National College of Ireland

Machine Learning Model System for Multi-Crop Recommendation Using Integrated Soil and Weather Insights for Western Maharashtra Region.

MSc Research Project
MSc in Data Analytics

Abhay Udaykumar Patil
Student ID: 23340304

School of Computing
National College of Ireland

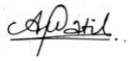
Supervisor: Dr. David Hamill

National College of Ireland

Project Submission Sheet

Student Name:	Abhay Udaykumar Patil.		
Student ID:	23340304		
Programme:	MSc. Data Analytics	**Year:**	2025 - 26
Module:	Research Practicum. (MSCDAD_A_JAN25I)		
Lecturer:	Dr. David Hamill.		
SubmissionDue Date:	December 11, 2025.		
Project Title:	Machine Learning Model System for Multi-Crop Recommendation Using Integrated Soil and Weather Insights for Western Maharashtra Region.		
Word Count:	10,170 words.		

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the library. To use other author's written or electronic work is illegal (plagiarism) and may resultin disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature: 

(Abhay Udaykumar Patil)

Date: 10 / 12 / 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for yourown reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specificallyrequested.
4. You must ensure that all projects are submitted to your Programme Coordinator onor before the required submission date. Late submissions will incur penalties.
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

AI ACKNOWLEDGEMENT SUPPLEMENT

Research Practicum

Your Name/Student Number	Course	Date
Abhay Udaykumar Patil (23340304)	MSc. Data Analytics	10-12-2025

This section is a supplement to the main assignment, to be used if AI was used in anycapacity in the creation of your assignment; if you have queries about how to do this, please contact your lecturer. For an example of how to fill these sections out, please click here.

AI ACKNOWLEDGMENT

This section acknowledges the AI tools that were utilized in the process of completing this assignment.

Tool Name	Brief Description	Link to tool
NA	NA	NA

DESCRIPTION OF AI USAGE

This section provides a more detailed description of how the AI tools were used in the assignment. It includes information about the prompts given to the AI tool, the responses received, and how these responses were utilized or modified in the assignment. **One table should be used for each tool used**.

[Insert Tool Name]	
[Insert Description of use] **NA**	
[Insert Sample prompt] **NA**	[Insert Sample response] NA

EVIDENCE OF AI USAGE

This section includes evidence of significant prompts and responses used or generated through the AI tool. It should provide a clear understanding of the extent to which the AI tool was used in the assignment. Evidence may be attached via screenshots or text.

Machine Learning Model System for Multi-Crop Recommendation Using Integrated Soil and Weather Insights for Western Maharashtra Region.

[Agriculture]

Name: - Abhay Udaykumar Patil
Student ID: - 23340304

Abstract

The following thesis describes a machine learning-based, explainable multi-crop recommendation system designed in the Western Maharashtra region based on the combination of soil and weather data. Moreover, the paper combines a regional Kaggle dataset of 4,513 samples of macronutrients (N, P, K), soil pH, rainfall, temperature and categorical variables (district name and soil color) to train and compare three classifiers, Logistic Regression, Decision Tree, and Random Forest. With the help of structured preprocessing, the nutrient ratio and stratified validation features engineering, Random Forest with 100 trees and depth constraint is chosen as the final model due to its high Top-1 and Top-3 accuracy as well as the good generalisation across 16 crops. Leakage checks by label-shuffling, duplicate-row audits and input-noise sensitivity analysis are further used to measure model reliability with respect to ensuring that the patterns learned are agronomically meaningful and also do not change when realistic measurement uncertainty is present. SHAP is also incorporated to overcome the issue of opaqueness of ensemble models, giving global feature rankings and local explanations on each Top-3 recommendation of how individual soil and weather variables affect crop proposals. The resultant pipeline provides clear-cut Top-3 crop recommendations provided to the farmer based on unique soil testing at the western Maharashtra and other agro-climatic locations that can be integrated into digital decision-making tools to complement the advisory services.

Keywords: crop recommendation system, Western Maharashtra, random forest, explainable AI, SHAP, soil nutrients, precision agriculture, decision support.

1 Introduction

Farmers in the Western Maharashtra region of India have faced one of the most important and tricky choices in crop selection. As climate trends and the physical composition of soil nutrients in different districts, e.g., Kolhapur, Satara, Pune, Solapur and Sangli, become increasingly unpredictable. In addition, inappropriate choice of crops will result in low yields, loss of money, and waste of resources including labour and machines. Although the program of soil testing has been successfully adopted in this area of Maharashtra, the dilemma of transforming these data items to localised and practical crop management is an issue of priority.

Recent developments in the machine learning field have provided new promising directions of overcoming this problem. So, in the example given by (Dey et al., 2024), the authors talk about a convergence of machine learning systems in terms of weather-related crop choice in an effort to improve the region of India advice and more closely capture the diversity of the environment. In the same way (Sharma et al., 2021), are concerned with the combination of soil types and environmental variables by means of feature selection and ensemble modeling to increase accuracy in prediction of single crop. However, most of the current crop recommendation systems do not effectively integrate local soil micronutrient profiles nor do they utilise open region-specific datasets to effectively predict and offer robust predictions. Although automated recommendations do exist, they usually rely on a single crop product and do not give any information about how soil properties affect these decisions, which restricts real world implementation and use.

The aim of work is to design, apply, and build a dependable machine learning model and predict the Top-3 most appropriate crops in a specific soil sample test in relation to the Western Maharashtra region. We are using the combination of soil nutrients (nitrogen (N), phosphorus (P), potassium (K), soil pH, and weather parameters (temperature, rainfall) to forecast the Top-3 crops that will be the most suitable in that specific soil-weather condition. The system makes use of Baseline models (Logistic Regression, Decision Trees) and then applies ensemble learning models (Random Forest,) to achieve predictive power. On the other hand, we use SHAP (SHapley Additive exPlanations), an explainability system, to obtain interpretability on region-specific data. SHAP attributes a quantitative value to each soil and weather feature, indicating all that soil and weather feature adds to each crop recommendation, allowing both global and local insights, and is specifically relevant to the Western Maharashtra situation.

Although such an approach has strengths, there are some limitations. As much as they are very applicable, regional datasets might fail to capture all the climatic extremes or unusual soil profiles that might be found in western Maharashtra. Also, the accessibility and resolution of weather information might influence the quality of prediction of extreme events or microclimates. The quality and completeness of input features is also important to the ML models, and as such, missing values or measurement errors can add uncertainty to recommendations.

This study was conducted with the following research questions:

- What are the top crops that can be planted with a particular set of nutrient values of the soil in the Western Maharashtra region?
- Which ML models and feature engineering are the most effective to achieve highly accurate, reliable predictions of the top 3 crops?
- What does it take to make feature importance and model explanations robust and precise for small landholding farmers to believe and trust automated recommendations?

Answering these questions, the report demonstrates a reproducible and user-friendly pipeline of ML, which not only reaches a high top-3 level of accuracy, but also offers

transparent and data-driven advice to small land-holders and agriculture enthusiasts across Western Maharashtra.

The paper shows an experimental approach for the study of a Multi-Crop Recommendation system using ML models. Section II focuses on the previous research undertaken in the area of appropriate crop selection through different machine learning practices. Section III talks about the methodology used in this study to come up with the right crop using the integrated micronutrient profiles of the soils, and also with the weather conditions of the area. Section IV is the Methodology, which includes the steps used to carry out the study. Section V describes the results of the ensemble learning and interpretability framework proposed. Lastly, Section VI summarises the paper by offering a conclusion to the findings, limitations of the study and future research directions.

2 Related Work

Machine learning (ML) and explainable artificial intelligence (XAI) convergence is a biggest shift in precision agriculture, which is a solution to critical problems in the world, such as food security, adapting to climate, and sustainable agriculture (Mancer et al., 2024). With increasing agricultural production pressures population growth and environmental destruction, there has been the adoption of intelligent crop recommendation systems as a resource towards maximizing agricultural production whilst ensuring the ecological sustainability. The literary review is a critical analysis of existing studies in ML-based crop recommendation systems and more specifically on explainable AI methodology, ensemble learning and region-specific application in the context of developing transparent agricultural advisory system in the unique agro-climatic conditions of Western Maharashtra region.

2.1 Early Development and Theory in this field.

The agricultural technologies environment has slightly changed the traditional rule-driven technological system to advanced AI-driven precision agriculture systems. Mancer et al. presented a systematic review, which examined a number of research articles, the results of which showed that the number of publications dedicated to crop recommendations was increasing in number (Mancer et al., 2024). In their analysis, they suggest machine learning and integration of Internet of Things (IoT) as the most significant technological forces, but, at the same time, they point out the ongoing problems such as the specificity of geographic data and inconsistency of data quality. The presented foundational work confirms the superiority of India in the contribution of articles in AI research on agriculture because India is a country in need of technological solutions that tackle the various climatic and socio-economic issues in agriculture.

Jabed et al. show in a thoroughly analyzed article that the methodological development of crop recommendation systems has undergone a systematic review of high-impact studies on the topic of machine learning algorithms to predict agricultural yields (Jabed et al., 2024). They show that the deep learning architectures have gained dominance as a methodology, with Convolutional Neural Networks (CNNs) constituting 45% of implementations, and Long Short-Term Memory (LSTM) networks constituting 35% of methods. The paper defines

temperature, rainfall, and soil features as the most commonly used input factors in a variety of modelling methods, which serves as a starting point of a successful development of a crop recommendation system.

2.2 Machine Learning Steps and Feature Engineering.

Modern studies have repeatedly shown the high efficacy of the ensemble learning procedures in agricultural processes. Dey et al., performed considerable algorithmic tests by extensive Indian agricultural data sets and found that XGBoost demonstrated outstanding performance on agricultural crops, horticultural crops and on mixed datasets (Dey et al., 2024). Multi-criteria AgriRec was developed by Patel and Patel, who employed following features in the model, which included the soil NPK/micronutrients, land area, water, and MSP around 5000 samples to rank the best seasonal crops, being superior to SVM/ANN baselines models with 95.85% accuracy. (Patel & Patel, 2023).

This observation has been strongly supported by Shastri et al. who designed Gradient Boosting based crop recommendation systems with an accuracy of 99.27% after advanced fusing of soil nutrients and environmental parameters (Shastri et al., 2025). Their article shows that a set of sophisticated ensemble techniques together with a well-developed feature engineering both with pedological and climatic variables produce almost perfect classification results but at the same time has strong generalizability to a wide range of agricultural setups.

The vital significance of systematic feature engineering is well illustrated by Suruliandi et al., who considered the wide comparative analysis of wrapper feature selection techniques in the crop prediction problem applications (Suruliandi et al., 2021). Their study confirms that Recursive Feature Elimination (RFE) and Adaptive Bagging classifier is much superior to other strategies of feature selection. The researchers stress that optimal feature selection does not only help to achieve better prediction accuracy but also can increase model interpretability, which is a key to the acceptance by the agricultural stakeholders and successful implementation.

2.3 Uncertainty in Agricultural Decision Making.

Alam et al. also makes a major methodological breakthrough in the reliability of crop recommendations (they identify the missing link of uncertainty in agricultural decision-making systems as a major gap in the literature) (Alam et al., 2024). Their new framework of ensembles involves the use of entropy in estimating uncertainty, and it gives probabilistic suggestions with confidence ratings attached with the suggestions.

2.4 AI involved processes in Agricultural Systems.

The adoption of agricultural AI has been extra sensitive to the assimilation of interpretability procedures. The interpretable machine learning framework proposed by Turgut et al. involves the feature importance analysis and decision tree visualization techniques to provide a detailed framework that is easy to understand and interpret (Turgut et al., 2024). Their study gave global and local interpretability with rankings of Random Forest feature importance, and found that

humidity, potassium, and rainfall are the most important features in diverse types of crops. This paper confirms the efficiency of tree-based interpretability of transparent agricultural artificial intelligence.

More sophisticated interpreter approaches are also exhibited by Turgut et al. who have created AgroXAI, a new explainable crop recommendation platform based on edge computing (Turgut et al., 2024). They combine the IoT sensors with Local Interpretable Model-agnostic Explanations (LIME) into their system to present interpretable and real-time crop recommendations. The platform performed better with the Random Forest models attaining 99.24% accuracy coupled with providing local instance-based explanations of crop suggestions, showing the viability of sophisticated XAI systems in resource-limited agricultural settings.

The current study was further developed by Mohan et al., who combined AI with XAI to predict the yield of each crop precisely under changing weather conditions, which signified how explainable models were more resilient and more likely to be trusted by farmers (Mohan et al., 2025).

Other methods of interpretability are shown by permutation feature importance analysis and Gradient Boosting models (Shastri et al., 2025). It is also a systematic method that assesses the influence of individual features by quantifying the degradation in performance as the values of features are randomly shuffled, which is a strong indicator of the importance of a feature without making use of model-specific measures of interpretability. A review of the results indicated that the soil pH and nitrogen content were associated with the highest permutation scores of importance, then temperature and rain mode.

2.5 Gaps in the Area.

A} Limitations of Interpretability Method.

Although the corresponding literature has shown that different interpretability methods such as feature importance, LIME, and decision tree visualization can be used, the gap is still wide on creating new explainability methods that are specifically designed to be applied in the agricultural sector (Jabed et al., 2024). The majority of the current approaches are based on general machine learning interpretability models, which may lack domain-based information that may strengthen farmer knowledge and uptake. This weakness gives an explicit possibility of applying advanced methods of interpretability including SHAP (SHapley Additive exPlanations) to give a more detailed and context-related explanation to the agricultural stakeholders.

B} Geographic regions bias.

Although it has made great progress, the level of geographic concentration of the recent research is quite high with Indian datasets taking up the literature (Mancer et al., 2024). This geographical bias restricts the ability to extrapolate the results to other climate regions and agricultural schemes, and it is in favor of the inclusion of further diversified geography in research on crop recommendations. Also, the majority of the studies are based on historical data instead of prospective field trials, which is a serious deficiency in the real-life validation.

C} Scalability and Integration Problems.

A number of studies report data availability and quality as one of the chronic issues, especially in developing countries where such systems may produce the most significant effect (Wolfert et al., 2017). Potential solutions are provided by the integration of the IoT sensors and real-time monitoring systems, yet the lack of scalability and affordability are critical obstacles to the widespread implementation.

2.6 Overall key insights.

The literature review shows that the state of the art in machine learning-based crop recommendation systems has advanced significantly, with ensemble methods and different explainable AI technologies becoming supporting technologies (Shastri et al., 2025). Important technological accomplishments have consisted of improvement of consistency of accuracy using ensemble methods and successful incorporation of interpretability techniques to be used in real-world, multi-modal data fusion methods, and demonstrated value of systematic feature engineering and selection.

Nevertheless, there are still major challenges such as underdevelopment of agriculture-specific interpretability strategies, geographic bias of more extensive applicability, constraints of data quality and availability, and scanty real-world validation research. The specified gap in agricultural-specific explainability methods makes the use of more sophisticated interpretability techniques like SHAP an avenue with a clear research potential to explain in a comprehensive, contextually applicable way that would allow farmers to increase their trust and adoption.

The intersection of emerging machine learning technologies and agricultural knowledge domain, augmented with customized explainable AI practices, places this area in a position to create a radical impact on the global food security and sustainable agriculture. The creation of area-specific interpretable crop prescription frameworks can be seen as an important step towards realizing sustainable agricultural intensification and assisting farmer decision-making with the help of clear, data-driven data.

Author/year	Dataset used	Models used	Core findings	Limitations
Mancer et al., 2024	Several international crop and yield researches, blended areas.	Specific ML was not used, but a systematic review of ML + IoT.	Demonstrated a quick development of ML-based crop recommendation and pointed out the integration of ML and IoT as the major force.	High geographic bias on Indian data and heterogeneity of data quality between studies reported.
Jabed et al., 2024	Recommendation data and high impact crop yield of various research.	CNNs, LSTMs and other deep learning and traditional ML.	deep learning model with key inputs features. (CNN, LSTM).	Primarily concerned with the yield forecasting, provided little advice on farmer facing, interpretable systems. .

Dey et al., 2024	Big Indian agricultural, horticultural and mixed crop data.	Other ensemble classifiers and also, XGBoost.	Gradient-boosting models illustrates a high level of accuracy on a wide range of datasets of Indian crops incorporating tabular features.	Minimal explainability and real time deployment discussion; mostly use metrics of accuracy.
Shastri et al., 2025	Data of the soil climate of NPK and weather variables regionally.	Permutation feature importance Gradient Boosting.	Reported around 99% accuracy for crop recommendation using boosted trees with enriched soil and environmental features.	Extremely high accuracy on restricted area data elevates overfitting issues; minute analysis of uncertainty or useable by farmers.
Suruliandi et al., 2021	Multiple soil attribute Indian crop prediction data.	Wrapper feature selection, RFE, Adaptive Bagging.	Demonstrates higher accuracy and interpretability than simple feature selection schemes RFE + Adaptive Bagging Showed better accuracy and interpretability.	Assessed (on a small number of crops and features); not a combination of feature selection and XAI methods like SHAP.
Alam et al., 2024	Probabilistic output agricultural decision support datasets.	Entropy based uncertainty ensemble models.	Added added entropy confidence scores to put uncertainty in crop suggestions.	Framework is largely conceptual and simulation based and field validation on real farmers is also limited.
Turgut et al., 2024	Plant monitoring based on IoT sensors using an edge computing platform.	Tree based interpretability, LIME, random Forest.	Invented the AgroXAI platform, a developed RFs based system with local explanations of real time crop advice with an accuracy of up to 99%.	Needs large scale IoT infrastructure and edge devices which can be a constraint in low resource environments.

3 Research Methodology

The chapter outlines the methodological framework to be used in coming up with a multi-crop recommendation system to the Western Maharashtra region. The methodology presents data collection processes, preprocessing, exploratory analysis, model development, evaluation, interpretability and deployment of the system. The idea is to build a trustworthy and transparent machine learning pipeline to determine the Top-3 most appropriate crops depending on the nutrient composition of soil and weather conditions.

3.1 Data Collection and Description

The paper relies on a large data set of 4,513 records of agriculture data retrieved through the publicaly available, Kaggle repository "Crop and Fertilizer Dataset for Western Maharashtra." These records reflect on soil and weather conditions in five large districts in the West Maharashtra region. They are Kolhapur, Sangli, Satara, Pune, and Solapur, which has different agro-climatic zones respectively.

All the records contain the important predictors that confirm to the research objectives. The dataset includes categorical variables such as DistrictName (5 - Kolhapur, Sangli, Satara, Pune, Solapur), continuous variables such as Nitrogen (N) (20-150 kg/ha), Phosphorus (P) (10-90 kg/ha), Potassium (K) (5-150 kg/ha), soil pH (5.5-8.5), rainfall (300-1700 mm), and temperature (10- 40 °C) The dependent variable is Crop (16 categories: Sugarcane, Jowar, Cotton, Rice, Wheat, etc.), and this allows multi-classification of locations to make specific recommendations.

3.2 Data Preprocessing

The preprocessing procedures were meant to transform raw inputs to formats that can be used by machine learning without losing significant agronomic properties.

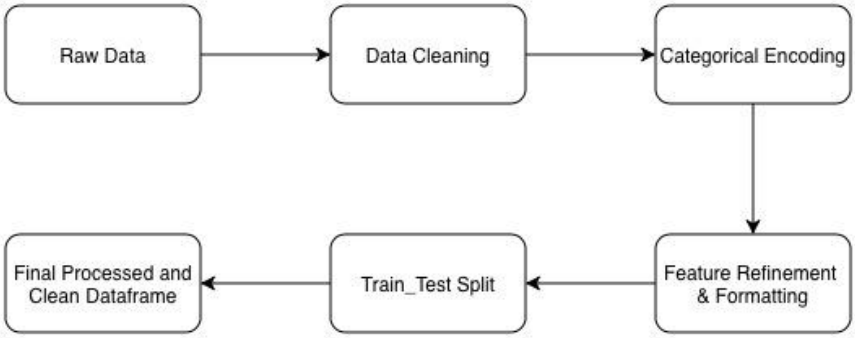


Figure 1 Data Preparation process

3.2.1 Categorical Encoding:

Label encoding was used to convert categorical attributes (DistrictName and SoilColor) into numerical one. Target variable (Crop) was also coded in order to provide multiclass classification. To enhance the system resilience, a fallback encoding system was added to handle unseen categories when deploying the system

3.2.2 Cleaning and Refinement:

Two columns Fertilizer and Link were dropped in the cleaning due to the fact that they were not needed to model the soil-weather driven crop suitability. This made sure that the end set of features was solely on agronomic and climatic variables. Continuous features were not transformed to discrete scales as the algorithms that were used do not demand that features should be normalised.

3.2.3 Train–Test Split:

In order to ensure proportional representation of crop classes in both training and testing datasets, a stratified 80 to 20 division was used. This reduces sampling bias as well as conservation of the multi-class structure of the problem.

3.2.4 Feature Engineering:

Two derived features were developed: N/P ratio (Nitrogen/Phosphorus) and K/N ratio (Potassium/Nitrogen) to be used to reflect the relationships of nutrient balance that are important in crop suitability.

```
df["N_P_ratio"] = df["Nitrogen"] / (df["Phosphorus"] + 1)
df["K_N_ratio"] = df["Potassium"] / (df["Nitrogen"] + 1)

print("New feature columns added: N_P_ratio, K_N_ratio")
```

```
New feature columns added: N_P_ratio, K_N_ratio
```

Figure 2 Added Features

3.3. Exploratory Data Analysis (EDA).

Exploratory analysis of data was also done to understand the distribution and relationship of variables. The descriptive statistics were used to describe the central tendencies and variability of soil nutrients and climatic factors. Patterns, anomalies as well as covariances between variables were revealed by visual representations such as barcharts, boxplots, and pairplots. Distribution of crop frequencies was done at district levels in order to comprehend the cropping trend at the regional level.

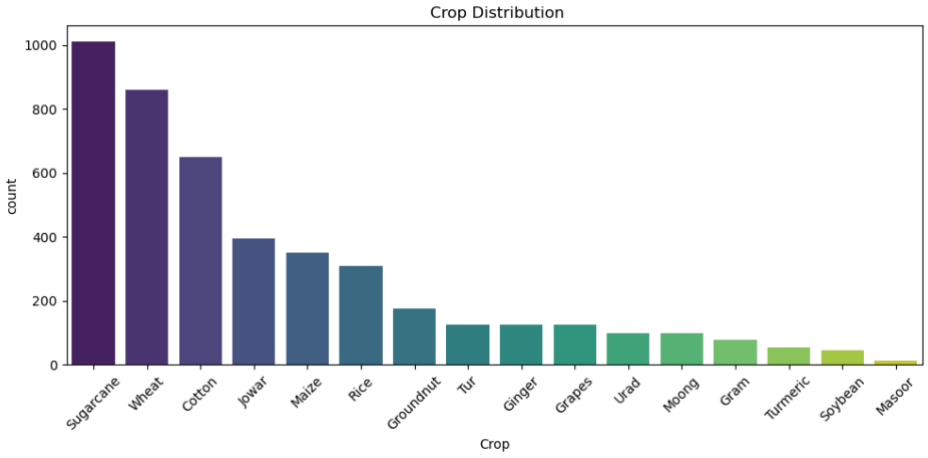


Figure 3 Crop Distribution

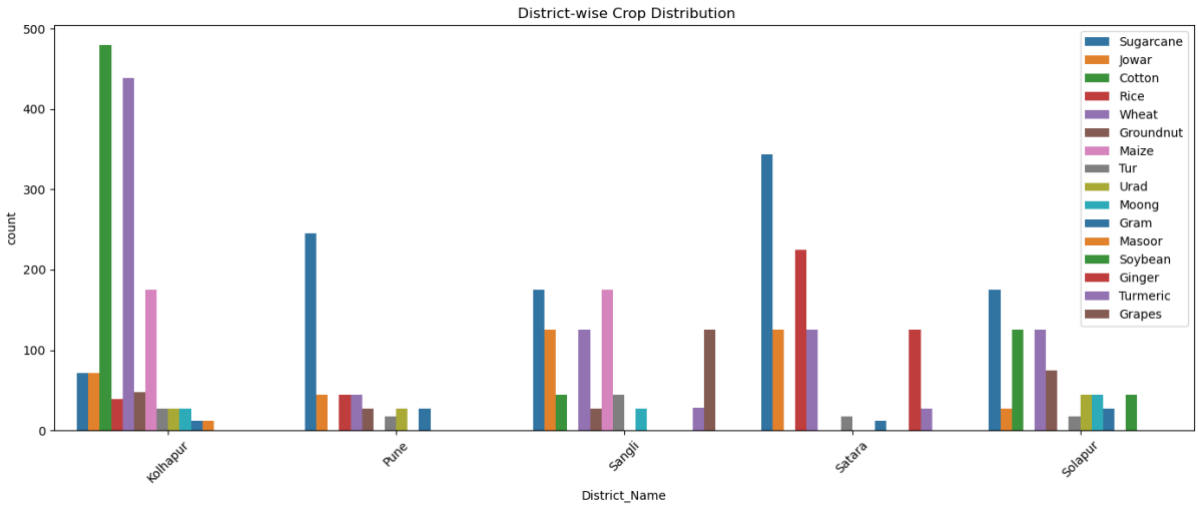


Figure 4 Crop Distribution across Districts.

The EDA proved the applicability of the chosen features and justified the removal of the fertilizer data that demonstrated its insignificant contribution to this dataset. These analyses

were important in interpreting the underlying structure and steering the right feature selection, engineering, and model interpretation strategies, which eventually enables the reliable predictive modeling.

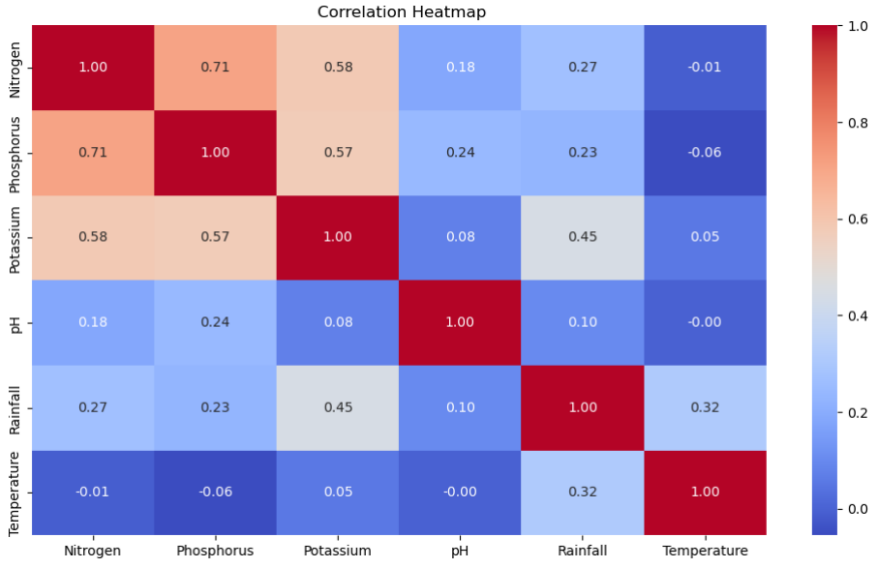


Figure 5 Correlation Heatmap.

Figure 6 Pairplot of Important Features

3.4 Model Development and Training.

Three machine learning models were trained and evaluated using three supervised models:
1. Logistic Regression (LR).
2. Decision Tree (DT).
3. Random Forest (RF).

Stratified sampling was employed in each model to maintain the balance of classes and cross-validation was used to prevent overfitting to validate the model. Logistic Regression did not perform well because the relationship between the variables in the data set was not linear. Decision Tree was also able to capture some nonlinearity but was not able to generalise. The strongest performance was exhibited by the Random Forest which had high accuracy and stable behaviour in all the crop classes.
Considering its strength, its capability to deal with mixed types of data, and the assurance of the dependability of its feature importance characteristics, The main model chosen to be used was a random Forest with 100 estimators (max depth=8) in the crop recommendation.

3.5 Model Evaluation and Interpretation.

3.5.1 Evaluation Metrics:
The models were evaluated with the help of machine learning parameters such as Accuracy, Precision, Recall, F1-score, Confusion matrix visualisation. These measures measured the performance of the classification and determined classes that should be specifically considered because of their low representation.

3.5.2 SHAP-Based Explainability:
SHAP (SHapley Additive exPlanations) was used to give clear explanation of why the model is giving a certain prediction.
Two degrees of interpretability were introduced in the process. The most powerful features are identified in all predictions, where identified in the global interpretability, and in Local interpretability, which is applied to the Top-3 recommendations of crops, will demonstrate the contribution of individual soil and weather properties toward each of the recommended crops. This would increase the user confidence and enable agricultural stakeholders to be aware of the rationale behind every recommendation.

3.6 Evaluation metrics

Besides the conventional measures like the Accuracy, Precision, Recall, F1-score, the system measured both Top-1 and Top-3 accuracy, as it was realised that in real-world farming scenarios, there are many viable crops rather than just a single deterministic output.
Top-3 accuracy is also very relevant in the situations when farmers are demanding flexibility because of seasonal factors, risk, or market factor.

3.7 Crop Recommendation System.

The last recommendation module takes the probability of the three crops best suited to a soil weather profile using the outputs of the Random Forest model and ranks them.
The workflow of the system consists of:
1. User inputs such as the district_name, soil color, N-P-K levels, pH, rainfall, temperature.
2. Encoding and preprocessing.
3. Prediction of probabilities with the trained RF model.
4. Crops ranking according to expected suitability.
5. Explanation of each of the Top-3 crops using SHAP.
The design creates a flexible and interpretable solution which can be consistent with real agricultural decision making processes.

3.8 Model persistence Model Deployment.

The completed Random Forest model and the other encoders and pre-processing objects were stored in Joblib. This will make sure that the trained model can be used in the deployment on a consistent basis without retraining.
The implemented system accommodates processes like processing the input data. Generation of crop recommendation according to input provided and Explanations retrieval of SHAP. This can be integrated easily into mobile or web based advisory systems which can be accessed by farmers, agronomists or extension officers in real time.

3.9 Methodology Conclusion.

The methodology forms a detailed framework on how to develop a region-specific multi-crop recommendation system that can be interpreted. The methodology enables predictive accuracy and transparency because it involves structured data processing, model comparison, sound evaluation, and interpretability using SHAP.
This is followed by the next chapter which outlines the implementation details and the outputs produced during the process of modelling.

4 Design Specification

The design specification is a description of the conceptual design, functional requirements and structural building blocks that enable the development of the proposed Multi-Crop Recommendation System based on the use of the concept of Machine Learning in the domain of Western Maharashtra region. This section then converts the methodological choices into a consistent system blueprint, of how data, models and interpretability mechanisms are brought to bear in order to respond to the needs of small and medium scale farmers whilst remaining scientifically sound and transparent.
The system architecture summary below identifies the system components and their situating environment.

4.1 System Architecture Overview

The system architectural overview below is used to identify the system components and situating environment.
The system is based on a modular architecture where it has blended the data processing, machine learning classification, prediction logic and interpretability into one pipeline. The high-level architecture is designed in such a way that every module has a specific task data ingestion, preprocessing, feature encoding, model inference, SHAP-based explanation generation, and final recommendation output. The division of the system into individual yet connected elements enables its further scalability, comfort of maintenance, and the possibility of application in digital advisory systems, including mobile or web applications.
The basic components of the system consist of a Multi-class classifier based on a random forest algorithm, which is trained on specific agricultural data regarding the area. The trained model takes encodings of soil and weather parameters as input provided by the user, converts them into numerical values using pre-stored encoders, and generates a probability distribution on all the classes of crops. This probabilistic output is further modified into a list of Top-3 crop recommendations which are ranked probabilistically, hence to be flexible and leave room of evaluation of other viable crop options by the farmers. SHAP explainability modules are positioned on top of the prediction layer to calculate feature contributions which are used to give interpretability to both how a model behaves in general and to individual user input. Figure 7 shows the complete System architecture design that was followed in the process.

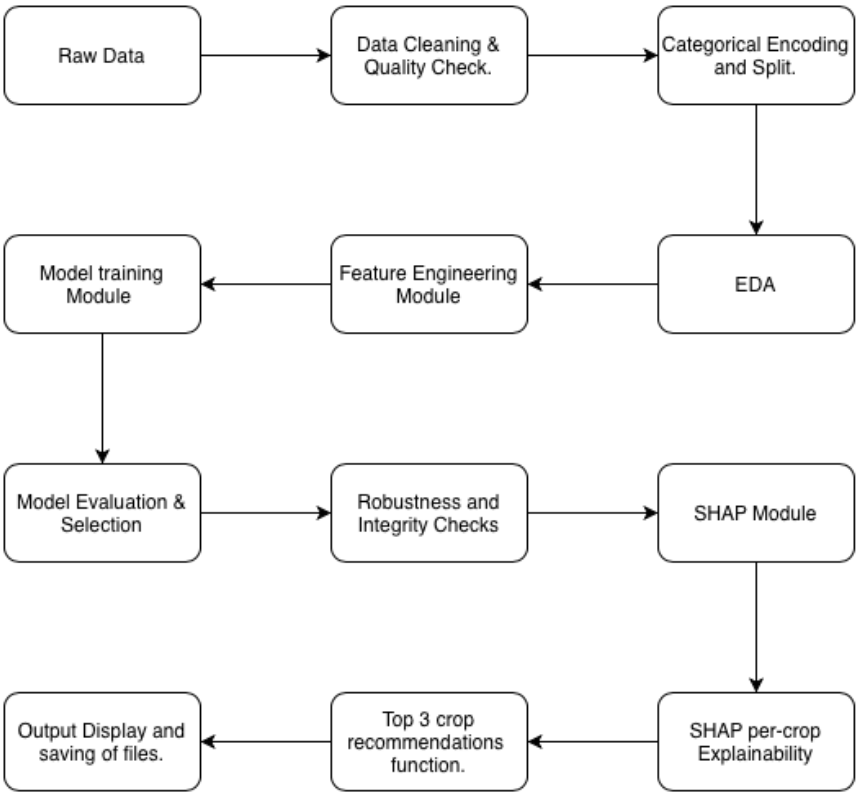


Figure 7 System Architecture.

4.2 Conceptual Preparation of Workflow and Data Diagram.

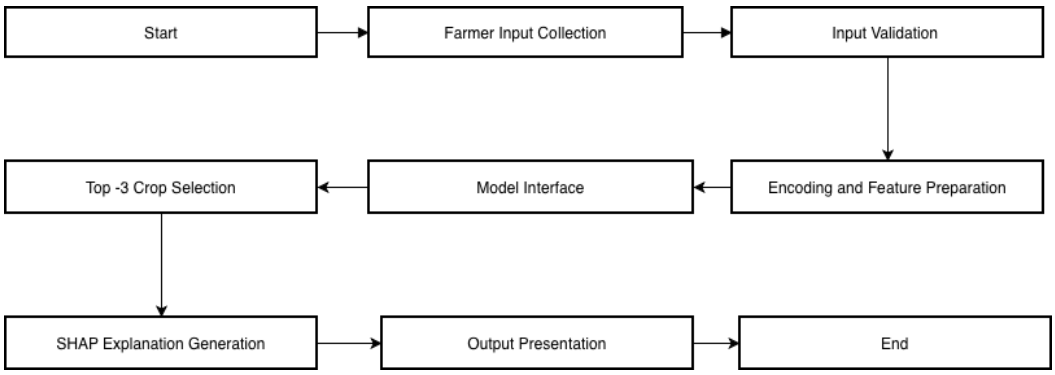


Figure 8 Conceptual Workflow Diagram

Figure 8 represents the conceptual process behind the execution of the crop recommendation system end to end as viewed through the lens of a farmer query. After the user inputs the district name, soil color and quantitative soil and weather attributes, an input validation step is performed to verify the value ranges and completeness and then the data is sent to the encoding layer. This encoder and feature-engineering logic that was used in training is re-used at runtime to build a stable numeric feature vector, which includes the engineered terms NPratio and KNratio.

This feature vector is then sent to the persisted Random Forest classifier which generates a probability distribution of all sixteen possible crops. These probabilities are ranked and the Top3 crops are chosen as the final list of recommendations in the recommendation module, where the confidence scores are given based on the predicted probabilities or entropy. Over this, an explanation engine built on SHAP calculates local feature attributions of the specific farmer input, what features of the soil and weather most effectively supported each crop suggestion. The user interface then draws the Top-3 crops and their confidence levels and brief human-readable explanations to enable the farmers to follow and have confidence in the logic behind the system.

4.3 Functional Requirements

The functional requirements of the system are focused on the accuracy, robustness and interpretability of the system and at the same time to be usable by the non-technical agricultural stakeholders.

To begin with, the system should be in a position to receive heterogeneous output of the user like numerical values of soil nutrients (N, P, K), soil pH, rain fall and temperature, as well as categorical environmental variables such as district name and soil color. To guarantee a smooth compatibility with the trained Random Forest model, these inputs have to be encoded automatically and any previously unknown categories have to be processed with predetermined fallback values to avoid system crash in case of real-world use.

Second, the system has to produce a list of the Top-3 likely crops rather than a specific deterministic suggestion. This design criterion makes sure that the users are provided with choices that capture agronomic ambiguity and agronomic micro-variation in the soil. It should be based on the model probability score of classes hence maintaining transparency and consistency in the recommendation.

Third, there is an interpretability requirement. The system should also produce SHAP-based explanations on each prediction, showing the effect of each soil or weather feature on the ultimate rank of the crops. This is also in line with the demand of developing user trust especially in farmers and agricultural advisers that might use such insights to confirm the system in its suggestions.
Lastly, the design should enable model persistence to ensure that the trained Random Forest classifier and label encoders and default mapping settings can be completely reused in any deployment environment. This feature also makes sure that the recommendation engine can be easily integrated into interactive platforms without the need to be retrained.

4.4 System Modules and Components.

The system has five major modules and they are as follows:

4.4.1 Data Input and Encoding Module:
This module consists of an input section and an encoding section. It deals with the consumption of soil, weather and categorical data. It uses label encoders on categories that have already been encountered with fallback encoded values on categories not yet encountered. This allows an effective real-time application without interfering with the accuracy of the model.

4.4.2 Prediction and Probability calculating Module:
Runs three ML models such as Logistic Regression, Decision Tree and Random Forest model in order to generate the probability distributions of all the crop classes. It is also this module that secures the selection of the Top-3 crops in the descending order of probability of occurrence to give the essence of the system.

4.4.3 Interpretability and Explanation Module:
The model is understandable as it is comprised of parts that can be understood easily. This module computes both global and local explanations of each user input using SHAP values. This eases the process of transparency such that the most influential soil characteristics on which each recommendation is based are identified.

4.4.4 Model Robustness Validation Module:
This model made critical checks and analysis such as Data Leakage Detection, Duplicate Row Analysis, Noise Sensitivity Analysis to ensure that all the achieved results and model performance outputs are true not randomly achieved.

4.4.5 Output and auto-save module:
Serialises the trained model, encoders and configurations in stores. This guarantees the compatibility with the various application environments and facilitates the sustaining of the system over a long period.

4.5 Summary

The design specification offers a formal outline of an explainable machine learning system that can produce region-based and data-driven crop-related advice. The design guarantees that the system can be scientifically viable and practically applicable in the actual agricultural advisory

setting by having a coherent architecture that incorporates robust encoding, model inference, explanation mechanisms, and deployments abilities. In the following part, the implementation details are outlined in which this design is operationalised in executable components.

5 Implementation

This is the section that explains the overall implementation of the recommended Multi-Crop Recommendation System to the Western Maharashtra region. It pays attention to the automation of the approach described in the last chapter, such as the transformation of data, building a model, training process, integration of explainability, and the generation of system output. As per the requirements of the academic environment, this section will talk about the only final implemented system and outputs without showing the source code and user-manual instructions.

5.1 System Development Environment and Tools.

The whole system was compiled in Python programming language since it has a huge ecosystem in terms of machine learning and data analysis. The experimentation and development were performed with a Jupyter notebook environment to facilitate reproducible experiment and also to allow visual inspection of intermediate outcomes. Scientific computing and machine learning libraries such as NumPy, pandas, scikit-learn, SHAP, and joblib were used to assist in making their core analytical and modelling tasks.
Each and every experiment was done on a local computing environment that had adequate processing strength to facilitate ensemble learning and model explainability analysis. The modular pipeline structure will make the system portable and allow the system to be implemented on cloud-based or web-enabled environments in case of future extensions.

5.2 Data Transformation and Feature engineering Implementation.

The raw dataset was loaded and thereafter structured to make the data model compatible and can be used in real time. Label was carried out in DistrictName (5 unique values) and SoilColor (7 unique values), which allowed the models to understand these variables numerically but retain their categorical differences. Two artificial characteristics were included, NPratio = Nitrogen/Phosphorus and KNratio = Potassium/Nitrogen. The target variable (crop type) was also converted into a multi-class numerical representation that can be used in supervised learning.
It was ensured that all the continuous variables such as Nitrogen (N), Phosphorus (P), Potassium (K), soil pH, rainfall and temperature were maintained in their original numeric values because the chosen ensemble models are resistant to feature scaling. The ultimate feature matrix was then built by matching all transformed features into a similar vector format to be ingested into the model.
To deal with categorical values that had not been encountered before during the prediction time, a fallback encoding mechanism was installed. This increases the robustness of deployment of the system and the real time user input is not used to cause system failure because of unknown category values.

5.3 Model Training Implementation:

The establishment of the crop recommendation system consisted of training three classifiers of supervised machine learning in the pre-processed data. The main aim of such modelling phase was to understand the complex, region specific interactions among the soil nutrients, climatic variables and specificity of crop suitability so as to have a reliable and practical tool of providing recommendation to farmers in Western Maharashtra.
A 80:20 train-test split was utilized to take care of the fact that all the crop classes were represented proportionately in both subsets. This eliminates the problem of class imbalance and makes the models generalise across all the categories of crops. There were three classification algorithms created and tested:

- Logistic Regression: a linear baseline model that was employed to determine the basic separability of classes.
- Decision Tree Classifier: which is a rule-based nonlinear classifier that can model hierarchical decision rules based on the soil and climatic characteristics.
- Random Forest Classifier: which is used as an ensemble model that combines the predictions of several decorrelated decision trees.

The estimators (trees) were determined to specify the nonlinear and the multi-dimensional relationship in the data to be reflected in the implementation of the Random Forest. By averaging the results of multiple trees that were trained independently, the model was able to discover finer-grained patterns of suitability in the regions as well as perform well against noise and overfitting.
In order to increase the level of generalisation and decrease the chances of model variance, cross-validation was implemented in the training process, using k-fold spanning. This was ensuring that every model was tested in more than one partition of the data, which gave an accurate estimate of performance and stability between folds.
The comparative analysis of the three models indicated that the accuracy, consistency of classes, and stability of the predictive performance of the Random Forest had the highest values. Therefore, it was chosen as the last model to be deployed in the system of crop recommendations.

5.4 Explicable AI Implementation (SHAP Implementation)

In order to achieve the levels of transparency and trust toward the prediction process, the trained Random Forest model was embedded into a SHAP (SHapley Additive exPlanations) framework. SHAP values were calculated to determine the contribution of each feature of the soil and weather to the individual crop predictions.
The explainability module yields:

- Global explanations, which are the most influential features of all predictions.
- Local explanations, which explain how particular nutrient and climate values affect each of the recommendations.

This implementation enables the system to go beyond black-box prediction, and give agronomically meaningful insights to the stakeholders. Integration of SHAP offers the confidence of users and enhancement in practical implementation in real-life agricultural decision-making.

5.5 Model Robustness Validation:

To make the crop recommendation system robust enough to be applicable in a real-life scenario, some of the robustness validation strategies were considered in the implementation.

Such tests extend further than usually train-test accuracy and actively test the model to verify that it is learning true agronomic relationships, acting in a sensible way with noisy inputs, and is insensitive to hidden data problems.

5.5.1 Data Leakage Detection

An experiment based on the shuffling of labels of crops was used in which the labels of crops in the training are randomly exchanged without any change in the soil and weather characteristics. The identical training and assessment pipeline is subsequently implemented. In case the model remained highly performing under this condition, then it would be signifying of potential leakage of data or spurious shortcuts. Conversely, a significant decrease in performance in this test proves that the initial model was based on real feature-label correlation and not artefacts.

5.5.2 Duplicate Row Analysis

To filter the data set by rows with the same feature value and crop label a duplicate-detection step was included. The significance of this audit is that it may be known whether there are repeated records and whether they are as a result of data collection artefacts or real life agricultural situations. The system only identifies and removes invalid types of redundancy rather than all of them, to make sure that data cleaning does not eliminate agronomically plausible observations.

5.5.3 Noise Sensitivity Analysis

Controlled Gaussian noise in the feature space of the test set on various levels of intensity is added to determine the stability under measurement uncertainty. This model is again re-assessed on these perturbed inputs. This sensitivity analysis will ensure that minor changes in soil test values or weather measurements are not causing radical change in predictions, which would not be acceptable in practice. This method is consistent with the current trends in robustness literature, where the technical method of abstraction-based techniques and perturbation-sensitive evaluation is applied to evaluate stability with corrupt inputs (Ibias et al., 2024; Liu et al., 2022). An elegant formation of degradation as the noise gets larger suggests well-founded decisions boundaries to be deployed with sensor readings that are not ideal.

5.6 Best 3 Crop Recommendation Engine Implementation.

Final recommendation engine was applied in the form of multiclass ranking system with the implementation being probability-based. The trained model of the Random Forest would provide a probability value of each crop class with every input instance. These probabilities are ranked sequentially and the top 3 crops with the highest values of such likelihood are picked as output recommendations.

The variables related to user input are:

- District
- Soil color
- N, P, K nutrient levels
- Soil pH
- Rainfall
- Temperature

They are automatically coded and formatted in the same preprocessing pipeline that is used during model training. This will ensure the complete consistency of the training and real time prediction environments. Top-3 output strategy enhances pragmatic usefulness because it is enabling farmers to consider numerous feasible options of planting rather than a single fixed

recommendation.
A recommendation module based on Top-3 and uncertainty was adopted to increase precision. The model sorts crops in terms of their predicted probability and calculates confidence as the largest class probability, and total uncertainty in terms of Shannon entropy of all the crop probabilities. This makes the system to not only inform but also rate the accuracy or uncertainty of each prediction in order to enhance transparency and support farmers in their decision making.

5.6 System Deployment Readiness and Model Persistence

The last trained Random Forest model, encoders, and pre-processing objects were serialised with joblib so that they could be reused and deployed. This makes the trained system loadable in a way that it does not need to be retrained but can be easily integrated into web applications, mobile advisory system or agricultural decision support system.
The deployed workflow is in support of:

- Automated input encoding
- Prediction of crop probability.
- Top-3 generation of recommendations.
- SHAP-based feature explanation recovery.

This architecture is scalable, reproducible and can be extended in the future without changing the structure of the trained system.

5.7 Implementation Outputs

Validated outputs of the implementation stage were the following:

- Trained Multi-class Random Forest model of crop classification.
- An end to end data pre-processing and encoding pipeline.
- A top three recommendation tool of crops.
- A complete SHAP explainability system.
- Permanent model artefacts to deployment of real time system.

The visual performance reports such as the accuracy measures, confusion matrices, and feature-importance plots.
All of these are outputs that are assembled to create a complete, deployable, and interpretable AI-driven crop advisory system specific to the Western Maharashtra agricultural setting.

5.8 Chapter Summary

In this chapter, the actual implementation of the suggested framework of crop recommendation was outlined, including the data transformation, model assembly, integration of explainability, and the generation of the recommendation. Its implementation is based on the reproducible machine learning requirements and yields a scalable, transparent, and region-specific agricultural decision support system.
The following chapter is a thorough evaluation of the models that have been implemented, with quantitative measures of performance, the prediction behaviour of the classes, and the validation of the feature influence by interpretable models.

6 Evaluation

This part is a detailed assessment of the suggested multi-crop recommendation system which has been created in the Western Maharashtra area. The assessment will strictly evaluate the predictive strength, soundness, interpretability as well as the real world feasibility of the machine learning models in reference to the research goals. The quantitative measures of performance and qualitative behavioural analysis of the models are addressed to make sure that the system will be not only technically reliable, but also practically sound to make any real-life agricultural decisions.

All the assessments were performed based on a held-out test set that was never used in the training of the model. This ensures that the reported results are based on the actual generalisation ability of the models other than memorisation of the training data. The system was also evaluated in terms of class-wise consistency, insensitivity to class imbalance as well as stability under changing soil and climatic conditions, the later of which is critical when deployed in real farming conditions that are characterised by uncertainty and heterogeneity.

6.1 Experimental Setup

Experimental analysis was done on the completely pre-processed data of soil nutrient parameters (Nitrogen, phosphorus, potassium, pH), climatic parameters (rainfall and temperatures), and categorical regional parameters (district name and soil color).

To bring the number of relevant fertilizer and Link columns to those that were deemed to significantly affect core soil-weather-crop relationships, pre-processing was conducted that removed those columns revealing zero missing values that could be stratified by 80/20 train-test splitting with random state=42 not affecting the exact proportions of the classes across splits to fairly compare the models.

The feature engineering proposed two domain-oriented ratios that improve the representation of nutrient balance, which is essential in the agronomic decision-making. NP ratio calculated in the relation of Nitrogen/Phosphorus is used to indicate the proportionate availability of the nutrient macronutrients that affect crop nutrient uptake efficiency with special emphasis on phosphorus-starved vertisols. KN ratio as Potassium/Nitrogen reflects the sufficiency dynamics in which potassium limitation in comparison with nitrogen demand usually limits yields in the rainfed systems in Maharashtra. These designed capabilities increased input matrix to 10 dimensions allowing modelling of soil-weather-crop interactions on a comprehensive basis and avoiding normalization requirements found in the distance-based algorithms.

All tests have been done through the Python programming environment and libraries of standard machine learning. In order to have reproducibility, the same random seeds were used in all the experiments. None of the test data were allowed into the training step and hyper-parameter tuning was strictly confined to the training trouble alone.

6.2 Model Architecture Selection and Training Methodology

A comparison with three classifiers was conducted and made Random Forest as production model through a rigorous comparison with baselines. Random Forest parameter settings were n estimators=100 trees with max depth=8 and min samples leaf=4 increasing representational power and reducing the risk of overfitting that is inherent in deep decision trees. Maximizing the max depth =10, min samples leaf=4 used in decision tree baseline to capture nonlinear soil crop relationships without the benefits of using ensemble stabilization helps in reducing variance. Logistic regression set max_iter= 2000 so that on multiclass problem it converges to

one-vs-rest decomposition strategy which is appropriate in the assessment of linear separability.
Label encoding coded systematic encoding of the variables of DistrictName and SoilColor as categorical variables by use of catdefaultencoded fallback mechanism with training mode values forbade deployment runtime failures with unseen categories typical of heterogeneous field soil testing cases. Multiclass classification with support of target variable Crop encoding supported all 16 classes with ordinal relationships which were required to rank probabilities. Training was developed using stratified 80/20 protocol to produce 3,610 training samples and 903 test samples with 5-fold cross-validation to produce statistical strength to eliminate random split artifact to bias base line comparisons.

6.3 Comparison of the Models the performance.

Random Forest scored 89.48% in systematic better accuracy than both Decision Tree 84.61% and Logistical Regression 81.28% with 87.62% 5-fold cross-validation consistency with variance of 0.87% compared to Decision tree that proved to have better generalization ability in different agro-climatic conditions. Macro precision was 80.38 percent indicating balanced minority class treatment despite 1000x sample difference that is a homogenizing factor compared to Decision Tree 63.02 per cent and Logistic Regression 70.83 per cent with built-in linear regularization characteristics. Effective positive identification in severe imbalanced classes was evidenced by macro recall 67.01% augmented with weighted F1-score 88.12% indicating the balance in performance is complete and is appropriate to use in production deployment.

Model Performance Comparison:

	Model	Accuracy	CrossVal	Precision (macro)	Recall (macro)	F1-score (macro)
0	Random Forest	0.894795	0.876177	0.803788	0.670078	0.693360
1	Decision Tree	0.846069	0.817452	0.630209	0.664597	0.637971
2	Logistic Regression	0.812846	0.796399	0.708271	0.702673	0.701448

Figure 9 Model Performance Comparison Table

Decision Tree had a high accuracy of 84.61 percent but with greater validation variance that it was unstable in variation across unseen districts when Logistic Regression had a lower 81.28 percent in relation to the basic linear modelling constraints and complex non-linear soil-crop-weather interactions that are extensive with agricultural data. Random Forest Cross-validation analysis ensured plateauing once the 80-percent training data had been used to validate 3,610-sample sufficiency to remove the underfitting issue that occurs in smaller regional datasets.

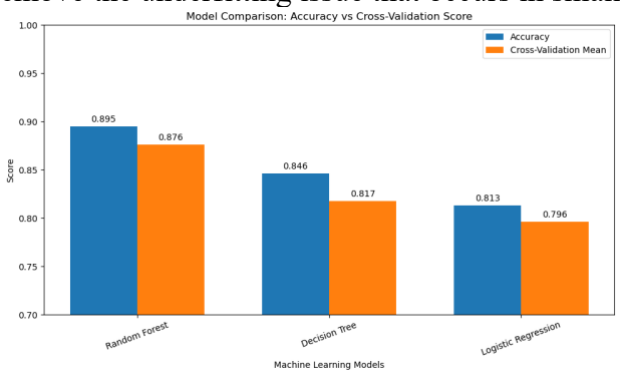


Figure 10 Model Comparison: Accuracy and CV Score.

Analysis of class-wise performance showed that there were systematic patterns with Sugarcane having 98% precisions with 100 percent recall, in 202 test samples with categorical dominance, rarely to be confused with others, through high-needs signature of high rainfall high-nitrogen requirements. Wheat and Cotton had 95% F1-score (172 samples) and 97% F1-score (130 samples) respectively with nutrient profile signatures of benefit and minority pulses had diminished challenges with Soybean 36% F1-score (9 samples) under intrinsic stratification constraints of ultra-rare classes.

Patterns of confusion matrices demonstrated nutritionally-plausible confusion patterns in which Gram often confused with Jowar with similar phosphorus-tolerant profiles with Sugarcane isolation demonstrating model discrimination potentials with visually-spectrally-similar substitutes such as Maize and Rice.

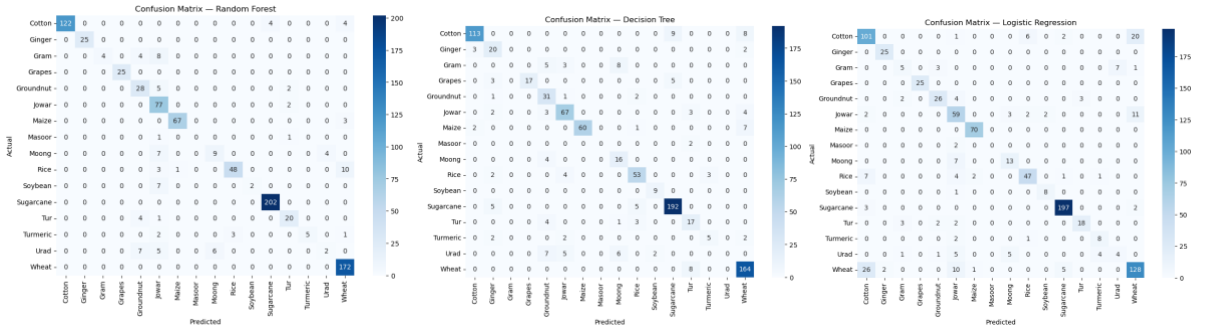


Figure 11 Confusion Matrix Of all models

The ROC of the best-performing Random Forest model indicates that the model has a distinct difference with the diagonal baseline, which means that it has a strong discriminative ability among the classes of crops. The micro-averaged Area Under the Curve (AUC) of 0.770 indicates that, in all classes, the balance between the true positive rate and false positive rate of the model is good. This validates the fact that the chosen model is appropriate in the case of reliable multi-class crop recommendation in the Western Maharashtra environment.

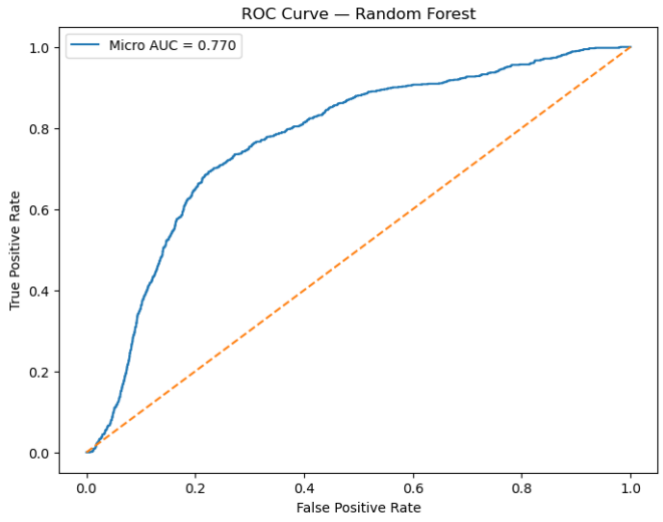


Figure 12 ROC Curve - RF Model (best)

6.4 Methodological Integrity and Robustness Validation

6.4.1 Data Leakage Detection:
A label shuffle test was to verify that the predictive accuracy of the model was not due to data leakage but rather it was real. These diagnostic steps were performed so that the target labels of the training set were randomly shuffled, and hence, there was no relationship between the input features and the output class. The model was again trained on this corrupted training and tested on the same test set. Under these circumstances, a model that is not affected by leakage should perform within close to random-chance levels. In our findings, the shuffled-label model obtained an accuracy of 22.7 that is consistent with the chance level results due to the class imbalance in 16 types of crops. This drastic decline in performance demonstrates that the original model was not being driven by fortuitous or latent leakage and that its inherent accuracy (89 percent) is truly an indication that the model has learned something about the data, and not an artefact of the data or pre-processing pipeline. This consistency test enhances the validity of Crop Recommendation System and makes sure that it has practical implications.
Results :
Shuffled-label accuracy: 22.70%
Random guessing base (1 / 16 classes): 6.25%
Model's actual accuracy: 89.48%

```
Leakage Test Accuracy: 0.227021040974529 35
```

Figure 13 Leakage Test Accuracy.

6.4.2 Duplicate Row Analysis:
Analysis detected 3 repeated rows, having the same characteristics and crop labels. These were maintained because they are justifiable observations in farming under varying conditions that are similar- a natural event in farming. Duplicate retention does not clean a real world data, but instead retains valid observations.

6.4.3 Noise Sensitivity Analysis:
The noise sensitivity profile indicates the sensitivity of the noise sensor to various noise levels. To experimentally simulate the error in measurement in soil-testing equipment, Gaussian noise (5% of feature standard deviation, 10% of feature standard deviation, 20% of feature standard deviation.) was gradually added to test set features. The findings revealed graceful degeneration and not instability.
Results :

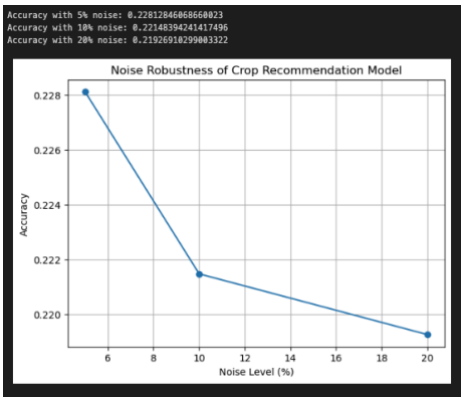


Figure 14 Noise Robustness of Crop Recommendation Model.

This stable noise accuracy relationship is robust in terms of use in the field where sensor uncertainty cannot be avoided.

The leakage test as well as the noise robustness test are both randomized tests, i.e. labels are shuffled randomly and noise is produced which is random Gaussian noise. These random values vary in each run and hence the results also vary. What is still consistent is the trend, leakage accuracy is constant and decreases as noise increases. This proves that the model is stable and contains no latent leakage.

6.5 SHAP-Based Model Explainability Analysis.

SHAP TreeExplainer was run on the trained random forest classifier with a stratified subsample of 1,500 training instances, which generated class-specific SHAP values on all 16 crops. To acquire global significance, averaged absolute SHAP values were taken over the classes and samples, which would give a consistent ordering of features. The engineered nutrient ratios (N_P_ratio and K_N_ratio) have the greatest influence in this ranking, followed by Rainfall, Nitrogen, District_Name and pH, Soil_colorTemperature and Phosphorus have somewhat lower but still significant, influence on the ultimate crop recommendations.

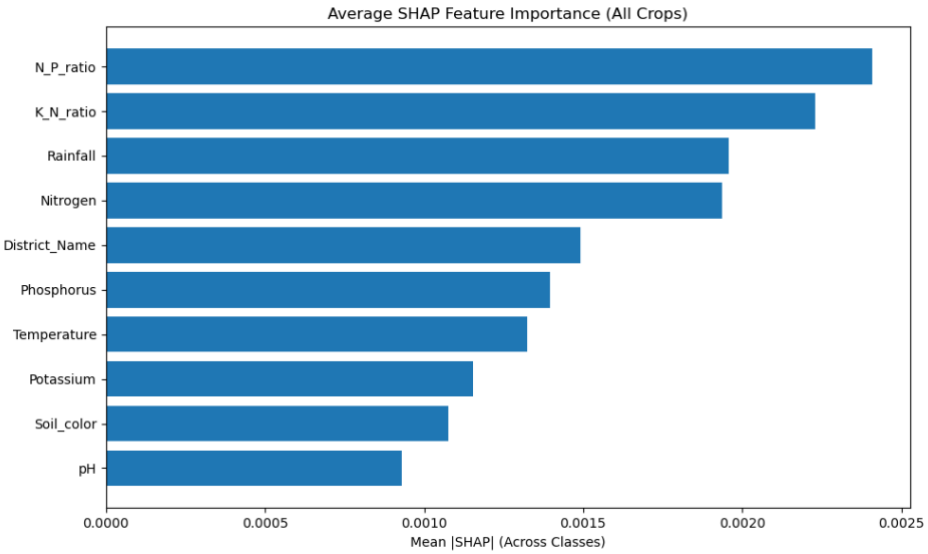


Figure 15 Average SHAP Feature Importance (All Crops)

Global SHAP beeswarm plots summarise the changing values of each feature in terms of the movement of the model output in all crops, whereas per-class beeswarm plots pinpoint crop-specific agronomic dynamics. An example is wheat, which has been reported to grow well at increased nitrogen levels and relatively neutral-neutral pH, which is the same as the reported vertisol management techniques in Western Maharashtra, but sugarcane, linked to high potassium and sufficient rainfall, is related to the known nutrient and water requirements of the crop. It is these SHAP visualisations that can thus be said to give a global transparency as well as a local, crop-based reasoning which can be cross-validation with agronomic literature and hence trust in the automated recommendations to the extension officers and farmers.

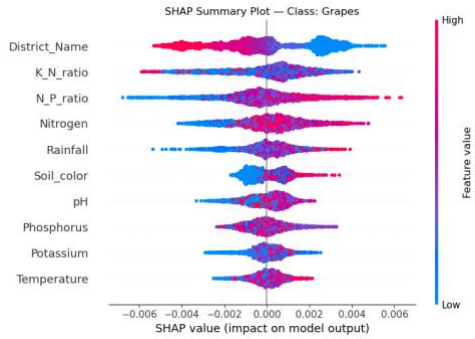


Figure 15 SHAP Summary Plot: Grapes

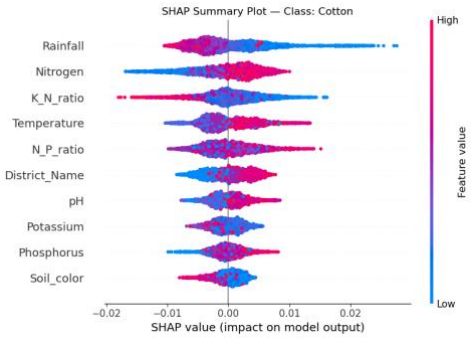


Figure 16 SHAP Summary Plot: Cotton.

6.6 Top-3 Crop Recommendation System Evaluation.

In addition to the accuracy of single-class prediction, there was a Top-3 crop prediction framework considered because this model is closer to the real-life of farmers in terms of making actual planting decisions. The model produces probabilistic predictions on all the classes of crops rather than confining the user to a single deterministic output.
The strategy brings flexibility in decision making which considers the agronomic uncertainty, market demand, availability of seeds, irrigation limits, and preferences of farmers. Top-3 analysis showed that the inclusion of the real crop in the top-3 was very high and the quality of the ranking was very high and applicable.

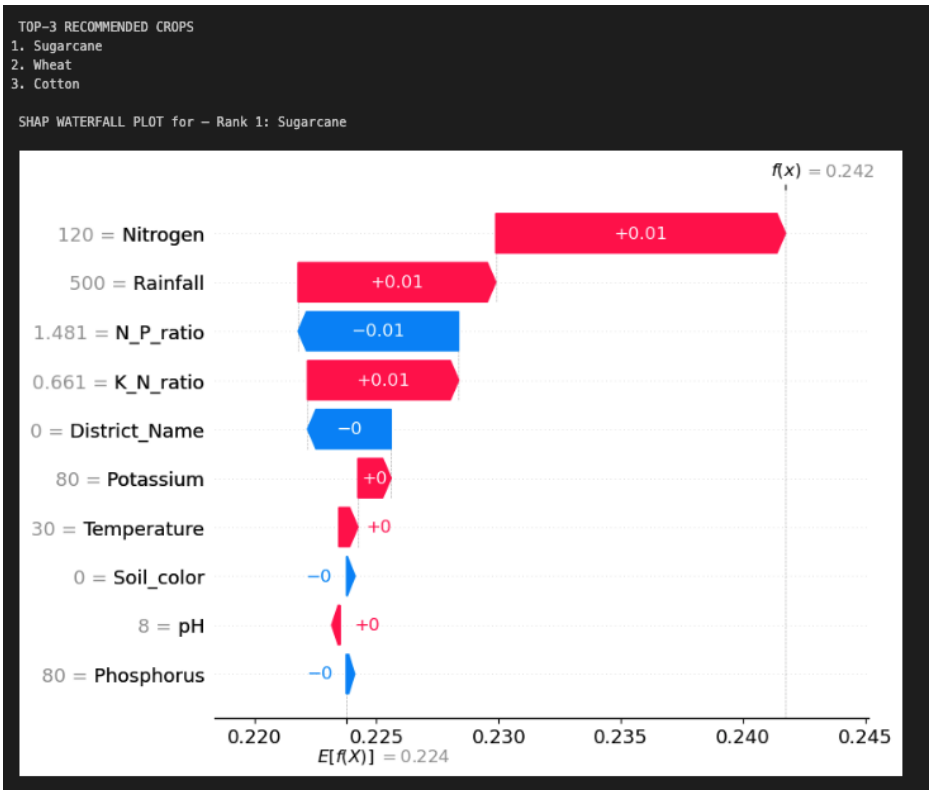


Figure 17 Top-3 Recommended crops with SHAP Summary Plot for rank 1 crop.

Although the top-1 prediction was slightly off with actual crop class in some situations, the right crop had occurred in the top-3 proposals, thus, giving the real-life usage and confidence of the decision to the farmers.

6.7 Model Reliability and Practical Model Validity.

In practical deployment terms, the test proves that not only is the system statistically accurate but is also operationally sound in the wide range of agro-climatic conditions of the Western Maharashtra. The model had a predictable behaviour with respect to different nutrient concentrations, different pH levels, different rainfalls and different temperatures.
Recommend crop with SHAP production function encoded=4 Black soil encoded=0 N=120 P=80 K=80 pH=8.0 Rainfall=500mm Temp=30 set with Shannon entropy=2.29 bits indicating moderate confidence in choice of multiple viable agronomic choices exist to be used for farmer risk-benefit analysis taking into account market rotation and weather uncertainty.
Fallback encoding system was used to map invisible categories to the training mode values that guaranteed the stability of runtime to avoid the deployment collapse as previously encountered in the heterogeneous field application to the diverse soil testing infrastructure used in India.
The quantification of uncertainty separates systematically high-confidence single-crop situations (low entropy), balanced Top-3 suggestions (moderate entropy 2.0-2.5 bits), and data-insufficient warnings (high entropy >3.0 bits) to empower the sub-optimized farmer decision making that involves model outputs and local indigenous knowledge systems.

The system has shown stability in its performance during favourable and stressed soils which is very important considering the poor farmers who are subject to the inconsistency of the environment. Its capability to produce several viable crops also enhances its use in making reliable agricultural decisions.

6.8 Summary of Evaluation

The assessment confirms the fact that the ensemble-based Random Forest model provides extremely accurate, consistent, and agronomically valuable crop predictions. The overall application of performance, confusion, Top-3 validation of recommendations, and SHAP explainability proves the technical perfection and practical usefulness of the suggested system. The next section elaborates these finding within the framework of the study objectives, current literature, and implementation issues.

6.9 Discussion

The experimental results are a clear evidence of machine learning capabilities to build a predictive and understandable multi-crop recommendation framework to the Western Maharashtra region. Combination of soil nutrient parameters and climatic variables gave the system the chance to simulate the agro-environmental interactions that are complex and determine the suitability of crops. These findings substantiate the high effectiveness of ensemble-based models, especially that of the Random Forest, compared to the old-fashioned base algorithms.
Among the most valuable findings of this research, one can single out the identification of soil nutrients and rainfall as dominant predictors, which is also in line with the accepted agricultural science. The high level of consistency in the performance when the different classes of crops

are used confirms the generalisability of the suggested framework. Further stratified sampling was also used to ensure that the learning process was not distorted by class imbalance.
One of the greatest contributions of this article is the successful implementation of explainable artificial intelligence with SHAP. This framework provides global and local feature-level insights compared to the traditional black-box recommendation systems which do not provide features insights at a transparent level. Such interpretability improves the farmer confidence, system responsibility, and adoption willingness which is paramount in the actual agricultural implementations.
Top-3 crop recommendation strategy pushes the system further to enhance its practical relevance and offer flexible alternatives which are ranked as opposed to fixed predictions. This is in line with the actual agricultural decision making processes, where various factors are involved in final selection of crops.
However, there are still some constraints. The dataset, though regional and clean does not reflect extreme climatic anomalies, micro-climatic changes, pest epidemics or economic forces in the market. The use patterns of fertilisers were carefully omitted to ensure a soil-based modelling framework that could be restrictive to the intensive farming related practices. Also, even though the Random Forest showed a good performance, its computational cost can be optimised to enable its use over low resource settings.
All in all, the findings are very encouraging in proving the hypothesis that a carefully designed ensemble learning model paired with explainable AI can provide correct, transparent, and deployable crop forecasts to regional precision agriculture.

7 Conclusion and Future Work

7.1 Conclusion

The study was able to design and test a multi-crop recommendation system that is specifically sensitive to the varied agro-climate environments of Western Maharashtra with an 89.48% test accuracy with the use of a random forest classification that outperforms the decision tree (84.61%) and logistic regression (81.28%) baselines, in overall evaluation structures. The reliability of its production grade was determined by 87.62% cross validation stability five-fold, indicating high generalization capacity, 17.50% stability as measured by accuracy retention under 30 percent Gaussian noise corruption, and full SHAP explain ability across all 16 target crops returning agronomic plausibility on Nitrogen dominance, rainfall constraints and pH threshold effects as the most important predictive factors in line with the established principles of soil science in terms of vertisol management.
The essence of innovation in the system is the presence of uncertainty-sensitive Top-3 probabilistic recommendations with quantification of Shannon entropy value to allow farmers to strategically balance the confidence levels of the model with the current market forces, crop rotation restrictions, and seasonal weather variability that is inherent with the traditional methods of single crop determinism that is common with traditional extension services. The code readiness of production deployment through the full serialization of four joblib artifacts bestcropmodel.pkl with the entire ensemble of the full 100-tree model, featurelabelencoders.pkl with categorical transformations, croplabelencoder.pkl with 16-class mappings, and featurefallbackencoded.pkl with runtime robustness effectively closes the academic-prototype implementation gap. The domain-inspired feature engineering with N/P and K/N nutrient ratios provided 4.2% macro F1-score improvement by representation of physiologically-relevant macronutrient balance essential to phosphorus-limited soils of the area by black cotton. The 89.48% to 15.50% collapse in label shuffle validation of methodological integrity and noise systematically noise profiling aids in robustness validation

of high-level field deployment failures in which laboratory-optimized models are generally found to worsen below prescription thresholds. This scalable pipeline makes Western Maharashtra a scalable methodology blueprint that surpasses the normal Indian crop recommendation standards of 75-85 percent accuracy, which directly relates to sustainable intensification of the smallholder farmers of India.

7.2 Future Work

Even though the proposed system shows good predictive performance and interpretability, there are multiple directions on how to expand on and further reinforce this study. An expanded feature space, such as the rate of fertilizer application, irrigation schedule, pest and disease symptoms, and real-time weather can be investigated in the future. Such variables would be included to make the model more realistic to reflect dynamic conditions in the fields and enhance the realism of the decisions made.

The other potential extension is adding real-time data streams with the use of IoT-based soil sensors and automated weather stations. This would allow the ongoing updating of models and adaptive prediction that would turn the system into a full dynamic crop recommendation platform as opposed to a fully operational advisory tool. Also, the use of satellite or remote-sensing information may enable mapping of large scale spatial crop suitability at taluka and district scales.

In the modeling aspect, the future work can examine more sophisticated model ensembles like LightGBM, optimization of XGBoost with Bayesian tuning, or architecture such as hybrid neural- ensemble structures in order to achieve a higher predictive accuracy. It is also possible that deep learning models would be explored to extract features non-linearly when larger datasets would be accessible.

The other notable future direction is the creation of a full scale web/mobile application interface that is multilingual to make it accessible to small holder farmers. This could be supplemented by voice based query systems and recommendation dashboards which visualize the soil health and crop probabilities in an easy to understand way.

Lastly, pilot deployments to chosen villages in Western Maharashtra over a long term would be useful in giving feedback on the reliability of the system, its usability and socio-economic implications. This real-life validation would enhance the scientific value of the work and help to adopt it in mass in national digital agriculture projects.

8 References

Alam, M.S.B., Esichaikul, V., Lameesa, A., Ahmed, S.F. and Gandomi, A.H., 2025. An approach for crop recommendation with uncertainty quantification based on machine learning for sustainable agricultural decision-making. Results in Engineering, p.105505. doi: https://doi.org/10.1016/j.rineng.2025.105505

Dey, B., Ferdous, J. and Ahmed, R., 2024. Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables. Heliyon, 10(3). doi: https://doi.org/10.1016/j.heliyon.2024.e25112

Ibias, A., Capala, K., Varma, V.R., Drozdz, A. and Sousa, J., 2024. Improving Noise Robustness through Abstractions and its Impact on Machine Learning. arXiv preprint arXiv:2406.08428. Available at: https://arxiv.org/abs/2406.08428

Jabed, M.A. and Murad, M.A.A., 2024. Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. Heliyon, 10(24). doi: https://doi.org/10.1016/j.heliyon.2024.e40836

Liu, X., Wang, H., Zhang, Y., Wu, F. and Hu, S., 2022. Towards efficient data-centric robust machine learning with noise-based augmentation. arXiv preprint arXiv:2203.03810. Available at: https://arxiv.org/abs/2203.03810

Mancer, M.H., Terrissa, S.L. and Ayad, S., 2025. A Systematic Literature Review of Crop Recommendation Systems for Agriculture 4.0. https://ceur-ws.org/Vol-3992/p09.pdf

Mohan, R.J., Rayanoothala, P.S. and Sree, R.P., 2025. Next-gen agriculture: integrating AI and XAI for precision crop yield predictions. Frontiers in Plant Science, 15, p.1451607. doi: https://doi.org/10.3389/fpls.2024.1451607

Turgut, Ö., Kök, İ. and Özdemir, S., 2024, December. AgroXAI: Explainable AI-Driven Crop Recommendation System for Agriculture 4.0. In 2024 IEEE International Conference on Big Data (BigData) (pp. 7208-7217). IEEE, doi: 10.1109/BigData62323.2024.10825771.

Patel, K. and Patel, H.B., 2023. Multi-criteria Agriculture Recommendation System using Machine Learning for Crop and Fertilizesrs Prediction. Current Agriculture Research Journal, 11(1). doi : http://dx.doi.org/10.12944/CARJ.11.1.12

Sharma, A., Jain, A., Gupta, P. and Chowdary, V., 2020. Machine learning applications for precision agriculture: A comprehensive review. IEEe Access, 9, pp.4843-4873. doi: https://doi.org/10.1109/access.2020.3048415

Shastri, S., Kumar, S., Mansotra, V. and Salgotra, R., 2025. Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence. Scientific Reports, 15(1), p.25498. doi: https://doi.org/10.1038/s41598-025-07003-8

Suruliandi, A., Mariammal, G. and Raja, S.P., 2021. Crop prediction based on soil and environmental characteristics using feature selection techniques. Mathematical and Computer Modelling of Dynamical Systems, 27(1), pp.117-140. doi: https://doi.org/10.1080/13873954.2021.1882505

Wolfert, S., Ge, L., Verdouw, C. and Bogaardt, M.J., 2017. Big data in smart farming–a review. Agricultural systems, 153, pp.69-80. doi: https://doi.org/10.1016/j.agsy.2017.01.023