ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 2, pp: (1-13), Month: October 2024 – March 2025, Available at: www.paperpublications.org

Evaluation of Selected Machine Learning Techniques in Feature Extraction based Fraud Detection System in Online Transactions

[1]Ademola Sefiu A., [2]Ismaila W. Oladimeji, [3]Omotosho I. O., [4]Ismaila Folasade M.

[1,2]Department of Computer Science, Ladoke Akintola University of Technology, Nigeria

[3]Department of Cyber Security Science, Ladoke Akintola University of Technology, Nigeria

[4]Department of Computer Science, Osun State Polytechnic, Iree, Nigeria

DOI: https://doi.org/10.5281/zenodo.13897705

Published Date: 07-October-2024

Abstract: **The recent advances of e-commerce and e-payment systems have sparked an increment in financial fraud cases such as credit card fraud. Several classification techniques have been employed to detect credit card frauds in online transactions but their performances were affected by high cardholder's data dimensionality. Thus, work employed Ant Colony Optimization for features extraction and evaluate its effectiveness using three selected classifiers. to detect fraud in credit cards online transactions. 3200 cardholders data (real and simulated) dataset with mix of genuine and fraudulent transactions. Ants Colony Optimization technique was used to extract features from the transactional data. Then, fraud detection system was designed with the three selected machine learning techniques (Back Propagation Neural Network, BPNN, Support Vector Machine, SVM and Naïve Bayes, NB) for classification. The results revealed that without features selection technique, NB, BPNN and SVM produced 86.4%, 88.7%, 93.6%, for accuracy respectively and while with ACO technique, the results or NB, BPNN and SVM produced 95.3%, 96.8%, and 97.6%.**

Keywords: **Fraud Detection System, Back Propagation Neural Network, Support Vector Machine, Naïve Bayes, Ants Colony Optimization.**

I. INTRODUCTION

Increasing dependence on technology in e-Commerce and the m-Commerce domains has open room for serious threats in terms of security and privacy at alarming rate. The security threats are noticeable in banks transactions, internet shopping, insurance, etc. Hence, the proliferation of internet technology nowadays has led to the increasing in number of Online Identity Thefts, Credit Card Fraud, Insurance Fraud, Banking Fraud, and Money Laundering for illegal activities etc. [43], [28] [31]. The different types of methods for committing credit card frauds are application fraud (assumed identity, Financial fraud, Not-received items (NRIs); lost or stolen cards; account takeover; fake and counterfeit cards [4].

Machine Learning is a subset of Artificial Intelligence (AI) that enables computers to learn from data and make decisions or predictions without being explicitly programmed to do so. Several of Machine Learning techniques have been deployed by the researchers to curb the proliferation of fraudulent activities on online fraudsters. [40] [6]. These techniques include Naïve Bayes (NB) [7, 21 20], K–nearest neighbours (KNN) [26], Support Vector Machine (SVM), [54, 62, 53, 36; 25], Back Propagation Neural Network (BPNN) [5, 63, 1,54], Random Forest (RF) [66, 63, 47; 44, 41; 37,30], Local Outlier Factor and Isolation Forest [67, 42, 57, 51, **27**], Logistic Regression,(LR) by [15, 8, 22, 2], XGBoost [65], Decision Tree (DT) [33], Self-Organized Map (SOM) by [23], Hidden Markov model (HMM) [25, 58, 59; 38, 17], Deep learning [60, 41,49], Convolutional Neural Networks [19], Fisher discriminant analysis [45], K-means [17], generative adversarial networks [10]; Graph Neural Network [11] multi-layer neural networks [18, 16, 4], C4.5 decision tree [33],; fuzzy rough

nearest neighbor [22, Bayesian belief network [24], AdaBoost [46], Deep Learning Neural Network [33, 51, 24; 12,68], boosted stacking [55] etc.

However, in order to produce an optimum results using the above techniques, several efforts have been made to compare the effectiveness of these algorithms. The machine learning techniques that were compared in the literature include NB, KNN, and LR by [56], LR, DT, SVM and RF by [13], DT, k-Nearest Neighbor, LR, RF and NB by [32] and authors in [69] used genetic algorithm (GA) for feature selection and compared DT, RF, LR, Artificial Neural Network (ANN), and NB. Also [60] employed GA for feature selection and compared GA-RF, GA-ANN and GA-DT, [60] evaluated the performances of random forest algorithm and decision trees. [62] worked on DT, RF, KNN, and Multilayer Perceptron Neural network (MLP).

In 2020 [48] highlighted that inefficiency of these classifiers are caused by imbalance dataset (problem of skewed distributions of legal and fraudulent transactions in available data); problem of noise (irrelevant features) and problem of overlapping data, all these limits the accuracy of detection systems. Thus, this work employed Ant Colony Optimization as feature selector of attributes cardholders and then evaluate the performance of selected machine learning techniques in Fraud Detection System in online transaction system. The selected machine learning techniques are Back-Propagation Neural networks (BPNN), Support Vector Machines (SVM) and Naïve Bayes (NB).

Ant Colony Optimization (ACO) is an algorithm inspired by the foraging behavior of ants and based on positive feedback mechanism. ACO has advantages of strong robustness, good global optimization ability and very useful in path planning and features extraction. BPNN is fast, simple and easy to program, has no parameters to tune apart from the numbers of input and is a flexible method as it does not require prior knowledge about the network. SVM model is a supervised machine-learning method that is used to perform classification and regression analysis. SVM has high accuracy, and ability to deal with high-dimensional data, ability to generate non-linear decision boundaries. NB helps to calculate relationship for each predictor variable [32].

II. RELATED WORKS

The authors in [34] presents an automated credit card fraud detection system based on the unsupervised neural network technology (Self-Organizing Map algorithm). The results were evaluated with performance metrics to determine its effectiveness. , It was deduced that the developed system produced false-positive rates that decrease as the rate of transactions intermix increase at malicious distribution of 0.5. Also, the developed fraud system produces 0.95 and 0.83 for precision and accuracy respectively at profile 95 3 2. The researches in [35] proposed a probabilistic based model to serve as a basis for mathematical derivation for adaptive threshold algorithm for detecting anomaly transactions. The model was optimized with Baum-Welsh and hybridized posterior-Viterbi algorithms. The model used 3200 simulated data for training and 800 for prediction. The results obtained from the evaluation showed the overall average of accuracy and precision are about 84% and 86% respectively. Also, The ROC curve revealed that this research falls in conservative performance region which implies that classifiers in this region commit few false positive errors.

[7] presented a comparison analysis of different ML methods on the European cardholders credit card fraud dataset. The authors used an hybrid sampling technique to deal with the imbalanced nature of the dataset. The NB, KNN, and LR techniques were implemented using a Python based ML framework. The experimental results demonstrated that the NB, LR, and KNN achieved the following accuracies, respectively: 97.92%, 54.86%, and 97.69%. [55]. implemented a credit card fraud detection system using LR, DT, SVM and RF. These classifiers were evaluated using a credit card fraud detection dataset generated from European cardholders charaterised with highly skewed and imbalanced dataset. The experimental outcomes showed that the LR, DT, SVM and RF obtained the following accuracy scores: 97.70%, 95.50%, 97.50% and 98.60%, respectively. In 2019, the authors in [54] developed a novel fraud detection system which contain two steps. In the first Step, sequences of transactions are added in to the system using HMM to categorize expenditure behavior of card holders as low, medium and high. In the second step, the fraud is detected in the credit card using the cluster comparison of the transactions. The proposed system has been developed for detection of credit card frauds which gave about 87% accuracy.

In 2020, the researchers in [63] proposed a machine learning based credit card fraud detection engine using the GA for feature selection. After the optimized features are chosen, the proposed detection engine uses the following ML

classifiers: DT, RF, LR, ANN, and NB. The proposed credit card fraud detection engine is evaluated using European cardholders and synthetic dataset were used. The proposed system gave average of 99.75% accuracy. [58] conducted a performance analysis of ML techniques for credit card fraud detection. In this research, the authors considered DT, KNN, LR, RF and NB to classify European cardholders dataset. The experimental outcomes showed that the DT, KNN, LR, RF and NB obtained precisions of 85.11%, 91.11%, 87.5%, 89.77%, 65.2%, respectively.

In 2021, Authors in [59] implemented an intelligent payment card fraud detection system using the GA for feature selection and aggregation. The authors implemented several machine learning algorithms to validate the effectiveness of their proposed method. The results demonstrated that the GA-RF obtained an accuracy of 77.95%, the GA-ANN achieved an accuracy of 81.82%, and the GA-DT attained an accuracy of 81.97%. [65] researchers designed a model to detect the fraud activity in credit card transactions with important features required to detect illegal and illicit transactions. The acquired credit card usage data-set were trained and tested using a RF and DT techniques. The results indicated concerning the best accuracy for RF is 98.6% respectively.

[56] authors in 2022 presented an innovative sensing method by employing SVM hyperparameter optimization using grid search cross-validation and separating the hyperplane using the theory of reproducing kernels. An online data science site called http://Kaggle.com provided the dataset used. Each data entry has 30 fields. Only 492 or 0.17 percent of the 284,807 credit card transactions in the database are fraudulent. The results showed that the developed system with SVM gave 98.36% accuracy. [3] researchers in 2024 introduced a state-of-the-art hybrid ensemble dependable Machine learning model that intelligently combines multiple algorithms with proper weighted optimization using Grid search, including DT, RF, KNN, and MLP, to enhance fraud identification. The Instant Hardness Threshold was used to address the data imbalance issue, technique in conjunction with LR, surpassing conventional approaches. They achieves impressive accuracy rates of 99.66%, 99.73%, 98.56%, and 99.79%, and a perfect 100% for the DT, RF, KNN, MLP and ENS models, respectively.

III. MATERIALS

This section discusses the feature selection technique, the selected machine learning classifiers and the performance metrics employed.

A. Ant Colony Algorithm

The authors in [38] stated that Ant Colony Optimization (ACO) is a metaheuristic technique based on the research of ant group behaviours in the natural world. It imitates ant's behaviours in establishing shortest paths from their nest to feeding sources and back. Individual ants are supposed to interact with each other by some chemical pheromone released by them. When an ant finds a food source, it releases chemical pheromone on the ground. The quantity of pheromone depends on the quantity and quality of the food source. The pheromone vanishes over time. When another ant looks for food, it moves in a random manner basically. However, if it detects chemical pheromone in the environment around, it will have higher probability to move toward the direction with denser pheromone. This in turns reinforces the trail with more pheromone, and attracts more ants to the food source. The indirect communication between the ants via the pheromone trail allows them to find shortest paths between their nest and food sources. The collective behaviour brings ACO with characteristics of positive feedback, parallel computing, robustness, and global optimization [38, 6].

Algorithm 1. Simple ACO Algorithm [67]

Initialize $\tau_{ij}(0)$ to small random values;
Place n_k ants on the origin node;
repeat
Let $t = 0$;
  for each ant $k = 1, \ldots, n_k$ **do**
  //Construct a path $x^k(t)$;
  $x^k(t) = \emptyset$;
  repeat
    Select next node based on the probability
    Add link (i, j) to path $x^k(t)$;
  until destination node has been reached;

Remove all loops from $x^k(t)$;
Calculate the path length $f(x^k(t))$;
end
for each link (i, j) of the graph **do**
//pheromone evaporation;
Reduce the pheromone, $\tau_{ij}(t)$,
end
for each ant k = 1, . . . ,n_k **do**
for each link (i, j) of $x^k(t)$ **do**
$\Delta\tau^k = \frac{1}{f(x^k(t))}$
Update τ^{ij} using equation (2.3);
end
end
$t = t + 1$;
until stopping condition is true;
Return the path $x^k(t)$with smallest $f(x^k(t))$ as the solution;

B. BN

The Bayesian network was first introduced by Cooper and Herskovits (1992). Bayesian belief networks are statistical techniques in data mining. Bayesian networks are very effective for modeling situations where some information is already known and incoming data is unsure or partially unavailable. The goal of using Baye rules is to correctly predict the value of a designated discrete class variable given a vector of predictors or attributes.

Also, Bayesian classifier is a statistical method calculating probability that feature belongs to class based on applying Bayes' theorem. Because it assumes that the probabilities of individual features are independent of each other which is quite hard to happen in real world it is reason to be called naive. Considering that another event has already occurred to calculate the likelihood that an event will occur. It can be written as:

$$\Pr(c / X) = \frac{\left(\Pr(c / X) * \Pr(c)\right)}{\left(\Pr(X)\right)} \quad (1)$$

Where posterior probability of target class c P(c|X) is calculated from P(c), P(X|c) and P(X) [33].

C. SVM

The SVM is statistical learning techniques and has successful application in a range of problems including classification tasks. They are closely related to neural networks and through the use of kernel functions, they can be considered an alternative way to obtain neural network classifiers. SVM algorithm is a supervised machine learning algorithm that has been applied to anomaly detection in the one-class setting. The basic idea of SVM classification algorithm is to construct a hyper plane as the decision plane which making the distance between the positive and negative mode maximum.

This model has been demonstrated that it possess a higher accuracy of detection compared with other algorithms. It also has a better time efficiency and generalization ability. The primal formulation of SVM is shown as following equation (2)

$$\min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}\xi_i \quad (2)$$

Subject to $y_i\left(\omega^T.\varphi(x_i) + b\right) \geq 1 - \xi_i \quad \xi_i \geq 0, i = 1, 2, ..., n$

There are mainly four kinds of hyperparameters for SVM to decide, which regularization factor C, kernel function type, kernel efficient γ and degree of kernel functions. In particular, the degree is for polynomial kernel functions. Therefore, with different kernel functions, different hyper-parameters need to be tuned to find the best model [29].

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 2, pp: (1-13), Month: October 2024 – March 2025, Available at: www.paperpublications.org

D. BPNN

Backpropagation neural network (as shown in figure 1) **i**s the essence of neural network training. It is the method of fine-tuning the weights of a neural network based on the error rate obtained in the previous epoch (i.e., iteration). Proper tuning of the weights allows you to reduce error rates and make the model reliable by increasing its generalization. Backpropagation in neural network is a short form for "backward propagation of errors." It is a standard method of training artificial neural networks. This method helps calculate the gradient of a loss function with respect to all the weights in the network.

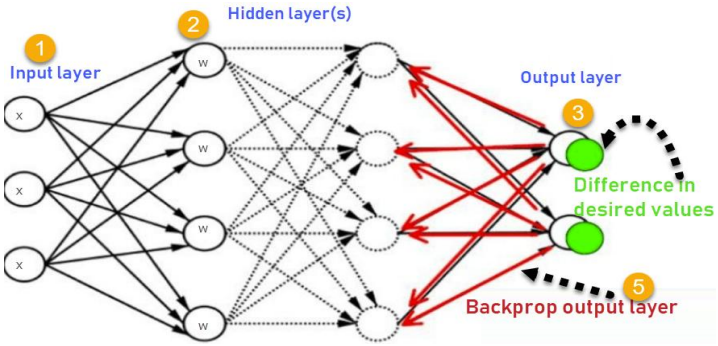

Figure 1: The architecture of the ANN-BP [14]

E. Performance Metrics

The performance metrics employed are Accuracy (ACC) and False positive rate (FPR). The mathematical formulation of these indicators is as follows:

$$ACC = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (3) \qquad FPR = \frac{FP}{(FP + TN)} \quad (4)$$

where False positive – FP; False Negative- FN: True positive - TP: True Negative -TN

IV. METHODOLOGY

This work delves into the captivating world of fraud detection, utilizing the power of machine learning to unveil hidden patterns and safeguard individuals and institutions from financial deception. The developed fraud system comprises of data acquisition, data pre-processing, feature extraction/selection, data classification and evaluation. The work flow diagram of the designed fraud detection system is shown in figure 2.

A. Data Acquisition

Real and simulated dataset with genuine and fraudulent were acquired for the experiment. Three thousand and two hundred transactional cardholders data (real and simulated) dataset with mix of genuine and fraudulent transactions. Ten real cardholders with twenty transactions each and one hundred and fifty simulated cardholders with twenty transactions each.

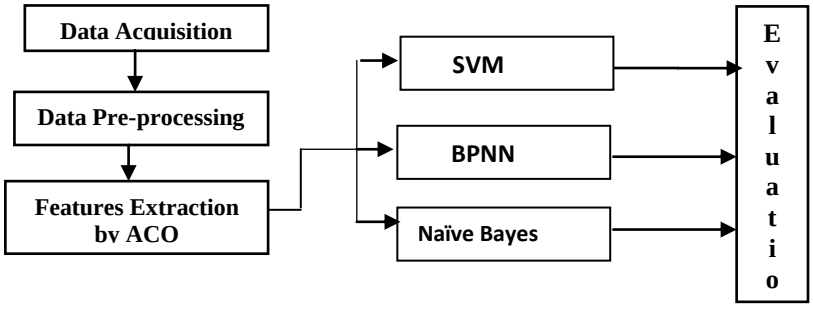

Figure 2. Block Diagram of developed Fraud Detection System

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 2, pp: (1-13), Month: October 2024 – March 2025, Available at: www.paperpublications.org

B. Data Preprocessing

Pre-processing corresponds to the modifications made to the dataset before feeding the algorithm. Several algorithms of machine learning make assumptions about data. It is often a very good idea to plan the data in such a way that the problem structure is better presented to the machine learning algorithms. Data pre-processing is a way of transforming original data into a clean dataset.

C. Credit Card Features Extraction by ACO

Feature extraction involves removing redundant or unnecessary data from the input dataset to facilitate their analysis using data mining techniques. The essential attributes like Card ID, Card Holder Name, Location, Date and amount was extracted by ACO from about 25 attributes of the dataset. The ACO algorithm is shown in Algorithm 1.

Algorithm 1 Simple ACO Algorithm

```
Function ACOFeatureSelection(Transactions, Features):
  // Initialization
 Set ACO parameters [N, F, T, P, H, F, τ_min, τ_max]
   // ACO Algorithm
 For each iteration:
   For each ant:
    f = randomFeature()
    FI = computeFeatureImportance(f, Transactions)
    updatePheromoneTrail(f, FI, P)
   updatePheromoneMatrix(P)
   updateHeuristicInformationMatrix(H)
   // Feature Selection
 For each feature:
   FSP = computeFeatureSelectionProbability(f, P, H)
   If FSP > θ:
    SelectedFeatures.append(f)
   Return SelectedFeatures
```

D. Data Classification

This is the last stage of the fraud detection process. The features extracted (ACO variables) were fed into the SVM, BPNN and NB classifier and the transactions were trained and classified to the corresponding transactions (legal or illegal).

i. SVM

SVM model intricate decision boundaries between fraudulent and non-fraudulent transactions. SVM incorporates a regularization parameter that helps prevent overfitting, ensuring better generalization to unseen data. The primal formulation of SVM is shown as following equation (5)

$$\min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}\xi_i \qquad (5)$$

Subject to $$y_i\left(\omega^T.\varphi(x_i)+b\right)\geq 1-\xi_i \qquad \xi_i \geq 0, i=1,2,...,n$$

This formulation is called the soft-margin SVM. In essence, SVM picks a hyperplane to separate the data points from two classes with the largest margin, which means the smallest distance from all training points to the hyperplane. In this work, the kernel function for the validation data is Radial basis function kernel. The SVM algorithm is shown in Algorithm 2.

Algorithm 2: Pseudocode of SVM

1. Collect and preprocess dataset:
   - Features (F): {transaction_amount, location, time, merchant_category, etc.}
   - Target variable (T): {fraudulent (1) or legitimate (0)}
2. Split dataset into training (60%) and testing sets (40%)
3. Choose SVM kernel (e.g., linear, polynomial, radial basis function (RBF))

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 2, pp: (1-13), Month: October 2024 – March 2025, Available at: www.paperpublications.org

4. Train SVM model:
   - Initialize weights (w) and bias (b)
   - Minimize loss function using optimization algorithm (e.g., gradient descent)

Classification Phase

1. Input new transaction data (D)
2. Compute decision value using trained SVM model:
   - decision_value = w^T * D + b
3. Determine class label:
   - If decision_value >= 0, classify as legitimate
   - Else, classify as fraudulent

ii. BPNN

The BPNN model consists of an Input layer Hidden layers and an Output layer. The BPNN pseudocode is shown in Algorithm 3.

Algorithm 3. BPNN Pseudo Code

Assign all network inputs and output Initialize all weights with small random numbers
. Load cardholder dataset (features and target variable)
repeat
for every pattern in the training set
Present the pattern to the network
// Propagated the input forward through the network:
for each layer in the network
for every node in the layer
1. Calculate the weight sum of the inputs to the node
2. Add the threshold to the sum
3. Calculate the activation for the node
End
End
// Propagate the errors backward through the network for every node in the output layer calculate the error signal end
for all hidden layers for every node in the layer
1. Calculate the node's signal error
2. Update each node's weight in the network
End
End
// Calculate the Error Function End while ((maximum number of iterations < = specified)

iii. NB

The goal of Naïve Bayes is to correctly predict the value of a designated discrete class variable given a vector of predictors or attributes. For the purpose of fraud detection, two Bayesian networks are constructed that describe the behavior of user. First, a Bayesian network is constructed to model behavior under the assumption that the user is fraudulent (F) and another model under the assumption that the user is a non- fraudulent (NF).

Considering that another event has already occurred to calculate the likelihood that an event will occur [17]. It can be written as:

$$\Pr(c / X) = \frac{\left(\Pr(c / X) * \Pr(c)\right)}{\left(\Pr(X)\right)} \quad (3)$$

Where posterior probability of target class c P(c|X) is calculated from P(c), P(X|c) and P(X). Admel [33, 72]. The NB pseudocode is shown in Algorithm 4.

Algorithm 4: NB pseudocode

Step 1. Collect and preprocess dataset:
Features (F): {transaction_amount, location, time, merchant_category, etc.}
Target variable (T): {fraudulent (1) or legitimate (0)}

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 2, pp: (1-13), Month: October 2024 – March 2025, Available at: www.paperpublications.org

Step 2. Split dataset into training (80%) and testing sets (20%)
Step 3. Calculate prior probabilities:
P(Fraud) = number of fraudulent transactions / total transactions
P(Legitimate) = number of legitimate transactions / total transactions
Step 4. Calculate likelihood probabilities for each feature (F):
P(F|Fraud) = probability distribution of feature values for fraudulent transactions
P(F|Legitimate) = probability distribution of feature values for legitimate transactions
****Classification Phase****
Step 5. Input new transaction data (D)
Step 6. Calculate posterior probabilities using Bayes' theorem:
P(Fraud|D) = P(D|Fraud) * P(Fraud) / P(D)
P(Legitimate|D) = P(D|Legitimate) * P(Legitimate) / P(D)
Step 7. Compare posterior probabilities:
If P(Fraud|D) > P(Legitimate|D), classify as fraudulent
Else, classify as legitimate

V. SYSTEM IMPLEMENTATION

The system configuration include window10 operation system, MATLAB2016 version and SQL software. The developed system is illustrated by a Graphical User Interface shown in Figures 2. The dataset used in this study consisted of 3200 transactions data, with 1920 transactions for training and 1280 transactions for testing. The dataset were classified using the three aforementioned techniques at 0.2, 0.35, 0.5 and 0.85 thresholds, and their performances were measured using False Positive Rate (FPR), and accuracy.

VI. DISCUSSION OF RESULTS

The results, as shown in Table 1, gave the results of the classifiers at threshold of 0.85 with and without ACO. The results of FPR obtained by the techniques without ACO feature extractor, at threshold of 0.85 (which gave the lowest values), showed that NB, SVM and BPNN produced 18.9%, 19.5% and 23.3%

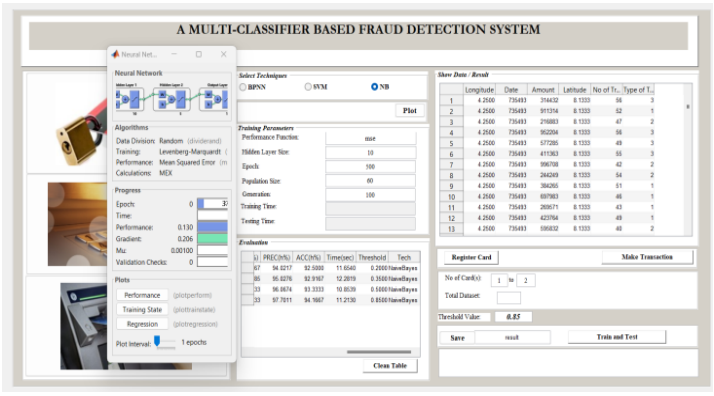

Figure 2: Graphical User Interface showing training and testing phase

respectively. While the results of FPR obtained by the techniques with ACO feature extractor showed that SVM, NB and BPNN achieved 9.8%, 10.9% and 12.3% respectively. The results of accuracy produced by the techniques without ACO feature extractor, at threshold of 0.85 (which gave the highest values), showed that SVM, BPNN and NB gave 93.6%, 88.7%, and 96.4% respectively. While the results of accuracy obtained by the techniques with ACO feature extractor at threshold of 0.85, showed that SVM BPNN and NB achieved 97.3%, 96.8%, and 95.3% respectively. The graphs of these results are shown in figure 3, 4, 5 and 6.

Table 1: Results of Classifiers with and without ACO at 0.85 threshold.

Metrics	Classifiers without ACO			Classifies with ACO		
	NB%	BPNN%	SVM%	NB%	BPNN%	SVM%
FPR	18.9	23.3	10.9	10.9	12.3	9.8
Accuracy	86.4	88.7	93.6	95.3	96.8	97.3

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 11, Issue 2, pp: (1-13), Month: October 2024 – March 2025, Available at: www.paperpublications.org

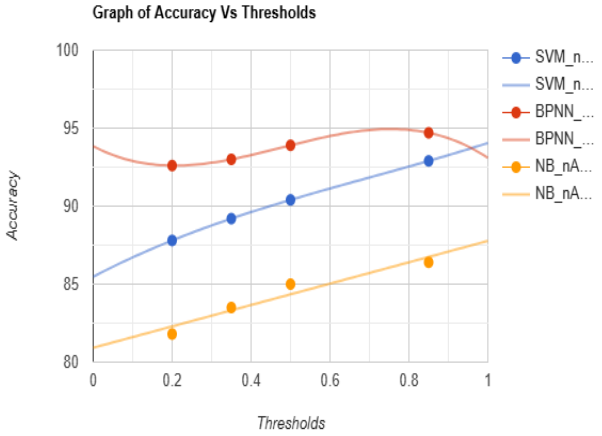

Figure 3: Graph of Accuracies across all classifiers without ACO

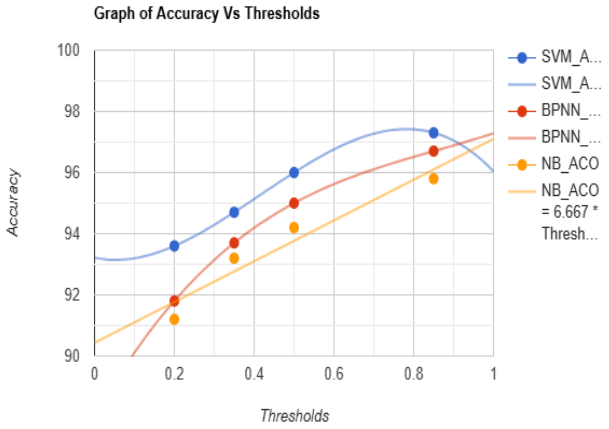

Figure 4: Graphs showing Accuracies across all classifiers with ACO

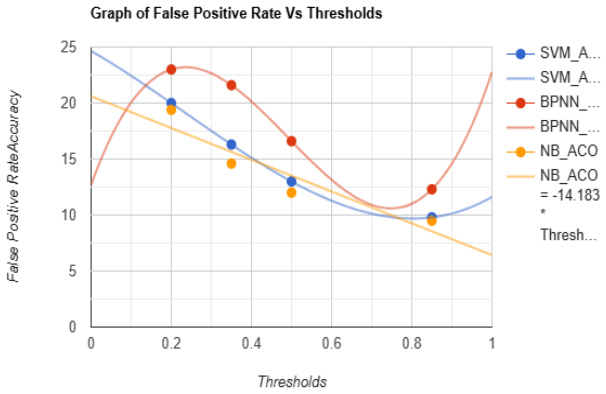

Figure 5: Graphs showing False Positive Rate across all classifiers without ACO

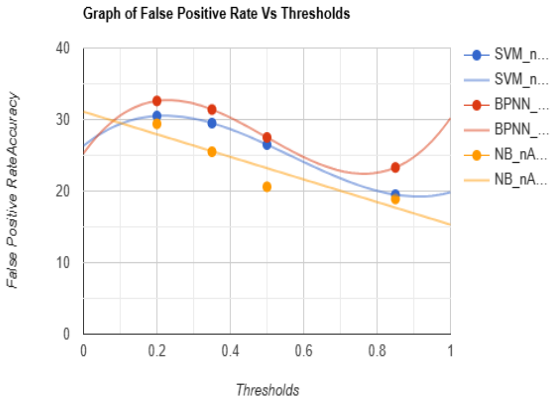

Figure 6: Graphs showing False Positive Rate across all classifiers with ACO

VII. CONCLUSION

This study was able to evaluate the effectiveness of extracting features by ACO and then applied SVM, BPNN and NB algorithms to detect fraudulent and non-fraudulent transactions in an online transactions. A real and simulated cardholder dataset consisting of 3200 transactions were classified by three classification techniques with different threshold values. Under the two conditions, SVM demonstrated optimum performance in terms of accuracy and false positive rate, making it the most effective classifier for detecting fraudulent transactions and its ability to handle low and high-dimensional data, robustness against overfitting, and capacity to capture non-linear relationships. A lower FPR indicates fewer false alarms, making SVM the most effective in minimizing non-fraudulent transactions misclassified as fraudulent. Naive Bayes classifier, while providing acceptable results, showed slightly lower performance compared to SVM but better than BPNN when features were extracted.. In conclusion, extracting features improve the abilities of classifiers viz; SVM, BPNN and NB to detect fraudulent and non-fraudulent transactions more accurately.

ACKNOWLEDGEMENT

I appreciate TETFUND for her financial support toward the successful completion of this research.

REFERENCES

[1] B. A. Abdulsalami, A. A Kolawole, M. A Ogunrinde, M. A. Lawal, R. A, Azeez, A. Z. Afolabi (2019). Comparative analysis of back-propagation neural network and K-means clustering algorithm in fraud detection in online credit card transactions. *Fountain Journal of Natural and Applied Sciences*; 8(1):315.

[2] H. Z Alenzi, N. O .Aljehane (2020). Fraud detection in credit cards using logistic regression. *International Journal of Advanced Computer Science and Applications.* 11(12):5570.

[3] Alamin, Rakib, Ashraf, Mohammed and Uzzal (2024). Securing Transactions: A Hybrid Dependable Ensemble Machine Learning Model using IHT-LR and Grid Search. *2024 in the Cybersecurity, Springer Open Journal*, 3-4.

[4] R. B. Asha (2021) Suresh Kumar K. R. (2021). Credit card fraud detection using artificial neural network. *Global Transitions Proceedings.*2(1):35–41.

[5] R. Amardeep and T. Swamy (2017): Training Feed forward Neural Network With Backpropogation Algorithm, *International Journal Of Engineering And Computer Science*. 6(1): 19860-19866.

[6] A. P. Engelbrecht(2007): Computation Intelligence, John Wiley & Sons Ltd, Southern Gate, Chichester, West Sussex PO19 8SQ, England.

[7] J. O, Awoyemi, A. O, Adetunmbi, S. A. Oluwadare (2017). Credit card fraud detection using machine learning techniques: a comparative analysis. In: International conference on computer networks and Information (ICCNI); 2017. 1-9.

[8] R. B. Asha, K. R. Suresh Kumar (2021). Credit card fraud detection using artificial neural network. *Global Transitions Proceedings.* 2(1):35–41.

[9] S. Esra and A.O Selma (2014). An Ant Colony Optimization Based Feature Selection for Web Page Classification.

[10] H. Ba (2019). Improving detection of credit card fraudulent transactions using generative adversarial networks. *ArXiv preprint.* 2019 doi: 10.48550/arXiv.1907.03355.

[11] D Cheng, X. Wang, Y. Zhang, L. Zhang. (2020). Graph neural network for fraud detection via spatial-temporal attention. *IEEE Transactions on Knowledge and Data Engineering.*34(8):3800–3813.

[12] Cheon M. J., Lee D. H., Joo H. S., Lee O. (2021). Deep learning based hybrid approach of detecting fraudulent transactions. *Journal of Theoretical and Applied Information Technology.* 2021;99(16):4044–4054.

[13] K. Campus (2018). Credit card fraud detection using machine learning models and collating machine learning models. Int J Pure Appl Math. 118(20):825–38.

[14] J. Daniel (2022): Back Propagation Neural Network: What is Backpropagation Algorithm in Machine Learning?

[15] I. Emmanuel. S. Yanxia and W. Zenghui (2020). A machine learning based credit card fraud detection using the GA algorithm for feature selection. *Journal of Big Data. 9:24. 1-17.*

[16] S. A. Ebiaredoh-Mienye, E, Esenogho, T. G. Swart. (2021). Artificial neural network technique for improving prediction of credit card default: a stacked sparse autoencoder approach. *International Journal of Electrical and Computer Engineering.* 11(5):4392.

[17] S. Esakkiraj and S. Chidambaram (2013). "A predictive approach for fraud detection using hidden markov model" International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 1.

[18] M. Faridpour, A. Moradi (2020). A novel method for detection of fraudulent bank transactions using multi-layer neural networks with adaptive learning rate. *International Journal of Nonlinear Analysis and Applications.*11(2):437–445.

[19] K. Fu, D. Cheng, Y. Tu, L. Zhang (2016) Credit card fraud detection using convolutional neural networks. In: International conference on neural information processing. Springer, Cham, 483–490.

[20] A. Gupta, M. C. Lohani, M. Manchanda M. (2021) Financial fraud detection using naive Bayes algorithm in highly imbalance data set. *Journal of Discrete Mathematical Sciences and Cryptography.*24(5):1559–1572.

[21] A. Husejinovic (2020) Credit card fraud detection using naive Bayesian and c4. 5 decision tree classifiers. *Periodicals of Engineering and Natural Sciences.*(4),1–5.

[22] A. S. Hussein, R. S. Khairy, S. M. Najeeb, Al-Rikabi H. T.. (2021). Credit card fraud detection using fuzzy rough nearest neighbor and sequential minimal optimization with logistic regression. *International Journal of Interactive Mobile Technologies.* 15(5):24–42.

[23] A. Husejinović (2020). Credit card fraud detection using naive Bayesian and C4.5 decision tree classifiers. *http://pen.ius.edu.ba*

[24] W.O. Ismaila, B. K. Alese, O. O. Adeosun., O. T. Arulogun (2012). *Performance Evaluation of Unsupervised Neural Network in Fraud Detection.* Journal of Physical Sciences and Innovations. Nigeria, 1(4), 54-62.

[25] M. Habibpour, H. Gharoun, Mehdipour M, Tajally A, Asgharnezhad H, Shamsi A, Khosravi A, Shafie-Khah M, Nahavandi S, Catalao J. P. (2021). Uncertainty-aware credit card fraud detection using deep learning. *ArXiv preprint.* 2021.

[26] W. O. Ismaila, S. O. Falaki, B. K Alese, Adewale O. S., Ayeni J. O., Aderounmu G. A. (2013): Probabilistic Credit Card Fraud Detection System In Online Transactions. *International Journal of Emerging Technologies in Computational And Applied Science*. 6(4), 586-596. India.

[27] F. Itoo, Singh S. (2021) Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. *International Journal of Information Technology.* 2021;13(4):1503–1511.

[28] S. Jaiswal, R. Brindha, Lakhotia S. (2021). Credit card fraud detection using isolation forest and local outlier factor. *Annals of the Romanian Society for Cell Biology.* 2021; 25(5):4391–4396.

[29] C. Jiang, D. Broby (2021). Mitigating cybersecurity challenges in the financial sector with artificial intelligence.

[30] Jianglin Xia (2022). Credit Card Fraud Detection Based on Support Vector Machine. *https://www.researchgate.net /publication/366261463*

[31] V. Jonnalagadda, P. Gupta, Sen E. (2019). Credit card fraud detection using random forest algorithm. *International Journal of Advance Research, Ideas and Innovations in Technology*;5(2):1–5.

[32] A. B. Kamalu (2022). Fraud Prevention And Detection System In Nigeria Banking Industries, Computer Science & It Research Journal 3(2):52-65, 2022.

[33] S. Khatri, A. Arora, A. P. Agrawal (2020). Supervised machine learning algorithms for credit card fraud detection: a comparison. In: 10th international conference on cloud computing, data science & engineering (Confluence); 680-683.

[34] E. Kim, J. Lee, H. Shin, Yang H, Cho S, Nam S. K., Song Y, Yoon J. A., Kim J. I.. (2019). Champion-challenger analysis for credit card fraud detection: hybrid ensemble and deep learning. *Expert Systems with Applications.*, 128(3):214–224.

[35] M. D. Kumar, A. Mubarak, M. S. Dhanush. (2020). Credit card fraud detection using Bayesian belief network. *International Journal of Research in Engineering, Science and Management.* 3(7):316–319.

[36] Y. Kumar, S. Saini, R. Payal. (2020). Comparative analysis for fraud detection using logistic regression, random Forest and support vector machine. *SSRN Electronic Journal.* 2020(18), 2020.

[37] C. Li, N. Ding, Y. Zhai, Dong H. (2021). Comparative study on credit card fraud detection based on different support vector machines. *Intelligent Data Analysis*; 25(1):105–119.

[38] T. H. Lin, J. R. Jiang. (2021). Credit card fraud detection with autoencoder and probabilistic random forest. *Mathematics.* 9(21):2683.

[39] Liu O., Ma J., Pak-Lok P., and Zhang J. (2009): On an Ant Colony-Based Approach for Business Fraud Detection. *http://www.researchgate.net/publication/220777993*

[40] Makolo A, Adeboye T. Credit card fraud detection system using machine learning. *International Journal of Information Technology and Computer Science.* 2021(4):24–37.

[41] M.K.R Mallidi, Y. Zagabathuni (2021). Analysis of credit card fraud detection using machine learning models on balanced and imbalanced datasets. *International Journal of Emerging Trends in Engineering Research.* 9(7):2972021.

[42] B. D. Meenakshi, B. Janani., S. Gayathri., Indira N.. (2019). Credit card fraud detection using random forest. *International Research Journal of Engineering and Technology (IRJET)* 2019;6(3):2019.

[43] M. Meenu, S. Gupta, Patel S, Kumar S, Chauhan G. (2020). Anomaly detection in credit card transactions using machine learning. *International Journal of Innovative Research in Computer Science & Technology (IJIRCST)* ;8(3):2020.

[44] S. Morgan.. Global cybersecurity spending predicted to exceed $1 trillion from 2017–2021. 2019. [1 June 2021].

[45] S. Monika, K. Venkataramanamma, P.P Paul, Usha M. (2019). Credit card fraud detection using random forest algorithm. *International Journal of Research in Engineering, Science and Management.*2(3):2019.

[46] N. Mahmoudi, E. Duman (2015) Detecting credit card fraud by modified Fisher discriminant analysis. Expert Syst Appl 42(5):2510–2516.

[47] B. Nithin, R. Ravula, Sulthana S. G. (2020).. Credit card fraud detection using AdaBoost. *International Journal of Scientific Research & Engineering Trends.* 6(2).

[48] G. Niveditha, K. Abarna, G. V. Akshaya (2019). Credit card fraud detection using random forest algorithm. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology.* 2019; 5(2):301–306.

[49] O. C. Okeke, N. C. Ezenwegbu (2020). A Machine Learning Algorithm for ATM Card Fraud Detection, COOU Journal of Physical CJPS Sciences 3(1), 2020.

[50] A. A Ojugo, O. Nwankwo (2021). Spectral-cluster solution for credit-card fraud detection using a genetic algorithm trained modular deep learning neural network. *JINAV: Journal of Information and Visualization.* 2(1):15–24.

[51] B. G. Pratap, P. Vijayaraghavulu (2021). A hybrid method for credit card fraud detection using machine learning algorithm. *International Journal of Computers, Electrical and Advanced Communication Engineering (IJCEACE)* 2021 boosted stacking;10(19):46–50.

[52] D. Phua,. Alahakoon and V Lee, (2005). "Minority report in fraud detection: classification of skewed data," ACM SIGKDD Explorations Newsletter, 6(1), 50-59.

[53] V. Palekar, S. Kharade, H. Zade, S. Ali, K. Kamble, Ambatkar S. (2020). Credit card fraud detection using isolation forest. *International Research Journal of Engineering and Technology (IRJET)*;7(3):1–6.

[54] B. G. Pratap, P. Vijayaraghavulu (2021). A hybrid method for credit card fraud detection using machine learning algorithm. *International Journal of Computers, Electrical and Advanced Communication Engineering (IJCEACE)* 2021 boosted stacking;10(19):46–50.

[55] N. Rtayli, N. Enneya. (2020) Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization. *Journal of Information Security and Applications.*;55(3):102596.

[56] Y. Saad,, and K. Shaker, (2017). Support vector machine and Back propagation neural network approach for text classification. *Journal of University of Human Development*, *3*(2), 869-876.

[57] M. Sathyapriya., V. Thiagarasu (2019). A Cluster Based Approach for Credit Card Fraud Detection System using Hmm with the Implementation of Big Data Technology, International Journal of Applied Engineering Research, 14(2) 393-396

[58] G. Sasikala., M. Laavanya., B. Sathyasri., C. Supraja., Mahalakshmi V., Mole S. S., Jaison M., S. Chidambaranathan, C. Arvind, K. Srihari, and Minilu D. (2022). An Innovative Sensing Machine Learning Technique to Detect Credit Card Frauds in Wireless Communications. vol. 2022, Article ID 2439205, 1-12.

[59] G. K. Singh, A. Bhayye, S. Dhamnaskar, S. Patil, Phulari S. V.. (2021). Credit card fraud detection using isolation forest. *International Journal of Recent Advances in Multidisciplinary Topics.* 2(6):118–119.

[60] G. K Sandhu,., and Kaur, R. (2019). Plant disease detection techniques: A Review. *In 2019 International Conference on Automation, Computational and Technology Management (ICACTM)*, 34-38.

[61] M. Singh, S.. Kumar, , Garg T. (2019). Credit card fraud detection using hidden Markov model. *International Journal of Engineering and Computer Science.* 8(11):24878–24882.

[62] A. Srivastava., A. Kundu, S. Sural, and A. K. Majumdar, (2008). "Credit card fraud detection using hidden markov model", IEEE transactions on dependable and secure computing, 5(1), 2008.

[63] Sobana Devi V, Dr, Ravi G. (2020). Real time deep learning based credit card fraud detection. *International Journal of Scientific & Technology Research.* 2020;9(3):2020.

[64] M. J. Simi (2019) Credit card fraud detection: a comparison using random forest, SVM and ANN. *International Research Journal of Engineering and Technology.* 6(3): 225–228.

[65] S. L. Sarvani and M. Markandeyulu (2021). Artificial intelligence framework for credit card fraud detection using supervised random forest. *Open Access International Journal of Science and Engineering.* 6(3):2021.

[66] M. Seera, C. P Lim, Kumar A, Dhamotharan L, K. H. Tan.(2021). An intelligent payment card fraud detection system. Ann OperRes.1–23.

[67] M. Thirunavukkarasu, N. Achutha, Adusumilli J. (2021), Credit Card Fraud Detection Using Machine Learning, International Journal Of Computer Science And Mobile Computing, 10(4), 71-79.

[68] D. Trisanto, N. Rismawati, M. F. Mulya, Kurniadi F. I.. (2021). Modified focal loss in imbalanced XGBoost for credit card fraud detection. *International Journal of Intelligent Engineering and Systems.* 4(4):350–358.

[69] J. M. Vadakara, D. V.. Kumar (2019). Aggrandized random forest to detect the credit card frauds. *Advances in Science, Technology and Engineering Systems Journal.* 4(4):121–127.

[70] V. Vijayakumar, N. S. Divya, P. Sarojini, Sonika K. (2020). Isolation forest and local outlier factor for credit card fraud detection system. *International Journal of Engineering and Advanced Technology (IJEAT)* 2020; 9(4):261–265.

[71] X. Zhang, Y. Han, W. Xu, Wang Q. (2021) . HOBA: a novel feature engineering methodology for credit card fraud detection with a deep learning architecture. *Information Sciences.*557 (10):302–316.

[72] S. I. Yanxia and W. Zenghui (2020).A machine learning based credit card fraud detection using the GA algorithm for feature selection. *Journal of Big Data (2022) 9:24. 1-17.*