CROP RECOMMENDATION SYSTEM USING ML ALGORITHM

Mr. Goutam Kamble[1], Vaishnavi Pravin Gondkar[2], Prajakta Anil Pasare[3], Vaishnavi Umesh Bhoite[4]

[1] *Lecturer, Computer Science and Engineering, Sharad Institute of Technology College of Engineering, Yadrav, Maharashtra, India*
[2] *Student, Computer Science and Engineering, Sharad Institute of Technology College of Engineering, Yadrav, Maharashtra, India*
[3] *Student, Computer Science and Engineering, Sharad Institute of Technology College of Engineering, Yadrav, Maharashtra, India*
[4] *Student, Computer Science and Engineering, Sharad Institute of Technology College of Engineering, Yadrav, Maharashtra, India*

ABSTRACT

Agriculture is the backbone of the global economy and plays a crucial role in ensuring food security. However, the choice of crops to cultivate is a complex decision that depends on various factors, including soil quality, climate, and market demand. Traditional crop selection methods often rely on anecdotal knowledge or outdated practices, leading to suboptimal yields and environmental concerns. To address these challenges, we propose a novel Crop Recommendation System (CRS) that leverages advanced Machine Learning (ML) algorithms to provide personalized and data-driven crop recommendations.

Keyword: - *Machine Learning Algorithms, Crop Yield Prediction, Crop Recommendation System, etc.*

1. INTRODUCTION

Agriculture plays a vital role in feeding the global population and ensuring food security. However, the agricultural sector faces numerous challenges, one of the most critical being the selection of the right crops for a given region. Crop selection involves considering a multitude of factors, including soil quality, climate conditions, and market demand. Traditionally, farmers have relied on anecdotal knowledge and historical practices to make these decisions, often resulting in suboptimal yields, resource inefficiencies, and environmental concerns. Today, we are on the cusp of a significant change in agriculture, thanks to Machine Learning (ML) algorithms. These algorithms have the power to analyze vast amounts of data and extract valuable insights. It's within this context that the Crop Recommendation System (CRS) emerges. The Crop Recommendation System is a modern solution that combines the best of human wisdom with the analytical prowess of machines. It processes a range of data, such as soil composition, historical weather patterns, past crop performance, and market trends. Using this data, the CRS generates personalized crop recommendations.

2. LITERATURE SURVEY

Dhruv Piyush Parikh, Jugal Jain, Tanishq Gupta and Rishit Hemant Dabhade, "Machine Learning Based Crop Recommendation System" [1], The three most basic amenities required for the survival of a human being are food, shelter and clothing. In today's tech-savvy generation, the latter two have witnessed a huge scientific boost. Unfortunately, even today, agriculture is considered as more of a man-power oriented field. Most of the farmers are untutored and have little to no scientific knowledge of farming. So, they have to rely on the hit and trial method to learn from experience which leads to wastage of time and resources. Our system focuses on building a

predictive model to recommend the most suitable crops to grow in a particular farm based on various parameters. This can be helpful for the farmers to be more productive and competent without wasting any resources by farming the most competent crops.

G. Buvaanyaa, Dr. S. Radhimeenakshi, "Crop Recommendation System Using Random Forest Algorithm" [3], India is the place where there is agribusiness and it is the significant wellspring of economy 70% of the Indian populace straightforwardly depends on farming. The regular issue existing among the youthful Indian ranchers is to pick the correct yield dependent on the dirt prerequisites. Because of this, they face a genuine difficulty in efficiency. Arising advancements can be utilized to further develop the efficiency of the harvests by changing conventional cultivating completely to accurate cultivating. The significant issue yet to be settled is developing the exact harvest at the exact time. This should be possible with the assistance of machining learning calculations which is viewed as a powerful strategy for anticipating a reasonable harvest. The crop recommendation parameters such as NPK, temperature, humidity, rainfall, and pH are collected from the benchmark repository. The framework created utilizing Machine Learning greatly assists the farmers to make valuable decisions.

Thomas van Klompenburga, Ayalew Kassahuna, Cagatay Catalb, "Crop yield prediction using machine learning: A systematic literature review" [2], Machine learning is an important decision support tool for crop yield prediction, including supporting decisions on what crops to grow and what to do during the growing season of the crops. Several machine learning algorithms have been applied to support crop yield prediction research. In this study, we performed a Systematic Literature Review (SLR) to extract and synthesize the algorithms and features that have been used in crop yield prediction studies. Based on our search criteria, we retrieved 567 relevant studies from six electronic databases, of which we have selected 50 studies for further analysis using inclusion and exclusion criteria. We investigated these selected studies carefully, analyzed the methods and features used, and provided suggestions for further research. According to our analysis, the most used features are temperature, rainfall, and soil type, and the most applied algorithm is Artificial Neural Networks in these models. After this observation based on the analysis of machine learning-based 50 papers, we performed an additional search in electronic databases to identify deep learning-based studies, reached 30 deep learning-based papers, and extracted the applied deep learning algorithms. According to this additional analysis, Convolutional Neural Networks (CNN) is the most widely used deep learning algorithm in these studies, and the other widely used deep learning algorithms are Long-Short Term Memory (LSTM) and Deep Neural Networks (DNN).

2. PROBLEM STATEMENT

Traditional farming lacks data-driven precision, hindering optimal crop choices. The absence of adaptive solutions and limited access to technology leave farmers vulnerable to environmental risks. The need for a Crop Recommendation System using Machine Learning algorithms is evident, aiming to enhance precision, adaptability, and accessibility for informed crop selection and resource optimization. To develop a system to decide a crop selection for crop growing based on random forest and logistic regression algorithm.

3. METHODOLOGY

3.1 Data Collection and Preprocessing:

Gather and clean data from various sources, including soil quality, weather, and crop performance. Clean and preprocess the data, handling missing values and outliers, and ensuring that it's in a format suitable for ML algorithms.

3.2 Dataset Details:

A sample dataset is collected from Kaggle. The dataset considered for usage in the given proposed work is a crop recommendation dataset primarily comprising soil properties, along with the temperature, humidity, rainfall, pH value, N, P, and K details. In this dataset, there are different types of attributes such as:

- NPK of the soil
- pH value
- Humidity
- Rainfall
- Temperature

3.3 Algorithm Selection:

Choose ML algorithms suitable for the CRS. Common choices include collaborative filtering, regression models, decision trees, random forests, support vector machines, and neural networks. The selection may depend on the specific requirements and nature of the dataset. Algorithms are used in this project such as:

- **Random Forest Regression:**

It generates multiple decision trees from which each decision tree uses a part of the data sample and predicts the result. Then, the result which was achieved by a maximum number of trees is considered as the final prediction. This is a supervised learning algorithm that uses an ensemble learning method for classification and regression. It is a bagging technique and the trees in random forests run in parallel without any interactions.

- **Logistic Regression:**

Logistic regression is typically used for binary classification tasks, where the goal is to predict one of two classes or outcomes (e.g., Yes/No, True/False). In the context of a crop recommendation system, where the goal is to recommend the best crop for a specific scenario, logistic regression may not be the most suitable algorithm. Crop recommendation is more of a multiclass classification or regression problem, as you are predicting from a range of possible crop options. Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables

3.4 Model Development:

Build and train the ML models using the training data. Depending on the algorithm, this step may involve iterative training, parameter tuning, and optimization.

3.5 Testing and Validation:

Evaluate the performance of the model on the test dataset to assess its real-world predictive capabilities. Ensure that the CRS delivers accurate and useful recommendations.

4. UML DIAGRAMS

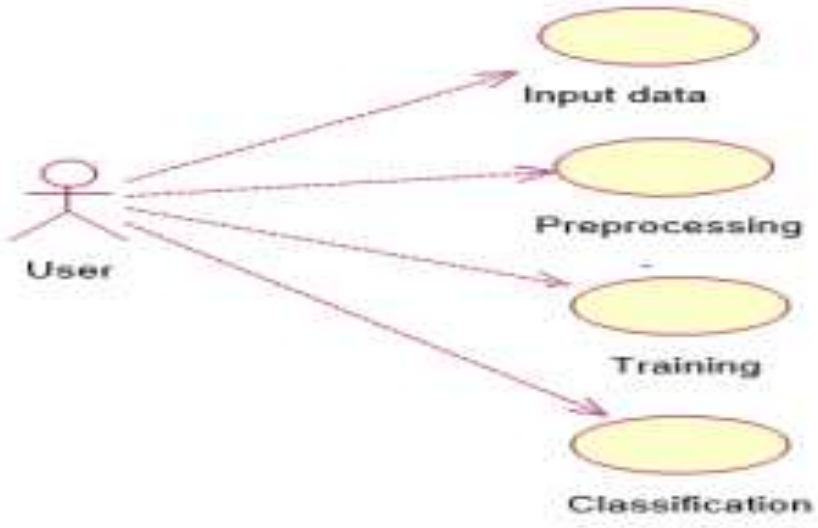

Fig -1 Use case Diagram

5. IMPLEMENTATION

	A	B	C	D	E	F	G	H
1	N	P	K	temperatu	humidity	ph	rainfall	label
2	90	42	43	20.87974	82.00274	6.502985	202.9355	rice
3	85	58	41	21.77046	80.31964	7.038096	226.6555	rice
4	60	55	44	23.00446	82.32076	7.840207	263.9642	rice
5	74	35	40	26.4911	80.15836	6.980401	242.864	rice
6	78	42	42	20.13017	81.60487	7.628473	262.7173	rice
7	69	37	42	23.05805	83.37012	7.073454	251.055	rice
8	69	55	38	22.70884	82.63941	5.700806	271.3249	rice
9	94	53	40	20.27774	82.89409	5.718627	241.9742	rice
10	89	54	38	24.51588	83.53522	6.685346	230.4462	rice
11	68	58	38	23.22397	83.03323	6.336254	221.2092	rice
12	91	53	40	26.52724	81.41754	5.386168	264.6149	rice
13	90	46	42	23.97898	81.45062	7.502834	250.0832	rice
14	78	58	44	26.8008	80.88685	5.108682	284.4365	rice
15	93	56	36	24.01498	82.05687	6.984354	185.2773	rice
16	94	50	37	25.66585	80.66385	6.94802	209.587	rice
17	60	48	39	24.28209	80.30026	7.042299	231.0863	rice
18	85	38	41	21.58712	82.78837	6.249051	276.6552	rice
19	91	35	39	23.79392	80.41818	6.97086	206.2612	rice
20	77	38	36	21.86525	80.1923	5.953933	224.555	rice
21	88	35	40	23.57944	83.5876	5.853932	291.2987	rice
22	89	45	36	21.32504	80.47476	6.442475	185.4975	rice
23	76	40	43	25.15746	83.11713	5.070176	231.3843	rice
24	67	59	41	21.94767	80.97384	6.012633	213.3561	rice
25	83	41	43	21.05254	82.6784	6.254028	233.1076	rice

Fig -1 Dataset

```
In [220]: data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2200 entries, 0 to 2199
Data columns (total 8 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   N            2200 non-null   int64
 1   P            2200 non-null   int64
 2   K            2200 non-null   int64
 3   temperature  2200 non-null   float64
 4   humidity     2200 non-null   float64
 5   ph           2200 non-null   float64
 6   rainfall     2200 non-null   float64
 7   label        2200 non-null   object
dtypes: float64(4), int64(3), object(1)
memory usage: 137.6+ KB
```

Fig - 2 Data Preprocessing

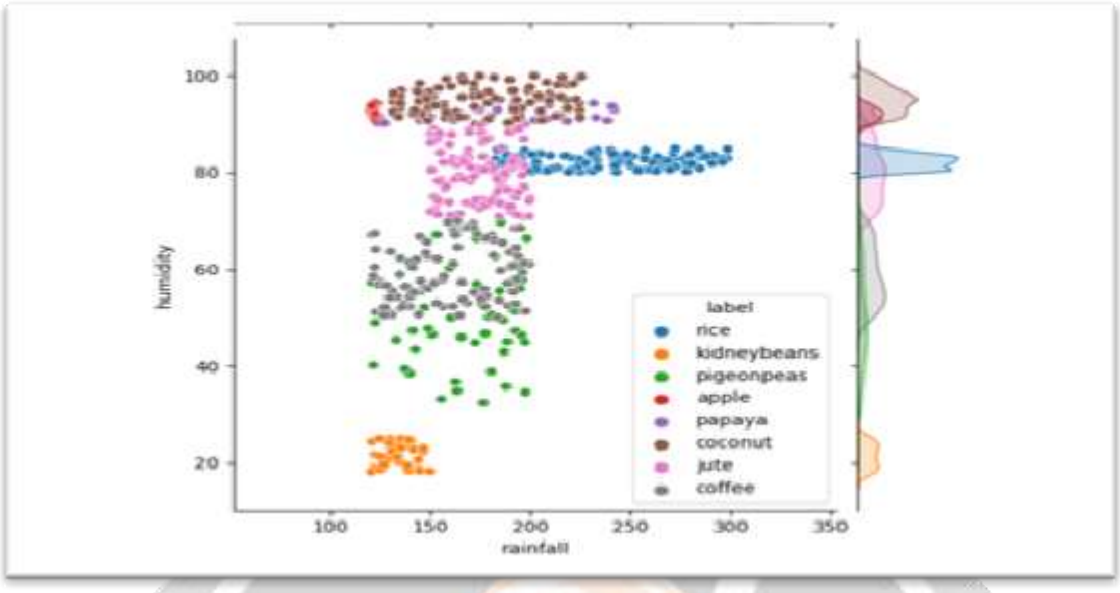

Fig - 3 Data Visualization Using Joint Plot

```
In [452]: # final result.
          prediction = lg.predict((np.array([[10,
                                              70,
                                              40,
                                              60,
                                              80,
                                              15,
                                              35]])))
          print("The Suggested crop for given climatic condition is := ", prediction)

          The Suggested crop for given climatic condition is :=  ['mothbeans']

          c:\users\abc\appdata\local\programs\python\python39\lib\site-packages\sklearn\base.py:450: UserWarning: X d
```

Fig - 4 Final Result

6. CONCLUSION

This system gives recommendations for crops to stay more profitable in agriculture and get better yields. The system assists the agronomists in picking an appropriate crop for their farming land based on the essential variables. The system is to plan and grow a recommendation model to create recommendations for crops relying on geological and climatic attributes using machine learning procedures.

7. ACKNOWLEDGEMENT

It gives us immense pleasure to express our gratitude to our guide for the project Mr. Goutam Kamble, who provided us constructive criticism and positive feedback during this project work. Our sincere thanks to the teaching and supporting staff of Computer Science & Engineering of Sharad Institute of Technology College of Engineering, Yadrav, without their help, we could not have even conceived the accomplishment of this report. This work is virtually the result of their inspiration and guidance. We would also like to thank the entire Library Staff and all those who directly or indirectly were part of this work. We are also thankful to our parents whose best wishes are always with us.

8. REFERENCES

[1]. Dhruv Piyush Parikh, Jugal Jain, Tanishq Gupta, and Rishit Hemant Dabhade, "Machine Learning Based Crop Recommendation System" in IJARSCT 2021.
[2]. G. Buvaanyaa, Dr. S. Radhimeenakshi, "Crop Recommendation System Using Random Forest Algorithm".
[3]. Thomas van Klompenburga, Ayalew Kassahuna, Cagatay Catalb, "Crop yield prediction using machine learning: A systematic literature review".
[4]. https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset