

Crop Suggestion Using Machine Learning Based on Soil Conditions

M.S. Roobini¹, R. SivaSangari², L. Sujihelen³, T. Ananthi⁴, G. Nagarajan⁵

Assistant Professor^{1,2,3,4}, Sathyabama Institute of Science and Technology, Chennai.

Professor⁵, Sathyabama Institute of Science and Technology, Chennai.

¹roobinims@gmail.com

Abstract –

Crop growth and production depend on multiple circumstances such as soil conditions, climatic changes and more importantly the usage of artificial fertilizers which used excessively can cause permanent damage to the farmland. In order to reuse these excess fertilizers. There are several types of soil and they differ based on how soil is nourished the climate they type of soil and due to that different crops are being grown in different states and different countries. In order to determine the yield rate, we need to know the soil conditions. Machine learning is a technology which can predict an outcome based on previous dataset and the idle dataset. The larger the data gathered the accurate the data is. To suggest suitable crop there are various machine learning algorithms such as k – nearest neighbour (KNN), Gaussian support vector machine (SVM), braggged tree. These are good algorithms present compared to others.

Keywords - Machine learning, k – nearest neighbour (KNN), Gaussian support vector machine (SVM), braggged tree.

I. INTRODUCTION

1.1. Importance of agriculture in India's economy

Agriculture plays an important role in Indian's economy while considering that. the methods used for crops to produce are harsh and not suitable for present conditions. The excessive usage of harmful chemicals and pesticides can cause permanent damage to the field which can result in improper growth in agriculture field. Based on the data acquired from the sample of the soil taken from the field we can determine which crop might be the best option for the soil[1,2,3,4].

This paper consists of all the detailed methods and working of this project with accuracy. Initially to predict something there has to be a certain valuation, upper limits and lower limits which determine the contents precisely without any complex quantities. Generally, when a farmer consults an agriculture industry, they are some requirements they need to carry such as percentage of potassium (K), phosphorous (P), Iron, Magnesium, calcium etc. This data provides a view over the soil and can show what minerals the soil is lacking for the growth. Instead of consulting a large industry our objective is to create a shorter and faster way to determine what crop is best suitable for agriculture in that field. To do so a web application is created as interface where one can give the data of minerals of their field and comparing that data the machine learning program sorts a better solution for the problem not only on the basis of minerals but also seasonal basis this classification is done by sorting them as Kharif crops and Rabi crops, Kharif season is on the start of summer rainfalls (Autumn), Rabi season typically comes during winter season. Usually, this classification helps to predict crops based on the seasons. Considering a state Tamilnadu, Soil test and yield target-based fertilizers are prescribed under Integrated Plant Nutrition System (STCR-IPNS) have been developed

for 31 agricultural and high valued crops and 10 major cropping sequences on 17 soil types covering 6agri-climatic zones of Tamil Nadu[5,6,7,8].

Agro climatic Zone	Sample size	Zn	Fe	Mn	Cu	B	S
North Eastern Zone	13,741	46.0	9.25	11.7	17.0	17.0	6.39
North western zone	5,392	48.2	6.81	3.96	4.15	20.1	13.0
Western zone	3,389	38.8	15.9	11.8	21.4	52.7	5.45
Cauvery Delta zone	8,991	33.8	7.64	2.84	13.9	14.1	7.68
Southern zone	13,898	56.8	13.4	4.36	24.6	15.7	19.2
High rainfall zone	617	18.3	6.36	0.43	15.9	18.9	7.25
Hilly zone	152	51.1	0.00	0.00	0.37	18.4	5.83
State Overall	46,180	42.0	10.0	5.84	16.7	19.9	11.2

Fig 1.1 Collected and analysed the soils samples (46,180) from all the districts in Tamil Nadu at village level for secondary and micronutrients.

Above figure represents a simple survey over the sample all over Tamilnadu. Rest of the paper is organized in according sequence. Second section is all about the Related Work through which this work of ours have been developed with the reference of works done in the past based on the same field. Third section shows the methods used in this work that might help in producing a new model. Next section is all about the work proposed which consists different methods used. In final section, analysis of developed product and results were displayed[9,10].

II. METHODOLOGY AND MATERIALS

2.1. Technology and algorithms implemented

Machine learning (ML) as the term defines it is a concept where a machine can learn what to do based on simple algorithms it is a part of Artificial Intelligence (AI). A machine with a proper algorithm can deduce a problem for a better solution on its own based on the previous datasets and instructions given.

Weighted K-NN Equation

Weighted k-NN is derived from k-NN algorithm which helps to weight all the k neighbours corresponding to the distance from query point x_q , resulting closest neighbours with respect to weight w_i .

$$F(x_q) = \frac{\sum_{i=1}^k w_i f(x_i)}{\sum_{i=1}^k w_i}$$

i.e., the weight is

$$w_i = \frac{1}{d(x_q, x_i)^2}$$

If x_q equals to x_i where the denominator is equal to zero, then $F(x_q)$ equals $f(x_i)$. It makes sense to use all training problems not only k if weighting is used, the algorithm now turns in a global one. Only down-fall is the algorithm can run slower than anticipated.

“Support Vector Machine” (SVM) is an algorithm that will be used for classification and regression problems in machine learning. although [11,12], this is often frequently utilized in classification problems. this algorithm uses an n -dimensional space in which each data item is plotted with a value of every assert being a specific coordinate. Then, we perform classification based on finding the hyper-plane that differentiates the two classes well.

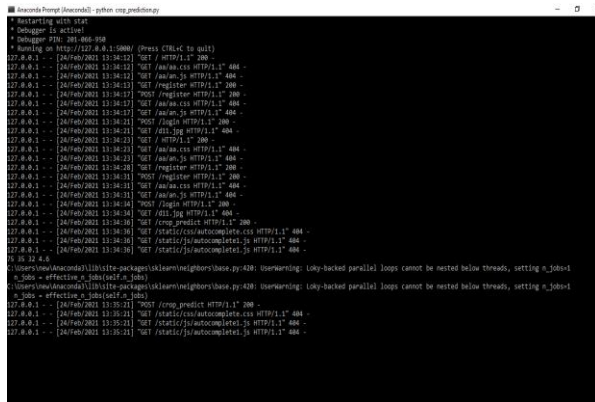


Figure 2.1: System log of datasets being transferred to web page.

This python program helps form a bridge between all the datasets stored in excel sheets to the web page created. Backend as mentioned is to be python and as for the front-end, HTML and CSS were the best options to create a good interface.

Data stored in excel sheets are refined data for different types of crop and their attributes such as data required mineral resources, pH values.

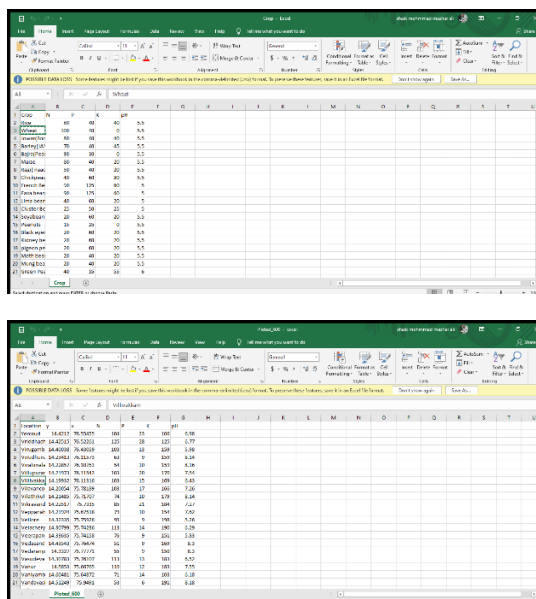


Figure 2.2. Above created datasets determine the outcome for the request by the user.

These datasets contain all the value such as Nitrogen (N), potassium (K), phosphorous (P) and also pH values. Based on these values given by the user of their samples. The comparison can determine if the user requires either addition in the minerals or dilution using a different plantation.

The question of dilution may sound complex but it is possible by planting a different attributed crop which absorbs the excess amounts regulating the soil perfect for next plantation frequently.

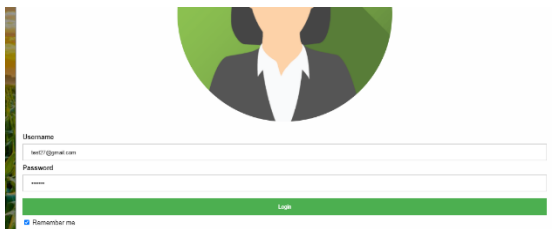
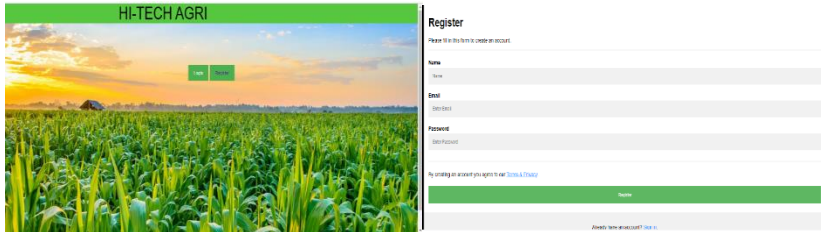


Figure 2.3. Homepage

This page contains both login and registration page. Registration page allows you to sign-up for the site.



Figure 2.4. Selection page

This page has an interface which has several options crop recommender where you can also get suggestions on basis of locations.



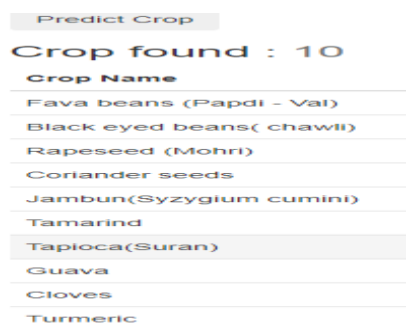


Figure 2.5. PredictedOutcome.

The above prediction is on the basis of manual data of minerals given by the user. So, the favourable crops selected are shown as a result of the given data.



Figure 2.6: Above crop prediction by user's data for a required crop of his choice and his parameters

This page can show in what attributes your soil is lacking or have excess. Using this one can determine the quantities of minerals required for better growth.

III. RELATED WORK

B. Vardhan and D. Ramesh users of DM (data mining) [2] were dealing with the challenges for predicting yield of various different crops in various fields. One's solution to this kind of problem is to implement data mining methods. In present paper work, various types of data mining techniques were utilized and examined based on the datasets prepared. These datasets are collective data of mineral consistency in a geographical location. S. Pudumalar discussing the hardships that producers such as farmers suffer because irregular and inadequate crop selection in order to cultivate in a particular kind of the soil available to them. Reason behind this is the lack of knowledge [8]. They have designed some sort of system for recommending crop that can work as a reference for farmers for selecting the best and desired crops for farming. After a several experiments on these algorithms, they ended up using these algorithms such as Gaussian SVM and k-NN, and came up with a concluding system designed to predict outcome resulting approximately 88% accurately. Even though they ended up examining models in one particular location, this model has barely any changes in the data and gives us at most accurate recommendations for large areas.

Padikkal Pramod Tiwary, Padikkal Duraisamy Vasu, Surendra Kumar Singh, Chandran [3] on the importance of soil and effects of its degradation. Being essential for functioning of all the terrestrial ecosystems, soil is deliberate in food production and for feeding the rapidly increasing population of the world. Nevertheless, food-production and storage are facing threat from soil

degradation occurring worldwide. Soil degradation takes place because of the force from various reasons such as urbanization and industrialization. The major forces that cause soil degradation are deforestation, changing land usage, soil erosion and chemical waste disposal. Worldwide, 24% of the present land is facing degradation resulting 50–100 lakh hectares a year and poses threat to the livelihood of millions of people.

Machine learning is a part of Artificial Intelligence (AI) which focuses on learning based on the previous data and gives us a better result, is a practical approach which can help us to predict better yield based on various considerations. Machine learning (ML) can identify patterns and learn from datasets. Chlingaryan and Sukkarieh experimented this concept and resulted in measuring the nitrogen contents which shows that other minerals can also be included. Crop rotation [5] is a technique where a course of different types of crops are cultivated over a year's time period, usually a rice crop takes up to 90-110 days which is around 3-4 months of Kharif season (summer season) a year. Whether or not it leads to profit or loss, the other 9 months can be used for other crops in that year such as winter crops like cauliflower, Onion, Green peas, etc. this concept of multiple harvests a year is crop rotation it plays an important role in the present concept.

Eswari 2018 [4] presented that there are various factors that can affect production of crop such as temperature, humidity, rainfalls, precipitation. She has added an extra attribute called crop evapo-transpiration. The crop evapo-transpiration is a combine result of atmosphere and development of the plant growth. In aridest regions of Europe. A method has allowed to predict wheat and barley growth for a few months before crop cultivation. In addition to that, the predictive models such as prediction on crop production by using drought indicators at different time periods have proved that at least 88% and 82% of flexibility for wheat and barley productions, respectively [7].

Soil condition [1] is an important entity that determines agriculture production about less than 8% of land has high organic matter while 60% of agricultural lands have low organic matter and 32% has medium organic matter in Nilgiris in Tamilnadu due to imbalance fertilizers usage this influenced in developing program focusing on soil health.

IV. RESULTS AND DISCUSSIONS

The final product provides solution regarding crop selection based on the mineral content in the soil sample these sample data can be obtained from any convenient government agriculture centre who can provide data from sample. Usually, farmers require that data to condition the field for next plantation. Using that data our program can predict a better solution for different crop that is suitable to that field in its present state to regulate the usage of artificial fertilizers. Not only the circumstances of soil conditions but also the seasonal conditions are also to be taken not of for accurate results. In that case this program has rabi and kharif classification of crops on seasonal basis.

This program has two major features:

- 1) Prediction based on the minerals required
- 2) Prediction based on selected crop's required mineral values to the values the sample of soil actually have.

These features can not only suggest a crop for you but also can help you regulate mineral contents for user's desired crop selection.

V. CONCLUSION

This program in its final stage is a combination of large datasets of minerals and pH values of actually crops collected to be as accurate as possible so that the users aren't supposed to face issue which can cause Hugh loss for an average farmer keeping that in observation there is to be a constant observation on the conditions for the crop for better results and also the nature conditions changing the entities of soil can also cause improper growth in crop to avoid such circumstances constant sample checks to maintain data accurate in order to predict a better solution is necessary. Main cause of this program is to suggest a best option. Based on the suggestions it is up to the user's choice to decide whether to rely on suggestion or to go with their own priority but in either cause this program gives it solution as accurate as possible.

REFERENCES

- 1) Raj Kumar R, Deepthi Kaviraj, Soil Fertility Suggestions in Relation to Crop Productivity Across Various Tea Growing Regions of the Nilgiris, International Journal of Science and Research, Volume 8 Issue 3, March 2019.
- 2) D Ramesh, B Vishnu Vardhan. "Data Mining Techniques and Applications to Agricultural Yield Data". International Journal of Advanced Research in Computer and Communication Engineering Vol.2, Issue 9, September 2013, pp.3477-3480.
- 3) Surendra Kumar Singh, Duraisamy Vasu, Padikkal Chandran, Pramod Tiwary, Soil Quality for Sustainable Agriculture, Nutrient dynamics for sustainable crop production, Chapter 2, pages 41-66.
- 4) L.Vinitha, K.E. Eswari "Crop Yield Prediction in States of Tamil Nadu Using Bayesian Network ", International Journal of Intellectual Advancements and Research in Engineering Computations, Volume-5, Issue-1, January-February 2019.
- 5) Rasulova F.G., Saribaeva N.N., Khalikov B.M. Influence of the Properties of Soil Bulk and Porosity on Repeated Crops in the Vegetable Crop Rotation System, International Journal of Science and Research (IJSR), Volume 9, Issue 4, April 2020.
- 6) M.A.Hearst, S.T. Dumais, E. Osuna, J. Platt, Support Vector Machine(10), IEEE Intelligent System and their application, volume 13, issue-4, July-august 1998.
- 7) Sergio M. Vicente-Serrano, Jose M. Cuadrat-Prats, Alfredo Romo, Sergio M. Vicente-Serrano. Early prediction of crop production using drought indices at different time-scales and remote sensing data: application in the Ebro Valley, International journal of remote sensing, Volume 27, Issue-3, 2006.
- 8) S. Pudumalar, E. Ramanujam, R. Harine Rajashree, C. Kavya, T. Kiruthika, J. Nisha, "Crop Recommendation System for Precision Agriculture", IEEE International Conference on Advanced Computing, Chennai, India, 2017, pp. 32-36.
- 9) Nagarajan, G., and R. I. Minu. "Fuzzy ontology based multi-modal semantic information retrieval." *Procedia Computer Science* 48 (2015): 101-106.
- 10) Nagarajan, G., and R. I. Minu. "Wireless soil monitoring sensor for sprinkler irrigation automation system." *Wireless Personal Communications* 98, no. 2 (2018): 1835-1851.
- 11) Nagarajan, G., R. I. Minu, and A. Jayanthila Devi. "Optimal nonparametric bayesian model-based multimodal BoVW creation using multilayer pLSA." *Circuits, Systems, and Signal Processing* 39, no. 2 (2020): 1123-1132.

- 12) Indra, Minu Rajasekaran, Nagarajan Govindan, Ravi Kumar Divakarla Naga Satya, and Sundarsingh Jebaseelan Somasundram David Thanasingh. "Fuzzy rule based ontology reasoning." *Journal of Ambient Intelligence and Humanized Computing* (2020): 1-7.