**“DETERMINING CUSTOM FERTILIZER PROFILE**

**BASED ON SOIL CHEMISTRY”**

*Minor project report submitted*

*in partial fulfillment of the requirement for award of the degree of*

**Bachelor of Technology**

**in**

**Computer Science & Engineering**

**By**

**K.DRUVA KUMAR** (20UECS0472) **(VTU 16815)**

**K.BHARGAV** (20UECS0394) **(VTU 17701)**

*Under the guidance of*

*Dr. P. J. Beslin Pajila, M.E., Ph.D.,*

*ASSISTANT PROFESSOR*



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**SCHOOL OF COMPUTING**

**VEL TECH RANGARAJAN DR. SAGUNTHALA R&D INSTITUTE OF SCIENCE & TECHNOLOGY**

**(Deemed to be University Estd u/s 3 of UGC Act, 1956)**

**Accredited by NAAC with A++ Grade**

**CHENNAI 600 062, TAMILNADU, INDIA**

**May, 2023**

**“DETERMINING CUSTOM FERTILIZER PROFILE**

**BASED ON SOIL CHEMISTRY”**

*Minor project report submitted*

*in partial fulfillment of the requirement for award of the degree of*

**Bachelor of Technology**

**in**

**Computer Science & Engineering**

**By**

**K.DRUVA KUMAR** (20UECS0472) **(VTU 16815)**

**K.BHARGAV** (20UECS0394) **(VTU 17701)**

*Under the guidance of*

*Dr. P. J. Beslin Pajila , M.E., Ph.D.,*

*ASSISTANT PROFESSOR*



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING SCHOOL OF COMPUTING**

**VEL TECH RANGARAJAN DR. SAGUNTHALA R&D INSTITUTE OF SCIENCE & TECHNOLOGY**

**(Deemed to be University Estd u/s 3 of UGC Act, 1956)**

**Accredited by NAAC with A++ Grade**

**CHENNAI 600 062, TAMILNADU, INDIA**

**CERTIFICATE**

It is certified that the work contained in the project report titled “DETERMINING CUSTOM FER-

TILIZER PROFILE BASED ON SOIL CHEMISTRY” by ”K.DRUVA KUMAR (20UECS0472) , K.BHARGAV (20UECS0394)” has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

**Signature of Supervisor Dr. P. J. Beslin Pajila**

**Assistant Professor**

**Computer Science & Engineering**

**School of Computing**

**Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science & Technology**

**May , 2023**

**Signature of Head of the Department**

**Dr.M.S.MURALIDHAR**

**Associate Professor & Head of the Department**

**Computer Science & Engineering**

**School of Computing**

**Vel Tech Rangarajan Dr. Sagunthala R&D**

**Institute of Science & Technology**

**May , 2023**

**Signature of the Dean Dr. V. Srinivasa Rao**

**Professor & Dean**

**Computer Science & Engineering**

**School of Computing**

**Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science & Technology**

**May , 2023**

**DECLARATION**

We declare that this written submission represents our ideas in our own words and where others’ ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(K.DRUVA KUMAR)

Date: / /

(K.BHARGAV)

Date: / /

**APPROVAL SHEET**

This project report entitled “DETERMINING CUSTOM FERTILIZER PROFILE BASED ON SOIL CHEMISTRY” by K.DRUVA KUMAR (20UECS0472), K.BHARGAV (20UECS0394) is approved

for the degree of B.Tech in Computer Science & Engineering.

**Examiners**

**Supervisor**

Dr. P. J. Beslin Pajila, M.E., Ph.D.,

Assistant professor

**Date: / /**

**Place:**

**ACKNOWLEDGEMENT**

We express our deepest gratitude to our respected **Founder Chancellor and President Col. Prof. Dr. R. RANGARAJAN B.E. (EEE), B.E. (MECH), M.S (AUTO),D.Sc., Foundress President Dr. R. SAGUNTHALA RANGARAJAN M.B.B.S.** Chairperson Managing Trustee and Vice President.

We are very much grateful to our beloved **Vice Chancellor Prof. S. SALIVAHANAN,** for provid- ing us with an environment to complete our project successfully.

We record indebtedness to our **Professor & Dean, Department of Computer Science & Engi- neering, School of Computing, Dr. V. SRINIVASA RAO, M.Tech., Ph.D.,** for immense care and encouragement towards us throughout the course of this project.

We are thankful to our **Head, Department of Computer Science & Engineering, Dr. M. S. MURALI DHAR, M.E., Ph.D.,** for providing immense support in all our endeavors.

We also take this opportunity to express a deep sense of gratitude to our Internal Supervisor **Dr. S. PRIYA, M.E., Ph.D.,** for her cordial support, valuable information and guidance, she helped us in completing this project through various stages.

A special thanks to our **Project Coordinators Mr. V. ASHOK KUMAR, M.Tech., Ms. C. SHYAMALA KUMARI, M.E.,** for their valuable guidance and support throughout the course of the project.

We thank our department faculty, supporting staff and friends for their help and guidance to com- plete this project.

**K.DRUVA KUMAR K.BHARGAV**

(20UECS0472) (20UECS0394)

**ABSTRACT**

Collecting soil samples from different locations and conducting analysis to deter- mine their chemical composition. The data obtained from the soil analysis will be used to train machine learning models that can predict the best fertilizer profiles for a given soil type. The system will also include a user interface that allows farm- ers to input soil sample data and receive fertilizer recommendations.The benefits of this project include increased crop yield, reduced fertilizer usage, and improved soil health. By providing customized fertilizer recommendations based on soil chemistry, the proposed system will help farmers optimize their agricultural production and re- duce the environmental impact of excessive fertilizer use.The prediction of fertilizer profiles of the soil-based upon various factors and tries to find the best fit offertiliz- ers profile for the crop which satisfies all the nutrient requirements, optimal growth and yield of food crops that require the availability of multiple essentials in soil like N, P, K Can, Mg, Cu, Fe, etc. Decision trees, Random forest and Navie bayes are used for predicting the most suitable fertilizer according to the given attributes

**Keywords: Crop Yield, Decision Tree, Fertilizers Recommendations, Humidity, Irrigation, KNN- K-nearest neighbour, Moisture, Naivebayes, Random Forest.**

**LIST OF FIGURES**

4.1 **Architecture Diagram** [11](#_bookmark1)

4.2 **Data Flow Diagram for Fertilizer prediction** [12](#_bookmark2)

4.3 **Use Case Diagram** [13](#_bookmark3)

4.4 **Class Diagram** [14](#_bookmark4)

4.5 **Sequence Diagram** [15](#_bookmark5)

4.6 **Collection of Data** [16](#_bookmark6)

4.7 **Processing The Data** [17](#_bookmark7)

4.8 **Visualization of Data** [18](#_bookmark8)

4.9 **Applying Machine Learning Algorithms** [18](#_bookmark9)

4.10 **Accuracy and Error Detection** [21](#_bookmark10)

4.11 **Getting the Output** [21](#_bookmark11)

5.1 **Input design** [23](#_bookmark12)

5.2 **Text Image** [24](#_bookmark13)

5.3 **Figure Soil type** [27](#_bookmark14)

5.4 **Test Image** [28](#_bookmark15)

6.1 **Validated Output** [31](#_bookmark16)

6.2 **Predicted Output** [32](#_bookmark17)

**LIST OF ACRONYMS AND**

**ABBREVIATIONS**

|  |  |
| --- | --- |
| ANN | Artificial Neural Network |
| Cu | Copper |
| DFD | Data Flow Diagram |

|  |  |
| --- | --- |
| ELM | Environmental Liabilities Management |
| Fe | Iron |
| GIS | Geographic Information System |
| GPS | Global Positioning System |
| GDP | Gross Domestic Product |

K Potassium

KNN K-Nearest Neighbour

|  |  |
| --- | --- |
| Mg | Magnesium |
| ML | Machine Learning |
| N | Nitrogen |

|  |  |
| --- | --- |
| NDVI | Normalized Difference Vegetation Index |
| OMT | Ocean Mean Temperature |
| pH | Potential of Hydrogen |
| PLSR | Partial Least Square Regression |
| RF | Random Forest |
| RMSE | Root Mean Square Error |
| RUP | Rational Unified Process |

|  |  |
| --- | --- |
| SVM | Support Vector Machine |
| UML | Unified Modeling Language |
| UAV | Unmanned Aerial Vehicle |
| VRA | Variable Rate Application |

**TABLE OF CONTENTS**

**Page.No**

**ABSTRACT v**

**LIST OF FIGURES vi**

**LIST OF TABLES vii**

**LIST OF ACRONYMS AND ABBREVIATIONS vii**

**1 INTRODUCTION 1**

1.1 Introduction [1](#_bookmark18)

1.2 Aim of the Project [2](#_bookmark19)

1.3 Project Domain [2](#_bookmark20)

1.4 Scope of the Project [3](#_bookmark21)

**2 LITERATURE REVIEW 4**

**3 PROJECT DESCRIPTION 7**

3.1 Existing System [7](#_bookmark22)

3.2 Proposed System [7](#_bookmark23)

3.3 Feasibility Study [8](#_bookmark24)

3.3.1 Economic Feasibility [8](#_bookmark25)

3.3.2 Technical Feasibility [8](#_bookmark26)

3.3.3 Social Feasibility [9](#_bookmark27)

3.4 System Specification [9](#_bookmark28)

3.4.1 Hardware Specification [9](#_bookmark29)

3.4.2 Software Specification [9](#_bookmark30)

3.4.3 Standards and Policies [9](#_bookmark31)

**4 METHODOLOGY** [**11**](#_bookmark32)

4.1 General Architecture [11](#_bookmark33)

4.2 Design Phase [12](#_bookmark34)

4.2.1 Data Flow Diagram [12](#_bookmark35)

4.2.2 Use Case Diagram [13](#_bookmark36)

4.2.3 Class Diagram [14](#_bookmark37)

4.2.4 Sequence Diagram [15](#_bookmark38)

4.3 Module Description [16](#_bookmark39)

4.3.1 Module1 : Data Collection and PRE-PROCESSING 16

4.3.2 Module2 : Data exploration [18](#_bookmark40)

4.4 Algorithm & Pseudo Code [19](#_bookmark41)

4.4.1 KNN Description [19](#_bookmark42)

4.4.2 Random Forest [19](#_bookmark43)

4.4.3 Decision Tree [19](#_bookmark44)

4.4.4 Pseudo Code [20](#_bookmark45)

4.4.5 Module3 [21](#_bookmark46)

**5 IMPLEMENTATION AND TESTING** [**23**](#_bookmark47)

5.1 Input and Output [23](#_bookmark48)

5.1.1 Input Design [23](#_bookmark49)

5.1.2 Output Design [24](#_bookmark50)

5.2 Testing [24](#_bookmark51)

5.3 Types of Testing [24](#_bookmark52)

5.3.1 Unit testing [24](#_bookmark53)

5.3.2 Integration testing [26](#_bookmark54)

5.3.3 Functional testing [26](#_bookmark55)

5.3.4 Test Result [28](#_bookmark56)

5.4 Testing Strategy [28](#_bookmark57)

**6 RESULTS AND DISCUSSIONS** [**29**](#_bookmark58)

6.1 Efficiency of the Proposed System [29](#_bookmark59)

6.2 Comparison of Existing and Proposed System [29](#_bookmark60)

6.3 Sample Code [30](#_bookmark61)

**7 CONCLUSION AND FUTURE ENHANCEMENTS** [**33**](#_bookmark62)

7.1 Conclusion [33](#_bookmark63)

7.2 Future Enhancements [33](#_bookmark64)

**8 PLAGIARISM REPORT 35**

**9 SOURCE CODE & POSTER PRESENTATION 36**

9.1 Source Code [36](#_bookmark65)

9.2 Poster Presentation [43](#_bookmark66)

**References** [**43**](#_bookmark67)



**Chapter 1**

**INTRODUCTION**

**1.1 Introduction**

Artificial intelligence is that the most rapidly growing area integrated into most aspects of human life.Due to the boom in the population, there is a rapid increase in the demand for food and crops. The goal is not only to improve crop yield but also with good quality. This completely depends on the nature of soil and fertilizers that farmers Nature has experienced drastic changes due to global warming.

Fertilizers play a crucial role in the growth and development of plants. How- ever, the effectiveness of a fertilizer largely depends on the specific needs of the plants and the soil in which they grow. Soil chemistry can have a significant impact on the growth of plants, as different soils contain varying levels of essential nutri- ents and minerals.

Determining a custom fertilizer profile based on soil chemistry involves analyzing the composition of the soil and identifying which nutrients are lacking or in excess. This information can then be used to create a customized fertilizer blend that is tai- lored to the specific needs of the soil and the plants growing in it. By providing the right balance of nutrients, a custom fertilizer profile can help optimize plant growth and yield, while minimizing waste and environmental impact.

Frequent droughts, heavy rainfall are the biggest challenges that farmers are fac- ing. Crop condition is getting varied within a limited time due to the climatic changes. Fertilizer recommendation should be in a way that plant must become capable enough to adapt and face any drastic condition. The mechanism that drives its edge technology is machine learning. It gives machines the power to find out without being explicitly programmed. The areas that are focused are a prediction of soil parameters like organic carbon and moisture content, crop yield prediction, disease and weed detection in crops, and species detection.



**1.2 Aim of the Project**

The main aim of the project is to determine the custom fertilizer profiles based on the soil chemistry using machine learning algorithms based on its properties like nitrogen, phosphorous, and potassium pH level content of the soil along with the rainfall accuracy. Apart from the primary goal we also work on predicting the ap- propriate crop that must be grown in the field based on the soil properties.

By analyzing the composition of the soil, identifying nutrient deficiencies or ex- cesses, and creating a custom fertilizer blend, farmers and growers can improve the efficiency and effectiveness of their fertilization practices.

To improve the efficiency and effectiveness of fertilization practices, minimize waste, and promote sustainable agriculture practices. By optimizing the use offertil- izers, farmers and growers can maximize their crop yields and reduce their environ- mental impact,which can lead to increased profitability and long-term sustainability. Additionally, such a project can help to address global food security challenges by increasing the availability and accessibility of nutritious crops.

**1.3 Project Domain**

Agriculture plays an important role within the economic process of any country. With the rise of population, frequent changes in weather conditions and restricted resources, it becomes a difficult task to fulfil the food demand of the current pop- ulation. Exactitude agriculture additionally called good farming have emerged as associate degree innovative tool to deal with current challenges in agricultural prop- erty.

The mechanism that drives the leading edge technology is machine learning (ML). It provides the machine ability to find out while not being expressly programmed. During this article, authors gift a scientific review of millilitre applications within the field of agriculture. The square measures that square measure centred are pre-

diction of soil parameters like organic carbon and wet content, crop yield prediction, sickness and weed detection in crops and species detection.

Millilitre with laptop vision square measure reviewed for the classification of a distinct set of crop pictures so as to watch the crop quality and yield assessment. This approach is integrated for increased eutherian mammal production by predicting fer- tility patterns, designation feeding disorders, bovine behaviour supported millilitre models victimisation information collected by collar sensors,etc.

Intelligent irrigation which has drip irrigation and intelligent gather techniques

also are reviewed that reduces human labour to an excellent extent. this text demon-

strates however knowledge-based agriculture will improve the property productivity

and quality of the merchandise.

**1.4 Scope of the Project**

Determining a custom fertilizer profile based on soil chemistry is a complex pro- cess that requires a range of skills and knowledge. The scope of a project focused on this area is broad and may include several components such as soil sampling, chem- ical analysis, and fertilizer formulation.

To demonstrate that recent evolving technologies are mature enough to provide cost-effective and end-to-end Precision Farming services and showing farmers what technology can do to help their business and how machine learning can make them challenge the frequent droughts and heavy rainfall.

The project can have a significant impact on the productivity and sustainability of agricultural practices, and can help to address global food security challenges.

**Chapter 2**

**LITERATURE REVIEW**

**Abhinav Sharma , et al., [1]**described the precision agriculture also known as smart farming have emerged as an innovative tool to address current challenges in agri- cultural sustainability. The areas that are focused are prediction of soil parameters such as organic carbon and moisture content, crop yield prediction, disease and weed detection in crops and species detection. ML with computer vision are reviewed for the classification of a different set of crop images in order to monitor the crop quality and yield assessment. This approach can be integrated for enhanced livestock pro- duction by predicting fertility patterns, diagnosing eating disorders, cattle behaviour based on ML models using data collected by collar sensors.

**Aman Kumar Dewangan, et al .,[2]** in soil nature is monitored using various de- tectors and toxicity is measured by considering various paramet

**Halima Sadiyah. Abdullahi,Et al .,[3]** precision Agriculture is focusing on tech- nical as well as traditional methods to improve the yield quality and quantity.Soil nature and its requirements changes from place to place so,various soil frameworks are taken into consideration and examined according to the requirements.



**Janez Trontelj ml. and Olga Chambers ,et al ., [4]** the researchers presented the hypothesis that the machine learning approach improves the accuracy of soil prop- erties prediction. The correlations obtained in this research are important for under- standing the overall strategy for soil properties prediction using optical spectroscopy sensors. Several research results have been stated and investigated. A comparison is made between six commonly used techniques: Random Forest, Decision Tree, Naive Bayes, Support Vector Machine, Least-Square Support Vector Machine and Artifi- cial Neural Network, showing that the best prediction accuracy cannot always be achieved by the most common and complicated method. The influence of the chosen category for nutrient characterization was investigated, indicating better prediction when a multi-component strategy was used. In contrast, the prediction of single-

component soil properties was less accurate. In addition, the influence of category levels was not as significant as expected when choosing between 3-level, 5-level or 13-level nutrient characterization for some nutrients, which can be used for a more precise nutrient characterization strategy. A comparative analysis was performed be- tween soil from a local farm with similar texture and soils collected from different locations in Slovenia, which gave a better prediction for a local farm. Finally, the influence of principal component analysis was validated using 5, 10, 20 and 50 first principal components, indicating the better performance of machine learning when using the 50 principal components.

**Kingsley John,et al.,[5]** the complete focus is kept on Environmental variables and their impact on the soil and crop.

**Max Grell,et al.,[6]** described the Soil is a complex, living organism which is con- stantly evolving, physically, chemically and biologically. Standard laboratory testing of soil to determine the levels of nitrogen (mainly NH4 + and NO3 -) is infrequent as it is expensive and slow and levels of nitrogen vary on short timescales. Current testing practices, therefore, are not useful to guide fertilization. Over-fertilization with nitrogen fertilizers has damaged the environment and health of soil; yields are declining, while the population continues to rise.

**M.S. Suchithra ,et al., [7]** precision Farming tries to increase productivity of culti- vation of fields by a more efficient utilisation of different agricultural resources (land, seeds, fertilisers and water mainly). Up to now, available systems and tools concen- trate on a specific area of precision farming, requiring an enormous investment by farmers. The objective of the activity is to demonstrate that technology is mature enough to provide cost-effective end-to-end Precision Farming services while show- ing farmers what space technology can do to help their business. Targeted users are arable farms. Traditionally, farming was an activity dedicated to the production of food for human purposes. Nevertheless, in the last years Energy Industry has also become an active player in the field of agriculture due mainly to bio-combustibles. Precision Farming requires an open mind regarding new technologies and their ap- plications, and here is where the Energy Industry is taking some advantage. Also, younger farmers are starting to be aware of the advantages to their business derived of this concept.

**Varshitha D N,et al.,[8]** agriculture is the main occupation of India and more than 50 percentage of people are dependent on agriculture. Research on agriculture will strengthen the economic growth of the country. Technologies play a vital role to bolster the agriculture. Since soil is the main fount of agriculture, there is a need for significant approach to help the farmer to test and monitor the soil and its properties, which will boost the fertility of the soil thereby intensifying the crop growth, also if crop recommendations are imparted to farmers in a proper way, crop yield can be enhanced to meet the growing demand for the food. Proper awareness on soil will benefit the farmers to grow the right and healthy crop. To overcome the disadvan- tages of traditional soil testing practices we are proposing an approach which has Deep learning, an artificial intelligence (AI) technique and IOT features. This helps in getting fast and accurate result. Soil fertility can be calculated by parameters like pH level, temperature, Moisture content of the soil, temperature, humidity and NPK (nitrogen, phosphorus, and potassium) ,organic matter, carbon level. Weather and Climatic conditions along with the soil parameters will help to evaluate the soil fer- tility. The lacking nutrients in the soil and needed nutrients/fertilizers to boost the soil fertility can be suggested to the farmers and also the crops which can be suitably

grown from the given soil sample and nutrients required for all the recommended crops to enhance the yield can be suggested to the farmers.

**Chapter 3**

**PROJECT DESCRIPTION**

**3.1 Existing System**

For fertilizer prediction using soil chemistry there are several existing systems.

There are dynamic model to predict nutrient requirements based on crop growth stages and soil characteristics. It considers the soil nutrient supply, crop nutrient demand, and fertilizer management practices to provide recommendations. Nutrient recommendations based on soil test results, crop nutrient requirements, and the nu- trient content of fertilizers. It considers the nutrient availability in the soil and the nutrient requirements of the crop to provide recommendations. Real-time soil and weather data is taken to optimize fertilizer applications. It considers the nutrient re- quirements of the crop, the nutrient content of the fertilizer, and the soil moisture and temperature to provide recommendations.

**3.2 Proposed System**

The proposed system The proposed system involves the training and testing of data using various algorithms that results better accuracy. The complete system is designed using Python. Different data sets like crop, crop yield dataset, Location, soil and crop nutrients, fertilizer data sets are gathered from other sources like agri- cultural books, agricultural websites. The soil classification using Random Forest algorithm and Support Vector Machine permits a network to be trained from scratch with an out sized information set or fine standardisation Associate in Nursing exist- ing model or creating use of off the shelf K-nearest neighbour features. The random forest with grid search is opted finally for model generation.

**Advantages**

The level of soil is unknown which helps the plant to be healthy growth. As

the best crops are recommended, this can help growers to choose the right crops according to him resources and needs.As fertilizers or nutrients are recommended in all crop proposals, this will help farmers to get the right knowledge to grow the crops and improve productivity and income.

**3.3 Feasibility Study**

Natural resources, including land, soil, water, climate and land suitability. The information is used to model crop production expectations as well as the shape, size and visual features of the project.Agronomy, which includes relevant species, pest and disease issues, sugarcane planning and distribution, fertilizer and pesticide requirements as well as research and development areas. Agricultural production, which includes an agricultural development program and the identification of sus- tainable production programs suitable for the project.

**3.3.1 Economic Feasibility**

The target of the activity is to demonstrate that technology is mature enough to produce efficient end-to-end exactness Farming services whereas showing farm- ers what house technology will do to assist their business. Targeted user’s square measure productive farms. Historically, farming was AN activity dedicated to the assembly of food for human and animal functions. Even so, within the last years Energy business has conjointly become a vigorous player within the field of agri- culture due primarily to bio-combustibles. Exactness Farming needs AN open mind concerning new technologies and their applications, and here is wherever the Energy business is taking some advantage. Also,younger farmer’s square measure beginning to bear in mind of the benefits to their business derived of this idea.

**3.3.2 Technical Feasibility**

Agriculture is the backbone of every farmer. The vast majority of people depen- dent on subsistence farming-related livelihoods. Over the past 30 years, the effects of floods and droughts estimated to cost the country 13.8 billion. If there are no steps to betaken climate change is expected to slow GDP growth by up to 0.9 percent per annum. Farmers especially women are not partners affected by climate-related risks,

which are often detrimental to food security and sustainable livelihoods opportuni- ties. Farmers in the target regions also face limited access water and usually get less. Climate change is a serious threat to agricultural productivity, health and poverty.

**3.3.3 Social Feasibility**

Field equipment, which includes large sums of money and estimates of operating costs for appropriate area development, crop cultivation, harvesting and transport vessels. Communities, including local farmers’ development models, community development projects and appropriate donor funding strategies. Environment, which includes lending compliance with environmental government standards and environ- mental management systems.

**3.4 System Specification**

**3.4.1 Hardware Specification**

•Windows 7/8/10(32-bit or 64-bit).

• Intel Core i3 or later

• 2GB RAM minimum, 8 GB RAM-recommended.

• 2GB of available disk space minimum, 4 GB-Recommended.

**3.4.2 Software Specification**

• Python

• Jupyter Notebook

**3.4.3 Standards and Policies**

**Jupyter Notebook**

Similar to an open source web application, it enables us to exchange and generate documents with real-time code, equations, visuals, and text. It can be applied to machine learning, statistical modelling, data visualisation, and data cleansing and transformation.

**Standard Used: ISO/IEC 27001**

**Google Colab**

Colab is a completely cloud-based Jupyter notebook environment that is free to use. The notebooks you create can be simultaneously modified by your team members, exactly like you edit documents in Google Docs, and most significantly, it doesn’t require any setup. Many well-known machine learning libraries are supported by Colab and are simple to load in your notebook.

**Chapter 4**

**METHODOLOGY**

**4.1 General Architecture**



Figure 4.1: **Architecture Diagram**

The figure 4.1 shows the project general Architecture. Soil samples are collected from various locations within a field, and crop yield data is collected from previous growing seasons. The soil and yield data are cleaned, normalized, and transformed to make them suitable for use in machine learning algorithms. Relevant features, such as soil pH, nutrient content, and crop type, are selected for use in the machine learning models. The selected features and yield data are used to train a machine learning model, such as a decision tree, random forest, or neural network, to predict fertilizer requirements. The trained model is evaluated on a test data set to assess its accuracy and generalization ability. Based on the soil and crop characteristics input by the user, the machine learning model generates a fertilizer recommendation, which is displayed on the user interface.

**4.2 Design Phase**

**4.2.1 Data Flow Diagram**

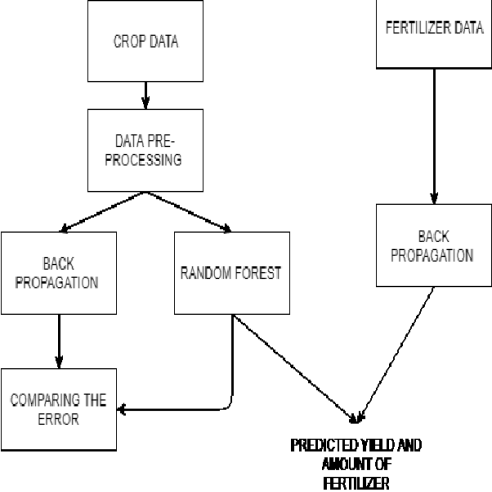


Figure 4.2: **Data Flow Diagram for Fertilizer prediction**

Data Flow Diagram (DFD) in the figure 4.2 shows the flow of the model. This data flow diagram illustrates the main data inputs, processes, and outputs involved in a fertilizer prediction system using machine learning.Soil samples and crop yield data are collected and stored in a database. Data preprocessing where the data is cleaned, normalized, and transformed to make it suitable for use in machine learning algorithms. Model training includes selected features and yield data are used to train a machine learning model to predict fertilizer requirements.Model evaluation evaluates the trained model is evaluated on a test dataset to assess its accuracy and generalization ability.Fertilizer recommendation is finally based on the soil and crop characteristics input by the user, the machine learning model generates a fertilizer recommendation.

**4.2.2 Use Case Diagram**

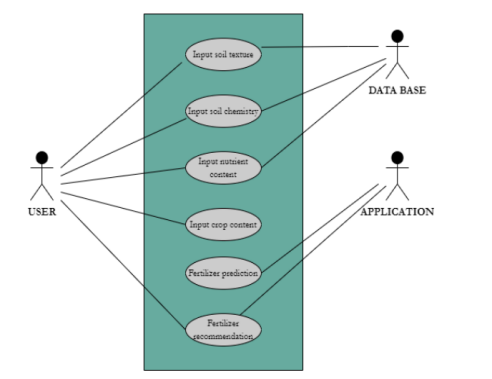


Figure 4.3: **Use Case Diagram**

The Use Case Diagram in figure 4.3 represents the all necessary usecases.The user is the primary actor, who interacts with the system to input soil and crop data, and re- ceive fertilizer recommendations. Fertilizer recommendation allows the user to view the fertilizer recommendation generated by the machine learning model, based on the input soil and crop data.Enter soil data allows the user to input soil data, such as pH, nutrient content, and soil type.Enter crop data allows the user to input crop data, such as crop type, yield, and planting date. Train machine learning model allows the system administrator to train the machine learning model using historical soil and crop data. This use case allows the system administrator to update the machine learning model with new soil and crop data, to improve its accuracy over time. This use case diagram illustrates the main interactions between the user and the system, as well as the key use cases involved in generating fertilizer recommendations using machine learning.

**4.2.3 Class Diagram**

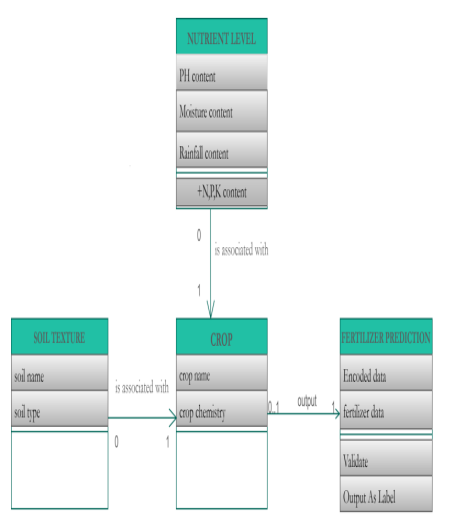


Figure 4.4: **Class Diagram**

Class diagram in figure 4.4 shows classes and their relationships involved in the sys-

tem User Interface soil texture class represents the user interface component of the system, which allows the user to input soil and crop data, and view fertilizer recom- mendations. Soil data represents the soil data input by the user, such as pH, nutrient content, and soil type. Crop Data represents the crop data input by the user, such as crop type, yield, and planting date. Fertilizer Recommendation represents the fertil- izer recommendation generated by the machine learning model, based on the input soil and crop data. Model Evaluation represents the model evaluation component of the system, which assesses the accuracy and generalization ability of the trained machine learning model.

**4.2.4 Sequence Diagram**

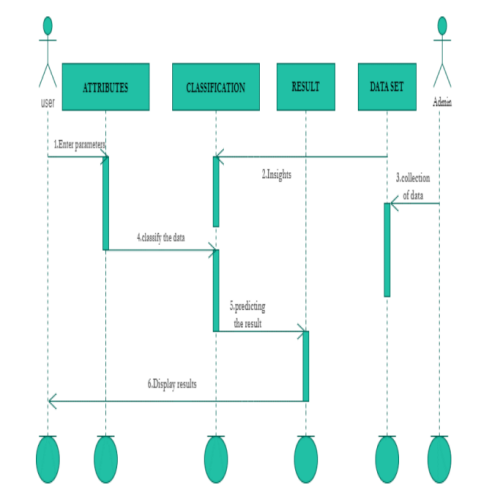


Figure 4.5: **Sequence Diagram**

Fig 4.5 Sequence Diagram or System Sequence Diagram (SSD) shows the The dia- gram shows the sequence of interactions between the user and the system. The user interface component displays input fields for soil and crop data, and prompts the user to input this data. The user enters the soil and crop data into the input fields and sub- mits the data to the system. The data preprocessing component receives the soil and crop data and performs cleaning, normalization, and transformation on the data to make it suitable for use in the machine learning model. The featureselection compo- nent selects relevant features, such as soil pH, nutrient content, and crop type, for use in the machine learning model. The model training component receives the selected features and historical yield data, and trains the machine learning model to predict fertilizer requirements. The machine learning model generates a fertilizer recom- mendation based on the input soil and crop data. The fertilizer recommendation is displayed on the user interface component, along with any additional information,

such as the recommended amount and type of fertilizer.

**4.3 Module Description**

**4.3.1 Module1 : Data Collection and PRE-PROCESSING**

**STEP1:** COLLECTION OF DATA

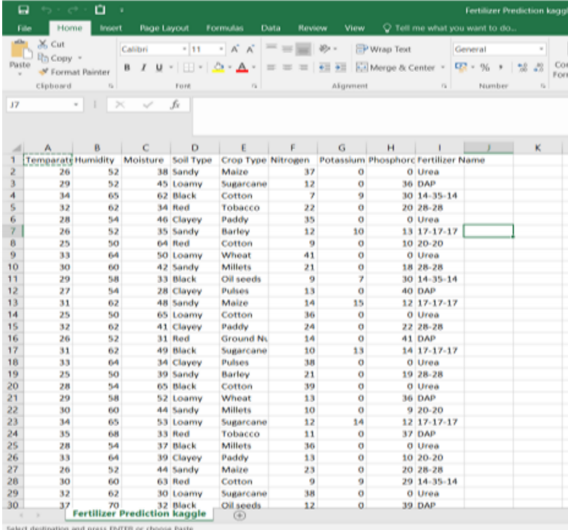


Figure 4.6: **Collection of Data**

Data collection is the process of collecting and measuring information about vari- ables targeted at a fixed system, which allows one to answer relevant questions and evaluate results. Data collection is part of research across all fields of study, includ- ing physical and social science, personality, and business.

**STEP2:** PROCESSING THE DATA

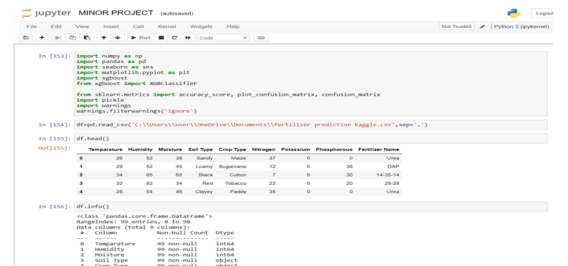


Figure 4.7: **Processing The Data**

Processing the data includes converting raw data into machine-readable form, moving CPU data and memory to output devices, and formatting or converting out- put. Any use of computers to perform descriptive tasks in data can be included under data processing.

**4.3.2 Module2 : Data exploration**

**STEP 1:** VISUALIZATION OF DATA

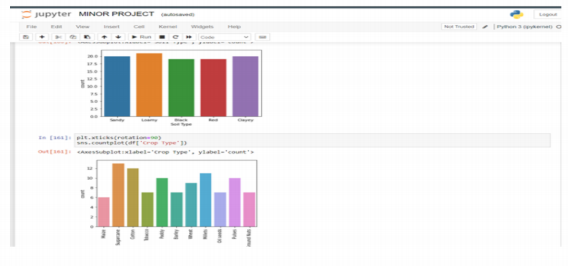
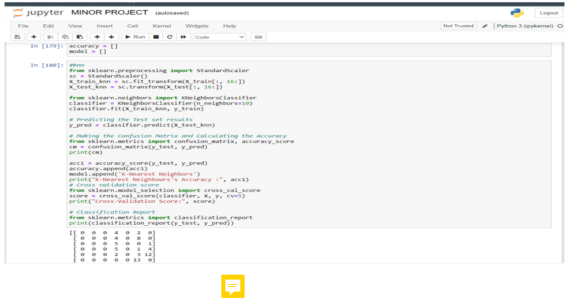


Figure 4.8: **Visualization of Data**

Visualization of data is the process of translating big data sets and metrics into charts, graphs and more. The visual representation of the resulting data makes it easier to identify and share real-time trends, outliers, and new details about the data to be represented.

**STEP2:** APPLYING MACHINE LEARNING ALGORITHMS

Figure 4.9: **Applying Machine Learning Algorithms**

**4.4 Algorithm & Pseudo Code**

**4.4.1 KNN Description**

• K-Nearest Neighbors is a type of division where work is measured only locally and all calculations are reversed until the work is evaluated.

• Import required libraries and) Define a function to that the algorithm will from group of data class

• Identify the class of nearest neighbours and evaluate model

• And the data split was compared with pressent data set

• And make the predictions and calculate the accuracy.

• stop

**4.4.2 Random Forest**

• Random forest is a supervised machine learning algorithm based on ensemble learning

• Ensemble learning is a type of learning where join different types of algorithms or same algorithm multiple times to form a more powerful pre- diction model.

• the algoritham combines multiple algorithms of the same type

• The algorithm was used for classification and regression problems.

• stop

**4.4.3 Decision Tree**

• Decision Tree is a decision support tool that uses a model of decision-like deci- sions and their potential consequences

• Decision tree is a flow-like structure where each internal node represents the ”test” in the attribute

• This algorithm used to make decision for better output

• Basing on the attributes that are predicated decision tree algorithm works

• stop

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

**4.4.4 Pseudo Code**

|  |
| --- |
| Input s :  − Tr a ining d at a s e t ( X tr a in , y tr a in )  − Number of d e c i s i o n tr e e s in th e f o r e s t ( num tree s )  − Maximum depth of each d e c i s i o n tr e e ( max depth )  1 . For each d e c i s i o n tr e e in th e f o r e s t :  a . Sample a random s ub s e t of th e tr a ining d at a s e t with r epl a c em ent  b . Tr ain a d e c i s i o n tr e e on th e s ub s e t of d at a u sing th e CART alg orithm , limiting th e depth of th e tr e e t o max depth  2 . To make a pr e di c ti o n for a new input v e c t or x :  a . For each d e c i s i o n tr e e in th e f o r e s t :  i . Tr a v er s e th e d e c i s i o n tr e e t o d etermin e th e pr e di c t e d c l a s s for x  b . Combine th e pr e di c ti o n s from a ll d e c i s i o n tr e e s in th e f o r e s t t o g e n er a t e a fin a l pr e di c ti o n : i . For c l a s s ifi c a t i o n : u se m aj ority v ote  ii . For r e gr e s s i o n : u se mean or median of pr e di c ti o n s  Output :  − Random f o r e s t model , c o n s i s ting of multipl e d e c i s i o n tr e e s , each tr a in e d on a random s ub s e t of th e tr a ining d at a s e t . |

**4.4.5 Module3**

**STEP 1:** ACCURACY AND ERROR DETECTION

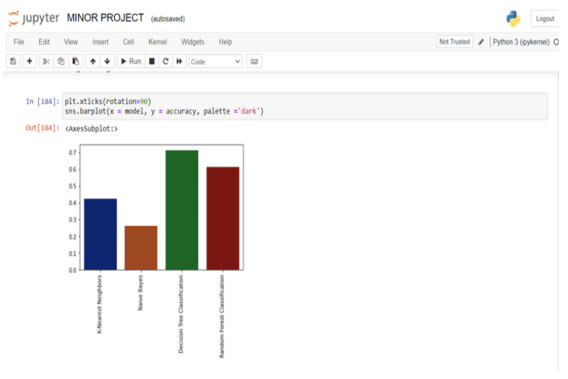


Figure 4.10: **Accuracy and Error Detection**

**STEP 2:** GETTING THE OUTPUT

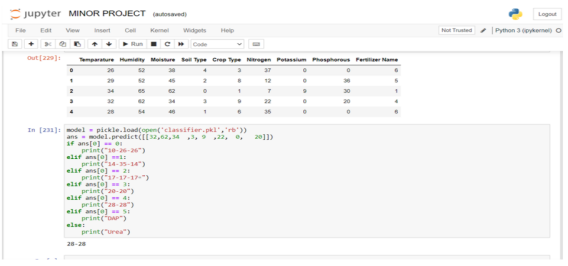


Figure 4.11: **Getting the Output**

Based on the analysis, you need to select the right combination of fertilizers that will address any nutrient deficiencies and improve the overall fertility of the soil.

There are several types of fertilizers available in the market, including organic and inorganic options, and need to choose the one that suits the customer’s needs.Based on the analysis,need to select the right combination of fertilizers that will address any nutrient deficiencies and improve the overall fertility of the soil. There are several types of fertilizers , including organic and inorganic options, and you need to choose the one that suits the customer’s needs

**Chapter 5**

**IMPLEMENTATION AND TESTING**

**5.1 Input and Output**

**5.1.1 Input Design**

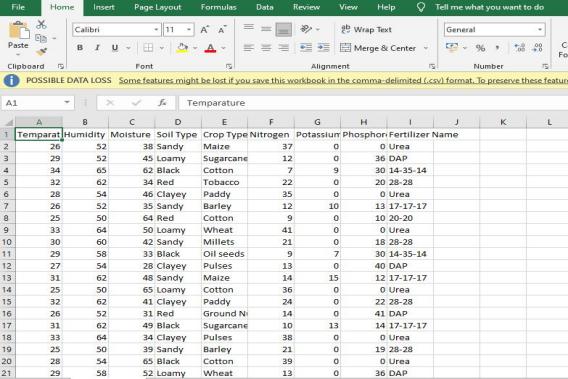


Figure 5.1: **Input design**

In the Input design we had taken the parameters of Soil moisture, Humidity, Temper- ature, Soil type, Crop type, and Fertilizers. These things we considered and given as a input to the design. These are the main aspects for the finding the best results for the crops. Need to collect soil samples from the customer’s field and get them tested for various parameters such as pH, organic matter content, nutrient levels, etc. You may also need to consider other factors such as climate, crop type, and irrigation.

1

2

3

4

**5.1.2 Output Design**

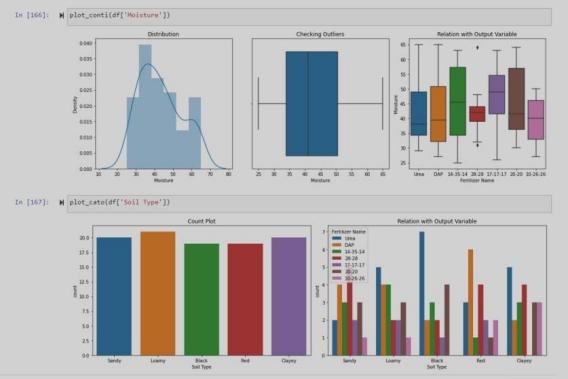


Figure 5.2: **Text Image**

**5.2 Testing**

The complete system is designed using Python. Different data sets like crop, crop yield dataset,Location, soil and crop nutrients, fertilizer data sets are gathered from other sources like agricultural books, agricultural websites. The soil classification using Random Forest algorithm and Support vector Machine. The output of these algorithms shows confusion matrix as summary of algorithms different parameters like Precision, Recall averages and accuracy in percentage.

**5.3 Types of Testing**

**5.3.1 Unit testing**

**Input**

|  |
| --- |
| # n aiv e baye s  from s k l e a rn . n a i v e b ay e s import GaussianNB  c l a s s i fi e r = GaussianNB ( )  c l a s s i fi e r . fi t ( X tr a in , y tr a in ) |

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

|  |  |
| --- | --- |
|  | # Pr e di c ting th e Te st s e t r e s ul t s  y p re d = c l a s s i fi e r . pr e di c t ( X t e s t )  # Making th e Confu sion Matrix and C a l c ul a ting th e Accuracy  from s k l e a rn . m e tri c s import c o nfu s i o n m atrix , a c c ur a c y s c or e  cm = c o nfu s i o n m a trix ( y t e s t , y p re d )  print ( cm )  acc 3 = a c c ur a c y s c or e ( y t e s t , y p re d )  accur acy . append ( acc 3 )  model . append ( ’ Naive Bayes ’ )  print ( ” Naive Bayes ’ s Accuracy : ” , acc 3 )  # Cro s s v a li d a ti o n s c or e  from s k l e a rn . m o d e l s e l e c ti o n import c r o s s v a l s c o r e  s c or e = c r o s s v a l s c o r e ( c l a s s i fi e r , X, y , cv = 5 )  print ( ” Cross − V a lid a ti o n Score : ” , s c or e )  # Cl a s s ifi c a t i o n Report  from s k l e a rn . m e tri c s import c l a s s ifi c a t i o n r e p o r t |
| print ( c l a s s ifi c a t i o n r e p o r t ( y t e s t , y p re d ) ) | |

**Test result**

**Input**

|  |
| --- |
| [ [ 1 0 1 3 0 1 0 ]  [ 0 1 0 4 0 7 0 ]  [ 0 0 0 4 0 2 0 ]  [ 0 0 0 8 0 2 0 ]  [ 0 0 0 10 0 4 3 ]  [ 0 0 0 6 0 6 1 ]  [ 0 0 0 9 0 2 5 ] ]  Naive Bayes ’ s Accuracy : 0 . 2 6 2 5  Cross − V a lid a ti o n Score : [ 0 . 6 0 . 5 5 0 . 9 0 . 4 5 1 . ]  pr e c i s i o n r e c a ll f1 − s c or e supp ort  10 −26 −26 1 . 0 0 0 . 1 7 0 . 2 9 6  14 −35 −14 1 . 0 0 0 . 0 8 0 . 1 5 12  17 −17 −17 0 . 0 0 0 . 0 0 0 . 0 0 6  20 −20 0 . 1 8 0 . 8 0 0 . 3 0 10  28 −28 0 . 0 0 0 . 0 0 0 . 0 0 17  DAP 0 . 2 5 0 . 4 6 0 . 3 2 13  Urea 0 . 5 6 0 . 3 1 0 . 4 0 16  accur acy 0 . 2 6 80 |

1

2

3

4

5

6

7

8

9

10

11

1

2

3

4

5

6

1

2

3

4

5

6

7

8

**5.3.2 Integration testing**

**Input**

|  |
| --- |
| import numpy a s np  import panda s a s pd  import se ab orn a s sn s  import m a t pl o t lib . py pl ot a s pl t  import xg boo st  from xg boo st import X GB Cl a s s i fier  from s k l e a rn . m e tri c s import a c cur a cy s c or e , pl o t c o nfu s i o n m a trix , c o nfu s i o n m a trix import pi c kl e  import w arning s  w arning s . filt e r w a rning s ( ’ ign or e ’ ) |

**Test result**

|  |
| --- |
| Temp ara ture Humidity Moi sture Soil Type CropType Nitr o gen Pota s sium Pho sphorou s F e r t iliz e r Name 0 26 52 38 Sandy Maize 37 0 0 Urea  1 29 52 45 Loamy Sugarcane 12 0 36 DAP  2 34 65 62 Black Cotton 7 9 30 14 3514  3 32 62 34 Red Tobacco 22 0 20 28 28  4 28 54 46 Clayey Paddy 35 0 0 Urea |

**5.3.3 Functional testing**

**Input**

|  |
| --- |
| def pl o t c o nti ( x ) :  fig , axe s = pl t . s ubpl o t s ( nrows = 1 , n c o l s = 3 , fig s iz e = ( 1 5 , 5 ) , t ight l a y o u t = True )  axe s [ 0 ] . s e t t i t l e ( ’ Di s tributi o n ’ )  sn s . di s t pl o t ( x , ax = axe s [ 0 ] )  axe s [ 1 ] . s e t t i t l e ( ’ Checking Outli e r s ’ )  sn s . b ox pl ot ( x , ax = axe s [ 1 ] )  axe s [ 2 ] . s e t t i t l e ( ’ R e l a ti o n with Output V ari abl e ’ )  sn s . b ox pl ot ( y = x , x = df [ ’ F e r t iliz e r Name ’ ] ) |

9

10

11

12

13

14

15

|  |
| --- |
| def pl o t c a t o ( x ) :  fig , axe s = pl t . s ubpl o t s ( nrows = 1 , n c o l s = 2 , fig s iz e = ( 1 5 , 5 ) , t ight l a y o u t = True )  axe s [ 0 ] . s e t t i t l e ( ’ Count Pl o t ’ )  sn s . c o unt pl o t ( x , ax = axe s [ 0 ] )  axe s [ 1 ] . s e t t i t l e ( ’ R e l a ti o n with Output V ari abl e ’ )  sn s . c o unt pl o t ( x = x , hue = df [ ’ F e r t iliz e r Name ’ ] , ax = axe s [ 1 ] ) |

**Test Result**

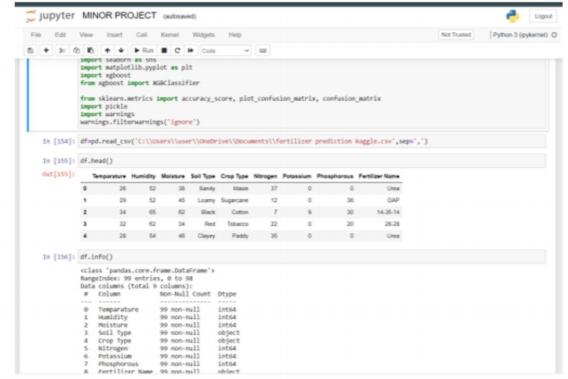


Figure 5.3: **Figure Soil type**

**5.3.4 Test Result**

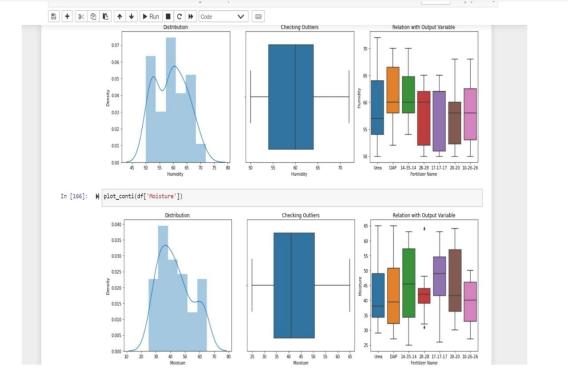


Figure 5.4: **Test Image**

**5.4 Testing Strategy**

A test strategy is a framework that defines the test method of a software development cycle. The purpose of the evaluation strategy is to provide realistic revenue from organizational, high-level objectives to actual assessment activities in order to meet those objectives in terms of quality assurance. The development and writing of an assessment strategy should be done systematically to ensure that all objectives are fully integrated and understood by all stakeholders. It should also be updated fre- quently, challenged and updated as the organization and product change over time. In addition, the evaluation strategy should also aim to align the various stakeholders to ensure quality in terms of names, tests and integration levels, roles andresponsi- bilities, tracking, resource planning, etc. Evaluation strategies describe how product risk for stakeholders is reduced at the level of testing, what types of testing should be performed, and what entry and exit methods work

**Chapter 6**

**RESULTS AND DISCUSSIONS**

**6.1 Efficiency of the Proposed System**

The application of ML algorithms highly depends on the agriculture cycle and the dataset concerned.This section discusses the benefits and limitations of assorted cu- bic centimetre and decilitre algorithms like regression and classification algorithms supported the agriculture cycle concerned. K-nearest neighbour permits a network to be trained from scratch with an out sized information set of fine standardisation As- sociate in Nursing existing model or creating use of off the shelf K-nearest neighbour features. Fine standardisation involves transferring weights of the primary ‘n’ layers learned from a previous base network onto the new network. The dataset obtained for the network is then trained with the transferred connections to perform the spec- ified tasks. KNN will expeditiously learn generic image options and these options is used with classifiers just like the k-means clump etc. to resolve most laptop vision issues. method the method involves setting out the last layer of the trained KNN and exploitation the activation’s of the last connected layer as features the KNN is employed during this stage as a feature extractor rather than a classifier and a clas- sifier is employed for the sorting process. Analysis has shown that this approach is effective for a data set with tiny range of pictures and is additionally according to outmatch each the fine standardisation and coaching from the scratch approach.

**6.2 Comparison of Existing and Proposed System**

Existing fertilizer recommendation systems based on soil chemistry typically use rule-based or expert system approaches, which rely on a set of predefined rules or heuristics to generate recommendations. These systems can be effective for simple scenarios, but may not be able to handle complex interactions between soil, crops, and fertilizers, and may not be able to adapt to new data or changing conditions. In contrast, a fertilizer prediction system using machine learning can learn from histor-

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

ical data and adapt to new conditions, making it more flexible and robust. It can also take into account a wider range of factors, such as weather, irrigation, and crop genet- ics, that may affect fertilizer requirements. Furthermore, machine learning models can often achieve higher accuracy than rule-based systems, especially for complex scenarios.However, the main disadvantage of a machine learning-based fertilizer pre- diction system is that it requires a large amount of high-quality data for training and validation. This data must be representative of the target population and must include a wide range of conditions and scenarios. In addition, machine learning models can be computationally expensive and require specialized hardware and software to train and run.Overall, a fertilizer prediction system using machine learning has the poten- tial to outperform existing rule-based systems in terms of accuracy, flexibility, and adaptability, but requires significant investment in data collection, preprocessing, and model development. In practice, the choice of system depends on the specific needs and resources of the user, and may involve a trade-off between accuracy, complexity, and cost.

**6.3 Sample Code**

|  |
| --- |
| import numpy a s np  import panda s a s pd  import se ab orn a s sn s  import m a t pl o t lib . py pl ot a s pl t  import xg boo st  from xg boo st import X GB Cl a s s i fier  from s k l e a rn . m e tri c s import a c cur a cy s c or e , pl o t c o nfu s i o n m a trix , c o nfu s i o n m a trix import pi c kl e  import w arning s  w arning s . filt e r w a rning s ( ’ ign or e ’ )  df . head ( )  df =pd . r e a d c s v ( ’C:\\ U ser s \\ u s er \\One Drive \\Documents \\ f e r t iliz e r pr e di c ti o n k a gg le . c sv ’ , sep = ’ , ’ ) df . info ( )  df . d e s c rib e ( )  f . apply ( lambda x : l en ( x . unique ( ) ) )  df . i s null ( ) . sum ( )  sn s . c o unt pl o t ( df [ ’ S o il Type ’ ] )  pl t . x ti c k s ( r o t a ti o n = 9 0 )  sn s . c o unt pl o t ( df [ ’ Crop Type ’ ] )  pl t . figur e ( fig s iz e = ( 1 6 , 8 ) )  sn s . c o unt pl o t ( x= ’ F e r t iliz e r Name ’ , d at a = df )  def pl o t c o nti ( x ) :  fig , axe s = pl t . s ubpl o t s ( nrows = 1 , n c o l s = 3 , fig s iz e = ( 1 5 , 5 ) , t ight l a y o u t = True )  axe s [ 0 ] . s e t t i t l e ( ’ Di s tributi o n ’ ) |

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

|  |
| --- |
| sn s . di s t pl o t ( x , ax = axe s [ 0 ] )  axe s [ 1 ] . s e t t i t l e ( ’ Checking Outli e r s ’ )  sn s . b ox pl ot ( x , ax = axe s [ 1 ] )  axe s [ 2 ] . s e t t i t l e ( ’ R e l a ti o n with Output V ari abl e ’ )  sn s . b ox pl ot ( y = x , x = df [ ’ F e r t iliz e r Name ’ ] )  y= df [ ’ F e r t iliz e r Name ’ ] . copy ( )  def pl o t c a t o ( x ) :  fig , axe s = pl t . s ubpl o t s ( nrows = 1 , n c o l s = 2 , fig s iz e = ( 1 5 , 5 ) , t ight l a y o u t = True )  axe s [ 0 ] . s e t t i t l e ( ’ Count Pl o t ’ )  sn s . c o unt pl o t ( x , ax = axe s [ 0 ] )  axe s [ 1 ] . s e t t i t l e ( ’ R e l a ti o n with Output V ari abl e ’ )  sn s . c o unt pl o t ( x = x , hue = df [ ’ F e r t iliz e r Name ’ ] , ax = axe s [ 1 ] )  from s k l e a rn . m o d e l s e l e c ti o n import c r o s s v a l s c or e  s c or e = c r o s s v a l s c o r e ( c l a s s i fi e r , X, y , cv = 5 )  print ( Cro s s V a lid a ti o n Score : , s c or e )  # Making th e Confu sion Matrix and C a l c ul a ting th e Accuracy  from s k l e a rn . m e tri c s imp or c o nfu s i o n m atrix , a c c ur a c y s c or e  cm = c o nfu s i o n m a trix ( y t e s t , y p re d )  print ( cm ) |

**Output**

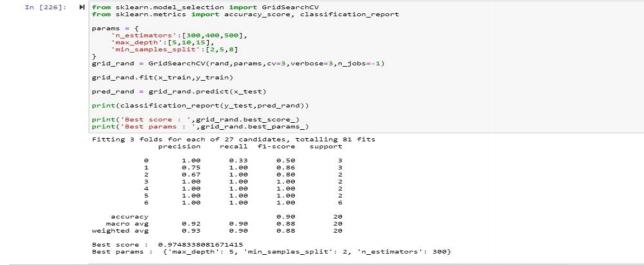


Figure 6.1: **Validated Output**

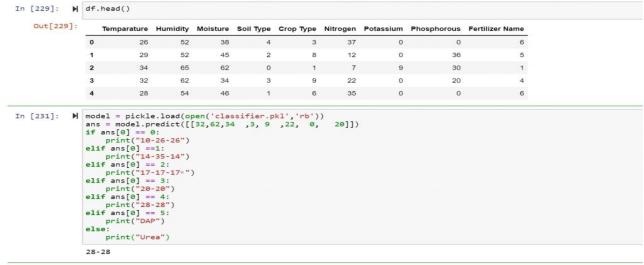


Figure 6.2: **Predicted Output**

**Chapter 7**

**CONCLUSION AND FUTURE**

**ENHANCEMENTS**

**7.1 Conclusion**

This project focuses on soil parameters mostly as fertilizer prediction involves all aspects like type of soil, crop selected and other factors. These are done by using machine learning algorithms. The use of machine learning for fertilizer prediction in agriculture has the potential to revolutionize the way we approach crop management. By leveraging historical data and advanced algorithms, machine learning models can generate accurate and adaptable recommendations for fertilizer application, taking into account a wide range of factors such as soil chemistry, weather, and crop genet- ics. While there are some challenges to overcome, such as the need for large, high- quality data sets and the computational resources required to train and run machine learning models, the benefits of this approach are significant. By providing farmers with accurate and timely recommendations for fertilizer application, we can improve crop yields, reduce waste and environmental impact, and contribute to a more sus- tainable and efficient agricultural industry. Overall, the use of machine learning for fertilizer prediction represents a promising direction for future research and develop- ment in the field of precision agriculture, and has the potential to transform the way we approach crop management for years to come.

**7.2 Future Enhancements**

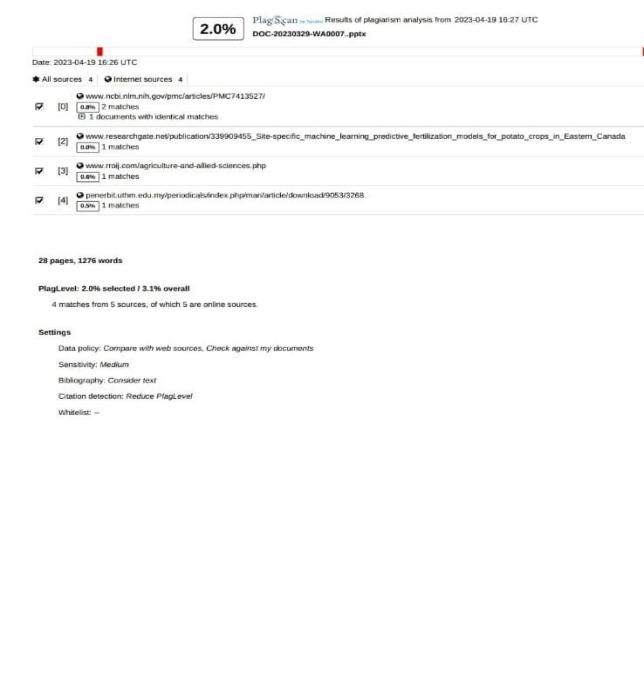
There are several potential enhancements and improvements that could be made to a fertilizer prediction system using machine learning. Currently, most fertilizer prediction systems rely primarily on soil chemistry data. In the future, it may be possible to incorporate additional data sources, such as weather, climate, satellite imagery, and IoT sensor data, to further improve accuracy and adaptability. While

random forest models are effective for many applications, there are more advanced machine learning models, such as deep learning and neural networks, that may pro- vide even higher accuracy and flexibility. These models could be explored and devel- oped further for fertilizer prediction. Currently, fertilizer prediction systems typically require some manual input or oversight from farmers or agronomists. In the future, it maybe possible to automate the entire process, from data collection to recommen- dation generation, using robotics, IoT devices, and other advanced technologies.

**Chapter 8**

**PLAGIARISM REPORT**

ATTACH ONLY SUMMARY PAGE OF PLAGIARISM REPORT



1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

**Chapter 9**

**SOURCE CODE & POSTER**

**PRESENTATION**

**9.1 Source Code**

|  |
| --- |
| import numpy a s np  import panda s a s pd  import se ab orn a s sn s  import m a t pl o t lib . py pl ot a s pl t  import xg boo st  from xg boo st import X GB Cl a s s i fier  from s k l e a rn . m e tri c s import a c cur a cy s c or e , pl o t c o nfu s i o n m a trix , c o nfu s i o n m a trix  import pi c kl e  import w arning s  w arning s . filt e r w a rning s ( ’ ign or e ’ )  df =pd . r e a d c s v ( ’C:\\ U ser s \\hp \\Downloads \\MINOR PROJECT\\ min or d a t a s e t . c sv ’ , sep = ’ , ’ )  df . info ( )  df . d e s c rib e ( )  df . apply ( lambda x : l en ( x . unique ( ) ) )  df . i s null ( ) . sum ( )  sn s . c o unt pl o t ( df [ ’ S o il Type ’ ] )  pl t . x ti c k s ( r o t a ti o n = 9 0 )  sn s . c o unt pl o t ( df [ ’ Crop Type ’ ] )  pl t . figur e ( fig s iz e = ( 1 6 , 8 ) )  sn s . c o unt pl o t ( x= ’ F e r t iliz e r Name ’ , d at a = df )  def pl o t c o nti ( x ) :  fig , axe s = pl t . s ubpl o t s ( nrows = 1 , n c o l s = 3 , fig s iz e = ( 1 5 , 5 ) , t ight l a y o u t = True )  axe s [ 0 ] . s e t t i t l e ( ’ Di s tributi o n ’ )  sn s . di s t pl o t ( x , ax = axe s [ 0 ] )  axe s [ 1 ] . s e t t i t l e ( ’ Checking Outli e r s ’ ) |

sn s . b ox pl ot ( x , ax = axe s [ 1 ] )

36

axe s [ 2 ] . s e t t i t l e ( ’ R e l a ti o n with Output V ari abl e ’ )

37

38

sn s . b ox pl ot ( y = x , x = df [ ’ F e r t iliz e r Name ’ ] )

39

def pl o t c a t o ( x ) :

40

fig , axe s = pl t . s ubpl o t s ( nrows = 1 , n c o l s = 2 , fig s iz e = ( 1 5 , 5 ) , t ight l a y o u t = True )

41

42

axe s [ 0 ] . s e t t i t l e ( ’ Count Pl o t ’ )

sn s . c o unt pl o t ( x , ax = axe s [ 0 ] )

43

axe s [ 1 ] . s e t t i t l e ( ’ R e l a ti o n with Output V ari abl e ’ )

44

sn s . c o unt pl o t ( x = x , hue = df [ ’ F e r t iliz e r Name ’ ] , ax = axe s [ 1 ] )

45

46

pl o t c o nti ( df [ ’ Temp ara ture ’ ] )

47

48

49

pl o t c o nti ( df [ ’ Humidity ’ ] )

50

pl o t c o nti ( df [ ’ Moi sture ’ ] )

51

52

53

pl o t c a t o ( df [ ’ S o il Type ’ ] )

54

# R e l a ti o n of S o il Type and Temperature with Output V ari abl e

55

56

pl t . figur e ( fig s iz e = ( 1 5 , 6 ) )

sn s . b ox pl ot ( x= df [ ’ S o il Type ’ ] , y= df [ ’ Temp ara ture ’ ] , hue = df [ ’ F e r t iliz e r Name ’ ] )

57

58

# R e l a ti o n of Crop Type with Humidity

59

60

pl t . figur e ( fig s iz e = ( 1 5 , 8 ) )

sn s . b ox pl ot ( x= df [ ’ Crop Type ’ ] , y= df [ ’ Humidity ’ ] )

61

62

# R e l a ti o n of Nitr o gen w rt t o Crop Type

63

64

pl t . figur e ( fig s iz e = ( 1 5 , 8 ) )

sn s . b ox pl ot ( x= df [ ’ Crop Type ’ ] , y= df [ ’ Nitr o gen ’ ] )

65

66

#EDA − Pota s sium v a ri a bl e

67

pl o t c o nti ( df [ ’ Pot a s sium ’ ] )

68

69

#EDA − Pho sphorou s v a ri a bl e

70

pl o t c o nti ( df [ ’ Pho sphorou s ’ ] )

71

72

# pr e pr o c e s s ing u sing encoder

73

74

y = df [ ’ F e r t iliz e r Name ’ ] . copy ( )

75

X = df . drop ( ’ F e r t iliz e r Name ’ , a xi s = 1 ) . copy ( )

76

77

from s k l e a rn . compose import Column Transformer

78

from s k l e a rn . pr e pr o c e s s ing import OneHotEncoder

79

c t = Column Transformer ( tr a n s fo rm e r s = [ ( ’ encoder ’ , OneHotEncoder ( ) , [ 3 , 4 ] ) ] , rem ainder = ’ p a s s thr ough ’ ) X = np . arr a y ( c t . fit tr a n s f o rm (X) )

80

81

82

X [ 0 ]

83

84

from s k l e a rn . m o d e l s e l e c ti o n import t r a in t e s t s p li t

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

X tr a in , X t e st , y tr a in , y t e s t = t r a in t e s t s p li t (X, y , t r a in s iz e = 0 . 2 , s huffl e = True , r a nd o m st at e = 4 2 )

# f e a u tr e s c a ling

from s k l e a rn . pr e pr o c e s s ing import St a nd ard S c a l e r

sc = St a nd ard S c a l e r ( )

X tr a in = sc . fit tr a n s f o rm ( X tr a in )

X t e s t = sc . tr a n s form ( X t e s t )

X tr a in [ 0 ]

accur acy = [ ]

model = [ ]

# knn

from s k l e a rn . pr e pr o c e s s ing import St a nd ard S c a l e r

sc = St a nd ard S c a l e r ( )

X tr a in k nn = sc . fit tr a n s f o rm ( X tr a in [ : , 1 6 : ] )

X t e s t kn n = sc . tr a n s form ( X t e s t [ : , 1 6 : ] )

from s k l e a rn . n eighb or s import K N e ighb or s Cl a s s i fi er

c l a s s i fi e r = K N e ighb or s Cl a s s i fi er ( n n e ighb or s = 1 0 )

c l a s s i fi e r . fi t ( X tr a in k nn , y tr a in )

# Pr e di c ting th e Te st s e t r e s ul t s

y p re d = c l a s s i fi e r . pr e di c t ( X t e s t kn n )

# Making th e Confu sion Matrix and C a l c ul a ting th e Accuracy

from s k l e a rn . m e tri c s import c o nfu s i o n m atrix , a c c ur a c y s c or e

cm = c o nfu s i o n m a trix ( y t e s t , y p re d )

print ( cm )

acc 1 = a c c ur a c y s c or e ( y t e s t , y p re d )

accur acy . append ( acc 1 )

model . append ( ’K− N e ar e st Neighbor s ’ )

print ( ”K− N e ar e st Neighbour s ’ s Accuracy : ” , acc 1 )

# Cro s s v a li d a ti o n s c or e

from s k l e a rn . m o d e l s e l e c ti o n import c r o s s v a l s c o r e

s c or e = c r o s s v a l s c o r e ( c l a s s i fi e r , X, y , cv = 5 )

print ( ” Cross − V a lid a ti o n Score : ” , s c or e )

# Cl a s s ifi c a t i o n Report

from s k l e a rn . m e tri c s import c l a s s ifi c a t i o n r e p o r t

print ( c l a s s ifi c a t i o n r e p o r t ( y t e s t , y p re d ) )

# n aiv e baye s

from s k l e a rn . n a i v e b ay e s import GaussianNB

c l a s s i fi e r = GaussianNB ( )

c l a s s i fi e r . fi t ( X tr a in , y tr a in )

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

# Pr e di c ting th e Te st s e t r e s ul t s

y p re d = c l a s s i fi e r . pr e di c t ( X t e s t )

# Making th e Confu sion Matrix and C a l c ul a ting th e Accuracy

from s k l e a rn . m e tri c s import c o nfu s i o n m atrix , a c c ur a c y s c or e

cm = c o nfu s i o n m a trix ( y t e s t , y p re d )

print ( cm )

acc 3 = a c c ur a c y s c or e ( y t e s t , y p re d )

accur acy . append ( acc 3 )

model . append ( ’ Naive Bayes ’ )

print ( ” Naive Bayes ’ s Accuracy : ” , acc 3 )

# Cro s s v a li d a ti o n s c or e

from s k l e a rn . m o d e l s e l e c ti o n import c r o s s v a l s c o r e

s c or e = c r o s s v a l s c o r e ( c l a s s i fi e r , X, y , cv = 5 )

print ( ” Cross − V a lid a ti o n Score : ” , s c or e )

# Cl a s s ifi c a t i o n Report

from s k l e a rn . m e tri c s import c l a s s ifi c a t i o n r e p o r t

print ( c l a s s ifi c a t i o n r e p o r t ( y t e s t , y p re d ) )

# d e c i s i o n tr e e

from s k l e a rn . tr e e import D e c i s i o n Tr e e Cl a s s i fi e r

c l a s s i fi e r = D e c i s i o n Tr e e Cl a s s i fi e r ( c ri t e ri o n = ’ entr opy ’ , r a nd o m st at e = 0 )

c l a s s i fi e r . fi t ( X tr a in , y tr a in )

# Pr e di c ting th e Te st s e t r e s ul t s

y p re d = c l a s s i fi e r . pr e di c t ( X t e s t )

# Making th e Confu sion Matrix and C a l c ul a ting th e Accuracy

from s k l e a rn . m e tri c s import c o nfu s i o n m atrix , a c c ur a c y s c or e

cm = c o nfu s i o n m a trix ( y t e s t , y p re d )

print ( cm )

acc 4 = a c c ur a c y s c or e ( y t e s t , y p re d )

accur acy . append ( acc 4 )

model . append ( ’ Deci si on Tree Cl a s s ifi c a t i o n ’ )

print ( ” Deci si on Tree Cl a s s ifi c a t i o n ’ s Accuracy : ” , acc 4 )

# Cro s s v a li d a ti o n s c or e

from s k l e a rn . m o d e l s e l e c ti o n import c r o s s v a l s c o r e

s c or e = c r o s s v a l s c o r e ( c l a s s i fi e r , X, y , cv = 5 )

print ( ” Cross − V a lid a ti o n Score : ” , s c or e )

# Cl a s s ifi c a t i o n Report

from s k l e a rn . m e tri c s import c l a s s ifi c a t i o n r e p o r t

print ( c l a s s ifi c a t i o n r e p o r t ( y t e s t , y p re d ) )

# random f o r e s t

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

from s k l e a rn . en semble import R a nd o m F or e s t Cl a s s i fi er

c l a s s i fi e r = R a nd o m F or e s t Cl a s s i fi er ( n e s tim a t o r s = 1 0 0 , c ri t e ri o n = ’ gin i ’ , r a nd o m st at e = 4 2 ) c l a s s i fi e r . fi t ( X tr a in , y tr a in )

# Pr e di c ting th e t e s t s e t r e s ul t s

y p re d = c l a s s i fi e r . pr e di c t ( X t e s t )

# Making th e Confu sion Matrix and C a l c ul a ting th e Accuracy

from s k l e a rn . m e tri c s import c o nfu s i o n m atrix , a c c ur a c y s c or e

cm = c o nfu s i o n m a trix ( y t e s t , y p re d )

print ( cm )

acc 5 = a c c ur a c y s c or e ( y t e s t , y p re d )

accur acy . append ( acc 5 )

model . append ( ’Random F or e s t Cl a s s ifi c a t i o n ’ )

print ( ”Random F or e s t Cl a s s ifi c a t i o n ’ s Accuracy : ” , acc 5 )

# Cro s s v a li d a ti o n s c or e

from s k l e a rn . m o d e l s e l e c ti o n import c r o s s v a l s c o r e

s c or e = c r o s s v a l s c o r e ( c l a s s i fi e r , X, y , cv = 5 )

print ( ” Cross − V a lid a ti o n Score : ” , s c or e )

# Cl a s s ifi c a t i o n Report

from s k l e a rn . m e tri c s import c l a s s ifi c a t i o n r e p o r t

print ( c l a s s ifi c a t i o n r e p o r t ( y t e s t , y p re d ) )

pl t . x ti c k s ( r o t a ti o n = 9 0 )

sn s . b a r pl o t ( x = model , y = accuracy , p a l e tt e = ’ dark ’ )

# l a b e l encoding

from s k l e a rn . pr e pr o c e s s ing import Label Encoder

# encoding S o il Type v a ri a bl e

e n c o d e s o il = Label Encoder ( )

df [ ’ S o il Type ’ ] = e n c o d e s o il . fit tr a n s f o rm ( df [ ’ S o il Type ’ ] )

# c r e a ting th e Data Frame

S oil Type = pd . Data Frame ( zip ( e n c o d e s o il . c l a s s e s , e n c o d e s o il . tr a n s form ( e n c o d e s o il . c l a s s e s ) ) , columns = [ ’ Origin a l ’ , ’ Encoded ’ ] )

S oil Type = S oil Type . s e t ind e x ( ’ Origin a l ’ )

S oil Type

enc ode crop = Label Encoder ( )

df [ ’ Crop Type ’ ] = encode crop . fit tr a n s f o rm ( df [ ’ Crop Type ’ ] )

# c r e a ting th e Data Frame

Crop Type = pd . Data Frame ( zip ( enc ode crop . c l a s s e s , enc ode crop . tr a n s form ( encode crop . c l a s s e s ) ) , columns = [ ’ Origin a l ’ , ’ Encoded ’ ] )

Crop Type = Crop Type . s e t ind e x ( ’ Origin a l ’ )

Crop Type

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

e n c o d e fe rt i = Label Encoder ( )

df [ ’ F e r t iliz e r Name ’ ] = e n c o d e fe rt i . fit tr a n s f o rm ( df [ ’ F e r t iliz e r Name ’ ] )

# c r e a ting th e Data Frame

F e r t iliz e r = pd . Data Frame ( zip ( e n c o d e fe rt i . c l a s s e s , e n c o d e fe rt i . tr a n s form ( e n c o d e fe rt i . c l a s s e s ) ) , columns = [ ’ Origin a l ’ , ’ Encoded ’ ] )

F e r t iliz e r = F e r t iliz e r . s e t ind e x ( ’ Origin a l ’ )

F e r t i l i z e r

# s pli t t in g th e d at a int o tr a in and t e s t

from s k l e a rn . m o d e l s e l e c ti o n import t r a in t e s t s p li t

x tr a in , x t e s t , y tr a in , y t e s t = t r a in t e s t s p li t ( df . drop ( ’ F e r t iliz e r Name ’ , a xi s = 1 ) , df [ ’ F e r t iliz e r

Name ’ ] , t e s t s iz e = 0 . 2 , r a nd o m st at e = 1 )

print ( ’ Shape of Spli t t ing : ’ )

print ( ’ x tr a in = {} , y tr a in = {} , x t e s t = {} , y t e s t = {} ’ . form at ( x tr a in . shape , y tr a in . shape ,

x t e s t . shape , y t e s t . shape ) )

x tr a in . info ( )

# random f o r e s t c l a s s i fi e r

rand = R a nd o m F or e s t Cl a s s i fi er ( r a nd o m st at e = 4 2 )

rand . fi t ( x tr a in , y tr a in )

pr e d r and = rand . pr e di c t ( x t e s t )

from s k l e a rn . m o d e l s e l e c ti o n import Grid Search CV

from s k l e a rn . m e tri c s import a c cur a cy s c or e , c l a s s ifi c a t i o n r e p o r t

params = {

’ n e s tim a t o r s ’ : [ 3 0 0 , 4 0 0 , 5 0 0 ] ,

’ max depth ’ : [ 5 , 1 0 , 1 5 ] ,

’ min s a mpl e s s plit ’ : [ 2 , 5 , 8 ]

}

grid r a nd = Grid Search CV ( rand , params , cv = 3 , verb o se = 3 , n j ob s = −1)

grid r a nd . fi t ( x tr a in , y tr a in )

pr e d r and = grid r a nd . pr e di c t ( x t e s t )

print ( c l a s s ifi c a t i o n r e p o r t ( y t e s t , pr e d r and ) )

print ( ’ Be st s c or e : ’ , grid r a nd . b e s t s c o r e )

print ( ’ Be st params : ’ , grid r a nd . b e s t p ar a m s )

y tr a in [ 6 ]

import pi c kl e

pi c kl e o ut = open ( ’ c l a s s i fi e r . p kl ’ , ’wb ’ )

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

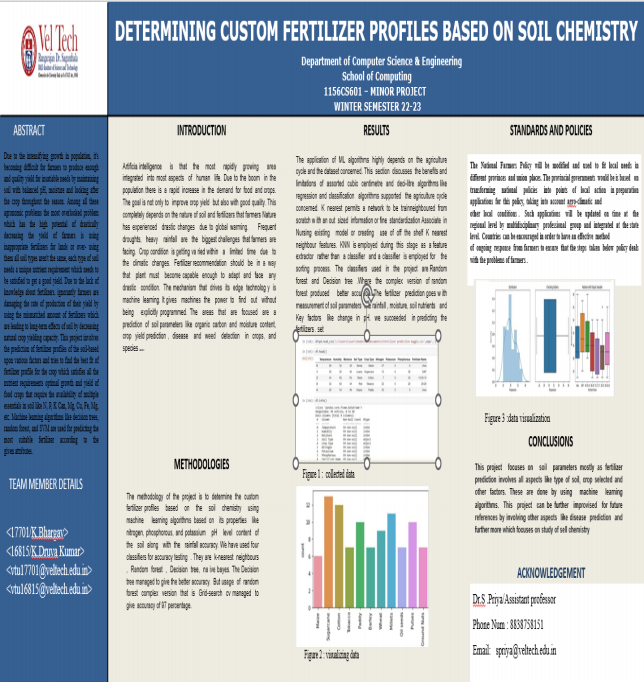
298

299

300

|  |
| --- |
| pi c kl e . dump ( grid r a nd , pi c kl e o ut )  pi c kl e o ut . c l o s e ( )  df . head ( )  model = pi c kl e . l o ad ( open ( ’ c l a s s i fi e r . p kl ’ , ’ rb ’ ) )  an s = model . pr e di c t ( [ [ 3 2 , 6 2 , 3 4 , 3 , 9 , 2 2 , 0 , 2 0 ] ] )  i f an s [ 0 ] == 0 :  print ( ” 10 −26 −26 ” )  e li f an s [ 0 ] = = 1 :  print ( ” 14 −35 −14 ” )  e li f an s [ 0 ] == 2 :  print ( ” 17 −17 −17 ” )  e li f an s [ 0 ] == 3 :  print ( ” 20 −20 ” )  e li f an s [ 0 ] == 4 :  print ( ” 28 −28 ” )  e li f an s [ 0 ] == 5 :  print ( ”DAP” )  e l s e :  print ( ” Urea ” ) |

**9.2 Poster Presentation**



**References**

[1] Abhinav sharma, “Machine Learning Applications for Precision Agriculture: A Comprehensive Review”, 1Department of Electrical and Electronics Engineer- ing, School of Engineering, University of Petroleum and Energy Studies (UPES), Dehradun 248007, india,2021

[2] Aman Kumar Dewangan,” Application of IoT and Machine Learning in Agri- culture”, International Journal of Engineering Research Technology (IJERT) http://www.ijert.org ISSN: 2278- 0181 IJERTV9IS070080 (This work is licensed under a Creative Commons Attribution 4.0 International License.) Published by : www.ijert.org Vol. 9 Issue 07,July-2020

[3] Halima Sadiyah. Abdullahi” Advances of image processing in Precision Agri- culture: “Journal of Multidisciplinary Engineering Science and Technology (JMEST),ISSN: 2458-9403Vol. 4 Issue 8,August - 2017

[4] Janez Trontelj ml. and Olga Chambers:”Machine Learning Strategy for Soil Nu- trients Pre- diction Using Spectroscopic Method”Faculty of Engineering, Nan- jing Agricultural University, Nanjing 210031, PR China,2019

[5] Kingsley JOHN ” Using Machine Learning Algorithms to Estimate Soil Organic Carbon Variability with Environmental Variables and Soil Nutrient Indicators in an Alluvial Soil”, Land, 2020, 9,

[6] Max Grell, Firat Guder,” Determining and Predicting Soil Chemistry with a Pointof-Use Sensor ¨ Toolkit and Machine Learning Model”, Department of Bio- engineering, Imperial College London, London SW7 2AZ, UK ,2021.

[7] M.S. Suchithra , Maya L. Pai,2019,” Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters” Shri Mata Vaishno Devi University, Katra 182320,India,2020

[8] Varshitha D N, Savita Chowdary: “An AI solution for Soil Fertility and Crop Friendliness Detection and Monitoring,Quebec Research and Development Cen- tre, Agriculture and Agri-Food Canada, 2020