Machine learning for the detection of soil pH, macronutrients, and micronutrients with crop and fertilizer recommendations

John Joshua Montañez[1,2], Jeffrey Sarmiento[2]
[1]College of Engineering, Bicol State College of Applied Sciences and Technology, Naga City, Philippines
[2]College of Engineering, Batangas State University, the National Engineering University, Batangas City, Philippines

Article Info

Article history:

Received Mar 19, 2024
Revised Oct 10, 2024
Accepted Oct 18, 2024

Keywords:

Colorimetry
Crop and fertilizer
Machine learning
Soil parameters
Soil test kit

ABSTRACT

The study aims to determine the levels of soil parameters such as soil pH, macronutrients, and micronutrients. After determining said parameters, the system appropriately recommends crops and fertilizers suitable for the soil samples. For soil pH and macronutrient levels, i.e., nitrogen, phosphorus, and potassium, these parameters can be detected using the soil test kit. Meanwhile, for soil micronutrients, i.e., copper, iron, and zinc, there is a need for the development of appropriate assays for colorimetric processes that can be done for the appropriate determination of said micronutrients. Comparison of available machine learning such as support vector machine algorithm, naïve Bayes algorithms, and K-nearest neighbor algorithm is a must to determine the well-fit algorithm that is considered fast and has high predictive power in classification and regression. The outputs of the colorimetric and spectrometric processes are the inputs in the machine learning activities intended for crop and fertilizer recommendation.

Corresponding Author:

John Joshua Montañez
College of Engineering, Bicol State College of Applied Sciences and Technology
Peñafrancia Avenue, Naga City, Camarines Sur, Philippines
Email: jjfmontanez@astean.biscast.edu.ph

1. INTRODUCTION

In the worldwide setup, agriculture is considered a vital industry that continues to thrive considerably with us in the foreseeable future. The incorporation of technology in agriculture is inevitable since the role of technology is seen as favorable to the advancement of humankind, and this is visible in advances in monitoring soil parameters like humidity, temperature, and moisture, leading to an increase in the production of high-value crops. Technically, the integration of the advancement of technology, various protocols, and advances in computational paradigms in agriculture that seek to increase yields in crop production is called smart agriculture. Smart agriculture, smart farming, and agriculture 4.0 are used interchangeably [1]–[4]. Implementing various agricultural processing controlled by the internet, internet of things (IoT) represents smart agriculture, robotics, big data analytics, unmanned aerial vehicles, and artificial intelligence via machine learning and deep learning algorithms [5]–[7]. Smart agriculture, through the inculcation of information communication technology manifested in the management of farms and other areas of implementation of clean and efficient agro-industry processes, paves the way for a sustainable future of food production for the growing population worldwide. In terms of the sustainable development goals (SDGs) of the United Nations (UN), smart agriculture, in its goal to have a high and clean increase in yields, directly and certainly addresses SDG number two, which is zero hunger [6]–[8].

Success in intelligent agriculture, especially in crop production, can be attributed to proper maintenance of the soil as it is considered a core component of the environment. Thus, a substantial effort must be made to preserve it. Soil productivity and soil fertility must be studied and analyzed as this is proportional to the crop and yields. The soil paraments such as soil pH, macronutrients, i.e., nitrogen, phosphorus, and potassium, and micronutrients, e.g., iron, copper, and zinc, and their levels should be monitored and determined so that appropriate intervention must be deployed in order to avoid polluting or abusing the soil [9], [10]. Determination of said soil parameters can be done in several ways, and one efficient means is using machine learning algorithms, which are efficient tools that can be used ranging from classification up to the performance of regression of various sets of thousands of data.

In the proposed model of Reddy *et al.* [11], the decision tree algorithm, a machine learning algorithm, was implemented to address the crop's water requirement in the advancement of smart irrigation. The three essential parameters that were used in the system are humidity, temperature, and moisture. Corresponding sensors for the said parameters were deployed for testing. A microprocessor controlled the said sensors. In Balne's investigation, multiple linear regression was utilized. Developing an application was deemed appropriate which concentrates on several smart agricultural parameters like soil dampness, the temperature of the soil surface, weather viewpoint, i.e., rainfall and temperature, and soil nutrients. With all the necessary parameters, the information investigation method was the research methodology implored in the study [12].

A review study on machine learning and IoT about agricultural management was conducted by Maduranga and Abeysekera [13]. The review study focuses on machine learning applications in IoT-based agriculture: plant management, crop and yield management, disease management, weed management, water management, and animal tracking. These applications can be done using widely used algorithms like the naïve Bayes algorithm, support vector regression, and K-nearest neighbor. The said review study further stated that the hybrid application of machine learning and IoT will benefit the advances of agriculture worldwide.

Salim and Mitton [14] proposed a system incorporating machine learning and wireless sensor networks for smart agriculture. The large amount of data produced by the network sensors causes a decrease in the system's life span; therefore, the machine learning-based data reduction algorithm was made. The algorithm was based on data from the environment that benefitted agriculture in the long run. Moreover, the said machine learning algorithm aided in reducing the energy consumption of the wireless sensor network. MATLAB was used to validate the developed algorithm, resulting in a data reduction of 70% being sent. The system was rated to have excellent accuracy.

In the study conducted by Abraham *et al.* [15], two machine learning algorithms were compared and implemented in the application-based monitoring of farm conditions and in the classification of animals in the agricultural field. Convolutional neural networks (CNN) and support vector machine (SVM) algorithms were compared in the matter of classification, especially for image inputs. The CNN was concluded to perform well compared to the SVM algorithm. The inputs from the classifications done by machine learning were incorporated into the water and sprinkler monitoring system.

Maximizing irrigation for crop production using various machine-learning algorithms was the main focus of the study performed by Fredj *et al.* [16]. The machine learning algorithms gradient boost regressor, decision tree, XGBRegressor, and random forest, which were programmed through Python, were compared to predict the amount of water needed by the crop for optimal productivity of irrigation. The performance of the machine learning algorithm was evaluated through the use of mean squared error (MSE), root mean square error (RMSE), and mean absolute error (MAE). The study concluded that XGBRegression was the most efficient algorithm, while random forest was the least efficient.

Another study concentrating on the development of intelligent irrigation was accomplished by Bhanu *et al.* [17]. The intelligent irrigation was implemented through the ThingSpeak cloud, an IoT Cloud. It utilized the naïve Bayes algorithm to classify the sensed data from parameters like atmospheric temperature data, humidity data, soil moisture, and soil temperature. When the sensed data is below the threshold, the system will email the user and prompt an appropriate response. The naïve Bayes algorithm gave 76.47% accuracy in terms of classification.

Padarian *et al.* [18] reviewed several works of literature on the application of machine learning as a tool in the progress of agriculture, especially in soil science. The review is comprehensive in that it determined the performance of the basic machine learning algorithm up to hybrid machine learning, with a focus on soil analysis. Neural networks, SVM, and random forests performed better than simpler algorithms such as partial least square regression, principal component regression, multiple linear regression, and K-nearest neighbor. The machine learning algorithm aided the advancement of soil science in modeling categorical and continuous soil properties. The non-linear relationship among presented parameters was determined through the high-performing machine learning algorithms, thus making them more accurate than other basic machine learning algorithms.

Estimation of the soil moisture from remote sensing data through machine learning was the focus of the study of Adab *et al.* [19]. Determination of the soil moisture content is tantamount to the production of more crops and yields. The study compared the performance of several machine learning algorithms, such as elastic net regression, SVM, and random forest, in retrieving information on soil moisture from thermal and optical sensors. The random forest machine learning algorithm regarding the nash-sutcliffe efficiency was noted and interpreted as high with a value of 0.73.

With the aid of a machine learning algorithm, Dash *et al.* [20] conducted a study on classifying and mapping rice, wheat, and sugar cane-based on several soil parameters and meteorological data such as soil radiation, soil temperature, soil humidity, soil pH level, and rainfall data. The study compared machine learning algorithms: SVM with linear kernels, linear SVM, and decision trees. Among the listed machine learning algorithms, SVM with linear kernels obtained a 92% accuracy level in developed model forecasts. Performing non-linear curve fitting in the training data will increase the percent of accuracy of the model.

Farwa *et al.* [21] conducted a study predicting the soil macronutrient using machine learning. The study utilized several regression models, such as linear regression, ridge regression, lasso regression, elastic net regression, and Bayesian ridge regression. The soil parameters tested were the pH level, cation exchange capacity, soil macronutrients like nitrogen, phosphorus, potassium, soil moisture, and temperature. The Bayesian and ridge regression were the highest-performing machine learning algorithms in the concluded soil analysis.

Several studies have incorporated various techniques in machine learning algorithms that have been applied in different industries and combined with other advanced existing technologies. The internet, big data analytics, robotics, wireless sensor networks, and unmanned aerial vehicles utilize machine learning. It maximized the use of machine learning algorithms, improving the efficiency and accessibility of data analysis [22]–[28].

This study aims to detect soil pH level, soil macronutrients, i.e., nitrogen, phosphorus, potassium, and soil micronutrients, e.g., iron, copper, and zinc, using a machine learning algorithm. For soil pH and soil macronutrients, the soil test kits given by the Department of Agriculture provide the inputs for the machine learning algorithm based on the successful performance of the procedures. Developing an appropriate assay that implores a colorimetric process for soil micronutrients is a must for iron, copper, and zinc. The detection of the said parameters primarily relies on the image produced from the colorimetric processes. After successfully performing the chemical processes, the inputs determine the appropriate crop and fertilizer recommendation based on the soil sample.

2. MACHINE LEARNING

Developing an appropriate assay that implores a colorimetric process for soil micronutrients is a must for iron, copper, and zinc. The image produced from the colorimetric processes plays a pivotal role in detecting the parameters. The inputs determining the appropriate crop and fertilizer recommendation, considering the soil sample, are followed after successfully performing the chemical processes. This section discusses the advantages of using machine learning algorithms compared to the conventional method of determining the level of soil parameters. Furthermore, several machine learning algorithms can be employed to achieve this study's goals are compared.

2.1. Machine learning vs conventional method

The conventional method of determining the soil parameter levels depends on the color-matching ability of the users of soil test kits subjected to different lighting conditions, thus relying on the optical perception of the farmers or other interested individuals. After performing the prescribed process from the soil test kit, the individual must compare the color produced by the process to the color chart, as color charts are standard in soil testing. This can be attributed to a subjective process of determination, which can be eliminated by optimizing the machine. Any misinterpreted values of the level of soil parameters can cause soil pollution or abuse, thus lessening crop and yield production. Producing a considerable amount of data set to be trained is a requirement to ensure the validity and accuracy of classification and regression upon implementation. It is implied to have images, considered the significant data subjected to training, from various lighting environments to eliminate considerations for issues from lighting. The device will independently determine the said levels using machine learning and take into consideration the best among the selected machine learning algorithms.

2.2. Prospected machine learning algorithm

2.2.1. Naïve Bayes algorithm

The naïve Bayes algorithm is a machine learning technique that classifies data using statistical models to estimate the likelihood of a class having a particular attribute. It applies Bayes' Theorem, assuming independence among features, to predict outcomes. Additionally, the algorithm seeks optimal classification across all classes and their corresponding attributes [29]–[32].

2.2.2. Support vector machine algorithm

SVM is a machine learning algorithm that uses kernel functions for regression and classification tasks. It solves quadratic optimization problems to identify the best hyperplane for separating data into distinct classes. The algorithm focuses on a subset of support vectors, a small portion of the training data, for decision-making [33]–[36].

2.2.3. Decision tree algorithm

The decision tree algorithm is a nonparametric supervised learning method widely used for classification and regression tasks. It makes no assumptions about the data distribution and builds a tree structure by recursively splitting the data based on feature values. This algorithm is trained on labeled data to classify new, unseen instances accurately [37]–[40].

3. MATERIALS AND METHODS

This section outlines the research methodologies that will be employed throughout the study. It details the specific approaches and techniques used to ensure the accurate achievement of research objectives. Additionally, it highlights the materials and tools essential for meeting the study's goals.

3.1. Research methodology

The research methodology implored in this study is mainly quantitative and coupled with an experimental research design. The approach in experimental research design started with data collection and was followed by data pre-processing, e.g., data cleansing. Several data were involved in the model training, followed after the data had been pre-processed since it must capture the necessary data features. The model used in classification and regression will be the primary product of model training. The developed model was tested to assess its accuracy, given the other data to be tested for classification and regression. The experimental research design concludes in the model evaluation as it is assessed through several key performance indicators set by the researchers and other standards available. Optimal feature selection is one of the trending research methodologies used in machine learning and is usually done under the data pre-processing potion of experimental research. Through this research methodology, incorporating this selection, the research can observe the performance of the machine learning algorithm in an optimized fashion.

3.2. Methodology process flowchart

The methodology process flowcharts present the direction in the achievement of the objectives of the study. Chemical and electronic design were the main processes in the flowchart, as reflected in Figures 1 and 2. These main processes were observed in determining soil pH, macronutrients, and micronutrients.

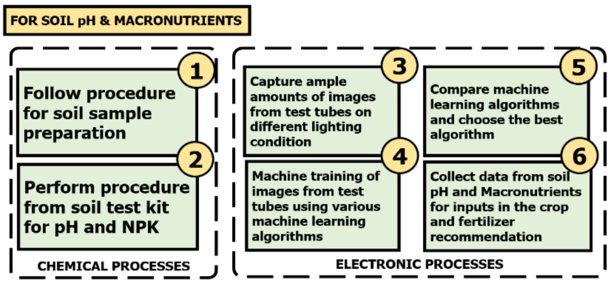

Figure 1. Methodology process flowchart for determination of soil pH and macronutrients

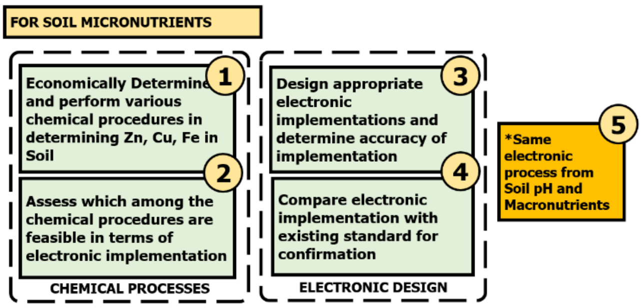

Figure 2. Methodology process flowchart for determination of soil macronutrients

3.2.1. Soil pH and macronutrients

The soil test kit is essential in determining soil pH and macronutrients. This specifies the procedure to be followed to prepare soil samples. After successfully gathering soil samples, the soil tests followed, which concluded the chemical processes, i.e., pH, nitrogen, potassium, and phosphorus tests. Capturing the images generated from the soil test performance observed in the test tubes initiates the electronics processes. The data pre-processing portion includes cleaning and assessing the data to see if it suits the machine learning algorithm parameters. Several images must meet the necessary amount needed for the model training of the machine learning algorithm. Model training utilizes various machine learning algorithms, i.e., naïve Bayes algorithm, SVM algorithm, and decision tree algorithm, where the data were subjected to after successful data pre-processing. A comparison of the machine learning algorithm's performance was observed to determine which of the three machine learning algorithms is considered the highest-performing algorithm in classification and regression. Lastly, upon proper determination of the soil pH, nitrogen, phosphorus, and potassium levels, the levels served as input in the crop and fertilizer recommendations of the study.

3.2.2. Soil micronutrients

In determining soil micronutrients, i.e., iron, copper, and zinc, no readily available tools or devices incorporate colorimetric processes. Observing various chemical processes requires the economical determination of the said elements or compounds in the soil sample. After the chemical processes have been done, an assessment of the end product from the chemical processes must be done to determine whether it is applicable for electronic implementation. After finishing the necessary adjustment of the electronic parameters, the evaluation of the electronic implementation followed, which was more comprehensive than comparing existing standards found in the industry and governmental instrumentalities.

3.3. Performance metrics of machine learning algorithm

The assessment of the machine learning algorithm can be achieved using four performance metrics: accuracy, precision, recall, and F1-score. These are the usual performance metrics to assess the classification capabilities of the three selected machine learning algorithms. Listed as follows is the corresponding equation:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

$$Precision = TP/(TP + FP) \quad (2)$$

$$Recall = TP/(TP + FN) \quad (3)$$

$$F1 - Score = TP/(TP + 1/2(FP + FN)) \quad (4)$$

where TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives.

3.4. Materials used

Soil test kit. It is a kit developed by the Bureau of soils and water of the Department of Agriculture, which contains chemicals that field technicians and farmers use to study the soil samples from the farmland. Using these soil test kits, time-efficient determination of the nutrients will be observed. The soil test kit includes the procedure of proper soil sampling, the procedures for the various soil tests, and crop and fertilizer recommendations, as depicted in Figure 3. It is cost-efficient since the farmers can readily determine the levels of soil parameters without paying laboratory fees.

Figure 3. Soil test kit from the Department of Agriculture

4. RESULTS AND DISCUSSION

The comparative analysis of the different machine learning algorithms utilizing the abovementioned four performance metrics is showcased in Figure 4. Regarding the performance metrics, the SVM was considered the highest-performing machine learning, with garnered scores of 0.862 for accuracy, 0.889 for precision, 0.864 for recall, and 0.891 for F1-score. The SVM outranked the other two machine learning in all performance metrics. The naïve Bayes algorithm ranked second and last for the decision tree. Therefore, the SVM algorithm was considered the highest-performing machine learning algorithm among the three. The system correctly mapped the soil sample's appropriate crop and fertilizer recommendations.

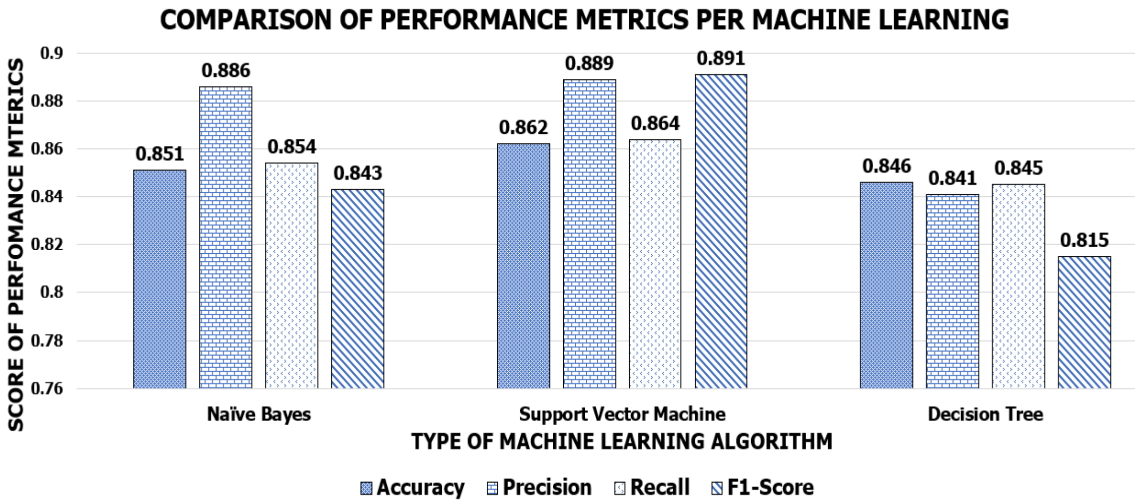

Figure 4. Comparison of the performance metrics of the three-machine learning algorithm

5. CONCLUSION

Identification of soil pH and macronutrients, i.e., nitrogen, phosphorus, potassium, and micronutrients, i.e., zinc, iron, and copper, is essential in maintaining soil productivity. Colorimetry and spectrometry are chemical processes that help determine the level of soil parameters. Creating "in-situ" devices is vital for farmers and other interested parties in soil analysis. Crop and fertilizer recommendations are essential in maximizing crop yields and soil status. Machine learning is optimally used in several applications, especially in image processing, to determine the status of soil parameters. It is subjected to several training images for proper identification of the level. Various machine-learning techniques can be implored, but the SVM is considered best.

ACKNOWLEDGEMENTS

The authors express their gratitude to the Bicol State College of Applied Sciences and Technology for publishing this paper.

REFERENCES

[1] X. Yang *et al.*, “A survey on smart agriculture: development modes, technologies, and security and privacy challenges,” *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 2, pp. 273–302, Feb. 2021, doi: 10.1109/JAS.2020.1003536.
[2] H. Azadi *et al.*, “Rethinking resilient agriculture: from climate-smart agriculture to vulnerable-smart agriculture,” *Journal of Cleaner Production*, vol. 319, Oct. 2021, doi: 10.1016/j.jclepro.2021.128602.
[3] O. Friha, M. A. Ferrag, L. Shu, L. Maglaras, and X. Wang, “Internet of things for the future of smart agriculture: a comprehensive survey of emerging technologies,” *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 4, pp. 718–752, Apr. 2021, doi: 10.1109/JAS.2021.1003925.
[4] P. K. R. Maddikunta *et al.*, “Unmanned aerial vehicles in smart agriculture: applications, requirements, and challenges,” *IEEE Sensors Journal*, vol. 21, no. 16, pp. 17608–17619, Aug. 2021, doi: 10.1109/JSEN.2021.3049471.
[5] A. S. Kumar, G. Suresh, S. Lekashri, L. G. Babu, and R. Manikandan, “Smart agriculture system with E – carbage using IoT,” *International Journal of Modern Agriculture*, vol. 10, no. 1, pp. 928–931, 2021.
[6] R. K. Goel, C. S. Yadav, S. Vishnoi, and R. Rastogi, “Smart agriculture – urgent need of the day in developing countries,” *Sustainable Computing: Informatics and Systems*, vol. 30, Jun. 2021, doi: 10.1016/j.suscom.2021.100512.
[7] A. R. D. A. Zanella, E. D. Silva, and L. C. P. Albini, “Security challenges to smart agriculture: current state, key issues, and future directions,” *Array*, vol. 8, Dec. 2020, doi: 10.1016/j.array.2020.100048.

[8] N. Singh and A. N. Singh, "Odysseys of agriculture sensors: current challenges and forthcoming prospects," *Computers and Electronics in Agriculture*, vol. 171, Apr. 2020, doi: 10.1016/j.compag.2020.105328.
[9] B. B. Sinha and R. Dhanalakshmi, "Recent advancements and challenges of internet of things in smart agriculture: a survey," *Future Generation Computer Systems*, vol. 126, pp. 169–184, Jan. 2022, doi: 10.1016/j.future.2021.08.006.
[10] A. Lytos, T. Lagkas, P. Sarigiannidis, M. Zervakis, and G. Livanos, "Towards smart farming: systems, frameworks and exploitation of multiple sources," *Computer Networks*, vol. 172, May 2020, doi: 10.1016/j.comnet.2020.107147.
[11] S. S. Reddy, N. Sethi, and R. Rajender, "Evaluation of deep belief network to predict hospital readmission of diabetic patients," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, IEEE, Jul. 2020, pp. 5–9. doi: 10.1109/ICIRCA48905.2020.9182800.
[12] S. Balne, "Smart agriculture using advanced machine learning algorithms," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 9, no. 7, pp. 6836–6840, 2020.
[13] M. W. P. Maduranga and R. Abeysekera, "Machine learning applications in iot based agriculture and smart farming: a review," *International Journal of Engineering Applied Sciences and Technology*, vol. 4, no. 12, pp. 24–27, May 2020, doi: 10.33564/IJEAST.2020.v04i12.004.
[14] C. Salim and N. Mitton, "Machine learning based data reduction in wsn for smart agriculture," in *Advances in Intelligent Systems and Computing*, Springer, Cham, 2020, pp. 127–138. doi: 10.1007/978-3-030-44041-1_12.
[15] G. Abraham, R. Raksha, and M. Nithya, "Smart agriculture based on iot and machine learning," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, IEEE, Apr. 2021, pp. 414–419. doi: 10.1109/ICCMC51019.2021.9418392.
[16] M. Fredj, R. Grati, and K. Boukadi, "CropWaterNeed: a machine learning approach for smart agriculture," in *Advances in Intelligent Systems and Computing*, Springer, Cham, 2021, pp. 33–38. doi: 10.1007/978-3-030-70416-2_5.
[17] K. N. Bhanu, H. J. Jasmine, and H. S. Mahadevaswamy, "Machine learning implementation in iot based intelligent system for agriculture," in *2020 International Conference for Emerging Technology (INCET)*, IEEE, Jun. 2020, pp. 1–5. doi: 10.1109/INCET49848.2020.9153978.
[18] J. Padarian, B. Minasny, and A. B. McBratney, "Machine learning and soil sciences: a review aided by machine learning tools," *SOIL*, vol. 6, no. 1, pp. 35–52, 2020, doi: 10.5194/soil-6-35-2020.
[19] H. Adab, R. Morbidelli, C. Saltalippi, M. Moradian, and G. A. F. Ghalhari, "Machine learning to estimate surface soil moisture from remote sensing data," *Water*, vol. 12, no. 11, Nov. 2020, doi: 10.3390/w12113223.
[20] R. Dash, D. K. Dash, and G. C. Biswal, "Classification of crop based on macronutrients and weather data using machine learning techniques," *Results in Engineering*, vol. 9, Mar. 2021, doi: 10.1016/j.rineng.2021.100203.
[21] U. E. Farwa, A. U. Rehman, S. Q. Khan, and M. Khurram, "Prediction of soil macronutrients using machine learning algorithm," *International Journal of Computer*, vol. 38, no. 1, pp. 1–14, 2020.
[22] P. Angin, M. H. Anisi, F. Göksel, C. Gürsoy, and A. Büyükgülcü, "AgriLoRa: a digital twin framework for smart agriculture," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 11, no. 4, pp. 77–96, 2020, doi: 10.22667/JOWUA.2020.12.31.077.
[23] D. Vimalajeewa, C. Kulatunga, D. Berry, and S. Balasubramaniam, "A service-based joint model used for distributed learning: application for smart agriculture," *IEEE Transactions on Emerging Topics in Computing*, vol. 10, no. 2, 2021, doi: 10.1109/TETC.2020.3048671.
[24] P. Samuel S., K. Malarvizhi, S. Karthik, and M. Gowri S.G., "Machine learning and internet of things based smart agriculture," in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, IEEE, Mar. 2020, pp. 1101–1106. doi: 10.1109/ICACCS48705.2020.9074472.
[25] T. Hengl *et al.*, "African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning," *Scientific Reports*, vol. 11, no. 1, Mar. 2021, doi: 10.1038/s41598-021-85639-y.
[26] C. Bayudan-Dacuycuy and R. B. Serafica, "Harnessing the potential of the philippines' agricultural sector: an assessment using the product space," *PIDS - Philippine Institute for Development Studies*, Quezon City, pp. 1-20, 2019.
[27] G. J. Morales, R. K. Villaronte, M. C. Yap, and M. A. Rosete, "The relationship between rural-urban migration and the agricultural output of the philippines," *International Journal of Social and Management Studies*, vol. 3, no. 1, pp. 62–74, 2022, doi: 10.5555/ijosmas.v3i1.88.
[28] I. Parlina *et al.*, "Naive bayes algorithm analysis to determine the percentage level of visitors the most dominant zoo visit by age category," *Journal of Physics: Conference Series*, vol. 1255, no. 1, Aug. 2019, doi: 10.1088/1742-6596/1255/1/012031.
[29] G. Kaur and A. Chhabra, "Improved j48 classification algorithm for the prediction of diabetes," *International Journal of Computer Applications*, vol. 98, no. 22, pp. 13–17, Jul. 2014, doi: 10.5120/17314-7433.
[30] K. K. Napa, A. K. Kumar, S. Murugan, K. Mahammad, and T. A. Assegie, "Early prediction of chronic heart disease with recursive feature elimination and supervised learning techniques," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 1, pp. 730-736, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp730-736.
[31] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, Sep. 2020, doi: 10.1016/j.neucom.2019.10.118.
[32] Y. Dani and M. A. Ginting, "Comparison of iris dataset classification with gaussian naïve bayes and decision tree algorithms," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 2, pp. 1959-1968, Apr. 2024, doi: 10.11591/ijece.v14i2.pp1959-1968.
[33] P. Mallioris, E. Diamantis, C. Bialas, and D. Bechtsis, "Predictive maintenance framework for assessing health state of centrifugal pumps," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 1, pp. 850-862, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp850-862.
[34] J. Cervantes, F. G. Lamont, A. López-Chau, L. R. Mazahua, and J. S. Ruíz, "Data selection based on decision tree for SVM classification on large data sets," *Applied Soft Computing*, vol. 37, pp. 787–798, Dec. 2015, doi: 10.1016/j.asoc.2015.08.048.
[35] A. Z. A. Zainuddin, W. Mansor, K. Y. Lee, and Z. Mahmoodin, "Machine learning and deep learning performance in classifying dyslexic children's electroencephalogram during writing," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 6, pp. 6614-6624, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6614-6624.
[36] S. Rajayyan and S. M. M. Mustafa, "Prediction of dementia using machine learning model and performance improvement with cuckoo algorithm," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 4, pp. 4623-4632, Aug. 2023, doi: 10.11591/ijece.v13i4.pp4623-4632.
[37] S. Fletcher and M. Z. Islam, "Decision tree classification with differential privacy," *ACM Computing Surveys*, vol. 52, no. 4, pp. 1–33, Jul. 2020, doi: 10.1145/3337064.

[38] F. A. Vadhil, M. L. Salihi, and M. F. Nanne, "Machine learning-based intrusion detection system for detecting web attacks," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 1, pp. 711-721, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp711-721.

[39] J. Liang, Z. Qin, S. Xiao, L. Ou, and X. Lin, "Efficient and secure decision tree classification for cloud-assisted online diagnosis services," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 4, pp. 1632–1644, Jul. 2021, doi: 10.1109/TDSC.2019.2922958.

[40] A. R. Zarzoor, N. A. S. Al-Jamali, and D. A. A. Qader, "Intrusion detection method for internet of things based on the spiking neural network and decision tree method," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 2, pp. 2278-2288, Apr. 2023, doi: 10.11591/ijece.v13i2.pp2278-2288.

BIOGRAPHIES OF AUTHORS

John Joshua Montañez obtained his Bachelor of Science in Electronics Engineering at the Ateneo de Naga University cum laude. He finished his Master of Engineering at the Bicol State College of Applied Sciences and Technology (BISCAST), where he is also a faculty member of the electronics engineering department of the College of Engineering (CEng). He is a student of the Straight Masters-Doctorate in Electronics Engineering at the Batangas State University, the National Engineering University. He is designated as the Intellectual Property Unit Manager and CEng Research Coordinator. His research interests are electronics engineering, signal processing, computer vision, artificial intelligence, engineering education, and intellectual property. He is also an associate member of the Engineering and Industrial Research Division of the National Research Council of the Philippines. He can be contacted at email: jjfmontanez@astean.biscast.edu.ph.

Jeffrey Sarmiento finished his Bachelor of Science and Master of Science in Computer Engineering at the Batangas State University. He finished his Doctor of Engineering major in Computer Engineering at the Technological Institute of the Philippines. He is the program chair for multiple programs, including MSCpE (Master of Science in Computer Engineering), MSAI (Master of Science in Artificial Intelligence), and MEng-CpE (Master of Engineering in Computer Engineering) at the Batangas State University. His research interests are computer engineering, applied machine learning, Python (programming language), and applied artificial intelligence. He is also an associate member of the Engineering and Industrial Research Division of the National Research Council of the Philippines. He can be contacted at email: jeffrey.sarmiento@g.batstate-u.edu.ph.