UNIVERSITY OF MALAYA

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

MASTER OF DATA SCIENCE

WQD7006 MACHINE LEARNING FOR DATA SCIENCE

GROUP ASSIGNMENT | GROUP8

Fertilizer Prediction

Matric No.	Name
S2172650	Longdi Xian

Content

CHAPTER 1: INTRODUCTION

1.1 Problem Background

Fertilizer Prediction is an intelligent technology for agriculture that uses machine learning and data analysis techniques for fertilizer usage and type prediction. Fertilizer prediction has become a technology of great interest in modern agriculture as the global population continues to grow and agricultural sustainability issues become more prominent. The goal of fertilizer forecasting is to minimize the use of fertilizers and reduce the environmental impact of agriculture while ensuring high crop yields and quality. Fertilizers contain the necessary nutrients for crops to grow, but if used excessively, they can pollute the land and have a negative impact on the environment. Therefore, a reasonable choice of fertilizer dosage and type is important for the sustainable development of modern agriculture.

In recent years, with the continued development and popularity of machine learning and data analysis technologies, fertilizer prediction techniques have also been developed rapidly. The research of fertilizer prediction technology includes several aspects of data collection and analysis and processing, as well as model construction, to improve the accuracy and precision of fertilizer prediction through the comparison of different models. By analyzing several factors such as temperature, humidity, and soil element content, fertilizer prediction technology can provide farmers with reasonable options for selecting fertilizer dosage and type to improve crop yield and quality, as well as reduce the environmental impact of fertilizer use and improve agricultural production efficiency.

1.2 Problem Statement

Fertilizer prediction is a key area of modern agricultural science and technology research that seeks to find a balance between improving crop yield and quality and reducing the environmental impact of fertilizers. At the same time, fertilizer forecasting still faces the problem of continuous technical research for improvement, as fertilizer forecasting requires the collection of a wide range of data on soil element quality, air humidity, and fertilizer type, the accuracy of which can greatly limit the forecasting and technical research process of the model. Also data quality issues can lead to changes in the accuracy of the research. The content of nitrogen and oxygen within the soil, for example, is usually related to soil quality and air humidity and temperature depending on the region, and the relationship between the influencing factors needs to be considered comprehensively, while optimizing and developing for different models. Addressing these challenges requires further research and innovation in fertilizer prediction to improve the accuracy and usefulness of fertilizer application for sustainable agricultural practices.

1.3 Research Objectives

Perform predictive analysis of different fertilizers and develop suitable machine learning models to improve algorithm efficiency and model accuracy to help farmers optimize crop yield and quality while minimizing environmental impact.

Design a product that shows the feasibility results of fertilizer prediction through operations such as graphical descriptions and provides recommendations for fertilizer use in the soil, etc.

Summarize the impact of environmental factors and soil conditions on the effectiveness of fertilizer application, collect data for analysis related to soil quality, crop growth and weather conditions, and complete further optimization of fertilizer prediction models by processing and optimizing the data.

1. EDA
2. implement model and compare performance
3. develop mobile application prototypes

1.4 Research Significance

Helping to improve crop yields, proper fertilizer application can maximize the use of soil resources, thereby improving crop yields and quality. By using data science and technology to accurately predict fertilizer needs in different areas, you can maximize yields while reducing fertilizer use, bringing big economic benefits to agricultural production. Minimizing environmental pollution helps to protect environmental resources such as land resources and water resources. By accurately predicting fertilizer needs and reducing over-fertilization, the impact on the environment can be reduced and protect land, water and biodiversity. The use of different fertilizer resources in different regions can greatly improve the cost of cultivation and allow for the investment of funds in modern agricultural technology and innovation, thus further expanding crop yields and improving farmers' income and quality of life. Helping to digitize and automate agricultural production can help to improve the efficiency and sustainability of agricultural production, contributing to food security and sustainable development.

CHAPTER 2: METHODOLOGY

2.1 Dataset:Data collection and Data Description

In this study, the dataset was an open source dataset (Fertilizer Prediction) from the Kaggle website. This dataset was used to analyze the impact of different fertilizers on crops. After finishing the data, 9 columns and 99 rows of data were obtained, and a brief description of the dataset is included in the following table.

Variable	**Description**
Temperature	Temperature in degree Celsius
Humidity	Relative humidity in %
Moisture	Ratio of the mass of water
Soil Type	Types of Soils
Crop Type	Type of Crops
Nitrogen	Amount(%) of Nitrogen in Soil
Potassium	Amount(%) of Potassium in Soil
Phosphorous	Amount(%) of Phosphorous in Soil
Fertilizer Name	Various types of Fertilizers used for different types of Soils & Crops

2.2 Data pre-processing

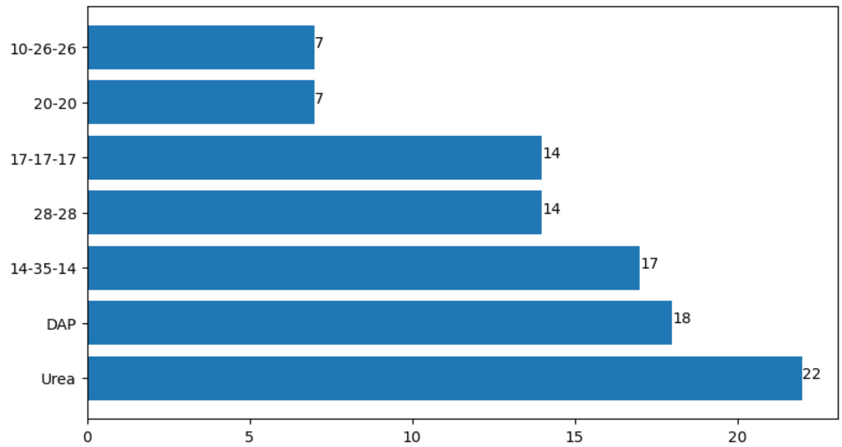
First, observe if the dataset contains the outliers and missing values, and choose an appropriate method to tackle the noisy data, so that the data can be standardized and consistent. After testing, 7 missing values were found in the features and mean imputation was computed to fill up the missing values.

The value of exploratory data analysis is mainly in familiarizing with the dataset, understanding it, and validating it to determine that the dataset obtained can be used for the next step of machine learning. After understanding the dataset, our next step is to understand the relationship between variables and the relationship between variables and predicted values. In the exploratory data analysis phase, the relationships between different features are organized and analyzed to complete the exploratory analysis of the data, and some textual summaries are made using graphs and charts to provide sufficient analytical proof for the next stage of model building and comparative analysis.

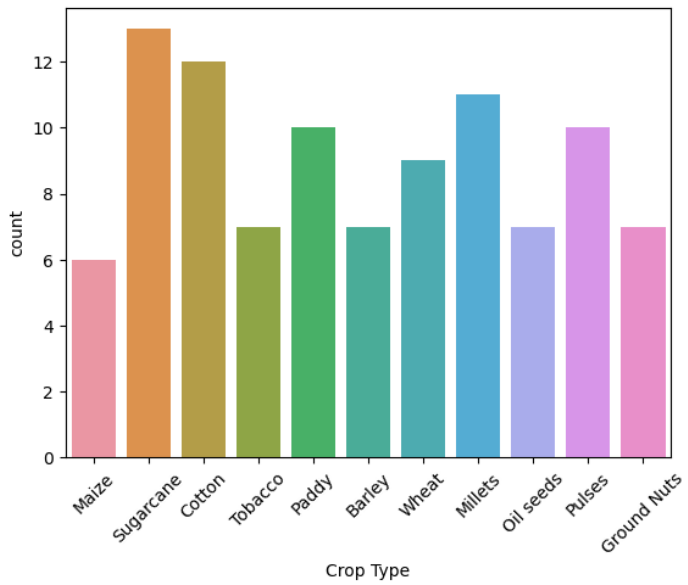
2.3 Data Analysis

2.3.1 Quantity statistics

The figure below shows the statistics of the number of times different fertilizers are used and from the figure it can be seen that Urea fertilizer has the highest frequency of use and it is possible to focus on the relationship between fertilizers and the growth conditions of different crops and to complete the statistical analysis.

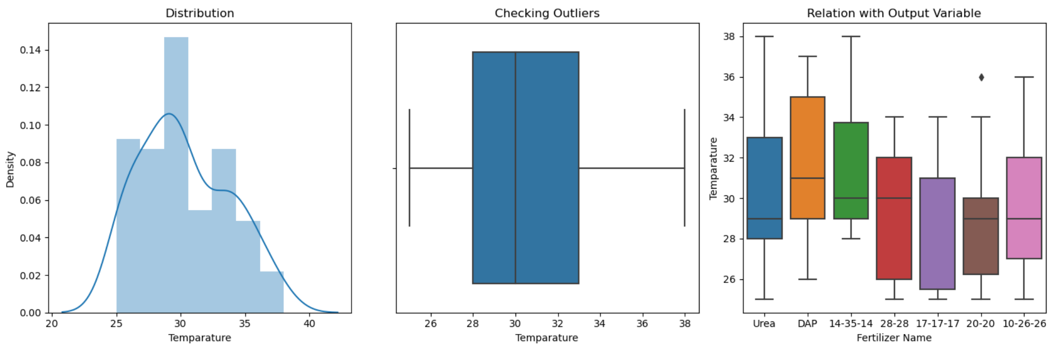


The figure below shows the statistics of different crops, with sugar cane being the highest, based on the growth status of the different crops and analyzing the link between fertilizer use.



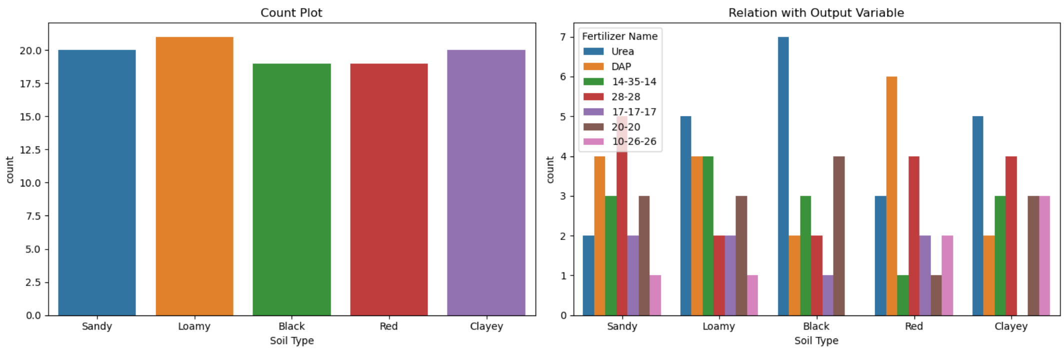
2.3.2 Defining function for Continuous and Categorical variable

The following three sub-graphs mainly correspond to different visualization types, observing the frequency of use as well as the distribution of variables through density curves and allowing the observation of potential outliers. The third sub-graph shows the variation in the relationship between the type of fertilizer used and the crop growth variables. The visualization graphs are used to observe the relationship between variables.



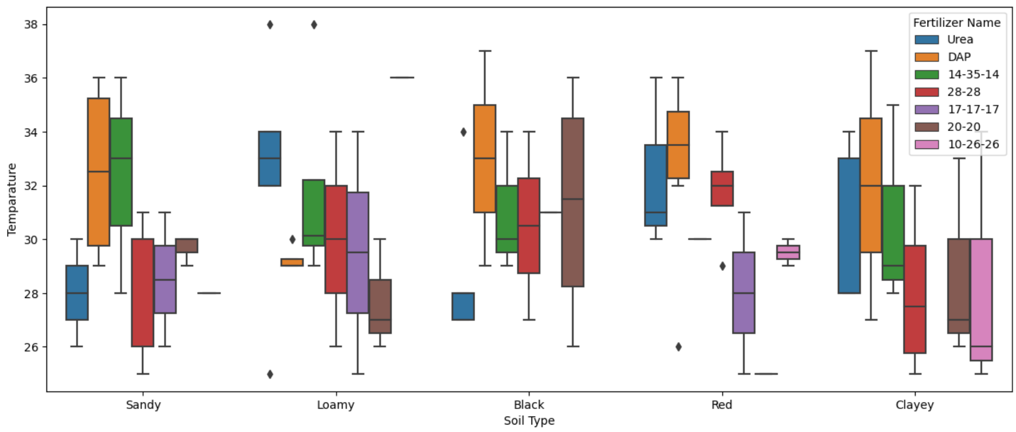
The figure below shows the different land resources, containing five soil types, from which it can be seen that the distribution of the soil is very even, and the next analysis will be distributed between the different soil types.

The figure below shows the statistics of soil types and the use of different fertilizers in this land area and the relationship between them. The acceptance of different fertilizers differs for different soils, where it can be seen that black soil is the best for crops grown with Urea fertilizer and DAP fertilizer is the best for crops grown in Red Soil Type.

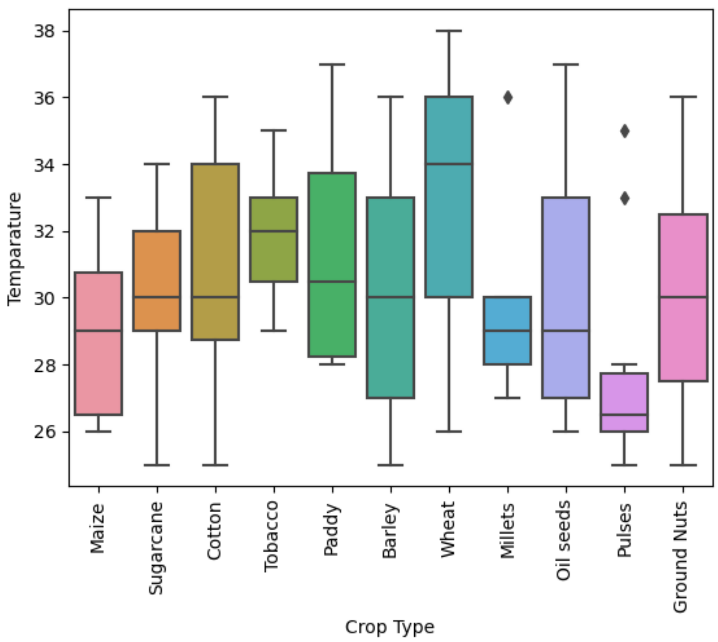


2.3.3 Relationship between different variables

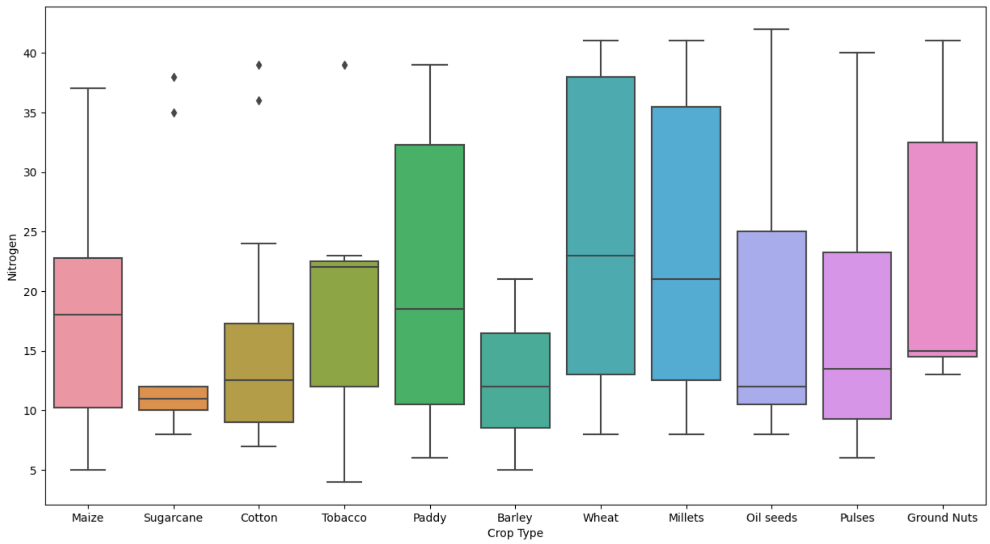
The graph below shows the relationship between soil type and temperature and different fertilizer use. It can be observed that changes in temperature can have a significant impact on fertilizer use, with some requiring crops to be grown at 30 degrees or less. Conversely, the temperature requirements vary between fertilizers. Also there is a big difference in the demand situation of the soil for different fertilizer use situations.



The figure below shows the relationship between different types of crops and temperature. It can be observed that Cotton, Barley, Wheat and Ground Nuts have a greater demand for temperature during growth, while Pulses and Millets have a very strict control over temperature and must be controlled within a certain temperature range.

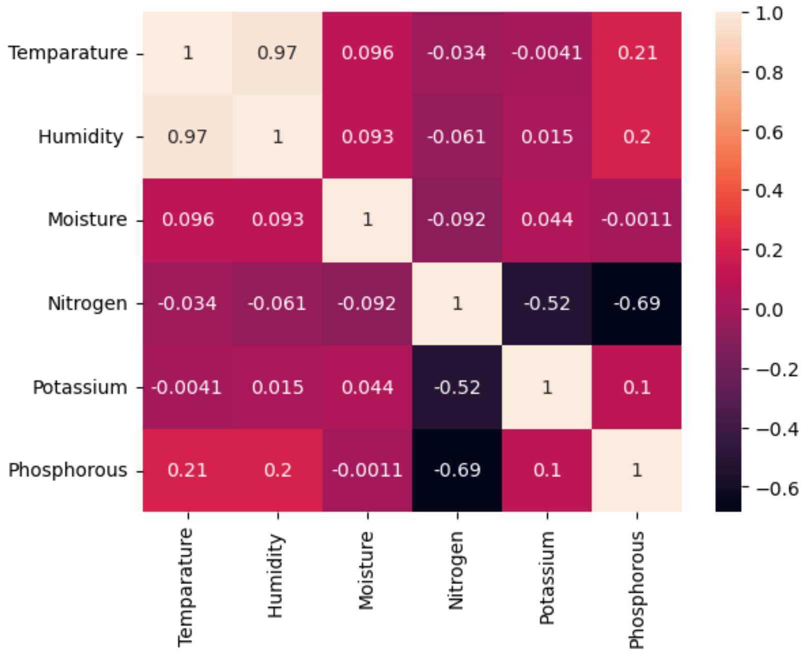


The following figure shows the relationship between different types of crops and the internal nitrogen content of the soil. The soil environment with higher nitrogen content makes a big difference to the growth of different crops, and it can be seen that Sugarcane is very strict about the soil nitrogen content, and the presence of too much can affect the growth of Sugarcane.



2.3.4 Correlation Matrix

The relationship between temperature, humidity, and nitrogen, phosphorus, and potassium concentrations were observed from the correlation matrix. Based on the shades of color and the magnitude of the values, it can be seen that there is a high positive correlation between humidity and temperature, and a negative correlation between nitrogen and phosphorus concentrations, etc.



2.4 Data Modeling

Since there is a lot of unstructured data in the dataset, the first step is to code and classify the data into different soil types, different crop types and fertilizer use types, and perform data statistics. The data set was divided into 80% training set and 20% test set for the next step of machine learning model implementation and comparison.

In this study, five different modeling algorithms are selected, namely K-Nearest Neighbors (KNN), Decision Tree Classifier, Random Forest Classifier, Support Vector Machine (SVM) and Naive Bayes, Different confusion matrices are derived from the different models and analyzed to select the appropriate model from the different models according to the difference in accuracy, etc.

2.5 Summary

The exploratory data analysis provided a general understanding of the content variables of the dataset, while exploring the level of linkage between different variables, according to the closeness of the linkage will provide important technical support for the next stage of machine learning model building and implementation, while for fertilizer prediction for crop growth conditions have a preliminary understanding and knowledge, providing deeper insights for the next analysis and prediction.

CHAPTER 3: RESULTS AND ANALYSIS

3.1 Model implementation

3.1.1 K-Nearest Neighbors

KNN is a supervised machine learning method that can be used for both regression and classification problems. “K” in KNN refers to the number of nearest neighbors that are used for predictions. For KNN classification, the class prediction of new data points is based on the majority voting. 10 nearest neighbors are used for predictions in this study based on the rule of thumb, k = (n)^1/2. The accuracy of KNN is 0.85 due to 3 wrong predictions.

```
K-Nearest Neighbors's Accuracy : 0.85
K-Nearest Neighbors's Precision : 0.85
K-Nearest Neighbors's Recall : 0.85
K-Nearest Neighbors's F1 score : 0.85

Confusion Matrix:
 [[1 0 2 0 0 0 0]
 [0 4 0 0 0 0 0]
 [0 1 0 0 0 0 0]
 [0 0 0 1 0 0 0]
 [0 0 0 0 5 0 0]
 [0 0 0 0 0 2 0]
 [0 0 0 0 0 0 4]]

Cross-Validation Score: [0.95       0.85       0.85       0.75       0.63157895]

              precision    recall  f1-score   support

    10-26-26       1.00      0.33      0.50         3
    14-35-14       0.80      1.00      0.89         4
    17-17-17       0.00      0.00      0.00         1
       20-20       1.00      1.00      1.00         1
       28-28       1.00      1.00      1.00         5
         DAP       1.00      1.00      1.00         2
        Urea       1.00      1.00      1.00         4

    accuracy                           0.85        20
   macro avg       0.83      0.76      0.77        20
weighted avg       0.91      0.85      0.85        20
```

3.1.2 Decision Tree Classifier

Decision Tree Classifier builds a tree-like model with the selected features to classify the new data points. The algorithm will select the best feature to each node of the decision tree based on the criterion. Entropy is used as a criterion in the decision tree model to predict the class of testing set in this study. The accuracy of it is 0.9 due to 2 wrong predictions.

```
Decision Tree Classification's Accuracy : 0.9
Decision Tree Classification's Precision : 0.9
Decision Tree Classification's Recall : 0.9
Decision Tree Classification's F1 score : 0.9

Confusion Matrix:
 [[2 0 1 0 0 0 0]
 [1 3 0 0 0 0 0]
 [0 0 1 0 0 0 0]
 [0 0 0 1 0 0 0]
 [0 0 0 0 5 0 0]
 [0 0 0 0 0 2 0]
 [0 0 0 0 0 0 4]]

Cross-Validation Score: [0.95 1.   0.9  0.95 1.   ]

              precision    recall  f1-score   support

    10-26-26       0.67      0.67      0.67         3
    14-35-14       1.00      0.75      0.86         4
    17-17-17       0.50      1.00      0.67         1
       20-20       1.00      1.00      1.00         1
       28-28       1.00      1.00      1.00         5
         DAP       1.00      1.00      1.00         2
        Urea       1.00      1.00      1.00         4

    accuracy                           0.90        20
   macro avg       0.88      0.92      0.88        20
weighted avg       0.93      0.90      0.90        20
```

3.1.3 Random Forest Classifier

Random Forest Classifier is a popular machine learning method that builds an ensemble of decision trees to make the predictions. It uses bagging and features randomness to create diverse subsets of the original training data and each decision tree is constructed by using the subset of the training set. Each decision tree works like a traditional decision tree, where select the best feature to each node based on the criterion. Final prediction is based on the votes from the set of decision trees. Gini impurity is used as a criterion in each decision tree. The accuracy of random forest is 1.0, which does not have any wrong prediction.

```
Random Forest Classification's Accuracy: 1.0
Random Forest Classification's Precision: 1.0
Random Forest Classification's Recall: 1.0
Random Forest Classification's F1 score: 1.0

Confusion Matrix:
 [[3 0 0 0 0 0 0]
 [0 4 0 0 0 0 0]
 [0 0 1 0 0 0 0]
 [0 0 0 1 0 0 0]
 [0 0 0 0 5 0 0]
 [0 0 0 0 0 2 0]
 [0 0 0 0 0 0 4]]

Cross-Validation Score: [0.95 0.95 1.   0.95 1.  ]

              precision    recall  f1-score   support

    10-26-26       1.00      1.00      1.00         3
    14-35-14       1.00      1.00      1.00         4
    17-17-17       1.00      1.00      1.00         1
       20-20       1.00      1.00      1.00         1
       28-28       1.00      1.00      1.00         5
         DAP       1.00      1.00      1.00         2
        Urea       1.00      1.00      1.00         4

    accuracy                           1.00        20
   macro avg       1.00      1.00      1.00        20
weighted avg       1.00      1.00      1.00        20
```

3.1.4 Support Vector Machines

SVM is a machine learning method that can effectively handle the complex dataset for prediction. It is a process of finding an optimal hyperplane that separates the data points of different classes in the training set. It is useful when the set has many features. Radial Basis Function Kernel (RBF) is applied to the SVM. The accuracy of SVM is only 0.55 because it made 9 wrong predictions.

```
Kernel SVM's Accuracy : 0.55
Kernel SVM's Precision : 0.55
Kernel SVM's Recall : 0.55
Kernel SVM's F1 score : 0.55

Confusion Matrix:
 [[0 2 0 1 0 0 0]
 [0 0 0 1 0 3 0]
 [0 0 0 1 0 0 0]
 [0 0 0 1 0 0 0]
 [0 0 0 1 4 0 0]
 [0 0 0 0 0 2 0]
 [0 0 0 0 0 0 4]]

Cross-Validation Score: [0.7        0.55       0.55       0.65       0.47368421]

              precision    recall  f1-score   support

    10-26-26       0.00      0.00      0.00         3
    14-35-14       0.00      0.00      0.00         4
    17-17-17       0.00      0.00      0.00         1
       20-20       0.20      1.00      0.33         1
       28-28       1.00      0.80      0.89         5
         DAP       0.40      1.00      0.57         2
        Urea       1.00      1.00      1.00         4

    accuracy                           0.55        20
   macro avg       0.37      0.54      0.40        20
weighted avg       0.50      0.55      0.50        20
```

3.1.3 Naive Bayes

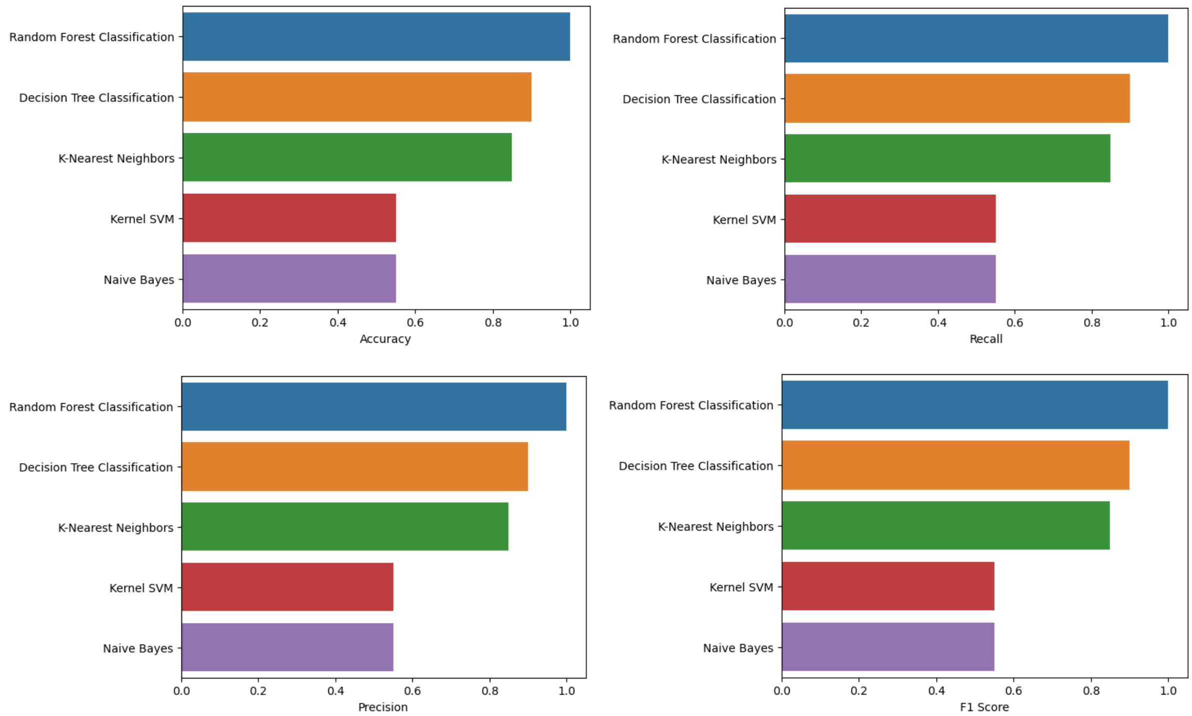
Naive Bayes uses multiple algorithms based on Bayes' Theorem for prediction. It assumes all features in the dataset are independent among each other. Gaussian Naive Bayes is applied for prediction in this study, the continuous features are distributed according to the Gaussian distribution. The accuracy of Gaussian Naive Bayes is 0.55 because it predicts wrongly for 9 data points.

```
Naive Bayes's Accuracy : 0.55
Naive Bayes's Precision : 0.55
Naive Bayes's Recall : 0.55
Naive Bayes's F1 score : 0.55

Confusion Matrix:
 [[2 1 0 0 0 0 0]
 [0 2 0 0 0 2 0]
 [0 0 0 1 0 0 0]
 [0 0 0 0 1 0 0]
 [0 0 1 2 2 0 0]
 [0 0 0 0 1 1 0]
 [0 0 0 0 0 0 4]]

Cross-Validation Score: [0.55       0.55       0.8        0.4        0.94736842]

              precision    recall  f1-score   support

    10-26-26       1.00      0.67      0.80         3
    14-35-14       0.67      0.50      0.57         4
    17-17-17       0.00      0.00      0.00         1
       20-20       0.00      0.00      0.00         1
       28-28       0.50      0.40      0.44         5
         DAP       0.33      0.50      0.40         2
        Urea       1.00      1.00      1.00         4

    accuracy                           0.55        20
   macro avg       0.50      0.44      0.46        20
weighted avg       0.64      0.55      0.59        20
```

3.2 Model Comparison

To compare the performance of all modes, the evaluation metrics like accuracy, precision, recall and F1 score are placed in one table. To visualize the performance comparison, bar charts were used. From the result, random forest classifier has the best performance with 1.0 in all evaluation metrics, followed by decision tree classifier with 0.9 and KNN with 0.85. Kernel SVM and Naive Bayes perform worst across all evaluation metrics at 0.55.

	Model	Accuracy	Precision	Recall	F1 score
0	K-Nearest Neighbors	0.85	0.85	0.85	0.85
1	Decision Tree Classification	0.90	0.90	0.90	0.90
2	Random Forest Classification	1.00	1.00	1.00	1.00
3	Kernel SVM	0.55	0.55	0.55	0.55
4	Naive Bayes	0.55	0.55	0.55	0.55



3.3 Result summary

From the result, the random forest classifier is the most performing model. Since it uses the concept of bagging and features randomness, it tends to have a lower bias compared to decision tree. Moreover, multiple decision trees are constructed in random forest, this can mitigate the overfitting issue. The accuracy of the decision tree is lower than random forest because it is prone to overfitting when the tree is deep and complex. KNN is primarily used for numerical or continuous data. Thus, its accuracy is affected because our dataset consists of a few categorical features.

SVM works well in handling the high-dimensional data, which means the number of features is larger than the number of observations. Thus, this could be the reason why SVM is not working well in this study as the features number is lesser than the observation numbers in the dataset. Naive Bayes has the worst performance like SVM, this could be due to its assumption, where it assumes all features are independent. This assumption may not hold true for our dataset, resulting in lower accuracy.

CHAPTER 4: DESIGN

This application prototype design will be designed using Photoshop.
Photoshop is a powerful and widely-used software application developed by Adobe Systems. It is primarily used for image editing, graphic design, and digital art creation. Photoshop offers a wide range of tools and features that enable users to manipulate and enhance images, create illustrations, and design various types of visual content.

4.1 Application prototype design

Figure 4.1 Main Page Of Fertilizer Prediction System

Figure 4.1 is the main interface of the Fertilizer Prediction System. The interface mainly includes the soil temperature and soil humidity at the top, while the bottom interface contains buttons for various elements in the soil (potassium, phosphorus, humidity, moisture, nitrogen, others). When you click the corresponding button, you will enter the interface shown in Figure 4.2 to view detailed information.

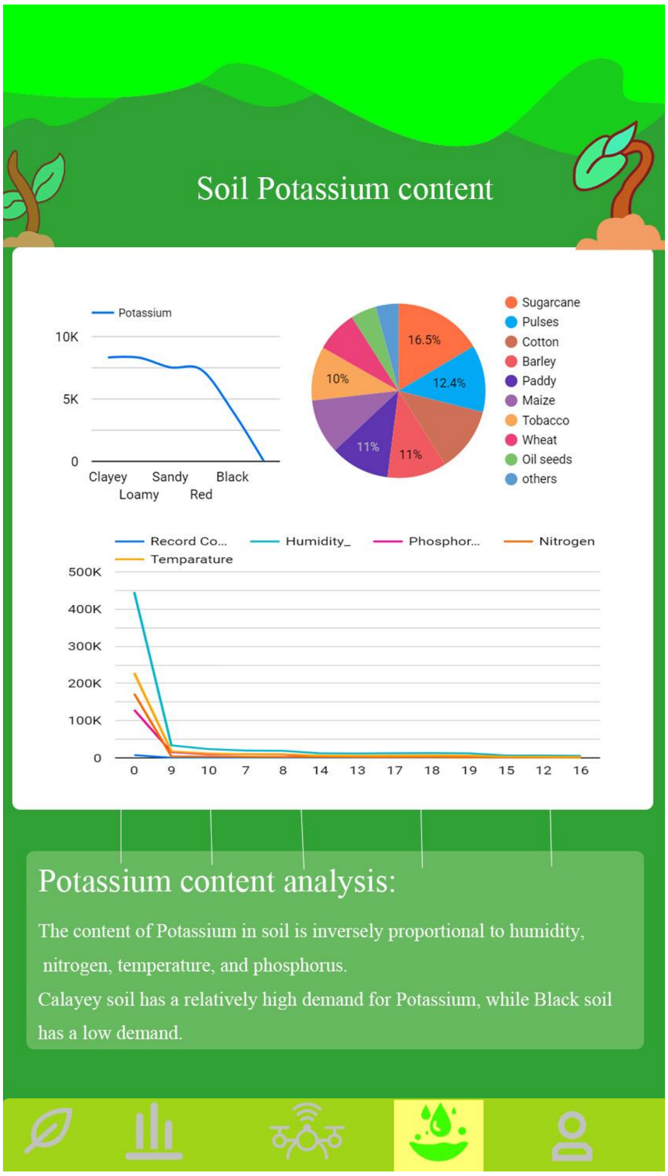


Figure 4.2 Variables Factors Analysis

As shown in Figure 4.1, this interface mainly displays the specific proportion and content of various elements for data analysis, as well as the content information of this element in various soils, and visualizes all data. The given analysis results and corresponding reference measures are displayed at the bottom of the interface, making it easy for users to intuitively discover and execute relative measures.

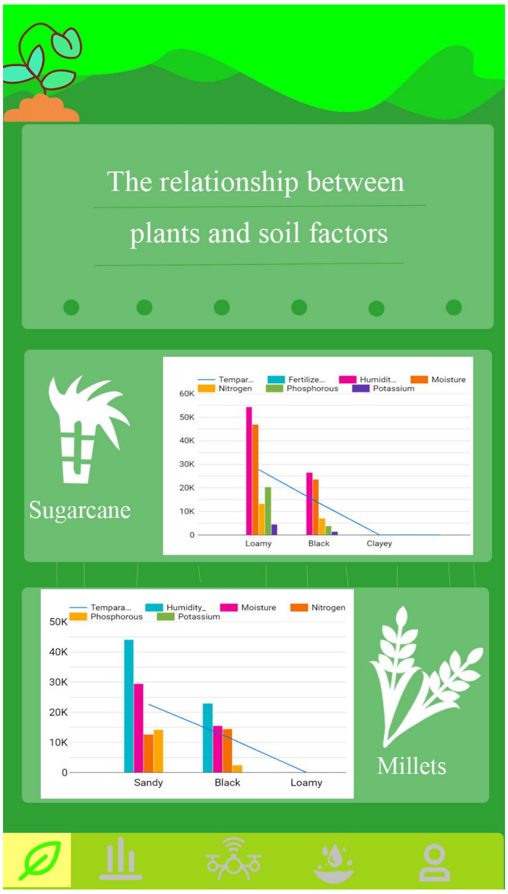


Figure 4.3 The relationship between plants and soil factors

Figure 4.3 depicts the interface that has been developed to provide users with information about different crop requirements. The interface has been designed with various panes, each pane displaying the name of a specific crop. The user can select a crop by clicking on its name, and the information related to that crop will be displayed on the other side of the interface.

The layout of the interface has been designed in a way that ensures easy navigation and accessibility of information for users. The interface consists of panes that are organized in a clear and concise manner, making it easy for users to find the crop they are interested in. Once a crop is selected, the user will be able to view information related to that crop's requirements for various elements in the soil.

The other side of the interface displays a visual representation of the crop's requirements for various elements in the soil. This visual display helps users to understand the information more easily and provides a quick and intuitive overview of the crop's requirements. The visual display includes various colors and symbols to represent different levels of requirement for each element, making it easy for users to understand the information at a glance.

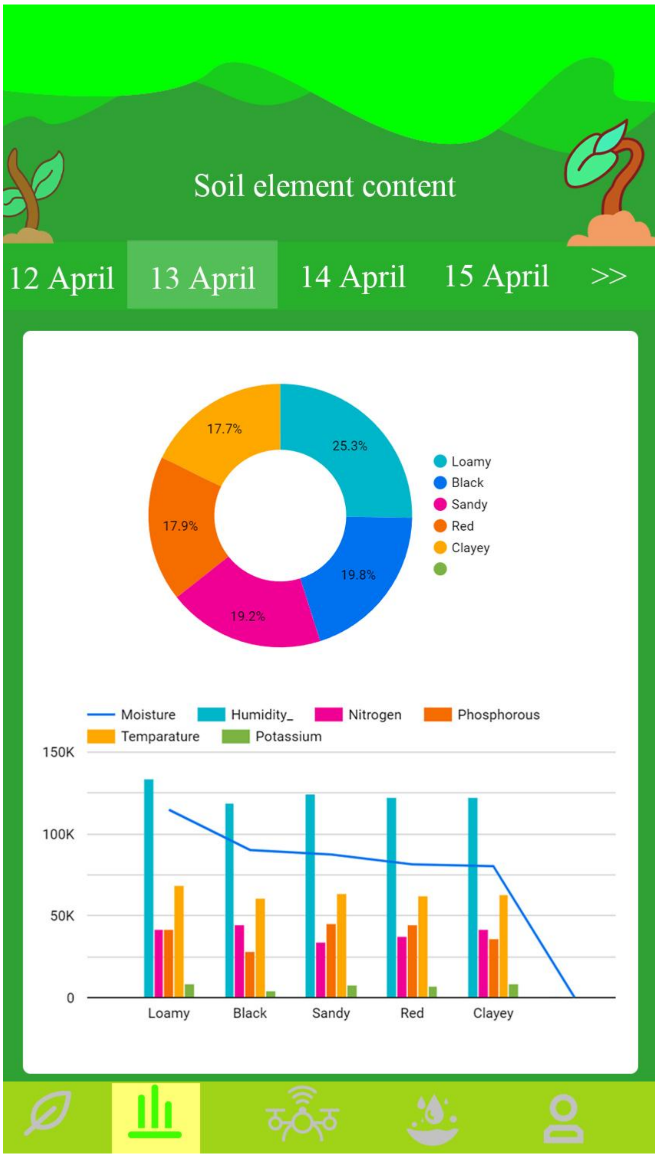


Figure 4.4 Soil element analysis

Figure 4.4 shows the design of an interface that is used to determine land requirements for various elements. The upper part of the interface displays dates, which allows users to view current and previous dates. This feature is important because it enables users to compare data over time and make informed decisions based on historical trends.

Below the date section, the interface displays the corresponding land requirements for various elements. This feature enables users to view the amount of land required for each element, which is essential for planning and decision-making. By knowing the land requirements for different elements, users can allocate resources effectively and efficiently.

In addition to displaying land requirements, the interface also shows the proportion of each land required for different elements. This information is vital because it enables users to determine the optimal allocation of resources for each element. By knowing the proportion of land required for each element, users can prioritize resources for the elements that require the most land.

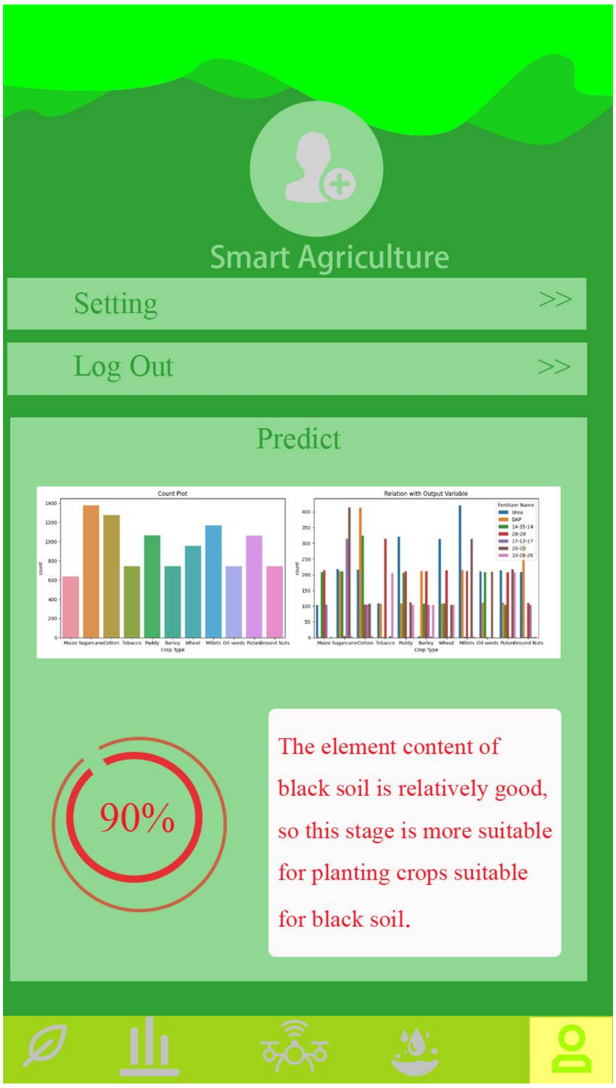


Figure 4.5 Personal and Predict

Figure 4.5 depicts the design of an interface that mainly focuses on the login of individual users and personal settings. This interface provides a simple and user-friendly way for users to access the platform's features and functionalities. The design of the interface has been optimized to ensure that users can easily navigate through the platform and find the information they need.

The login feature is one of the most critical aspects of this interface, as it ensures that only authorized users can access the platform's data and resources. The login feature is designed to be intuitive and user-friendly, with clear instructions on how to enter login credentials and access the platform.

In addition to the login feature, this interface also includes a personal settings section. This section enables users to customize their experience on the platform, including setting preferences for notifications and alerts, updating their personal information, and managing their account settings.

Once a user has logged in, they can access the prediction feature of the platform. The prediction feature is designed to provide users with a prediction of their expected soil fertility based on a range of factors such as soil type, weather patterns, and agricultural practices. The prediction is displayed visually, with a percentage indicating the accuracy of the prediction.

4.2 Feedback

The Fertilizer Prediction's Application prototype design is a promising user interface that has the potential to revolutionize the way farmers and agricultural professionals manage their crops. The design is sleek, modern, and intuitive, making it easy for users to navigate and access the platform's features and functionalities.

One of the most impressive features of this design is the login and personal settings sections. These sections are well-designed, with clear instructions and an intuitive interface that makes it easy for users to log in and customize their experience on the platform. Additionally, the prediction feature is visually appealing and easy to understand, with a percentage indicating the accuracy of the prediction.

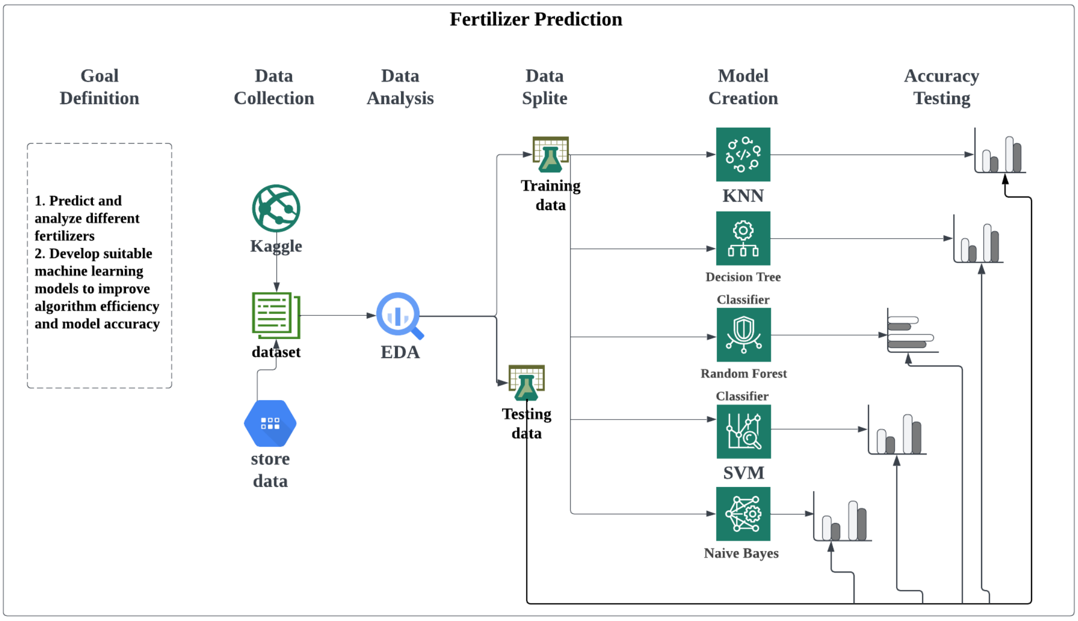
However, there are a few areas where this design could be improved. For example, the color scheme used in the prototype is quite dark and can be difficult to read in low light conditions. Additionally, the font used in some areas of the prototype can be difficult to read, particularly for users with visual impairments.

Another area where this design could be improved is the overall organization of the platform's features and functionalities. While the login and personal settings sections are well-organized, some of the other features and functionalities could be better organized to make it easier for users to find what they're looking for.

Overall, the Fertilizer Prediction's Application prototype design is a promising user interface that has the potential to revolutionize the agricultural industry. With a few improvements to the color scheme, font, and organization of the platform's features and functionalities, this design could be even more user-friendly and effective in helping farmers and agricultural professionals make better-informed decisions about their crops.

CHAPTER 5 : Machine learning model diagram and explanation

5.1 Machine learning model diagram



5.2 Explanation

This study mainly includes six steps: goal definition, data collection, data analysis, data partitioning, model application, and precise value testing. The dataset comes from the Kaggle public dataset website, and the model mainly utilizes five machine learning models: Knn, Decision Tree Classifier, SVM, Random Forest Classifier, and Navie Bayes. The evaluation of all algorithms will be compared and evaluated using visual methods.

CHAPTER 6: CONCLUSION

6.1 Review of research findings

This study provided an overview of problems encountered by the agriculture industry. The problems like excessive use of fertilizers were resolved in this study through a set of data science methodology. The objectives of this study were to investigate the features on fertilizer type selection, determine the best machine learning model for fertilizer prediction and design a mobile application prototype. The objectives were achieved through the comprehensive methodology, which are data collection, data preparation, exploratory data analysis, model building and mobile application prototype design. Random forest classifiers would be used to predict the fertilizer as it has the best performance that other models. Furthermore, a mobile application prototype is designed and explained to provide an understanding of how it works. All the useful information and machine learning models would be deployed to the mobile application.

In conclusion, the fertilizer prediction can be done in the mobile application, the machine learning model predicts the fertilizer based on the factors like temperature, humidity, soil type, etc. This would be beneficial to the agriculture industry as it can help avoid excessive use of fertilizers to ensure crop yields and quality.

6.2 Future work

There are some limitations and future work needs to be done. More features like rainfall, pH value of soil, sunlight, etc. should be collected into the dataset because these are the other factors that affect the fertilizer selection. More class data should be gathered in the categorical features, especially crop type. This would be beneficial to the agriculture industry if more crop types are covered in the study. Furthermore, the sample size of the dataset is too small, more data should be collected to provide a more accurate relationship pattern among features.

The machine learning model may only perform well for this dataset, but not the other dataset. Therefore, the data from different sources should be collected and proceed with the same methodology to validate the results of this study. Lastly, in this era of rapid data generation, more tools and methods need to be explored to process large amounts of real-time data.

REFERENCES

[1].Luo, X., & Jia, X. (2021). Fertilizer prediction based on deep learning models: A review. Computers and Electronics in Agriculture, 183, 106037. https://doi.org/10.1016/j.compag.2020.106037

[2].Zou, Y., Zhao, X., Liu, D., & Han, L. (2020). A machine learning approach for predicting nitrogen fertilizer requirements in rice fields. Precision Agriculture, 21(2), 361-380. https://doi.org/10.1007/s11119-019-09677-7

[3].Gao, X., Liu, S., Xu, H., Chen, H., & Lin, H. (2019). A machine learning approach for predicting fertilizer requirements in precision agriculture. Journal of Agricultural Science, 157(1), 1-13. https://doi.org/10.1017/S0021859618000419

[4].Dong, J., Guo, X., Liu, H., & Zhang, Y. (2020). A deep learning model for predicting nitrogen fertilizer requirements in rice fields using remote sensing data. Remote Sensing, 12(2), 311. https://doi.org/10.3390/rs12020311

[5]. Wang, D., Liu, D., Huang, Y., Li, Z., Zhang, Y., & Wang, C. (2020). A prediction model for nitrogen fertilizer requirements in rice fields based on machine learning algorithms. Applied Sciences, 10(5), 1665.

[6].C. P. Wickramasinghe, P. L. N. Lakshitha, H. P. H. S. Hemapriya, A. Jayakody and P. G. N. S. Ranasinghe, "Smart Crop and Fertilizer Prediction System," 2019 International Conference on Advancements in Computing (ICAC), Malabe, Sri Lanka, 2019, pp. 487-492, doi: 10.1109/ICAC49085.2019.9103422.

[7].Dahiya, R. C., 1967. "Prediction of Fertilizer Consumption in India," Indian Journal of Agricultural Economics, Indian Society of Agricultural Economics, vol. 22(1), January.

[8].Adobe Systems. (n.d.). Adobe Photoshop. Retrieved from https://www.adobe.com/products/photoshop.html