e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

Volume:04/Issue:09/September-2022 Impact Factor- 6.752 www.irjmets.com

HYBRID MACHINE LEARNING ALGORITHM FOR CREDIT CARD FRAUD DETECTION

Shishobitveer Singh[*1], Vinay Chopra[*2]

[*1,2]DAV Institute of Engineering & Technology, Jalandhar, India.

DOI: https://www.doi.org/10.56726/IRJMETS29605

ABSTRACT

The prediction analysis (PA) refers to an approach using which the upcoming prospect can be predicted from the previous occurred events. This technique has two phases namely feature extraction and classification process. The detection of fraud in credit card using PA becomes challenging because of the complexity in datasets. There are numerous classification methods that are implemented in state-of-art schemes in order to detect the frauds in credit card. The prediction analysis is a DM (data mining) method which is useful for future forecasting on the basis of current information. This research is carried out to perform the CCFD (credit card fraud detection) on the basis of recent information. The data of credit card is available in an enormous volume form. Consequently, it becomes difficult to establish association among diverse features that have impact on the predictive accuracy. The KNN (K nearest Neighbor) is deployed to extract the attributes and the PA (prediction analysis) is performed by the means of NB (naïve bayes) algorithm.

Keywords: Credit Card Fraud, KNN, Naïve Bayes, Data Mining.

I. INTRODUCTION

Fraudsters and detectors of credit card fraud transactions have long been maintaining a dynamic approach. Particularly in the current cyberspace age, incidents of transactional fraud are more frequent than ever and bring about significant financial losses. The Nilsson Report provided extensive research into the state of card fraudulent cases around the world. The total economic loss due to credit card fraud in 2017 was $28.65 billion. Worse yet, global card fraud losses will continue to grow annually and will likely reach $34 billion in 2022. Consequently, banks and financial institutions need an efficient fraud detection system to trace or monitor payments that take place over web. Most fraud detection frameworks have a similar purpose, that is, mining distrustful payment patterns from multiple payment records and using them to trace or monitor incoming payments. Machine learning has shown outstanding efficiency in mining such patterns that be considered as a function of supervised binary classification. Put differently, a high-performance classifier can be trained to identify deceitful transactions with plenteous transaction logs. While machine learning has done a tremendous job in detecting fraudulent transactions, there should be no disruption in improving the existing fraud detection models as a small step forward can reduce the economic losses to a great extent.

Cost-sensitive learning is an elective methodology of managing this by applying different misclassification error costs to different classes, and the minority class is normally doled out with a more prominent cost. A few works treat the information dynamic variety issue as idea float. Their definitive objective is to early distinguish the presence of idea float and to adaptively update a classifier to plan for new presumptions. There are characteristic distinctive features among fraudulent and legitimate exchanges. Thus, it is likewise essential to have a strong portrayal equipped for recognizing fraudulent payments from honest ones, while the techniques for fraud are continually evolving. Building a proficient credit card fraud detection model includes a couple of fundamental steps that significantly influence detection cycle. The initial step is feature engineering which means to separate enlightening features of the exchange conduct of clients. Raw features like time and date of transaction and worth of payments may not depict the conditional way of behaving of card proprietors and shams proficiently. The most well-known approach is to take on an payment collection procedure to extricate a few new characteristics. Exchanges are gathered in view of a chose meeting, card number, payment type and dealer code to remove collected features. The subsequent stage is to process the quantity of payments and the general cash spent on those payment. A payment with rough attributes is transformed into a feature matrix with more infomercial conglomeration traits, trailed by a pattern of transaction or payment aggregation. The following after feature engineering is to train a classifier as a binary classification function. By and by, the learned classifier perceives most extreme number of fraudulent payments as genuine ones when class imbalance issue isn't dealt with. This is on the grounds that larger part of classifiers is dependent upon the default hypothesis of a reasonable informational index, and thusly, the learned choice limit show tendency towards the class with additional cases. Subsequently, handling the issue of class imbalance has transformed into a basic step before training a fraud detection system. Data sampling is one of the most often utilized techniques to manage the issue of class imbalance. Specifically, the inactivity of genuine payments can be decreased by utilizing the under-sampling strategy which speeds up the model training process. Irregular under-sampling is one of the most notable under-sampling procedures for its simplicity and effectiveness. By and by, these sampling procedures don't address the spatial distribution of models from unique classes.

An under-sampling strategy called gaussian mixture can be executed to test more helpful models and, subsequently, upgrade the efficiency of the classification engineering. Be that as it may, assuming the data set contains fundamentally less fraudulent transactions than legitimate, an up-sampling technique, as SMOTE, ought to be utilized to feature the fraudulent transactions. Subsequent to handling the class imbalance issue, training a fraudulent exchange detection system as a binary classifier should be possible with a nearly adjusted data set. Many AI strategies, as SVM RF, CNN, and Recurrent Neural Networks (RNNs), have been productively utilized for fraudulent transactions' detection. The greater part of them are connected with representation learning. Their goal is to find a superior portrayal of the contribution through learning the modifications of the data that seclude the parts of changeover in the data and hold most extreme data. Specifically, deep representation learning along deep neural networks has made unprecedent progress in different fields in the beyond couple of years due to a few imaginative pieces.

Since very unbalanced data and diffuse examples impact the prediction accuracy of standard ML algorithms, and a few non-static data break the standards of classic clustering and classification strategies, there has been an expanded exploration interest in utilizing new procedures to manage this test in current years. Both supervised and unsupervised methods have been advanced for recognizing credit card frauds. Non-supervised techniques incorporate exception/peculiarity detection strategies that treat any exchange as phony that doesn't conform to the larger part. Supervised procedures are without a doubt the most well-performing strategies in fraud detection, which influence labeled transactions to train a classifier. The feature vectors of genuine transactions are classified, or sometimes the posterior of the classifier is analyzed to detect frauds. Researchers have tested many classification algorithms in terms of credit card transactions for fraud detection. Random Forest is related to supervised learning algorithm. Its principle is to create a number of decision trees through the selection of random samples and random attributes. Lastly, the classification results of many decision trees are obtained based on the rule that the minority is inferior to the majority. On contrary to a single decision tree, random forest is able to efficiently mitigate the risk of overfitting, successfully balance the error for imbalanced data, and rapidly decide the significance of features. A random forest is a classifier comprising an array of tree-configured classifiers $\{h(\mathbf{x},\ominus_k), k = 1, \ldots.\}$where {⊖_k} are independently uniformly distributed random vectors and each tree has cast one unit vote for the most well-known class on the input. For each tree in the random forest a new training set is produced by drawing it with replacements from the new training set. Then, features are selected randomly at each node for growing tree on the new training set. The resultant trees are not pruned. SVM is a non-probabilistic linear classification algorithm which is able to learn for discriminating the data comes under two classes. For this, the linear boundary recognized as hyperplane is searched due to which the margin amid two known classes is increased. In case, the input denotes an array x that has n attributes which indicated that it is a point in n-dimensional space. A linear surface of dimension n-1 is discovered using Support vector machines for dividing two clouds of n-dimensional points that comes in the two classes. The hyperplane parameters are optimized for increasing its distance from the closest point which is a problem whose solution is required to mitigate a quadratic error function. Though, SVM is a linear classifier, the nonlinear kernel trick can be utilized to carry out the nonlinear classification process. Furthermore, the adjustment of this model can be easily done in case the separation of two classes is not performed clearly in n-dimensional space by relaxing the hard margin constraint in the context of a soft margin. The Support Vector Machine is adaptable for dealing with multiclass problems in various ways, generally with the integration of a bank of SVM classifiers. AdaBoost algorithm refers to an ensemble method. This method casts vote to the weighted predictions of the weak learner to construct powerful classifiers. This algorithm has performed outstandingly in many applications such as credit card fraud detection and intrusion detection systems. Most machine learning algorithms have a similar problem of overfitting which results in lower productivity of classifiers. Nevertheless, there is usually less chance for the classifiers trained with the AdaBoost technique to be overfit and, at the same time, have a lower risk of high false-positive predictions. In the AdaBoost application, an elected algorithm uses the primary input data for the base classifier's training. Also, an adjustment is made in the weights of patterns, and more weight is assigned to the misclassified instances. Moreover, the changed occurrences are utilized to train the ensuing base learner, which endeavours to address the misclassifications from the past models. The iteration goes on until the predefined number of models is fabricated, or data has no more misclassified instances.

II. LITERATURE REVIEW

C. Wang, et.al (2018) suggested a WOA-BP (whale algorithm optimized back propagation) to detect the credit card fraud so that the issues related to slow convergence rate, local optimum, network defects and lower stability were tackled [14]. WSO (whale swarm optimization) algorithm was implemented for optimizing the weight of BP network. Initially, this approach made the deployment of WOA algorithm for acquiring an optimal primary value. Subsequently, the error values were corrected using BPNN algorithm for acquiring the optimal value. Eventually, MATLAB was applied to simulate the suggested algorithm. The simulation outcomes depicted that the suggested algorithm yielded accuracy and fast convergence speed while detecting the fraud in credit card.

E. Esenogho, et.al (2022) projected an effectual technique for detecting the fraud in credit card in which a NN (neural network) ensemble algorithm and a hybrid data resampling technique was implemented [15]. The initial algorithm incorporated the LSTM (long short-term memory) in the AdaBoost (adaptive boosting) method. In the meantime, the latter algorithm was planned on the basis of SMOTE-ENN (synthetic minority oversampling technique and edited nearest neighbor). A real time dataset was employed to compute the projected technique. The experimental results indicated the supremacy of the initial algorithm over other. Moreover, this algorithm offered the sensitivity of 99.6% and specificity of 99.8% to detect CCF (credit card fraud).

A. A. Taha, et.al (2020) investigated an intelligent technique called OLightGBM (optimized light gradient boosting machine) to detect fraud in credit card [16]. The components of LightGBM were refined by integrating BHO (Bayesian-based hyper-parameter optimization) algorithm. The investigated technique was evaluated on 2 real time datasets in which frauds and authentic transactions were comprised. The experimental outcomes revealed that the investigated technique performed well as compared to other methods with regard to an accuracy of 98.40%, AUC (Area under receiver operating characteristic curve) of 92.88%, precision of 97.34% and F1-score of 56.95% for detecting the CCF.

D. Cheng, et.al (2022) introduced a STAGN (spatial-temporal attention-based graph network) in order to detect fraud in credit card [17]. Generally, a GNN (graph neural network) for learning graph attributes. After that, STA (spatial-temporal attention) was exploited on the top of learned tensor representations which a 3D ConvNet (convolution network) employed later on. The E2E (end-to-end) scheme with 3D convolution and detection network was presented to learn the attentional weights. A real time dataset was applied in the experimentation. The experimental outcomes depicted that the introduced algorithm outperformed the traditional algorithms concerning AUC and precision-recall curves. Furthermore, the introduced algorithm was able to detect the suspected transactions.

G. K. Arun, et.al (2020) developed a new DL (deep learning) based C-LSTM (convolutional long short term memory) approach with the objective of detecting the CCF (credit card fraud) [18]. This algorithm had diverse stages such as to pre-process, and classify the data. The transactions were classified relied on pre-processed data for detecting whether the fraud was conducted or not. German Credit and Kaggle datasets were executed to quantify the developed approach. The experimental results confirmed that the developed approach provided an accuracy of 94% on first dataset and 94.65% on second.

G. K. Arun, et.al (2022) formulated BEPO-OGRU algorithm in which the BEPO (binary emperor penguin optimization) was employed with OGRU (optimal gated recurrent unit) to detect the frauds in credit cards [19]. This algorithm emphasized on detecting and classifying the possible credit card frauds. Moreover, the BEPO algorithm was assisted in selecting a promising set of attributes. Thereafter, the OGRU algorithm was adopted to select the hyper-parameters of the GRU (gated recurrent unit). For this, HHO (Harris Hawks Optimization) algorithm was implemented. The simulations exhibited that the accuracy of the formulated algorithm was calculated 94.78% on the German Credit dataset and 94.16% on other dataset.

A. A. El Naby, et.al (2021) presented an effectual framework for detecting and preventing the fraud in credit card in advance [20]. The fraudulent transactions were detected using DL (deep learning) methods. A framework was put forward to predict the transactions as authentic or fraud on Kaggle dataset. The presented framework was known as OSCNN (Over Sampling with Convolution Neural Network) that was planned on the basis of over-sampling pre-processing and CNN (convolution neural network). Furthermore, it also deployed MLP (Multi-layer perceptron) on the dataset. The outcomes reported that the presented framework attained the accuracy of 98% to detect frauds in CC.

Y. Ling, et.al (2021) devised a CSHIM (cost-sensitive heterogeneous integration model) in order to detect the fraud in credit card [21]. Various costs of every transaction were taken into consideration and the higher efficacy of diverse classification algorithm was put together. To achieve this, CSWDSF (cost-sensitive weighted Dempster-Shafer fusion) theory was applied for generating promising outcomes. The fundamental objective of the devised model was to alleviate the monetary losses and enhance the accuracy. The open-source datasets were employed to conduct the experiments. The experimental results demonstrated that the devised model led to save the cost around 74.69% in contrast to the conventional techniques for detecting the frauds.

H. Zhou, et.al (2019) suggested a risk control algorithm to achieve the integrated learning of SMOTE (synthetic minority oversampling technique) and PCA (principal component analysis) to process the data. A single-layer DT (decision tree) and the enhanced Adaboost algorithm along with the trained risk control system had exploited to individual data so that the fraud was recognized [22]. The testing results on data sample of a commercial bank confirmed that the suggested algorithm attained the accuracy of96.50% and F-Measure value is 97.3%. Moreover, the suggested algorithm performed more effectively as compared to other method for detecting the fraud in credit card.

D. Devi, et.al (2019) established a CSWRF (cost-sensitive weighted random forest) algorithm to detect the fraud in credit card [23]. This algorithm was trained on the basis of CF (cost-function) for every tree in bagging. Moreover, this algorithm focused on assigning more weight to the minority instances. The ranking of the trees was done in accordance with their predictive potential of the minority class instances. A comparative analysis was conducted on the established algorithm against two other methods on two binary datasets. The experimental results demonstrated that the established algorithm was effectual and offered enhanced G-mean, F-measure and AUC values.

E. M. S. W. Balagolla, et.al (2021) presented an approach that allowed to model the credit card transactions on a blockchain to decentralize the credit card processing and verify it with an accredited set of computing nodes [24]. This approach had potential for diminishing the fraud. Additionally, a scaling system was implemented to blockchain as the existing projects were not scalable. A proactive anomaly detection was suggested for detecting the fraudulent transactions in credit card. The results validated the applicability of the presented approach for preventing frauds.

III. RESEARCH METHODOLOGY

A huge development in the scams has been noticed which causes the misuse of credit cards, in last few years. In recent time days, distinct technologies are deployed by the criminals to conduct these kinds of frauds. Particularly, the card of authentic user is stolen to conduct such scams. A huge financial loss is faced y the cardholder and card issuing organization in case the cardholder is not ware about the stealing of its credit card. The attacker needs only little information to carry out any fraud transaction to do online payment. The internet or telephone devices are useful to purchase any product and service online. In some scenarios, the pattern, which assisted in making the transactions, is employed to know about the stealing of card. Thus, a technique to detect the fraud has required in order mitigate such scams. This research work emphasizes on forecasting the fraud transactions of credit card. This technique is also called PA (prediction analysis) as it is applicable to forecast the future potentials with regard to the existing data. This method has contained two phases in which attributes are extracted and classification is performed. The KNN and NB algorithms are utilized in this method. The K-Nearest Neighbor is implemented as base classification model and a basic and significant ML (machine learning) model. The classification process is often carried out using KNN algorithm. This classifier assists in detecting the intrusion, DM (data mining) and recognizing the pattern. The KNN is a non-parametric due to which is cannot be re-utilized in the practical conditions. Thus, this algorithm is unable to make any basic assumptions related to the data distribution. Some given data named training data is employed in this algorithm. This data is applied to classify the coordinates into groups whose detection is done through a feature [12]. The distance among the data points is evaluated using KNN. This distance is recognized as the Euclidean Distance. The formula utilized for the distance measurement is mentioned as:

$$P(A\backslash B)=(P(B\backslash A)P(A))/(P(B))$$

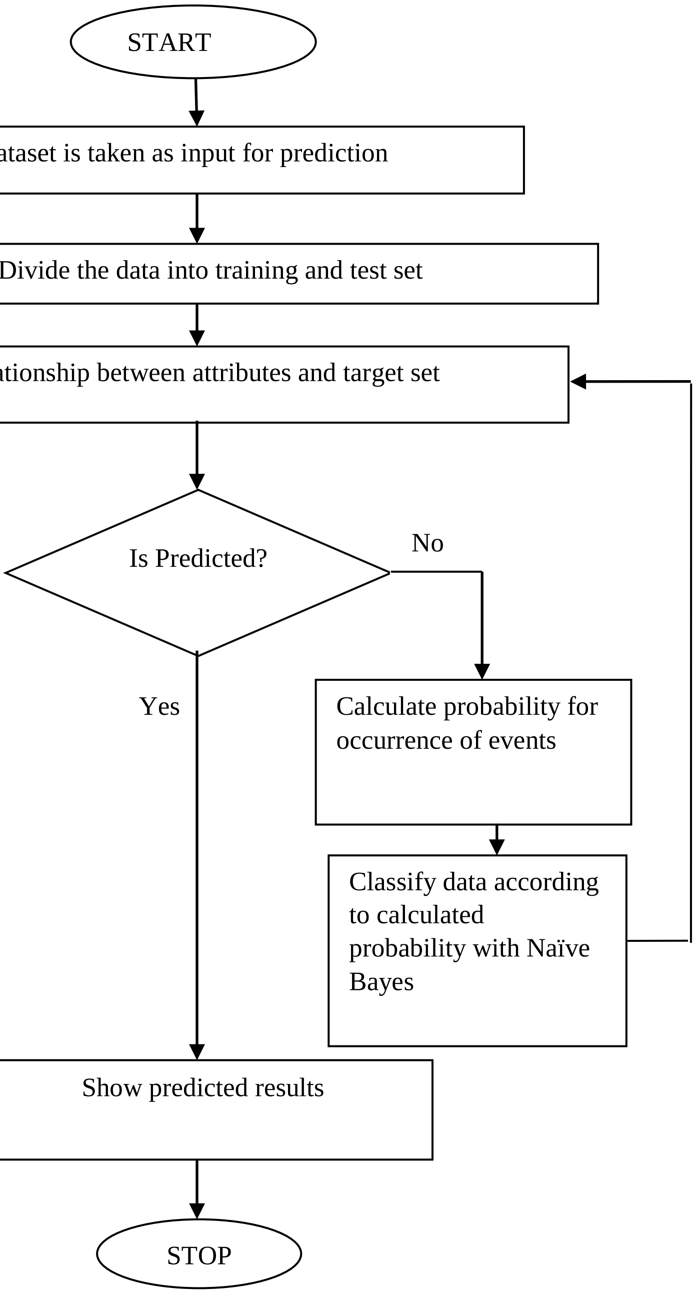

START
Dataset is taken as input for prediction
Divide the data into training and test set
Establish relationship between attributes and target set
Is Predicted?
No
Yes
Calculate probability for occurrence of events
Classify data according to calculated probability with Naïve Bayes
Show predicted results
STOP

Figure 1: Proposed Methodology

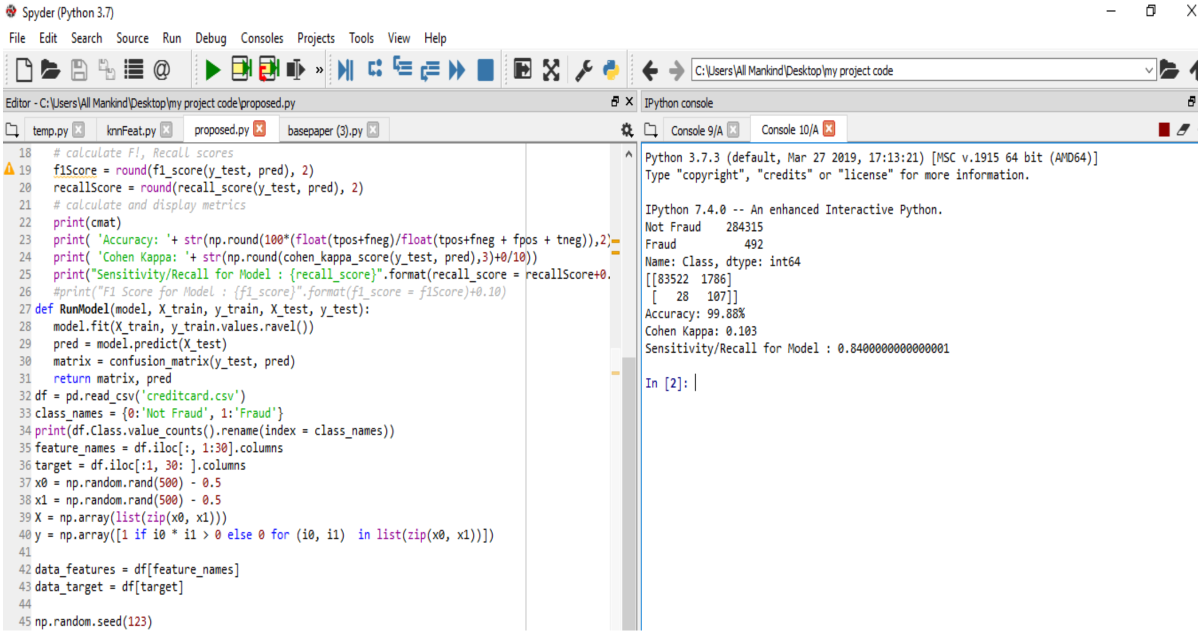
In this, n is the number of dimensions or attributes in the ML (machine learning). The hypothesis is that the data point having position at the minimal distance from the test point is related to the similar category. The same process is executed in n number of dimensions using this formula. Thus, it is executed with n number of attributes. The results are acquired using this classification method with regard to the forecasting. The result is fed as an input in NB (Naive Bayes) algorithm. This probabilistic classifier is based on the ML. The NB algorithm is adopted to perform the classification based on Bayes theorem. This classification algorithm is described as the set of classification algorithms whose collection is done in accordance with the of Bayes' Theorem. A common concept is employed in this group of algorithms. This implies that, the selection of every feature pair is performed in an independent way to carry out the classification and no attribute is depending upon the other attributes. The possibility of occurrence of A, offered that B has happened and Bayes theorem is executed to determine it. The resulting output is created through this NB algorithm. The suggested approach is expected to generate result accurately to detect the scams of credit card. The suggested approach is compared with other algorithms in order to perform its quantification.

IV. RESULT AND DISCUSSION

This work implements the proposed and the existing techniques of Credit Card Fraud Detection (CCFD). The technique initiated in this work is hybrid in nature and an amalgamation of KNN and naïve Bayes algorithms. The data used for this work has been described in the table below:

Table 1: Description of dataset

Attribute	Value
Dataset Characteristics	Multivariate
Number of Instances	30000
Attribute Characteristics	Integer, Real
Associate Tasks	Classification
Number of Attributes	24
Missing Values	No
Area	Business
Date Donated	2016-01-26
Number of web Hits	368365


Spyder (Python 3.7)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\All Mankind\Desktop\my project code
Editor - C:\Users\All Mankind\Desktop\my project code\proposed.py
temp.py
knnFeat.py
proposed.py
basepaper (3).py
IPython console
Console 9/A
Console 10/A

```
    # calculate F!, Recall scores
    f1Score = round(f1_score(y_test, pred), 2)
    recallScore = round(recall_score(y_test, pred), 2)
    # calculate and display metrics
    print(cmat)
    print( 'Accuracy: '+ str(np.round(100*(float(tpos+fneg)/float(tpos+fneg + fpos + tneg)),2)
    print( 'Cohen Kappa: '+ str(np.round(cohen_kappa_score(y_test, pred),3)+0/10))
    print("Sensitivity/Recall for Model : {recall_score}".format(recall_score = recallScore+0.
    #print("F1 Score for Model : {f1_score}".format(f1_score = f1Score)+0.10)
def RunModel(model, X_train, y_train, X_test, y_test):
    model.fit(X_train, y_train.values.ravel())
    pred = model.predict(X_test)
    matrix = confusion_matrix(y_test, pred)
    return matrix, pred
df = pd.read_csv('creditcard.csv')
class_names = {0:'Not Fraud', 1:'Fraud'}
print(df.Class.value_counts().rename(index = class_names))
feature_names = df.iloc[:, 1:30].columns
target = df.iloc[:1, 30: ].columns
x0 = np.random.rand(500) - 0.5
x1 = np.random.rand(500) - 0.5
X = np.array(list(zip(x0, x1)))
y = np.array([1 if i0 * i1 > 0 else 0 for (i0, i1)  in list(zip(x0, x1))])

data_features = df[feature_names]
data_target = df[target]

np.random.seed(123)
```

```
Python 3.7.3 (default, Mar 27 2019, 17:13:21) [MSC v.1915 64 bit (AMD64)]
Type "copyright", "credits" or "license" for more information.

IPython 7.4.0 -- An enhanced Interactive Python.
Not Fraud    284315
Fraud           492
Name: Class, dtype: int64
[[83522  1786]
 [   28   107]]
Accuracy: 99.88%
Cohen Kappa: 0.103
Sensitivity/Recall for Model : 0.8400000000000001

In [2]: 
```

Figure 2: Proposed Methodology Executions

The figure 2 shows execution of proposed methodology for the CFF prediction. The suggested technique is hybrid system of K-NN and naïve bayes algorithms. The Naïve Bayes and KNN are used for the tasks of classification and the feature extraction respectively.

Table 2: Table of Comparison

Classification approach	**Parameters**			
	Accuracy	Recall	Execution time	Precision
Voting	89.94%	0.76	00.795 second	0.76
Hybrid	99.88%	0.84	0.103 second	0.84
KNN Classifier	72.5 %	0.72	0.11 second	0.72
Naïve Bayes	61.5 %	0.61	0.80 second	0.61

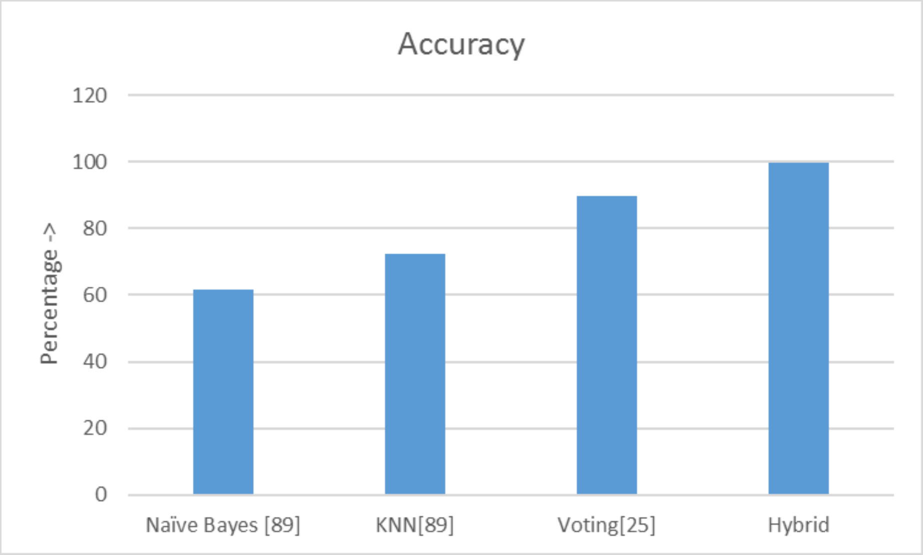


Figure 3: Accuracy Analysis

Figure 3 shows comparison between the voting-based classification and hybrid classification approaches in terms of accuracy. The hybrid classification scheme generates more accurate results than voting classification for detecting credit card related frauds.

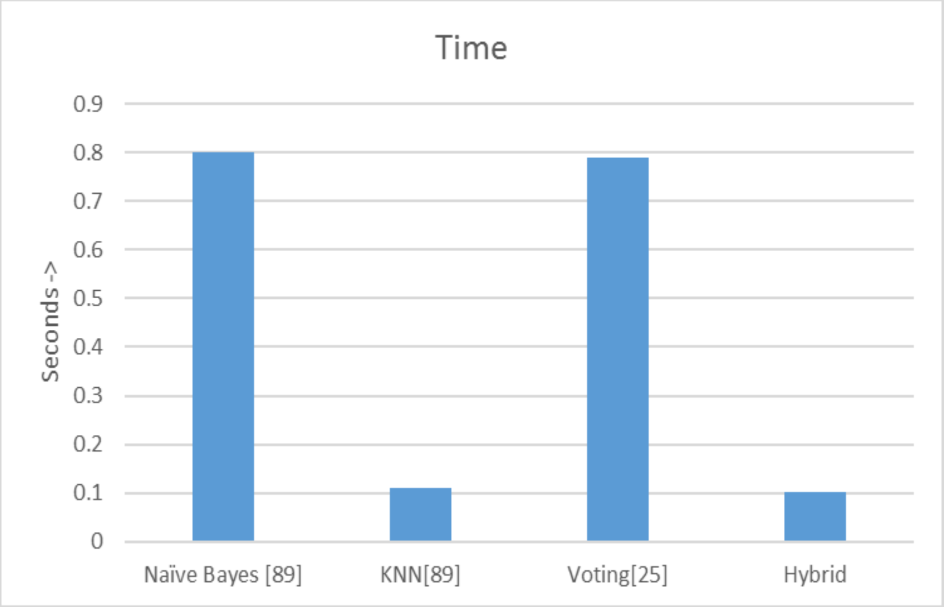


Figure 4: Execution Time

Figure 4 shows comparison between the voting-based classification and hybrid classification approaches in terms of execution time. The execution time of hybrid classification scheme is lower than voting classification in CCFD.

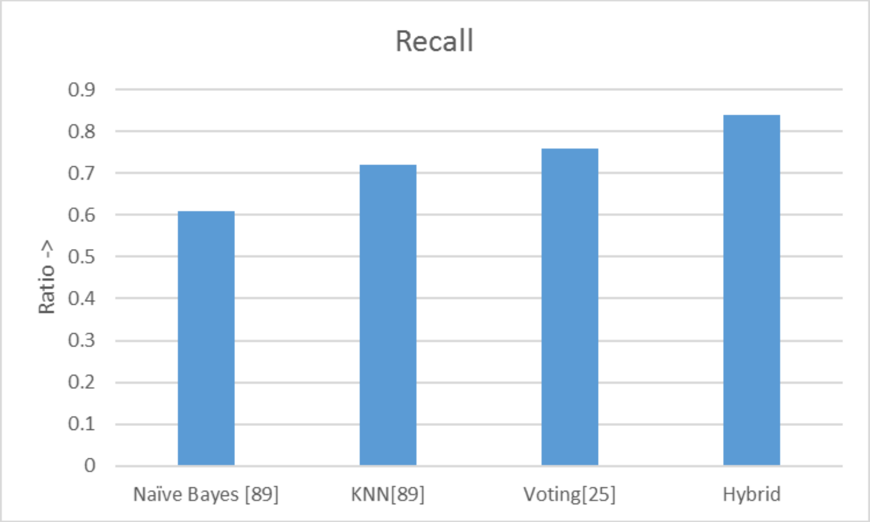


Figure 5: Recall Analysis

Figure 5 shows the comparison between the voting-based classification and hybrid classification approaches in terms of recall. The hybrid classification scheme provides a better recall rate than voting classification for detecting credit card-related frauds.

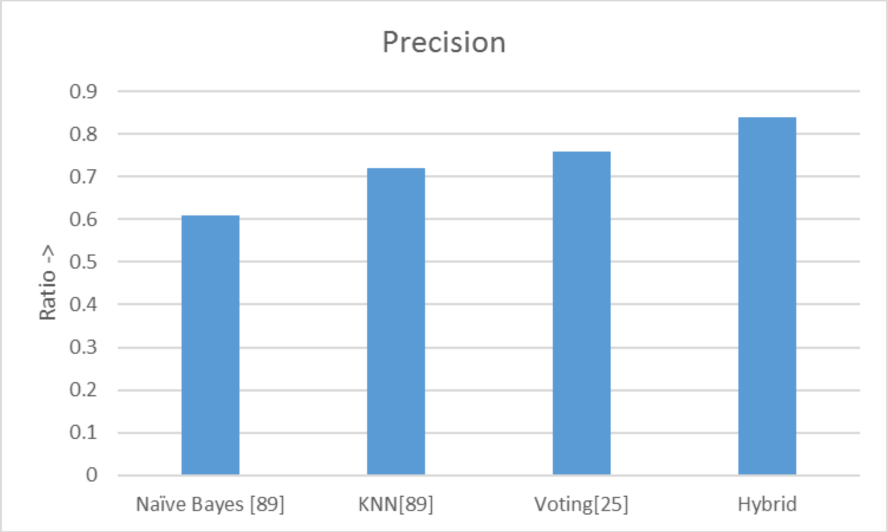


Figure 6: Precision Analysis

Figure 6 shows the comparison between the voting-based classification and hybrid classification approaches in terms of precision. The hybrid classification scheme provides a better precision rate than voting classification for detecting credit card-related frauds.

V. CONCLUSION

The prediction, on the other hand, aims to forecast the future possibilities on the basis of the historic episodes. The two main tasks in data mining are feature extraction and classification. Over the past few years, researchers have put forward many classification schemes for the detection of credit card related scams. The existing methods of the feature extraction are not efficient enough to establish relationships among features. This factor causes the employed classifier's performance to be degraded. The current work makes use of KNN and Naïve Bayes algorithms for the feature extraction and the classification respectively. This work implements proposed approach in python software. The results of the tests depict about 99% of improvement in the detection of the credit card related scams based on proposed approach.

VI. REFERENCES

[1] A. Agrawal, S. Kumar and A. K. Mishra, "A Novel Approach for Credit Card Fraud Detection," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015, pp. 8-11

[2] Y. Lucas et al., "Dataset Shift Quantification for Credit Card Fraud Detection," 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), 2019, pp. 97-100

[3] N. Carneiro, G. Figueira and M. Costa, “A data mining-based system for credit-card fraud detection in e-tail”, Decision Support Systems, vol. 95, no. 7, pp. 91-101, March 2017

[4] F. Carcillo, Y. L. Borgne, O. Caelen, Y. Kessaci and G. Bontempi, “Combining unsupervised and supervised learning in credit card fraud detection”, Information Sciences, vol. 2, no. 12, pp. 568-572, 16 May 2019

[5] M. Nur-E-Arefin and M. S. Islam, "Application of Computational Intelligence to Identify Credit Card Fraud," 2018 International Conference on Innovation in Engineering and Technology (ICIET), 2018, pp. 1-6

[6] A. Srivastava, M. Yadav, S. Basu, S. Salunkhe and M. Shabad, "Credit card fraud detection at merchant side using neural networks," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), 2016, pp. 667-670

[7] A. Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams and P. Beling, "Deep learning detecting fraud in credit card transactions," 2018 Systems and Information Engineering Design Symposium (SIEDS), 2018, pp. 129-134

[8] T. Choudhury, G. Dangi, T. P. Singh, A. Chauhan and A. Aggarwal, "An Efficient Way to Detect Credit Card Fraud Using Machine Learning Methodologies," 2018 Second International Conference on Green Computing and Internet of Things (ICGCIoT), 2018, pp. 591-597

[9] N. Kalaiselvi, S. Rajalakshmi, J. Padmavathi and J. B. Karthiga, "Credit Card Fraud Detection Using Learning to Rank Approach," 2018 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC), 2018, pp. 191-196

[10] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang and C. Jiang, "Random Forest for credit card fraud detection," 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 2018, pp. 1-6

[11] C. V. Priscilla and D. P. Prabha, "Influence of Optimizing XGBoost to handle Class Imbalance in Credit Card Fraud Detection," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 1309-1315

[12] R. R. Popat and J. Chaudhary, "A Survey on Credit Card Fraud Detection Using Machine Learning," 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), 2018, pp. 1120-1125

[13] J. O. Awoyemi, A. O. Adetunmbi and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," 2017 International Conference on Computing Networking and Informatics (ICCNI), 2017, pp. 1-9

[14] C. Wang, Y. Wang, Z. Ye, L. Yan, W. Cai and S. Pan, "Credit Card Fraud Detection Based on Whale Algorithm Optimized BP Neural Network," 2018 13th International Conference on Computer Science & Education (ICCSE), 2018, pp. 1-4

[15] E. Esenogho, I. D. Mienye, T. G. Swart, K. Aruleba and G. Obaido, "A Neural Network Ensemble With Feature Engineering for Improved Credit Card Fraud Detection," in IEEE Access, vol. 10, pp. 16400-16407, 2022

[16] A. A. Taha and S. J. Malebary, "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine," in IEEE Access, vol. 8, pp. 25579-25587, 2020

[17] D. Cheng, X. Wang, Y. Zhang and L. Zhang, "Graph Neural Network for Fraud Detection via Spatial-Temporal Attention," in IEEE Transactions on Knowledge and Data Engineering, vol. 34, no. 8, pp. 3800-3813, 1 Aug. 2022

[18] G. K. Arun and K. Venkatachalapathy, "Convolutional Long Short Term Memory Model for Credit Card Detection," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1168-1172

[19] G. K. Arun and P. Rajesh, "Design of Metaheuristic Feature Selection with Deep Learning Based Credit Card Fraud Detection Model," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), 2022, pp. 191-197

[20] A. A. El Naby, E. El-Din Hemdan and A. El-Sayed, "Deep Learning Approach for Credit Card Fraud Detection," 2021 International Conference on Electronic Engineering (ICEEM), 2021, pp. 1-5

[21] Y. Ling, R. Zhang, M. Cen, X. Wang and M. Jiang, "Cost-sensitive Heterogeneous Integration for Credit Card Fraud Detection," 2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), 2021, pp. 750-757

[22] H. Zhou, L. Wei, G. Chen, P. Lin and Y. Lin, "Credit Card Fraud Identification Based on Principal Component Analysis and Improved Adaboost Algorithm," 2019 International Conference on Intelligent Computing, Automation and Systems (ICICAS), 2019, pp. 507-510

[23] D. Devi, S. K. Biswas and B. Purkayastha, "A Cost-sensitive weighted Random Forest Technique for Credit Card Fraud Detection," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2019, pp. 1-6

[24] E. M. S. W. Balagolla, W. P. C. Fernando, R. M. N. S. Rathnayake, M. J. M. R. P. Wijesekera, A. N. Senarathne and K. Y. Abeywardhana, "Credit Card Fraud Prevention Using Blockchain," 2021 6th International Conference for Convergence in Technology (I2CT), 2021, pp. 1-8