

Soil Nutrient Analysis Using Machine Learning Techniques

Madhumathi R¹, Arumuganathan T², Sneha Iyer R³, Shruthi R⁴, Shruthi K⁵

^{1,3,4,5}Department of Computer Science and Engineering,

Sri Ramakrishna Engineering College, Coimbatore, India

Email: madhumathi.r@srec.ac.in, snehaiyer.1801224@srec.ac.in

²Principal Scientist, Sugarcane Breeding Institute, Coimbatore, India

Email: arumuganathan@gmail.com

Abstract: Agriculture is an important aspect in determining the overall growth of a nation, both in terms of economic and also in terms of progress in the people's lives. Nutrient analysis forms the main criteria for producing a healthy crop. Hence, adopting different measures for soil analysis has become the need of the hour. Primarily, the soil analysis prediction consists of determining the composition of nitrogen(N), phosphorus(P) and the potassium(K) content present in the soil. These are the major nutrients that the crop requires for its growth. Machine learning algorithms are very effective in predictive analytics and are widely used in the field of agriculture. It has a set of well-defined models like classification, regression, Support vector machine(SVM) or clustering that collect specific data and apply specific algorithms to achieve expected results. The system aims at providing a feasible solution which can be obtained through the machine learning model. For determining the N, P, K composition in the soil the multiple linear regression(MLR) algorithm is used. Thus, this method would help the farmers in choosing the right crop and fertilizer to have better harvest of crops. The prediction technique of soil nutrient content would thereby increase the country's overall economy among other nations.

Keywords- Multiple linear regression, Soil nutrients, Nitrogen, Phosphorus, Potassium.

1. INTRODUCTION

Agriculture is a vital aspect for a nation in determining its overall progress among the other nations. It is the art of living in the country like India. Even with the establishment of the technological era, agriculture still continues to be the one of the major areas wherein technical advancements are highly appreciated. To have an effective production of the crop and adding fertilizers in the right ratio in the soil, it is important for the farmers to know the soil nutrient composition. Hence, determination of soil nutrient analysis has become the need of this today's world. The machine learning algorithm plays a major role and gives faster and accurate results. The quantity of the N, P, K present in soil can be determined using the laboratory methods. Here the machine learning algorithm is used for the prediction for soil nutrients.

It is a booming technology in the field of

agriculture. Smart farming powered by machine learning with its high-precision algorithms is an emerging concept. It includes crop quality identification using image processing, yield prediction using various algorithms, disease detection using sensors and algorithms, field conditions monitoring like soil management, water management and crop management. There are various algorithms that are used for addressing different algorithms. Machine learning algorithms like regression, classification are mainly used for prediction and it proves to give valuable results. There are various other techniques as well that are used in agricultural field. Sensor, devices with smart optimization techniques are also employed in this field. The main motto is to predict the results accurately through the implementation of the MLR technique

The artificial neural network(ANN)model was developed for determining the soil nutrients for a specific plant [1]. Machine learning algorithms like multiple linear regression and classification algorithms were used for comparing their results in soil nutrient analysis[2].A study was proposed which focuses on the various types of the soil,crop types and the soil test reports. A suitable machine learning model was developed for prediction using the implementation of the k-nearest neighbor's technique[3].The determination of the soil nutrient analysis using recurrent neural network,SVM and MLR models were developed. The accuracy of each model was compared[4].

This paper discusses the methodology and solution employed to solve the problem for detecting soil nutrients. Henceforth, it would be very useful for the farmers for growing better crops.

II.RELATED WORK

The soil nutrients were analyzed using the regression algorithm and the root mean square error is calculated for the prediction of N, P and K[5]. A consecutive comprehensive evaluation models has been presented for the soil nutrient analysis using the SVM, ANN and MLR[6]. A study has been conducted on the various machine learning techniques which are available for crop determination and other agriculture related practices[7].The soil fertility and the plant nutrients were analyzed using Supervised Machine learning, Back Propagation Neural Network(BPN)[8].The reference crop's growth properties are trained to find the actual values. BPN will find the correlation percentage of the properties trained and its performance is evaluated using the test data. A system was proposed which integrates a sensing module with an image processing setup to determine the essential characteristics for plant growth. The image processing mainly helps in the capturing the soil images. Features extraction is performed for determining the results[9].The feature values obtained were analyzed by comparison with the database.SVM was used to classify the ratio level of N, P, K. The soil moisture, pH, temperature,

humidity and N, P, K values are displayed in a mobile application. The soil pH, Electrical Conductivity, Moisture, temperature and N, P, K values were analyzed and based on the data obtained the accurate crop to be grown in that soil along with the suggestion of fertilizers were predicted using Supervised Machine learning algorithms[10].

A study was proposed where the soil features like village wise soil fertility indices of Available Phosphorous, Available Potassium (K), Organic Carbon (OC), Boron (B) and Soil Reaction(pH) are classified by the index analysis test report values[11]. The classification is done using the Extreme Learning Machine (ELM) with different activation functions. This learning helps in the better production of better results.A system was designed where the Soil fertility index for the nutrients is predicted at different places. The prediction is done automatically using regression methods specifically like support vector regression and more[12].

The data obtained is useful for making decision of amount of fertilizers to be used. A system was proposed to determine the pH value of soil and based on the pH value obtained the NPK values are estimated[13]. The SVM algorithm and the artificial neural networks play a major role in the analysis of soil nutrient[14].Using classification algorithms, the suitable crops and the amount of fertilizer is recommended for that land in a handheld device.

III.PROPOSED SYSTEM

The proposed system predicts the amount of soil nutrients namely N,P,K using the trained dataset which has various parameters and the multiple linear regression algorithm. Thus, knowing these values of soil nutrients, it would be easier for the farmers for adding fertilizer in a right range and produce better yield.

A.Methodology

The aim is to predicted the values of N,P,K using the machine learning algorithm. The dataset is

being imported in RStudio platform, it consists of parameters like the electrical conductivity, pH, soil moisture, soil temperature and the corresponding values of N,P,K in the soil. These parameters play a significant role in development and growth of a plant. Thus, these parameters are used as variables where the nitrogen, potassium and phosphorous forms the target variable respectively.

The dataset is trained using the MLR model. This model is used because of its efficiency to handle multiple variables. The scatterplot and pairs function which is used for the creation of graph of the model that shows the contribution of each parameter.

There are different values of each nitrogen, potassium and phosphorous values predicted after giving the parameters pH, moisture, electrical conductivity, temperature as an input. The predicted values are then used for creating graphs for N, P, K respectively.

Figure.1 is the block diagram of the proposed system which shows the process of predicting the nutrient values. It starts from importing of the dataset that has all the parameters and the respective N,P,K values. Train the model with the multiple linear regression algorithm. The model is being tested with different values as input to check its efficiency. The graph is being plotted for the respective target values by taking predicted values as x-axis and the parameters as y-axis.

B. RStudio

The RStudio is a platform that is an integrated development environment for R that allows to perform various functions related to the predictive, statistical and graphical analysis. Various packages can be used for different applications and its implementation. Thus, RStudio is a great platform for performing machine learning algorithms, predicting values and for graphical analysis.

C. Multiple Linear Regression Algorithm

It is the most common form of Linear Regression. It basically describes how a single response variable Y depends linearly on a number of predictor variables. The model is mainly used when there are multiple parameters and one target variable to be found.

The target variable is the one that would be depended on the other parameters. Thus, the variables like the pH, electrical conductivity, soil moisture and soil temperature are considered. The dataset is trained to get the predicted values of N,P,K. The general formula of multiple linear regression algorithm:

$$Y = \beta_0 + \beta_1X_1 + \dots + \beta_kX_k + e$$

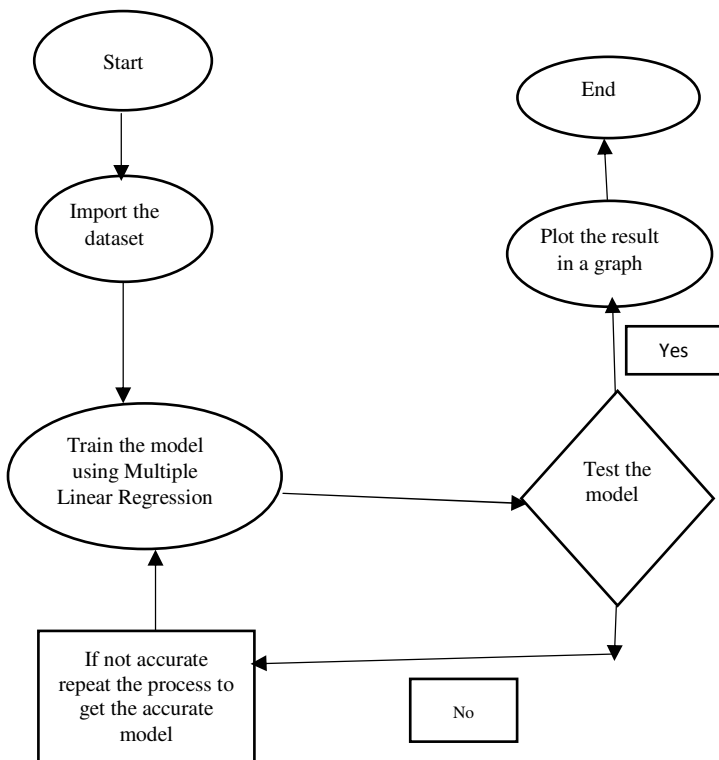


Fig.1 Block Diagram of the Proposed System

IV.RESULT ANALYSIS

The table1 shows that the dataset’s variables have huge significance the highest one goes to the moisture [***] and then the pH [**] has its second most significance and then comes the conductivity [**] which has third most significance. It has a p-value of 2.847e-09 which tells that the model is quite efficient and thus most of the data contributes for the prediction.

**TABLE 1
CONTRIBUTION OF EACH VARIABLE**

	Estimated standard error t value pr(<lt)				
Intercept	530.009	161.064	3.291	0.00141	**
Temperature	5.737	3.364	1.706	0.09137	*
Moisture	6.826	1.36	5.02	2.44-e06	***
pH	33.796	10.483	3.224	0.00174	**
Conductivity	338.417	127.663	2.651	0.00942	**

The amount of nitrogen,potassium and phosphorus is predicted using the multiple linear regression and gives an accuracy of the model for about 78%.

**TABLE 2
THE PREDICTED VALUES OF NITROGEN**

Temperature	Moisture	pH	Conductivity	Nitrogen
24	60	5	0.045	320.27
17	85	4.7	0.035	312.41
19	75	5	0.035	329.92
15	95	5.2	0.25	381.63
25	55	4	0.031	310.16

**TABLE 3
THE PREDICTED VALUES OF PHOSPHOROUS**

Temperature	Moisture	pH	Conductivity	Phosphorous
25	99	5	0.31	353.25
21	84	5.1	0.36	214.37
24	93	5.2	0.32	309.94
18	73	5.1	0.34	128.85
20	83	4	0.24	205.25

**TABLE 4
THE PREDICTED VALUES OF POTASSIUM**

Temperature	Moisture	pH	Conductivity	Potassium
24	90	4.7	0.025	16.11
17	68	4.1	0.05	21.54
25	98	4.2	0.35	8.38
24	92	5.2	0.31	9.72
19	75	4.2	0.28	14.96

Thus, the table2 shows the predicted value of nitrogen when the other parameters temperature, moisture, pH and conductivity given as input. Thetable3 shows predicted value of potassium when the other parameters like the temperature, moisture,pH and conductivity given as input.Thetable4shows the predicted the value of phosphorous aswhen the other parameters temperature, moisture,pH and conductivity given as input.Henceforth, the predicted values could be compared with the dataset values to the efficiency of the model.

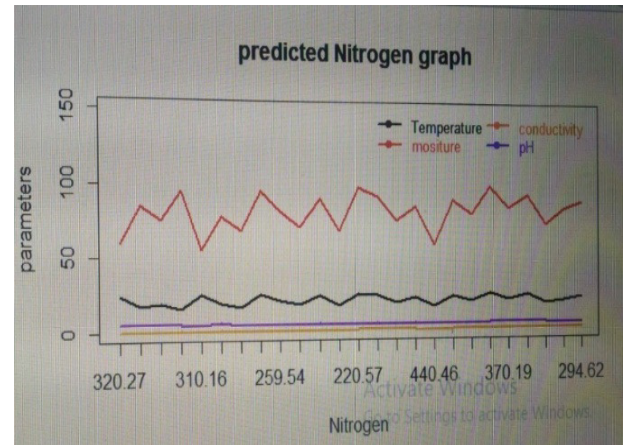


Fig.1 Predicted graph of Nitrogen

V. CONCLUSION

This system predicts the values of N,P,K using the multiple linear regression model and gives the composition of the soil nutrient i.e. N, P, K content present in the soil. The multiple linear regression model used for prediction gives an accuracy of 78%. This accuracy makes it almost a suitable approach as well as a trustworthy technique in determining the fertility of the crop with a good amount of efficiency maximum possible. Hence machine learning technique proves to be an effective method for prediction of soil nutrient and as well as its analysis.

This could further be extended by developing a mobile application where the values of N, P, K that is obtained from the model could be displayed in the application developed. The values can be stored in a database and displayed in the app. Hence the system benefits the farmers and helps them taking right decisions for growing crops.

ACKNOWLEDGMENT

The authors sincerely thank the Science and Engineering Research Board (SERB) for the financial support, the Director, ICAR-SBI, Coimbatore, and the Management and Principal of Sri Ramakrishna Engineering College, Coimbatore for extending the required facilities for this study.

REFERENCES

- [1] C.Wu, et al "Evaluating soil nutrients of Dacrydium pectinatum in China using machine learning techniques", International Journal of forest ecosystems, Volume 6, Issue 7, 2020.
- [2] G.Forkuor et al "High resolution mapping of soil properties using remote sensing variables", Plos one publishers, 2017.
- [3] P.Yadav, "Soil health analysis for crop suggestions using machine learning", International Journal of research and development, Volume 6, Issue 12, 2019.
- [4] S.Prakash, "Soil moisture prediction using machine learning", Institute of electrical electronics engineering, Volume 8, Issue 7, 2018.
- [5] K.Kumar, "Random forest algorithm for soil fertility prediction and grading using machine learning "International journal of innovative technology and exploring engineering, Volume 1, Issue 9, 2019.
- [6] H.Li, "Evaluation models for soil nutrient based on support vector machine and artificial neural networks", Hindawi publishers, 2014.
- [7] Konstantinos and et al, "Machine learning in agriculture", International Journal of Innovative Science and Research, Volume 6,

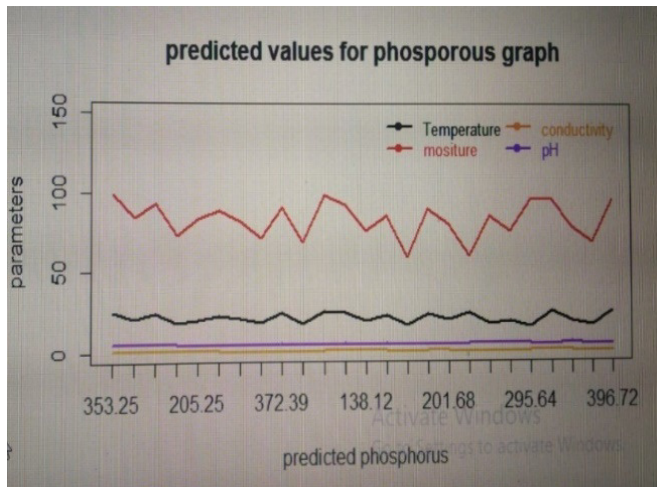


Fig.2 Predicted graph of Phosphorous

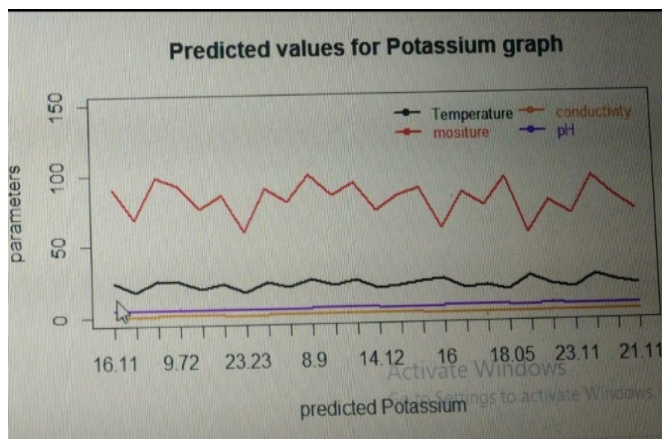


Fig.3 Predicted graph of Potassium

Hence these are the predicted values of N,P,K where the line red depicts moisture, black depicts temperature, blue depicts pH and orange depicts conductivity. The fig.1 shows the graph of predicted values of nitrogen to the parameters. The fig.2 shows the graph of predicted values of phosphorous to the parameters. The fig.3 shows the predicted values of potassium to the parameters.

Issue 7,2018.

- [8] S.Ghosh and S.Koley, "Machine Learning for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks", International Journal on Recent and Innovation Trends in Computing and Communication(IJRITCC), Volume 2, Issue 2, pp.292-297, 2014.
- [9] S.Rahaman, et al "Detection of NPK Ratio Level Using SVM Algorithm and Smart Argo Sensor System", International Journal of Latest Research in Engineering and Technology (IJLRET), Volume 3, Issue 7, pp.11-15, 2017.
- [10] S. Panchamurthi, M. D. Perarulalan, A. Syed Hameeduddin and P. Yuvaraj, "Soil Analysis and Prediction of Suitable Crop for Agriculture using Machine Learning", International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 7, Issue 3, pp. 2328-2335, 2019.
- [11] M.S. Suchitra and Maya L. Pai,"Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters", Information Processing in Agriculture, Elsevier, 2019.
- [12] M.S. Sirsat, E. Cernadas, M. Fernández-Delgado and S. Barro, "Automatic prediction of village-wise soil fertility for several nutrients in India using a wide range of regression methods", 2018, Computers and Electronics in Agriculture, Volume 154, pp.120-133.
- [13] K.Abhang, S.Chaughule, P.Chavan and S.Ganjave, "Soil Analysis and Crop Fertility Prediction", International Research Journal of Engineering and Technology (IRJET), 2018, Volume 5, Issue 3, 2018.
- [14] J.C Puno, E.Sybingo, E.Dadios, I.Valenzuela and J.Cuello, "Determination of Soil Nutrients and pH level using Image processing and Artificial neural networks", IEEE,2017.