Creative Research Publishers
https://www.crossref.org/06members/50go-live.html

Journal of Behavioral Informatics
Digital Humanities & Development Research
Vol 10, No 1, March, 2024 Series

Quest for Empirical Solution to Runoff Prediction in Nigeria via Random Forest Ensemble: Pilot Study

Malasowe, B.O.[1], Edim, E.B.[2], Adigwe, W.[3], Okpor, M.D.[4], Ako, R.E[5], Okpako, A.E[6], Ojugo, A.A.[7] & Ojei, E.O.[8]
[1]Department of Computer Science, University of Delta, Agbor, Nigeria.
[2]Department of Computer Science, University of Calabar, Nigeria.
[3,4&8]Dept of Computer Science, Delta State University of Science and Technology, Ozoro, Nigeria.
[5&7]Dept of Computer Science, Federal University of Petroleum Resources, Effurun, Nigeria.
6Department of Cyber Security, University of Delta, Agbor, Nigeria.
E-mails: bridget.malasowe@unidel.edu.ng[1], etimb@unical.edu.ng[2], adigwew@dsust.edu.ng[3], okpormd@dsust.edu.ng[4], ako.rita@fupre.edu.ng[5], ejaita.okpako@unidel.edu.ng[6], ojugo.arnold@fupre.edu.ng[7], ojeieo@dsust.edu.ng[8]

ABSTRACT

Environmental factors and features often change to result in either harsh weather conditions or rainfall, which often calms the weather as well as provides fast, significant downstream hydrology known as runoff with a variety of implications such as erosion, water quality, and infrastructures. These, in turn, impact the quality of life, sewage systems, agriculture, and tourism of a nation, to mention a few. It chaotic, complex and dynamic nature has necessitated studies in the quest for future direction of such runoff via prediction models. With little successes in use of knowledge driven models – many studies have now turned to data-driven models. Dataset is retrieved from Metrological Center in Lagos, Nigeria for the period 1999–2023. The retrieved dataset was split: 70% for train dataset, and 30% for test dataset. Our study used the Random Forest ensemble. Result yields a sensitivity of 0.9, specificity 0.19, accuracy of 0.74, and improvement rate of 0.12. Other ensembles underperformed as compared to proposed model. Study reveals annual rainfall is an effect of variation cycle. Models will help simulate future floods and provide, lead time warnings in flood management.

Keywords: Runoff, Random Forest, flood resources, hydrology management, Nigeria

Journal Reference Format:
Malasowe, B.O, Edim, E.B, Adigwe, W, Okpor, M.D., Ako, R.E, Okpako, A.E, Ojugo, A.A. & Ojei, E.O. (2024): Quest for Empirical Solution to Runoff Prediction in Nigeria via Random Forest Ensemble: Pilot Study. Journal of Behavioural Informatics, Digital Humanities and Development Rese Vol. 10 No. 1. Pp 73-90 . dx.doi.org/10.22624/AIMS/BHI/V10N1P8
https://www.isteams.net/behavioralinformaticsjournal

1. INTRODUCTION

Prediction cum forecast of rainfall can help refocus, reshape and repurpose a society toward agriculture, economics and other advancements [1]. The forecast of rainfall seeks to explore and use rainfall runoff and its associated feats to estimate rain-runoff quantification [2].

The increased awareness and interest, and the dynamic nature of environmental feats of the coverage area [3] – have continued to yield the added impetus for the continual study to model hydrological processes [4]. Such new models are today, tasked with innovative solutions to resolve challenges and meet the rising new requirements [5] that will effectively deal with associated feats in runoff simulation expressed as erosion, land degradation, runoff resources management, leaching, land-use consequences, and climate changes [6], [7]. Rainfall runoff as a classification task has become critical especially in Nigeria [8]. Reasons for this may include (not limited to): (a) the deluge and tsunami from 2015 to 2018 – resulting in the displacement of residents in Southern Nigeria [9], (b) the adoption of runoff as an imperative component in the planning processes vis-à-vis the execution of farming directives cum initiatives at both the local and federal government strata [10], [11], (c) use of runoff models in the estimation of runoff application in farming etc [12], and (d) use of rain-runoff prediction for pollutant leaching, dissolution of chemical processes for land-degradation [13] and as modes for management of rainfall runoff resources [14].

Thus, it is today – both critical and imperative to design rain-runoff ensembles that will adequately forecast rainfall features such as humidity, sunshine, runoff etc – all of which is possible via the use of mathematical models and heuristics that can are grouped into knowledge-driven and data-driven models [15]. These have been useful in the provision of early warning with runoff situations, which in turn – is critical in water resources management [16]. The inherent dynamic, complex and chaotic nature of environmental conditions cum atmospheric processes that yields rainfall – makes runoff prediction modeling a tedious task [17]. Even with the numerous advances in weather forecast, an accurate prediction of runoff stream remains challenging due to its significance and influence in downstream hydrology and water resources management [18]. These, often ripple across the society – a range of implications on flood cum runoff as erosion, water quality, and the designs for both residential and industrial structure(s) [19]. These, in turn – also impacts on the quality of life, agriculture, sewage system, and tourism etc [20].

The birth and use of machine learning (ML) to successfully model and train inherent selected features – so that the model can effectively recognize a variety of domain task patterns [21] and yield a low-cost optimal solution [22]. These, are explored to learn the underlying features of interest via feature selection and extraction for either classification [23] and regression [24] tasks. With training, the model can effectively detect anomalies or unusual activities in its use as profiled patterns [25]. A variety of ML models that have been successfully used in a variety of endeavors and task to include: Logistic Regression [26], [27], Deep Learning [28], [29], Bayesian model [30], Naive Bayes [31], Support Vector Machine [32], [33], K-Nearest Neighbors [34], [35], Random Forest [36], [37], and others [38], [39]. Its inherent drawback is with their choice and flexibility of feature selection [40] and importance in relation to the target parameters for the discovery of ground-truth [41], [42]. Thus, our study adopts a Random Forest (RF) with synthetic minority oversampling feature selection techniques used on the Kaggle dataset. Our choice for RF is due to its ability to reduce overfitting, to address imbalanced datasets, and yield a vigorous prediction accuracy [43]–[45]. Our study explores use of a tree-based Random Forest ensemble to forecast rainfall runoff in downstream hydrological operations especially in Delta State as a point of focus. Range of complications present with hydrology modeling – are challenges that the study wishes to address [46].

Section I introduces the study with a view to unveiling the meaning of rainfall runoff and others. Section II details the related literatures, problem formulation, data gathering and proposed ensemble. Section III details result found as evidence to support the decision during discussions of the findings, and conclusion.

2. METHODS AND MATERIALS

2.1 Literature Review

Studies and many research have continued in their quest to develop stochastic models that yield cost-effective solutions to rainfall runoff modeling. [47] investigated the adoption of the TOPModel (a knowledge-driven ensemble used at the Benin-Owena Rivar Basin Development Agency with dataset retrieved from the Metrological centre at Oshodi in Lagos State, Nigeria. having noted the inherent drawbacks – modeled a gravitational search trained neural network algorithm (HGANN) that sought to enhance accurate rainfall prediction(s). Results showed high accuracy with its COE as 58%, 24%, 56% and 42% respectively for the various stations. Extended by [48] – it observed annual rainfall variations from long-term runoff, is an effect of variation cycle with significant correlation between rainfall and runoff – as indicated in the adopted cum adapted dataset [49]. The study effectively simulated a range of future runoff values – providing lead time warning especially with flood cum runoff resources management. This have also noted to be used to effectively manage irrigation for smart and precision agriculture.

[50] used ANN to investigate runoff in India just as [51] did same in Cyprus. Their efforts agreed with [52]. Also, efforts are intensified with the adoption of autoregressive moving average (ARMA) model with an exogenous variable (ARMAX) used to investigate hydrological data. Justified with its results as merely theoretical, [53] investigated runoff using the 3-function AR(3) model – which yielded significant association to establish that rainfall significantly impacts relative humidity, cloud cover and temperature difference.

They further noted that sunshine was not selected as a feat of interest due to the resulting impulse response functions. They identified TF(3,2,2,2) as the 4-TF models that predicted rainfall with a root mean square error of 0.023 as the most appropriate function – noting that model outperformed the multiple regression/univariate SARIMA$(1,0,1)*(1,0,1)_{12}$ model.

Also, [54] compared a hybrid gravitational search algorithm trained neural network with historical data for the Chad River Basin in Nigeria using dataset from 1996 - 2007. It developed a GARCH model to forecast rainfall using historical data from National Metrological Centre at Oshodi. Dataset used rainfall, temperature difference, relative humidity, sunshine and cloud cover to establish the significant association therein for a variety of parameters vis-à-vis ground-truth. It successfully predicted runoff with a mean square error of 0.012 as the most appropriate.

3. DATA GATHERING/SAMPLE POPULATION

The dataset was retrieved from the Nigerian Metrological Centre Oshodi in Lagos State of Nigeria. The dataset consist of the following features: (a) year, (b) mean rain, (c) temperature, (d) relative humidity, (e) mean sunshine, (f) mean windspeed, and (g) wind direction for the period under coverage (i.e. 1999-2022). Delta State has a land-mass of 22045km^2, with an annual mean rain of 1354mm with perennial discharge of 3.8m/1.5m^3/s for its dry/peak periods respectively. Figure 1 shows coverage area under study; while, the Figure 2 time-plot for the period; and Table 1 details dataset description with its various features.

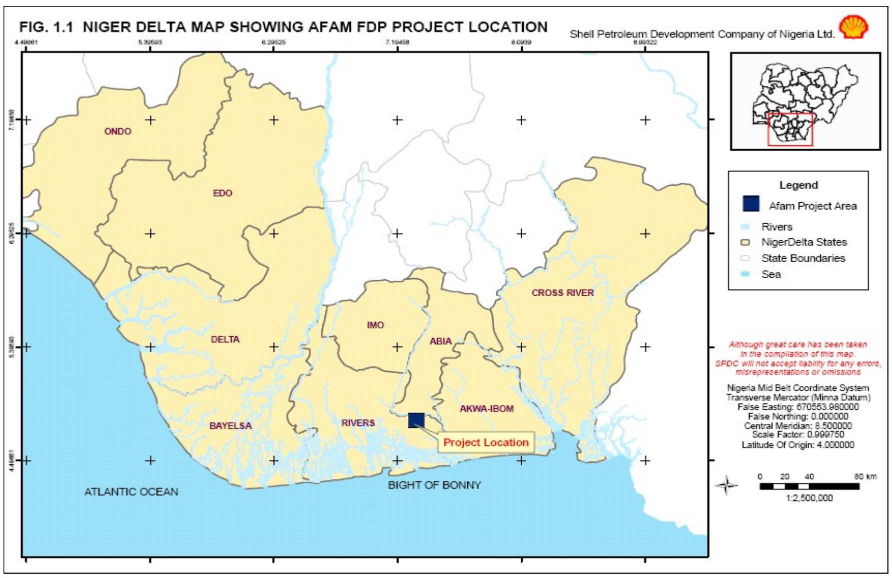

Figure 1. Landmass of the geographical area considered

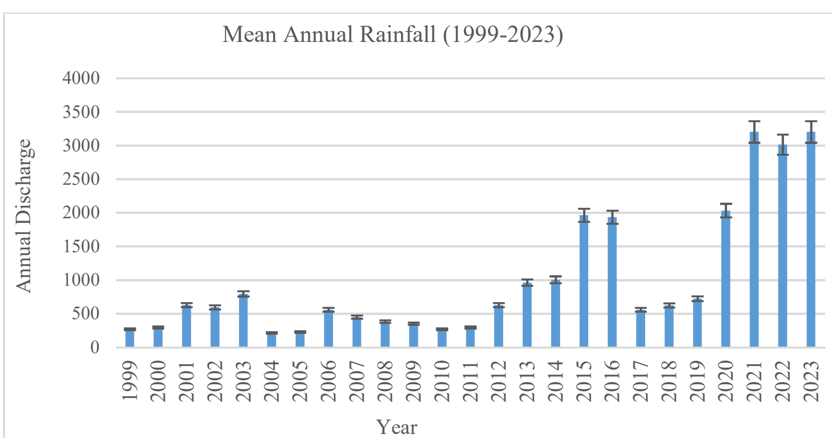

Figure 2. Clustered time plot of Annual Rain Discharge

TABLE I. DATASET DESCRIPTION, DATA TYPES, AND FORMAT

Year	Rain	Temp	Mean Humidity	Mean Sunshine	Wind-Speed	Wind Direction
1999	271.4	31.57	78.901	3.256	2.902	SW
2000	295.1	32.10	76.902	3.761	3.508	S
2001	628.9	31.53	83.000	3.021	2.892	W
2002	594.4	32.02	85.200	2.994	2.858	SW
2003	795.7	31.58	83.134	5.012	2.917	W
2004	216.4	31.73	79.013	4.561	3.375	SW
2005	229.4	31.57	85.301	4.092	2.935	SW
2006	558.8	32.12	79.34	4.432	3.451	SW
2007	449.6	31.92	81.211	3.895	3.209	S
2008	383.4	32.04	83.120	4.501	3.021	S
2009	351.7	31.58	83.753	4.458	3.508	NE
2010	271.4	31.73	83.917	5.067	2.892	W
2011	295.1	31.57	83.751	4.433	2.858	SW
2012	628.0	32.12	83.667	3.850	2.917	S
2013	963.0	31.53	83.667	4.042	3.375	SW
2014	1005.0	32.17	83.583	3.883	3.733	SW
2015	1963.1	31.58	81.501	2.933	3.3	S
2016	1934.1	32.42	84.751	4.358	3.058	SW
2017	558.8	32.17	85.167	4.001	2.825	S
2018	623.9	32.42	83.001	4.158	2.983	S
2019	723.1	31.58	81.333	4.575	3.15	W
2020	2031.8	32.86	85.231	2.994	2.858	SW
2021	3201.9	32.92	84.909	3.895	3.218	SE
2022	3012.3	33.01	85.200	2.994	2.858	SW
2023	3201.8	31.72	89.342	4.432	3.451	SW

Proposed Random Forest (RF) Ensemble

RF is a widely-used, tree-based, supervised ML heuristics – which achieves accuracy by successfully combining as output, the multiple majority voting of weak decision trees to yield a single outcome. Its flexibility have necessitated its adoption in both classification and regression tasks [55]. It is constructed from the various decision trees as in figure 2. With same nodes, and different inputs to yield distinct leaves – it uses labeled data and a voting scheme that assumes all its base classifiers have the same weight. Due to randomized bootstrap sampling, some trees will relatively yield higher weights, and the selected attribute(s) cannot guarantee that all trees will yield the same ability to make decisions. Thus, it mitigates model overfit, poor generalization, and handle(s) complex continuous/categorical datasets (in both regression and classification tasks) [56] – by leveraging on the decisions of many weak trees/learners to yield a single stronger learner [57], [58].

Creative Research Publishers
https://www.crossref.org/06members/50go-live.html
Journal of Behavioral Informatics
Digital Humanities & Development Research
Vol 10, No 1, March, 2024 Series

The steps involved includes [59]: (a) first, we split the original dataset (into subsets for training and testing) using row sampling and feature sampling – so that each partition consists of selected rows/columns with replacements, (b) second, we create individual decision tree for each subset selected and assigned, (c) third, each decision tree will yield an output, and (d) finally, our ensemble will use the majority voting scheme to yield its final outcome. Figure 3 depicts the structural flow diagram of the Random Forest Tree-based heuristic with prediction style and voting mechanism to yield either the classification or regression task output.

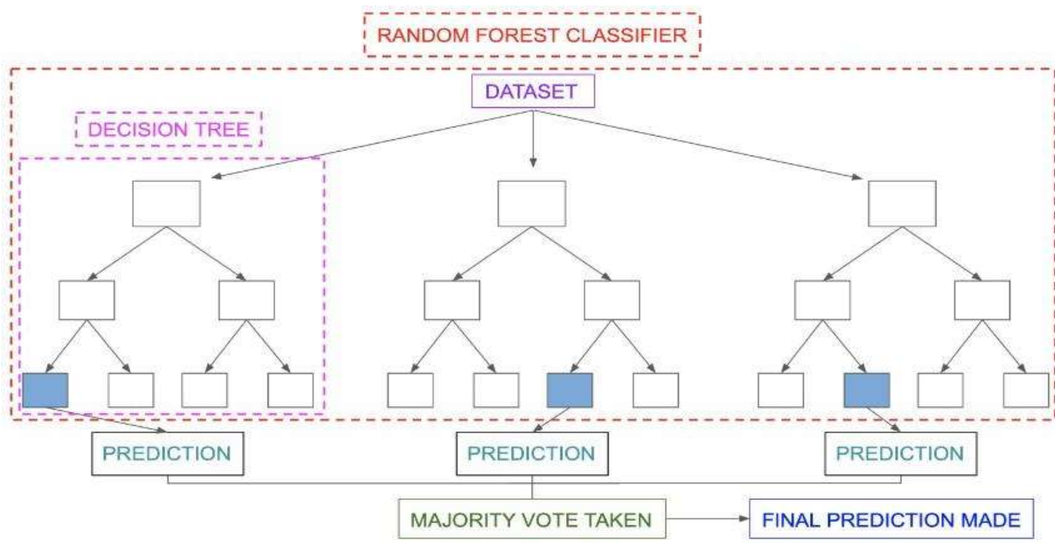

Figure 3. Depicts the Random Forest Classifier

The dynamic, stochastic, complex, chaotic and non-linear nature of runoff implies that we may experience an incomplete dataset as well as in its unstructured form. Thus, we explore the use of Synthetic Minority Over-sampling Technique (SMOTE) to fasten data pre-processing and cleaning of sampled dataset. SMOTE is a resampling strategy that creates artificial instances of a target class (i.e. runoff) to resolve all imbalance. It uses the oversampling scheme to generate amended data points for the underlying features of interest (prior the application of feature selection and extraction for the construction and training of the model). This aims to help balance target representation.

We use SMOTE to: (a) identify minority data-points in the original dataset, (b) it select instances of minority representations in the data point, adjusting the number of its closest neighbors, (c) it then interpolates data point ranges to yield an amended dataset that can be adopted for training with instances and its chosen neighbors to create synthetic instances (i.e., added data-points that links all generated data-points using the minority instances to its closest neighbors), (d) it adds the synthetic instances to the dataset – to yield an oversampled dataset with balanced picture of both classes, and (e) it splits dataset into train and test as used in the construction, and generalization to assess the ensemble.

Creative Research Publishers
https://www.crossref.org/06members/50go-live.html
Journal of Behavioral Informatics
Digital Humanities & Development Research
Vol 10, No 1, March, 2024 Series

Some benefits for applying SMOTE includes: (a) prevents bias and skewness with imbalanced dataset that normally can distort model's prediction, (b) it enhances an ensemble's performance via balanced datasets as ensemble can adequately learn features and patterns from all classes even with majority or minority voting with the balanced dataset as well as detect anomalies during testing, and (c) the characteristics linked to the majority class often have a greater significance than other features in an unbalanced dataset – so that by balancing the dataset, the model is better able to understand the significance of each feature for every class, yielding more insightful results.

Experimental Ensemble Training

Some reasons for choosing RF includes: (a) ensemble learning that allows it to leverage the decision of many weak learners fused into a single strong classifier, (b) its ability to handle complex dataset, (c) its decreased risk in poor generalization and overfitting of model, (d) its capability to understand the relative contribution of various features to prediction, especially when attempting to identify fraudulent activities, and (e) its resilience to noise especially in real-world applications where dataset is often unstructured and there are no ground truths. Using the dataset produced via SMOTE, the Random Forest model was trained as follows:

1. Data Splitting: The dataset was divided into training and testing sets once it had been balanced using SMOTE. By using the training set just for model training [60], the oversampled data allowed the Random Forest algorithm to identify patterns. Conversely, the testing set, which consisted of hypothetical cases, functioned as a specific assessment subset, enabling a thorough examination of the model's ability to identify credit card fraud. This division made sure that the trained model had a strong framework for assessment, which enhanced its usefulness in practical situations as in Figure 2 and 3 respectively.
2. Model Initialization: The default hyperparameters were used to initialize the Random Forest model. During this, no hyperparameter adjustment was done [61]. This is because the experimental RF-ensemble remains unaffected by, and is less susceptible to hyper-parameters tuning as with other models/heuristics. An acceptable results can be obtained via its default configurations [62].
3. Feature Selection: As a pre-processing step – FS seeks to select features in relations to the target variable. We adopt the filter scheme to ascertain how relevant a selected feats is, in support to the output via statistical test [63]. We use chi-square to test if the occurrence of a selected feature correlates to target (runoff) class [64] using their frequency distribution. Thus, FS extracts only those parametric feats that highly correlates with the output-class [65]. For this section, we use the Python sklearn (and set as 0 the value for features that have no mutual information; and set as 1 for those that does correlate) of the chosen feats in relation with the target feature/class. All feats are ranked by chi-squared using the threshold value as in Eq.1 [66].

$$X = \frac{\sum x_i}{n} \quad (1)$$

Creative Research Publishers
https://www.crossref.org/06members/50go-live.html
Journal of Behavioral Informatics
Digital Humanities & Development Research
Vol 10, No 1, March, 2024 Series

A total of 22-feats was contained [67] therein the original dataset. With FS, only six(6) features were selected and used to extract from the original dataset. Using chi-square approach, we compute the threshold value using Eq. 1 for each attribute to yield the scores, in lieu of each attribute's correlation with the target class 1 as in the table 2. With the computed threshold of 2.1874, a total of six (6) feats were selected. These were examined to help us gain insights into the contribution of different features to the classification process [68].

4. Training: The RF ensemble learns from scratch via a pre-designated training set, and is expanded to include both the original and artificial data-points via SMOTE [69]. We use iterative construction to create several decision trees that results in our ensemble [70]. Each tree is then trained using bootstrap sampling to yield resampled subset [71] from the enhanced train dataset). The trees' collective knowledge is enhanced by this iterative process, which in turn also helps it to identify the intricate patterns present in each record [72]. Training set is thus, a blend of both the synthetic and actual samples guaranteed of RF-ensemble comprehensive learning experience [73], and by extension improving its flexibility and adaptability to the various settings inside the dataset [74].

TABLE II. RANKING ATTRIBUTES SCORES USING CHI-SQUARE

Feature	Selected (Yes/No)	X^2-Value
Year	No	0.4920
Mean_Annual_Rainfall	Yes	3.0298
Max_Temperature_Difference	No	1.3029
Average_Temperature	Yes	18.006
Min_Temperature_Difference	No	1.2093
Humidity	Yes	23.092
Cloud_cover	No	0.9837
Sunshine	Yes	6.0929
Windspeed	Yes	38.389
Wind-Direction	Yes	41.902

I. Results and Findings Discussion

Performance Evaluation of the Framework

Table 3 shows confusion matrix prior applying of SMOTE, which agrees with the results therein [75], [76] in that ensemble adequately impacts outlies [77]–[80]. In addition, our proposed, experimental ensemble outperformed other benchmark models as it was best in its ability to successfully balance ensemble accuracy, recall, and precision [81]. It supports effectiveness and efficiency of the ensemble – offering a detailed perspective of ensemble's performance in differentiating between genuine positives, true negatives, false positives, and false negatives.

Creative Research Publishers
https://www.crossref.org/06members/50go-live.html

Journal of Behavioral Informatics
Digital Humanities & Development Research
Vol 10, No 1, March, 2024 Series

TABLE III. ENSEMBLE RESULT BEFORE FEATURE SELECTION

Heuristics	F1	Accuracy	Precision	Recall
Logistic Regression	92.19	97.18	93.57	95.82
KNN	94.35	77.47	92.64	66.57
Naïve Bayes	95.08	83.03	83.62	82.45
Support Vector Machine	90.08	50.00	94.57	33.98
Random Forest	98.02	98.02	96.89	99.01

Table 3 shows the proposed ensemble's performance prior to its application of the features selection technique. The results show that the Random Forest approach ensemble outperforms other models yielding an accuracy of 98.02%, and a F1-score of 98.02%.

TABLE IV. ENSEMBLE RESULT AFTER FEATURE SELECTION APPLIED

Heuristics	F1	Accuracy	Precision	Recall
Logistic Regression	98.05	98.05	98.05	98.05
KNN	92.10	92.28	90.18	94.48
Naïve Bayes	91.25	90.74	96.16	85.90
Support Vector Machine	81.45	80.32	85.41	75.81
Random Forest	99.19	98.19	98.28	98.10

Our experimental ensemble was found to outperform other ensembles. Prior to the application of chi-squared FS approach, ensemble yields accuracy of 98.02%; while, other ensembles such as Logistic Regression, KNN, Naïve Bayes and Support Vector Machine respectively resulted cum yielded an accuracy of 0.9219, 0.9435, 0.9508 and 0.9008 respectively. In addition, our proposed ensemble yields an F1-score of 0.9919 with the application of chi-squared FS-approach and SMOTE; while, other ensemble (i.e. Logistic Regression, KNN, Naïve Bayes and Support Vector Machine) yielded F1-score of 0.9805, 0.9210, 0.9125 and 0.8.45 respectively.

We observed that the adaption of both the chi-square filter feature selection approach, and use of the synthetic minority oversampling technique (SMOTE) data balancing ensures that improved accuracy when compared with the results yielded in the studies [82]–[85]; This is as effectively seen in Table 4, and also in agreement with [86]–[88].

4. DISCUSSION OF FINDINGS

The adoption of chi-square, filter-based FS approach with Random Forest and the consequent application of SMOTE [89] – have successfully shown a variety of benefits to include [90]: (a) it yields fewer features with dataset balancing for use during the construction of a model as well as in its inherent training [91], [92], (b) the training time for the experimental ensemble is greatly shortened by the impact of the feature selection approach used, as it is predominantly significant for real-time prediction, where quick response times are critical for provision of early warning of flood management resources when compared with [93], [94], and (c) the ensemble excellent accuracy of 99.19% holds that adopted ensemble feature selection did not degrade its performance – as compared with [95].

The figure shows that the ensemble accurately classified and predicted runoff with a 98.19% accuracy for 1300-correctly classified instance with only 21-incorrectly classified instances for the dataset used. Thus, in reality, our ensemble holds true to have successfully minimized the false-positive errors, accurately. Our results indicates that the Random Forest ensemble can effectively be used to yield runoff prediction accuracy with data balancing technique.

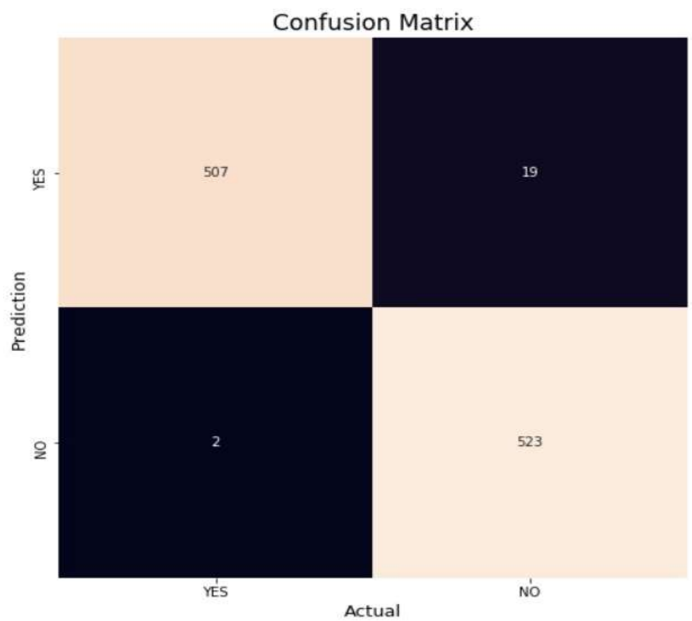

Figure 4. Ensemble Confusion Matrix

5. CONCLUSION

Ensembles are quite challenging to implement due to a variety of conflicts namely [96]: (a) that data encoding conflict from one algorithm to another within the proposed ensemble, (b) issue of the underlying features of interest generated for the candidate solution is easily resolved with tree-based algorithm [97], and (c) conflicts arising from structural dependencies imposed on the ensemble by dataset features not contained from the outset [98]. All these, ensures that the ensemble yield its optimal solution [99]. Modelers must then, select the requisite, appropriate parameter(s) to avoid overtraining and overfit of the ensemble [100].

Furthermore, the effects of such ensemble is to prevent agents within a multi-goal tasks such as this [101] – from creating and enforcing their own behavioral rules on the dataset at training [102]. Our resultant confusion matrix for the proposed, experimental ensemble yielded a sensitivity value and precision of 0.83, a specificity or recall value 0.08, prediction accuracy of 0.991, and a misclassification error rate of 0.018 for hyper-parameter tuning and data inclusion (that were not originally used) during the model's training phase [103].

REFERENCES

[1] Z. Abda, B. Zerouali, M. Chettih, C. A. Guimarães Santos, C. A. S. de Farias, and A. Elbeltagi, "Assessing machine learning models for streamflow estimation: a case study in Oued Sebaou watershed (Northern Algeria)," *Hydrol. Sci. J.*, vol. 67, no. 9, pp. 1328–1341, Jul. 2022, doi: 10.1080/02626667.2022.2083511.

[2] 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, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp623-635.

[3] S. L. James *et al.*, "Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017," *Lancet*, vol. 392, no. 10159, pp. 1789–1858, Nov. 2018, doi: 10.1016/S0140-6736(18)32279-7.

[4] A. Awotwi, G. K. Anornu, J. A. Quaye-Ballard, and T. Annor, "Monitoring land use and land cover changes due to extensive gold mining, urban expansion, and agriculture in the Pra River Basin of Ghana, 1986–2025," *L. Degrad. Dev.*, vol. 29, no. 10, pp. 3331–3343, Oct. 2018, doi: 10.1002/ldr.3093.

[5] N. N. Wijaya, D. R. I. M. Setiadi, and A. R. Muslikh, "Music-Genre Classification using Bidirectional Long Short-Term Memory and Mel-Frequency Cepstral Coefficients," *J. Comput. Theor. Appl.*, vol. 2, no. 1, pp. 13–26, 2024.

[6] A. A. Ojugo and I. P. Okobah, "Computational Solution for Modeling Rainfall Runoff Using Intelligent Stochastic Model: A Case of Warri in Delta State Nigeria," *J. Digit. Innivations Contemp. Res. Sci. Eng. Technol.*, vol. 5, no. 4, pp. 45–58, 2017, doi: 10.22624.

[7] R. E. Yoro and A. A. Ojugo, "An Intelligent Model Using Relationship in Weather Conditions to Predict Livestock-Fish Farming Yield and Production in Nigeria," *Am. J. Model. Optim.*, vol. 7, no. 2, pp. 35–41, 2019, doi: 10.12691/ajmo-7-2-1.

[8] 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, doi: 10.38124/ijisrt20aug526.

[9] 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, pp. 108–115, Apr. 2024, doi: 10.11591/ijict.v13i1.pp108-115.

[10] S. Ju, H. Lim, and J. Heo, "Machine learning approaches for crop yield prediction with MODIS and weather data," *40th Asian Conf. Remote Sensing, ACRS 2019 Prog. Remote Sens. Technol. Smart Futur.*, no. Acrs, pp. 1–4, 2020.

[11] J. A. Aper, S. A. Iorkua, and N. J. Akegh, "Relationship between Rainfall, Surface Runoff and Soil Loss on an Experimental Farm in the Middle Belt of Nigeria," *Niger. Ann. Pure Appl. Sci.*, vol. 1, pp. 122–139, 2019, doi: 10.46912/napas.116.

[12] R. Alfred, J. H. Obit, C. P.-Y. Chin, H. Haviluddin, and Y. Lim, "Towards Paddy Rice Smart Farming: A Review on Big Data, Machine Learning, and Rice Production Tasks," *IEEE Access*, vol. 9, pp. 50358–50380, 2021, doi: 10.1109/ACCESS.2021.3069449.

[13] O. A. Obiora-Okeke, J. R. Adewumi, and O. M. Ojo, “Impact of Climate Change on Runoff Prediction in Ogbese River Watershed,” *FUOYE J. Eng. Technol.*, vol. 6, no. 4, pp. 392–396, 2021, doi: 10.46792/fuoyejet.v6i4.721.
[14] H. N. Najm, “Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics,” *Annu. Rev. Fluid Mech.*, vol. 41, no. 1, pp. 35–52, Jan. 2009, doi: 10.1146/annurev.fluid.010908.165248.
[15] A. A. Ojugo and O. Nwankwo, “Forging a Spectral-Clustering Multi-Agent Hybrid Deep Learning Model To Predict Rainfall Runoff In Nigeria,” *Int. J. Innov. Sci. Eng. Technol.*, vol. 8, no. 3, pp. 140–147, 2021.
[16] A. . Adegede, K. . Ogbu, V. Ogwo, and C. . Mbajiorgu, “Comparison of ACRU and HEC-HMS models in runoff prediction in a watershed, Southern Nigeria,” *J. Agric. Eng. Technol.*, vol. 22, no. 4, pp. 142–149, 2014.
[17] A. Oludapo Olusola, O. Olumide, O. Adeola Fashae, and S. Adelabu, “River sensing: the inclusion of red band in predicting reach-scale types using machine learning algorithms,” *Hydrol. Sci. J.*, vol. 67, no. 11, pp. 1740–1754, Aug. 2022, doi: 10.1080/02626667.2022.2098752.
[18] H. Chen, S. Guo, C. Xu, and V. P. Singh, “Historical temporal trends of hydro-climatic variables and runoff response to climate variability and their relevance in water resource management in the Hanjiang basin,” *J. Hydrol.*, vol. 344, no. 3–4, pp. 171–184, Oct. 2007, doi: 10.1016/j.jhydrol.2007.06.034.
[19] A. A. Ojugo, A. O. Eboka, R. E. Yoro, M. O. Yerokun, and F. N. Efozia, “Framework design for statistical fraud detection,” *Math. Comput. Sci. Eng. Ser.*, vol. 50, pp. 176–182, 2015.
[20] P. Panjala, M. K. Gumma, H. A. Ajeigbe, M. M. Badamasi, K. C. Deevi, and R. Tabo, “Identifying Suitable Watersheds across Nigeria Using Biophysical Parameters and Machine Learning Algorithms for Agri–Planning,” *ISPRS Int. J. Geo-Information*, vol. 11, no. 8, 2022, doi: 10.3390/ijgi11080416.
[21] F. O. Aghware, R. E. Yoro, P. O. Ejeh, C. C. Odiakaose, F. U. Emordi, and A. A. Ojugo, “Sentiment analysis in detecting sophistication and degradation cues in malicious web contents,” *Kongzhi yu Juece/Control Decis.*, vol. 38, no. 01, p. 653, 2023.
[22] 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. 2, no. 2, pp. 190–203, 2024, doi: 10.62411/jcta.10323.
[23] 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, doi: 10.1007/978-981-33-6893-4_20.
[24] 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, doi: 10.1186/s40537-021-00541-8.
[25] 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. 2, no. 1, pp. 126–137, 2024, doi: 10.62411/jcta.9539.
[26] E. Ileberi, Y. Sun, and Z. Wang, “A machine learning based credit card fraud detection using GA algorithm for feature selection,” *J. Big Data*, vol. 9, no. 1, p. 24, Dec. 2022, doi: 10.1186/s40537-022-00573-8.

[27] 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, doi: 10.18178/ijmlc.2021.11.1.1011.

[28] E. . Ihama, M. I. Akazue, E. U. Omede, and D. V. Ojie, "A Framework for Smart City Model Enabled by Internet of Things (IoT)," *Int. J. Comput. Appl.*, vol. 185, no. 6, pp. 6–11, 2023, doi: 10.5120/ijca2023922685.

[29] 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, doi: 10.14569/IJACSA.2023.0140610.

[30] L. E. Mukhanov, "Using bayesian belief networks for credit card fraud detection," *Proc. IASTED Int. Conf. Artif. Intell. Appl. AIA 2008*, no. February 2008, pp. 221–225, 2008.

[31] V. Filippov, L. Mukhanov, and B. Shchukin, "Credit card fraud detection system," in *2008 7th IEEE International Conference on Cybernetic Intelligent Systems*, IEEE, Sep. 2008, pp. 1–6. doi: 10.1109/UKRICIS.2008.4798919.

[32] 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)*, IEEE, Mar. 2019, pp. 1–5. doi: 10.1109/INFOTEH.2019.8717766.

[33] E. R. Altman, "Synthesizing Credit Card Transactions," *PeerJ Comput. Sci.*, vol. 14, Oct. 2019, [Online]. Available: http://arxiv.org/abs/1910.03033

[34] E. A. L. Marazqah Btoush, X. Zhou, R. Gururajan, K. C. Chan, R. Genrich, and P. Sankaran, "A systematic review of literature on credit card cyber fraud detection using machine and deep learning," *PeerJ Comput. Sci.*, vol. 9, p. e1278, Apr. 2023, doi: 10.7717/peerj-cs.1278.

[35] A. A. Ojugo and R. E. Yoro, "Predicting Futures Price And Contract Portfolios Using The ARIMA Model: A Case of Nigeria's Bonny Light and Forcados," *Quant. Econ. Manag. Stud.*, vol. 1, no. 4, pp. 237–248, 2020, doi: 10.35877/454ri.qems139.

[36] 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)*, IEEE, Mar. 2018, pp. 1–6. doi: 10.1109/ICNSC.2018.8361343.

[37] M. I. Akazue, C. Asuai, A. E. Edje, and E. Omede, "Cybershield: harnessing ensemble feature selection technique for robust distributed denial of service," *Kongzhi yu Juece/Control Decis.*, vol. 38, no. August, pp. 1211–1224, 2023.

[38] Y. Abakarim, M. Lahby, and A. Attioui, "An Efficient Real Time Model For Credit Card Fraud Detection Based On Deep Learning," in *International Conference on Intelligent Systems*, New York, NY, USA: ACM, Oct. 2018, pp. 1–7. doi: 10.1145/3289402.3289530.

[39] M. Zareapoor and P. Shamsolmoali, "Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier," *Procedia Comput. Sci.*, vol. 48, pp. 679–685, 2015, doi: 10.1016/j.procs.2015.04.201.

[40] 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, Jun. 2021, doi: 10.30534/ijatcse/2021/851032021.

[41] F. Shirbani and H. Soltanian Zadeh, “Fast SFFS-Based Algorithm for Feature Selection in Biomedical Datasets,” *Amirkabir Int. J. Sci. Res.*, vol. 45, no. 2, pp. 43–56, 2013, [Online]. Available: https://eej.aut.ac.ir/article_434_6d0cf5e07bb414dc6eec6c82785c0149.pdf

[42] Y. Kang, M. Ozdogan, X. Zhu, Z. Ye, C. Hain, and M. Anderson, “Comparative assessment of environmental variables and machine learning algorithms for maize yield prediction in the US Midwest,” *Environ. Res. Lett.*, vol. 15, no. 6, p. 064005, Jun. 2020, doi: 10.1088/1748-9326/ab7df9.

[43] B. Gaye and A. Wulamu, “Sentimental Analysis for Online Reviews using Machine learning Algorithms,” pp. 1270–1275, 2019.

[44] Maya Gopal P S and Bhargavi R, “Selection of Important Features for Optimizing Crop Yield Prediction,” *Int. J. Agric. Environ. Inf. Syst.*, vol. 10, no. 3, pp. 54–71, Jul. 2019, doi: 10.4018/IJAEIS.2019070104.

[45] D. O. Oyewola, E. G. Dada, N. J. Ngozi, A. U. Terang, and S. A. Akinwumi, “COVID-19 Risk Factors, Economic Factors, and Epidemiological Factors nexus on Economic Impact: Machine Learning and Structural Equation Modelling Approaches,” *J. Niger. Soc. Phys. Sci.*, vol. 3, no. 4, pp. 395–405, 2021, doi: 10.46481/jnsps.2021.173.

[46] Y. Ampatzidis, V. Partel, and L. Costa, “Agroview: Cloud-based application to process, analyze and visualize UAV-collected data for precision agriculture applications utilizing artificial intelligence,” *Comput. Electron. Agric.*, vol. 174, p. 105457, Jul. 2020, doi: 10.1016/j.compag.2020.105457.

[47] A. A. Ojugo, R. E. Yoro, E. O. Okonta, and A. O. Eboka, “A Hybrid Artificial Neural Network Gravitational Search Algorithm for Rainfall Runoffs Modeling and Simulation in Hydrology,” *Prog. Intell. Comput. Appl.*, vol. 2, no. 1, pp. 22–34, 2013, doi: 10.4156/pica.vol2.issue1.2.

[48] 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, doi: 10.11591/ijict.v9i3.pp185-194.

[49] A. A. Ojugo, A. O. Eboka, E. O. Okonta, R. E. Yoro, and F. O. Aghware, “Predicting Behavioural Evolution on a Graph-Based Model,” *Adv. Networks*, vol. 3, no. 2, p. 8, 2015, doi: 10.11648/j.net.20150302.11.

[50] B. U. Ngene and N. Obianigwe, “Nigerian Rain Gauge Station Optimization and National Development: The importance of Head Count,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 413, no. 1, 2018, doi: 10.1088/1757-899X/413/1/012022.

[51] A. T. Abdulrahim *et al.*, “Rainfall-runoff modeling for Challawa and Jakara catchment areas of Kano city, Nigeria,” *Arid Zo. J. Eng. Technol. Environ.*, vol. 17, no. June, pp. 211–220, 2021.

[52] O. A. Adeola, O. Adeyemi, and V. Onyemaenu, “Rainfall–Runoff in Conterminous Tropical River Basins of Southwestern Nigeria,” *African Geogr. Rev.*, vol. 43, no. 5, pp. 1–16, Nov. 2021, doi: 10.1080/19376812.2021.1996250.

[53] D. Eni and F. J. Adeyeye, “Seasonal ARIMA Modeling and Forecasting of Rainfall in Warri Town, Nigeria,” *J. Geosci. Environ. Prot.*, vol. 03, no. 06, pp. 91–98, 2015, doi: 10.4236/gep.2015.36015.

[54] A. A. Ojugo *et al.*, “Hybrid neural network models for rainfall runoffs: Comparative study,” *Adv. Sci. Eng. Res.*, vol. 1, no. 2, pp. 22–34, 2013.

[55] M. Gratian, S. Bandi, M. Cukier, J. Dykstra, and A. Ginther, "Correlating human traits and cyber security behavior intentions," *Comput. Secur.*, vol. 73, pp. 345–358, Mar. 2018, doi: 10.1016/j.cose.2017.11.015.
[56] I. P. Okobah and A. A. Ojugo, "Evolutionary Memetic Models for Malware Intrusion Detection: A Comparative Quest for Computational Solution and Convergence," *Int. J. Comput. Appl.*, vol. 179, no. 39, pp. 34–43, 2018, doi: 10.5120/ijca2018916586.
[57] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, "A Comparative Analysis of XGBoost," no. February, 2019, doi: 10.1007/s10462-020-09896-5.
[58] 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, doi: 10.46481/jnsps.2021.308.
[59] A. A. Ojugo and A. O. Eboka, "Memetic algorithm for short messaging service spam filter using text normalization and semantic approach," *Int. J. Informatics Commun. Technol.*, vol. 9, no. 1, p. 9, 2020, doi: 10.11591/ijict.v9i1.pp9-18.
[60] A. A. Ojugo *et al.*, "Dependable Community-Cloud Framework for Smartphones," *Am. J. Networks Commun.*, vol. 4, no. 4, p. 95, 2015, doi: 10.11648/j.ajnc.20150404.13.
[61] P. . Maya Gopal and Bhargavi R, "Feature Selection for Yield Prediction Using BORUTA Algorithm," *Int. J. Pure Appl. Math.*, vol. 118, no. 22, pp. 139–144, 2018.
[62] G. Cho, J. Yim, Y. Choi, J. Ko, and S. H. Lee, "Review of machine learning algorithms for diagnosing mental illness," *Psychiatry Investig.*, vol. 16, no. 4, pp. 262–269, 2019, doi: 10.30773/pi.2018.12.21.2.
[63] A. A. Ojugo and R. E. Yoro, "Extending the three-tier constructivist learning model for alternative delivery: ahead the COVID-19 pandemic in Nigeria," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 21, no. 3, p. 1673, Mar. 2021, doi: 10.11591/ijeecs.v21.i3.pp1673-1682.
[64] 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, Mar. 2022, doi: 10.22214/ijraset.2022.40702.
[65] A. A. Ojugo and E. O. Ekurume, "Predictive Intelligent Decision Support Model in Forecasting of the Diabetes Pandemic Using a Reinforcement Deep Learning Approach," *Int. J. Educ. Manag. Eng.*, vol. 11, no. 2, pp. 40–48, Apr. 2021, doi: 10.5815/ijeme.2021.02.05.
[66] A. E. Ibor, E. B. Edim, and A. A. Ojugo, "Secure Health Information System with Blockchain Technology," *J. Niger. Soc. Phys. Sci.*, vol. 5, no. 992, pp. 1–8, 2023, doi: 10.46481/jnsps.2022.992.
[67] A. A. Ojugo and O. D. Otakore, "Improved Early Detection of Gestational Diabetes via Intelligent Classification Models: A Case of the Niger Delta Region in Nigeria," *J. Comput. Sci. Appl.*, vol. 6, no. 2, pp. 82–90, 2018, doi: 10.12691/jcsa-6-2-5.
[68] 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, doi: 10.22624/AIMS/DIGITAL/v11N4P3.
[69] A. A. Ojugo and O. D. Otakore, "Intelligent cluster connectionist recommender system using implicit graph friendship algorithm for social networks," *IAES Int. J. Artif. Intell.*, vol. 9, no. 3, p. 497~506, 2020, doi: 10.11591/ijai.v9.i3.pp497-506.

[70] O. D. Voke, A. Maureen, and I. Anthony, “A Framework for Feature Selection using Data Value Metric and Genetic Algorithm,” vol. 184, no. 43, pp. 14–21, 2023.
[71] D. A. Oyemade *et al.*, “A Three Tier Learning Model for Universities in Nigeria,” *J. Technol. Soc.*, vol. 12, no. 2, pp. 9–20, 2016, doi: 10.18848/2381-9251/CGP/v12i02/9-20.
[72] A. A. Ojugo and D. O. Otakore, “Redesigning Academic Website for Better Visibility and Footprint: A Case of the Federal University of Petroleum Resources Effurun Website,” *Netw. Commun. Technol.*, vol. 3, no. 1, p. 33, Jul. 2018, doi: 10.5539/nct.v3n1p33.
[73] K. Vaishnavi, U. N. Kamath, B. A. Rao, and N. V. S. Reddy, “Predicting Mental Health Illness using Machine Learning Algorithms,” *J. Phys. Conf. Ser.*, vol. 2161, no. 1, 2022, doi: 10.1088/1742-6596/2161/1/012021.
[74] A. A. Ojugo, M. I. Akazue, P. O. Ejeh, C. Odiakaose, and F. U. Emordi, “DeGATraMoNN: Deep Learning Memetic Ensemble to Detect Spam Threats via a Content-Based Processing,” *Kongzhi yu Juece/Control Decis.*, vol. 38, no. 01, pp. 667–678, 2023.
[75] G. TekalignTujo, G. Dileep Kumar, D. ElifeneshYitagesu, and B. MeseretGirma, “Predictive Model to Predict Seed Classes using Machine Learning,” *Int. J. Eng. Res. \& Technol.*, vol. 6, no. 08, pp. 334–344, 2017.
[76] Q. Li *et al.*, “An Enhanced Grey Wolf Optimization Based Feature Selection Wrapped Kernel Extreme Learning Machine for Medical Diagnosis,” *Comput. Math. Methods Med.*, vol. 2017, pp. 1–15, 2017, doi: 10.1155/2017/9512741.
[77] C. C. Odiakaose, F. U. Emordi, P. O. Ejeh, O. Attoh, and N. C. Ashioba, “A pilot study to enhance semi-urban tele-penetration and services provision for undergraduates via the effective design and extension of campus telephony,” *FUPRE J. Sci. Ind. Res.*, vol. 7, no. 3, pp. 35–48, 2023.
[78] F. Mustofa, A. N. Safriandono, A. R. Muslikh, and D. R. I. M. Setiadi, “Dataset and Feature Analysis for Diabetes Mellitus Classification using Random Forest,” *J. Comput. Theor. Appl.*, vol. 1, no. 1, pp. 41–48, 2023, doi: 10.33633/jcta.v1i1.9190.
[79] A. R. Muslikh, I. D. R. 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, doi: 10.52436/1.jutif.2023.4.6.1529.
[80] E. U. Omede, A. 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. 2, no. 1, pp. 43–53, 2024, doi: 10.33633/jcta.v2i1.9541.
[81] A. Bahl *et al.*, “Recursive feature elimination in random forest classification supports nanomaterial grouping,” *NanoImpact*, vol. 15, p. 100179, Mar. 2019, doi: 10.1016/j.impact.2019.100179.
[82] 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, Oct. 2023, doi: 10.33633/jcta.v1i2.9259.
[83] C. Jiang, J. Song, G. Liu, L. Zheng, and W. Luan, “Credit Card Fraud Detection: A Novel Approach Using Aggregation Strategy and Feedback Mechanism,” *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3637–3647, Oct. 2018, doi: 10.1109/JIOT.2018.2816007.
[84] A. Ali *et al.*, “Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review,” *Appl. Sci.*, vol. 12, no. 19, p. 9637, Sep. 2022, doi: 10.3390/app12199637.

[85] A. Shaji, S. Binu, A. M. Nair, and J. George, “Fraud Detection in Credit Card Transaction Using ANN and SVM,” 2021, pp. 187–197. doi: 10.1007/978-3-030-79276-3_14.
[86] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0.
[87] J. K. Oladele *et al.*, “BEHeDaS: A Blockchain Electronic Health Data System for Secure Medical Records Exchange,” *J. Comput. Theor. Appl.*, vol. 2, no. 1, pp. 1–12, 2024, doi: 10.33633/jcta.v2i19509.
[88] C. C. Odiakaose *et al.*, “DeLEMPaD: Pilot Study on a Deep Learning Ensemble for Energy Market Prediction of Price Volatility and Direction,” *Comput. Inf. Syst. Dev. Informatics Allied Res. J.*, vol. 15, no. 1, pp. 47–62, 2024, doi: 10.22624/AIMS/CISDI/V15N1P4.
[89] R. E. Yoro, F. O. Aghware, B. O. Malasowe, O. Nwankwo, and A. A. Ojugo, “Assessing contributor features to phishing susceptibility amongst students of petroleum resources varsity in Nigeria,” *Int. J. Electr. Comput. Eng.*, vol. 13, no. 2, p. 1922, Apr. 2023, doi: 10.11591/ijece.v13i2.pp1922-1931.
[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, doi: 10.5815/ijitcs.2018.10.07.
[91] J. Li *et al.*, “Feature selection: A data perspective,” *ACM Comput. Surv.*, vol. 50, no. 6, 2017, doi: 10.1145/3136625.
[92] 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, Dec. 2023, doi: 10.11591/ijict.v12i3.pp205-213.
[93] S. Wang, J. Tang, H. Liu, and E. Lansing, “Encyclopedia of Machine Learning and Data Science,” *Encycl. Mach. Learn. Data Sci.*, no. October 2017, pp. 1–9, 2020, doi: 10.1007/978-1-4899-7502-7.
[94] A. Jović, K. Brkić, and N. Bogunović, “A review of feature selection methods with applications,” *2015 38th Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2015 - Proc.*, pp. 1200–1205, 2015, doi: 10.1109/MIPRO.2015.7160458.
[95] Y. Bouchlaghem, Y. Akhiat, and S. Amjad, “Feature Selection: A Review and Comparative Study,” *E3S Web Conf.*, vol. 351, pp. 1–6, 2022, doi: 10.1051/e3sconf/202235101046.
[96] S. Nosratabadi, F. Imre, K. Szell, S. Ardabili, B. Beszedes, and A. Mosavi, “Hybrid Machine Learning Models for Crop Yield Prediction,” Mar. 2020, [Online]. Available: http://arxiv.org/abs/2005.04155
[97] R. Nasir, M. Afzal, R. Latif, and W. Iqbal, “Behavioral Based Insider Threat Detection Using Deep Learning,” *IEEE Access*, vol. 9, pp. 143266–143274, 2021, doi: 10.1109/ACCESS.2021.3118297.
[98] P. K. Kosamkar and V. Y. Kulkarni, “Agriculture crop simulation models using computational intelligence,” *Int. J. Comput. Eng. Technol.*, vol. 10, no. 3, pp. 3–19, May 2019, doi: 10.34218/IJCET.10.3.2019.015.
[99] 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, Jun. 2022, doi: 10.1155/2022/2439205.

[100] 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, Nov. 2021, doi: 10.35877/mathscience614.

[101] B. O. Malasowe, M. I. Akazue, E. A. 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, doi: 10.14569/IJACSA.2023.0140816.

[102] A. Taravat and F. Del Frate, "Weibull Multiplicative Model and Machine Learning Models for Full-Automatic Dark-Spot Detection From Sar Images," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XL-1/W3, no. September 2013, pp. 421–424, 2013, doi: 10.5194/isprsarchives-xl-1-w3-421-2013.

[103] J. Jung, M. Maeda, A. Chang, M. Bhandari, A. Ashapure, and J. Landivar-Bowles, "The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems," *Curr. Opin. Biotechnol.*, vol. 70, pp. 15–22, Aug. 2021, doi: 10.1016/j.copbio.2020.09.003.