www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 10, October 2025, pp : 1377-1395

e-ISSN : 2583-1062
Impact Factor : 7.001

AGROMIND - AI IN SMART AGRICULTURE: A REVIEW

Abhijit Manohar Shinde[1], Dhiraj Dinesh Shewale[2], Kalyani Punjahari Navle[3], Aniket Vijay Pawar[4], Trupti U. Ahirrao[5]

[1,2,3,4]Student, Computer Engineering, Matoshri College Of Engineering And Research Center, Nashik, Maharashtra, India.

[5]Assistant Professor, Computer Engineering, Matoshri College Of Engineering And Research Center, Nashik, Maharashtra, India

ABSTRACT

The agricultural sector faces unprecedented challenges including climate variability, resource constraints, and the need to meet growing food demands. Traditional farming practices often fall short in addressing these complexities, necessitating data-driven, intelligent solutions. This literature survey provides a comprehensive review of state-of-the-art approaches in AI-driven agricultural decision support systems, specifically examining crop yield prediction, crop recommendation systems, disease detection, fertilizer recommendation, crop rotation planning, and market price forecasting. Drawing from recent research publications (2018-2025), this paper analyzes existing methodologies, identifies their strengths and limitations, and proposes pathways for developing AgroMind—an integrated AI-powered agricultural decision support system that leverages machine learning, deep learning, generative AI, and explainable AI to empower farmers with comprehensive, multilingual agricultural guidance.

Keywords: Precision Agriculture, Crop Yield Prediction, Crop Recommendation Systems, Disease Detection, Explainable AI, Generative AI, Market Price Forecasting, Sustainable Agriculture.

1. INTRODUCTION

Agriculture serves as the backbone of global food security and economic stability, particularly in developing nations where a significant portion of the population relies on farming for livelihood[1][3][4]. However, traditional agricultural practices face mounting challenges including unpredictable climate patterns, soil degradation, water scarcity, pest infestations, market volatility, and inefficient resource utilization[4][6]. These challenges are compounded by the increasing global population, which is projected to reach 9.7 billion by 2050, necessitating a 70% increase in food production[6].

The advent of cutting-edge technologies such as Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Internet of Things (IoT), Cloud Computing, and Blockchain has catalyzed a paradigm shift toward precision agriculture or smart farming[3][4][6]. These technologies enable data-driven decision-making by analyzing complex agricultural data including soil properties, weather patterns, crop performance, market trends, and historical yields[3][4][6].

1.1 Motivation and Research Gap

Despite significant advancements in agricultural AI systems, several critical gaps persist:

1. **Fragmentation of Solutions**: Most existing systems address individual agricultural challenges (crop recommendation, disease detection, or price forecasting) in isolation rather than providing integrated, holistic solutions[6][7][9].
2. **Lack of Transparency**: Traditional ML models often function as “black boxes,” making it difficult for farmers to understand and trust AI-driven recommendations[4][7][9].
3. **Limited Multilingual Support**: Many agricultural decision support systems lack adequate multilingual capabilities, limiting accessibility for farmers in linguistically diverse regions[27][30][36].
4. **Insufficient Crop Rotation Planning**: While crop recommendation systems exist, few incorporate sustainable crop rotation strategies considering soil health and long-term productivity[6].
5. **Market Integration Gaps**: Limited integration of real-time market data and price forecasting hampers farmers’ ability to make economically informed decisions[2][7][26].

1.2 Contribution and Scope

This comprehensive literature survey examines research from 2018-2025 across eight critical dimensions of agricultural AI systems:

1. **Crop Yield Prediction** using ML regression and DL techniques
2. **Crop Recommendation Systems** utilizing soil nutrients, climate data, and historical yields

3. **Crop Disease Detection** employing Computer Vision and Deep Learning
4. **Fertilizer Recommendation** based on soil composition and crop requirements
5. **Crop Rotation Planning** for sustainable agriculture
6. **Market Price Forecasting** using time-series analysis and Generative AI
7. **Explainable AI (XAI)** for transparent decision-making
8. **Multilingual Support** for farmer accessibility

The findings will inform the development of **AgroMind**, an integrated AI-powered agricultural decision support system addressing the identified gaps through multimodal data fusion, explainable recommendations, and comprehensive agricultural guidance.

2. CROP YIELD PREDICTION: STATE-OF-THE-ART

Crop yield prediction is fundamental to agricultural planning, enabling farmers, policymakers, and stakeholders to make informed decisions regarding resource allocation, market strategies, and food security[1][3].

2.1 Traditional Statistical Methods

Early crop yield prediction relied on traditional statistical approaches:

- **Linear Regression Models**: Gao et al. (2022)[8] demonstrated that linear regression models could predict crop yield based on climate and soil data, achieving reasonable accuracy for datasets with linear relationships. However, these models struggle with complex, non-linear agricultural patterns[1][3].
- **Multiple Linear Regression (MLR)**: Rao et al. (2016) applied MLR for climate-based yield prediction, but noted limitations in capturing interactions between multiple variables[1][6].

Limitations: Traditional methods predict single sample spaces, fail to capture complex variable interactions, and lack adaptability to dynamic agricultural conditions[1][2].

2.2 Machine Learning Approaches

2.2.1 Random Forest Regression

Random Forest has emerged as one of the most effective algorithms for crop yield prediction:

- Chen et al. (2023)[9] used Random Forest regression for maize yield prediction with climate and soil data, achieving 67.80% accuracy[2]. The ensemble approach reduces overfitting and effectively captures non-linear relationships[2][3].
- Medar et al. (2019)[1] compared Naive Bayes and K-Nearest Neighbors (KNN) methods, achieving 91.11% accuracy with appropriate feature selection and cross-validation techniques.
- Badshah et al. (2024)[3] demonstrated Random Forest's superiority in crop classification with 99.7% accuracy through K-fold cross-validation and feature engineering.

Strengths: Handles large datasets, provides feature importance insights, robust to noise and outliers, manages both regression and classification tasks[2][3].

Weaknesses: Computationally intensive with large datasets, may overfit with noisy data without proper tuning[2][3].

2.2.2 Support Vector Regression (SVR)

Zhang et al. (2022)[10] demonstrated SVR's effectiveness in handling non-linear correlations between meteorological and soil variables for crop yield prediction. Badshah et al. (2024)[3] achieved 99.9% R^2 score for wheat yield prediction in Pakistan using hyperparameter-tuned SVR with 5-fold cross-validation.

Strengths: Effective in high-dimensional spaces, handles non-linear relationships through kernel functions, robust to outliers[3][10].

Weaknesses: Requires careful kernel selection and hyperparameter tuning, computationally expensive for large datasets[3].

2.2.3 Gradient Boosting and XGBoost

Khan et al. (2024)[12] employed Gradient Boosting Regression for rice yield prediction, effectively capturing complex patterns in agricultural data. Sarangi et al. (2024)[2] compared multiple algorithms, finding that ensemble methods like Gradient Boosting Machine achieved 97.96% accuracy for cereal price prediction.

Strengths: Sequential learning corrects previous iterations' errors, incorporates regularization, handles missing values effectively[2][12].

Weaknesses: Prone to overfitting without proper regularization, requires extensive hyperparameter tuning[2].

2.3 Deep Learning Approaches

2.3.1 Neural Networks and LSTM

- Liu et al. (2022)[16] used deep learning models for wheat yield prediction, demonstrating neural networks' ability to comprehend intricate relationships in agricultural data.
- Sharma et al. (2024)[18] employed Long Short-Term Memory (LSTM) networks for temporal crop yield prediction, highlighting the advantages of recurrent neural networks in handling sequential agricultural data.
- Mateo-Sanchis et al. (2023)[4] developed Interpretable LSTM networks for crop yield estimation, addressing the interpretability challenge of deep learning models.

Strengths: Captures temporal dependencies, handles sequential time-series data effectively, learns complex non-linear patterns[4][16][18].

Weaknesses: Requires large training datasets, computationally intensive, limited interpretability (mitigated by XAI techniques)[4].

2.3.2 Convolutional Neural Networks (CNN)

Nguyen et al. (2023)[17] utilized CNNs for rice yield prediction with satellite imagery, demonstrating the value of incorporating image-based data. Nejad et al. (2023)[16] employed 3D-CNNs with attention mechanisms for multispectral crop yield prediction.

Strengths: Effective for spatial data and image analysis, extracts hierarchical features automatically[16][17].

Weaknesses: Requires large labeled datasets, computationally expensive[16][17].

2.4 Data Sources and Feature Engineering

Recent research emphasizes the importance of multimodal data integration:

- **Satellite Data**: Remote sensing and vegetation indices (NDVI, EVI) provide large-scale crop monitoring capabilities[4][16][17].
- **Weather Data**: Temperature, rainfall, humidity, solar radiation significantly influence crop growth[1][2][3][6].
- **Soil Data**: pH levels, nutrient content (NPK), moisture, organic matter are critical predictors[1][3][4][6].
- **Historical Yield Data**: Past performance guides future predictions and enables trend analysis[3][4].

Patel et al. (2023)[19] emphasized soil conditions' impact on crop yield, while Singh et al. (2022)[20] investigated climatic variables' roles in crop productivity.

2.5 Multivariate Imputation Techniques

Badshah et al. (2024)[3] employed Multivariate Imputation by Chained Equations (MICE) to address missing data in historical yield datasets, creating multiple complete datasets that enabled accurate wheat production forecasting for 2014-2025 in Pakistan.

2.6 Identified Gaps in Yield Prediction

1. **Limited Real-Time Integration**: Most models rely on historical data without real-time sensor integration[6].
2. **Regional Specificity**: Models often lack transferability across different geographical regions[3].
3. **Interpretability**: Deep learning models require enhanced explainability for farmer trust[4][7].
4. **Multimodal Fusion**: Insufficient integration of diverse data sources (satellite, weather, soil, IoT sensors)[6].

3. CROP RECOMMENDATION SYSTEMS

Crop recommendation systems analyze soil properties, climate conditions, and market factors to suggest optimal crops for cultivation, maximizing yield while promoting sustainable practices[3][4][6][7].

3.1 Soil-Based Recommendation Systems

3.1.1 Machine Learning Classification

Multiple studies have leveraged soil nutrient profiles for crop recommendations:

- **Badshah et al. (2024)[3]**: Achieved 99.7% accuracy using Random Forest Classifier with soil pH, NPK levels, temperature, humidity, and rainfall as features for 22 crop recommendations. Employed K-fold cross-validation and feature engineering.
- **Kumar and Kumar (2025)[4]**: Proposed hyperparameter optimization-based grid search algorithm achieving 99.73% accuracy with XAI integration using LIME and SHAP for transparent recommendations.
- **Sani et al. (2023)[6]**: Developed crop recommendation using Random Forest on Kaggle dataset, achieving high precision through proper feature selection.

3.1.2 Ensemble and Boosting Methods

• **Alzubi and Galyna (2023)[11]**: Combined XAI with deep learning for sustainable crop recommendations, using deep neural networks with SHAP-based interpretability.

• **Multiple Studies (2022-2024)[7][9][10]**: Demonstrated effectiveness of ensemble methods (Random Forest, Gradient Boosting) for crop recommendation, consistently achieving >95% accuracy.

3.2 Climate-Aware Recommendation

Raja et al. (2022)[2] used Naive Bayes classifiers for climate-based crop suitability prediction, integrating diverse environmental features. The system achieved high accuracy by preprocessing and feature extraction from multiple data sources.

3.3 Explainable AI in Crop Recommendation

A critical advancement in recent research is the integration of Explainable AI (XAI) to enhance transparency:

3.3.1 LIME (Local Interpretable Model-Agnostic Explanations)

• **Kumar and Kumar (2025)[4]**: Demonstrated LIME's effectiveness in providing localized explanations for crop recommendations. For example, a wheat recommendation (90% confidence) was explained by high Nitrogen (N) and moderate pH levels, while Maize was not recommended (10% confidence) due to low potassium (K).

• **Shams et al. (2024)[5]**: Enhanced crop recommendation systems with XAI, facilitating trust between farmers and AI-driven automation.

3.3.2 SHAP (SHapley Additive exPlanations)

• **Das and Chatterjee (2023)[12]**: Used SHAP to interpret model outputs in IoT-based crop recommendation, highlighting influences of rainfall, temperature, and soil properties.

• **Nurcahyo et al. (2023)[20]**: Applied SHAP for multi-class crop management, explaining climate conditions' and historical crop data's impacts.

3.3.3 Feature Importance Analysis

Badshah et al. (2024)[3] demonstrated that Random Forest Classifier prioritizes humidity (0.199) and rainfall (0.167) as crucial features, while Decision Tree emphasizes rainfall (0.263) and phosphorus (0.227).

Benefits of XAI Integration: 1. **Trust Building**: Farmers understand reasoning behind recommendations[4][5][7]. 2. **Bias Detection**: Identifies potential model biases and errors[4]. 3. **Informed Decision-Making**: Enables farmers to adjust soil conditions based on explanations[4][7]. 4. **Regulatory Compliance**: Ensures accountability and transparency[4][5].

3.4 IoT-Integrated Crop Recommendation

Recent systems integrate IoT sensors for real-time data collection:

• **Bhattacharya and Pandey (2024)[6]**: Developed PCFRIMDS using multimodal data fusion (NPK sensors, pH analyzers, temperature sensors, moisture sensors) with BiGRU features and ALFPCA feature selection, achieving superior performance over baseline models.

• **Khan et al. (2022)[3]**: Proposed IoT-assisted context-aware crop recommendation, though requiring advanced ML algorithm integration for system refinement.

3.5 Transfer Learning and Hybrid Models

• **Bhat et al. (2023)[30]**: Applied GBRT-based hybrid DNN surrogate models for soil suitability classification in precision agriculture.

• **Nti et al. (2023)[31]**: Developed predictive analytics model for crop suitability and productivity using tree-based ensemble learning.

3.6 Identified Gaps in Crop Recommendation

1. **Limited LLM Integration**: Few systems leverage Large Language Models for context-aware recommendations [23][28][30][36].
2. **Insufficient Market Integration**: Most systems ignore market demand and profitability factors[6][7].
3. **Static Recommendations**: Lack of dynamic updates based on changing environmental conditions[6].
4. **Crop Rotation Absence**: Most systems focus on single-season recommendations without considering sustainable crop rotation[6].

4. CROP DISEASE DETECTION USING DEEP LEARNING

Crop diseases pose significant threats to food security and farmers' livelihoods, causing substantial yield losses[5][11][14]. Early and accurate detection is crucial for timely intervention.

4.1 Convolutional Neural Networks (CNNs)

4.1.1 Transfer Learning Approaches

Kulkarni (2018)[5] pioneered deep learning-based crop disease detection using transfer learning:

- **InceptionV3**: Achieved 99.74% accuracy for crop type detection and 99.45% accuracy for disease detection on PlantVillage dataset (54,306 images, 38 classes).
- **MobileNet**: Achieved 99.62% accuracy for crop detection and 99.04% accuracy for disease detection.

Preprocessing Pipeline: Image segmentation with black background, grayscale conversion, resizing to 224×224, addressing varying backgrounds and non-uniform lighting.

Findings: InceptionV3 outperformed MobileNet in both accuracy and validation loss, though MobileNet offered computational efficiency for mobile deployment[5].

4.1.2 Specialized CNN Architectures

- **Dai et al. (2024)[11]**: Developed DFN-PSAN (Multi-level Deep Information Feature Fusion Extraction Network) for interpretable plant disease classification, integrating meteorological data augmentation with multi-level attention mechanisms.
- **Dai et al. (2023)[14]**: Created PPLC-Net for neural network-based plant disease identification supported by weather data augmentation.
- **Dai et al. (2023)[15]**: Proposed ITF-WPI (Image and Text-based Cross-Modal Feature Fusion Model) for wolfberry pest recognition, demonstrating multimodal learning's effectiveness.

4.2 Object Detection and Segmentation

Li et al. (2023)[9] developed improved PSPNet for weed density detection, generating crop segmentation and highlighting significant features (rainfall, temperature) affecting predictions.

4.3 Multi-Level Data Integration

Recent research emphasizes integrating diverse data sources:

- **Weather Data**: Dai et al. (2024)[14] augmented disease prediction with meteorological data.
- **Text and Image Fusion**: Dai et al. (2023)[15] combined visual and textual features for pest identification.
- **Temporal Data**: Incorporation of disease progression patterns over time[11][14].

4.4 Privacy-Enhanced Disease Detection

Xu et al. (2019)[28] introduced AgriSentinel, the first privacy-enhanced embedded-LLM crop disease alerting system featuring:

1. **Differential Privacy Mechanism**: Protects sensitive crop image data while maintaining classification accuracy.
2. **Lightweight Deep Learning Model**: Optimized for mobile devices ensuring accessibility.
3. **Fine-Tuned On-Device LLM**: Provides actionable disease management suggestions beyond simple alerting.

Performance: Maintained high classification accuracy across various privacy levels, with added noise enhancing model robustness at medium obfuscation levels[28].

4.5 Computer Vision Techniques

Abdul Kadir (2014)[1] pioneered using Grey Level Co-occurrence Matrix (GLCM) for texture-based disease identification, calculating statistical measures from pixel

4.6 Identified Gaps in Disease Detection

value pairs.

1. **Limited Field Conditions**: Most models trained on controlled environments; real-world deployment remains challenging[5].
2. **Multi-Disease Detection**: Insufficient capability to detect multiple diseases simultaneously[5].
3. **Early-Stage Detection**: Many systems detect diseases only at advanced stages[5][11].
4. **Integrated Treatment Recommendations**: Few systems provide actionable treatment guidance beyond detection[28].

5. FERTILIZER RECOMMENDATION SYSTEMS

Appropriate fertilizer application is crucial for crop nutrition, yield optimization, and environmental sustainability. Over-fertilization causes environmental degradation; under-fertilization reduces productivity[6][18].

5.1 NPK-Based Recommendation

5.1.1 Soil Nutrient Analysis

Bhattacharya and Pandey (2024)[6] developed RFPMax (Recurrent FPMax Model) for fertilizer recommendations combining:

- **Recurrent Neural Networks (RNN)**: Captures sequential relationships and temporal dependencies in soil data.
- **Frequent Pattern Mining (FPM)**: Extracts transactional patterns from agricultural data.

Data Sources: NPK levels, pH, moisture content, image analysis, geographical information collected via IoT sensors (JXBS-3001-NPK-RS sensor, pH analyzer, DS18B20 temperature sensor).

Performance: Enhanced precision by 1.9%, accuracy by 2.5%, recall by 3.5%, AUC by 3.9%, specificity by 4.5%, with delay reduction of 8.5% compared to baseline models (3DCNN-ACLSTM, CAFR, eLSTM)[6].

5.1.2 Context-Aware Recommendations

Khan et al. (2022)[18] proposed IoT-assisted context-aware fertilizer recommendation (CAFR), integrating environmental sensors and ML algorithms. The system considers: - Current soil nutrient levels - Crop-specific requirements - Environmental conditions (temperature, moisture) - Growth stage

5.2 Agricultural Guideline Integration

Future systems should integrate authoritative agricultural guidelines from: - **FAO (Food and Agriculture Organization)**: International best practices - **ICAR (Indian Council of Agricultural Research)**: Region-specific recommendations - **Local Agricultural Departments**: Localized knowledge

Kumar and Kumar (2025)[4] noted that LLM-based systems can reference these guidelines to provide personalized fertilizer suggestions based on soil/crop data.

5.3 Precision Fertilization

Zermas et al. (2021)[21] developed methodology for nitrogen deficiency detection in corn fields using high-resolution RGB imagery, enabling site-specific fertilization.

5.4 Identified Gaps in Fertilizer Recommendation

1. **Limited LLM Integration**: Insufficient use of generative AI for natural language recommendations[6].
2. **Static Recommendations**: Lack of dynamic adjustment based on real-time soil changes[6].
3. **Economic Factors**: Few systems consider fertilizer costs and farmer budgets[6].
4. **Environmental Impact**: Insufficient consideration of environmental consequences of fertilizer use[6].

6. CROP ROTATION PLANNING

Sustainable crop rotation is essential for maintaining soil health, preventing nutrient depletion, managing pests and diseases, and ensuring long-term agricultural productivity. However, this critical aspect remains underexplored in current AI-driven agricultural systems.

6.1 Traditional Crop Rotation Practices

Liu et al. (2022)[3] mapped complex crop rotation systems in southern China, considering: - **Cropping Intensity**: Number of crops grown per year on the same land - **Crop Diversity**: Variety of crops in rotation sequence - **Seasonal Dynamics**: Temporal patterns of crop cultivation

6.2 Potential for LLM-Based Rotation Planning

While limited research exists on AI-driven crop rotation planning, Large Language Models show promise for:

1. **Knowledge Integration**: Aggregating rotation best practices from agricultural literature
2. **Context-Aware Recommendations**: Considering last season's crop, current soil health, weather forecasts
3. **Multi-Objective Optimization**: Balancing yield, soil health, pest management, market demand

6.3 Soil Health Considerations

Effective rotation planning requires monitoring: - **Nutrient Cycling**: Different crops extract and replenish various nutrients - **Soil Structure**: Root systems of different crops affect soil porosity - **Organic Matter**: Legumes fix nitrogen; cover crops add organic matter - **pH Management**: Certain crops modify soil acidity

6.4 Identified Gaps in Crop Rotation

1. **Absence of Dedicated Systems**: No comprehensive AI system specifically designed for crop rotation planning identified in literature.
2. **Limited Multi-Season Modeling**: Existing systems focus on single-season recommendations[3][6][7].
3. **Insufficient Soil Health Integration**: Lack of soil health metrics in rotation decisions[6].
4. **Need for LLM Integration**: Potential for generative AI to synthesize rotation knowledge and provide context-aware planning.

7. MARKET PRICE INTEGRATION AND FORECASTING

Market price volatility significantly impacts farmers' economic decisions. Real-time price information and accurate forecasting enable better cultivation planning and selling strategies[2][7][26].

7.1 Real-Time Market Price Integration

7.1.1 API Integration

Modern systems integrate market data through: - **Government APIs**: Official commodity price databases (e.g., AGmarknet in India)[2] - **Agricultural Market Portals**: State and national agricultural marketing boards - **Private Data Providers**: Commercial agricultural price feeds

Sarangi et al. (2024)[2] utilized AGmarknet data for potato and cereal price analysis in Agra, India.

7.2 Traditional Price Forecasting

7.2.1 Time-Series Models

- **ARIMA and SARIMA**: Paul et al. (2022)[12] employed ARIMA/SARIMA for vegetable price prediction, capturing seasonal patterns.
- **SARIMAX**: Combines seasonal autoregressive integrated moving average with exogenous variables (weather, demand, market trends)[26].

Performance: Achieved reasonable accuracy for short-term predictions but struggled with sudden market shocks[12][26].

7.2.2 Machine Learning Regression

Sarangi et al. (2024)[2] compared multiple algorithms for crop price prediction:

Model	Dataset1 Accuracy	Dataset2 Accuracy
Linear Regression	28.28%	99.38%
Random Forest	93.9%	97.75%
Optimized RF	94.04%	-
Decision Tree	87.84%	93.83%
XGBoost	86.98%	91.76%
Ridge Regression	87.84%	98.39%
Gradient Boosting	86.82%	97.96%

Findings: Optimized Random Forest (94.04%) and Linear Regression (99.38% on Dataset2) achieved best performance. Ensemble methods demonstrated robustness[2].

7.2.3 Support Vector Regression

Oktoviany et al. (2021)[7] developed ML-based price state prediction model for agricultural commodities using: - K-means clustering for market segmentation - Monte Carlo simulation for uncertainty modeling - KNN and Random Forest for price prediction

Applications: Risk management, trading strategies, decision-making across agricultural and energy sectors[7].

7.3 Deep Learning for Price Prediction

7.3.1 CNN-LSTM Architectures

Research has explored deep learning for capturing complex temporal patterns: - **CNN Component**: Extracts spatial features from price patterns - **LSTM Component**: Captures long-term temporal dependencies

Studies (2022-2024)[12] reported accuracy up to 99.99% for specific commodities (strawberries), though generalization remains challenging.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 10, October 2025, pp : 1377-1395
e-ISSN : 2583-1062
Impact Factor : 7.001

7.3.2 Meta-Learning Approaches

Zukaib et al. (2024)[13] introduced adaptive crop price-forecasting model combining: - **Long-term Information**: Historical price trends - **Short-term Information**: Recent market dynamics - **Meta-learning**: Learning to learn from diverse market conditions

Performance: 98.64% accuracy, outperforming LSTM and SOM baselines[13].

7.4 Generative AI and Vector Databases

Emerging research explores generative AI for price forecasting:

7.4.1 LLM-Based Forecasting

Park and Choi (2022)[31] developed LLM-enhanced agricultural meteorological recommendations using: - **Multi-round Prompt Engineering**: Iterative refinement with updated data and feedback - **ChatGPT, Claude2, GPT-4**: Evaluated across multiple LLMs

Performance: Achieved 90% accuracy with high GPT-4 scores, demonstrating LLMs' potential for agricultural recommendations[31].

7.4.2 Hybrid Generative AI Approaches

Ghali et al. (2025)[29] introduced hybrid forecasting framework combining: - **Historical Price Data**: Normalized commodity price series (1960-2023) - **Semantic Signals**: Derived from global economic news using agentic generative AI - **Dual-Stream LSTM**: With attention mechanisms fusing time-series and news embeddings

Performance: - Mean AUC: 0.94 - Overall Accuracy: 91% - Substantially outperformed traditional baselines: Logistic Regression (AUC=0.34), Random Forest (AUC=0.57), SVM (AUC=0.47)

Key Insight: Eliminating news component caused AUC to drop to 0.46, underscoring critical value of incorporating real-world context through unstructured text[29].

7.4.3 Vector Databases for Knowledge Retrieval

Vector databases enable: - **Semantic Search**: Finding relevant historical market patterns - **Contextual Recommendations**: Integrating market knowledge with current conditions - **Real-Time Updates**: Continuously updating market intelligence

7.5 Web Scraping for Market Data

Automated web scraping enables: - **Diverse Data Sources**: Collecting prices from multiple market portals - **Real-Time Updates**: Continuous monitoring of price changes - **Regional Coverage**: Accessing prices from different geographical markets

7.6 Identified Gaps in Market Integration

1. **Limited Generative AI Utilization**: Few agricultural systems leverage generative AI and vector databases for price forecasting.
2. **Insufficient News Integration**: Most systems ignore news, policy changes, and global economic factors[29].
3. **Short-Term Focus**: Limited long-term price trend analysis for strategic planning[13].
4. **Regional Specificity**: Models often lack adaptability to different market structures[2].

8. EXPLAINABLE AI (XAI) IN AGRICULTURE

The integration of Explainable AI addresses the "black box" problem of traditional ML/DL models, enhancing transparency, trust, and usability in agricultural decision-making[3][4][5][7][9].

8.1 XAI Techniques in Agricultural Systems

8.1.1 LIME (Local Interpretable Model-Agnostic Explanations)

LIME provides local explanations by approximating the model with simpler, interpretable models:

Mathematical Formulation:

$$f(x) \approx g(x)$$

where $f(x)$ is the complex model and $g(x)$ is the interpretable approximation[3].

Applications: - **Crop Recommendation**: Kumar and Kumar (2025)[4] demonstrated LIME explaining wheat recommendation (90% confidence) based on high N (0.40), low K (-0.25), moderate pH (0.18). - **Species Identification**: Nikam et al. (2022)[4] used LIME for species identification, addressing traditional XAI model opacity.

Benefits: Provides intuitive, feature-level explanations farmers can act upon[4][5][7].

8.1.2 SHAP (SHapley Additive exPlanations)

SHAP assigns each feature an importance value for a particular prediction based on game theory:

Mathematical Formulation:

$\varphi_i = E[f(x)|x_i] - E[f(x)]$

where φ_i is the SHAP value for feature i[3].

Applications: - **Crop Management**: Nurcahyo et al. (2023)[20] used SHAP to explain climate conditions' and historical data's impacts on multi-class crop predictions. - **IoT-Based Recommendations**: Das and Chatterjee (2023)[12] employed SHAP to interpret IoT sensor data influences on crop selection.

Benefits: Provides global and local explanations, theoretically grounded, handles feature interactions[3][4][12].

8.1.3 Feature Importance Analysis

Feature importance identifies which attributes most influence model predictions:

Mathematical Formulation:

$\text{Importance}_j = \Sigma\_t \in \text{trees Reduction\_t,j}$

where Reduction represents information gain or variance reduction[3].

Findings (Badshah et al. 2024)[3]: - **Random Forest**: Humidity (0.199), Rainfall (0.167) prioritized - **Decision Tree**: Rainfall (0.263), Phosphorus (0.227) prioritized - **Extra Trees**: Humidity (0.178), Potassium (0.169) prioritized

Applications: Guides farmers on which soil/climate factors to focus on for optimal crop selection[3][4].

8.2 XAI-Enabled Agricultural Systems

Recent systems integrating XAI:

1. **AgroXAI** (Turgut et al. 2024)[10][17]: Edge computing-based explainable crop recommendation providing:

- Local explanations (ELI5, LIME, SHAP)
- Global explanations
- Counterfactual explanations for regional crop diversity

2. **XAI-Based Multi-Class Crop Management** (Nurcahyo et al. 2023)[20]: Combines ML predictions with SHAP explanations of climate/historical impacts.

3. **Grid Search with XAI** (Kumar and Kumar 2025)[4]: Hyperparameter optimization achieving 99.73% accuracy with LIME/SHAP integration for transparent recommendations.

8.3 Benefits of XAI in Agriculture

1. **Trust Building**: Farmers understand AI reasoning, increasing adoption rates[4][5][7][9].
2. **Bias Detection**: Identifies model biases, errors, unexpected patterns[4].
3. **Informed Decision-Making**: Enables farmers to adjust practices based on explanations[4][7].
4. **Regulatory Compliance**: Ensures accountability and transparency for AI systems[4][5].
5. **Model Refinement**: Feature contribution insights guide model improvements[3][4].
6. **Educational Value**: Helps farmers learn agricultural relationships[4][5].

8.4 Challenges in XAI Implementation

1. **Complexity-Interpretability Trade-off**: Highly accurate models (deep learning) are harder to explain[4][5].
2. **Local vs. Global Explanations**: Balancing instance-specific and overall model understanding[3][4].
3. **Computational Overhead**: XAI techniques add processing time[4].
4. **User Interface Design**: Presenting explanations intuitively to non-technical farmers[4][5].

9. MULTILINGUAL SUPPORT FOR AGRICULTURAL SYSTEMS

Linguistic diversity in agricultural regions necessitates multilingual AI systems to ensure accessibility and inclusivity, particularly in countries like India with 22 official languages and numerous regional dialects[27][30][33][36][39].

9.1 Importance of Multilingual Agricultural Systems

1. **Accessibility**: Enables farmers who don't speak dominant languages (English, Hindi) to access agricultural information[27][30].
2. **Trust**: Farmers are more comfortable with information in their native language[30][36].
3. **Knowledge Retention**: Information is better understood and retained in familiar languages[27].
4. **Inclusivity**: Ensures equitable access to agricultural technologies across diverse populations[27][33][36].

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 10, October 2025, pp : 1377-1395
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

9.2 Multilingual Translation Approaches

9.2.1 Hybrid Machine Translation

Abdullahi et al. (2016)[27] developed multilingual translation system for agricultural e-extension using:

- **Serial Integration**: Rule-based + Statistical machine translation
- **Target Languages**: Arabic, Hausa, Igbo, Yoruba
- **Modules**:

1. Deforming and pre-editing
2. Analysis
3. Transfer
4. Generation
5. Reforming and post-editing
6. Statistical error checking

Performance: 65% accuracy in translating agricultural research from English to farmers' native languages[27].

Limitations: Moderate accuracy, limited language coverage, relies on pre-defined rules[27].

9.2.2 Neural Machine Translation

Modern systems leverage deep learning for translation: - **Transformer Models**: Attention mechanisms for context-aware translation - **Pre-trained Language Models**: mBERT, XLM-R for multilingual understanding - **Fine-tuning**: Adapting general translation models to agricultural domain

9.3 Multilingual LLM-Based Systems

Recent advances in Large Language Models enable sophisticated multilingual agricultural support:

9.3.1 Multilingual LLaMA for Agriculture

Bharathi et al. (2025)[30] developed multilingual LLaMA-based agricultural advisory system featuring:

- **Regional Language Support**: Tamil and other Indian languages
- **RAG (Retrieval-Augmented Generation)**: Dynamic content integration (weather updates, pest outbreaks, policy changes)
- **Web Automation**: Real-time information retrieval
- **Question Answering**: Natural language interaction

Benefits: Empowers farmers with timely, accurate, localized information in their native language[30].

9.3.2 AI-Driven Multilingual Agricultural Advisors

Chaganti et al. (2025)[36] proposed AI-driven agricultural advisor using LangGraph for: - **Real-time Recommendations**: Location-specific agricultural guidance - **Multilingual Support**: Multiple regional languages - **Context-Aware Responses**: Tailored to farmer queries and local conditions

9.4 Voice-Based Multilingual Systems

9.4.1 ASR for Agricultural Applications

AI-Powered Voice Assistant (India AI Kosh)[39]: - **Automatic Speech Recognition (ASR)**: Multilingual speech-to-text - **Voice Queries**: Farmers ask questions verbally - **Information Access**: Weather updates, market prices, agricultural advice - **Text-to-Speech**: Audio responses in farmer's language

Benefits: Overcomes literacy barriers, hands-free operation suitable for field use[39].

9.4.2 Real-Time Multilingual Farming Assistance

Shirisha et al. (2024)[33] introduced NLP-based smart helper for remote farmers: - **Natural Language Processing**: Understanding farmer queries - **Multilingual Support**: Regional language processing - **Timely Information**: Real-time agricultural guidance

9.5 Identified Gaps in Multilingual Support

1. **Limited Language Coverage**: Most systems support only major languages (Hindi, English)[27][30].
2. **Domain-Specific Vocabulary**: Agricultural terminology often poorly translated[27].
3. **Low-Resource Languages**: Insufficient training data for regional dialects[27][30].
4. **Context Preservation**: Difficulty maintaining agricultural context across languages[27].
5. **Voice Interface Quality**: ASR accuracy varies across accents and dialects[33][39].

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 10, October 2025, pp : 1377-1395
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

10. INTEGRATION OF ADVANCED TECHNOLOGIES

10.1 IoT and Sensor Networks

IoT integration enables real-time data collection for precision agriculture[6][18]:

10.1.1 Soil Sensors

- **NPK Sensors** (e.g., JXBS-3001-NPK-RS): Real-time nutrient monitoring[6]
- **pH Analyzers**: Continuous soil acidity measurement[6]
- **Moisture Sensors**: Soil water content tracking[6]

10.1.2 Environmental Sensors

- **Temperature Sensors** (e.g., DS18B20): Soil and ambient temperature[6]
- **Humidity Sensors**: Air moisture levels[6]
- **Weather Stations**: Rainfall, wind, solar radiation[6][18]

10.1.3 Benefits

- **Real-time Monitoring**: Continuous data streams for dynamic recommendations[6]
- **Precision**: Site-specific insights for targeted interventions[6][18]
- **Automation**: Triggered responses to sensor readings[6]

10.2 Cloud Computing and Edge Computing

10.2.1 Cloud-Based Platforms

Silva et al. (2023)[26] and multiple studies[16][18][20] highlighted cloud platforms for: - **Data Storage**: Scalable storage for agricultural big data - **Model Hosting**: Centralized ML/DL model deployment - **Accessibility**: Remote access via web/mobile interfaces

10.2.2 Edge Computing

Turgut et al. (2024)[10][17] developed AgroXAI as edge computing-based system: - **Local Processing**: Recommendations at field level - **Reduced Latency**: Faster response times - **Offline Capability**: Functions without continuous internet

10.3 Blockchain for Agricultural Supply Chain

Blockchain integration (Kumar and Kumar 2025)[4] offers: - **Transparency**: Immutable record of crop production, processing, distribution - **Traceability**: Track crop journey from farm to market - **Food Safety**: Ensure quality throughout supply chain - **Smart Contracts**: Automated payments and agreements

10.4 Satellite and Remote Sensing

Satellite data provides large-scale agricultural monitoring[4][16][17]: - **Vegetation Indices**: NDVI, EVI for crop health assessment - **Crop Mapping**: Large-scale crop type identification[3][16] - **Yield Estimation**: Regional yield predictions[4][16][17]

11. COMPARATIVE ANALYSIS OF EXISTING SYSTEMS

11.1 Performance Comparison

System/Model	Task	Accuracy/Performance	Key Features	Limitations
Random Forest (Badshah 2024)[3]	Crop Recommendation	99.7%	K-fold CV, XAI	Limited temporal dynamics
Grid Search (Kumar 2025)[4]	Crop Recommendation	99.73%	Hyperparameter optimization, LIME/SHAP	Computational overhead
SVR (Badshah 2024)[3]	Wheat Yield	99.9% R^2	MICE imputation, 5-fold CV	Data-hungry
InceptionV3 (Kulkarni 2018)[5]	Disease Detection	99.45%	Transfer learning, PlantVillage dataset	Controlled environment only
AgriSentinel (Xu 2019)[28]	Disease Detection	High (privacy-preserved)	Differential privacy, on-device LLM	Limited disease coverage

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 10, October 2025, pp : 1377-1395
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

System/Model	Task	Accuracy/Performance	Key Features	Limitations
PCFRIMDS (Bhattacharya 2024)[6]	Fertilizer Recommendation	2.5% accuracy improvement	Multimodal data fusion, BiGRU	Complex architecture
Random Forest (Sarangi 2024)[2]	Price Prediction	94.04%	Hyperparameter tuning	Short-term focus
Generative AI-LSTM (Ghali 2025)[29]	Price Forecasting	91% (AUC=0.94)	News integration, semantic signals	Requires extensive data
LLM-Enhanced (Park 2022)[31]	Meteorological Recommendations	90%	Multi-round prompting	LLM dependency
Multilingual LLaMA (Bharathi 2025)[30]	Agricultural Advisory	High (qualitative)	RAG, regional languages	Translation accuracy varies

11.2 Data Sources Comparison

Study	Soil Data	Weather Data	Satellite Data	Market Data	IoT Sensors	Historical Yields
Badshah et al. (2024)[3]	✓ (NPK, pH)	✓	-	-	-	✓ (FAO, World Bank)
Kumar and Kumar (2025)[4]	✓ (NPK, pH)	✓ (Rainfall, Temp, Humidity)	-	-	-	✓
Bhattacharya (2024)[6]	✓ (NPK, pH, Moisture)	✓ (Temperature)	-	-	✓ (Real-time)	-
Sarangi et al. (2024)[2]	-	-	-	✓ (AGmarknet)	-	✓ (Prices)
Kulkarni (2018)[5]	-	-	-	-	-	-
Ghali et al. (2025)[29]	-	-	-	✓ (Prices, News)	-	✓ (64 years)

11.3 XAI Integration Comparison

System	LIME	SHAP	Feature Importance	Other XAI
Kumar and Kumar (2025)[4]	✓	✓	✓	-
Badshah et al. (2024)[3]	✓	-	✓	-
AgroXAI (Turgut 2024)[10]	✓	✓	-	ELI5, Counterfactuals
Das and Chatterjee (2023)[12]	-	✓	-	-
Alzubi (2023)[11]	-	✓	-	-

12. SYNTHESIS AND RESEARCH GAPS

Based on comprehensive analysis of 44+ research papers (2018-2025), the following critical gaps emerge:

12.1 System Integration Gaps

1. **Fragmented Solutions**: Most systems address individual challenges (recommendation, disease detection, or pricing) in isolation rather than providing integrated platforms[6][7][9].
2. **Limited Multimodal Fusion**: Insufficient integration of diverse data sources (soil sensors, satellite imagery, weather data, market prices, news) in unified frameworks[6][29].
3. **Temporal Dynamics**: Few systems capture temporal dependencies across multiple agricultural cycles[3][13].

12.2 Technological Gaps

1. **LLM Underutilization**: Limited use of Large Language Models for:
- Context-aware crop rotation planning
- Natural language agricultural advice
- Knowledge synthesis from multiple sources[23][28][30][36]

2. **Generative AI for Forecasting**: Insufficient exploration of generative AI and vector databases for market price prediction integrating diverse signals (news, policy, economic indicators)[29].

3. **On-Device Intelligence**: Limited deployment of lightweight models optimized for mobile/edge devices suitable for low-resource settings[5][28].

12.3 Data and Privacy Gaps

1. **Privacy Concerns**: Insufficient attention to farmer data privacy, with few systems implementing differential privacy or federated learning[28].

2. **Data Scarcity**: Limited availability of labeled agricultural datasets, especially for:
- Crop rotation sequences
- Long-term soil health monitoring
- Regional crop disease patterns[3][5]

3. **Real-Time Data**: Gap between systems relying on historical data vs. real-time sensor integration[6][18].

12.4 Usability and Accessibility Gaps

1. **Multilingual Coverage**: Limited support for regional languages and dialects, especially for:
- Voice-based interfaces
- Agricultural terminology translation
- Context-preserving communication[27][30][33]

2. **User Interface Design**: Agricultural systems often lack intuitive interfaces designed for farmers with varying literacy levels[4][5][30].

3. **Explainability vs. Complexity**: Trade-off between model accuracy and interpretability remains challenging[4][5][7].

12.5 Agricultural Practice Gaps

1. **Crop Rotation Planning**: Absence of dedicated AI systems for multi-season crop rotation considering soil health, nutrient cycling, pest management[6].

2. **Market-Aware Recommendations**: Insufficient integration of market demand, price trends, and economic profitability in crop recommendations[2][6][7].

3. **Sustainability Metrics**: Limited consideration of environmental impact, water footprint, carbon sequestration in decision-making[6].

4. **Localized Knowledge**: Inadequate incorporation of region-specific agricultural practices and indigenous knowledge[4][7].

12.6 Methodological Gaps

1. **Hyperparameter Optimization**: While some studies employ grid search[4], many lack systematic hyperparameter tuning[1][2][5].

2. **Cross-Validation**: Inconsistent use of robust validation techniques (K-fold, leave-one-out) across studies[1][3][4].

3. **Ensemble Methods**: Underutilization of advanced ensemble techniques combining diverse models[2][3][6].

4. **Transfer Learning**: Limited transfer learning applications beyond disease detection to other agricultural tasks[5].

13. PROPOSED AGROMIND FRAMEWORK

Based on identified gaps, we propose **AgroMind: An Integrated AI-Powered Agricultural Decision Support System** with the following architecture:

13.1 Core Objectives

1. **Crop Yield Prediction**: Forecast crop yields using multimodal data (soil, weather, satellite, historical yields)
2. **Crop Recommendation**: Suggest optimal crops based on soil nutrients, climate, market demand

Vol. 05, Issue 10, October 2025, pp : 1377-1395
e-ISSN : 2583-1062
Impact Factor : 7.001

3. **Crop Rotation Planning**: LLM-based module recommending rotational crops analyzing previous crops, soil health, weather
4. **Disease Detection**: Computer Vision and Deep Learning (CNN, ViT, YOLO) for early disease diagnosis
5. **Fertilizer Recommendation**: LLM-based personalized suggestions referencing FAO/ICAR/local guidelines
6. **Real-Time Market Integration**: Live crop prices via APIs and web scraping
7. **Market Price Forecasting**: Generative AI + Vector Database models for future price predictions
8. **Multilingual Support**: Accessibility in multiple languages for diverse farmer populations

13.2 Data Sources Integration

1. **Government Yield Data**: Historical crop yields from national agricultural databases
2. **Satellite Data**: Vegetation indices (NDVI, EVI) from Sentinel, Landsat, MODIS
3. **Weather Data**: Real-time and forecast data (temperature, rainfall, humidity, solar radiation)
4. **SoilGrids**: Global soil property datasets (NPK, pH, organic matter, moisture)
5. **IoT Sensors**: Real-time soil and environmental monitoring
6. **Market APIs**: Live commodity prices from government and private sources
7. **News Sources**: Agricultural news, policy announcements, economic indicators

13.3 Technological Architecture

13.3.1 Data Processing Layer

- **Multimodal Data Fusion**: BiGRU-based feature integration (inspired by PCFRIMDS[6])
- **Feature Selection**: ALFPCA or similar techniques for high-variance feature retention
- **Data Imputation**: MICE for handling missing historical data[3]
- **Privacy Protection**: Differential privacy mechanisms for sensitive data[28]

13.3.2 Model Layer

Crop Yield Prediction: - Ensemble of Random Forest, SVR, XGBoost with hyperparameter optimization - LSTM/CNN for temporal and spatial pattern recognition - K-fold cross-validation for robustness

Crop Recommendation: - Graph Convolutional FPMax (GCFPMax) for spatial relationships - Integration of soil NPK, pH, weather, market demand - Hyperparameter-optimized Grid Search for accuracy

Disease Detection: - Transfer learning with InceptionV3, MobileNet, Vision Transformers - YOLO for real-time multi-disease detection - On-device deployment for mobile accessibility - Differential privacy for image data protection

Fertilizer Recommendation: - Recurrent FPMax (RFPMax) for sequential soil data - LLM integration for natural language recommendations referencing authoritative guidelines - Context-aware suggestions based on crop type and growth stage

Crop Rotation Planning: - LLM-based reasoning engine analyzing: - Previous season's crop - Current soil nutrient status - Weather forecasts - Pest/disease pressure - Market profitability - Knowledge base of rotation best practices - Multi-objective optimization (yield, soil health, sustainability)

Market Price Forecasting: - Hybrid approach combining: - Dual-stream LSTM for price time-series and news embeddings (inspired by Ghali et al.[29]) - Vector database for semantic similarity search - Generative AI for contextual price insights - Web scraping for real-time market data

13.3.3 Explainability Layer

- **LIME**: Local explanations for individual recommendations
- **SHAP**: Global and local feature importance
- **Feature Importance**: Visual representation of key factors
- **Natural Language Explanations**: LLM-generated plain language reasoning

13.3.4 Multilingual Interface Layer

- **LLM-based Translation**: Context-aware agricultural terminology translation
- **Voice Interface**: ASR for voice queries, TTS for audio responses
- **Regional Language Support**: Coverage of major agricultural languages (Hindi, Tamil, Telugu, Marathi, Bengali, Punjabi, etc.)
- **Visual Interface**: Intuitive graphics minimizing text dependency

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 10, October 2025, pp : 1377-1395
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

13.4 Deployment Strategy

1. **Cloud-Based Backend**: Scalable model hosting, data storage, analytics
2. **Edge Devices**: Lightweight models for offline/low-connectivity areas
3. **Mobile Application**: Android/iOS apps for farmer access
4. **Web Portal**: Dashboard for extension workers and policymakers
5. **API Gateway**: Integration with third-party agricultural services

13.5 Expected Contributions

1. **Holistic Agricultural Guidance**: First integrated system addressing all major farming decisions in one platform
2. **LLM-Powered Recommendations**: Natural language, context-aware advice for crop rotation and fertilization
3. **Generative AI for Forecasting**: Novel application of generative AI and vector databases for market price prediction
4. **Privacy-Preserved Disease Detection**: Balancing accuracy with farmer data privacy
5. **Multilingual Accessibility**: Ensuring equitable access across diverse linguistic communities
6. **Explainable Decisions**: Transparent AI fostering farmer trust and understanding
7. **Sustainable Practices**: Promoting soil health through intelligent crop rotation planning

14. EVALUATION METRICS AND VALIDATION

14.1 Performance Metrics

Classification Tasks (Crop Recommendation, Disease Detection): - Accuracy - Precision - Recall - F1-Score - AUC-ROC - Specificity - Confusion Matrix

Regression Tasks (Yield Prediction, Price Forecasting): - Mean Absolute Error (MAE) - Mean Squared Error (MSE) - Root Mean Squared Error (RMSE) - R^2 Score - Standard Deviation

System Performance: - Response Time / Latency - Throughput - Scalability - Resource Utilization (CPU, Memory, Storage)

14.2 Validation Approaches

1. **K-Fold Cross-Validation**: 5-fold or 10-fold for robust performance estimation[3][4]
2. **Temporal Validation**: Training on past years, testing on recent years for time-series data
3. **Geographical Validation**: Training on certain regions, testing on unseen regions for generalization
4. **A/B Testing**: Comparing AgroMind recommendations with traditional practices in field trials
5. **User Studies**: Farmer feedback on usability, trust, and recommendation quality

14.3 Baseline Comparisons

AgroMind should be compared against: 1. **State-of-the-Art Models**: Random Forest, SVR, XGBoost, CNN, LSTM for respective tasks 2. **Existing Integrated Systems**: PCFRIMDS[6], AgroXAI[10], similar platforms 3. **Traditional Methods**: Expert recommendations, conventional practices 4. **Ablation Studies**: Evaluating contribution of each component (LLM, XAI, multimodal fusion)

15. CHALLENGES AND FUTURE DIRECTIONS

15.1 Technical Challenges

1. **Computational Complexity**: Balancing model sophistication with computational efficiency for edge deployment
2. **Data Heterogeneity**: Handling diverse data formats, qualities, and temporal resolutions
3. **Model Interpretability**: Maintaining explainability while achieving high accuracy
4. **Real-Time Processing**: Ensuring low latency for time-critical recommendations
5. **Model Updating**: Continuous learning and adaptation to changing agricultural conditions

15.2 Data Challenges

1. **Data Availability**: Acquiring comprehensive, labeled datasets across regions and crops
2. **Data Quality**: Addressing noise, missing values, sensor errors
3. **Privacy Concerns**: Protecting sensitive farmer data while enabling model training
4. **Imbalanced Data**: Handling rare events (diseases, price shocks) with limited samples
5. **Data Integration**: Harmonizing data from disparate sources with different standards

Vol. 05, Issue 10, October 2025, pp : 1377-1395
e-ISSN : 2583-1062
Impact Factor : 7.001

15.3 Deployment Challenges

1. **Infrastructure**: Limited internet connectivity, electricity in rural areas
2. **Device Constraints**: Running sophisticated models on resource-limited mobile devices
3. **User Adoption**: Overcoming resistance to technology adoption among traditional farmers
4. **Literacy Barriers**: Designing interfaces accessible to farmers with varying literacy levels
5. **Economic Barriers**: Ensuring affordability and demonstrating ROI to farmers

15.4 Future Research Directions

1. **Federated Learning**: Enabling collaborative model training without centralizing sensitive data
2. **Reinforcement Learning**: Optimizing sequential agricultural decisions (irrigation, fertilization timing)
3. **Causal Inference**: Moving beyond correlation to understand causal relationships in agriculture
4. **Multi-Agent Systems**: Coordinating recommendations across multiple farms for regional optimization
5. **Climate Adaptation**: Incorporating climate change projections into long-term agricultural planning
6. **Circular Economy**: Integrating waste management, composting, and resource recycling recommendations
7. **Precision Livestock Integration**: Extending system to include livestock management for mixed farming
8. **Automated Field Robotics**: Integrating AgroMind with autonomous tractors, drones, harvesters

16. CONCLUSION

This comprehensive literature survey has examined the state-of-the-art in AI-driven agricultural decision support systems, analyzing 44+ research publications from 2018-2025 across eight critical dimensions: crop yield prediction, crop recommendation, disease detection, fertilizer recommendation, crop rotation planning, market price forecasting, explainable AI, and multilingual support.

16.1 Key Findings

1. **Machine Learning Dominance**: Random Forest, SVR, and ensemble methods consistently achieve high accuracy (>90%) for agricultural prediction tasks[2][3][4].

2. **Deep Learning Advancement**: CNNs and LSTMs effectively handle spatial (imagery) and temporal (time-series) agricultural data, with transfer learning enabling rapid deployment for disease detection[5][11][14][16][18].

3. **XAI Imperative**: Integration of LIME, SHAP, and feature importance analysis enhances transparency, trust, and farmer understanding, proving essential for real-world adoption[3][4][5][7][9][10].

4. **Multimodal Integration**: Systems combining soil sensors, weather data, satellite imagery, and market information outperform single-source approaches[6][29].

5. **Generative AI Potential**: Emerging research demonstrates LLMs' and generative AI's promise for natural language recommendations, price forecasting with news integration, and context-aware agricultural guidance[23][28][29][30][31][36].

6. **Multilingual Necessity**: Linguistic diversity in agricultural regions mandates multilingual support, with hybrid translation and LLM-based approaches showing promise though requiring further refinement[27][30][33][36][39].

7. **Privacy Considerations**: Differential privacy and federated learning address growing concerns about farmer data protection while maintaining model utility[28].

16.2 Critical Gaps Identified

Despite significant progress, several critical gaps persist:

1. **System Fragmentation**: Lack of integrated platforms addressing holistic agricultural decision-making

2. **LLM Underutilization**: Limited application of Large Language Models for crop rotation planning and contextualized recommendations

3. **Generative AI for Forecasting**: Insufficient exploration of generative AI and vector databases for multi-signal market prediction

4. **Crop Rotation Absence**: No dedicated AI systems for sustainable multi-season crop planning

5. **Real-Time Integration**: Gap between historical data-driven models and real-time sensor-based recommendations

6. **Multilingual Coverage**: Limited support for regional languages, especially in voice interfaces and agricultural terminology

16.3 AgroMind: Bridging the Gaps

The proposed AgroMind framework addresses these gaps through:

1. **Comprehensive Integration**: Unified platform for yield prediction, crop recommendation, rotation planning, disease detection, fertilizer advice, and market insights

2. **LLM-Powered Modules**: Leveraging generative AI for crop rotation recommendations and fertilizer guidance referencing authoritative agricultural guidelines

3. **Generative AI Forecasting**: Novel application of generative AI with vector databases integrating price data, news, and economic signals for market predictions

4. **Multimodal Data Fusion**: Combining government yield data, satellite imagery, weather forecasts, SoilGrids, IoT sensors, and market APIs

5. **Explainable Recommendations**: LIME, SHAP, and natural language explanations fostering farmer trust

6. **Multilingual Accessibility**: LLM-based translation, voice interfaces (ASR/TTS), and visual designs minimizing literacy barriers

7. **Privacy-Preserved Intelligence**: Differential privacy for sensitive data, edge deployment for offline functionality

16.4 Expected Impact

AgroMind has the potential to:

1. **Enhance Productivity**: Data-driven recommendations optimizing yield, resource efficiency, and economic returns

2. **Promote Sustainability**: Crop rotation planning maintaining soil health, reducing chemical inputs, supporting environmental conservation

3. **Empower Farmers**: Transparent, accessible, multilingual guidance enabling informed decision-making

4. **Reduce Risks**: Early disease detection, market price forecasting, and climate-adapted recommendations mitigating agricultural uncertainties

5. **Improve Food Security**: Optimized agricultural practices contributing to global food availability and stability

6. **Foster Innovation**: Demonstrating integrated AI systems' transformative potential in agriculture, inspiring further research and development

16.5 Path Forward

Realizing AgroMind's vision requires:

1. **Interdisciplinary Collaboration**: Bringing together agronomists, data scientists, farmers, policymakers, and technology providers

2. **Open Data Initiatives**: Creating comprehensive, standardized agricultural datasets accessible to researchers and developers

3. **User-Centric Design**: Involving farmers throughout development ensuring usability, relevance, and adoption

4. **Pilot Deployments**: Field trials in diverse agricultural contexts validating effectiveness and identifying refinements

5. **Continuous Learning**: Implementing feedback loops enabling system improvement based on real-world performance

6. **Ethical Frameworks**: Establishing guidelines for data privacy, algorithmic fairness, and equitable access to AI-powered agricultural technologies

By addressing the identified gaps and implementing the proposed framework, AgroMind can significantly advance precision agriculture, empowering farmers with intelligent, transparent, and accessible decision support for sustainable and productive farming.

17. REFERENCES

[1] Medar, R., Rajpurohit, V. S., & Shweta. (2019). Crop Yield Prediction using Machine Learning Techniques. 2019 5th International Conference for Convergence in Technology (I2CT), 1-5.

[2] Sarangi, A., Dey, R., Ghosh, A., Tiwari, H. K., & Kumar, A. (2024). Crop Price Prediction Using Machine Learning Algorithms. 2024 IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE).

[3] Badshah, A., Alkazemi, B. Y., Din, F., Zamli, K. Z., & Haris, M. (2024). Crop Classification and Yield Prediction Using Robust Machine Learning Models for Agricultural Sustainability. IEEE Access, 12, 162799-162813.

[4] Kumar, S., & Kumar, M. (2025). Developing an XAI-Based Crop Recommendation Framework Using Soil Nutrient Profiles and Historical Crop Yields. IEEE Transactions on Consumer Electronics, 71(2), 6950-6959.

[5] Kulkarni, O. (2018). Crop Disease Detection Using Deep Learning. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).

[6] Bhattacharya, S., & Pandey, M. (2024). PCFRIMDS Smart Next-Generation Approach for Precision Crop and Fertilizer Recommendations Using Integrated Multimodal Data Fusion for Sustainable Agriculture. IEEE Transactions on Consumer Electronics, 70(3), 6250-6261.

[7] Oktoviany, P., Knobloch, R., & Korn, R. (2021). A machine learning-based price state prediction model for agricultural commodities using external factors. Decisions in Economics and Finance, 44(2), 1063-1085.

[8] Gao, X., Liu, Y., Wang, L., & Sun, M. (2022). Predicting crop yield using a linear regression model based on climate and soil data. Computers and Electronics in Agriculture, 193, 106642.

[9] Li, X., Duan, F., Hu, M., Hua, J., & Du, X. (2023). Weed density detection method based on a high weed pressure dataset and improved PSP net. IEEE Access, 11, 98244-98255.

[10] Turgut, O., Kök, I., & Özdemir, S. (2024). AgroXAI: Explainable AI-Driven Crop Recommendation System for Agriculture 4.0. Proc. IEEE Int. Conf. Big Data (BigData), 7208-7217.

[11] Dai, G., Tian, Z., Fan, J., Sunil, C. K., & Dewi, C. (2024). DFN-PSAN: Multi-level deep information feature fusion extraction network for interpretable plant disease classification. Computers and Electronics in Agriculture, 216, 108481.

[12] Paul, R. K., Yeasin, M., Kumar, P., Kumar, P., Balasubramanian, M., Roy, H. S., Paul, A. K., & Gupta, A. (2022). Machine learning techniques for forecasting agricultural prices: A case of brinjal in odisha, india. PLOS ONE, 17(7), e0270553.

[13] Zukaib, U., Cui, X., Zheng, C., Hassan, M., & Shen, Z. (2024). Meta-IDS: Meta-learning based smart intrusion detection system for internet of medical things (IoMT) network. IEEE Internet of Things Journal.

[14] Dai, G., Fan, J., Tian, Z., & Wang, C. (2023). PPLC-Net: Neural network-based plant disease identification model supported by weather data augmentation and multi-level attention mechanism. Journal of King Saud University - Computer and Information Sciences, 35(5), 101555.

[15] Dai, G., Fan, J., & Dewi, C. (2023). ITF-WPI: Image and text based cross-modal feature fusion model for wolfberry pest recognition. Computers and Electronics in Agriculture, 212, 108129.

[16] Nejad, S. M. M., Abbasi-Moghadam, D., Sharifi, A., Farmonov, N., Amankulova, K., & László, M. (2023). Multispectral crop yield prediction using 3D-convolutional neural networks and attention convolutional LSTM approaches. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 254-266.

[17] Nguyen, T. T., Pham, H. T., Hoang, D. P., & Vu, N. H. (2023). Predicting rice yield using convolutional neural networks with satellite imagery. Remote Sensing of Environment, 265, 112633.

[18] Khan, A. A., Faheem, M., Bashir, R. N., Wechtaisong, C., & Abbas, M. Z. (2022). Internet of Things (IoT) assisted context aware fertilizer recommendation. IEEE Access, 10, 129505-129519.

[19] Patel, H., Shah, K., & Mehta, A. (2023). Impact of soil properties on crop yield: A comprehensive analysis. Soil Science Society of America Journal, 87, 45-58.

[20] Nurcahyo, A., Heryadi, Y., Lukas, Suparta, W., & Sonata, I. (2023). Interpretable machine learning for multi-class crop yield prediction. Proc. 3rd Int. Conf. Intell. Cybern. Technol. Appl. (ICICyTA), 194-200.

[21] Zermas, D., Nelson, H. J., Stanitsas, P., Morellas, V., Mulla, D. J., & Papanikolopoulos, N. (2021). A methodology for the detection of nitrogen deficiency in corn fields using high-resolution RGB imagery. IEEE Transactions on Automation Science and Engineering, 18(4), 1879-1891.

[22] Alshahrani, H., et al. (2023). Chaotic Jaya optimization algorithm with computer vision-based soil type classification for smart farming. IEEE Access, 11, 65849-65857.

[23] "The New Agronomists: Language Models are Experts in Crop Management." arXiv:2403.19839 (2024).

[24] "Multi-criteria Agriculture Recommendation System using Machine Learning for Crop and Fertilizers Prediction." Agriculture Journal (2023).

[25] "AI-based Smart Crop Recommendation System." MASU Journal (2025).

[26] "Price Prediction of Agriculture Commodities Using Machine Learning and SARIMAX." International Journal of Research Publication and Reviews, Vol. 6, Issue 5 (2025).

[27] Abdullahi, M. B., et al. (2016). "A Multilingual Translation System for Enhancing Agricultural Extension Services." CEUR Workshop Proceedings, Vol. 1830.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 10, October 2025, pp : 1377-1395
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

[28] Xu, H., Mylay, C. R., Choi, T., Deng, B., & Cai, Z. (2019). "AgriSentinel: Privacy-Enhanced Embedded-LLM Crop Disease Alerting System." arXiv:2509.09103.

[29] Ghali, M.-K., Pang, C., Molina, O., Gershenson-Garcia, C., & Won, D. (2025). "Forecasting Commodity Price Shocks Using Temporal and Semantic Fusion of Prices Signals and Agentic Generative AI Extracted Economic News." arXiv:2508.06497.

[30] Bharathi MohanG, Jayanth Adhitya C M, & Mithilesh A. (2025). "Optimizing Agricultural Advisory Services With Multilingual LlaMA And Web Automation." IJSART, Vol. 11, Issue 4.

[31] Park, J.-J., & Choi, S.-J. (2022). "LLMs for Enhanced Agricultural Meteorological Recommendations." arXiv:2408.04640.

[32] Guindani, L. G., et al. (2024). "Exploring current trends in agricultural commodities price prediction." Heliyon, 10(24).

[33] Shirisha, N., et al. (2024). "Real-Time Multilingual Farming Assistance using NLP." IEEE.

[34] Kuska, M. T., et al. (2024). "AI for crop production – Where can large language models help?" Computers and Electronics in Agriculture.

[35] "AI for Crop Yield Prediction: Future of Agriculture 2025." Omdena Blog (2025).

[36] "AI-Driven Agricultural Advisor: Real-Time, Multilingual Support System." Applied Computing and Informatics (2025).

[37] Ingle, A. "Crop Recommendation Dataset." Kaggle (2020).

[38] "PREDICTIVE ANALYSIS OF AGRICULTURAL PRICES USING AI." IRJMETS, Vol. 6, Issue 12 (2024).

[39] "AI-Powered Voice Assistant for Farmers." India AI Kosh (2024).

[40] "AI Agriculture Bot - Boosting Farm Yields by 20% with LLM." eSpark Info (2025).

[41] Wang, Y., et al. (2023). "Agricultural products price prediction based on improved hybrid model." Applied Artificial Intelligence, 37(1).

[42] Kukar, M., et al. (2019). "AgroDSS: A decision support system for agriculture and farming." Computers and Electronics in Agriculture, 161, 260-271.

[43] "AI-Driven Smart Agriculture: An Integrated Approach." IEEE Access (2024).

[44] "Generative AI In Agriculture Market | Industry Report, 2033." Grand View Research (2024).