Fed-GRU-XGBoost: A Federated Learning-Based Crop Yield Prediction with Improved GRU-XGBoost Framework for Precision Agriculture

Sarowar Morshed Shawon1, a), Ifaz Ahmed Adan1, b), Shah Nawaz Haider2, c), Arnab Barua1, d), Steve Austin2, e) and H.T. Zubair3, 4, f)

*1Department of Electrical and Electronic Engineering, University of Science and Technology Chittagong, Chattogram 4202, Bangladesh*

*2Department of Computer Science and Engineering, University of Science and Technology Chittagong, Chattogram 4202, Bangladesh*

3Faculty of Computing and Informatics (FCI), Multimedia University, Cyberjaya 63100, Selangor, Malaysia.

4Center for Fiber Networking and Communication, CoE for Intelligent Network, Multimedia University, 63100 Cyberjaya, Selangor, Malaysia.

*f) Corresponding author: zubair@mmu.edu.my*

*a) shawon@ustc.ac.bd  
b) adan@ustc.ac.bd  
c)**nawazhaider60@gmail.com  
d) arnab.me@outlook.com  
e) steve\_austin\_@outlook.com*

**Abstract.** Accurate crop yield prediction is essential for precision agriculture, enabling efficient resource use and improved food security. However, traditional centralized approaches often require data sharing, raising concerns about privacy, data ownership and regulatory compliance. To address this, we propose Fed-GRU-XGBoost, a novel hybrid framework that integrates Gated Recurrent Units (GRU) for temporal feature extraction with XGBoost for regression, within a Federated Learning (FL) environment to preserve data privacy. The model was trained on global data (1990–2008) and tested on five South Asian countries (2009–2013) using key environmental and agronomic features. The centralized GRU-XGBoost model achieved the highest predictive accuracy (R² = 0.9787, MAE = 7111.39, RMSE = 15186.09), while the proposed Fed-GRU-XGBoost offered slightly lower performance (R² = 0.9676, MAE = 8758.08, RMSE = 15329.38), highlighting a trade-off between model accuracy and data security. Despite this trade-off, Fed-GRU-XGBoost demonstrates strong potential as a scalable and privacy-preserving solution for real-world agricultural forecasting applications.

# Introduction

Agriculture serves as a cornerstone of the global economy, providing employment and sustaining livelihoods, particularly in developing nations. Agriculture dominates employment and output in many low-income nations and remains a steady pillar of the world economy. As of 2022, agriculture, forestry and fishing contributed roughly 4 % to global GDP, a share largely unchanged since 2000 [1]. The value produced by the agriculture sector grew from USD 3.0 trillion in 2013 to USD 3.8 trillion in 2022, increasing at an average rate of 2.9% per year [2]. AI, including machine and deep learning, improves agricultural efficiency, precision and sustainability by reducing labor and enhancing decision-making [3]. It supports real-time monitoring, yield prediction and disease detection. Crop yield depends on environment, soil and management, with minor long-term effects often overlooked [4]. Limitations of manual methods have led to ML and remote sensing using UAVs and satellites for accurate yield forecasting. Reliable forecasting helps in agricultural smart farming and food security with planning and good utilization of resources. IoT is now combined with machine learning. This allows for innovations such as solar-powered multi-layer soil moisture sensing systems [5], rooftop soil profiling for urban agriculture [6] and frameworks for recommending suitable crops based on environmental data and soil nutrient parameters [7]. These technologies serve precision agriculture by enhancing decision support, resource efficiency and data accessibility throughout a wide range of farming situations. While many agricultural data needed to predict crop production are exchanged between farms, companies and agencies, data-sharing is often inhibited by concerns over privacy. Such fragmentation prevents classic machine learning models based on centralized datasets from being effective [8]. However, Federated Learning (FL) provides a solution for decentralized model training without moving raw data and preserving privacy and security. FL allows stakeholders to jointly train models on local data maintained on the farms, government servers or both, which improves prediction accuracy but keeps data ownership and privacy intact [9].

In this study, Fed-GRU-XGBoost, a hybrid framework employing XGBoost is developed for enhanced predictive accuracy on multi-crop yield prediction and Gated Recurrent Units (GRU) for temporal relationship modelling of crop yield data within the framework of Federated Learning (FL) for distributed multi-crop data held across multiple devices and evaluated on real-world agricultural datasets with a focus on decentralized training to ensure data privacy.

# Literature review

Accurate forecasting crop yields is vital for agricultural planning because it determines what we grow, at what price to set and also the policy that is made. Subsequent early methods like copula models for risk assessment [10] and the Weibull distribution for weather effect on corn yield [11] set the stage, but struggled with complex high dimensional data. As machine learning started to pick up, more powerful models emerged. Choudhary et al. [12] used Random Forest (RF) with Sentinel-2 imagery and GEE, achieving over 85% classification accuracy and rice yield estimates between 0.40 to 1.01 t/ha. Khan et al. [13] proposed a Geographically Weighted RF Regression (GWRFR), addressing spatial non-stationarity and achieving R² = 0.90, RMSE = 0.764 MT/ha in the U.S. Corn Belt.

Deep learning models like Gated Recurrent Units (GRUs) became popular for handling temporal data and missing values, offering efficient, real-time forecasting. XGBoost, known for its speed and scalability, also proved effective for real-time agricultural predictions [14]. Sanchez et al. [15] applied RF, SVR and XGBoost for wheat yield prediction using features like NDVI, HI and biomass, improving accuracy. Huber et al. [16] used XGBoost with satellite-derived features for soybean yield prediction, highlighting near-infrared reflectance as a key factor and offering a simpler alternative to deep learning. Autoencoders are unsupervised learning models. They minimize data size and extract essential features maintaining crucial information as well. They are used in tasks like computer vision, NLP, anomaly detection for their ability to learn simplified illustrations of complex data [17]. The GRU-Autoencoder captures temporal patterns in data, enabling more effective detection of abnormalities and monitoring of operating conditions [18]. To predict crop yield from environmental and soild data, Shawon et al. [19] proposed machine learning models, including XGBoost and Random Forest. Both models performed exceptionally well, with R² = 0.999, highlighting their effectiveness for precision agriculture in tropical regions.

Federated Learning (FL) emerged to address data privacy and distribution. Aggarwal et al. [20] conducted a systematic review of 92 studies on Federated Learning (FL) in IoT, highlighting its growing use in areas like smart agriculture. This paper highlights the unique strengths of FL for IoT applications due to features such as privacy utilization and resource efficiency, making a compelling case for its superiority over centralized DL in real-world scenarios. Li et al. [21] proposed FL with network pruning to compress the model size (84%) and communication cost (64.7%) while improving the performance (15.5%-20%) at the same time. Durrant et al. [22] applied FL to soybean yield prediction, preserving data privacy and outperforming individual models, showcasing FL’s potential for secure, collaborative agriculture. Shawon et al. [23] conducted a systematic literature review of crop yield prediction studies from 2017 to 2024, identifying temperature, soil type and vegetation as key features. The study highlighted that ML algorithms like RF, GBT and deep learning models such as CNN and LSTM are most frequently applied with RMSE and R² as common evaluation metrics.

# Methodology

This section compares three more advanced models for the prediction of multi- crop yields in tropical agricultural regions: a baseline XGBoost model, a centralized GRU-Autoencoder-enhanced XGBoost model and a federated hybrid model (Fed-GRU-XGBoost). Historical agricultural and climatic data from 101 countries were utilized to train the models, with a focus on tropical South Asian regions for model validation.

## Dataset and Input Features

The dataset consisted of crop yield data identified by three main input attributes: Area (covering 101 countries), Item (type of crop) and Year (spanning from 1990 to 2003). The key features used for modeling included average annual rainfall (mm), pesticide usage (tonnes) and average temperature, all of which influence crop yield. The target variable was crop yield measured in hectograms per hectare (hg/ha\_yield). For model training, data from all countries between 1990 and 2008 was used, while testing was conducted on data from 2009 to 2013 for five South Asian countries: Bangladesh, India, Nepal, Pakistan and Sri Lanka. Data was complete with no missing values. Area and Item were label encoded and all features were normalized using MinMax scaling to ensure compatibility across all models.

## Model 1: Raw XGBoost (Baseline, Centralized)

The first model employed a centralized XGBoost regressor. Hyperparameter tuning was performed using Bayesian optimization, minimizing Mean Squared Error (MSE). The best configuration found is shown in Table 1.

|  |  |
| --- | --- |
| **TABLE** **1**. XGBoost best hyperparameter values | |
| **Parameter** | **Value** |
| n\_estimators | 615 |
| learning\_rate | 0. 293 |
| max\_depth | 11 |
| min\_child\_weight | 9 |
| Subsample | 0. 697 |
| colsample\_bytree | 0. 936 |
| Gamma | 3. 685 |
| reg\_alpha | 0. 00064 |
| reg\_lambda | 0. 0091 |

Table 1 shows the optimized XGBoost hyperparameters: 615 estimators, learning rate 0.293, max depth 11, min child weight 9, subsample 0.697, colsample\_bytree 0.936, gamma 3.685, reg\_alpha 0.00064, reg\_lambda 0.0091 and early stopping 50 rounds. This configuration served as the benchmark for subsequent comparisons.

## Model 2: Centralized GRU Autoencoder + XGBoost (Hybrid)

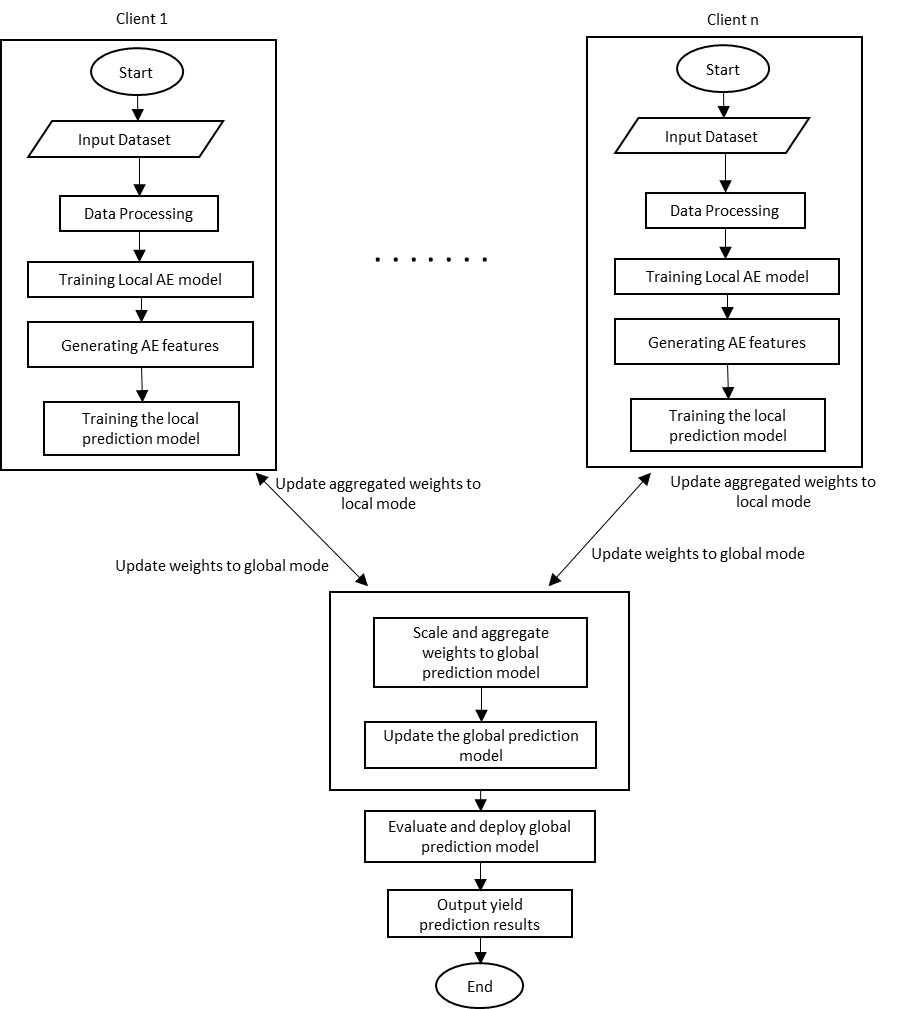
The GRU Autoencoder used a symmetric encoder-decoder with GRU layers (32, 16 units) and a 2-neuron bottleneck, trained using MSE loss, Adam (0.001), batch size 64, for 100 epochs. Bottleneck features (gru\_ae\_feature\_0, gru\_ae\_feature\_1) were appended to inputs and fed into a tuned XGBoost mode.

Table 2 presents the tuned hyperparameters for the XGBoost model. It uses 358 estimators with a learning rate of 0.214 for faster training. A max depth of 3 and min child weight of 4 help prevent overfitting. Subsample (0.826) and colsample\_bytree (0.991) add randomness to improve generalization. Gamma (2.284) controls split regularization, while L1 (0.0514) and very small L2 (≈0.000076) regularization further constrain model complexity.

|  |  |
| --- | --- |
| **TABLE** **2**. Best hyperparameter values for XGBoost model | |
| **Parameter** | **Value** |
| n\_estimators | 358 |
| learning\_rate | 0. 214 |
| max\_depth | 3 |
| min\_child\_weight | 4 |
| subsample | 0. 826 |
| colsample\_bytree | 0. 991 |
| gamma | 2. 284 |
| reg\_alpha | 0. 0514 |
| reg\_lambda | 0. 7591e-05 |

## Model 3: Federated GRU Autoencoder + XGBoost (Fed-GRU-XGB)

The third model used the GRU-Autoencoder + XGBoost pipeline within a Federated Learning (FL) setup. As shown in Figure 1, a global model is shared with multiple clients, each training locally on private data. Clients send updated weights, not raw data, to the server, which aggregates them using Federated Averaging. This process repeats until convergence, resulting in a privacy-preserving, scalable model for crop yield prediction.



**FIGURE 1.** Federated learning process

In this study, the training process was implemented using a federated learning framework, wherein the dataset was evenly and randomly partitioned across 20 simulated clients. Each client operated independently, performing local data preprocessing and training a GRU-based Autoencoder in conjunction with an XGBoost model. Rather than sharing raw data, each client transmitted only the learned model parameters to a central server. The server aggregated these parameters using the Federated Averaging (FedAvg) algorithm and redistributed the updated global model to all clients. This process was repeated over three communication rounds. The GRU Autoencoder architecture used in each client was consistent. All training and aggregation procedures were simulated using CPU-based computation and model evaluation was conducted centrally after each round to assess global performance.

# RESULTS AND DISCUSSION

This section presents a detailed evaluation of the proposed and baseline models for multi-crop yield prediction. It comprises two subsections: the first analyzes correlations between key environmental and agronomic features and crop yield, while the second compares model performance using standard regression metrics. Together, these analyses offer insights into feature significance and the relative effectiveness of centralized versus federated learning approaches in precision agriculture.

## Feature Correlation Analysis

To assess the influence of environmental and agronomic variables on crop yield, Pearson correlation matrices were computed for ten key tropical crops: soybeans, sweet potatoes, wheat, yams, cassava, maize, plantains, potatoes, rice and sorghum. These matrices reveal varying degrees of correlation between crop yield and features like average temperature, annual rainfall and pesticide usage. From Figure 2, (a) to (j), the correlation matrices illustrate the relationship between crop yield and environmental factors for various tropical crops.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| (a) | | (b) | (c) | |
| (d) | (e) | | | (f) |
| (g) | (h) | | | (i) |
| (j) | | | | |

**FIGURE 2.** Correlation Matrix for (a) Cassava, (b) Maize, (c) Plantains, (d) Potatoes, (e) Rice, (f) Sorghum,   
(g) Soybeans, (h) Sweet Potatoes, (i) Wheat and (j) Yams

The correlation analysis indicated that maize (r = -0.55), yams (r = -0.53) and sweet potatoes (r = -0.41), were most negatively affected by temperature yields, while cassava (r = 0.35) yields were positively influenced by temperature yields. We found varied impacts of rainfall: early rainfall had small positive effects on plantains and yams (r ≈ 0.06 - 0.21) and negative effects on maize, potatoes and rice (indicating declines that may have occurred through waterlogging or disease) (Figure 2 yields of soybeans (r = 0.37), potatoes and sorghum (r = 0.20) benefited moderately from pesticide use, reaffirming the key role of pest control. For wheat or plantains, with negligible or negative pesticide correlations, productivity is driven by other agronomic factors. In summary, the analysis shows that crop-by-crop feature influence is very different, with temperature being the most consistent influence. This trend suggests that crop-specific tuning will enhance predictive model performance and will affect feature weighting during model training.

## Comparative Evaluation of Predictive Models

This section provides a systematic comparison of centralized model, hybrid model and federated model for multi-crop yield prediction based on five widely accepted regression metrics, R², RMSE, MAE, MAPE and MSE. In addition to this, visual accuracy assessments were performed.

As indicated in Table 3, the best overall performance was achieved by the XGBoost-GRU-AE hybrid model with the highest R² (0.979) and the lowest MAE, RMSE and MAPE, which proves its efficacy in capturing both temporal patterns as well as nonlinear feature interactions. The performance of the centralized XGBoost (R²=0.975) was also good as it has full access to data over the world. Although accuracy was slightly lower (R² = 0.968), the proposed Fed-GRU-XGBoost model still performed well, while being trained in a decentralized way, indicating that it is well-suited to privacy-sensitive environments. On the other hand, we observed significant performance degradation (R² = 0.886) of the Federated-CatBoost-GRU-AE model, which might be attributed to the non-IID data sensitivity of GRU-AE and its fitness of optimization during federated context.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **TABLE** **3**. Performance comparison of different models | | | | | |
| **Model** | **MAE** | **MAPE** | **MSE** | **R2** | **RMSE** |
| XGBoost | 9397.225 | 0.187 | 264821231 | 0.975 | 16273.329 |
| CatBoost | 8338.738 | 0.132 | 281837466 | 0.974 | 16788.016 |
| CatBoost-GRU-AE | 10067.161 | 0.184 | 497733583 | 0.954 | 22309.944 |
| Federated-Catboost-GRU-AE | 22400.458 | 0.662 | 1231768565 | 0.886 | 35088.927 |
| XGBoost-GRU-AE | 7111.392 | 0.129 | 230617222 | 0.979 | 15186.086 |
| **Proposed Fed-XGBoost-GRU-AE** | **8758.084** | **0.252** | **234989967** | **0.968** | **15329.383** |

Overall, Fed-GRU-XGBoost achieves a good trade-off between performance and data privacy, which is significant for pragmatic implementation of distributed agricultural forecasting.

# CONCLUSION

Crop yield forecasting plays a vital role in maximizing agricultural production and ensuring food security, but traditional centralized methods usually need large amounts of data to be shared, which leads to privacy, ownership and regulatory compliance issues. To tackle these issues, Fed-GRU-XGBoost is triangulates GRU based temporal feature extraction with XGBoost regression and federated learning, allowing localized data with decentralized learned model. On five South Asian countries, it yielded well above average predictive ability (R² = 0.9676, MAE = 8758.08, RMSE = 15329.38), performing not much worse compared to its centralized counterpart (R² = 0.9787). However, this small loss in accuracy is compensated by major improvements in scalability, privacy protection and compliance. Importantly, through a proper trade-off between the performance of the model and the ethical management of the data, Fed-GRU-XGBoost represents a promising approach to build secure, scalable and compliant with regulations crop yield prediction model, contributing to precision agriculture in the data-theoretic world.

# References

1. “World Bank Open Data,” World Bank Open Data, (n.d.).
2. “Gross domestic product and agriculture value added 2013–2022. Global and regional trends,” Statistics, (n.d.).
3. V.G. Dhanya, A. Subeesh, N.L. Kushwaha, D.K. Vishwakarma, T. Nagesh Kumar, G. Ritika, and A.N. Singh, “Deep learning based computer vision approaches for smart agricultural applications,” Artificial Intelligence in Agriculture **6**, 211–229 (2022).
4. S. Sunoj, K. C. McRoberts, M. Benson, and Q. M. Ketterings, “Digital image analysis estimates of biomass, carbon, and nitrogen uptake of winter cereal cover crops,” Computers and Electronics in Agriculture **184**, 106093 (2021).
5. S. M. Shawon, A. N. Adnan, K. M. S. Noor, Z. Kabir, and A. Chakraborty, “Design and Development of Solar Powered IoT Enabled Multi-layer Soil Moisture Probe,” in *2024 27th International Conference on Computer and Information Technology (ICCIT)*, (2024), pp. 3456–3461.
6. A. N. Jui, N. Dey, K. N. -E. -A. Siddiquee, and S.M. Shawon, “Botanika: An IoT-Enabled Solar Powered Soil Profiling Sensor Probe for Rooftop Gardening,” in *2024 3rd International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)*, (2024), pp. 1–6.
7. S. M. Shawon, N. I. Neha, A. N. Jui, N. Dey, and Md.M.S. Raihan, “AgriAI: Machine Learning Frameworks for Tailored Crop Recommendations using Soil Nutrient Parameters,” in *2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS)*, (2024), pp. 1–6.
8. A. Cravero, S. Pardo, S. Sepúlveda, and L. Muñoz, “Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review,” Agronomy **12**(3), 748 (2022).
9. S. M. S. Bukhari, M. H. Zafar, M. A. Houran, S. K. R. Moosavi, M. Mansoor, M. Muaaz, and F. Sanfilippo, “Secure and privacy-preserving intrusion detection in wireless sensor networks: Federated learning with SCNN-Bi-LSTM for enhanced reliability,” Ad Hoc Networks **155**, 103407 (2024).
10. B. K. Goodwin, and A. Hungerford, “Copula‐Based Models of Systemic Risk in U.S. Agriculture: Implications for Crop Insurance and Reinsurance Contracts,” American J Agri Economics **97**(3), 879–896 (2015).
11. J. D. Woodard, “Impacts of Weather and Time Horizon Selection on Crop Insurance Ratemaking: A Conditional Distribution Approach,” North American Actuarial Journal **18**(2), 279–293 (2014).
12. K. Choudhary, W. Shi, Y. Dong, and R. Paringer, “Random Forest for rice yield mapping and prediction using Sentinel-2 data with Google Earth Engine,” Advances in Space Research **70**(8), 2443–2457 (2022).
13. S. N. Khan, D. Li, and M. Maimaitijiang, “A Geographically Weighted Random Forest Approach to Predict Corn Yield in the US Corn Belt,” Remote Sensing **14**(12), 2843 (2022).
14. Y. Li, H. Zeng, M. Zhang, B. Wu, Y. Zhao, X. Yao, T. Cheng, X. Qin, and F. Wu, “A county-level soybean yield prediction framework coupled with XGBoost and multidimensional feature engineering,” International Journal of Applied Earth Observation and Geoinformation **118**, 103269 (2023).
15. J. C. M. Sánchez, H. G. A. Mesa, A. T. Espinosa, S. R. Castilla, and F. G. Lamont, “Improving wheat yield prediction through variable selection using Support Vector Regression, Random Forest, and Extreme Gradient Boosting,” Smart Agricultural Technology **10**, 100791 (2025).
16. F. Huber, A. Yushchenko, B. Stratmann, and V. Steinhage, “Extreme Gradient Boosting for yield estimation compared with Deep Learning approaches,” Computers and Electronics in Agriculture **202**, 107346 (2022).
17. K. Berahmand, F. Daneshfar, E.S. Salehi, Y. Li, and Y. Xu, “Autoencoders and their applications in machine learning: a survey,” Artif Intell Rev **57**(2), 28 (2024).
18. Y. Ma, and H. Li, “GRU-Auto-Encoder neural network based methods for diagnosing abnormal operating conditions of steam drums in coal gasification plants,” Computers & Chemical Engineering **143**, 107097 (2020).
19. S. M. Shawon, F. Barua Ema, A.K. Mahi, and Md. Mohsin Sarker Raihan, “Crop Yield Prediction: Robust Machine Learning Approaches for Precision Agriculture,” in *2023 26th International Conference on Computer and Information Technology (ICCIT)*, (IEEE, Cox’s Bazar, Bangladesh, 2023), pp. 1–6.
20. M. Aggarwal, V. Khullar, S. Rani, T. A. Prola, S. B. Bhattacharjee, S. M. Shawon, and N. Goyal, “Federated Learning on Internet of Things: Extensive and Systematic Review,” CMC **79**(2), 1795–1834 (2024).
21. A. Li, M. Markovic, P. Edwards, and G. Leontidis, “Model pruning enables localized and efficient federated learning for yield forecasting and data sharing,” Expert Systems with Applications **242**, 122847 (2024).
22. A. Durrant, M. Markovic, D. Matthews, D. May, J. Enright, and G. Leontidis, “The role of cross-silo federated learning in facilitating data sharing in the agri-food sector,” Computers and Electronics in Agriculture **193**, 106648 (2022).
23. S. M. Shawon, F. B. Ema, A. K. Mahi, F. L. Niha, and H. T. Zubair, “Crop yield prediction using machine learning: An extensive and systematic literature review,” Smart Agricultural Technology **10**, 100718 (2025).