International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 11s, 2025
https://theaspd.com/index.php

Deep Learning-Based Crop Prediction Using LSTM And GRU For Sustainable Agriculture

Mr.Ravi Kumar Munaganuri[1], Dr. Pavani Gonnuri[2],Dr Vishnu Vardhana Devi V[3],Ramakrishna Badiguntla[4], Aravinda Kasukurthi[5], Dr. Seethiraju L.V.V.D.Sarma[6], Shaik Mahaboob Basha[7],Dr. Narasimha Rao Yamarthi[8,]

[1]School of Computer Science and Engineering, VIT-AP University, Amaravati, Guntur, Andhra Pradesh, India, ravikumar.m@vitap.ac.in

[2]Assistant Professor of English, Velagapudi Ramakrishna Siddhartha School of Engineering ,Siddhartha Academy of Higher Education(Deemed to be University),Vijayawada,AP,India, pawanigonnuri@gmail.com

[3]Assistant Professor of English, Velagapudi Ramakrishna Siddhartha School of Engineering ,Siddhartha Academy of Higher Education(Deemed to be University), Vijayawada,AP,India, vandanaenglish10@gmail.com

[4]Department of Computer Science and Engineering (Data Science), RVR&JC College of Engineering, Guntur, Andhra Pradesh, India, Ramakrishna.badiguntla@gmail.com.

[5]Department of Computer Science and Engineering (Data Science), RVR&JC College of Engineering, Guntur, Andhra Pradesh, India, aravindakasukurthi@gmail.com

[6]Department of Computer Science and Engineering, Vasireddy Venkatdri institute of Technology,Namburu, Guntur, Andhra Pradesh, India, slvvdsarma@gmail.com

[7]School of Computer Science and Engineering, VIT-AP University, Amaravati, Guntur, Andhra Pradesh, India, mahaboob.23phd7064@vitap.ac.in

[8*]School of Computer Science and Engineering, VIT-AP University, Amravati, Guntur, Andhra Pradesh, 522237, India, y.narasimharao@vitap.ac.in

Abstract

The rapid development of technology in agriculture has opened the door to data-driven solutions to increase crop productivity and sustainability. This paper introduces a machine learning-based method for crop prediction through deep learning models such as long short-term memory (LSTM) and gated recurrent units (GRU). By exploiting environmental factors such as soil content, temperature, moisture, and nutrient levels, the suggested model accurately suggests the most suitable crops for a specific area. Preprocessing operations such as feature scaling, principal component analysis (PCA), and one-hot encoding are applied to the dataset derived from Kaggle, which enhance the prediction effectiveness. Experimental results reflect the model's ability to optimize crop choices and, subsequently, reduce the associated risks from unknown climate changes, soil degradation, and less efficient use of resources. This work highlights the impact of deep learning in agricultural decision-making that supports higher productivity and sustainability.

Index Terms–Crop Prediction, Machine Learning, Deep Learning, LSTM, GRU, Precision Agriculture, Sustainable Farm- ing, Climate Resilience, Resource Optimization, AI in Agricul- ture.

INTRODUCTION

Agriculture is perhaps the most important industry for human survival, providing livelihoods, raw materials, and food for billions of people around the world. With the world popula- tion projected to reach nearly 10 billion by 2050, maintaining food security has become a major concern [1], [2]. Traditional farming methods are becoming increasingly unreliable due to challenges such as climate change, soil degradation, water scarcity, and pest threats. To address these challenges, modern agriculture is looking for data-intensive solutions that increase productivity and sustainability. Crop forecasting is perhaps the most significant development in the field, as it plays a significant role in increasing agricultural productivity [3].Crop forecasting uses scientific and technological tools to identify suitable crops for planting in a given area based on soil type, weather patterns, temperature, moisture, and nutrient

availability. Accurate crop forecasting helps farmers make smart choices, reduce the chance of crop failure, and increase overall productivity [4]. By using past data and current environmental models, farmers are able to select the most ideal crops for the land, avoiding the risks associated with inappropriate crop selection. This allows for the establishmentof optimal sowing dates, irrigation rates and fertilizer use, ensuring favorable conditions for crops to grow [5].Climate change has made agriculture more uncertain, with extreme weather patterns such as heat waves, floods and droughts threatening crop production. Traditional farming methods are unable to cope with these rapid changes and many farmers are suffering great losses [6]. However, crop forecasting models use weather data and weather forecasts to enable farmers to predict adverse conditions and avoid losses [7]. This allows them to switch to more resistant crops, modify planting calendars and use better water management practices to reduce risks.Pest and disease infestations are also a significant agri- cultural problem. Once an infestation is detected, farmers adopt response strategies such as the use of pesticides, but this leads to crop loss and excessive chemical consumption [8]. Crop forecasting models based on machine learning and deep learning can predict potential infestations by examining environmental conditions and past pest trends. Farmers can take steps such as biological pest control or changing planting plans to reduce the risk of infestation while conserving the use of toxic chemicals.

Another benefit of crop forecasting is efficient resource management. One of the most valuable assets in agriculture is water, yet many farmers operate under inefficient irrigation strategies. By combining crop forecasting with smart irrigation technologies, farmers can maximize water use with real-time monitoring of soil moisture and weather forecasts. Fertilizer use can also be optimized to provide plants with the nutrients they need without overuse, which can have a negative impact on the soil after years of cultivation [9].The economic contribution of crop forecasting is also high because it helps farmers reduce losses and increase profits. Farmers can accurately predict their crop yields, so that they can devise better strategies for harvesting, storage, and marketing. This reduces post-harvest losses and protects them from selling their produce at the wrong time for low prices. Furthermore, governments and policymakers can use crop forecasting data to design more effective agricultural policies, allocate resources more efficiently, and help farmers make better decisions [10].

Over time, crop forecasting will become more sophisticated,incorporating advanced tools such as satellite imagery, Inter- net of Things (IoT) sensors, and real-time analytics. These new technologies will continue to increase the precision of agriculture, allowing it to be more resilient, more efficient, and more sustainable. The future of agriculture lies in using the power of data and technology to make better choices and ensure food security for future generations while protecting the environment.

METHODOLOGY

Dataset

In this paper, we have used the crop recommendation dataset obtained from Cagle and its aim is to help predict the most suitable crop for planting given environmental and soil conditions. It includes a wide range of attributes such as important agronomic parameters such as nitrogen (N), phosphorus (P), potassium (K), temperature, moisture, soil pH and rainfall. In addition, the data includes information about crop water requirements, market price per kilogram and average water requirements, thereby increasing its value for crop planning and decision-making. A row in the data represents a specific combination of such input attributes and a target variable, which indicates the recommended crop type, e.g., rice, wheat, chickpea, pigeon pea or watermelon. The dataset is well-structured and balanced, making it suitable for training machine learning models such as LSTM and GRU for crop recommendation purposes. Its practical applicability and rich feature set make it a useful tool for farmers, researchers and data scientists to optimize farming practices using data- driven insights.

Pre-processing data

Feature cleaning and selection help to refine the dataset by keeping only the most important information. The process begins by separating the target variable or crop type from the rest of the dataset. This is done to ensure that the model can learn from independent features without being biased by the output labels while training. Another column related to water requirements is checked for deletion with a

conditional check. If the column exists, it is dropped along with the target variable. This process ensures that only the necessary numeric and categorical features are left, eliminating duplicates or unnecessary information that would add noise to the model.The crop type classification is in text form, it needs to be converted to numeric format before being input into the model. This is achieved through label encoding, where each distinct crop name gets a specific integer value. But treating the categorical data as numeric values makes the model assume that these have an ordinal relationship, which is not the case. To avoid this, one-hot encoding is used. This process converts the numerical labels into vectors of binary values, where each crop type gets a different set of 0s and 1s. Using one-hot encoding, the model treats all crop types as individual classes without assuming any ranking among them internally.Feature scaling ensures that all numerical features affect the model equally and prevents some variables from overshadowing others due to their large values. Standardization is used here using a standard scaler, where the data is transformed so that each feature has a mean value of zero and a standard deviation of one. This process helps to reduce the differences between scales for different features, so that the model can learn efficiently. Standardization is mostly used in deep learn- ing models such as LSTMs and GRUs because it speeds up training convergence and increases the stability of the overall model.

Feature Extraction

Feature extraction is an important preprocessing step that helps reduce the complexity of a dataset while preserving key information. Here, feature extraction is done using Principal Component Analysis (PCA). PCA is a dimensionality reduc- tion method that reduces the number of new features, called principal components, from the original set of features. These elements select the greatest variation in the data, that is, they retain the most complex patterns and remove unnecessary or less useful information.PCA is used in this process to handle 95% of the variation in the dataset. That is, instead of using all the original features, the algorithm selects the most informative ones, filtering out noise and redundancy. The number of components selected depends on how well they explain the variation in the dataset. With this, the model gains advantages in several aspects: it reduces computational complexity, avoids overfitting, and makes training more efficient.

After PCA transforms the dataset, the extracted features are reshaped to fit the input format of deep learning models such as LSTM and GRU. Since these models require a three- dimensional input structure, the data is reshaped before being fed into the network. Feature extraction using PCA provides the most compact and relevant representation of the dataset to the model, resulting in improved performance and efficiency.

LSTM-GRU

The LSTM-GRU model is designed to analyze agricultural and environmental data to recommend the most suitable crop for a given set of conditions. By utilizing deep learning, this model processes time-dependent factors such as soil nutrients, rainfall, temperature, and humidity, ensuring that historical trends and real-time data contribute to accurate predictions. The combination of Long Short-Term Memory (LSTM) net- works and Gated Recurrent Units (GRU) allows for effective handling of both long-term and short-term dependencies in the dataset. LSTM layers capture long-term trends in climate and soil conditions, while GRU layers refine these predictions by efficiently managing recent fluctuations in environmental parameters.LSTM networks use a memory cell mechanism that controls information flow through three primary gates: the forget gate, input gate, and output gate. The forget gate determines how much past information should be retained by computing a

Algorithm 1 Principal Component Analysis (PCA)

1: **Load the dataset**
2: Extract feature matrix X (excluding target variable)

3: **Standardize the features**
4: Compute mean μ and standard deviation σ for each feature
5: Normalize: $X_{scaled} = \frac{X - \mu}{\sigma}$
6: **Compute the covariance matrix**

The updated memory cell is computed as:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

where C_t stores information that influences future predic- tions. The final output from the LSTM layer is determined by the output gate:

7: Covariance matrix: $\Sigma = \frac{1}{n}(X^T X_{scaled})$
8: **Compute eigenvalues and eigenvectors**
9: Solve: $\Sigma v = \lambda v$
10: λ (eigenvalues) represent variance, v (eigenvectors) define principal components
11: **Sort eigenvalues and eigenvectors**
12: Rank eigenvalues in descending order
13: Select top k eigenvectors corresponding to the highest

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

and updates the hidden state as:

$$h_t = o_t * \tanh(C_t) \quad (6)$$

This hidden state is then passed to the next layer of the model, helping in learning the relationship between different

eigenvalues

$\Sigma \lambda$ environmental factors and crop yield.

14: Retain components that preserve 95% variance: $\frac{\Sigma}{\lambda_{total}} \geq 0.95$
15: **Transform the dataset**
16: Project X_{scaled} onto selected principal components:
17: $X_{pca} = X_{scaled} V_k$
18: **Output transformed dataset**
19: Use X_{pca} for model training

On the other hand, GRU layers streamline this process by replacing the three gates of LSTM with just two: the reset gate and the update gate. The reset gate is defined as:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (7)$$

determining how much past information should be for- gotten. If the reset gate value is small, older information is ignored, making GRUs particularly useful when recent

sigmoid activation function on the concatenation of the previ- ous hidden state and the current input. This is mathematically represented as:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

conditions, such as sudden rainfall or drought conditions, are more relevant than past data. The update gate is given by:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (8)$$

which balances the retention of old information with new

where W_f represents trainable weights, h_{t-1} is the hidden inputs, ensuring that the model adapts dynamically to changing environmental factors. The hidden state is then computed as:
state from the previous step, and x_t is the current input data.
If the forget gate outputs a value close to 1, the previous information is retained; if it is close to 0, it is discarded. This ensures that only relevant historical trends, such as seasonal changes in soil fertility or temperature fluctuations, are remembered for crop recommendation.
The input gate then regulates the addition of new informa- tion by computing another sigmoid function:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

and a candidate memory content

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (9)$$

where $\tilde{h}_t$ represents the candidate activation. Unlike LSTM, GRU reduces the computational complexity by eliminating the memory cell and instead directly updating the hidden state, making it more efficient for real-time crop prediction.
Once the LSTM and GRU layers extract relevant features from the environmental data, the output is passed through fully connected dense layers. These layers apply an activation func- tion, typically ReLU (Rectified Linear Unit) for hidden layers and softmax for the output layer, to classify and recommend the most suitable crop. The softmax function converts the final output into probability values, ensuring that the model assigns
a confidence score to each possible crop type. The equation for softmax is given by:

$$P(y_i) = \frac{e^{z_i}}{\sum_j} \quad (10)$$

Algorithm 2 LSTM-GRU Model for Crop Classification
1: **Input:** Preprocessed feature matrix X_{train}, X_{test}, and one-hot encoded labels y_{train}, y_{test}
2: **Output:** Trained LSTM-GRU model and classification predictions
where z_i is the output of the last dense layer before activa- tion, and $P(y_i)$ represents the probability of recommending a particular crop. The crop with the highest probability is selected as the final recommendation.
This entire process is optimized through categorical cross- entropy loss, which is computed as:

$$L = -\sum_i y_i \log(\hat{y}_i) \quad (11)$$

where y_i is the actual label, and $\hat{y}_i$ is the predicted probability from the softmax function. The model is trained using the Adam optimizer, which updates weights based on the computed loss, ensuring that the recommendation system improves over multiple training epochs. By integrating both LSTM and GRU, the model effectively learns patterns in soil fertility, climate variations, and previous crop yields. It ensures that historical trends (captured by LSTM) and recent changes (managed by GRU) collectively influence the final crop recommendation. This makes the system robust, adaptive, and efficient in recommending the best crops for sustainable agriculture.

EXPERIMENTAL RESULTS

The results demonstrate that the model effectively learns from the dataset and achieves high classification accuracy in recommending suitable crops. The feature correlation heatmap provides valuable insights into relationships between agricul- tural parameters, which can be useful for further optimization. The

confusion matrix highlights the model's strengths and weaknesses, helping refine the recommendation system by addressing misclassifications. The loss and accuracy trends indicate stable training with no significant overfitting, con- firming that the model is well-optimized for predicting suit- able crops based on environmental and soil conditions. The results presented in the visualizations provide insights into the model's performance and data relationships in the crop recommendation system. Here is a detailed explanation of each visualization:

Feature Correlation Heatmap

The feature correlation heat map shows the correlation between different variables in a dataset. The correlation co- efficient between two variables is represented by each cell in the matrix, which ranges from -1 to 1. A positive correlation (close to 1) means that as one variable increases, the other also increases, while a negative correlation (close to -1) means that there is an inverse relationship. For example, nitrogen (N) has a high positive correlation with water requirements, which indicates that crops with high nitrogen demands also have relatively high water demands. Temperature and humidity also have significant correlations with other environmental factors,

3: **Begin**
4: Initialize Sequential Model
5: **Add LSTM Layer**
6: Set number of units to 64
7: Define input shape as (*timesteps* = 1, *features* = *number of PCA components*)
8: Set return_sequences = True
9: **Add Dropout Layer** (Drop probability = 0.2)
10: **Add GRU Layer**
11: Set number of units to 64
12: Set return_sequences = False
13: **Add Dropout Layer** (Drop probability = 0.2)
14: **Add Fully Connected Dense Layer**
15: Set number of neurons to 128
16: Use ReLU activation function
17: **Add Dropout Layer** (Drop probability = 0.2)
18: **Add Output Dense Layer**
19: Set number of neurons = number of classes
20: Use softmax activation function
21: **Compile Model**
22: Set optimizer = Adam
23: Set loss function = categorical cross-entropy
24: Set metric = accuracy 25: Display Model Summary 26: **Train Model**
27: Fit model using X_{train} and y_{train}
28: Set batch size = 32
29: Set epochs = 100
30: Use 20% validation split
31: **Evaluate Model**
32: Compute test loss and accuracy using X_{test} and y_{test}
33: **Make Predictions**
34: Use trained model to predict labels on X_{test}
35: Convert predicted probabilities to class labels
36: **End**

which demonstrate their control over crop growth. A heat map can help determine redundant traits and how different traits affect crop yield and suitability.

Confusion Matrix

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 11s, 2025
https://theaspd.com/index.php

The confusion matrix evaluates the performance of a trained model by comparing the true and predicted labels. The row represents the actual class and the column represents the predicted class. Diagonal values show correct classifications, while off-diagonal values show incorrect classifications. A large number of correct classifications (dark blue along the diagonal) shows that the model is classifying most crops correctly. But some incorrect classifications can be seen, for example, mungbean with other crops. This means that

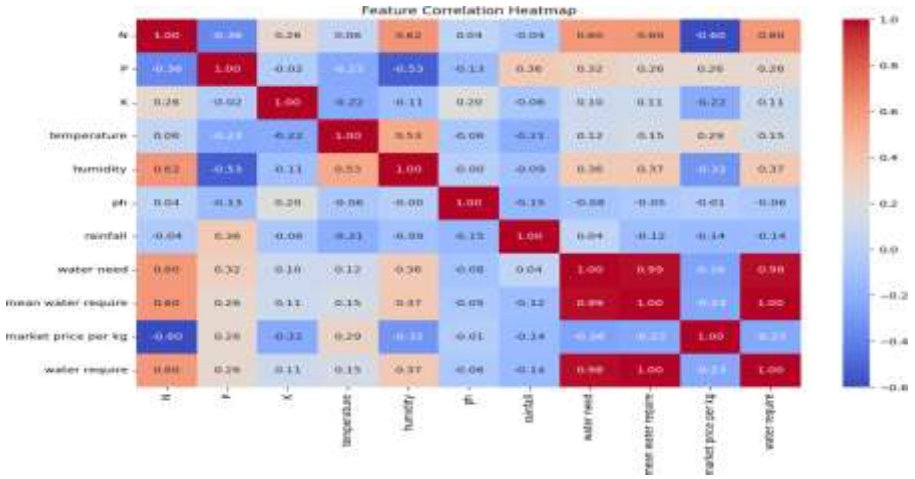

Fig. 2. Feature Correlation Heatmap

some crop classes have similar characteristics and are difficult to distinguish. The confusion matrix helps to find specific categories where the model performs poorly and therefore may have specific improvements, such as feature engineering or more training data.

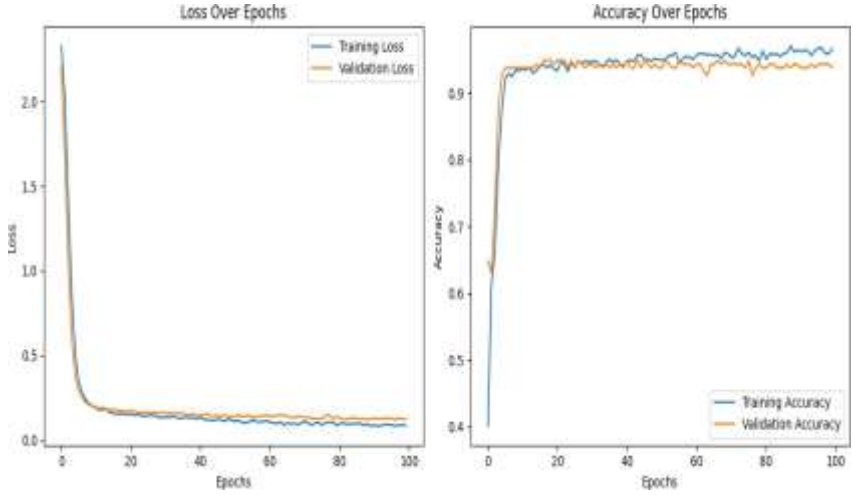

Fig. 4. Training and Validation Loss & Accuracy

CONCLUSION

In this paper, we propose a crop prediction model based on LSTM and GRU to examine environmental variables and suggest suitable crops. The proposed method effectively ad- dresses contemporary agricultural problems such as climate change, resource control, and infestation. The high accuracy and ability to generalize from different agricultural situations prove the effectiveness of the model as a decision support tool for farmers and policy makers. Future updates will in- clude real-time data streams, satellite imagery, and IoT-based monitoring to further enhance accuracy. The use of artificial intelligence in agriculture is a bright direction for sustainable agriculture, which will provide food security for the growing world population.

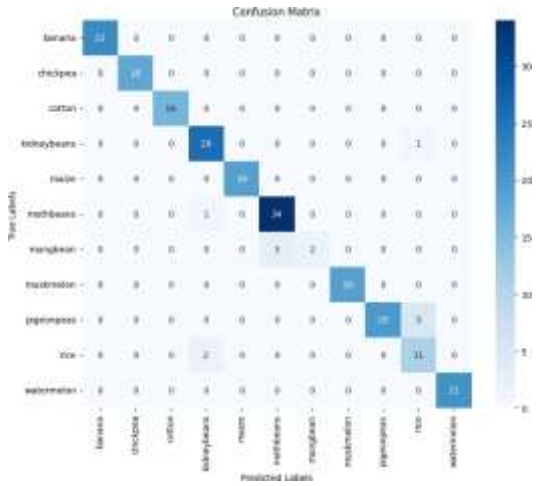

Fig. 3. Confusion Matrix

Training and Validation Loss & Accuracy

The loss and accuracy plots give us an idea of the learning process of the model over several epochs. In the loss plot, the training loss gradually decreases, which means that the model is reducing the error over time. The validation loss also tends to follow the same direction, which means that the model generalizes well to new data without overfitting.

In the accuracy plot, both the training and validation accura- cies increase strongly in the early epochs, then reach a plateau above 90%. This shows that the model learns well from the input features and can make good crop recommendations. The closeness of the training and validation accuracy curves to each other indicates that the model is not overfitting and has good generalization ability.

References

[1] M. Venkatanaresh and I. Kullayamma, "Deep learning based concurrent excited gated recurrent unit for crop recommendation based on soil and climatic conditions," *Multimedia Tools and Applications*, vol. 83, no. 24, pp. 64 109–64 138, 2024.

[2] S. Poonkuzhali, K. Akilesh, and R. Arya, "Development of efficient forecasting models for climate-resilient crop rotation based on deep learning techniques," in *International Conference on Smart Computing and Communication*. Springer, 2024, pp. 445–462.

[3] J. Padmapriya and T. Sasilatha, "Deep learning based multi-labelled soil classification and empirical estimation toward sustainable agriculture," *Engineering Applications of Artificial Intelligence*, vol. 119, p. 105690, 2023.

[4] X.-B. Jin, N.-X. Yang, X.-Y. Wang, Y.-T. Bai, T.-L. Su, and J.-L. Kong, "Hybrid deep learning predictor for smart agriculture sensing based on empirical mode decomposition and gated recurrent unit group model," *Sensors*, vol. 20, no. 5, p. 1334, 2020.

[5] E.-S. M. El-Kenawy, A. A. Alhussan, N. Khodadadi, S. Mirjalili, and M. M. Eid, "Predicting potato crop yield with machine learning and deep learning for sustainable agriculture," *Potato Research*, pp. 1–34, 2024.

[6] J. Padmapriya and T. Sasilatha, "Deep learning based multi-labelled soil classification and empirical estimation toward sustainable agriculture," *Engineering Applications of Artificial Intelligence*, vol. 119, p. 105690, 2023.

[7] M. Eftekhari, C. Ma, and Y. L. Orlov, "Applications of artificial intelligence, machine learning, and deep learning in plant breeding," p. 1420938, 2024.

[8] D. Elavarasan and P. D. Vincent, "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications," *IEEE access*, vol. 8, pp. 86 886–86 901, 2020.

[9] M. Mohamed, "Agricultural sustainability in the age of deep learning: current trends, challenges, and future trajectories," *Sustainable Machine Intelligence Journal*, vol. 4, pp. 2–1, 2023.

[10] K. Kumari and A. M. Nafchi, "Sustainable agriculture with ai, machine learning, deep learning, and iot for future farming," in *2024 ASABE Annual International Meeting*. American Society of Agricultural and Biological Engineers, 2024, p. 1.