Research Article / Review Article FUPRE JOURNAL 7(3): 20-34 (2023)

FUPRE Journal

of

Scientific and Industrial Research

ISSN: 2579-1184(Print) ISSN: 2578-1129 (Online)

http://fupre.edu.ng/journal

Reinforcement Deep Learning Memetic Algorithm For Detection of Short Messaging Service Spam Using Filters To Curb Insider Threats in Organizations

ADISHI, E. [1,*], ***EJEH, P.*** [2], ***OKORO, E.*** [3] ***JISU, A.*** [4]

[1,3,4]*Department of Intelligence and Security Studies, College of Management and Social Sciences, Novena University, Ogume, Delta State, Nigeria*

[2]*Department of Computer Science, College of Computing & Telecommunications, Novena University, Ogume, Delta State, Nigeria*

ARTICLE INFO

Received: 07/04/2023
Accepted: 08/06/2023

Keywords

Heuristic-processing, Memetic algorithm, Semantic approach,Spam filters, Text normalization, strings

ABSTRACT

Today's popularity of the short messages services (SMS) has created a propitious environment for spamming to thrive. Spams are unsolicited advertising, adult-themed or inappropriate content, premium fraud, smishing and malware. They are a constant reminder of the need for an effective spam filter. However, SMS limitations of 160-charcaters and 140-bytes size as well as its being rippled with slangs, emoticons and abbreviations further inhibits effective training of models to aid accurate classification. The study proposes Genetic Algorithm Trained Bayesian Network solution that seeks to normalize noisy feats, expand text via use of lexicographic and semantic dictionaries that uses word sense disambiguation technique to train the underlying learning heuristics. And in turn, effectively help to classify SMS in spam and legitimate classes. Hybrid model comprises of text preprocessing, feature selection as well as training and classification section. Study uses a hybrid Genetic Algorithm trained Bayesian model for which the GA is used for feature selection; while, the Bayesian algorithm is used as classifier.

1. INTRODUCTION

The advent of short messaging services by Neil Papworth since 1992, has seen great penetration and a tremendous growth rate of the service. Advent of mobile phones with enhance features has contributed to the large-scale adoption of SMS by users. The portability, mobility, ubiquity of services and its low cost continues to promote text messages to become the most used means of electronic communication in the world today (Ojugo & Eboka, 2018; Ojugo & Yoro, 2021a). Short Message Service (SMS) is text service component of mobile communication system or phones with standardized protocols that allow the exchange of short text messages between fixed line or mobile phone devices. An estimated 23-billion SMS is sent daily worldwide in 2014 (Ezpeleta et al., 2016, 2020; Ojugo & Oyemade, 2021); While, a total of 8.3 trillion SMS was sent worldwide in the same year with net market revenue of over $128Billion in 2011 (Ojugo & Eboka, 2020b; Sahmoud & Mikki, 2022).

In 2016, the revenue was forecasted to be over $153Billion; And, evidence has shown that 3.39billion SMS was sent and received

*Corresponding author, e-mail: *kemmyadesanya@gmail.com*
DIO

in Nigeria alone in the year 2013 (Iiloani, 2015). The increased popularity and the consequent proliferation of SMS platforms, has also seen a corresponding rise in unsolicited SMS called spams (G. Bhati, 2019; Paliwal et al., 2022; Verma et al., 2020). The ITU 2005 campaign witnessed a rise in the unsolicited commercial adverts as sent to mobile phones via SMS. Recent drift from email to SMS spams is attributed to the availability of effective email filters, user awareness and industry collaboration (Akazue et al., 2023; Yoro, Aghware, Akazue, et al., 2023; Yoro, Aghware, Malasowe, et al., 2023).

Spams are unsolicited electronic messages that include, and not limited to, emails, SMS, Voice over Internet Protocols (VoIP) and instant messaging from chats. Spams are unsolicited or unwanted messages from a sender, sent indiscriminately with no prior relationship to a user mostly for commercial reasons (Ojugo et al., 2021; Ojugo & Nwankwo, 2021d; Ojugo & Obruche, 2021). SMS Spams ranges from adult-themed and inappropriate contents, unsolicited adverts, smishing and mobile malware etc (Ibor et al., 2023; Ojugo et al., 2015; Ojugo & Otakore, 2018a; Udeze et al., 2022).

SMS spams have since become enormous challenge – causing great loss of revenue to Internet Service Providers, Mobile Network Operators and users in general. The overall growth of spams grew by 300% from 2011 to 2012 from millions of SMS received worldwide; And 33.3% attributed to spam-related messages (Ojo et al., 2021; Oyewola et al., 2021). Also, in Nigeria alone, an estimated 334,857,685 SMS spam were received daily in 2015 (Hong, 2018; Kumaraguru et al., 2010; Zhang et al., 2007). This implies that lots of mobile phone users are handicapped in the control of the number of spams they receive (Chaminda etal., 2013). Besides being distractive and annoying, users need a certain degree of privacy with their phones and free from Spam and viruses invasions (Alsowai & Al-Shehari, 2021; Gaye & Wulamu, 2019; Mahajan & Sharma, 2015). Mobile network operators are geared towards reducing the number of spams over their network as such flooding makes the SMS channel more invasive and less secure (Ahmad et al., 2016; Ojugo & Eboka, 2020b; Ojugo & Otakore, 2020b; Ojugo & Yoro, 2020).

The tremendous rise in the usage of SMS is attributed to (Ojugo & Eboka, 2020b; Ojugo & Oyemade, 2021):

1. Trust in SMS channel: SMS is a private communication between two parties only has created some level of trust and acceptance all over the world such that financial institution has adopted its use in payment authorization
2. High open rate*:* Average time it takes to respond to SMS is faster than email and voice call – making it a preferred choice. Statistics have shown SMS has an average open rate of 99% and opens within 15-minutes; While, an email has an open rate of 20-25% and open with 24-hours
3. Low cost of transaction: Average cost per SMS is almost negligible, and free for some networks – affording mobile phone users the opportunity to send as many without recourse to cost. Marketers and many other institution has embrace bulk SMS a medium for advertising and interact with customers.
4. Ease and Convenience of texting enables its use in nearly every environment without disrupting people around phone users; Unlike in voice call, SMS can be in absolute silence without inconveniencing people around. Aided by the portable size of most mobile devices, communication can be done almost everywhere and any position.

SMS has great benefit for both subscribers and operators in diverse ways centered on convenience, flexibility, seamless integration of messaging services and data access. Others may include (Ojugo & Otakore, 2018b, 2018a, 2020a): (a) delivery of notifications, (b) guaranteed delivery, (c) reliable, low-cost for concise data, (d) capability to screen messages and return calls, (e) increases productivity, (f) more sophisticated functions provides enhanced user benefits, (g) delivery to multiple users at same time, (h) receive diverse data, (i) create user groups, (l) e-mail generation, (m) integrate with other data and Internet-based application, and (n) increase in revenue for MNOs (Ojugo et al., 2012; Ojugo, Eboka, Yerokun, et al., 2013; Ojugo, Yoro, Eboka, et al., 2013; Ojugo et al., 2014).

The tremendous rise in the usage of SMS is attributed to its ease of use, ubiquity in nature, high open rates, low cost of transaction and inherent trust in the channel. The ease of use, portability, ubiquity, low open rate, and low cost of SMS are major factors for its popularity and usage. This growth rate has equally attracted spamming to the channel. Spammers are often, well organized businesses seeking to make money via the use of email, short messages (SMS), Instant message, UseNet newsgroup, Social network and internet telephony channel without the consent of subscriber (Ojugo, Yoro, Okonta, et al., 2013; Wemembu et al., 2014).

Their merchandise is unsolicited advertising, inappropriate or adult-themed content, premium fraud, smishing and even distribution of malware generally called spam. SMS spams are thus, unsolicited and unwanted messages sent to mobile phone users. Spam trend is on the rise and its toll on subscribers and even MNO is getting intensive and proven to be of great concern to all (Eboka & Ojugo, 2020; Ojugo, Abere, Eboka, et al., 2013; Ojugo, Abere, Orhionkpaiyo, et al., 2013).

1.1 Sources of Spam

SMS spam is generated from various sources. A typical spam sources is number harvesting, which is carried out by Internet sites offering “free” services (Ojugo, Yoro, Oyemade, et al., 2013; Ojugo, Yoro, Yerokun, et al., 2013); While, users can also receive mobile spam from the following sources (Artikis et al., 2017; Sohony et al., 2018; Wang et al., 2019):

1. Organizations and individuals that pay MNO to deliver SMS to the subscribers: They are responsible for the highest number of spams received on subscriber’s mobile phones. Although, MNOs have adopted and enforced use of opt-out, or even opt-in processes for the user to stop receiving promos or ads.
2. Organizations that do not pay for the SMS that are delivered to the subscribers: they are usually worse and considered as fraud because it damages MNO brands.
3. Individual originated messages that disturb recipients.

Besides being distracting and annoying, spams have serious consequences it generates. There is the issue of competition for resources between millions of illegitimate and legitimate messages being transmitted. These messages consume network resources that could have otherwise been allocated to other legitimate services by MNO (Tingfei et al., 2020). Spamming attracts extra cost for mobile operators to adequately maintain their communication infrastructures for effective service delivery. Flooding it with illegitimate massages can

cause legitimate users to suffer denial of service. Huge amount of spam also concerns the cellular carriers as the messages traverse via the network, causing congestion and degrade in performance (Akazue et al., 2022, 2023). Communication industries are also faced with threat from virus propagated by spam SMS. Fraudulent messages such as phishing and other fraudulent acts that were common with email, are now ported on SMS platform. Financial loss, damage to mobile user's reputation and that of the MNO are issues to be considered (Ojugo & Ekurume, 2021a, 2021b; Yoro & Ojugo, 2019b, 2019a).

2.2 Types of Spam

SMS spam filters shares similar features and challenges with email spam filters. They are both saddled with the task of real-time filtering efficiency and the option to decide between client-side and or server-side filtering. The mobile space is faced with the challenge of overcoming misclassification cost and eliminate false positives (genuine SMS incorrectly classified as spam by filter), and issue of concept drift in order to evade detection by system filters. Thus, most existing approaches of combating SMS spam are imported from successful email solutions (Nivedha & Raja, 2022; Ojugo & Eboka, 2019, 2020a). But not all solutions to email spam are applicable to SMS due to the fact that established email spam filters are unable to tackle SMS spam because the performance of email spam filters is seriously degraded when used to filter spam. This is attributed to its limited 160-character of 140-bytes sized messages. Also, these messages are rife with slang, symbols, emoticons, and abbreviations that inhibit proper classification (Rathi & Pareek, 2013).

To overcome the shortfall of email filters in handling SMS spam successfully, a combination of filtering techniques to reduce noise in SMS and expands the message size – is the focus of this research. Spam filters can be divided into a number of broad categories based on the method used to filter Spam. They include (Redondo-Gutierrez et al., 2022; Sahmoud & Mikki, 2022) the following methods of filtering namely:

1. List-based Filters – simply blocks spam (i.e. unwanted messages) using an already created list of senders. It then seeks to create a whitelist (i.e. a list of authentic senders), or a blacklist (i.e. a list of blocked sender records). Thus, when an incoming message is received – the spam filter simply checks if the IP, email address and/or mobile number is on the list. If the list is on the whitelist, the sender is accepted; otherwise, the message is not delivered to the user/subscriber. Vice versa in the case of a blacklist.
2. Challenge and Response Filters – forces a message sender to prove themselves via series of tests. It blocks undesirable messages by forcing the sender to perform a task before their message is delivered. If the task is successful, the message (and future messages) are delivered to the recipient; While, a failure to complete the challenge leads to message rejection.
3. Content-based Filters – evaluates words or phrases found in messages to determine if the message is spam or not. It analyzes the message header, subject and body to discover any distinctive characteristic. It performs such feats via two-modes: word-based and heuristic filters.

   Word-based filters use a set of rules to detect genuine from spam SMS. Also known as rule-filters, they use rules about actual word(s) or phrase(s) in a message to classify messages into genuine and spam classes. Rule features include word type, frequency of occurrence, structure

of text (e.g., font size, colour etc.), presence of many periods between letters (e.g., F.R.E.E), existence of image, etc. Rules are filter-dependent and can vary from simple to very complex. A demerit of rule-based filters is that: (a) they are knowledge intensive, (b) time consuming process in reviewing spam messages to determine the rules, and (c) needs regular update of rules as spammers changes their tactics.

Heuristic-filter examines data content via various algorithms and resources, and assigns points to words or phrases. Words commonly found in spams such as “FREE” or "SEX," receive higher scores. Terms commonly found in normal messages receive lower scores. The filter then adds up total scores. If the message receives a certain score or higher (determined by anti-spam application's administrator), the filter identifies it as spam and blocks it. Messages with score(s) lower than the target number are delivered to the user. Using a heuristic filter allows many spam filtering methods to be used, resulting in better performance than any single method by itself.

1.3. Study Motivation

To re-investigate prediction of oil futures price – our statement of problem is thus:

1. The increased adoption of smartphones and its popularity, easy access, mobility, portability – have consequently led to the increase of spam and phishing activities.
2. Companies are today still dealing with the many issue(s) of SMS spam, and existing approaches are quite unable to effectively tackle SMS spam successfully – as their performance is seriously hampered and degraded by the parametric feats used to filter spams.
3. Spams have continued to rise at an alarming growth rate – causing financial loss and emotional stress to smartphone users, businesses, and network operator(s).
4. Formulation and design of an effective SMS filter has continued to suffered setback(s) due to the inherent reason that SMS filters by design are not as simple as email filters due to its limited size of 160-characters of 140bytes sized data. These amongst other constraints, continue to create rippled impediment in size of feature to be selected for training and consequently contributing to poor learning and classification of learning algorithm.
5. In addition, SMS are rippled with slangs, abbreviations, symbols, and emoticons that inhibit proper classification of words.

To overcome these amongst many other shortfalls inherent in the adoption of email filters as adapted to handling SMS spam successfully, a hybrid filtering technique that reduces noise in form of slangs, emoticons, abbreviations in SMS as well as expand message size must be employed to enhance adequate classification. Thus, our research goal(s) is to propose a hybrid deep learning neural network model for text normalization and semantic expansion in SMS spam filtering.

2. MATERIALS AND METHODS

2.1 Dataset Used & Proposed Framework

The dataset used for this study is the knowledge discovery in databases (KDD 2022). Diagram of the proposed framework is seen in figure 1 with the framework grouped into the sections namely (Ojugo & Eboka, 2018, 2019, 2020b; Okobah & Ojugo, 2018):

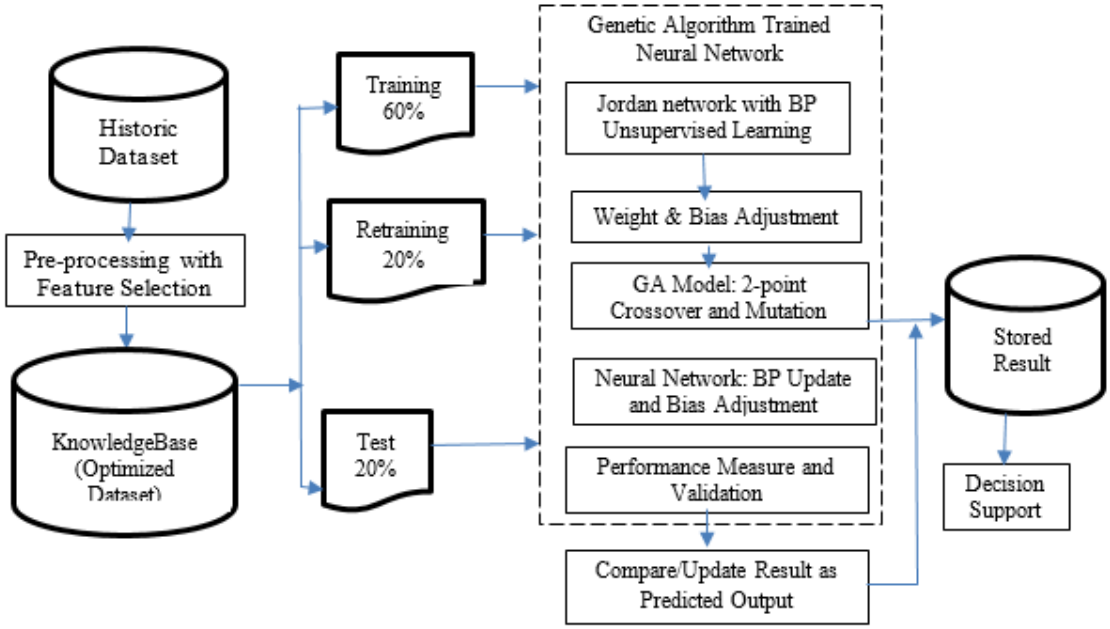

Figure 1: Schematic diagram of the Memetic Algorithm

2.2. Experimental Memetic Model

The proposed model cum ensemble consists of these parts (Ojugo and Yoro, 2020; 2021; Wheeler and Aitkens, 2019):

1. Knowledgebase – consists of observed, historic structured data feats. The dataset is a record of fraudulent malware intrusion transactions stored and converted as fuzzy if-then ruleset using optimized membership functions. The rule-based system consists of *classifier* to propagate the IF-THEN rule values of selected data, enhanced them as predefined variables classification into intrusion types for fraud detection. Its houses the optimized universe discourse values as represented by fuzzy-if-then, linguistic variables (rule-based) as selected data feats.
2. Inference engine – consists of the memetic algorithm (i.e. the hybrid, rule-based genetic algorithm trained neural network model). The neural network is constructed using the Jordan network, and seeks to provide a self-learning ability, optimized by the CGA optimizer that recombines and mutates the rule-based fuzzy dataset to train and test the system so that it autonomously classifies diabetes into its class types. Conversely, the Genetic Algorithm helps train the neural network so that combined – they can effectively optimize our collated-answers within the tuned fuzzy dataset linguistic variable (symptoms) universe of discourse values in other to yield a centralized, fuzzy-scaled function boundary in determining high/low degree membership function. Thus, the inference engine infers conclusion derived from genetic algorithm trained neural network from the selected data feats encoded as fuzzy-if-then conditions with possible outcomes and consequent action upon criteria being met.
3. Decision support– consists of the predicted output and the output database that is updated automatically in time as patients are diagnoses as long as it encounters and read sin new data. The decision support predicts system output based on the cognitive and the emotional filers as display by the output device. This is seen in figure 1.

*Corresponding author, e-mail: *kemmyadesanya@gmail.com*
DIO

The experimental ensemble is initialized with the if-then rules as individuals, whose fitness is computed. 30-individuals are then selected via tournament method as new pool. It then determines mating individuals to yield solutions. We use a multi-point crossover and mutation to help the network to learn all the dynamic and non-linear feats in the dataset (as feats of interest). With mutation, suspicion score for each rule between 1-to-30 is then randomly generated using Gaussian distribution corresponding to crossover points (all genes are from single parent). As new parents contribute the rest to yield new individuals whose genetic makeup is a combination of both parents, mutation is also applied to yield 3-random genes. These further undergo mutation and are then allocated new random values that still conform to the belief space. These random values will range between 0 and 1, which yields the suspicion score for each transaction as generated for each account holder (Syeda et al, 2002; Sylla and Wild, 2011; Vooshoghi et al, 2019).

The number of mutations applied depends on how far CGA is progressed on the network (how fit is the fittest individual in the pool), which equals fitness of the fittest individual divided by 2. New individuals replace old with low fitness so as to create a new pool. Process continues until individual with a fitness value of 0.8 is found – indicating that the solution has been reached (Ojugo and Ekurume, 2020; Ojugo and Otakore, 2018; 2020). Initialization/selection via ANN ensures that first 3-beliefs are met; mutation ensures fourth belief is met. Its influence function influences how many mutations take place, and the knowledge of solution (how close its solution is) has direct impact on how algorithm is processed. Algorithm stops when best individual has fitness of 0 (Ojugo and Otakore, 2021; Ojugo and Oyemade, 2021; Phua et al, 2007; Stolfo et al, 2000).

3. RESULT FINDINGS & DISCUSSION

3.1. Ensemble Evaluation

Here, to compute accuracy, recall, error rate (ER) and specificity are used to evaluate the performance of the detection models (Brofman Epelbaum & Garcia Martinez, 2014; Gokarn & Choudhary, 2021; Zawislak et al., 2022). The formulas of the above criteria are calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Error\ Rate = \frac{FP + FN}{FP + TP + TN + FN} \quad (3)$$

$$SPEC = \frac{TN}{TN + FP} \quad (4)$$

A true positive (TP) is a case (rule) that correctly distinguishes spam from genuine text. The true negative (TN) shows normal text message data classified correctly as normal. The false negative (FN) denotes a case in which a text is classified as normal data, and a false positive (FP) means that a normal text is classified as a spam. The accuracy rate shows the overall correct detection accuracy of the dataset, ER refers to the robustness of the classifier, recall indicates the degree of correctly detected attack types of all cases classified as attacks,

and specificity shows the percentage of correctly classified normal data. In the above, higher accuracy and recall and lower ER indicate good performance.

4.2. Result Findings and Discussion

We measure effectiveness and accuracy, we measure their rate of misclassification and corresponding improvement percentages in both training and test data sets as summarized (Brofman Epelbaum & Garcia Martinez, 2014; Gokarn & Choudhary, 2021; Zawislak et al., 2022) as in Table 1.

$$Misclassification\ Rate = \frac{No.of\ Incorrectly\ Classified\ Rules}{No.of\ Sample\ set} \quad (5)$$

Table 1: Model misclassification rates

	Classification Errors	
Model	Training	Testing
Naïve Bayes	52.5%	23.2%
Genetic Algorithm	48.4%	4.7%
Proposed Memetic Algorithm	19.6%	1.02%

Table 1 shows misclassification error rate with Naïve Bayes, GA and GABN at 23.2%, 4.7% and 1.02% (that is, error rate in false-positive and true-negative) respectively.

Conversely, table 2 shows improvement rate for the ensemble/model as:

$$Improvement = \frac{MR(A) - MR(B)}{MR(A)}\ x\ 100 \quad (6)$$

Table 4: Improvement Percentage

	Improvement %	
Model	Training	Testing
Naïve Bayes	2.11%	3.6%
Genetic Algorithm	2.32%	4.02%
Proposed Memetic Algorithm	0.09%	0.12%

The various benchmark ensembles promise an improvement rate as of 3.6%, 4.02% and 0.12% respectively (Ibor et al., 2023).

Hybrids are difficult to implement and its accompanying data must be appropriately encoded so that model can exploit numeric data and efficiently explore the domain space to yield an optimal solution void of over-fitting, over-training and over-parameterization. Models serve as educational tools to compile knowledge about a task, serve as new language to convey ideas as we gain better insight to investigate input parameter(s) crucial to a task; while, its sensitivity analysis helps to reflect on theories of systems functioning. Simple model may not yield enough data; and, complex model may not be fully understood. A detailed model helps develop reasonably applicable models even when not operationally applicable in a larger scale Their implementation should seek its feedback as more critical rather than seeking an accurate agreement with historic data. Since, a balance in the model's complexity will help its being understood and its manageability, so that the model can be fully explored as in (Ojugo & Nwankwo, 2021a, 2021b, 2021c, 2021d; Ojugo & Yoro, 2021b).

4. CONCLUSION

SMS spam filters can have the capacity, and be granted capability to transcribe slangs, emoticons and abbreviations into standard terms as well as expand the message size to enhance better feature extraction for classification algorithms and approaches. The study will also serve to reduce orthographic error found in SMS, chat groups and other social network communication medium that impedes machine learning algorithm. This is because from the various approaches adopted to SMS spam filters – the content-based models with text pre-processing has shown to perform better. Machine translation performs better when applied to normalized text messages. We can equally combined multiple approaches in noisy data, text normalization to create a better output. But, extracting only relevant feats to train the classifier has been reported to contribute to the efficiency of SMS spam filters. Thus, we propose text pre-processing SMS spam filter model with the capability of normalizing, expanding text messages and extracting suitable features as dataset input parameters for training the adopted classification algorithm and model.

Conflict of Interest

The authors declare that there is no conflict of interest.

References

Ahmad, H. W., Zilles, S., Hamilton, H. J., and Dosselmann, R. (2016). Prediction of retail prices of products using local competitors. *International Journal of Business Intelligence and Data Mining*, *11*(1): 19–30. https://doi.org/10.1504/IJBIDM.2016.076418

Akazue, M. I., Ojugo, A. A., Yoro, R. E., Malasowe, B. O., and Nwankwo, O. (2022). Empirical evidence of phishing menace among undergraduate smartphone users in selected universities in Nigeria. *Indonesian Journal of Electrical Engineering and Computer Science*, *28*(3):1756–1765. https://doi.org/10.11591/ijeecs.v28.i3.pp1756-1765

Akazue, M. I., Yoro, R. E., Malasowe, B. O., Nwankwo, O., and Ojugo, A. A. (2023). Improved services traceability and management of a food value chain using block-chain network : a case of Nigeria. *Indonesian Journal of Electrical Engineering and Computer Science*, *29*(3): 1623–1633. https://doi.org/10.11591/ijeecs.v29.i3.pp1623-1633

Alsowai, R. A., and Al-Shehari, T. (2021). A multi-tiered framework for insider threat prevention. *Electronics (Switzerland)*, *10*(9). https://doi.org/10.3390/electronics10091005

Artikis, A., Katzouris, N., Correia, I., Baber, C., Morar, N., Skarbovsky, I., Fournier, F., and Paliouras, G. (2017). A Prototype for Credit Card Fraud Management. *Proceedings of the 11th ACM International Conference on Distributed and Event-Based Systems*, 249–260. https://doi.org/10.1145/3093742.3093912

Brofman Epelbaum, F. M., and Garcia Martinez, M. (2014). The technological evolution of food traceability systems and their impact on firm sustainable

performance: A RBV approach. *International Journal of Production Economics*, *150*: 215–224. https://doi.org/10.1016/j.ijpe.2014.01.007

Eboka, A. O., and Ojugo, A. A. (2020). Mitigating technical challenges via redesigning campus network for greater efficiency, scalability and robustness: A logical view. *International Journal of Modern Education and Computer Science*, *12*(6): 29–45. https://doi.org/10.5815/ijmecs.2020.06.03

Ezpeleta, E., de Mendizabal, I. V., Gómez Hidalgo, J. M., and Zurutuza, U. (2020). Novel email spam detection method using sentiment analysis and personality recognition. *Logic Journal of the IGPL*, *28*(1): 83–94. https://doi.org/10.1093/jigpal/jzz073

Ezpeleta, E., Zurutuza, U., and Gómez Hidalgo, J. M. (2016). *Does Sentiment Analysis Help in Bayesian Spam Filtering?* (pp. 79–90). https://doi.org/10.1007/978-3-319-32034-2_7

G. Bhati, R. (2019). A Survey on Sentiment Analysis Algorithms and Datasets. *Review of Computer Engineering Research*, *6*(2), 84–91. https://doi.org/10.18488/journal.76.2019.62.84.91

Gaye, B., and Wulamu, A. (2019). *Sentimental Analysis for Online Reviews using Machine learning Algorithms*. 1270–1275.

Gokarn, S., and Choudhary, A. (2021). Modeling the key factors influencing the reduction of food loss and waste in fresh produce supply chains. *Journal of Environmental Management*, *294*: 113063. https://doi.org/10.1016/j.jenvman.2021.113063

Hong, I. B. (2018). Social and personal dimensions as predictors of sustainable intention to use facebook in Korea: An empirical analysis. *Sustainability (Switzerland)*, *10*(8). https://doi.org/10.3390/su10082856

Ibor, A. E., Edim, E. B., and Ojugo, A. A. (2023). Secure Health Information System with Blockchain Technology. *Journal of the Nigerian Society of Physical Sciences*, *5*(992): 1–8. https://doi.org/10.46481/jnsps.2022.992

Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F., & Hong, J. (2010). Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology*, *10*(2), 1–31. https://doi.org/10.1145/1754393.1754396

Mahajan, A., and Sharma, S. (2015). The Malicious Insiders Threat in the Cloud. *International Journal of Engineering Research and General Science*, *3*(2): 246–256. www.ijergs.org

Nivedha, M. A., and Raja, S. (2022). Detection of Email Spam using Natural Language Processing Based Random Forest Approach. *International Journal of Computer Science and Mobile Computing*,*11*(2):7–22. https://doi.org/10.47760/ijcsmc.2022.v11i02.002

Ojo, O. E., Gelbukh, A., Calvo, H., and Adebanji, O. O. (2021). Performance Study of N-grams in the Analysis of Sentiments. *Journal of the Nigerian Society of Physical Sciences*, *3*(4), 477–

483. https://doi.org/10.46481/jnsps.2021.201

Ojugo, A. A., Abere, R. A., Eboka, A. O., Yerokun, M. O., Yoro, R. E., Onochie, C. C. and Oyemade, D. A. (2013). Hybrid neural network models for rainfall runoffs: Comparative study. *Advancement in Scientific and Engineering Research*, *1*(2), 22–34.

Ojugo, A. A., Abere, R. A., Orhionkpaiyo, B. C., Yoro, R. E., and Eboka, A. O. (2013). Technical Issues for IP-Based Telephony in Nigeria. *International Journal of Wireless Communications and Mobile Computing*, *1*(2), 58. https://doi.org/10.11648/j.wcmc.20130102.11

Ojugo, A. A., Ben-Iwhiwhu, E., Kekeje, O. D., Yerokun, M. O., and Iyawa, I. J. (2014). Malware Propagation on Social Time Varying Networks: A Comparative Study of Machine Learning Frameworks. *International Journal of Modern Education and Computer Science*, *6*(8), 25–33. https://doi.org/10.5815/ijmecs.2014.08.04

Ojugo, A. A., and Eboka, A. O. (2018). Comparative Evaluation for High Intelligent Performance Adaptive Model for Spam Phishing Detection. *Digital Technologies*, *3*(1), 9–15. https://doi.org/10.12691/dt-3-1-2

Ojugo, A. A., and Eboka, A. O. (2019). Signature-Based Malware Detection Using Approximate Boyer Moore String Matching Algorithm. *International Journal of Mathematical Sciences and Computing*, *5*(3): 49–62. https://doi.org/10.5815/ijmsc.2019.03.05

Ojugo, A. A., and Eboka, A. O. (2020a). An Empirical Evaluation On Comparative Machine Learning Techniques For Detection of The Distributed Denial of Service (DDoS) Attacks. *Journal of Applied Science, Engineering, Technology, and Education*, *2*(1), 18–27. https://doi.org/10.35877/454ri.asci2192

Ojugo, A. A., and Eboka, A. O. (2020b). Memetic algorithm for short messaging service spam filter using text normalization and semantic approach. *International Journal of Informatics and Communication Technology (IJ-ICT)*, *9*(1), 9. https://doi.org/10.11591/ijict.v9i1.pp9-18

Ojugo, A. A., Eboka, A. O., Okonta, E. O., Yoro, R. E., and Aghware, F. O. (2012). Genetic Algorithm Rule-Based Intrusion Detection System (GAIDS). *Journal of Emerging Trends In Computing Information Systems*, *3*(8), 1182–1194. http://www.cisjournal.org

Ojugo, A. A., Eboka, A. O., Yerokun, M. O., Iyawa, I. J., and Yoro, R. E. (2013). Cryptography: Salvaging Exploitations against Data Integrity. *American Journal of Networks and Communications*, *2*(2), 47. https://doi.org/10.11648/j.ajnc.20130202.14

Ojugo, A. A., Eboka, A. O., Yoro, R. E., Yerokun, M. O., and Efozia, F. N. (2015). Hybrid Model for Early Diabetes Diagnosis. *2015 Second International Conference on Mathematics and Computers in Sciences and in Industry (MCSI)*, *50*, 55–65. https://doi.org/10.1109/MCSI.2015.35

Ojugo, A. A., & Ekurume, E. O. (2021a). Predictive Intelligent Decision Support Model in Forecasting of the Diabetes Pandemic Using a Reinforcement Deep Learning Approach. *International Journal of Education and Management Engineering*, *11*(2), 40–48. https://doi.org/10.5815/ijeme.2021.02.05

Ojugo, A. A., and Ekurume, E. O. (2021b). Deep Learning Network Anomaly-Based Intrusion Detection Ensemble For Predictive Intelligence To Curb Malicious Connections: An Empirical Evidence. *International Journal of Advanced Trends in Computer Science and Engineering*, *10*(3): 2090–2102. https://doi.org/10.30534/ijatcse/2021/851032021

Ojugo, A. A., and Nwankwo, O. (2021a). Forging a Spectral-Clustering Multi-Agent Hybrid Deep Learning Model To Predict Rainfall Runoff In Nigeria. *International Journal of Innovative Science, Engineering and Technology*, *8*(3), 140–147.

Ojugo, A. A., and Nwankwo, O. (2021b). Spectral-Cluster Solution For Credit-Card Fraud Detection Using A Genetic Algorithm Trained Modular Deep Learning Neural Network. *JINAV: Journal of Information and Visualization*, *2*(1): 15–24. https://doi.org/10.35877/454RI.jinav274

Ojugo, A. A., and Nwankwo, O. (2021c). Modeling Mobility Pattern for the Corona-Virus Epidemic Spread Propagation and Death Rate in Nigeria using the Movement-Interaction-Return Model. *International Journal of Emerging Trends in Engineering Research*, *9*(6), 821–826. https://doi.org/10.30534/ijeter/2021/3096962021

Ojugo, A. A., and Nwankwo, O. (2021d). Tree-classification Algorithm to Ease User Detection of Predatory Hijacked Journals: Empirical Analysis of Journal Metrics Rankings. *International Journal of Engineering and Manufacturing*, *11*(4), 1–9. https://doi.org/10.5815/ijem.2021.04.01

Ojugo, A. A., and Obruche, C. O. (2021). Empirical Evaluation for Intelligent Predictive Models in Prediction of Potential Cancer Problematic Cases In Nigeria. *ARRUS Journal of Mathematics and Applied Science*, *1*(2), 110–120. https://doi.org/10.35877/mathscience614

Ojugo, A. A., Obruche, C. O., and Eboka, A. O. (2021). Quest For Convergence Solution Using Hybrid Genetic Algorithm Trained Neural Network Model For Metamorphic Malware Detection. *ARRUS Journal of Engineering and Technology*, *2*(1), 12–23. https://doi.org/10.35877/jetech613

Ojugo, A. A., and Otakore, O. D. (2018a). Improved Early Detection of Gestational Diabetes via Intelligent Classification Models: A Case of the Niger Delta Region in Nigeria. *Journal of Computer Sciences and Applications*, *6*(2), 82–90. https://doi.org/10.12691/jcsa-6-2-5

Ojugo, A. A., and Otakore, O. D. (2018b). Seeking Intelligent Convergence for Asymptotic Stability Features of the Prey / Predator Retarded Equation Model Using Supervised Models. *Computing, Information Systems,*

Development Informatics & Allied Research Journal, *9*(2): 13–26.

Ojugo, A. A., and Otakore, O. D. (2020a). Computational solution of networks versus cluster grouping for social network contact recommender system. *International Journal of Informatics and Communication Technology (IJ-ICT)*, *9*(3): 185. https://doi.org/10.11591/ijict.v9i3.pp185-194

Ojugo, A. A., and Otakore, O. D. (2020b). Investigating The Unexpected Price Plummet And Volatility Rise In Energy Market: A Comparative Study of Machine Learning Approaches. *Quantitative Economics and Management Studies*, *1*(3), 219–229. https://doi.org/10.35877/454ri.qems12119

Ojugo, A. A., and Oyemade, D. A. (2021). Boyer moore string-match framework for a hybrid short message service spam filtering technique. *IAES International Journal of Artificial Intelligence*, *10*(3), 519–527. https://doi.org/10.11591/ijai.v10.i3.pp519-527

Ojugo, A. A., and Yoro, R. E. (2020). Predicting Futures Price And Contract Portfolios Using The ARIMA Model: A Case of Nigeria's Bonny Light and Forcados. *Quantitative Economics and Management Studies*, *1*(4), 237–248. https://doi.org/10.35877/454ri.qems139

Ojugo, A. A., and Yoro, R. E. (2021a). Forging a deep learning neural network intrusion detection framework to curb the distributed denial of service attack. *International Journal of Electrical and Computer Engineering*, *11*(2), 1498–1509. https://doi.org/10.11591/ijece.v11i2.pp1498-1509

Ojugo, A. A., and Yoro, R. E. (2021b). Extending the three-tier constructivist learning model for alternative delivery: ahead the COVID-19 pandemic in Nigeria. *Indonesian Journal of Electrical Engineering and Computer Science*, *21*(3), 1673. https://doi.org/10.11591/ijeecs.v21.i3.pp1673-1682

Ojugo, A. A., Yoro, R. E., Eboka, A. O., Iyawa, I. J., and Yerokun, M. O. (2013). A Hybrıd Partıcle Swarm Genetıc Algorıthm (PSO-GA) For N-Queen Problem. *ARPN Journal of Science and Technology*, *3*(5), 538–546. http://www.ejournalofscience.org538

Ojugo, A. A., Yoro, R. E., Okonta, E. O., and Eboka, A. O. (2013). A Hybrid Artificial Neural Network Gravitational Search Algorithm for Rainfall Runoffs Modeling and Simulation in Hydrology. *Progress in Intelligent Computing and Applications*, *2*(1), 22–34. https://doi.org/10.4156/pica.vol2.issue1.2

Ojugo, A. A., Yoro, R. E., Oyemade, D. A., Eboka, A. O., Ugboh, E., and Aghware, F. O. (2013). Robust Cellular Network for Rural Telephony in Southern Nigeria. *American Journal of Networks and Communications*, *2*(5), 125. https://doi.org/10.11648/j.ajnc.20130205.12

Ojugo, A. A., Yoro, R. E., Yerokun, M. O., and Iyawa, I. J. (2013). Implementation Issues of VoIP to Enhance Rural Telephony in Nigeria. *Journal of Emerging Trends in Computing and Information Sciences ©2009-2013*,

4(2), 172–179. http://www.cisjournal.org

Okobah, I. P., and Ojugo, A. A. (2018). Evolutionary Memetic Models for Malware Intrusion Detection: A Comparative Quest for Computational Solution and Convergence. *International Journal of Computer Applications*, *179*(39), 34–43. https://doi.org/10.5120/ijca2018916586

Oyewola, D. O., Dada, E. G., Ngozi, N. J., Terang, A. U., and Akinwumi, S. A. (2021). COVID-19 Risk Factors, Economic Factors, and Epidemiological Factors nexus on Economic Impact: Machine Learning and Structural Equation Modelling Approaches. *Journal of the Nigerian Society of Physical Sciences*, *3*(4), 395–405. https://doi.org/10.46481/jnsps.2021.173

Paliwal, S., Mishra, A. K., Mishra, R. K., Nawaz, N., and Senthilkumar, M. (2022). XGBRS Framework Integrated with Word2Vec Sentiment Analysis for Augmented Drug Recommendation. *Computers, Materials and Continua*, *72*(3), 5345–5362. https://doi.org/10.32604/cmc.2022.025858

Rathi, M., and Pareek, V. (2013). Spam Mail Detection through Data Mining – A Comparative Performance Analysis. *International Journal of Modern Education and Computer Science*, *5*(12), 31–39. https://doi.org/10.5815/ijmecs.2013.12.05

Redondo-Gutierrez, L. Á., Jáñez-Martino, F., Fidalgo, E., Alegre, E., González-Castro, V., & Alaiz-Rodríguez, R. (2022). Detecting malware using text documents extracted from spam email through machine learning. *Proceedings of the 22nd ACM Symposium on Document Engineering*, 1–4. https://doi.org/10.1145/3558100.3563854

Sahmoud, T., and Mikki, D. M. (2022). *Spam Detection Using BERT*. https://doi.org/10.48550/arXiv.2206.02443

Sohony, I., Pratap, R., and Nambiar, U. (2018). Ensemble learning for credit card fraud detection. *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, 289–294. https://doi.org/10.1145/3152494.3156815

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

Udeze, C. L., Eteng, I. E., and Ibor, A. E. (2022). Application of Machine Learning and Resampling Techniques to Credit Card Fraud Detection. *Journal of the Nigerian Society of Physical Sciences*, 769. https://doi.org/10.46481/jnsps.2022.769

Verma, S., Bhatia, A., Chug, A., and Singh, A. P. (2020). *Recent Advancements in Multimedia Big Data Computing for IoT Applications in Precision Agriculture: Opportunities, Issues, and Challenges* (pp. 391–416). https://doi.org/10.1007/978-981-13-8759-3_15

Wang, D., Chen, B., and Chen, J. (2019). Credit card fraud detection strategies with consumer incentives. *Omega*, *88*, 179–195. https://doi.org/10.1016/j.omega.2018.07.001

Wemembu, U. R., Okonta, E. O., Ojugo, A. A., and Okonta, I. L. (2014). A Framework for Effective Software Monitoring in Project Management. *West African Journal of Industrial and Academic Research*, *10*(1): 102–115. http://www.ajol.info/index.php/wajiar/article/view/105798

Yoro, R. E., Aghware, F. O., Akazue, M. I., Ibor, A. E., and Ojugo, A. A. (2023). Evidence of personality traits on phishing attack menace among selected university undergraduates in Nigerian. *International Journal of Electrical and Computer Engineering (IJECE)*, *13*(2), 1943–1953. https://doi.org/10.11591/ijece.v13i2.pp1943-1953

Yoro, R. E., Aghware, F. O., Malasowe, B. O., Nwankwo, O., and Ojugo, A. A. (2023). Assessing contributor features to phishing susceptibility amongst students of petroleum resources varsity in Nigeria. *International Journal of Electrical and Computer Engineering*, *13*(2), 1922–1931. https://doi.org/10.11591/ijece.v13i2.pp1922-1931

Yoro, R. E., and Ojugo, A. A. (2019a). An Intelligent Model Using Relationship in Weather Conditions to Predict Livestock-Fish Farming Yield and Production in Nigeria. *American Journal of Modeling and Optimization*, *7*(2), 35–41. https://doi.org/10.12691/ajmo-7-2-1

Yoro, R. E., and Ojugo, A. A. (2019b). Quest for Prevalence Rate of Hepatitis-B Virus Infection in the Nigeria: Comparative Study of Supervised Versus Unsupervised Models. *American Journal of Modeling and Optimization*, *7*(2), 42–48. https://doi.org/10.12691/ajmo-7-2-2

Zawislak, P. A., Reichert, F. M., Barbieux, D., Avila, A. M. S., and Pufal, N. (2022). The dynamic chain of innovation: bounded capabilities and complementarity in agribusiness. *Journal of Agribusiness in Developing and Emerging Economies*. https://doi.org/10.1108/JADEE-04-2021-0096

Zhang, Y., Egelman, S., Cranor, L. F., and Hong, J. (2007). Phinding Phish: Evaluating Anti-Phishing Tools. *In Proceedings of the Network & Distributed System Security Symposium (NDSS 2007)*, *March*, 1–16.