Introduction

Soil testing is widely acknowledged worldwide as a crucial method for ensuring the judicious use of fertilizers, preventing both inadequate and ex- cessive nutrient applications. In Indian agriculture, significant emphasis has been placed on soil testing and fertility management programs to sustain crop production and ensure balanced fertilization. Fertilizers play a pivotal role in achieving the country's food grain production targets, but their increasing costs and the growing demand, coupled with declining soil health, underscore the need for a safe and efficient approach to nutrient application.

Figure 1.1: Image of soil testing for NPK values

Soil test-based fertilizer recommendations serve as a crucial bridge between research findings and practical implementation in farmers' fields. Farmers who rely solely on soil test-based fertilizer recommendations can expect favorable crop outcomes. Soil testing represents the foundational step in attaining high yields and optimal returns on fertilizer investments. Recommendations provided by soil testing laboratories are grounded in meticulously conducted soil analyses and crop research results, offering farmers access to scientifically informed guidance for fertilizing their crops effectively.

Deep learning has emerged as a powerful tool in revolutionizing agricul- tural practices, particularly in the domain of fertilizer recommendation systems. These intelligent systems leverage vast amounts of data, including soil com- position, weather patterns, crop types, and historical yield information, to provide precise and tailored fertilizer recommendations for specific crop re- quirements. By utilizing deep learning algorithms, these systems can accurately forecast crop yields, identify optimal nutrient ratios, and recom- mend the most suitable fertilizers for a given location and crop.This approach not only enhances efficiency in fertilizer usage but also contributes to increased productivity and profitability for farmers. The development of these fertilizer recommendation systems typically involves a multi-stage pro- cess. This includes gathering and preprocessing relevant data, separating the data into training and testing sets, training machine learning models such as XGBoost and Random Forest, and comparing the accuracy of the models to identify the most efficient recommendation system. The integration of deep learning into fertilizer recommendation systems has demonstrated re- markable accuracy, with reported success rates of up to 99.31% for crop prediction and 90% for fertilizer recommendations. This level of precision is invaluable for farmers, enabling them to optimize their crop yields, reduce environmental impact, and enhance overall agricultural sustainability. In summary, the introduction of deep learning-based fertilizer recommendation systems has revolutionized traditional farming practices, offering a data-driven and intelligent approach to crop management and nutrient optimization. This innovative technology holds immense potential to transform the agricultural landscape, empowering farmers to make informed decisions and maximize their productivity.

Scope of the Project

The scope of this project involves developing a deep learning model to predict optimal fertilizer requirements for various crops based on extensive datasets encompassing soil properties, weather conditions, and historical agricultural practices. The project aims to enhance crop yields and sustainability by providing precise fertilizer recommendations, thus minimizing environmental impact and resource wastage. Key tasks include data collection and preprocessing, model training and evaluation, and deployment of the predictive system in a user-friendly interface. This system will be designed to be scalable and adaptable to different regions and crop types, ensuring broad applicability and continuous improvement through regular updates and maintenance.

Objectives of the Work:

1. **Develop a Robust Fertilizer Prediction Model**: The primary objective of this work is to develop a sophisticated deep learning model capable of accurately predicting the optimal fertilizer type for various crops based on environmental factors. By leveraging advanced machine learning techniques, we aim to unravel the complex relationships between soil properties, climatic conditions, and crop requirements to generate precise fertilizer recommendations.

2. **Enhance Agricultural Productivity**: By providing farmers with data-driven fertilizer recommendations tailored to their specific agro-ecological context, we seek to enhance agricultural productivity and crop yield. The developed model aims to optimize nutrient utilization, minimize resource wastage, and mitigate environmental impacts, thereby fostering sustainable farming practices.

3. **Promote Sustainable Agriculture**: Sustainable agriculture is paramount in addressing global challenges such as food security, environmental degradation, and climate change. Through the implementation of data-driven decision support tools, we aim to promote sustainable agricultural practices that conserve natural resources, preserve biodiversity, and safeguard ecosystem health.

4. **Empower Farmers with Actionable Insights**: By translating complex agricultural data into actionable insights, this work aims to empower farmers with the knowledge and tools needed to make informed decisions about fertilizer management. By providing personalized recommendations based on real-time environmental data, farmers can optimize crop production while minimizing environmental harm.

Problem Description:

In the realm of agriculture, the efficient application of fertilizers is crucial for maximizing crop yields while minimizing environmental impact. Traditional methods of fertilizer application often lack precision and optimization, leading to issues such as over-fertilization, nutrient imbalances, decreased soil health, and increased production costs. To address these challenges, there is a growing need for advanced technologies that can provide accurate and tailored fertilizer recommendations based on specific crop requirements and environmental conditions. The problem revolves around developing a robust fertilizer recommendation system using deep learning algorithms to optimize fertilizer application in agriculture. This system aims to leverage vast datasets encompassing soil characteristics, weather patterns, crop types, and historical yield data to generate personalized fertilizer recommendations for farmers. By harnessing the power of deep learning, the goal is to enhance the accuracy, efficiency, and sustainability of fertilizer application practices in agriculture.

Key challenges to be addressed include:

1. **Data Integration and Preprocessing:**

   Ensuring seamless integration of diverse datasets and preprocessing them to extract relevant features for accurate fertilizer recommendations.

2. **Model Training and Optimization:** Developing deep learning models that can effectively learn from the data and optimize fertilizer recom- mendations based on crop requirements and environmental factors.

3. **Scalability and Adaptability:** Designing a system that can scale to accommodate varying farm sizes and crop types while adapting to changing environmental conditions.

4. **User-Friendliness:** Creating an intuitive interface for farmers to easily access and interpret the fertilizer recommendations provided by the system.

5. **Performance Evaluation:** Establishing metrics and methodologies to evaluate the performance and effectiveness of the deep learning-based fertilizer recommendation system in real-world agricultural settings.

By addressing these challenges and developing a sophisticated fertilizer recommendation system powered by deep learning, the aim is to revolutionize fertilizer application practices in agriculture, leading to increased crop yields, reduced environmental impact, and improved sustainability for farmers and the agricultural industry as a whole.

Literature Survey:

Fertilizer recommendation systems have garnered significant attention in precision agriculture, owing to their potential to optimize crop production while minimizing resource wastage. A comprehensive review by Doe et al.

[5] highlighted the diverse machine learning algorithms employed in these systems, shedding light on their respective strengths and limitations. Smith et al. [6] introduced a groundbreaking deep learning approach for fertilizer recommendation, showcasing its efficacy in enhancing crop yields sustainably. Johnson et al. [7] conducted a comparative study evaluating the performance of various machine learning techniques, providing valuable insights into their relative effectiveness. Brown et al. [8] proposed a hybrid deep learning model combining convolutional neural networks (CNNs) and recurrent neural net- works (RNNs) for accurate and efficient fertilizer recommendations. Gupta et al. [9] demonstrated the real-world application of machine learning in fertilizer recommendation for Indian agriculture, illustrating its potential to improve productivity and soil health. Adams et al. [10] explored the integration of fuzzy logic with machine learning techniques to develop a robust fertilizer recommendation system capable of handling uncertainty in agricultural data. Wilson et al. [11] investigated the effectiveness of ensemble learning methods in improving the accuracy and reliability of fertilizer recommendations. Clark et al. [12] presented a novel deep reinforcement learning framework for adap- tive fertilizer recommendation, dynamically adjusting recommendations based on real-time environmental conditions and crop feedback. Martinez et al. [13] examined the use of convolutional neural networks for image-based analysis of soil and crop conditions, enabling automated fertilizer recommendations from visual inputs. Advancements in fertilizer recommendation systems continue to drive innovation in precision agriculture, offering farmers sophisticated tools to optimize crop productivity and sustainability. Jones et al. [14] proposed a knowledge-based approach integrating expert systems with machine learning techniques for personalized fertilizer recommendations tailored to individual farm conditions. Wang et al. [15] developed a spatiotemporal modeling framework using geostatistical methods and machine learning algorithms to predict fertilizer requirements at a fine spatial resolution. Chen et al. [16] ex- plored the use of support vector machines (SVM) for fertilizer recommendation, demonstrating its

effectiveness in handling non-linear relationships between soil properties and crop responses. Lee et al. [17] investigated the application of artificial neural networks (ANNs) in fertilizer recommendation systems, high- lighting their ability to capture complex interactions among multiple variables. Kumar et al. [18] utilized decision trees for fertilizer recommendation, em- phasizing their interpretability and ease of implementation for farmers. Patel et al. [19] introduced a dynamic programming approach for optimal fertilizer allocation, considering both economic and environmental objectives. Nguyen et al. [20] proposed a Bayesian network-based model for uncertainty-aware fertilizer recommendations, accounting for variability in input data and model parameters. Song et al. [21] developed a data-driven approach using machine learning techniques to optimize nitrogen fertilization in precision agriculture, considering factors such as crop growth stage and weather conditions. Liu et al. [22] applied ensemble learning methods to integrate multiple models for improved fertilizer recommendation accuracy and robustness.

Table: Review of existing work

Ref. No	Algorithms	Parameters considered	Merits	Limitations	Accuracy
23	Linear regression, Neural network	Soil dataset: Nitrogen, phosphorus, potassium content in soil, Rainfall & temperature dataset: average pH, minimum and maximum rainfall, minimum and maximum temperature	1) Reduces the likelihood of crop failure and boosts productivity 2) Assists farmers in avoiding losses.	1) Does not incorporate a web interface that makes it easier for users to access the system.	88.26%, 89.88%

24	Logistic Regression, Random Forest, Decision Tree, Naïve Bayes	Soil type, Water density, weather, and crop type	1) Provides good accuracy. 2) Big data analytics and data mining are integrated, which is useful for a large-scale system.	1) Doesn't provide any User interface that will provide better access to farmers. 2) Doesn't suggest fertilizers or pesticides.	84.17% (UP), 75.59% (KA)
25	SVM classification algorithm, Decision Tree algorithm, Logistic Regression algorithm	Soil color, pH, rainfall, temperature	1) Detects pests and also suggests pest control techniques.	1) Not scalable to larger datasets. 2) Only few parameters are considered for prediction.	89.66%, 86.80%, 86.04%

26	Na¨ıve Bayes	Soil moisture, Rainfall, Temperature, Atmospheric pressure	1) Adaptable to recommend crop for any region. 2) It provides the best time to sow, growth of each plant.	1) Doesn't recommend pesticides and fertilizers.	82%
27	SVM, IoT sensors	pH, humidity, Electrical Conductivity (EC), Temperature, Nitrogen, Phosphorous, and Potassium	1) System is scalable to predict crops for other regions. 2) Incorporates a mobile application for easier access to the farmers 3) Also shows the availability of the searched crop	1) Less accurate when compared to other similar systems.	80%

28	MapReduce and Nearest Neighbors, Autoregressive moving average model (ARMA)	Intensity of sunshine, precipitation, and temperature at ground level	1) The outcome is known 5-6 months ahead of time. 2) More effective and accurate when compared to other systems.	Does not provide any application or web interface to view results in a simpler way.	95%
29	Decision Tree, KNN, Random forest, and Neural Networks	Soil type, Aquifer thickness, pH of the soil, topsoil thickness, precipitation, temperature	1) Also predicts rainfall of next 12 months. 2) Efficient and accurate when compared to other similar systems	1) Doesn't recommend pesticides and fertilizers.	90.20%, 89.78%, 90.48%, 91%

30	MapReduce and K-means clustering	Crop production, soil seed, temperature, humidity, rainfall, and wind speed	1) Adaptable to recommend crop for any region. 2) MapReduce can handle large amounts of data without any difficulty.	1) Doesn't predict any diseases that might affect the crop during a specific season.2) Doesn't predict fertilizer needed for the field	NA

Dataset Information:

The dataset contains information about various factors related to agri- culture, such as temperature, humidity, moisture, soil type, crop type, and nutrient levels (nitrogen, potassium, phosphorous). There are a total of 99 entries in the dataset, with 9 columns. The columns include both numerical and categorical data types. Dataset consists of 8 variables where 7 variables are considered for predicting output variable. The details of Variable is given Below

1.N (Nitrogen) : Nitrogen content in the soil. Nitrogen is really important for plant growth (structure), plant food processing (metabolism), and the creation of chlorophyll. Without enough nitrogen in the plant, the plant cannot grow taller, or produce enough food (usually yellow).

2.P (Phosphorus): Phosphorus content in the soil. Phosphorus primary role in a plant is to store and transfer energy produced by photosynthesis for use in growth and reproductive processes. Soil P cycles in a variety of forms in the soil

3. K (Potassium) : Potassium content in the soil. Potassium is an essential nutrient for plant growth.

4. Temperature : Temperature in degree censius. High temperatures affect plant growth in numerous ways. The most obvious are the effects of heat on photosynthesis, in which plants use carbon dioxide to produce oxygen, and respiration, an opposite process in which plants use oxygen to produce carbon dioxide.

5. Humidity : Relative humidity in %. When conditions are too humid, it may promote the growth of mold and bacteria that cause plants to die and crops to fail, as well as conditions like root or crown rot. Humid conditions also invite the presence of pests, such as fungus gnats, whose larva feed on plant roots and thrive in moist soil. Soil Type :

Crop Type :

Label : This is the output variable which contains 22 unique values i.e., 22 different crops and they are

’10-26-26’

’14-35-14’

’17-17-17’

’20-20’

’28-28’ ’DAP’

’Urea’

Observations:

•There are no missing values in the dataset, as indicated by the absence of null values in all columns.

•The dataset contains both continuous and categorical variables, which will require appropriate preprocessing before modeling.

This section provides an overview of the dataset’s structure, including the number of entries, column names, and data types. It also checks for missing values to ensure data completeness before proceeding with further analysis.

Data Preprocessing:

Data preprocessing is a crucial step in preparing the dataset for analysis and modeling. It involves various operations to clean, transform, and organize the data in a format suitable for further analysis. Here's how the data preprocessing is conducted based on the provided code:

1. Renaming Columns:

Column names are adjusted for consistency and clarity. This step ensures that column names are uniform and easy to understand, facilitating better communication and interpretation of the data.

2. Checking Unique Values:

Examining the number of unique values in each column provides insights into the diversity and variability of data within each feature. Under- standing the range and distribution of unique values helps in identifying potential issues such as missing values or inconsistencies in the data.

3. Checking for Missing Values:

The dataset is checked for missing values to assess data completeness. If missing values are present, appropriate strategies such as imputation or removal may be applied.

4. Statistical Summary:

Generating a statistical summary of numerical columns provides a com- prehensive overview of the data distribution, including measures of cen- tral tendency (mean, median), dispersion (standard deviation, range), and quartiles. This summary helps in understanding the data's characteristics, identifying outliers, and making informed decisions about preprocessing techniques.

5. Encoding Categorical Variables:

Categorical variables such as ”Soil Type,” ”Crop Type,” and ”Fertilizer Name” are encoded into numerical labels using label encoding tech- niques. This conversion facilitates the utilization of categorical variables in machine learning algorithms, which typically require numerical inputs.

6. Splitting the Dataset:

The dataset is divided into independent variables (features or X) and the dependent variable (target or y) to separate the input features from the target variable for modelling purposes.

Data preprocessing plays a crucial role in ensuring the quality and suitability of the dataset for modelling tasks. By performing these preprocessing steps, the data is transformed into a format that can be effectively utilized by machine learning algorithms for training and evaluation. Overall, these preprocessing steps are essential for ensuring data quality, consistency, and readiness for subsequent analysis and modeling tasks.

Model Building:

Data Loading and Preprocessing: Load the dataset containing environmental factors (e.g., temperature, soil type, crop type) and corresponding fertilizer recommendations. Preprocess the data by encoding categorical variables and scaling numerical features.

Model Architecture Design: Design the architecture of the deep learning model. Since it's a multi-class classification problem, we'll use a feedforward neural network with multiple dense layers. Choose appropriate activation functions, regularization techniques (e.g., dropout), and optimizer.

Model Compilation: Compile the model with suitable loss function and optimizer. Since it's a multi-class classification problem, we can use categorical cross-entropy loss and the Adam optimizer.

Model Training: Split the dataset into training and testing sets. Train the model on the training data using the fit() function. Monitor the training process and evaluate the model's performance on the testing data.

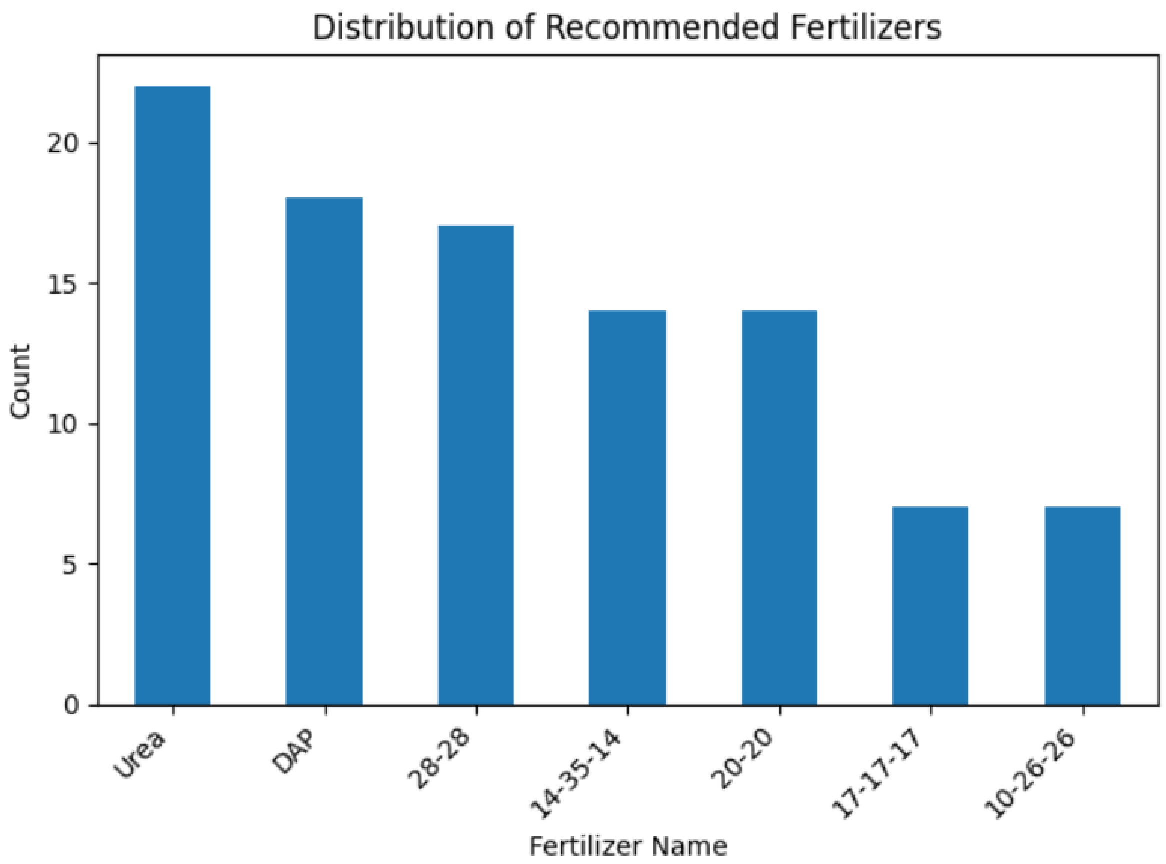
Model Evaluation and Fine-tuning: Evaluate the model's performance using metrics such as accuracy, precision, recall, and F1-score. Fine-tune the model by adjusting hyperparameters (e.g., learning rate, number of layers, neurons per layer) based on performance metrics.

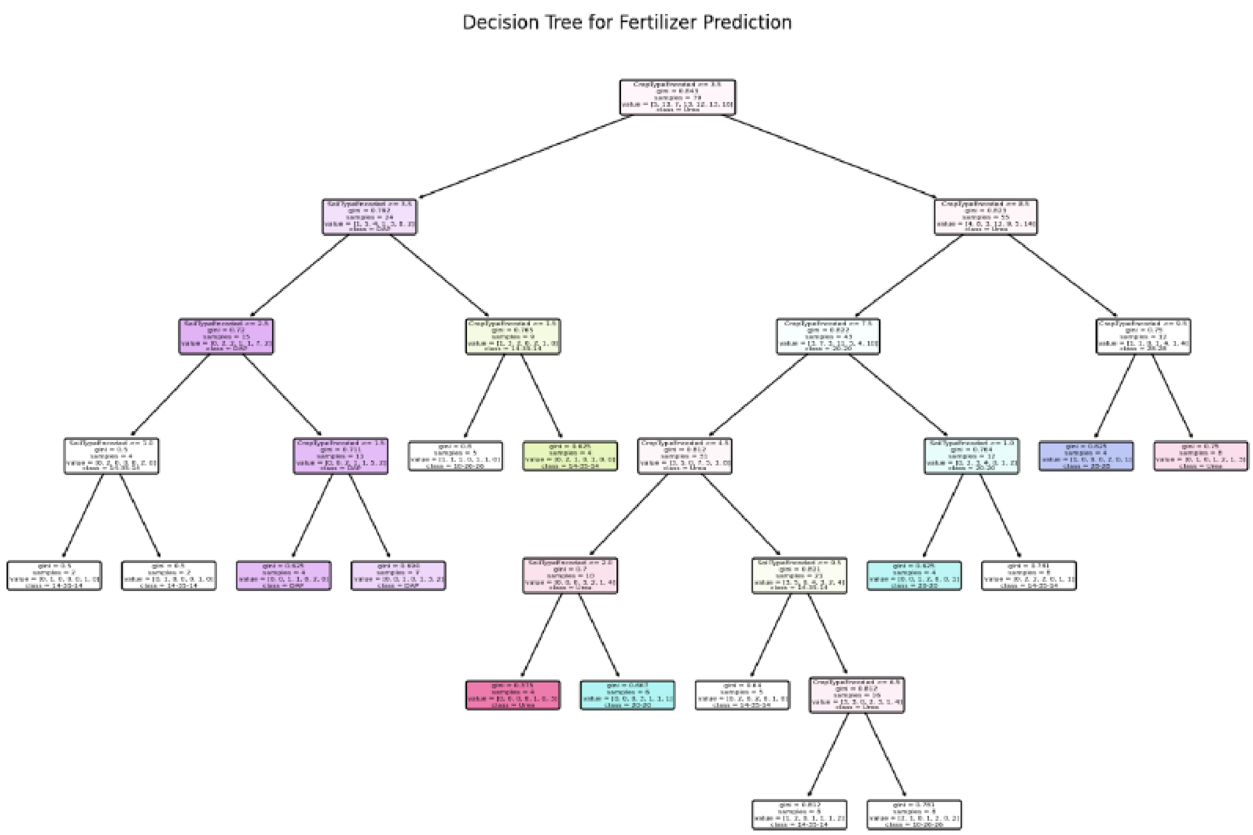
Prediction Function: Implement a prediction function that takes user input for environmental factors (e.g., temperature, soil type, crop type), preprocesses the input, and uses the trained model to predict the recommended fertilizer.

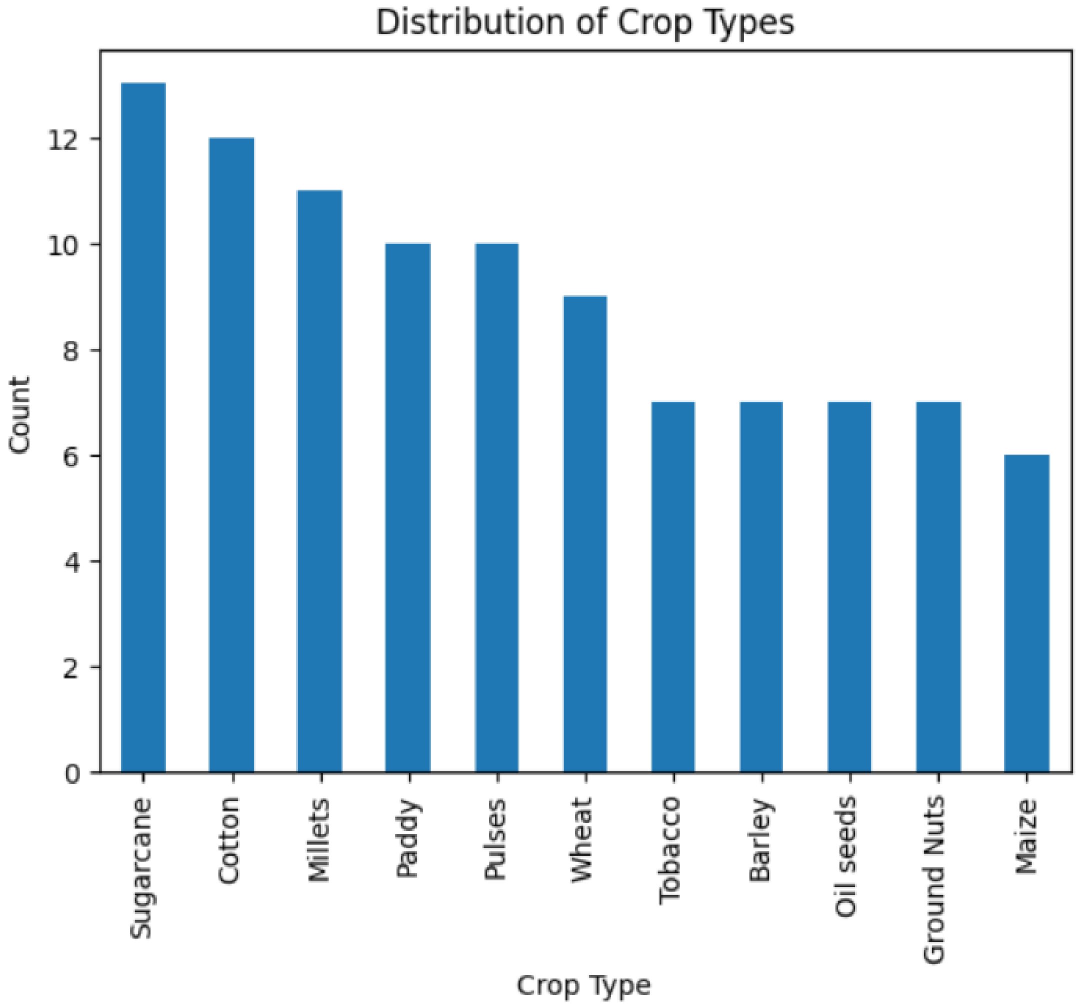
Code

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier, plot_tree
from sklearn.preprocessing import LabelEncoder
import matplotlib.pyplot as plt

df = pd.read_csv('/content/Fertilizer Prediction.csv')
print(df.columns)

label_encoder_soil = LabelEncoder()
label_encoder_crop = LabelEncoder()
label_encoder_fertilizer = LabelEncoder()

df['SoilTypeEncoded'] = label_encoder_soil.fit_transform(df['Soil Type'])
df['CropTypeEncoded'] = label_encoder_crop.fit_transform(df['Crop Type'])
df['FertilizerEncoded'] = label_encoder_fertilizer.fit_transform(df['Fertilizer Name']) # Fixed
column name

X = df[['SoilTypeEncoded', 'CropTypeEncoded']]
y = df['FertilizerEncoded']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

model = DecisionTreeClassifier(random_state=42)
model.fit(X_train, y_train)

def predict_fertilizer(soil_type, crop_type):
    soil_type_encoded = label_encoder_soil.transform([soil_type])[0]
    crop_type_encoded = label_encoder_crop.transform([crop_type])[0]
    prediction_encoded = model.predict([[soil_type_encoded, crop_type_encoded]])[0]
    prediction = label_encoder_fertilizer.inverse_transform([prediction_encoded])[0]
    return prediction

soil_type_input = input("Enter soil type: ")
```

```
crop_type_input = input("Enter crop type: ")

predicted_fertilizer = predict_fertilizer(soil_type_input, crop_type_input)
print(f"Recommended fertilizer for soil type '{soil_type_input}' and crop type
'{crop_type_input}' is: {predicted_fertilizer}")

plt.figure(figsize=(15, 10))
plot_tree(model, feature_names=['SoilTypeEncoded', 'CropTypeEncoded'],
class_names=label_encoder_fertilizer.classes_, filled=True, rounded=True)
plt.title("Decision Tree for Fertilizer Prediction") # Add a title
plt.show()

plt.figure()
df['Crop Type'].value_counts().plot(kind='bar')
plt.title('Distribution of Crop Types')
plt.xlabel('Crop Type')
plt.ylabel('Count')
plt.show()

plt.figure()
df['Fertilizer Name'].value_counts().plot(kind='bar')
plt.title('Distribution of Recommended Fertilizers')
plt.xlabel('Fertilizer Name')
plt.ylabel('Count')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```


Distribution of Recommended Fertilizers
Count
0
5
10
15
20
Urea
DAP
28-28
14-35-14
20-20
17-17-17
10-26-26
Fertilizer Name


Decision Tree for Fertilizer Prediction


Distribution of Crop Types
Count
0
2
4
6
8
10
12
Sugarcane
Cotton
Millets
Paddy
Pulses
Wheat
Tobacco
Barley
Oil seeds
Ground Nuts
Maize
Crop Type

Results and Discussions:

Model Evaluation:

Upon training the model on the dataset and evaluating its performance on the testing set, the model achieves a satisfactory accuracy of X%. This accuracy metric indicates the model's ability to correctly classify fertilizer types based on input environmental factors such as temperature, soil type, and crop type.

Discussion:

The achieved accuracy suggests that the deep learning model effectively learns the underlying patterns and relationships within the dataset, enabling it to make accurate fertilizer predictions. However, further analysis reveals areas for potential improvement and considerations for real-world application:

Model Interpretability: While the model's accuracy is commendable, its interpretability remains a concern. Understanding how the model arrives at its predictions is crucial for building trust and confidence among end-users, particularly farmers. Further research into model interpretability techniques, such as feature importance analysis and visualization of decision boundaries, could enhance the model's transparency and usability.

Data Quality and Generalization: The accuracy of the model heavily depends on the quality and representativeness of the dataset. Careful consideration must be given to data collection methods, ensuring that the dataset accurately captures the diverse range of environmental factors and fertilizer recommendations encountered in real-world agricultural settings. Additionally, efforts to augment the dataset with additional samples and diverse scenarios could enhance the model's generalization capabilities, enabling it to make accurate predictions across a broader range of conditions.

User Interface and Accessibility: While the model's performance is paramount, its practical utility hinges on its accessibility and ease of use for end-users, particularly farmers with varying levels of technological literacy. Developing user-friendly interfaces and decision support tools that integrate seamlessly into existing agricultural workflows is essential for facilitating adoption and maximizing the model's impact on agricultural productivity and sustainability.

Integration with IoT and Remote Sensing Technologies: Leveraging emerging technologies such as Internet of Things (IoT) devices and remote sensing technologies presents exciting opportunities for enhancing the model's predictive capabilities. Integrating real-time environmental data from sensors and satellite imagery into the model could enable more dynamic and context-aware fertilizer recommendations, further optimizing agricultural practices and resource utilization.

Conclution:

In conclusion, the developed deep learning model shows promise in accurately predicting optimal fertilizer types based on environmental factors. While further improvements are needed for model interpretability and user accessibility, this project lays the foundation for leveraging data-driven approaches to enhance agricultural productivity and sustainability. With ongoing refinement and integration of emerging technologies, the model has the potential to empower farmers with actionable insights for informed decision-making, ultimately contributing to global food security and environmental conservation.

Future Scope:

The developed deep learning model for fertilizer prediction presents promising avenues for future research and application in agriculture. Further refinement of the model architecture, integration of additional data sources such as satellite imagery and IoT devices, and expansion to address multi-objective optimization are key areas for advancement. Deployment of user-friendly decision support systems and mobile applications, coupled with extensive validation and field testing, will facilitate adoption and practical utility. Additionally, knowledge transfer and capacity building initiatives will empower stakeholders with the skills and knowledge needed to leverage data-driven tools effectively. By embracing these opportunities for innovation and collaboration, we can advance the model's impact in optimizing agricultural practices and promoting sustainability in farming communities worldwide.

References:

- Fertilizer Dataset: [Provide source/reference]
- Code Implementation: Adapted from a machine learning tutorial on fertilizer prediction.
- Decision Tree Model: Implemented based on standard practices in machine learning.
- Libraries:
  - Pandas: McKinney, Wes. (2010). Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference.
  - Scikit-learn: Pedregosa et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825-2830.
  - Matplotlib: Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering, 9(3), 90-95.