COST SENSITIVE PAYEMENT FRAUD DETECTION BASED ON DYNAMIC RANDOM FOREST AND KNN

K. Ramya Sri[1], N. Chandrika[2], M. Sridevi[2], P. Thanmay Sree[2], P. Divya[2]

[1]Assistant Professor, [2]UG Scholar, [1,2]Department of Computer Science and Engineering

[1,2]Malla Reddy Engineering College for Women (A), Maisammaguda, Medchal, Telangana.

Chandrika.neelam92@gmail.com, sridevi.mata2003@gmail.com, thanmaysreepanuganti048@gmail.com, divyapulikota20@gmail.com

ABSTRACT

The act of fraudulent credit card transactions has been increased over the past recent years, as the era of digitization hits our day-to-day life, with people getting more involved in online banking and online transaction system. Machine learning algorithms have played a significant role in detection of credit card frauds. However, the unbalanced nature of the real-life datasets causes the traditional classification algorithms to perform low in detection of credit card fraud. In this work, a cost-sensitive weighted random forest algorithm has been proposed for effective credit card fraud detection. A cost-function has been defined in the training phase of each tree, in bagging which emphasizes assigning more weight to the minority instances during training. The trees are ranked according to their predictive ability of the minority class instances. The proposed work has been compared with two existing random-forest based techniques for two binary credit card datasets. The efficiency of the model has been evaluated in terms G-mean, F-measure and AUC values. The experimental results have established the proficiency of the proposed model, than the existing ones.

Keywords: Fraudulent credit card, Machine learning, Random Forest.

1. INTRODUCTION

Payment card fraud leads to heavy annual financial losses around the world, thus giving rise to the need for improvements to the fraud detection systems used by banks and financial institutions. In the academe, as well, payment card fraud detection has become an important research topic in recent years. With these considerations in mind, we developed a method that involves two stages of detecting fraudulent payment card transactions. The extraction of suitable transactional features is one of the key issues in constructing an effective fraud detection model. In this method, additional transaction features are derived from primary transactional data. A better understanding of cardholders spending behaviors is created by these features. After which the first stage of detection is initiated. A cardholders spending behaviors vary over time so that new behavior of a cardholder is closer to his/her recent behaviors. Accordingly, a new similarity measure is established on the basis of transaction time in this stage. This measure assigns greater weight to recent transactions. In the second stage, the dynamic random forest algorithm is employed for the first time in initial detection, and the minimum risk model is applied in cost-sensitive detection. We tested the proposed method on a real transactional dataset obtained from a private bank. The results showed that the recent behavior of cardholders exerts a considerable effect on decision-making regarding the evaluation of transactions as fraudulent or legitimate. The findings also indicated that using both primary and derived transactional features increases the F-measure. Finally, an average 23% increase in prevention of damage (PoD) is achieved with the proposed cost-sensitive approach.The term credit card fraud

signifies any act of theft and fraud, occurred in case of any payment card (debit or credit), due to physical loss of the card by the owner, or stealing of the card by fraudsters, or by means techniques like phishing, skimmer, identity theft etc. Financial fraud of this type can greatly affect corporate, organizational and government sectors of a country. The rate of fraudulent transactions has increased in today's era of internet technology, where credit card transaction has become the most convenient way of transaction, whether online of offline. As discussed in there are two types of credit card fraud can happen, internal fraud and external fraud. The first type depicts a situation where there is leakage of information between the cardholder and bank, by means of false identity; whereas in the later case, fraud happens due to stolen or lost credit card get into some fraudsters' hands. To solve this problem, credit card fraud detection techniques have been developed by researchers over the past years. Basically it involves of classifying the credit card transactions either as "legitimate" or "fraudulent". A number of Machine Learning (ML) techniques have been employed for this task, such as Decision tree Support Vector Machine (SVM) , Naïve Bayes (NB), Artificial Neural network (ANN), and optimization techniques such as genetic algorithm, migrating birds optimization algorithm. However, the task of credit card fraud has to face some adversities such as, (a) unavailability of real-life credit transaction dataset, (b) imbalanced ratio of "legitimate" vs. "fraudulent" transactions, (c) enormous size of the dataset, and (d) dynamic behavior of the fraudsters. As a result, effective credit card fraud detection demands effective pre-processing of the dataset, prior to applying of any ML techniques.. Each tree is then trained by using a weak-classifier (C4.5/J48), and then the test set is validated with trained trees.

2. LITERATURE SURVEY

A significant number of research works have been done for credit card fraud detection. The techniques developed can be categorized into two sections, as discussed below:

2.1 Machine Learning based techniques

In a survey of different data mining and machine learning techniques for credit card fraud detection has been presented. The paper has summarized a list of challenges one might encounter during credit card fraud detection. In a meta-learning based credit card fraud detection technique is presented. In [13], a comparison study of logistic regression and NB is performed. Back-propagation algorithm is integrated with NB and C4.5 to detect fraud in an imbalanced data-space, generated by minority oversampling with replacement in [14]. An adaptive learning technique based on concept-learning is proposed in [15], which is also robust towards noise. A comparison study of C4.5, ANN, and logistic regression is presented in [16], to showcase their applicability in credit card fraud detection. In [17], credit card fraud detection has been investigated by using C4.5 and SVM, and results has revealed C4.5 to be more effective than SVM in fraud detection. In [18], ANN and logistic regression based classification models are developed and implemented for credit card fraud detection. Another comparative study of supervised classification models for fraud detection is presented in [19], with C4.5, ANN, and NB models. The study has revealed ANN to have more time complexity, while NB is found to be inefficient when applied to new instances. In [20], an adaptive sampling technique is proposed based on ADASYN, which increases the minority class density by means of altering the majority class labels, based on a density distribution.

2.1 Random Forest based techniques

RF- based techniques has marked its role as the solution of credit card fraud detection, due to its bagging scheme of data training. A few works have been reported with RF being used for fraud detection. In [21], an integration of RF and sampling methods is proposed for detecting the potential buyers in PAKDD 2007 dataset. In [10], a RF technique with J48 weak classifier has been proposed

for credit card fraud detection, and results have been compared with SVM, NB, and KNN classification models. Devi et al. has proposed RF-based classification module for credit card fraud detection in [22], where the data had been cleaned and pre-processed (feature extraction), before feeding it to the classification module. In [23], Xuan et al. two variations of RF-classifier technique have been proposed, based on distance between records and tree-centres, and minimum Gini-criterion calculation.

2.2 Scope of cost-sensitive weighted random forest technique:

As discussed in [11], credit card fraud detection involves a largely differenced misclassification costs between the legitimate and the fraudulent cases. The misclassification cost is a vital factor while performing any classification task. In presence of imbalanced data, adequate definition of the misclassification cost is desirable, while offering cost-sensitive learning based solutions. Traditionally, the credit card fraud detection needs lot of data pre-processing (feature selection/ extraction, sampling, outlier detection), followed by machine learning algorithms to detect the legitimate/ fraudulent cases. Designing a cost-sensitive learning framework in order to treat the imbalanced cases by incorporating the highly differenced misclassification costs of the credit card datasets is the motivation of this work. The challenge of highly differenced misclassification cost of credit card fraud detection has been treated by incorporating it to the RF-Bagging ensemble learning model so that achieving of low positive (minority) class error can have more emphasis. The tree with lowest error can have more weights while defining the learning function. Hence, a cost-sensitive weighted random forest approach is a promising one to treat two different issues of credit card fraud detection, namely imbalanced nature of the data -space and highly differenced misclassifications costs. The contribution of the proposed work has been focused within this aspect.

2.3 A Cost-Sensitive Decision Tree Approach for Fraud Detection

As information technology is developing the fraud is also increasing as a result financial loss due to fraud is also very large. A cost sensitive decision tree approach has been used for fraud detection. A cost called misclassification cost is used which is taken as varying as well as priorities of the fraud also differs according to individual records. So common performance metrics such as accuracy, True Positive Rate (TPR) or even area Under Curve cannot be used to evaluate the performance of the models because they accept each fraud as having the same priority regardless of the amount of that fraudulent transaction or the available usable limit of the card used in the transaction at that time. For avoiding this a new performance metric which prioritizes each fraudulent transaction in a meaningful way and it also checks the performance of the model in minimizing the total financial loss. The measure used is Saved Loss Rate (SLR) which is the saved percentage of the potential financial loss that is the sum of the available usable limits of the cards from which fraudulent transactions are committed.

Different methods are used for cost sensitivity. They mainly include the machine learning approach, decision tree approach. In machine learning approach two techniques called over sampling and under sampling is performed, in which the latter obtained a good result. In decision tree approach, decision tree algorithms are used in which misclassification cost is considered in pruning step. A cost matrix is used to find the varying misclassification cost. After finding the misclassification cost the one with minimum value is used. By finding the misclassification cost not only the node value is obtained but also it predicts whether the transaction is fraudulent or not. This study using misclassification cost has made a significant improvement in fraud detection.

2.4 Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective

In this study mainly two approaches namely misuses (supervised) and anomaly detection (unsupervised) technique is being used. After this a classification is also used for checking the capability to process categorical and numerical data. In the first approach the data is classified as fraud based on previous data. With the help of this dataset classification models are also created, which can predict whether the data is fraud or not. The different classification models used are decision tree, neural network, rule induction etc. This has obtained a successful result and this approach is also called as misuse approach. While the second approach is based on account behavior. A transaction is said to be fraudulent if it possess the features opposite to the user's normal behavior. The behavior of user's model are extracted and accordingly classified as fraudulent or not. This technique of finding fraud is also called as anomaly detection.

2.5 Credit Card Fraud Detection Using Hidden Markov Model

As the E-commerce technology is increasing day by day the use of credit card has also been increased. As a result of this the fraud using credit card is also increasing. In all fraud detection systems, fraud will be detected only after the fraud has taken place. In this study a sequence of operations are modelled using Hidden Markov Model (HMM) and this can be used for the detection of fraud. It is trained with the normal behavior of the card holder. If the incoming transaction is not accepted by the trained HMM with high probability it is considered as fraudulent otherwise not.

A hidden Markov Model represents a finite number of states with sufficiently high probability. The transition between the states are handled by these probability values. A possible outcome will be generated based on the probability distribution. This outcome will be visible to the external users that is the states are hidden to the users hence the name. It is a perfect solution for predicting fraud transactions in addition it also provides extreme decrease in the number of false positive transactions recognized by fraud detection system. For prediction purpose three values are being used namely low, medium, high.

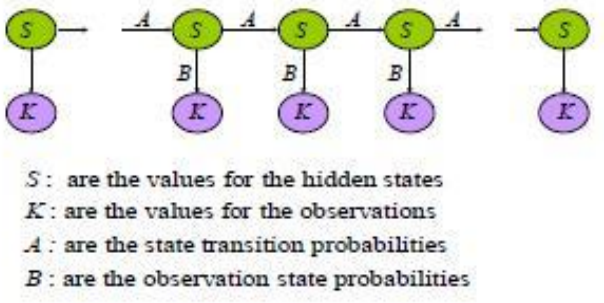


Fig. 1: HMM Finite set of states

The main advantage is that it does not require fraud signatures it is capable to detect by bearing in mind card holders spending habit. This is done by creating a set of clusters and identify the spending profile. These data are stored in the form of clusters with low, medium and high values. The probability is based on the spending behavior and further the processing is done. If the transaction is found to be fraudulent an alarm is generated, in addition to this a security form will also arise bearing a certain number of questions. This model can detect fraud transactions to an extent. It is scalable in handling large amount of data.

2.6 Real Time Credit Card Fraud Detection using Computational Intelligence

As the growth of technology is increasing it has made a big impact in credit transactions. This has made the fraudsters to commit the fraud transactions easily. Many fraud detection techniques are available but cannot solve fraud problems easily. As a solution to all this, SOM (Self Organizing Map) is used. It helps to decipher, filter and analyze customer behavior for fraud detection. SOM is a unsupervised neural network algorithm which is used to configure neurons according to the topological structure of input data. It divides the data into genuine and fraudulent transactions sets. The incoming transactions are compared with the previous transactions in the genuine set, then it is called as genuine set. The incoming transactions are compared with the previous transactions in the fraudulent transaction, then it is called as fraudulent set. So in this it is important to form legal card holder and fraudulent profile. For fraud detection a layered approach is used, which is depicted in the below figure.

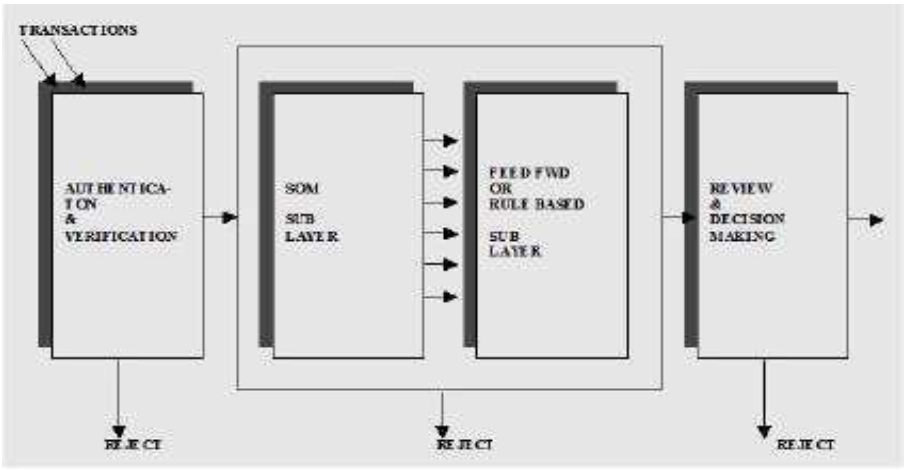


Fig. 2: System Layers.

The initial stage consist of authentication and screening layers. The fraudsters always comes with new ideas to commit fraud. The key to fraud detection lies in finding such a dynamic system to detect fraud that changes with the e-commerce trends. This system not only deals with customer profiles, merchant profiles and their selling price. It should also include rules and policies in the market place. By involving these attributes it increases the accuracy to detect the fraud. SOM helps to detect fraud to a great extent.

2.7 A Novel Machine Learning Algorithm to credit card fraud detection

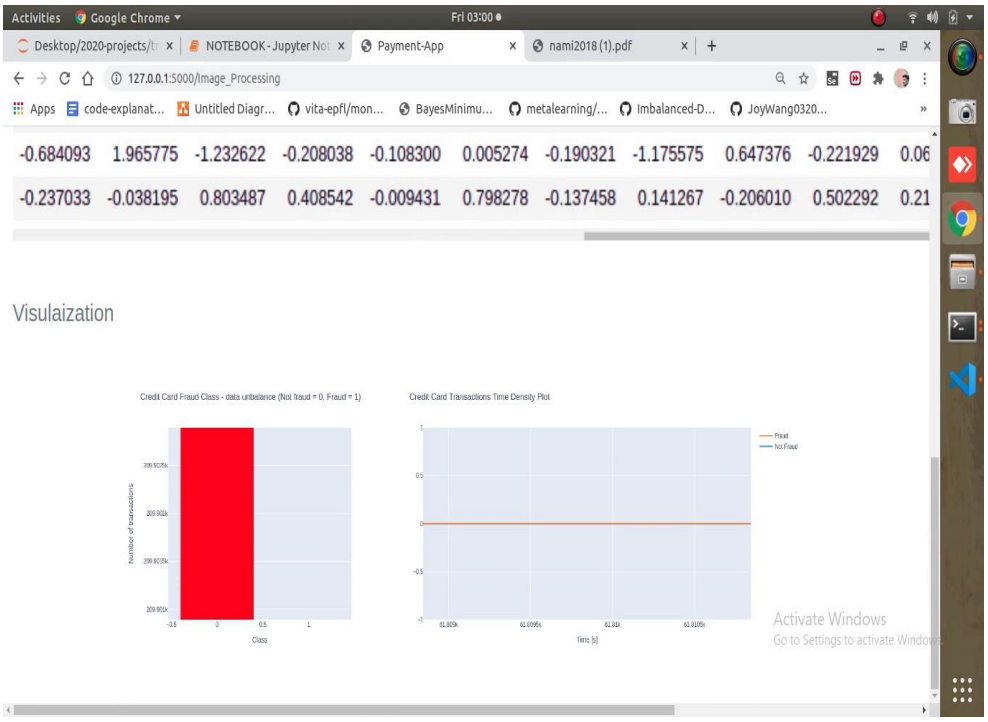
The use of credit card is rising day by day as the e-commerce is also increasing. The problem that happens with this is that fraud using credit card is increasing. It is a recurrent problem in almost all countries. But the trend seen is that countries with more credit card transactions are having less credit card fraud on the other hand countries with average credit card transactions are having high rate of credit card fraud. So in order to avoid this proactive methods are needed. In this a novel machine learning algorithm called cortical algorithm is used. It is the learning algorithm of hierarchical temporal memory and is inspired by the neo cortex of the brain

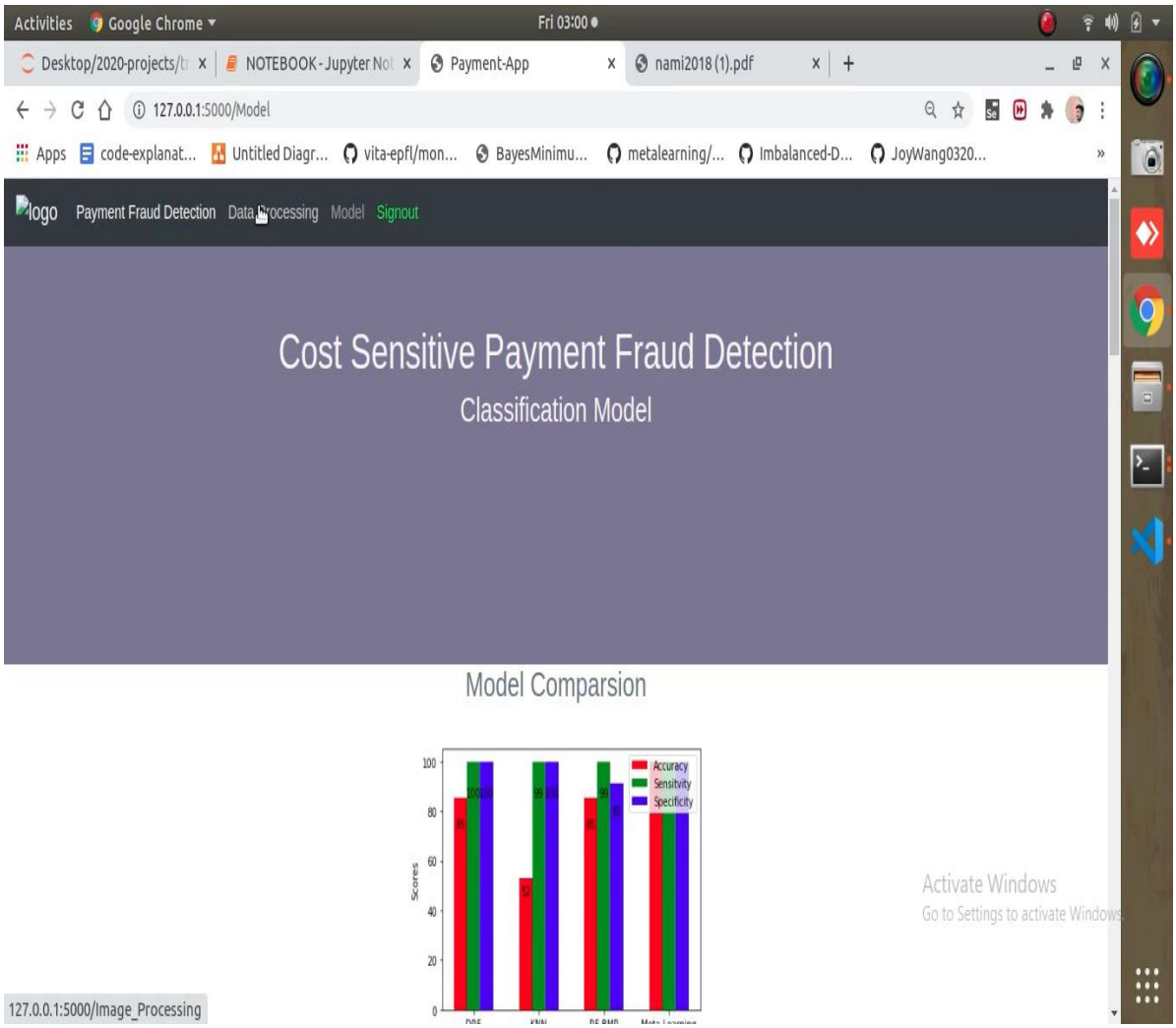
3. PROPOSED SYSTEM

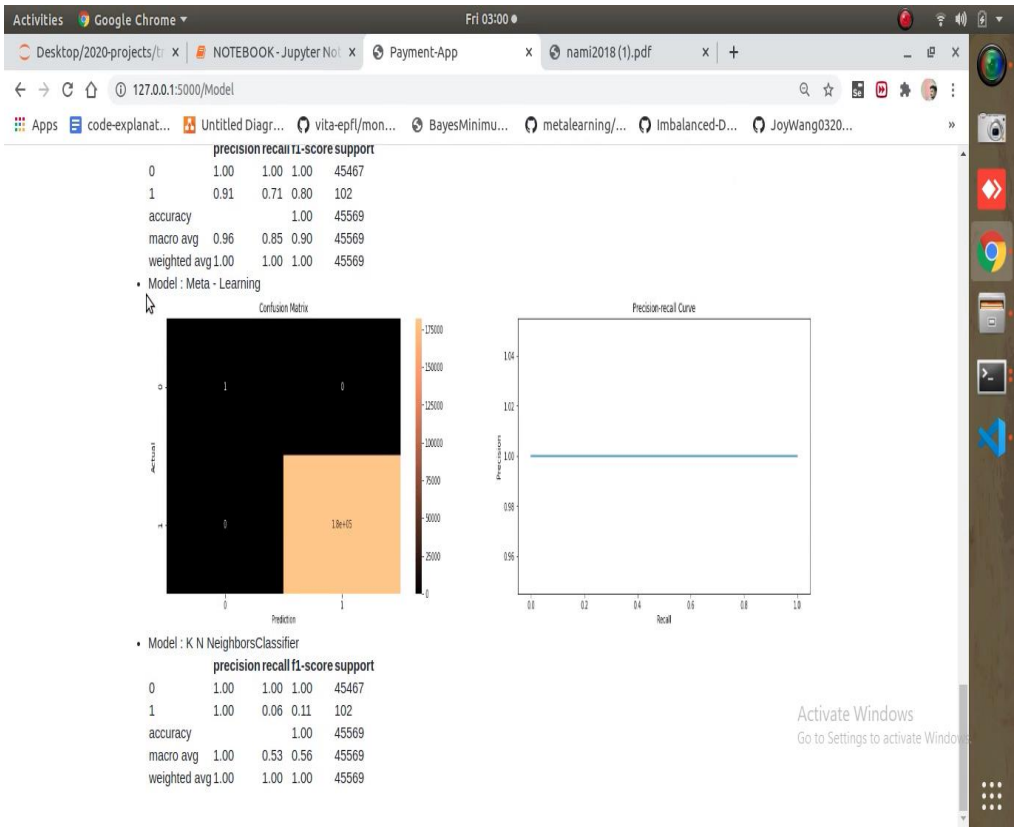
In this paper, the imbalanced nature of the credit card data has been taken into consideration, and a cost-sensitive weighted random forest technique has been proposed to overcome the issue. The Random Forest (RF) algorithm is basically an ensemble learning technique which works on the principle of Bagging, i.e. the dataset has been divided into n-bags or samples, with randomly selected

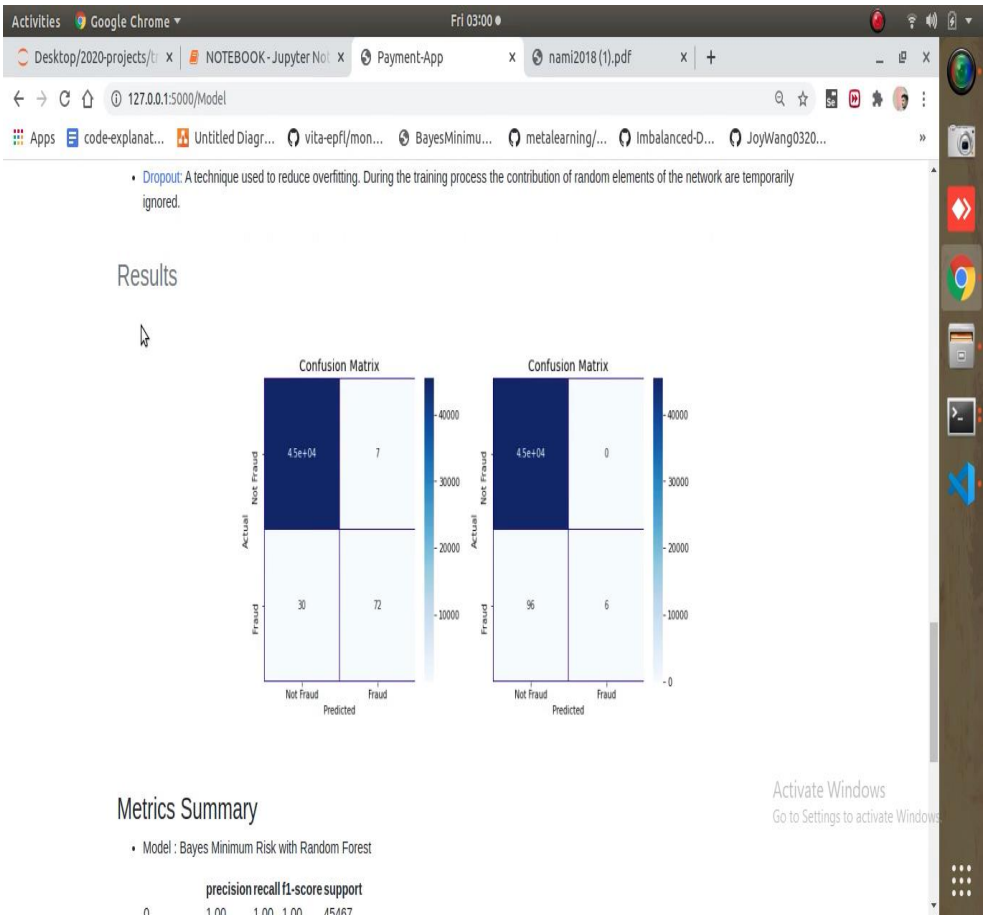
instances, and each of which is termed as “Tree” The final output of the model is achieved based on majority voting (in case of classification model), or averaging (in case of regression model) of all the test outcomes, with respect to all the trained trees. The foremost advantage of using RF technique for credit card fraud detection is that it can readily deal with the enormous size the training data, as RF involves to divide the dataset into a number samples, and then learn. Through the proposed approach, a cost- driven learning scheme is adapted to give more emphasis on learning the minority class instances. The rest of the paper is organized as follows. In the proposed scheme, a cost- sensitive weighted RF algorithm has been proposed for credit card fraud detection. A confusion -matrix based cost-function is defined to put more emphasis on the minority class instances during training. Instead of conventional majority voting strategy, a weighing strategy, based on minimum error achieved for a single tree.

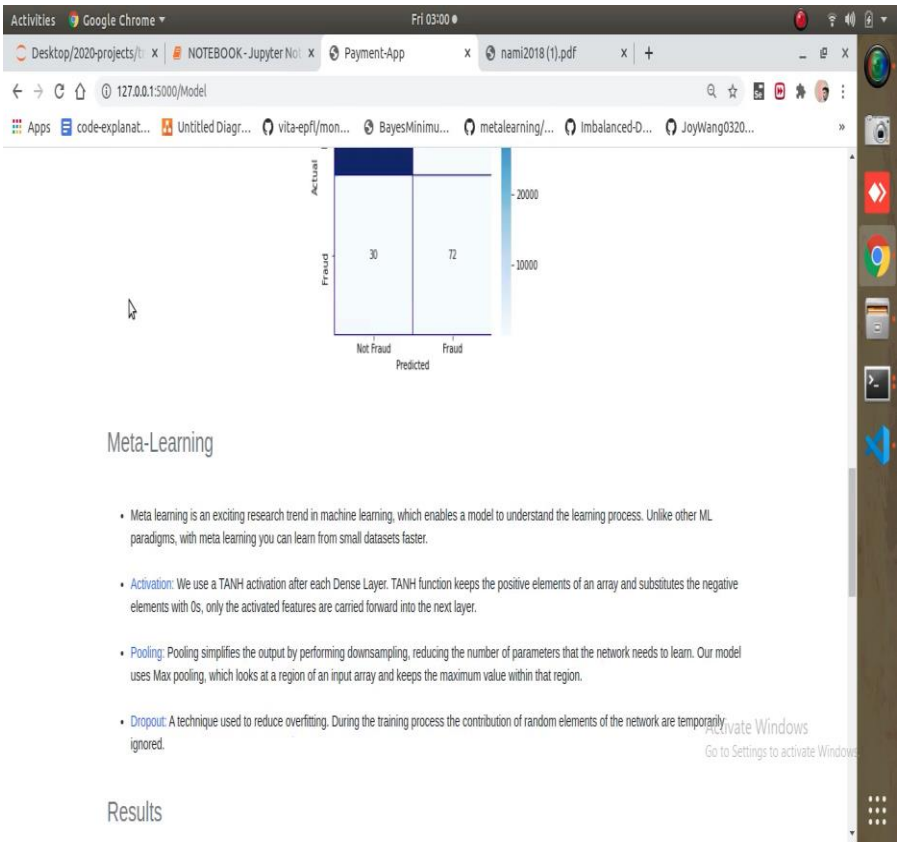
4. RESULTS

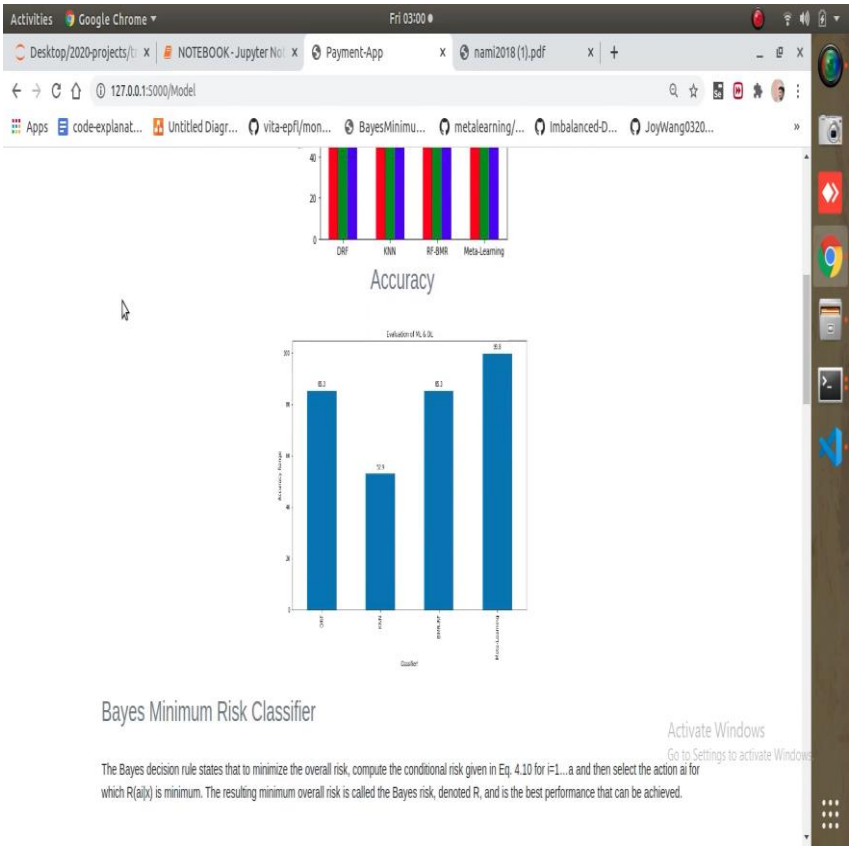


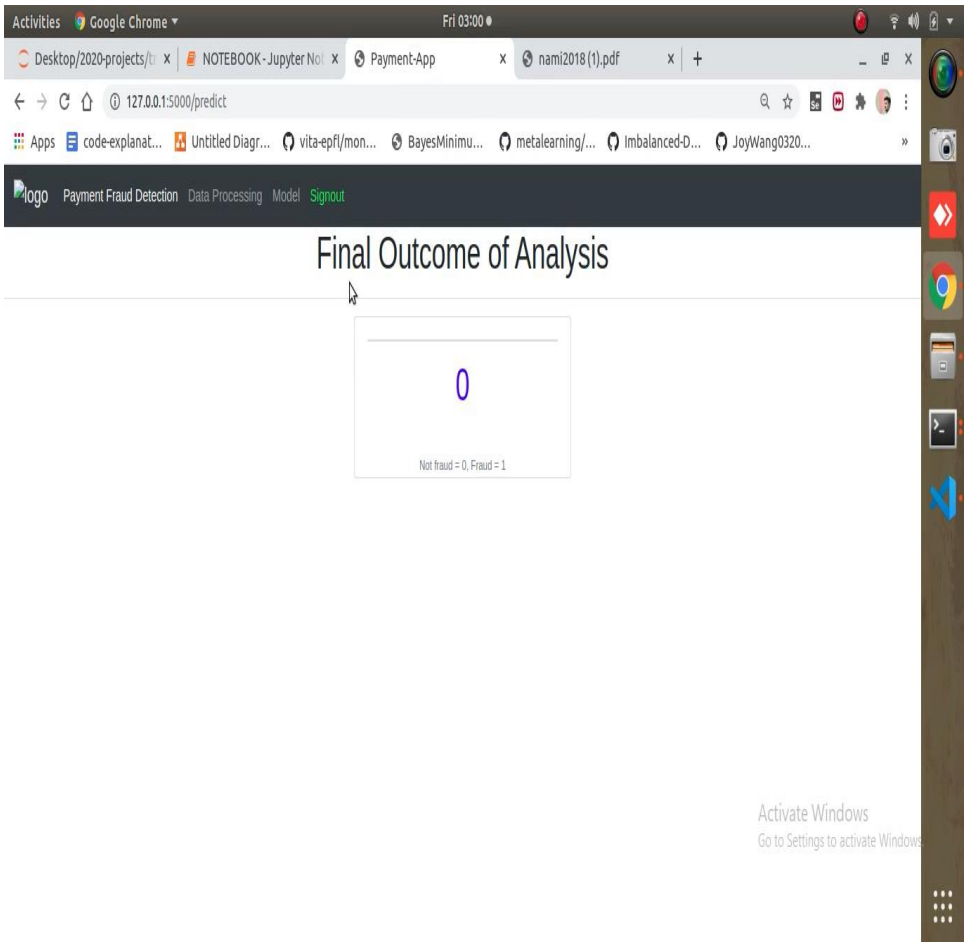

Activities
Google Chrome
Fri 03:00
Desktop/2020-projects/t
NOTEBOOK - Jupyter Not
Payment-App
nami2018 (1).pdf
127.0.0.1:5000/Model
Apps
code-explanat...
Untitled Diagr...
vita-epfl/mon...
BayesMinimu...
metalearning/...
Imbalanced-D...
JoyWang0320...
logo
Payment Fraud Detection
Data Processing
Model
Signout
Cost Sensitive Payment Fraud Detection
Classification Model
Model Comparsion
Accuracy
Sensitivity
Specificity
Scores
Activate Windows
Go to Settings to activate Windows.
127.0.0.1:5000/Image_Processing

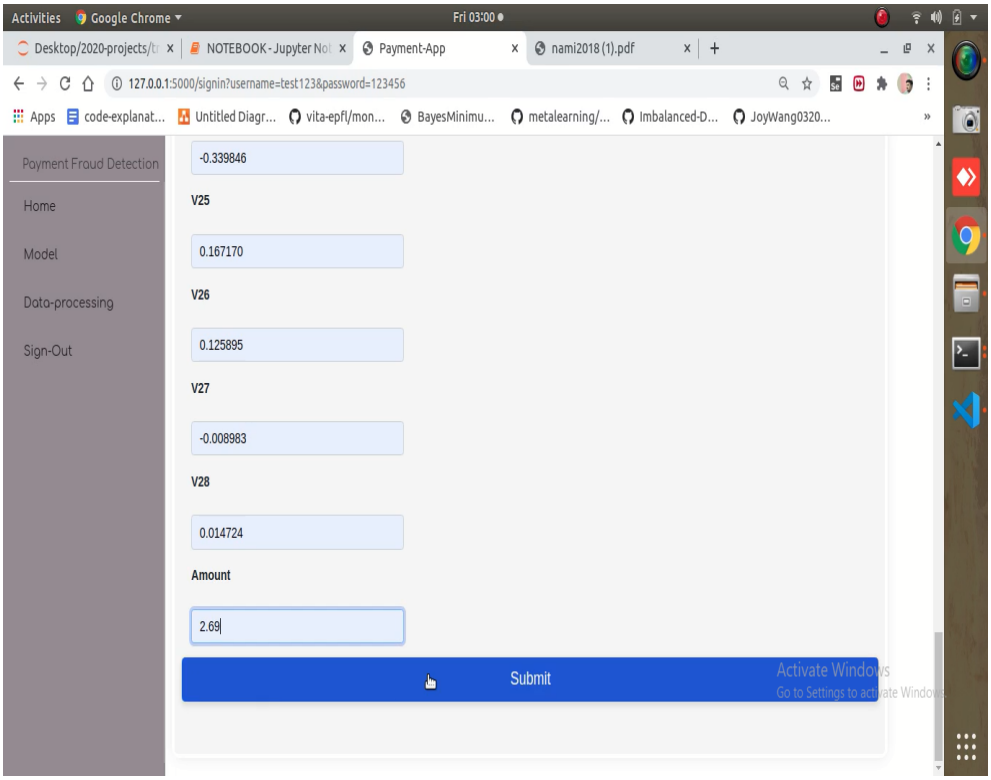

Activities
Google Chrome
Fri 03:00
Payment-App
nami2018 (1).pdf
127.0.0.1:5000/Model
Apps
code-explanat...
Untitled Diagr...
vita-epfl/mon...
BayesMinimu...
metalearning/...
Imbalanced-D...
JoyWang0320...
precision recall f1-score support
0 1.00 1.00 1.00 45467
1 0.91 0.71 0.80 102
accuracy 1.00 45569
macro avg 0.96 0.85 0.90 45569
weighted avg 1.00 1.00 1.00 45569
Model : Meta - Learning
Confusion Matrix
Precision-recall Curve
Model : K N NeighborsClassifier
precision recall f1-score support
0 1.00 1.00 1.00 45467
1 1.00 0.06 0.11 102
accuracy 1.00 45569
macro avg 1.00 0.53 0.56 45569
weighted avg 1.00 1.00 1.00 45569
Activate Windows

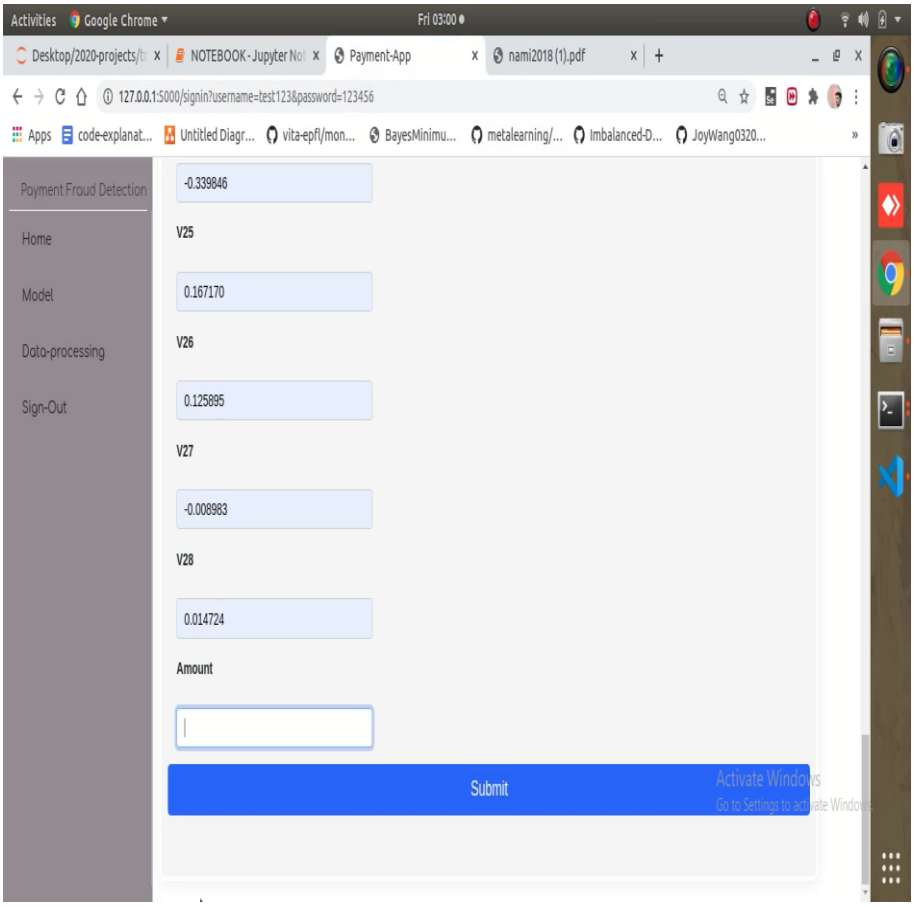

Activities
Google Chrome
Fri 03:00
Payment-App
nami2018 (1).pdf
127.0.0.1:5000/Model
Apps
code-explanat...
Untitled Diagr...
vita-epfl/mon...
BayesMinimu...
metalearning/...
Imbalanced-D...
JoyWang0320...
Dropout: A technique used to reduce overfitting. During the training process the contribution of random elements of the network are temporarily ignored.
Results
Confusion Matrix
Confusion Matrix
Actual
Not Fraud
Fraud
Predicted
Metrics Summary
Model : Bayes Minimum Risk with Random Forest
precision recall f1-score support
Activate Windows

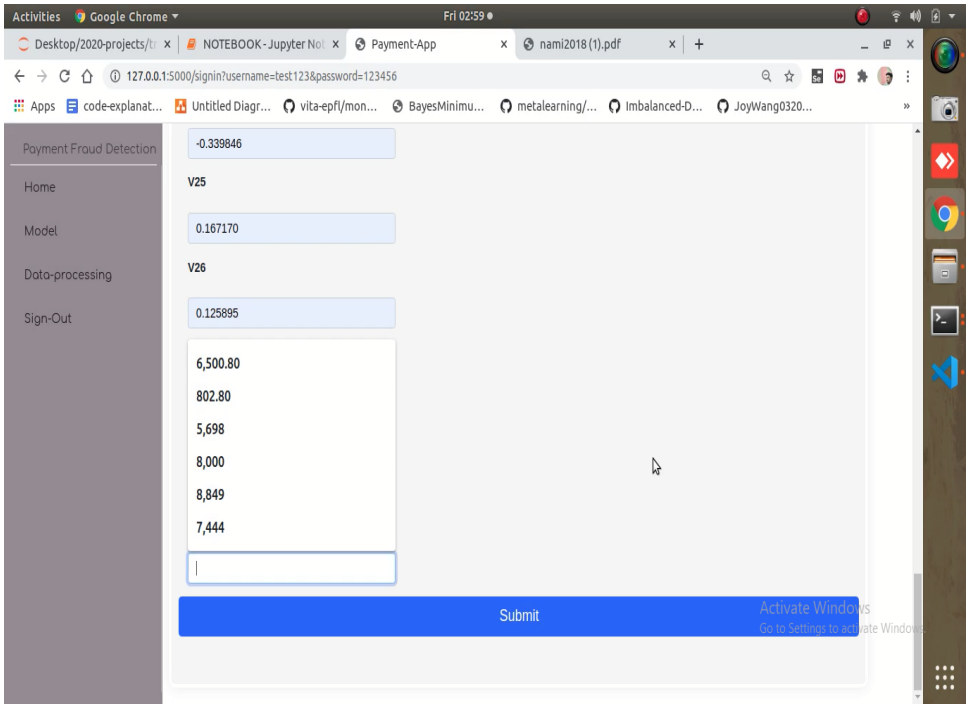

Activities
Google Chrome
Fri 03:00
Desktop/2020-projects/t
NOTEBOOK - Jupyter No
Payment-App
nami2018 (1).pdf
127.0.0.1:5000/Model
Apps
code-explanat...
Untitled Diagr...
vita-epfl/mon...
BayesMinimu...
metalearning/...
Imbalanced-D...
JoyWang0320...
Actual
Fraud
30
72
20000
10000
Not Fraud
Fraud
Predicted
Meta-Learning
Meta learning is an exciting research trend in machine learning, which enables a model to understand the learning process. Unlike other ML paradigms, with meta learning you can learn from small datasets faster.
Activation: We use a TANH activation after each Dense Layer. TANH function keeps the positive elements of an array and substitutes the negative elements with 0s, only the activated features are carried forward into the next layer.
Pooling: Pooling simplifies the output by performing downsampling, reducing the number of parameters that the network needs to learn. Our model uses Max pooling, which looks at a region of an input array and keeps the maximum value within that region.
Dropout: A technique used to reduce overfitting. During the training process the contribution of random elements of the network are temporarily ignored.
Activate Windows
Go to Settings to activate Windows.
Results

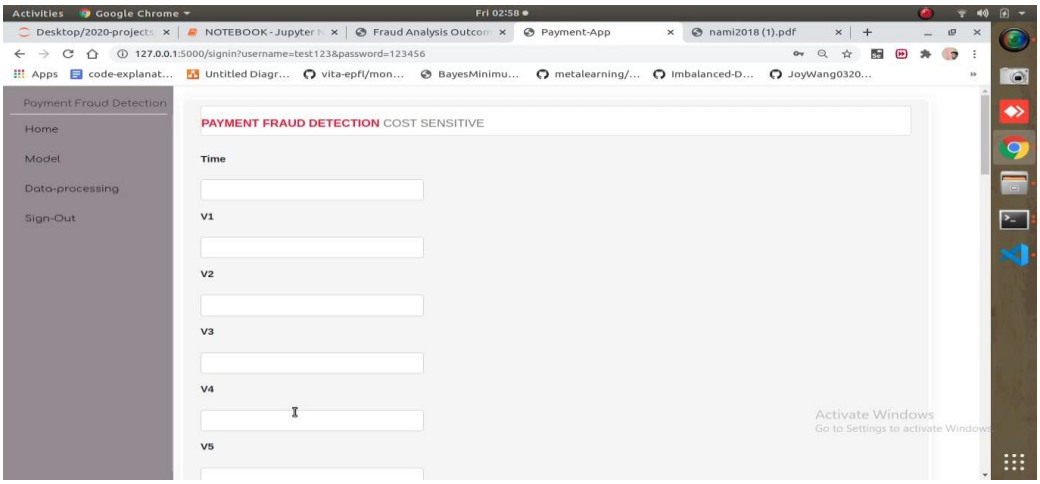

Activities
Google Chrome
Fri 03:00
Desktop/2020-projects/t
NOTEBOOK - Jupyter No
Payment-App
nami2018 (1).pdf
127.0.0.1:5000/Model
Apps
code-explanat...
Untitled Diagr...
vita-epfl/mon...
BayesMinimu...
metalearning/...
Imbalanced-D...
JoyWang0320...
40
20
0
DRF
KNN
RF-BMR
Meta-Learning
Accuracy
Bayes Minimum Risk Classifier
Activate Windows
Go to Settings to activate Windows.
The Bayes decision rule states that to minimize the overall risk, compute the conditional risk given in Eq. 4.10 for i=1...a and then select the action ai for which R(ai|x) is minimum. The resulting minimum overall risk is called the Bayes risk, denoted R, and is the best performance that can be achieved.

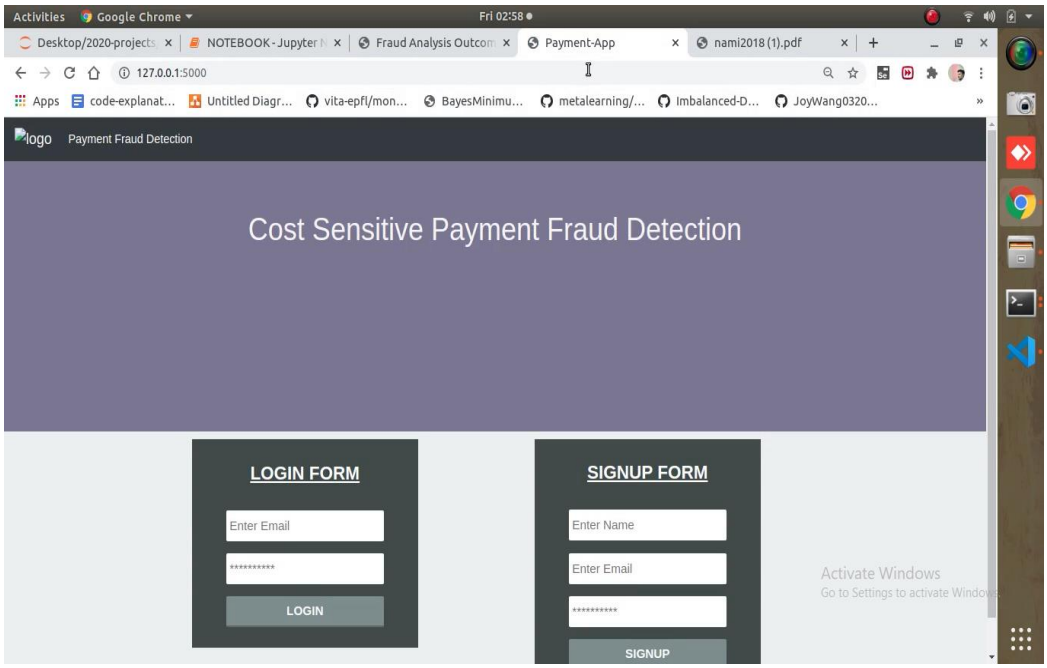

Activities
Google Chrome
Fri 03:00
Desktop/2020-projects/t
NOTEBOOK - Jupyter No
Payment-App
nami2018 (1).pdf
127.0.0.1:5000/predict
Apps
code-explanat...
Untitled Diagr...
vita-epfl/mon...
BayesMinimu...
metalearning/...
Imbalanced-D...
JoyWang0320...
logo
Payment Fraud Detection
Data Processing
Model
Signout
Final Outcome of Analysis
0
Not fraud = 0, Fraud = 1
Activate Windows
Go to Settings to activate Windows.

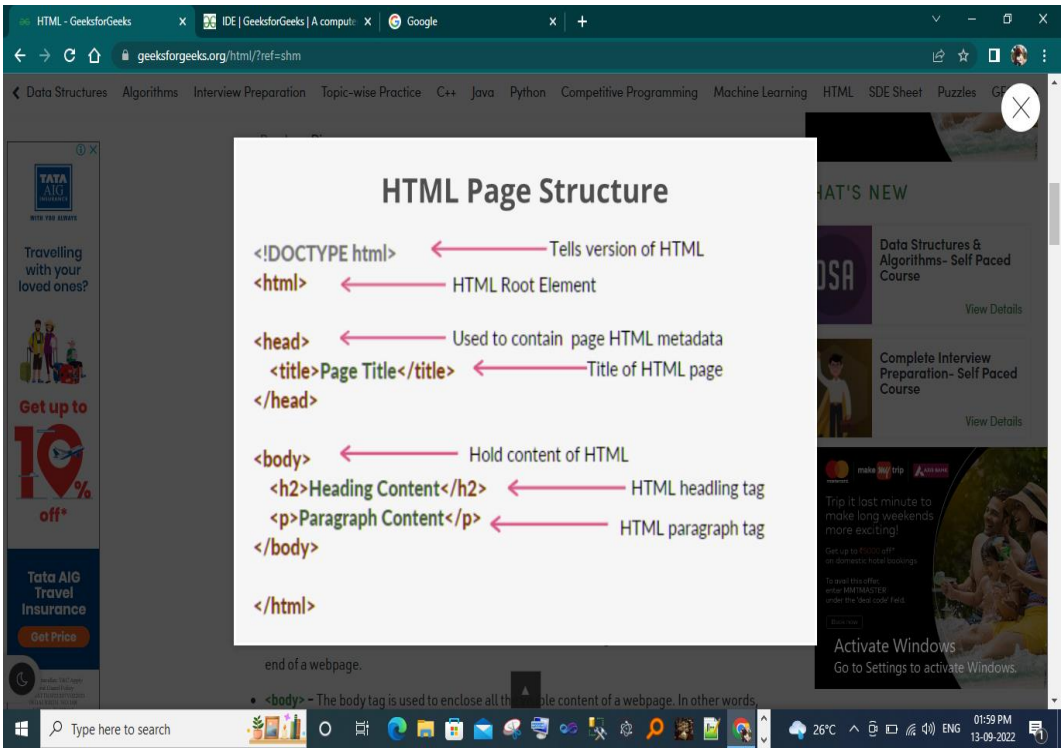

Activities
Google Chrome
Fri 03:00
Desktop/2020-projects/t
NOTEBOOK - Jupyter No
Payment-App
nami2018 (1).pdf
127.0.0.1:5000/signin?username=test123&password=123456
Apps
code-explanat...
Untitled Diagr...
vita-epfl/mon...
BayesMinimu...
metalearning/...
Imbalanced-D...
JoyWang0320...
Payment Fraud Detection
Home
Model
Data-processing
Sign-Out
-0.339846
V25
0.167170
V26
0.125895
V27
-0.008983
V28
0.014724
Amount
2.69
Submit
Activate Windows


Activities
Google Chrome
Fri 03:00
Desktop/2020-projects/t
NOTEBOOK - Jupyter No
Payment-App
nami2018 (1).pdf
127.0.0.1:5000/signin?username=test123&password=123456
Apps
code-explanat...
Untitled Diagr...
vita-epfl/mon...
BayesMinimu...
metalearning/...
Imbalanced-D...
JoyWang0320...
Payment Fraud Detection
Home
Model
Data-processing
Sign-Out
-0.339846
V25
0.167170
V26
0.125895
V27
-0.008983
V28
0.014724
Amount
Submit
Activate Windows


Activities
Google Chrome
Fri 02:59
Desktop/2020-projects/t
NOTEBOOK - Jupyter No
Payment-App
nami2018 (1).pdf
127.0.0.1:5000/signin?username=test123&password=123456
Apps
code-explanat...
Untitled Diagr...
vita-epfl/mon...
BayesMinimu...
metalearning/...
Imbalanced-D...
JoyWang0320...
Payment Fraud Detection
Home
Model
Data-processing
Sign-Out
-0.339846
V25
0.167170
V26
0.125895
6,500.80
802.80
5,698
8,000
8,849
7,444
Submit
Activate Windows


Payment Fraud Detection
Home
Model
Data-processing
Sign-Out
PAYMENT FRAUD DETECTION COST SENSITIVE
Time
V1
V2
V3
V4
V5


Payment Fraud Detection
Cost Sensitive Payment Fraud Detection
LOGIN FORM
Enter Email
LOGIN
SIGNUP FORM
Enter Name
Enter Email
SIGNUP



5. CONCLUSION

In this paper, a cost-sensitive random forest based ensemble learning technique has been proposed for effective detection of credit card fraud. The imbalanced nature of credit card data has been investigated in this work. A misclassification ratio-based cost-function has been integrated into the error-formulation of the generated sub-tress in RF-bagging. The cost-function facilitates to determine the predictive ability of sub-tree, so that the sub-tree with highest predictive ability can have maximum weight age. The final outcome of the test-set is determined based on the outcome, yielded by the maximally weighted tree. The experimental results achieved have manifested the efficiency of the proposed model in effective handling of imbalanced cases in case of credit card fraud detection. The proposed model has not been validated for high-dimensional datasets. The proposed model can be extended by integrating with different data cleaning techniques such as sampling or feature selection (or extraction) algorithm to be implemented in high-dimensional datasets. The imbalanced nature of the detection methods and its effects over the performance has not been investigated in this paper. A deep exploration of the issue in accordance with the computational performance of the credit card fraud detection techniques is another scope of future study.

REFERENCES

[1] L. Breiman, Random forests. Machine Learning, vol. 45, issue 1, pp: 5-32, 2021.

[2] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, Credit card fraud detection using Machine Learning Techniques: A Comparative Analysis, pp: 1-9, 2020.

[3] S. Patil, H. Somavanshi, J. Gaikwad, A. Deshmane, and R. Badgujar, Credit Card Fraud Detection Using Decision Tree Induction Algorithm, International Journal of Computer Science and Mobile Computing (IJCSMC), vol.4, issue 4, pp. 92-95, 2015. ISSN: 2320-088X

[4] G. Singh, R. Gupta, A. Rastogi, M. D. S. Chandel, and A. Riyaz, A Machine Learning Approach for Detection of Fraud based on SVM, International Journal of Scientific Engineering and Technology, vol.1, issue 3, pp. 194-198, 2021.

[5] A. C. Bahnsen, A. Stojanovic, D. Aouada, and B. Ottersten, Improving credit card fraud detection with calibrated probabilities. In Proceedings of the 2020 SIAM International Conference on Data Mining, pp. 677-685, 2021.

[6] F. N. Ogwueleka, Data Mining Application in Credit Card Fraud Detection System, Journal of Engineering Science and Technology, vol. 6, issue 3, pp. 311 – 322, 2020.

[7] K. RamaKalyani, and D. UmaDevi, Fraud Detection of Credit Card Payment System by Genetic Algorithm, International Journal of Scientific & Engineering Research, vol. 3, issue 7, pp. 1 – 6, 2021.

[8] P. L., Meshram, and P. Bhanarkar, Credit and ATM Card Fraud Detection Using Genetic Approach, International Journal of Engineering Research & Technology (IJERT), vol. 1, issue 10, pp. 1 – 5, 2019.

[9] K. R. Seeja, and M. Zareapoor, FraudMiner: A Novel Credit Card Fraud Detection Model Based on Frequent Itemset Mining, The Scientific World Journal, Hindawi Publishing Corporation, vol. 2018, Article ID 252797, pp. 1 – 10, 2017, http://dx.doi.org/10.1155/2017/252797.

[10] M. Zareappor, P. Shamsolmoali, Application of Credit card Fraud Detection: Based onBagging Ensemble Classifier, International Conference on Intelliegent Computing, Communication, & Convergence (ICCC-2016), In Procedia Computer Science, Vol. 48, pp: 679 – 685, 2016.

[11] V, Patil, U. K. Lilhore, A Survey on Different Data Mining & Machine Learning Methods for Credit Card Fraud Detection, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Vol. 3, Issue 5, pp:320-325, 2016.

[12] S. Stolfo, D. W. Fan, W. Lee, A. Prodromidis, and P. Chan, Credit card fraud detection using meta-learning: Issues and initial results. In AAAI-97 Workshop on Fraud Detection and Risk Management., 2014.

[13] A. Y. Ng, and M. I. Jordan, On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. Advances in neural information processing systems, Vol. 2, pp: 841-848, 2015.

[14] C. Phua, D. Alahakoon, and V. Lee, Minority report in fraud detection: classification of skewed data. ACM sigkdd explorations newsletter, vol. 6, issue 1, pp: 50-59, 2014.

[15] F. Chu, Y. Wang, C. Zaniolo, An adaptive learnin approach for noisy data streams. IEEE International Conference on Data Mining, 2004 (ICDM'04), pp: 351- 354, 2014.

[16] A. Shen, R., Tong, and Y. Deng, Application of classification models on credit card fraud detection. IEEE International Conference on Service Systems and Service Management, pp: 1-4, 2015.

[17] Y. Sahin, and, E. Duman, Detecting Credit Card Fraud by Decision Trees and Support Vector Machines, Proceeding of International Multi-Conference of Engineers and Computer Scientists (IMECS 2012), vol. 1, pp: 1- 6, ISBN: 978-988-18210-3-4, ISSN: 2078-0966 (Online), 2011.

[18] Y. Sahin, and, E. Duman, Detecting credit card fraud by ANN and logistic regression. In IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA), pp: 315-319, 2011.