IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CROP RECOMMENDATION SYSTEM USING MACHINE LEARNING ALGORITHM

Mr. Omkar Kulkarni, Mr. Sushil Buragute, Mr. Ashwin Khode, Mr. Melvin Fernando, Prof. Nilesh Korade,

Department of Computer Engineering,
Pimpri Chinchwad College of Engineering and Research,
Ravet, Pune

Abstract: India is land of farmers. A major role in India's economy is played by farming. As much as 60% of the land is used for farming in India. This feeds a total of 1.2 billion in population. In general, agriculture is the backbone of India and conjointly plays a crucial role in the Indian economy by providing a definite share of domestic products to ensure food security. However, nowadays, food production and prediction are depleted because of unnatural climatic changes, which is able to adversely affect the economy of farmers by obtaining a poor yield and conjointly facilitate the farmers to stay less acquainted in prognostication in the long run. Using appropriate parameters like rain patterns, temperature patterns, soil structures, and other factors such as crop diseases makes it possible to yield accurate crop prediction results. We make use of Random Forest and other machine learning/artificial intelligence algorithms to generate an accurate model. With the help of our system, we will be able to recommend the best possible crop for the given conditions. This system would also help reduce the financial loss incurred by farmers caused by planting the incorrect crop.

Index Terms **– Crop recommendation system, machine learning, crop prediction**

I. INTRODUCTION

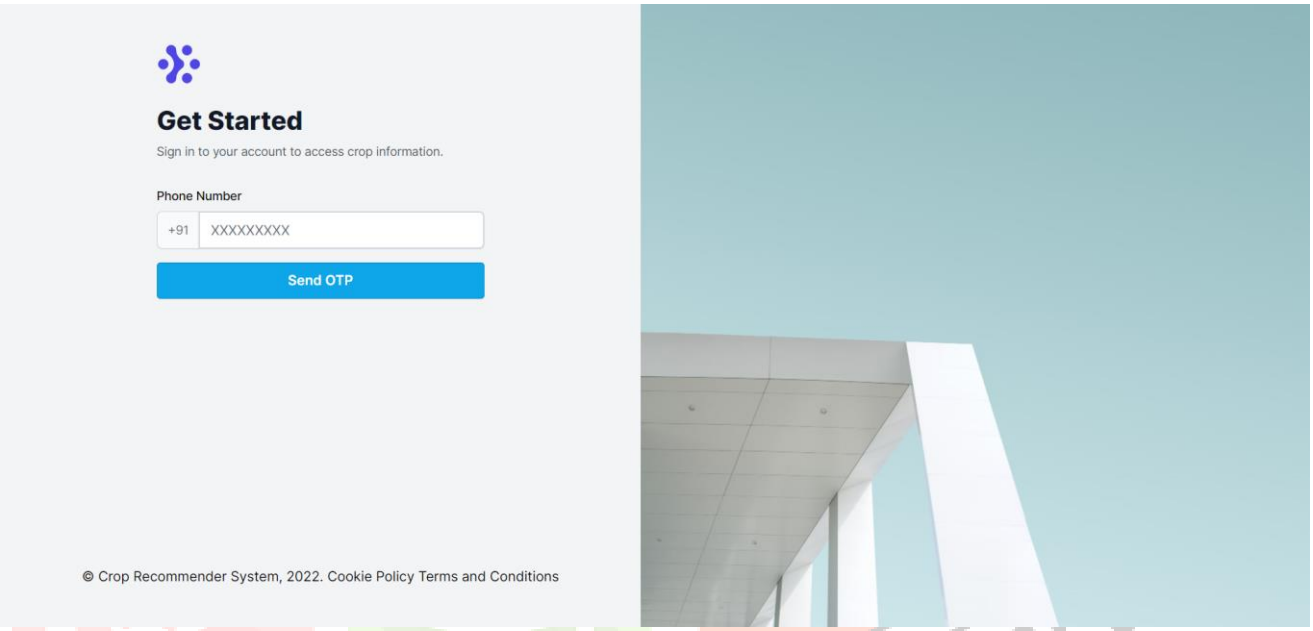
A major contribution to India's economy is made by Agriculture. A lot of families in India directly or indirectly depend on farming in general. Hence, we can say farming is an integral part of India and define who we are as a country. A larger part of the Indian farming population trusts in relying upon their instinct to conclude which yield to plant in a specific season. They usually take historical parameters and ancestral farming patterns into consideration without knowing that crop depends on weather, present-day, and soil conditions. Thus, it is difficult in a sense to blame a farmer for all of their decisions because of innumerable factors that are involved. A wrong or misguided decision by a farmer can have a lot of impact on the agricultural economy of that region and India in a broader sense. A combination of machine learning algorithms, historical and scientifically collected soil parameters, and weather data can help predict the most appropriate crop to be sown on the farm.

II. PROBLEM STATEMENT

There are very few platforms that help farmers with their farming strategy. Intuition-based decisions may not prove beneficial in the long run. Farmers often underestimate/overestimate the fertility of the soil on their farms. They often find it difficult to notice plant diseases that directly affect the production rate. Using appropriate parameters like rain patterns, temperature patterns, soil structures, and other factors such as crop diseases makes it possible to yield accurate crop prediction results. Not only that, but it is also possible to identify what disease a crop has beforehand. A lot of existing systems have many flaws and make them non-intuitive to use or are very difficult.

III. PROPOSED DESIGN

- The system consists of a web portal used to get farm soil details and show results.
- The website will be powered by OTP (One Time Password) based end-to-end authentication system.
- The portal has a location picker that can be used to detect the exact location of the farming land. It also has a form that the user needs to fill out. It requires them to input basic information about the soil such as Nitrogen, Phosphorous, and Potassium contents of the soil.
- The website also has a feature that allows users to upload images of a diseased leaf to identify the disease and possible preventive measures.
- After the user submits details, the system predicts a suitable crop, and users are redirected to the page showing the crop's details



IV. APPROACH:

A. Dataset Gathering:

In this step, all the required dataset is gathered. This mainly constitutes the past 20 years of soil, and crop dataset of India. This data consists of over 2000 observations and over 20 different crops. This is achieved through Python Jupyter Notebook on Google Colab. This data is a fundamental resource to train machine learning models. Some part of the data is used for training the model and some part of it is used for testing purposes. Soil is essentially the anchor of crops. Its texture and nutrient values like NPK determine the fertility of the soil. This data also constitutes the weather pattern history of India.

B. Pre-processing:

Pre-processing involves the elimination of disturbances and outliers present in the CSV dataset. A lot of times data can go missing that directly affects the efficiency and accuracy of the final machine learning model. This needs to be addressed through various techniques such as median, and mean of the whole column. Using SKLearn, we can conveniently clean the dataset. It provides an imputer class that addresses and fixes missing values. The imputer class accepts parameters like missing values and strategies that the imputer internally uses. This strategy involves replacing missing values using the mean on-axis.

C. Feature Selection:

Not all columns contribute majorly towards favorable prediction. Feature selection is a specific process in the pipeline where such features are selected which contribute most towards the prediction variable. Feature selection reduces the overfitting of data. An optimized dataset means a machine learning model makes decisions against a low amount of noise. It directly increases the accuracy of the model.

D. Choosing Machine Learning Model:

When choosing a machine learning algorithm, Random Forest is one of the most popular and widely accepted supervised learning techniques. It consists of a number of decision trees for different subsets of data, rather than using the whole data as a single unit. This helps maximize the accuracy of the prediction of each variable. It makes sense to use Random Forest on a large dataset as it can provide results with maximum accuracy in a short span of time.

E. Application of Random Forest Algorithm:

We chose to use Google Colab for executing python code that trained our machine learning model. A lot of crucial libraries such as pandas, Sklearn were imported to perform a proper analysis of the dataset. Using pandas, a data frame was created that helped read the CSV file. The dataset was split into a specific ratio to train and test respectively. We performed training of the model using random forest and tested its output with testing data. We found the accuracy to be 99%.

```
RF's Accuracy is:  0.990909090909091
              precision    recall  f1-score   support

       apple       1.00      1.00      1.00        13
      banana       1.00      1.00      1.00        17
   blackgram       0.94      1.00      0.97        16
    chickpea       1.00      1.00      1.00        21
     coconut       1.00      1.00      1.00        21
      coffee       1.00      1.00      1.00        22
      cotton       1.00      1.00      1.00        20
      grapes       1.00      1.00      1.00        18
        jute       0.90      1.00      0.95        28
 kidneybeans       1.00      1.00      1.00        14
      lentil       1.00      1.00      1.00        23
       maize       1.00      1.00      1.00        21
       mango       1.00      1.00      1.00        26
   mothbeans       1.00      0.95      0.97        19
    mungbean       1.00      1.00      1.00        24
   muskmelon       1.00      1.00      1.00        23
      orange       1.00      1.00      1.00        29
      papaya       1.00      1.00      1.00        19
  pigeonpeas       1.00      1.00      1.00        18
 pomegranate       1.00      1.00      1.00        17
        rice       1.00      0.81      0.90        16
  watermelon       1.00      1.00      1.00        15

    accuracy                           0.99       440
   macro avg       0.99      0.99      0.99       440
weighted avg       0.99      0.99      0.99       440
```

Fig: Model's accuracy using Random Forest Algorithm

F. Serving Machine Learning Prediction as a Services through Web App:

We leveraged modern web infrastructures such as Python Flask API, Node.JS, Next.JS, MDX, PostgreSQL, Prisma 2 ORM, and Google Firebase PaaS to serve an online web app that directly enables users to submit soil information along with required location data. We made use of the OpenWeather API to accurately forecast and send rainfall data and pH information. The web app is written in JavaScript and Python Language. It is end-to-end protected through OTP-based authentication.

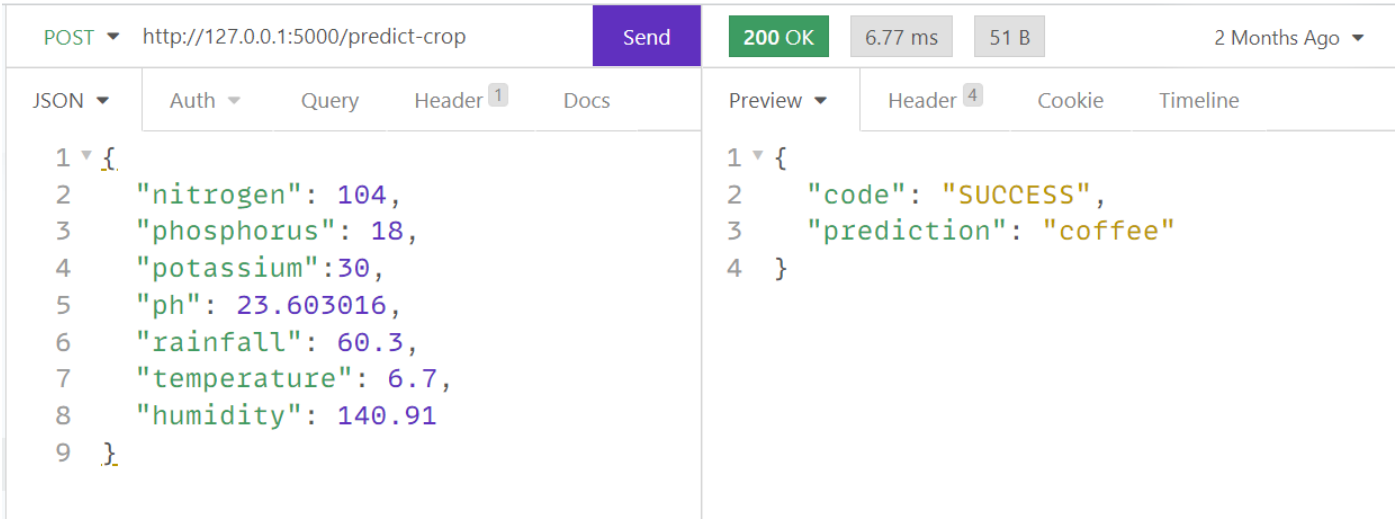


Fig: API Services returns predicted crop result in a JSON format

V. CONCLUSION:

The proposed crop recommendation system is a great helper for farmers in India. The aim is to make farmers aware of modern tools and infrastructure and promote precision farming. A well-informed decision can directly affect their profits. Such a system is especially useful for newbie farmers who lack knowledge about soil and crop specifics. With more research going into this system, it can expand to provide more accurate recommendations. This system is secure and does not involve any specific hardware to get started with. With machine learning, such a system can keep evolving to support a wide variety of crops, and identify a lot of weather patterns. The GUI provides an intuitive dashboard, easy-to-understand navigation, and availability in more Indian native languages.

VI. REFERENCES:

[1] Dhruv Piyush Parikh, Jugal Jain, Tanishq Gupta and Rishit Hemant Dabhade, "Machine Learning Based Crop Recommendation System" in IJARSCT 2021.

[2] D.Jayanarayana Reddy, Dr M. Rudra Kumar, "Crop Yield Prediction Using Machine Learning Algorithm" in 2021 Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021).

[3] Annapoorna. S, Apoorva Herle, C M Sushma, Neha.Y.Jain, Mr.Vijay Kumar.S, "CRAPCSS-Crop Recommendation and Pest Control Suggestion System", IRJECT 2020.