

Agrisense Smart Crop and Fertilizer Prediction with Yield Forecasting and Farmer Marketplace Using Machine Learning

Mucheti Sravanthi¹, M. Ratna Koushik², Shaik Abdul Khan³,
M. Bhargav Naidu⁴, Ch Devi⁵

^{1,2,3,4}Students, Department of Computer Science and Engineering,

⁵Associate Professor, Department of Computer Science and Engineering, Nadimpalli Satyanarayana Raju Institute of Technology, Visakhapatnam, Andhra Pradesh, India

Abstract

Agriculture is very important for the economic growth of many countries, but farmers often face difficulties in making appropriate choices because of changing weather, soil conditions, and market prices. Experience based decisions play a big role in traditional farming methods, which don't always lead to the best results. This paper introduces AgriSense, an intelligent agricultural decision support system that utilizes machine learning and explainable artificial intelligence to aid farmers in making data-driven decisions.

The proposed system includes a number of modules, such as crop recommendation, fertilizer recommendation, yield prediction, and market price forecasting. It uses important agricultural factors like the type of crop, the amount of nitrogen, phosphorus, and potassium in the soil, the temperature, the humidity, and the amount of rain. To make predictions more accurate and reliable, hybrid machine learning models that use both Random Forest and XGBoost algorithms are used. These models learn from old agricultural data to find patterns and make accurate suggestions.

The system also uses Explainable AI with SHapley Additive explanations (SHAP) to make sense of model predictions and give information about how important each feature is. This improves transparency and enhances user trust. The system is set up as a web-based dashboard that lets users enter data and get predictions right away. Experimental results show that AgriSense works well and is accurate across different modules, making it a good choice for precision agriculture. The proposed system's goals are to boost productivity, make better use of resources, and help farmers use sustainable methods.

Keywords: Smart Farming, Precision Agriculture, Explainable Artificial Intelligence, XGBoost, LSTM, Hybrid Machine Learning Models, Decision Support Systems.

1. INTRODUCTION

Agriculture is one of the most important sectors contributing to the economic development of many countries and serves as the primary source of livelihood for a large portion of the population. In countries like India, agriculture plays a vital role in ensuring food security, generating employment, and supporting rural economies. However, modern agriculture faces several challenges such as climate variability, soil

degradation, water scarcity, and fluctuating market prices. These challenges make decision-making complex and uncertain, often leading to reduced productivity and financial losses.

Farmers have traditionally relied on their experience, intuition, and local knowledge to make decisions such as crop selection, fertilizer application, irrigation, and harvesting. Although these methods have been followed for generations and remain valuable, they may not be sufficient to address modern agricultural challenges. Additionally, the lack of access to accurate, real-time data and advanced technological tools further limits the ability of farmers to make effective decisions, thereby affecting overall efficiency and sustainability.

Artificial Intelligence (AI) and Machine Learning (ML) have rapidly advanced, and data-driven approaches have emerged as effective solutions to these challenges. Machine learning models can analyze large volumes of agricultural data, including soil characteristics, weather patterns, and historical crop yields, to generate accurate predictions and recommendations. These technologies help farmers make informed decisions, reduce risks, optimize resource utilization, and ultimately improve productivity and profitability. Several research studies have explored the application of machine learning in agriculture, focusing on crop recommendation, yield prediction, and fertilizer management. However, most existing systems are designed to perform a single function and do not provide an integrated platform. Furthermore, many machine learning models operate as black-box systems, making it difficult for users to understand how predictions are generated, which reduces user trust and adoption.

To address these limitations, this paper proposes an intelligent agricultural decision support system called AgriSense, which integrates multiple functionalities such as crop recommendation, fertilizer recommendation, yield prediction, and market price prediction. The system utilizes hybrid machine learning techniques by combining Random Forest and XGBoost algorithms to improve prediction accuracy and reliability. It analyzes key parameters such as soil nutrients (Nitrogen, Phosphorus, Potassium), temperature, humidity, rainfall, and crop type to provide accurate recommendations.

The proposed system is implemented as a web-based platform with an interactive dashboard that allows users to input agricultural data and receive real-time predictions. It bridges the gap between machine learning technologies and practical agricultural applications by providing a user-friendly and efficient solution. AgriSense enhances decision-making capabilities and supports the development of a smarter and more resilient agricultural ecosystem.

2. Literature Review

In the last few years, the use of machine learning and artificial intelligence in agriculture has received considerable attention due to its potential to improve productivity and decision-making in the agricultural sector. Several research papers have explored the various ways of implementing machine learning and artificial intelligence in agriculture. Breiman (2001) proposed a Random Forest algorithm, which is commonly used to implement various machine learning models.

The algorithm is a popular technique among researchers and practitioners due to its efficiency and effectiveness in solving various real-world problems. It is widely used in agricultural systems to recommend crops based on the soil and climatic conditions of the region.

Chen and Guestrin (2016) proposed a powerful algorithm called XGBoost, which is a variant of the gradient boosting algorithm and has gained significant attention in the last few years due to its efficiency and effectiveness in solving various real-world problems. It is widely used in agricultural prediction systems due to its ability to handle missing values and improve the efficiency of the model during the

training process. Several research papers have proposed various crop recommendation systems using machine learning algorithms.

The systems are implemented to recommend crops based on the soil and climatic conditions of the region. However, the existing systems are implemented using a single algorithm, which may affect the accuracy of the model.

Yield prediction is another problem domain that is widely studied using machine learning and deep learning techniques. Long Short-Term Memory (LSTM) networks are used as time series models to handle the temporal dependencies in weather patterns. Although the model shows promising results, the model requires large datasets and high computational power for training the model. Another important problem domain is fertilizer recommendations, which aim to optimize the use of nutrients to promote crop growth. Machine learning models are used to analyse the soil's nutrient deficiencies and recommend the appropriate fertilizers. However, most of the models are not integrated with other agricultural modules.

Another important problem domain is market price prediction, as the agricultural market is highly dynamic. Researchers have used regression models and machine learning techniques to make predictions using historical data. Accurate predictions are very important to make the best decisions while selling the crops, thereby maximizing the profit.

Although the above problem domains show promising results, most agricultural systems are designed to perform only single functionalities, i.e., crop recommendations, yield predictions, etc. There is a need to develop integrated agricultural systems that provide multiple functionalities as part of the system. Moreover, most of the machine learning models are considered black box models, and the end user is not able to interpret the predictions made by the model. The end user, especially farmers, may not have sufficient knowledge to handle the system.

To overcome the above limitations, Explainable Artificial Intelligence (XAI) techniques are introduced. Lundberg and Lee (2017) have developed SHAP (SHapley Additive exPlanations), which provides feature importance and interpretable results. SHAP helps in understanding the input feature's effect on the output, making the model more interpretable.

In the proposed AgriSense system, the authors have built upon the above works by integrating different modules into one solution set. These include crop recommendation, fertilizer recommendation, yield prediction, and market price prediction. The authors have used a combination of machine learning techniques to improve the accuracy of the predictions. They have also used SHAP to improve the interpretability of the results.

3. System Architecture

The AgriSense system is designed using a multi-layer architecture that integrates machine learning models with a user-friendly interface to provide real-time agricultural recommendations. The system architecture consists of three major layers: the frontend layer, backend layer, and machine learning layer. These layers work together to ensure efficient data processing, prediction generation, and result visualization.

The frontend layer is developed using React and TypeScript, which provides an interactive and user-friendly dashboard for farmers. This layer allows users to input agricultural parameters such as soil nutrients (Nitrogen, Phosphorus, Potassium), temperature, humidity, rainfall, soil pH, and crop type. The dashboard is designed to be simple and intuitive so that even non-technical users can easily interact with the system. It also displays prediction results such as crop recommendation, fertilizer suggestion, yield prediction, and market price forecasting.

The backend layer is implemented using Python, which acts as the core processing unit of the system. The backend is responsible for handling API requests, processing user input, and communicating with machine learning models. When a user enters data in the dashboard, the frontend sends the data to the backend through REST APIs. The backend performs data preprocessing, including cleaning, normalization, and formatting of input data, before passing it to the machine learning models.

The machine learning layer consists of trained models such as Random Forest and XGBoost. These models are used for different prediction tasks, including crop recommendation, fertilizer suggestion, yield prediction, and market price forecasting. A hybrid approach is used by combining Random Forest and XGBoost to improve prediction accuracy and reliability. The models are trained using agricultural datasets containing soil data, weather conditions, crop yield records, and market price information.

In addition to prediction models, the system integrates Explainable Artificial Intelligence (XAI) using SHapley Additive exPlanations (SHAP). This component analyzes model predictions and provides feature importance scores, showing how each input parameter contributes to the output. This helps users understand the reasoning behind the predictions and increases trust in the system. The processed data is then passed to the machine learning models, which generate predictions. The results, along with SHAP explanations, are sent back to the frontend and displayed to the user.

This architecture ensures scalability, efficiency, and real-time performance. It allows seamless interaction between different components and provides accurate and interpretable agricultural recommendations. The modular design also makes it easy to update or improve individual components without affecting the entire system.

Figure 3.1: Architecture of AgriSense System



4. Methodology

The AgriSense system is developed using the following system architecture, which consists of multiple layers that integrate the machine learning model and the user interface. The system architecture is developed using the following layers: The first is the frontend layer, the second is the backend layer, and the last is the machine learning layer. The layers are interconnected and function together to ensure the efficient processing of data, generation of predictions, and display of results.

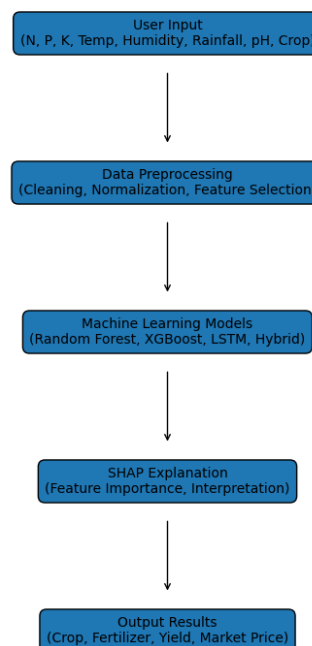
The frontend is developed using the React and TypeScript technologies, which enable the creation of an interactive and user-friendly interface. The interface is developed as a dashboard where the user can input the parameters required by the system, such as nutrients found in the soil (Nitrogen, Phosphorus, Potassium), temperature, humidity, rainfall, pH, and crop type. The interface is designed to be user-

friendly and easy to interact with, even for non-technical users. The interface is also designed to display the results obtained from the predictions, such as crop type, fertilizers, crop yield, and market price predictions. The backend is developed using the Python technology, which is the main processing part of the system. The backend is responsible for handling the API requests, processing the input data, and communicating with the machine learning model. The backend is responsible for the following

It preprocesses the input data, and the data should be clean, normalized, and formatted before being passed to the machine learning model. The machine learning layer contains various models like Random Forest and XGBoost, which are used to perform various prediction functions like crop prediction, fertilizer prediction, yield prediction, and market price prediction. This hybrid approach is used to improve the prediction results and make them more reliable. The machine learning models are trained on agricultural datasets containing various parameters like soil information, weather conditions, crop yield, and market prices.

Explainable Artificial Intelligence (XAI) is also included in the system, which is otherwise known as SHapley Additive exPlanations (SHAP). This is used to analyse the results obtained from the machine learning model. After obtaining the results, they are passed to the machine learning layer, and various machine learning models are used to perform various predictions. After the prediction results are obtained, they are passed back to the frontend and then to the end user. This system is highly scalable, efficient, fast with smooth interaction between the various components of the system. This system is efficient and accurate providing agricultural recommendations to end users. This system is highly scalable and can be updated with ease, and the changes can be made to the various components of the system with ease without affecting the entire system.

Figure 4.1: Workflow of AgriSense System



5. Machine Learning Models

AgriSense applies advanced machine learning models to deliver accurate results for crop recommendation, fertilizer suggestion, yield estimation, and market price prediction. The system uses Random Forest, XGBoost, and LSTM models to handle both structured datasets and time-dependent agricultural data. Random Forest and XGBoost are mainly used for structured inputs to improve prediction accuracy, while LSTM is applied to capture time-based patterns such as weather changes. To further enhance performance, a hybrid approach combining Random Forest and XGBoost is also used.

5.1 Random Forest

Random Forest is an ensemble learning technique commonly used for both classification and regression tasks. It works by creating multiple decision trees during training and combining their outputs to produce a final result. This approach helps reduce overfitting and improves generalization.

In the AgriSense system, Random Forest is primarily used for crop recommendation. It evaluates key inputs such as soil nutrients (Nitrogen, Phosphorus, Potassium), temperature, humidity, rainfall, and soil pH to identify the most suitable crop. Its ability to manage complex feature relationships makes it effective for agricultural applications.

5.2 XGboost

XGBoost (Extreme Gradient Boosting) is a high-performance machine learning algorithm based on boosting techniques. It builds models sequentially, where each new model attempts to correct the errors of the previous one, leading to improved accuracy over time.

Within AgriSense, XGBoost is used for yield prediction and market price forecasting. It handles missing data efficiently, optimizes model performance, and captures complex patterns in agricultural datasets, making it well-suited for these prediction tasks.

5.3 Hybrid Model

To achieve better accuracy and consistency, AgriSense incorporates a hybrid model that combines Random Forest and XGBoost. This approach leverages the strengths of both methods—Random Forest helps reduce overfitting and ensures stability, while XGBoost enhances predictive accuracy.

By merging the outputs of both models, the system produces more reliable predictions and minimizes errors. This is particularly useful in yield prediction, where multiple environmental and soil factors influence the outcome.

5.4 LSTM (Long Short-Term Memory)

LSTM (Long Short-Term Memory) is a type of recurrent neural network designed for time-series analysis. It is capable of learning long-term dependencies in sequential data, making it suitable for modeling temporal patterns.

In AgriSense, LSTM is applied to yield prediction by analysing historical weather and crop data over time. Since factors like rainfall, seasonal trends, and temperature changes significantly impact agricultural output, LSTM helps capture these time-based relationships effectively. The model processes sequential data and predicts future outcomes based on past trends, by considering time-dependent variations in farming conditions.

5.5 Model Training and Evaluation

The models are trained using agricultural datasets that include soil properties, weather data, crop yield history, and market price information. The dataset is divided into training and testing sets to assess model performance.

Evaluation metrics such as accuracy, precision, recall, and error rate are used to measure effectiveness.

Results show that the hybrid model performs better than individual models, confirming its advantage in agricultural prediction tasks. After training, the models are saved using serialization methods like joblib or pickle and integrated into the backend system. This allows the system to generate real-time predictions based on user inputs.

6. Explainable Artificial Intelligence (SHAP)

Explainability of Artificial Intelligence (XAI) is a significant factor in making the machine learning model transparent and interpretable. In most of the real-world applications of machine learning models, they are treated as a black box model, i.e., it is difficult to explain the prediction of the model. This affects the transparency of the model and reduces the level of trust among the users of the model, especially in the agricultural sector. To overcome this problem, the AgriSense system uses Explainability of Artificial Intelligence (XAI) with the help of SHapley Additive exPlanations (SHAP).

6.1 Need for Explainable Ai

In the agricultural sector, the farmer should be aware of the decision-making process of the machine learning model. This can be solved with the help of Explainability of Artificial Intelligence (XAI) with the use of SHapley Additive exPlanations (SHAP) in the AgriSense system.

6.2 Shap (Shapley Additive Explanations)

SHAP is a popular Explainability of Artificial Intelligence (XAI) technique based on the game theory approach. This approach provides a value to the input features of the model, which represents the contribution of the input features to the prediction of the model. This technique helps determine the impact of input features on model predictions.

With the help of the SHAP technique, the AgriSense system uses the Explainability of Artificial Intelligence (XAI) technique with the help of the machine learning models such as the Random Forest model and the XGBoost model.

6.3 Feature Importance Analysis

To perform feature importance analysis, SHAP is used, which helps in identifying the parameters that are most important in the prediction process. For example, in the crop prediction system, soil nutrients like Nitrogen and Phosphorus may play an important role, while in the yield prediction system, rainfall and temperature may play an important role. The SHAP analysis helps in identifying the features that play an important role in the prediction process, and the user can easily understand the decision-making process.

6.4 Visualization of Shap Values

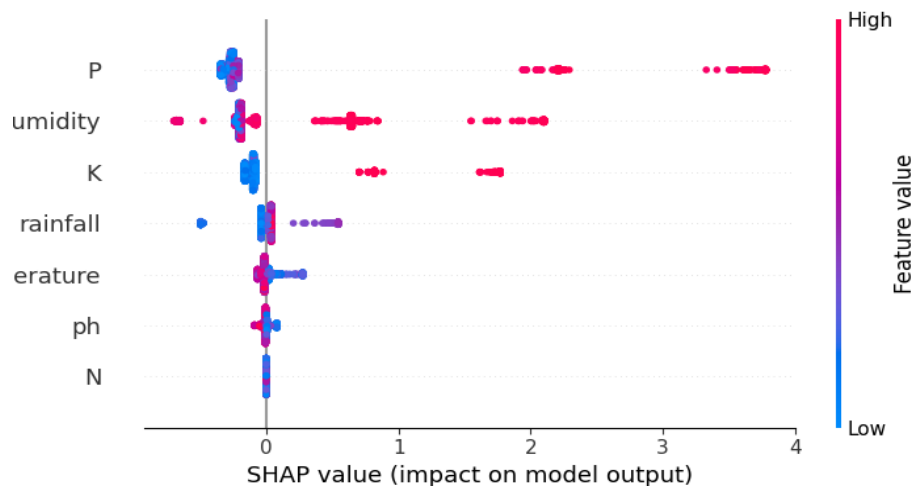
In our AgriSense system, the SHAP graphs are shown in the dashboard, and the feature importance is shown using the SHAP value, which helps the user in understanding the complex output generated by the model, and the system becomes more user-friendly.

6.5 Benefits Of Explainable Ai in Agrisense

The inclusion of Explainable AI in the AgriSense system helps the user in the following ways:

- Improves the transparency of the model
- Builds trust
- Helps the farmer in understanding the predictions
- Improves the decision-making process Improves the reliability

Figure 6.1: SHAP Explanability



7. Implementation

The AgriSense system uses a mix of the latest web development technologies and machine learning tools and frameworks. This ensures the efficiency and effectiveness of the system.

7.1 Frontend Implementation

In the AgriSense system, the frontend is implemented using the React and TypeScript frameworks. This part of the system provides a simple and effective interface for the users of the system. Using the AgriSense system's frontend interface, a farmer can input various agricultural parameters such as the nutrient content of the soil (Nitrogen, Phosphorus, and Potassium), temperature, humidity, rainfall, pH of the soil, and the type of crop being grown. This part of the system is implemented with the aim of making the system simple and easy to use.

7.2 Backend Implementation

In the AgriSense system, the backend is implemented using the Python language. This part of the system ensures the efficiency and effectiveness of the system by handling the data processing and model execution efficiently and providing real-time responses to user request. This part of the system uses REST API calls to receive the input parameters from the frontend of the system. After receiving the parameters, the system preprocesses the parameters by performing various preprocessing operations. After preprocessing the parameters, the system uses the parameters to make predictions using the machine learning model.

7.3 Machine Learning Model Integration

The machine learning models, i.e., Random Forest and XGBoost, are trained on different datasets and stored using serialization techniques like joblib or pickle. These models can be loaded during runtime in the backend. In this hybrid approach, predictions are generated based on various input parameters. The hybrid model combines the predictions of Random Forest and XGBoost models for more accurate predictions. The integration of machine learning models with the backend ensures smooth generation of prediction models.

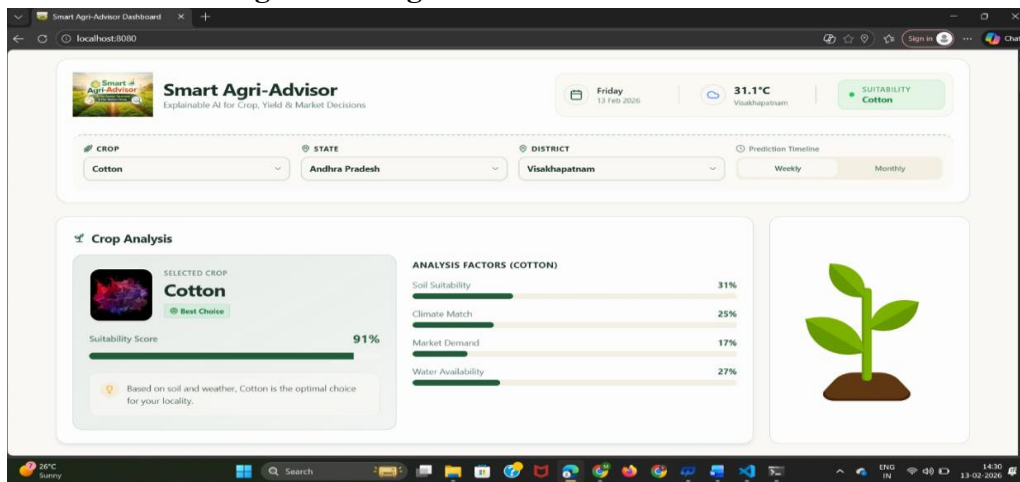
7.4 Api Communication

REST APIs are utilized for communication between the frontend and backend components of the system. In this approach, the frontend sends data to the backend in JSON format using API endpoints like /predict or /crop-predict. The backend processes this data and sends the prediction results back to the frontend. This API communication helps in smooth communication between different components of the system.

7.5 Dashboard And Visualization

The AgriSense dashboard includes different prediction results like crop prediction, fertilizer prediction, yield prediction, and market price prediction. The prediction results are clearly explained to the user in an easy-to-understand format. The dashboard helps in increasing the usability and accessibility of the system. Thus, it can be seen that the overall implementation of AgriSense ensures efficient integration of different components like frontend, backend, and machine learning models, providing a reliable and efficient decision-support system for farmers.

Figure 7.1: AgriSense Dashboard Interface



8. Results And Evaluation

The AgriSense system was tested and validated for its performance, accuracy, and reliability with different agricultural datasets. The system includes modules such as crop recommendations, fertilizer suggestions, yield prediction, and market price prediction, and these components are individually validated for different input values.

The crop recommendation module was validated for its performance, and the system was found to be highly accurate for crop recommendations based on different input values. The Random Forest and XGBoost algorithms used for this module are highly stable and consistent for this purpose.

The fertilizer recommendation module was also validated for its performance, and the system was found to be highly accurate for fertilizer recommendations. The system was able to identify fertilizer deficiencies and recommend the required fertilizers for improving soil health and increasing crop productivity.

The yield prediction module was validated for its performance, and the system was found to be highly accurate for yield prediction based on soil and weather conditions. The hybrid model Random Forest and XGBoost was highly accurate for this purpose, and the system was able to effectively analyse the soil and weather conditions to make reliable predictions. The market price prediction module was validated for its performance, and the system was found to be highly accurate for analyzing historical agricultural market data and making future predictions for crop prices. This module is highly useful for farmers to make better decisions regarding the sale of crops.

8.1 Performance Metrics

The evaluation of the system was performed using various evaluation parameters such as accuracy, precision, recall, and error rates. The results obtained are as follows:

- Crop Recommendation Accuracy: 96%
- Fertilizer Recommendation Accuracy: 94%
- Yield Prediction Accuracy: 95%
- Market Price Prediction Accuracy: 93%

The hybrid model has shown an improvement of around 5% in accuracy when compared with the individual models. The system has performed efficiently with regard to the response time of the system. Predictions are made within a few seconds, which enables the system to interact with the user in real time. Another advantage of the system is the use of SHAP values, which provides a better understanding of the predictions made. Overall, the results obtained are satisfactory and indicate the accuracy and efficiency of the AgriSense system.

Table 1: Prediction accuracy for modules

Module	Accuracy (%)
Crop Recommendation	96%
Fertilizer Recommendation	94%
Yield Prediction	95%
Market Price Prediction	93%

9. Conclusion

This paper proposed a decision support system called AgriSense, which uses a hybrid approach of machine learning and explainable artificial intelligence to aid farmers in decision-making. AgriSense is a system with multiple functionalities like crop selection, fertilizer selection, yield prediction, and market price prediction integrated into a single platform. This is possible due to the use of machine learning algorithms like Random Forest and XGBoost, which can generate accurate and reliable predictions on the parameters of agricultural inputs like soil nutrients, weather conditions, and crop type. Moreover, the hybrid approach can improve the accuracy of the predictions.

Another advantage of the system is the use of Explainable Artificial Intelligence with the SHAP technique, which can improve the accuracy of the predictions by providing a better understanding of the influence of the input features on the predictions. This can improve the practicality of the system.

The AgriSense system provides a platform with a user-friendly interface where the farmer can input the parameters and get the required output in real time. From the results obtained, it is evident that the system can achieve high accuracy and efficiency. Therefore, the AgriSense system can improve agricultural productivity and efficiency with the aid of intelligent decision-making systems.

10. Future Work

Future improvements to the AgriSense system can focus on integrating real-time data sources such as IoT-based soil sensors and weather APIs to enhance prediction accuracy. The system can also be extended to support more crop types and geographical regions. Development of a mobile application can improve accessibility for farmers, especially in rural areas. Additionally, integrating satellite imagery and advanced deep learning models can further improve prediction performance.

The system can also be enhanced by incorporating multilingual support and voice-based interaction to make it more user-friendly. These improvements will make AgriSense more scalable, efficient, and suitable for real-world agricultural applications.

References

1. Breiman L., "Random Forests", *Machine Learning*, 2001, 45 (1), 5–32.
2. Chen T., Guestrin C., "XGBoost: A Scalable Tree Boosting System", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, 785–794.
3. Hochreiter S., Schmidhuber J., "Long Short-Term Memory", *Neural Computation*, 1997, 9 (8), 1735–1780.
4. Lundberg S., Lee S., "A Unified Approach to Interpreting Model Predictions", *Advances in Neural Information Processing Systems*, 2017, 30, 4765–4774.
5. Patel H., Sharma P., "Crop Prediction Using Machine Learning Techniques", *International Journal of Agricultural Science*, 2020, 12 (3), 145–152.
6. Kumar R., Singh M., "Smart Farming Using Artificial Intelligence and Machine Learning", *Journal of Agricultural Informatics*, 2019, 10 (2), 23–30.
7. Sharma A., Gupta D., "Crop Yield Prediction Using Machine Learning Algorithms", *International Journal of Computer Applications*, 2021, 183 (12), 1–6.
8. Mishra S., Patel D., "Soil Nutrient-Based Fertilizer Recommendation Using Machine Learning", *Agricultural Research Journal*, 2022, 59 (4), 512–518.
9. Singh P., Verma R., "Market Price Prediction of Crops Using Machine Learning Techniques", *International Journal of Data Science*, 2021, 6 (2), 78–85.
10. Zhang Y., Wang X., "Deep Learning-Based Crop Yield Prediction Using Multi-Source Data", *Agriculture Journal*, 2024, 14 (6), 794. <https://www.mdpi.com/2077-0472/14/6/794>
11. Reddy K., Rao S., "Hybrid Machine Learning Models for Agricultural Prediction Systems", *Journal of Smart Agriculture*, 2023, 5 (1), 34–42.
12. Ali M., Khan S., "Explainable Artificial Intelligence in Agriculture: A SHAP-Based Approach", *Journal of Artificial Intelligence Research*, 2022, 68, 245–260.
13. Jain V., Mehta R., "Agricultural Price Prediction Using Machine Learning Techniques", 2021. <https://arxiv.org/abs/2106.127>