Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 01, January-2022

An Evaluation of Machine Learning Methods to Predict Fraud in Mobile Money Transactions

Francis Effirim Botchey [1,2], Zhen Qin [1], Kwesi Hughes-Lartey [1,2]
[1]School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 610051, China
[2]Department of Computer Science, Koforidua Technical University, Koforidua EN-112-2188

***Abstract*:- Mobile Money transactions continue to be a dominant force in elevating the unbanked in developing countries into the FinTech community. Mobile Money transactions have seen tremendous growth over the years despite a worldwide economic downturn due to the corona virus pandemic. The positive projections in Mobile Money transactions face a lot of challenges, prominent among them being fraud-related crimes. Methods that have been fundamentally implemented by both mobile network operators and governments (central banks) who are responsible for regulating all issues related to money in a country have been the education of customers on how to protect their digital wallets. These instituted methods have had little impact in resolving this problem. This article performs an evaluation on four different approaches, specifically deep learning (artificial neural network), ensemble machine learning methods and anomaly detection using python. The experiments showed the above-mentioned approaches hold great potential in combating this growing canker with optimal accuracies of 99.95% for artificial neural network, 99.78% for anomaly detection, 99.62% for XGBoost and 99.69% for random forest. Comparatively random forest obtained superior results in 3 out of the 4 evaluation criteria.**

Keywords— Mobile Money; artificial neural network; ensemble machine learning; anomaly detection

I. INTRODUCTION

There is no doubt about the numerous benefit that mobile handsets have brought about. In the 21st century, mobile handsets can be considered as a lifesaving tool. Virtually everything conceivable can be done with the aid of mobile handset; from virtual medical appointments, monitoring agricultural environments to performing all sorts of financial transactions. One of the key benefit of mobile money transactions (MMTs) if the offer of formal monetary instruments through multifaceted innovation. These ecosystems work on existing versatile technologies, and are overseen by telecommunication companies, banks, and outsider programming organizations and allows individuals with no financial history to establish financial accounts [1].

Apple Pay, Google Pay, Samsung Pay, WeChat and Alipay are examples of electronic payment systems that are enjoyed in the developed world but are essentially unheard of in sub-Saharan Africa and other developing countries due to the unavailability or poor internet infrastructure and services that these payments systems rely on. MMTs started as M-Pesa in 2007 in Kenya when it was introduced by Safaricom [2] and have quickly spread to other developing worlds. It is an alternative electronic payment system that can be operated using simple unstructured supplementary service data (USSD) or short message services (SMS) without necessarily the need of internet as compared to the other modes used in the developed world.

According to [3], a third of all account holders which is 12% of the adult population reported having a mobile money account in sub-Saharan Africa. Furthermore about half of this population held accounts at both financial institutions and mobile money accounts and the other half having only mobile money accounts. In Ghana, for example, the total value of mobile money transactions increased from about US$ 3.6 Billion in February 2019 to US$ 5.2 Billion within the same period in 2020, an increase of about 1.5 billion United States dollars [4]. The same period saw an increase from 32.7 million to 33.4 million in the number of mobile money registered accounts with an accompanying transaction volume of US$ 138 million to US$ 193 million [5]. Ghana's mobile money industry attained a transaction value of US$ 36.1 Billion in 2018 and is expected to reach a transaction value of US$ 204.3 Billion by 2024, growing at a compound annual growth rate (CAGR) of 32.5% during 2019-2024 [4]. However, as technology develops, the vulnerabilities in such technologies are exploited by crooks to perform various forms of nefarious activities [6, 7].

The success story of MMTs faces a bleak future if adequate measure to combat fraud in the industry are not vigorously pursued as [8] discovered in their research that MM users are reliant on structural soundness and perceived low risks. Castle et al. [1] asserted to the fact that prior work has identified discouraging vulnerabilities in the current ecosystem and also gave the assurance that all is not lost as the circumstance isn't as despairing as it might appear since many of the issues could be resolved by security best practices. Their work examined 197 Android applications and took a deeper look at 71 products in an attempt to assess specific organizational practices and concluded that although attack vectors are present in

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 01, January-2022

many of the applications, service providers are making security conscious decisions. Darvish and Husain [9] performed security analysis of mobile banking and mobile payment systems and found 80% of the selected applications were not adhering to best security practices.

Reports on MM fraud has become virtually a daily occurrence in most of these developing countries that have embraced this service [10–13]. The most perturbing trend according to the Ghana chamber of telecommunications is the ability of some of these fraudster targeting the bank accounts linked to mobile money accounts [14].

One of the methods used as a solution to this problem is the manual verification of the person initiating a transaction which usually happens at the withdrawal stage. This method is unreliable as humans cannot match the productivity of intelligent systems as they work unfailingly and continuously 24/7 without fatigue or other external influences. [15].

Several machine learning methods have been proposed in solving the problem of fraud in the financial sector particularly credit cards [16] and recently mobile money transactions [17, 18]. This article, however, looks at a different approach in providing a solution to this problem as very little research has been done in this area. The authors experimented with a combination of shallow and deep learning methods as well as anomaly detection. The rest of the paper is organized as follows. Review of similar work is presented in section II. Section III, present the experimental setup and methods. Performance Evaluation metrics for the experiments are presented in section IV. The results as well as the evaluations are discussed in section V. The paper is concluded in section VI.

The main contributions in this paper are:

1. Shallow and deep learning models based on four different classification algorithms for fraud prediction in mobile money are presented
2. Contributing greatly to literature in the domain of mobile money fraud prediction based on shallow and deep learning
3. An aid to mobile network operators enhance the security of their operations with regards to mobile money fraud by adopting machine learning approaches to protect subscribers of mobile money transactions.

II. RELATED WORK

Zhdanova et al. [19] used micro-structuring as a method of detecting fraud chains in MMTs. Their work shifted from the classical fraud detection methods such as machine learning and data mining and focused on extension to predictive security analysis at runtime which is a model based approach for event-driven process analysis. They reported a precision of 99.81% and 90.18% for recall. Gaber et al. [20] studied fraud detection in mobile payment systems. The work highlighted the limit of security information and events management (SIEM) to the problem of mobile based money transfer systems. They further explained the challenges associated with such systems, methods of identifying fraud schemes and finally proposing security features.

Reaves et al. [21] performed an automated analysis of mobile money apps and uncovered pervasive vulnerabilities which spanned from botched certification validation, do-it-yourself cryptography and other forms of information leakage. Their work concluded that majority of the apps used for these mobile money transactions failed to provide protection needed by mobile services. Novikova and Kotenko [22] proposed the use of RadViz visualization as representation for MMTs users' behavior; a technique that helps to identify groups with similar behavior and outliers. Lebeck et al. [23] proposed the use of Braavos, a system that combines existing primitive models in novel ways to secure branchless banking and also enables new functionality such as secure offline transactions.

Rieke et al. [24] used predictive security analyzer as a tool for predictive security analysis at runtime to observe processes and behavior in money transfers and tries to match it with an expected behavior. Seeja and Zareapoor [25] proposed a novel credit card fraud detection model which was based on frequent itemset mining. They handled the class imbalanced problem by obtaining legitimate as well as fraudulent transactions patterns for each customer by using frequent itemset mining. An algorithm was then developed to allocate to which class every transaction whether legitimate or illegal. The performance of their model was tested on UCSD data mining contest 2009 dataset and reported a very high fraud detection rate.

Halvaiee and Akbari [26] presented a credit card fraud detection system based on artificial immune called artificial immune system based fraud detection model and reported an increase in accuracy by 25% a cost reduction of 85% and a decrease system response time of up to 40% in comparison to the base algorithm. Huang et al. [27] pro- posed a novel detection framework, CoDetect that uses both network features to provide complementary information and feature information to simultaneously detect financial fraud activities as well as the feature patterns associated with them. Carcillo et al. [28] proposed a scalable real-time fraud finder that integrates big data tools with machine learning which deals with data imbalance, non-stationarity and feedback latency and reported its scalability, efficiency and accuracy over a big stream of transactions. Alam et.al [29] experimented with different datasets and data resampling after they have used Min-Max normalization is used to scale the features within one range. They also used different machine learning methods to predict loan default for both financial and analogous institutions. Ayeb et.al [30] proposed the use of community detection for detecting fraud in mobile money transactions. The work is an interesting proposal that seeks to study different community detection approaches and proceed to use them in their detection process. One of the key fraud schemes is a vulnerability in USSD known as SMShing which is a phishing attack performed through SMS. The work is however in its initial stages and many conclusions cannot be drawn about it.

Mubalaike et al. [31] experimented with deep learning approach to detecting fraud in financial transactions. Several other methods have been proposed [32]. It can be concluded from the above reviews that fraud prediction in mobile money transfer using both deep learning and shallow learning

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 01, January-2022

methods have been overwhelmingly neglected. Ahmed et.al [33] performed a structured survey on clustering-based fraud detection. The paper outlined the types of fraudulent activities and concluded the unavailability of a universal technique in the domain of fraud detection.

It is in this spirit that this article, seeks to experiment with three different approaches in providing solution to fraud in mobile money transactions namely, artificial neural network, ensemble machine learning methods and anomaly detection. These approaches were selected taking into consideration the need for huge dataset for other deep learning methods. The principal objective of this work was to ascertain the performance of these three different models on the imbalanced dataset and use appropriate evaluation metrics to establish their performance as none has been proposed to the best of our knowledge.

III. Experimental Setup And Methods

The dataset for the experiments as well as the experimental setup and methods are presented in this section.

A. Dataset description

The dataset for this paper was acquired from Kaggle [34]. The dataset is based on real world transactions from an international mobile money service provider in Africa. The dataset is made up of 6354407 data points with 10 attributes presented in Table 1 below. The description of the attributes are sourced from the same source.

TABLE I. MOBILE MONOEY DATASET DESCRIPTION

Attribute	Description
type	ASH-IN, CASH-OUT, DEBIT, PAYMENT, and TRANSFER, which are the main types of transactions under MMT.
amount	amount of the transaction in local currency
nameOrig	customer who started the transaction
oldbalanceOrg	initial balance before the transaction newbalanceOrig
newbalanceOrig	new balance after the transaction
nameDest	customer who is the recipient of the transaction
oldbalanceDest	initial balance recipient before the transaction
newbalanceDest	new balance recipient after the transaction
isFraud	This is the transaction made by the fraudulent.
isFlaggedFraud	The business model aims to control massive transfers from one account to another and flags illegal attempts. An illegal attempt in this dataset is an attempt to transfer more than 200.000 in a single transaction

B. Data preprocessing

Analysis of individual attributes in the dataset were performed. The attribute "type" consist of multiple attributes (CASH-IN, CASH-OUT, DEBIT, PAYMENT, and TRANS-FER) and were decomposed to its individual constituents using one-hot encoding. The new augmented numeric attributes were analyzed further to ascertain their influence on the dataset using Spearman's rank correlation coefficient to check for the independence between attributes. The p-values were also found to be accepted with the exception of "Debit" and "Payment" whose values were above the acceptable limit of 0.5 [35]. Five of the attributes were finally selected for the development of our experiments.

C. Artificial Neural Network

Artificial neural network (ANN) belongs to the larger field of machine learning with its inspiration from the structure and functions of the human brain. ANNs are based on a collection of connected artificial neurons. This article made use of multilayer perceptron. It also made use of an input layer 2 hidden layers and an output layer represented in Figure 1. The numbers for the hidden layers were obtained by performing a RandomizedSearchCV. The first step after the input layer was the summation of all the input variables and the weights that is:

$$w_t x_t = \sum_{i=1}^{n} w_i x_i \quad (1)$$

Where x_i are the input variables and w_i are the weights. The second step was to apply the activation function after the biases have been added which is given by:

$$\mathrm{A_t} = \text{ Activation } \left(\sum_{i=1}^{n} w_i x_i + \text{bias }\right) \quad (2)$$

This article used *he* uniform for the weight initialization. At the output stage, the loss was determined by;

$$\text{Loss} = (y - \hat{y}), \quad (3)$$

And the cost function, C_t is given by:

$$C_t = \sum_{i=1}^{n} (y - \hat{y})^2 \quad (4)$$

Where y is the real instance and $\hat{y}$ is the prediction from the artificial neural network. The weights were adjusted for backpropagation using adam [36–38] as the optimizer. Forward and backward propagation were performed for 100 epochs. The time complexity of neural network is . . . where n is the number of training samples and p the number of features [39]. Keras was used with TensorFlow as the backend.

D. Ensemble Machine learning methods

Fundamentally, ensemble machine learning are basically meta-algorithms that put together a number of machine learning algorithms to improve on the predictive power of the final model. There are two basic types of which include,

1. Bagging, which is also known as Bootstrap Aggregation with random forest as an example and
2. Boosting, with AdaBoost, Gradient boosting and XGBoost as examples.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 01, January-2022

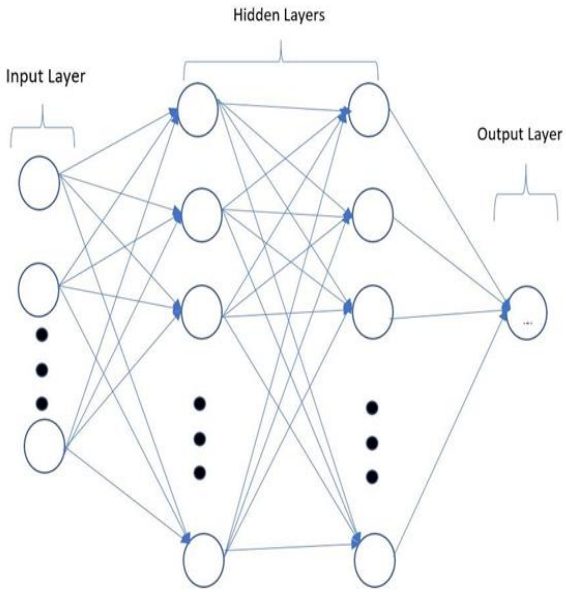

Figure 1. An ANN with an input layer, 2 hidden layers and an output layer.

This article made use of both bagging and boosting to enhance the quality of the evaluation process. Random forest was used as the bagging method and extreme gradient boosting (XGBoost) was used for the boosting method.

E. XGBoost

XGBoost is a member of the family of decision-tree-based ensemble machine learning algorithms which is based on the gradient boosting framework [40–42]. XGBoost uses decision trees which has a time complexity of $O(mn)^2$ where n denotes the number of instances and m denotes the number of attributes [43] cited by [44]. XGBoost makes use of optimized gradient boosting algorithm that attempts to prevent overfitting and bias. It attempts to improve on the standard boosting algorithms by providing the following from [42]:

1. A regularized model that attempts to prevent the problem of overfitting
2. A sparsity-aware split searching algorithm which deals efficiently with the different sparsity patterns in datasets.
3. A distributed weighted quantile sketch algorithms that effectively finds the optimal splits points in weighted dataset
4. A block structure that support parallelization of tree construction
5. Cache-area perfecting algorithm to fetch and store gradient statistics
6. Blocks for out-of-core computers

F. Random Forest

Random forest (RF) is a tree-based supervised machine learning algorithm that can be used for both classification and regression problems. It was used for classification in this paper. Random forest uses a number of decision trees in it workings. It then aggregates the outcomes from the individual trees to produce an output based on majority vote. The ability of random forest to handle nonlinear classification and the efficient handling of imbalanced data has made it one of the most accepted classification algorithms [45–47]

Random forest overcomes the problems of high variance and low bias. Since decision trees are created to its complete depth, it gets properly trained which reduces the training error. Again as many decision trees are used, the majority votes converts the high variance into low variance. This paper made use of random forest with row sampling and feature sampling with replacement.

G. Anomaly detection

Anomaly detection (AD) is a procedure used in the detection or identification of events in a dataset which defers from the norm [48]. Anomaly detection is gaining interest both in the fields of machine learning and statistics [49]. AD can be used for supervised (labeled). Due to the consistent perception of normality and abnormality, supervised anomaly detection is theoretically superior in overall accuracy [50], semisupervised (both labeled and unlabeled data), and unsupervised. Unsupervised anomaly detection is equivalent to semi-supervised anomaly detection in terms of merits, but it is often criticized for the validity of assumptions made in relevant activities. [50] (Unlabeled data) [51]. Oneclass classification eliminates the issue of an imbalanced dataset due to the single type of samples [50]. Anomaly detection has a time complexity of On^2 where n is the number of nodes [52].This article, however, made use of labeled data to predict (classify) fraudulent mobile money transactions. Anomaly detection operates on 2 postulates;

1. Anomalies in data are rare occurrences
2. The characteristics of anomalies in a dataset differ significantly from normal in- stances.

The Anomaly detection methods used in this article comprised the following;

1. Isolation forest
2. One-Class support vector machine
3. Local outlier factor

a. Isolation Forest

This is a tree based model. It is an algorithm used to detect outliers in a dataset and returns the score for each sample. It operates on the fact that anomalies in the dataset are rare and dissimilar [48, 53–56].

b.One-Class support vector machine (OCSVM)

One-Class support vector machine is an unsupervised learning algorithm whose training is based on the “data of interest”. In this case the data of interest is the fraudulent data points in the dataset. It observes the boundaries of these data

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 01, January-2022

points and are therefore able to detect other data points that lies outside the boundary [57].

c. Local outlier factor (LOF)

Local outlier factor (LOF) is a score that indicates how likely a particular data point is an anomaly or and outlier [58, 59]. Mathematically, it is given by:

$$\text{if } \begin{cases} \text{LOF} & \approx 1, \text{Not an outlier} \\ \text{LOF} & >> 1, \text{Outlier} \end{cases} \quad (5)$$

H. Performance Evaluation Metrics

The fundamental criteria for the evaluation of machine learning algorithms are; True Positive *(TP)*, False Positive *(FP)*, True Negative *(TN)*, and False Negative *(FN)*. The experiments were evaluated based on accuracy, Precision and Recall, and F- Measure, and whose definitions are given below: Accuracy: It is defined as the ratio of total of TP and TN to the total of TP, TN, FP, and FN. It shows in total how often a classification algorithm is right.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Precision: It is the ratio of *TP* to the total of *TP* an *FP*. It indicates the proportion of positive instances that are correctly predicted by the classification algorithm with regards to all the instances predicted as positives by the classifier.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

Recall: It is the ratio of *TP* to the total of *TP* and *FN*. It is a measure of how the classification algorithm accurately identifies the *TP's*

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

F-Measure: It is a representation of the harmonic mean of the Precision and Recall. The F-Measure combines both the properties of Precision and Recall.

$$F - \text{Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

I. Result and Discussion

The experimental results are presented and same are discussed. Table 2 presents the results for model accuracy. The accuracy of the experiments ranged from 99.95% for ANN, 99.78% for AD, 99.69% for Rf and 99.62% for XGBoost. ANN achieved its optimal accuracy after 35 epochs as shown in Figure 2. This is an indication of an overall good

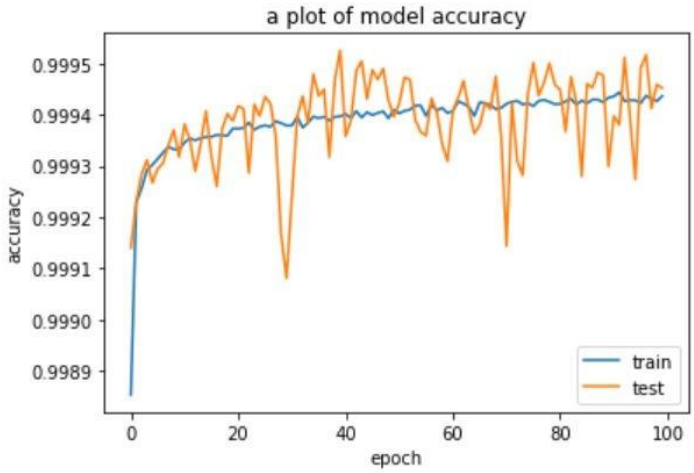

Figure 2. A plot of model accuracy for ANN

performance by all the classifiers as they all performed pretty well. ANN was the best classifier with regards to the model accuracy with AD coming in second RF third and XGBoost fourth.

The scores for precision are presented in Table 3. The scores are in the range of 16% for AD to 95% for RF. In terms of precision, RF was superior having obtained a value of 95%. XGBoost followed with a value of 94% then ANN with 93.4%. AD performed poorly in terms of precision. The values obtained for precision by the classifier is an indication that all classifiers with the exception of AD had a high proportion of positive instances correctly predicted with regards to all instances predicted as positive.

The result for recall are presented in Table 4. The values ranged from 16% for AD to 80% for RF. With RF been the best in this category, XGBoost was second, ANN third and AD coming in fourth with a poor performance of only 16%. Overall, the recall shows that the classifiers were above average in their ability to accurately identify true positives.

Table 5 presents the results for the F-Measure. The result for the F-Measure fell in the range of 16% for AD to 87% for RF. RF continued its dominance in this experiment obtaining the best score. XGBoost followed with ANN coming in third and AD been fourth. This is the harmonic mean of the precision and recall presented

Table 2: Report for Model Accuracy

Classifier	Accuracy (%)
Artificial Neural Network	99.95
Anomaly Detection	99.78
XGBoost	99.62
Random Forest	99.69

Table 3: Report for Precision

Classifier	**Precision (%)**
Artificial Neural Network	93.4
Anomaly Detection	16
XGBoost	94
Random Forest	95

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 01, January-2022

Table 4: Report for Recall

Classifier	Recall (%)
Artificial Neural Network	64.18
Anomaly Detection	16
XGBoost	75
Random Forest	80

Table 5: Report for F-Measure

Classifier	F-Measure (%)
Artificial Neural Network	76.08
Anomaly Detection	16
XGBoost	83
Random Forest	87

J. *Performance evaluation*

This section performs an evaluation of the results from the experiments. Considering the model accuracy, ANN was superior in this regard having achieved 99.95%. All models performed excellently in this criteria. However, due to the nature of our dataset which is highly imbalanced less emphasis is laid ion this performance criteria.

Precision which is an important evaluation metric produced pretty good results having 3 of the classifiers performing above 93%. AD however performed poorly in this criteria.

Recall is considered as an important metric in this experiments. It towed the path of Precision with 3 of the classifiers having values above 64%.

F-Measure, been the harmonic mean between Precision and Recall is one of the most important metric in evaluating the performance of ML algorithms. The F-Measure, just like Precision and Recall, had 3 of the classifiers having values greater than 76%. AD was the only classifier that perform poorly.

In comparison with other works, Xuan et.al [16] on credit card fraud detection using random forest and reported an accuracy of 98.67%, precision of 32.68%, and recall of 59.62%. Liu et.al [60] used data from a Chinese listed company to detect financial fraud based on random forest with four models with four statistical methods and reported the best model of obtaining an accuracy of 88%. Vidanelage et.al [61] conducted a study on machine learning techniques with conventional tools for payment and came out with an accuracy of 99.33% for KNN, 99.41% for MLP, 98.43% for GNB and 93.48% for MNB. Sa'adah et.al [62] worked on the same dataset (paysim) but used only 100 and 150 test data to classify the class between 0 and 1. Their work concluded that a combination between Probabilistic Neural Network and binary classification are good enough in classify those classes. Lu et.al [63] experimented on the same dataset with a different approach to the data preprocessing. The work also used g Random Forest, Decision Tree, Logistic Regression, Support Vector Machine, and Shallow Neural Network. The researchers. The research produced optimal accuracy of 99.2% for RF, 99.2% for decision tree and 87.9% for logistic regression. Pambudi et al. [64] worked on using the paysim dataset to improve detection in money laundering using optimized support vector machine with n precision of 40.82% and f1-score of 22.79%. Mubalaike et.al [31] experimented using with the paysim and produced an accuracy of 90.49% for decision tree, 80.52% for stacked auto encoder, and 91.53% for restricted Boltzmann machines.

Comparatively, our experiments produced superior results in comparison from accuracy, precision, recall and f-score with regards to the surveyed related works.

K. *Conclusion*

This article performed an evaluation analysis of 4 classification algorithms in predicting fraud in mobile money transactions. Ensemble machine learning with both bootstrap aggregation (Bagging) using RF and boosting using XGBoost were considered. Deep learning using ANN and AD were the other two classifiers. The performance of the classifiers was evaluated based on accuracy, precision, recall, and F-Measure. The experiment began by performing a critical examination of the dataset to determine it's relevance for the work at hand. Of the 10 in the original dataset, only 5 were found to be suitable for our experiments. The result are discussed and presented with ANN performing better than the other 3 models slightly in terms of model accuracy. Considering the fact of the high imbalance nature of the dataset, other 3 evaluation metrics were pursued for the analysis of the obtained results. Overall, RF was considered to be a better classifier for predicting fraud in mobile money transactions with regards to the other 3 classifiers in all the other 3 evaluations. AD apart from its good performance in model accuracy did poorly in the other 3.

This might be due to the fact the dataset used for the classification was not appropriate for anomaly detection [50] and will serve as a future direction RF is considered the best classifier in this paper. The limitation of our work points to the fact only one deep learning approach that is artificial neural network was considered. In future work, the dataset shall be resampled using both undersampling and oversampling to obtain a bigger picture and more information deduced. Again other deep learning methods shall be considered.

DATA AVAILABILITY STATEMENT

The dataset for this work is available at "https://www.kaggle.com/ntnu- testimon/paysim1".

ACKNOWLEDGMENTS

This work was supported in part by the Frontier Science and Technology Innovation Projects of National Key R&D Program (No.2019QY1405), the National Natural Science Foundation of China (No.61672135), the Sichuan Science-Technology Support Plan Program (No.2018GZ0236 and No.2017FZ0004), the Fundamental Research Funds for the Central Universities (No.2672018ZYGX2018J057), and the CERNET Innovation Project (No.NGII20180404).

REFERENCES

[1]. Castle, S.; Pervaiz, F.; Weld, G.; Roesner, F.; Anderson, R. Let's talk money: Evaluating the security challenges of mobile money in the developing world. Proceedings of the 7th Annual Symposium on Computing for Development, 2016, pp. 1–10.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 01, January-2022

[2] Buku, M.W.; Meredith, M.W. Safaricom and M-Pesa in Kenya: financial inclusion and financial integrity. Wash. Jl tech. & arts 2012, 8, 375.
[3] Demirgüç-Kunt, A.; Klapper, L.F.; Singer, D.; Van Oudheusden, P. The global findex database 2014: Measuring financial inclusion around the world. World Bank Policy Research Working Paper 2015.
[4] businesswire. Ghana Mobile Money Market: Industry Trends, Share, Size, Growth, Opportunity and Forecast 2019-2024. http://https://www.businesswire.com/news/home/20190919005505/en/Ghana-Mobile-Money-Market-Report-2019-2024---ResearchAndMarkets.com, 2021.
[5] ghanatalksbusiness. Mobile money transactions grew by GH10 billion over the last year- BoG report. http://ghanatalksbusiness.com/2020/03/mobile-money-transactions-grew-by-gh%E2%82%B510-billion-over-the-last-year-bog-report/, 2021.
[6] Shaukat, K.; Luo, S.; Varadharajan, V.; Hameed, I.A.; Chen, S.; Liu, D.; Li, J. Performance comparison and current challenges of using machine learning techniques in cybersecurity. Energies 2020, 13, 2509.
[7] Shaukat, K.; Luo, S.; Varadharajan, V.; Hameed, I.A.; Xu, M. A Survey on Machine Learning Techniques for Cyber Security in the Last Decade. IEEE Access 2020, 8, 222310–222354.
[8] Baganzi, R.; Lau, A.K. Examining trust and risk in mobile money acceptance in Uganda. Sustainability 2017, 9, 2233.
[9] Darvish, H.; Husain, M. Security analysis of mobile money applications on android. 2018 IEEE International Conference on Big Data (Big Data). IEEE, 2018, pp. 3072–3078.
[10] Myjoyonline. the earlier you report momo fraudsters to us the better mtn. http://www.myjoyonline.com/the-earlier-you- report-momo-fraudsters-to-us-the-better-mtn/, 2020.
[11] Mercy, W.B.; Mazer, R. fraud mobile financial services protecting consumers providers system. http://responsiblefinanceforum. org/publications/fraud-mobile-financial-services-protecting-consumers-providers-system/, 2017.
[12] Akomea-Frimpong, I.; Andoh, C.; Akomea-Frimpong, A.; Dwomoh-Okudzeto, Y. Control of Fraud on Mobile money services in Ghana: An exploratory study. Journal of Money Laundering Control 2019.
[13] Gilman, L.; Joyce, M. Managing the risk of fraud in mobile money. GSMA: Mobile Money for Unbanked (MMU) 2012.
[14] of Telecommunications, G.C. Mobile Money Fraudsters Now Target Bank Accounts Linked To MoMo Accounts. https: //telecomschamber.com/news-media/industry-news/mobile-money-fraudsters-now-target-bank-accounts-linked-to- momo-accounts, 2020.
[15] Shaukat, K.; Iqbal, F.; Alam, T.M.; Aujla, G.K.; Devnath, L.; Khan, A.G.; Iqbal, R.; Shahzadi, I.; Rubab, A. The Impact of Artificial intelligence and Robotics on the Future Employment Opportunities. Trends in Computer Science and Information Technology 2020, 5, 050–054.
[16] Xuan, S.; Liu, G.; Li, Z.; Zheng, L.; Wang, S.; Jiang, C. Random forest for credit card fraud detection. 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC). IEEE, 2018, pp. 1–6.
[17] Botchey, F.E.; Qin, Z.; Hughes-Lartey, K. Mobile Money Fraud Prediction—A Cross-Case Analysis on the Efficiency of Support Vector Machines, Gradient Boosted Decision Trees, and Naïve Bayes Algorithms. Information 2020, 11, 383.
[18] Adedoyin, A.; Kapetanakis, S.; Samakovitis, G.; Petridis, M. Predicting fraud in mobile money transfer using case-based reasoning. International Conference on Innovative Techniques and Applications of Artificial Intelligence. Springer, 2017, pp. 325–337.
[19] Zhdanova, M.; Repp, J.; Rieke, R.; Gaber, C.; Hemery, B. No smurfs: Revealing fraud chains in mobile money transfers. 2014 Ninth International Conference on Availability, Reliability and Security. IEEE, 2014, pp. 11–20.
[20] Gaber, C.; Gharout, S.; Achemlal, M.; Pasquet, M.; Urien, P. Security challenges for security information and event management systems in mobile money transfer services, 2012.
[21] Reaves, B.; Bowers, J.; Scaife, N.; Bates, A.; Bhartiya, A.; Traynor, P.; Butler, K.R. Mo (bile) money, mo (bile) problems: Analysis of branchless banking applications. ACM Transactions on Privacy and Security (TOPS) 2017, 20, 1–31.
[22] Novikova, E.; Kotenko, I. Visual analytics for detecting anomalous activity in mobile money transfer services. International Conference on Availability, Reliability, and Security. Springer, 2014, pp. 63–78.
[23] Lebeck, K.; Oluwafemi, T.; Kohno, T.; Roesner, F. Rethinking Mobile Money Security for Developing Regions. Technical report, Technical Report. University of Washington, 2015.
[24] Rieke, R.; Zhdanova, M.; Repp, J.; Giot, R.; Gaber, C. Fraud detection in mobile payments utilizing process behavior analysis. 2013 International Conference on Availability, Reliability and Security. IEEE, 2013, pp. 662–669.
[25] Seeja, K.; Zareapoor, M. Fraudminer: A novel credit card fraud detection model based on frequent itemset mining. The Scientific World Journal 2014, 2014.
[26] Halvaiee, N.S.; Akbari, M.K. A novel model for credit card fraud detection using Artificial Immune Systems. Applied soft computing 2014, 24, 40–49.
[27] Huang, D.; Mu, D.; Yang, L.; Cai, X. CoDetect: financial fraud detection with anomaly feature detection. IEEE Access 2018, 6, 19161–19174.
[28] Carcillo, F.; Dal Pozzolo, A.; Le Borgne, Y.A.; Caelen, O.; Mazzer, Y.; Bontempi, G. Scarff: a scalable framework for streaming credit card fraud detection with spark. Information fusion 2018, 41, 182–194.
[29] Alam, T.M.; Shaukat, K.; Hameed, I.A.; Luo, S.; Sarwar, M.U.; Shabbir, S.; Li, J.; Khushi, M. An Investigation of Credit Card Default Prediction in the Imbalanced Datasets. IEEE Access 2020, 8, 201173–201198.
[30] El Ayeb, S.; Hemery, B.; Jeanne, F.; Cherrier, E. Community Detection for Mobile Money Fraud Detection. 2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS). IEEE, 2020, pp. 1–6.
[31] Mubalaike, A.M.; Adali, E. Deep learning approach for intelligent financial fraud detection system. 2018 3rd International Conference on Computer Science and Engineering (UBMK). IEEE, 2018, pp. 598–603.
[32] Mahmoudi, N.; Duman, E. Detecting credit card fraud by modified Fisher discriminant analysis. Expert Systems with Applications
2015, 42, 2510–2516.
[33] Ahmed, M.; Mahmood, A.N.; Islam, M.R. A survey of anomaly detection techniques in financial domain. Future Generation Computer Systems 2016, 55, 278–288.
[34] Lopez-Rojas, E.; Elmir, A.; Axelsson, S. PaySim: A financial mobile money simulator for fraud detection. 28th European Modeling and Simulation Symposium, EMSS, Larnaca. Dime University of Genoa, 2016, pp. 249–255.
[35] Sullivan, G.M.; Feinn, R. Using effect size—or why the P value is not enough. Journal of graduate medical education 2012, 4, 279.
[36] Bock, S.; Weiß, M. A proof of local convergence for the Adam optimizer. 2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019, pp. 1–8.
[37] Bock, S.; Goppold, J.; Weiß, M. An improvement of the convergence proof of the ADAM-optimizer. arXiv preprint arXiv:1804.10587 2018.
[38] Zhang, Z. Improved adam optimizer for deep neural networks. 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). IEEE, 2018, pp. 1–2.
[39] Arora, S.; Barak, B. Computational complexity: a modern approach; Cambridge University Press, 2009.
[40] Chen, Z.; Jiang, F.; Cheng, Y.; Gu, X.; Liu, W.; Peng, J. XGBoost classifier for DDoS attack detection and analysis in SDN-based cloud. 2018 IEEE international conference on big data and smart computing (bigcomp). IEEE, 2018, pp. 251–256.
[41] Dhaliwal, S.S.; Nahid, A.A.; Abbas, R. Effective intrusion detection system using XGBoost. Information 2018, 9, 149.
[42] Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016, pp. 785–794.
[43] Oliveto, P.S.; He, J.; Yao, X. Time complexity of evolutionary algorithms for combinatorial optimization: A decade of results. International Journal of Automation and Computing 2007, 4, 281–293.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 01, January-2022

[44] Shaukat, K.; Luo, S.; Chen, S.; Liu, D. Cyber Threat Detection Using Machine Learning Techniques: A Performance Evaluation Perspective. 2020 International Conference on Cyber Warfare and Security (ICCWS). IEEE, 2020, pp. 1–6.

[45] Paul, A.; Mukherjee, D.P.; Das, P.; Gangopadhyay, A.; Chintha, A.R.; Kundu, S. Improved random forest for classification. IEEE Transactions on Image Processing 2018, 27, 4012–4024.

[46] Khoshgoftaar, T.M.; Golawala, M.; Van Hulse, J. An empirical study of learning from imbalanced data using random forest. 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007). IEEE, 2007, Vol. 2, pp. 310–317.

[47] Liu, X.Y.; Wu, J.; Zhou, Z.H. Exploratory undersampling for class-imbalance learning. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 2008, 39, 539–550.

[48] Liu, F.T.; Ting, K.M.; Zhou, Z.H. Isolation-based anomaly detection. ACM Transactions on Knowledge Discovery from Data (TKDD) 2012, 6, 1–39.

[49] Cabrera, J.B.; Gutiérrez, C.; Mehra, R.K. Ensemble methods for anomaly detection and distributed intrusion detection in mobile ad-hoc networks. Information fusion 2008, 9, 96–119.

[50] Kamran, S.; Talha, M.A.; Suhuai, L.; Shakir, S.; Ibrahim, A.H.; Jiaming, L.; Syed, K.A.; Umair, J. A Review of Time-Series Anomaly Detection Techniques: A Step to Future Perspectives. In Advances in Intelligent Systems and Computing; Springer, 2021.

[51] Bandaragoda, T.R.; Ting, K.M.; Albrecht, D.; Liu, F.T.; Zhu, Y.; Wells, J.R. Isolation-based anomaly detection using nearest-neighbor ensembles. Computational Intelligence 2018, 34, 968–998.

{52] Faroughi, A.; Javidan, R. CANF: Clustering and anomaly detection method using nearest and farthest neighbor. Future Generation Computer Systems 2018, 89, 166–177.

[53] Ding, Z.; Fei, M. An anomaly detection approach based on isolation forest algorithm for streaming data using sliding window. IFAC Proceedings Volumes 2013, 46, 12–17.

[54] Togbe, M.U.; Barry, M.; Boly, A.; Chabchoub, Y.; Chiky, R.; Montiel, J.; Tran, V.T. Anomaly Detection for Data Streams Based on Isolation Forest Using Scikit-Multiflow. International Conference on Computational Science and Its Applications. Springer, 2020, pp. 15–30.

[55] Elnour, M.; Meskin, N.; Khan, K.; Jain, R. A dual-isolation-forests-based attack detection framework for industrial control systems.
IEEE Access 2020, 8, 36639–36651.

[56] Liu, F.T.; Ting, K.M.; Zhou, Z.H. Isolation forest. 2008 eighth ieee international conference on data mining. IEEE, 2008, pp. 413–422.

[57] Zhang, R.; Zhang, S.; Muthuraman, S.; Jiang, J. One class support vector machine for anomaly detection in the communication network performance data. Proceedings of the 5th conference on Applied electromagnetics, wireless and optical communications. Citeseer, 2007, pp. 31–37.

[58] Xu, L.; Yeh, Y.R.; Lee, Y.J.; Li, J. A hierarchical framework using approximated local outlier factor for efficient anomaly detection.
Procedia Computer Science 2013, 19, 1174–1181.

[59] Paulauskas, N.; Bagdonas, A. .F. Local outlier factor use for the network flow anomaly detection. Security and Communication Networks 2015, 8, 4203–4212.

[60] Liu, C.; Chan, Y.; Alam Kazmi, S.H.; Fu, H. Financial fraud detection model: Based on random forest. International journal of economics and finance 2015, 7.

[61] Vidanelage, H.M.M.H.; Tasnavijitvong, T.; Suwimonsatein, P.; Meesad, P. Study on machine learning techniques with conventional tools for payment fraud detection. 2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE). IEEE, 2019, pp. 1–5.

[62] Sa'adah, S.; Pratiwi, M.S. Classification of Customer Actions on Digital Money Transactions on PaySim Mobile Money Simulator using Probabilistic Neural Network (PNN) Algorithm. 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). IEEE, 2020, pp. 677–681.

[63] Lu, C.; Lee, C.T.; Qiu, H.; Liu, M. Compare Shallow Neural Network and Conventional Machine Learning in Predicting Money Laundering Crimes.

[64] Pambudi, B.N.; Hidayah, I.; Fauziati, S. Improving Money Laundering Detection Using Optimized Support Vector Machine. 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). IEEE, 2019, pp. 273–278.

Identify applicable sponsor/s here. If no sponsors, delete this text box (*sponsors).*