International Journal of
AgroChemistry
ISSN: 2456-7000
Volume 10, Issue 1, 2024
DOI (Journal): 10.37628/IJA
https://chemical.journalspub.info/index.php?journal=IJCPD
Review IJA

Estimation of Crop Yield Prediction Using Machine Learning Technique.

Abhijith D.A., Ajay V.G., Adimulam Gayathri, Sai Sushma Shree K.S. Zeba S. Khan

Abstract

Agriculture serves as the primary source of livelihood for a significant portion of India's population, but it faces a myriad of challenges, including climate variability, soil fertility issues, and reliance on traditional farming practices. To address these challenges, researchers are increasingly turning to machine learning (ML) methodologies to predict crop yields more accurately. Studies have delved into the effectiveness of various ML algorithms such as Random Forest, Gaussian Process Regression, and Support Vector Machines. These algorithms analyze factors like weather data, soil parameters, and historical yield data to provide precise predictions. The ultimate goal is to empower farmers with the information they need to make informed decisions about crop selection and agronomic practices, thereby contributing to food security and economic stability in agricultural-dependent regions like India. Furthermore, researchers are exploring techniques like Association Rule Mining and Genetic Algorithms to further refine crop yield prediction models. As climate change continues to impact agricultural productivity, the adoption of advanced technologies and data-driven approaches becomes increasingly crucial for sustainable farming practices. These approaches not only aid in mitigating the effects of climate variability but also ensure food security for India's growing population. By leveraging ML and other advanced technologies, agriculture in India can transition towards more sustainable practices while maximizing yields. This shift towards data-driven decision-making not only benefits farmers but also plays a vital role in ensuring the overall economic stability and food security of the nation. With continued research and implementation of these innovative techniques, India can navigate the challenges posed by climate change and sustainably meet the food demands of its populace.

Keywords: Crop yield prediction, SVM, LR, RF, DT.

INTRODUCTION

Crop recommendation is a critical aspect of modern agriculture, aiming to assist Assisting farmers in making well-informed decisions about which crops to cultivate on their land. Maximizing yields, optimizing resource allocation, and ensuring sustainable agricultural practices all hinge on the critical task of selecting the most suitable crops. However, the vast array of available crop options, coupled with the complex interplay of environmental factors, makes the decision-making process challenging for farmers. In recent years, advancements in machine learning and data analytics have unveiled fresh avenues for enhancing crop recommendation systems. These systems leverage Different data sources, including soil characteristics, climate data, historical crop performance, and expert knowledge, to generate tailored recommendations based on specific farming conditions. Machine learning algorithms play a

***Author for Correspondence**
Abhijith D.A.
E-mail: abijith.da@saividya.ac.in

[1]Assistant Professor, Department of CSE, Sai Vidya Institute of Technology Bangalore, Indi
[2]Associate Professor, Department of CSE, Sai Vidya Institute of Technology Bangalore, India
[3-5]Student, Department of CSE, Sai Vidya Institute of Technology Bangalore, India

Received Date: May 09, 2024
Received Date: May 14, 2024
Published Date: August 06, 2024

Citation: Abhijith D.A., Ajay V.G., Estimation of Crop Yield Prediction Using Machine Learning Technique International Journal of Agrochemistry. 2024; 10(2): 7–12p.

pivotal role in crop recommendation, enabling the analysis of large datasets and the extraction of meaningful patterns and insights. These algorithms can learn from historical data to identify correlations between crop performance and environmental variables, providing valuable guidance for future planting decisions. By integrating machine learning techniques with domain knowledge, crop recommendation systems have the potential to revolutionize agricultural practices. They can help farmers overcome challenges such as unpredictable weather patterns, changing soil conditions, and market demands. Moreover, these systems enable precision agriculture by optimizing resource utilization, reducing the use of fertilizers and pesticides, and minimizing environmental impact the system aims to provide accurate and personalized recommendations to farmers, ultimately leading to improved crop yields, cost-effectiveness, and sustainability in agricultural practices. The subsequent sections of this paper will delve into the methodologies, techniques, and datasets utilized in crop recommendation systems. We will discuss the challenges faced in crop selection and how machine learning algorithms can effectively address these challenges. Additionally, we will present experimental results and performance evaluations to showcase the efficacy of the suggested approach. Finally, we will discuss the potential benefits and future directions of crop recommendation systems in transforming the agriculture industry.

Literature Survey

Utilizing Machine Learning for Predicting Crop Yields Techniques: case of Irish potato and Maize, 2023 Martin Kuradusenge, Eric Hitimana - The literature Regarding the prediction of crop yields employing machine learning models, with a focus on Irish Potato and Maize, reveals a growing interest in leveraging advanced computational techniques to enhance agricultural productivity. Numerous studies have explored the application of various machine learning algorithms such as support vector machines, random forests, neural networks, and ensemble methods to predict crop yields based on a multitude of input factors, including weather conditions, soil properties, and agricultural practices.

Researchers emphasize the significance of accurate yield predictions for optimizing resource allocation, improving decision making processes for farmers, and addressing food security challenges [1].

Approach for prediction of Regarding the prediction of crop yields employing machine learning using ML and Big Data Techniques, 2019 Kodimalar Palanivel and Chellammal Surianaratanam - The literature pertaining to the prediction of crop yield using a combined approach of machine learning and big data techniques underscores a paradigm shift in agricultural practices toward data-driven decision making. Numerous studies have explored the integration of advanced machine learning algorithms, such as regression models, decision trees, and ensemble methods, with big data technologies to harness the vast amounts of agricultural data generated from diverse sources, including satellite imagery, weather stations, and sensor networks [2].

Crop Yield Prediction using Machine Learning Algorithm, 2020 Nikita S. Sapike and S. Sambare The literature on crop prediction using machine learning reflects a burgeoning field that holds significant promise for revolutionizing agricultural practices. Researchers have extensively explored the application of diverse machine learning algorithms, including support vector machines, random forests, neural networks, and ensemble methods, to predict crop yields based on a multitude of input factors such as climate variables, soil characteristics, and agronomic practices. These models offer the potential to enhance decision-making processes for farmers, optimize resource allocation, and contribute to overall food security. Moreover, the integration of remote sensing data, satellite imagery, and sensor networks has bolstered the precision and timeliness of crop prediction models. Despite notable achievements, challenges persist, including issues related to data quality, model interpretability, and scalability [3].

Data Set

In this project, we employ a comprehensive crop recommendation dataset that covers a wide variety of 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, among others. The dataset is meticulously analyzed to discern insights into crop attributes and performance. The number of instances of each crop within the training dataset is meticulously portrayed, ensuring a robust understanding of crop representation.

The factors being considered include Nitrogen (N), Potassium (K), Phosphorus (P), Temperature, Humidity, pH levels, and Rainfall. Nitrogen plays a pivotal role in fostering leaf growth within plants. Phosphorus, on the other hand, primarily facilitates root development and enhances flower and fruit maturation. Potassium serves as a vital nutrient that optimizes overall plant functionality. Temperature emerges as a critical factor influencing plant growth dynamics. In conjunction with light levels, carbon dioxide levels, air humidity, water availability, and nutrient levels, temperature exerts a profound impact on plant growth trajectories and eventual crop yields.

Humidity serves as a direct influencer on plant water relations, indirectly impacting leaf expansion, photosynthesis rates, pollination processes, disease occurrences, and ultimately, economic yields. The Ph level of soil is a pivotal variable affecting soil nutrient availability, microbial activity, and exchangeable aluminum levels. Rainfall patterns are instrumental in determining crop growth rates from seedling stages to maturity, significantly influencing harvest timelines. Achieving an optimal balance between natural rainfall and proper irrigation practices fosters accelerated plant growth, minimizes germination periods, and optimizes the time span between seeding and harvest, thereby enhancing overall agricultural productivity.

Methodology

To address the aforementioned limitations, we introduce an Intelligent Crop Recommendation system. This system considers all relevant parameters, including temperature, rainfall, geographical location, and soil conditions, which are essential for accurately predicting crop suitability. This system primarily functions akin to an Agricultural Consultant by offering tailored crop recommendations to farmers algorithms. Additionally, we offer a profit analysis for crops cultivated in various states, providing users with accessible and dependable insights to facilitate decision-making and crop planning. This forms the basis of our implementation plan.

System implementation entails several steps: beginning with the acquisition of a training dataset. The accuracy of any machine learning algorithm relies heavily on the quantity of parameters and the accuracy of the training data set..

Data Pre-processing

It involves replacing null and zero values for yield with -1 to ensure they do not impact the overall prediction. Further we had to encode the data-set so that it could be fed into the neural network.

Training Model and Crop Recommendation

Following pre-processing, we proceed to train various machine learning models, including logistic regression, decision trees, SVM, etc., using the dataset. Our aim is to achieve the highest possible accuracy in crop recommendation.

The proposed system's workflow, as depicted in **Figure 1**, encompasses multiple steps for sentiment extraction, outlined below:

Data Collection

The dataset includes parameters such as Nitrogen (N), Phosphorus (P), Potassium (K), soil pH value, humidity, temperature, and rainfall.

Pre-Processing

Pre-processing is essential for the successful application of the model. Data obtained from various sources often come in raw form, potentially containing incomplete, redundant, or inconsistent entries. Thus, in this phase, redundant data is filtered out, and normalization techniques are applied to ensure consistency and accuracy in the dataset.

Feature Extraction

Feature extraction refers to the process of Converting raw data into numerical features allows for processing while retaining the essential information from the original dataset. This approach typically yields superior results compared to directly applying machine learning algorithms to the raw data.

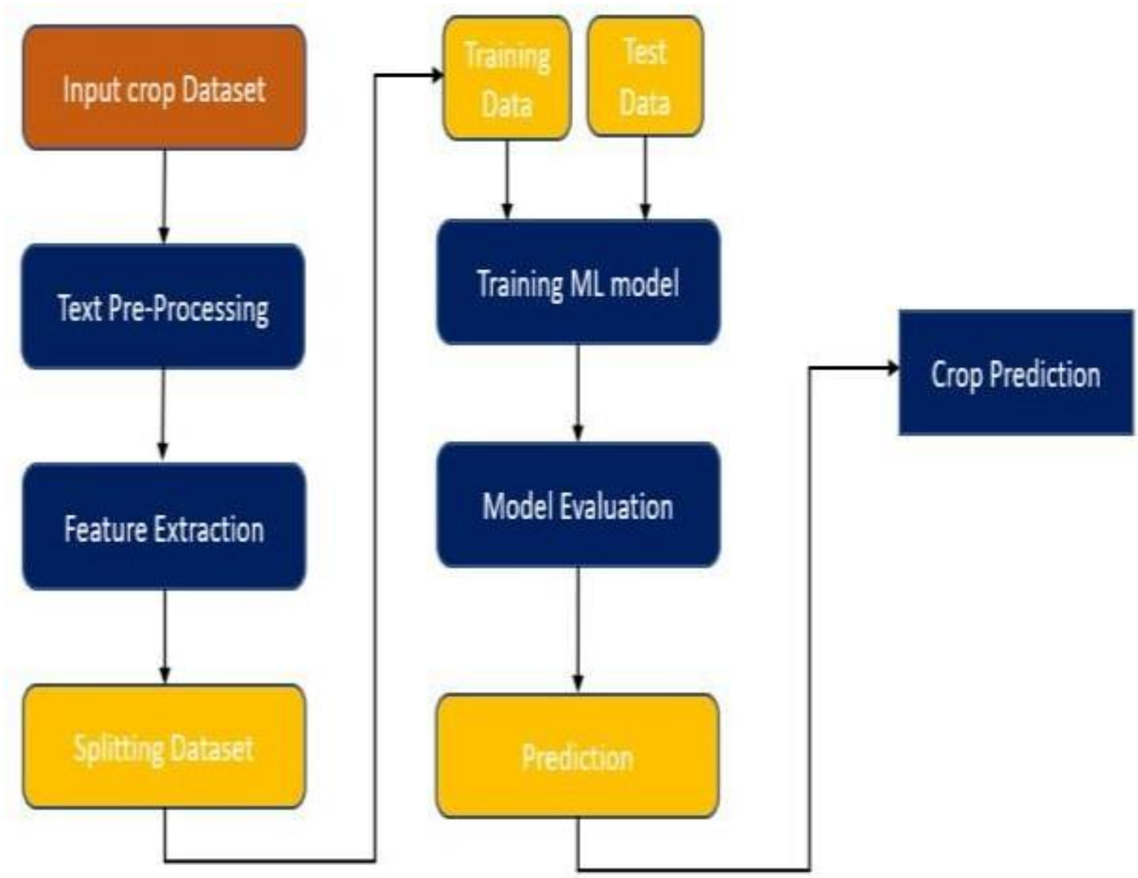

FIGURE 1: **Proposed System**

Applied Machine Learning Algorithm

Decision Tree

Decision tree classifiers operate with a greedy methodology, functioning as a supervised learning algorithm that represents attributes and class labels within a tree structure [4].

The primary objective of employing Decision Trees is to construct a training model, enabling the anticipation of class or target variable values by learning decision rules derived from historical data (training data). Within the structure of a Decision Tree, two distinct types exist: decision nodes and leaves. Leaves represent the final outcomes, while each node serves as a test case for an attribute, with edges corresponding to potential answers. This recursive process iterates for each sub-tree rooted at new nodes [5].

Logistic Regression

Logistic regression serves as a statistical analysis method aimed at predicting a binary outcome. It involves utilizing a model to forecast a dependent data variable by examining the relationship among one or more independent variables, drawing from previous observations of a dataset [6].

Support Vector Machine (SVM)

Support Vector Machine (SVM) stands as a robust technique for both classification and regression tasks, aiming to maximize predictive accuracy while avoiding overfitting to training data. It excels especially in handling datasets with a high number of predictor fields, often reaching into the thousands. SVM finds applications across various domains, including customer relationship management (CRM), image recognition (such as facial recognition), bioinformatics, text mining for concept extraction, intrusion detection, protein structure prediction, and voice/speech recognition [7.

Random Forest Classifier

The Random Forest Classifier stands as a widely adopted machine learning algorithm within supervised learning, adept at handling both classification and regression tasks. It functions on the principle of ensemble learning, which combines multiple classifiers to address complex problems and improve model efficacy. Essentially, a Random Forest classifier comprises numerous decision trees, each trained on distinct subsets of the dataset. By aggregating predictions from these trees, it boosts the predictive accuracy. Unlike depending on a single decision tree, Random Forest utilizes collective predictions from all trees, ultimately determining the final output through a majority vote mechanism [8].

Recommendation System

To forecast crop yield, a chosen Machine Learning algorithm is employed. Various machine learning algorithms, including Support Vector Machine (SVM), Logistic Regression, Decision Tree, and Random Forest (RF), were employed. Notably, Random Forest exhibited superior performance, achieving an impressive accuracy rate of 95% [9].

Crop Recommendation

In this step only, we predict the crop by giving the climate credentials like nitrogen level, Rain fall level, Potassium level etc. [10].

Future Enhancement

High-resolution Satellite Imagery

Integrating high resolution satellite imagery into crop prediction models allows for detailed analysis of crop health, growth patterns, and identifying potential issues such as pest infestations or nutrient deficiencies. This The data is available for utilization to generate more accurate predictions and enable timely interventions.

Weather data integration

Including real-time weather data in predicting crop outcomes models helps account for weather patterns and their impact on crop growth. By incorporating variables like temperature, precipitation, humidity, wind speed, and solar radiation, machine learning algorithms can capture the relationships between weather conditions and crop performance more effectively [11].

Crop-specific Models

Developing crop-specific prediction models can improve accuracy by accounting for the unique characteristics and growth patterns of different crops. Models tailored to specific crops can consider factors like planting and harvesting schedules, crop varieties, and regional conditions, leading to more precise predictions [12].

CONCLUSION

Utilizing machine learning for crop recommendation presents a promising and impactful approach poised to transform agriculture and empower farmers in decision-making regarding crop selection. Through the utilization of advanced algorithms and analysis of extensive datasets, machine learning models offer precise and tailored recommendations considering factors like soil quality, climatic conditions, market demand, and available resources. The advantages of employing machine learning for crop recommendation are manifold. Firstly, it facilitates farmers in optimizing their crop choices based on objective and data-driven insights, thereby enhancing productivity and increasing yields. Furthermore, by accounting for the intricate interplay of multiple variables, machine learning models can discern patterns and provide predictions, aiding farmers in making informed decisions tailored to their specific farming environments. By integrating advanced technologies with traditional farming practices, machine learning can optimize productivity, resource utilization, and decision making in agriculture. However, a collaborative approach that combines the power of machine learning with human expertise is necessary to maximize the benefits of crop recommendation systems and ensure their successful implementation in diverse farming contexts.

REFERENCES

1. R. Ghadge, J. Kulkarni, P. More, S. Nene, and R..L. Priya, “Prediction of crop yield using machine learning,” Int. Res. J. Eng. Technolgy, “(vol. 5, 2018)”.
2. F. H. Tseng, H. H.Cho, and H. T. Wu, “Applying big data for intelligent agriculture-based crop
selection analysis,” IEEE Access,”(vol. 7, pp. 116965- 116974, 2019)”.
3. A. Suresh, N. Manjunathan, P. Rajesh, and E. Thangadurai, “Crop Yield Prediction Using Linear Support Vector Machine,” European Journal of Molecular & Clinical Medicine, “(vol. 7, no. 6, pp. 2189- 2195, 2020)”
4. M. Alagurajan, and C. Vijayakumaran, “ML Methods for Crop Yield Prediction and Estimation: An Exploration,” International Journal of Engineering and Advanced Technology, “(vol. 9 no. 3, 2020)”.
5. P. Kumari, S. Rathore, A. Kalamkar, and T. Kambale, “Predicition of Crop Yeild Using SVM Approch with the Facility of E-MART System” Easychair 2020.
6. S. D. Kumar, S. Esakkirajan, S. Bama, and B. Keerthiveena, “A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier,” Microprocessors and Microsystems, “(vol. 76, pp.103090, 2020)”.
7. P. Tiwari, and P. Shukla, “Crop yield prediction by modified convolutional neural network and geographical indexes,” International Journal of Computer Sciences and Engineering, “(vol. 6, no. 8, pp. 503-513, 2018)”.
8. P. Sivanandhini, and J. Prakash, “Crop Yield Prediction Analysis using Feed Forward and Recurrent Neural Network,” International Journal of Innovative Science and Research Technology,”(vol.5, no. 5, pp. 1092-1096, 2020)”.
9. N. Nandhini, and J. G. Shankar, “Prediction of crop growth using machine learning based on seed,” Ictact journal on soft computing,”(vol. 11, no. 01, 2020)”.
10. A. A. Alif, I. F. Shukanya, and T. N. Afee, “Crop prediction based on geographical and climatic data using machine learning and deep learning”,”(Doctoral dissertation, BRAC University) 2018)”.
11. A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, “A robust deep learning-based detector for real- time tomato plant diseases and pests’ recognition,” Sensors,”(vol. 17, no. 9, pp. 2022, 2017)”.
12. J. Sun, L, Di, Z. Sun, Y. Shen, and Z. Lai, “Countylevel soybean yield prediction using deep CNNLSTM model,” Sensors,”(vol. 19, no. 20, pp. 4363,2019)”.

