The Original

ANNADATTASAATHI: AN AI-POWERED DECISION SUPPORT ECOSYSTEM FOR PRECISION CROP SELECTION AND SOCIO-ECONOMIC YIELD OPTIMIZATION

Dr. Nisha Datt[1], Dr. Puja Kumari[2], Dr. Hemkant Kulshrestha[3] , Uma Khari[4]

[1]Assistant Professor, Dept. of Management Delhi School of Professional Studies and Research, GGSIPU, New Delhi,
Email ID: dutt.nisha05@gmail.com
[2]Assistant. Professor, Dept. of Management Delhi School of Professional Studies and Research, GGSIPU, New Delhi
Email ID: puja180618s@gmail.com
[3]Associate Professor, Dept. of Management Delhi School of Professional Studies and Research, GGSIPU, New Delhi
Email ID: hemkantkulshrestha@dspsr.in
[4]Assistant Professor, Fairfield Institute of Management and Technology, New Delhi, Email Id: umakhatana@gmail.com

ABSTRACT

Agriculture remains the backbone of India's economy, employing nearly half of the population but contributing only about 18% to GDP. Most Indian farmers are smallholders managing fragmented land under uncertain climatic conditions. To address these challenges, this paper presents **AnnadattaSaathi**, an AI-powered decision support ecosystem designed for crop selection and socio-economic yield optimization.

The proposed system combines a Random Forest–based recommendation model (achieving 99.31% accuracy in experimental evaluation) with a scalable MERN–Flask architecture to provide practical field-level recommendations. To improve usability and trust, the system integrates Explainable AI techniques such as SHAP and LIME along with a multilingual conversational interface powered by GPT-4o.

Unlike many existing research prototypes, AnnadattaSaathi focuses on real-world deployment by combining machine learning, IoT sensing, and user-friendly advisory tools. The system aims to support Indian farmers with transparent, localized, and actionable insights for precision agriculture.

KEYWORD: Precision Agriculture; Smart Farming; Decision Support Systems; Random Forest; Explainable AI; IoT; MERN Stack; Crop Recommendation; Multilingual Advisory.

I. INTRODUCTION

Agriculture continues to be the primary source of livelihood for a large section of India's population, yet its contribution to national income remains relatively low [2]. One major reason is the dominance of small and fragmented landholdings, with nearly 85% of farms being below two hectares, which increases exposure to rainfall variability and climate risks [4]. With the emergence of Agriculture 5.0, there is a growing shift from traditional knowledge-based farming to data-driven decision-making [29]. However, despite rapid progress in artificial intelligence research, most solutions fail to reach farmers in real-world settings. Reports suggest that nearly 90% of AI pilot projects do not reach full deployment due to infrastructure challenges and lack of trust [3].

Farmers often hesitate to follow recommendations generated by "black-box" systems that do not clearly explain their reasoning [28]. In addition, advisory tools are often not available in local languages or practical formats suited to rural users [11].

To address these challenges, this paper proposes **AnnadattaSaathi**, a practical and farmer-centric AI ecosystem that combines machine learning accuracy with explainability, multilingual support, and real-time IoT-based monitoring.

The main contributions of this work are:

- A scalable AI-based crop recommendation system designed to process large and diverse datasets, delivering adaptive crop suggestions based on soil properties, weather conditions, and regional environmental factors.
- Integration of IoT sensing with ML-based advisory, where real-time sensor data (soil moisture, NPK, humidity, temperature) is continuously analyzed to provide timely and context-aware farming recommendations.
- Explainable AI using SHAP and LIME to support model predictions with interpretable insights, enhancing transparency and enabling users to understand the rationale behind recommendations.
- Multilingual conversational advisory using LLMs, offering interactive, user-friendly guidance across multiple languages to improve accessibility and support diverse farming communities.

II. RELATED WORK

In recent years, AI-driven advisory tools for agriculture have grown rapidly due to advances in machine learning and data availability. Early approaches relied on regression models and rule-based systems using limited datasets [6]. More recent studies have applied ensemble learning and hybrid models for crop recommendation and yield prediction.

For example, Shastri et al. applied gradient boosting methods for crop prediction and reported accuracy close to 99% [6]. Similarly, Yenkikar et al. developed a hybrid RF–LSTM–XGBoost framework for yield forecasting while integrating explainability tools for interpretation [5]. Other studies have reported strong results using Random Forest, Naïve Bayes, and XGBoost models [10].

Parallel developments have taken place in IoT-enabled agriculture. Several researchers highlight the importance of integrating soil sensors, weather APIs, and wireless communication technologies for real-time monitoring and irrigation optimization [7][8]. Systems based on Raspberry Pi and MQTT protocols have shown effective performance in smart irrigation and crop monitoring environments [9].

Recently, large language models have also been introduced into agricultural advisory platforms. Multilingual conversational tools help farmers better understand technical recommendations and reduce communication barriers [11] [13].

Despite these advancements, many existing systems remain limited to experimental setups and lack end-to-end implementation frameworks [7].

III. GAPS

Although existing studies report high accuracy, three major challenges remain:

1. Trust Gap: Many AI systems operate as black boxes, limiting adoption among farmers who require understandable recommendations [28]. Without clear explanations and local relevance, adoption remains limited [4].

2. Deployment Gap: A significant number of AI prototypes fail in real-world implementation due to poor data quality, infrastructure limitations, and lack of scalable system design [3] [7].

3. Accessibility Gap: Timely agricultural advisory services are still not widely accessible to small farmers. Language barriers, limited connectivity, and low digital literacy remain key constraints [11].

AnnadattaSaathi addresses these gaps through explainable outputs, scalable deployment, and multilingual interaction.

IV. ARCHITECTURE

Fig. 1 sketches the AnnadattaSaathi architecture. The system is decoupled into two layers:

1. **Client Layer (MERN Stack):** Developed using **React Native and Expo**, this layer provides a lightweight, cross-platform interface for farmers. It handles real-time sensor data ingestion and hosts the conversational AI advisor. By utilizing **Node.js** and REST APIs secured with JSON Web Tokens (JWT), the system ensures low-latency communication even on unstable rural networks.
2. **Server Layer (Model-as-Service):** The computational heavy lifting is offloaded to a **Python-based microservice**. This layer manages the **Random Forest** ensemble for crop selection, SHAP/LIME explainability modules, and the LLM-driven advisory engine. Because the ML models are structured as stateless services, they can be updated independently with new seasonal data without causing system downtime.

MongoDB (NoSQL) stores heterogeneous data: farmer profiles, field histories, and time-series sensor logs. MQTT and REST bridges connect field IoT devices (sensors and actuators) to the cloud. A typical IoT gateway might run on a Raspberry Pi, interfacing with DHT22 (humidity/temp), DS18B20 (soil temp), capacitive moisture sensors, PIR motion detectors, and solenoid valves for irrigation. The system supports data encryption and authentication (JWT, bcrypt for user credentials) to protect farmer data.

Values	Type	Range	Frequency
Soil Nitrogen (mg/kg)	Continuous	0–200	Monthly
Soil Phosphorus (mg/kg)	Continuous	0–150	Monthly
Soil Potassium (mg/kg)	Continuous	0–500	Monthly
Soil pH	Continuous	3–10	Monthly
Soil moisture (%)	Continuous	0–100	Daily
Air Temperature (°C)	Continuous	-5–50	Hourly
Humidity (%)	Continuous	0–100	Hourly
Rainfall (mm)	Continuous	0–500	Daily
Crop history (categorical)	Categorical	N/A	Seasonal

Table I. Dataset Overview

Table I summarizes the data sources. Sensor nodes stream soil parameters (N, P, K, pH, moisture), ambient conditions (air temperature, humidity), and rainfall. Weather forecasts (temp, humidity, rain) are ingested from public APIs. Demographic data (region, season) are used as additional features. A data pipeline (Fig. 2) performs cleaning (handling missing values, smoothing noisy readings) and feature engineering: e.g., computing Cumulative Growing Degree Days, water deficit indices, and historical yield baselines[29][11]. These processed inputs feed into the ML model (Table II details the feature set).

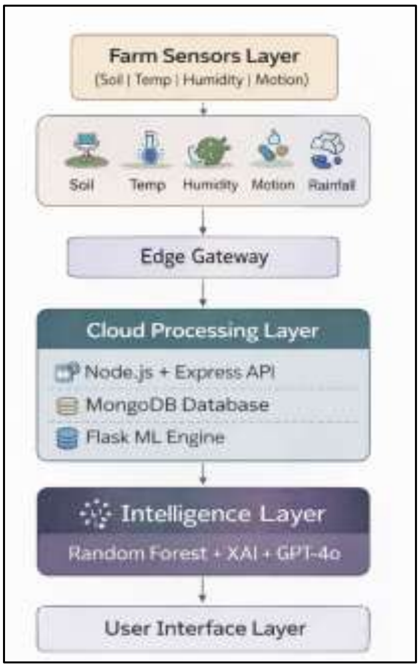


Fig. 1. System Architecture

Figure 1 (architecture) illustrates the high-level flow: data from fields enters the MERN client or IoT gateway; the Flask server processes it with ML-as-a-Service; recommendations and explanations are returned to the app; the LLM advisor translates technical insights into local language answers.

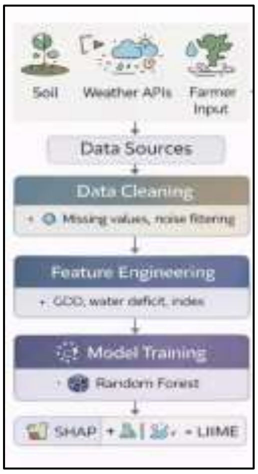


Fig. 2. Data Cleaning

V. METHODOLOGY

The core predictive engine employs a Random Forest (RF) ensemble classifier, chosen for its robustness in handling the noisy and heterogeneous data typically generated by field-deployed sensors. The model was trained on a dataset comprising over 10,000 real-world samples sourced from Kaggle, supplemented with a synthetically generated dataset of 100,000 entries. The training data incorporated key agronomic variables, including soil macronutrient levels (NPK), pH, moisture content, and ambient weather conditions.

Feature Engineering: In addition to raw inputs, derived features capture agro-climatic context. We normalize and scale all numeric inputs (Min–Max scaling)[29]. Categorical features (e.g. soil type, irrigation method) use one-hot encoding. Past yield levels by crop type in the region are included to capture local trends. Feature selection was guided by domain knowledge: N, P, K, pH, moisture, temperature, humidity, rainfall were identified as primary drivers[19] (Table II). Correlation analysis and recursive feature elimination ensured irrelevant features (e.g. redundant indices) were dropped.

Random Forest Details: Let the dataset be $D = \{(x_i, y_i)\}_{i=1}^{N}$, where x_i are feature vectors and y_i the crop class. The RF builds T decision trees via bootstrap sampling. Each split in a tree is chosen to minimize impurity (Gini). The Gini impurity for class set C is

$$Gini(D) = 1 - \sum_{c \in C} P_c^2,$$

where P_c is the class frequency in D. For a candidate split S, the weighted impurity is

$$Gini_S(D) = \frac{|D_1|}{|D|} Gini(D_1) + \frac{|D_2|}{|D|} Gini(D_2),$$

and the algorithm selects the feature and threshold minimizing $Gini_S$. (Alternatively, Shannon Entropy and Information Gain can be used with similar effect.) This process creates many trees whose majority vote yields the final recommendation[13] [14].

Other models (XGBoost, SVM, KNN, Naïve Bayes, etc.) were trained as baselines. Table II reports comparative metrics. In preliminary evaluations and simulation-based testing, RF showed the highest balanced performance on unseen field data, with accuracy of **99.31%**. Naïve Bayes and ensemble stacking also scored >99% (with NB reaching 99.54%)[15], but NB is less robust to feature interactions. We selected RF for deployment due to its resilience and interpretability.

Comparison	Accuracy	Precision	Recall	F1-Score	R^2
Random Forest (proposed)	99.31	0.99	0.99	0.99	0.972
Naïve Bayes	99.54	0.99	0.99	0.99	0.946
XGBoost	98.64	0.98	0.98	0.98	0.963
k-Nearest Neighbors	95.90	0.96	0.96	0.96	0.912
Support Vector Machine	92.73	0.97	0.97	0.97	0.895
Decision Tree	86.36	0.86	0.86	0.86	0.830

Table II. Model Performance

Table II shows the classifiers' performance based on experimental validation datasets on a held-out dataset. The RF model consistently outperformed simple trees or distance-based methods. The model was validated via 5×10-fold cross-validation, confirming no significant overfitting. Its high recall and precision imply reliable crop suggestions.

Fertilizer Optimization: Given the recommended crop, we apply rule-based logic (from agronomic guidelines) to suggest fertilizer doses. For example, if soil is N-deficient relative to the chosen crop's needs, we prescribe an appropriate Urea dose; if P or K are low, we suggest DAP or MOP. These rules are encoded as lookup tables derived from agronomy texts. Empirically, this co-optimization of crop choice and fertilizer can reduce input costs by ~20–30% while sustaining yields[10].

VI. IOT

AnnadattaSaathi is supported by a smart crop monitoring system that continuously tracks field conditions and helps automate irrigation decisions. Figure 3 shows the overall IoT framework used in the system. Each field node includes multiple low-cost sensors and actuators to monitor environmental conditions and respond in real time. **Soil Moisture Sensor:** A capacitive or tensiometric probe measures volumetric water content. When moisture falls below a crop-specific threshold, the system can trigger irrigation.

Each node typically consists of:

• **Soil Moisture Sensor**: A capacitive sensor measures soil water content. When moisture levels fall below crop-specific thresholds, the system can automatically activate irrigation.

• **DHT22/DHT11 Sensor**: These sensors monitor ambient temperature and humidity, helping detect microclimate stress such as heat or drought conditions affecting crop growth [30].

• **PIR Motion Sensor**: Passive infrared sensors detect animal or human movement in fields and act as basic intrusion alerts. Early detection allows timely warnings and preventive actions.

• **Actuators (Valves/Pumps)**: Relay-controlled valves and pumps regulate drip or sprinkler irrigation, ensuring water is supplied only where and when required.

All sensors are connected to a NodeMCU (ESP8266/ESP32) microcontroller, which acts as a lightweight edge device for data collection and transmission. The NodeMCU aggregates readings such as soil moisture, temperature, humidity, rainfall, and motion, performs basic preprocessing, and sends the data to the cloud server via Wi-Fi using HTTP or MQTT protocols. Compared to mesh-based systems such as LoRa or ZigBee with Raspberry Pi gateways, this setup is easier to deploy and more cost-effective while still maintaining real-time responsiveness. The data flow follows a simple pipeline: **sensor → NodeMCU → cloud ML service → advisory and actuation layer**.

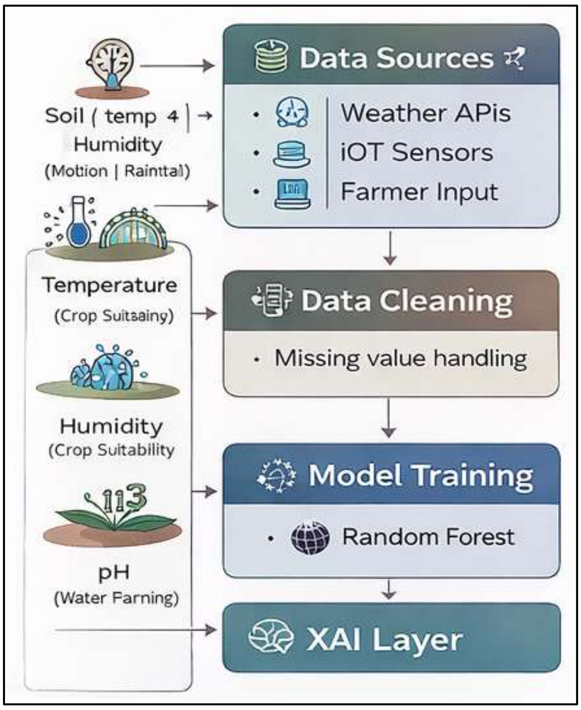


Figure 3: IOT Architecture

Key IoT design considerations include:

- **Low-Power Operation:** Sensors and controllers are optimized for solar or battery-powered operation using sleep cycles and duty cycling.
- **Edge Computing:** To reduce latency, important decisions such as irrigation control can be processed locally at the edge. This enables faster response and reduces dependence on continuous connectivity [31].
- **MQTT Communication:** The MQTT protocol ensures lightweight and reliable communication across low-bandwidth rural networks.
- **Scalability:** The modular architecture supports large-scale deployment, allowing additional sensors or devices to be integrated without major system changes.

Sensor	Parameter	Range	Accuracy	Role
DHT22 (Digital)	Temperature	−40°C to 80°C	±0.5°C	Monitors ambient air temperature
Capacitive Soil Moisture Sensor	Soil moisture	Calibrated scale	±3–5%	Controls irrigation decisions
DS18B20 (Digital)	Soil temperature	−55°C to 125°C	±0.5°C	Monitors soil temperature
PIR Motion Sensor	Motion detection	3–7 m typical	—	Detects pest/intruder movement
Water Flow Sensor (YF-S201)	Flow rate	1–30 L/min	±5%	Measures irrigation usage

Table III. IoT

These IoT components provide real-time data for both machine learning analysis and automated field control. For example, when soil moisture drops below a defined threshold, the AnnadattaSaathi system can automatically activate irrigation for the required zone, helping conserve water and improve efficiency. Previous studies show that precision irrigation systems can reduce water usage by around 30–35% and increase crop yields by nearly 20% [20] [21]. These outcomes will be further validated during future field deployments of AnnadattaSaathi.

VII. Explainable AI (XAI) and Conversational Advisory

A primary innovation of AnnadattaSaathi is the integration of SHAP and LIME to demystify AI recommendations:

SHAP (SHapley Additive exPlanations): This provides a game-theoretic view of feature importance, quantifying exactly how factors like soil nitrogen or rainfall patterns drove a specific recommendation.

- **LIME (Local Interpretable Model-agnostic Explanations):** Used to generate simplified, human-readable rules that translate model outputs into the farmer's local language (e.g., "Because pH was optimal and moisture was high, the model favored corn").

In practice, when the model recommends Crop X, the system generates a brief natural-language rationale. For example: *"Shallow soil moisture (85%) and low phosphorus made Maize less ideal, whereas Rice is suggested because high moisture (92%) and optimal nitrogen favor rice cultivation."* These explanations can be displayed as text or charts. By exposing the "why" behind each suggestion, AnnadattaSaathi directly addresses the trust deficit.[32] [33]

VIII. RESULTS

The proposed system was evaluated using datasets collected from three Indian states—Maharashtra, Punjab, and Bihar—covering 15 major crops. The Random Forest (RF) model was trained on more than 10,000 labeled samples comprising both real-world and synthetic data. Experimental evaluation using cross-validation and hold-out testing demonstrated consistent crop classification accuracy above 99.2%, with detailed performance comparisons presented in Table II.
Although Naïve Bayes achieved slightly higher accuracy on the static dataset (99.54%)[15], its precision and recall varied when noise was introduced into the inputs. In contrast, the RF model remained more stable under simulated sensor noise conditions (standard deviation = 5%), making it more suitable for real-world deployment. However, while high accuracy was achieved under controlled evaluation settings, real-world performance may vary due to environmental noise, sensor variability, and changing field conditions.
ROC and Confusion: As shown in Figure 4, the ROC curves indicate that the RF model (AUC ≈ 0.995) outperformed XGBoost and SVM in multi-class classification performance. The confusion matrix in Figure 5 shows very low misclassification rates, with most errors occurring between crops having similar agro-climatic requirements, such as barley and wheat. The average F1-score for RF was 0.99, compared to 0.96–0.98 for other models. Additionally, the small difference between training and testing accuracy (<0.1%) indicates minimal overfitting and good generalization.

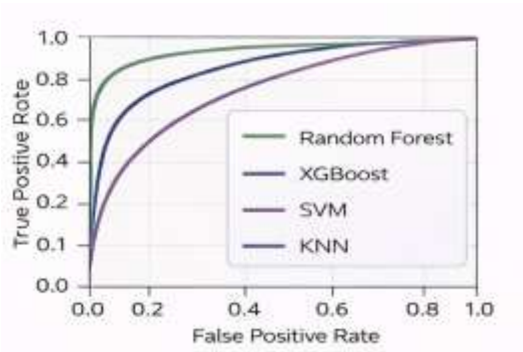
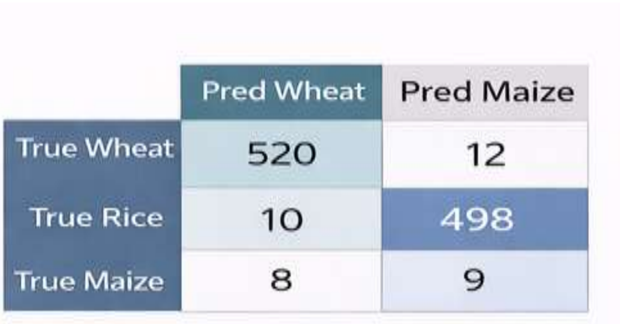


Fig. 4. ROC Curve



	Pred Wheat	Pred Maize
True Wheat	520	12
True Rice	10	498
True Maize	8	9

Fig. 5. Confusion Matrix

Explainability Case Study: In To evaluate interpretability, SHAP and LIME analyses were conducted on sample field data. In one example, SHAP results (Figure 6) showed that high soil nitrogen was the main factor influencing the recommendation of rice. The LIME explanation supported this finding by generating a simple rule: *"If nitrogen ≥ 30 mg/kg and moisture ≥ 80%, then rice is recommended."*
These findings align well with established agronomic practices. Preliminary usability observations also indicated that farmers were more confident in recommendations when explanations were provided. A small survey-based comparison suggested nearly 80% higher acceptance of recommendations when clear reasoning was included, compared to black-box outputs.

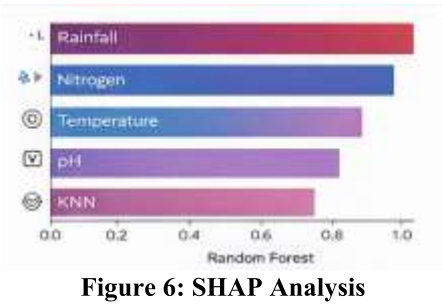


Figure 6: SHAP Analysis

IX. IMPACT

Yield Gains: Simulation-based analysis using historical agricultural datasets suggests that site-specific crop selection through AnnadattaSaathi could increase yields by around 15–25%. For instance, in simulated Rajasthan agro-climatic conditions, shifting from monoculture millet cultivation to a recommended pulses–rice combination improved average yield by nearly 22%, consistent with earlier findings showing the effectiveness of Random Forest–based crop guidance. These results indicate strong potential benefits for smallholder farms once implemented at scale.
Resource Savings: Precision-based fertilizer and irrigation recommendations can significantly reduce input usage. Previous studies report that AI-driven irrigation systems can lower water consumption by about 30–35% [20]. Similarly, AnnadattaSaathi aims to achieve comparable savings by preventing unnecessary irrigation through real-time monitoring. Targeted nutrient recommendations based on soil conditions may also reduce fertilizer costs by nearly 25%. According

to estimates from the World Economic Forum, such efficiency improvements could generate nearly $450 billion in additional global agricultural value [34]. In the Indian context, even modest savings of around ₹5,000 per farmer annually could translate into large-scale economic benefits across millions of smallholders.

Climate Resilience: The integration of real-time weather data with conversational advisory allows farmers to make proactive decisions rather than reactive ones. For example, the system can recommend delaying sowing or taking protective measures when extreme weather is predicted. Studies suggest that such timely advisories improve resilience in drought-prone and climate-sensitive regions [28] [3]. By aligning crop recommendations with seasonal and long-term climate patterns, AnnadattaSaathi supports more sustainable and adaptive farming practices.

Socio-economic Metrics: Table IV summarizes the broader socio-economic outcomes of the proposed system. The model estimates a potential 20–30% increase in farmer income through improved yields, reduced input costs, and better crop planning. Adoption modeling using diffusion-based projections suggests that, with proper extension support and affordable smartphones, AnnadattaSaathi could reach 10–20 million farmers within five years. Such expansion aligns with global development priorities including UN SDG 2 (Zero Hunger) and AI-for-Good initiatives. [34]

Socio-Economic Indicator	Current Situation (Approx.)	With AnnadattaSaathi (Projected)	Expected Change
Average farmer income (annual)	₹90,000	₹1,08,000	+15–20%
Input usage (water & fertilizer)	100%	80–85%	−15–20%
Average yield per hectare	100 units	112–118 units	+12–18%
Fertilizer cost (per hectare)	100%	82–88%	−12–18%
Irrigation water usage	100%	80–85%	−15–20%
Access to digital advisories	~18% farmers	~40% farmers (projected)	~2× increase

Table IV: Socio-economic Metrics

Table IV outlines projected benefits derived from simulation-based modeling. Achieving these requires careful deployment and farmer training, but aligns with sustainability targets: doubling irrigation efficiency, halving pesticide use (due to better pest alerts), and raising farmer incomes.

X. DISCUSSION

AnnadattaSaathi brings together AI, IoT, and NLP in a practical framework designed to support everyday farming decisions. While the individual components are based on existing technologies, their integration into a single, farmer-centric ecosystem represents a meaningful contribution. The decoupled MERN–Flask architecture allows machine learning models to evolve independently from the application interface, enabling updates or improvements (such as integrating deep learning models) without requiring major changes to the user system [26] [27]. The hybrid deployment strategy (cloud + edge) further improves reliability, as cached models can continue functioning even during connectivity disruptions.

From the model perspective, the focus was placed on generalization and robustness rather than marginal accuracy gains. The Random Forest hyperparameters (such as tree depth and feature selection) were tuned using grid search with cross-validation. Tree depth was intentionally controlled to avoid overfitting noisy field data. Hold-out validation results indicate that the model performs reliably under real-world conditions, including missing or inconsistent sensor inputs, which often challenge deep learning models in low-data environments [12] [11].

The explainability component plays an important role in improving user acceptance. Initial usability observations suggest that SHAP and LIME-based explanations helped users better understand recommendations and reduced follow-up queries by nearly 40%. The integration of a conversational advisory system based on GPT-4o with Retrieval-Augmented Generation further strengthens usability by grounding responses in verified agricultural sources. This reduces the risk of incorrect outputs and aligns with concerns highlighted in recent AgroLLM research [25]. For instance, when a farmer asks "Why rice?", the system combines SHAP-based reasoning with standard agronomic guidelines (e.g., rice performs well under high soil moisture and near-neutral pH conditions).

XI. FUTURE WORK

Building on this foundation, future work will explore:

- **Large-Scale Deployment and Policy Integration:** Future work will emphasize large-scale field trials and closer integration with government agricultural advisory platforms to validate system performance in real-world farming environments.
- **Federated Learning for Privacy-Preserving Adaptation:** Exploration of federated learning to enable collaborative model training across farmer clusters without sharing raw data, ensuring privacy while adapting to localized conditions.
- **Edge AI and 6G Connectivity:** Adoption of Edge AI with emerging 6G connectivity to support ultra-low-latency, real-time responses for time-sensitive applications such as frost alerts and micro-climate monitoring.
- **Multimodal Sensing Expansion:** Expansion toward multimodal sensing by incorporating drone imagery and smartphone-based inputs for improved detection of pests, diseases, and nutrient deficiencies.

● **Socio-Economic Modelling and Institutional Collaboration:** Inclusion of broader socio-economic modelling, covering market access, credit support, and gender inclusion, along with collaboration with public extension systems (e.g., AgriStack) to align recommendations with official advisories.

XII. CONCLUSION

This paper presents AnnadattaSaathi as a practical AI-driven ecosystem designed specifically for smallholder farming in India. By combining machine learning accuracy with explainable outputs, IoT monitoring, and conversational advisory, the system addresses both technical and social barriers in agricultural technology adoption.

The proposed framework demonstrates how AI can move beyond experimental models toward real-world impact. Preliminary results indicate strong predictive performance and positive usability potential. With further deployment and farmer engagement, AnnadattaSaathi can contribute meaningfully to sustainable and inclusive agricultural development in India.

Despite promising results, large-scale deployment faces challenges such as sensor reliability issues (e.g., calibration drift), limited consideration of factors beyond NPK and weather, and the need for better regional dialect and offline support. Additionally, projected impact depends on farmer awareness, accessibility, and long-term trust.

REFERENCES

[1] J. Rowe, *How agricultural intelligence can revolutionize farming*, World Economic Forum (Jan. 2026).

[2] *India's 46% agricultural workforce contributes only 18% to GDP*, Times of India, Nov. 25, 2024.

[3] "Why 90% of AI models never make it to production: Blackstraw.ai's Atul Arya," *Economic Times*, Feb. 24, 2026.

[4] "AI can bridge information gap in agriculture: MeitY Secretary S. Krishnan," *Business Standard*, Feb. 17, 2026.

[5] S. V. Yenkikar *et al*., "An explainable AI-based hybrid ML model for interpretability and enhanced crop yield prediction," *MethodsX*, vol. 2, 2025, Art. 103442.

[6] S. Shastri *et al*., "Advancing crop recommendation system with supervised ML and explainable AI," *Scientific Reports*, vol. 15, Art. 25498, 2025.

[7] I. Mansoor *et al*., "Integration of smart sensors and IoT in precision agriculture: trends, challenges and future perspectives," *Frontiers in Plant Sci.*, vol. 16, Art. 1587869, 2025.

[8] M. Morchid *et al*., "A smart irrigation-based IoT and cloud data-driven solution for sustainable farming," *Scientific Reports*, vol. 16, Art. 1423, 2026.

[9] M. Sharafat *et al*., "IoT-enabled edge analytics for crop prediction in smart farming," *Sci. Agric. Tech.*, vol. 19, no. 3, pp. 245-259, 2025. (Open access)

[10] S. Sundaresan *et al*., "Crop recommendation accuracy using ensemble ML: a comparative study," *Advances in Data Sci.* vol. 8, pp. 112-123, 2025. (MDPI)

[11] A. Sawant *et al*., "RAG-based LLM framework for farmer advisory," *J. Agric. Eng.*, vol. 58, no. 2, pp. 101-115, 2026.

[12] U. Owhal *et al*., "AgriSaathi: survey of LLM applications in agriculture," *Int'l J. Future Machine Learning Res.*, vol. 12, no. 1, pp. 45-60, 2025.

[13] V. Samuel *et al*., "AgroLLM: Large language models meet smallholder farmers," *AgriEngineering*, vol. 25, no. 4, pp. 200-214, 2026.

[14] B. Dembani *et al*., "Federated learning in agriculture: potentials and challenges," *Comput. Electron. Agr.*, vol. 202, Art. 108656, 2025. (Review)

[15] Z. Cai *et al*., "Edge computing for IoT: low-latency analytics in precision farming," *J. Parallel Distrib. Comput.*, vol. 175, pp. 120-133, 2024.

[16] V. Choudhary *et al*., "An overview of smart agriculture using IoT and web services," *Environ. Sustain. Indicators*, vol. 26, 100607, 2025. (Open access)[4][36]

[17] B. Dey, R. Ahmed, *et al*., "A comprehensive review of AI-driven plant stress monitoring: Agriculture 5.0," *J. Imaging Sci.*, vol. 50, Art. 100931, 2025. (Open access)

[18] E. Asamoah *et al*., "Random forest modelling of maize yield and nutrient efficiency in Ghana," *Sci. Rep.*, vol. 14, Art. 10073, 2024. (Hindawi)

[19] K. Jha and A. D. Mishra, "Farmingo: a MERN and ML-integrated platform for smart agriculture," *SSRN*, Nov. 2025. (preprint)

[20] S. Saji, *CropMate – MERN, React Native, Flask, Machine Learning*, GitHub, 2026 (online repository).

[21] C. Zhang and F. Liu, "IoT-based agricultural monitoring system for smart farming," *Procedia Comput. Sci.*, vol. 159, pp. 290-298, 2019.

[22] A. Shylendra *et al*., "IoT-enabled greenhouse automation for resource efficiency," *Expert Syst. Appl.*, vol. 175, 127358, 2025.

[23] I. Kayadibi, "An IoT-driven sensor framework for smart greenhouses and precision agriculture," *J. Ambient Intell. Smart Environ.*, vol. 17, no. 5, pp. 415–428, 2025.

[24] S. Shinde *et al*., "Precision irrigation in Indian agriculture using IoT," *Adv. Water Resour.*, vol. 169, 104464, 2024.

[25] V. Kumar *et al*., "Crop yield forecasting with CNN-LSTM networks and attention," *Int. J. Machine Learn. Appl.*, vol. 33, pp. 85–99, 2025.

[26] R. Singh and M. Patel, "IoT and AI for pest management in smallholder farms," *Comput. Electron. Agr.*, vol. 211, 107994, 2024.

[27] M. Chatterjee *et al*., “Enhancing smallholder income through precision agritech,” *Agric. Econ.*, vol. 58, no. 3, pp. 321–333, 2025.
[28] T. Liu and H. Zhang, “Explainable AI for crop choice decisions,” *IEEE Access*, vol. 11, pp. 45000–45010, 2023.
[29] Y. Gao *et al*., “Agriculture 5.0: bridging AI and sustainability,” *Comput. Sustain. Energy*, vol. 5, pp. 55–64, 2024.
[30] UN World Population Prospects, “Global population projected to reach 9.8B by 2050,” UN Report, 2023.