Biological Forum – An International Journal 15(3): 417-421(2023)

ISSN No. (Print): 0975-1130
ISSN No. (Online): 2249-3239

Machine Learning-based Crop Recommendation System in Biswanath District of Assam

Samikhya Bhuyan[1], Dilip Kumar Patgiri[1], Simanta Jyoti Medhi[2], Roopali Patel[3] and Tabuiliu Abonmai[4]*
[1]*Department of Soil Science, Assam Agricultural University, Jorhat (Assam), India.*
[2]*Senior Software Engineer, Volvo Group, India.*
[3]*Department of Horticulture, Assam Agricultural University, Jorhat (Assam), India.*
[4]*Department of Agronomy, College of Agriculture, CAU-Imphal (Manipur), India.*

(Corresponding author: Samikhya Bhuyan)*
(Received: 20 January 2023; Revised: 15 February 2023; Accepted: 21 February 2023; Published: 22 March 2023)
(Published by Research Trend)

ABSTRACT: The present study makes attempt to formulate a crop recommendation system based on soil physical property. The study was undertaken in the Biswanath district of Assam which is situated 26°35' to 27°00' of latitude and 92°50' to 93°50' of longitude, with a total area of 14,15,000 ha. The 180 soil samples were collected from different location and physical properties were analyzed in laboratory using standard protocol. The selected physical properties includes soil texture, bulk density, particle density, total porosity, hydraulic conductivity, maximum water holding capacity, volume expansion, field capacity, permanent wilting point, available water, microaggregate, macroaggregate and mean weight diameter. The coefficient of variation (CV) was used for interpretation of variability of the soil properties. Silt, clay, hydraulic conductivity, were identified to be the most variable soil indicators (CV > 35%). Available water was found as the moderately variable parameters (CV 15–35%). The least variation (CV < 15) were found in bulk density, particle density, total porosity, field capacity, mean weight diameter, microaggregate, macroaggregte and volume expansion. The results of the physical properties were fit in the machine learning model and a farmer's friendly crop recommendation system was developed.

Keywords: Crop recommendation system, Physical property, coefficient of variation, Mean weight.

INTRODUCTION

One of the most significant industries in the Indian economy is the agricultural industry. Almost 151 million people were depends on the industry. The sector employs over 60% of the Indian population and accounts for about 18% of India's GDP. Therefore, the progress of the country's economic growth relies greatly on the agriculture industry. The population of India accounts for 17.7% of the entire global population. The populations were increasing day by day but the sizes of holding of the farmers were decreasing. The soil which is the basic need of agriculture has been under cultivation for thousands of years, that has resulted in depletion of soil health, which lowers agricultural productivity. As farmers are the main users of the word "soil health", it is often referred to as a farmer's term

The examination of soil health is crucial for Indian soil, but however technologies have overlooked it. High-level computing systems may automate the procedure, which also increases its objectivity. The use of artificial intelligence (AI) technologies like machine learning and deep learning to agricultural operations can increase crop yield and soil health maintenance (Patil, 2022). It contains a variety of well defined models, including classification, regression, Support vector machines, and clustering, which collect specific data and use specific techniques to get desired conclusions. A practical answer that can be found using a machine learning model is what the system aims to provide. To forecast soil health, the support vector machine, random forest, and linear regression approaches are used. The lifespan of agricultural land and good crop production are both steadily enhanced by scientific management of soil health. The theory, performance, and attributes of learning systems and algorithms are the formal emphasis of the machine learning in research field. It is a very multidisciplinary field built on a variety of fields, including artificial intelligence, optimisation theory, information theory, statistics, cognitive science, optimal control, and many others (Cherkassky and Mulier 2007). Almost every scientific field has been covered by machine learning because to its numerous applications, having a tremendous influence on both research and society (Rudin and Wagstaff 2014). It is used in autonomous control systems, informatics and data mining, recognition systems, and recommendation drivers (Qiu *et al.,* 2016). Machine learning offers helpful and best-in-class algorithms for tracking soil health and classifying it into categories of healthy and bad soil. The interest in utilizing machine learning techniques to predict the values of various soil attributes helps to reduce needless fertilizer input spending by studying soil health and environmental quality. Hence, this approach would help farmers choose the best crop and production forecast. Therefore, in the present study we have developed crop recommendation system based on soil physical property will enhance the nation's total economy to expand relative to other nations.

METHODOLOGY

Study area: The current research was carried out in the district of Biswanath, which is 48-849 m above mean sea level (Fig. 1). The region covers 26°35' to 27°00' of latitude and 92°50' to 93°50' of longitude, with a total area of 14,15,000 ha. The area has a monsoon climate because it is located in the Subtropical climatic area. With an average temperature of 28.5°C, summers are hot and humid. The region has on average 253.5mm of rainfall and 79.65% relative humidity. At a depth of 5 cm, 10 cm, and 20 cm, the average yearly soil temperature is 29.16°C, 26.75°C and 27.7°C, respectively, while the average vapor pressure is 18.97 Pascal.

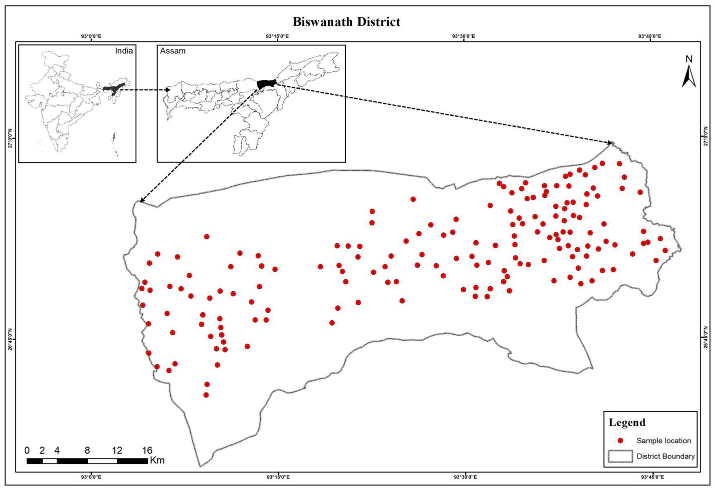

Fig. 1. Location map of study area (Biswanath District, Assam

Soil sample collection: Five primary cropping systems, including (1) rice-fallow (RF), (2) rice-potato (RP), (3) rice-rabi vegetables (RR), (4) fallow-rapeseed (FR), and (5) fallow-blackgram, were assessed during a survey of the Biswanath District (FB). The rice crop was planted (June-October/ November) and succeeding crops; i.e. potato, rapeseed and blackgram were produced as Rabi crop (November–February). The soil samples were taken from each location following the harvest of the agricultural system between march and april. At random places covering an area of 4.5 to 5 km^2, soil samples were taken at a depth of 0 to 30 cm. 180 soil samples in total were taken for testing, and their locations are shown on the map of soil sampling sites (Fig. 1).

Analysis of soil sample. The physical properties of soils were determined following standard procedures (Table 1). The crop recommendation system is developed using three modules namely the pre-processing, feature engineering and model fitting-evaluation subsystem. The pre-processing module is intended to clean the raw datasets by eliminating duplicates, reducing the dimension of the datasets and by eliminating irrelevant samples. The feature engineering module performs the two major tasks of feature generation and feature selection. The model fitting and evaluation module is responsible for training-testing models and identifying a model with the best accuracy. Decision Tree Classifier Model is used for recommendation of crop in the present study. Fig. 2 captures the flow of the model fitting-evaluation module.

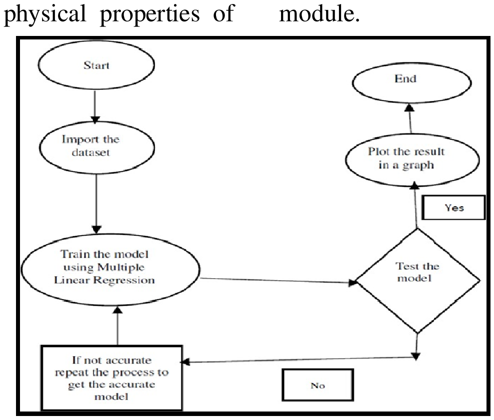

Fig. 2. Block Diagram of Proposed System.

Table 1: Procedures used for analyzing physical properties of soils.

Sr. No.	Parameters	Methods	References
1.	Particle size analysis (%)	International pipette method	Piper (1966)
2.	Hydraulic conductivity (cm hr^{-1})	Constant head method	Klute (1965)
3.	Field capacity (%)	Pressure plate technique	Richards (1948)
4.	Permanent wilting point (%)	Pressure plate technique	Richards (1948)
5.	Available water content (%)	Field capacity-Permanent wilting point	Richards (1948)
6.	Water holding capacity (%)	Keen-Rackzowski box method	Piper (1966)
7.	Bulk density (g cm^{-3})	Core method	Bodman (1942)
8.	Particle density (g cm^{-3})	Pycnometer (water displaced method)	Black (1965)
9.	Soil aggregate (%)	Wet sieving method	Yoder (1936)
10.	Mean weight diameter (mm)	Summation method	Van Bavel (1949)

RESULTS AND DISCUSSION

The soil texture varied from sandy loam to clay loam (Table 2). The variation in soil texture in the North Bank Plain Zone of Assam is due to the soil forming factors characterized by recent and old alluvium soil (Hazarika *et al.,* 2016). The sand content ranged (Table 2) from 24.12 to 77.6 per cent with the mean value of 56.60 per cent. The presence of higher sand content indicates better water flow leading to a better drainage pattern, which facilitates the multiple cropping systems. The silt content ranged from 7.90 to 39.50 per cent with the mean value of 21.53 per cent (Table 2). The clay content ranged from 5.6 to 39.81 per cent with the mean value of 21.81 per cent (Table 2). The distribution of clay content bears a significant aspects of scientific study on soils. The variation of clay content is due to the high rainfall area and more so in alluvial soils. The bulk density ranged from 1.14 to 1.46 g/cm^3 with the mean value of 1.31 g/cm^3 (Table 2). The observation indicates that the bulk density in the soils of Biswanath District does not pose any restriction in agricultural activities as majority of the soil had medium bulk density. The soils showing a higher bulk density (>1.30 g/cm^3) which might be due to continuous use of farm machinery like power tiller and tractor and lesser incorporation of organic inputs (Gayan *et al.*, 2020). The particle density ranged from 2.22 to 2.67 g/cm^3 with the mean value of 2.49 g/cm^3 (Table 2). The observation on particle density stated that the soils of the district were mainly formed from parent materials originated from light minerals and also soils have adequate component of organic carbon in it.

Table 2: Descriptive statistical parameters of soil properties.

Soil properties	Minimum	Maximum	Mean	S.D.	C.V. (%)
Sand (%)	24.12	77.60	56.60	16.31	28.82
Silt (%)	7.90	39.50	21.53	8.37	38.87
Clay (%)	5.60	39.81	21.81	9.25	42.42
Hydraulic conductivity (cm/hr)	1.28	6.90	3.73	1.58	42.49
Particle density (g/cm^3)	2.22	2.67	2.49	0.11	4.76
Bulk density (g/cm^3)	1.14	1.46	1.31	0.07	6.02
Total porosity (%)	41.22	53.84	47.16	2.73	5.80
Maximum water holding capacity (%)	26.10	42.90	33.69	4.42	13.13
Volume expansion (%)	6.18	10.85	8.12	0.86	10.69
Field capacity (%) at 0.3 bar	20.40	36.68	28.54	4.05	14.20
Permanent wilting point (%) at 15 bar	4.45	8.98	6.73	1.11	16.49
Available water (%)	14.78	31.86	21.81	4.23	19.39
Microaggregate (%)	30.24	59.54	43.13	5.34	12.38
Macroaggregate (%)	40.45	69.75	56.87	5.34	9.39
Mean weight diameter (mm)	1.42	2.48	1.87	0.17	9.09

The total porosity ranged from 41.22 per cent to 53.84 per cent with the mean value of 47.16 per cent (Table 2). The increase in soil porosity was accomplished by decrease in soil bulk density as reported by Werner (1997). Similar results were reported by Swezey *et al.* (1998); Glover *et al.* (2000). The hydraulic conductivity ranged from1.28 cm/hr to 6.90 cm/hr with the mean value of 3.734 cm/hr (Table 2). This range of hydraulic conductivity indicating better porosity and lower compaction of soil due to medium to high organic carbon. The maximum water holding capacity ranged from 26.10 to 42.90 per cent with the mean value of 47.168 per cent (Table 2). The volume expansion ranged from 6.18 per cent to 10.85 per cent with the mean value of 8.12 per cent (Table 2). The water content at field capacity at 0.33 bar suction ranged from 20.40 per cent to 36.68 per cent with the mean value of 28.54 per cent, indicating slightly higher water content in surface soil which is due to effect of porosity (Table 2) mediated by the organic matter content of the soils. The differences in soil water characteristics might be attributed to variation in clay and organic carbon content (Kumar *et al*., 2020; Rajpoot *et al.*, 2021). The water content at permanent wilting point at 15 bar suction ranged from 4.45 per cent to 8.98 per cent with the mean value of 6.73 per cent (Table 2). Therefore, the available water ranged from 14.80 to 31.86 per cent with the mean value of 21.81 per cent (Table 2). The microaggregate ranged from 30.24 per cent to 59.54 per cent with the mean value 43.13 per cent (Table 2). The macroaggregate ranged from 40.45 per cent to 69.75 per cent with the mean value of 56.87 per cent. The mean weight diameter ranged from 1.42 mm to 2.48 mm with the mean value of 1.88 mm (Table 2). The result justifies the variation of structural stability of different cropping system.

Crop recommendation system. Soil is a nonrenewable dynamic natural resource that is necessary for life. Previously, crop cultivation was done by farmers who had hands-on experience acquired from forefathers. Farmers are no longer able to choose the best suitable crop based on soil physical properties. Therefore, recommendation system has been developed that employs machine learning algorithms to recommend the crop that can be grown in particular soil. There are several machine learning algorithms including decision tree, random forest, naive bayes and gradient boosting to recommend the crop. In the present study "decision tree classifier model" (Fig. 3) was used for crop recommendation system. In supervised learning for classification and regression, decision trees (DTs) are included. A decision tree is built using the independent variables, with a condition over a feature at each node. To determine the class of the given dataset, the procedure starts at the root node of the tree.

	precision	recall	f1-score	support
Blackgram	0.60	1.00	0.75	3
Potato	0.67	0.67	0.67	3
Rabi_vegetables	1.00	0.50	0.67	2
Rapseed	0.00	0.00	0.00	2
Rice	1.00	1.00	1.00	55
accuracy			0.94	65
macro avg	0.65	0.63	0.62	65
weighted avg	0.94	0.94	0.93	65

Fig. 3. Decision tree classifier model summary.

Fig. 4. Crop recommendation application for user interface (UI).

Decision trees are, to put it simply, a collection of if-else expressions. It determines if the condition is met and, if so, advances to the following node in the decision chain. The nodes choose which node to visit next based on the criteria. After the leaf node is reached, output is anticipated. When the circumstances are in the right sequence, the tree functions effectively. Therefore in the present study we used this model which has produced an accuracy around 94 percent . Similar results were also found by Kamatchi and Parvathi (2019) on crop yield prediction using a data mining techniques. A web application was developed to connect with this model. Using the web application model is exposed to end user (Fig. 4).

CONCLUSIONS

In Assam's Biswanath District, crop recommendation systems was forecasted using machine learning techniques. The crop recommendation system utilized application of the "decision tree classifier model". The model's accuracy was almost 94%. A web application is being created to link the model. This application presents the model to the user. The study recommends that farmers can change the type of crop planted in their field during various season depending on the physical properties of the soil in order to minimize crop failure.

FUTURE SCOPE

From the outcome of the present investigation, it could be inferred that future studies may be carried out with the images of soil samples and implement various deep learning techniques, so that soil health assessment can become easier which would improve the economy of the state.

Acknowledgements. This work was a part of the Ph.D. research of department of Soil Science at Assam Agricultural University (AAU), Jorhat, India. The first author is appreciative to Mr. Simanta Jyoti Medhi, Senior Software Engineer at Volvo Group, India for aiding the application of machine learning techniques in the research.
Conflict of Interest. None.

REFERENCES

Black, C. A. (1965). Methods of Soil Analysis: Part I, Physical and Mineralogical Properties. American Society of Agronomy, Madison, Wisconsin.

Bodman, G. B (1942). Nomograms for Rapid Calculation of Soil Density, Water Content, and Total Porosity Relationships. *Agron. J., 34*(10), 883-893.

Cherkassky, V. and Mulier, F. (2007). *Learning from Data: Concepts, Theory, and Methods*; John Wiley & Sons: Hoboken, NJ, USA.

Gayan, D. J., Nath, B., Bhattacharyya and Dutta, N. (2020). Assessment of soil quality indicators under rice ecosystem of Assam. India. *J. Envt. Biol., 41*, 1655-1664.

Glover, J. D., Reganold, J. P. and Andrews, P. K. (2000). Systematic method for rating soil quality of conventional, organic, and integrated apple orchards in Washington State. *Agric. Ecosyst. Envt., 80*, 29-45.

Hazarika, M., Sarma, P. K., Borah, P., Baruah, H. C., Deka, B., Moral, G., Sarma, D., Bhattacharjee, M., Sarma, M. K. and Rao, S. C. (2016). Long Term Effect of Integrated Nutrient Management on Soil Properties and Productivity of Toria (*Brassica campestris*) in an Inceptisol under Rainfed Upland Condition of North Bank Plains Zone of Assam. *Indian J. Dryland Agric. Res. Dev., 31*(1), 9-14.

Kamatchi, S. B. and Parvathi, R. (2019). Improvement of Crop Production Using Recommender System by Weather Forecasts. *Procedia Computer Sci., 165*, 724-732.

Klute, A. (1965). Laboratory measurement of hydraulic conductivity of saturated soil. In: Methods of Soil Analysis Part-I. American Society of Agronomy, Inc.

Kumar, U., Mishra, V. N., Kumar, N., Srivastava, L. K. and Bajpai, R. K. (2020). Soil physical and chemical quality under long-term rice-based cropping system in hot humid eastern plateau of India. *Commun. Soil Sci. Plant Anal., 51*(14), 1930-1945.

Patil, P. (2022). Soil Health Prediction Using Supervised Machine Learning Technique. *Intl. J. Res. Appl. Sci. Engg. Technol. (IJRASET), 10.*

Piper, C.S. (1966). Soil and Plant Analysis. University of Adelaide, Australia.

Qiu, J., Wu, Q., Ding, G., Xu, Y. and Feng, S. (2016) A survey of machine learning for Big Data processing. *EURASIP J. Adv. Signal Process, 1*, 1–16

Rajpoot, R. S., Bajpai, R. K., Shrivastava, L. K., Kumar, U., Tedia, K. and Mishra, V. N. (2021). Evaluation of soil physical and chemical properties under rice-based cropping system in alfisols of Northern Hill Region of Chhattisgarh. *Intl. J. Curr. Microbiol. Appl. Sci., 10*(1), 2748-2761.

Richards, L. A. (1948). Porous plate apparatus for measuring moisture retention and transmission by soils. *Soil Sci., 66*, 105-110

Rudin, C. and Wagstaff, K. (2014). Machine learning for science and society. *Mach. Learn.*, *95*, 1–9.

Swezey, S. L., Werner, M. R., Buchanan, M. and Allison, J. (1998). Comparison of conventional and organic apple production systems during three years of conversion to organic management in coastal California. *American J. Altern. Agric., 13*(4), 162-180.

van Bavel, C. H. M. (1949). Mean Weight-Diameter of Soil Aggregates as a Statistical Index of Aggregation. *Soil Sci. Soc. America J., 14*, 20-23

Werner, M. R. (1997). Soil quality characteristics during conversion to organic orchard management. *Appl. Soil Ecol., 5*(2), 151-167.

Yoder, R. E. (1936). A direct method of aggregate analysis of soils and a study of the physical nature of erosion losses. *J. American Soc. Agric., 28*, 337-351.

How to cite this article: Samikhya Bhuyan, Dilip Kumar Patgiri, Simanta Jyoti Medhi, Roopali Patel and Tabuiliu Abonmai (2023). Machine Learning-based Crop Recommendation System in Biswanath District of Assam. *Biological Forum – An International Journal, 15*(3): 417-421.