[Mesopotamian Journal of Computer Science](https://map.researchcommons.org/mjcsc)

Volume 5 | Issue 1 Article 1

1-1-2025

Leveraging Artificial Intelligent for Optimized Crop Production: An ANN-Based Approach

Yahya Layth Khaleel
Department of Computer Science, Computer Science and Mathematics College, Tikrit University, Tikrit 34001, Iraq, yahya@tu.edu.iq

Fadya A. Habeeb
Department of Mathematics, College of Education for Women, Tikrit University, Tikrit 34001, Iraq.

Mustafa Abdulfattah Habeeb
Department of Computer Science, Computer Science and Mathematics College, Tikrit University, Tikrit 34001, Iraq .

See next page for additional authors

How to Cite This Article

Khaleel, Yahya Layth; Habeeb, Fadya A.; Habeeb, Mustafa Abdulfattah; and Ameen, Fatimah N. (2025) "Leveraging Artificial Intelligent for Optimized Crop Production: An ANN-Based Approach," *Mesopotamian Journal of Computer Science*: Vol. 5: Iss. 1, Article 1.
DOI: https://doi.org/10.58496/MJCSC/2025/001
Available at: https://map.researchcommons.org/mjcsc/vol5/iss1/1

Leveraging Artificial Intelligent for Optimized Crop Production: An ANN-Based Approach

Authors

Yahya Layth Khaleel, Fadya A. Habeeb, Mustafa Abdulfattah Habeeb, and Fatimah N. Ameen

Mesopotamian journal of Computer Science
Vol. 2025, **pp**. 1–16
DOI: https://doi.org/10.58496/MJCSC/2025/001 ; ISSN: 2958-6631
https://mesopotamian.press/journals/index.php/cs

Research Article

Leveraging Artificial Intelligent for Optimized Crop Production: An ANN-Based Approach

Yahya Layth Khaleel[1, *], Fadya A. Habeeb[2], Mustafa Abdulfattah Habeeb[1], Fatimah N. Ameen [3]

[1] *Department of Computer Science, Computer Science and Mathematics College, Tikrit University, Tikrit 34001, Iraq*
[2] *Department of Mathematics, College of Education for Women, Tikrit University, Tikrit 34001, Iraq.*
[3] *Institute of Automation and Info-Communication, Faculty of Mechanical Engineering and Informatics, University of Miskolc, Miskolc, Hungary.*

ARTICLE INFO

Article History
Received 18 Oct 2024
Revised: 19 Nov 2024
Accepted 18 Dec 2024
Published 11 Jan 2025

Keywords
Crop
Sustainability
Agriculture
Artificial Intelligent
Machine Learning
Deep Learning
Analysis

ABSTRACT

To incite modern day crop production and ensure sustainability, exact crop recommendations are key to the process. This study pays significant attention to the need for the use of big data tools in studies involving comprehensive data sets that contain information on soil and other environmental characteristics. The set of data used in this research includes Nitrogen, Phosphorus, and Potassium content coordinated with Temperature, Humidity, pH Value, and Rainfall. Knowing these factors is to make a favorable decision about improving agricultural products yield, availability and management of the resources, as well as general well-being of the crops. Specialized advisory on crop type according to the specific data of the soil is likely to click the concept of yield by dealing with nutrient gaps and variable environments. This approach also guarantees reasonable distribution of the limited resources as well as cutting the cost of production and the impact on the environment through preserving the soils and keeping off the faster pace of nutrient leaching. This method of producing plants also helps in minimizing the utilization of fertilizers and enhanced inputs hence minimizing the possible negative impacts on the environment as well as increasing the lifespan of the field. Furthermore, specific advice also assists the farmers in handling changes in climate patterns in a way that increases crop's ability to withstand the impacts of more frequent episodes of the extremities and manage the risks implicated in those circumstances. Since climate issue plays significant role in threatening the global food security, the strategic maneuvering capacity which enables to analyze the environment and respond to the changes in a timely manner has gained added importance. This also eliminates the losses likely to be incurred as a result of adverse weather factors through the use of data and empirical analysis, which provides management measures to be adopted in order to protect crops and ensure steady food production. In the end, enhanced decision making regarding the types of crop that do well on the soil can easily enhance farmers' revenues via; improved yields, improved crop quality, lower input expenses etc. This does not only positively affect the economics of the farming enterprises but it is also a way of supporting the stability of agricultural businesses globally. This has informed a build of an Artificial Neural Network (ANN) model which is capable of analyzing the large and rich dataset of soil information to specify accurate crop recommendation. In the validation of the model, this achieved an accuracy level of 99% as it manifested its ability to provide succor in decision making analysis in the sector of agriculture. In comparison to the conventional algorithms in Machine Learning like Logistic Regression, K-Nearest Neighbors (KNN), Multinomial Naive Bayes algorithms, our ANN model exhibits higher accuracy and consistent performance. This study thus establishes the importance of adopting the higher inclusive machine learning with improved data about soil to enhance a more efficient, resilient, and sustainable agriculture era.

*Corresponding author. Email: yahya@tu.edu.iq

1. INTRODUCTION

The use of artificial intelligence (AI) can change lot of things because of the advanced capabilities in data analysis [1], [2], processing, and decision making [3]–[5]. With a wide range of applications across the fields of healthcare [6]–[8], finance [9], autonomous vehicles [10], environmental studies [11], natural language processing [12], [13], production [14]–[16], education and more to drive operational improvements [17]–[20], user experiences and decision making [21]–[23], it is no surprise that its presence is being felt. AI has helped on forensic evidence analysis [24], preventing the spread of fake news [25], as well as improving of public safety by smart technologies [26], [27]. It also plays an important role both in cybersecurity to protect information systems from adversarial attacks and building secure digital environments [28], [29].
AI is critical for improving agriculture through crop management and crop productivity in agriculture [30]. AI through technologies of predictive analytics, smart irrigation and disease detection systems helps farmers to maximize resource usage, maximize yields and cope with climate condition challenges. AI's potential to transform this is actually a key reason why it's such an important contribution to tackling global food security [31]; supporting better sustainable farming practices; and accelerating the pace of technological innovation in ag. As a whole, AI serves as a powerful well, helping to solve the challenges, helping to stimulate creativity and create the future of different areas.
Now, agriculture stands situated right between two opposite poles – on one side there is the pursuit of the highest yield as well as on the other side, there, is sustainability which is, hand in hand with data science principles [32], [33]. According to this therefore it will be important to lay down emphasis on a number of facets that may be relevant to the present instance and may not be limited only to the type of the soil that is in the selected land but also may include a large number of factors that may be supposed to have impact on the growth of the plant [34]–[37]. This is where this work undertakes a comprehensive and systematic endeavor at trying to unravel all those elements that seem to define these factors and the implications they have for the productivity and well-being of crops as well as the general agricultural system.
Again, in the strategies of innovations in emerging agri-food systems, there is the cardinal principle of a circumscribed or selective intervention in the socio-cultural environment of the agri-food systems to be transformed, in ways that the depth, spatial coverage/ intensity and specific areas of intervention correspond to the opportunities assessment of the ground realities [38]. It becomes important therefore to consider the factors that are associated with soil quality these include Nitrogen, Phosphorus, Potassium these are nutrients and Temperature, Relative Territory, pH_Value and RAINFALL these are environmental features [39]–[43]. One is therefore able to ask these questions and make rational decisions that steer agriculture in the desired direction, this in essence, facilitates its direction towards utter sustainability and attaining higher yields.
The main emphasis in this work lies on the idea of individual recommendations of crops – practice that goes beyond the traditional frameworks and relies on the advantages provided by big data. These recommendations aim at enhancing per crops the possibility of yields as well as decreasing the necessities of input and unfavorable effects on soil, water, and air. This does not only include a more expressive economic effect with lower costs for the farmers but also leads to its more sustainable forms of farming not damaging the soil health and having less impacts on the contamination of waters.
However, this research considers it crucial to have climate resilience be part of the decisions made in the agricultural sector. Several changes in the environment have been observed in the recent past especially due to effects of climate change and therefore management of food production has required this ability [44]–[46]. Due to climate change farmers can reduce risk factors which are associated with climate shocks, thus ensuring crop productivity to be stable under any climate conditions by incorporating crop advice to changing climatic conditions which are recommended by experts [47], [48].
Closely connected to these objectives is a proper application of the latest technologies, specifically Artificial Neutral Networks (ANNs), to bring data analysis to their utmost potential. These complex mathematical models have helped in the scientific analysis of various parameters in the course of agriculture and the combining of several sets of data felt to a very high level. In a way, through their ability to analyze significant quantities of data, ANNs enable farmers and other specialists in the sphere of agriculture to make informed decisions in real life, thus increasing crop yield and protecting it against failures.
Therefore, in a nutshell, this research calls for the integration of ISFM strategies inclusive of information technology in to the conventional farming practices. Therefore, it is evident that innovations and use of analytics should be embraced so that the farmers are able to deal with the challenges associated with the contemporary farming practices. In this psychology of moving towards a better and more productive world of agriculture, the findings of this study are expected to form the fundamental paths for the generation of the next generation of agricultural innovations.
At the core of these objectives, it is possible to identify the aim to apply advanced data analytics tools, with a particular emphasis placed on the Artificial Neural Networks (ANNs) approach. They are complex computational procedures that have transformed the sciences and practices of agriculture because they assist with the amalgamation and interpretation of multifarious data sets in an efficient and very precise manner. Due to the ability of ANNs to handle massive data and

*Corresponding author. Email: yahya@tu.edu.iq

coming up with accurate results, farmers and agricultural specialists can make appropriate decisions in real-time to increase crop yields and their resistance [49]–[52].
In other words, this study can be viewed as a plea for including the findings from advanced methods into the practices applied within the given field of agriculture. Farming has become tremendously challenging due to factors such as global warming that has led many farmers to seek the support of innovation and the power of big data so that they be equipped to face the challenges that modern farming presents. As we progress on the quest of achieving sustainable and enhance agriculture, we have been able to dissect and discover the knowledge that will reform the course of agriculture in the future.

2. RELATED WORKS

The importance of accurate crop recommendations in today's farming practice be overemphasized with the increasing awareness towards non-heuristic decision making for crop production. From different studies, it has been established that the use of extensive databases of environmental metrics is essential for the improvement of agricultural yield and crop quality. For example, the paper [53] describes a new approach to control the crop irrigation process through the Phyto indication system that implements the computer vision algorithms. It has customizable for center pivot irrigation and includes 8 IP cameras in conjunction with DVR and a laptop. Image preprocessing and classification stages, usage of the neural networks in plant identifying and mapping stages provide high accuracy up to 93% in plant identification and 92% in growth stage identification; the system shows high results compared to similar systems. The measured benefits are enhanced productivity and overall water utilization without hesitation on the farm environment.
While in [54] addresses the critical need for accurate crop yield prediction in India, focusing on major kharif crops in Andhra Pradesh's Visakhapatnam district. Using modular artificial neural networks (MANNs) to predict monsoon rainfall and support vector regression (SVR) to estimate crop yield, the proposed methodology enables the development of effective agricultural strategies. Comparative analysis demonstrates the superior predictive performance of the MANNs-SVR approach in kharif crop production forecasting.
In addition, the authors in [55] introduces a Bayesian Neural Network (BNN) model for county-level corn yield prediction in the U.S. Corn Belt. Utilizing multiple publicly available data sources, including satellite products, climate observations, and historical yield records, the model achieves high prediction accuracy (R^2 of 0.77) and effectively quantifies predictive uncertainty. Timeliness of prediction and factors influencing uncertainty are also analyzed, offering insights for improved crop yield estimation and agricultural decision-making.
In the same way, in [56] assesses the efficacy of combining UAV-based RGB, multispectral, and thermal sensors for soybean grain yield prediction using Deep Neural Network (DNN) models. Findings reveal that multimodal data fusion enhances prediction accuracy, with DNN-based models, particularly DNN-F2, achieving the highest accuracy (R^2 of 0.720 and RMSE% of 15.9%). These models are revealed to have a low tendency to saturation and are thus more versatile across the numerous types of soybean genotypes including the potential for efficient and spatially adaptive crop yield prediction.
Finally, [57] explores the use of deep learning architectures for estimating crop yield in field images, vital for modern agricultural technologies and ensuring food security. Utilizing ground-mounted cameras capturing plant images at half-hour intervals, the study employs intermediate outputs of deep learning models to develop a relative measure of crop yield estimation. Experimentation on sunflower image sequences from four parcels yielded promising results, achieving an accuracy of 87%.

3. METHODOLOGY

The research methodology includes five basic stages shown in Figure 1:

Fig. 1. Steps of the Methodology

3.1 Data Collection

The data employed in the “Crop Recommendation” was obtained from Kaggle [58], which is a platform for data science and machine learning. In it there are data about different parameters of agriculture and its associated crops recommendations.

The data contains:

- Nitrogen: The immediate physical and actual measurement of the concentration of nitrogen in the soil.
- Phosphorus: The total and available phosphorus in the soil.
- Potassium: Potassium K is considered to be available at this given level of the soil.
- Temperature: Thermal energy is the amount of heat contained in the region.
- Humidity: About the climate, the percentage of humidity in the region.
- pH_Value: The pH value of the specific type of soil.
- Rainfall: The rate of rainfall in the region if any.
- Crop: Some of the recommended crops are Rice, Maize, Chick Pea, Kidney Beans, Pigeon Peas, Moth Beans, Mung Bean, Black Gram, Lentil, Pomegranate, Banana, Mango, Grapes, Water Melon, Muskmelon, Apple, Orange, Papaya, Coconut, Cotton, Jute and Coffee.

Below are the first five rows of the dataset in Table 1, followed by the descriptive statistics for the variables in Table 2

TABLE I. FIRST FIVE ROWS OF THE DATASET

	Nitrogen	Phosphorus	Potassium	Temperature	Humidity	pH_Value	Rainfall	Crop
0	90	42	43	20.879744	82.002744	6.502985	202.935536	Rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	Rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	Rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	Rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	Rice

TABLE II. DESCRIPTIVE STATISTICS FOR THE VARIABLES

	Nitrogen	Phosphorus	Potassium	Temperature	Humidity	pH_Value	Rainfall
count	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000
mean	50.551818	53.362727	48.149091	25.616244	71.481779	6.469480	103.463655
std	36.917334	32.985883	50.647931	5.063749	22.263812	0.773938	54.958389
min	0.000000	5.000000	5.000000	8.825675	14.258040	3.504752	20.211267
25%	21.000000	28.000000	20.000000	22.769375	60.261953	5.971693	64.551686
50%	37.000000	51.000000	32.000000	25.598693	80.473146	6.425045	94.867624
75%	84.250000	68.000000	49.000000	28.561654	89.948771	6.923643	124.26750
max	140.00000	145.00000	205.00000	43.675493	99.981876	9.935091	298.56011

The compiled dataset includes 2200 lines, where each line describes a particular type of soil and environmental conditions as well as the recommended crop. Every record in the dataset is confined to such features as pH level, nutrient availability, humidity, precipitation rates, and other characteristics of the environment. Such a diverse set of data is particularly useful in the machine learning approaches to developing a model to predict the suitable crop for a particular environment. With the help of the above stated models, it is possible to enter different parameters concerning the environment and the soil type to arrive at accurate recommendations for the type of crop to be grown.

The main use is in helping farmers and those involved in agricultural decisions to choose the right crops for farming. In the case of farmers, this means that the application of such predictive models will greatly contribute to the improvement of productivity. In this way farmers are able to make informed decisions related to cropping practices that will not likely result in crop failure, increase crop yield and therefore promote sustainable farming practices given their soil and environmental

conditions. Moreover, these models may useful to delineate the area that requires addition of soil additives or other treatments to obtain the right conditions for plant growth.
Therefore, the detailed information regarding the soil and environmental condition and the corresponding crop recommendation from the dataset under consideration make the catalog quite suitable for developing the machine learning models. These models, in their turn, help farmers and decision-makers in agriculture to make choices for improving yields and the outcomes of the agricultural industry.

3.2 Preprocessing

The crop variables which were used into the equation include nitrogen, phosphor, potassium, temperature, humidity, ph, and rain fall. The following data pre-processing took place as part of preparing the data for analysis and use in the models (see Figure 2):

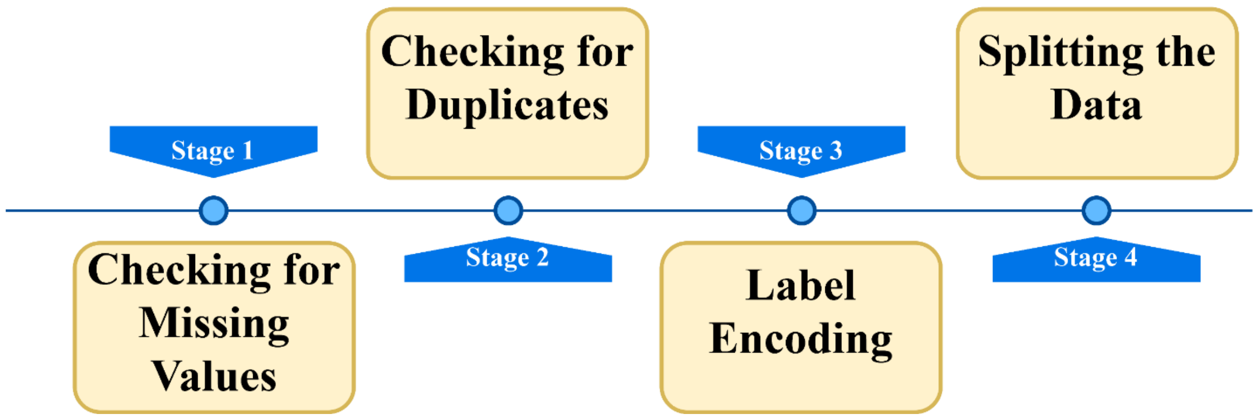

Fig. 2. Steps of Preprocessing

- **Checking for Missing Values:** It is possible to manage with missing values in various ways but it is significant to take into consideration reasonable balance between it and do not bring quality of data to the worse state. It also checked the entries of the dataset for omitted values not contained in the lists and arrays. This issue cited that it was necessary to show which values should have been omitted to avoid distorting the results of the analysis and the creation of predictive models based on the use of the data.
- **Checking for Duplicates:** It can manifest in such aspects as overlapped instances which occurs when the data set has been a copy of data set X, and above data set x using the same data will provide a skewed, or else repeating the result. The data also checked were whether the data set contains any of the duplicate entry in it or not? This issue was minimized by performing constant checks for any duplicated record, which entailed deleting such records to ensure that each input was different in improving the analysis of the model.
- **Label Encoding:** Also, it is necessary to emphasize that categorical variables can be transformed into numerical ones and that is why the target variable will be utilized in this form, the ‘Crop’ variable.
- **Splitting the Data:** Splitting of the data et resulted in attainment of 75% of training data and 24% test data in the process of measuring the model performance.

Therefore, the preprocessing of the dataset was rather strict, and after the given step, it is possible to state that the structure of the dataset is appropriate for further analysis and modeling. It is equally so when it comes to preparing for the enhancement of the crop prediction model’s effectiveness. These steps include dealing with missing values, scaling the features, the process of converting the categorised values to numerical and feature reduction and if these steps are not taken, then the model is likely to be very unreliable and non-predictive. Through this elaborate preprocessing, the dataset is prepared in a manner suitable for training high performing Machine learning algorithms and, therefore, arrives at efficient recommendations for crops selections.

3.3 Feature Engineering

In this section, we performed feature engineering to enhance the dataset used for predicting crop types based on various factors such as Nitrogen, Phosphorus, Potassium, Temperature, Humidity, pH Value, and Rainfall. Before knowing the relationship between the features, we presented the individual characteristics of each feature as numerical columns (see Figures 3 and 4).

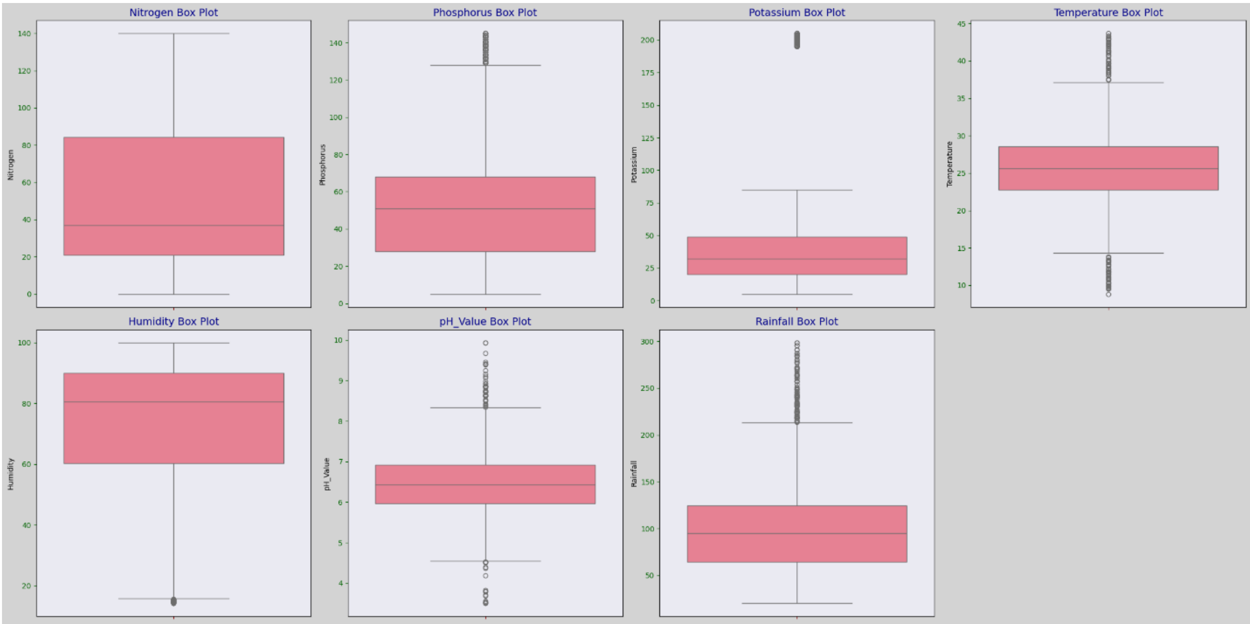

Fig. 3. Plot the Numerical Columns of Features as a box plot

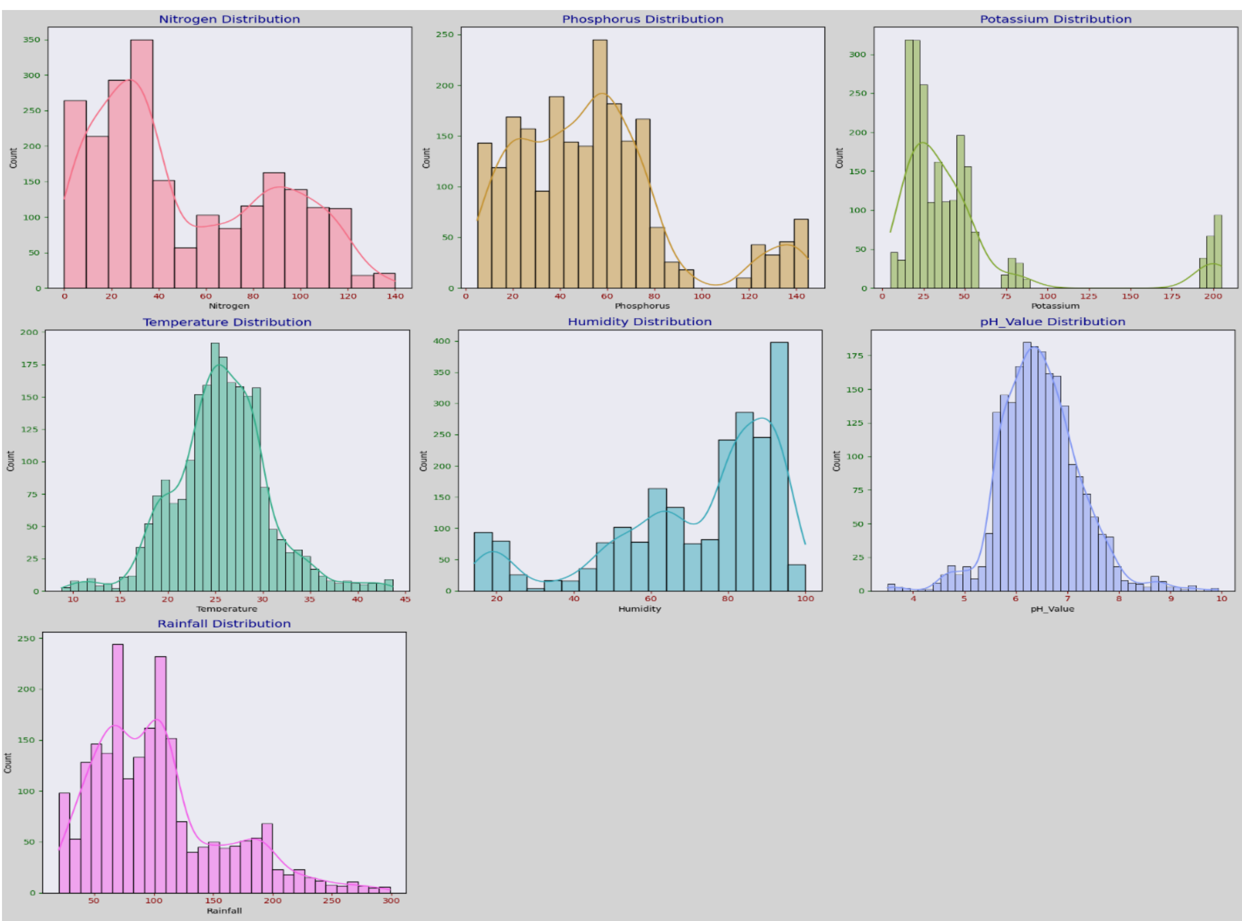

Fig. 4. Plot the Numerical Columns of Features as a Histogram

Pair Plot of Features: To visually explore the relationships between different features, we generated a pair plot as presented in **Figure 5**. a pair plot helps in understanding the distribution of individual features and the interactions between them.

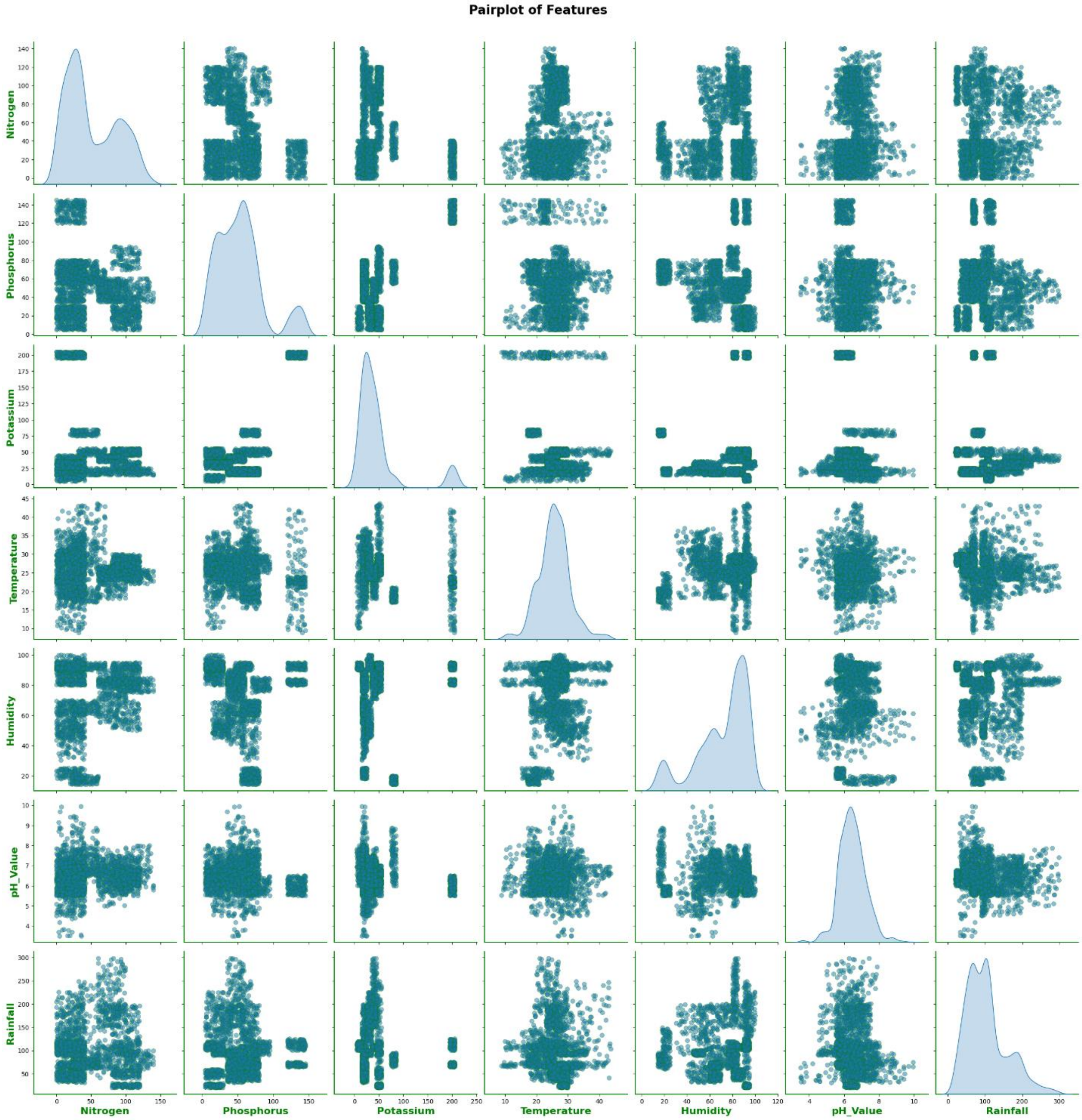

Fig. 5. Pair Plot of Features

Finding Correlation: Next, we calculated the correlation matrix of the dataset to identify the strength and direction of the relationships between the features. The correlation matrix helps in understanding which features are strongly correlated with each other. Features with high correlation might be redundant and can be considered for removal to reduce multicollinearity. Additionally, identifying highly correlated features with the target variable (crop type) can help in selecting important features for the model. Figure 6 below shows the correlation matrix

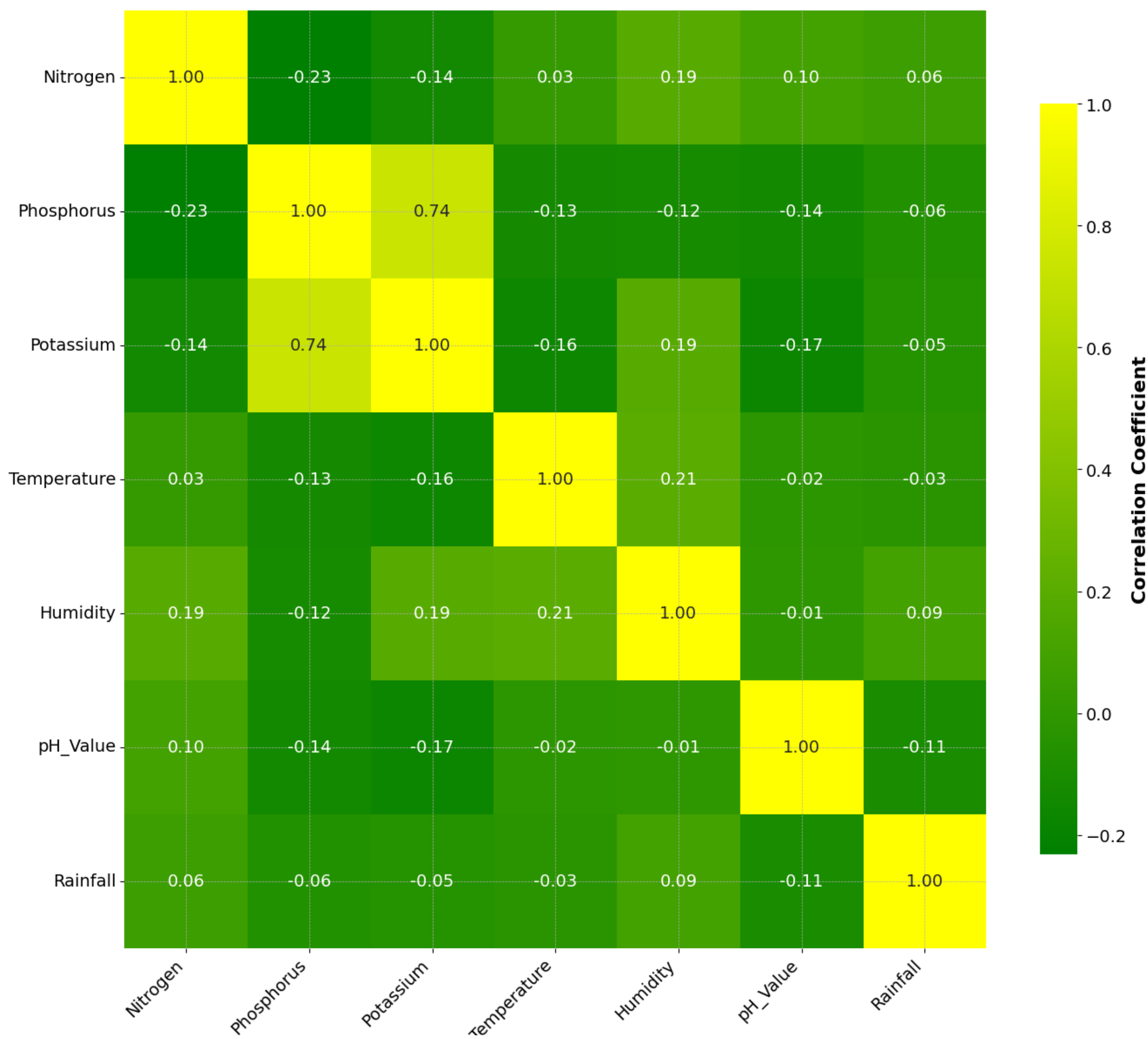

Fig. 6. Correlation Matrix

By carefully analyzing the pair plot and correlation matrix, we can make informed decisions about feature creation, transformation, and selection, which are essential steps in building a robust predictive model.
Furthermore, by continuing further with the analysis of the data as tabled above, it can be seen that variability of nitrogen content and rainfall in the different crops is high. .For example, in growing banana and coffee crops, excessive nitrogen content and proportional rainfall are among the requirements identified. This may point out to the fact that these crops need nitrogen in adequate amounts and that large quantities of rainfall may be needed by such crops to fulfill their nitrogen needs. Other crops such as lentil crop or pomegranate crop requires comparatively smaller nitrogen input, which may be due to the limited figures mentioned in the aspect of rainfall. Representing this trend is some crops that have been instilled to grow in the regions that experience low rainfall and low nitrogen values. In the same way, sturdy nitrogen requirement for crops like jute as well as moderate rainfall for Maize suggested that the crops might need the nutrient as well as the water in moderate amounts. All in all, it is evident from the data provided here there is no general trend whereby certain crops prefer to grow where other crops are grown based on the nitrogen and intensity of rainfall, but rather, based on what the ecology of each crop requires- if it is the nitrogen or the intensity of rainfall it requires- the ecology of the different crops again. This is a very important relationship and helpful in relation to crop calendar for it defines link between rates of nutrient in the soil, irrigation regimes and the amount of rainfall available (see Figure 7).

Fig. 7. Scatter Plot of Nitrogen vs Rainfall

3.4 Model Development and training

An Artificial Neural Network (ANN) is a computational model that mimics the neural structure of the human brain to process complex data inputs. It consists of layers of interconnected nodes, or neurons, where each connection has an associated weight. The primary objective of an ANN is to learn patterns and relationships within the input data through training, allowing it to make accurate predictions or classifications. Each neuron receives inputs, processes them with an activation function, and passes the output to the next layer. The training process of a neural network entails setting up weight values of these connections such that a minimum error rate is achieved through backpropagation and gradient descent, among others. In this study, we have proposed the ANN model to select the suitable crops according to the soil and environmental characters like Nitrogen, Phosphorus, Potassium, Temperature, Humidity, pH Value, Rainfall. The model architecture and training process are detailed below (see Figure 8).

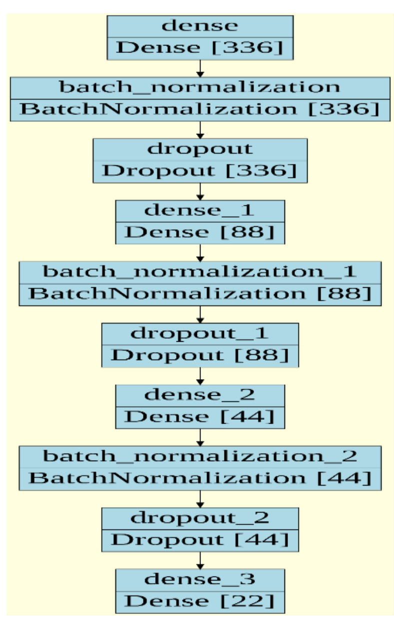

Fig. 8. plotting model

- Input Layer: The input layer accepts features corresponding to the soil and environmental conditions. The number of neurons in this layer equals the number of features in the dataset.
- First Hidden Layer: Contains 336 neurons with a ReLU (Rectified Linear Unit) activation function. ReLU helps in learning complex patterns by introducing non-linearity.
- Batch Normalization: Applied after the first hidden layer to normalize the output and accelerate the training process.
- Dropout: A dropout rate of 0.2 is used to prevent overfitting by randomly setting 20% of the neurons to zero during each training step.
- Second Hidden Layer: Contains 88 neurons with a ReLU activation function.
- Batch Normalization: Applied after the second hidden layer.
- Dropout: A dropout rate of 0.1 is used to prevent overfitting.
- Third Hidden Layer: Contains 44 neurons with a ReLU activation function.
- Batch Normalization: Applied after the third hidden layer.
- Dropout: A dropout rate of 0.1 is used to prevent overfitting.
- Output Layer: Contains 22 neurons with a softmax activation function, which outputs a probability distribution over the 22 possible crop types, enabling the model to predict the most suitable crop.

The model is compiled and trained with the following components:

- Optimizer: Adam optimizer is used for its efficiency and adaptive learning rate capabilities.
- Loss Function: Categorical cross-entropy is used as the loss function since this is a multi-class classification problem.
- Early Stopping: Monitors the validation loss and stops training if it does not improve for 10 consecutive epochs, restoring the best weights to avoid overfitting.
- Model Checkpoint: Saves the best model during training based on validation loss.
- Reduce Learning Rate on Plateau: Reduces the learning rate by a factor of 0.2 if the validation loss does not improve for 5 consecutive epochs, with a minimum learning rate of 0.0000005.

The model is trained for a maximum of 150 epochs and split the data into validation set to track on the performance. This setup contributes to obtaining a broader model that enables to guess the most appropriate crop as per the given soil and other environmental conditions.

3.5 Evaluation and Comparison

Proper metrics to assess the crop recommendation performance of the developed model are the confusion matrix, accuracy, and model fitting in a held-out test set. The confusion matrix would help understand the model's effectiveness in classifying the recommended crop correctly while the accuracy would be an overall measure of the model. Furthermore, checking the model fitting on the test set would be useful in ascertaining the capacity of the ANN model to identify the existing patterns in the data set.

In other words, the study aimed at comparing the performance of the ANN-developed model with other methods such as ordinary statistical methodologies or other machine learning algorithms including Logistics Regression, K-Nearest Neighbor, and Multinomial Naive Bayes, So, the idea would be to ensure that significantly better performance of the ANN model in the aspect of crop recommendation could be shown. The comparison is done by training and testing the two or more proposed models on the same set of data to reduce prejudice.

The study, therefore, can provide a detailed analysis and comparison of the model formulated through ANN with the other methods to point out the benefits and enhanced precision in recommending the suitable crop via inputting the right type of soil and environment.

4. RESULTS

The analysis of the crop prediction using the ANN model was conducted, and the results show high performance and accuracy. Outlined below in detail are aspects that include accuracy, training history, and the confusion matrix to support the result of the proposed machine learning model.

The success rate of the ANN model was measured to be 99% and this proves that the model can recognize the types of crops rightly. An even greater indication of this high accuracy is given by the other parts of the confusion matrix as the one presented in figure 9, that depicts the capacity of the model to classify most of the thing's instances with minimal misclassifications. High accuracy of a similar nature was achieved for other classes, proving the efficiency of the used model.

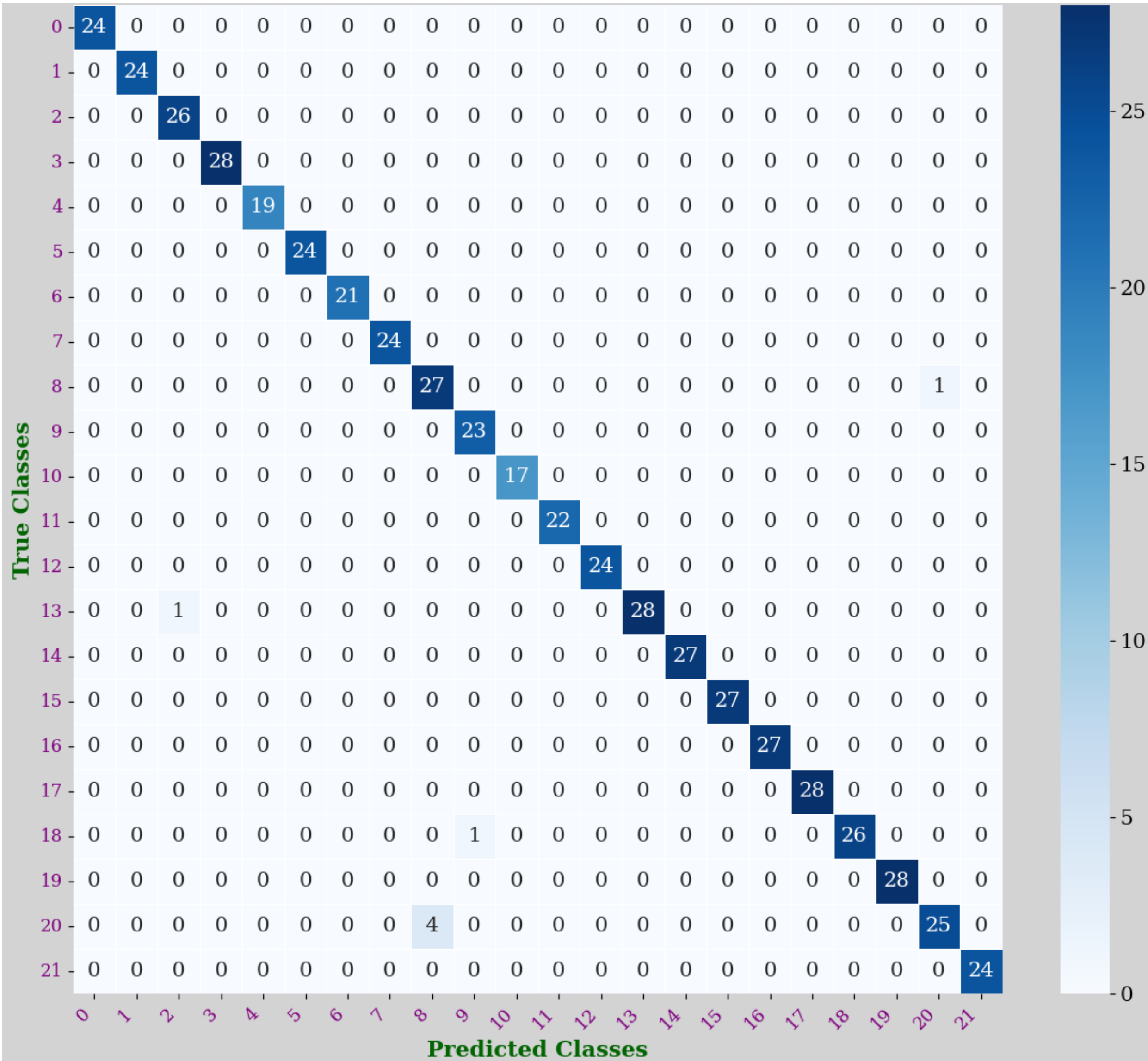

Fig. 9. Confusion Matrix

In fitting the ANN model for crop prediction, the total epochs of training were set to 150, and the transformations of the training and validation metrics were observed. The performances of the model were very much promising as the depicted good convergence and a good fit to the training set. As it can be viewed from Figure 10, the fitting of the proposed model is appropriate.

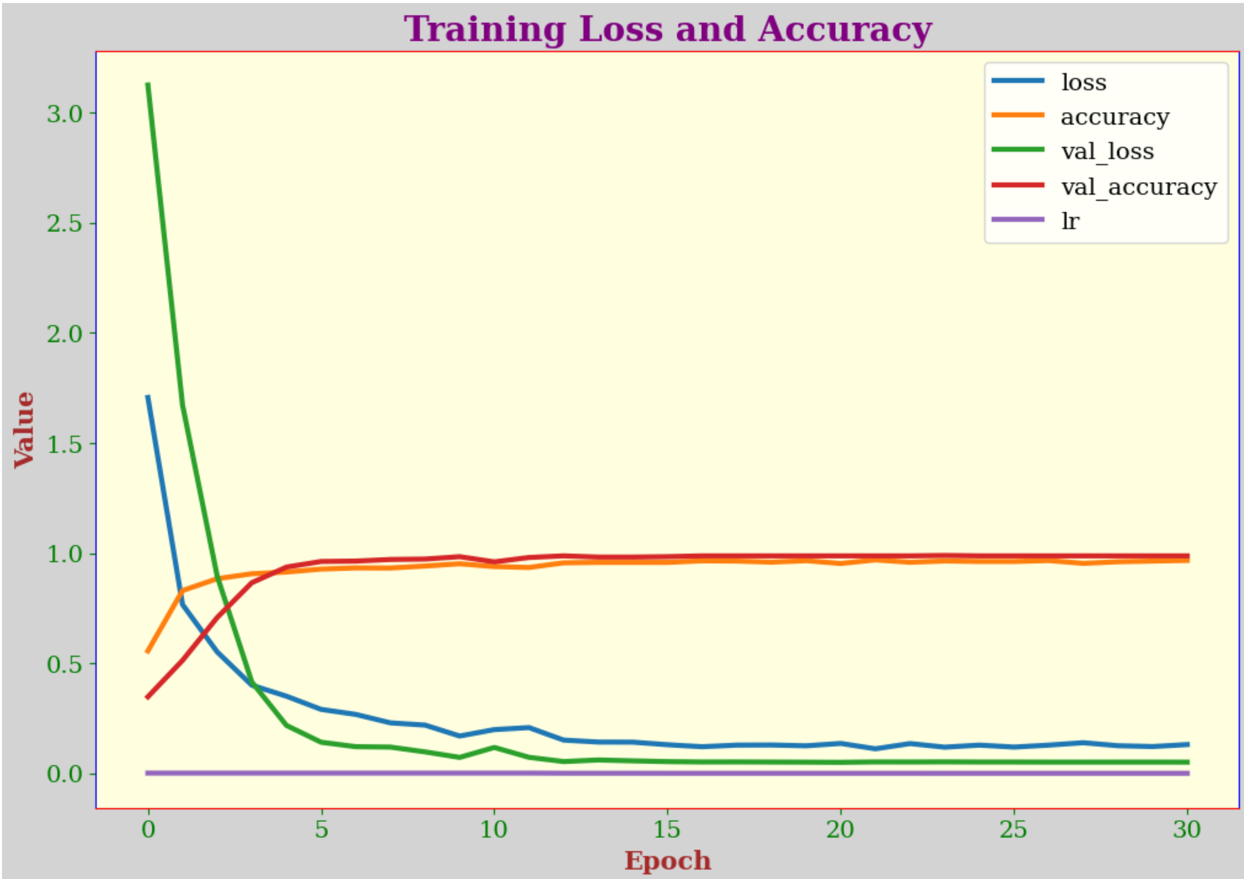

Fig. 10. ANN fitting

In the first 5 epochs, the model had sharp increases in learning as envisaged during the training phase. Training accuracy rose to 55. 58% to 91. And validation accuracy increased it from 34. 73% to 93. 64%. At the same time, the loss showed the quantifiable reduction, which meant the model was learning the patterns in data sets promptly.

From epoch 6 to epoch 20, it is seen that the model is fine tuned progressively. After epoch 10, training Accuracy was at 95.15 percent. and hence, the validation accuracy was 98. 36% and the validation loss dropped to 0. 0721. These results indicated that the model has successfully learnt features of training data and moreover, it is adequately valid in recognition of unseen validation data.

The overall trend in the accuracies obtained from the trained model, showed that the the model has reached a good fit during the later epochs particularly from epoch 21 to epoch 30. The training accuracy was at par with about a 96% level to 96. 55% and in validation accuracy the percentage was remains very high 98.73%. The validation loss did not continue to decrease and reached around 0. 0509 that depicts that the model has reduced overfitting to an extremity.

The values to adjust learning rate were also instrumental in the aspect of fitting. Initially set at 0001, the learning rate contributed to faster learning. With regards to the training process, the learning rate was subsequently decreased throughout the model's learning process to ensure that the weights could be adjusted accurately in trying to achieve the optimal minima without overshooting it. Thus, by epoch 30, it dropped a few notches to settle at a learning rate of 2.44e-07, this allows the model to fine tune in order to obtain the ideal values.

The analysis of results also included comparison with several standard ML methods, namely, the Logistic Regression, K-Nearest Neighbors, and Multinomial Naive Bayes. The comparison was made using accuracy on the training set and on the test set, as depicted in the figure 11 below.

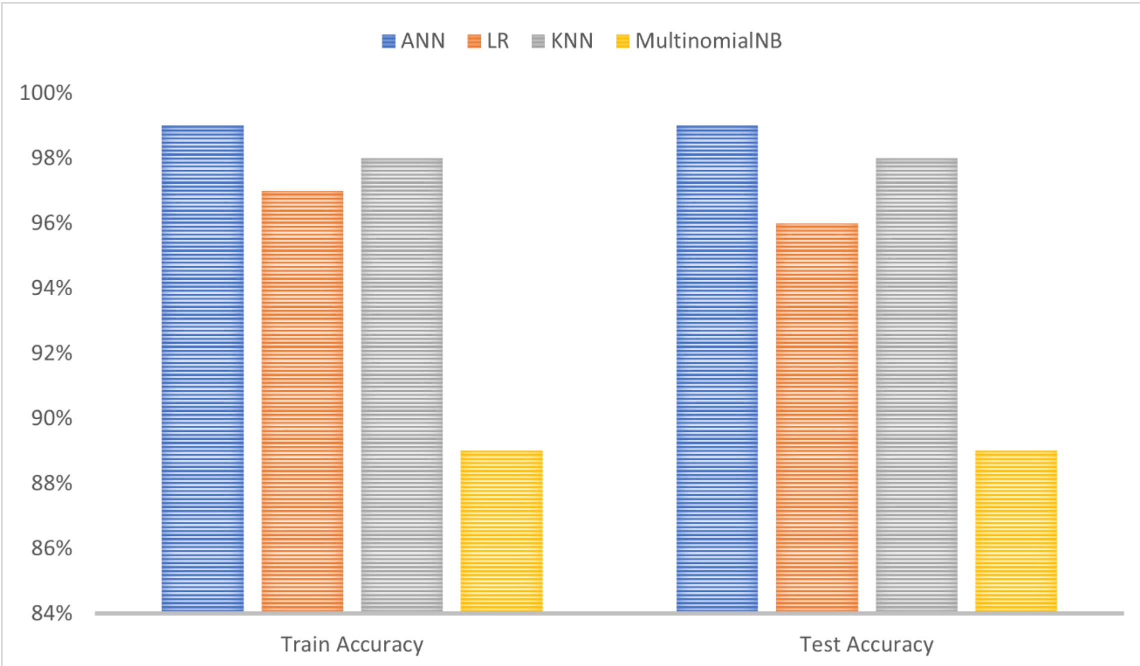

Fig. 11. Comparison Results of ANN Model with Traditional ML Methods.

5. CONCLUSION

This paper provides a sound method for improving crop advice and agricultural sustainability using ANNs and big data analysis. With the help of exact data and sets of input variables that include Nitrogen, Phosphorus, Potassium, Temperature, Humidity, pH value, and Rainfall the ANN model proposed in this paper has enough potential to transform agricultural practices.

The ANN model came out with an accuracy rate of 99% which proves that the model is efficient in giving accurate crop recommendation. Such high accuracy level means that this model is capable of handling relations that are present in the data and could be helpful for farmers and anyone who has to make decisions regarding agriculture. The measures included the confusion matrix which showed that the model had a high capability of classifying the crops correctly hence reducing misclassification.

The paper also discusses data preparation steps like missing value handling, duplicate data removal, and label encoding since these steps significantly affect the quality of the data and the model. Pair plots and correlation matrices also proved useful in feature engineering in an attempt to refine the understanding of the relationships between different variables so as to come up with a better predictive model.

In the comparison with other traditional machine learning algorithms like Logistic Regression, K-Nearest Neighbors, and Multinomial Naive Bayes, ANN model outperforms all of them. This is an indication of the ANN's high capacity for handling complex and non-linear data as opposed to conventional approaches to crop recommendation.

Apart from enhancing the yield and quality of crops, the ANN model of recommendation contributes to the optimum utilization of resources, low expenses, and the protection of the environment. Due to improved nutrient utilization and reduced nutrient pollution, the approach encourages sustainable use of soil resources for production hence increased production.

Furthermore, the incorporation of climate considerations into the decision process helps the farmers to effectively manage climate change, thus improving crop yields and reducing vulnerability to climate shocks. This proactive adaptation is specifically highly important within the scope of the global climate change, contributing to food security and sustainability of the agricultural systems.

Therefore, this work proves the significance of enhancing existing machine learning approaches, especially ANNs, with detailed soil and environmental data, thus enabling a more effective, secure, and sustainable future of the agricultural industry. The findings and approaches of this research can serve as a basis for a continuous advancement of agriculture, enabling farmers to use science-based information for the optimization of yields, revenues, and impacts.

Conflicts Of Interest

The author's disclosure statement confirms the absence of any conflicts of interest.

Funding

No financial contributions or endorsements from institutions or sponsors are mentioned in the author's paper.

Acknowledgment

The author’s acknowledges the support and resources provided by the institution in facilitating the execution of this study.

References

[1] A. Ghosh, D. Chakraborty, and A. Law, “Artificial intelligence in Internet of things,” *CAAI Trans. Intell. Technol.*, vol. 3, no. 4, pp. 208–218, 2018, doi: https://doi.org/10.1049/trit.2018.1008.

[2] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Kojima, “Big Data Analytics, Machine Learning, and Artificial Intelligence in Next-Generation Wireless Networks,” *IEEE Access*, vol. 6, pp. 32328–32338, 2018, doi: 10.1109/ACCESS.2018.2837692.

[3] A. Alamoodi *et al.*, “Evaluating agriculture 4.0 decision support systems based on hyperbolic fuzzy-weighted zero-inconsistency combined with combinative distance-based assessment,” *Comput. Electron. Agric.*, vol. 227, p. 109618, 2024, doi: https://doi.org/10.1016/j.compag.2024.109618.

[4] A. H. Alamoodi, M. S. Al-Samarraay, O. S. Albahri, M. Deveci, A. S. Albahri, and S. Yussof, “Evaluation of energy economic optimization models using multi-criteria decision-making approach,” *Expert Syst. Appl.*, vol. 255, p. 124842, 2024, doi: 10.1016/j.eswa.2024.124842.

[5] O. S. Albahri *et al.*, “Selection of smartphone-based mobile applications for obesity management using an interval neutrosophic vague decision-making framework,” *Eng. Appl. Artif. Intell.*, vol. 137, p. 109191, 2024, doi: https://doi.org/10.1016/j.engappai.2024.109191.

[6] Y. L. Khaleel, M. A. Habeeb, and B. Rabab, “Emerging Trends in Applying Artificial Intelligence to Monkeypox Disease: A Bibliometric Analysis,” *Appl. Data Sci. Anal.*, vol. 2024, pp. 148–164, 2024, doi: 10.58496/ADSA/2024/012.

[7] R. D. Ismail, Q. A. Hameed, M. A. Habeeb, Y. L. Khaleel, and F. N. Ameen, “Deep Learning Model for Hand Movement Rehabilitation,” *Mesopotamian J. Comput. Sci.*, vol. 2024, no. SE-Articles, pp. 134–149, Oct. 2024, doi: 10.58496/MJCSC/2024/011.

[8] M. E. Alqaysi, A. S. Albahri, and R. A. Hamid, “Evaluation and benchmarking of hybrid machine learning models for autism spectrum disorder diagnosis using a 2-tuple linguistic neutrosophic fuzzy sets-based decision-making model,” *Neural Comput. Appl.*, 2024, doi: 10.1007/s00521-024-09905-6.

[9] L. Cao, “AI in Finance: Challenges, Techniques, and Opportunities,” *ACM Comput. Surv.*, vol. 55, no. 3, Feb. 2022, doi: 10.1145/3502289.

[10] H. M. Abdulfattah, K. Y. Layth, and A. A. Raheem, “Enhancing Security and Performance in Vehicular Adhoc Networks: A Machine Learning Approach to Combat Adversarial Attacks,” *Mesopotamian J. Comput. Sci.*, vol. 2024, pp. 122–133, 2024, doi: 10.58496/MJCSC/2024/010.

[11] A. S. Albahri *et al.*, “A systematic review of trustworthy artificial intelligence applications in natural disasters,” *Comput. Electr. Eng.*, vol. 118, p. 109409, 2024, doi: 10.1016/j.compeleceng.2024.109409.

[12] Y. L. Khaleel, “Fake News Detection Using Deep Learning,” University of Miskolc, 2021. doi: http://dx.doi.org/10.13140/RG.2.2.31151.75689.

[13] M. A. Habeeb, “Hate Speech Detection using Deep Learning Master thesis,” University of Miskolc, 2021. [Online]. Available: http://midra.uni-miskolc.hu/document/40792/38399.pdf

[14] R. Cioffi, M. Travaglioni, G. Piscitelli, A. Petrillo, and F. De Felice, “Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions,” *Sustain.*, vol. 12, no. 2, 2020, doi: 10.3390/su12020492.

[15] M. A. Habeeb, Y. L. Khaleel, and A. S. Albahri, “Toward Smart Bicycle Safety: Leveraging Machine Learning Models and Optimal Lighting Solutions,” in *Proceedings of the Third International Conference on Innovations in Computing Research (ICR’24)*, K. Daimi and A. Al Sadoon, Eds., Cham: Springer Nature Switzerland, 2024, pp. 120–131.

[16] V. Dohale, M. Akarte, A. Gunasekaran, and P. Verma, “Exploring the role of artificial intelligence in building production resilience: learnings from the COVID-19 pandemic,” *Int. J. Prod. Res.*, vol. 62, no. 15, pp. 5472–5488, 2024.

[17] A. H. Alamoodi *et al.*, “A Novel Evaluation Framework for Medical LLMs: Combining Fuzzy Logic and MCDM

for Medical Relation and Clinical Concept Extraction," *J. Med. Syst.*, vol. 48, no. 1, p. 81, 2024, doi: 10.1007/s10916-024-02090-y.

[18] M. A. Fadhel *et al.*, "Navigating the metaverse: unraveling the impact of artificial intelligence—a comprehensive review and gap analysis," *Artif. Intell. Rev.*, vol. 57, no. 10, p. 264, 2024, doi: 10.1007/s10462-024-10881-5.

[19] L. Alzubaidi *et al.*, "Comprehensive review of deep learning in orthopaedics: Applications, challenges, trustworthiness, and fusion," *Artif. Intell. Med.*, vol. 155, p. 102935, 2024, doi: https://doi.org/10.1016/j.artmed.2024.102935.

[20] A. S. Albahri, Y. L. Khaleel, and M. A. Habeeb, "The Considerations of Trustworthy AI Components in Generative AI; A Letter to Editor," *Appl. Data Sci. Anal.*, vol. 2023, pp. 108–109, 2023, doi: 10.58496/adsa/2023/009.

[21] Z. T. Al-qaysi, A. S. Albahri, M. A. Ahmed, and M. M. Salih, "Dynamic decision-making framework for benchmarking brain–computer interface applications: a fuzzy-weighted zero-inconsistency method for consistent weights and VIKOR for stable rank," *Neural Comput. Appl.*, vol. 36, no. 17, pp. 10355–10378, 2024, doi: 10.1007/s00521-024-09605-1.

[22] A. S. Albahri *et al.*, "Prioritizing complex health levels beyond autism triage using fuzzy multi-criteria decision-making," *Complex Intell. Syst.*, 2024, doi: 10.1007/s40747-024-01432-0.

[23] G. G. Shayea *et al.*, "Fuzzy Evaluation and Benchmarking Framework for Robust Machine Learning Model in Real-Time Autism Triage Applications," *Int. J. Comput. Intell. Syst.*, vol. 17, no. 1, 2024, doi: 10.1007/s44196-024-00543-3.

[24] F. K. H. Mihna, M. A. Habeeb, Y. L. Khaleel, Y. H. Ali, and L. A. E. Al-Saeedi, "Using Information Technology for Comprehensive Analysis and Prediction in Forensic Evidence," *Mesopotamian J. CyberSecurity*, vol. 4, no. 1, pp. 4–16, 2024, doi: 10.58496/MJCS/2024/002.

[25] S. Dadvandipour and Y. L. Khaleel, "Application of deep learning algorithms detecting fake and correct textual or verbal news," *Prod. Syst. Inf. Eng.*, vol. 10, no. 2, pp. 37–51, 2022, doi: 10.32968/psaie.2022.2.4.

[26] Y. L. Khaleel, M. A. Habeeb, A. S. Albahri, T. Al-Quraishi, O. S. Albahri, and A. H. Alamoodi, "Network and cybersecurity applications of defense in adversarial attacks: A state-of-the-art using machine learning and deep learning methods," *J. Intell. Syst.*, vol. 33, no. 1, 2024, doi: 10.1515/jisys-2024-0153.

[27] R. S. Abdul Ameer, M. A. Ahmed, Z. T. Al-Qaysi, M. M. Salih, and M. L. Shuwandy, "Empowering Communication: A Deep Learning Framework for Arabic Sign Language Recognition with an Attention Mechanism," *Computers*, vol. 13, no. 6, p. 153, 2024, doi: 10.3390/computers13060153.

[28] L. A. E. Al-saeedi *et al.*, "Artificial Intelligence and Cybersecurity in Face Sale Contracts: Legal Issues and Frameworks ," *Mesopotamian J. CyberSecurity*, vol. 4, no. 2 SE-Articles, pp. 129–142, Aug. 2024, doi: 10.58496/MJCS/2024/0012.

[29] Y. L. Khaleel, H. M. Abdulfattah, and H. Alnabulsi, "Adversarial Attacks in Machine Learning: Key Insights and Defense Approaches," *Appl. Data Sci. Anal.*, vol. 2024, pp. 121–147, 2024, doi: 10.58496/ADSA/2024/011.

[30] L.-J. Gardiner and R. Krishna, "Bluster or Lustre: Can AI Improve Crops and Plant Health?," *Plants*, vol. 10, no. 12, 2021, doi: 10.3390/plants10122707.

[31] Y. L. Khaleel, M. A. Habeeb, and G. G. Shayea, "Integrating Image Data Fusion and ResNet Method for Accurate Fish Freshness Classification," *Iraqi J. Comput. Sci. Math.*, vol. 5, no. 4, p. 21, 2024.

[32] S. O. Araújo, R. S. Peres, J. Barata, F. Lidon, and J. C. Ramalho, "Characterising the agriculture 4.0 landscape—emerging trends, challenges and opportunities," *Agronomy*, vol. 11, no. 4, p. 667, 2021.

[33] H. F. Williamson and S. Leonelli, "Accelerating agriculture: Data-intensive plant breeding and the use of genetic gain as an indicator for agricultural research and development," *Stud. Hist. Philos. Sci.*, vol. 95, pp. 167–176, 2022.

[34] E. Oleńska, W. Małek, M. Wójcik, I. Swiecicka, S. Thijs, and J. Vangronsveld, "Beneficial features of plant growth-promoting rhizobacteria for improving plant growth and health in challenging conditions: A methodical review," *Sci. Total Environ.*, vol. 743, p. 140682, 2020.

[35] T. Zenda, S. Liu, A. Dong, and H. Duan, "Revisiting sulphur—The once neglected nutrient: It's roles in plant growth, metabolism, stress tolerance and crop production," *Agriculture*, vol. 11, no. 7, p. 626, 2021.

[36] D. Kapoor, S. Bhardwaj, M. Landi, A. Sharma, M. Ramakrishnan, and A. Sharma, "The impact of drought in plant metabolism: How to exploit tolerance mechanisms to increase crop production," *Appl. Sci.*, vol. 10, no. 16, p. 5692, 2020.

[37] A. Grzyb, A. Wolna-Maruwka, and A. Niewiadomska, "Environmental factors affecting the mineralization of crop residues," *Agronomy*, vol. 10, no. 12, p. 1951, 2020.

[38] M. Rut, "ICT mediated food sharing and food transitions towards sustainable food systems in Singapore." Trinity College Dublin, 2021.

[39] S. Maurya, J. S. Abraham, S. Somasundaram, R. Toteja, R. Gupta, and S. Makhija, "Indicators for assessment of soil quality: a mini-review," *Environ. Monit. Assess.*, vol. 192, pp. 1–22, 2020.

[40] S. M. YAHAYA, A. A. MAHMUD, M. ABDULLAHI, and A. HARUNA, "Recent advances in the chemistry of nitrogen, phosphorus and potassium as fertilizers in soil: a review," *Pedosphere*, vol. 33, no. 3, pp. 385–406, 2023.

[41] M. Kijowska-Strugała, A. Baran, M. Szara-Bąk, Ł. Wiejaczka, and P. Prokop, "Soil quality under different agricultural land uses as evaluated by chemical, geochemical and ecological indicators in mountains with high rainfall (Darjeeling Himalayas, India)," *J. Soils Sediments*, vol. 22, no. 12, pp. 3041–3058, 2022.

[42] Z. Feng, L. Wang, Q. Peng, J. Li, and T. Liang, "Effect of environmental factors on soil properties under different land use types in a typical basin of the North China Plain," *J. Clean. Prod.*, vol. 344, p. 131084, 2022.

[43] F. Samaei, H. Emami, and A. Lakzian, "Assessing soil quality of pasture and agriculture land uses in Shandiz county, northwestern Iran," *Ecol. Indic.*, vol. 139, p. 108974, 2022.

[44] A. Gomez-Zavaglia, J. C. Mejuto, and J. Simal-Gandara, "Mitigation of emerging implications of climate change on food production systems," *Food Res. Int.*, vol. 134, p. 109256, 2020, doi: 10.1016/j.foodres.2020.109256.

[45] M. G. Muluneh, "Impact of climate change on biodiversity and food security: a global perspective—a review article," *Agric. Food Secur.*, vol. 10, no. 1, pp. 1–25, 2021.

[46] C. M. Godde, D. Mason-D'Croz, D. E. Mayberry, P. K. Thornton, and M. Herrero, "Impacts of climate change on the livestock food supply chain; a review of the evidence," *Glob. Food Sec.*, vol. 28, p. 100488, 2021.

[47] S. H.-K. Tui *et al.*, "Climate change impacts and adaptation for dryland farming systems in Zimbabwe: a stakeholder-driven integrated multi-model assessment," *Clim. Change*, vol. 168, no. 1, p. 10, 2021.

[48] L. Murken and C. Gornott, "The importance of different land tenure systems for farmers' response to climate change: A systematic review," *Clim. Risk Manag.*, vol. 35, p. 100419, 2022.

[49] K. U. Singh *et al.*, "An Artificial Neural Network-Based Pest Identification and Control in Smart Agriculture Using Wireless Sensor Networks," *J. Food Qual.*, vol. 2022, 2022, doi: 10.1155/2022/5801206.

[50] K. Batra and P. Gandhi, "Artificial Neural Network-Based Model for the Prediction of Evaporation in Agriculture," *Lect. Notes Electr. Eng.*, vol. 836, pp. 561 – 569, 2022, doi: 10.1007/978-981-16-8542-2_46.

[51] D. Popescu, A. Dinca, L. Ichim, and N. Angelescu, "New trends in detection of harmful insects and pests in modern agriculture using artificial neural networks. a review," *Front. Plant Sci.*, vol. 14, 2023, doi: 10.3389/fpls.2023.1268167.

[52] M. K. Mohanty, P. K. G. Thakurta, and S. Kar, "Efficient sensor node localization in precision agriculture: an ANN based framework," *OPSEARCH*, 2023, doi: 10.1007/s12597-023-00625-4.

[53] G. Kamyshova *et al.*, "Artificial neural networks and computer vision's-based phytoindication systems for variable rate irrigation improving," *IEEE Access*, vol. 10, pp. 8577–8589, 2022.

[54] E. Khosla, R. Dharavath, and R. Priya, "Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression," *Environ. Dev. Sustain.*, vol. 22, no. 6, pp. 5687–5708, 2020.

[55] Y. Ma, Z. Zhang, Y. Kang, and M. Özdoğan, "Corn yield prediction and uncertainty analysis based on remotely sensed variables using a Bayesian neural network approach," *Remote Sens. Environ.*, vol. 259, p. 112408, 2021.

[56] M. Maimaitijiang, V. Sagan, P. Sidike, S. Hartling, F. Esposito, and F. B. Fritschi, "Soybean yield prediction from UAV using multimodal data fusion and deep learning," *Remote Sens. Environ.*, vol. 237, p. 111599, 2020.

[57] H. Yalcin, "An approximation for a relative crop yield estimate from field images using deep learning," in *2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, IEEE, 2019, pp. 1–6.

[58] "Crop Recommendation Dataset." https://www.kaggle.com/datasets/varshitanalluri/crop-recommendation-dataset/code (accessed Jun. 19, 2024).