Journal of Applied Sciences, Information and Computing
Volume 5, Issue 1, June 2024
School of Mathematics and Computing, Kampala International University

ISSN: 1813-3509 **https://doi.org/10.59568/JASIC-2024-5-1-02**

A review of soil analysis and classification to improve crops in agriculture using machine learning algorithm

[1]V.S Manjula, [2]Fatou Marega

[1]Professor, Department of Computer Science, Kampala International University, Kampala, Uganda, East Africa, **manjusunil.vs@gmail.com,** ***Orcid ID: (0000-0003-0308-3289)***

[2]Research Student, Master of Science and Information System, Kampala International University, Kampala, Uganda, East Africa, maregaf34@gmail.com

Abstract

One of the most crucial aspects of our society is agriculture. For agriculture to be successful, soil is essential. Each type of soil has a unique composition. These chemical properties of the soil have an impact on crop growth. It is important to choose the right crop for that specific type of soil. Machine learning algorithms are able to classify the data from the soil series. to predict which crops are best suited to the soil type and climate of a certain area, the findings of this categorization can also be paired with crop datasets. Both the crop and the soil datasets are used. The files include geographical and chemical characteristics of soil and crop features. The classification of soil series data and the prediction of appropriate crops can be done using algorithms like SVM and the assembling approach. For this study, 15 villages in the near area of Kampala, Uganda, provided 1200 soil samples with chemical, physical, and biological characteristics. Overall, the Random Forest model produced results with 99% accuracy for cellulose enzyme activity and 94% accuracy for N-acetyl-glucosaminidase enzyme activity. Based on evaluation performance parameters and the best performance accuracy algorithms, recommendations for crops have been made.

Keywords: Soil Examination, Machine Learning, Random Forest Model, Crop, Agriculture, Soil Classification

1. Introduction:

Investment in Agricultural Research and Development: Recognizing the importance of agriculture, there is a growing emphasis on investing in research and development to improve agricultural practices and technologies. This includes promoting scientific advancements, conducting studies on soil analysis, crop classification, and exploring innovative solutions to address agricultural challenges specific to Uganda's context.

Capacity Building and Training: To enhance the performance of agriculture, there is a need to focus on capacity building and training programs for farmers, extension workers, and agricultural professionals. These programs aim to provide them with the necessary knowledge and skills to implement modern farming techniques, adopt sustainable practices, and effectively utilize technologies such as machine learning for soil analysis and crop classification.

Policy and Institutional Support: Creating an enabling environment through supportive policies and institutions is crucial for the growth of agriculture. This involves developing policies that prioritize agriculture, ensure sustainable land use, provide access to credit and agricultural inputs, and foster collaboration between government agencies, research institutions, and farmers' organizations. Additionally, strengthening agricultural extension services and establishing data collection systems for monitoring and evaluation are important steps in supporting the implementation of machine learning for soil analysis and classification.

By focusing on these elements, Uganda can harness the potential of agriculture to drive economic growth, improve farmers' livelihoods, and reduce poverty. Soil analysis and crop classification using Machine learning techniques play a vital role in providing data-driven insights and decision support to farmers, enabling them to be researchers, and farmers is crucial for successful implementation. It makes informed choices about crop selection, soil management practices, and resource allocation. As agriculture takes center stage in the development agenda, it is essential to integrate technological advancements like machine learning to maximize the sector's potential and ensure sustainable agricultural growth in Uganda.

Machine learning for soil analysis and classification involves collecting soil samples from different regions, conducting laboratory tests to measure nutrient levels, pH, organic matter content, and texture. These soil samples are then combined with climate data to create a comprehensive dataset. Feature engineering techniques are applied to extract relevant features and reduce dimensionality. Machine learning models, such as support vector machines, decision trees, and neural networks, are trained on the data to predict soil properties and classify soil types.

Using artificial intelligence for soil analysis and classification offers several benefits. It provides more accurate and faster results compared to traditional methods. Farmers can make crop selection informed decisions, fertilization strategies, irrigation techniques based on the soil characteristics of their land. This leads to improved crop productivity, reduced input costs, and increased overall agricultural sustainability.

In Uganda, where agriculture is a vital part of the economy, implementing machine learning for soil analysis can have significant positive impacts. By better understanding the soil properties and selecting the appropriate crops for specific soil types, farmers can optimize their yields and improve their livelihoods. Moreover, machine learning can aid in the identification of soil degradation or nutrient deficiencies, allowing farmers to take proactive measures to address these issues.

However, there are challenges in implementing machine learning for soil analysis in Uganda. Access to reliable and diverse datasets, technological infrastructure, and awareness about the benefits of machine learning among farmers are some of the factors that need to be considered. Collaboration between agricultural institutions, researchers, and farmers is crucial for successful implementation.

2. Literature Survey:

2.1: Machine Learning-Based Soil Classification and Crop Suggestions Based on Soil Series:

The research included in the literature review intends to create a model that can forecast soil series based on the kind of land and offer suitable crop recommendations. The soil series is classified using machine learning methods, notably weighted K-Nearest Neighbor (KNN), bagged trees, and support vector machines (SVM) using a Gaussian kernel. The soil classification approach utilized in the project is based on existing knowledge and practical circumstances. By classifying soil samples found on different land surfaces, a connection is established between the soil types and various natural entities. This soil classification process helps in understanding the characteristics and properties of different soil series present in a region.

 https://doi.org/10.59568/JASIC-2024-5-1-02

Figure 1. Urban Farmers in Uganda

Once the soil series are accurately classified and mapped, the model can provide suggestions for suitable crops to be cultivated in a specific region. By leveraging the knowledge gained from the soil classification, the model can recommend crops that are well-suited to the identified soil types. This information is valuable for farmers and agricultural experts in making informed decisions about crop selection and cultivation practices.

Overall, the project focuses on using machine learning algorithms to classify soil series based on land type and provide crop suggestions based on the identified soil classifications. The aim is to establish a connection between soil samples and different natural entities through soil classification and utilize this information to make appropriate crop recommendations for specific regions.

2.2 System Overview:

The project discussed in the literature survey aims to develop a soil prediction model that makes crop recommendations based on the type of land. The soil series is classified using machine learning methods, notably weighted K-Nearest Neighbor (KNN), bagged trees, and support vector machines (SVM) using a Gaussian kernel.

The soil classification approach utilized in the project is based on existing knowledge and practical circumstances. By classifying soil samples found on different land surfaces, a connection is established between the soil types and various natural entities. This soil classification process helps in understanding the characteristics and properties of different soil series present in a region.

Once the soil series are accurately classified and mapped, the model can provide suggestions for suitable crops to be cultivated in a specific region. By leveraging the knowledge gained from the soil classification, the model can recommend crops that are well-suited to the identified soil types. This information is valuable for farmers and agricultural experts in making informed decisions about crop selection and cultivation practices.

Overall, the project focuses on using machine learning algorithms to classify soil series based on land type and provide crop suggestions based on the identified soil classifications. To create a connection between soil samples is the goal. and various natural entities through soil classification and utilize this information to make appropriate crop recommendations for specific regions.

Weighted K-Nearest Neighbor (KNN) is a machine learning algorithm used for classification tasks. In the context of soil classification, KNN is applied to predict the soil series based on the features of the soil samples. In this particular study, the researchers employed a weighted version of the KNN algorithm.

The basic principle of KNN is to find the k nearest neighbors to a given data point based on a distance metric (e.g., Euclidean distance). In the weighted version, the neighbors that are closer to the data point are given more weightage or importance compared to the farther neighbors. This means that the closer neighbors have a stronger influence on the final classification decision.

Once the k nearest neighbors are identified, the majority class among them is considered as the predicted class for the given data point. This approach is based on the assumption that data points with similar features tend to belong to the same class.

In the study, the accuracy of the classification using weighted KNN was reported to be 92.93%. This means that the model achieved a high level of accuracy in predicting the soil series based on the given soil features. The accuracy indicates the percentage of correct predictions made by the model.

Overall, the weighted KNN algorithm was used in the study to classify soil samples into their respective soil series based on the features of the soil. The approach of giving more weight to closer neighbors improved the classification's accuracy, resulting in a 92.93% accuracy rate.

Gaussian kernel-based Support Vector Machines (SVM) is another machine learning algorithm used for classification tasks. SVM aims to separate objects of different classes by constructing decision boundaries or hyperplanes in a higher-dimensional space. The objects of one class are distinguished from those of another class by the decision boundary. The data points that are closest to the decision boundary or hyperplane are known as support vectors.

These support vectors play a crucial role in defining the decision boundary and determining the class labels of new, unseen data points. To handle non-linearly separable data, SVM uses a kernel function to map the input data into a higher-dimensional space, where it becomes easier to find a linear decision boundary. In this project, the researchers employed a Gaussian kernel function, also

Bagged Trees is another machine learning algorithm used in the project. Bagging, short for bootstrap aggregating, is a technique that generates an ensemble of models by training them on random samples of the data. All the models in the ensemble forecast the class labels and classify the data. The final forecast is calculated by summing the predictions of various models when a new case is presented. In the project, Bagged Trees algorithm was used to classify the soil series based on the given soil samples and their features. The accuracy obtained using this algorithm was reported to be 90.91%. This accuracy represents the percentage of correct predictions made by the Bagged Trees model. After classifying the soil series, the study also involved suggesting suitable crops for

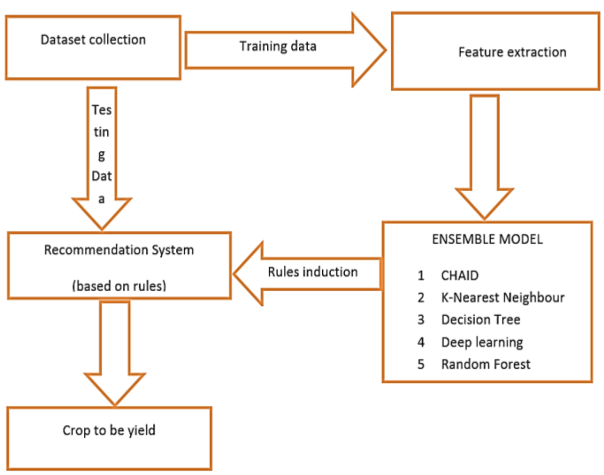

Figure 2. The Soil Classification and Crop Suggestion System Architecture

2.3. Big Data Analytics with Crop Prediction Mode Utilizing Optimization Technique

Big data analytics to forecast crops in the agriculture sector is the main emphasis of the project that is being detailed. Using numerous soil and environmental characteristics, including averages for temperature, humidity, total precipitation, and output yield, it is hoped to make predictions about crop yields, including whether they will be excellent or bad. One uses a hybrid classifier model to improve the prediction model. Three stages make up the project: pre-processing, feature selection, and SVM_GWO. To assure the data's reliability and integrity, it is prepped and cleansed during the pre-processing stage. For accurate analysis and prediction, it is imperative to complete this stage. The most important features are found and chosen from the dataset during the feature selection process. By doing this, the prediction model's accuracy is increased and noise is reduced.

Grew Wolf Optimization Based Feature Selection: A method for choosing the best subset of features for classification purposes is Grey Wolf Optimization (GWO) based feature selection. This method uses optimization to reduce training error and assign features relative weights based on their relative information.

GWO successfully creates a relationship between features and draws out important data from them. In doing so, it aids Support Vector Machine (SVM) classification by lowering the number of support vectors. This decrease of support vectors ultimately results in a drop in the classifier's training error. The testing error of the classification process directly reflects the effects of GWO-based feature selection. It improves the accuracy, precision, and memory of by choosing the most pertinent and instructive elements.

Machine Learning Classifier: For easier and more accurate data visualization, a feature selection algorithm selects a subset of essential features while eliminating extraneous, related, and noisy features.

Proposed Algorithms: SVM_GWO's proposed algorithm. The goal of this project is to use several parameters to improve the prediction model's accuracy for precision agriculture in the future. The data set was obtained from the Food Agriculture Organization and used in the map reduction processing model. Map reduction is used to process huge data in order to shorten execution time and do further classification.

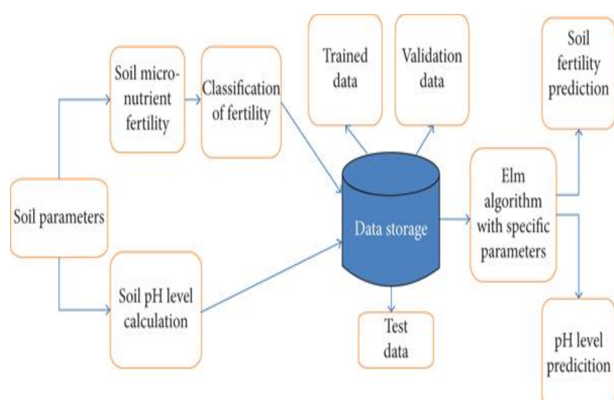

Figure 3. The system architecture of Soil Analysis and Micronutrient Classification using Machine Learning

In classification systems, feature selection and extraction are crucial phases. This study offers the SVM_GWO hybrid model, which employs a combinational technique to determine the best SVM parameter values in order to increase classification accuracy, recall, precision, and f-measure. In this classification, we have built SVM_GWO for choosing the features with the lowest error and convergent.

2.4: Enhancing Crop Productivity with an Assembling Technique Crop Recommendation System:

Using the assembling technique, a model that appropriately suggests the best crop depending on the unique kind and characteristics of the soil by combining the predictions of numerous machine learning models. Random Forest, Naive Bayes, and Linear SVM serve as the ensemble model's independent base learners. Every classifier offers a unique set of class labels with a respectable level of accuracy. The majority voting method is used to integrate the class labels of distinct base learners. The input soil dataset is divided into two categories by the crop recommendation system: Kharif and Rabi. The dataset includes samples of the average rainfall and surface temperature as well as the physical and chemical properties particular to the soil. The typical categorization.

Proposed Algorithm: Naive Bayes, Linear SVM, Random Forest, Ensemble Framework, and Majority Voting.

The soil dataset taken into consideration for use in the specific proposed study is made up mostly of soil physical and chemical parameters as well as climatic information. An open-source dataset is obtained from the Government of India's data repository website, data.gov.in.

The dataset is 5MB in size, contains 9000 rows, and has 15 crucial properties. Wheat, rice, cotton, and sugarcane are the crops taken into account. The following dataset attributes are of the utmost significance: soil type, pH value, NPK concentration, porosity, average rainfall, surface temperature, and sowing season.

With respect to the soil dataset, a crop recommendation system has been developed that

2.5: Machine Learning-Based Intelligent Farming Forecasting:

One of the key game changers and a significant source of income in India is agriculture. Crop output is influenced by the seasons, the market, and biological cycles, yet disruptions to these patterns cause farmers to lose a lot of money. By employing an appropriate strategy based on an understanding of soil types, pressure, ideal weather, and crop type, these factors can be reduced. In contrast, crop kinds and weather may be predicted utilizing a valuable information that provides farmers with a forecast of the most profitable crops to grow. These publications mostly focus on crop yield and crop cost prediction algorithms. All of these traits can be used to implement smart farming. Datasets from Koggle.com are used in the implementation to feed

Research Work: Research is the first step in collecting machine-level data. For these, they just used the train datasets and pre-processed them. Test and training portions of the data are separated. The required data's attributes, including soil type, temperature, humidity, and others, are all extracted.

Feature Extraction: Given the quantity of feature extraction is required for the data to be processed. There will only be an extraction of the pertinent data (about 25 attributes) from the test and train segments. These characteristics will enable farms to advance across all levels.

Classification Technique: The most crucial step in the process is the classification approach because this is where the algorithm is put into practice. In order to produce thorough dataset results, the Random Forest technique is employed in the procedure. To calculate the system size, the approach employs 80% of the train data and 20% of the test data (random data). We receive two outputs as a result of applying the classification techniques: Dataset outcomes that will take the shape of a matrix, such as true positive, true negative, etc., as a result of the algorithm (accuracy of the datasets). The matrix itself can be used to determine the predicted data. Real-time predictions of the best land to plant a certain crop are made using the matrix's values.

2.6. Using classification methods and soil attribute prediction, analyze the soil data:

Automation and data mining are examples of technological advancements that have benefited agricultural research. Despite the widespread usage of data mining today and the availability of several commercially available off-the-shelf data mining systems and domain-specific data mining application software, research into data mining in agricultural soil datasets is still in its infancy. The massive volumes of data that are now practically gathered alongside the crops must be analyzed and utilized to the fullest degree possible. The examination of a soil dataset using data mining techniques is the goal of this study. It focuses on classifying soil using several algorithms that are available. Regression analysis is used to predict untested properties and automated soil sample classification is put into practice.

3. Research Methodology:

 https://doi.org/10.59568/JASIC-2024-5-1-02

The surveys that were frequently conducted in Pune District provided the data for this study. The information was specifically gathered from a private soil testing facility in Pune. The number of soil samples is disclosed in the dataset. The significance of a soil classification system for precisely defining soil attributes was acknowledged by the researchers. Using tables or flowcharts to classify soil the old-fashioned way might take a lot of time. Consequently, a quick and trustworthy automated solution is required to increase the effectiveness of soil categorization and make better use of technicians' time. The researchers created an automated technique for soil classification based on fertility in response to this demand. The automated system can be swiftly and precisely by utilizing the strength of expert systems.

3.1 A Comparison of Classification Schemes for Soils

The significance of soil type in selecting appropriate crops and fertilizer applications. In order to make informed decisions regarding agriculture, it is essential to classify the soil since it contains important information about the fertility and characteristics of the soil. Experts in the field can choose the best crops by dividing the soil into classes according to its fertility. In terms of soil fertility, pH levels, nutrient content, and drainage features, different crops have particular needs. Experts can suggest suitable crops that are likely to grow and generate high yields in that specific soil type by knowing the fertility class of the soil.

The right fertilizers to employ can also be determined by the soil classification. Because different soils have different levels of nutrients, understanding the fertility class aids specialists in determining the precise

3.2 Proposed Algorithm:

Naive Bayes is a straightforward but powerful probabilistic classifier built on the Bayes theorem. Given the class title, it is assumed that every feature is independent of every other feature. The algorithm determines the likelihood that a sample will fall into each class, then chooses the class with the highest likelihood to be the predicted class. Naive Bayes is renowned for being straightforward,

A decision tree algorithm called J48 (C4.5) is based on the C4.5 method. Recursively dividing the dataset based on the most informative attribute creates a decision tree. J48 employs information gain or entropy to find the most advantageous attribute to

Naive Bayes: A simple probabilistic classifier known as a naive Bayes classifier utilizes the Bayes theorem on the basis of strong (naive) independence. The advantage of using the naive Bayes classifier is that less training data is needed to estimate the parameters.

J48: Using the assumption of strong (naive) independence, a simpleton A straightforward probabilistic classifier that makes use of Bayes' theorem is the Bayes classifier. . Since less training data is needed to estimate the parameters, the naive Bayes classifier has this advantage.It is a development of Quinlan's original ID3 algorithm. Rule derivation, decision tree pruning, continuous attribute value ranges, and so forth.

JRip: This technique implements the propositional rule learner known as Repeated Incremental Pruning to Produce Error Reduction (RIPPER), which was proposed by William W. Cohen as an enhanced version of IREP. The speed, accuracy, mistake rate, true positive rate, and false positive rate of three data mining classification methods were evaluated and compared in this study. Tenfold cross validation was applied in the experiment. In the last few years, soil samples used to investigate human visual attention have proven the J48 model to be the most accurate classifier, according to our research.

4. Results and Discussion:

Categorized the micronutrients are the soil samples in four districts of Uganda based on the analysis that was done.

Algorithm: PSEA–MLR (Predict the Soil Enzyme Activity using Multiple Linear Regression)

Goal: Finding the best possible balance between soil enzyme activity and response factors
Enter: N = (PH, Silt, Sand, Clay, N, P, SOM, SOC, Depth, and Fertility Level)
K = (Predict soil enzyme activity) is the output.
1: Setting up the parameters for the N and K soil data
2: Using N and K parameters for pre-processing the soil dataset
3: Choose 80% of the soil dataset at random for training and 20% of the soil dataset at random for testing.
4: Use MLR-supervised machine learning methods on a specified dataset.
5: Determine the model's Accuracy, MSE, RMSE, and MAE; 6: Forecast the activity of soil enzymes
7: End

4.1 Soil Fertility:

The four sites in Uganda yielded a totally calculated 190 soil samples. There are four various districts: Iganga (71), Tororo (72), Kamuli (24), and Soroti District (20) (Annex 1). Only the soil chemical analysis is mostly taken into account when evaluating The PH of the soil is compared to crop-recommended agronomic standards, and the soil

 https://doi.org/10.59568/JASIC-2024-5-1-02

fertility from the sites and in various land uses suitable levels for certain components (Tables 1 and 2). Although texture is a very important factor in how soil behaves in terms of hydrology or erodibility, it is also taken into account in terms of soil water holding capacity in another area. The Mehlich et al. (1964) and Hinga et al. (1980) modified classes of nutrient availability were employed to evaluate nutrient availability.

Table 1: Status of Soil fertility P, K, Ca, Mn, Mg, and Fe

Nutrient	Deficiency level	Adequate level	Excessive or reactionary level	Remarks
Sodium me%	Seldom applies	0.0-2.1	>2.1	Salinity possible
Potassium %	<0.3	0.3-1.6	>1.6	In Calculation soil
Calcium %	<2.1	2.0-15.0	>16	
Magnesium %	<1.1	1.0-3.1	>2.0	
Manganese %	<0.11	0.11-2.0	>80	
Fe ppm	<9	>9		

Output: K = (Predict soil enzyme activity)

1: Initialization of all N and K soil data parameters
2: Pre-processing of the soil dataset with N and K parameters
3: Randomly select 80% soil dataset for training and 20% soil dataset for testing purposes
4: Apply MLR-supervised ML algorithms on a given data set
5: Compute the Accuracy, MSE, RMSE, and MAE of the model
6: Predict soil enzyme activity
7: End

4.1 Soil Fertility:

The four sites in Uganda yielded a totally calculated 190 soil samples. There are four various districts: Iganga (71), Tororo (72), Kamuli (24), and Soroti District (20) (Annex 1). Only the soil chemical analysis is mostly taken into account when evaluating The PH of the soil is compared to crop-recommended agronomic standards, and the soil fertility from the sites and in various land uses suitable levels for certain components (Tables 1 and 2). Although texture is a very important factor in how soil behaves in terms of hydrology or erodibility, it is also taken into account in terms of soil water holding capacity in another area. The Mehlich et al. (1964) and Hinga et al. (1980) modified classes of nutrient availability were employed to evaluate nutrient availability.

Table 2: Status of Soil fertility P, K, Ca, Mn, Mg, and Fe

Nutrient	Deficiency level	Adequate level	Excessive or reactionary level	Remarks
Sodium me%	Seldom applies	0.0-2.1	>2.1	Salinity possible
Potassium %	<0.3	0.3-1.6	>1.6	In Calculation soil
Calcium %	<2.1	2.0-15.0	>16	
Magnesium %	<1.1	1.0-3.1	>2.0	
Manganese %	<0.11	0.11-2.0	>80	
Fe ppm	<9	>9		
Zn ppm	<5	>5		
Cu ppm	<1.0	>1.0		

Changes in soil PH, NPK, and organic carbon due to different land usage at four different sites in Uganda. In Kamuli, Soroti, Tororo, and Iganga, the change in soil PH across different land uses is depicted in (Figure 3-5), correspondingly. The findings demonstrate that sweet potatoes, maize, and coffee were being cultivated at the Kamuli site on nearly appropriate PH soils. However, to elevate the PH above the current PH, lime must be applied to at least lt/ha of maize and coffee. This crop has been deemed uneconomical (AIC tech. handbook), and cassava will flourish under the current PH. However, Lime must be applied at least lt/ha if the grass is to be produced commercially. Bushland, grazing, grass, and fallow will all naturally modify their PH. In Soroti

4.2 Random Forest for soil enzyme activity:

This multivariate method was created to improve the accuracy and efficiency of regression and classification trees. To increase the filling of each classification and regression tree, it combines bagging and random variable selection with numerous classification and regression tree techniques. At the same time, random feature extraction and bagging techniques lead to a decrease in correlation for each factor in the Random Forest. Compute information gain (IG) using the entropy method of all splitting feature data provided and Algorithm. Based on the analysis conducted, the nutrients are classified and the ideal soil conditions are suggested for various resources.

 https://doi.org/10.59568/JASIC-2024-5-1-02

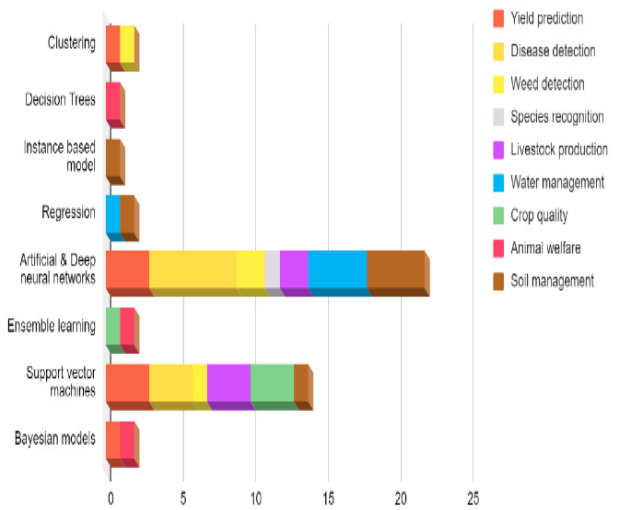

Figure 4. Cross-validation accuracy (%) for determining the pH level in the soil is plotted against the number of hidden neurons.

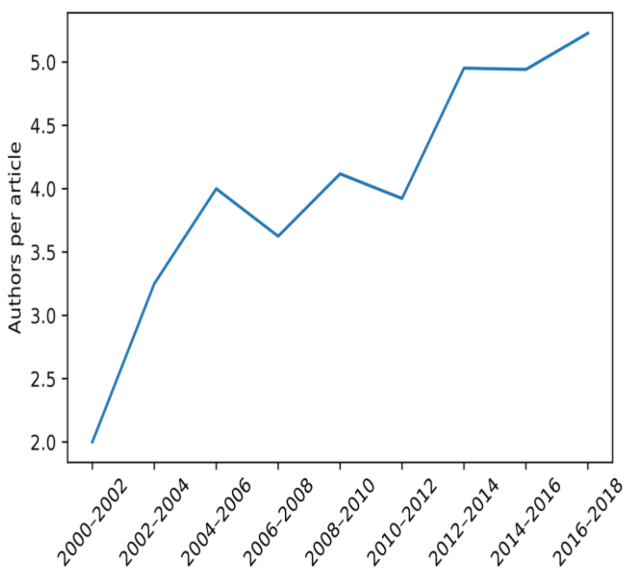

Figure 5. Soil- A review aided by machine learning tools

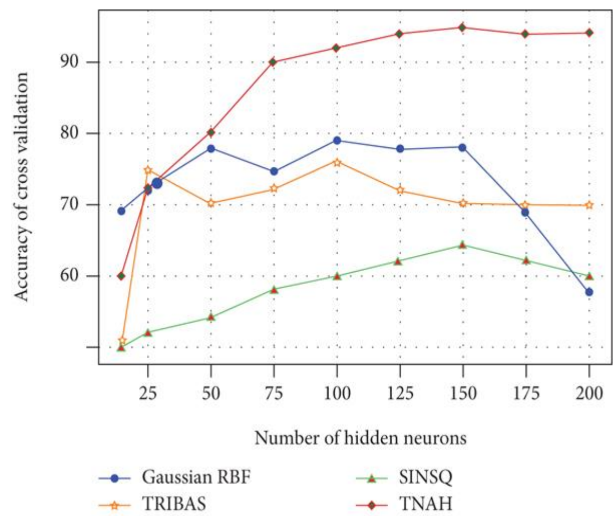

Figure 6. ML algorithm for soil categorization and evaluation

5. CONCLUSION

The Machine Learning algorithm-based application in soil analysis and classification has the potential to revolutionize agriculture in Uganda. By leveraging advanced algorithms and data-driven insights, farmers can make informed decisions that lead to improved crop productivity and sustainable agricultural practices. Further research, investment in infrastructure, and capacity-building initiatives are essential to harness the full potential of machine learning in Ugandan agriculture. They have suggest various algorithms and ed a study of the soil data utilizing prediction techniques in this paper. Even though it has been demonstrated that least median squares regression produces results that are superior to the traditional

Data Availability

The article has all the **Data Availability** supporting information for the study's findings.

Disclosure

This research's publishing declaration was made solely for academic purposes by Uganda's Kampala International University of Science and Technology.

Conflicts of Interest

The authors claim that the publication of this research does not involve any conflicts of interest.

6. References:

[1] Gholap, J., Ingole, A., Gohil, J., Gargade, S. and Attar, V., 2012. "Soil data analysis using classification techniques and soil attribute prediction", ArXiv preprint ar Xiv: 1206.1557.

[2] Sofianita Mutalib, S-N-Fadhlun Jamian, Shuzlina Abdul Rahman, Azlinah Mohamed, 2010, "Soil Classification: An Application of Self Organising Map and k-means".

[3] Nabulongo Ali, V. S. Manjula and Fatou Marega, "Impact of Digitization of Sustainable Agriculture in Uganda: A Case Study", Journal of Applied Science, Information and Computing(JASIC),Volume , Issue 1, Published 10th July 2023, ISSN: 1813-3509| Journal article DOI: 10.59568/JASIC-2023-4-1-05, https://doi.org/10.59568/JASIC-2023-4-1-01

[4] Prachi Sharma Dr. D.V. Padole, 2017, "Design and Implementation Soil Analyzer using IoT".

[5] S. Pudumalar, E. Ramanujam, R.Harine Rajashreeń, C.Kavyań, T.Kiruthikań, J.Nishań, 2016,"Crop Recommendation System for Precision Agriculture".

[6] Avinash Kumar, Sobhangi Sarkar and Chittaranjan Pradhan, 2019," Recommendation System for Crop Identification and Pest Control Technique in Agriculture".

[7] Dr.V.S. Manjula, Image Edge Detection and Segmentation by using Histogram Thresholding method Journal of Engineering Research and Application, ISSN : 2248-9622, Vol. 7, Issue 7, (Part -6) July 2017, pp.76-83[**Google Scholar**]

[8] Nidhi H Kulkarni , Dr. G N Srinivasan , Dr. B MSagar, Dr.N K Cauvery,2018, "Improving Crop Productivity through a Crop Recommendation System using Ensembling Technique".

[9] Meiqin Zhang, Shanqin Wang, Shuo Li, Jing Yi, Peng Fu, 2011, "Prediction and Map-making of Soil Organic Matter of Soil Profile Based on Imaging Spectroscopy": A Case in Hubei China.

[10] Dr. V.S. Manjula, "New Trends and Developments in Educational Technology Research", International Journal of Research & Development Organization (IJRDO), Volume 2, Issue 3, March 2016. Impact Factor: 2.303

[11] Parin M. Shah, 2012, "Face Detection from Images Using Support Vector Machine".

[12] James O. Hortle, 2018, "Alzheimer's Disease and Support Vector Machines, an Introduction to Machine Learning".

[13] Amos Baranes, Rimona Palas, 2019, "Earning Movement Prediction Using Machine Learning-Support Vector Machines (SVM)".

[14] S. K. A. L. Z. Rahman, K. C. Mitra, and S. M. M. Islam, "Soil classification using machine learning methods and crop suggestion based on soil series," in *2018 21st International Conference of Computer and Information Technology (ICCIT)*, pp. 1–4, Dhaka., Bangladesh, 2018.
View at: Google Scholar

[15] E. U. Eyo, S. J. Abbey, T. T. Lawrence, and F. K. Tetteh, "Improved prediction of clay soil expansion using machine learning algorithms and meta-heuristic dichotomous ensemble classifiers," *Geoscience Frontiers*, vol. 13, no. 1, article 101296, 2022.

1. [16] Maheswari, M.U.; Ramani, R. A Comparative Study of Agricultural Crop Yield Prediction Using Machine Learning Techniques. In Proceedings of the 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 17–18 March 2023; IEEE: Piscataway, NJ, USA, 2023; Volume 1, pp. 1428–1433. [**Google Scholar**]