Crop Recommendation System with XAI

K. Kiran, V. Pavan Kumar, K. Chakri Pavan Sesh Kumar, K. Harsha Vardhan
students, dept of CSM, Raghu Engineering College, dakamarri, Bheemunipatnam Visakhapatnam Dist. pin code: 531162
Mr. K.L Ganapathi Reddy (PH.D), Associate Professor,
dept of CSE, Raghu engineering college, dakamarri(V), Bheemunipatnam Visakhapatnam Dist. pin code:531162
***DR. Akula Siva Prasad, lecturer, Dr. V.S. Krishna** Government Degree College, Vishakhapatnam*

Abstract: Modern agriculture is increasingly leveraging advanced technologies like XAI for improving crop recommendation systems. This paper presents a unique approach that integrates XAI methodologies into crop recommendation frameworks to improve transparency and interpretability. By harnessing machine learning models such as classification & regression techniques and ensemble techniques alongside XAI explanations, our system offers personalized crop suggestions based on environmental and geographic data.
Our study contributes to the agricultural sector by enhancing the transparency of crop recommendation systems using XAI, allowing farmers to grasp the underlying factors influencing each recommendation.

We investigate various machine learning algorithms and their interpretability within the context of crop selection, emphasizing user-friendly design and actionable insights. This research aims to advance agricultural decision-making by advocating for the adoption of XAI-driven crop recommendation systems, empowering farmers with accessible and understandable information to optimize crop yield and sustainability. This approach not only supports informed decision-making but also fosters trust and acceptance of AI-driven solutions in agriculture.

Keywords Crop recommendation, ML algorithms, XAI, Agriculture, Decision support, Precision farming

I. INTRODUCTION

The agricultural sector is fundamental to global food security and economic stability, with farmers facing complex decisions when selecting which crops to cultivate based on environmental and geographical factors. Traditionally, farmers have relied on personal experience and local knowledge for crop selection, but this approach may not always optimize productivity or resource utilization.
In recent years, the integration of machine learning & AI has revolutionized agriculture by enabling data-driven crop recommendation systems. However, one significant challenge with AI models is their inherent complexity and lack of transparency, which can hinder farmers' trust and understanding of the recommendations provided.

Explainable AI has surfaced as a remedy to tackle this issue by amplifying the interpretability of AI models. By incorporating XAI techniques into crop recommendation systems, we can provide farmers with transparent and interpretable explanations for the AI-generated recommendations. This enhances trust and empowers farmers to make decisions to their specific environmental conditions and preferences.

In this paper, we explore the intersection of XAI methodologies with crop recommendation systems, examining various machine learning algorithms and their interpretability within the context of crop selection. Our goal is to demonstrate the potential of XAI in transforming crop recommendation practices, promoting sustainability, and supporting farmers in making optimal agricultural decisions based on transparent and understandable AI insights.

II. LITERATURE REVIEW

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Exploring the literature on XAI reveals a burgeoning field that has gained significant attention in recent years. Here's a concise review of some key studies:

1. "Explainable AI: Guidotti et al. (2018) conducted a comprehensive review covering various XAI techniques, ranging from rule-based methods to local interpretable model-agnostic explanations (LIME). Their study emphasizes the significance of explainability across diverse domains, including healthcare, finance, and cyber security.

2. "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI" by Adadi and Berrada (2018): This paper presents a detailed taxonomy of XAI techniques and categorizes them based on their interpretability levels, model-agnostic vs. model-specific approaches, and intrinsic vs. post-hoc explanations. It also discusses the opportunities and challenges associated with implementing XAI in real-world applications.

3. "Explainable AI for Trees: From Local Explanations to Global Understanding" by Craven et al. (2020): Focusing on decision tree-based models, this study introduces an XAI framework that provides both local and global explanations. It proposes methods for extracting human-understandable rules from decision trees and visualizing feature importance.

4.. "Interpretable Machine Learning: A Guide for Making Black Box Models Explainable" by Molnar (2019): This book offers a comprehensive overview of interpretable machine learning methods, including LIME, PDP plots etc. It provides practical guidance on how to use these techniques to interpret and debug complex machine learning models.

5. "Towards Transparent AI Systems: Interpreting Visual Question Answering Models" by Zhou et al. (2018): This study introduces an XAI framework tailored for interpreting Visual Question Answering (VQA) models, particularly in computer vision tasks. It introduces an attention-based explanation method and demonstrates its effectiveness in providing insights into model decision-making processes.

These studies highlight the growing interest in XAI and its potential applications across various domains. They underscore the importance of developing transparent and interpretable AI systems to enhance trust, accountability, and user understanding.

III. DATA SET DESCRIPTION

Precision agriculture facilitates informed decision-making for farming strategies. This dataset enables users to construct predictive models for recommending optimal crops based on diverse parameters. Compiled by augmenting datasets on rainfall, climate, and fertilizer data specific to India, it includes key fields:

- N: Nitrogen
- P: Phosphorous
- K: Potassium
- Temperature
- Humidity
- pH: pH value
- Rainfall: Rainfall

	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	rice

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

IV. PROPOSED SYSTEM

The proposed "CROPX" system represents an innovative application of Explainable Artificial Intelligence (XAI) to improve the transparency and interpretability of Crop Recommendation Systems. This approach encompasses five distinct phases aimed at providing clear insights into the recommendation process.

1. Data Collection and Preprocessing: The data preprocessing module assumes a pivotal role, orchestrating the acquisition, cleansing, and transformation of the collected data. Leveraging an array of cutting-edge technologies such as sensors, cameras, and Internet of Things (IoT) devices, this module captures real-time data pertaining to soil type, weather conditions, and historical crop yields. Once the data is collected, it undergoes a series of steps to ensure its quality and usability.

Firstly, the module gathers input data from diverse sources relevant to agriculture, including soil databases, weather monitoring stations, and historical crop databases. Next, the collected data undergoes thorough cleaning to eliminate duplicates and handle missing values. This step is essential to ensure the accuracy and reliability of the dataset used for analysis. Following data cleaning, the module transforms the dataset into a suitable format for analysis. This transformation involves converting categorical variables into numerical representations, which are easier for machine learning algorithms to interpret.

Data integration is another key aspect where different datasets from multiple sources are combined into a cohesive dataset that provides a comprehensive view of the agricultural environment. To ensure uniformity and consistency in the dataset, data normalization techniques are applied. This process standardizes the range of values across different variables, preventing biases in the model training process.

Finally, the dataset is split into two subsets. One is train dataset and another one test dataset. The train dataset is used to build and optimize the crop recommendation model, while the test dataset evaluates its performance and generalization ability.

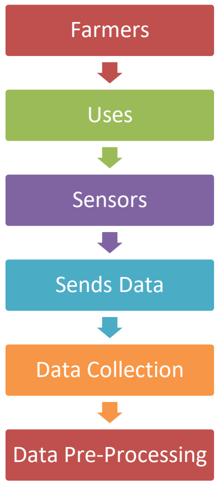

Fig1

2. Exploratory Data Analysis (EDA):

Performing exploratory data analysis (EDA) is essential to comprehend the variable distributions, detect correlations, and extract insights regarding the interrelationships among diverse features and crop yields. Visualize the data using plots and charts to uncover patterns and anomalies.

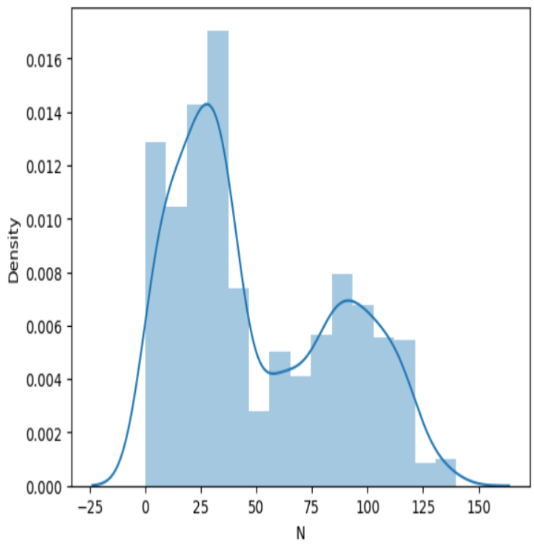

Fig2

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

3. Model Training: Model training is a pivotal phase in the development of any machine learning system, including crop recommendation with XAI. This process involves feeding the algorithm with labeled train data to enable it to learn patterns and relationships within the data. Each machine learning model, such as logistic regression, naïve bayes, decision trees, or support vector machines, employs distinct algorithms and techniques during training to optimize its performance. For instance, logistic regression utilizes gradient descent to minimize the error between predicted and actual outcomes, iteratively adjusting model parameters to improve accuracy.

In the model training phase, we employ a diverse set of machine learning models tailored for classification tasks. These models are instrumental in predicting the most suitable crops based on various input parameters.

Machine Learning Models with their Accuracies:

Logistic_with accuracy : 0.9636363636363636
naive_bayes with accuracy : 0.9954545454545455
KNN with accuracy : 0.9590909090909091
SVM with accuracy : 0.9681818181818181
DecisionTree with accuracy : 0.9818181818181818
RandomForest with accuracy : 0.9931818181818182
Bagging with accuracy : 0.9863636363636363
Adaboost with accuracy : 0.1409090909090909
GradientBoosting with accuracy : 0.981818181818

It's evident that most models perform well, with Naive Bayes and Random Forest achieving particularly high accuracies. However, Adaboost seems to perform significantly worse compared to the other models. It's important to further analyze the performance metrics and possibly fine-tune the models to improve overall performance. Furthermore, it's advisable to incorporate methods like cross-validation to ensure more robust evaluations of model efficiency. Among the models listed, the "naive_bayes" model stands out with the highest accuracy, potentially making it a prime candidate for integration into the CROPX system. Nevertheless, crucial factors like model interpretability, computational demands, and the contextual needs of the application warrant careful consideration before arriving at a conclusive decision.

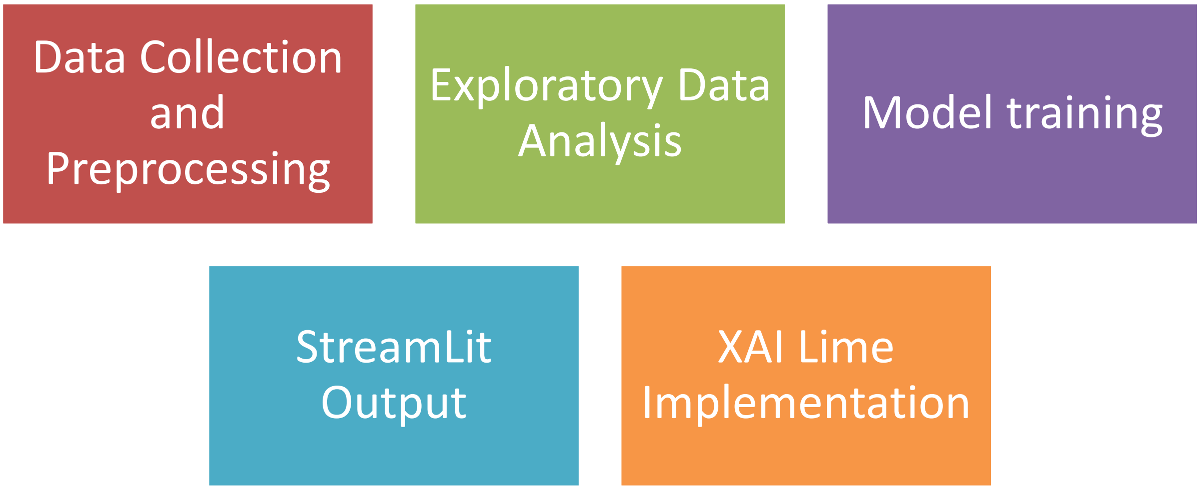

Fig3

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

4. XAI LIME Implementation : LIME is a powerful technique in ML interpretability, providing insights into the decisions made by complex models on individual instances. LIME's core objective is to provide local interpretability by approximating the behavior of opaque models concerning individual predictions. This is executed through training an interpretable model, typically a linear one, utilizing perturbed samples generated around the specific instance of interest. These perturbations introduce small changes to the input features and observe the corresponding changes in predictions, enabling LIME to discern which features are most influential in a particular prediction.

Model Agnosticism: LIME offers local interpretability for machine learning models, irrespective of their complexity or training algorithm. Its agnostic nature enables application across diverse models, including deep learning and ensemble methods, without requiring modifications to the underlying model structure.

Local Interpretability: LIME prioritizes explaining individual predictions over understanding global model behavior. It achieves this by approximating the model's decision boundary around a specific instance using a simpler, interpretable model like linear regression.

Simplicity and Transparency: LIME produces intuitive and human-understandable explanations by representing complex model behaviors using simpler, transparent models. This transparency enhances trust and interpretability, crucial for stakeholders who may not be machine learning experts but need to understand model predictions.

Quantitative Explanations: LIME quantifies the significance of individual input features by allotting weights, to explain their impact on the model's prediction. at the local level.

This feature importance ranking helps prioritize and understand which features are most influential for specific predictions.

Practical Application: LIME enjoys widespread adoption across diverse domains like healthcare, finance, and natural language processing, where it serves as a pivotal tool for interpreting complex black-box models and enhance decision-making processes. Its practical utility in real-world applications underscores its significance in the field of XAI.

Open-Source Implementation: LIME is available as an open-source library in Python, making it accessible to researchers and practitioners for experimentation, customization, and integration into existing workflows. This availability fosters collaboration and encourages further development of XAI techniques.

LIME explainer can be set up using two main steps: (1) import the lime module, and (2) fit the explainer using the training data and the targets. During this phase, the mode is set to classification, which corresponds to the task being performed.

LIME Explainations:

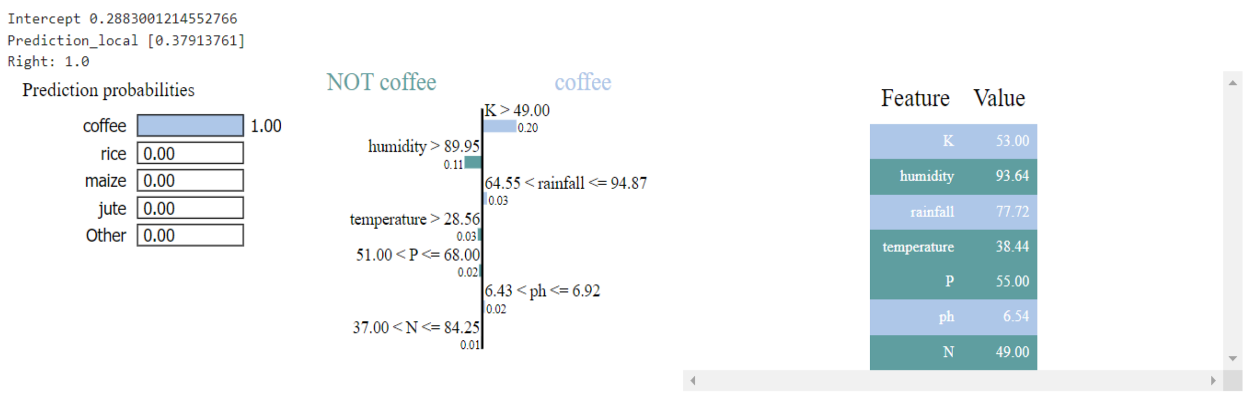

Fig4

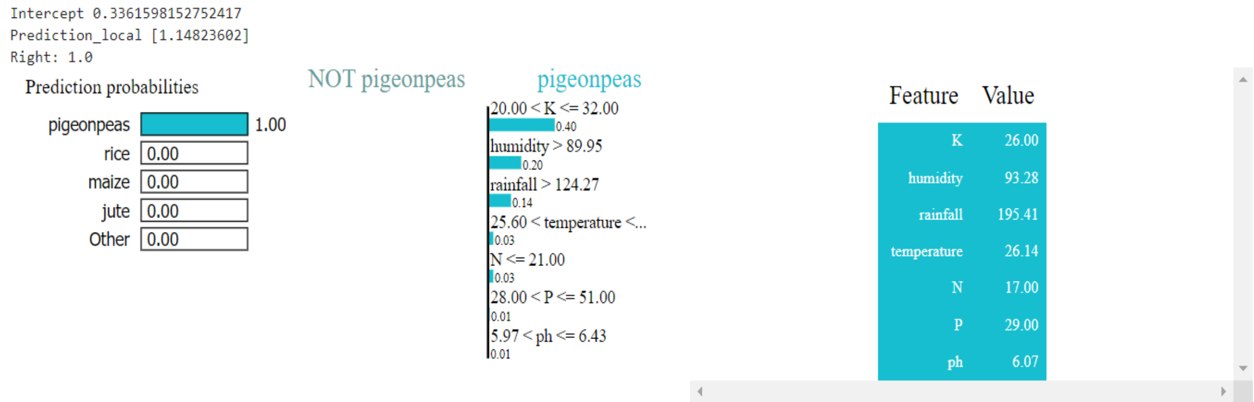

Fig5

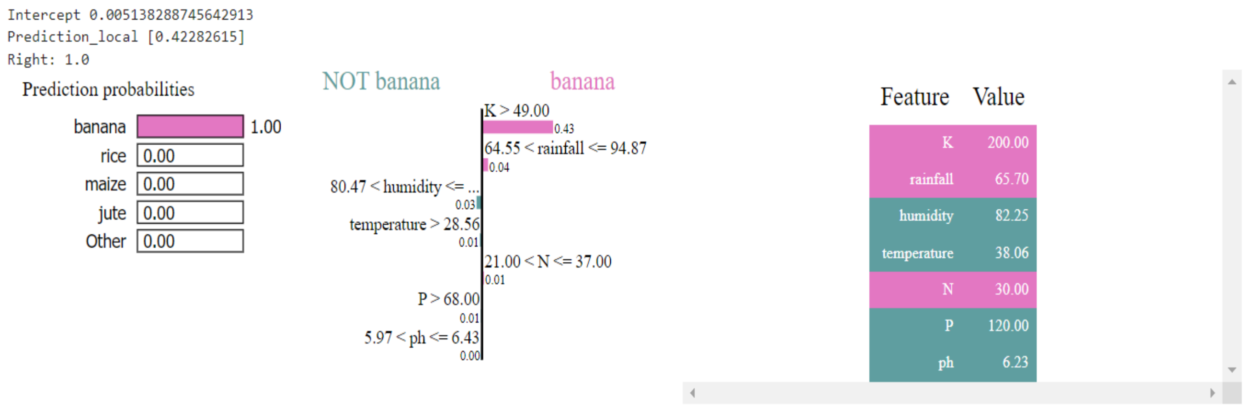

Fig6

Above the visualizations, there are three key values:
Right value: This represents the prediction made by our model for the given test vector.
Prediction_local value: It indicates the value generated by a linear model trained on perturbed samples, obtained by sampling around the test vector. This model uses only the top k features outputted by LIME.
Intercept value: This signifies the constant component of the prediction provided by the linear model for the given test vector.
Regarding the visualizations, in Fig4, the colors dark blue and light blue represent negative and positive associations of features, respectively. These associations aid in predicting that coffee is suitable based on the feature values.

5. Output Analysis:

Streamlit is a powerful and user-friendly Python library used for building interactive web applications for machine learning, data visualization, and other data-intensive tasks. It allows developers to create web-based interfaces directly from Python scripts, making it easier to share and demonstrate data-driven applications without extensive web development experience.

In this project, Streamlit is used to make the web based interface for the crop recommendation system with XAI. The model recommends coffee as the best crop to be cultivated based on certain input features in the below shown fig7 which is streamlit web based userinterface.

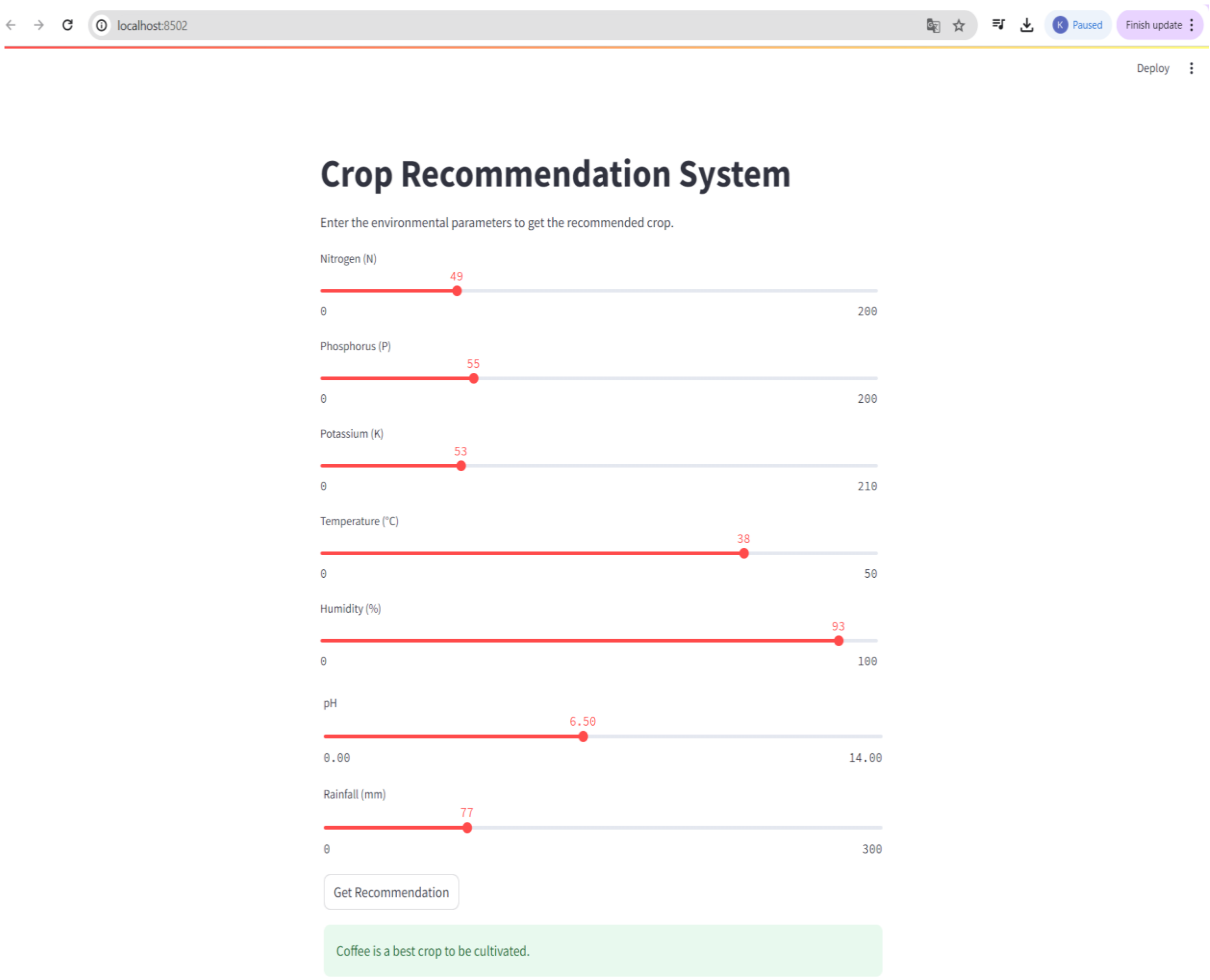

Fig7

5.

CONCLUSION:

This project underscores the pivotal role of explainable artificial intelligence (XAI) in revolutionizing crop recommendation systems. By harnessing a wealth of data encompassing soil composition, climate parameters, and rainfall patterns, the XAI model provides transparent and interpretable insights crucial for optimizing crop selection strategies. Such transparency not only enhances decision-making processes in precision agriculture but also empowers farmers with actionable recommendations tailored to their specific needs and environmental conditions. Moreover, the integration of XAI fosters a deeper understanding of the underlying factors driving crop suitability, paving the way for more informed and sustainable agricultural practices.

The implementation of XAI in crop recommendation systems holds immense promise for the agricultural sector, offering transformative benefits that extend beyond conventional approaches. By leveraging advanced analytics and machine learning algorithms, the system can analyze complex datasets and identify intricate patterns that may elude traditional methodologies. This comprehensive approach empowers farmers to make informed decisions based on data, leading to optimized crop yields, minimized resource wastage, and improved resilience against environmental factors.

REFERENCES:

1. Rashi Agarwal, "Crop Recommendation System Using Machine Learning Algorithms", International Journal of Innovative Technology and Exploring Engineering (IJITEE), Volume 9, Issue 2, December 2019.

2. Priyadharshini A, "Machine Learning Techniques for Crop Recommendation: A Review", International Journal of Computer Applications (IJCA), Volume 182, Issue 27, February 2019.

3. Shilpa Mangesh Pande, "A Comprehensive Study of Crop Recommendation Systems for Precision Agriculture", International Journal of Research in Engineering and Technology (IJRET), Volume 8, Issue 2, February 2020.

4. Mahmoud Y. Shams 1 ,Samah A. Gamel 2 &Fatma M. Talaat "Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making", Original Article, Open access, Published: 11 January 2024Volume 36 , pages 5695–5714, (2024)