Detecting Fraudulent Credit Card Transactions while countering Precision-Recall Trade-off using Machine Learning

Abhinav

April, 2023

1 Introduction

This is introduction

2 Related Work

Carcillo et al. [5] presented a novel approach based on machine learning algorithms which used a combination of both supervised and unsupervised learning approaches in order to detect fraud patterns in credit card transactions and the results showed positive outcomes. Lakshmi et al. [11], in another paper, introduced another novel approach for predicting fraud transactions in credit card system using different machine learning algorithms and achieved a remarkable result with highest accuracy rate of 95.5 percent. In another paper, Taha et al. [18] proposed an intelligent approach for detecting credit card transactions related frauds using a Bayesian-based hyperparameter optimization algorithm and intelligently integrated it to tune the parameters of a light gradient boosting machine (LightGBM). To demonstrate the effectiveness of their proposed novel approach, they took two real world public credit card transaction data sets consisting of both fraudulent and benign samples and achieved remarkable results in terms of 98.40 percent accuracy, 92.88 percent area under receiver operating characteristic curve (AUC), 97.34% precision and 56.95% F1-score.

Sarker et al. [16] proposed a machine learning based network security model, named as IntruDTree, to detect suspicious and/or potentially harmful behavior over a network using various cyber-security based standard datasets. The effectiveness of the model was determined by conducting a wide range of experiments on various relevant datasets and also comparing the results of the model with several other traditional and popular machine learning based models. Injadat et al. [9], in another paper, proposed an optimized machine learning based framework model for network intrusion detection by choosing the most suitable subset of features and optimizing the parameters to enhance their performance. The entire framework model was divided into three stages with the first being doing Z-score normalization and applying SMOTE technique. In the second stage, a feature selection process was done to reduce the number of features needed for the classification model. In the third stage, framework optimization of the considered hyper-parameters of the different classification models was done and the results were combined to build the multi-stage optimized ML classification model that reduced the computational complexity while maintaining considerable detection performance. In another paper, Larriva-Novo et al. [12] proposed another IDS framework that used machine learning and mainly emphasized on proper data preprocessing for increase in precision. As a first point of analysis, study was carried out using the proposed datasets, without pre-processing, to obtain a base precision measure, which later allowed to determine the increase in precision when various types of preprocessing techniques were applied. The results showed that the NSL-KDD dataset offered a better accuracy, with 95.5 percent, compared to 87.68 percent of the UGR16 and 55.80 percent of the UNSW-NB15 datasets.

Jain et al. [10] developed a machine learning based phishing detector model, named as Phish-Safe, using a large dataset from phishtank.com. The model showed more than 90 percent accuracy in detecting phishing websites using SVM and Naive Bayes classifiers.

Fang et al. [8] presented a novel approach to detect XSS attacks based on deep learning, named as DeepXSS. Word2vec was used to extract the feature of XSS payloads which captured word order information and mapped each payload to a feature vector and then, the detection model was trained and tested using Long Short Term Memory (LSTM) recurrent neural networks. Experimental results showed that the proposed XSS detection model based on deep learning achieved a precision rate of 99.5 percent and a recall rate of 97.9 percent in real dataset.

The work for Android malware detection, proposed by Darus et al. [6] was able to achieve 84.14 percent detection accuracy by using Random Forest machine learning algorithm on image features generated from APK samples. The images were generated from 483 APK samples which were made up of 183 malware samples and 300 benign samples, and their features were extracted using GIST descriptor. In another paper, Yeboah-Ofori et al. [19] used ML and pipelining techniques to propose an intelligent cyber threat detection technique that can predict which nodes on a system are vulnerable to attacks to be able to predict future attacks using the Microsoft Malware Prediction dataset. Five-fold cross-validation technique and GridsearchCV were used to test the parameter estimation and estimate the best optimization to cross-validated the GridsearchCV in the parameter tuning. `AUC_ROC` curve was then used to predict the dimensions of the graph for the true positive and false negative rates. The result showed that ML algorithms in Decision Trees methods can be used in cyber supply chain predictive analysis to detect and predict future cyber attack trends.

3 Methodology

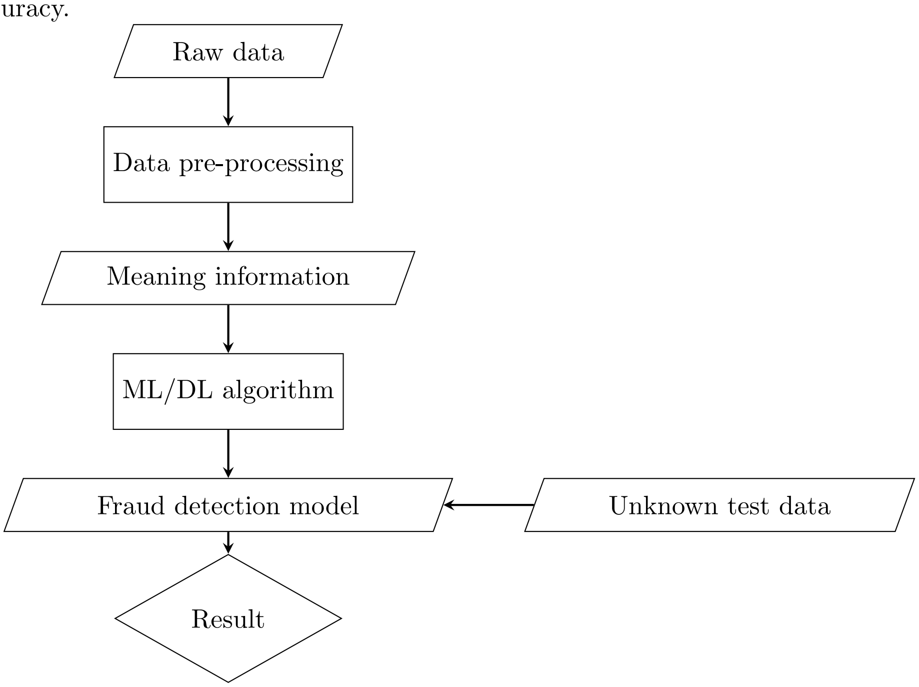
3.1 Background

3.1.1 Definition and purpose of credit card fraud detection

Credit card fraud detection is the process of classifying transactions made using credit cards as being fraudulent or non-fraudulent. The main purpose of such detection technique is identifying the purchase attempts initiated using credit cards that are fraudulent and informing the administrator or authority about the same for rejecting the transaction rather than processing them.

3.1.2 Working mechanism and diagrammatic explanation

Such detection models, in their most general form, are built using some machine learning and deep learning algorithms and their efficiency varies greatly as per the type or combination of algorithms chosen and data preprocessing techniques used. Initially, raw data concerning credit card transactions are collected from various sources and then they are processed using different techniques to extract meaningful information from them. This is called data pre-processing. Then, an appropriate algorithm is selected or a hybrid technique is implemented as per requirements and a model is trained for the detection purpose using the pre-processed data. Finally, efficiency of the model is calculated by testing it against unknown data containing both fraudulent and benign samples in varying proportions and obtaining result in the form of precision, recall, F1-score and accuracy.



3.1.3 Tools and technologies used

While carrying out the implementation part of the novel architecture for credit card fraud detection purpose (discussed in detail in section 3.3), Visual Studio Code [1] with an extension for Jupyter Notebook [13], [14] was used as a code editor tool and different machine learning techniques and algorithms were implemented to convert raw data into meaningful information and train the fraud detection model in order to obtain insightful results.

3.1.4 Role of credit card fraud detection in cyber security world

With rapid advancement in new technologies and growth in e-commerce sector, use of credit cards for online purchases and other transactions has grown exponentially. Consequently, this has resulted in increasing the graph of frauds related to it. Attackers are frequently coming up with new techniques to exploit this opportunity for their unethical purposes and consequently, credit card related fraudulent activities are alarmingly increasing day-by-day. This has created a worrisome situation in the world of cyber security. Hence, this paper introduces a novel architecture model to detect credit card frauds effectively using machine learning to assist the cyber world.

3.2 Problem statement

As per a report published by MyNCR [3], a 2018 study by the Federal Reserve showed that the amount of loss from card-not-present (CNP) fraud jumped from $3.4 billion to $4.57 billion in 2016. Another study by Javelin Strategy & Research [2] revealed that CNP fraud is 81% more likely to occur than card present fraud. As per U.S.News & World report [7], it is estimated that Global financial losses related to card payments may reach $34.66 billion in 2022. One major way of tackling this worrisome situation is by building credit card fraud detection model using machine learning. The use of machine learning primarily focuses on building and increasing the accuracy of models to detect frauds. Many researchers [5], [11] have worked in this regard. But, main issue with this focused approach is that the data present in real world is highly imbalanced. In such conditions, if the data has 90% legal and 10% fraud samples and the model fails to detect any fraud transaction, still the accuracy will be 90%. Hence, only accuracy measurement becomes a poor factor for determining the performance of any model when data is imbalanced and other factors like precision, recall and F1-score becomes important. Though many researchers like [18] have touched upon this issue, the precision-recall rate is still not very satisfactory and often, a trade-off between the two occurs. Hence, there is still scope for improvement. Therefore, this paper proposes a novel architecture for credit card fraud transaction detection model using machine learning which gives an accuracy of x% along with precision and recall of p% and r% respectively. The model gives an F1-score of F. Hence, this model not only increases precision and recall rates but also minimizes the precision-recall trade-off considerably.

3.3 Proposed solution

The overall framework of the proposed novel approach to efficiently detect credit card fraud transactions is illustrated in figure 1 in section 3.3.1.

The proposed intelligent approach for credit card fraud transactions detection which consists of five major steps along with the hardware and software requirements, system setup on which the implementation is performed and the algorithm used, are explained in the following subsections.

3.3.1 Architecture of proposed model

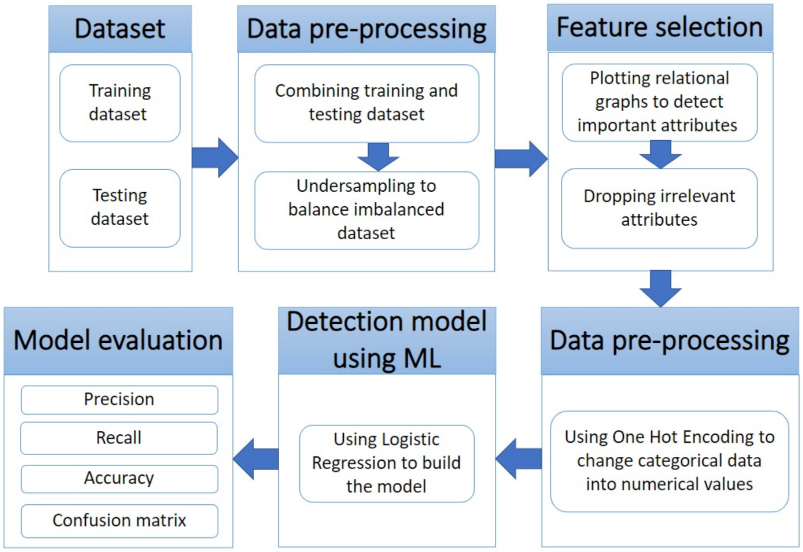

Dataset
Training dataset
Testing dataset
Data pre-processing
Combining training and testing dataset
Undersampling to balance imbalanced dataset
Feature selection
Plotting relational graphs to detect important attributes
Dropping irrelevant attributes
Data pre-processing
Using One Hot Encoding to change categorical data into numerical values
Detection model using ML
Using Logistic Regression to build the model
Model evaluation
Precision
Recall
Accuracy
Confusion matrix

Figure 1: Overall framework for proposed architecture

Dataset

To develop the proposed model and evaluate its performance, I consider a simulated credit card transaction dataset [17] containing both legal and fraud transactions from the duration 1st Jan 2019 to 31st Dec 2020. It covers credit cards of 1,000 customers doing transactions with a pool of 800 merchants. Out of a total of 18,52,394 transaction samples, the dataset contains 18,42,743 legitimate and 9,651 fraudulent transactions with the fraudulent ones being only 0.52% of the total samples. The dataset contains 23 feature columns namely unnamed:0, trans_date_trans_time, cc_num, merchant, category, amt, first, last, gender, street, city, state, zip, lat, long, city_pop, job, dob, trans_num, unix_time, merch_lat, merch_long and is_fraud. The is_fraud feature is the classification variable, which is 1 in case of fraud transactions and 0 otherwise. The dataset is summarized in table 1, which depicts the total number of samples, number of legal samples, number of fraudulent samples, number of features in the dataset and the reference for accessing the data set.

Total number of samples	Number of legal samples	Number of fraudulent samples	Number of features	Ref.
18,52,394	18,42,743	9,651	23	[17]

Table 1: Dataset summary

Data pre-processing

Data pre-processing is a kind of data preparation technique which is used to transform the raw data obtained from dataset into meaningful information for the machine learning model to properly train upon.

In order to test the proposed model with different sub-parts of the dataset each time to obtain a convincing result, the training and testing datasets were first combined to form a larger group of dataset. Now, because the total number of fraudulent transactions is much less than the number of legitimate transactions, the data distribution is unbalanced, i.e., skewed towards the legitimate observations and it is a well known fact that the performance of various machine learning techniques tends to decrease when the analyzed dataset is unbalanced. Hence, to obtain results with better precision, undersampling technique is used to randomly select a set of 48,255 legitimate samples (5 times that of total fraud samples). This sample is then combined with the total fraudulent samples and then the entire new combination of fraud and benign samples were divided in 8:2 ratio for the training and testing purpose of the model. In between, feature selection is also performed and unnecessary noise are removed. One Hot Encoding is used to convert important categorical data into corresponding numerical values before training and testing the model.

Feature selection

Selecting significant and important features is critical for the effective detection of credit card fraud transactions when the number of features is large. In the proposed model, a manual feature selection procedure is applied. The significance of various features were determined by plotting their relational curve or graph against the classification attribute and then, their importance were drawn accordingly. Fig. 2,3,4,5,6 shows some of the relational plots which were used to draw important conclusions about the significance and importance of the corresponding attributes.

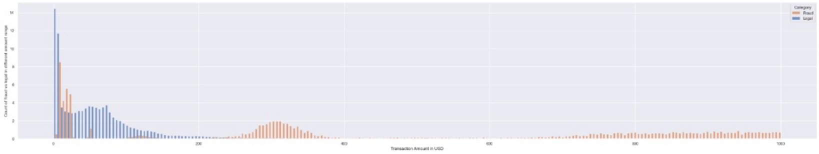


Figure 2: This figure shows fraud distribution in various transaction amount ranges

Fig. 2 shows a clear relation. While most legal transactions occur near price range of $200 and less, illegal transactions rises at around $250 to $300 and then again at around $650 to $700 and more. Also, illegal transactions, also occur below $200 in some small amount range. Hence, amount is a significant feature or indicator.

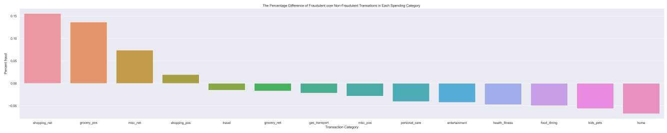


Figure 3: This figure shows fraud distribution in various expense categories

Fig 3 shows a clear pattern that most fraud happens in case of shopping_net category and least in home category. Hence, it is also a significant feature.

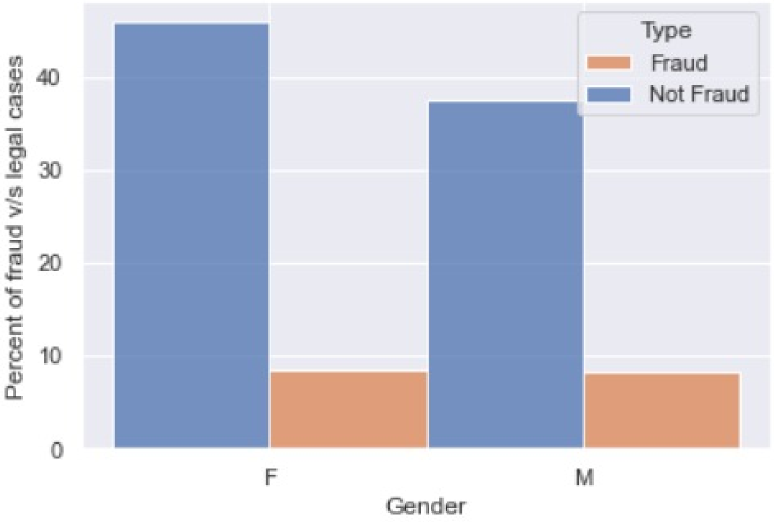


Figure 4: This figure shows which gender is more prone to being trapped in a credit card transaction fraud

The above histograph in fig. 4 suggests that gender is neither a significant nor an important feature for fraudulent transactions detection as both the genders are equally likely to come across both legal as well as illegal transactions.

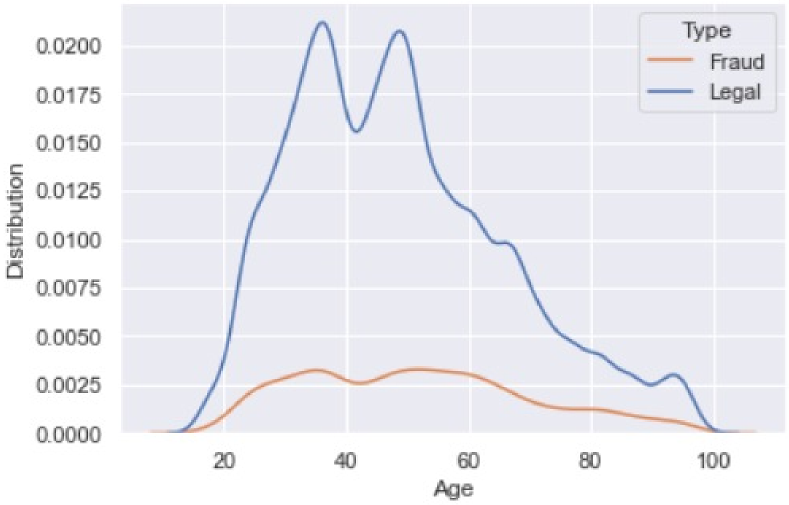


Figure 5: This figure shows which age group is more prone to being trapped in a credit card transaction fraud

The above plot in fig. 5 clearly indicates that older people are more prone to frauds as compared to the younger ones. Although it cannot be stated as a significant feature as fraud can happen with anyone belonging to any age group, but still it gives an important information that attackers tend to target older people more often than the younger ones.

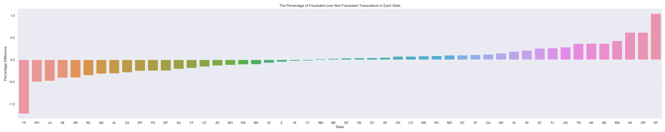


Figure 6: This figure depicts the percent difference in fraud transactions in different states of a county

Although the above fig. 6 cannot be used to draw any significant conclusion as fraud can happen with anyone in any state but still it gives important insights about the credit card fraud transactions trend in different states of a country and helps to distinguish between the states on the basis of high and low credit card fraud transactions cases.

Detection model using ML

This section explains the machine learning algorthm used to propose the novel approach for credit card fraud transactons detection. In the proposed approach, logistic regression is used to develop the novel framework. It is a supervised algorithm in machine learning and used when the data is linearly separable and the classification should be binary or dichotomous in nature.

A classification problem is said to be binary in nature when the nature of output is discrete in two values or classes.

Model evaluation

To evaluate the performance of the proposed approach for credit card fraud transactions detection model, the modified dataset (discussed in detail under data pre-processing part above) is divided in 8:2 ratio for training and testing purpose respectively.

To assess the performance of the proposed approach, several measures are considered, including the confusion matrix, precision, recall, accuracy, AUC and F1-score. The metrics are defined based on the confusion matrix. It is a performance measurement technique for machine learning classification problem where output can be two or more values or classes. It is a table with 4 different combinations of values namely true positive (TP), false positive (FP), false negative (FN) and true negative(TN).

TP refers to the number of fraudulent credit card transactions correctly classified.
FP denotes the number of legitimate credit card transactions classified as fraud.
FN indicates the number of fraudulent credit card transactions classified as normal.
TN defines the number of normal credit card transactions correctly classified.

The measures that were used are defined as follows.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$F1 - score = \frac{2(Precision \times Recall)}{Precision + Recall}$$

Although accuracy sounnds to be the most promising factor but precision and recall plays a bigger role when the data is highly imbalanced. A high recall score reflects a low false negative (FN) rate, while high precision reflects a low false positive (FP) rate. High scores for precision and recall indicates that the classifier restores results with higher accuracy and recovers most of the positive results successfully [15]. Therefore, the Precision-Recall curve reveals a complete picture of the accuracy of the classifier and is robust even in an imbalanced data set [18]. The AUC value is considered as a general performance measure in addition to the above measures. AUC is a graphical plot of the false positive rate (FPR) and the true positive rate (TPR) at different possible levels [4]. A model with better overall performance has both AUC value and F1-score close to one.

3.3.2 Implementation

System setup

The experiment was performed in a system using an Intel Core i7 processor configuration with a RAM of 8GB in Microsoft Windows 10 Home environment.

Software used

The proposed approach and other machine learning techniques were implemented and tested using the Python programming language and Visual Studio code [1] with some extensions for Jupyter Notebook [13], [14] was used as the code editor tool.

Algorithms and techniques used

The proposed model is trained using a credit card fraud transactions detection dataset obtained from Kaggle.com [17]. Undersampling and One Hot Encoding methods are used for data pre-processing, feature selection is done by plotting relation between various attributes and the classification attribute and the model is finally trained using a maachine learning based binary classification algorithm known as Logistic Regression.

4 Result

To investigate the effectiveness of the proposed approach for credit card fraud transactions detection model, it is trained using Logistic Regression, a Machine Learning classification algorithm, upon a simulated real-world credit card transactions data set [17]. Figure 7 and 8 depicts the performance evaluation of the proposed approach using confusion matrix and a classification report respectively.

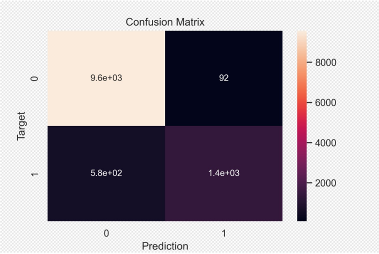
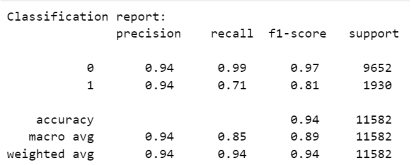


Figure 7: Confusion Matrix



```
Classification report:
              precision    recall  f1-score   support

           0       0.94      0.99      0.97      9652
           1       0.94      0.71      0.81      1930

    accuracy                           0.94     11582
   macro avg       0.94      0.85      0.89     11582
weighted avg       0.94      0.94      0.94     11582
```

Figure 8: Performance evaluation of the proposed approach using classification report

The proposed approach achieves an AUC value of 89.51% which indicates the ability of the proposed approach to distinguish between legitimate and fraudulent credit card transactions efficiently. In addition, the proposed approach achieved an accuracy of 94%, which is the ratio of correctly predicted credit card transactions to the total number of transactions. The proposed approach also obtained recall scores of 99% and 71% for legal and fraudulent transactions respectively, which indicates the ability of the proposed approach to correctly detect more than 71% of the suspicious credit card transactions with a low false negative rate. In addition, the proposed approach achieved precision scores of 94% each for legal and fraudulent transactions, which is the ratio of correctly classified fraudulent transactions to total classified fraudulent transactions. The proposed approach also achieved F1-score of 97% for legitimate and 81% for fraudulent transactions respectively indicating the balance between two important measures, namely precision and recall. Therefore, this score takes both false negatives and false positives into account. It is an important measure, particularly if the number of legitimate and fraudulent credit card transactions in the data set are highly imbalanced.

For better visualization of the performance evaluation results, the ROC curve and the P-R curve is shown in fig. 9 and 10 respectively.

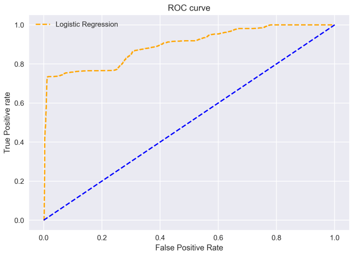


Figure 9: ROC curve obtained from the proposed model implementation

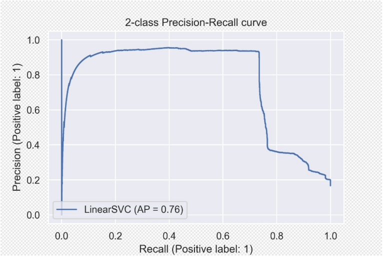


Figure 10: Precision-Recall curve of the proposed model

Table 2 depicts a comparative analysis of the results obtained in this proposed approach with other proposed approaches by some other authors.

Proposed work	Precision	Recall	F1-score	Accuracy	AUC value
Approach proposed in [5]	-	-	-	-	-
Approach proposed in [11]	-	-	-	95.5%	-
Approach proposed in [18]	97.34%	-	56.95%	98.4%	92.88%
Proposed approach	94%	71%	81%	94%	89.51%

Table 2: Comparative analysis of results of proposed approach with other renowned approaches

The overall performance of the proposed approach proves to be highly consistent based on the train test split procedure using the simulated real-world data set.

5 Conclusion and Future Work

References

[1] Visual studio code. https://code.visualstudio.com/.

[2] Identity fraud hits all time high with 16.7 million u.s. victims in 2017, according to new javelin strategy research study. https://javelinstrategy.com/press-release/identity-fraud-hits-all-time-high-167-million-us-victims-2017-according-new-javelin, February 6, 2018.

[3] Credit card transaction fraud continues to climb to new heights. https://www.ncr.com/blogs/payments/credit-card-fraud-detection, March 10, 2021.

[4] A. Bhandari. Auc-roc curve in machine learning clearly explained. https://www.analyticsvidhya.com/blog/2020/06/auc-roc-curve-machine-learning/, June 16, 2020.

[5] F. Carcillo, Y.-A. Le Borgne, O. Caelen, Y. Kessaci, F. Oblé, and G. Bontempi. Combining unsupervised and supervised learning in credit card fraud detection. *Information sciences*, 557:317–331, 2021.

[6] F. M. Darus, N. A. A. Salleh, and A. F. M. Ariffin. Android malware detection using machine learning on image patterns. In *2018 Cyber Resilience Conference (CRC)*, pages 1–2. IEEE, 2018.

[7] J. Egan. What is card-not-present fraud? https://money.usnews.com/credit-cards/articles/what-is-card-not-present-fraud, February 1, 2019.

[8] Y. Fang, Y. Li, L. Liu, and C. Huang. Deepxss: Cross site scripting detection based on deep learning. In *Proceedings of the 2018 international conference on computing and artificial intelligence*, pages 47–51, 2018.

[9] M. Injadat, A. Moubayed, A. B. Nassif, and A. Shami. Multi-stage optimized machine learning framework for network intrusion detection. *IEEE Transactions on Network and Service Management*, 18(2):1803–1816, 2020.

[10] A. K. Jain and B. Gupta. Phish-safe: Url features-based phishing detection system using machine learning. In *Cyber Security*, pages 467–474. Springer, 2018.

[11] S. Lakshmi and S. D. Kavilla. Machine learning for credit card fraud detection system. *International Journal of Applied Engineering Research*, 13(24):16819–16824, 2018.

[12] X. Larriva-Novo, V. A. Villagrá, M. Vega-Barbas, D. Rivera, and M. Sanz Rodrigo. An iot-focused intrusion detection system approach based on preprocessing characterization for cybersecurity datasets. *Sensors*, 21(2):656, 2021.

[13] Microsoft. Jupyter. https://marketplace.visualstudio.com/items?itemName=ms-toolsai.jupyter.

[14] Microsoft. Jupyter notebook renderers. https://marketplace.visualstudio.com/items?itemName=ms-toolsai.jupyter-renderers.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[16] I. H. Sarker, Y. B. Abushark, F. Alsolami, and A. I. Khan. Intrudtree: a machine learning based cyber security intrusion detection model. *Symmetry*, 12(5):754, 2020.

[17] K. Shenoy. Credit card transactions fraud detection dataset. https://www.kaggle.com/datasets/kartik2112/fraud-detection, 2020.

[18] A. A. Taha and S. J. Malebary. An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. *IEEE Access*, 8:25579–25587, 2020.

[19] A. Yeboah-Ofori and C. Boachie. Malware attack predictive analytics in a cyber supply chain context using machine learning. In *2019 International Conference on Cyber Security and Internet of Things (ICSIoT)*, pages 66–73. IEEE, 2019.