Agricultural Crop Recommendation System using Machine Learning: Recommending based on the Environmental Aspects of a Particular land

Thriloke N[1], Vaishnav Kumar G[2], Suriya Sundaram K S[3]
Computer Science and Engineering Department,
Rajalakshmi Engineering College Chennai
210701291@rajalakshmi.edu.in[1]
210701297@rajalakshmi.edu.in[2]
210701272@rajalakshmi.edu.in[3]

ABSTRACT - Agriculture is an important field worldwide, where there are many challenges in solving problems in conditionally estimating crops. Many solutions have been proposed regarding this problem using IOT-based services and mechanical technology to reduce manual work. These methods are mainly useful in the case of minimizing manual labour but not in the prediction process. It is necessary to be able to predict the optimal crop to plant based on the soil condition to minimize losses, harm and maximize profits. We build machine learning models to recommend optimal crops to growers based on many parameters and help them make informed decisions before farming. Dataset prepared with values of nitrogen, phosphorus, potassium and soil pH, temperature and rainfall required for a particular crop. In this system, we will provide records approximately
The specific traits of the soil and the temperature, humidity, autumn conditions of the region, primarily based totally on which we can make appropriate crop prediction. the research finding is 320 in 2024

Keywords: Random Forest, Decision Tree, KNN Neighbours, Flask, Html, CSS.

I. INTRODUCTION

Agriculture is the basic source of food supply of all the countries of the world—whether underdeveloped, developing or even developed. The world population is estimated to be about 9.7 billion by 2025. This added with unpredictable weather conditions makes it difficult to ensure food sustainability. Fortunately there is a solution for this problem as for many others. Crop Recommendation System takes the N-P-K (Nitrogen, Phosphorous and Potassium) and pH values along with the temperature, humidity values as input and recommends the optimal crop to the farmer, hence ensuring that the farmer takes an informed decision before cultivation. In this system, we train the model using Random Forest, Decision Tree and KNN neighbour. We compare the accuracy of this model and choose the best out of it and store that specific model using pickle Module and deploy the machine learning model using Flask. A user can input the various parameters like Nitrogen, Phosphorous, potassium, PH value, Rainfall and Location by interacting with user interface to predict the appropriate crop. The research gap is above 1 year.

II. MATERIALS AND METHODS

Collect a broad data (dataset) of environmental aspects like Nitrogen, Phosphorous and Potassium and pH values along with the temperature, humidity values as input and recommends the optimal crop to the farmer, then the dataset is manipulated and the dataset is spited into trained and testing data. Then In this system, we train the model using Random Forest, Decision Tree and KNN neighbour. We compare the accuracy of this model and choose the best out of it and store that specific model using pickle Module and deploy the machine learning model using Flask and a UI/UX is designed in html, css where the user inputs the data for predicting the crop based on the user's data and the crop is recommended.

Hardware requirements of the project include

- Laptop or Personal Computer

Software requirements include

- Internet browser (Chrome/Edge/Mozilla Firefox)

- Internet connection

III. EXISTING SYSTEM

We aim to create a system that can accurately tell the farmer the suitable crop to be grown based on the input features of the soil and the temperature conditions of the region. In this system, the models will be trained on a textual data set which will be engineered carefully after performing the feature engineering. The user can interact with the model through a website which takes the necessary inputs and loads the trained model. Based on the input data, the model makes a prediction on the optimum crop to be cultivated. The result is then displayed to the user. Instead of directly taking the temperature and humidity values from the user, the website asks the user for their location. By using a weather API, the system automatically retrieves the temperature and humidity values for that region.

Our System has the merits like which is User friendly and Very easy to use as data is submitted using forms and then the Information sharing is easy. disadvantages of existing system efficiency is low, the existing system which recommends crop yield is either hardware-based being costly to maintain, or not easily accessible, despite many solutions that have been recently proposed, there are still open challenges in creating a user-friendly application with respect to crop recommendation, more number of repeated work.

IV. UML DIAGRAMS:

4.1 Architecture Diagram

The architecture diagram is a blueprint showing how a system's parts connect and work together, like a map for a complex building. The crop recommendation system architecture diagram is shown in the below fig 4.1.1.

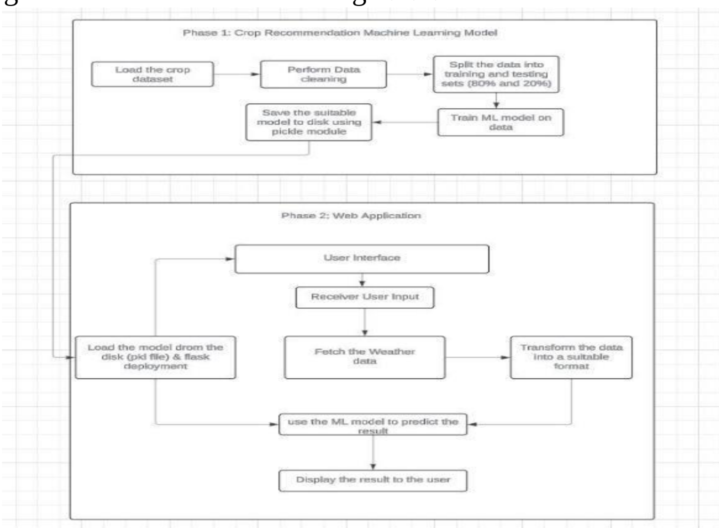

Fig 4.1.1 architecture diagram

4.2 Sequence Diagram

A sequence diagram simply depicts the interaction between the objects in a sequential order. An sequence diagram is used to show the interactive behaviour of a system. The sequence diagram for Agricultural Crop Recommendation System is attached in the below figure 4.2.1.

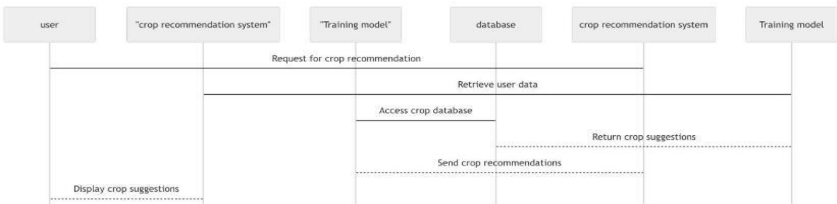

Fig 4.2.1 sequence diagram

V. MATHEMATICAL JUSTIFICATION

To evaluate the performance of the proposed algorithm for predicting the Spending Score, we use three common regression metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared.

i. Mean Squared Error (MSE): MSE measures the average squared difference between the predicted and actual spending scores. MSE can be calculated as in Eq. (1).

$$\mathrm{MSE} = \left(\frac{1}{\mathrm{n}}\right) * \mathrm{sum}\left(\left(\mathrm{y}_{\mathrm{pred}} - \mathrm{y}_{\mathrm{actual}}\right)^2\right)$$

ii. Mean Absolute Error (MAE): MAE measures the average absolute difference between the predicted and actual spending scores. MAE can be calculated as in Eq. (2).

$$\mathrm{MAE} = (1/\mathrm{n}) * \mathrm{sum}\left(\mathrm{abs}\left(\mathrm{y}_{\mathrm{pred}} - \mathrm{y}_{\mathrm{actual}}\right)\right)$$

iii. R-squared: R-squared is a statistical measure representing the proportion of variance in the dependent variable explained by the independent variables. R-squared can be calculated as in Eq. (3).

$$\mathrm{R}^2 = 1 - \left(\frac{\mathrm{sum}\left(\left(\mathrm{y}_{\mathrm{actual}} - \mathrm{y}_{\mathrm{pred}}\right)^2\right)}{\mathrm{sum}\left(\left(\mathrm{y}_{\mathrm{actual}} - \mathrm{y}_{\mathrm{mean}}\right)^2\right)}\right)$$

VI. PROPOSED SYSTEM

6.1. Dataset

The main data used in the data set are initialized with the number to use in the algorithm it is like initializing all the details. In this metadata, we are going to initialize all the crop names with the numbers. This data makes us use the data easily in the algorithm. Hear the metadata of all the crops is given with a particular number. This number is not duplicated that is one number is given to one crop; the same number is not given to the other crop. This metadata consists of more than a hundred crops that grown all over India. The described version of dataset is where it contains attributes like N - ratio of Nitrogen content in soil, P - ratio of Phosphorous content in soil, P - ratio of Potassium content in soil, temperature - temperature in degree Celsius, humidity - relative humidity in %, ph - ph value of the soil, rainfall - rainfall in mm. Based on this the crops 22 unique crops which is cultivated in India, are recommended.

Table 6.1.1 Crops classes dataset

	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	rice

```
'rice': 1,
'maize': 2,
'jute': 3,
'cotton': 4,
'coconut': 5,
'papaya': 6,
'orange': 7,
'apple': 8,
'muskmelon': 9,
'watermelon': 10,
'grapes': 11,
'mango': 12,
'banana': 13,
'pomegranate': 14,
'lentil': 15,
'blackgram': 16,
'mungbean': 17,
'mothbeans': 18,
'pigeonpeas': 19,
'kidneybeans': 20,
'chickpea': 21,
'coffee': 22
```

Table 6.1.2 Crops Classes data

6.2 Training and Testing

In order to solve the classification problem, we must split the data into training and testing datasets. The training data set is the data on which our models will be trained on, and the testing set is the data in which we evaluate the performance of our model using various metrics. We use the sklearn module, which has the necessary implementation of how the split is done and we also import classification report of sklearn.

This will be useful in order to evaluate our model later on. From sklearn.model_selection import train_test_split from sklearn.metrics import classification report, confusion matrix In the above step, we have split the data such that 20% of the dataset is testing data and the remaining 80% of the data is training data. As the split is done in a random order, we set the random state to 2.

VII. LITERATURE SURVEY

[1] **"An Acquisition Based Optimised Crop Recommendation System with Machine Learning Algorithm" by Choudhury, S. S., Pandharbale, P. B., Mohanty, S. N., & Jagadev, A. K.,2024,** The agricultural sector makes a significant economic impact in India. It contributes 19.9% to the national GDP. The prosperity of the country's economy greatly affects the country's progress and the quality of life for Indian citizens. The vast majority of farms still use antiquated methods rather than adopting a data-driven strategy to increase output and earnings. It is considered a cornerstone of India's financial structure. Since achieving independence, increasing output through the implementation of cutting- edge technologies has been a top priority.

[2] “**Intelligent Crop Recommender System for Yield Prediction Using Machine Learning Strategy.,2024” by Maheswary, A., Nagendram, S., Kiran, K. U., Ahammad, S. H., Priya, P. P., Hossain, M. A., & Rashed, A. N. Z.,** For most developed nations, agriculture is a significant economic force. The realm of contemporary agriculture is consistently growing with evolving farming techniques and agricultural innovations. Farmers face challenges in keeping pace with the evolving demands of the planet and meeting the requirements of profitable initiatives, characters, and various other stakeholders.

[3] **“Red fox optimization with ensemble recurrent neural network for crop recommendation and yield prediction model.," by Gopi, P. S. S., and M. Karthikeyan,2024**; Precision agriculture concentrates on monitoring (sensing technologies), management information system, variable rate technologies, and responses to inter- and intravariability in cropping systems. The advantages of precision agriculture involve improving crop productivity and crop quality with minimum environmental impact. Crop yield prediction (CYP) is one of the challenging tasks in agriculture, which mainly depends upon soil, meteorological, environmental, and crop-related variables.

[4] **“Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation.,2023”,by Hasan, M., Marjan, M. A., Uddin, M. P., Afjal, M. I., Kardy, S., Ma, S., & Nam, Y.** the authors Agriculture is the most critical sector for food supply on the earth, and it is also responsible for supplying raw materials for other industrial productions. Currently, the growth in agricultural production is not sufficient to keep up with the growing population, which may result in a food shortfall for the world‟s inhabitants. As a result, increasing food production is crucial for developing nations with limited land and resources. It is essential to select a suitable crop for a specific region to increase its production rate. Effective crop production forecasting in that area based on historical data, including environmental and cultivation areas, and crop production amount, is required.

[5] **"Data-Driven Analysis and Machine Learning-Based Crop and Fertilizer Recommendation System for Revolutionizing Farming Practices,2023" by Musanase, C., Vodacek, A., Hanyurwimfura, D., Uwitonze, A., & Kabandana,** Agriculture plays a key role in global food security. Agriculture is critical to global food security and economic development. Precision farming using machine learning (ML) and the Internet of Things (IoT) is a promising approach to increasing crop productivity and optimizing resource use. This paper presents an integrated crop and fertilizer recommendation system aimed at optimizing agricultural practices in Rwanda. The system is built on two predictive models: a machine learning model for crop recommendations and a rule-based fertilization recommendation model.

VIII. RESULTS AND DISCUSSION

8.1. Confusion Matrix

The proposed model is evaluated and the confusion matrix for the trained model is attached in below figure 8.1.1

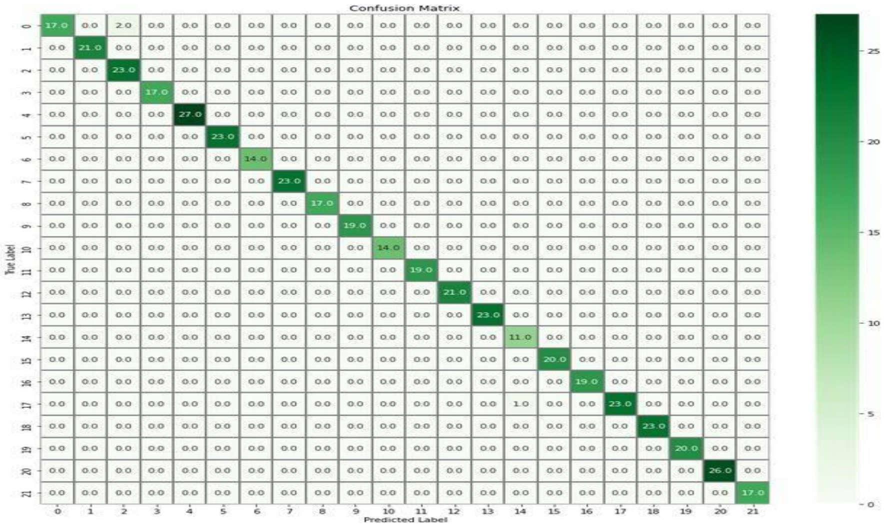

Figure 8.1.1

8.2. Classification Report

The classification report is a detailed performance evaluation of a classification algorithm. In the context of a crop recommendation system, such a report would provide insights into how well the system can predict the appropriate crops for given conditions (like soil type, weather, pH level, etc.). Here's an example of how you can generate and interpret a classification report for a crop recommendation system using common metrics like precision, recall, F1-score, and support. Below figure 8.2.1

	precision	recall	f1-score	support
1	1.00	0.89	0.94	19
2	1.00	1.00	1.00	21
3	0.92	1.00	0.96	23
4	1.00	1.00	1.00	17
5	1.00	1.00	1.00	27
6	1.00	1.00	1.00	23
7	1.00	1.00	1.00	14
8	1.00	1.00	1.00	23
9	1.00	1.00	1.00	17
10	1.00	1.00	1.00	19
11	1.00	1.00	1.00	14
12	1.00	1.00	1.00	19
13	1.00	1.00	1.00	21
14	1.00	1.00	1.00	23
15	0.92	1.00	0.96	11
16	1.00	1.00	1.00	20
17	1.00	1.00	1.00	19
18	1.00	0.96	0.98	24
19	1.00	1.00	1.00	23
20	1.00	1.00	1.00	20
21	1.00	1.00	1.00	26
22	1.00	1.00	1.00	17
accuracy			0.99	440
macro avg	0.99	0.99	0.99	440
weighted avg	0.99	0.99	0.99	440

Figure 8.2.1

IX. CONCLUSION

The Crop Recommendation System is mainly used to recommend the optimal crop to the Farmer. With many industries going digital, it is important that the agricultural sector also take advantage of the various technologies by using them to solve problems faced by the farmers. Using the approach that we discussed, farmers can expect greater yields. This system can be integrated with other smart agricultural systems that already exist. It is cost efficient and helps the farmers make an informed decision. Finding of working numbers is 350 in year 2024.

REFERENCES

[1] Choudhury, S. S., Pandharbale, P. B., Mohanty, S. N., & Jagadev, A. K. (2024). An acquisition based optimised crop recommendation system with machine learning algorithm. *EAI Endorsed Transactions on Scalable Information Systems*, *11*(1).

[2] Hasan, M., Marjan, M. A., Uddin, M. P., Afjal, M. I., Kardy, S., Ma, S., & Nam, Y. (2023). Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation. *Frontiers in Plant Science*, *14*, 1234555.

[3] Kiruthika, S., & Karthika, D. (2024). Crop recommendation system and pest classification using weighted support vector machine on climate data. *Salud, Ciencia y Tecnología-Serie de Conferencias*, *3*, 757-757.

[4] N. N. Li, T. S. Li, Z. S. Yu, Y. Rui, Y. Y. Miao, and Y. S. Li, "Factors influencing farmers' adoption of new technology based on Logistic-ISM model-a case study of potato planting technology in Dingxi City, Gansu Province," Progress in Geography, vol. 33, no. 4, pp. 542-551, 2014.

[5] Rawat, P., Bajaj, M., Vats, S., & Sharma, V. (2023, March). An Analysis of Crop Recommendation Systems Employing Diverse Machine Learning Methodologies. In *2023 International Conference on Device Intelligence, Computing and Communication Technologies,(DICCT)* (pp. 619-624). IEEE.