iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60788

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Agricultural Decision Support System

Neelima Gurrapu[1], Adupa Vivek[2], Neeripelly Vaibhavi[3], Koleti Mahalaxmi[4], Mohammed Shareeq[5], Thungani Vishnu Sri[6]

[1]*Assistant Professor, Department of Computer Science & Artificial Intelligence, SR University, Warangal, Telangana, India*
[2, 3, 4, 5, 6]*Department of Computer Science & Artificial Intelligence, SR University, Warangal, Telangana, India*

Abstract: ***In India, agriculture stands as the backbone of the economy, yet farmers encounter persistent challenges rooted in conventional farming methods that result in suboptimal yields and environmental degradation. To address these issues, a transformative shift towards precision agriculture is imperative. Our Agricultural Decision Support System (ADSS) represents a pioneering initiative harnessing advanced machine learning algorithms to confront these challenges head-on. By scrutinizing soil characteristics alongside critical weather parameters encompassing temperature, humidity, pH levels, and precipitation, our system furnishes tailored recommendations for crop selection and optimal fertilizer application. Through the implementation of our innovative solution, farmers stand poised to diversify their crop cultivation, augment yields, and pave the way for sustainable agricultural practices. The integration of our system offers farmers the opportunity to explore new crop varieties, potentially leading to heightened profitability and the mitigation of soil pollution concerns.***

Keywords: Precision agriculture, Decision making, Machine learning, crop yield prediction, resource optimization, sustainable farming practices, soil analysis, weather patterns, data-driven recommendations, agricultural innovation

I. INTRODUCTION

In our research paper, we will explore the transformative potential of our Agricultural Decision Support System (ADSS) in revolutionizing India's agricultural landscape. Acknowledging the foundational role of agriculture in India's economy, we recognize the pressing need to enhance productivity, profitability, and sustainability in the sector. Inspired by seminal works in the field, including studies by Sagar et al. [1], Kushwaha et al. [2], and Girish et al. [3], our ADSS integrates cutting-edge machine learning algorithms and extensive datasets to address the multifaceted challenges faced by farmers.

Our ADSS represents a paradigm shift in agricultural decision-making, leveraging insights from soil composition, weather dynamics, and historical yield data to predict crop outcomes with precision. Building upon the pioneering research of Katarya et al. [4] and Iniyan et al. [5], our system not only forecasts crop yields but also recommends suitable crops and fertilizers tailored to each agricultural context. This comprehensive approach, influenced by the groundbreaking work of Shivani et al. [6], Morve et al. [7], and Ramya et al. [8], empowers farmers to make informed decisions regarding crop selection, cultivation practices, and resource management.

Furthermore, our commitment to precision agriculture, as highlighted by the contributions of Kulkarni et al. [9], Shedthi et al. [10], and Krupa Patel [11], underpins the development of our recommendation system. By analyzing intricate soil properties, climatic conditions, and crop requirements, our ADSS offers actionable insights that optimize agricultural practices, foster sustainable growth, and mitigate environmental impact. Our ADSS essentially represents the fusion of agriculture and technology, bringing in a new era of resilience, sustainability, and productivity.

Through informed decision-making in crop yield prediction, crop recommendation, and fertilizer selection, our system empowers farmers to navigate the complexities of modern agriculture with confidence and foresight, clearing the path for improved agricultural prospects in India.

II. LITERATURE REVIEW

Sagar et al. [1] employ Decision Tree and Random Forest models to predict crop yield and offer recommendations, stressing the importance of factors like temperature, rainfall, and area in crop decision-making. They emphasize the necessity of robust algorithms and comprehensive datasets to understand crop yield determinants. Addressing soil exhaustion concerns, they propose a recommendation system based on rainfall, temperature, soil content, and type. Overall, their study integrates machine learning, datasets, and recommendations to enhance yield prediction accuracy and advocate sustainable farming.

Kushwaha et al. [2] introduced a method utilizing big data, including soil and weather data, along with the Hadoop platform and agro algorithms for crop yield prediction. Their approach not only predicts crop yield but also recommends suitable crops based on data repository, thereby enhancing both farmer profitability and agricultural quality.

©IJRASET: All Rights are Reserved

Girish et al. [3] focused on predicting crop yield and rainfall using machine learning methods. Through their study, they compared the efficiency of different algorithms such as linear regression, SVM, KNN, and decision trees. Their findings highlighted SVM's superiority, particularly in rainfall prediction.

Katarya et al. [4] discussed diverse machine learning methods aimed at accelerating crop yield. They emphasized AI techniques and precision agriculture, detailing a hybrid approach for agricultural yield estimation that outperforms existing methods like Decision Tree and Random Forest.

Iniyan et al. [5] presented a hybrid approach for crop yield prediction, employing a Random Forest classifier and the Random Search method. Their methodology yielded superior accuracy and also explored neural network models for specific regions of India, along with implementing the Agro algorithm on the Hadoop platform.

Shivani et al. [6] proposed a machine learning-based approach for crop yield and success rate prediction. Their work involved the development of a crop recommendation system utilizing the Random Forest algorithm. This system aims to provide farmers with informed decisions to optimize crop selection and enhance productivity.

Morve et al. [7] introduced a crop recommendation system utilizing machine learning techniques to offer precise recommendations to farmers. Their system aims to reduce monetary losses by identifying suitable crops for different seasons and improving overall agricultural productivity.

Ramya et al. [8] presented a smart agriculture solution focusing on tracking agricultural fields and assisting farmers in enhancing productivity. Their work emphasizes the importance of sensor technologies in agriculture and forestry, contributing to technological advancements in the field.

Kulkarni et al. [9] proposed a supervised learning approach employing the Random Forest algorithm for crop recommendation. Their technique is designed to help farmers choose crops, seed types, and amounts of fertilizers that will be utilized during cultivation. The technology improves agricultural outcomes by offering reliable and useful recommendations through the integration of machine learning techniques.

Shedthi et al. [10] introduced a crop and nutrient recommendation system tailored for precision agriculture. The system provides individualized crop suggestions by utilizing machine learning algorithms to analyze soil parameters, including pH levels, organic matter content, and nutrient levels. This method contributes to sustainable agriculture practices by helping with nutrient management and crop selection optimization.

Krupa Patel [11] introduced the AgriRec algorithm, tailored for precision agriculture. By leveraging machine learning techniques, AgriRec predicts crops for different seasons and recommends suitable fertilizers based on soil type, water requirement, and temperature conditions. This personalized approach not only optimizes crop selection but also aids in fertilizer management, contributing to sustainable agricultural practices.

Nihar Patel et al. [12] focused on predicting crop yields using machine learning algorithms. Utilizing crop data from six states in India, they identified the Extra Trees Regressor as the best-performing model. This approach provides a practical implementation for estimating crop yields, which can be beneficial for farmers, governments, economists, and banks, thus contributing to improved agricultural productivity and food security.

Thombare et al. [13] addressed the critical need for agricultural yield forecasting and fertilizer suggestion amidst climate change. Their research deployed machine learning technologies such as linear clustering, Artificial Neural Networks, and Decision Making to improve agricultural production prediction. The proposed technique has been rigorously assessed for errors, demonstrating promising operational outcomes and contributing to sustainable agricultural practices.

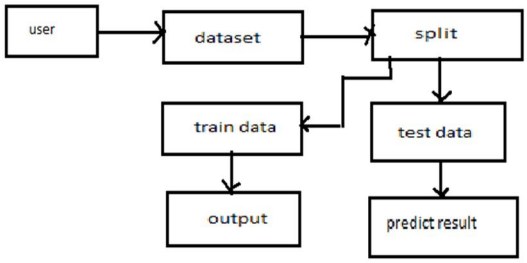
Fiadjoe [14] conducted a study focusing on predicting agricultural yield in Karnataka state using neural network regression. The dataset encompassed parameters such as agricultural area, crop, taluka, year, season, district-wise rainfall, and temperature. Employing a Multilayer Perceptron Neural Network with ReLu Activation and Adam Optimizer, trained for 50 epochs with a batch size of 200, the model achieved a notable test accuracy of 96.43%. Fiadjoe (2023) also compared it with other regression algorithms, demonstrating its superior performance in mean absolute error and accuracy.

III. PROPOSED METHODOLOGY

1) *Data Collection and Integration:* The initial phase involves collecting a variety of datasets from multiple sources, such as historical crop yield records, weather observations, agricultural research organizations, and soil databases. These datasets include vital information about crop performance, weather patterns, soil composition, and other relevant factors affecting agricultural results. Strict quality controls and integration ensure the completeness and consistency of the data.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

2) *Data Preprocessing and Cleaning:* The gathered datasets are then performed preprocessing and cleaning in order to address missing values, outliers, and inconsistencies. To guarantee the quality and integrity of data, procedures like imputation, outlier identification, and normalizing are used. In order to prepare the data for additional analysis and model building, this step is essential.
3) *Feature Selection and Engineering:* This stage focuses on identifying important features influencing crop growth and yield through domain knowledge and statistical analysis. Feature engineering techniques transform raw data into informative features capturing underlying patterns and relationships. This aids in selecting the most relevant variables for model development.
4) *Model Development and Training:* The core of the methodology involves developing predictive models using supervised learning algorithms such as regression and classification techniques. These models are trained on preprocessed data to predict crop yields, recommend suitable crops, and optimize fertilizer usage. Multiple models are evaluated to identify the most accurate and reliable ones for integration into the ADSS.
5) *Integration and Deployment:* Following model development, they are integrated into the ADSS platform, which provides a user-friendly interface for farmers to access recommendations and insights. This integration ensures seamless deployment across different devices and environments, with provisions for web and mobile applications. User feedback and testing refine the system before full-scale deployment. Evaluation and Validation: The performance of the ADSS is rigorously evaluated and validated using real-world data and field trials. Model predictions are compared against ground-truth observations to assess accuracy, precision, and reliability. Cross-validation techniques ensure the robustness and generalization of models across different regions and cropping systems. Continuous monitoring and feedback loops iteratively improve the system's performance.
6) *Scaling and Sustainability:* Finally, the ADSS is designed to scale seamlessly to accommodate growing data volumes and user base. Cloud-based infrastructure and distributed computing technologies ensure scalability and sustainability. Regular maintenance and updates address evolving user needs and technological advancements, ensuring the long-term viability and effectiveness of the system.



IV. DATASETS

A. *Crop Yield Dataset*

1) Introduction

The crop yield dataset provides information on the historical yields of various crops cultivated in different agricultural regions. It serves as a valuable resource for understanding crop performance and identifying factors influencing yield variability.

2) Columns:

a) Crop Type: Categorical variable indicating the type of crop harvested.
b) Region: Geographic region or location where the crop was cultivated.
c) Season: Season of harvest.
d) Yield: Yield of the crop, measured in units per hectare or other relevant metrics.
e) Yield Result: The results of yield whether it is good or bad.

3) Data Summary

- This dataset offers insights into crop yield trends over time and across different regions.
- It provides a historical perspective on crop performance, enabling the identification of factors contributing to yield fluctuations.
- Weather conditions may be included as supplementary variables to assess their impact on crop productivity.

4) Data Preprocessing

- Appropriate handling of missing values was done in order to assure data integrity.
- Numerical features were scaled as needed for modeling, and categorical variables were encoded.
- The dataset's dependability was improved by identifying and addressing anomalies and outliers.

5) Exploratory Data Analysis (EDA)

- EDA techniques were applied to analyze yield distributions, identify trends over time, and explore regional variations in crop productivity.
- Correlation analyss may be conducted to assess the relationships between weather variables and crop yields.

B. Crop Recommendation Dataset

1) Introduction

This dataset encompasses various environmental and soil parameters crucial for agricultural decision-making. It is derived from on-site measurements and laboratory analyses conducted at agricultural fields.

2) Columns

a) Temperature: Recorded temperature at the agricultural site, measured in Celsius.
b) Humidity: Humidity level at the agricultural site, expressed as a percentage.
c) Moisture: Soil moisture content, indicating the water content in the soil, measured as a percentage.
d) Soil Type: Categorical variable representing the type of soil present at the agricultural site. Examples include sandy loam, clay, silt loam, etc.
e) Crop Type: Categorical variable denoting the type of crop cultivated in the agricultural field. Examples include wheat, rice, maize, soybean, etc
f) Nitrogen: Nitrogen content in the soil, measured in parts per million (ppm).
g) Potassium: Potassium content in the soil, measured in ppm.
h) Phosphorous: Phosphorous content in the soil, measured in ppm.

3) Data Summary

- This dataset provides a comprehensive overview of the environmental and soil conditions prevalent in agricultural fields.
- It includes both numerical and categorical variables, allowing for a detailed analysis of the factors influencing crop growth and fertilizer recommendations.
- The dataset exhibits variations in temperature, humidity, soil type, and nutrient content, reflecting the diverse agricultural landscapes.

4) Data Preprocessing

- Missing values were addressed using suitable imputation techniques, ensuring data completeness.
- Categorical variables were encoded to numerical format for compatibility with machine learning algorithms.
- The dataset was split into training and testing sets to facilitate model development and evaluation.

5) Exploratory Data Analysis (EDA)

- EDA techniques such as histograms, box plots, and correlation analyses were employed to explore the distribution and relationships between variables.
- Insights gained from EDA informed the development of predictive models for fertilizer recommendation and crop management.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

C. Fertilizer Recommendation Dataset

1) Introduction

The fertilizer recommendation dataset contains information on recommended fertilizers for different crops and soil types. It serves as a guide for optimizing nutrient management practices and enhancing crop yield and quality.

2) Columns

a) Crop Type: Categorical variable representing the type of crop for which the fertilizer recommendation is provided.

b) Soil Type: Categorical variable indicating the type of soil for which the fertilizer recommendation is applicable.

c) Moisture: Soil moisture content, indicating the water content in the soil, measured as a percentage.

d) Nutrient Composition: Composition of the recommended fertilizer, specifying the nutrient content (e.g., nitrogen, phosphorus, potassium) and application rates.

3) Data Summary

- This dataset facilitates informed decision-making regarding fertilizer selection and application.
- It provides tailored recommendations based on specific crop and soil characteristics, promoting efficient nutrient management practices.
- Nutrient composition details enable users to understand the nutritional requirements of crops and select appropriate fertilizers accordingly.

4) Data Preprocessing

- The dataset underwent preprocessing steps to handle missing values, encode categorical variables, and ensure data consistency.
- Fertilizer recommendations were standardized and categorized to facilitate analysis and application.

5) Exploratory Data Analysis (EDA)

EDA techniques were employed to examine fertilizer usage patterns, assess the prevalence of different fertilizer types across crops and soil types, and identify potential correlations between fertilizer composition and crop performance.

D. Algorithms

In order to solve our prediction difficulties we deploy several different kinds of Machine Learning models, all of which are distinct Classification models. These models are as follows:

- Decision Tree
- Random Forest
- Logistic Regression
- Gradient Boosting
- Support Vector Machine
- K Nearest Neighbor

1) *Decision Tree:* A DT serves as a supervised algorithm suitable for working both bracket and retrogression problems, although it's primarily favored for diving bracket tasks. This classifier adopts a tree structure where internal bumps correspond to dataset features, branches represent decision rules, and each splint knot signifies an outgrowth. In the DT, two crucial bumps are present the Decision knot and the Leaf Node. Decision bumps grease decision- making and give rise to multiple branches, while splint bumps represent the results without any farther branching. opinions or tests are executed grounded on the characteristics within the handed dataset.

2) *Random Forest:* It's an ensemble literacy system, a conception embedded in the idea that multiple models working together can produce further accurate results than a single model. The RF algorithm is a testament to this principle, combining the strength of individual DT to deliver robust and dependable prognostications. This essay delves into the armature, performing, and operations of the RF algorithm, pressing its significance in colorful disciplines. At its core, the RF algorithm is a collection of DT.

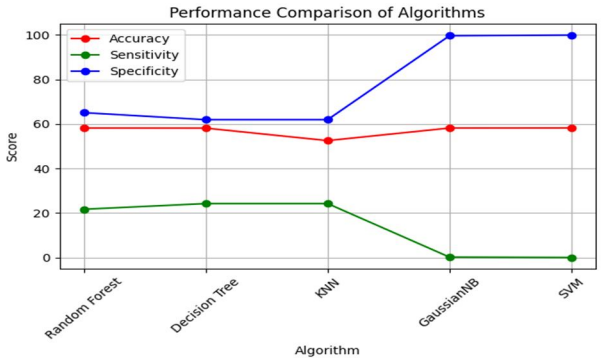
©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

3) *Logistic Regression:* Logistic retrogression is a statistical system that falls under the marquee of generalized direct models(GLMs). The logistic retrogression model begins by establishing a direct relationship between the independent variables(features) and the log- odds of the probability of an event being. The logistic function latterly maps this direct combination to the probability range. The model is trained using labeled data, and the thing is to learn the portions of the direct equation that stylish separates the classes. Once trained, logistic retrogression can prognosticate the probability of an outgrowth belonging to a specific class and classify it grounded on a predefined threshold.
4) *Gradient Boosting:* Grade Boosting is an ensemble literacy system, and like its counterpart, RF, it capitalizes on the conception that combining multiple models can produce superior prognostications. still, grade Boosting follows a different approach. It successionally constructs an ensemble of weak learners, generally DT, to ameliorate prophetic accuracy. GBDT finds operations across different disciplines Web Hunt Ranking It's necessary in ranking hunt machine results by prognosticating stoner click-through rates. Anomaly Discovery In cybersecurity and fraud discovery, it identifies unusual patterns in data. Healthcare Used for complaint opinion and prognosticating patient issues.
5) *Support Vector Machine:* SVMs are deeply embedded in fine optimization and direct algebra. SVMs have left a profound impact on multitudinous fields, including Image Recognition SVMs are necessary in classifying objects within images, face recognition, and indeed handwritten number recognition. Text Bracket They're employed in spam dispatch filtering, sentiment analysis, and document categorization. Biomedical lore's SVMs help in prognosticating complaint issues, protein structure vaticination, and gene bracket. Finance They play a part in stock request vaticination, credit scoring, and fraud discovery. Speech Recognition SVMs contribute to speech- to- textbook conversion and speaker identification.
6) *K Nearest Neighbor:* KNN is an case- grounded ,non-parametric algorithm. It operates on the principle that objects with analogous characteristics tend to be close to each other in the point space. Chancing the optimal' k' is frequently a matter of trial and error and sphere-specific knowledge. Distance Metrics Beyond Euclidean utmost introductory accoutrements on KNN emphasize the use of the Euclidean distance metric. still, KNN is incredibly flexible when it comes to distance criteria . Experimenters and interpreters constantly acclimatize and design custom distance criteria acclimatized to the specific characteristics of their datasets.
7) *Gaussian Naive Bayes:* Gaussian Naive Bayes (GNB) is a classic classification technique known for its simplicity and efficiency. It operates under the assumption that features are independent and follow a Gaussian (normal) distribution. By calculating the probabilities of each class given the feature values of a data point, GNB makes predictions based on the class with the highest probability. This algorithm finds applications in diverse domains like text categorization, medical diagnosis, fraud detection, customer segmentation, and image recognition. Its straightforward nature and computational efficiency make it a popular choice, especially for tasks with limited computational resources and relatively small datasets.

V. COMPARISON AND RESULTS

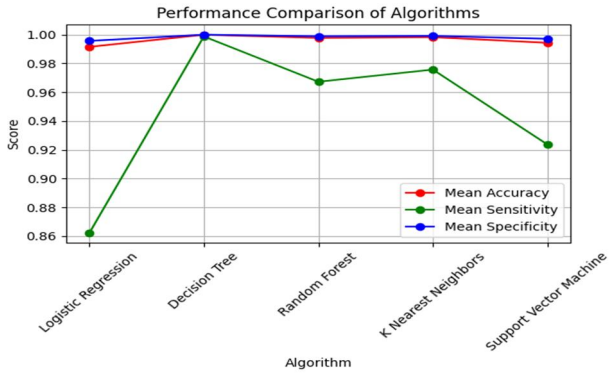
A. *Crop Yield Prediction*

Model	Accuracy
Random Forest	0.8067
Decision Tree	0.8016
K-Nearest Neighbors	0.7880
-Gaussian Naive Bayes	0.8726
SVM	0.8731



B. Crop Recommendation System

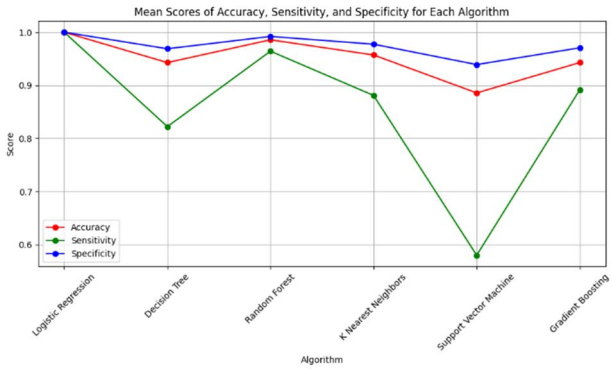
Model	Accuracy
Random Forest	0.9995
Decision Tree	1.0000
K-Nearest Neighbors	0.9620
Support Vector Machine	0.9987
Logistic Regression	0.8653



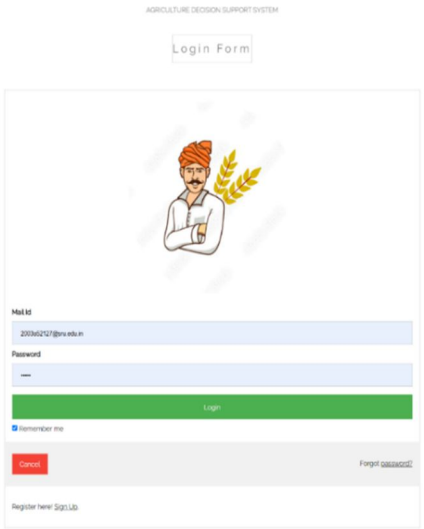
C. Fertilizer Recommendation System

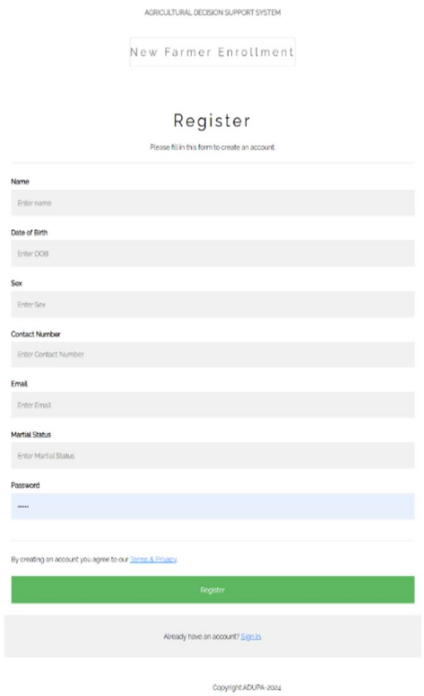
Model	Accuracy
Random Forest	0.9857
Decision Tree	0.9482
K-Nearest Neighbors	0.9571
Support Vector Machine	0.8857
Logistic Regression	1.0000
Gradient Boosting	0.9428

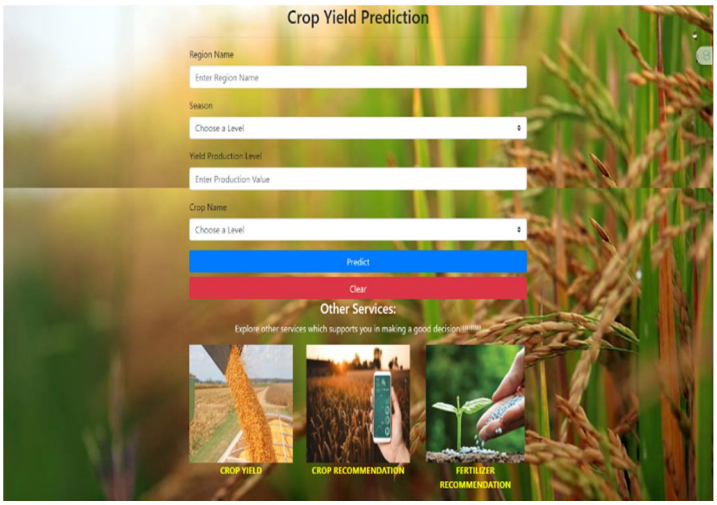
©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

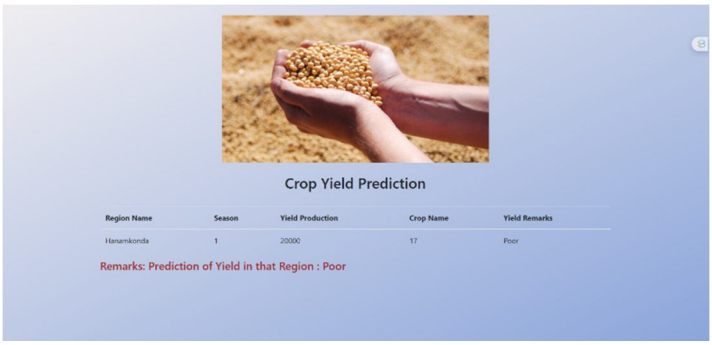


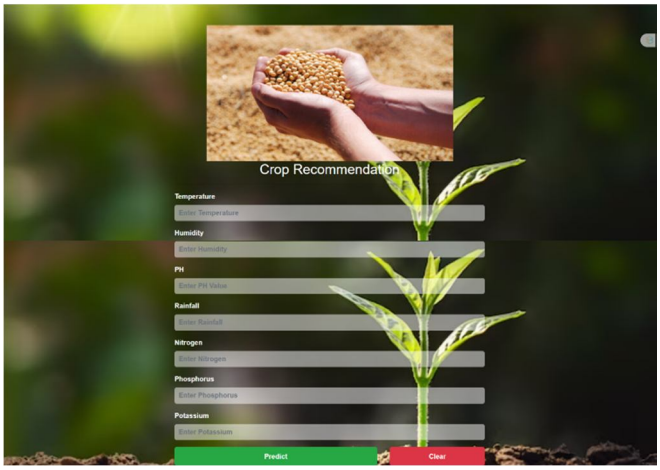
D. Interface Images

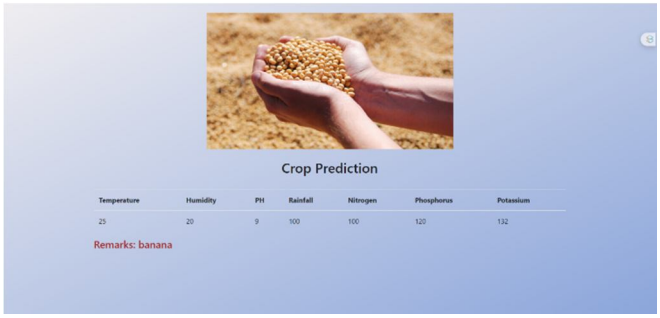




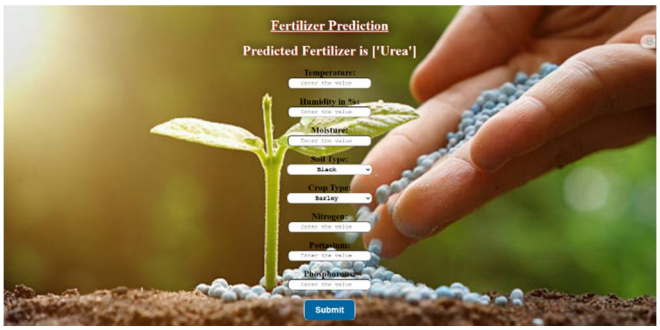








©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



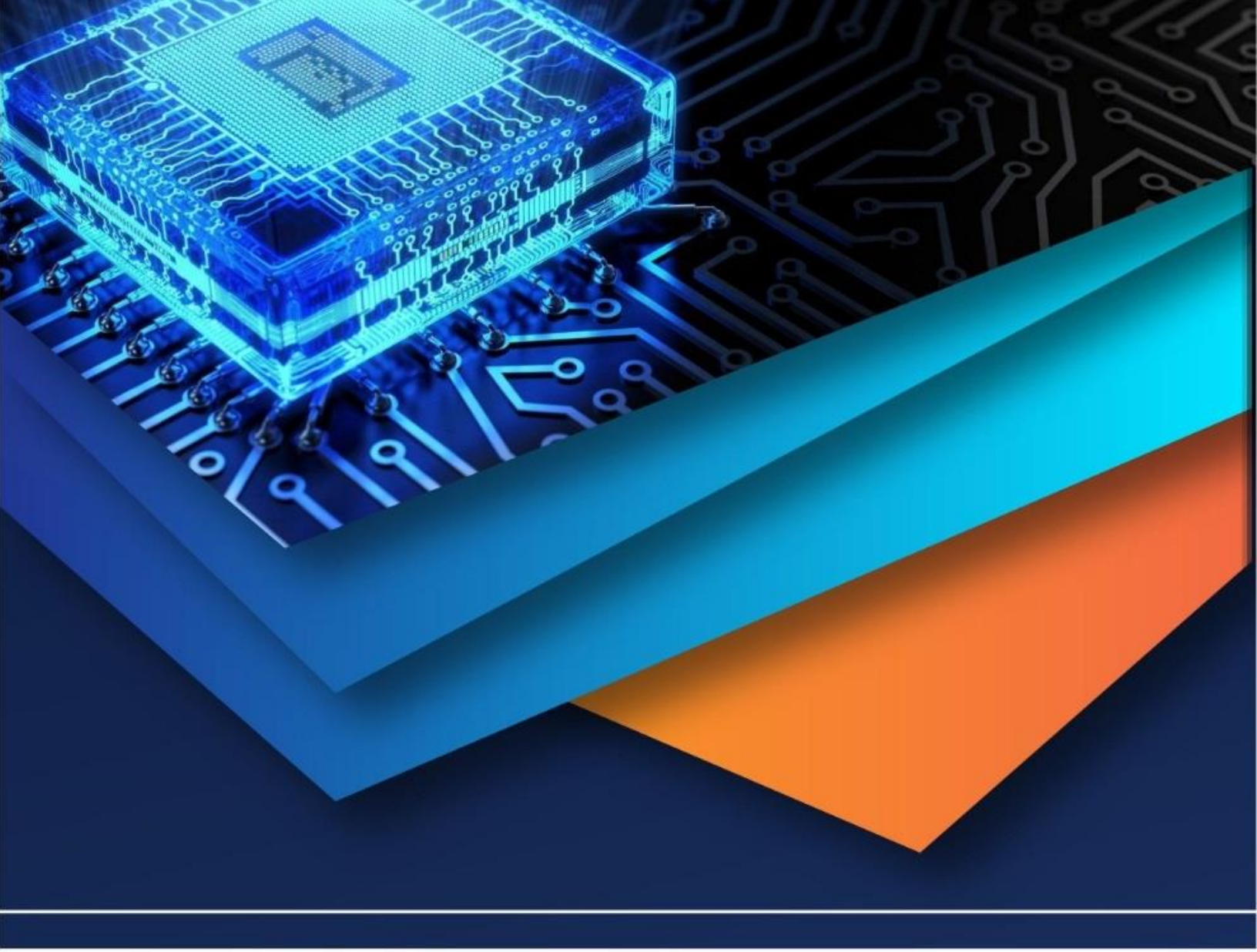
E. Experimental Outcome

In our project, the system has been designed to recommend the most suitable crop for a particular land based on key parameters such as annual rainfall, temperature, humidity, and soil pH. The system utilizes machine learning algorithms, particularly Support Vector Machine (SVM), to predict annual rainfall based on historical data. Additionally, users input data related to temperature, humidity, and soil pH to generate personalized crop recommendations. Furthermore, the system allows users to input NPK (Nitrogen, Phosphorus, Potassium) values, which are crucial nutrients for crop growth. Based on this input, the system provides recommendations for the required NPK fertilizers tailored to the recommended crop.

REFERENCES

[1] Raskar, S., Abhang, A., Pethe, N., & Attar, V.Z. (2022). Crop Yield Prediction and Recommendation System. *2022 IEEE Pune Section International Conference (PuneCon)*, 1-6.
[2] Kushwaha, A.K. (2015). Crop yield prediction using Agro Algorithm in Hadoop.
[3] Girish L, Gangadhar S, Bharath T R, Balaji K S, Abhishek K T (2018)"Crop Yield and Rainfall Prediction in Tumakuru District using Machine Learning".
[4] Katarya, R., Raturi, A., Mehndiratta, A., & Thapper, A. (2020). Impact of Machine Learning Techniques in Precision Agriculture. *2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE)*, 1-6
[5] Iniyan, S., Varma, V.A., & Naidu, C.T. (2023). Crop yield prediction using machine learning techniques. *Adv. Eng. Softw., 175*, 103326.
[6] Shivani, S., Kale., Preeti, Patil. (2019). A Machine Learning Approach to Predict Crop Yield and Success Rate. doi: 10.1109/PUNECON46936.2019.9105741
[7] Morve, P., Khandagle, K., Donurkar, V., SameershaFakeerProf.N., G., & Bhojne (2023). CROP RECOMMENDATION SYSTEM USING MACHINE LEARNING. *International Research Journal of Modernization in Engineering Technology and Science*.
[8] Ramya, J., Scholar, M., & Sankar, D.M. (2023). Crop Recommendation Using Supervised Learning Techniques. *2023 4th International Conference on Smart Electronics and Communication (ICOSEC)*, 1076-1084.
[9] Kulkarni, N., Srinivasan, G., Sagar, B., & Cauvery, N.K. (2018). Improving Crop Productivity Through A Crop Recommendation System Using Ensembling Technique. *2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS)*, 114-119.
[10] Shedthi, B.S., Shetty, V., Anusha, Shetty, R.R., Shetty, A., & Alva, B.D. (2022). Crop and Nutrient Recommendation System using Machine Learning for Precision Agriculture. *2022 International Conference on Artificial Intelligence and Data Engineering (AIDE)*, 230-235.
[11] Krupa, Patel. (2023). Multi-criteria Agriculture Recommendation System using Machine Learning for Crop and Fertilizesrs Prediction.. Current agriculture research journal, 11(1):137-149. doi: 10.12944/carj.11.1.12
[12] Nihar, Patel., Deep, Patel., Samir, B., Patel., Vibha, Patel. (2020). Crop Yield Estimation Using Machine Learning. 328-342. doi: 10.1007/978-981-16-0708-0_27
[13] Thombare, Shubhangi, Radhakisan., Zaware, Payal, Eknath., Dongare, Snehal, Anil., Kasar, Yogesh, Shivaji. (2022). Farmer's Friend: An AI Model to Predict Crop Yield and Provide Suggestion. International Journal of Advanced Research in Science, Communication and Technology, 124-130. doi: 10.48175/ijarsct-402
[14] John, Fiadjoe. (2023). Crop Yield Prediction Using Deep Learning. 93-102. doi: 10.1007/978-981-19-5936-3_9

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)