Available online @ https://jjem.jnnce.ac.in
https: www.doi.org/10.37314/JJEM.SP0240
Indexed in International Scientific Indiexing (ISI)
Impact factor: 1.395 for 2021-22
Published on: 08 December 2024

Detecting credit card fraud using machine learning and deep learning techniques.

Tarun C L[1], Mr. Arun Kumar K L[2]

[1*]Student, [2]Assistant Professor,Dept of MCA JNNCE
tarun.cl07@gmail.com, arunkumarkl@jnnce.ac.in

Abstract

The internet gets utilized a lot more these days and has become a necessity in the modern world the simplicity and flexibility of purchasing and selling goods online has expanded along with e-commerce the advent of modern technologies such as online banking and credit card payments has led to a significant rise in the utilization of online shopping and bill having to pay the risk to utilizing a credit card increased as e-commerce and shopping became more popular this could be that credit cards were used in greater domains making it more difficult for any business related organizations to orient between legitimate and fraudulent transactions furthermore a disguise in the transaction may arise the prototype what we have been discussing shows how this study uses machine learning to detect fraud and block payments in real-time. It demonstrates how abnormalities may be found in an unsupervised way with greater accuracy. This anomaly detection approach has several uses, including identifying non-legal banking and overbilling in telecoms. security monitoring network traffic health care and a variety of manufacturing industries

This work is a critical application of machine learning and neural networks, aimed at "detecting fraudulent activities. from a vast number of legitimate ones.However, one of the highest machine learning approaches such as logistic regression and Decision trees are often usedDeep learning techniques.Credit card companies use machine learning to reduce false positives in fraud detection,Real-time fraud detection systems help prevent unauthorized credit card use.

Keywords:*detecting fraud in credit cards, fraudulent transactions.*

1. Introduction

e-commerce is growing progressively integral to everyday life for individuals vast quantities of data are produced these days as a result of the significant increase in internet and iot device execution as further information is collected in various formats such as structured and semi-structured.[4] Digitalisation is becoming increasingly popular due to the seamless, accessible, and convenient use of e-commerce. With the exponential increase in internet usage, numerous organisations,including the financial industry, have operationalised online services. People choose online payment and e-shopping; because of their own convenience. Evidence shows commonly referred to as big data the amount of data advances with the banks in particular needing to watch transactions in real time this enormous volume of data, it becomes much more important for the institutions to understand the increasing number of online transactions being reviewedto detect fraud and the vast variety of internetusage have made credit card security an acute threat in the current world we require some features based on which we can help the predictive model learn and decide about the frauds these features are extracted based on the application being used for they help to understand and bring about a pattern which is followed by the model for prediction As the quantity of characteristics increasesthe convolution,the model also grows along with it thus we need to fetch the specific features that contribute substantially towards the output. The prototype

generates probabilities indicating the likelihoodof a transaction being fraudulent assisting in the detectionof illigitemate activities to integrate all components effectively the predictive model leverages a specific algorithm occasionally false transactions go unnoticed leading to some fraud losses not being classified correctly after evaluating output efficiency and other relevant factors neural networks are recognized as highly effective classifiers for this prediction task because of their capacity to learn and identify transaction patterns we use the keras library to Improve our accuracy. deep learning network despite the small number of fraudulent transactions compared to genuine ones. the financial impact of fraud can be significant This results in a dataset. biased towards genuine transactions to address this we employ techniques to reduce data skewness and improve prediction efficiency as outlined in reference 1 suitable optimizers and loss functions are also used to enhance model efficiency this paper details our chosen classifier its comparison with others and methods to reduce dataset skewness our work is summarized as follows 1 a four-layer neural network was developed to we utilized a credit card fraud detection dataset from reference 1 3 the datasets skewness was adjusted using appropriate techniques 4 the neural network was trained and tested with this dataset the The paper is composed as follows section ii reviews related work and classifier comparisons section iii introduces our approach section iv presents experimental results and discussions and section v Wraps up by highlighting directions for additional research.

2.Literature review

Hariteja Bodepudi,"Credit Card Fraud Detection Using Unsupervised Machine Learning Algorithms"[1]. In this study, Hariteja Bodepudi explores Transaction fraudRecognition using unsupervised machine learning algorithms, specifically focusing on Local Outlier Factor (LOF), Isolation Forest, and One-Class SVM. The research highlights methods to identify Unauthorized transactions block payments in real-time. It provides a detailed explanation of how anomalies can be detected using an unsupervised approach to achieve high accuracy.

Dr anju pratap , anu maria babu [2] techniques made up of deep learning are created for recognizing electronic fraud using credit cards in theSituation of the fraudulent activity investigation situation we assume a neural artificial convolution network cnn model in the findings of this study, it can be accomplished by utilizing cnn classification to properly classify documents and enable efficient data retrieval across layers misclassifications and high mistaken positive rates of positives destroy the fraud coping system with such situations this research study gives use of the convolutional neural network that was used layers cooperation in order to establisha prototype that gives a higher level of success percentage for carrying credit card fraudlent

Pranali Shenvi et al [3]this paper by Pranali Shenvi use deep learning to detect credit card fraud the deep learning techniques used include under-sampling which lowers the percentage of genuine data and over-sampling which keeps the greatest number of fictitious individual class observations the writers of the publications execute deep learning for the purpose of recognizing Unauthorized credit card transactions were made using artificial neural networks the goal of the mentioned paper concentrate on the family issue

Abhilash Sharma M et al [4] "Credit Card Fraud Detection Using Deep Learning Based on Auto-Encoder". We try to compare the performance of Theprototype built upon three dif-

ferent datasets from three parts of the world viz. European cardholders' data, Australian credit card data and Asian(Taiwanese) credit card data.
The datasets what we have considered in this work explainsconsumer spending behavioris expected to vary significantly in multiple parts of the world. A comparison of the resultant metrics helps us to identify what data this model is fit for since any of themodel-fits-all approach in this area as different data require different kinds of models

Akshat Shah, Yogeshwari Makwana [5] "Credit Card Fraud Detection".One the methods most frequently used for credit card identifying illegal transactions isrule-based systems. These systems use predefined rules to identify transactions that are deemed suspicious.However,rule-based
systems have limitations, as they are only agood as the rules that have beenpredefined, and be incapable of detect new types of fraud.

Himani Ranpariya, Nidhi Musale and Rajan Singh Kushwaha [6] "credit card fraud detection system". Data was obtained from the Kaggle website. Different training and testing methods used when working on a dataset consist of logistic regression, random forest with a decision tree of classification on xgboost, isolation forest, confusion matrix.

Emmanuel Ileberi1,Yanxia Sun1 and Zenghui Wang[7] "A machine learning based credit card fraud detection using the GA algorithm for feature selection".This research,the emphasis is on applying the followingsupervised algorithms for identifying illegitimate banking transactions detection: Decision Tree (DT) , Random Forest (RF), Artificial Neural Network (ANN) ,Naive Bayes (NB) and Logistic Regression (LR) .ML systems are trained and tested using large datasets. In this work, a credit card fraud dataset generated from European credit cardholders is utilized.

Deepak Gwale, Prof. Sumit Sharma [8], "Credit Card Fraud Detection using Machine Learning", This paper therefore what it entails to use machine learning in the detection of Credit Card Fraud. To build the predictive models of the credit card fraud a study is made on the supervised methodologies for learning such as logistic regression and decision trees., random forest, and support vector machines. These models are developed from transactional data with incorporated labeled authentic and fraudulent records of transactions.

Bhattacharyya, Debashis, and Manish Jha. "A Study on Risk Management Strategies of Small and Medium Sized Enterprises in India." [9]" Credit Card Fraud Detection: As with the previous study by Xu et al., "A Machine Learning Approach". This working paper focuses on the following question:
Machine learning methods designed for credit card fraud detection. It embraces a couple of them and analyses how proficient they are in identifyingfraudulent transactions. Dal Pozzolo, A. Cae-len, O. Le Borgne, Y. A. ,Waterschoot, S. , Bon-tempi, G, (2015). Penetration of identifying the not original credit card transactions mistakes from a practitioner perspective. From a practical standpoint.

Aaron Rosenbaum [10]. Detecting Credit Card Fraud with Machine Learning by Aaron Rosenbaum [10]. Four primary model categories were considered: basic logistic regression model, logistic regression using nonlinear expressions and LASSO penalization, randomized forest, and optimization neural networks. When fitting the models, I used a trend-lending tournament approach at every one of the pro-

totype categories: I trained several models within each category. Using the parameters for each model, k-fold or simple cross validation tune the features was used. From every one of the categories, the model that had the highest accuracy on the validation set was the only model taken forward to become a finalist. At the end of the process, the four finalists were fine-tunedvia the jointly offered up training and validation datasets and tested on the remaining test set, which had been untouched throughout the training process

3.Proposed Method:

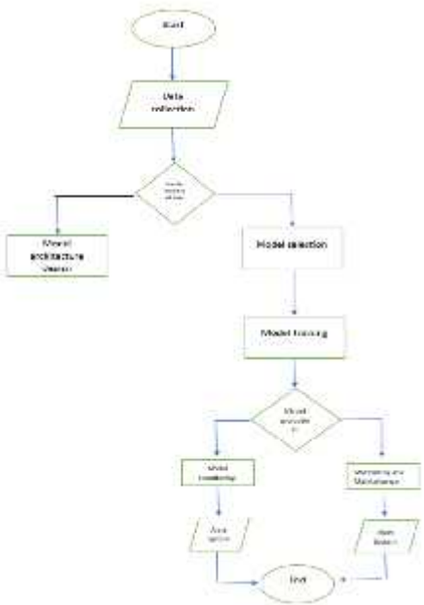

Figure 1: Project flow chart

The proposed methodology explains about comprehensive approach to detect fraudulent transactions in credit card data using machine learning techniques. The methods include Isolation Forest, Local Outlier Factor, and One-Class SVM. The proposed system aims to identify anomalies effectively and reduce the number of false positives, ensuring robust fraud detection.

The dataset contains transactions made by credit cards of European cardholders. It comprises 284,807 transactions, including 492 fraud cases. Each transaction has 30 features: 'Time', 'Amount', and 28 anonymized features resulting from a PCA transformation. The class label indicates whether a transaction is fraudulent (Class = 1) or normal (Class = 0).

The methodology involves several steps: data preprocessing, exploratory data analysis, sampling, model training, and evaluation.

3.1 Convolution Neural Network(CNN):

Identifying of the card transactions that are not legitimate detection techniques, Convolutional Neural Networks (CNNs) are not directly applicable. CNNs are primarily used in tasks involving image recognition and processing where spatial relationships in data are crucial.

However, if we incorporate CNNs into a fraud detection system, typically, it would involve converting non-image data (such as transaction data in your case) into a format that CNNs can process. This might involve techniques like converting sequential data (such as transaction sequences over time) into image-like formats (e.g., spectrograms) or using 1D convolutional layers to process temporal sequences directly.

Convolutional Neural Networks (CNNs) are deep learning models generally utilized for tasks that image recognition and computer vision. They are characterized by their ability to automatically learn hierarchical patterns in data through convolutional layers.

Convolutional Layers,Pooling Layers,Fully Connected Layers

Result:

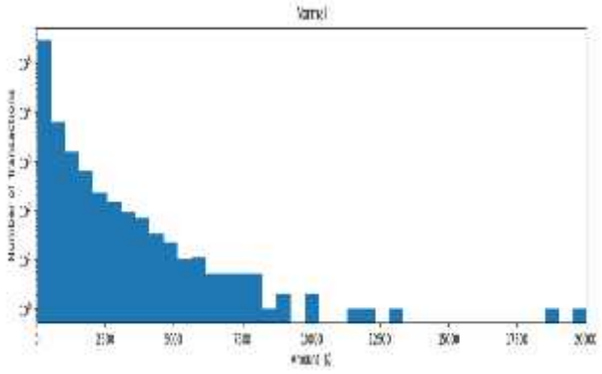

Figure 2:Distribution:

The histogram shows the distribution of transaction amounts. Lower transaction amounts are more frequent, while higher amounts are less common.

Anomalies: This pattern can be utilized as a baseline for identifying unusual occurrenc-es.. Unusual transactionsmay indicate fraud.

Fraud Detection: Algorithms compare new transactions against this typical distribution to identify potential fraudulent activities.

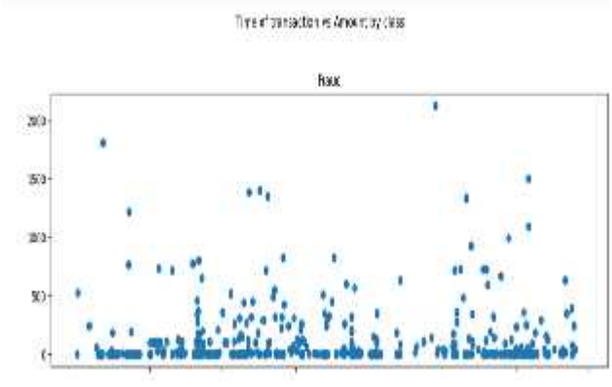

Figure 3: The scatter plot consists of individual data points.

Most of These markers are concentrated to-ward the bottom of the graph.

This suggests that most fraudulent transactions involve smaller amounts**.**

However, Additionally, there are multiple points. spread out above the cluster.

These outliers indicate that some fraudulent transactions involve larger amounts.

```
(28481, 31)
(284807, 31)
0.001723410241980866
Fraud Cases : 49
Valid Cases : 28432
(28481, 30)
(28481,)
Isolation Forest: 73
Accuracy Score:
0.9974368877497279
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     28432
           1       0.26      0.27      0.26        49

    accuracy                           1.00     28481
   macro avg       0.63      0.63      0.63     28481
weighted avg       1.00      1.00      1.00     28481
```

Figure 4: The above results are being taken from the project ,these results explains the F1 score , precision ,recall that are obtained by running the project by the Isolation forest method.

```
Local Outlier Factor: 97
Accuracy Score:
0.9965942207085425
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     28432
           1       0.02      0.02      0.02        49

    accuracy                           1.00     28481
   macro avg       0.51      0.51      0.51     28481
weighted avg       1.00      1.00      1.00     28481

Support Vector Machine: 8516
Accuracy Score:
0.7009936448860644
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.70      0.82     28432
           1       0.00      0.37      0.00        49

    accuracy                           0.70     28481
   macro avg       0.50      0.53      0.41     28481
weighted avg       1.00      0.70      0.82     28481
```

Figure 5: These above results explains about the F1 score, precision and recall that are obtained when we use Local outlier factor and support vector machine.

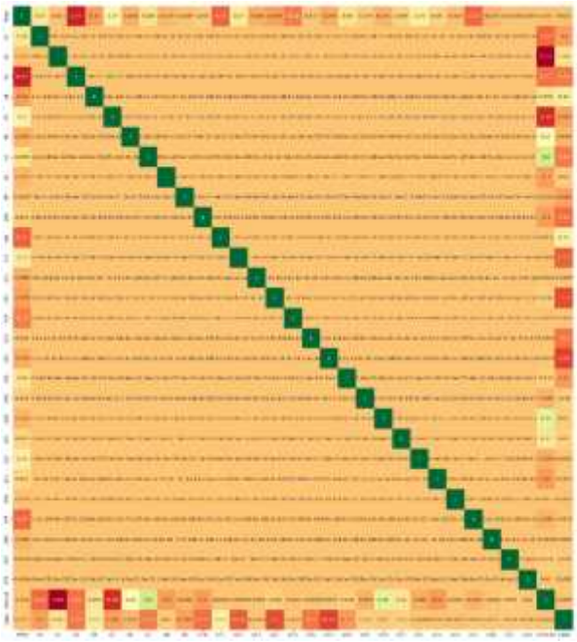

Figure 6: The above figure is the correlation heatmap

Conclusion:

In summary we have seen the importance of the detection of the fraudlent transactions of credit card that are taking place.All these techniques are tested based on accuracy andprecision. We have selected supervised learning techniqueRandom Forest to classify the alert as fraudulent orauthorized. This classifier will be trained using feedbackand delayed supervised samples. Next it will aggregateeach probability to detect alerts.

Future enhancement:

The implementation of this work is mainly a plus point for the society , but the recent works are trying to increase the productivity and accuracy of the results. This work is being used by the banking systems, it should be used by all the people of society so that an awareness can be created among them to identify the fraudlents that are happening in the real world. It should also be used by all the financial institutions, schools, business place and so on.Develop a real-time fraud detection system capable of analyzing transactions as they occur. This requires efficient streaming data processing and immediate anomaly detection to prevent fraudulent activities in real-time.

References:

1. Bodepudi, H. (2021). Credit Card Fraud Detection Using Unsupervised Machine Learning Algorithms. International Journal of Computer Trends and Technology, 69(8), 1-3. DOI: 10.14445/22312803/IJCTT-V69I8P101.

2. Anu Maria Babu, & Dr. Anju Pratap. (2020). Credit Card Fraud Detection Using Deep Learning. IEEE Recent Advances in Intelligent Computational Systems (RAICS), 1-4.

3. Ranali Shenvi, Neel Samant, Shubham Kumar and Dr. Vaishali Kulkarni (2019). Credit Card Fraud Detection using Deep Learning. 2019 5th International Conference for Convergence in Technology (I2CT)

4. Abhilash Sharma M, Ganesh Raj B R , Ramamurthy Band Hari Bhaskar R(2022).Credit Card Fraud Detection Using Deep Learning Based on Auto-Encoder.ITM Web of Conferences 50, 01001 (2022) ICAECT.

5. Akshath Shah,Yogeshvari Makwana(2023). Credit Card Fraud Detection.Url:https://www.researchgate.net/publication/369857378_Credit_Card_Fraud_Detection

6. Emmanuel Ileberi1, Yanxia Sun and ZenghuiWang(2022).A machine learning based credit card fraud detection using the GA algorithm for feature

selection.Journal of Big Data.Url: https://doi.org/10.1186/s40537-022-00573-8
7. Himani Ranpariya, Nidhi Musale, Anushka More, Sarthak Salunke and Prof. Sujit Tilak(2022). Credit card fraud detection system.Procedia Computer Science Url:https://www.researchgate.net/publication/359005632
8. Deepak Gwale,Prof Sumit Sharma(2014).Credit card Fraud Detection using Machine Learning. Journal of emerging technologies and innovative research(JETIR).www.jetir.org (ISSN-2349-5162)
9. Aaron Rosenbaum(2015). Detecting Credit Card Fraud with Machine Learning. Stanford University, Stanford, CA, 94305, USA.
10. Ibtissam Benchaji,Samira Douzi, Bouabid El Ouahidiand Jaafar Jaafari(2021) Enhanced credit card fraud detection based on attention mechanism and LSTM deep model.Benchaji et al. Journal of Big Data (2021) 8:151 Url:https://doi.org/10.1186/s40537-021-00541-8
11. Arunkumar K L, Ajit Danti, Manjunatha H T (2022). Classification of Vehicle Make Based on Geometric Features and Appearance-Based Attributes Under Complex Background. Springer 1035 (CCIS), pp 41-48.
12. Arunkumar K L, Ajit Danti (2018). A Novel Approach for Vehicle Recognition Based on the Tail Lights Geometrical Features in the Night Vision. International Journal of Computer Engineering and Applications, Volume XII, Issue I, Jan. 18. www.ijcea.com ISSN 2321-3469.
13. Manjunatha H T, Arunkumar K L, Ajit Danti (2022). A Novel Approach for Detection and Recognition of Traffic Signs for Automatic Driver Assistance System Under Cluttered Background. Springer 1035 (CCIS), pp 407-419.
14. Arunkumar K L, Ajit Danti, Manjunatha H T (2019). Estimation of Vehicle Distance Based on Feature Points Using Monocular Vision. IEEE 8816996 (2019), pp 1-5.
15. CM Nrupatunga, KL Arunkumar (2020). Peruse and Recognition of Old Kannada Stone Inscription Characters. Springer, Singapore.