Fraudulent Transaction Detection in Credit Card by Applying Ensemble Machine Learning techniques

Debachudamani Prusti
Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela, India
debaprusti@gmail.com

Santanu Kumar Rath
Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela, India
skrath@nitrkl.ac.in

Abstract—**In the present day scenario, fraudulent activities associated with financial transactions, particularly while using credit card, are observed to be occurring in a fast rate. Hence, a fraud detection system involving various detection techniques is very much essential for the financial institutions to sustain the goodwill from the customers. Several fraud detection techniques have been proposed by researchers as well as practitioners with application of various algorithms to find the pattern of fraud. In this study, the application of various classification models are proposed by implementing machine learning techniques to find out the accuracy and other performance parameters to identify the fraudulent transaction. Classification algorithms such as K-Nearest Neighbor (K-NN), Extreme Learning Machine (ELM), Random Forest (RF), Multilayer Perceptron (MLP) and Bagging classifier have been implemented to critically assessed their performances and the performances are evaluated. We have proposed a predictive classification model by ensemble of five individual algorithms, as it provides a better predictive performance.**

Index Terms—**Credit card fraud, Fraud detection, Classification Model, Predictive Performance**

I. Introduction

One of the easiest payment method and popular way of financial transaction is the credit card payment system. But it is observed that there occurs a good number of fraudulent transactions with credit card. Fraudulent transactions using credit card can be defined as an unauthorized or illegal usage of card, unusual or suspicious transaction activity, or transaction on a dead card [1]. With rapid increase in credit card transactions, the card counterfeit is also growing in parallel. A good number of models are available in literature on various fraud detection techniques, which categorize the transaction as fraudulent or legit one. Credit card fraud is a pertinent and significant issue both in online and offline transaction systems as the fraudsters are able to overcome existing security parameters [18].

In the trigger of fraudsters, various type of financial frauds such as credit card fraud, corporate fraud, insurance fraud, bank fraud and money laundering have fetched a major concern and point of attention [1]. Fraud can be described as one of the leading method for the ill-use of organization's financial system with no direct approach of legal consequences. It is a sensitive issue of wrong activities against the laws, rules and policies with an intention to incur unauthorized or illegal financial gain. In a BBC news report, by the year 2015, fraudulent credit card operation had claimed a total cost of 21 billion dollars worldwide and it is expected to reach 31 billion dollars by 2020 [16] . The overall losses made by financial frauds are incalculable as there is no standard constraint to curb the issue. The fraudsters frequently change their techniques and patterns to perpetrate the illegal activities and able to access the cardholders account as well as from the financial institutions.

Credit Card Fraud Detection (CCFD) system mainly involves in distinguishing fraudulent financial data from the authentic data. By applying machine learning algorithms, the models help to identify the fraud pattern in the databases. Various challenges like non-availability of real dataset, size of dataset, determining the appropriate evaluation parameters and dynamic behavior of the fraudsters are associated with credit card fraud detection and they hinder the path of fraud detection.

Basically, credit card fraud is categorized as two types: behavioral and application fraud [2]. Behavioral fraud are considered as stolen or lost card, counterfeit card, mail theft and inactive cardholder fraud. But in application fraud a new card is obtained from the issuing companies or the agencies by submitting fake information or other cardholders information. In both type of frauds, the fraudsters obtain the card details without the knowledge of cardholders and later use them for carrying out various fraudulent activities to hijack the money from account.

Objective of the study

The objective of this study is to identify whether the transaction is fraudulent or not by using credit card transaction data and ultimately to find the detection rate at which the model is able to classify. The testing on the dataset has been performed on the input transaction data by choosing the classification algorithms. Machine learning algorithms provide the applications and frameworks to identify fraudulent activities with greater predictive accuracy. This study aims to:

- provide a systematic and comprehensive review of the existing articles in the application of machine learning classification algorithms to CCFD.
- Detect the fraudulent transactions in the credit card Efficiently with faster detection rate (true positive rate).

- Develop an efficient classification model (accurate fraud detection system) based on the machine learning classification algorithms in order to classify the instances in the dataset.
- Process the high dimensional data and to select the desired features using the feature selection method by removing irrelevant and redundant features.

II. Related Works

In the domain of credit card fraud detection technique, a good number of techniques are being implemented with the support of various machine learning algorithms by various researchers [17]. C. Whitrow et al., has addressed various common problems of presenting the data in the most possible way to a fraud classifier by formulating a framework for aggregating all the transaction records and investigated as to how this affects the detection of various performance parameters [3]. S. Maes et al., also have compared the performances of neural network classifiers with Bayesian networks for the credit card transaction classification models [4]. Boltan and Hand have made a study on fraudulent activities in credit card that makes interchange of ideas and keeps back with the potential innovation in credit card fraud detection [5].

M. Zareapoor et al., have presented a standard comparative performance study of several machine learning algorithms for fraud detection methods based on credit card [6]. The main idea behind this study is to review the methodology of various credit card detection methods. A. Srivastava et al., have presented Hidden Markov Model (HMM) application for credit card fraud detection [7]. The aggregation for the credit card transaction procedure are mapped out as a random process of an HMM. In this, They have used a particular range of transaction amount and others are considered as HMM. Also they have suggested a technique to find the expenditure profile of all cardholders to estimate the models performance parameters.

Research motivation on predictive analytics for fraud detection

The rapid increase of fraudulent transactions in the present day society helped to motivate in detecting the cause of fraud. Research on predictive analytics in machine learning intends to identify the hidden patterns and trends from the large volume of data [8]. It improvises the transaction process by enhancing the decision making for the prediction of various performance parameters. It has the ability to optimize and automate the decision to operate in real time transaction analysis. It provides an efficient way to learn the patterns of cardholders such as legal or fraudulent patterns. The study on predictive analytics includes various statistical methods to evaluate the outcomes of a future predictive class analysis. The most obvious advantage of having a proper fraud detection system is to maintain the goodwill among the customers.

III. Machine Learning Algorithms for Fraud detection

Various researchers and practitioners have implemented a good number of machine learning classifiers to develop an efficient fraud detection model that can identify the frauds accurately and quickly with better prediction accuracy. In this paper, we have considered five important classification machine learning models such as K-Nearest Neighbor (K-NN), Extreme Learning Machine (ELM), Random Forest (RF), Multilayer Perceptron (MLP) and Bagging classifier for investigating the performance parameters.

A. K-Nearest Neighbor Classifier

One of the standard classification technique which is implemented in CCFD is K-nearest neighbor classification algorithm, where the output of the new training sample is classified based on majority of K-NN class [6] [20]. Among all the active CCFD techniques of supervised learning pattern recognition and K-NN achieves very good performance consistently, without any analytical assumption related to the distributions in which the training samples are fetched properly. Fraud detection techniques based on K-NN technique in credit card require various distance measures is defined in between two data instances. While implementing K-NN technique, we classify any of the input transaction by estimating the nearest point to the new input transactions. If the nearest neighbor is identified to be a fraudulent transaction, then it is termed as a fraud one.

B. Extreme Learning Machine Technique

Extreme learning machine (ELM) is also a feedforward neural network system for classification, clustering, regression, compression, feature learning and sparse approximation with a single layer or numerous layers of hidden neurons or nodes, where the parameters of different hidden nodes require not be tuned [15]. These models can deliver a generalized performance and learn a number of times quicker than systems trained utilizing backpropagation. Generally, every parameter of the feedforward neural networks should be tuned and in this manner, there exists the dependency between various layers of parameters (loads and biases).

C. Random Forest Method

Random forest classification technique is an ensemble learning technique both for classification as well as regression. It is suitable for solving problems involves in the dataset into classes [11] [12]. In it, prediction value is achieved by using a set of Decision Trees (DT).

In Figure 1, the training process is done by taking a set of decision trees. A group of Decision Trees (DT) are built, which are then used for predicting the class. By considering the majority voting technique of all the single trees, we can achieve the final predictive output class from the highest voted class. RF is computationally efficient because every tree is constructed independently by others [19]. With a series of trees in the ensemble technique are robust to overfitting and noise in the data. Predictive class for new instances are obtained by aggregating every individual trees output in the ensemble method.

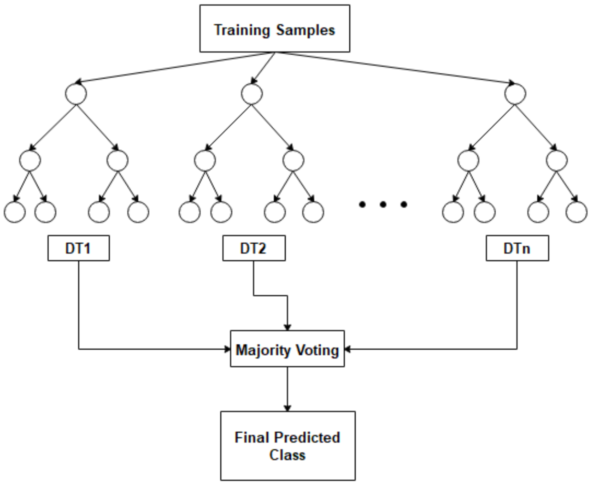

Fig. 1. Block Diagram of random Forest classifier

D. Multilayer Perceptron Model

Multilayer Perceptron (MLP) neural network is capable of approximating any type of non-linear function into a higher degree of accuracy value [10]. It is also known as multilayer feed forward neural network. MLP classifier uses supervised learning known as backpropagation method and is used for training purpose of the classification model.

MLP consists of at least three layers as shown in Figure 2. The input layer of MLP represents the problem of input variables with one neuron for every input variable. The hidden layer traces the non-linear relationships among all the variables. The last or output layer provides final predictive class value. In a completely connected MLP, every node of the lower layer is linked to all nodes of the next layer through a set of connected weights. The weighted sum of all the outputs of neurons in the lower layer are calculated, then it is passed through a nonlinear function. During training, the input pattern is applied to MLP classification model and the output of nodes of all three layers are computed.

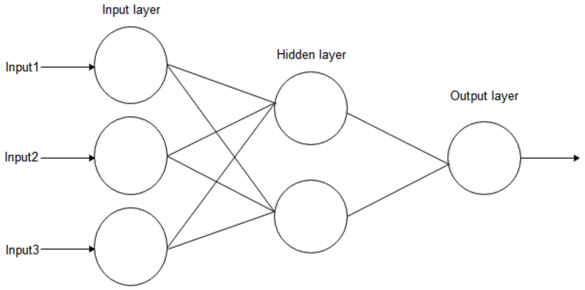

Fig. 2. Common architecture of an MLP model

E. Bagging Classifier

Bagging classifier is a popular ensemble technique, very much suitable for classification and regression purposes [13]. The design methodology is to improve both stability and accuracy in classification as well as regression. It operates by hybridizing classification models of the training sets till getting the final prediction. For variance reduction method we use it through randomization in its building procedure and then to generate an ensemble from it. Because of simple implementation and improved accuracy, it is used broadly. The basic principle of ensemble method is, a set of weak learners are iteratively added to build a strong learner. Every single decision tree is a weak learner, but when they ensemble, it becomes strong learner. Each tree votes for a class and the class having maximum votes gains the final predicted class.

Ensemble of Classification Algorithms

Ensemble classification method is a supervised learning technique that combines a number of weak learners iteratively to form an efficient learner that classifies the given training samples in a more accurate way [14] [18]. Weak learners are comparatively better than random guessing and when they are added iteratively one by one, their performance increases gradually. This method is able to create a better predictive model in which the classification accuracy is close to the correct value. An ensemble technique uses the same base classifier to create multiple possibilities.

It is observed from some empirical studies, that a specific algorithm can outperform all others for a unique set of problems, but a single classifier cannot achieve best accuracy for every situation. This leads to a growing research interest area to combine a set of multiple learning algorithms into a single system. The ensemble of learners increases the scalability to classify the normal and fraudulent class efficiently after the proper model training.

The classifiers predictions are combined to a meta-classifier to provide an effective performance with the help of majority voting technique [9]. It combines the nominal outputs to predict the class label for a set of possible class label. In this technique each weak learner votes to a specific class label and the class label, which receives more than half of the votes is the final class label. In the binary classification, the majority of the votes fetched by the classification model is considered as the better predictive model. With the help of majority voting, the hybridized classification model is developed by applying ensemble technique. By this operation, it achieves a very strong generalization ability.

IV. Proposed Model

In this study, five individual machine learning algorithms have been implemented and their accuracy and other performance parameters are investigated. We have proposed a predictive classification model by hybridizing the individual models, since the ensemble of machine learning algorithms help to improve the performances significantly [10] [14]. In the proposed model, five classification models Extreme Learning machine, K-Nearest Neighbor, Random Forest, Multilayer Perceptron and Bagging classifier are hybridized. The result is investigated and observed that the accuracy value of the proposed model is significantly high as compared to individual classification models.

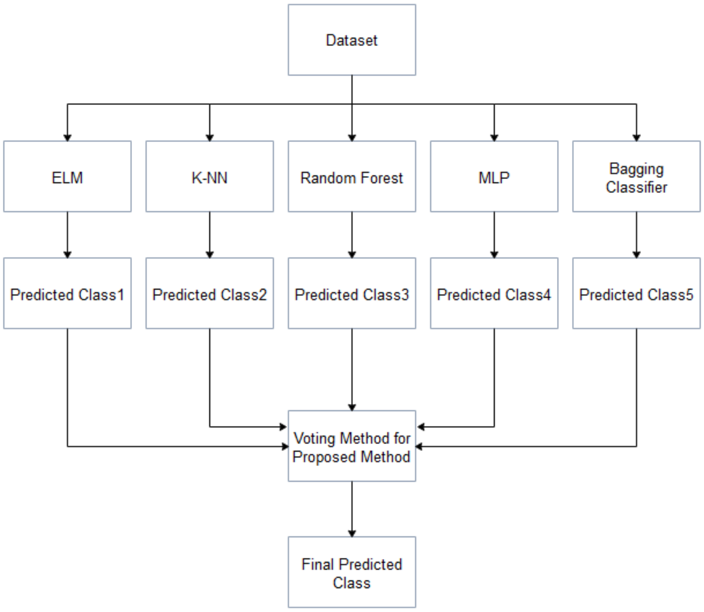

Fig. 3. Block diagram for the proposed classification model

As compared to a single classification model, the predictive result for the hybridized classification model is much more robust and better accuracy value can be achieved. A hybridized model reduces the variance and bias with increasing the prediction accuracy. In Figure 3, five models are trained with the dataset and their predicted classes are combined with majority voting method to obtain final predicted class. To design the combined model we have chosen five standard classification models from heterogeneous family such as ELM, K-NN, RF, MLP and Bagging classifier. In these classification models, their individual predictive accuracy is observed to be good and when they are combined, the resultant accuracy and other performance are observed to be better than individual models. With the stacking of classifiers, the decision boundary is optimized and error is reduced. Ensemble of machine learning algorithms with majority voting yields a better hybridized model that can correctly classify the fraudulent and non-fraudulent transactions.

V. Result and Performance Analysis

In order to test and compare the performances we have implemented the classification models such as ELM, K-NN, RF, MLP and Bagging classification models. Later, these five classification models are hybridized and then its performance was critically examined.

A. Experimental setup

Five machine learning classification algorithms have been implemented in Jupyter Notebook version 5.5.0 using Python 3.6 version. The system configuration is of i7 processor with 3.4 GHz clock speed. The secondary memory space and main memory space in each system is 1TB and 4GB respectively.

B. Dataset used for the Experiment

The optimized use of dataset is an important requirement for performing the classification technique. The dimension of the dataset can affect both training and testing of a model. Large volume of default credit card fraud classification data have been used for our proposed classification model is given (https://archive.ics.uci.edu/ml/machine-learning-databases/00350/). It has a total 690000 data with a dimension of 23 columns and 30000 rows. From the dataset 80% data samples are used for training and 20% for testing. The accuracy percentage is being optimized when 80% training data and 20% testing data are used.

C. Performance Parameters

Various performance parameters such as accuracy, sensitivity, specificity, precision, F1-score, Matthews correlation coefficient (MCC) are evaluated using confusion matrix and their importance for the model development is explained. The parameter values are compared with every individual models and predictive accuracy is calculated for the proposed classification model.

Confusion Matrix

A confusion Matrix is a representation technique to implement classification models. It shows correctly and incorrectly classified samples with actual result in the test data. Its framework for the two class classification is a 2 X 2 table designed to count the quantity of all four results of a binary classifier and denoted as TP, FP, TN and FN.

- True positive (TP): Total number of fraud transactions predicted as fraud
- False positive (FP): Total number of legal transactions predicted as fraud
- True negative (TN): Total number of legal transactions predicted as legal
- False negative (FN): Total number of fraud transactions predicted as legal

Accuracy

It is calculated to find out upto which extent the classification model can predict the samples correctly. We calculate it by dividing total number of correct predictions with the total samples. The best accuracy is assumed 1.0, whereas 0.0 is the worst. We have calculated the predictive accuracy for all the five classifiers by using the confusion matrix.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

We can also calculate by using

$$1 - ErrorRate \quad (2)$$

TABLE I
CONFUSION MATRIX FOR THE PROPOSED MODEL

	Actual class	
Predicted class	True Positive 4519	False positive 263
	False Negative 707	True Negative 511

TABLE II
PERFORMANCE RESULTS FOR VARIOUS CLASSIFIERS

Classifiers Parameters	ELM	MLP	Bagging	K-NN	Random Forest	Proposed Model
Accuracy	78.75	80.38	80.87	81.43	81.92	83.83
Sensitivity	87.26	88.38	88.55	89.16	89.32	86.47
Specificity	47.00	44.98	44.46	43.96	43.62	66.02
Precision	86.00	87.68	88.31	88.52	89.12	94.50
F1-Score	86.63	88.03	88.43	88.84	89.22	90.31
MCC	34.98	33.82	33.19	33.60	33.09	43.74

Sensitivity

It is calculated to identify the true positive rate for the error estimation. It indicates how good the test is to identify the fraud rate. By using confusion matrix, we have calculated sensitivity. It is also known as true positive rate (TPR) or recall. Sensitivity is best when the value is 1.0 and worst when it is 0.0.

$$Sensitivity = TP/(TP + FN) \tag{3}$$

Specificity

It is evaluated to find how efficiently and accurately it recognizes the false alarm rate. It is also known as true negative rate or TNR. Specificity is best when the value is 1.0 and worst when it is 0.0.

$$Specificity = TN/(TN + FP) \tag{4}$$

Precision

It is evaluated to identify the total pertinent positively classified instances from the retrieved instances. It is also known as positive predicted value (PPV). Precision is best when the value is 1.0 and worst when it is 0.0.

$$Precision = TP/(TP + FP) \tag{5}$$

F1-Score

F1-Score or F1-Measure is calculated to find the testing accuracy of the fraud detection model. To calculate the score it takes the harmonic mean of both recall and precision of the test into consideration.

$$F1 - score = 2TP/(2TP + FP + FN) \tag{6}$$

Matthews correlation coefficient (MCC)

It is applied as a measure of quality parameter in the binary classification. It is used as a balanced measure when the classes are of different sizes. It returns the value between -1 and +1.

MCC =

$$\frac{(TP * TN - FP * FN)}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \tag{7}$$

The parameters of the confusion matrix for our proposed model to classify the samples as true positive class, true negative class, false positive class and false negative class are shown in Table I. The objective is to maximize the correct prediction and to minimize the false alarms.

In Table II, the performance parameters have been presented. They are evaluated by considering 20% of testing data. The performance parameters are evaluated for accuracy, sensitivity, specificity, Precision, F1-Score and Matthews Correlation Coefficient. The prediction accuracy for the proposed model is observed to be 83.83%, which is significantly improved as compared to other single classification models. The fraud prediction error has been reduced with the margin of separation between the optimal boundary values. By using this proposed method, higher value of prediction accuracy has been achieved and false alarm rate is reduced.

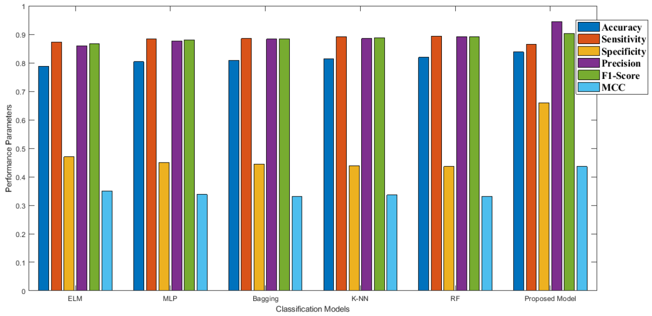

Fig. 4. Performance Parameters of various classification models

In Figure 4, the performance parameters such as accuracy, sensitivity, specificity, Precision, F1-Score and Matthews Correlation Coefficient are shown with respect to the classification models. We observed that the accuracy value for the proposed model is highest as compared to the individual classification models.

VI. CONCLUSION

Five standard machine learning classification models are investigated and compared with each other. The performance evaluation for our proposed model has been developed by using 20% testing data from the datset. Ensemble of the machine learning algorithm is one of the novel approach for the credit card fraud detection technique. Although there is a marginal difference in accuracy among the individual models, but the predictive accuracy percentage for the proposed classification model is observed to be 83.83%, which is significantly improved. The fraud detection error for the proposed model is reduced and the fraud prediction rate is improved.

The limitation to the fraud detection is subject to the availability of real-time dataset. If at all the sensitive dataset of various fraudulent activities are made available from the financial institutions, the research outcome will be more efficient and qualitative.

References

[1] Ngai, Eric WT, Yong Hu, Yiu Hing Wong, Yijun Chen, and Xin Sun. ”The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature.” *Decision support systems*, vol. 50, no. 3, pp. 559-569, 2011.

[2] Bolton, Richard J., and David J. Hand. ”Unsupervised profiling methods for fraud detection.” *Credit Scoring and Credit Control VII*, pp. 235-255, 2001.

[3] Whitrow, Christopher, David J. Hand, Piotr Juszczak, D. Weston, and Niall M. Adams. ”Transaction aggregation as a strategy for credit card fraud detection.” *Data mining and knowledge discovery*, vol. 18, no. 1, pp. 30-55, 2009.

[4] Maes, Sam, Karl Tuyls, Bram Vanschoenwinkel, and Bernard Manderick. ”Credit card fraud detection using Bayesian and neural networks.” *In Proceedings of the 1st international naiso congress on neuro fuzzy technologies*, pp. 261-270. 2002.

[5] Bolton, Richard J., and David J. Hand. ”Statistical fraud detection: A review.” *Statistical science*, pp. 235-249,2002.

[6] Zareapoor, Masoumeh, K. R. Seeja, and M. Afshar Alam. ”Analysis on credit card fraud detection techniques: based on certain design criteria.” *International journal of computer applications*, vol. 52, no. 3, 2012.

[7] Srivastava, Abhinav, Amlan Kundu, Shamik Sural, and Arun Majumdar. ”Credit card fraud detection using hidden Markov model.” *IEEE Transactions on dependable and secure computing*, vol. 5, no. 1 pp. 37-48, 2008.

[8] Mishra, Nishchol, and Sanjay Silakari. ”Predictive analytics: A survey, trends, applications, oppurtunities & challenges.” *International Journal of Computer Science and Information Technologies*, vol. 3, no. 3, pp. 4434-4438, 2012.

[9] Chan, Philip K., and Salvatore J. Stolfo. ”Toward Scalable Learning with Non-Uniform Class and Cost Distributions: A Case Study in Credit Card Fraud Detection.” *In KDD*, vol. 1998, pp. 164-168, 1998.

[10] Mishra, Mukesh Kumar, and Rajashree Dash. ”A comparative study of chebyshev functional link artificial neural network, multi-layer perceptron and decision tree for credit card fraud detection.” *In 2014 International Conference on Information Technology*, pp. 228-233, 2014.

[11] Akinyelu, Andronicus A., and Aderemi O. Adewumi. ”Classification of phishing email using random forest machine learning technique.” *Journal of Applied Mathematics*, vol. 2014, 2014.

[12] Bhattacharyya, Siddhartha, Sanjeev Jha, Kurian Tharakunnel, and J. Christopher Westland. ”Data mining for credit card fraud: A comparative study.” *Decision Support Systems*, vol. 50, no. 3, pp. 602-613, 2011.

[13] Zareapoor, Masoumeh, and Pourya Shamsolmoali. ”Application of credit card fraud detection: Based on bagging ensemble classifier.” *Procedia computer science*, vol. 48, no. 2015 pp. 679-685, 2015.

[14] Dietterich, Thomas G. ”Ensemble methods in machine learning.” In International workshop on multiple classifier systems, pp. 1-15. Springer, Berlin, Heidelberg, 2000. *Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 255-261. IEEE, 2015.

[15] Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. ”Extreme learning machine: a new learning scheme of feedforward neural networks.” *Neural networks*, vol. 2, pp. 985-990, 2004.

[16] http://www.bbc.com/capital/story/20170711-credit-card-fraud-what-you-need-to-know

[17] Kou, Yufeng, Chang-Tien Lu, Sirirat Sirwongwattana, and Yo-Ping Huang. ”Survey of fraud detection techniques.” *In IEEE International Conference on Net-working, Sensing and Control*, vol. 2, pp. 749-754, IEEE, 2004.

[18] Adewumi, Aderemi O., and Andronicus A. Akinyelu. ”A survey of machine-learning and nature-inspired based credit card fraud detection techniques.” *International Journal of System Assurance Engineering and Management*, vol. 8, no. 2, pp.937-953, 2017.

[19] Carneiro, Nuno, Goncalo Figueira, and Miguel Costa. ”A data mining based system for credit-card fraud detection in e-tail.” *Decision Support Systems*, vol. 95, pp.91-101, 2017.

[20] Kiran, Sai, N. Kumar, J. Guru, D. Katariya, R. Kumar, and Maheshwar Sharma. ”Credit card fraud detection using Nave Bayes model based and KNN classifier.” *International Journal of Advance Research, Ideas and Innovations in Technoloy*, vol. 4, no. 3, 2018.