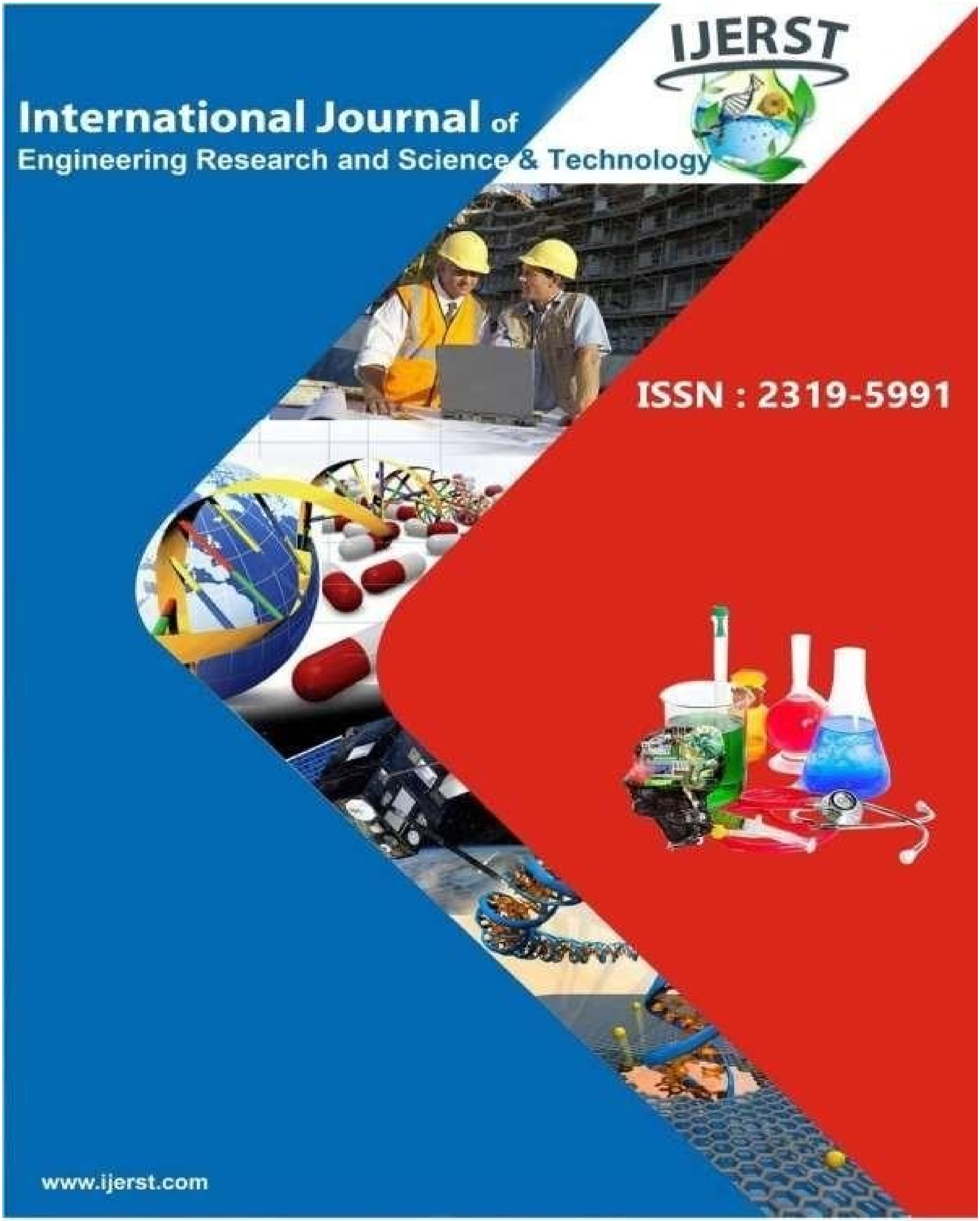


Email: editor@ijerst.com or editor.ijerst@gmail.com

https://doi.org/10.62643/ijerst.2024.v20.i3.pp131-136

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 3, 2024

# CROP YIELD PREDICTION USING MACHINE LEARNING

[1]**Mr. Y. Nagendra Kumar, Associate Professor, Department of computer science and engineering, Ramachandra College of Engineering, Eluru, Andhra Pradesh-521230, India**

[2]**K. Shanmukh Jagan Chowdary,** [3] **M. Sruthi Keerthi Harshini,** [4] **V. Praveen Srikar,** [5] **K. Shanmukha Krishna**

**Department of computer science and engineering,**

**Ramachandra College of Engineering, Eluru, Andhra Pradesh-521230, India**

1 kakarlashanmukhchowdary12@gmail.com, [2] sruthimuthoju@gmail.com,[3] praveensreekar05@gmail.com, [4] shanmukhshannu89@gmail.com

## Abstract

Agricultural productivity hinges on a myriad of factors, each playing a vital role in shaping crop growth and yield. This abstract offers an insight into the significance of key factors such as nitrogen, phosphorus, potassium, pH, temperature, humidity, rainfall, and type of soil in agricultural practices. Nitrogen, phosphorus, and potassium are the indispensable nutrients critical for plant metabolic functions, growth, and productivity. Their availability in soil profoundly impacts crop health, necessitating precise fertilization strategies. Soil pH, governing nutrient uptake and microbial activity, underscores the importance of maintaining optimal levels for crop development. Temperature dictates crucial growth stages, with crops exhibiting specific temperature requirements for germination, growth, and reproduction. Humidity influences transpiration rates, disease susceptibility, and overall plant vigor, necessitating judicious irrigation and crop management practices. Rainfall patterns profoundly influence agricultural outcomes. Adequate rainfall ensures soil moisture, sustains plant growth, and supports crop yields. Soil type, defined by texture, structure, and composition, shapes water retention, nutrient availability, and root growth. Tailored soil management, including amendments and rotations, optimizes fertility and productivity.

**Keywords:** Predictive analytics, Agriculture industry, Fertilization strategies, Soil pH, Temperature requirements, Humidity, Rainfall patterns, Soil type, Nutrient uptake, Flask deployment, Weather conditions, Soil quality, climate data, Yield prediction, Agricultural productivity, Random Forest Classification, Model training, Crop selection

## 1.Introduction:

Crop yield prediction is an essential predictive analytics technique in the agriculture industry. It is an agricultural practice that can help farmers and farming businesses predict crop yield in a particular season when to plant a crop, and when to harvest for better crop yield. Predictive analytics is a powerful tool that can help to improve decision-making in the agriculture industry. It can be used for crop yield prediction, risk mitigation, reducing the cost of fertilizers, etc. The crop yield prediction using ML and flask deployment will find analysis on weather conditions, soil quality, fruit set, fruit mass, etc.

Agriculture plays a vital role in Indian economy. Our Project helps farmers to get more yields of crops can be achieved by analyzing agro - climate data using machine learning techniques. Machine Learning is emerging research field in crop yield analysis. Yield prediction is a very important

https://doi.org/10.62643/ijerst.2024.v20.i3.pp131-136

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 3, 2024

issue in agricultural. Any farmer is interested in knowing how much yield he is about to expect. In the past, yield prediction was performed by considering farmer's experience on particular field and crop. The yield prediction is a major issue that remains to be solved based on available data.

Machine learning model are trained using the datasets, and the expected result are based on previous experience. We can determine the parameter for the model during the training phase by evaluating past data. A portion of the previous data is considered in the evaluation.

Weather prediction is the important responsibility in agriculture; however, it is the tough process to predict manually. Too many elements affect agriculture productivity since it is not dependent on a single factor. The logistic regression is the simple supervised machine learning algorithm which train the dataset and that dataset will predict the suitable crop for the yield.

Thus, the main objective of our paper is to predict the suitable crop based on weather condition.Do not even comply to a small percentage of the requirements mentioned above, the failure rate of startups is as high as 97.6% in India as per the survey data released by concerned department of Government of India.[4,5,6]

So far, "research" has been considered as an outcome in the form of many research papers published in journals. The current requirements are publication in highly reputable journals, maintaining the shelf life of published paper over a long period of time, and getting good citations. The publications in Scopus Indexed, Web of Science (WoS), Science Citation Index Expanded (SCIE) journals with good impact factor and cite-score fetch good accolades. Faculty and students must work together with other national and international researchers to publish in reputable journals and also to maintain high citations for published articles. The term "research" is not limited to publications, but extends to obtaining research grants from government and private organizations. Therefore, the rubric proposed by the Ranking Agencies for Higher Education Institutions focuses on the research and consultancy grant received by higher education institutions.

A good research ambience should be created by higher education institutions to encourage the faculty and students to make research one of the core activities of the institution. The faculty and students must be encouraged with incentives to publish research papers, seed grant to undertake preparatory research work. Reducing the teaching burden of accomplished researchers is another good measure by higher education institutions' authorities. Encouraging the registration of Junior Research Fellow (JFR) and Senior Research Fellow (SFR), attracting full-time researchers with good research scholarships and many such initiatives help to improve the research rubric.

Hence, all higher education institutions must promote the culture of research and innovation among their faculty and students and contribute to the scientific community in particular and to society as a whole.

## 2. Discussion

Crop yield prediction is an essential predictive analytics technique in the agriculture industry. It is an agricultural practice that can help farmers and farming businesses predict crop yield in a particular season when to plant a crop, and when to harvest for better crop yield. Predictive analytics is a powerful tool that can help to improve decision-making in the agriculture industry. It can be used for crop yield prediction, risk mitigation, reducing the cost of fertilizers, etc. The crop yield prediction using ML and flask deployment will find analysis on weather conditions, soil quality, fruit set, fruit mass, etc.

Agriculture plays a vital role in Indian economy. Our Project helps farmers to get more yields of crops can be achieved by analyzing agro-climate data using machine learning techniques. Machine Learning is emerging research field in crop yield analysis.Yield prediction is a very important issue in agricultural. Any farmer is interested in knowing how much yield he is about to expect. In the past,

https://doi.org/10.62643/ijerst.2024.v20.i3.pp131-136
ISSN 2319-5991 www.ijerst.com
Vol. 20, Issue 3, 2024

yield prediction was performed by considering farmer's experience on particular field and crop. The yield prediction is a major issue that remains to be solved based on available data.

Machine learning model are trained using the datasets, and the expected result are based on previous experience. We can determine the parameter for the model during the training phase by evaluating past data. A portion of the previous data is considered in the evaluation.

Weather prediction is the important responsibility in agriculture; however, it is the tough process to predict manually. Too many elements affect agriculture productivity since it is not dependent on a single factor. The logistic regression is the simple supervised machine learning algorithm which train the dataset and that dataset will predict the suitable crop for the yield. Thus, the main objective of our paper is to predict the suitable crop based on weather condition.

The Crop Yield Prediction uses the various farm inputs, weather conditions, type of soil is used to predict the suitable crop in the particular scenario and the predicted crop yield in terms of kilograms per hectare. The Farm inputs include the seeds, fertilizers, pesticides, machinery, and labour. In our project we are considering only the fertility of the soil that is the ability of soil to sustain plant growth and optimize crop yield.

There are two types of nutrients in the soil are macro and micro nutrients. Macronutrients are Nitrogen (N), Phosphorus (P), Potassium (K) and Micronutrients are Manganese (Mn), Boron (B), Copper (Cu), Iron (Fe), Zinc (Zn), Nickel (Ni), Molybdenum (Mo) and Chlorine (Cl). The availability of N, P and K in soil should be sufficient, but not too high. Too low availabilities will lead to hampered growth and low yields, while too high availabilities of one or more nutrients may lead to disturbed plant growth and adverse effects for yield and/or quality of harvested products.

Moreover, the N, P and K availability should be balanced, so the availability of the other nutrients should be taken into account while the availability of the considered nutrient is adjusted. It is important for farmers to know the NPK content in their soil. For the optimal growth of crops, sufficient amounts of nutrients should be available in the root zone of the crops.

## Architecture

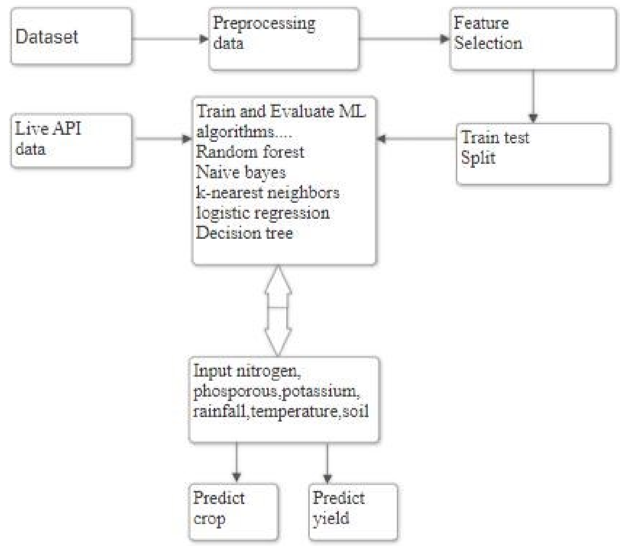


**Fig: Crop Yield Prediction Architecture**

### Model

In the Crop Yield Prediction System, the training and validation process is crucial for developing accurate predictive models. The process begins with the collection of historical agricultural data from various sources, including weather stations, agricultural databases, and government agencies. This data encompasses a wide range of variables, such as weather conditions (temperature, humidity, rainfall), soil fertility indicators (nitrogen, phosphorus, potassium levels), and crop yields.

Once the data is collected, it undergoes preprocessing to ensure its quality and suitability for model training. Preprocessing tasks include handling missing values, scaling numerical features, and encoding categorical variables. Exploratory data analysis (EDA) techniques are then applied to gain insights into the relationships between different variables and identify potential predictors of crop yields.

https://doi.org/10.62643/ijerst.2024.v20.i3.pp131-136
ISSN 2319-5991 www.ijerst.com
Vol. 20, Issue 3, 2024

Next, suitable machine learning algorithms are selected for building predictive models. Decision trees, random forests, and gradient boosting machines are commonly employed due to their ability to capture complex relationships in the data. The selected models are trained using the preprocessed data, with the training process involving feeding the models with input data and adjusting their parameters iteratively to minimize prediction errors.

Validation of the trained models is performed using techniques such as holdout validation or cross-validation. A portion of the dataset is held out as a validation set, and the models are evaluated on this set to ensure their generalization ability and detect overfitting. Performance metrics such as accuracy, precision, recall, and F1 score are calculated to evaluate the effectiveness of the trained models.

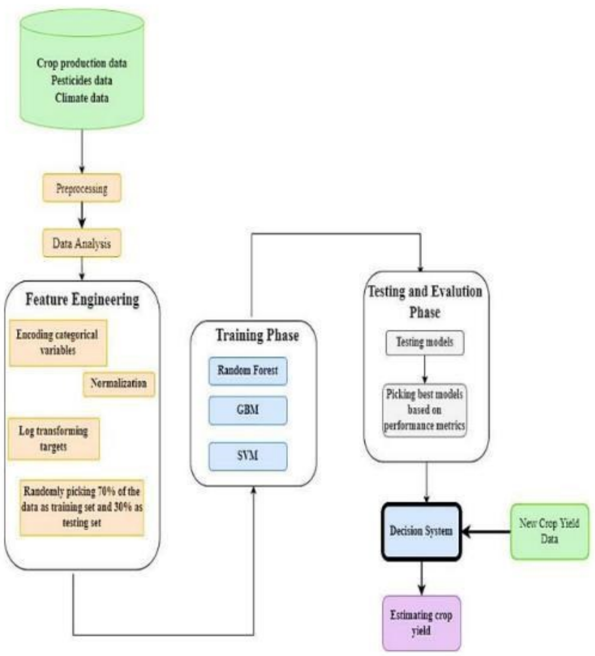


**Fig. 2 – Model Diagram for Crop Yield Prediction**

In the testing and evaluation phase of the Crop Yield Prediction System, the focus shifts to assessing the performance of the trained models in real-world scenarios. This phase is critical as it determines the practical applicability and reliability of the predictive models developed during the training phase. The primary objective here is to validate the efficacy of the models in making accurate predictions on unseen data, simulating the conditions that the models will encounter when deployed in production environments.

The trained models, which have successfully passed validation during the training phase, are then deployed to a server environment, making them accessible to end users for real-time crop yield prediction. End users, including farmers, agricultural advisors, and policymakers, interact with the deployed models through user-friendly interfaces, providing inputs such as current weather conditions and soil fertility indicators.

Upon receiving input data, the deployed models generate predictions for crop yields based on their learned patterns and relationships derived from the training data. These predictions are compared against ground truth values from the test dataset to assess the accuracy and reliability of the models predictions.

## 3. Result

The project has significantly enhanced agricultural planning, enabling farmers and stakeholders to make informed decisions about planting schedules, resource allocation, and crop management. By predicting optimal planting and harvesting times, the project ensures better use of time and resources, leading to improved agricultural productivity. This precision in planning allows farmers to align their activities with favourable weather conditions, thereby maximizing crop yields.

In terms of resource allocation, the machine learning model has proven invaluable in optimizing the use of inputs such as water, fertilizers, and pesticides. By providing data-driven insights, the model helps farmers avoid the costs associated with overuse or underuse of these inputs. This not only reduces expenses but also promotes sustainable farming practices, contributing to long-term soil health and ecosystem balance.

https://doi.org/10.62643/ijerst.2024.v20.i3.pp131-136

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 3, 2024

The accurate crop yield predictions offered by the model empower farmers to maximize yields by tailoring their farming practices to current and forecasted weather conditions. This adaptability enhances overall productivity, ensuring that interventions such as irrigation and fertilization are timely and effective. Sustainable practices fostered by these predictions support the long-term viability of agricultural operations.

Predictive models also play a crucial role in mitigating risks related to adverse weather conditions. By providing early warnings, the models enable farmers to take preventive measures, such as adjusting planting dates or implementing protective structures. This proactive approach helps protect crops from extreme weather events, ensuring more stable and predictable agricultural outputs.

Food security is another critical area supported by predictive analytics. Reliable crop production forecasts help maintain a balance between supply and demand, preventing shortages and price spikes. This stability is vital for meeting the nutritional needs of the growing global population, especially in regions prone to food scarcity.

The project promotes the adoption of data-driven agricultural practices, encouraging farmers to select the most suitable crops based on soil and weather conditions. Better understanding of soil nutrient content and other characteristics allows for improved soil management and fertility optimization, leading to higher yields and more efficient land use.

Technological advancements are a significant outcome of the project. By integrating machine learning techniques and data analytics into agriculture, the project advances the development of robust predictive models based on historical and real-time data. These models enhance the accuracy and reliability of crop yield forecasts, driving continuous improvement and innovation in agricultural technology.

Economically, the project benefits farmers by maximizing crop yields and reducing input costs, directly contributing to increased profitability. Additionally, the insights derived from yield predictions help traders and policymakers forecast market trends, leading to informed decision-making and economic stability. This economic boost positively impacts the broader agricultural industry and national economies.

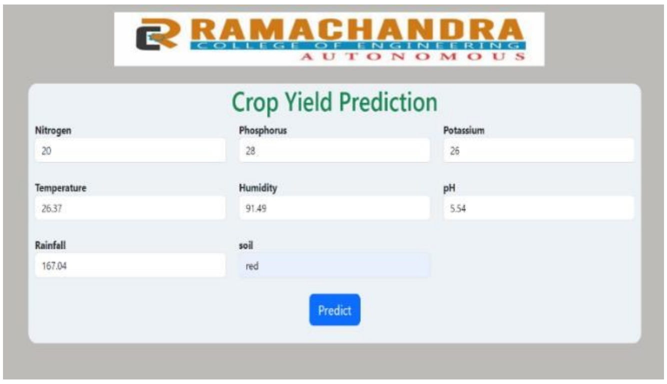


**Fig: Crop Yield Predicition Input Web Page**

The project also has a profound educational impact, raising awareness and promoting the use of modern agricultural technologies. It supports ongoing research and innovation in fields such as agronomy, climate science, and data analytics, fostering a culture of continuous learning and improvement in the agricultural sector.

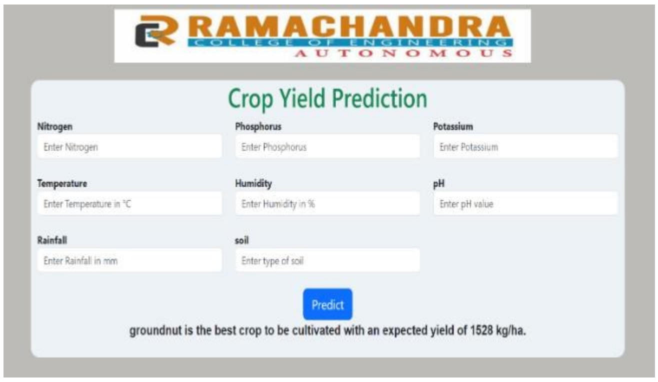


**Fig: Crop Yield Predicition Output Web Page**

Finally, the project emphasizes the importance of continuous monitoring and improvement. Regular updates and recalibrations ensure that the predictive models remain effective and relevant, adapting to evolving agricultural needs and challenges. This iterative process supports the long-term sustainability and usability of the system.

https://doi.org/10.62643/ijerst.2024.v20.i3.pp131-136

ISSN 2319-5991 www.ijerst.com

Vol. 20, Issue 3, 2024

## 4. Conclusion

In Conclusion, the crop yield prediction project underscores the transformative potential of integrating machine learning and data analytics into agricultural practices. The accurate and timely predictions provided by the Random Forest Classifier model facilitate informed decision-making, enabling farmers and policymakers to optimize planting schedules, resource allocation, and crop management. This data-driven approach leads to significant improvements in agricultural productivity and sustainability.

Resource allocation optimization emerges as a critical benefit, reducing the waste of inputs such as water, fertilizers, and pesticides while lowering costs. The model's predictive capabilities allow for efficient use of resources, ultimately enhancing the economic viability of farming operations. By maximizing yields and minimizing costs, farmers can achieve higher profitability and contribute to a more stable agricultural market.

Risk mitigation is another crucial outcome, with predictive models offering early warnings about adverse weather conditions. This proactive approach allows farmers to take preventive measures, safeguarding their crops and ensuring more stable yields. Such stability is essential for maintaining food security, particularly in regions vulnerable to climate variability.

The project also promotes the adoption of data-driven agricultural practices, encouraging the use of advanced technologies to enhance crop selection, soil management, and overall farm management. This emphasis on modern technology and continuous learning fosters innovation in the agricultural sector, driving technological advancements and research.

Economic benefits extend beyond individual farmers to impact the broader agricultural industry and national economies. Predictive analytics help stabilize market prices, reduce volatility, and ensure a reliable supply of agricultural produce, contributing to a more resilient and efficient agricultural sector.

## 5. References

[1] Jain A., "Analysis of growth and instability in the area, production, yield, and price of rice in India", Journal of Social Change and Development, vol. 2, pp.

[2] Sangeeta, Shruthi G. "Design and implementation of crop yield prediction model in " International Journal of Engineering Research & Technology (IJERT), vol. 9, no. 4, pp. 305-310, Apr. 2020.

[3] Johnson LK, Bloom JD, Dunning RD, Gunter CC, Boyette MD, Creamer NG, "Farmer harvest decisions and vegetable loss in primary production. Agricultural Systems", 176, pp. 102672, November 2019.

[4] Sharma A, Jain A, Gupta P, Chowdary V. Machine learning applications for precision agriculture: A comprehensive review. IEEE Access. 2020 Dec 31;9:4843-73.

[5] Meshram V, Patil K, Meshram V, Hanchate D, Ramkteke SD. Machine learning in agriculture domain: A state-of-art survey. Artificial Intelligence in the Life Sciences. 2021 Dec 1;1:100010.

[6] Reddy, D. J., & Kumar, M. R. (2021). Crop Yield Prediction using Machine Learning Algorithm. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). doi:10.1109/iciccs51141.2021.9432236.

[7] S. P. Raja, B. Sawicka, Z. Stamenkovic and G. Mariammal, "Crop prediction based on characteristics of the agricultural environment using various feature selection techniques and classifiers," IEEE Access, vol. 10, pp. 23625-23641, 2022, doi: 10.1109/ACCESS.2022.3154350