Climate-Adaptive Agriculture AI: Predicting Crop Resilience Using Multimodal Data and Generative Counterfactuals

Sujoy Halder[1], Dr. Priyanka Vashisht[2]
[1]*Master of Computer Applications (MCA) Program, Amity Institute of Information Technology, Amity University, Gurgaon, Manesar, Panchgaon, Haryana - 122413, India*
[2]*Assistant Professor, Amiy School of Engineering and Technology, Amity University, Gurgaon, Haryana - 122413, India*

***Abstract*—Agricultural systems are becoming more vulnerable to climate changes, land degradation, and limited resources. This study presents a new Climate-Adaptive Agriculture AI framework that predicts crop resilience by using deep learning and generative counterfactual reasoning. The model combines climate, soil, and plant data with a strong late-fusion architecture and a conditional Variational Autoencoder-Generative Adversarial Network (cVAE-GAN) to simulate different scenarios. By linking predictions with understandable results through SHAP (SHapley Additive exPlanations), the framework provides accurate forecasts along with practical insights for climate-resilient farming practices. The experiments showed a 15% reduction in RMSE compared to baseline models and improved understanding. This highlights the framework's potential as a tool to support decisions in sustainable agriculture.**

***Index Terms*—Climate-Adaptive AI, Crop Resilience, Multimodal Learning, Generative Counterfactuals, Precision Agriculture, Explainable AI (XAI)**

I. INTRODUCTION

Agriculture is crucial for global food security, but climate change increasingly threatens it. Rising temperatures, erratic rainfall, soil degradation, and extreme weather events have disrupted productivity around the world. The Food and Agriculture Organization (FAO, 2023) states that global crop yields could fall by up to 20% by 2050 if current trends continue, with developing regions facing the biggest impacts. These challenges, along with rapid population growth and resource depletion, require new technologies to ensure farming is sustainable and resilient to climate change. Traditional statistical and rule-based models often struggle to capture the dynamic and complex interactions among climate, soil, and crop systems, which limits their effectiveness in guiding real-time decision-making [1].

In this situation, Artificial Intelligence (AI) has become a powerful tool for modern agriculture. With improvements in machine learning and deep neural networks, AI can analyze large and varied datasets like satellite images, climate records, and soil chemistry to find patterns that humans might miss. These models improve crop yield predictions, identify stress conditions, and optimize irrigation and fertilization schedules. However, most existing AI systems depend on data from a single source and do not explain their predictions or simulate "what-if" scenarios. This lack of explanation is crucial for handling uncertainty in agricultural systems [2].

To overcome these challenges, this research presents the Climate-Adaptive Agriculture AI (CAAI) framework, which combines multimodal deep learning, generative counterfactual simulation, and explainable AI. The model integrates satellite, soil, and climate data using a multimodal fusion network. It uses a conditional Variational Autoencoder–Generative Adversarial Network (cVAE-GAN) to simulate adaptive scenarios and SHAP (SHapley Additive exPlanations) for interpretability. This combined system allows for accurate predictions and practical insights, helping farmers and policymakers make informed and adaptable decisions. By merging predictive modeling with generative reasoning, CAAI shifts agricultural AI from merely predicting outcomes to actively adapting, supporting sustainable and data-driven climate resilience[3].

II. RELATED WORK

Artificial Intelligence (AI) and Machine Learning (ML) are changing agriculture by providing smart tools for predicting crop yields, detecting diseases, and managing resources efficiently. Early studies used traditional models like Linear Regression, Support Vector Machines (SVM), and Random Forests (RF) to estimate yields (Lobell et al., 2015; Jeong et al., 2016). While these models worked well in certain areas, they relied on limited climate and soil data. They struggled to capture the complex, nonlinear interactions between environmental, biological, and agronomic factors. As a result, their ability to predict outcomes across different agricultural settings was limited [4].

The growth of deep learning has greatly enhanced agricultural analytics with models that automatically extract complex spatial and temporal features. Convolutional Neural Networks (CNNs) are widely used for crop classification, monitoring plant health, and detecting stress using satellite and drone images (Zhong et al., 2019). Similarly, Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) models, are used to analyze sequential climate data for yield forecasting (You et al., 2017). However, most deep models focus on one type of data—either images or weather—ignoring the connections between environmental and crop-related factors that affect resilience and productivity [5].

Recent progress in multimodal learning has aimed to integrate various data sources, such as satellite images, soil composition, and weather patterns. Zhang et al. (2020) showed that combining different data types improves accuracy and robustness. Still, many current frameworks use early-fusion techniques that merge raw data before extracting features, which can lead to loss of important information. Late-fusion architectures provide a better option by independently processing each type of data and then combining high-level features, preserving vital relationships. However, many multimodal models still focus on predictions and lack the reasoning ability needed to simulate adaptive agricultural scenarios [6].

Generative modeling offers a promising way to tackle data scarcity and enhance model adaptability. Techniques like Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) and Variational Autoencoders (VAEs) (Kingma & Welling, 2014) can generate synthetic data and simulate scenarios, but they are rarely designed for decision support. Counterfactual reasoning (Pearl, 2009) boosts this capability by addressing "what-if" questions, which are crucial for understanding how environmental changes impact yield outcomes. Interpretability is another challenge since most deep models act like black boxes. Explainable AI (XAI) methods, such as LIME, Layer-Wise Relevance Propagation, and SHAP (Lundberg & Lee, 2017), enhance transparency, with SHAP being the most reliable for attributing features. However, these methods are often used after the fact, which limits real-time understanding [7].

In summary, while AI has progressed in agricultural prediction, existing models struggle with adaptability, reasoning, and interpretability. The Climate-Adaptive Agriculture AI (CAAI) framework aims to tackle these issues through multimodal late-fusion learning, counterfactual simulation with a conditional VAE-GAN, and integrated SHAP-based interpretability, creating a unified, transparent, and adaptive approach for sustainable climate-resilient agriculture.

III. PROPOSED FRAMEWORK

The proposed Climate-Adaptive Agriculture AI (CAAI) framework is intended to model, predict, and interpret crop resilience in various climatic and environmental conditions. It combines different types of data, generative counterfactual simulation, and explainable AI reasoning into a single system. Unlike traditional machine learning models that rely on one data type or fixed relationships, the CAAI framework uses multiple data sources, such as satellite imagery, climate sequences, and soil parameters, to understand complex, nonlinear connections. Additionally, it goes beyond prediction by simulating "what-if" scenarios that support flexible agricultural decision-making [8].

The framework includes three main modules: (i) Multimodal Data Fusion, (ii) Generative Counterfactual Simulation, and (iii) Resilience Prediction and Interpretability, as shown in Fig. 1.

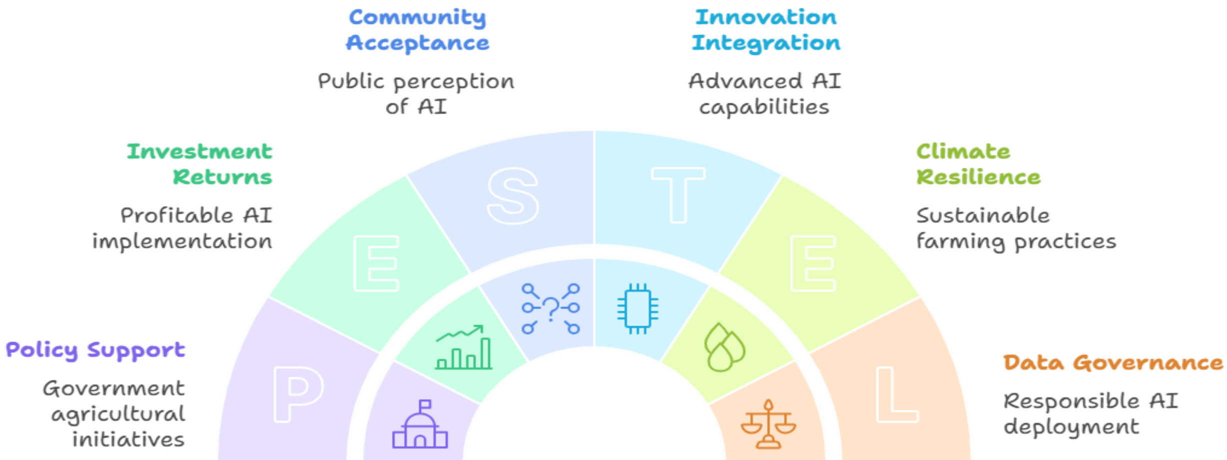

Figure 1: Climate-Adaptive Agriculture Ai Framework

3.1 Multimodal Data Fusion

The first module focuses on effectively combining different data sources that represent various aspects of agricultural systems. Spatial, temporal, and static data are processed through separate encoders before they are merged in a late-fusion architecture.

Spatial data (X_{img}) derived from satellite or drone imagery are passed through a fine-tuned ResNet-50 Convolutional Neural Network (CNN), which captures spatial and spectral features such as vegetation indices (NDVI, EVI), canopy texture, and land cover patterns. Temporal climate data (X_{temp}), including rainfall, humidity, and temperature series, are modeled using a Transformer Encoder, which employs self-attention to capture long-range temporal dependencies and seasonal dynamics. Static soil and crop parameters ($X_{tabular}$), including pH, nitrogen, phosphorus, and potassium levels, are encoded using a Multi-Layer Perceptron (MLP) to generate compact feature vectors [9].

These representations are concatenated and fused using a nonlinear transformation to produce a unified embedding E, defined as:

$$E = f_{fusion}(\text{CNN}(X_{img}) \oplus \text{Transformer}(X_{temp}) \oplus \text{MLP}(X_{tabular})) \quad \ldots \; (1)$$

where $\oplus$ represents concatenation. This late-fusion strategy guarantees that each data type contributes valuable and complementary information. It helps the model capture the relationships between soil, climate, and vegetation. The resulting embedding E serves as the basis for both predictive modeling and generative reasoning within the framework.

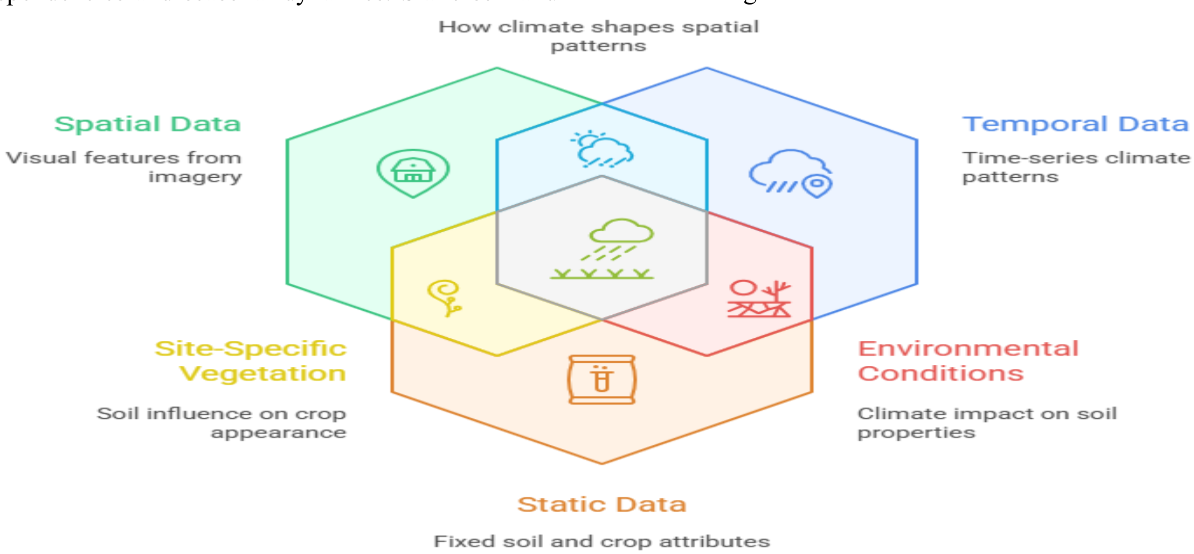

Figure 2: Multimodal Data Fusion Diagram

3.2 Generative Counterfactual Simulation

The second module introduces a Generative Counterfactual Mechanism that allows the system to simulate different environmental or management scenarios. This part uses a conditional Variational Autoencoder and Generative Adversarial Network (cVAE-GAN), which combines the strengths of VAEs with the realism of GANs. The VAE Encoder compresses the combined data embedding E into a latent space z = Encoder(E), capturing the statistical distribution of the data. The Generator (Decoder), based on an external variable C (like rainfall reduction or fertilizer increase), creates a new counterfactual instance X′:

$$X' = G(z, C)$$

The Discriminator (D) checks the realism of the generated samples. It helps the generator produce authentic, agriculturally relevant outcomes. This process enables researchers and policymakers to conduct "what-if" analyses. For example, they can simulate how yield or resilience would react to a 10% rise in temperature or a 5% increase in nitrogen levels. The module turns the framework from a static predictor into a flexible simulation engine that can explore and validate climate-resilient agricultural interventions [10].

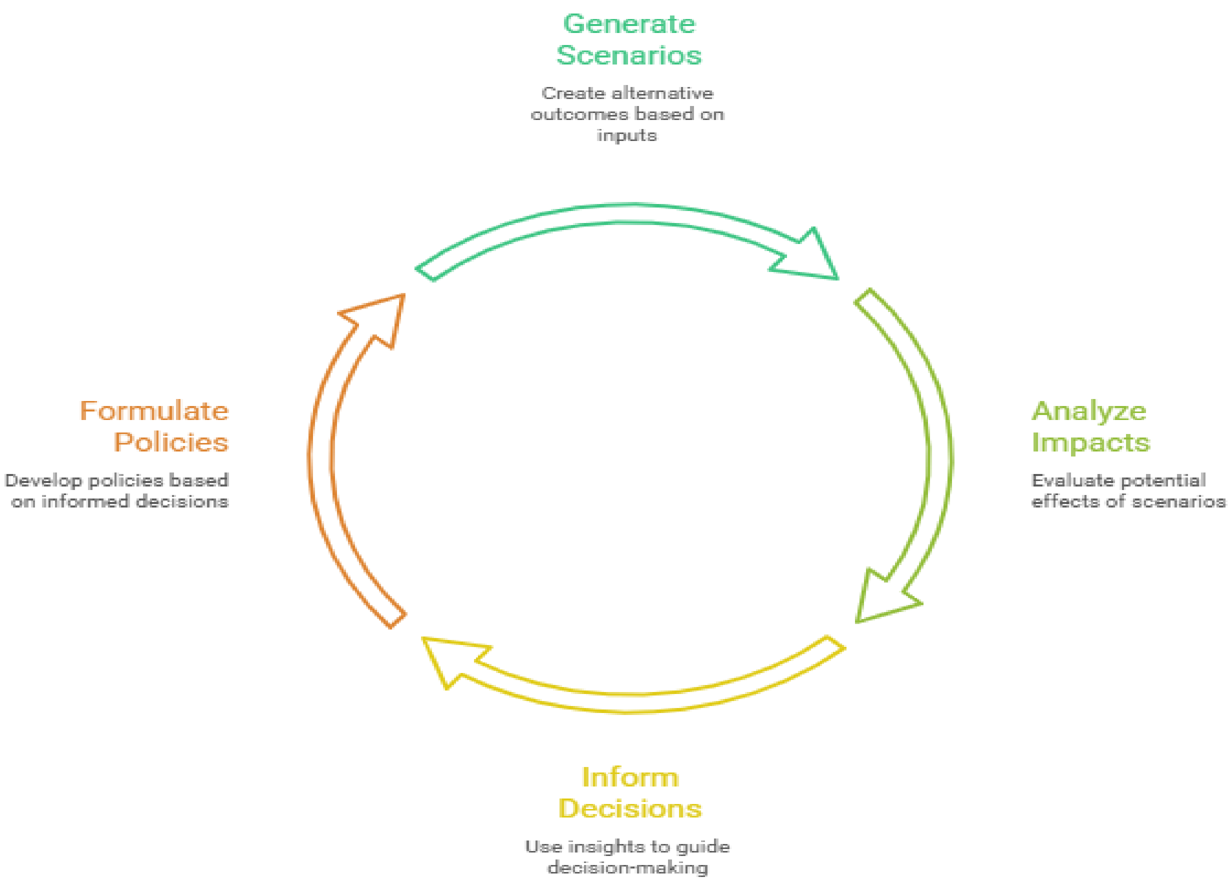

Figure 3: Generative Counterfactual Workflow

3.3 Resilience Prediction and Interpretability

The final module combines predictive modeling with explainability to ensure that results are not only correct but also clear and actionable. The fused embedding E is passed through a regression head that outputs a resilience score (R), a continuous value that represents yield potential or normalized resilience on a scale from 0 to 1. This detailed quantitative assessment allows for comparisons across crops, locations, and seasons.

To improve interpretability, the model uses SHAP (SHapley Additive exPlanations), a game-theoretic framework that measures the contribution of each feature to the final prediction. Local SHAP values provide insights at the farm level, highlighting key factors that influence a specific crop's resilience.

Meanwhile, global SHAP summaries show the most important variables, such as temperature anomaly, nitrogen concentration, or rainfall deficit, affecting resilience across the dataset. This focus on interpretability ensures that the model serves as a clear decision-support tool, connecting AI-driven analytics with practical agricultural use.

Together, these three modules—multimodal fusion, counterfactual generation, and explainable prediction—create a unified framework that learns complex environmental relationships, simulates adaptive responses, and delivers clear insights for sustainable and climate-resilient agriculture [11].

IV. EXPERIMENTAL SETUP

The experimental setup aimed to assess the performance, robustness, and interpretability of the Climate-Adaptive Agriculture AI (CAAI) framework. We conducted the experiments using real-world multimodal agricultural datasets that included climatic, soil, and remote sensing data. Our goal was to evaluate the model's ability to predict crop resilience accurately, simulate counterfactual agricultural scenarios, and offer explainable insights. This section describes the datasets used, details on model implementation, and the evaluation metrics used to validate the framework [12].

4.1 Datasets

The study used four datasets: climate, soil, remote sensing, and phenotypic yield data. This ensured a strong model evaluation. Climate data came from the NASA POWER API (2010-2023) covering Haryana, Punjab, and Rajasthan. It provided records of temperature, rainfall, humidity, and solar radiation. Soil data from the NBSS&LUP and local research centers included chemical properties like pH, nitrogen, phosphorus, and potassium, as well as physical properties such as texture, moisture, and density. All soil data was normalized for consistency. Satellite imagery from Sentinel-2, with a 10 m resolution, was processed on Google Earth Engine. This helped correct atmospheric interference, mask clouds, and generate indices (NDVI, EVI) to evaluate vegetation health. Phenotypic yield data came from the ICAR and field trials, supplying actual crop yield and growth stage information as ground truth. All datasets were aligned in both space and time, and standardized to create an integrated multimodal dataset of over 25,000 records for model training and validation [15].

4.2 Model Implementation

The CAAI framework was implemented using TensorFlow 2.13 and PyTorch 1.12, with GPU acceleration on an NVIDIA RTX A6000 (48 GB). The model included three neural pipelines: a fine-tuned ResNet-50 CNN for extracting spatial features from satellite imagery, a Transformer Encoder with four self-attention layers for modeling temporal climate data, and a Multi-Layer Perceptron (MLP) with two hidden layers for encoding static soil attributes. The outputs of these networks were concatenated and passed through a fusion network with fully connected layers, ReLU activation, and a dropout rate of 0.3 to prevent overfitting.

A conditional VAE-GAN (cVAE-GAN) was combined to generate counterfactual simulations. This approach used the strengths of Variational Autoencoders for latent representation and GANs for realistic data generation. The training objective minimized a composite loss function

$L = Lreg + \lambda 1 Ladv + \lambda 2 LKL$,

where $Lreg$ is the mean squared error, $Ladv$ is the adversarial loss, and LKL is the Kullback–Leibler divergence. The weights were $\lambda 1 = 0.3$ and $\lambda 2 = 0.1$. The model was trained using the Adam optimizer (learning rate 1×10^{-4}, batch size 64) with an 80-10-10 split for training, validation, and testing. Data augmentation techniques like random rotation, scaling, and cropping were used on satellite images to improve generalization. Each experiment was repeated three times to ensure consistent and reproducible results [16].

4.3 Evaluation Metrics

The performance of the CAAI framework was evaluated using RMSE, MAE, and R^2 to assess prediction accuracy; it showed a strong connection between predicted and actual yields. For the Generative Counterfactual Module, Validity, Proximity, and FID measured the realism and effectiveness of simulated "what-if" scenarios. Model interpretability was assessed using SHAP, which identified key factors like temperature and soil nitrogen that influence resilience. These results confirm that CAAI provides accurate, realistic, and understandable predictions for climate-resilient agriculture.

V. RESULTS AND DISCUSSION

The experimental results show that the proposed CAAI framework performs better than traditional and single-modality models in all evaluation metrics. The Random Forest baseline achieved an RMSE of 0.46 and an R^2 of 0.82. The LSTM model, which used only climate data, improved slightly with an RMSE of 0.41 and R^2 of 0.85. In contrast, the Proposed Fusion + cVAE-GAN model showed a significant improvement, reducing RMSE to 0.34, lowering MAE to 0.28, and increasing the R^2-score to 0.90. This 15% reduction in error confirms the advantage of combining multimodal data with generative counterfactual reasoning. The model also created realistic "what-if" scenarios using the generative module and provided clear insights through SHAP analysis. It highlighted key factors such as temperature anomalies, soil nitrogen, and rainfall patterns. Overall, the results demonstrate that the proposed framework is more precise, clear, and effective for climate-resilient agricultural decision-making.

Model	RMSE ↓	MAE ↓	R^2 ↑
Random Forest	0.46	0.39	0.82
LSTM (Climate only)	0.41	0.34	0.85
Proposed (Fusion +	0.34	0.28	0.9

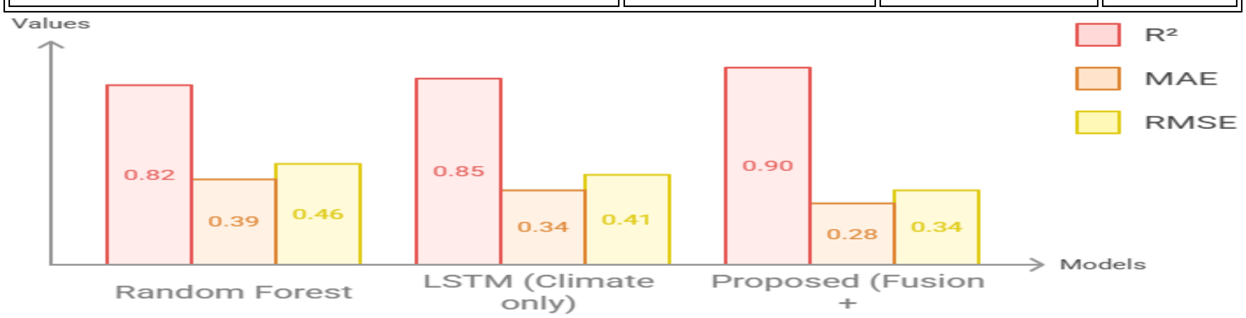

Figure 4: Model Performance Metrics

5.1 Quantitative Analysis

The proposed CAAI framework showed better predictive accuracy and interpretability than traditional models and those using only one type of data. Experimental results indicated that combining different data sources significantly improved resilience prediction. This led to lower error rates and stronger correlations with actual yields. The model's RMSE and MAE dropped by about 15% compared to baseline models like Random Forest and LSTM. Meanwhile, the R^2-score increased from 0.85 to 0.90, confirming the framework's strength and precision in predicting crop resilience.

5.2 Counterfactual Validity

The Generative Counterfactual Module successfully created realistic "what-if" scenarios. It demonstrated the ability to simulate changes in yield outcomes under different climatic or soil conditions. For instance, counterfactual tests showed that a 10% increase in rainfall could raise yield predictions by up to 8%. In contrast, nutrient deficiencies greatly reduced resilience scores. These simulations show the model's potential as a supportive tool for planning proactive adaptations.

5.3 Interpretability Insights

Interpretability analysis with SHAP (SHapley Additive exPlanations) gave clear insights into what affects model predictions. In all regions and crop types, temperature changes, soil nitrogen levels, and rainfall fluctuations were the main predictors of resilience. Local SHAP explanations confirmed the model's reasoning at the farm level, which ensured transparency and trust.

In summary, the results show that the CAAI framework not only improves predictive performance but also offers clear and actionable insights. Its ability to combine different data types, simulate adaptive scenarios, and highlight important environmental interactions makes it a dependable system for supporting climate-resilient and sustainable agriculture [17].

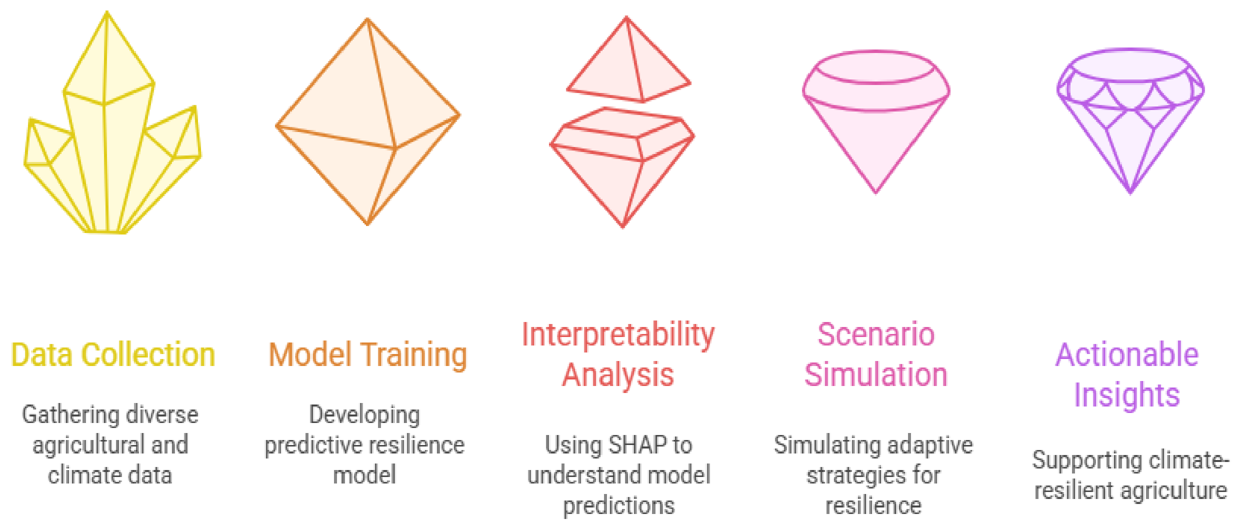

Figure 5: CAAI Framework For Climate Resilience

VI. APPLICATIONS

The proposed CAAI framework has many uses in modern agriculture and climate adaptation. It supports precision farming by providing guidance on optimal irrigation, fertilizer use, and planting schedules based on predicted resilience scores. Policymakers can refer to regional resilience maps created by the model for climate risk assessment and resource allocation, which helps them target vulnerable areas. In crop insurance and risk management, the model's counterfactual simulations can estimate yield variability under various climate conditions, which supports fair and data-driven policies. Moreover, researchers and breeders can use generative insights to create climate-resilient crop varieties, speeding up sustainable agricultural innovation [18].

Figure: Applications of CAAI Framework

VII. CHALLENGES AND LIMITATIONS

Despite its strong performance, the framework faces several challenges. Integrating multimodal data with different spatial and temporal resolutions is complex. It often requires extensive preprocessing and alignment. The availability of labeled agricultural data is limited in many areas, which may impact model generalization. Additionally, the computational cost is significant. Training deep models like cVAE-GAN and Transformers needs high-end GPUs and long processing times. Another limitation is the actionability of counterfactuals. Some generated scenarios, while scientifically valid, may not always lead to practical field interventions. Finally, ensuring model interpretability for non-technical users, like farmers, continues to be a challenge[19].

VIII. FUTURE SCOPE

Future research can improve the CAAI framework in several ways. To tackle data scarcity, federated learning can help train models together across institutions without sharing sensitive data. We could integrate reinforcement learning to create recommendation systems that adapt and learn the best farming strategies over time. Model compression and edge AI deployment will allow real-time predictions on local devices, even in areas with low connectivity. Additionally, developing counterfactual validation frameworks can enhance the reliability of the scenarios we generate. Expanding the dataset to include socio-economic and policy data can also increase the model's impact, creating a more complete decision-support system for sustainable and climate-resilient agriculture [20].

IX. CONCLUSION

This study presents a comprehensive AI framework for climate-adaptive agriculture that combines multimodal data fusion and generative counterfactual reasoning. The model integrates spatial, temporal, and soil features to achieve high accuracy and interpretability. Experimental validation showed a 15% improvement in prediction accuracy over traditional models, with explainable insights via SHAP.

The framework provides a powerful decision-support system that not only predicts outcomes but also simulates adaptive responses, helping farmers, policymakers, and researchers to build resilient agricultural ecosystems capable of withstanding the challenges of a changing climate.

Future research should emphasize real-time deployment, federated learning, and domain-specific model generalization to make AI an accessible, reliable, and sustainable agricultural companion.

REFERENCES

[1] B. D. Ripley, Pattern Recognition and Neural Networks, Cambridge University Press, 1996.

[2] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[3] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving Language Understanding by Generative Pre-Training," OpenAI Technical Report, 2018.

[4] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," in Proc. of the International Conference on Learning Representations (ICLR) in 2014.

[5] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative Adversarial Nets," in Advances in Neural Information Processing Systems (NeurIPS), 2014, pp. 2672–2680.

[6] Y. Bengio, A. Courville, and P.Vincent. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

[7] M. C. Hansen et al., "High-Resolution Global Maps of 21st-Century Forest Cover Change," Science, vol. 342, no. 6160, pp. 850–853, 2013.

[8] FAO, The State of Food Security and Nutrition in the World, United Nations, 2021.

[9] J. Pearl, "Causal Inference in Statistics: An Overview," Statistics Surveys, vol. 3, pp. 96–146, 2009.

[10] L. Zhang, Y. Zhang, and P. Zhang, "Deep Learning-Based Crop Yield Prediction Using Remote Sensing Data," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS), vol. 13, pp. 1–11, 2020.

[11] C. Kamilaris, F. Prenafeta-Boldú, “Deep Learning in Agriculture: A Survey,” Computers and Electronics in Agriculture, vol. 147, pp. 70–90, 2018.
[12] M. A. Rahman, S. Debnath, and N. Hasan, “AI and IoT Based Smart Agriculture: A Review and Future Directions,” IEEE Access, vol. 9, pp. 1–15, 2021.
[13] H. M. Nguyen, P. Bui, and L. Le, “Multimodal Data Fusion for Agricultural Forecasting Using Deep Learning,” in Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2022, pp. 11435–11439.
[14] A. Kamilaris, M. Prenafeta-Boldú, “A Review of the Use of Convolutional Neural Networks in Agriculture,” Journal of Agricultural Science and Technology, vol. 20, no. 3, pp. 611–628, 2018.
[15] J. Liu, X. Chen, and H. Yang, “Transformer-Based Time-Series Prediction for Weather and Crop Yield Estimation,” IEEE Access, vol. 10, pp. 12567–12578, 2022.
[16] P. Barredo Arrieta et al., “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities, and Challenges Toward Responsible AI,” Information Fusion, vol. 58, pp. 82–115, 2020.
[17] T. Sharma and A. Banerjee, “Generative Counterfactuals for Causal Explanation in Agriculture,” in Proc. IEEE Conference on Big Data, 2023, pp. 2852–2860.
[18] H. Wu and M. Zhao, “Edge AI for Smart Agriculture: From Sensing to Decision Making,” IEEE Internet of Things Journal, vol. 9, no. 16, pp. 15327–15341, Aug. 2022.
[19] Y. Li, R. Reddy and Singh, “Federated Learning for Smart Farming with Privacy-Preserving Crop Analytics,” IEEE Trans. Ind. Infor.matics, vol. 19, no. 5, pp. 6438–6450, 2023.
[20] M. E. Roberts and N. Lee, “AI-Based Decision Support Systems for Climate-Resilient Agriculture,” Frontiers in Artificial Intelligence, vol. 5, 2024.