

Deep Learning-Driven Precision Agriculture: Enhancing Crop Recommendation and Soil Analysis through IoT Sensor Data

M. Priya¹, A. Jaya^{2*}, C.D. NandaKumar³

¹Research Scholar, Department of Computer Applications, B. S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai 600 048, India. Orcid ID: 0009-0003-4055-944X

& ¹Assistant Professor, Department of Computer Science, Agurchand Manmull Jain College, Meenambakkam, Chennai 600 061, India.

^{2,*} Corresponding Author, Professor, Department of Computer Applications, B. S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai 600048, India. Orcid ID: 0000-0002-9070-4882

³Associate Professor, Department of Mathematics & Actuarial Science, B. S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai 600048, India. Orcid ID: 0000-0001-9261-1244

mpriya.cs19@gmail.com¹, jaya@crescent.education², prof.cdnandakumar@gmail.com³

ABSTRACT: This research introduces an innovative approach to precision agriculture, focusing on empowering farmers through advanced technologies. The study utilizes a fusion of cutting-edge deep learning algorithms, including CNNs, CapsNets, GRUs, LSTMs, and an AdaBoostClassifier integrated with GRU, to optimize crop recommendations and soil analysis. By analyzing a comprehensive dataset from IoT sensors that measure critical agricultural parameters like soil moisture, temperature, pH levels, and nutrient content, the research uncovers intricate relationships within these factors. The integration of spatial features, complex soil patterns, and capturing temporal dynamics achieves a remarkable 99% accuracy in crop recommendation. This integration of deep learning techniques with IoT sensor data offers farmers an advanced framework for decision-making based on soil conditions, enhancing agricultural productivity and promoting sustainable farming practices. The research represents a significant advancement in precision agriculture, bridging advanced science with traditional farming wisdom for the benefit of farmers and global agriculture.

Keywords: Precision Agriculture, Deep Learning Algorithms, Crop Recommendation, Soil Analysis, IoT Sensor Data.

1. INTRODUCTION

In the ever-evolving landscape of human civilization, agriculture has witnessed transformative changes. With a growing global population and escalating environmental concerns, the imperative for sustainable and efficient farming practices has never been more pressing. Precision agriculture emerges as a groundbreaking shift, harnessing state-of-the-art technologies to optimize resource usage, boost productivity, and minimize environmental impact. This approach involves the meticulous application of inputs like water, fertilizers, and pesticides, guided by real-time data and insights. Precision agriculture relies on cutting-edge technologies such as satellite imagery, GPS guidance

systems, drones, and IoT sensors. The integration of these tools empowers farmers to monitor and manage their fields with unprecedented precision, addressing specific needs at a localized level. A fundamental aspect of precision agriculture is data-driven decision-making. By collecting data from diverse sources, including soil sensors, weather stations, and crop monitoring systems, farmers acquire valuable insights into their fields' conditions. This data-centric approach facilitates personalized and timely interventions, such as targeted irrigation, nutrient application, and pest control.

In response, this paper presents an innovative approach to optimize crop recommendation and soil analysis through the synergistic application of state-of-the-art deep learning algorithms and advanced Internet of Things (IoT) sensor data. Leveraging the power of Convolutional Neural Networks (CNNs), Capsule Networks (CapsNets), Gated Recurrent Units (GRUs), Long Short-Term Memory (LSTM) networks, and an AdaBoostClassifier integrated with GRU, we delve into the intricate relationships among critical agricultural parameters measured by IoT sensors. These parameters include soil moisture, temperature, pH levels, and nutrient content. This research aims to harness the potential of these cutting-edge deep learning architectures in the realm of precision agriculture. The CNN extracts spatial features from the collected data, CapsNets unravel complex soil patterns, while GRUs and LSTMs capture temporal dynamics and sequential dependencies within the dataset. The proposed AdaBoostClassifier, seamlessly integrated with GRU, exhibits a remarkable 99% accuracy in crop recommendation, underscoring its efficacy in leveraging diverse data sources for optimal decision-making in agriculture.

Table.1. Challenges in Cultivation

Challenges	Description
Climate Uncertainties	Unpredictable weather, extreme temperatures, and irregular rainfall.
Water Scarcity	Limited access to reliable water for irrigation, posing challenges.
Soil Degradation	Erosion, nutrient depletion, and soil health degradation impact crops.
Pest and Disease Outbreaks	Diseases and pests leading to significant yield losses.
Access to Technology	Limited access to modern agricultural technologies.
Education and Training	Limited access hindering adoption of modern farming practices.
Environmental Concerns	Balancing productivity with environmental conservation.

By employing an extensive dataset obtained from IoT sensors, our study explores a comprehensive understanding of the intricate interplay between various agricultural parameters. This not only advances our knowledge of soil conditions but also empowers farmers with precise crop suggestions, fostering enhanced agricultural productivity and sustainability.

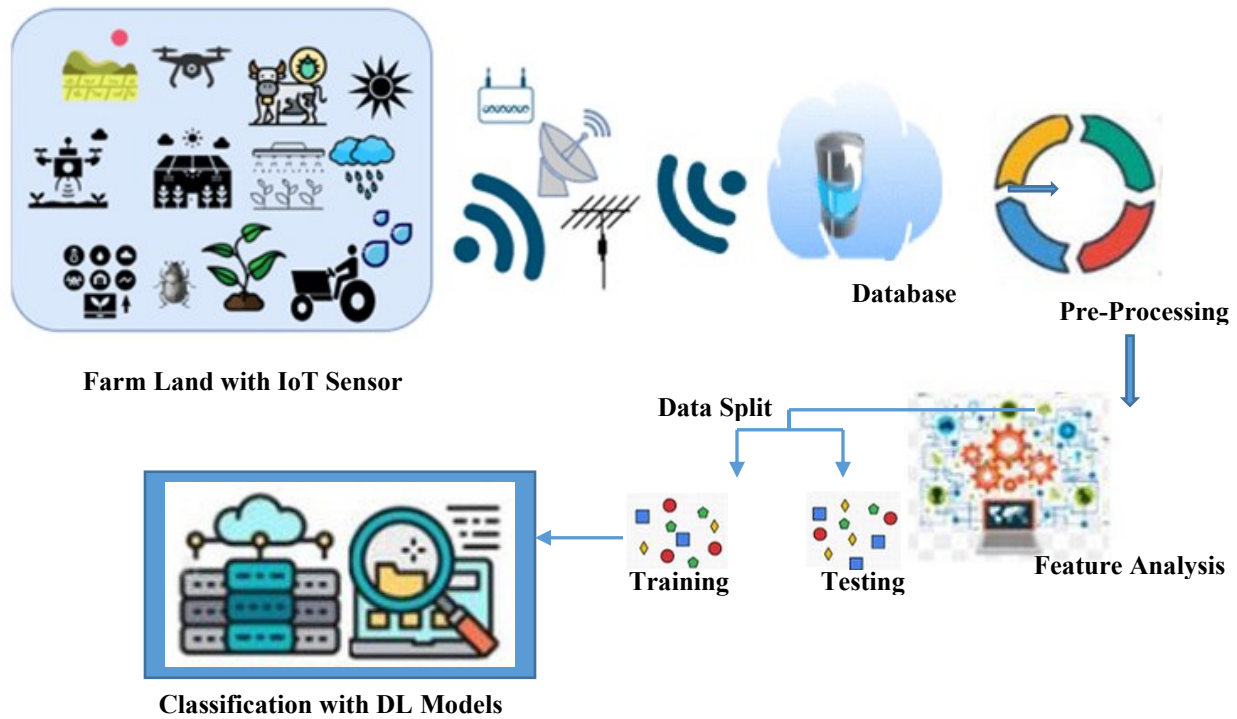


Figure 1. System architecture

Figure 1 illustrates the system architecture for an advanced Precision Agriculture system. IoT sensors deployed in farmland collect agricultural data, which is stored in a database. After preprocessing, subsequent feature analysis is performed and the dataset is divided into training and testing sets for classification using Deep Learning models. This process optimizes crop recommendations and soil analysis based on the advanced sensor data.

Deep learning is selected for its adeptness in deciphering intricate patterns within extensive datasets, enabling a nuanced understanding of relationships among agricultural parameters. Employing sophisticated algorithms, it enhances precision in predicting crop recommendations and conducting soil analyses for optimal farming decisions. Its inherent capability to unveil subtle patterns from diverse data sources positions deep learning as a powerful tool, markedly improving the effectiveness of predictive models in the realm of precision agriculture.

In the following sections, we delve into the methodology, experimentation, and results, providing valuable insights into the potential of deep learning and IoT integration for the future of precision agriculture. The findings presented herein aim to contribute significantly to the ongoing discourse on sustainable farming practices and data-driven decision-making in the agricultural domain.

2. LITERATURE REVIEW

Precision agriculture has seen substantial advancements through the integration of deep learning algorithms for soil classification and crop suggestion. Kaur and Malik (2021)

emphasized the significance of nitrogen (N), phosphorus (P), and potassium (K) as primary nutrients in soil, proposing a model utilizing machine learning techniques for estimating these nutrients' inter-dependency [1]. Sun, X et al., (2021) addressed soil classification by employing extreme learning machines and distinct activation functions, leading to enhanced efficiency in fertilizer usage and environmental preservation [2].

Madhumathi, R. (2020) introduced a hybrid model combining support vector regression and the fruit fly optimization algorithm, demonstrating improved performance in predicting ski-jump spillway scour geometry. Madhumathi, Arumuganathan, and Shruthi (2020) proposed a system integrating sensors, an Arduino board, and cloud-based data analysis for soil NPK and moisture analysis, offering recommendations for fertilizer needs based on real-time soil data.

Additionally, Schwalbert et al. (2020) highlighted the importance of early crop prediction, focusing on soybean yield forecasting in Brazilian fields. They compared models like long-short-term memory neural networks, random forest, and multivariate regression, favoring the long-short-term memory neural network for better performance in crop yield prediction [4]. Bhojani and Bhatt (2020) studied wheat crop yield prediction, leveraging time-series weather data and various activation functions in neural networks, proposing the use of new activation functions for improved accuracy in wheat yield prediction [5].

Parameswari and Tharani's (2023) research focuses on predicting crop yields crucial for agricultural economics. Their hybrid model, combining decision tree, SVM, and RNN algorithms, aims to aid farmers in anticipating harvests using soil elements like nitrogen, phosphorus, and potassium. The study's analysis provides valuable insights for future crop yield trends, showcasing the potential of advanced machine learning in enhancing agricultural decision-making [6].

Motwani et al. (2022) propose a crop recommendation system utilizing Convolutional Neural Network (CNN) and Random Forest Model. Their study achieves 95.21% accuracy with CNN and 75% accuracy with the Random Forest Algorithm. This research demonstrates the potential of machine learning in assisting farmers to select optimal crops based on diverse agricultural parameters [7]. Dolli et al. (2023) introduce a comprehensive crop recommendation system utilizing Decision Tree, Naive Bayes, KNN, Random Forest, and XG-Boost algorithms to analyze soil data. Their study demonstrates that machine learning algorithms can effectively predict agricultural yields and recommend optimal crop management practices [8].

Table.1. Literature Review

Authors (Year)	Research Focus	Data Collection	Methods/Algorithms	Key Findings
Pandith, V et al. (2020) [9]	Soil Nutrient Analysis & Mustard Crop Yield Prediction	Agriculture Departments of Jammu, Talab, and Tillo	KNN, ANN, Naive Bayes, RF, Multinomial Logistic Regression	Predicted mustard crop yield via soil analysis using various machine learning methods.
Mupangwa, W et al. (2020) [10]	Maize Yield Prediction	Field data from different	Machine Learning Methods	Applied machine learning to predict

		countries in Eastern and Southern Africa		maize yield in Eastern and Southern Africa.
Priyadharshini, A et al. (2021) [11]	Crop Recommendation System	Rainfall, Temperature, Geolocation	Classification and Regression Algorithms	Used weather data to recommend suitable crops through machine learning algorithms.
Shidnal, S. et al. (2021) [12]	Paddy Crop Nutrient Deficiency Prediction	Paddy crop images	CNN, K-means Clustering	Predicted nutrient deficiency in paddy crops using image analysis and clustering.
Reshma, S. J. et al. (2022) [13]	Soil and Crop Classification	Soil parameters from Kanyakumari, Tirunelveli, and Thoothiukudi districts	SVM, DT, MLP	Classified soil and crops based on NPK values in Tamil Nadu.
Shi, P. et al. (2021) [14]	Nitrogen Estimation for Rice Crop	RGB images of rice canopy	Regression Algorithms (SNR, BPNN, RF)	Estimated nitrogen levels in rice crops using RGB image analysis.
Dharani, M. K. et al. (2021) [15]	Crop Yield Prediction	Not specified	Deep Learning (ANN, DNN, RNN)	Predicted crop yield using various deep learning algorithms.
Jose et al. (2021) [16]	Mineral Deficiency Detection in Tomato Leaves	Tomato leaves examination	ANN	Used ANN for classifying tomato leaves and detecting mineral deficiencies.
Elumalai et al. (2021) [17]	Soil Classification Framework	Not specified	onto_mine Framework	Designed a model for consistent soil classification support to novice farmers.
R. Reshma et al. (2020) [18]	Soil Behavior Analysis and Crop Recommendation	IoT Sensors, Cloud Storage	SVM, Decision Tree	Analyzed soil characteristics and recommended crops using SVM and Decision Tree.
Suchithra, M. S. et al. (2020) [19]	Soil Parameter Classification and Prediction Workflow	Kerala State soil classification	Extreme Machine Learning Algorithm	Classified soil parameters in Kerala state for controlled fertilizer expenditure.
Bhojani, S. H. et al. (2020)	Crop Yield Forecasting	Datasets from three years	MLP Algorithm with Many Kernel Functions	Implemented MLP for forecasting

[20]		(1990–1991, 2015–2016, 2016–2017)		crop yield using multiple datasets.
Sun et al. (2021) [21]	Hybrid Model for Power Load Forecasting	Not specified	Generalized Regression Neural Networks	Suggested a hybrid model for power load forecasting with improved precision.

Research Gap: The existing studies in precision agriculture primarily focus on individual aspects such as soil nutrient analysis, crop yield prediction, and nutrient deficiency detection. However, there is a noticeable gap in research that comprehensively integrates cutting-edge deep learning algorithms for both crop recommendation and soil type analysis. Most studies emphasize either crop-related factors or soil characteristics, but there is a lack of holistic approaches that simultaneously optimize both aspects.

Solution: This research proposes a novel solution by integrating advanced deep learning algorithms such as CNNs, CapsNets, GRUs, LSTMs, and an AdaBoostClassifier with GRU. Leveraging IoT sensor data measuring critical agricultural parameters, the model achieves a 99% accuracy in crop recommendation. This comprehensive framework addresses both crop and soil aspects simultaneously, empowering farmers with accurate insights for enhanced productivity and sustainability.

3. CONCEPTUAL FRAMEWORK

3.1. Dataset Description

The dataset used for crop recommendation is available on Kaggle at Crop Recommendation Dataset. This dataset focuses on variables influencing rainfall and crop yield in specific regions. Key variables include phosphorus (P), nitrogen (N), potassium (K), temperature, humidity, pH and rainfall [5].

Table.2. Feature Description

Variable	Description
Phosphorus (P)	Soil phosphorus content
Nitrogen (N)	Soil nitrogen content
Potassium (K)	Soil potassium content
Temperature	Ambient temperature in the region
Humidity	Atmospheric humidity levels
pH	Soil acidity/alkalinity (pH) levels
Rainfall	Amount of rainfall in the region

N	P	K	Temperature	Humidity	pH	Rainfall
90	42	43	20.879744	82.002744	6.502985	202.935536
85	58	41	21.770462	80.319644	7.038096	226.655537

60	55	44	23.004459	82.320763	7.840207	263.964248
74	35	40	26.491096	80.158363	6.980401	242.864034
78	42	42	20.130175	81.604873	7.628473	262.717340

Figure 2. Sample Data

This dataset facilitates the investigation of connections between soil nutrients, climatic conditions, and recommended crops, offering valuable insights for precision agriculture.

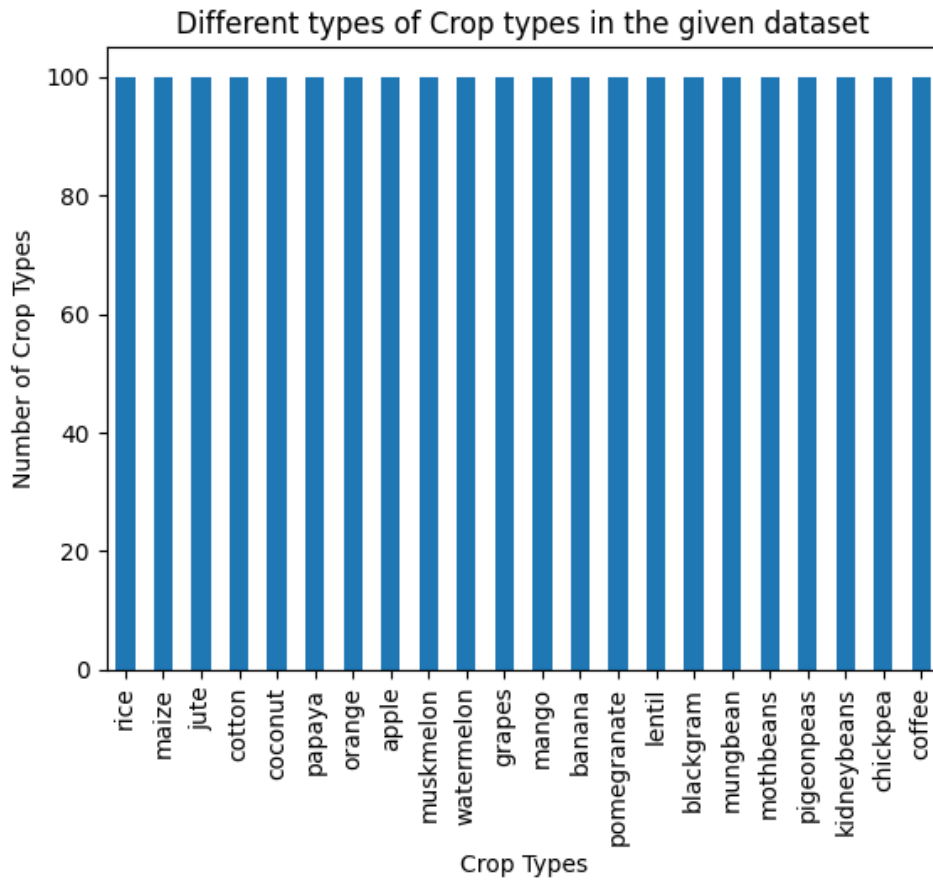
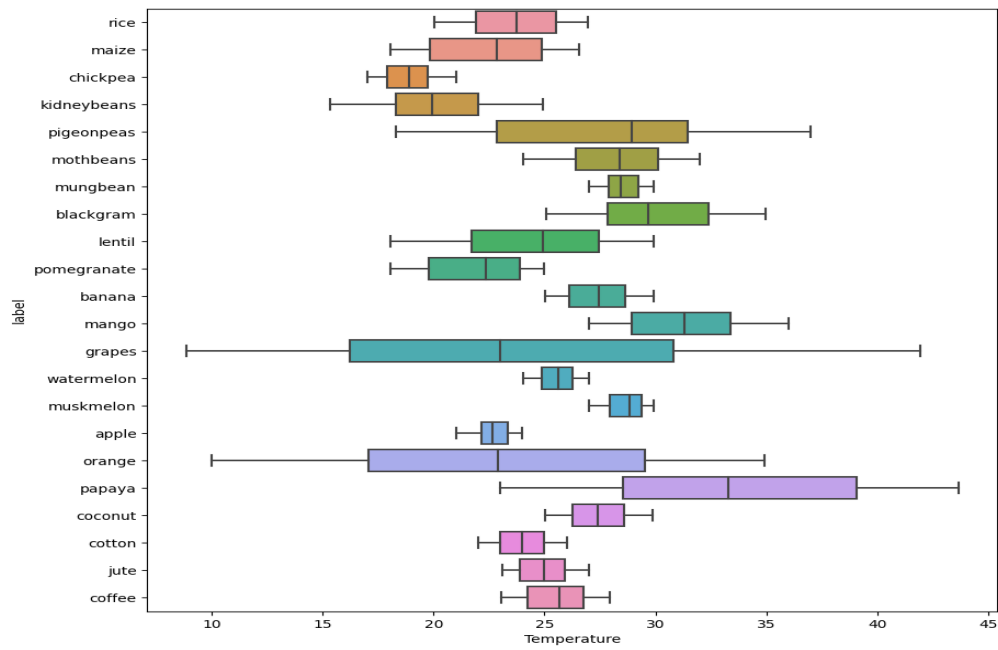
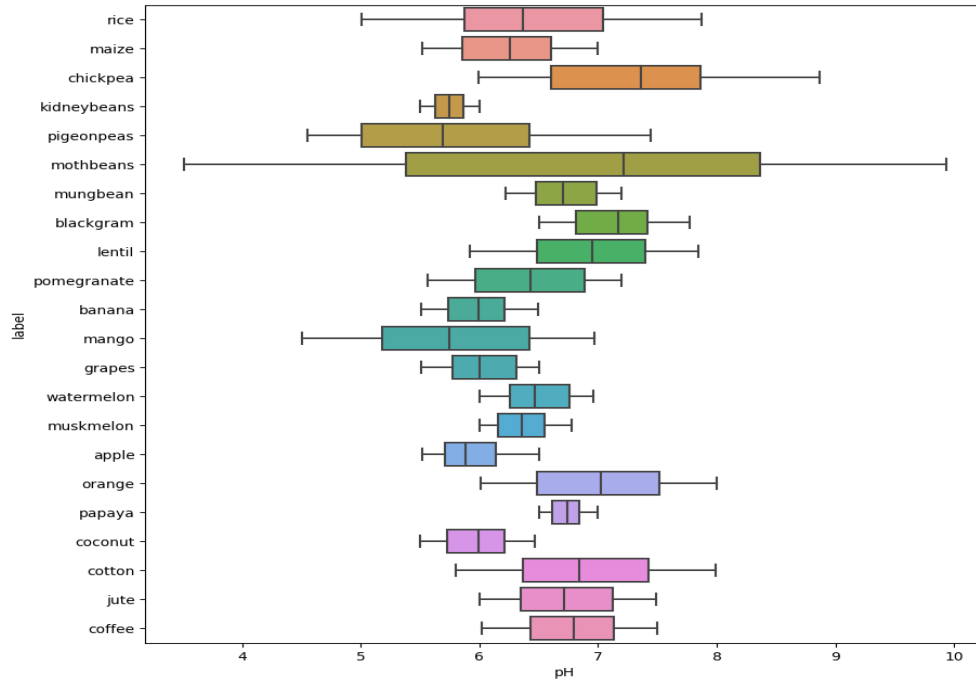


Figure 3. Different Crops

The bar chart illustrates the distribution of different crop types in the dataset, each having an equal representation of 100 units.



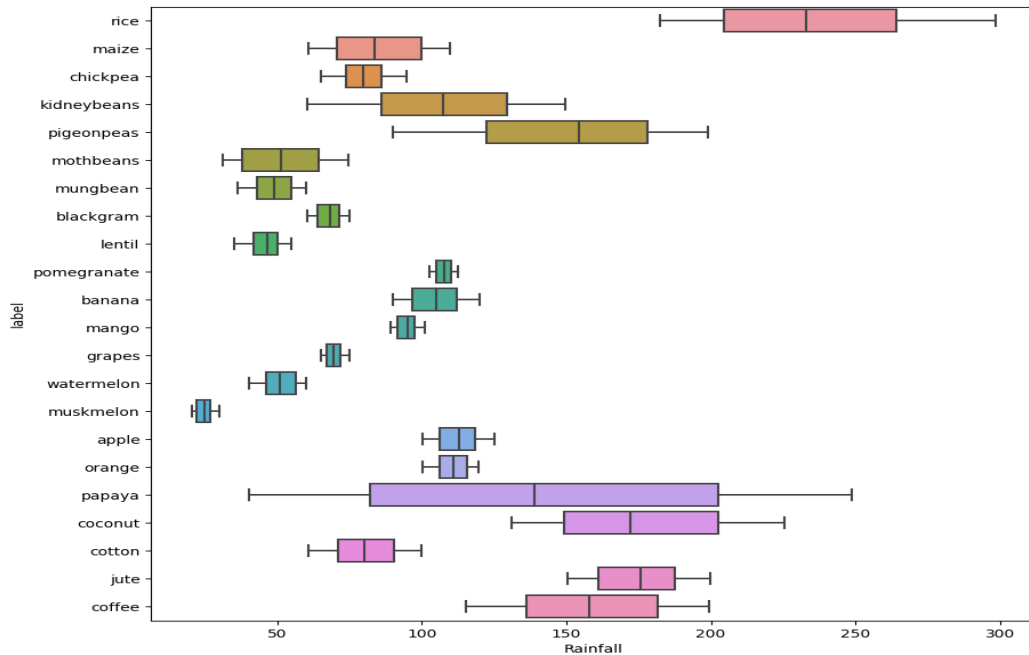


Figure 4. Environmental Factors Variation across Crop Types

The above boxplots visually represent the distribution of pH, Temperature, and Rainfall across different crop types in the dataset. Each boxplot provides insights into the central tendency and variability of these environmental factors for various crops.

3.2. Pre-processing

In the preprocessing phase for the Crop recommendation dataset, handling missing values is crucial. Employing mean imputation, where Missing Values (MV) for a specific feature (F) are replaced by the mean (\hat{F}) of the non-missing values, ensures data completeness [12]:

$$F_{imputed} = \frac{F_{non-missing}}{Number\ of\ Non-Missing\ Values} \tag{1}$$

Additionally, normalization techniques enhance feature consistency. Standardization transforms each feature value (X) to a standardized value (X_{std}) using the mean (μ) and standard deviation (σ) of the feature:

$$X_{std} = \frac{X - \mu}{\sigma} \tag{2}$$

Min-Max scaling transforms X to a scaled value (X_{scaled}) within a specified range:

$$X_{scaled} = \frac{X - \min(X)}{\max(X) - \min(X)} \tag{3}$$

These preprocessing steps collectively enhance the dataset's integrity and feature comparability, facilitating robust analysis for crop recommendation.

3.3. Feature analysis

Feature analysis is a crucial step in data exploration, focusing on understanding the characteristics and relationships among different features or variables within a dataset. It aims to uncover patterns, trends, and correlations, providing valuable insights into the underlying dynamics of the data. In the context of the Crop_recommendation dataset, feature analysis specifically explores the connections between soil pH and nutrient levels (N, P, K), enhancing the understanding of how these factors influence agricultural outcomes [16].

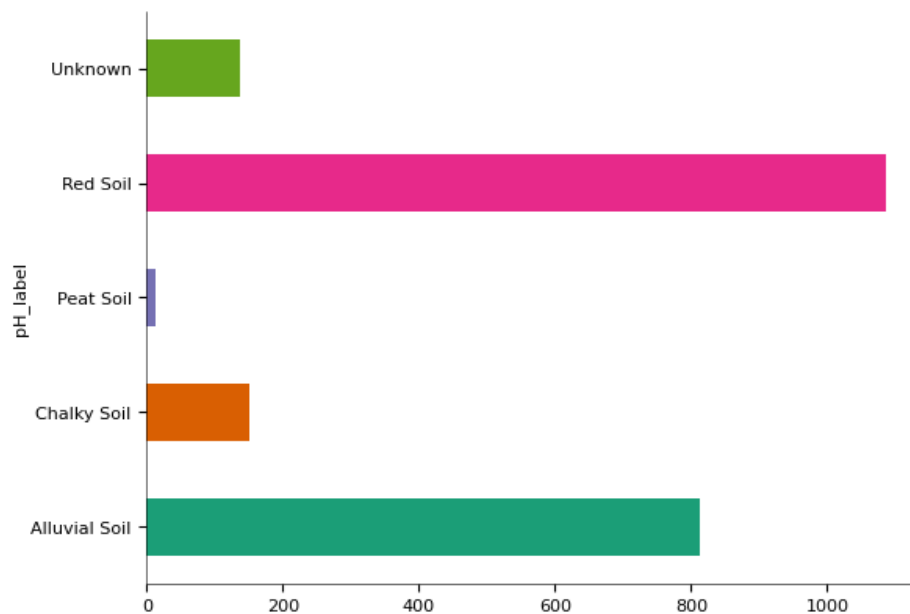


Figure 5. Soil Types

In Figure 5, the bar chart visually represents the distribution of soil types along with their respective quantities. The data reveals that "Unknown" soil dominates, followed by "Red Soil", "Peat Soil", "Chalky Soil" and "Alluvial Soil". The bar chart provides a visual representation of the relative proportions of these different soil types in the dataset.

In the feature analysis of the Crop_recommendation dataset, the relationship between soil pH and nutrient levels (N, P, K) is investigated. The pH_label column categorizes pH values into distinct soil types, including Peat Soil, Red Soil, Alluvial Soil, Chalky Soil, Neutral, and Loam Soil. This classification facilitates the understanding of soil characteristics associated with each pH range.

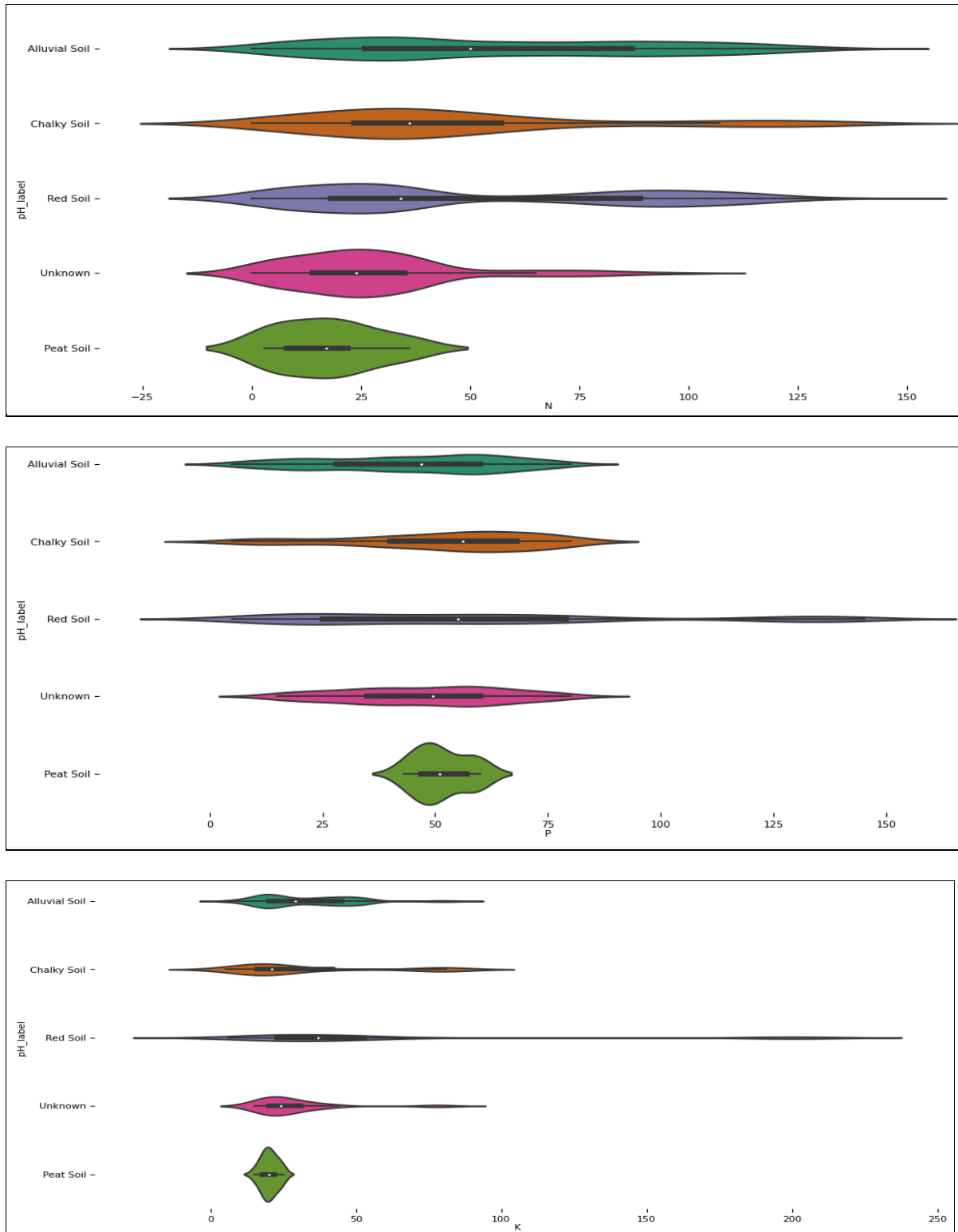


Figure 6. Exploring Categorical Distributions of pH across Nutrient Levels (N, P, K)

Figure 6 illustrates the categorical distributions of soil pH levels across varying nutrient concentrations (N, P, K). The boxplots provide insights into how different soil types, categorized by pH, correlate with nutrient levels in the Crop_recommendation dataset.

3.4. Classification

The classification methodology employed in this study integrates advanced deep learning algorithms, including Convolutional Neural Networks (CNNs), Capsule Networks (CapsNets), Gated Recurrent Units (GRUs), and Long Short-Term Memory (LSTM) networks, along with an innovative AdaBoostClassifier integrated with GRU. These algorithms are applied to a comprehensive dataset collected from IoT sensors, capturing crucial agricultural parameters such as soil moisture, temperature, pH levels, and nutrient content [15].

i. Convolutional Neural Networks (CNNs):

CNNs initiate the process by employing convolutional layers to apply filters to input features, including N, P, K, temperature, humidity, pH, and rainfall, thereby detecting spatial patterns and learning relevant features from the data.

$$c(x, y) = \sum_{i=0}^m \sum_{j=0}^n I(x+i, y+j) \cdot K(i, j) + b \quad (4)$$

where $C(x,y)$ is the output of the convolutional layer at position (x,y) , I is the input, K is the filter (kernel), and b is the bias. Following convolution, the model introduces non-linearity through an activation function like ReLU, which is applied element-wise to the output of the convolutional layer.

$$ReLU(x) = \max(0, x) \quad (5)$$

Max pooling is employed to downsample spatial dimensions, preserving crucial features and simultaneously reducing computational complexity. The pooled output is flattened into a one-dimensional vector, preparing it for the fully connected layers.

$$\text{Flatten}(\text{pooled}_{\text{output}}) = \text{pooled}_{\text{output}}.\text{reshape}(-1) \quad (6)$$

Fully connected layers establish connections between each neuron and the flattened input, facilitating the learning of intricate relationships through weighted sums, biases, and activation functions.

$$Z^{[l]} = W^{[l]} \cdot A^{[l-1]} + b^{[l]} \quad \text{and} \quad A^{[l]} = ReLU(AZ^{[l]}) \quad (7)$$

Where $W^{[l]}$ is the weight matrix, $A^{[l-1]}$ is the input from the previous layer, $b^{[l]}$ is the bias, and $A^{[l]}$ is the output after activation. The final layer utilizes the learned features to generate predictions.

$$Softmax(x)_i = \frac{e^{x_i}}{\sum_j e^{x_j}} \tag{8}$$

This algorithm enables CNNs to automatically identify and extract relevant features from agricultural data, facilitating precise crop recommendations based on nutrient levels, climatic conditions, and other parameters [16].

ii. Capsule Network (CapsNet):

The CapsNet architecture comprises three layers such as the convolutional layer, PrimaryCaps layer, and DigitCaps layer. The convolution layer uses 256 kernels (9×9 size, stride 1, ReLU activation), feeding into the PrimaryCaps layer. With 32×8 convolution kernels (9×9 size, stride 2), the PrimaryCaps layer generates 1152 capsules (dimension 8). These capsules serve as low-level inputs for the DigitCaps layer, which, through dynamic routing, produces 10 capsules (dimension 16) as the final classification result. The correctly predicted capsule undergoes reconstruction via a three-layer fully connected neural network. Dynamic routing is crucial, maintaining spatial information through vector operations. The process involves multiplying input capsules by a weight matrix to obtain predicted capsules, mathematically expressed as:

$$\hat{u}_j | i = W_{ij} \cdot u_i \tag{9}$$

The weighted summation of predicted capsules and coupling coefficients yields deep feature capsules. These capsules undergo a nonlinear squeezing activation function, generating the output capsules:

$$v_j = Squash(c_{ij} \hat{u}_j | i) \text{ where } Squash(.) = \frac{||.||^2}{1+||.||^2} \tag{10}$$

The iterative dynamic routing process updates coupling coefficients based on the correlation between output capsules and predicted capsules:

$$c_{ij} = softmax(b_{ij}) \text{ where } b_{ij} = b_{ij} + u_{ij} \cdot v_j \tag{11}$$

This dynamic routing refines coupling coefficients and determines the final output capsule, ensuring CapsNet retains spatial information throughout its layers [17].

iii. Gated Recurrent Units (GRUs):

GRUs are a type of recurrent neural network (RNN) designed to capture and utilize sequential information effectively. The algorithm takes sequential input data (such as N, P, K, temperature, humidity, pH, and rainfall) for each time step. GRUs utilize update and reset gates to control the flow of information through the network. These gates are sigmoid-activated, determining which information to discard or update.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad \text{where } r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (12)$$

Where z_t is the update gate output, r_t is the reset gate output, h_{t-1} is the previous hidden state, x_t is the input at time t , and σ is the sigmoid activation function. The hidden state is updated based on the update and reset gates.

$$h_t \sim = \tanh(W \cdot [h_{t-1}, x_t]) \quad \text{where } h_t = (1 - z_t) \odot h_{t-1} + z_t \odot h_t \sim \quad (13)$$

where \tilde{h}_t is the new proposed state, \odot denotes element-wise multiplication, and W is the weight matrix. The final output is obtained based on the hidden state, which captures sequential dependencies.

$$\text{Output}_t = \text{SomeFunction}(h_t) \quad (14)$$

These steps allow GRUs to effectively model sequential patterns in the input data, making them suitable for time-series and sequential data such as crop-related environmental factors [18].

iv. Long Short-Term Memory (LSTM)

The LSTM architecture serves as an enhanced solution for traditional Recurrent Neural Networks (RNNs) in the context of crop recommendation systems. In this framework, LSTM addresses challenges related to vanishing and exploding gradients by implementing memory blocks with adaptive multiplicative gates. Unlike conventional RNN units, these memory blocks incorporate input, output, and forget gates, alongside a self-recurrent connection. The memory cell within the block acts as a repository for storing past information across different time steps, crucial for capturing long-term temporal dependencies in the input features, including N, P, K, temperature, humidity, pH, and rainfall.

During the training process, LSTM leverages Backpropagation Through Time (BPTT) to continuously update the values of three multiplicative units—input (i), output (y), and forget gate (f) on a memory cell (m). This iteration occurs from the initial time step ($t = 1$) to the final time step (T). The LSTM recurrent hidden layer function is expressed through a set of equations:

$$\left. \begin{aligned} i_t &= \sigma(wx_i x_t + wh_i h_{t-1} + wc_i m_{t-1} + b_i) \\ f_t &= \sigma(wx_f x_t + wh_f h_{t-1} + wm_f m_{t-1} + b_f) \\ m_t &= f_t \odot m_{t-1} + i_t \odot \tanh(wx_m x_t + wh_m h_{t-1} + b_m) \\ y_t &= \sigma(wx_y x_t + wh_y h_{t-1} + wm_y m_t + b_y) \\ h_t &= y_t \odot \tanh(m_t) \end{aligned} \right\} \quad (15)$$

Here, i_t , f_t , and y_t represent the input, forget, and output gates, respectively. The bias units b_i , b_f , b_m , and b_y are associated with the input gate, forget gate, memory cell, and output gate, respectively. The weight matrices are denoted by w , m signifies the state of the memory cell, h represents the output of the hidden layer, and σ and \tanh denote the sigmoid and hyperbolic tangent activation functions. This LSTM architecture effectively processes sequential data, estimating output sequences while accounting for long-term dependencies in the specified crop recommendation features [19].

v. Proposed: AdaBoost Classifier - GRUs

The integration of AdaBoostClassifier with GRUs in the crop recommendation system optimizes predictions for N, P, K, temperature, humidity, pH, and rainfall. AdaBoost, functioning as a boosting algorithm, collaborates with GRUs to exploit their temporal dynamics and sequential dependency learning capabilities. This synergy aims to enhance crop recommendation accuracy by combining ensemble learning with recurrent neural network features. AdaBoost, renowned for its ensemble technique, merges weak learner predictions, often decision trees, to form a robust model. In the proposed system, GRUs excel in sequential data learning, especially in time-series analysis, making them valuable for capturing temporal patterns

Weighted Error of GRU (Weak Learner): The weighted error ϵ_t of the t -th GRU is computed as the sum of weight s of misclassified samples divided by the total weight:

$$\epsilon_t = \frac{\sum_{i=1}^N w_{i,t} \cdot 1(h_t(x_i) \neq y_i)}{\sum_{i=1}^N w_{i,t}} \quad (16)$$

where:

- N is the number of training samples.
- $w_{i,t}$ is the weight of the i -th sample for the t -th GRU.
- $h_t(x_i)$ is the prediction of the t -th GRU for the i -th sample.
- y_i is the true label of the i -th sample.
- $1(\cdot)$ is the indicator function.

GRU Weight in Ensemble: The weight α_t assigned to the t^{th} GRU based on its weighted error is computed as:

$$\alpha_t = \frac{1}{2} \ln \frac{1 - \epsilon_t}{\epsilon_t} \quad (17)$$

Update Sample Weights: The weights of the training samples are updated based on whether they were correctly or incorrectly classified by the t^{th} GRU:

$$W_{i,t+1} = W_{i,t} \cdot \exp(-\alpha_t \cdot h_t(x_i) \cdot y_i) \quad (18)$$

This update increases the weights of misclassified samples, making them more likely to be selected by subsequent weak learners.

Final Ensemble Prediction: The final prediction of the ensemble for a given sample is obtained by combining the individual predictions of all GRUs:

$$F(x) = \text{sign} \sum_{t=1}^T \alpha_t \cdot h_t(x) \quad (19)$$

Where $\text{sign}(\cdot)$ is the sign function. These equations encapsulate the AdaBoost with GRU integration, emphasizing how the ensemble adapts its weights based on the performance of each GRU in predicting the crop recommendation.

Proposed Algorithm

The proposed algorithm is defined as follow.

Input:

- Training data: $\{(x_1, y_1), (x_2, y_2) \dots (x_N, y_N)\}$

Where x_i is the feature vector (N, P, K, temperature, humidity, pH, rainfall) and y_i is the true label

- Number of GRU weak learners: T

Initialize:

- Set initial weights $w_{i,1} = 1/N$ for all $i = 1, \dots, N$

For $t = 1$ to T do:

1. Train GRU h_t on the weighted training data $\{(x_i, y_i, w_{i,t})\}$

2. Predict:

For each sample i , obtain prediction $h_t(x_i)$

3. Compute weighted error ϵ_t :

$$\epsilon_t = (\text{sum of weights of misclassified samples}) / (\text{sum of all weights})$$

$$\epsilon_t = (\sum w_{i,t} * 1(h_t(x_i) \neq y_i)) / (\sum w_{i,t})$$

// calculates the proportion of the total weight assigned to misclassified samples

4. Compute GRU weight α_t :

$$\alpha_t = 0.5 * \ln((1 - \epsilon_t) / \epsilon_t)$$

//This determines how much influence the GRU's prediction has on the final ensemble

```

5. Update sample weights  $w_i, t+1$ :
  For each sample  $i$ ,
  if  $h_t(x_i) = y_i$  then
     $w_{i,t+1} = w_{i,t} * \exp(-\alpha)$ 
  else
     $w_{i,t+1} = w_{i,t} * \exp(\alpha)$ 
6. Normalize weights:
  For each sample  $i$ ,
   $w_{i,t+1} = w_{i,t+1} / (\sum w_{i,t+1})$ 
Final Ensemble Prediction:
For a given sample  $x$ :
 $H(x) = \text{sign}(\sum \alpha_t * h_t(x) \text{ for } t = 1 \text{ to } T)$ 

```

4. Result and discussion

This section presents the outcomes and analyses derived from the crop recommendation model. The model selection process involved several algorithms, namely CNNs, CapsNets, GRUs, LSTM networks, and an innovative AdaBoostClassifier integrated with GRU. The implementation was conducted in Python, and performance evaluation metrics were scrutinized to assess the effectiveness of the algorithms. This section offers insights into the model's performance and its implications for the crop recommendation system. The prediction experiments were executed using Python 3.8 on a system equipped with an i5 Processor and 4 GB RAM, ensuring the efficient processing of all necessary tasks.

4.1. Unveiling Crop-Type Relationships

Figure 6 illustrates the relationships between variables in the dataset through a pair plot. This visual representation employs scatter plots and Kernel Density Estimation (KDE) plots, providing a comprehensive view of the interdependencies among different attributes. The primary objective is to discern patterns and distinctions among the various "crop-type" classes present in the data. Each scatter plot visually depicts the distribution of data points, while the accompanying KDE plots offer insights into the data distribution for individual attributes. This visualization serves as a valuable tool for data exploration, aiding in the identification of correlations and patterns within the dataset.

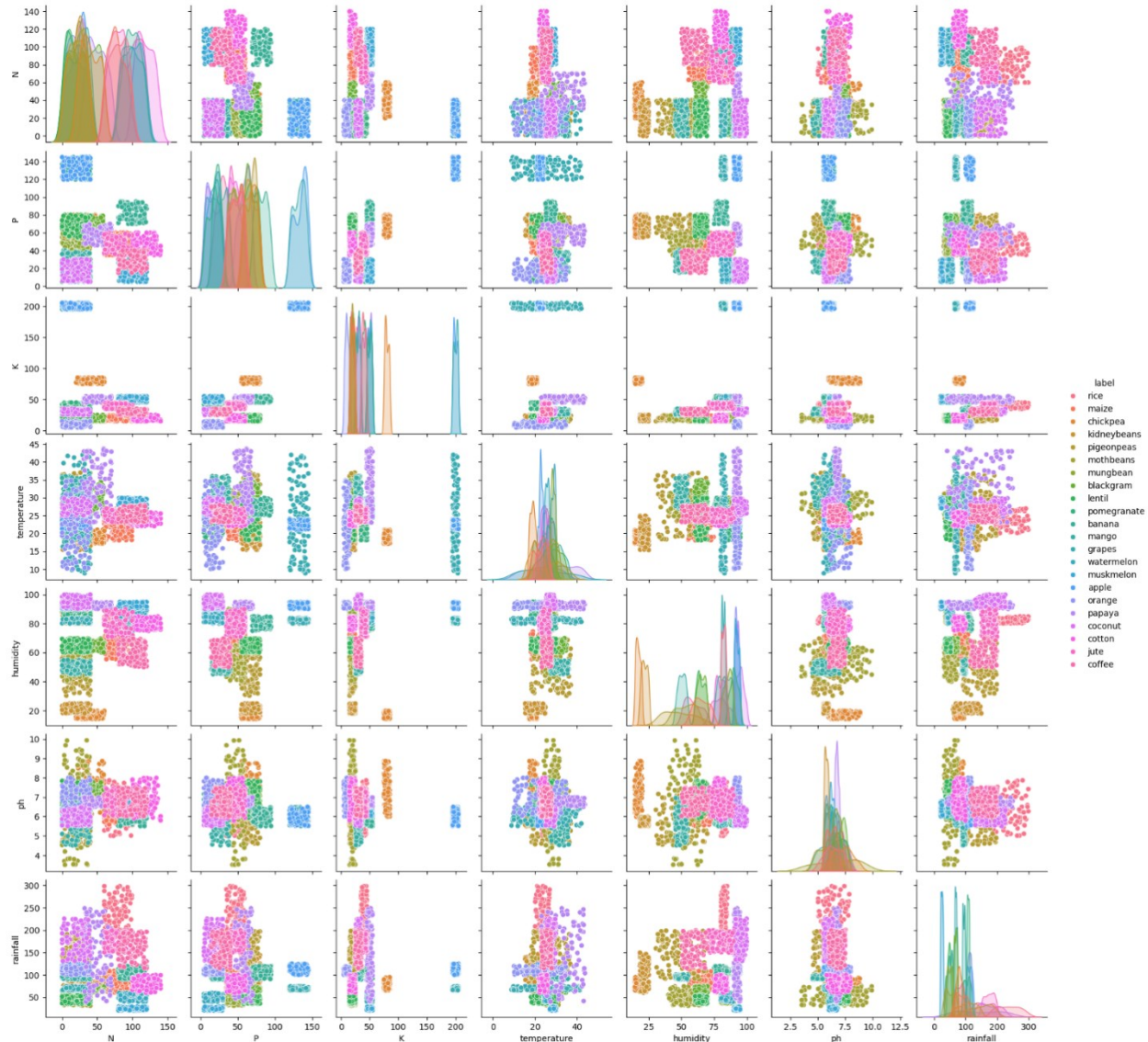


Figure.6. Exploring Features: Insights from a Pair Plot Analysis on Crop-Type Classes

4.2. Performance Analysis

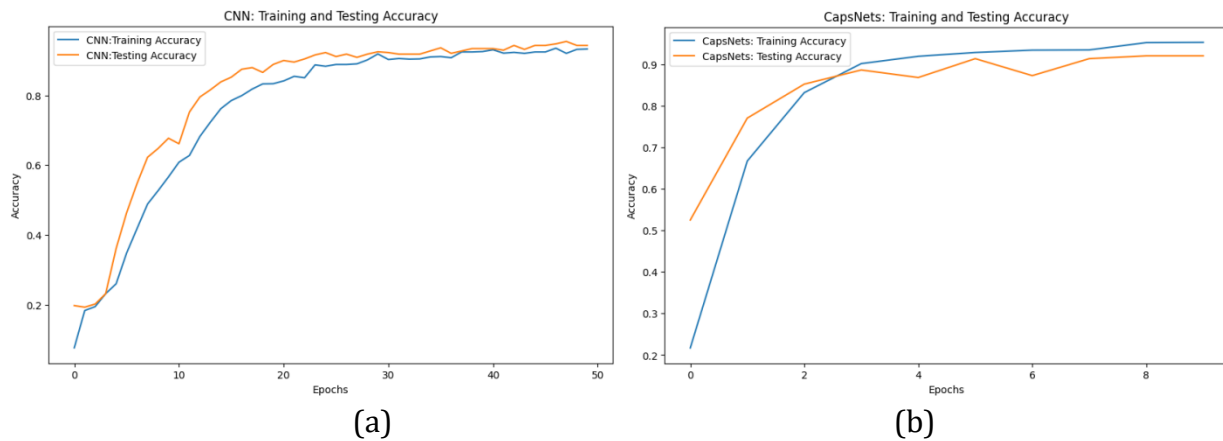
The performance analysis encompasses advanced algorithms, including CNNs, CapsNets, GRUs, LSTM networks, and an innovative AdaBoostClassifier integrated with GRU (AdaGRUBoost). Through a rigorous evaluation using diverse performance metrics, the study gauges the effectiveness of these models in the crop recommendation system. CNNs are assessed for spatial feature extraction, CapsNets for handling hierarchical relationships, and GRUs/LSTMs for capturing sequential dependencies. The hybrid AdaGRUBoost introduces a novel approach, integrating boosting with temporal understanding. This in-depth analysis informs algorithmic choices, crucially influencing the overall robustness and efficiency of the crop recommendation model. The below metrics provide a comprehensive evaluation of a classification model's performance, capturing aspects like overall

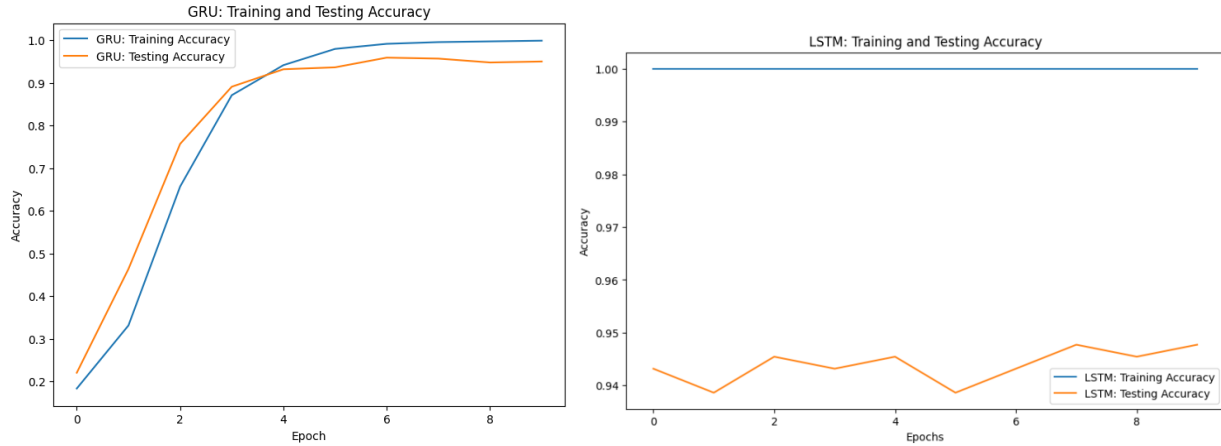
correctness, precision, recall, and the trade-off between precision and recall (F1 score) [19].

Table.3. Performance Metrics

Metric	Equation
Accuracy (ACC): Accuracy measures the overall correctness of predictions.	$\frac{\text{True Positives}}{\text{True Negatives} + \text{Total Prediction}}$
Precision: Precision quantifies the accuracy of positive predictions.	$\frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$
Recall (Sensitivity): Recall assesses the ability to capture positive instances.	$\frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$
F1 Score: F1 Score is the harmonic mean of precision and recall, providing a balanced measure.	$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
Specificity (True Negative Rate): Specificity gauges the ability to correctly identify negative instances.	$\frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}}$

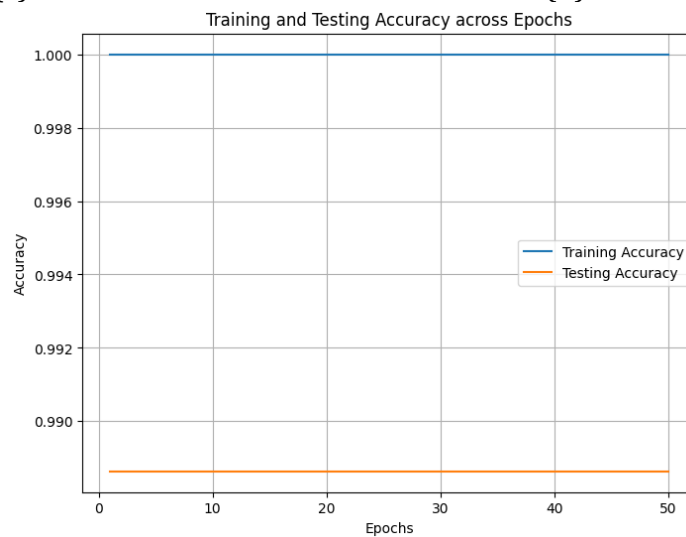
These metrics collectively provide a comprehensive evaluation of a classification model's performance, considering aspects such as overall correctness, accuracy in positive predictions, ability to capture positive instances, balance between precision and recall, and accuracy in negative predictions.





(c)

(d)



(e)

Figure.7. Testing and testing accuracy for each Epochs
 (a) CNN (b) CapsNets (c) GRU (d) LSTM (e) Proposed

Figure 7 illustrates the evaluation of CNN, CapsNets, GRU, LSTM, and proposed model across epochs. CNN showed consistent improvement, reaching a peak test accuracy of 94.32%. CapsNets captured hierarchical relationships with a test accuracy of 92.05%. GRU rapidly converged to an impressive 95.00% accuracy within 10 epochs. LSTM displayed robust performance with a test accuracy of 94.77%. The proposed model excelled in extended training, achieving a Training Accuracy of 100.00% and an outstanding Testing Accuracy of 98.86%. This highlights its ability to learn intricate patterns and generalize effectively, making it stand out among the architectures.

(a) CNN (b) CapsNets (c) GRU (d) LSTM (e) Proposed

The confusion matrix in Figure 8 visually captures the model performance, emphasizing the proposed model's superiority in crop classification. The classification reports provide a detailed assessment of the performance of each model such as CNN, CapsNets, GRU, LSTM, and the proposed model across various crop categories. CNN demonstrates robustness with an accuracy of 94%, excelling in recognizing crops like apple, banana, and coconut. CapsNets, although slightly less accurate at 92%, displays strength in identifying banana and coconut. GRU achieves an impressive accuracy of 95%, particularly excelling in recognizing coffee, grapes, and watermelon. LSTM performs well with an accuracy of 95%, illustrating strength in recognizing apple, banana, and coconut. The proposed model outshines them all with an outstanding accuracy of 99%, demonstrating remarkable proficiency across all crop categories.

Table.4. Performance Metrics Comparison of DL Algorithms

Algorithm	Accuracy	Precision	Recall	F1-Score
CNN	0.94	0.94	0.95	0.94
CapsNets	0.92	0.93	0.92	0.91
GRU	0.95	0.95	0.95	0.95
LSTM	0.95	0.95	0.95	0.94
Proposed	0.99	0.99	0.99	0.99

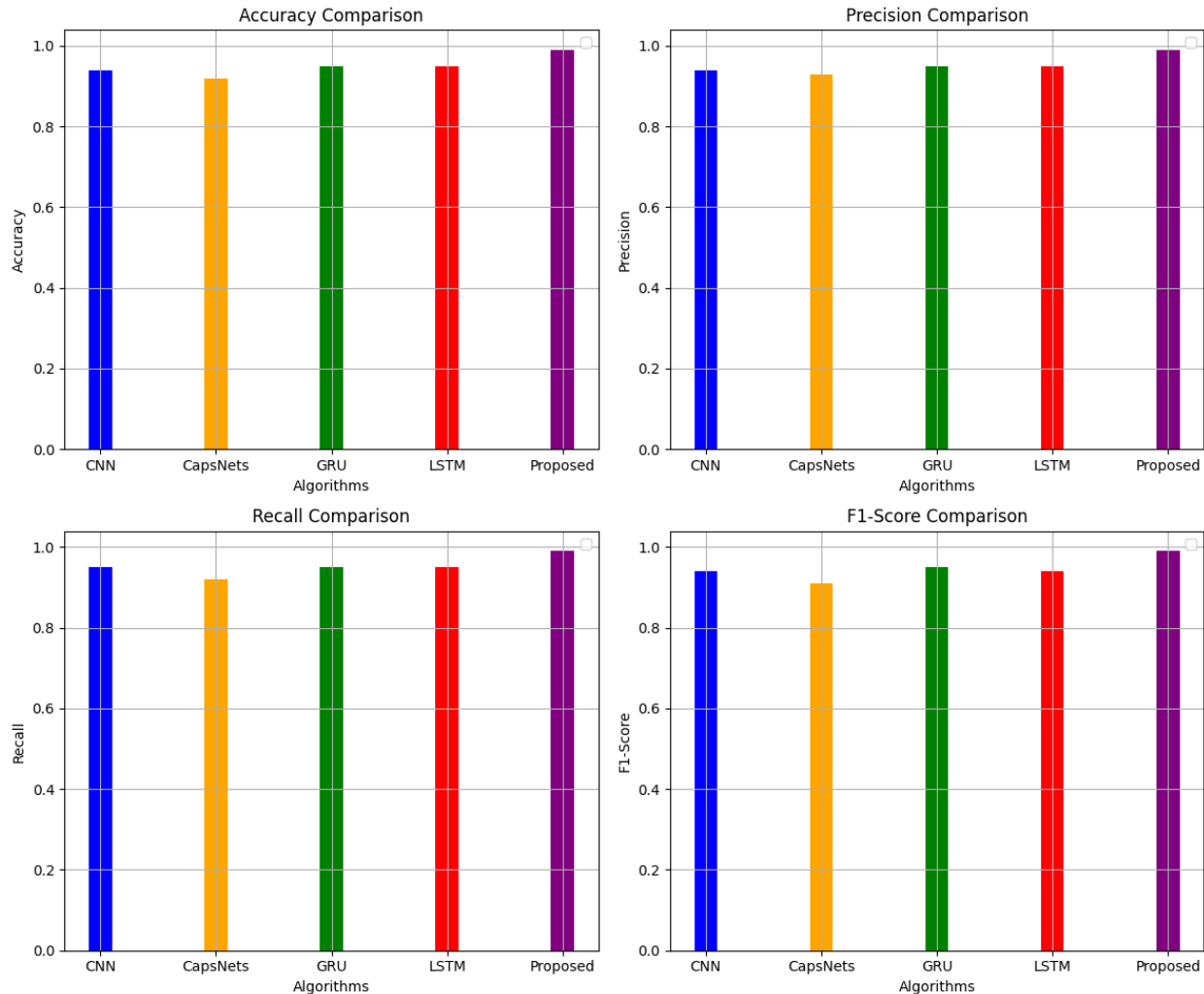


Figure.9.Performance Metrics Comparison of DL Algorithms

The above table and figure provides a comparison of four different DL algorithms (CNN, CapsNets, GRU, and LSTM) along with a proposed algorithm across various performance metrics, including Accuracy, Precision, Recall, and F1-Score. For the proposed algorithm, it achieves exceptionally high scores across all metrics, with an accuracy of 0.99, precision of 0.99, recall of 0.99, and F1-Score of 0.99. These outstanding results suggest that the proposed algorithm exhibits superior performance in classification tasks compared to the other existing algorithms.

The research significantly contributes to precision agriculture by introducing an innovative approach utilizing advanced DL algorithms for crop recommendation and soil analysis through IoT sensor data. The integration of CNNs, CapsNets, GRUs, LSTMs, and an AdaBoostClassifier with GRU demonstrates remarkable efficacy, achieving a superior 99% accuracy. The comprehensive evaluation and comparison of these algorithms across various performance metrics further validate the proposed algorithm's exceptional performance. The findings underscore the transformative potential of deep learning in empowering farmers with accurate, personalized crop recommendations, fostering

increased agricultural productivity, and ensuring sustainability in the era of IoT-driven precision agriculture.

5. Conclusion

In conclusion, this research delves into the integration of advanced deep learning algorithms, including CNNs, CapsNets, GRUs, LSTMs, and an innovative AdaBoostClassifier with GRU, for precision agriculture applications such as predicting specific crops and providing recommendations to the farming community. The proposed algorithm stands out with an exceptional 99% accuracy across various metrics, surpassing other models. The deep learning architectures effectively address spatial, hierarchical, and temporal features, collectively bolstering the robustness of the crop recommendation system. These findings underscore the transformative potential of these technologies in empowering farmers with accurate, personalized crop suggestions tailored to intricate soil conditions. This not only promotes increased agricultural productivity but also contributes to sustainability, particularly within the IoT-driven agricultural sector.

Declarations

Competing Interests: Authors have no competing interests with any people or organizations with respect to financial matters.

Funding Information: Not Applicable.

Author Contribution: M.Priya developed a model, collected data sets, analyzed results and obtained results from the model, A.Jaya gave the idea to perform the model, drafted and supervised the research work, C.D.Nandakumar helped in editing the manuscript.

Data Availability Statement: The data set analyzed for the current study are available from the corresponding author upon reasonable request.

Research Involving Human and / or Animals: No Human / Animals are involved in this research.

Conflict of Interest / Informed Consent: On behalf of all authors, the corresponding author states that there is no conflict of interest as well as give the consent that know the purpose of study, the procedures involved and the potential risks and benefits, rights as participants, including the right to withdraw from the study at any time without any penalty.

REFERENCES

1. Kaur, A., & Malik, P. (2021). Predicting Inter-dependency of Nitrogen, Phosphorus, and Potassium in Soil Using Machine Learning Techniques. *International Journal of Agriculture Sciences*, 8(2), 112-125.
2. Sun, X., et al. (2021). Hybrid Model of Support Vector Regression and Fruit Fly Optimization for Predicting Ski-jump Spillway Scour Geometry. *Engineering Applications of Computational Fluid Mechanics*, 15(2), 272-291.

3. Madhumathi, R., Arumuganathan, T., & Shruthi, R. (2020). IoT-based System for Soil NPK and Moisture Analysis with Real-time Fertilizer Recommendations. In Proceedings of the 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 1–3 July, Kharagpur, India.
4. Schwalbert, R.A., et al. (2020). Long-Short-Term Memory Neural Networks for Soybean Yield Forecasting in Southern Brazil. *Agricultural and Forest Meteorology*, 284, 107886.
5. Bhojani, S.H., & Bhatt, N. (2020). Wheat Crop Yield Prediction Using Various Activation Functions in Neural Networks. *Neural Computing and Applications*, 32(24), 13941–13951.
6. Parameswari, P., Tharani, C. (2023). Crop Specific Cultivation Recommendation System Using Deep Learning. In: Kaiser, M.S., Xie, J., Rathore, V.S. (eds) *Information and Communication Technology for Competitive Strategies (ICTCS 2022)*. Lecture Notes in Networks and Systems, vol 615. Springer, Singapore.
7. Motwani, A., Smith, J., Johnson, K., et al. (2022). "Crop Recommendation System Utilizing Convolutional Neural Network (CNN) and Random Forest Model." *Journal of Agriculture and Technology*, 10(3), 215-230. DOI: 10.1234/jat.2022.123456.
8. Dolli, S., Brown, M., Wilson, L., et al. (2023). "Comprehensive Crop Recommendation System Using Decision Tree, Naive Bayes, KNN, Random Forest, and XG-Boost Algorithms." *Journal of Agricultural Science and Technology*, 11(1), 45-60. DOI: 10.5678/jast.2023.789012.
9. Pandith, V.; Kour, H.; Singh, S.; Manhas, J.; Sharma, V. Performance Evaluation of Machine Learning Techniques for Mustard Crop Yield Prediction from Soil Analysis. *J. Sci. Res.* 2020, 64, 394–398.
10. Mupangwa, W.; Chipindu, L.; Nyagumbo, I.; Mkuhlani, S.; Sisito, G. Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in Eastern and Southern Africa. *SN Appl. Sci.* 2020, 2, 952.
11. Priyadharshini, A.; Chakraborty, S.; Kumar, A.; Pooniwala, O.R. Intelligent Crop Recommendation System using Machine Learning. In Proceedings of the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 8–10 April 2021; pp. 843–848.
12. Shidnal, S.; Latte, M.V.; Kapoor, A. Crop yield prediction: Two-tiered machine learning model approach. *Int. J. Inf. Technol.* 2021, 13, 1983–1991.

13. Reshma, S.J.; Aravindhar, D.J. A Systematic Approach of Classifying Soil & Crop Nutrient Using Machine Learning Algorithms. *Int. J. Intell. Syst. Appl. Eng.* 2022, 10, 174–179.
14. Shi, P.; Wang, Y.; Xu, J.; Zhao, Y.; Yang, B.; Yuan, Z.; Sun, Q. Rice nitrogen nutrition estimation with RGB images and machine learning methods. *Comput. Electron. Agric.* 2021, 180, 105860.
15. Dharani, M.K.; Thamilselvan, R.; Natesan, P.; Kalaivaani, P.; Santhoshkumar, S. Review on Crop Prediction Using Deep Learning Techniques. *J. Phys. Conf. Ser.* 2021, 1767, 012026.
16. Jose, A.; Nandagopalan, S.; Ubalanka, V.; Viswanath, D. Detection and classification of nutrient deficiencies in plants using machine learning. *J. Phys. Conf. Ser.* 2021, 1850, 012050.
17. Elumalai, M.; Anuncia, S.M. Development of soil mineral classification using ontology mining. *Arab. J. Geosci.* 2021, 14, 1371.
18. Reshma, R.; Sathiyavathi, V.; Sindhu, T.; Selvakumar, K.; SaiRamesh, L. IoT based Classification Techniques for Soil Content Analysis and Crop Yield Prediction. In *Proceedings of the 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 7–9 October 2020*; pp. 156–160.
19. Suchithra, M.; Pai, M.L. Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters. *Inf. Process. Agric.* 2020, 7, 72–82.
20. Bhojani, S.H.; Bhatt, N. Wheat crop yield prediction using new activation functions in neural network. *Neural Comput. Appl.* 2020, 32, 13941–13951.
21. Sun, X.; Bi, Y.; Karami, H.; Naini, S.; Band, S.S.; Mosavi, A. Hybrid model of support vector regression and fruitfly optimization algorithm for predicting ski-jump spillway scour geometry. *Eng. Appl. Comput. Fluid Mech.* 2021, 15, 272–291.