Investigating an Anomaly-based Intrusion Detection via Tree-based Adaptive Boosting Ensemble

Paul Avweresuo Onoma [1,*], Joy Agboi [2], Victor Ochuko Geteloma [3], Asuobite ThankGod Max-Egba [4], Andrew Okonji Eboka [5], Arnold Adimabua Ojugo [6], Christopher Chukwufunaya Odiakaoase [7], Eferhire Valentine Ugbotu [8], Tabitha Chukwudi Aghaunor [9], and Amaka Patience Binitie [10]
[1,3,4,6] Department of Computer Science, Federal University of Petroleum Resources Effurun, Nigeria
[2] Department of Computer Science, Delta State University Abraka, Nigeria
[5,10] Department of Computer Science, Federal College of Education (Technical) Asaba, Nigeria
[7,8] Department of Computer, Dennis Osadebay University Asaba, Nigeria
[9] Department of Data Intelligence and Technology, Robert Morris University, Pittsburg, Pennsylvania, USA
[10] Department of Data Science, University of Salford, United Kingdom
Email: [1] kenbridge14@gmail.com, [2] agboijoy0@gmail.com, [3] geteloma.victor@fupre.edu.ng, [4] max-egbaasuobite@fupre.edu.ng, [5] andrew.eboka@fcetasaba.edu.ng, [6] ojugo.arnold@fupre.edu.ng, [7] osegalaxy@gmail.com, [8] eferhire.ugbotu@gmail.com, [9] tabitha.aghaunot@gmail.com, [10] amaka.binitie@fcetasaba.edu.ng
*Corresponding Author

***Abstract*—The eased accessibility, mobility, and portability of smartphones have caused the consequent rise in the proliferation of users' vulnerability to a variety of phishing attacks. Some users are more vulnerable due to factors like personality behavioral traits, media presence, and other factors. Our study seeks to reveal cues utilized by successful attacks by identifying web content as genuine and malicious data. We explore a sentiment-based extreme gradient boost learner with data collected over social platforms, scraped using the Python Google Scrapper. Our results show AdaBoost yields a prediction accuracy of 0.9989 to correctly classify 2148 cases with incorrectly classified 25 cases. The result shows the tree-based AdaBoost ensemble can effectively identify phishing cues and efficiently classify phishing lures against unsuspecting users from access to malicious content.**

***Keywords*—Anomaly Detection; Intrusion; Machine Learning; Boosting Ensemble; Tree-based Algorithms**

I. INTRODUCTION

Phishing utilizes multiple means like man-in-the-middle chat, forged links, and spoofed emails, amongst other means, to convince a user to divulge confidential data [1][2]. A major variant of phishing is spear phishing, which targets user email with links that seek to cleverly nudge an unsuspecting user to access malicious contents [3][4]; Thus, compromises a targeted user or device via malware download [5]. Phishing redirect users by exploiting vulnerabilities such as malware that is installed on the network infrastructure by an adversary [6][7]. Users are often redirected to fraudulent, spoofed websites without their knowledge cum consent. The socially-engineered attack consists of 3 elements [8][9]: (a) a lure feature from an adversary targets the unsuspecting user as a message that originates from a legitimate user on the network and is strengthened to exploit the unsuspecting victim's fear, curiosity, and empathy [10], (b) a hook is an attachment cum compromised link component of the message [11], and (c) a catch feature of the malicious content is the exploit facet of how the adversary obtains the unsuspecting user's private data [12][13].

While this may seem quite simple [14] – the method is constantly evolving by adversaries to evade detection. Also, its continued proliferation has equipped adversaries to vary their attacks to varying degrees of diversity at sporadic rates that improve their rate of success [15]. Its features include: (a) the message makes unrealistic demands via intimidation [16], (b) the message stresses a catch intent [17], (c) the message is rippled with misspellings and poor grammar [18], (d) mismatch in URL that redirects to a spoofed site [19], and (e) message asks for user sensitive data [20], etc. Phishing utilizes 2 features that aid its understanding capability to identify malicious data [21][22]: (a) believability increases possibility of a user to believe the contents of a website and not identify the malicious cues [23], and (b) insidiousness ascertains the potency of each cue, and its rate of success in remaining undetectable [24]. Studies have even tackled phishing attacks on IoTs [25][26].

To minimize phishing, machine learning (ML) has been trained to effectively recognize phishing cues and lures as patterns. They learn to identify these cues as predictor features using anomaly detection from the behavior norms as outlier data, or detected as unusual activity signatures in valid transactions that indicate a fraudulent profile [27]. Some MLs used are: Random Forest [28], Deep Learning [29][30], Logistic Regression [31], SVM [32], etc. Tree methods are common, and from a decision tree [33] – a tree generates a series of if-else rules, a majority vote for prediction [34][35]. Using a recursive tree to partition predictor space, each tree groups predictor(s) into a distribution of dependent variable y and predictor class x as homogeneous [36]. It constructs individually-trained trees to aggregate results into a stronger classifier that outperforms any single tree [37]; achieved via bagging [38][39] and boost [40]–[42] modes.

In boosting – tree(s) are constructed to achieve accuracy and performance by sequentially training its weak learners to correct its weaknesses [43][44] and in turn, yield a stronger classifier [45][46]. Common boost models are the adaptive boost [47], gradient boost [48], boosted logistic [49], and stochastic gradient boost [50][51]. In bagging, each decision

tree is constructed and summed via bootstrap mode to independently train each tree so that it samples data via majority vote [52]. The Random Forest extends the bagging mode with an extra layer of randomness that alters how its RF trees are constructed so that, unlike the decision trees that have each node split amongst its predictor variables – the RF tree splits its nodes amongst randomly chosen predictors [53]; Thus, exploits its recursive nature to unveil intricate interactions in the dataset as predictor feat [54][55]. In all, tree-based ensembles have successfully proven to be better than other schemes [56] in traffic flow [57], churn prediction [58], and purchase intention [59]. They are known to reduce variance [60] and bias [61] inherent in a dataset. While some models easily get stuck at local minima [62], the weighted fusion of trees produced by ensemble methods [63][64] – often minimizes the inherent risk of choosing the wrong local minimum [65]. Table 1 lists ML contributions to phishing schemes so far:

Table 1. Related Literature Contributions

Authors	Efficient Selected Algorithm	Accuracy
Ref. [66]	Long Term Short Memory (LSTM)	99.58%
Ref. [67]	LR, KNN, SVM, PCA, QDA, ANN	98.45%
Ref. [68]	LR, LSTM, XGBoost	97.23%
Ref. [69]	Deep Learning Ensemble	95.76%
Ref. [70]	KNN, LR, SVM, DT and RF	82.60%

Knowledge gaps inherent from previous works [71] as:

1. **Lack of Datasets:** Data access to the right quality, and the right format of domain dataset is a requisite for ML training and performance generalization – due to the limited availability of the dataset, which in turn, yields high error rates [72].
2. **Imbalanced Datasets:** A major issue with ML tasks is the imbalanced nature of its many datasets, which is also true with the case of phishing (as the minor class) where the data labels in the phishing (minor) class lag behind in its distribution frequency [40],[73][74]. Thus, a model must harness the robust flexibility in a tree-ensemble tailored to mitigating the challenges of data-imbalance [75][76].
3. **Cross-Channel Detection:** The new model must be able to handle and deal with the increased amount of transaction channels [77][78] vis-à-vis integrating various transaction data for enhanced generalization. Thus, cross-detection is now an imperative design and a focal area for businesses and researches [79][80] as traditional phishing detection modes are limited in adapting the emergent fraud patterns as well as keeping up with novel tactics.

II. MATERIALS AND METHODS

Our proposed method uses the Adaptive boosting scheme as in Fig. 1.

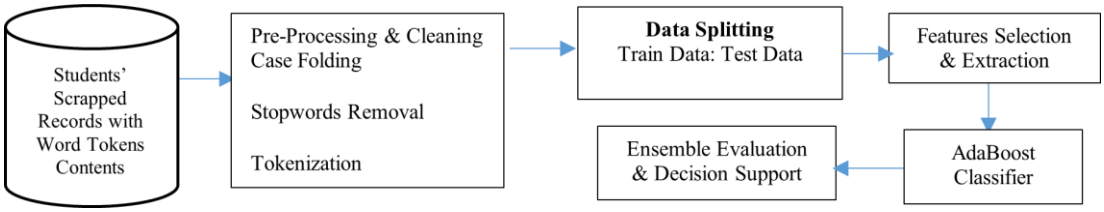

Fig. 1. Sentiment Analysis Process with Decision Support AdaBoost heuristics

1. **Step 1: Dataset –** were gathered via Google Play Scraper Library with a total of 8.693 records collected. Scrapped records include emails, compromised images-links-texts, posts, personal user data, likes/shares, etc – as reported and agreed by [81][82].
2. **Step 2: Preprocessing** cleans the dataset ensuring data integrity by removing redundant data, as well as data quality by removing missing data [83]. It restructures our dataset from its unstructured-to-normalized format [84] via the actions of word-stem, tokenization, and removal of stopwords [85] explained below as thus [22],[86]:
   - **Case Fold** converts all words onto tokens – and is achieved to avoid 2-or-more tokens to report the same meaning even when treated differently by the ML due to its written form (as in lowercase and uppercase) [87].
   - **Stopword** removes from the document – tokens that are common such as period, punctuation symbols, etc. This reduces the dataset dimension by resolving conjunctions, prepositions, and pronouns; And grouped as stopwords.
   - **Tokenization** splits a sentence into smaller units cum token elements so that each yields implicit meaning that when analyzed, grants insights into their placement order in the text corpus. Thus, they form input as normalized texts for tokens, stopwords, and special characters [88][89], some of which are removed by this unit or phase whose output yields input for the next.
   - **Normalization** extends/expands tokens to yield implicit meaning in their generic form. Thus, abridged words and slang/tokens are reverted to preserve their basic format. It expands tokens to their complete nature. For example, the term 'uwc' is expanded to 'you are welcomed'; while, 'notin' is expanded to 'nothing' [90]. Dictionaries used were as in the works [91].
   - **Word Stem** removes all word and token pre-fixes cum affixes – and converts each word to its root text [92].
3. **Step 3: Feature Extract** – identifies and determines the token to be used as input predictors for text classification via ranked selection of all underlying features grouping them in order of importance. This in turn, yields reduced dimensions for the adopted model inputs, yields fastened model construction, reduces the training time dynamics to yield a model devoid of overfit, decreases the required computational complexity, improves generalization of the

model, and enhances model accuracy [93][94]. We use **Term Frequency Inverse Document Frequency** to uncover relative probability scores for predictors; while ML schemes do not effectively recognize words as input – we use TD-IDF to convert each token into their numeric equivalence [95][96] by finding the occurrence in the frequency of each token in a document. A greater TF-value implies more frequency of such a token across the document [97]; while the IDF is the weighted sum of each token on the document so that the more a token appears, it yields a smaller IDF-value as expressed in (1), and in (2) [98] respectively:

$$IDF = \log\left(\frac{N}{DF}\right) \quad (1)$$

$$TF - IDF(d,k) = TF(d,k) * IDF(k) \quad (2)$$

4. **Step 4: Machine Learning** – We utilize the tree-based AdaBoost to effectively identify tokens into sentiments such as neutral, negative, and positive words-polarity according to the scrapped data. To train/test the dataset used [99] – our AdaBoost leverages the Gradient Boost scalability as it combines several weak learners to yield an optimal fit, that extends its objective function via minimized loss factor. This helps it to control its decision trees' complexity – and combines predictive processing so that each base learner contributes to ensure Adaboost is a stronger regressor [100] with train data x_i and its corresponding labels y_i – as in (3).

$$\hat{Y}_i^t = \sum_{k=1}^{t} f_k(x_i) = \hat{Y}_i^t + f_k(x_i) \quad (3)$$

For enhanced generalization – AdaBoost localizes its loss function $l(Y_i^t, \hat{Y}_i^t)$ as in its goal/objective function, which is fused with regularization term $\Omega(f_t)$. The loss factor ensures the tree is not over-trained or overfitted. Its regularization function suits fit model complexity so that with each tree tuned, the AdaBoost yields enhanced generalization with higher accuracy for a simplified ensemble that is devoid of over-parameterization, over-fit, and improved generalization as in (4) [101][102].

$$L^t = \sum_{i=1}^{n} l\left(Y_i^t, \hat{Y}_i^{t-1} + f_k(x_i)\right) + \Omega(f_t) \quad (4)$$

III. Result Findings and Discussion

A. Data Pre-Processing

As in Fig. 2 – positive sentiment words used to identify cues/lures [103] depicted as compromised links, images, etc.

B. Training Phase

We perform TF-IDF vectorization to help the ensemble convert retrieved texts into vectors via Python's ScikitLearn ***TfidfVectorizer*** function. With hyper-predictors, we train AdaBoost using the training dataset on a trial-and-error approach to yield its best-values generalization for learn_rate = 0.25, estimators_n = 250, and depth_max = 5 as in Table 2.

Table 2. Construction of Adaboost with Best Values

Features	Value	Description
estimators_n	250	Number of constructed decision tree
rate_learn	0.25	Learning stepsize to update the tree
depth_max	5	A decision tree's maximum depth
state_random	25	The seeds for reproduction
metrics_eval	["error', 'logloss']	Metrics to measure performance
eval_set	(x,val, y,val)	Train dataset to evaluate performance

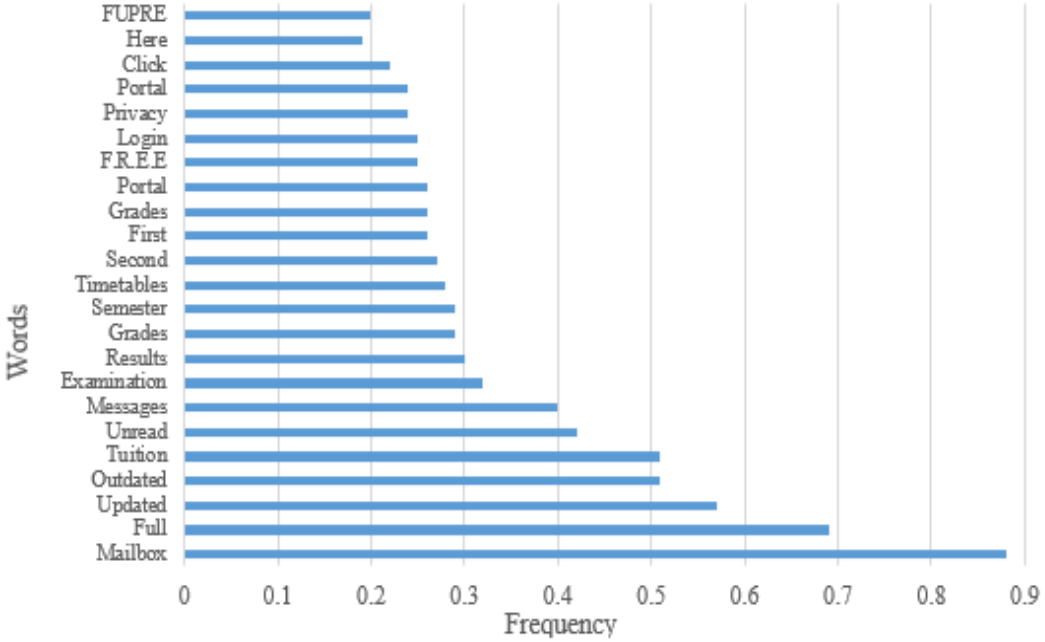

Fig. 2. Frequency Chart of Positive Sentiment Words

C. AdaBoost Learner Evaluation

Our word sentiments were evaluated in Table 3 [104]:

Table 3. AdaBoost Classifier Evaluation

Sentiments	F1-Score	Accuracy	Precision	Recall
Neutral	0.8992	0.9102	0.9082	0.9112
Negative	0.9829	0.9792	0.9789	0.9865
Positive	0.9981	0.9923	0.9795	0.9898

Table 3 notes that cues and lures for negative sentiments were detected with harmonic mean (F1-score) of 0.9829 with an accuracy of 0.9792; while the cues and lures for positive sentiments were detected with an accuracy of 0.9981, and agrees with [105]. Such disparities in accuracy are expected due to false-and-true positives cum negatives [106][107]. Testing yields an accuracy of 0.9923 to detect cues/lures for both the positive and negative sentiments [108]–[110].

D. Discussion of Findings

Fig. 3 confusion matrix clearly shows the proposed AdaBoost ensemble efficiently and correctly classified 2148 cases; 20 incorrectly-classified with an accuracy of 0.9989 as compared to [111]. Each user interface as *trust decision box* that allows a user to trust (*accept*) or not trust (*reject*) via content-specific decisions. Our use of the semantic-text normalization [112][113] yielded improved performance as compared with [114].

Fig. 3. Confusion Matrix for the Memetic Ensemble

Text normalization [115] did not degrade performance and agrees with [116]. Rather, it focuses on critical feats to successfully construct a model [117] to detect spoofed sites with reduced errors that will secure user resources and provide an enhanced user experience [118][119].

Table 4 shows sample cues in detecting malicious content such as emails, photo likes, posts, shares, etc as in agreement with [120][121].

Table 4. Website Content for Sample Lures and Cues

ID	Cues	Lures
S01	Legitimate site logos	Suspicious URL identifies sites
S02	The website looks/feels like a copy	Poor spelling/grammar issues
S03	Contextual/personal data required	Includes suspicious attachments in email
S04	Legitimate links hiding malicious data	Contains unnecessary warning messages
S05	Provides a sense of the previous trust	Directly requests the input of personal data
S06	No typos in Grammar and Writing style	References item prices too good to be true
S07	Uses official account usernames	Missing security designators, e.g. https padlock
S08	Identifies a known group of recipients	Appeals to emotion, e.g., urgency and greed
S09	Recognizes file types as attached	Unrecognized file types attached

Table 5 lists sophistication cues that make content harder to identify. The sample list is described below [101],[122] [123]:

S03: Contextual and Personal Data: user details known → *Content-Personal_information-IsObfuscated*

S04: Good Attachment: Generic boxes used → *Context-Image-Professional*

This study seeks to grasp how users make cum explore trusted decisions, identify cues and lures deficiencies in their decision/trust level, and adapt awareness campaigns cum training capabilities that will help prevent user susceptibility cum victimization, vis-à-vis associate such deficiencies with organizations where such users work [124]. The result yields mixed content for real-time user interactions with email, social media, and web-browser for which our participants' responses to malicious phishing content cues as insider tactics to keep unsuspecting victims engaged online and interested to yield increased online presence [125]. Also, our participants were provided a rich interaction platform with hover capabilities over attachments that explored and felt like natural browser-like schemes and behaviour [126][127].

E. Comparison

Improved generalization for our proposed ensemble as used in phishing demonstrates the flexibility, robustness, and adaptability of the tree-based AdaBoost ensemble as in [128] [129], which we benchmark across studies that explored the same. This is as seen in Table 5 [130][131].

Table 5. Benchmark / Comparative Testing of Method

Methods	F1	Accuracy	Precision	Recall
Ref [132]	0.8728	0.8500	0.8120	0.8925
Ref [133]	1.0000	1.0000	0.9999	1.0000
Ref [134]	0.7824	0.7631	0.7500	0.7732
Our Method	0.9928	0.9991	0.9792	0.9901

While some domain-chosen datasets are much easier to identify; Others, in turn, have proven to be more painstaking and tedious such as medical classification and image identification tasks. This requires that the explored ensemble yields a design generalization and performance evaluation that is strongly correlated to its error rates within the captured dataset. Thus, explored heuristics often measure sensitivity and specificity as crucial predictors that directly relate to all clinical outcomes.

IV. Conclusion

To protect users over social media platforms – designers have often explored the use of safeguards while utilizing ads to educate and campaign users against phishing. Where a case is reported, such is investigated, and where concluded that such is a potential phisher – the adversary is blacklisted. Thus, users are held accountable to report a case; while the platform is held responsible to investigate/blacklist potential adversaries vis-à-vis creating campaign awareness adverts that dissuades phishing attacks. Thus, platforms must have the capability to control and prevent attacks with actions or measures meted out to potential blacklisted users. This will help them stay ahead to limit such incidents. Furthermore, the consequent constant rise in technology and its widespread use to yield tech-rich business strategies with the proliferation of

smartphones has consequently resulted in improved user productivity, greater efficiency, and business profitability.

REFERENCES

[1] A. A. Ojugo *et al.*, "Forging a User-Trust Memetic Modular Neural Network Card Fraud Detection Ensemble: A Pilot Study," *J. Comput. Theor. Appl.*, vol. 1, no. 2, pp. 1–11, 2023, https://doi.org/10.33633/jcta.v1i2.9259.

[2] F. O. Aghware *et al.*, "BloFoPASS: A blockchain food palliatives tracer support system for resolving welfare distribution crisis in Nigeria," *Int. J. Informatics Commun. Technol.*, vol. 13, no. 2, p. 178, Aug. 2024, https://doi.org/10.11591/ijict.v13i2.pp178-187.

[3] H. Tingfei, C. Guangquan, and H. Kuihua, "Using Variational Auto Encoding in Credit Card Fraud Detection," *IEEE Access*, vol. 8, pp. 149841–149853, 2020, https://doi.org/10.1109/ACCESS.2020.3015600.

[4] A. Kiran, S. W. Prakash, B. A. Kumar, Likhitha, T. Sameeratmaja and U. S. S. R. Charan, "Intrusion Detection System Using Machine Learning," *2023 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1-4, 2023, https://doi.org/10.1109/ICCCI56745.2023.10128363.

[5] A. A. Ojugo and D. A. Oyemade, "Boyer moore string-match framework for a hybrid short message service spam filtering technique," *IAES Int. J. Artif. Intell.*, vol. 10, no. 3, pp. 519–527, 2021, https://doi.org/10.11591/ijai.v10.i3.pp519-527.

[6] R. R. Atuduhor *et al.*, "StreamBoostE: A Hybrid Boosting-Collaborative Filter Scheme for Adaptive User-Item Recommender for Streaming Services," *Adv. Multidiscip. Sci. Res. J. Publ.*, vol. 10, no. 2, pp. 89–106, 2024, https://doi.org/10.22624/AIMS/V10N2P8.

[7] G. Sasikala *et al.*, "An Innovative Sensing Machine Learning Technique to Detect Credit Card Frauds in Wireless Communications," *Wirel. Commun. Mob. Comput.*, vol. 2022, pp. 1–12, 2022, https://doi.org/10.1155/2022/2439205.

[8] S. V. S. . Lakshimi and S. D. Kavila, "Machine Learning for Credit Card Fraud Detection System," *Int. J. Appl. Eng. Res.*, vol. 15, no. 24, pp. 16819–16824, 2018, https://doi.org/10.1007/978-981-33-6893-4_20.

[9] A. M. Ifioko *et al.*, "CoDuBoTeSS: A Pilot Study to Eradicate Counterfeit Drugs via a Blockchain Tracer Support System on the Nigerian Frontier," *J. Behav. Informatics, Digit. Humanit. Dev. Res.*, vol. 10, no. 2, pp. 53–74, 2024, https://doi.org/10.22624/AIMS/BIJ/V10N1P6.

[10] B. O. Malasowe, F. O. Aghware, M. D. Okpor, B. E. Edim, R. E. Ako, and A. A. Ojugo, "Techniques and Best Practices for Handling Cybersecurity Risks in Educational Technology Environment (EdTech)," *J. Sci. Technol. Res.*, vol. 6, no. 2, pp. 293–311, 2024, https://doi.org/10.5281/zenodo.12617068.

[11] S. N. Okofu *et al.*, "Pilot Study on Consumer Preference, Intentions and Trust on Purchasing-Pattern for Online Virtual Shops," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 7, pp. 804–811, 2024, https://doi.org/10.14569/IJACSA.2024.0150780.

[12] B. O. Malasowe *et al.*, "Quest for Empirical Solution to Runoff Prediction in Nigeria via Random Forest Ensemble: Pilot Study," *Adv. Multidiscip. Sci. Res. J. Publ.*, vol. 10, no. 1, pp. 73–90, Mar. 2024, https://doi.org/10.22624/AIMS/BHI/V10N1P8.

[13] D. R. I. M. Setiadi, A. R. Muslikh, S. W. Irianandа, W. Warto, J. Gondohanindijo, and A. A. Ojugo, "Outlier Detection Using Gaussian Mixture Model Clustering to Optimize XGBoost for Credit Approval Prediction," *J. Comput. Theor. Appl.*, vol. 2, no. 2, pp. 244–255, 2024, https://doi.org/10.62411/jcta.11638.

[14] N. Vaughan, "Swapping algorithm and meta-heuristic solutions for combinatorial optimization n-queens problem," *2015 Science and Information Conference (SAI)*, pp. 102-104, 2015, https://doi.org/10.1109/SAI.2015.7237132.

[15] A. Algarni, Y. Xu, and T. Chan, "An empirical study on the susceptibility to social engineering in social networking sites: the case of Facebook," *Eur. J. Inf. Syst.*, vol. 26, no. 6, pp. 661–687, 2017, https://doi.org/10.1057/s41303-017-0057-y.

[16] R. E. Ako *et al.*, "Pilot Study on Fibromyalgia Disorder Detection via XGBoosted Stacked-Learning with SMOTE-Tomek Data Balancing Approach," *NIPES - J. Sci. Technol. Res.*, vol. 7, no. 1, pp. 12–22, 2025, https://doi.org/10.37933/nipes/7.1.2025.2.

[17] A. Basit, M. Zafar, A. R. Javed and Z. Jalil, "A Novel Ensemble Machine Learning Method to Detect Phishing Attack," *2020 IEEE 23rd International Multitopic Conference (INMIC)*, pp. 1-5, 2020, https://doi.org/10.1109/INMIC50486.2020.9318210.

[18] S. O. Dawodu, A. Omotosho, J. A. Odunayo, O. A. Abimbola, and S. K. Ewuga, "Cybersecurity Risk Assessment in Banking: Methodologies and Best Practices," *Comput. Sci. IT Res. J.*, vol. 4, no. 3, pp. 220–243, 2023, https://doi.org/10.51594/csitrj.v4i3.659.

[19] Y. Srivastava, P. Khanna and S. Kumar, "Estimation of Gestational Diabetes Mellitus using Azure AI Services," *2019 Amity International Conference on Artificial Intelligence (AICAI)*, pp. 321-326, 2019, https://doi.org/10.1109/AICAI.2019.8701307.

[20] J. Yao, C. Wang, C. Hu, and X. Huang, "Chinese Spam Detection Using a Hybrid BiGRU-CNN Network with Joint Textual and Phonetic Embedding," *Electronics*, vol. 11, no. 15, p. 2418, Aug. 2022, https://doi.org/10.3390/electronics11152418.

[21] P. O. Ejeh *et al.*, "Counterfeit Drugs Detection in the Nigeria Pharma-Chain via Enhanced Blockchain-based Mobile Authentication Service," *Adv. Multidiscip. Sci. Res. J. Publ.*, vol. 12, no. 2, pp. 25–44, 2024, https://doi.org/10.22624/AIMS/MATHS/V12N2P3.

[22] D. A. Oyemade and A. A. Ojugo, "A property oriented pandemic surviving trading model," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 5, pp. 7397–7404, 2020, https://doi.org/10.30534/ijatcse/2020/71952020.

[23] E. R. Altman, "Synthesizing Credit Card Transactions," In *Proceedings of the Second ACM International Conference on AI in Finance*, pp. 1-9, 2021, https://doi.org/10.1145/3490354.3494378.

[24] V. Umarani, A. Julian, and J. Deepa, "Sentiment Analysis using various Machine Learning and Deep Learning Techniques," *J. Niger. Soc. Phys. Sci.*, vol. 3, no. 4, pp. 385–394, 2021, https://doi.org/10.46481/jnsps.2021.308.

[25] M. I. Akazue *et al.*, "FiMoDeAL: pilot study on shortest path heuristics in wireless sensor network for fire detection and alert ensemble," *Bull. Electr. Eng. Informatics*, vol. 13, no. 5, pp. 3534–3543, 2024, https://doi.org/10.11591/eei.v13i5.8084.

[26] N. R. Pratama, D. R. I. M. Setiadi, I. Harkespan, and A. A. Ojugo, "Feature Fusion with Albumentation for Enhancing Monkeypox Detection Using Deep Learning Models," *J. Comput. Theor. Appl.*, vol. 2, no. 3, pp. 427–440, 2025, https://doi.org/10.62411/jcta.12255.

[27] L. R. Zuama, D. R. I. M. Setiadi, A. Susanto, S. Santosa, and A. A. Ojugo, "High-Performance Face Spoofing Detection using Feature Fusion of FaceNet and Tuned DenseNet201," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 4, pp. 385–400, 2025, https://doi.org/10.62411/faith.3048-3719-62.

[28] M. D. Okpor *et al.*, "Pilot Study on Enhanced Detection of Cues over Malicious Sites Using Data Balancing on the Random Forest Ensemble," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 2, pp. 109–123, 2024, https://doi.org/10.62411/faith.2024-14.

[29] I. Benchaji, S. Douzi, B. El Ouahidi, and J. Jaafari, "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model," *J. Big Data*, vol. 8, no. 1, p. 151, Dec. 2021, https://doi.org/10.1186/s40537-021-00541-8.

[30] F. O. Aghware, R. E. Yoro, P. O. Ejeh, C. C. Odiakaose, F. U. Emordi, and A. A. Ojugo, "DeLClustE: Protecting Users from Credit-Card Fraud Transaction via the Deep-Learning Cluster Ensemble," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 6, pp. 94–100, 2023, https://doi.org/10.14569/IJACSA.2023.0140610.

[31] D. Varmedja, M. Karanovic, S. Sladojevic, M. Arsenovic, and A. Anderla, "Credit Card Fraud Detection - Machine Learning methods," in *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)*, pp. 1–5, 2019, https://doi.org/10.1109/INFOTEH.2019.8717766.

[32] C. Li, N. Ding, H. Dong, and Y. Zhai, "Application of Credit Card Fraud Detection Based on CS-SVM," *Int. J. Mach. Learn. Comput.*, vol. 11, no. 1, pp. 34–39, 2021, https://doi.org/10.18178/ijmlc.2021.11.1.1011.

[33] V. O. Geteloma *et al.*, "Enhanced data augmentation for predicting consumer churn rate with monetization and retention strategies : a pilot study," *Appl. Eng. Technol.*, vol. 3, no. 1, pp. 35–51, 2024, https://doi.org/10.31763/aet.v3i1.1408.

[34] S. E. Brizimor *et al.*, "WiSeCart: Sensor-based Smart-Cart with Self-Payment Mode to Improve Shopping Experience and Inventory

Management," *Adv. Multidiscip. Sci. Res. J. Publ.*, vol. 10, no. 1, pp. 53–74, 2024, https://doi.org/10.22624/AIMS/SIJ/V10N1P7.

[35] T. Muralidharan and N. Nissim, "Improving malicious email detection through novel designated deep-learning architectures utilizing entire email," *Neural Networks*, vol. 157, pp. 257-279, 2023, https://doi.org/10.1016/j.neunet.2022.09.002.

[36] A. A. Ojugo and O. D. Otakore, "Forging An Optimized Bayesian Network Model With Selected Parameters For Detection of The Coronavirus In Delta State of Nigeria," *J. Appl. Sci. Eng. Technol. Educ.*, vol. 3, no. 1, pp. 37–45, Apr. 2021, https://doi.org/10.35877/454RI.asci2163.

[37] F. Jáñez-Martino, E. Fidalgo, S. González-Martínez, and J. Velasco-Mata, "Classification of Spam Emails through Hierarchical Clustering and Supervised Learning," *Natl. Cybersecurity Inst.*, vol. 24, pp. 1–4, 2020, https://doi.org/10.48550/arXiv.2005.08773.

[38] M. Kuradusenge *et al*., "Crop yield prediction using machine learning models: Case of Irish potato and maize," *Agriculture*, vol. 13, no. 1, p. 225, 2023, https://doi.org/10.3390/agriculture13010225.

[39] K. A. Egbe, A. Ike, and F. Egbe, "Knowledge and burden of hepatitis B virus in Nasarawa State, Nigeria," *Scientific African*, vol. 22, p. e01938, 2023, https://doi.org/10.1016/j.sciaf.2023.e01938.

[40] A. A. Ojugo and A. O. Eboka, "Extending Campus Network Via Intranet and IP-Telephony For Better Performance and Service Delivery: Meeting Organizational Goals," *J. Appl. Sci. Eng. Technol. Educ.*, vol. 1, no. 2, pp. 94–104, 2019, https://doi.org/10.35877/454RI.asci12100.

[41] A. A. Ojugo *et al.*, "Dependable Community-Cloud Framework for Smartphones," *Am. J. Networks Commun.*, vol. 4, no. 4, p. 95, 2015, https://doi.org/10.11648/j.ajnc.20150404.13.

[42] N. M. Shahani, X. Zheng, C. Liu, F. U. Hassan, and P. Li, "Developing an XGBoost Regression Model for Predicting Young's Modulus of Intact Sedimentary Rocks for the Stability of Surface and Subsurface Structures," *Front. Earth Sci.*, vol. 9, 2021, https://doi.org/10.3389/feart.2021.761990.

[43] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection," *Journal of Big Data*, vol. 9, no. 1, p. 24, 2022, https://doi.org/10.1186/s40537-022-00573-8.

[44] K. Muhamada, D. R. Ignatius, M. Setiadi, U. Sudibyo, B. Widjajanto, and A. A. Ojugo, "Exploring Machine Learning and Deep Learning Techniques for Occluded Face Recognition: A Comprehensive Survey and Comparative Analysis," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 2, pp. 160–173, 2024, https://doi.org/10.62411/faith.2024-30.

[45] F. Omoruwou, A. A. Ojugo, and S. E. Ilodigwe, "Strategic Feature Selection for Enhanced Scorch Prediction in Flexible Polyurethane Form Manufacturing," *J. Comput. Theor. Appl.*, vol. 1, no. 3, pp. 346–357, 2024, https://doi.org/10.62411/jcta.9539.

[46] S. Hemalatha, T. Kavitha, T. M. Saravanan, K. Chitra and N. Dinesh, "Forecasting Crop Using Machine Learning Model," *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pp. 783-788, 2022, https://doi.org/10.1109/ICESC54411.2022.9885377.

[47] [47] D. A. Al-Qudah, A. M. Al-Zoubi, P. A. Castillo-Valdivieso, and H. Faris, "Sentiment analysis for e-payment service providers using evolutionary extreme gradient boosting," *IEEE Access*, vol. 8, pp. 189930–189944, 2020, https://doi.org/10.1109/ACCESS.2020.3032216.

[48] T. Edirisooriya and E. Jayatunga, "Comparative Study of Face Detection Methods for Robust Face Recognition Systems," *5th SLAAI - Int. Conf. Artif. Intell. 17th Annu. Sess. SLAAI-ICAI 2021*, no. December, 2021, https://doi.org/10.1109/SLAAI-ICAI54477.2021.9664689.

[49] A. A. Ojugo, C. O. Obruche, and A. O. Eboka, "Quest For Convergence Solution Using Hybrid Genetic Algorithm Trained Neural Network Model For Metamorphic Malware Detection," *ARRUS J. Eng. Technol.*, vol. 2, no. 1, pp. 12–23, 2021, https://doi.org/10.35877/jetech613.

[50] M. G. Kibria and M. Sevkli, "Application of Deep Learning for Credit Card Approval: A Comparison with Two Machine Learning Techniques," *Int. J. Mach. Learn. Comput.*, vol. 11, no. 4, pp. 286–290, 2021, https://doi.org/10.18178/ijmlc.2021.11.4.1049.

[51] A. P. Binitie *et al.*, "Stacked Learning Anomaly Detection Scheme with Data Augmentation for Spatiotemporal Traffic Flow," *J. Fuzzy Syst. Control*, vol. 2, no. 3, pp. 203–214, 2024, https://doi.org/10.59247/jfsc.v2i3.267.

[52] A. Satpathi *et al.*, "Comparative Analysis of Statistical and Machine Learning Techniques for Rice Yield Forecasting for Chhattisgarh, India," *Sustainability*, vol. 15, no. 3, p. 2786, 2023, https://doi.org/10.3390/su15032786.

[53] A. Bahl *et al.*, "Recursive feature elimination in random forest classification supports nanomaterial grouping," *NanoImpact*, vol. 15, p. 100179, 2019, https://doi.org/10.1016/j.impact.2019.100179.

[54] A. Razaque *et al.*, "Credit Card-Not-Present Fraud Detection and Prevention Using Big Data Analytics Algorithms," *Appl. Sci.*, vol. 13, no. 1, p. 57, 2022, https://doi.org/10.3390/app13010057.

[55] A. A. Ojugo, C. O. Obruche, and A. O. Eboka, "Empirical Evaluation for Intelligent Predictive Models in Prediction of Potential Cancer Problematic Cases In Nigeria," *ARRUS J. Math. Appl. Sci.*, vol. 1, no. 2, pp. 110–120, 2021, https://doi.org/10.35877/mathscience614.

[56] B. P. Bhuyan, R. Tomar, T. P. Singh, and A. R. Cherif, "Crop Type Prediction: A Statistical and Machine Learning Approach," *Sustainability*, vol. 15, no. 1, p. 481, Dec. 2022, https://doi.org/10.3390/su15010481.

[57] E. U. Omede, A. E. Edje, M. I. Akazue, H. Utomwen, and A. A. Ojugo, "IMANoBAS: An Improved Multi-Mode Alert Notification IoT-based Anti-Burglar Defense System," *J. Comput. Theor. Appl.*, vol. 1, no. 3, pp. 273–283, Feb. 2024, https://doi.org/10.62411/jcta.9541.

[58] M. D. Okpor *et al.*, "Comparative Data Resample to Predict Subscription Services Attrition Using Tree-based Ensembles," *J. Fuzzy Syst. Control*, vol. 2, no. 2, pp. 117–128, 2024, https://doi.org/10.59247/jfsc.v2i2.213.

[59] B. O. Malasowe, M. I. Akazue, A. E. Okpako, F. O. Aghware, D. V. Ojie, and A. A. Ojugo, "Adaptive Learner-CBT with Secured Fault-Tolerant and Resumption Capability for Nigerian Universities," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, pp. 135–142, 2023, https://doi.org/10.14569/IJACSA.2023.0140816.

[60] A. A. Ojugo, E. Ugboh, C. C. Onochie, A. O. Eboka, M. O. Yerokun, and I. J. Iyawa, "Effects of Formative Test and Attitudinal Types on Students' Achievement in Mathematics in Nigeria," *African Educ. Res. J.*, vol. 1, no. 2, pp. 113–117, 2013, https://eric.ed.gov/?id=EJ1216962.

[61] D. R. I. M. Setiadi, K. Nugroho, A. R. Muslikh, S. W. Irianandа, and A. A. Ojugo, "Integrating SMOTE-Tomek and Fusion Learning with XGBoost Meta-Learner for Robust Diabetes Recognition," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 1, pp. 23–38, 2024, https://doi.org/10.62411/faith.2024-11.

[62] J. K. Oladele *et al.*, "BEHeDaS: A Blockchain Electronic Health Data System for Secure Medical Records Exchange," *J. Comput. Theor. Appl.*, vol. 1, no. 3, pp. 231–242, 2024, https://doi.org/10.62411/jcta.9509.

[63] M. Srividya, S. Mohanavalli, and N. Bhalaji, "Behavioral Modeling for Mental Health using Machine Learning Algorithms," *J. Med. Syst.*, vol. 42, no. 5, 2018, https://doi.org/10.1007/s10916-018-0934-5.

[64] N. C. Ashioba *et al.*, "Empirical Evidence for Rainfall Runoff in Southern Nigeria Using a Hybrid Ensemble Machine Learning Approach," *J. Adv. Math. Comput. Sci.*, vol. 12, no. 1, pp. 73–86, 2024, https://doi.org/10.22624/AIMS/MATHS/V12N1P6.

[65] C. Ren *et al.*, "Short-Term Traffic Flow Prediction: A Method of Combined Deep Learnings," *J. Adv. Transp.*, vol. 2021, pp. 1–15, 2021, https://doi.org/10.1155/2021/9928073.

[66] J. Femila Roseline, G. Naidu, V. Samuthira Pandi, S. Alamelu alias Rajasree, and D. N. Mageswari, "Autonomous credit card fraud detection using machine learning approach☆," *Comput. Electr. Eng.*, vol. 102, p. 108132, 2022, https://doi.org/10.1016/j.compeleceng.2022.108132.

[67] A. Ali *et al.*, "Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review," *Appl. Sci.*, vol. 12, no. 19, p. 9637, 2022, https://doi.org/10.3390/app12199637.

[68] N. Rtayli and N. Enneya, "Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization," *J. Inf. Secur. Appl.*, vol. 55, p. 102596, 2020, https://doi.org/10.1016/j.jisa.2020.102596.

[69] A. A. Ojugo and A. O. Eboka, "Inventory prediction and management in Nigeria using market basket analysis associative rule mining:

memetic algorithm based approach," *Int. J. Informatics Commun. Technol.*, vol. 8, no. 3, p. 128, 2019 https://doi.org/10.11591/ijict.v8i3.pp128-138.

[70] V. O. Geteloma *et al.*, "AQuaMoAS: unmasking a wireless sensor-based ensemble for air quality monitor and alert system," *Appl. Eng. Technol.*, vol. 3, no. 2, pp. 86–101, 2024, https://doi.org/10.31763/aet.v3i2.1536.

[71] F. Salahdine, Z. El Mrabet and N. Kaabouch, "Phishing Attacks Detection A Machine Learning-Based Approach," *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 0250-0255, 2021, https://doi.org/10.1109/UEMCON53757.2021.9666627.

[72] F. U. Emordi *et al.*, "TiSPHiMME: Time Series Profile Hidden Markov Ensemble in Resolving Item Location on Shelf Placement in Basket Analysis," *Digit. Innov. Contemp. Res. Sci.*, vol. 12, no. 1, pp. 33–48, 2024, https://doi.org/10.22624/AIMS/DIGITAL/V11N4P3.

[73] A. A. Ojugo and A. O. Eboka, "Empirical Evidence of Socially-Engineered Attack Menace Among Undergraduate Smartphone Users in Selected Universities in Nigeria," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 10, no. 3, pp. 2103–2108, 2021, https://doi.org/10.30534/ijatcse/2021/861032021.

[74] A. M. Almeshal, A. I. Almazrouee, M. R. Alenizi, and S. N. Alhajeri, "Forecasting the spread of COVID-19 in Kuwait using compartmental and logistic regression models," *Applied Sciences*, vol. 10, no. 10, p. 3402, 2020, https://doi.org/10.3390/app10103402.

[75] A. A. Ojugo and A. O. Eboka, "Empirical Bayesian network to improve service delivery and performance dependability on a campus network," *IAES Int. J. Artif. Intell.*, vol. 10, no. 3, p. 623, 2021, https://doi.org/10.11591/ijai.v10.i3.pp623-635.

[76] L. Shen, Y. Bao, N. Hasan, Y. Huang, X. Zhou, and C. Deng, "Intelligent crude oil price probability forecasting: Deep learning models and industry applications," *Computers in Industry*, vol. 163, p. 104150, 2024, https://doi.org/10.1016/j.compind.2024.104150.

[77] K. Deepika, M. P. S. Nagenddra, M. V. Ganesh, and N. Naresh, "Implementation of Credit Card Fraud Detection Using Random Forest Algorithm," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 3, pp. 797–804, 2022, https://doi.org/10.22214/ijraset.2022.40702.

[78] A. A. Ojugo, P. O. Ejeh, C. C. Odiakaose, A. O. Eboka, and F. U. Emordi, "Improved distribution and food safety for beef processing and management using a blockchain-tracer support framework," *Int. J. Informatics Commun. Technol.*, vol. 12, no. 3, p. 205, 2023, https://doi.org/10.11591/ijict.v12i3.pp205-213.

[79] L. De Kimpe, M. Walrave, W. Hardyns, L. Pauwels, and K. Ponnet, "You've got mail! Explaining individual differences in becoming a phishing target," *Telemat. Informatics*, vol. 35, no. 5, pp. 1277–1287, 2018, https://doi.org/10.1016/j.tele.2018.02.009.

[80] L. Loh *et al.*, "An ensembling architecture incorporating machine learning models and genetic algorithm optimization for forex trading," *FinTech*, vol. 1, no. 2, pp. 100-124, 2022, https://doi.org/10.3390/fintech1020008.

[81] A. A. Ojugo *et al.*, "Forging a learner-centric blended-learning framework via an adaptive content-based architecture," *Sci. Inf. Technol. Lett.*, vol. 4, no. 1, pp. 40–53, 2023, https://doi.org/10.31763/sitech.v4i1.1186.

[82] A. A. Ojugo and O. D. Otakore, "Computational solution of networks versus cluster grouping for social network contact recommender system," *Int. J. Informatics Commun. Technol.*, vol. 9, no. 3, p. 185, 2020, https://doi.org/10.11591/ijict.v9i3.pp185-194.

[83] A. A. Ojugo and I. P. Okobah, "Quest for an intelligent convergence solution for the well-known David, Fletcher and Powell quadratic function using supervised models," *Open Access J. Sci.*, vol. 2, no. 1, pp. 53–59, 2018, https://doi.org/10.15406/oajs.2018.02.00044.

[84] A. A. Ojugo *et al.*, "Evidence of Students' Academic Performance at the Federal College of Education Asaba Nigeria: Mining Education Data," *Knowl. Eng. Data Sci.*, vol. 6, no. 2, pp. 145–156, 2023, https://doi.org/10.17977/um018v6i22023p145-156.

[85] E. A. Otorokpo *et al.*, "DaBO-BoostE: Enhanced Data Balancing via Oversampling Technique for a Boosting Ensemble in Card-Fraud Detection," *Adv. Multidiscip. Sci. Res. J. Publ.*, vol. 12, no. 2, pp. 45–66, 2024, https://doi.org/10.22624/AIMS/MATHS/V12N2P4.

[86] A. A. Ojugo and O. Nwankwo, "Tree-classification Algorithm to Ease User Detection of Predatory Hijacked Journals: Empirical Analysis of Journal Metrics Rankings," *Int. J. Eng. Manuf.*, vol. 11, no. 4, pp. 1–9, 2021, https://doi.org/10.5815/ijem.2021.04.01.

[87] F. Jáñez-Martino, R. Alaiz-Rodríguez, V. González-Castro, E. Fidalgo, and E. Alegre, "A review of spam email detection: analysis of spammer strategies and the dataset shift problem," *Artif. Intell. Rev.*, vol. 56, no. 2, pp. 1145-1173, 2023, https://doi.org/10.1007/s10462-022-10195-4.

[88] A. A. Ojugo and O. D. Otakore, "Investigating The Unexpected Price Plummet And Volatility Rise In Energy Market: A Comparative Study of Machine Learning Approaches," *Quant. Econ. Manag. Stud.*, vol. 1, no. 3, pp. 219–229, 2020, https://doi.org/10.35877/454RI.qems12119.

[89] A. Jayatilaka, N. A. G. Arachchilage, and M. A. Babar, "Falling for Phishing: An Empirical Investigation into People's Email Response Behaviors," *arXiv preprint arXiv:2108.04766*, 2021, https://doi.org/10.48550/arXiv.2108.04766.

[90] A. A. Ojugo and A. O. Eboka, "Assessing Users Satisfaction and Experience on Academic Websites: A Case of Selected Nigerian Universities Websites," *Int. J. Inf. Technol. Comput. Sci.*, vol. 10, no. 10, pp. 53–61, 2018, https://doi.org/10.5815/ijitcs.2018.10.07.

[91] K. Afifah, I. N. Yulita, and I. Sarathan, "Sentiment Analysis on Telemedicine App Reviews using XGBoost Classifier," *2021 Int. Conf. Artif. Intell. Big Data Anal.*, pp. 22–27, 2022, https://doi.org/10.1109/ICAIBDA53487.2021.9689735.

[92] F. O. Aghware *et al.*, "Effects of Data Balancing in Diabetes Mellitus Detection : A Comparative XGBoost and Random Forest Learning Approach," *NIPES - J. Sci. Technol. Res.*, vol. 7, no. 1, pp. 1–11, 2025, https://doi.org/10.37933/nipes/7.1.2025.1.

[93] D. A. Obasuyi *et al.*, "NiCuSBlockIoT: Sensor-based Cargo Assets Management and Traceability Blockchain Support for Nigerian Custom Services," *Adv. Multidiscip. Sci. Res. J. Publ.*, vol. 15, no. 2, pp. 45–64, 2024, https://doi.org/10.22624/AIMS/CISDI/V15N2P4.

[94] A. N. Safriandono, D. R. I. M. Setiadi, A. Dahlan, F. Z. Rahmanti, I. S. Wibisono, and A. A. Ojugo, "Analyzing Quantum Feature Engineering and Balancing Strategies Effect on Liver Disease Classification," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 1, pp. 51–63, 2024, https://doi.org/10.62411/faith.2024-12.

[95] A. A. Ojugo and O. Nwankwo, "Multi-Agent Bayesian Framework For Parametric Selection In The Detection And Diagnosis of Tuberculosis Contagion In Nigeria," *JINAV J. Inf. Vis.*, vol. 2, no. 2, pp. 69–76, 2021, https://doi.org/10.35877/454RI.jinav375.

[96] A. A. Ojugo and A. O. Eboka, "Modeling the Computational Solution of Market Basket Associative Rule Mining Approaches Using Deep Neural Network," *Digit. Technol.*, vol. 3, no. 1, pp. 1–8, 2018, https://doi.org/10.11591/ijict.v8i3.pp128-138.

[97] R. G. Bhati, "A Survey on Sentiment Analysis Algorithms and Datasets," *Rev. Comput. Eng. Res.*, vol. 6, no. 2, pp. 84–91, 2019, https://doi.org/10.18488/journal.76.2019.62.84.91.

[98] A. R. Muslikh, D. R. I. M. Setiadi, and A. A. Ojugo, "Rice disease recognition using transfer xception convolution neural network," *J. Tek. Inform.*, vol. 4, no. 6, pp. 1541–1547, 2023, https://doi.org/10.52436/1.jutif.2023.4.6.1529.

[99] U. R. Wemembu, E. O. Okonta, A. A. Ojugo, and I. L. Okonta, "A Framework for Effective Software Monitoring in Project Management," *West African J. Ind. Acad. Res.*, vol. 10, no. 1, pp. 102–115, 2014, https://www.ajol.info/index.php/wajiar/article/view/105798.

[100] S. Paliwal, A. K. Mishra, R. K. Mishra, N. Nawaz, and M. Senthilkumar, "XGBRS Framework Integrated with Word2Vec Sentiment Analysis for Augmented Drug Recommendation," *Comput. Mater. Contin.*, vol. 72, no. 3, pp. 5345–5362, 2022, https://doi.org/10.32604/cmc.2022.025858.

[101] S. Carta, G. Fenu, D. R. Recupero, and R. Saia, "Fraud detection for E-commerce transactions by employing a prudential Multiple Consensus model," *Journal of Information Security and Applications*, vol. 46, pp. 13-22, 2019, https://doi.org/10.1016/j.jisa.2019.02.007.

[102] E. O. Okonta, A. A. Ojugo, U. R. Wemembu, and D. Ajani, "Embedding Quality Function Deployment In Software Development: A Novel Approach," *West African J. Ind. Acad. Res.*, vol. 6, no. 1, pp. 50–64, 2013, https://www.ajol.info/index.php/wajiar/article/view/87437.

[103] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, "A Comparative Analysis of XGBoost," *Artificial Intelligence Review*, *54*, 1937-1967, 2021, https://doi.org/10.1007/s10462-020-09896-5.

[104]H. J. V. L and D. Rajan, "Enhancing Customer Experience and Sales Performance in a Retail Store Using Association Rule Mining and Market Basket Analysis," *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1-5, 2023, https://doi.org/10.1109/ICCCNT56998.2023.10307411.

[105]C. C. Odiakaose *et al.*, "Hypertension Detection via Tree-Based Stack Ensemble with SMOTE-Tomek Data Balance and XGBoost Meta-Learner," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 3, pp. 269–283, 2024, https://doi.org/10.62411/faith.3048-3719-43.

[106]C. C. Odiakaose *et al.*, "Hybrid Genetic Algorithm Trained Bayesian Ensemble for Short Messages Spam Detection," *J. Adv. Math. Comput. Sci.*, vol. 12, no. 1, pp. 37–52, 2024, https://doi.org/10.22624/AIMS/MATHS/V12N1P4.

[107]E. O. Okonta, U. R. Wemembu, A. A. Ojugo, and D. Ajani, "Deploying Java Platform to Design A Framework of Protective Shield for Anti–Reversing Engineering," *West African J. Ind. Acad. Res.*, vol. 10, no. 1, pp. 50–64, 2014, https://www.ajol.info/index.php/wajiar/article/view/105790.

[108]B. O. Malasowe, D. V. Ojie, A. A. Ojugo, and M. D. Okpor, "Co-infection prevalence of Covid-19 underlying tuberculosis disease using a susceptible infect clustering Bayes Network," *Dutse J. Pure Appl. Sci.*, vol. 10, no. 2a, pp. 80–94, 2024, https://doi.org/10.4314/dujopas.v10i2a.8.

[109]R. Gangula, C. Sudha, K. Sreeveda, R. Bonagiri, B. C and S. Saturi, "Prediction and Prognosis of Diabetes Using Logistic Regression," *2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon)*, pp. 1-7, 2022, https://doi.org/10.1109/NKCon56289.2022.10126692.

[110]A. A. Ojugo and E. O. Ekurume, "Deep Learning Network Anomaly-Based Intrusion Detection Ensemble For Predictive Intelligence To Curb Malicious Connections: An Empirical Evidence," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 10, no. 3, pp. 2090–2102, 2021, https://doi.org/10.30534/ijatcse/2021/851032021.

[111]B. O. Malasowe, A. E. Okpako, M. D. Okpor, P. O. Ejeh, A. A. Ojugo, and R. E. Ako, "FePARM: The Frequency-Patterned Associative Rule Mining Framework on Consumer Purchasing-Pattern for Online Shops," *Adv. Multidiscip. Sci. Res. J. Publ.*, vol. 15, no. 2, pp. 15–28, 2024, https://doi.org/10.22624/AIMS/CISDI/V15N2P2-1.

[112]C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, no. 1, 2019, https://doi.org/10.1186/s40537-019-0197-0.

[113]A. A. Ojugo, A. O. Eboka, M. O. Yerokun, I. J. Iyawa, and R. E. Yoro, "Cryptography: Salvaging Exploitations against Data Integrity," *Am. J. Networks Commun.*, vol. 2, no. 2, p. 47, 2013, https://doi.org/10.11648/j.ajnc.20130202.14.

[114]M. Begum H., D. A. Janeera and A. Kumar. A.G, "Internet of Things based Wild Animal Infringement Identification, Diversion and Alert System," *2020 International Conference on Inventive Computation Technologies (ICICT)*, pp. 801-805, 2020, https://doi.org/10.1109/ICICT48043.2020.9112433.

[115]R. E. Ako *et al.*, "Effects of Data Resampling on Predicting Customer Churn via a Comparative Tree-based Random Forest and XGBoost," *J. Comput. Theor. Appl.*, vol. 2, no. 1, pp. 86–101, 2024, https://doi.org/10.62411/jcta.10562.

[116]F. O. Aghware *et al.*, "Enhancing the Random Forest Model via Synthetic Minority Oversampling Technique for Credit-Card Fraud Detection," *J. Comput. Theor. Appl.*, vol. 1, no. 4, pp. 407–420, 2024, https://doi.org/10.62411/jcta.10323.

[117]D. Nguyen *et al.*, "Adaptive Evaluation of LQR Control using Particle Swarm Optimization for Pendubot," *J. Fuzzy Syst. Control*, vol. 2, no. 2, pp. 58–66, 2024, https://doi.org/10.59247/jfsc.v2i2.203.

[118]S. A. Chowdhury and S. Aziz, "Financing Renewable Energy and Fossil Fuel Power Plants in Bangladesh: A Comparative Analysis," *2024 7th International Conference on Development in Renewable Energy Technology (ICDRET)*, pp. 1-6, 2024, https://doi.org/10.1109/ICDRET60388.2024.10503983.

[119]Z. Zhu, J. Peng, K. Liu, and X. Zhang, "A game-based resource pricing and allocation mechanism for profit maximization in cloud computing. *Soft Computing*, vol. 24, pp. 4191-4203, 2020, https://doi.org/10.1007/s00500-019-04183-0.

[120]A. A. Ojugo, R. E. Yoro, A. O. Eboka, M. O. Yerokun, C. N. Anujeonye, and F. N. Efozia, "Predicting Behavioural Evolution on a Graph-Based Model," *Adv. Networks*, vol. 3, no. 2, p. 8, 2015, https://doi.org/10.11648/j.net.20150302.11.

[121]A. Panthakkan, N. Valappil, M. Appathil, S. Verma, W. Mansoor and H. Al-Ahmad, "Performance Comparison of Credit Card Fraud Detection System using Machine Learning," *2022 5th International Conference on Signal Processing and Information Security (ICSPIS)*, pp. 17-21, 2022, https://doi.org/10.1109/ICSPIS57063.2022.10002517.

[122]M. S. Bhadriraju and K. Dasari, "SSDP DDoS Attacks Detection using Naïve Bayes Classifiers with Wrapper Feature Selection Methods," *2024 3rd Edition of IEEE Delhi Section Flagship Conference (DELCON)*, pp. 1-5, 2024, https://doi.org/10.1109/DELCON64804.2024.10866135.

[123]M. I. Akazue *et al.*, "Handling Transactional Data Features via Associative Rule Mining for Mobile Online Shopping Platforms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 3, pp. 530–538, 2024, https://doi.org/10.14569/IJACSA.2024.0150354.

[124]M. Rele and D. Patil, "Intrusive Detection Techniques Utilizing Machine Learning, Deep Learning, and Anomaly-based Approaches," *2023 IEEE International Conference on Cryptography, Informatics, and Cybersecurity (ICoCICs)*, pp. 88-93, 2023, https://doi.org/10.1109/ICoCICs58778.2023.10276955.

[125]C. Zoremsanga and J. Hussain, "Particle Swarm Optimized Deep Learning Models for Rainfall Prediction: A Case Study in Aizawl, Mizoram," in *IEEE Access*, vol. 12, pp. 57172-57184, 2024, https://doi.org/10.1109/ACCESS.2024.3390781.

[126]O. B. Chibuzo and D. O. Isiaka, "Design and Implementation of Secure Browser for Computer-Based Tests," *Int. J. Innov. Sci. Res. Technol.*, vol. 5, no. 8, pp. 1347–1356, 2020, https://doi.org/10.38124/IJISRT20AUG526.

[127]A. A. Ojugo *et al.*, "CoSoGMIR: A Social Graph Contagion Diffusion Framework using the Movement-Interaction-Return Technique," *J. Comput. Theor. Appl.*, vol. 1, no. 2, pp. 37–47, 2023, https://doi.org/10.33633/jcta.v1i2.9355.

[128]A. A. Ojugo, P. O. Ejeh, C. C. Odiakaose, A. O. Eboka, and F. U. Emordi, "Predicting rainfall runoff in Southern Nigeria using a fused hybrid deep learning ensemble," *Int. J. Informatics Commun. Technol.*, vol. 13, no. 1, p. 108, 2024, https://doi.org/10.11591/ijict.v13i1.pp108-115.

[129]A. A. Ojugo and O. D. Otakore, "Seeking Intelligent Convergence for Asymptotic Stability Features of the Prey / Predator Retarded Equation Model Using Supervised Models," *Comput. Inf. Syst. Dev. Informatics Allied Res. J.*, vol. 9, no. 2, pp. 13–26, 2018, https://doi.org/10.15406/oajs.2018.02.00044.

[130]A. A. Ojugo and R. E. Yoro, "Migration Pattern As Threshold Parameter In The Propagation of The Covid-19 Epidemic Using An Actor-Based Model for SI-Social Graph," *JINAV J. Inf. Vis.*, vol. 2, no. 2, pp. 93–105, 2021, https://doi.org/10.35877/454RI.jinav379.

[131]A. A. Ojugo and A. O. Eboka, "An Empirical Evaluation On Comparative Machine Learning Techniques For Detection of The Distributed Denial of Service (DDoS) Attacks," *J. Appl. Sci. Eng. Technol. Educ.*, vol. 2, no. 1, pp. 18–27, 2020, https://doi.org/10.35877/454RI.asci2192.

[132]A. O. Eboka *et al.*, "Pilot study on deploying a wireless sensor-based virtual-key access and lock system for home and industrial frontiers," *Int. J. Informatics Commun. Technol.*, vol. 14, no. 1, p. 287, 2025, https://doi.org/10.11591/ijict.v14i1.pp287-297.

[133]D. R. I. M. Setiadi, A. Susanto, K. Nugroho, A. R. Muslikh, A. A. Ojugo, and H. Gan, "Rice yield forecasting using hybrid quantum deep learning model," *MDPI Comput.*, vol. 13, no. 191, pp. 1–18, 2024, https://doi.org/10.3390/computers13080191.

[134]S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, "Random forest for credit card fraud detection," in *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, pp. 1–6, 2018, https://doi.org/10.1109/ICNSC.2018.8361343.