International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Crop Recommendation and Disease Prediction System

Sarvesh Lele[1], Nihal More[2], Raj Malekar[3], Rohit Londhe[4]

Prof Kiran Thoke (BE, ME, PhD (Pursuing)) [5]

[1-4]Dept. of Computer Engineering, KJCOEMR, Pune

[5]Professor, KJ College of Engineering and Management Research

---***---

Abstract - India is the seventh largest country in the world speaking of land area, covering a total of 3.288 million square kilometres. The agricultural exports from India for the FY 2022 is $50.21B. The government has launched various initiatives to help the farmers and the agricultural sector. With the advent of technology and the availability of large-scale agricultural data, there is an opportunity to leverage machine learning and data analytics techniques to promote better agricultural practices. With this project we aim to leverage the offerings of Machine Learning to help the agriculture sector in India, and help the farmers choose a crop to grow based on the weather conditions. We also aim to predict and determine whether a crop has caught a disease or illness and if yes, which disease, with the help of Machine Learning and Deep Learning algorithms.

Key Words - Crop Recommendation, Crop Disease Detection, Farmers, Machine Learning, Deep Learning, Agriculture

1. INTRODUCTION

Agriculture forms a major chunk of the Indian economy. A majority of the population depends on Agriculture for their livelihood. India is a leading producer of agriculture and food products in the world. India produces many crops and grains such as Wheat, Rice, Pulses, Coffee, Oilseeds, Jute, Tea, Sugarcane, Groundnut, Fruits, etc. India is also one of the largest agricultural products exporter in the world. With a 20% increase from USD 41.3 billion in 2020-21, the country recorded USD 49.6 billion in total agriculture exports in 2021-22. With the aforementioned statistics, it is necessary for the farmers and agricultural sector to be able to keep up and utilize the advancements in the technologies, to better be able to produce high quality yield and strengthen their own as well as the country's economy. With farmers being able to leverage the technological advancements to their advantage, it can help them plan the crop planting and agricultural activities better, helping them utilize their available land to the maximum. The farmers must be able to decide on a crop suitable for weather conditions in their region, and also the sand nutrient contents. With the various schemes launched by the Government bodies to help the farmers produce higher quality yield, it is also necessary for them to be able to use technology to its fullest potential. Machine Learning technology plays a major role in helping the agriculture sector using technology. With this project we have tried to leverage Machine Learning to help farmers and people in the agricultural field with narrowing down on the crop they can grow. We aim to be able to provide a very accurate recommendation for the crop to grow, based on the weather and soil data. We also aim to predict whether or not a crop has a disease, or in other words, is infected by a disease based on images of the leaves of the data. We have trained models to achieve the same, and have developed a front end using which the services can be offered in an interactive way to the users.

2. LITERATURE SURVEY

In this paper [1], the authors have emphasized on the usage of Decision Tree algorithms in various fields to solve real world problems, including in Image Processing, Medicine, Intelligent Vehicles, Web Applications and Remote Sensing.

In this paper [2], the authors have implemented recommendations for crops to grow in the fields using Naïve Bayes Classifier, Decision Trees, and K Nearest Neighbor Algorithms. The authors have found that KNN is superior to others for this specific application in terms of precision and specificity.

In this paper [3], the authors have compared various Machine Learning Models for predicting crops. The results showed that Gaussian Naive Bayes Classifier worked the best with the dataset used, achieving over 60% accuracy.

This paper [4] studies detection of diseases in plants using images. The authors have used Clustering to segment the images. For classification of diseases, a Support Vector Machine is used.

In Reference [5], various learners are used viz. Random Tree, CHAID, K Nearest Neighbours and Naive Bayes. If-then rules are induced from the Random tree model and CHAID, which are used to develop recommendation systems.

In this paper [6], the proposed system is divided into three parts: Crop Recommendation, Fertilizer Recommendation and Disease Detection. This system provides the best fit of the crop to grow as per the conditions, and also the top 5 possible crop options. Furthermore, the cause of a disease and preventive measures are also provided by the model.

Reference [7] employs similar techniques for prediction of crops. A cloud-based ML-driven crop recommendation platform is developed, with the intention of encouraging researchers who are engaged in this area towards the invention of novel solutions for revolutionizing agriculture.

In Reference [8], the author has used Random Forest and K Nearest Neighbours algorithms on a dataset of 2200+ records, achieving an accuracy of 0.995 for Random Forest algorithm.

Reference [9] proposes using GRU model. The proposed GRU model is shown to give the greatest accuracy and precision when compared to LSTM and CNN, among other deep learning algorithms.

In Reference [10], a dataset of 61000+ images was used to train a CNN model to detect diseases in plants from leaf images. A validation accuracy of 98.7 was achieved.

Reference [11] India Brand Equity Foundation (IBEF) , a Trust established by the Department of Commerce, Ministry of Commerce and Industry, Government of India. IBEF is a knowledge centre for global investors, international policy-makers and world media seeking updated, accurate and comprehensive information on the Indian economy, states and sectors.

3. PROPOSED SYSTEM

The proposed system is as the following diagram depicts. The diagram depicts the flow of using Machine Learning to recommend a crop to farmers based on the various nutrient contents of the soil.

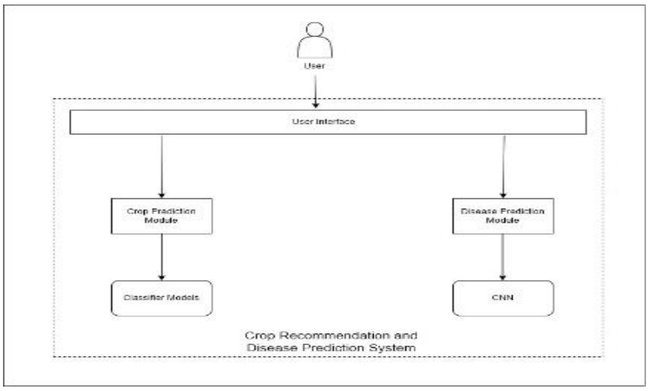


Fig -1. Proposed System

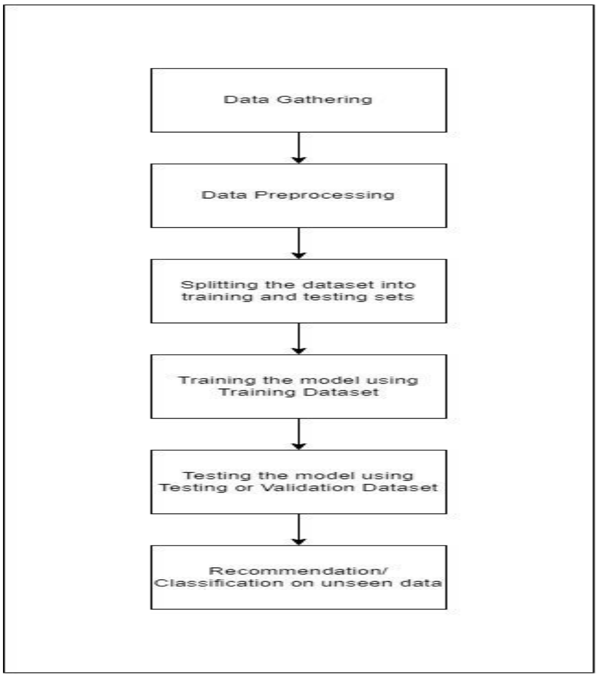


Fig -2. System Flow of Machine Learning model

For recommending a crop, we have used the dataset "Crop Recommendation Dataset" available on Kaggle. The dataset has a total of 7 features. They are:

- *N - ratio of content of Nitrogen in soil*
- *P - ratio of content of Phosphorus in soil*
- *K - ratio of content of Potassium in soil*
- *Temperature - temperature in degree Celsius*
- *Humidity - relative humidity in %*
- *ph – value of ph of the soil*
- *Rainfall - rainfall in mm*

The above dataset has a total of 2200 rows of data, with unique crops being Rice, Maize, Chickpea, Kidney beans, Pigeon peas, Moth beans, Mung bean, Black gram, Lentil, Pomegranate, Banana, Mango, Grapes, Watermelon, Muskmelon, Apple, Orange, Papaya, Coconut, Cotton, Jute, and Coffee, a total of 22 crops. The model can predict any of the above crop based on unseen data.

For Disease Detection, we have used the dataset "New Plant Diseases Dataset" which is available on Kaggle. The dataset has a collection of about 87000 RGB images, which are a collection of healthy as well as diseased crop leaves categorized in 38 distinct classes. The dataset is divided in the ratio of 80/20 for Training of Model and testing the same.

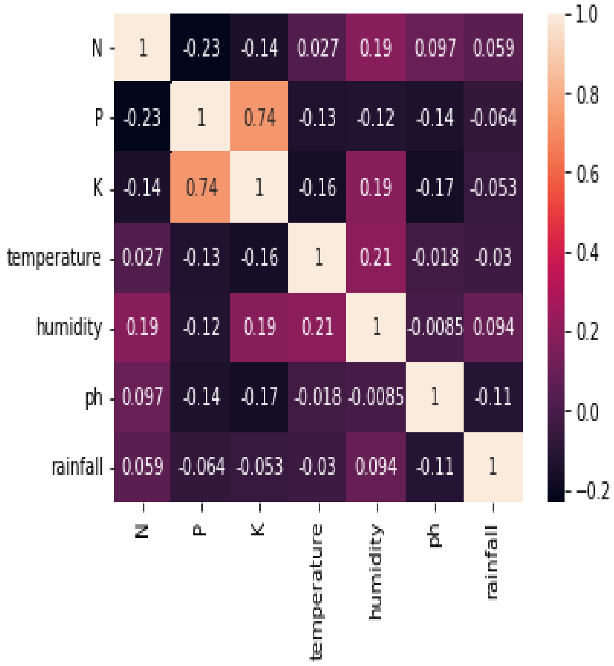
The basic steps followed to train Machine Learning and Deep Learning algorithms to achieve the above objectives are highlighted below:

A. Data Collection*:*

We gather the data from the aforementioned dataset. We use the Python library "Pandas" to access and work on the dataset, in a dataframe.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

B. Data Pre-processing*:*



After collection of the said data, we work on the data to make it fit for use. We pre-process it, in which we get rid of outliers, handle missing values, etc.

C. Splitting Dataset:

The complete dataset is divided into different parts viz, Training and Testing in 80/20 ratio. The training dataset is used to train the model to make recommendations of the crop. The testing part is used once the model is built, for validating and measuring accuracy of the model.

D. Training the Model:

After splitting the dataset into training and testing datasets, we use the training dataset to train the model. The model learns from the data in the training dataset in this phase.

E. Testing the Model:

Once the model is trained, we use the testing dataset to validate the model's output and accuracy. The model does predictions on data it has not yet seen, which belongs to the testing or validation dataset in this phase.

F. Recommendation on unseen data:

Once the model is tested using the Testing dataset, the model is ready to make recommendations on data that is not previously known to the model, or in other words, new data, which is collected in real time.

4. OUTCOME

The data heatmap for Crop Recommendation Dataset is as follows:

Fig -3. Heatmap for Crop Recommendation dataset

After training the models to recommend a crop, following is the outcome:

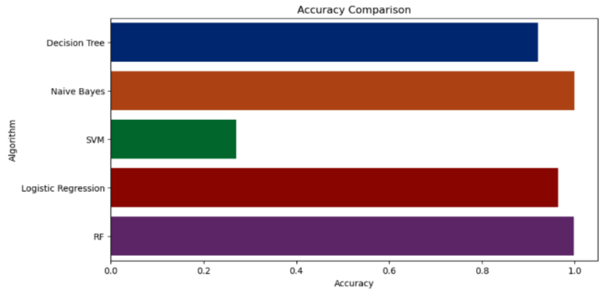


Fig -4. Accuracy Comparison for Crop Recommendation

```
#Testing with a prediction

data = np.array([[104,18, 30, 23.603016, 60.3, 6.7, 140.91]])
prediction = NaiveBayes.predict(data)
print(prediction)
```

```
['coffee']
```

Fig -5. Prediction for Crop Recommendation

```
[45]:
# getting all predictions (actual label vs predicted)
for i, (img, label) in enumerate(test):
    print('Label:', test_images[i], ', Predicted:', predict_image(img, model))
```

```
Label: AppleCedarRust1.JPG , Predicted: Apple___Cedar_apple_rust
Label: AppleCedarRust2.JPG , Predicted: Apple___Cedar_apple_rust
Label: AppleCedarRust3.JPG , Predicted: Apple___Cedar_apple_rust
Label: AppleCedarRust4.JPG , Predicted: Apple___Cedar_apple_rust
Label: AppleScab1.JPG , Predicted: Apple___Apple_scab
Label: AppleScab2.JPG , Predicted: Apple___Apple_scab
Label: AppleScab3.JPG , Predicted: Apple___Apple_scab
Label: CornCommonRust1.JPG , Predicted: Corn_(maize)___Common_rust_
Label: CornCommonRust2.JPG , Predicted: Corn_(maize)___Common_rust_
Label: CornCommonRust3.JPG , Predicted: Corn_(maize)___Common_rust_
Label: PotatoEarlyBlight1.JPG , Predicted: Potato___Early_blight
Label: PotatoEarlyBlight2.JPG , Predicted: Potato___Early_blight
```

Fig -6. Predicting diseases for Disease Prediction

As we can see from the Accuracy comparison, the Naïve Bayes and Random Forest models give the highest accuracy for Recommending a crop based on data. The accuracy is done on the Testing, or Validation dataset which we divided earlier.

Also as we can see from the Disease Prediction output, we are able to accurately predict and detect disease in a variety of crops based on leaf images.

We use the unified browser GUI to input the data and image to the model and network, and observe the output given by the model and network.

Should an error occur on the server-side, the GUI alerts the user about the same.

5. CONCLUSION AND FUTURE SCOPE

In conclusion, our research project has successfully developed a crop recommendation and disease detection system that utilizes soil data and image analysis techniques based on Convolutional Neural Networks (CNN). The system aims to assist farmers and agricultural practitioners in making informed decisions to optimize crop selection and manage crop health effectively.

Through the utilization of soil data, our system analyses various soil parameters such as pH levels, nutrient composition, and moisture content to generate accurate crop recommendations. By considering the specific requirements of different crops, the system suggests suitable crop options that align with the soil characteristics, maximizing productivity and yield potential.

Furthermore, the integration of CNN-based image analysis techniques enables the system to detect diseases in crops. By analysing images of plants, the CNN model identifies visual patterns and anomalies associated with common crop diseases. This allows for early disease detection, enabling prompt intervention and appropriate treatment measures to minimize crop losses and ensure better overall crop health.

The combination of soil data analysis and crop disease detection in a unified system provides comprehensive support to farmers and agricultural stakeholders. It facilitates data-driven decision-making, improves crop management practices, and enhances the efficiency and sustainability of agricultural operations.

Our project contributes to the advancement of precision agriculture and agricultural intelligence by harnessing the power of machine learning and image analysis. The integration of these technologies holds significant potential for revolutionizing farming practices, improving crop yields, and reducing environmental impact.

To further enhance our proposed system, we can consider the integration of reliable APIs to fetch weather data. This will reduce the manual entry for weather details in the system to recommend crops.

In addition, we can also analyse existing data for enabling yield prediction for the suggested crops. We can also offer various suggestions with recommendation for crops. These can include suggestions for irrigation scheduling, fertilizer application, crop rotation, and pest control strategies based on the specific crop, soil conditions, and disease prevalence.

6. ACKNOWLEDGMENT

It gives us immense pleasure to present a paper on the successful completion of our project and paper work on “Crop Recommendation and Disease Prediction System”. We express our deep sense of gratitude to our guide Prof. K. K. Thoke for his valuable guidance rendered in all phases of the project and paper. We are thankful for his wholehearted assistance, advice and expert guidance towards making our project and paper a success. Our special thanks to the respected Principal and Head of Department for their keen interest, encouragement and excellent support.

7. REFERENCES

[1] Himani Sharma, Sunil Kumar “A Survey on Decision Tree Algorithms of Classification in Data Mining” International Journal of Science and Research (IJSR) · April 2016

[2] Pavan Patil, Virendra Panpatil, Prof. Shrikant Kokate “Crop Prediction System using Machine Learning Algorithms” International Research Journal of Engineering and Technology (IRJET) - February 2020

[3] Prof. D.S. Zingade ,Omkar Buchade ,Nilesh Mehta ,Shubham “Crop Prediction using Machine Learning” International Research Journal of Engineering and Technology (IRJET) - June 2020

[4] R. Mounika, Dr. P.Shayamala Bharathi “Detection Of Plant Leaf Diseases Using Image Processing” Journal of Critical Reviews – June 2019

[5] S. Pudumalar, E. Ramanujam, R. Harine Rajashreen, C. Kavyan, T. Kiruthikan, J. Nishan “Crop Recommendation System for Precision Agriculture” IEEE Eighth International Conference on Advanced Computing (ICoAC) - 2016

[6] Rohini Jadhav, Dr. Pawan Bhaladhare “A Machine Learning Based Crop Recommendation System: A Survey” Publishoa – 2022

[7] Navod Neranjan Thilakarathne, Muhammad Saifullah Abu Bakar and Hayati Yassin, Pg Emerolylariffion Abas “A Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming” Sensors – 2022

[8] Kanika Bhatnagar , Mamilla Jaahnavi , Dr. B. U. Anu Barathi “Agriculture Crop Recommendation System using Machine-Learning” Mathematical Statistician and Engineering Applications – 2022

[9] E. Manjula, S. Djodiltachoumy “Efficient prediction of recommended crop variety through soil nutrients using deep learning algorithm” Journal of Postharvest Technology – 2022

[10] Goutami G Manvi, Gayana K N, G Ramya Sree, K Divyanjali, Dr Kirankumari Patil “Plant Disease Detection” Academia – 2022

[11]https://www.ibef.org/industry/agriculture-india/infographic