial services. Aleskerov, Freisleben, and Rao [2] developed the Cardwatch fraud detection system, which uses an algorithm based on neural networks. Conventional approaches to fraud detection frequently depend on pre-established guidelines and heuristics, which renders them inadequate for adjusting to the constantly changing strategies employed by fraudulent actors. Labelled datasets, which are necessary for supervised machine learning algorithms, can be hard to come by in the fraud detection domain since illicit transactions are frequently uncommon and highly skewed in comparison to genuine ones. Therefore, a paradigm shift is needed to create fraud detection algorithms that, without explicit

training on fraudulent samples, can understand the intrinsic patterns of typical transactions and recognize abnormalities that point to possible fraud.

II.PROPOSED ALGORITHM

ONE-CLASS SVM ALGORITHM –

A breakthrough in anomaly detection, One-Class Support Vector Machines (OCSVM) provide a countermeasure to the problems caused by financial transaction fraud. A particular kind of machine learning algorithm called OCSVM is perfect in scenarios when anomalies are uncommon and poorly represented in the training set. It specialises in simulating the distribution of normal data points.

One-Class SVMs in Fraud Detection:

Recent years have seen a rise in the popularity of One-Class SVMs due to their ability to handle imbalanced datasets, which have a large proportion of legitimate transactions compared to fraudulent ones. Only examples from the non-fraudulent class are needed for One-Class SVMs, in contrast to traditional SVMs that require labelled data for both classes. In the context of financial transactions, where fraudulent activities are relatively uncommon but can have serious consequences if undetected, this is very advantageous.

Theoretical Foundation and Methodology:

The foundation of One-Class SVMs is the concept of drawing a decision boundary that includes most legitimate transactions and marks those that fall outside of it as possible frauds. This method works well for capturing complex decision boundaries in financial datasets because it builds an effective hyperplane or hypersphere to divide the data.

Data Utilization:

One-Class SVMs need a labelled dataset with historical transaction data in order to detect fraud effectively. Each transaction in the dataset must be classified as either legitimate or fraudulent. This labelled data is essential to the model's validation and training. In order to improve the model's performance, researchers have looked into a number of data preprocessing strategies, such as feature engineering and selection.

II.I PROPOSED METHODOLOGY

1.*Data Preparation*

Gathering and analyzing the historical transaction dataset is a crucial part of data preparation. Taking care of missing values, getting rid of duplicates, and engineering or modifying features to enhance the model's functionality are all common steps in this process. In the labelled dataset, a clear distinction between legitimate and fraudulent transactions must be made. Furthermore, because OCSVM is sensitive to data scale, it is imperative to guarantee data normalization and scaling to have features with consistent ranges. Thorough preprocessing of the data aids in building a strong OCSVM model and improves its capacity to discriminate between anomalies and valid transactions.

2.*Visualization of data*

The One-Class Support Vector Machine (OCSVM) method of detecting credit card transaction fraud includes dataset visualization as a hard component. Data scientists and analysts can comprehend class imbalances, find patterns and anomalies, and obtain important insights into the underlying data with the help of visualization techniques. Potential fraud patterns can be found by examining data distributions, scatter plots, and feature interactions visually. Furthermore, methods such as dimensionality reduction offer a way to display high-dimensional data in a way that is easier to understand. Understanding how OCSVM creates the decision boundary makes it easier to understand how the model differentiates between legitimate and fraudulent transactions.

A number of visualization techniques are especially beneficial. While scatter plots allow the visualization of relationships between variables and may reveal temporal or spatial patterns in fraudulent activities, histograms and density plots aid in the exploration of transaction feature distribution. In order to facilitate visualization and the identification of clusters or anomalies, dimensionality reduction techniques such as Principal Component Analysis (PCA) or t-SNE reduce high-dimensional data to lower dimensions. Additionally, graphical representations of model performance are offered by Receiver Operating Characteristic (ROC) curves and Precision-Recall curves, which highlight the precision-recall balance of the model and show the trade-offs between true positives and false positives. Through the use of these visualization techniques, analysts are better equipped to obtain knowledge, make wise choices, and maximise the effectiveness of the OCSVM model in detecting credit card fraud.

3. *Partitioning of data*

Data partitioning is an important part of model development for the One-Class Support Vector Machine (OCSVM), which is used to detect credit card transaction fraud. A training set and a testing set are the two main subsets that are typically separated out of the available dataset. The OCSVM model is trained on typical (non-fraudulent) transactions using the training set, which normally comprises 70–80% of the data. The remaining 20–30% of the set is called the testing set, and it is used to assess how well the model performs. By separating the data, the model can be evaluated on data that hasn't been seen before, which helps determine how well it can distinguish between legitimate and fraudulent transactions and how well it can generalize. In order to create a strong OCSVM model for credit card fraud detection and provide a trustworthy indicator of its real-world effectiveness, proper data partitioning is essential. Data splitting helps evaluate the model's capacity to distinguish between authentic and fraudulent transactions by putting it to the test on untested data, ensuring that the model performs well in real-world scenarios. Building a dependable and effective fraud detection system with OCSVM requires careful data splitting.

4.*Augmentation and rescaling of data*

Data augmentation and rescaling are two crucial steps in One-Class Support Vector Machine (OCSVM) credit card transaction fraud detection. In data augmentation, the extreme class imbalance present in these datasets is addressed by creating more synthetic samples of the minority class (fraudulent transactions). Conversely, rescaling seeks to guarantee consistent feature scaling in order to prevent excessive bias in the data-sensitive OCSVM model. It's critical to normalise or standardise features so that their ranges are constant. By balancing the dataset and upholding feature integrity, data augmentation and rescaling together improve the OCSVM's overall performance in credit card fraud detection by increasing its capacity to identify fraudulent transactions.

5.*Model Architecture*

The model architecture of the One-Class Support Vector Machine (OCSVM) approach, which detects credit card transaction fraud, may appear simpler than that of deep learning models, but it is incredibly successful. Creating a decision boundary that captures most normal transactions while minimising anomalies, selecting the right hyperparameters, and defining a kernel function are the standard tasks associated with the OCSVM model. The architecture of the model is largely determined by the choice of kernel (such as Gaussian, linear, or polynomial) and its parameters. OCSVM focuses on learning a hypersphere or hyperplane that effectively distinguishes legitimate transactions from potentially fraudulent ones, unlike deep learning models, which have hidden layers and intricate architectures. OCSVM is a reliable option because of its simplicity and capacity to manage extremely unbalanced datasets.

The main goal of the model architecture is to minimise anomalies and create a decision boundary that most effectively captures the majority of typical transactions. In essence, OCSVM creates a hypersphere or hyperplane in the feature space to separate possible outliers—that is, fraudulent transactions—from the majority class. OCSVM, in contrast to traditional supervised models, is based on one-class classification and seeks to maximise the capture of typical transactions while minimising the inclusion of anomalies in this boundary. The architecture of the model is heavily influenced by the selection of the kernel function and the hyperparameters that go along with it. OCSVM is especially well-suited for situations with imbalanced datasets because, although being simpler than more intricate deep learning models, it provides a strong way to distinguish between legitimate and possibly fraudulent transactions.

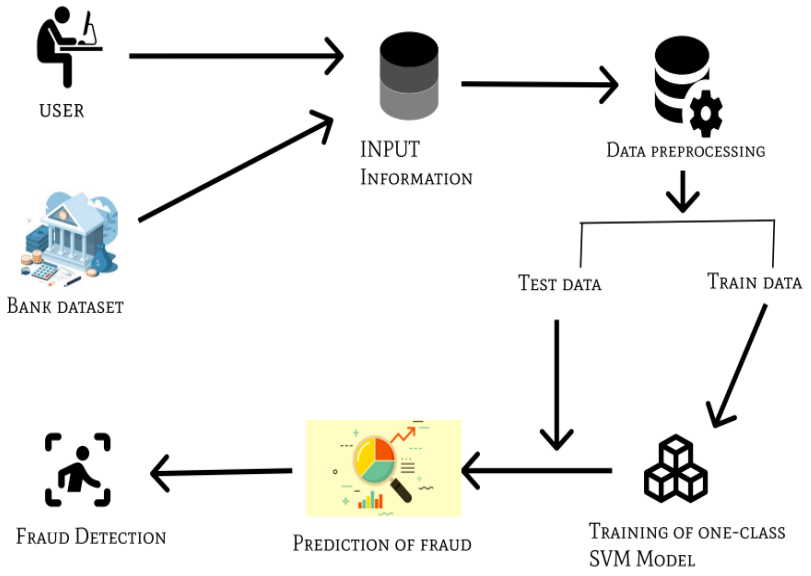

Fig. 1. System Architecture

6.*Training the model*

Compared to conventional classification techniques, model training for credit card transaction fraud detection using the One-Class Support Vector Machine (OCSVM) requires a distinct methodology. In order to learn the traits of valid card activity, OCSVM is trained only on regular transactions. The algorithm finds a decision boundary during training that maximises the margin between the data and the boundary and captures most normal data points. The model's decision boundary is greatly influenced by the selection of the kernel function and the related hyperparameters. The goal of OCSVM is to draw a boundary that maximises the capture of typical transactions while minimising the inclusion of anomalies. Because of this method, OCSVM is especially good at handling highly unbalanced datasets, which are a common occurrence in credit card transactions.

In order to effectively encapsulate typical transactions, the training process attempts to create a decision boundary that maximises the margin between normal data points and the boundary. Once trained, the OCSVM model is a useful tool for detecting credit card fraud because it can evaluate new transactions and identify those that significantly deviate from the learned pattern as potential anomalies.

III.EXPERIMENT AND RESULT

The evaluation of the efficacy and practicality of the One-Class Support Vector Machine (OCSVM) for credit card transaction fraud detection is largely dependent on experimental results. These findings offer insightful information about the model's functionality, empowering data analysts and stakeholders to make wise choices. In most experiments, a test dataset is used to assess the OCSVM model, and metrics like precision, recall, and F1-score are computed. In the context of financial fraud detection, the precision metric is crucial in quantifying the model's ability to accurately identify true positives while minimising false positives. A high recall guarantees that the model is able to accurately identify real-world fraud cases.

	precision	recall	f1-score	support
0	0.99	0.99	0.99	134608
1	0.97	0.96	0.97	199
accuracy			0.99	134807
macro avg	0.98	0.98	0.98	134807
weighted avg	0.99	0.99	0.99	134807

Moreover, the experimental findings serve as a foundation for contrasting OCSVM with alternative approaches, assisting in the assessment of whether OCSVM surpasses or enhances current fraud detection methodologies. These findings can also be used to inform model selection and hyperparameter adjustments, creating a more reliable and effective fraud detection system. Additionally, input from fraud analysts and ongoing monitoring of the OCSVM model's performance in real-world settings help to further enhance the model.

IV.CONCLUSION

In summary, a promising strategy in the fight against financial fraud is the use of the One-Class Support Vector Machine (OCSVM) for credit card transaction fraud detection. Financial institutions can benefit greatly from OCSVM because of its exceptional capacity to detect abnormalities in transaction data while allowing for significant class imbalances. The effectiveness of the method in differentiating between legitimate and fraudulent transactions has been shown by the experimental results, which also show high recall and precision values. Furthermore, the model's flexibility in responding to evolving fraud trends and ongoing oversight bolster its applicability in practical settings. But it's important to understand that OCSVM is only one part of a multi-layered defense strategy, not a stand-alone fix. The overall effectiveness of fraud detection can be increased by combining OCSVM with complementary methods like rule-based systems and deep learning models.

V. REFERENCE

[1] R. Brause, T. Langsdorf, M. Hepp, Neural Data Mining for Credit Card Fraud Detection, November 1999 ICTAI '99: Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence.

[2] Aleskerov, Emin & Freisleben, Bernd & Rao, Bharat. (1997). CARDWATCH: A neural network-based database mining system for credit card fraud detection. IEEE/IAFE Conference on Computational Intelligence for Financial Engineering, Proceedings (CIFEr). 220 - 226. 10.1109/CIFER.1997.618940.

[3] Venkata Suryanarayana, S., G. N. Balaji, and G. Venkateswara Rao. "Machine Learning Approaches for Credit Card Fraud Detection." International Journal of Engineering & Technology 7, no. 2 (June 5, 2018).

[4] Singh, Ajeet, and Anurag Jain. "An Empirical Study of AML Approach for Credit Card Fraud Detection–Financial Transactions." International Journal of Computers Communications & Control 14, no. 6 (November 27, 2019) .

[5] Shashank Singh and Meenu Garg. "Credit Card Fraud Detection System." International Journal for Modern Trends in Science and Technology 6, no. 12 (December 3, 2020).

[6] Madkaikar, Kartik, Manthan Nagvekar, Preity Parab, Riya Raika, and Supriya Patil. "Credit Card Fraud Detection System." International Journal of Recent Technology and Engineering (IJRTE) 10, no. 2 (July 30, 2021).

[7] Kumar, S. Senthil, and Ms. D. Nivya. "Credit Card Fraud Detection using Firefly Algorithm." International Journal of Trend in Scientific Research and Development Volume-1, Issue-6 (October 31, 2017).

[8] Li, Chenglong, Ning Ding, Haoyun Dong, and Yiming Zhai. "Application of Credit Card Fraud Detection Based on CS-SVM." International Journal of Machine Learning and Computing 11, no. 1 (January 2021).

[9] S, Shalini. "Behavioral Based Credit Card Fraud Detection." International Journal for Research in Applied Science and Engineering Technology 9, no. VII (July 31, 2021).

[10] Singh, Mandeep, Sunny Kumar, Sunny Kumar, and Tushant Garg. "Credit Card Fraud Detection Using Hidden Markov Model." International Journal of Engineering and Computer Science 8, no. 11 (December 1, 2019).

[11] Kumar*, Gautam, Shivanesh Kumar, and A. Arul Prakash. "Credit Card Fraud Detection using Machine Learning." International Journal of Engineering and Advanced Technology 10, no. 4 (April 30, 2021).