IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, December 2023

Crop Recommendation System Using ML Algorithms

Prof. Sonawane Meenakshi[1], Shaikh Moin[2], Dange Aniket[3], Salve Dnyaneshwari[4], Kaklij Rutuja[5]
Assistant Professor, Department of Computer Engineering[1]
Students, Department of Computer Engineering[2,3,4,5]
SND College of Engineering & Research Center, Yeola, India
sonawane.minakshi1@gmail.com[1], shaikhmoin7968@gmail.com[2], aniketmdange100@gmail.com[3],
dnyaneshwarisalve1506@gmail.com[4], kaklijr25@gmail.com[5].

Abstract*: This paper aims to give a comprehensive overview of crop recommendation systems, fastening on the application of machine literacy(ML) algorithms. Crop recommendation systems have gained significant attention in recent times due to their capability to help growers in making informed opinions regarding crop selection and optimization. This paper reviews the crucial generalities, ways, challenges, and advancements related to ML algorithms employed in crop recommendation systems. likewise, it discusses colourful datasets, evaluation criteria, and case studies available in the literature to illustrate the capabilities and limitations of being systems. The check concludes by relating implicit avenues for unborn exploration and pressing the significance of ML in revolutionizing agriculture*.

Keywords: Crop recommendation, Humidity, Rainfall, pH, Machine Learning (ML), Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), Logistic Regression (LR), and Naïve Bayes (NB),Data collection, Pre-processing, Feature extraction

I. INTRODUCTION

Crop recommendation systems play a vital part inultramodern husbandry by using technology and data- driven approaches to make informed crop selection opinions. These systems are of consummate significance as they enhance agrarian productivity, optimize resource application, and contribute to sustainable husbandry practices. By furnishing acclimatized recommendations grounded on environmental data and literal perceptivity, crop recommendation systems empower growers to maximize yields, minimize pitfalls, and ameliorate their overall profitable issues. In substance, these systems bridge the gap between traditional husbandry practices and the evolving demands of a changing world, icing food security and profitable stability for tilling communities. Machine Learning (ML) algorithms are instrumental in addressing the challenges of crop selection by analysing vast datasets and making data-driven recommendations.

In India today, agriculture has made significant advancements. Precision farming's secret weapon is "area-specific" cultivation. Although improvements have been made, there are still some problems with precision cultivation. Crop recommendations are significantly influenced by precision agriculture. Crop recommendations are determined by a variety of factors. Precision agriculture focuses on identifying these parameters in an area-specific way to identify issues. Not all the results given by precision agriculture are accurate to result but in agriculture, it is significant to have accurate and precise recommendations because in case of errors it may lead to heavy material and capital loss. Many research works are being carried out, to attain an accurate and more efficient model for crop prediction.

Machine Learning algorithms play a crucial role in various aspects.ML algorithms can process and interpret diverse data sources, including historical weather data, soil quality assessments, and geographical information, to extract meaningful insights. ML models can predict the performance of different crops under specific conditions, helping farmers choose the most suitable crops for their land.ML algorithms continuously learn and adapt, improving recommendations over time by incorporating real-time data and user feedback. ML-driven recommendations enable precise resource allocation, reducing waste and environmental impact while maximizing yields.

The purpose of this paper is to provide a comprehensive overview of the current state of Crop Recommendation Systems (CRS) that utilize Machine Learning algorithms. It aims to review existing research, technologies, and applications in this field to understand the advancements, challenges, and future directions. The scope of this paper covers various aspects of CRS, including data sources, ML algorithms, system architectures, case studies, challenges, and future prospects, offering valuable insights for researchers, practitioners, and stakeholders in agriculture and technology.

II. LITERATURE SURVEY

We've reviewed related work about this project in the past. For a crop recommendation system using ML algorithms, you should explore existing studies, research papers, and projects that are closely related to your work.[1] Kumar, Y. Jeevan Nagendra, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi. "Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector". In this proposed system crop yield prediction can be done from the past historical data which includes factors such as temperature, humidity, pH, rainfall, crop name. Under this system, maximum types of crops will be covered across different districts of India.[2] Suresh, G., A. Senthil Kumar, S. Lekashri, and R. Manikandan. "Efficient Crop Yield Recommendation System Using Machine Learning for Digital Farming". This proposed system is used to identify particular crop according to given particular data. By applying Support Vector Machine (SVM) acquired higher precision and productivity. This research paper mainly worked on two datasets: sample dataset of location data and sample dataset of crop data.

Agriculture plays a crucial role in the life of an economy. It is the backbone for developing countries like India as more than 70% of population depends on agriculture. To increase crop production many factors are responsible like soil, weather, rain, fertilizers and pesticides. They have used soil parameters to increase crop production because it is an essential key factor of agriculture. To maintain nutrient levels in the soil in case of deficiency, fertilizers are added to soil. The common problem existing among the Indian farmers is that they choose approximate number of fertilizers and add them manually. Excess or insufficient addition of fertilizer can harm the plant life and reduce the yield. The paper provides review of various data mining techniques used on agriculture soil dataset for fertilizer recommendation. Mainly focused on various soil parameters like Fe, S, Zn, Cu, N and Ph value etc. In this survey, authors also describe some Agriculture problems that can be solved by using data mining techniques such as Agriculture, Soil Fertility, Fertilizer Recommendation, Data Mining, Clustering, Classification, Neural Network. Algorithms used here are K-mean in Agriculture, K-nearest neighbour in Agriculture, SVMs in Agriculture, Decision Tree in agriculture.

III. PROPOSED SYSTEM

In this project, we have proposed a model that addresses the existing issues. The novelty of the proposed system is to guide the farmers to maximize the crop yield as well as suggest the most profitable crop for the specific region. The proposed model provides crop selection based on economic and environmental conditions, and benefit to maximize the crop yield that will subsequently help to meet the increasing demand for the country's food supplies. The proposed model predicts the crop yield by studying factors such as State, District, area, season. The system also helps to determine the best time to use fertilizers. The user provides a State, District, Season, Crop and Area as inputs for Production. The user provides a State, District, Season and Area as inputs for Crop Recommendation. According to the requirement, the model predicts the crop yield for a specific crop. The model also recommends the most profitable crop and suggests the right time to use the fertilizers. The main objective is to obtain a better variety of crops that can be grown over the season. The proposed system would help to minimize the difficulties faced by farmers in choosing a crop and maximize the yield.

As demonstrated in the figure, the methodology to extract the sentiment contains the several steps that are described below:

(1) Data Collection: The dataset consists of parameters like Nitrogen (N), Phosphorous (P), Potassium (K), PH value of soil, Humidity, Temperature and Rainfall. The datasets have been obtained from the Kaggle website. The data set has 2200 instance or data that have taken from the past historic data. This dataset includes twenty-two different crops such as 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. The dataset is separated in Train and Test sets in which 80% of the whole dataset is taken as Train and 20% as Test dataset.

(2) Pre-Processing (Noise Removal): For the successful application pre-processing is required. The data which is acquired from different resources are sometime in raw form. It may contain some incomplete, redundant, inconsistent data. Therefore, in this step such redundant data should be filtered. Data should be normalized. We also use Power BI to remove peak/downfall, local min-max, outliers, and junk values.

(3) Feature Extraction: This step is focus on identifying and using most relevant attribute from the dataset. Through this process irrelevant and redundant information is removed for the application of classifiers.

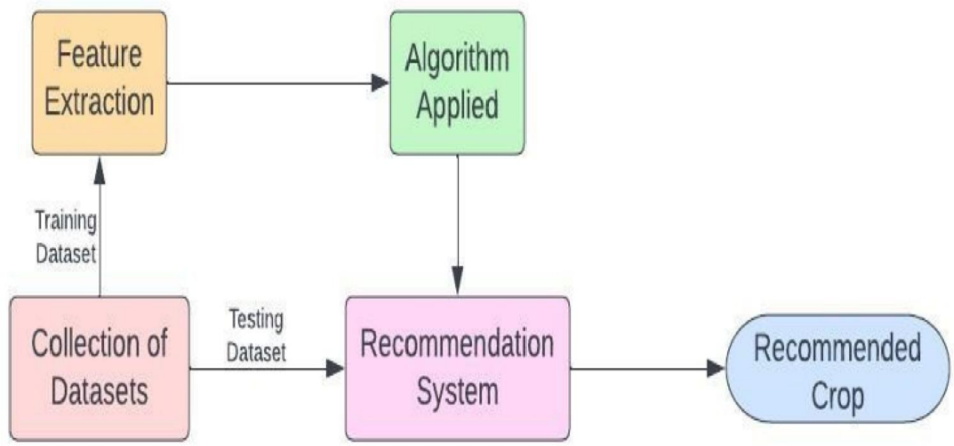

Fig 1.Block Diagram of Overall Methodology of Proposed System

IV. DESIGN AND ARCHITECTURE

The architecture diagram represents the overall design of the project. After taking the location as input from the user, the data get processed using soil attributes and weather attributes that includes crop details and all other trained data and finally the output that has maximum yield will be given to the user. This architecture ensures a seamless flow from data ingestion to user interaction, facilitating efficient crop recommendations and iterative model improvement. A crop recommendation system using machine learning (ML) algorithms involves creating a model that suggests the most suitable crops for a given area based on various factors. a crop recommendation system can provide valuable insights to farmers, helping them make informed decisions about the most suitable crops for their specific agricultural conditions. The data flow of the system architecture is as follow:

(1) User: The user interface in the system architecture serves as the entry point for farmers to input relevant data such as soil characteristics and geographical location. This user-provided information is then processed through machine learning algorithms, generating personalized crop recommendations. The intuitive design of the user interface ensures accessibility and ease of interaction for farmers seeking optimized agricultural decisions.

(2) User Interface: The user interface component in the crop recommendation system facilitates seamless interaction between farmers and the machine learning model. It provides an intuitive platform for farmers to input their specific soil and climate conditions, displays personalized crop recommendations generated by the ML algorithms, and ensures a user-friendly experience through clear visualization of results. Integration of user feedback mechanisms enhances the system's adaptability and usability for agricultural practitioners.

(3) Dataset: The dataset for the crop recommendation system includes comprehensive information on soil attributes, climate conditions, and historical crop performance. It encompasses variables such as soil Ph, temperature, precipitation, and crop yield, providing a rich source for training machine learning algorithms to make accurate crop recommendations.

(4) Data Pre-processing: In the data pre-processing stage of the crop recommendation system, raw datasets undergo cleaning to handle missing values and outliers. Feature scaling is applied to normalize variables, ensuring consistent influence during model training. Categorical variables are encoded, and the dataset is split

into training and testing sets for algorithm training and evaluation, respectively. Additionally, dimensionality reduction techniques may be employed to enhance computational efficiency.

(5) Algorithms: The system architecture incorporates machine learning algorithms such as Decision Trees, Random Forest, SVM, Gaussian NB, LR and K-NN to analyse the dataset and predict suitable crops. Ensemble techniques and cross-validation are employed to enhance the robustness and accuracy of the recommendations, ensuring adaptability to diverse agricultural scenarios.

(6) Recommended System: The recommendation system within the architecture employs machine learning algorithms like Decision Trees or Random Forest to analyse the input dataset and predict suitable crops based on soil and climate conditions. The system integrates user-friendly interfaces for farmers to input their specific parameters, facilitating personalized and optimized crop recommendations.

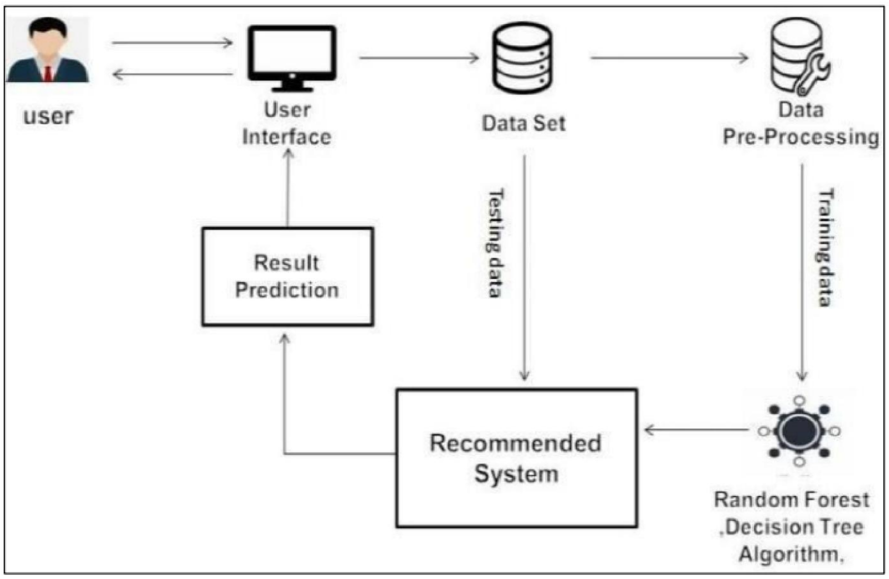

Fig. 2 System Architecture

V. IMPLEMENTATION AND ALGORITHMS

In regions where climate conditions frequently fluctuate, it's challenging to rely solely on weather data for crop cultivation. Technology is essential for collecting crop data and guiding farmers for better yields. Additionally, proper fertilizer usage is crucial, as excessive application can deplete soil fertility and result in subpar crop yields. In India, where agriculture significantly impacts the economy, accurate crop prediction is vital. Employing data mining techniques can provide a more reliable prediction tool, replacing inefficient and guesswork-based methods in farming decisions. In the prevailing device climatic conditions range very frequently. So, it's miles tough to extend flora with the useful aid of the use of facts weather situations. We need to use some era to locate o recognize the crop facts and guide the farmers to increase vegetation because of this and moreover fertilizer furthermore one of the important factors to boom flora as a forestall end result. If fertilizer is use more or less in the issue the soil might also moreover furthermore lose it fertility and crop may not supply the anticipated yield. So, fertilizer moreover becomes the number one element in it.

In this system we applied different Machine Learning algorithms like Random Forest, Decision Tree, Support Vector Machine (SVM), Logistic Regression (LR), and Gaussian NB, K-NN algorithm.

(1) Random Forest: Random Forest is an ensemble learning algorithm that builds multiple decision trees during training and merges their predictions. It operates by constructing a multitude of decision trees at training time and outputs the mode of the classes (classification) or the average prediction (regression) of the individual trees.

(2) Decision Tree: Decision tree classifiers utilize greedy methodology. It is a supervised learning algorithm where attributes and class labels are represented using a tree. The main purpose of using Decision Tree is to form a training

prototype which we can use to foresee class or value of target variables by learning decision rules deduced from previous data (training data).

(3) Support Vector Machine (SVM): Support Vector Machine (SVM) is a supervised machine learning algorithm or model which can be utilized for classification and as well as for regression challenges. However, we mainly use it in classification challenges. SVM is generally represented as training data points in space which is divided into groups by intelligible gap which is as far as possible.

(4) Logistic Regression (LR): The Logistic Regression model is a broadly used statistical model that, in its basic form, uses a logistic function to model a binary dependent variable; many more complex extensions exist. In Regression Examination, Logistic regression is predicting the parameters of a logistic model; it is a form of Binomial regression. Logistic Regression is a binary classification algorithm used for predicting the probability of an instance belonging to a particular class.

VI. CONCLUSION

In conclusion, the implementation of a Crop Recommendation System (CRS) utilizing Machine Learning (ML) algorithms offers immense promise in addressing the multifaceted challenges of crop selection in modern agriculture. This technology driven approach has the potential to revolutionize the agricultural landscape by providing farmers with data-driven insights and recommendations, thereby significantly enhancing agricultural productivity, sustainability, and economic outcomes. Through the synthesis of data from various sources, including historical weather data, soil quality assessments, and geographical information, ML algorithms enable the creation of predictive models capable of making accurate and tailored crop recommendations. These recommendations not only empower farmers to optimize resource allocation and mitigate risks but also contribute to the efficient and sustainable use of land and resources. Furthermore, the adaptability of ML algorithms ensures that CRS can continuously improve and evolve. By incorporating real-time data and user feedback, the system becomes increasingly precise and responsive, ultimately benefiting both farmers and the environment.

REFERENCES

[1]. Crop Recommendation Using Machine Learning Algorithms 2023 by Prashant Kumar, Keshav bhagat, Kusum Lata, Sushant Jhingran.

[2]. 2023 International Conférence on Disruptive Technologies (ICDT) | 979-8-3503-2388- 7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICDT57929.2023.10151325.

[3]. Crop Recommendation using Machine Learning and Plant Disease Identification using CNN and Transfer-Learning Approach 2022 by Shivesh Tiwari, Somesh Kumar, Sunil Tyagi, Minakshi Ponia.

[4]. Crop Recommender System Using Machine Learning Approach 2021 by Shilpa Mangesh Pande, Prem Kumar Ramesh, B. R. Aishwarya, Kumar Shourya.

[5]. Crop Recommendation System for Precision agriculture by S. Pudumalar, E. Ramanujam, R. Harine Rajashreen, C. Kavyan, T. Kiruthikan, J. Nishan.

[6]. International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 11 Issue II Feb 2023- Available at www.ijraset.com.