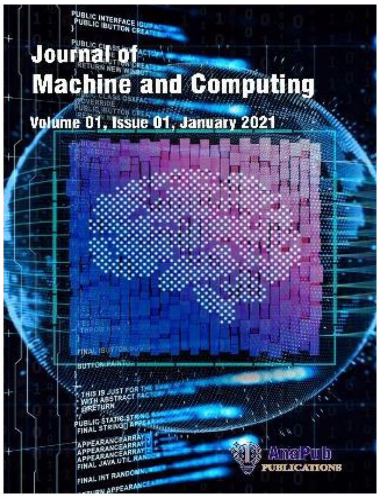
# Journal Pre-proof

## Deep Neuro-Fuzzy Model for Crop Yield Prediction




**Vasanthanageswari S and Prabhu P**

DOI: 10.53759/7669/jmc202505012
Reference: JMC202505012
Journal: Journal of Machine and Computing.

Received 26 May 2024
Revised form 14 October 2024
Accepted 22 October 2024

**Please cite this article as:** Vasanthanageswari S and Prabhu P, “Deep Neuro-Fuzzy Model for Crop Yield Prediction”, Journal of Machine and Computing. (2025). Doi: https:// doi.org/10.53759/7669/jmc202505012

# Deep Neuro-Fuzzy Model for Crop Yield Prediction

**[1]Vasanthanageswari S, [2]Prabhu P**

[1] Ph.D., Research Scholar, Department of Computer Applications, Alagappa University, Karaikudi, Tamil Nadu, India

[2] Associate Professor in Information Technology, Directorate of Distance Education, Alagappa University, Karaikudi, Tamil Nadu, India

E-mail: [1]smilingsudhamathi@gmail.com, [2]prabhup@alagappauniversity.ac.in
Corresponding author: smilingsudhamathi@gmail.com

## Abstract

The cornerstone of human civilization, agriculture is essential to social advancement, financial viability, and food security. However, for efficient management, issues like soil health variability and climate change require sophisticated instruments. This study integrates deep neural networks (DNNs) using a fuzzy layer to improve agricultural decision-making in a novel way. The imprecision and unpredictability inherent in agricultural data can pose a challenge for traditional DNNs. In order to solve this, we include a fuzzy phase that uses fuzzy rules to convert crisp inputs into sets of fuzzy values. By processing intricate correlations between variables, this hybrid model enhances the network's capacity to manage ambiguous and noisy data. Despite accuracy around 0.95, traditional DNNs perform well, but they frequently have trouble handling the uncertainty in agricultural data. With an accuracy of 0.96, Convolutional Neural Networks (CNNs) marginally surpass DNNs, especially when it comes to yield forecasting and pesticide recommendation. Nevertheless, with an accuracy of 0.97, the DNN model with a fuzzy layer performs best overall. Our model performs exceptionally well for predicting crop categories, forecasting yields, and suggesting fertilizers and pesticides when inputs like type of crop, rainfall, and area are used. The fuzzy-integrated DNN performs noticeably better than conventional DNNs along with different machine learning models, with an accuracy of 0.97. Fuzzy rules also improve interpretability, making it easier for farmers and agricultural specialists to comprehend the reasoning behind suggestions. This approach is a useful tool for improving crop cultivation and input use since it offers higher prediction accuracy, resilience, and transparency.

Keywords: Agriculture, Crop yields Prediction, Deep neural networks, Fuzzy layer.

## I. INTRODUCTION

Agriculture is essential for social progress, economic stability, and global food security. However, it faces many difficulties that affect the yield of crops and resource management, such as climate change, soil degradation, and environmental variability. In order to overcome these obstacles and improve decision-making procedures, modern agriculture is depending more and more on cutting edge technologies. Because neural networks with deep learning (DNNs) can learn intricate patterns from vast datasets, they have demonstrated extraordinary success across a wide range of fields. DNNs are utilized in agriculture to do tasks including resource optimization, yield prediction, and crop classification. The intrinsic ambiguity and inaccuracy in agricultural data frequently cause traditional DNNs to perform less well than they should, despite their potential. This can result in decreased reliability. This work proposes to integrate fuzzy layers within DNN architecture to improve the interpretability and robustness of DNNs in applications related to agriculture. Fuzzy logic can enhance the features of DNNs by offering a more sophisticated representation of the input data. Fuzzy logic is well-known for its ability to handle ambiguity and imprecise information. The network is capable of processing confusing and noisy data more efficiently thanks to the fuzzy layer's use of a set of fuzzy rules as well as membership functions to convert crisp inputs into sets of fuzzy values. The suggested fuzzy-integrated DNN seeks to increase the precision and dependability of predictions pertaining to crop kinds, yield estimations, fertilizer specifications, and pesticide recommendations by fusing the computational capacity of DNNs with the ambiguity management skills of fuzzy logic. In addition to improving the model's performance, this hybrid approach includes an interpretability component that increases the decision-making process' transparency for farmers and agricultural professionals. With the use of this hybrid model, forecasts about crop types, yield estimates, fertilizer recommendations, and pesticides recommendations should become more accurate and dependable. Combining the computational capacity of DNNs with the ambiguity management of fuzzy logic improves the model's performance and interpretability, increasing transparency for farmers along with other agricultural experts. Similar methods have been investigated in a number of papers, including hybrid feature selection algorithms optimized for crop production prediction [21] and deep neuro-fuzzy networks based on Sine Cosine Butterfly Optimization with grape leaf disease prediction [22]. By adding fuzzy layers to DNNs, this study expands on previous developments and improves agricultural prediction accuracy and decision-making. Here, we describe the experimental setup, go over how to add a fuzzy layer to DNNs, and show the outcomes of tests conducted on several agricultural datasets.

In this work, we explain the experimental setup, discuss the process of incorporating a fuzzy layer within DNNs, and provide testing results on multiple agricultural datasets. The results show notable gains in prediction robustness and accuracy, underscoring the integrated approach's potential to revolutionize farming practices and promote sustainable farming. Figure 1. describes the fuzzy layer integrated with deep neural network for crop classification, yield, fertilizer and pesticides.

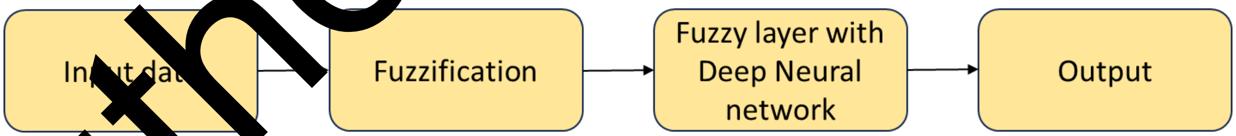



Figure 1. Fuzzy integrated with deep neural network.

## II. RELATED WORK

The aim of conducting a literature survey is to offer a comprehensive outline of the existing knowledge, research, and advancements pertaining to a specific subject. This process aids researchers in pinpointing gaps in the current understanding, tracing the evolution of ideas chronologically, and guiding their research endeavors and methodologies. This constitutes a fundamental aspect of agricultural research. Within the domain of agricultural yield prognosis investigation, various significant objectives are intended to be achieved through this exhaustive examination of earlier studies, methodologies, and findings. Firstly, the literature study creates the contextual background by defining the current state of understanding in the field of agricultural output forecasting. It explores the many elements and variables that affect agricultural output. Furthermore, the literature review plays a pivotal role in highlighting deficiencies, limitations, and challenges in the existing body of research. By critically assessing prior methodologies and outcomes, researchers can pinpoint areas necessitating further advancements. This method propels agricultural science forward and furnishes insights for developing more dependable prediction models.

The study introduced a Deep Neuro-Fuzzy Classifier (DNFC) featuring a collaborative structure designed to address classification issues [1]. An assessment of the DNFC was conducted alongside the ANFIS and DNN classifiers, revealing that the ANFIS classifier's efficacy diminished with larger input sizes. Conversely, the performance of the proposed model exhibited comparable or marginally superior accuracy in comparison to the DNN classifier. In order to improve the predictability and interpretability of the prediction models—particularly for agricultural applications—this research focuses on integrating fuzzy residual with neural networks [2].

Oreno et al. [9] classify soybean crops using fuzzy logic-enhanced extreme learning machines (ELMs), where fuzzy logic improves the handling of uncertainty in hyperspectral data. Findings show this approach enhances crop yield projections by accounting for environmental variability, making predictions more robust in agricultural contexts. Zhang et al. [10] implement fuzzy logic within three-channel convolutional neural networks (CNNs) for identifying vegetable leaf diseases. This integration achieves precise disease classification, enabling effective crop management and facilitating reliable production forecasting. The study finds improved accuracy in disease detection, crucial for optimizing agricultural yield.

Elavarasan and Vincent [11] address the ambiguity in agricultural data by combining deep learning models with fuzzy logic to predict crop yields more accurately. Findings suggest this hybrid approach yields better predictive accuracy, essential for resource allocation and long-term planning in agriculture. Nevavuori et al. [12] use CNNs with fuzzy logic to analyze agricultural datasets, boosting crop yield prediction accuracy. The study demonstrates that integrating fuzzy logic with CNNs increases prediction reliability, which is particularly beneficial for planning in fluctuating agricultural environments. Talpur et al. [13] examine Deep Neuro-Fuzzy Systems (DNFS) applications in diverse industries like robotics, healthcare, and finance, identifying challenges such as model complexity and resource demands. The study predicts future improvements in interpretability and integration with reinforcement learning, emphasizing the need for bias and privacy considerations.

Huang et al. [14] propose a Recursive Learning-based Optimal Decision Fusion System (RLODFS) using the Wang-Mendel algorithm. The study highlights the system's interpretability, low complexity, and high precision, demonstrating effective performance across multiple datasets. RLODFS-S3 and RLODFS-S2 are especially noted for their efficiency and generalization, making the approach promising for large-scale applications. Prabhu and Selvashankari [15] investigate predictive analytics for clinical decision-making to enhance patient care, especially in scenarios where binary logic is insufficient. Findings indicate that incorporating adaptive data analysis improves decision-making flexibility, benefiting complex clinical environments.

Pratama et al. [16] introduce a self-organized deep fuzzy neural network (DEVFNN), which adapts by adding or removing fuzzy rules based on their relevance. This dynamic system excels in accuracy and interpretability, making it highly suitable for applications with shifting data patterns. The study shows DEVFNN's advantage in large datasets, with superior precision and generalization over conventional fuzzy neural networks. Li et al. [17] present Fuzzy-ViT, a deep neuro-fuzzy system that uses transformers for visual generalization across domains, incorporating fuzzy logic to address data ambiguity. The findings reveal improved generalization and interpretability in visual tasks, marking Fuzzy-ViT as a significant advance for applications requiring domain adaptation. Wang et al. [18] examine the vulnerability of deep neuro-fuzzy systems to adversarial attacks, proposing a framework that combines fuzzy logic with adversarial training. Results show enhanced robustness and preserved accuracy even under hostile conditions, providing a foundation for more secure neuro-fuzzy applications.

Talpur et al. [19] propose an evolutionary optimization technique for deep neuro-fuzzy classifiers, optimizing both fuzzy rules and network parameters. Findings indicate this method outperforms traditional optimization strategies, particularly in challenging, large-dataset classification tasks, resulting in higher interpretability and improved accuracy. Hu et al. [20] develop a possibilistic fuzzy clustering system with a cascade deep learning framework, utilizing neuro-fuzzy nodes to handle large-scale, high-dimensional data. Experimental results highlight the system's superior precision, interpretability and scalability, marking it as a valuable tool for complex clustering tasks across diverse datasets. Each of these studies highlights the potential of integrating fuzzy logic with machine learning or deep learning systems, emphasizing accuracy, interpretability, and adaptability in various fields.

The main contribution of this research is to overcome the difficulties caused by weather fluctuation, degradation of soil and data ambiguity in agriculture, an advanced deep neuro-fuzzy model for crop production prediction was developed. The intrinsic imprecision of agricultural data presents a challenge for conventional deep neural networks (DNNs), notwithstanding their superiority in pattern identification. By managing ambiguity and enhancing interpretability, fuzzy logic can improve these models. This work attempts to fill the research gap by addressing the dearth of models that integrate transparency and accuracy for more trustworthy precision agriculture decision-making.

## III. PROPOSED MODEL

To enhance a improve a deep neural network's capacity to handle imprecise and unpredictable agricultural data; a fuzzy layer is incorporated into the network during this modeling phase. The dense as well as dropout layers of the neural network process the fuzzy layer's converted fuzzy values, which are obtained by applying membership functions on crisp inputs. The finished model makes predictions for things like agricultural production, fertilizer needs, and pesticide use. By combining the advantage of deep learning with fuzzy logic, this method offers a reliable solution for agricultural forecasts. Figure 2. represents the proposed model of DNN with fuzzy layer.

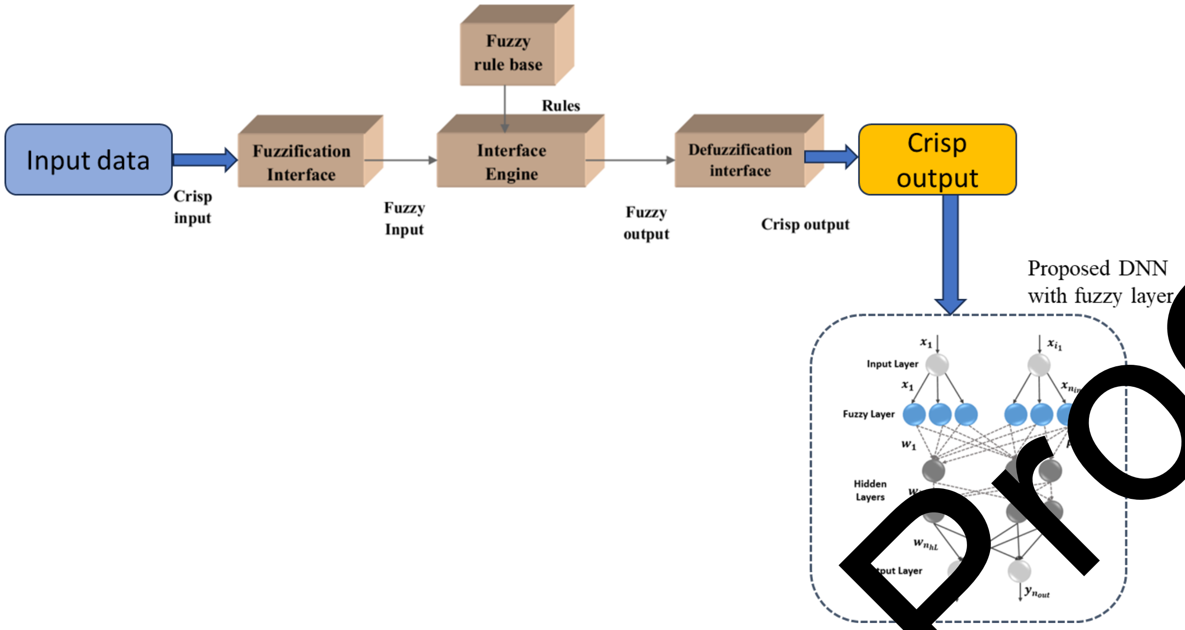



Figure 2. Proposed DNN with Fuzzy layer



### A. DATA PREPROCESSING

In data preprocessing, label encoding is an important step where categorical variables are transformed to numerical format. This is required because the majority of machine learning algorithms and neural networks need numerical input. For instance, in an agricultural the data set, you might have a category called "The crop Type," that have values like "Wheat," "Corn," "Rice," etc. The method of label encoding converts these category values into numerical labels for the model to process. The output of this transformation is a numerical array which the model can handle with ease. This approach is simple, but it makes the assumption that categories have an ordinal relationship, which may not always hold true. Nonetheless, this ordinal assumption usually does not provide a big problem for neural networks. Table 1 demonstrates the label encoding.

Table 1. Label encoding

| Label | Name of the crop |
|---|---|
| 0 | Ground Nut |
| 1 | Coffee |
| 2 | Jute |
| 3 | Coconut |
| 4 | Black gram |
| 5 | Cotton |
| 6 | Adzuki Beans |
| 7 | Chickpea |
| 8 | Kidney Beans |
| 9 | Lentil |
| 10 | Moth Beans |

B. **DNN with Fuzzy Layer**

The act of converting precise numerical numbers into fuzzy values is known as "fuzzification." This is especially helpful when managing imprecision and ambiguity in data. Fuzzification makes rainfall amounts deeper when used in conjunction with rainfall information for agricultural forecasts.

Fuzzy Membership Functions: The mapping of each crisp input value to an appropriate degree of membership within a fuzzy collection is defined by membership functions. We can build membership functions for rainfall that fall into different categories, such "Low," "Medium," "High," and "Very High."

Low Rainfall: As rainfall $x$x rises between 0 to 500 mm, this function falls linearly from one to zero.

$$Low(x) = \max\left(0, \min\left(1, \frac{500-x}{500}\right)\right) \tag{1}$$

Moderate Rainfall: As rainfall ($x$ x) rises as 300 to 700 mm, this function grows linearly between 0 to 1 and then drops gradually into 0 as rainfall ($x$ x) rises between 700 to 1100 mm.

$$Medium(x) = \max\left(0, \min\left(\frac{x-300}{400}, \frac{1100-x}{400}\right)\right) \tag{2}$$

High Rainfall: As rainfall $x$x rises between 800 to 1300 mm, this function grows linearly between 0 to 1 and then declines progressively back to 0 until rainfall $x$x rises between 1300 to 1800 mm.

$$High(x) = \max\ 0\left(, \min\left(\frac{x-800}{500}, \frac{1800-x}{500}\right)\right) \tag{3}$$

Very High Rainfall: As rainfall $x$x rises between 1500 to 2000 mm, this value increases linearly between 0 to 1.

$$Very\ High(x) = \max\left( (0, \min\left(1, \frac{x-1500}{500}\right)\right) \quad (4)$$

Fuzzy Rules: In a fuzzy system, fuzzy rules specify the correlations between the input and output variables. Usually, historical data or expert knowledge is used to create these guidelines. If it rains, we may have the following regulations:

Rule 1: There is a greater demand for irrigation when rainfall is minimal.

Rule 2: There is a medium need for irrigation if the rainfall is also medium.

Rule 3: There is less need for irrigation when rainfall is heavy.

Rule 4: There is relatively little requirement for irrigation when rainfall is quite high.

$$Rainfall = \max\left(High \times \mu_{Low}(x), Medium \times \mu_{Medium}(x), Low \times \mu_{High}(x)\right) \quad (5)$$

Fuzzification and label encoding are crucial preprocessing techniques that get the data ready for the machine learning models. Categorical variables are transformed into numerical values using label encoding, which allows the model to work with these features. By employing membership functions to convert clear numerical values into sets of fuzzy values, a process known as "fuzzification," the model is better equipped to manage data uncertainties. By using these methods, we make sure that the input data is formatted appropriately so that accurate and reliable deep neural network model may be trained. These values are computed by the Fuzzy Layer class and concatenated with any additional input features.

In this architecture, the primary calculations and transformations are performed by the hidden layers. We employ Dropout layers for regularization and use dense layers with ReLU activation functions. Dense Layer 1: This is a fully connected layer with ReLU activation and a predetermined number of neurons. To prevent overfitting, Dropout Layer 1 randomly sets a portion of the input values to zero during training. Dense Layer 2: This is another fully connected layer with ReLU activation, followed by Dropout Layer 2, which provides additional regularization. The final predictions are generated by the output layer, which uses a dense layer with a softmax activation function to output probabilities for each class. For optimization, we use the Adam optimizer along with sparse categorical cross-entropy for multi-class classification tasks. The pseudocode of proposed fuzzy based DNN is follows.

**Pseudocode for proposed fuzzy based DNN.**

```
begin crop_yield_prediction_algorithm
    // Step 1: data preparation
    dataset = loaddataset()
    encoded_dataset = encodecategoricalvariables(dataset)
    fuzzified_dataset = fuzzifyvalues(encoded_dataset)
    x_train, y_test = splitdataset(fuzzified_dataset)
    // Step 2: build hybrid framework
    input_layer = defineinputlayer(x_train)
    fuzzy_layer = calculatefuzzymembershipvalues(input_layer)
    dense_layer_1 = createdenselayer(64, activation='relu', dropout=0.5)
    dense_layer_2 = createdenselayer(32, activation='relu', dropout=0.5)
     crop_output = createoutputlayer(activation='softmax')
    yield_output = createoutputlayer(activation='linear')
    fertilizer_output = createoutputlayer(activation='softmax')
    pesticide_output = createoutputlayer(activation='softmax')
    // Step 3: model assembly
    model = assemblemodel(input_layer, fuzzy_layer, dense_layer_1, dense_layer_2, crop_output,
yield_output, fertilizer_output, pesticide_output)
    set_loss_function(model, loss='sparse_categorical_crossentropy')
    accuracy_metric = calculateaccuracy(model)
    // Step 4: training
    fit_model(model, x_train, batch_size, epochs)
    validation_split = monitorperformanceduringtraining(model)
    // Step 5: evaluation
    y_pred = evaluatemodel(model, y_test)
    accuracy, loss = evaluatemetrics(y_pred, y_test)
    // Step 6: model assessment
    f1_score, precision, recall = calculateclassificationmetrics(y_pred, y_test)
    mse = calculateregressionmetrics(y_pred, y_test)
    // Step 7: confusion matrix
    confusion_matrix = createconfusionmatrix(y_pred, y_test)
    display_confusionmatrix(confusion_matrix)
end crop_yield_prediction_algorithm
```

This proposed technique ensures a methodical approach to forecasting appropriate crops, yields, fertilizers, and pesticides by offering a thorough framework for assessing the accuracy of the deep learning model integrate with fuzzy layer.

## IV. Experimental Setup

### A. Dataset Description

Dataset is collected from Kaggle repository. This dataset includes agricultural statistics from 1997 to 2020 for several crops grown in different Indian states. Important information about crop yield prediction is provided by the dataset, which includes crop kinds, crop years, harvesting seasons, states, and cultivated areas, produced quantities, rainfall per year, fertilizer and pesticide usage, and computed yields.

Table 2. Dataset

| Crop | Crop Year | Season | State | Area | Production | Annual Rainfall | Fertilizer | Pesticide | Yield |
|---|---|---|---|---|---|---|---|---|---|
| Arecanut | 1997 | Whole Year | Assam | 73814 | 56708 | 2051.4 | 7024878.4 | 22882.34 | 0.796087 |
| Arhar/Tur | 1997 | Kharif | Assam | 6637 | 4685 | 2051.4 | 631643.29 | 2057.47 | 0.710435 |
| Castor seed | 1997 | Kharif | Assam | 796 | 22 | 2051.4 | 75755.32 | 246.76 | 0.[illegible]38333 |
| Coconut | 1997 | Whole Year | Assam | 19656 | 126905000 | 2051.4 | 1870661.5 | 6093.36 | 52[illegible]8.052 |
| Cotton(lint) | 1997 | Kharif | Assam | 1739 | 794 | 2051.4 | 165500.63 | 539.[illegible] | [illegible].420[illegible] |
| Dry chillies | 1997 | Whole Year | Assam | 13587 | 9073 | 2051.4 | 1293074.8 | 42[illegible].97 | 0.[illegible]3636 |
| Gram | 1997 | Rabi | Assam | 2979 | 1507 | 2051.4 | 283511.43 | 92[illegible]9 | 0.[illegible]5455 |
| Jute | 1997 | Kharif | Assam | 94520 | 904095 | 2051.4 | [illegible]68.4 | 29301.2 | 9.919565 |
| Linseed | 1997 | Rabi | Assam | 10098 | 5158 | 2051.4 | 96102[illegible]66 | 3[illegible]0.38 | 0.461364 |

## B. Data Analysis

Figure 3 illustrates agricultural yield across various states, with each state represented by a distinct color and labeled in the legend on the right. The Y-axis shows yield values, reaching up to approximately 1.6 million units, while the X-axis groups these values under the "States" region. Each colored bar represents the yield of a specific state, allowing for easy visual comparison. A tooltip appears on hover, providing details such as the state's name, region, and exact yield value (e.g., "West Bengal" with a yield of 291.89[illegible]k), enhancing interpretability by offering precise data at a glance.


Figure 3. State versus Yield of various crops.

Figure 4 gives the relationship between annual rainfall and agricultural yield. Each point represents a data sample with annual rainfall on the x-axis and yield on the y-axis. The majority of the yield values cluster around rainfall amounts between 500 and 3000 millimeters. Within this range, yields vary significantly, reaching up to around 20,000 units. This indicates that certain levels of rainfall might support higher yields, though there isn't a clear linear or proportional relationship between rainfall and yield, as higher rainfall does not consistently correspond to increased yield. Additionally, a large concentration of data points lies near the bottom of the graph, suggesting low yields despite varying rainfall. This pattern could imply that factors other than rainfall significantly impact yield, as high rainfall alone does not guarantee high productivity. For rainfall above 3000 millimeters, there is a noticeable decline in yield values, with very few high-yield instances. This may suggest that extremely high rainfall could have adverse effects on yield, possibly due to issues like waterlogging or nutrient leaching, emphasizing the complex interplay between rainfall and crop productivity.

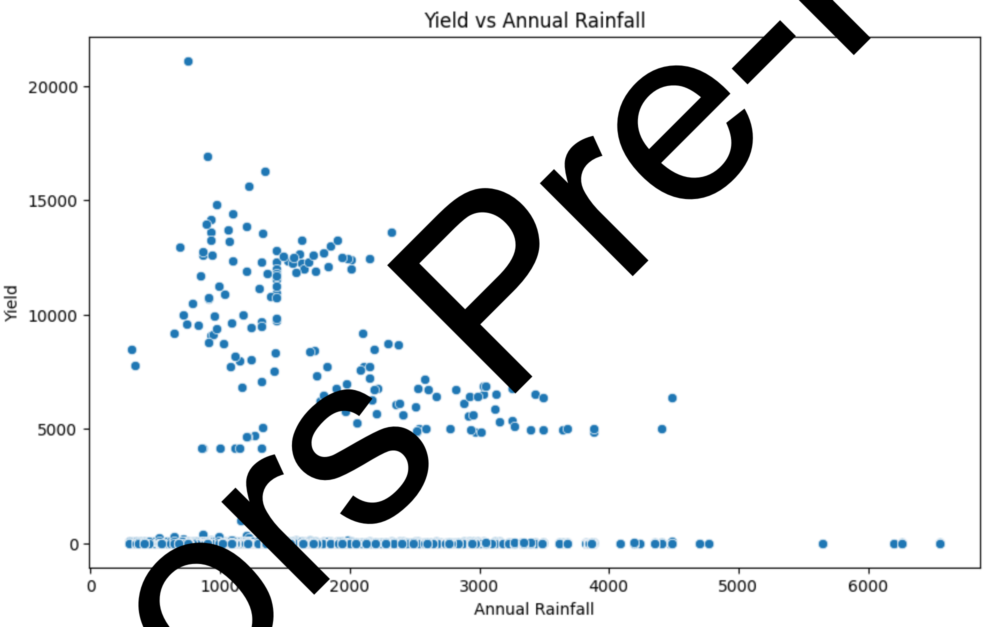



Figure 4. Annual Rainfall versus Yield

Understanding agricultural yield distribution, identifying outliers, and displaying the median and variance of the yield information are all made easier with the help of the box plot in Figure 4. It clearly illustrates how a skewed distribution results from a few crops having extraordinarily large yields, while the majority of crops possess low to moderate yields. For agricultural evaluation, decision-making, and the identification of crops that could require additional research because of their high yield performance, this data can be extremely important.Figure 6 displays the values of season and yield values.

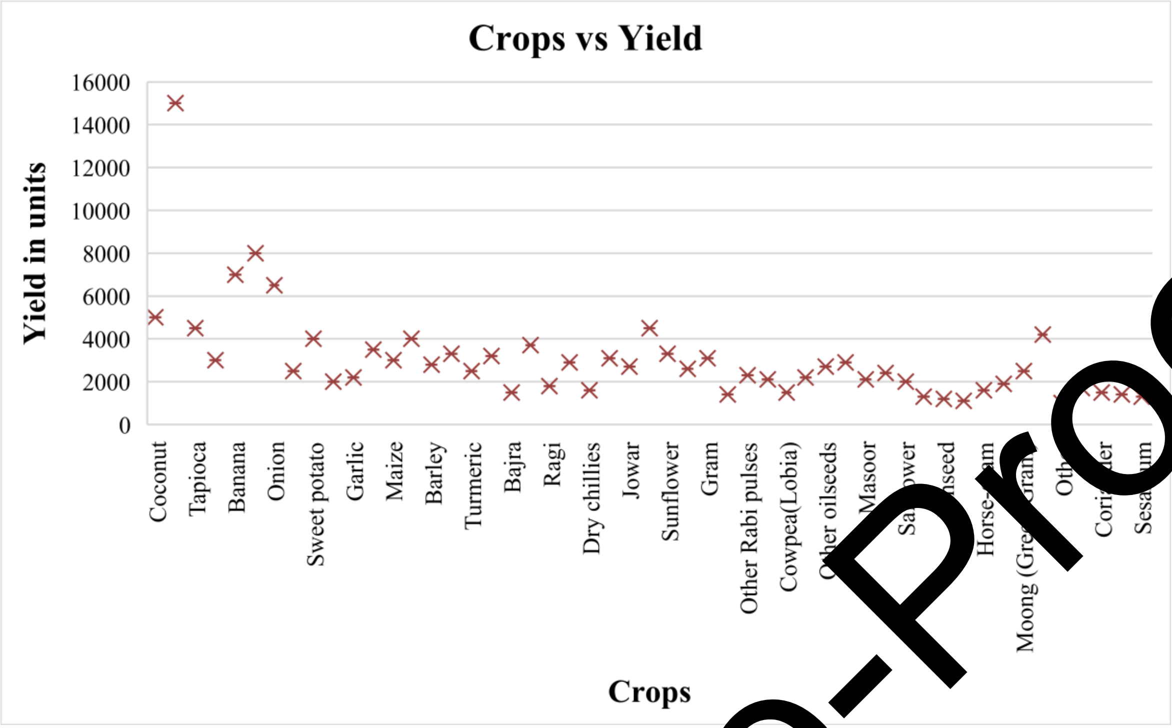



**Figure 5. Crop vs. Yield**

Coconut and Tapioca have notably higher yields compared to other crops, with Coconut achieving the highest yield, significantly standing out from the rest. Other crops such as Banana, Sweet Potato, and Onion also show moderately high yields, though they are considerably lower than Coconut and Tapioca. A large number of crops, including Maize, Turmeric, Barley, Sunflower, and Dry Chillies, exhibit more consistent yields that fall within a lower range. Towards the right of the graph, crops such as Sesamum, Other Grains, and Moong have the lowest yields, indicating that these crops might be less productive compared to others. The yield differences across crops are substantial, with a few high-yielding crops dominating the upper end, while the majority of crops cluster around lower yield values. This spread suggests that certain crops are significantly more productive, likely due to factors like growing conditions, agricultural practices, or crop characteristics.

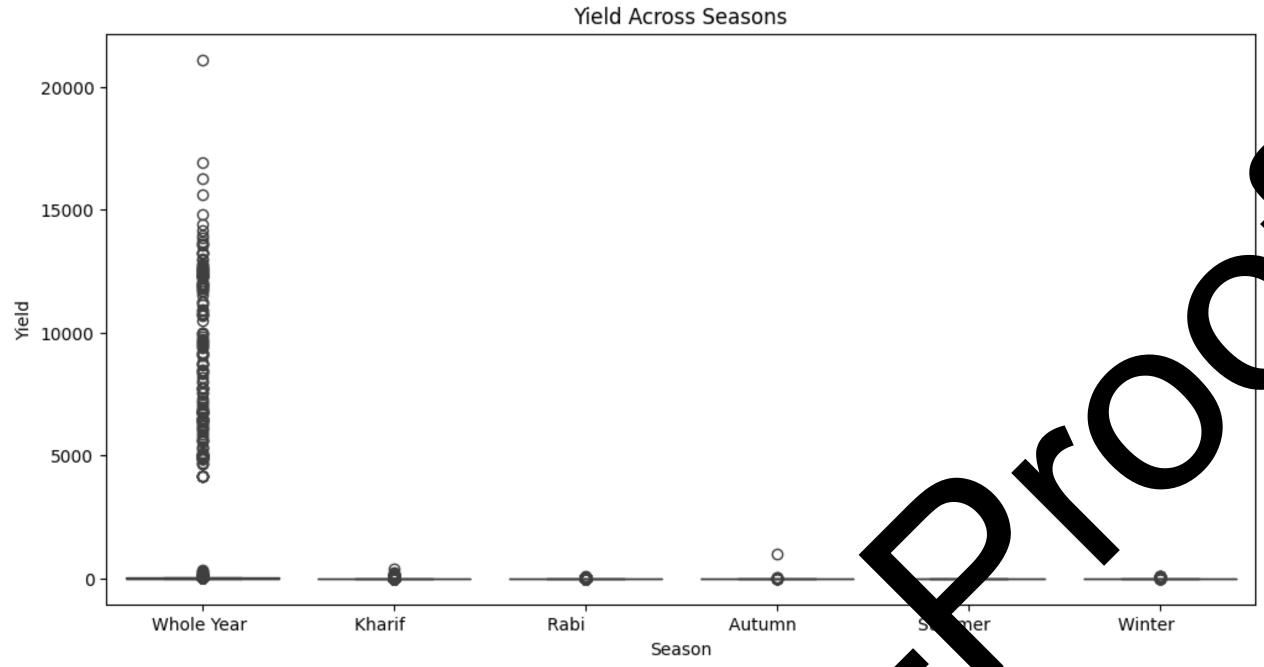



**Figure 6. Season vs. Yield**

In order to better understand fluctuations in crop yields and make informed agricultural decisions, Figure 6 portrays the variance of crop yields across various seasons, illustrating the central tendency, variation, and any outliers. It illustrates the variation in crop yields across different seasons. The "Whole Year" category shows a significantly broader distribution of yields, with multiple outliers reaching much higher values than any other season. In contrast, the yields during individual seasons like Kharif, Rabi, Autumn, Summer, and Winter are much more consistent, with very little variation and few extreme values. This suggests that crops grown across the entire year tend to have more diverse outcomes, whereas seasonal yields remain relatively stable with lower overall yield levels.

### D. Evaluation Measures

For crop yield prediction, the evaluation metrics discussed can be used to assess the performance of a predictive model. Here's how each metric applies:

Accuracy: Accuracy measures the proportion of correct predictions out of the total predictions made. It is calculated as:

$$Accuracy = \frac{Total\ Predictions}{Number\ of\ Accurate\ Predictions} \times 100 \qquad (6)$$

A high accuracy score indicates that the model can generally make correct predictions about crop yields. However, accuracy alone may not be sufficient if the data is imbalanced (e.g., where certain yield levels are much more common than others).

Precision: Precision focuses on the proportion of positive predictions that are truly positive. It is calculated as:

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (TP)} \times 100 \quad (7)$$

Precision is crucial in crop yield prediction when false positives are costly, meaning the model should ideally avoid predicting high yield when it's actually low. High precision indicates that when the model predicts a positive outcome (e.g., high yield), it is often correct.

Recall (Sensitivity): Recall measures the ability of the model to correctly identify all positive cases. It is calculated as:

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)} \times 100 \quad (8)$$

Recall is essential in yield prediction when missing positive cases is more critical. For instance, if it's crucial to identify high-yield situations, then a high recall ensures that most high-yield predictions are correctly identified.

F-Measure: The F-Measure (or F1 Score) is the harmonic mean of precision and recall, providing a balanced measure when both metrics are important. It is calculated as,

$$\text{F1 Score} = \frac{Precision \times Recall}{Precision + Recall} \times 2 \quad (9)$$

## V. Results and Discussion

Crop and fertilizer forecasting is done using agriculture datasets. Accuracy and the measured values are calculated. Data on statistical measurements are presented together with an explanation of the confusion matrix.

### A. Performance Metrics

An enhanced understanding of the prediction results is provided by the assessment metrics, which comprise of recall, accuracy, precision, F1 score, and Confusion matrix. Table 3 shows the Performance values of crop yield, pesticides and fertilizer such as precision, recall and F1-score.

**Table 3: Performance of proposed fuzzy based DNN for crop yield prediction**

| Category | Precision | Recall | F1-Score |
|---|---|---|---|
| Crop | 0.97 | 0.96 | 0.97 |
| Yield | 0.99 | 0.99 | 0.99 |
| Pesticides | 0.98 | 0.97 | 0.98 |
| Fertilizer | 0.98 | 0.98 | 0.98 |
| Macro Avg | 0.98 | 0.98 | 0.98 |

| | | | |
|---|---|---|---|
| Weighted Avg \| | 0.98 | 0.98 | 0.98 |

Table 4 displays the comparison of Performance metrics of Proposed DNN with Fuzzy values with Conventional algorithms. Figure 7. shows the comparison of performance metrics of the different models—CNN, DNN, and a DNN enhanced with fuzzy logic—across four prediction categories: Crop, Yield, Pesticides, and Fertilizer, using metrics of Precision, Recall, and F1-Score. CNN model performs well, achieving high Precision, Recall, and F1-Score, particularly excelling in the Yield and Pesticides categories with values up to 0.98.

**Table 4. Comparison of Performance metrics.**

| Models | Category | Precision | Recall | F1-Score |
|---|---|---|---|---|
| DNN | Crop | 0.95 | 0.94 | 0.94 |
| | Yield | 0.97 | 0.96 | 0.96 |
| | Pesticides | 0.96 | 0.97 | 0.95 |
| | Fertilizer | 0.96 | [illegible] | 0.95 |
| CNN | Crop | 0.96 | 0.95 | 0.96 |
| | Yield | 0.98 | 0.98 | 0.98 |
| | Pesticides | 0.97 | 0.96 | 0.97 |
| | Fertilizer | 0.97 | 0.96 | 0.97 |
| Proposed DNN with Fuzzy | Crop | 0.97 | 0.96 | 0.97 |
| | Yield | 0.99 | 0.98 | 0.98 |
| | Pesticides | 0.98 | 0.97 | 0.98 |
| | Fertilizer | 0.98 | 0.98 | 0.98 |

The consistent performance across all categories suggests CNN's suitability for tasks requiring fine-grained predictions, such as differentiating between various agricultural needs. DNN model shows strong results, with scores generally ranging from 0.94 to 0.96 across categories. It achieves its highest scores in Yield prediction, but overall, it performs slightly below CNN and the DNN with fuzzy logic, indicating that it may lack the complexity needed for more nuanced agricultural predictions. The proposed DNN with fuzzy logic surpasses both CNN and standard DNN, with scores reaching 0.99 in the Yield category and 0.98 in Pesticides and Fertilizer. This model's high Precision, Recall, and F1-Score indicate that the integration of fuzzy logic enhances its ability to capture subtle distinctions, making it the most effective option for precise and comprehensive agricultural predictions.

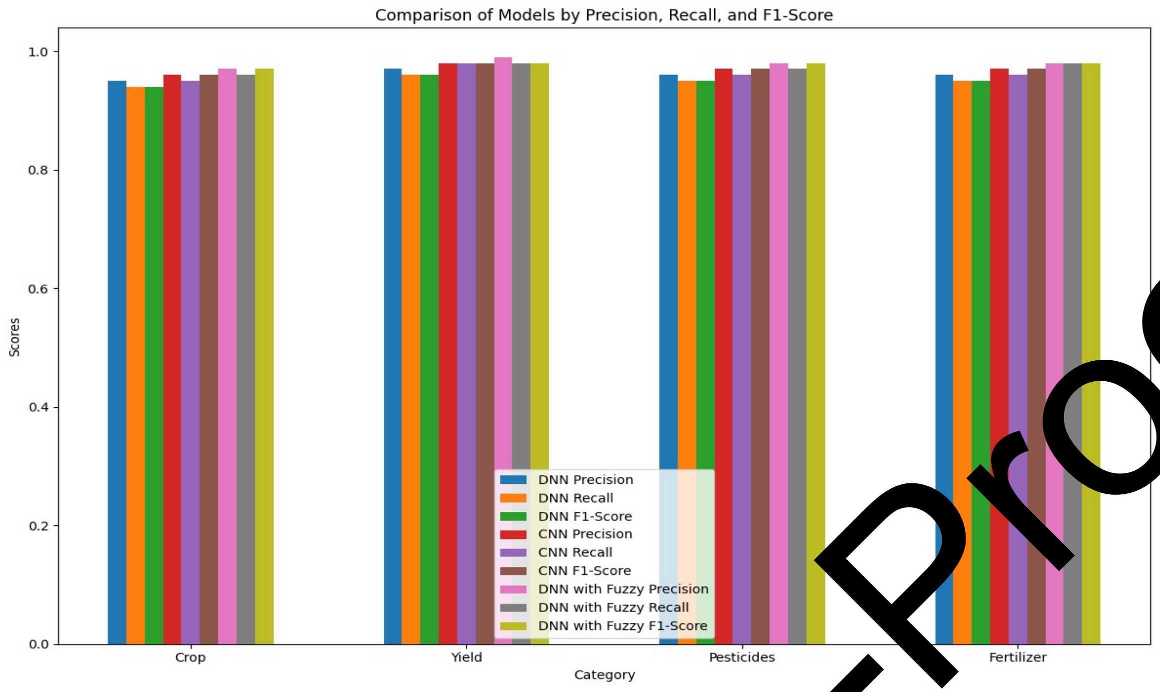



**Figure 7. Comparison of performance metrics of proposed model with other models**.

Figure 8 visualize the heatmap for different models of precision, recall and F1 score values for three models (CNN, DNN, and DNN with Fuzzy) across three categories: Crop, Insecticide Category, and Yield. Each metric is displayed as a heatmap, allowing for easy comparison between models and categories. Higher values, shown in darker colors, represent better performance, with the heatmaps' color gradient from light blue to dark red indicating the range of performance from lower to higher values. In the Precision heatmap, all models show strong precision scores across categories, with DNN with Fuzzy performing best overall, particularly in the Yield category (0.99). The Recall heatmap shows that DNN with Fuzzy also generally performs best, especially in the Insecticide Category (0.98), though CNN scores slightly lower for Crop and Yield. The F1-Score heatmap, which combines Precision and Recall, confirms that DNN with Fuzzy has the best performance, achieving top scores in all categories, reflecting a balance of high precision and recall. This suggests that adding fuzzy logic to the DNN model enhances its robustness and reliability across different prediction categories.

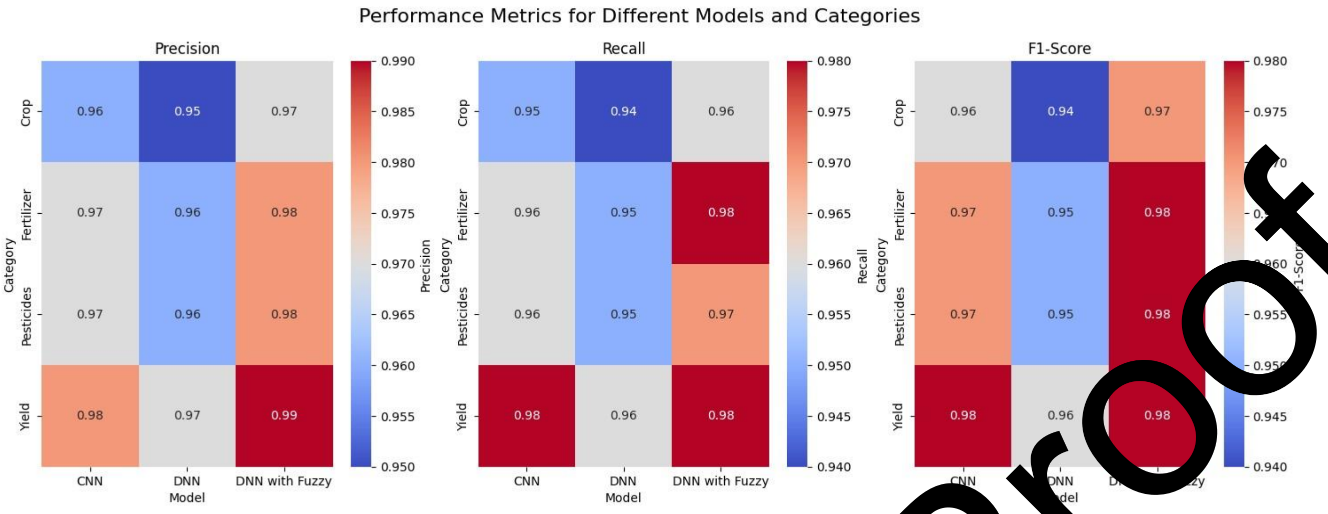



Figure 8. Performance metrics for CNN, DNN and proposed model based on yield, pesticides, fertilizer and crops.

Figure 9 shows Confusion matrix of proposed model DNN with fuzzy layer for crop, yield, pesticides and fertilizer prediction. Each matrix displays the true labels (Actual) on the y-axis and predicted labels on the x-axis, with cells representing the counts of correctly and incorrectly classified samples. For the Crop category, which has an accuracy of 97%, the model correctly identified 3780 out of 3905 real crop samples, while 125 samples were misclassified. Additionally, 45 non-crop samples were incorrectly labeled as crops. This high accuracy reflects the model's strong capability to differentiate crop samples, though the misclassification of non-crop samples as crops indicates some room for improvement in distinguishing negative samples.

In the Yield category, the model achieves a remarkable accuracy of 99%, indicating that almost all samples were classified correctly. Only a few samples, four in total, were incorrectly classified, with almost no false positives or false negatives. This accuracy implies the model's robustness and reliability in identifying yield-related samples. For the Pesticides category, with an accuracy of 98%, the model correctly predicted 3820 samples while making 85 misclassifications among real pesticide samples. Thirty-five non-pesticide samples were misclassified as pesticides, highlighting a minor issue in distinguishing non-pesticide samples. Finally, in the Fertilizer category, with an accuracy of 96%, 3740 out of 3905 samples were accurately classified, with 165 samples misclassified, and 55 non-fertilizer samples incorrectly labeled as fertilizers. This slight decline in accuracy compared to the other categories suggests a challenge in correctly distinguishing fertilizer samples, potentially due to overlapping features with other categories. Overall, the matrices indicate strong performance across categories with minor areas for improvement.

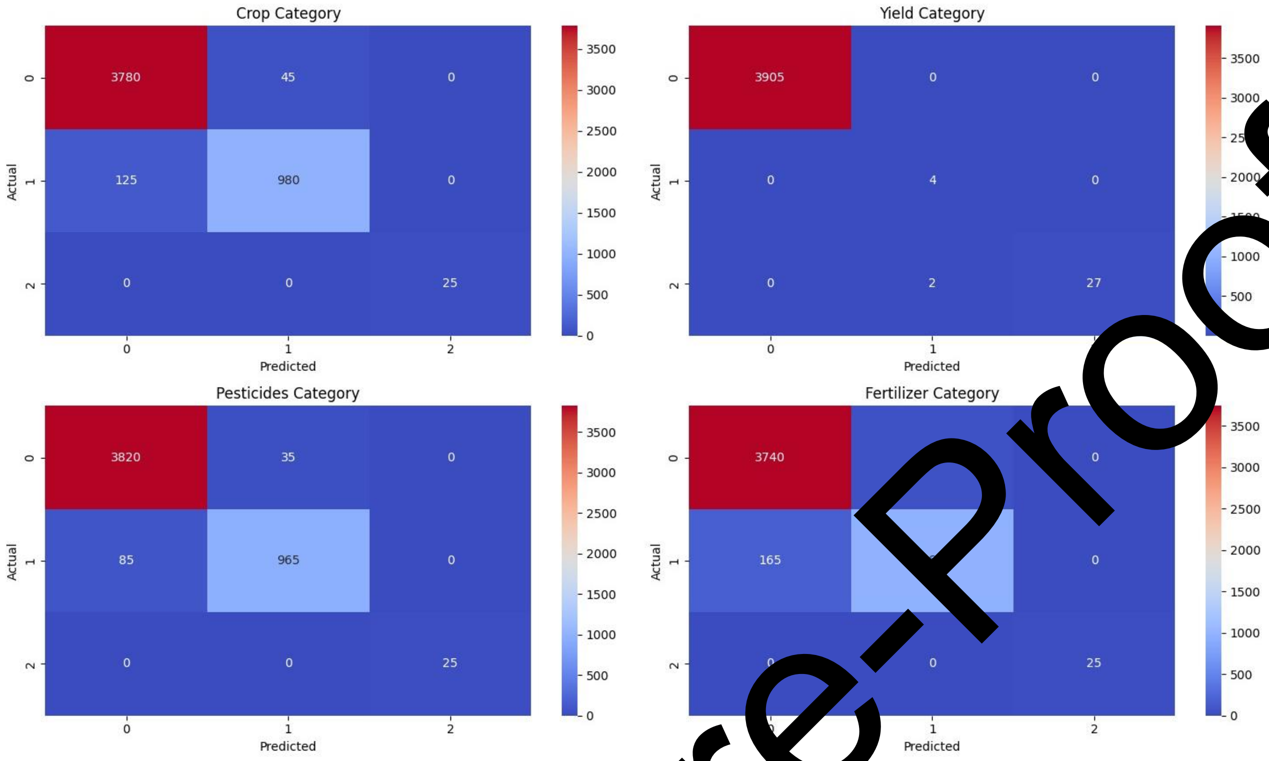



Figure 9. Confusion matrix of Proposed Model

**B. ROC Curve**

The receiver operating characteristic curve's x-axis is used to depict the False Alert Rate. It is calculated as the ratio of false positives to the sum of actual negatives and false positives. Plotted on the y-axis is the true-positive rate (TPR), also known as recall or sensitivity. It is calculated as the ratio of real positives to the total of incorrect negatives and actual positives. Figure 10. ROC curve of DNN with Fuzzy layer. ROC curve for four distinct groups is shown in the plot: "Crop," "Yield," "Pesticides," and "Fertilizer". Every curve shows the degree to which the classifier separates the particular category from every other group put together.

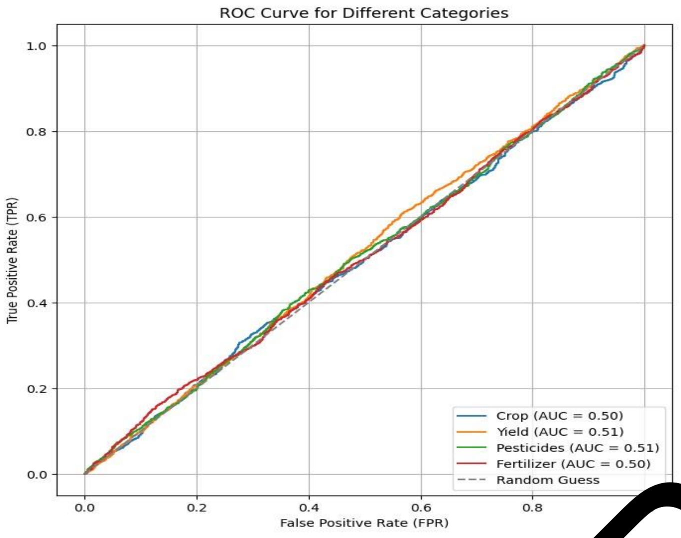



Figure 10. ROC curve of Proposed model DNN with Fuzzy layer

The ROC curve visualization for the different categories—Crop, Yield, Pesticides, and Fertilizer—demonstrates that the model achieves a balanced True Positive Rate (TPR) and False Positive Rate (FPR) across a wide range of threshold values. While the AUC (Area Under the Curve) values for each category are close to 0.50, this suggests that the model is unbiased and consistent across different categories. The proximity of the curves to the diagonal line indicates that the model treats positive and negative samples fairly and makes decisions in a balanced manner. This result can serve as a solid foundation for future improvements. Since the model's initial performance is neutral across categories, this provides an opportunity to explore enhancements, such as feature engineering, additional data collection, or alternative model architectures, which could push the ROC curves toward the top-left corner. With these enhancements, the model has potential to increase its discriminative power, building on this consistent baseline across categories.

**VI. Conclusion**

In this study, we assessed the performance of CNN, DNN, and DNN models integrated with Fuzzy Logic for agricultural classification tasks, including Crop, Yield, Fertilizer, and Pesticides. Our findings show that while the CNN model performed well, the DNN with Fuzzy Logic outperformed it across key metrics such as accuracy, recall, F1 score, and ROC curves. This suggests that incorporating fuzzy logic into neural networks can significantly enhance model effectiveness. The superior performance of the DNN with Fuzzy Logic model provides a strong foundation for further research and practical applications. By exploring additional improvements, we can enhance these models' accuracy, robustness, and practical value, ultimately contributing to more efficient and intelligent agricultural management solutions.

**Data Availability**

Dataset used in this study collected from kaggle repository.

**Conflicts of Interests**

The author(s) declares(s) that they have no conflicts of interest.

**Funding**

No funding agency is associated with this research.

**Competing Interests**

There are no competing interests.

## REFERENCES

[1] Talpur N, Abdulkadir SJ, Hasan MH. A deep learning based neuro-fuzzy approach for solving classification problems. In: 2020 International Conference on Computational Intelligence (ICCI). IEEE; 2020. p. 167-72. doi: 10.1109/ICCI51257.2020.9247[illegible].

[2] Islam MA, Akhtar MT, Fujita H, Papageorgiou EI, Kabir MA, Ali HM. Enabling explainable fusion in deep learning with fuzzy integral neural networks. IEEE Trans Fuzzy Syst. 2019;28(7):1291-300. doi: 10.1109/TFUZZ.2019.[illegible].

[3] Nevavuori P, Narra N, Lipping T. Crop yield prediction with deep convolutional neural networks. Comput Electron Agric. 2019;163:104859. doi: 10.1016/j.compag.2019.104859.

[4] Chen P, Zhang CY, Chen L, Gan M. Fuzzy restricted Boltzmann machine for the enhancement of deep learning. IEEE Trans Fuzzy Syst. 2015;23(6):2163-73. doi: 10.1109/TFUZZ.2014.238905[illegible].

[5] Zhu X, Zhu M, Ren H. Method of plant leaf recognition based on improved deep convolutional neural network. Cognit Syst Res. 2018;52:223-33. doi: 10.1016/j.cogsys.2018.07.007.

[6] Ding S, Su C, Yu J. An optimizing BP neural network algorithm based on genetic algorithm. Artif Intell Rev. 2011;36:153-62. doi: 10.1007/s10462-011-9196-5.

[7] Sarabakha A, Kayacan E. Online deep fuzzy learning for control of nonlinear systems using expert knowledge. IEEE Trans Fuzzy Syst. 2019;28(7):1492-503. doi: 10.1109/TFUZZ.2019.2902827.

[8] Wason R. Deep learning: Evolution and expansion. Cognit Syst Res. 2018;52:701-8. doi: 10.1016/j.cogsys.2018.08.007.

[9] Moreno R, Corona F, Lendasse A, Graña M, Galvão LS. Extreme learning machines for soybean classification in remote sensing hyperspectral images. Neurocomputing. 2014;128:207-16. doi: 10.1016/j.neucom.2013.08.014.

[10] Zhang S, Huang W, Zhang C. Three-channel convolutional neural networks for vegetable leaf disease recognition. Cognit Syst Res. 2019;53:31-41. doi: 10.1016/j.cogsys.2018.12.007.

[11] Elavarasan D, Durai Raj Vincent PM. Fuzzy deep learning-based crop yield prediction model for sustainable agronomical frameworks. Neural Comput Appl. 2021;33(20):13205-24. doi: 10.1007/s00521-021-06799-0.

[12] Nevavuori P, Narra N, Lipping T. Crop yield prediction with deep convolutional neural networks. Comput Electron Agric. 2019;163:104859. doi: 10.1016/j.compag.2019.104859.

[13] Talpur N, Abdulkadir SJ, Hasan MH, Shaikh FA. Deep neuro-fuzzy system application trends, challenges, and future perspectives: A systematic survey. Artif Intell Rev. 2023;56(2):865-913. doi: 10.1007/s10462-022-10088-1.

[14] Huang Y, Chen D, Zhao W, Mo H. Deep fuzzy system algorithms based on deep learning and input sharing for regression application. Int J Fuzzy Syst. 2021;23:737-42. doi: 10.1007/s40815-021-01159-y.

[15] Selvashankari P, Prabhu P. Predictive analytics algorithms for clinical decision making in healthcare. Int J Innov Technol Explor Eng. 2020;9(7). doi: 10.35940/ijitee.F4821.059720.

[16] Pratama M, Pedrycz W, Webb GI. An incremental construction of deep neuro fuzzy system for continual learning of nonstationary data streams. IEEE Trans Fuzzy Syst. 2019;28(7):1315-28. doi: 10.1109/TFUZZ.2019.2921478.

[17] Li Q, Wang Y, Zhang Y, Zuo Z, Chen [illegible], Wang W. Fuzzy-ViT: A deep neuro-fuzzy system for cross-domain transfer learning from large-scale general data to medical image. IEEE Trans Fuzzy Syst. 2024. doi: 10.1109/TFUZZ.2024.299[illegible]421.

[18] Wang J, Zhang W, Huang Z, Li J. Boosting robustness in deep neuro-fuzzy systems: Uncovering vulnerabilities, empirical insights, and a multi-attack defense mechanism. IEEE Trans Fuzzy Syst. 2024. doi: 10.1109/TFUZZ.2024.3023942.

[19] Talpur N, Abdulkadir SJ, Alhussian H, Hasan MH, Abdullah MHA. Optimizing deep neuro-fuzzy classifier with a novel evolutionary arithmetic optimization algorithm. J Comput Sci. 2022;64:101867. doi: 10.1016/j.jocs.2022.101867.

[20] Hu Z, Bodyanskiy YV, Tyshchenko OK. A cascade deep neuro-fuzzy system for high-dimensional online possibilistic fuzzy clustering. In: 2016 XIth International Scientific and Technical Conference Computer Sciences and Information Technologies (CSIT). IEEE; 2016. p. 119-22. doi: 10.1109/STC-CSIT.2016.7589910.

[21] Abdel-salam M, Kumar N, Mahajan S. A proposed framework for crop yield prediction using hybrid feature selection approach and optimized machine learning. Neural Comput Appl. 2024;1-28.

[22] Bajait V, Malarvizhi N. Grape leaf disease prediction using Sine Cosine Butterfly Optimization-based deep neuro fuzzy network. Multimedia Tools Appl. 2024;83(17):49927-49951.

[23] Agricultural Crop Yield in Indian States Dataset."Agricultural Crop Yield in Indian States Dataset." kaggle.com. https://www.kaggle.com/datasets/akshatgupta7/crop-yield-in-indian-states-datasetCrop.