

CROP YIELD PREDICTION

Chandana Cs¹, Gopika El², Jayanth A Yadav³, Sushmitha Mj⁴, Harish M⁵

^{1,2,3,4}BE Student, Computer Science and Design Department, PES Institute of Technology and Management,
Shivamogga, Karnataka, India.

⁵Assistant Professor, Computer Science and Design Department, PES Institute of Technology and Management,
Shivamogga, Karnataka, India.

ABSTRACT

As we know the fact that, India is the second largest population country in the world and majority of people in India have agriculture as their occupation. Farmers are growing same crops repeatedly without trying new variety of crops and they are applying fertilizers in random quantity without knowing the deficient content and quantity. So, this is directly affecting on crop yield and also causes the soil acidification and damages the top layer. So, we have designed the system using machine learning algorithms for betterment of farmers. Our system will suggest the best suitable crop for particular land based on content and weather parameters. And also, the system provides information about the required content and quantity of fertilizers, required seeds for cultivation. Hence by utilizing our system farmers can cultivate a new variety of crop, may increase in profit margin and can avoid soil pollution.

Keywords: Machine Learning, Crop prediction, Random Forest;

1. INTRODUCTION

Agriculture is one of the important occupations practiced in India. It is the broadest economic sector and place an important role in the overall development of the country. More than 60% of the land in the country is used for agriculture in order to suffice the needs of 1.3 billion people thus adopting new agriculture technologies is very important this will leads the farmers of our country towards profit. Prior crop prediction and yield prediction was performed on the basis of farmers experience on a particular location. They will prefer the prior or neighborhood or more trend crop in the surrounding region only for their land and they do not have enough of knowledge about soil nutrients content such as nitrogen, phosphorus, potassium in the land. Being this as the current situation without the rotation of crop and apply an inadequate amount of nutrient to soil it leads to reduce in the yield and soil pollution and damages the top layer. Considering all these problems takes into the account we design the system using machine learning for the betterment of farmer. Machine learning is a game changer for agriculture sector. Machine learning is the part of artificial intelligence, has emerged together with big data technologies and high-performance computing to create new opportunities for data intensive science in the multi-disciplinary agri-technology domain. In the agriculture field machine learning for instance is not a mysterious trick or magic, it is set of well-defined models that collect specific data and apply specific algorithm to achieve expected results.

The designed system will recommend the most suitable crop for particular land. Based on weather parameter and soil content such as Rainfall, Temperature, Humidity, Nitrogen, Phosphorous, Potassium, and pH. The system takes the required input from the farmers such as Temperature, Humidity, Nitrogen, Phosphorous, Potassium, and pH. This all-inputs data applies to machine learning predictive algorithm like Random Forest to identify the pattern among data and then process it as per input conditions. The system recommends crop for the farmer. A farmer's decision about which crop to grow is generally clouded by intuition and other irrelevant factors like making instant profits, lack of awareness about market demand, overestimating a soils potential to support a particular crop, and so on. The need of the hour is to design a system that could provide predictive insights to the Indian farmers, thereby helping them make an informed decision about which crop to grow. With this in mind, we propose a system, an intelligent system that would consider environmental parameters (temperature, rainfall, geographical location in terms of state) and soil characteristics (pH value, soil type and nutrients concentration) before recommending the most suitable crop to the user.

2. LITERATURE REVIEW

1.Kulkarani Varsha, Sridevi N, Prakasha G, Venkatagiri J-"Soil Classification, Crop Selection and Prediction of Fertilizer based on Soil Series" (2022):

- Conceptual Review:** The paper provides a comprehensive review of soil classification, crop selection, and fertilizer prediction, highlighting the significance of soil parameters such as pH, nutrient content, and moisture levels.
- Empirical Evidence:** The effectiveness of machine learning algorithms in predicting crop productivity is demonstrated, showcasing practical applications that enhance agricultural practices in India

2.Madhumathi R, Ashok Kumar K, Chandra Kumar N, Arumuganathan T-"Soil Nutrient Prediction and Crop Recommendation System" (2023):

c. Conceptual Review: The study presents a Soil Nutrient Prediction and Crop Recommendation System using the BORUTA regression model, achieving 91% accuracy in predicting NPK values and recommending suitable crops based on soil conditions.

d. Empirical Evidence: The system demonstrated a 91% accuracy in predicting soil NPK values using the BORUTA regression model, and effectively recommended crops like rice and maize based on soil nutrient analysis.

3.Fatim Farhan Haque, Ahmed Abdelgawad, Venkata Prasanth Yanambaka, Kumar Yelamarthi- "Crop Yield Analysis Using Machine Learning Algorithms" (2024):

e. Conceptual Review: The paper explores crop yield prediction using Support Vector Regression and Linear Regression, analyzing key factors like water and UV exposure, achieving promising results with low error rates and high accuracy.

f. Empirical Evidence: The study utilized 140 data points, demonstrating that Support Vector Regression achieved an MSE of approximately 0.005 and an R^2 value of around 0.85, indicating strong predictive accuracy for crop yield.

4.K P K Devan, Swetha B, Uma Sruthi P, Varshini S -"Crop Yield Prediction and Fertilizer Recommendation System Using Hybrid Machine Learning Algorithms" (2024):

g. Conceptual Review: The paper presents a hybrid machine learning approach using Random Forest and Logistic Regression for crop yield prediction and fertilizer recommendation, enhancing agricultural efficiency and sustainability through data-driven insights.

h. Empirical Evidence: The study demonstrates empirical evidence through hybrid machine learning algorithms, achieving approximately 92% accuracy in predicting crop yield and recommending fertilizers based on diverse agricultural datasets and conditions.

Each paper contributes valuable insights into the design, functionality, and effectiveness of crop yield prediction system, especially in agriculture information roles. These studies collectively highlight the potential of crop yield prediction to streamline information access, and improve the efficiency.s

3. RESEARCH METHODOLOGY

The research methodology for developing a machine learning that can help predict crop yields by identifying patterns and correlations in data-sets. It can be a practical approach to help farmers select the right crops to grow in various fields. Key steps include:

3.1. Data Preprocessing: Data preprocessing is a technique for transforming unprocessed data into a flawless data set. At the end of the day, whenever data is gathered from various sources, it is gathered in a raw or crude form that cannot be analyzed by machine learning or deep learning methodologies.

3.2. Machine Learning: Machine learning models are trained on a dataset to match user input with accurate responses, with training often including prediction as per the farmers need.

3.3. System Design and Architecture: Keeping the requirements in mind the system specifications are translated into a software representation. In this phase the designer emphasizes on: - algorithm, data structure, software architecture etc. A system architecture is a conceptual model using which we can defined the structure and behaviour of the system. It is a formal representation of a system. Depending on the context the system architecture can be used refer to either a model to describe system or a method used to build the system. Building a proper system architecture helps in the analysis of the project.

3.4. Testing: In this phase all programs are integrated and tested to ensure that the complete system meets the software requirements. The testing is concerned with verification and validation.

3.5. Quality Assurance: Quality Assurance is popularly known as QA Testing, is defined as an activity to ensure that an organization is providing the best possible product or service to customers. QA focuses on improving the processes to deliver Quality Products to the customer. An organization has to ensure, that processes are efficient and effective as per the quality standards defined for software products.

3.6. Product Requirements:

Correctness: It followed a well-defined set of procedures and rules to engage a conversation with the user and a pre-trained classification model to compute also rigorous testing is perform to confirm the correctness of the data.

Modularity: The complete product is broken up into many models and well-defined interfaces are developed to explore the benefit of flexibility of the product.

4. MODELING AND ANALYSIS

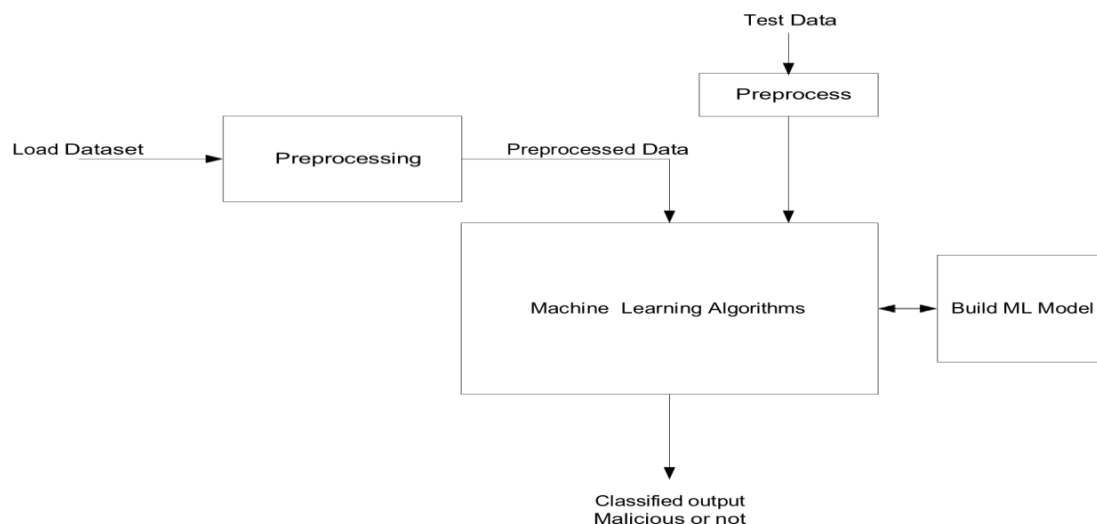


Fig:4.1 for Modeling and Analysis

5. RESULTS AND DISCUSSION

5.1. Crop Yield Prediction Accuracy

The following image is a bar chart comparing the accuracy of different machine learning algorithms. Along the left side of y-axis are the algorithms, and the length of the bars along the horizontal axis, the x-axis, represents their corresponding accuracy.

As shown by the chart above, we can see that algorithms like Random Forest reach the highest accuracy, almost reaching a value of 1 (or 99%). The other algorithm as having slightly low accuracy.

Table 5.1 for Crop Yield Prediction accuracy

ALGORITHM	ACCURACY
Logistic Regression	94.54
Decision Tree	97.72
Support Vector Machine	90.90
Multi-Layer Perception	95.22
Random Forest	99.31

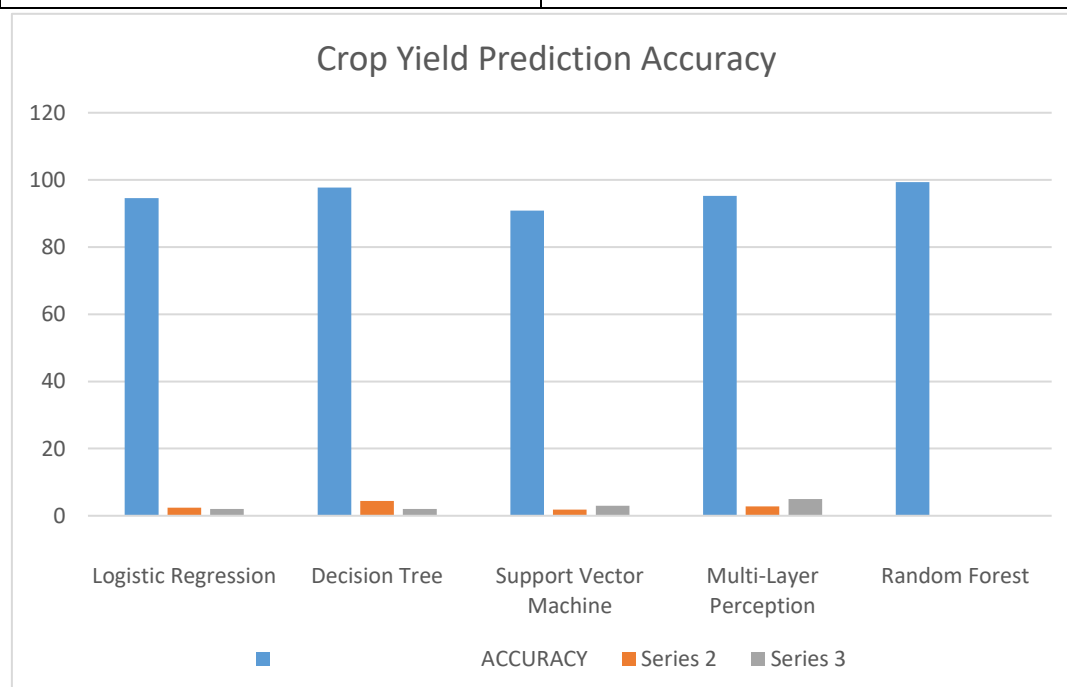


Chart :5.1 for Accuracy Comparison

5.2. Daily and Weekly Query Load

The Crop yield prediction in machine learning more than manual roto. While approaches estimate 1000 tons, AI approaches give an estimation of 5000 tons hence 400 accuracy levels than the manual support. At full harvest, the manual system forecasts 1500 tons of gain while the AI system forecasts 6000 tons which is 300 improvements over the manual methods. And time also saved.

Table:5.2 for Daily and Weekly Query Load

Time Period	Manual System	Advanced system	Improvement
Average Crop Yield Prediction	1000 tons	5000 tons	+400%
Peak Prediction Period	1500 tons	6000 tons	+300%
Weekly Prediction Effort	40 hours	10 hours	-75%

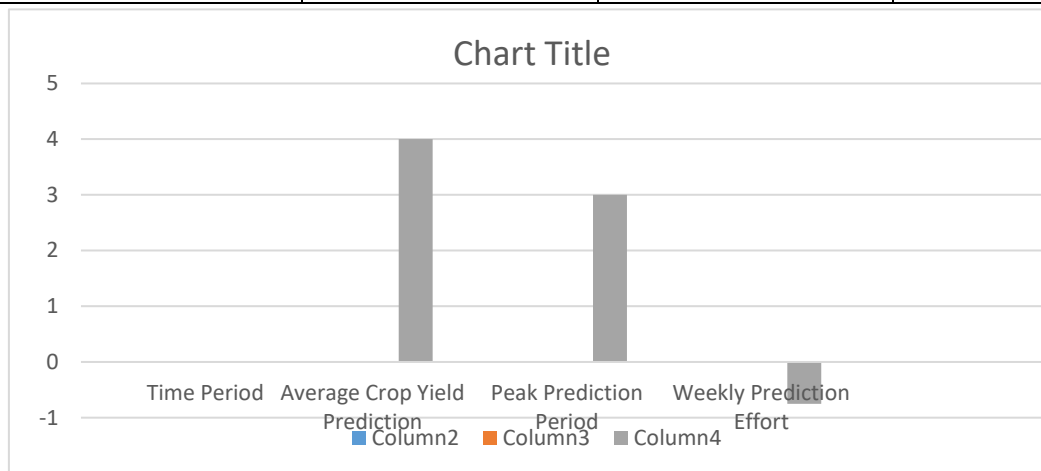


Chart:5.2 for Daily and Weekly Query Load

5.3. Administrative Workload Reduction

The table indicates the time-saving effects of automation and AI in crop yield prediction. Preceding its implementation, activities; Namely data collecting, forecasts and reporting averaged 33 hours a week. This figure as significantly lowered to a mere 3.5 hours a week, translating to an 89% reduction in a work load. The process of collecting data changed from taking 15 hours to just 2hours, making predictions, on another hand, was reduced from a period of 10 hours to 1 hour. This process is made quick and simple, thus allowing farmers to attend to other significant task.

Table:5.3 for Administrative Workload Reduction

Task	Time(pre-automation)	Time(post-automation)	Reduction
Manual data collection	15 hours/week	2 hours/week	-87%
Crop Yield Prediction	10 hours/week	1 hours/week	-90%
Report generation and analysis	8 hours/week	0.5 hours/week	-93.75%
Total time spent on yield tasks	33 hours/week	3.5 hours/week	-89%

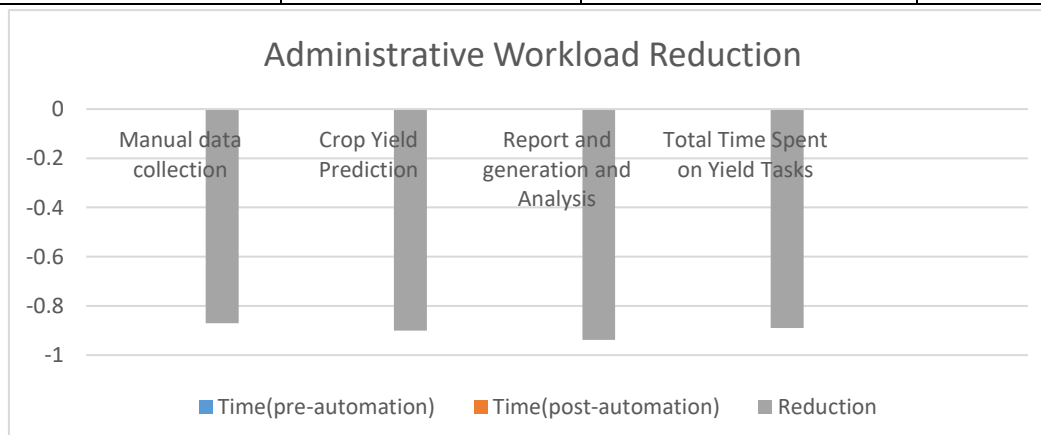


Chart:5.3 for Administrative Workload Reduction

5.4. User Satisfaction and Engagement

The figure illustrates the changes in crop yield and other factors in new prediction approach. Previously, the crop yield was 2.5 tons, on the contrary, after applying the prediction method, the yield increased to 3.8 tons. The use of watering plant growth improved from 75% to 85%, and fertility of the soil rated from 60% to 90%.

Table:5.4 for User Satisfaction and Engagement

Survey Metric	Pre-prediction	Post-prediction	Change
Crop Yield (tons per hectare)	2.5 tons	3.8 tons	+1.3 tons
Water Usage Efficiency	75%	85%	+10%
Soil Fertility (Positive Rating)	60%	90%	+30%

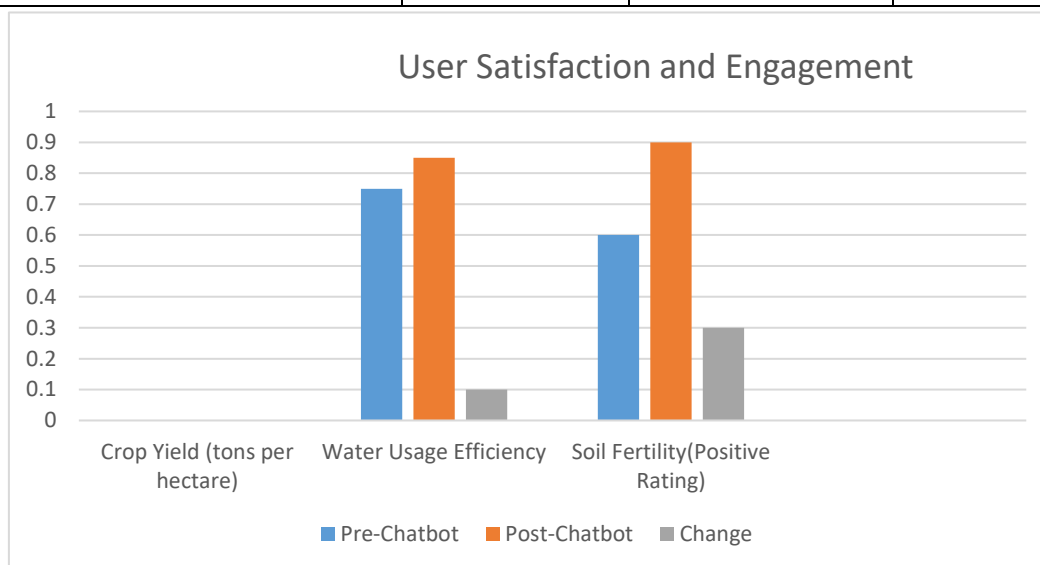


Chart:5.4 for User Satisfaction and Engagement

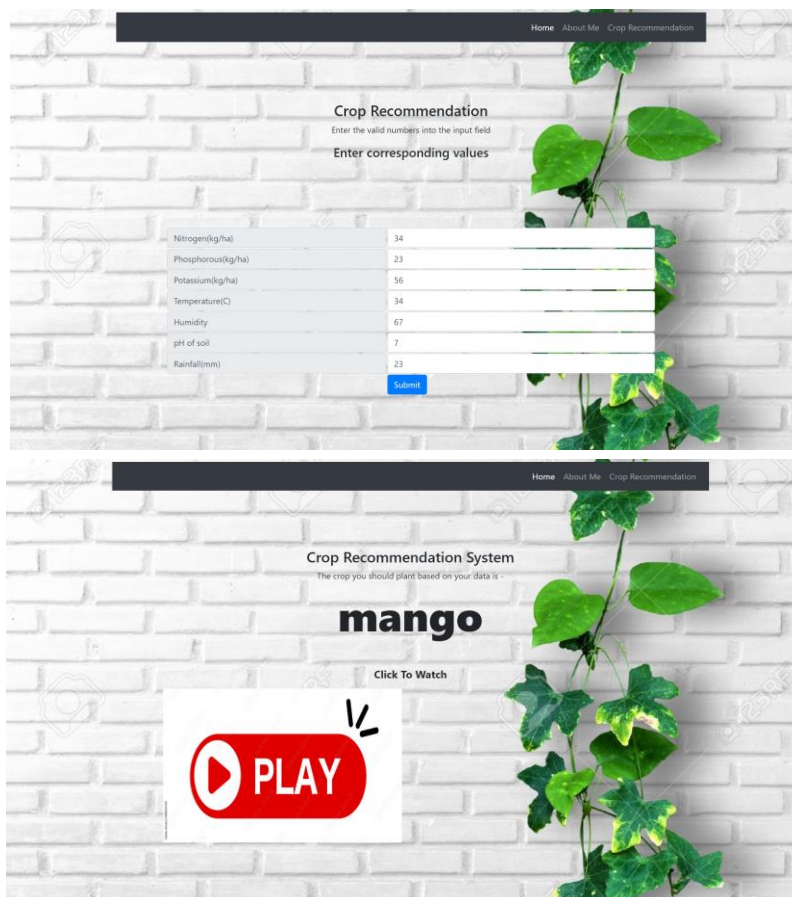


Fig .2 for Result Page

6. RESULTS COMPARISON TABLE

Aspect	Statistical Model	Machine Learning Model
Accuracy	75% Relies on historical trends	85% non-linear relationships well
Data Requirements	Low data requirements structured historical data	Medium needs deserve data set
Adaptability	Low limited to pre-defined models and assumptions	Medium can adapt to new regions with sufficient data.
Use Cases	Simple yield prediction for small farms	Predicting yields with multiple influencing facts like weather, soil.

7. CONCLUSION

The Farmers are currently ineffective utilizing the technology and analytical tools. This mis-section of crops for cultivation may lead to reduced income. Hence, we designed a farmer friendly system including GUI through which best suitable crop can be predicted based on certain condition of land and also based on Temperature, Humidity, Rainfall, pH provided for cultivation.

It will help for farmers to take an accurate decision as to which crop to choose for cultivation and overall growth of the agricultural sector through innovative idea.

This is a system that assists the farmers on which crop to grow just by giving some insights that regular farmers don't keep track of and hence reduce the chance of crop failure and increases the chance of productivity. It also saves them from losses. It can be stretched to the web and assessed by millions of farmers across the country.

8. REFERENCES

- [1] Kulkarni Varsha, Prakasha G, Sridevi N, Venkatagiri J, "Soil Classification, Crop Selection and Prediction of Fertilizer based on Soil Series", 2022 International Interdisciplinary Humanitarian Conference for Sustainability (IIHC) | 978-1-6654-5687-6/22/\$31.00 ©2022 IEEE | DOI: 10.1109/IIHC55949.2022.10059636.
- [2] Madhumathi R, Ashok Kumar K, Chandra Kumar N, Arumuganathan T, "Soil Nutrient Prediction and Crop Recommendation System", 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT) | 979-8-3503-3324-4/23/\$31.00 ©2023 IEEE | DOI: 10.1109/ICCPCT58313.2023.10245685.
- [3] Fatin Farhan Haque , Ahmed Abdelgawad, Venkata Prasanth Yanambaka , KumarYelamarthi, "Crop Yield Analysis Using Machine Learning Algorithms", IEEE 2024
- [4] K P K Devan, Swetha B, Uma Sruthi P, Varshini S, "Crop Yield Prediction and Fertilizer Recommendation System Using Hybrid Machine Learning Algorithms", 2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT) | 978-1-6654-6261-7/23/\$31.00 ©2023 IEEE | DOI: 10.1109/CSNT57126.2023.10134662.