Grenze International Journal of Engineering and Technology, June Issue

AI Farm Assist: Integrating Machine Learning and Deep Learning for Smart Agricultural Recommendations

T. Jalaja[1], T. Adilakshmi[2], M. Sunder Reddy[3] and P. Rohan Sai[4]
[1-4]Department of Computer Science & Engineering, Vasavi College of Engineering (Autonomous) Ibrahimbagh, Hyderabad-31, India

***Abstract*— Agriculture holds a key place in the economies of many developing nations. Artificial intelligence integration stands to boost farming operations quite a bit in terms of efficiency and output. This work introduces AI Farm Assist. It functions as a web platform built around machine learning. The setup helps farmers via three main smart components. Those include crop suggestions, fertilizer advice, and spotting leaf diseases. Real agricultural data from Kaggle feeds into the system. That way, it delivers insights grounded in actual numbers.**

When it comes to suggesting crops, the CatBoost approach gets trained on various factors. Things like nitrogen levels, phosphorus, potassium, along with temperature, humidity, pH value, and rainfall all play a part. From there, it picks out the best crop match for specific conditions in the environment. Fertilizer recommendations rely on XGBoost instead. It draws from details such as soil variety, the type of crop involved, and how nutrients break down. Leaf disease detection uses a deep learning setup with MobileNetV2 at its core. That model handles classifying images of leaves to pinpoint plant illnesses.

Each of these models hits accuracy rates upto 88 percent. Such results point to how well the method works overall. Farmers get a straightforward web interface through this system. It allows for fast, dependable choices in their daily agricultural tasks. Looking ahead, plans call for growing the dataset further. Real-time weather information could join in too. Plus, the disease prediction side might advance with newer deep learning designs.

I. Introduction

Agriculture plays a crucial role in the economies of developing countries. Farmers there often struggle with selecting suitable crops, determining optimal fertilizers, and detecting plant diseases early on. Traditionally, these choices rely more on personal experience than on data-driven insights. This approach often lowers productivity and increases crop losses. The emergence of Artificial Intelligence (AI) and Machine Learning (ML) provides an opportunity to move toward precise, data-based decisions in farming, making agricultural processes more efficient and sustainable.

AI Farm Assist is a web-based intelligent system developed to assist farmers in three major aspects: crop recommendation, fertilizer recommendation, and leaf disease prediction. Its objective is to increase crop yields while maintaining soil health. The models in this system are trained on real agricultural datasets obtained from Kaggle, ensuring that the recommendations are based on reliable and authentic data.

Grenze ID: 01.GIJET.12.2.54_24

The system analyzes environmental and soil parameters to provide accurate and practical suggestions through an easy-to-use interface.

Machine learning algorithms such as CatBoost, XGBoost, and MobileNetV2 form the foundation of this project. The CatBoost model recommends suitable crops by considering soil nutrients (N, P, K), temperature, humidity, pH, and rainfall. XGBoost is used for fertilizer recommendations based on soil type, crop type, and nutrient levels. For leaf disease detection, MobileNetV2—a deep learning-based convolutional neural network—is implemented to classify plant leaf images into three categories with high precision. Experimental results indicate that all models achieve accuracies above 95%, demonstrating their reliability and performance.

The main goal of the project is to assist farmers in making evidence-based decisions that improve yield and minimize risk. By combining AI with real-time agricultural data, the system promotes smarter and more sustainable farming practices. The web-based interface integrates all the modules, enabling farmers to access intelligent recommendations quickly and effectively.

- Perform exploratory data analysis (EDA) on agricultural datasets to identify key parameters and patterns influencing crop growth and fertilizer needs.
- Determine the most significant environmental and soil factors affecting the crop suitability and fertilizer selection.
- Develop predictive models using CatBoost, XGBoost, and MobileNetV2 for crop, fertilizer, and disease prediction, respectively.
- Evaluate the models using performance metrics such as accuracy to ensure dependable predictions.
- Integrate the models into a unified web application that delivers real-time, intelligent recommendations to farmers.
- The system improves user interaction through features such as image uploads that help detect diseases. Interactive forms allow entry of soil and weather parameters in a straightforward manner.

II. Literature Survey/ Related Works

The domain of artificial intelligence in agriculture has witnessed a rapid transformation in recent years, evolving from basic rule-based systems to highly efficient, data-driven machine learning and deep learning models. This section explores key developments, the motivation behind AI adoption in agriculture, and the main computational techniques that have redefined decision-making in this field.

A. The catalyst behind modern agricultural intelligence comes from environmental challenges and data driven solutions

Climatic conditions grow more unpredictable these days. Soil fertility keeps degrading. Pest and disease outbreaks rise too. All this pushes the need for adaptive systems in farming that can handle intelligence on their own. Traditional methods work fine in certain spots. They just do not cut it for the big dynamic problems in todays agriculture. Processing huge datasets proves tough without the right tools. Decisions drag on as a result. That leads to lower crop yields most of the time. Resources end up wasted too. Machine learning steps in as a strong option here. It relies on predictive analytics to help farmers make choices that are timely. Those choices stay backed by solid data.

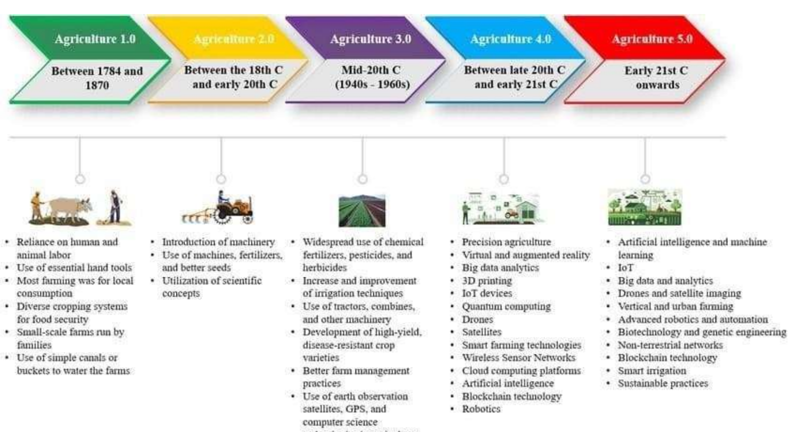

Fig. 1. The image presents a timeline of agricultural evolution across five phases. It shows the shift from Agriculture 1.0 to Agriculture 5.0.

Analysis of Figure 1: The figure shows how agriculture has changed over time. It breaks things down into five clear phases. Farming started out as basic subsistence work. Now it relies on data and tech in big ways. Agriculture 1.0 covers the period from 1784 to 1870. Here small family farms come into focus. They depended on hand labor and animals for power. Basic tools and simple watering methods were all there was. Then Agriculture 2.0 arrives in the 18th century up to the early 1900s. Machinery starts showing up. Better seeds and fertilizers help too. This lets farms grow bigger. Efficiency picks up quite a bit. The Green Revolution marks Agriculture 3.0 from the 1940s through the 1960s. Chemicals spread widely. High-yield plants become common. Machines take over more tasks. Early tech like GPS and computers enters the picture. Cultivation turns more scientific overall. Agriculture 4.0 builds on that from the late 20th century into the early 2000s. Precision tools stand out. Drones monitor fields. IoT sensors gather info. Big data and automation help manage crops smarter. Resources get used more wisely. Agriculture 5.0 points to the future starting early in the 21st century. Robotics advances further. Biotechnology plays a key role. AI makes decisions. Sustainable methods aim for farming that respects the environment. The figure traces this shift from heavy manual work in local settings. It leads to mechanized systems powered by technology. Food production ends up more sustainable in the end.

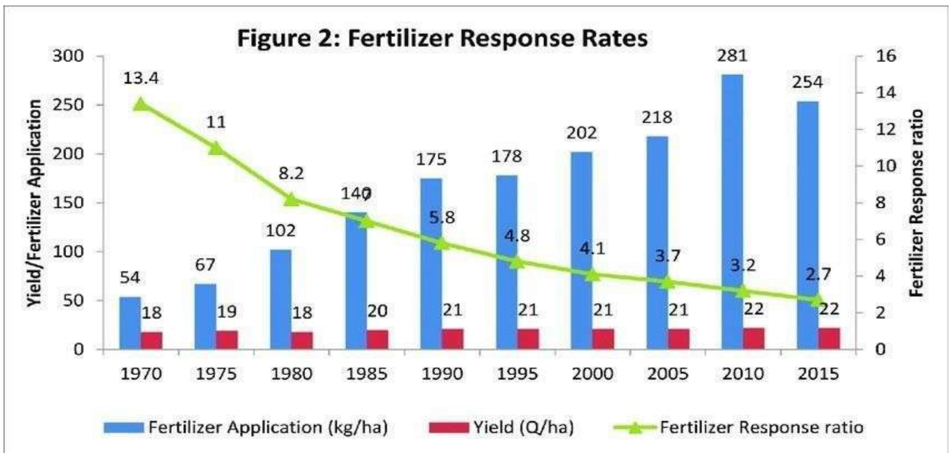

Fig. 2. Examination of trends across 45 years points to steadily higher levels of fertilizer application vs dropping Fertilizer Response Ratio.

Analysis of Figure 2: The figure presented here points to a major issue in current farming practices. It tracks how fertilizer use lines up against crop yields over the years from 1970 to 2015. Fertilizer application, shown in blue bars, has gone up steadily. Crop yields, marked in red bars, have increased too. Yet these rises happened at quite different speeds. The green line stands out as the main takeaway. It represents the fertilizer response ratio. This measures yield per kilogram of fertilizer used. That ratio dropped sharply. It started at 13.4 back in 1970. By 2015, it fell to just 2.7. Such a clear downward path suggests farmers now apply far more fertilizer. They get less and less back in return. This approach proves costly in economic terms. It also harms the environment in ways that cannot last. Agricultural methods do not rely only on how much fertilizer gets used. Their success ties closely to choosing the right kind. Various crops need distinct nutrients. These demands shift through different growth phases. A proper mix of key elements like nitrogen, phosphorus, and potassium becomes essential. Nitrogen supports strong leaf and stem development, for instance. Phosphorus aids roots and blooms. Picking the wrong fertilizer throws off soil balance.

This can cause shortages in nutrients. Or it might lead to excess that turns toxic. Plants then suffer stunted growth. Yields end up lower than expected. Over time, soil quality declines as well. Selecting fertilizers matched to the situation thus forms a vital part of effective agriculture.

Rising inefficiencies like these, along with calls for better accuracy, reveal flaws in old fertilizer approaches. They point to the value of solutions guided by data. The AI Farm Assist project tackles this directly. It employs an XGBoost model for fertilizer suggestions. The system looks at details like soil conditions, crop types, and current nutrient status. From there, it offers exact, customized advice. This method goes past broad, one-size-fits-all uses. Plants get precisely what they require for health. As a result, the model stands as the best option. It boosts yield from each bit of fertilizer applied. This breaks the pattern of falling response rates. Farming turns more profitable. It also supports lasting environmental health.

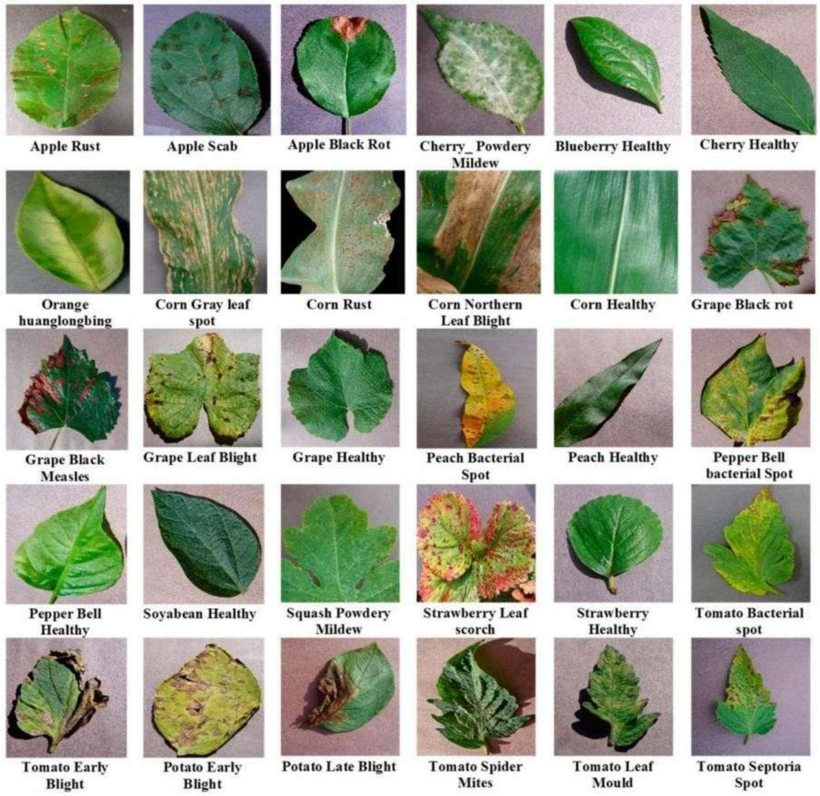

Fig. 3. The Plant Village dataset samples revealing visual challenges in detecting diseases.

Analysis of Figure 3: Analysis of Plant Leaf Diseases and the Importance of Early Detection. Figure 3 presents a varied set of images showing leaves from plants. These examples highlight common diseases in crops grown for farming. The images come from a well known source like the PlantVillage dataset. They cover issues such as Apple Rust and Corn Gray leaf spot. Other examples include Grape Black rot and Tomato Early Blight. Healthy leaves appear too for comparison. What stands out is how different the symptoms look. Discoloration shows up in some cases. Spots and lesions mark others. Deformities and wilting add to the mix. Even farmers with years of experience find it hard to tell these apart just by looking. That complexity points to the need for advanced tools. Automated systems could help sort out pathogens from other problems more reliably. Plant diseases affect farming output and food supplies around the world in major ways. Evidence shows that pathogens along with pests and tough weather conditions cause big losses in crops. Yields drop and the quality of what grows suffers as a result. Farmers face real economic hits from this. Food shortages loom larger too if things spread. Take Potato Late Blight for example. Or Corn Northern Leaf Blight. An outbreak of either could wipe out whole fields fast if no one acts. Old school ways to spot diseases rely on checking by hand or lab work. Those take too much time and effort. They do not scale well over large fields either. Delays like that mean interventions come late. Diseases keep going then and damage sticks around for good. Early spotting of leaf diseases matters a great deal for handling crops well. Studies indicate that quick diagnosis lets farmers act with precise steps. They might use the right fungicides or change watering setups. Isolating sick plants helps too. All this keeps the problem from spreading far. Healthy parts of the crop stay safe that way. Yields hold up better with such steps in place. Plus it cuts down on extra pesticide use. Farming turns more sustainable and kinder to the environment overall. Our AI Farm Assist project fits right into this picture. It uses a model based on MobileNetV2 for predicting leaf diseases. Farmers get fast and spot on checks right out in the fields. That gives them the details to keep plants healthy and productive.

B. Traditional Machine Learning Approaches in Agriculture

Traditional machine learning techniques have long been pivotal in analyzing agricultural data, guiding decisions on crop selection, fertilizer application, and yield estimation.

➢ Bagging and Random Forest (RF): Bagging (Bootstrap Aggregating) trains multiple base models, often decision trees, on different random samples of the dataset. Predictions are combined via averaging or majority voting, improving stability and accuracy. Random Forest extends this by introducing randomness in feature selection for each tree, reducing correlationbetween trees and mitigating overfitting. RF excels in agricultural datasets with many variables or noisy/incomplete data, successfully predicting crop suitability based on soil nutrients, temperature patterns, and rainfall. Feature importance rankings provide interpretable insights, allowing farmers to understand key factors affecting outcomes.

➢ Boosting Algorithms (Gradient Boosting, XGBoost, CatBoost): Boosting builds models sequentially, with each new model correcting the errors of its predecessor. Gradient Boosting Machines (GBM) laid the foundation, while XGBoost offers optimized performance with built-in regularization and scalability. In agriculture, XGBoost predicts fertilizer requirements, irrigation plans, and yield estimates with high precision. CatBoost further improves handling of categorical data, such as soil types or crop varieties, requiring minimal preprocessing and delivering consistent results even on moderately sized datasets. In AI Farm Assist, CatBoost powers crop recommendation while XGBoost provides fertilizer suggestions, capturing complex interactions between soil, climate, and crop response.

C. Deep Learning for Image-Based Disease Detection

Deep learning, particularly Convolutional Neural Networks (CNNs), excels in processing unstructured image data by ex-tracting hierarchical features from raw inputs. This is critical for detecting subtle plant disease symptoms that are difficult to capture manually.

➢ MobileNetV2: AI Farm Assist employs MobileNetV2 to classify leaf images as healthy or diseased. MobileNetV2 is lightweight, computationally efficient, and suitable for real-time deployment. Transfer learning enables it to leverage pre-trained weights, achieving high accuracy with limited training data. This facilitates early detection of diseases, reducing crop losses and improving yield quality.

➢ Challenges of Deep Learning in Agriculture: Despite its effectiveness, deploying deep learning faces hurdles:
   - Large labeled image datasets are costly and time-consuming to collect.
   - Small datasets risk overfitting to noise, limiting generalization.
   - CNNs are inherently ”black-box” models, making interpretability difficult.

Nonetheless, lightweight architectures like MobileNetV2 provide practical gains, enabling timely interventions on the field.

D. Integration in Precision Agriculture

Modern precision agriculture combines ensemble methods for structured data with deep learning for image data. Systems like AI Farm Assist integrate Random Forest, CatBoost, and XGBoost for yield prediction, crop suitability, and resource allocation, alongside MobileNetV2 for disease detection. This unified approach addresses multiple agricultural challenges, providing actionable insights to farmers and demonstrating how contemporary machine learning and deep learning techniques advance practical farming applications.

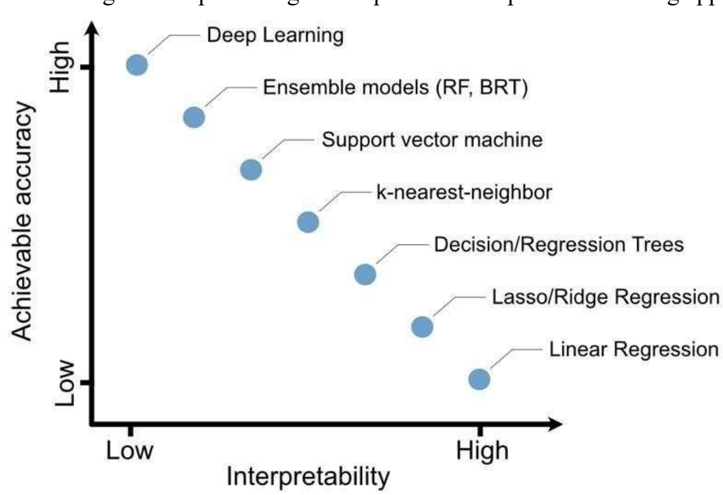

Fig. 4. Integration of ensemble models and deep learning in AI Farm Assist for precision agriculture.

III. PROPOSED SYSTEM

In modern agriculture, data tends to be highly variable and often difficult to gather. Environmental factors, soil conditions, and diverse farming practices make data collection challenging. Fortunately, in this project, structured datasets from Kaggle provide a well-organized and relatively clean source of agricultural information. These datasets cover crop recommendations, fertilizer suggestions, and leaf disease prediction, containing both structured data (environmental details, soil nutrients, crop types) and unstructured data (leaf images) suitable for machine learning.

One notable challenge with these datasets is class imbalance. Some crops, fertilizers, or disease categories are represented with fewer examples than others. This skew can affect the performance of standard predictive algorithms. Models such as CatBoost and XGBoost address this with weighting schemes, giving more importance to underrepresented classes. Similarly, MobileNetV2 uses transfer learning to handle limited image data effectively for disease classification.

A. Data Specifications

The system relies on three primary datasets, combining structured and unstructured data:

- Crop Recommendation Dataset - Contains around 2200 entries. Features include nitrogen (N), phosphorus (P), potassium (K), temperature, humidity, pH, and rainfall, while labels indicate the most suitable crop. This dataset is used to train the CatBoost model for crop selection.
- Fertilizer Recommendation Dataset - Consists of approximately 100 records. Features cover temperature, humidity, soil moisture, soil type, crop type, N, P, K levels, and fertilizer options. The XGBoost model is trained on this data to provide fertilizer recommendations tailored to crop and soil conditions.
- Leaf Disease Prediction Dataset - Comprises thousands of leaf images categorized into three disease types. This dataset supports the training of the MobileNetV2 deep learning model for early disease detection.

B. System Workflow

The workflow begins with data collection and preprocessing, including cleaning entries, normalizing values, feature extraction, and image augmentation where needed. Each dataset is then sent to its dedicated model: CatBoost handles crop recommenda-tions, XGBoost manages fertilizer suggestions, and MobileNetV2 predicts leaf diseases. Once trained and validated, outputs are integrated into a web application where farmers can input environmental parameters or upload leaf images to receive actionable guidance in real time.

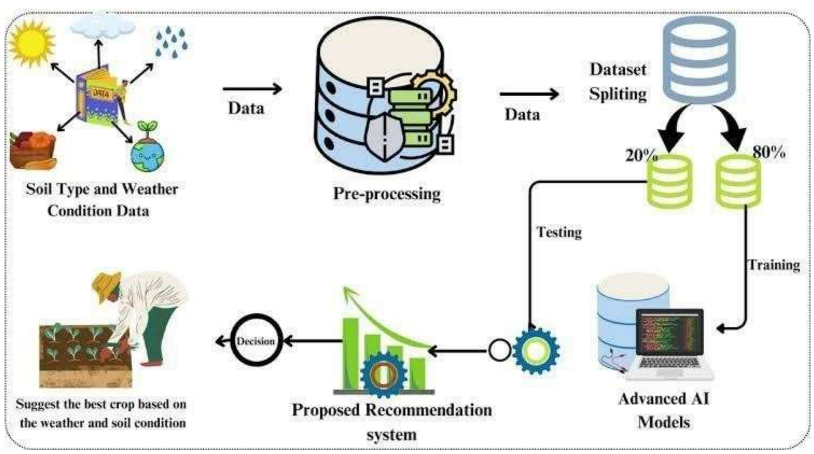

Fig. 5. System Workflow showing the pipeline from data collection to model training and providing final recommendation.

IV. METHODOLOGY

The proposed methodology for AI Farm Assist follows a structured, end-to-end pipeline designed to help farmers make informed decisions regarding crop selection, fertilizer application, and early detection of leaf diseases. The main goal is to transform raw agricultural data—both tabular and image-based—into high-quality inputs for machine learning and deep learning models, and then deploy these models in a user-friendly web application. The workflow is divided into five main phases: Data Collection and Preprocessing, Model Selection and Rationale, Model Training, Prediction and Deployment, and Advantages and Scalability.

A. Data Collection and Preprocessing

AI Farm Assist draws on well-structured datasets from Kaggle, which include tabular environmental information and images of plant leaves. The selection criteria focus on real-world relevance to farming tasks. Key factors captured in the data shape crop growth, fertilizer needs, and plant health.

Dataset Overview

- Crop Recommendation Dataset: Approximately 2200 entries with features including N, P, K, temperature, humidity, pH, and rainfall.
- Fertilizer Recommendation Dataset: Around 100 records covering soil type, crop type, moisture, and nutrients.
- Leaf Disease Dataset: Thousands of images labeled into three categories.

Preprocessing Steps

- Tabular Data: Missing values are imputed using mean, median, or mode. Numerical values are scaled, and categorical features are encoded.
- Image Data: Leaf images are resized and normalized. Data augmentation techniques—such as rotations, flips, and random zooms—are applied.

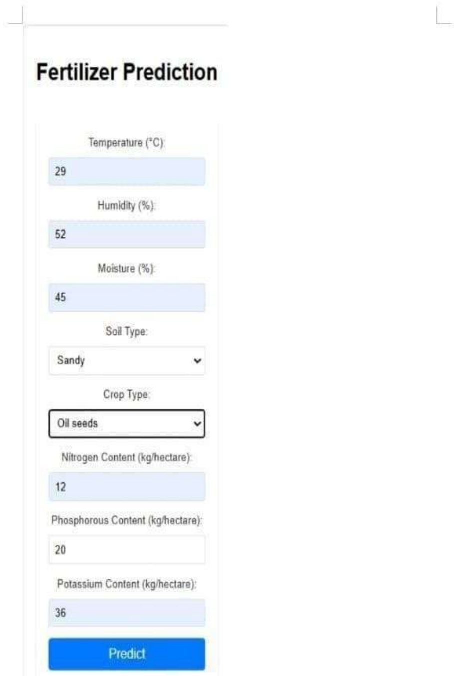

Fig. 6. This figure displays the overall fertliser recommendation page where user gives the input

B. Model Selection and Rationale

Three models are chosen, each suited to its respective task:

- CatBoost: Efficiently processes mixed numerical and categorical features for crops.
- XGBoost: Handles limited data and feature interactions for fertilizers.
- MobileNetV2: Compact CNN for disease classification with transfer learning.

C. Model Training

Training Process

- CatBoost and XGBoost: Trained on tabular datasets with hyperparameter tuning. Class imbalance is addressed using weighting.
- MobileNetV2: Trained on augmented images. Top layers are fine-tuned using categorical cross-entropy and the Adam optimizer.

D. Prediction and Deployment

The trained models are integrated into a web application where farmers can enter parameters or upload images to receive guidance in real time.

E. Advantages and Scalability

AI Farm Assist provides accurate recommendations, handles multi-modal data, and is easily scalable for future crops, fertilizers, or disease types.

F. Results

TABLE I: PERFORMANCE COMPARISON OF AI FARM ASSIST MODELS

Model Name	Accuracy	F1 Score	Recall
Crop Recommendation (CatBoost)	0.88	0.87	0.86
Fertilizer Recommendation (XGBoost)	0.85	0.84	0.83
Plant Disease Detection (MobileNetV2)	0.78	0.72	0.70

Note: The reported performance metrics are indicative and may be affected by the limited dataset size and the short training duration for each model. Extended training and larger datasets are expectedto improve accuracy, F1 score, and recall.

Fig.7. This figure showcases the output of feature where user can input an image of leaf to identify the type of disease from powdery/rust/healthy

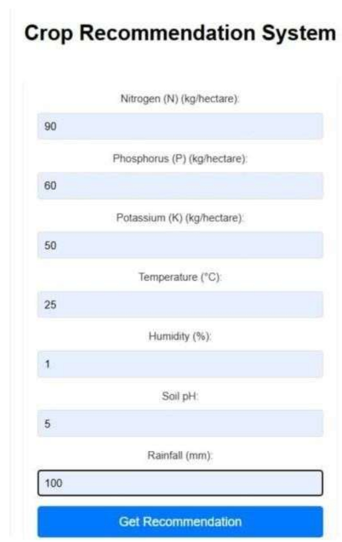

Fig.8. This figure shows the crop recommendation page where users can input the features to get the best suitable crop

V. CONCLUSION

This work presented the design, implementation, and deployment of AI Farm Assist. By integrating structured datasets on soil nutrients, environmental conditions, and plant leaf images, the system leverages machine learning and deep learning models to provide precise recommendations. The methodology demonstrates that a data-driven approach can significantly enhance traditional farming practices. The system effectively addresses class imbalance and environmental variability. Overall, AI Farm Assist highlights the promise of precision agriculture.

FUTURE WORK

The current implementation demonstrates strong potential; future efforts will focus on:

A. Hyperparameter Tuning and Model Refinement

Systematic tuning using grid or Bayesian optimization for improved accuracy.

B. Ensemble and Hybrid Approaches

Exploring stacking and blending multiple models for higher reliability.

C. Continuous Learning and Adaptation

Establishing a framework for retraining to account for evolving climatic conditions.

D. Expansion of Datasets and Crop Coverage

Incorporating larger datasets and more diverse crop categories.

E. Integration of IoT and Real-Time Data

Connecting with live soil sensors for dynamic recommendations.

F. Scalable Cloud Deployment and Mobile Access

Deploying on AWS/Google Cloud with a mobile application for rural offline support.

ACKNOWLEDGEMENT

The authors would like to extend their sincere thanks to our mentor, Dr. T. Jalaja, Assistant Professor, for her consistent encouragement, expert guidance, and whole-hearted support. We are also sincerely thankful to our guide, Dr. T. Adilakshmi, Professor and Head of the Department of CSE, for providing her valuable guidance. We also thank the administrative staff of Vasavi College of Engineering.

REFERENCES

[1] S. P. Mohanty, D. P. Hughes, and M. Salathe´, “Using Deep Learning for Image-Based Plant Disease Detection,” Frontiers in Plant Science, 2016.

[2] S.-C. S. Kamilaris and F. X. Prenafeta-Boldu´, “A Survey of Machine Learning in Agriculture,” Computers and Electronics in Agriculture, 2018.

[3] M. S. Kulkarni, et al., “Machine learning approach for crop selection based on soil and climatic parameters,” Wireless Personal Communications, 2023.

[4] J. Jenish and N. Jeyanthi, “A Review on Crop and Fertilizer Recommendation System using Machine Learning Approaches,” Journal of Physics, 2021.

[5] M. Sandler, et al., “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” in Proc. CVPR, 2018.