IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 4, April 2020

A Novel Classification of IOT-Enabled Soil Nutrients Data using Artificial Neural Networks

Aamo Iorliam[*1], Adekunle Adeyelu[1], Samera Otor[1], Idengeli Okpe[1], and Iveren Iorliam[2]

Department of Mathematics & Computer Science, BSU, Makurdi, Nigeria[1]

CEFTER, BSU, Makurdi, Nigeria[2]

Corresponding author: Aamo Iorliam

Abstract: A vast number of IoT devices have been fabricated or adapted into different aspects of smart and precision farming to carry out a variety of tasks one of which is soil nutrients detection. While performing these tasks that are often recurrent, these devices generate different datasets which are stored on local memories or communicated remotely to cloud servers. The analysis of these data is important in order to correctly classify and group such data for device identification and differentiation. This is very important because the productivity of crop yields has greatly reduced due to lack of knowledge of the appropriate nutrients in a particular soil. Our research focuses on Nitrogen, Phosphorus, and Potassium, for the fact that most inorganic fertilizers consists majorly of these. As such, in this paper, soil nutrients values (Nitrogen, Phosphorus, and Potassium) are used as input features into the neural network for the classification of IoT-enabled soil nutrients data. Experimental analysis proved that the classification of these soil samples based on nutrients can achieve good accuracies between the range of 81.33% to 97.13%.

Keywords: IoT; Agriculture; Artificial Neural Networks

I. INTRODUCTION

To determine the fertility of any soil, the nutrients are used as an important property. Soil nutrients are usually divided into macro-nutrients and micro nutrients (Havlin, Tisdale, Nelson, and Beaton, 2016). The macro nutrients include Nitrogen (N), Phosphorus (P), Potassium (K), Calcium (Ca), Sulfur (S), Magnesium (Mg), Carbon (C), Oxygen (O), and Hydrogen (H). Furthermore, the micro nutrients (or trace minerals) include Iron (Fe), Boron (B), Chlorine (Cl), Manganese (Mn), Zinc (Zn), Cupper (Cu), Molybdenum (Mo), and Nickel (Ni) (Sinfield, Fagerman, and Colic, 2010). Even though, the above mentioned nutrients are needed for the effective growth of crops, most inorganic fertilizers constituents are usually Nitrogen, Phosphorus, and Potassium (Belay, Claassens, and Wehner, 2002).
Climate change, declined crop yield, increased population, improper use of fertilizer to boost soil fertility and improper or lack of information communications technology (ICT) applications to farming are some of the indices that has made the agricultural sector less effective in food production (Adisa, Botai, Adeola, Hassen, Botai, Darkey, and Tesfamariam, 2019; Dahikar, Rode, and Deshmukh, 2015). In recent years, the Internet-of-Things (IoT) has played a significant role as a modern technology for the development of tools and applications that connect devices geared towards the realization of a smart world. These connected intelligent devices have yielded satisfying results to the world's demand for intelligence. Smart phones, smart cars, smart homes and even smart cities are obvious breakthroughs of IoT. Applications developed based on IoT enabled devices have produced outstanding results in monitoring and control of different systems including agriculture, manufacturing, health and fitness, logistics, micro computing, engineering, medicine and the industries (Dan, Xin, Chongwei, and Liangliang,, 2015). The industrial IoT has disrupted several industrial activities of which agriculture is not an exception. Modern agriculture attempts to manage crops in controlled environments such as green houses in order to plan production or to duplicate specific weather conditions obtained in different regions locally. With a comprehensive analysis of internet of things soil nutrients data, it is possible to achieve highly accurate crops information that forms a background for informed decisions about duplicated climate factors for precise management and improved crop production while reducing environmental hazards.

Specifically, agricultural IoT has gained much recognition on the count of precision - the precision of the mechanisms, tools or gadgets that controls the crops and animal production (Stočes, Vaněk, Masner, and Pavlík, 2016). Smart planting, irrigation and harvesting systems as well as the enhancement of soil nutrients are very crucial towards achieving high yields.
It is no secret that Neural Networks (NNs) has achieved great success with numerous applications in agriculture, biometrics etc. (Patil, Al-Gaadi, Biradar, and Rangaswamy, 2012; Iorliam, Ho, Waller, and Zhao, 2016). Basically,

 DOI 10.17148/IJIREEICE.2020.8418

an ANN is a network of parallel distributed information processing system that relates an input vector to an output vector (De Coninck, Verbelen, Vankeirsbilck, Bohez, Simoens, Demeester, and Dhoedt, 2016). Some few researchers have classified IoT devices using machine learning models like linear support vector machine and artificial neural networks (ANN) (Patil, et al., 2012). Recently, Muangprathub, Boonnam, Kajornkasirat, Lekbangpong, Wanichsombat, and Nillaor, (2019), Ayaz, Ammad-Uddin, Sharif, Mansour, and Aggoune, (2019) and Junnarkar, (2020) have carried out some research concerning IOT. Muangprathub, et al., (2019) designed and developed a control system using node sensors in the crop field with data management via smartphone and a web application. This work was concerned with developing a system optimally watering agricultural crops based on a wireless sensor network. Furthermore, Ayaz, et al., (2019) carefully reviewed a state-of-the-art status of IOT and its current applications. While Junnarkar, (2020) studied the benefits of IoT in agriculture and provided awareness of the important role of IoT in agriculture well-tailored for India. Therefore, in this paper, the ANN is used for the classification of IoT-enabled soil nutrients data based on the soil nutrients.

We have used dataset from soil nutrients values of kaggle which is believed to be analysed by an IOT device available at: https://www.kaggle.com/surabhiremix/soilset. We used MATLAB programming language for this research. The important contribution of this study is the analysis of the effectiveness and efficiency of NN for the classification of IoT-enabled soil nutrients datasets. In section 2.0, we reviewed how neural networks has been applied to the IoT domain. The Materials and Methods is presented in section 3.0. Section 4.0 presents the Results and Discussions. The Conclusion and Future work are presented in section 5.0.

II. RELATED LITERATURE REVIEW

Patil et al. (2012) imagined IoT as a vision where “things” be it furniture, clothes, vehicles or home appliances etc. are readable, recognizable and located or controlled through the internet (Patil et al., 2012). The researchers looked at the concept of IoT beyond connecting just smart devices but also a possibility of placing sensors and chips on even non electronic devices to make them smart enough to communicate via the internet.

In another research inspired by the potential attacks targeting IoT networks, Hodo, Bellekens, Hamilton, Dubouilh, Iorkyase, Tachtatzis, and Atkinson (2016) conducted a research to analyze threats on IoT networks. The researchers used ANN as an offline IDS to gather and analyze information from various parts of an IoT network in an attempt to track an attack or possible flaw in the IoT network.

In a separate view, Brewster, Roussaki, Kalatzis, Doolin, and Ellis (2017), analyzed the applications of IoT on a large-scale pilot in agriculture, taking Europe into consideration. The researchers outlined the challenges and constraints that a large scale deployment of IoT in the agricultural domain is likely to face (Brewster, et al, 2017). Alam, Mehmood, Katib, and Albeshri (2016) conducted a study to explore and find out whether the conventional algorithms for data classification will work also with IoT datasets. A total of eight different classifiers were analyzed with datasets gotten from IoT devices. They classified human activities, robot navigation, body postures and movements based on IOT data. The researchers observed that ANNs are extremely efficient in solving data mining tasks with higher accuracy (Alam et al., 2016). The research however observed that despite the high accuracy demonstrated by artificial neural networks, ANN-based algorithms are “complex” requiring a widespread amount of computation before arriving at solutions with such higher accuracies. Suma, Samson, Saranya, Shanmugapriya, and Subhashri (2017) created a novel IOT device well suited for smart agriculture. Khadse, Mahalle, and Biraris (2018) analysed K-Nearest Neighbor, Naive Bayes, Decision Tree, Random Forest and Logistic Regression on IoT datasets. The decision tree algorithm gave the best accuracy of 99% among all the algorithms for all datasets. Foley, Moradpoor, and Ochen (2020) recently employed machine learning techniques to detect attacks in IoT devices using their developed novel dataset. This paper proposes a novel use of ANN for the classification of IoT-enabled soil nutrients data.

III. MATERIALS AND METHODS

The researchers used three (3) constituents (Nitrogen (N), Phosphorus (P) and potasium(K)) of the IoT-enabled soil nutrients datasets. These datasets are publicly available at: https://www.kaggle.com/surabhiremix/soilset. The collected soil samples from India are usually analysed in the soil testing laboratory which contains devices such as atomic absorption spectroscopy, and other several IOT devices (Soilhealth.dac.gov.in, 2020). Even though the description of this dataset is not properly documented, a careful search of other IoT-enabled soil nutrients dataset online shows that, we could not find any better one as compared to the publicly available one from kaggle. Hence, the reason for using this dataset for our experiment. Figure 1 shows a sample of the input data used for this experiment. Columns A, B and C indicate Nitrogen, Phosphorus and Potassium, respectively.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 4, April 2020

	A	B	C
1	Nitrogen	Phosphorus	Potassium
2	191.000	19.0000	366.000
3	191.000	19.0000	366.000
4	191.000	19.0000	366.000
5	191.000	19.0000	366.000
6	141.000	38.0000	249.000
7	141.000	38.0000	249.000
8	192.000	38.0000	432.000
9	192.000	38.0000	432.000
10	192.000	38.0000	432.000

Figure 1: Sample Input Data

In order to classify the data gotten from the agricultural based IoT-enabled devices, an artificial neural network model was used. The NN works by accepting input data, extracting rules based on the accepted inputs and then making classification decisions (Iorliam, et al., 2016). The Multi-layer perceptron (MLP) which is used as a feed-forward type of NN has one or more hidden layers between the input layer and output layer.

Training in this case is carried out using back-propagation learning algorithm. MLP has the advantage of modelling functions that are highly non-linear and when trained, it classifies the soil nutrients in an agricultural farm based on input data from IoT device datasets.

Figure 2 is a pictorial representation of an MLP with two inputs, one hidden layer and two MLP with two inputs, one hidden layer and two outputs.

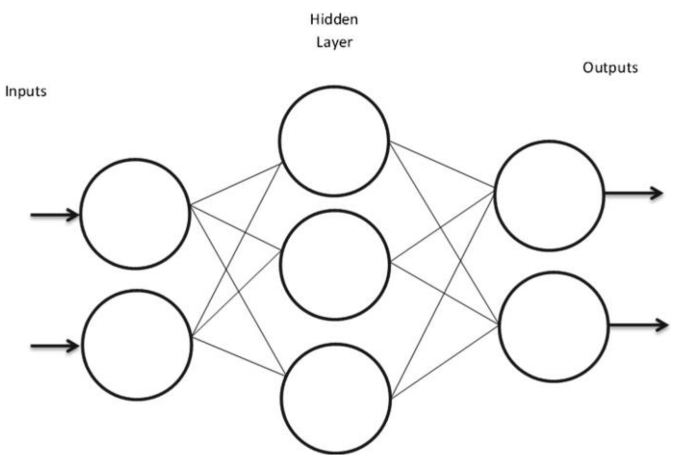

Figure 2: A Multilayer perception with one hidden layer (Iorliam, et al., 2016)

The MLP typically uses the sigmoid hidden neurons which are described by the formula shown in equation 1.

$$y = \frac{1}{1+\exp(-s)} \qquad \text{(eqn. 1)}$$

Where s is an integration function defined by $s = f(x; \theta)$, y is the output value of the trained network utilizing sigmoid hidden neurons. After the neural network is trained and tested, performance evaluation metrics such as the cross entropy (CE) and the percentage error (%E) are used. These evaluation metrics are used to determine how well the network has achieved with respect to classification. The CE and %E are used to determine the performance accuracy at each set (training, validation and testing). The classification accuracy is gotten as 100 – CE or 100 - %E when considering CE or %E.

In summary, the experimental methodology is shown in the steps below:

i. Collect raw IoT-enabled soil nutrients values from any soil type.
ii. Extract Nitrogen, Phosphorus and Potassium values from the raw data in (i).
iii. The values from (ii) are fed into an artificial neural network for simulation.
iv. The simulation in (iii) is performed using MATLAB 2019b.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 4, April 2020

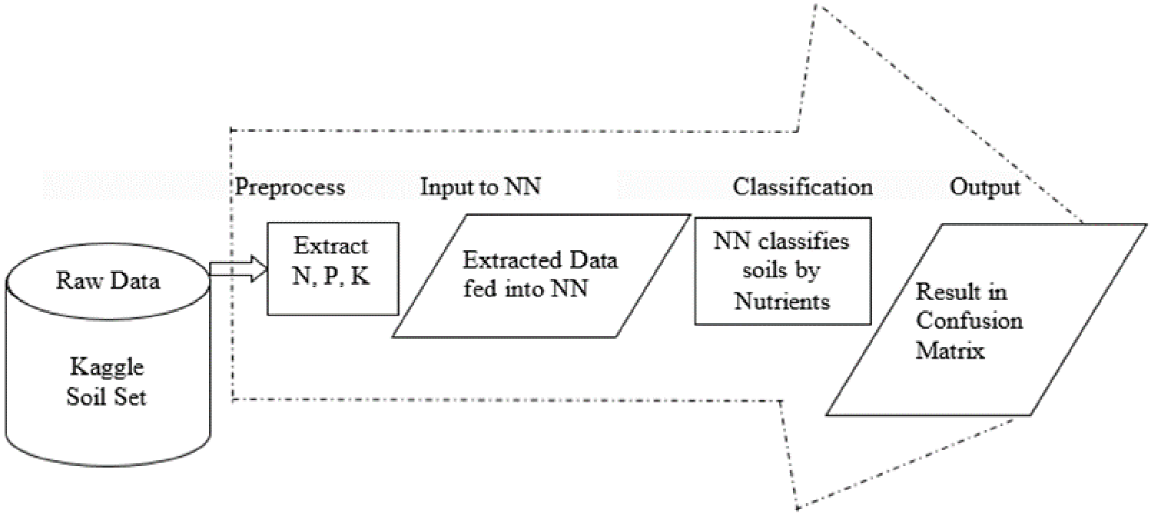

Figure 3: Artificial Network Methodology

v. The performance evaluation is shown in terms of confusion matrix, CE, %E, and classification accuracy.
The artificial neural network classification methodology is depicted in Figure 3.

IV. RESULTS AND DISCUSSION

Datasets gotten from nutrients monitoring IoT devices were classified by the neural network classifier. The hidden layer in the network was able to classify the data into three distinct classes based on soil nutrients for the training, validation and testing phases of the classification. The neural network randomly divided 5820 samples of the datasets into the proportion of 70% (4074) for the training, 15% (873) for the validation and 15% (873) for the testing. The training samples were presented during training such that the neural network can adjust the error that occurs as a result of the deviations (inconsistent patterns) in the data samples. The validation samples are a measure of the generalizations done by the training sample and serve as sentinel value or flag to stop or halt the training if training generalizations fail to improve during the training phase of the neural network.

The neural network architecture does pattern recognition by taking an input and doing classification in the hidden layer, i.e training, validating and testing based on the specified number of hidden layers. The hidden layer after classsification produces a processed output that looks familiar to the input that has been processed by the hidden layer. As earlier stated, 4074 (70%) data samples of the total 5820 data samples were separated for training the neural network. The remaining 1744 (30%) data samples were evenly divided into two; 15% for validation and 15% for testing which means that for each of validation and testing, 873 data samples equivalent to 15% of the total sample of 5820 datasets were assigned. The training automatically stops when the generalizations stop improving as an indication that there is an increase in the cross entropy error (CE) in validation of the samples. This means that minimizing the cross entropy results in good classification and increasing the cross entropy results in a poor classification. The percent error (%E) indicates a fraction of the data samples that has been misclassified. An error of 100 indicates maximum misclassification of the data by the neural network model while an error of 0 indicates no misclassification by the neural network model. Basically, multiple training of dataset will generate different results due to different initial conditions and sampling but a consistency will always exist in the cross entropy error and the percentage error indicating the extent of correct classification or misclassifications.

The separation (classification) of the soil nutrients is explained by the neural network confusion matrix. The training, validation, test confusion matrices as well as the overall all confusion matrix are divided into three classes; class 1, 2 and 3 for each of the soil nutrients per village. These nutrients are Nitrogen, Phosphorus and Potassium contained in each soil type. Figure 4 shows the percentage of correct classifications and misclassification of every class in each of the training, validation and testing performed. The confusion matrix corresponds to three classes; class 1 represent soil nutrients 1 (soil containing Nitrogen, Phosphorus or Potassium), class 2 represent soil nutrients 2 (soil containing Nitrogen, Phosphorus or Potassium) and class three represent soil nutrients 3 (containing either Nitrogen, Phosphorus or Potassium). These classes are classified at every stage of training, validation and testing with an overall classification performance presented by the all confusion matrix in the neural network.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 4, April 2020

In Figure 4, for the training set, the confusion matrix correctly classified 359 as class 1, 16 are misclassified as class 2 and 67 are misclassified as class 3. 128 soil in class 2 are correctly classified, whereas 16 are wrongly classified as class 1 and 56 are wrongly classified as class 3.

Furthermore, 2739 are correctly classified as class 3, where 67 are wrongly classified as class 1 and 549 are wrongly classified as class 2. The overall classification accuracy for the training set is 79.2%. For the validation confusion matrix, 81 are correctly classified as class 1, while 3 are wrongly classified as class 2, and 9 are wrongly classified as class 3. Furthermore, 34 are correctly classified as class 2, 3 are wrongly classified as class 1 and 18 are wrongly classified as class 3. Again, 570 are correctly classified as class 3, whereas 9 are wrongly classified as class 1, and 118 are wrongly classified as class 2. The overall classification accuracy for the validation set is 78.5%.

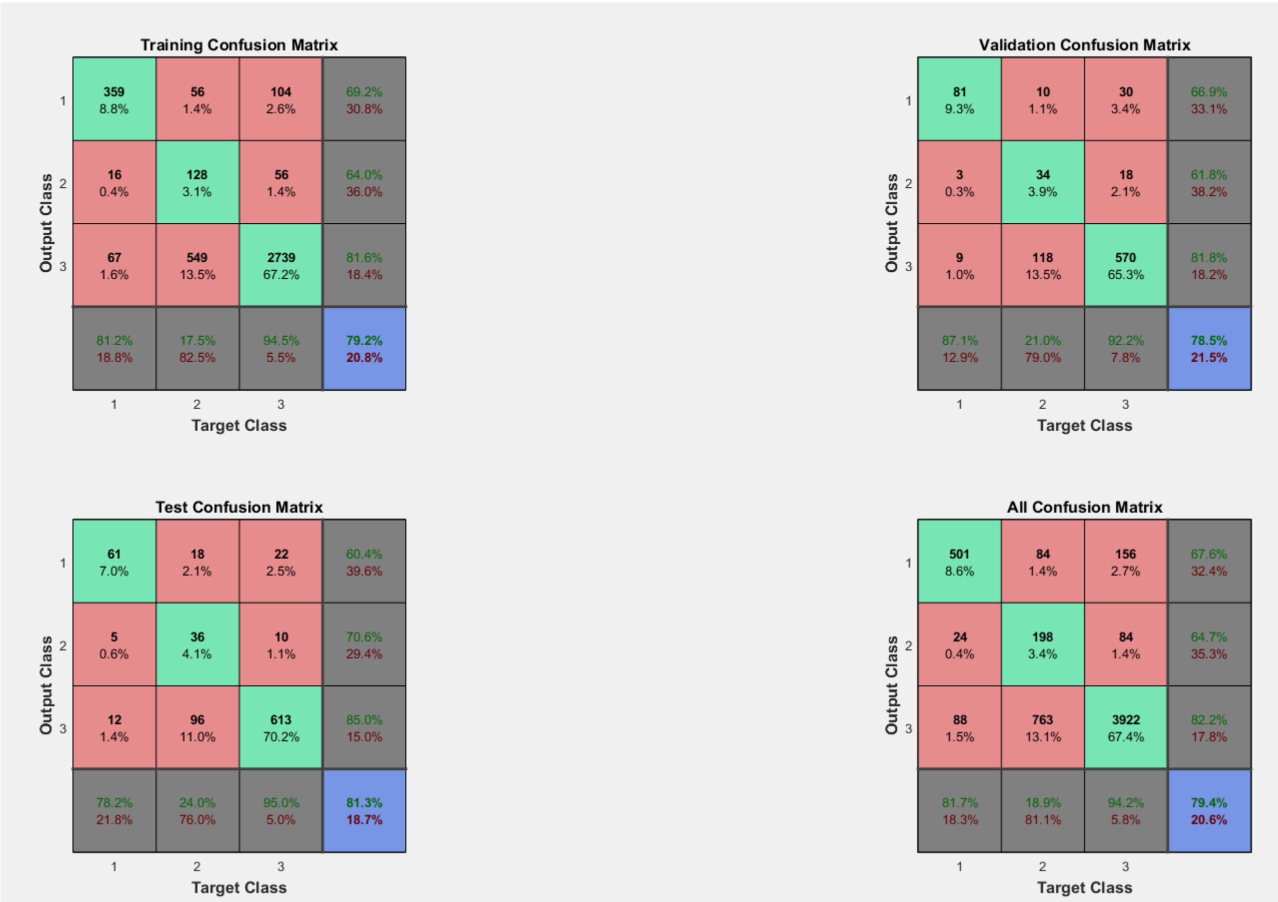

Figure 4: Classification Confusion Matrix

For the test confusion matrix, 61 are correctly classified as class 1, 5 are wrongly classified as class 2, and 12 are wrongly classified as class 3. Furthermore, 36 are correctly classified as well 2, 5 are wrongly classified as class 1 and 10 are wrongly classified as class 3. Again, 613 are correctly classified as class 3, 12 are wrongly classified as class 1 and 96 are wrongly classified as class 2. The overall accuracy for the test confusion matrix is 81.3%.
Overall, the all confusion matrix achieves an accuracy for the classification of three (3) soils based on their nutrients up to 79.4%.
This implies that between the first, second and third classes at each stage of the classification, there is always a variation in the classification performance (correctly classified or misclassified) of either training, validation or testing which is an indication that the soil nutrients classified in the first class is not same as the soil nutrients classified in the second class and likewise the third class. In general, the soil nutrients can be used to correctly classify different IOT devices that collected such nutrients.
The variation in the cross entropy error (CE) and percentage error (%E) for training, validation and testing as summarized in Table 1 is another indication that the neural network classifier performed classification of data produced by IoT-enabled soil nutrients data.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 4, April 2020

Table 1: CE and %E Results

	Samples	CE	%E
Training	4074	1.11685e-0	20.81492e-0
Validation	873	2.85720e-0	21.53493e-0
Testing	873	2.86707e-0	18.67124e-0

From Table 1, it can be seen that the neural network (NN) has correctly classified approximately 98.88% (when considering CE) and about 79.19% (when considering %E) for the training set. For the validation set, the NN has classifies correctly about 97.14% (when considering CE) and about 78.47% (when considering %E). Similarly, for the testing set, the NN classifies correctly about 97.13% (when considering CE) and about 81.33% (when considering %E).
Based on the %E, it can be observed that the highest and possibly best classification result is achieved at the testing set with a classification accuracy of 81.33% while the least classification accuracy of 79.19% is achieved at the training set. Based on cross entropy error (CE) a total of 4074 data samples produced an error as low as 1.12 at the testing set. 873 data samples of the same data produced an approximate error of 2.9 at the validation and testing sets. Thus, it can be inferred that the more the data, the less the cross entropy.

V. CONCLUSION

This research is a novel approach for the modeling of IoT-enabled soil nutrients data. The research is able to correctly classify three (3) agricultural IoT-enabled data based on soil nutrients using neural networks. The neural network model used was trained, validated and tested using an IoT sample data size of 5820 sets. The IoT generated dataset was divided into three proportions of 70%:15%:15% for the training, validation and testing sets respectively. The neural network model was able to classify correctly with an accuracy of 81.33% based on the %E and 97.13% based on the CE at the testing phase of the experiment. Based on the analysis carried out in this research, it can be concluded that IoT-enabled data generated by agricultural devices can be correctly classified using neural networks. In this case three classes of soil nutrients were distinctively and accurately classified to belong to class 1, class 2 and class 3. For our future work, the researchers will extend this research to classify more agricultural IoT data. Furthermore, we will use other machine learning techniques such as linear regression, logistic regression, and decision tree to classify different agricultural IoT data. Lastly, there is need for collection of IoT-enabled soil nutrients data in all the 36 states and FCT in Nigeria for similar investigations and analysis.

REFERENCES

[1]. Adisa, O., Botai, J., Adeola, A., Hassen, A., Botai, C., Darkey, D., & Tesfamariam, E. (2019). Application of Artificial Neural Network for Predicting Maize Production in South Africa. *Sustainability*, *11*(4), 1145.
[2]. Alam, F., Mehmood, R., Katib, I., & Albeshri, A. (2016). Analysis of eight data mining algorithms for smarter Internet of Things (IoT). *Procedia Computer Science*, *98*, 437-442.
[3]. Ayaz, M., Ammad-Uddin, M., Sharif, Z., Mansour, A., & Aggoune, E. H. M. (2019). Internet-of-Things (IoT)-Based Smart Agriculture: Toward Making the Fields Talk. IEEE Access, 7, 129551-129583.
[4]. Belay, A., Claassens, A., & Wehner, F. C. (2002). Effect of direct nitrogen and potassium and residual phosphorus fertilizers on soil chemical properties, microbial components and maize yield under long-term crop rotation. Biology and Fertility of Soils, 35(6), 420-427.
[5]. Brewster, C., Roussaki, I., Kalatzis, N., Doolin, K., & Ellis, K. (2017). IoT in agriculture: Designing a Europe-wide large-scale pilot. *IEEE communications magazine*, *55*(9), 26-33.
[6]. Dahikar, S. S., Rode, S. V., & Deshmukh, P. (2015). An Artificial Neural Network Approach for Agricultural Crop Yield Prediction Based on Various Parameters. *published in International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, *4*(1).
[7]. Dan, L. I. U., Xin, C., Chongwei, H., & Liangliang, J. (2015). Intelligent agriculture greenhouse environment monitoring system based on IOT technology. In *2015 International Conference on Intelligent Transportation, Big Data and Smart City* (pp. 487-490). IEEE.
[8]. De Coninck, E., Verbelen, T., Vankeirsbilck, B., Bohez, S., Simoens, P., Demeester, P., & Dhoedt, B. (2015). Distributed neural networks for Internet of Things: the Big-Little approach. In *International Internet of Things Summit* (pp. 484-492). Springer, Cham.
[9]. Foley, J., Moradpoor, N., & Ochen, H. (2020). Employing a Machine Learning Approach to Detect Combined Internet of Things Attacks against Two Objective Functions Using a Novel Dataset. *Security and Communication Networks*, *2020*.
[10]. Havlin, J. L., Tisdale, S. L., Nelson, W. L., & Beaton, J. D. (2016). *Soil fertility and fertilizers*. Pearson Education India.
[11]. Hodo, E., Bellekens, X., Hamilton, A., Dubouilh, P. L., Iorkyase, E., Tachtatzis, C., & Atkinson, R. (2016). Threat analysis of IoT networks using artificial neural network intrusion detection system. In *2016 International Symposium on Networks, Computers and Communications (ISNCC)* (pp. 1-6). IEEE.
[12]. Iorliam, A., Ho, A. T., Waller, A., & Zhao, X. (2016). Using benford's law divergence and neural networks for classification and source identification of biometric images. In *International Workshop on Digital Watermarking* (pp. 88-105). Springer, Cham.
[13]. Junnarkar, S. (2020). An Empirical Study of Applications of Internet of Things (IOT) in Agriculture. Studies in Indian Place Names, 40(51), 76-80.
[14]. Khadse, V., Mahalle, P. N., & Biraris, S. V. (2018). An empirical comparison of supervised machine learning algorithms for internet of things data. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)* (pp. 1-6). IEEE.
[15]. Muangprathub, J., Boonnam, N., Kajornkasirat, S., Lekbangpong, N., Wanichsombat, A., & Nillaor, P. (2019). IoT and agriculture data analysis for smart farm. Computers and electronics in agriculture, 156, 467-474.

[16]. Patil, V. C., Al-Gaadi, K. A., Biradar, D. P., & Rangaswamy, M. (2012). Internet of things (IoT) and cloud computing for agriculture: An overview. *Proceedings of Agro-Informatics and Precision Agriculture (AIPA 2012), India*, 292-296.
[17]. Sinfield, J. V., Fagerman, D., & Colic, O. (2010). Evaluation of sensing technologies for on-the-go detection of macro-nutrients in cultivated soils. *Computers and Electronics in Agriculture*, *70*(1), 1-18.
[18]. Soilhealth.dac.gov.in (2020). Technical Specifications of the Instrument/Equipment. https://soilhealth.dac.gov.in/Content/UserManual/Instrument%20Specification_WB.pdf.
[19]. Stočes, M., Vaněk, J., Masner, J., & Pavlík, J. (2016). Internet of things (iot) in agriculture-selected aspects. *Agris on-line Papers in Economics and Informatics*, *8*(665-2016-45107), 83-88.
[20]. Suma, N., Samson, S. R., Saranya, S., Shanmugapriya, G., & Subhashri, R. (2017). IOT based smart agriculture monitoring system. *International Journal on Recent and Innovation Trends in computing and communication*, *5*(2), 177-181.