British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 8, Issue 1, 2025 (pp. 42-54)
www.abjournals.org

FINANCIAL FRAUD PREDICTION IN CREDIT ADMINISTRATION: AN ENSEMBLE APPROACH

Ibukun O. Eweoya[1*], Taiwo O. Adigun[1,2], Oluwabamise Adeniyi[1], Amos Awoniyi[1], Alfred Udosen[1], Felix Idepefo[1], and Moradeke Adewumi[2,3]

[1]Department of Software Engineering, Babcock University, Nigeria.

[2]University of Lay Adventist of Kigali, Rwanda.

[3]Bamidele Olumilua University of Education, Science and Technology, Ikere-Ekiti, Nigeria.

Emails:

eweoyai@babcock.edu.ng, adigunt@babcock.edu.ng, adeniyi0416@pg.babcock.edu.ng, aawoniyi@babcock.edu.ng, audosen@babcock.edu.ng, idepefof@babcock.edu.ng, mgdewumi@gmail.com

[*]Corresponding Author's Email: eweoyai@babcock.edu.ng

Cite this article:

Eweoya, I. O., Adigun, T. O., Adeniyi, O., Awoniyi, A., Udosen, A., Idepefo, F., Adewumi, M. (2025), Financial Fraud Prediction in Credit Administration: An Ensemble Approach. British Journal of Computer, Networking and Information Technology 8(1), 42-54. DOI: 10.52589/BJCNIT-HRFN7ZOE

Manuscript History

Received: 11 Jan 2025

Accepted: 14 Feb 2025

Published: 6 Mar 2025

ABSTRACT: *The rate at which banks lose funds to loan beneficiaries due to loan default is alarming. As a result of this, subsequent applications to get loans are declined for paucity of funds while job loss is also a resultant effect. Due to the volatility, volume, and variety of data, the way human beings judge credit history has proven inefficient; including statistical approaches but the big data involved cannot be efficiently dealt with. This research uses past loan records based on employment of ensemble learning for fraud prediction in bank credit transactions in order to avoid credit. It evolves an ensemble learning approach to predict fraud in credit administration. AdaBoost ensemble approach was used for the work; MATLAB was employed for training, testing, validation, and to make fraud predictions. The result obtained was benchmarked with Naïve Bayes, Sequential Minimal Optimization (SMO), and decision tree; based on accuracy. The adopted approach attained an accuracy of 80.9% in 2.09 seconds being the highest accuracy compared to all learners used for the evaluation.*

KEYWORDS: Credit default, ensemble, fraud, machine learning, prediction.

Article DOI: 10.52589/BJCNIT-HRFN7ZOE
DOI URL: https://doi.org/10.52589/BJCNIT-HRFN7ZOE

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 8, Issue 1, 2025 (pp. 42-54)
www.abjournals.org

INTRODUCTION

In an attempt to effectively coordinate funds to credit beneficiaries, a large volume and variety of data are involved, with pronounced volatility. However, fraudsters leverage on diverse circumstances and loopholes to perpetrate fraud unhindered. This has financial implications on affected businesses and also a threat to the integrity of such organizations.

A fraud is said to have occurred in a situation where an intention to steal for one's benefit is established (Oloidi and Ajinaja, 2014; Rawte and Anuradha, 2015; Naik, and Laximinarayana, 2017; Akomolafe et al., 2017); a deceptive data usage is in this category too. There are existing fraud detection approaches in the context of discussion but a prediction is much desired in credit administration for enhanced efficiency, accuracy, and time benefits (Abdelhamid et al., 2017; Rohit, and Patel, 2015; Bagul et al., 2016).

Machine learning is an effective approach to reveal hidden patterns in data for fraud prediction (Bagul et al., 2016). Supervised or unsupervised techniques do exist (Brockett et al., 2002). Case-based reasoning failed in efficiency. However, this work employs ensemble approach for financial fraud prediction in bank credit default to enhance accuracy. It is evaluated by benchmarking with existing approaches using precision, recall and f-measures metrics, accuracy, and miscalculation rate.

LITERATURE REVIEW

There are many frauds in bank credit administration which need to be countered through intelligent technology (Bagul et al., 2016; Hameed et al., 2016). The fraud detection techniques that exist in this domain have accuracy shortcomings that need improvements; none has a focus on credit default fraud. Also, duplicates from fraudulent attempts, missing data, and yet to be identified fraud patterns have their tolls on prediction accuracy (Demla and Aggarwal, 2016; Fahmi et al., 2016; Vaishali, 2014; Abid et al., 2014, Sharma and Choudhury, 2016; Agaskar et al., 2017; Rawate and Tijare, 2017; Rimiru et al., 2017; Boateng and Oduro, 2018).

No existing supervised or unsupervised learning approach can singularly satisfy the required criteria of accuracy in solving a fraud detection problem effectively (Hetal and Amit, 2012). Problem of accuracy does arise because of missing data when using a single unsupervised learning approach. Similarly, a single supervised learning cannot discover fraudulent duplicates. It is therefore essential to explore ensembles in order to enhance credit fraud prediction accuracy.

A. FINANCIAL FRAUD

Any illegal action from men or machines in an attempt to derive personal financial benefits to the detriment of the legitimate human or institutional beneficiaries when it is devoid of errors is a financial fraud (Rawte and Anuradha, 2015; Bagul et al., 2016; Naik and Laximinarayana, 2017; Akomolafe et al., 2017; Kose et al., 2015). According to (Akomolafe et al., 2017), greed, gambling, debts, poor investments are the prompting situations. Financial fraud destroys systems, limits access to loan, causes death, and halts national economic growth due to credit defaults. Credit fraud could be via credit card, bankruptcy or credit application fraud (Delamaire et al., 2011; Laleh and Azgomi, 2009). Credit default is a situation where a loaned person cannot fulfill the payment promise as earlier documented.

Article DOI: 10.52589/BJCNIT-HRFN7ZOE
DOI URL: https://doi.org/10.52589/BJCNIT-HRFN7ZOE

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 8, Issue 1, 2025 (pp. 42-54)
www.abjournals.org

B. APPLICATION FRAUD

The act of making an application for a credit card with false information is referred to as application fraud. In an attempt to detect this type of fraud, a classification of two scenarios is carried out. Applications from same user with the same details are referred to as duplicates, when it comes from different individuals with similar details, then fraudsters are at play. Phua et al. (2006) describes it as "identity crime, leveraging on synthetically generated identities." Generally the types of fraud are as represented in Fig. 1 (Potamitis, 2013).

C. MACHINE LEARNING AND ITS TECHNIQUES

Machine learning employs data mining techniques, learning algorithms for the process of building models as it concerns hidden patterns in data to make predictions or detections.

1) CLASSIFICATION

Classification is a supervised machine learning method where datasets are labeled; predicting data in a predefined group. Neural networks, and support vector machine are examples. Ability to handle big data and attain high accuracy level; ability to provide results that industry experts can understand; and performance metrics are the basis for selection in fraud detection. Facial recognition, age estimation, stock prediction, fraud detection, are fertile areas of application (Guo et al., 2008; Kumar et al., 2010; Kirlidog and Asuk, 2012; Anwar et al., 2017). Table 1 (Bhavsar and Ganatra, 2012) offers a comparison of common classification techniques.

2) HYBRID Vs ENSEMBLE LEARNING

According to (Kazienko et al., 2013), both ensemble models and hybrid methods do employ information fusion concept but with a slight difference. Ensemble classifiers do combine multiple but homogeneous, weak models (Kajdanowiczet al., 2010), the output of each weak model are now subjected to merging approaches that do a grouping together, for instance, using majority voting and then train the combined output for instance, using decision templates (Kuncheva, 2002). In contrast, hybrid methods do combine completely different, heterogeneous machine learning approaches (Castillo et al., 2007; Corchado et al., 2010).

However, the hybrid and ensemble learning can make the complete solution more adaptive and enhance reasoning. This fact has endeared ensemble and hybrid approaches to a large field of study to solve problems for mankind, for example; facial recognition, medicine, age estimation, weather forecast, stock prediction, bioinformatics, text mining, and music classification (Castillo et al., 2007; Okun, 2011; Bergstra et al., 2006; Kempa et al., 2011).

D. ENSEMBLE LEARNING

Ensemble learning is a machine learning paradigm that is based on the training of multiple learners to solve a particular problem. Machine learning approaches do learn one hypothesis from training data. However, ensemble learning constructs a set of hypotheses and combine them for use. Combination of strengths in each base learner and fine-tuning the weaknesses gives an ensemble; enhancing reliability (Hansen and Salamon, 1990; Schapire, 1990).

Enhancing comprehensibility in ensembles is a research gap (Zhou et al., 2003). Boosting, bagging, and stacking are the foundational ensemble methods (Schapire, 1990; Wolpert, 1992;

Article DOI: 10.52589/BJCNIT-HRFN7ZOE
DOI URL: https://doi.org/10.52589/BJCNIT-HRFN7ZOE

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 8, Issue 1, 2025 (pp. 42-54)
www.abjournals.org

Breiman, 1996; Freund and Schapire, 1997) in literature. Adaboost is an example of the boosting ensemble approach, and its algorithm is an effective ensemble method.

The weighing strategy of Adaboost is equivalent to resampling the data space (Sun, 2007), this is also found in many classification approaches without a change to the method of learning. However, it minimizes information loss, reduces overfitting, as well as bias error of classifiers, with learning cost reduction (Ali et al., 2015).

E. CREDIT DEFAULT

There are numerous fraudulent bank credit operations resulting in credit defaults due to siphoned financing, double collateral, and dummy borrowers (Potamitis, 2013; NIBSS, 2024; CBN, 2023; CBN, 2024; Pollio and Obuobie, 2010; World Bank, 2024; O'Sullivan, and Sheffrin, 2003) . However, errors should not be mistaken for fraud.

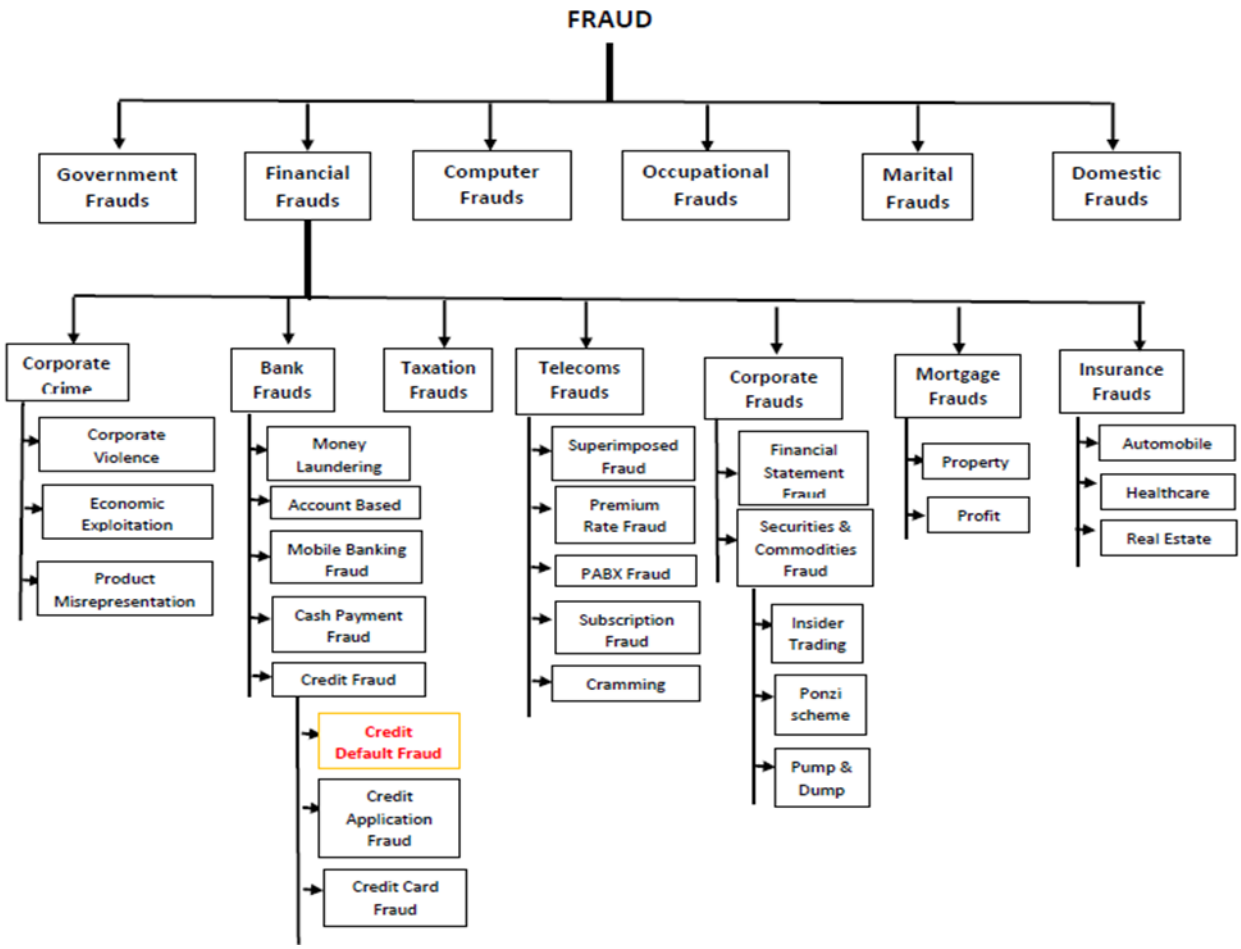

Figure 1: Types of fraud (Potamitis, 2013)

Article DOI: 10.52589/BJCNIT-HRFN7ZOE
DOI URL: https://doi.org/10.52589/BJCNIT-HRFN7ZOE

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 8, Issue 1, 2025 (pp. 42-54)
www.abjournals.org

REVIEW OF RELATED WORKS

In literature, many approaches were employed for fraud detection. Furthermore, machine learning approaches have been deployed in this domain too, and their capabilities, comparison explored (Kirkos et al., 2007; Abbasi et al., 2012; West and Bhattacharya, 2016).

In (Seeja and Zareapoor, 2014), a credit card fraud detection model was developed based on frequent item set mining. It employed a matching algorithm to discover the closeness of an incoming transaction to either legitimate or fraudulent for a decision to be made and it recorded high fraud detection. Furthermore, the work of (Rao and Singh, 2013) employed ensemble tree learning methods and genetic algorithm to discover financial fraud. This concentrated on credit card fraud using the UCI machine learning repository. The implemented decision tree was enhanced by Adaboost, using WEKA as can be done t other classifiers (Witten et al., 2011; Freund and. Schapire, 1998).

As reported by (Bian et al, 2016), previous fraud detection works disregarded the crucial imbalance nature of fraud data. It is imbalance because the number of valid records is largely smaller than the number of illegal fraud records. Ensembles (bagging and boosting) were concluded as the best.

Table 1: Comparison of common classification techniques (Bhavsar and Ganatra, 2012).

	Decision Trees	**Neural Networks**	**Naïve Bayes**	**K-NN**	**SVM**
General accuracy	**	***	*	**	****
Speed of learning	***	*	****	****	*
Speed of classification	****	****	****	*	****
Missing values tolerance	***	*	****	*	**
Tolerance to irrelevant attributes	***	*	**	**	****
Tolerance to redundant attributes	**	**	*	**	***
Tolerance to highly interdependent attributes	**	***	*	*	***
Dealing with discrete/binary/continuous attributes	All	Not discrete	Not continuous	All	Discrete
Tolerance to noise	**	**	***	*	**
Dealing with danger of overfitting	**	*	***	***	**
Attempt for incremental learning	**	***	****	****	**
Explanation ability, knowledge transparency, classification	****	*	****	**	*
Support multi-classification	****	Naturally extended	Naturally extended	****	Binary classifier

*** = Average, ** = Good, *** = Very good, **** = Excellent**

Article DOI: 10.52589/BJCNIT-HRFN7ZOE
DOI URL: https://doi.org/10.52589/BJCNIT-HRFN7ZOE

The work in (Rahmawati, et al., 2018) affirmed a 55% rise in bank frauds in 2016 and this happened because fraud could hardly be noticed traditionally except it has been committed. Therefore, the work proposed a fraud detection method in bank credit administration using the credit operation event log in combination with the hidden Markov models. The publication claimed an accuracy of 94% in its experimental fraud detection. Ninety event logs were utilized in the work. However, for enhanced research robustness, the event log quantity could be increased.

In (Huang et al., 2018), the complexity of networks and financial transactions involved in a financial fraud like money laundering was exploited as a reason for difficulty in detecting financial fraud. According to the work, the complement of fraud network and its features can improve fraud detection performance but many of existing works in that domain concentrate on either network or features but not the two in their fraud detection methods. Therefore, the work proposed "CoDetect" as a fraud detection framework that combines network information and feature information in its detection of financial fraud using the anomaly detection approach. It utilized both primary and secondary datasets but the specific interest was on money laundering.

Intelligent analysis of human behavior was used to detect financial fraud through a proposed framework called "FruadFind" in (Sanchez et al., 2018). It leveraged on the popular financial audit model called the fraud triangle theory to finally evolve the proposed conceptual framework. FraudFind is meant to discover, recognize, and be a pointer to fraudulent bank staff using semantic approaches.

The proposal for building a classification model in the detection of credit loan fraud based on individual level utility was evolved in (Choi et al., 2013), differentiating it from other works. A financial institution's dataset, spanning six months was used. However, more dataset could enhance its result.

In (Ajah and Inyiama, 2011), the loan assessment system in Nigeria was analyzed due to a high rate of credit default in bank loan administration. The researchers did a critical study of bank loan fraud detection and IT-based strategies in the Nigerian credit market in banks. In times past, credit default by corporate and individual clients had almost crashed banks before the intervention of the central bank. The work explored various IT options to mitigate loan fraud in Nigeria's banks credit administration. However, this work failed to employ a data mining approach to stop fraud, moreso the available data is huge nowadays with the embraced IT platforms of operation.

Odeh et al. (2011) evolved a multi-objective approach for the prediction of loan defaults using the fuzzy simplex algorithm which is a multi-objective optimization algorithm to generate decision rules in the prediction of loan default in a typical credit institution or bank. It reflected that pointers to default status are low equity and repayment of owners, too low or high capital.

Article DOI: 10.52589/BJCNIT-HRFN7ZOE
DOI URL: https://doi.org/10.52589/BJCNIT-HRFN7ZOE

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 8, Issue 1, 2025 (pp. 42-54)

www.abjournals.org

METHODOLOGY

This current work used a real life data from a primary source specifically the bank customers credit data from a reliable source where confidentiality is a special force to reckon with.

The data employed for this work has 16 attributes and 5000 instances with numerical and categorical attributes. The training, testing, modelling, and simulation of this research was achieved by MATLAB 2017b with cross-validation employed.

Data Analysis

The data was trained and tested at the rate of 75 percent for training and 25 percent for testing using MATLAB, with cross validation employed. This is to achieve the desired flexibility and desired results. There are many attributes, out of which the most important were selected to discover legitimate or fraudulent transactions, leading to reliable predictions. Some of these attributes are binary; others are categorical, or numerical.

Research Tools and Technology for Fraud Prediction

The fraud prediction process started with a pre-processing which was done through the principal component analysis in MATLAB and the prediction of fraudulent transaction based on outliers observed in the trained and tested dataset. Supervised learning approaches were employed using the decision trees, Naïve Bayes, and sequential minimal optimization (SMO). The model employs other concepts and basic tools which are: Features selection, model integration, proof-of-concept implementation using MATLAB.

The Boosting Ensemble Prediction Based on Adaboost

This work employs the Adaboost algorithm as an ensemble approach to predict fraud in credit default. In this domain, fraud prediction has not been carried out before or any ensemble approach used and specifically using Adaboost. The results of Adaboost on real life credit dataset to predict fraud in credit default is benchmarked against the results of base learners like Naïve Bayes, decision tree, logistic regression, and the sequential minimal optimization. Also, the result was validated against other existing ensemble approaches. The theoretical background of the Adaboost algorithm is explored next.

A. ADABOOST ALGORITHM

Adaptive boosting as a boosting algorithm was proposed by (Freund and Schapire, 1996). It focuses on classification problems and it converts a set of weak classifiers into a strong one and it works as presented below:

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxx

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 8, Issue 1, 2025 (pp. 42-54)
www.abjournals.org

The Pseudo-code for Adaboost

$*\ Input: a\ set\ S, of\ m\ labeled\ example: S = x_i,\ ((x_{i,} y_i), i = (1, 2, ..., m)), with\ labels\ y_I \in Y$

$*\ Learn\ (a\ learning\ a\lg orithm)$

$*\ A\ constant\ L.$

[1] $Initialize\ for\ all\ i: w_1(i) := \frac{1}{m}$ *initialize the weights*

[2] $for\ l = 1\ to\ L\ do$

[3] $for\ all\ i: \quad p_l(i) := \frac{w_l(i)}{\sum_{i wl(i)}}$ *compute normalized weigths*

[4] $h_l := Learn(S, p_l)$ *call Learn with normalized weights*

[5] $\varepsilon_l := \sum_i p_l(i)[h_l(x_i) \neq y_i]$ *Calculate the error of* h_l

[6] $if\ \varepsilon_l > 1/2\ then$

[7] $L := l - 1$

[8] $goto\ 12$

[9] $\beta_l := \frac{\varepsilon_l}{(1-\varepsilon_l)}$

[10] $for\ all\ i: w_{l+1}(i) := w_l(i)\beta_l^{1-[h_l(x) \neq y_i]}$ *compute new weights*

[8] $end\ for$

[9] $Output: \quad h_f(x) := \arg\max_{y \in Y} \sum_{l=1}^{L} (\log \frac{1}{\beta_l})[h_l(x) = y]$

RESULTS

The results obtained with the use of Adaboost on the dataset employed by this work is shown in Table 2. AdaBoost has the accuracy of 80.9%. Fig.2 shows a confusion matrix of positive predictive values and false discovery rate, while Fig.3 is a confusion matrix showing ROC curve which is a function of the true positive rate and the false positive rate. Furthermore, Fig.4 depicts the prediction flow for ensemble prediction based on AdaBoost as coined out of a Weka knowledge flow to achieve enhanced explanation.

Table 1: Fraud prediction using Adaboost

	TP Rate	FP Rate	Precision	Recall	F-measure	ROC Area	Class
	0.961	0.721	0.823	0.961	0.887	0.722	N
	0.279	0.039	0.674	0.279	0.394	0.722	Y
Av	0.809	0.569	0.79	0.809	0.777	0.722	

Instances correctly classified = 80.94%; Instances incorrectly classified = 19.06%

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxx

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 8, Issue 1, 2025 (pp. 42-54)
www.abjournals.org

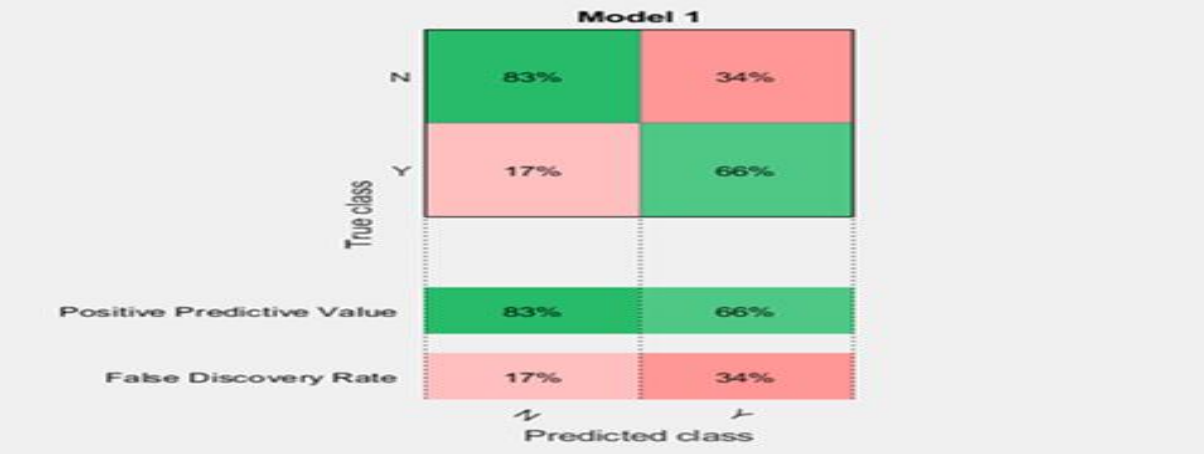

Figure 2: Positive predictive values and false discovery rate

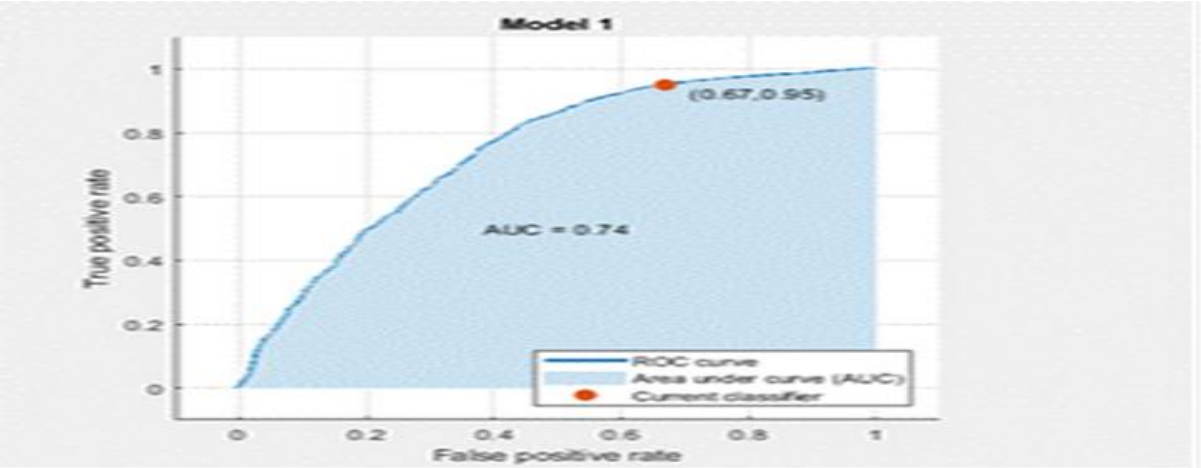

Figure 3: The ROC curve

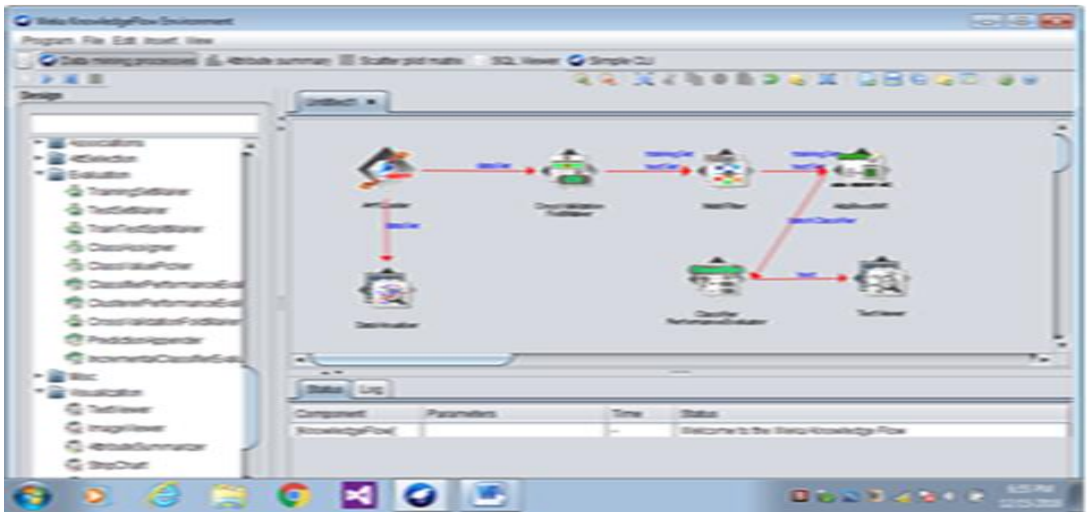

Figure 4: The knowledge flow

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxx

CONCLUSION

The ensemble learning approach in this work employs data pre-processing, base learners training and testing, and also ensemble training and testing based on a real life bank credit dataset with required consciousness of privacy and confidentiality of information. The proof-of- concept of the model has a high accuracy and fast speed.

The AdaBoost ensemble approach adopted in this research was found to attain an accuracy of 80.9% with the employed dataset being the highest accuracy compared to three base learners engaged for the evaluation and also other existing ensembles. SMO has an accuracy of 77.7%, Naïve Bayes attained 78%, decision tree has 77.8% accuracy. The adopted approach in this current paper attained the highest accuracy and within a training time of 2.09 seconds compared to random forest's 20.55 seconds. It also has the least number of incorrectly classified instances of 19.06% compared to random forest with 22.52% and bagging of 19.22%.

Therefore, Adaboost being an ensemble approach has been able to use a voting process to discover the shortcomings in each iteration. It thereafter used its computed normalized weights, calling the result initially available for learning based on normalized weights, and then the error was calculated. With the errors identified and a further computation of new weights to eliminate the errors; it was able to arrive at an enhanced accuracy.

REFERENCES

A. A. Hameed, B. Karlik, and M. S. Salman, "Back propagation algorithm with variable adaptive momentum," Knowledge-based Systems, vol. 114, pp. 79-87, 2016.

A. Abbasi, C. Albrecht, A. Vance, and J. Hansen, "Metafraud: A meta-learning framework for detecting financial fraud," MIS Quarterly, vol. 36, no. 4, pp. 1293-1327, 2012.

A. Ali, S. M. Shamsuddin, and A. L. Ralescu, "Classification with class imbalance problem: A review," International Journal of Advanced Soft Computing Appllications, vol. 7, no. 3, pp. 176-204, 2015.

A. O'Sullivan, and S. M. Sheffrin, "Economics: Principles in Action," Upper Saddle River, New Jersey, Pearson Prentice Hall, 2003, pp. 272.

B. Hetal, and G. Amit, "A comparative study of training algorithms for supervised machine learning," International Journal of Soft Computing and Engineering, (IJSCE), vol. 2, no. 4, pp. 74-81, 2012.

C. Phua, R. Gayler, V. Lee, and K. Smith, "On the approximate communal fraud scoring of credit applications," Proceedings of Credit Scoring and Credit Control IX, pp: 1-10, 2006.

Central Bank of Nigeria, "Financial stability report - December 2016," Available from: https://www.cbn.gov.ng/out/2017/fprd/fsr%20december%202016%20(2).pdf. Retrieved December, 2023.

Central Bank of Nigeria, "Financial stability report - June 2017," Available from: https://www.cbn.gov.ng/Out/2018/FPRD/FSR%20June%202017.pdf . Retrieved December, 2023.

D. Abdelhamid, S. Khaoula, and O. Atika, "Automatic bank fraud detection using support vector machines," in Proc., ICCTIM, Dubai, UAE, 2014, pp. 10-17.

D. H. Wolpert, "Stacked generalization," Neural Networks, vol. 5, no. 2, pp. 241–260, 1992.

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxx

D. Huang, D. Mu, L. Yang, X. Cai, "Codetect: Financial fraud detection with anomaly feature detection," IEEE, vol. 6, pp. 19161-19174, 2018.
E. Corchado, A. Abraham, A. de Carvalho, "Editorial: Hybrid intelligent algorithms and applications," Information Sciences, vol. 180, no. 14, pp. 2633–2634, 2010.
E. Kirkos, C. Spathis, and Y. Manolopoulos, "Data mining techniques for the detection of fraudulent financial statements," Expert Systems with Applications, vol. 32, no. 4, pp. 995-1003, 2007.
E. Y. Boateng, and F. T. Oduro, "Predicting microfinance credit default: a study of Nsoatreman rural bank, Ghana," Journal of Advances in Mathematics and Computer Science (JAMCS), vol. 26, no 1, pp. 1-9, 2018.
G. A. Oloidi, and O. T. Ajinaja, "Bank frauds and forgeries in Nigeria: A study of the causes, types, detection and prevention," IOSR Journal of Economics and Finance, vol. 4, no. 2, pp. 41-5, 2014.
G. Guo, Y. Fu, C. R Dyer, and T. S. Huang, "Image-based human age estimation by manifold learning and locally adjusted robust regression," Transactions on Image Processing, IEEE, vol. 17, no. 7, pp. 1178-1188, 2008.
G. Pollio, and J. Obuobie, "Microfinance default rates in Ghana: Evidence from individual-liability credit contracts" Micro-Banking Bulletin, vol. 20, pp. 8-13, 2010.
G. Potamitis, "Design and Implementation of a fraud detection expert system using ontology-based techniques," unpublished dissertation, University of Manchester, 2013.
H. Bhavsar, and A. Ganatra, "A comparative study of training algorithms for supervised machine learning," International Journal of Soft Computing and Engineering, vol. 2, no. 4, pp. 2231-2307, 2012.
I. Ajah, and C. Inyiama, "Loan fraud detection and IT-based combat strategies," Journal of Internet Banking and Commerce, vol. 16, no. 2, 2011.
I. H. Witten, E. Frank, and M. A. Hall, "Data mining: Practical machine learning tools and techniques" 3rd ed., Elsevier, 2011, pp. 664.
I. Kose, M. Gokturk, and K. Kilic, "An interactive machine learning-based electronic fraud and abuse detection system in healthcare insurance," Applied Soft Computing, vol. 36, pp. 283–299, 2015.
J. A. Akomolafe, D. F. Eluyela, S. O. Ilogho, J. W. Egharevba, and O. Aina, "Financial crime in Nigeria public sector: A study of Lagos state ministries," International Journal of Innovative Research in Social Sciences & Strategic Management Techniques, vol. 4, no. 1, pp. 13-21, 2017.
J. Bergstra, N. Casagrande, D. Erhan, D. Eck, B. Kegl, "Aggregate features and Ada-Boost for music classification," Machine Learning, vol. 65, pp. 473–484, 2006.
J. Naik, and J. A. Laximinarayana, "Designing hybrid model for fraud detection in insurance," IOSR Journal of Computer Engineering, vol. 1, 24-30, 2017.
J. West, and M. Bhattacharya, (2016). Intelligent financial fraud detection: A comprehensive review. Computers and Security, vol. 57, pp. 47-66, 2016.
K. Choi, G. Kim, Y. Suh, "Classification model for detecting and managing credit loan fraud based on individual-level utility concept", Data Base for Advances in Information Systems, vol. 44, no. 3, pp. 49-67, 2013.
K. D. Rohit, and D. B. Patel, "Review on detection of suspicious transaction in anti-money laundering using data mining framework," International Journal for Innovative Research in Science & Technology, vol. 1, no. 8, pp. 129-133, 2015.

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxx

K. R. Rawate, and P. A. Tijare, "Review on prediction system for bank loan credibility," International Journal of Advance Engineering and Research Development, vol. 4, no. 12, pp. 860-867, 2017.
K. R. Seeja, and M. Zareapoor, "FraudMiner: A novel credit card fraud detection model based on frequent itemset mining," The Scientific World Journal, vol. 1, pp, 1-10, 2014.
L. Abid, A. Masmoudi, and S. Zouari-Ghorbel, "The consumer loan's payment default predictive model: an application in a Tunisian commercial bank," Asian Economic and Financial Review, vol. 6, no. 1, pp. 27-42, 2016.
L. Breiman, "Bagging predictors," Machine Learning, vol. 24, no. 2, 123–140, 1996.
L. Delamaire, H. Abdou, and J. Pointon, (2011) "Credit card fraud and detection techniques: A review" Banks and Bank Systems, vol. 4, no 2, pp. 57-68.
L. K. Hansen, and P. Salamon, "Neural network ensembles," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no 10, pp. 993–1001, 1990.
L. Kuncheva, "Combining pattern classifiers, methods and algorithms," Hoboken, New Jersey: John Wiley & Sons, Inc.; 2004.
M. Fahmi, A. Hamdy, and K. Nagati, "Data mining techniques for credit card fraud detection: empirical study," Sustainable Vital Technologies in Engineering and Informatics, pp.1-9, 2016.
M. Kirlidog, and C. Asuk, "A fraud detection approach with data mining in health insurance," Procedia-Social and Behavioral Sciences, vol. 62 pp. 989-994, 2012.
M. Kumar, R. Ghani, and Z. S. Mei, "Data mining to predict and prevent errors in health insurance claims processing," in Proc. ACM 16th ICKDDM, 2010, pp. 65-74.
M. Sanchez, J. Torres, P. Zambrano, "FraudFind: Financial fraud detection by analyzing human behavior," in Proc. ICCC, Las Vegas, USA, 2018, pp. 78 – 90.
N. Demla, and A. Aggarwal, "Credit card fraud detection using SVM and reduction of false alarms," International Journal of Innovations in Engineering and Technology, vol. 7, no 2, pp. 176-182, 2016.
N. Laleh, and A. M. Azgomi, "A Taxonomy of Frauds and Fraud Detection Techniques," Information Systems Technology and Management, vol. 31. pp. 256-267, 2009.
Nigeria Interbank Settlement System, "2014-E-payment fraud landscape in Nigeria: A summary and analysis of reported e-payment frauds," 2015. Available from: www.nibss-plc.com.ng. Retrieved: March, 2018.
O. Castillo, P. Melin, W. Pedrycz, "Hybrid Intelligent Systems: Analysis and Design, Studies in Fuzziness and Soft Computing," Springer, Berlin, Heidelberg, 2007, pp. 431
O. Kempa, T. Lasota, Z. Telec, B. Trawiński, "Investigation of bagging ensembles of genetic neural networks and fuzzy systems for real estate appraisal," Lecture Notes in Artificial Intelligence, 2011, pp. 323–332.
O. Odeh, P. Koduru, A. Featherstone, S. Das, S. M. Welch, "Expert Systems with Applications," vol. 38, no. 7, pp. 8850-8857, 2011.
O. Okun, "Feature selection and ensemble methods for bioinformatics: Algorithmic classification and implementations," IGI Global, Hershy, PA: U.S.A, 2011, pp. 460
P. Brockett, R. Derrig, L. Golden, A. Levine, and M. Alpert, "Fraud classification using principal component analysis of RIDITs," Journal of Risk and Insurance, **69**(3): 341-371, 2002.
P. D. Bagul, S. Bojewar, and A. Sanghavi, "Survey on hybrid approach for fraud detection in health insurance," International Journal of Innovative Research in Computer and Communication Engineering, vol. 4, no. 4, pp. 6918-6922, 2016.

Article DOI: 10.52589/BJCNIT-HRFN7ZOE
DOI URL: https://doi.org/10.52589/BJCNIT-HRFN7ZOE

P. Kazienko, E. Lughofer, and B. Trawiński, "Hybrid and ensemble methods in machine learning," Journal of Universal Computer Science, vol. 19, no. 4, pp. 457–461, 2013.
R. E. Schapire, "The strength of weak learnability," Machine Learning, vol. 5, no. 2, 197–227, 1990.
R. Rimiru, S. W. Wa, and C. Otienoc, "A hybrid machine learning approach for credit scoring using PCA and logistic regression," International Journal of Computer, vol. 27, no. 1, pp. 84-102, 2017.
Rahmawati, D., Sarno, R., Fatichah, C., Sunaryono, D. "Fraud Detection on Event Log of Bank Financial Credit Business Process using Hidden Markov Model Algorithm," in Proc. 3rd ICSITech, IEEE, 2018, pp. 35-40.
Rawte, V., and Anuradha, G., "Fraud detection in health insurance using data mining techniques," in Proc. ICCICT, 2015, pp. 1-5, Mumbai.
S. Anwar, J. M. Zain, M. F. Zolkipli, Z. Inayat, S. Khan, B. Anthony, and V. Chang, "From intrusion detection to an intrusion response system: Fundamentals, requirements, and future directions," MDPI Algorithms, vol. 10, no. 2, pp. 1-24, 2017.
S. Sharma, and A. R. Choudhury, "Fraud analytics: A survey on bank fraud and fraud prediction using unsupervised learning based approach," International Journal of Innovations in Engineering Research and Technology, vol. 3, no. 3, pp. 1-9, 2016.
T. Kajdanowicz, P. Kazienko, J. Kraszewski,."Boosting algorithm with sequence-loss cost function for structured prediction," in Proc. HAIS LNAI 6076, Heidelberg, 2010, pp. 573–580.
V. Agaskar, M. Babariya, S. Chandran, and N. Giri, "Unsupervised learning for credit card fraud detection," International Research Journal of Engineering and Technology (IRJET), vol. 4, no. 3, pp. 2343-2346, 2017.
V. M. Rao, and Y. P. Singh, "Decision tree induction for financial fraud detection using ensemble learning techniques," in Proc. ICAICS and ICT, 2013, pp. 25-26.
Vaishali, V. (2014). Fraud detection in credit card by clustering approach, International Journal of Computer Applications, vol. 98, no. 3, pp. 29-32, 2014.
World Bank.(2018). Economic indicators for over 200 countries, https://www.theglobaleconomy.com/Nigeria/Nonperforming_loans/.Retrieved March, 2018.
Y. Bian, M. Cheng, C. Yang, Y. Yuan, and Q. Li, "Financial fraud detection: A new ensemble learning approach for imbalanced data," in Proc. PACIS, Chianyi, Taiwan, 2016, pp. 1-11.
Y. Freund, and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences, vol. 55, no. 1, pp. 119–139, 1997.
Y. Freund, and R. E. Schapire, "Experiments with a new boosting algorithm," in Proc. 13th ICML, 1996, Bari, Italy, pp. 148 – 156.
Y. Sun, M. S. Kamel, A. K. Wong, and Y. Wang, "Cost-sensitive boosting for classification of imbalanced data," Pattern Recognition, vol. 40, no. 12, pp. 3358-3378, 2007.
Z. H. Zhou, Y. Jiang, and S. F. Chen, "Extracting symbolic rules from trained neural network ensembles," AI Communications, vol. 16, no. 1, pp. 3–15, 2003.

Article DOI: 10.52589/BJCNIT-HRFN7ZOE
DOI URL: https://doi.org/10.52589/BJCNIT-HRFN7ZOE