AGRI AI: REVOLUTIONISING FARMING PRACTICES THROUGH MACHINE LEARNING ENABLED CROP GUIDANCE AND YIELD ESTIMATION

A Major Project Report Submitted
In partial fulfillment of the requirements for the award of the degree of

Bachelor of Technology
In
Computer Science and Engineering

By

M.KARTHIK KUMAR (20N31A05F9)
M.SUKANYA (20N31A05G0)
MD ARSHAD BASHA (20N31A05E7)
P.VINAY REDDY (20N31A05H4)

Under the esteemed guidance of
Mr.SANDEEP AGARWALLA
Assistant Professor

Department of Computer Science and Engineering

Malla Reddy College of Engineering and Technology

(Autonomous Institution-UGC, Govt. of India)

(Affiliated to JNTUH, Hyderabad, Approved by AICTE, NBA&NAAC with ‘A’ Grade) Maisammaguda, Kompally, Dhulapally, Secunderabad – 500100
Website: www.mrcet.ac.in

2020-2024

Malla Reddy College of Engineering and Technology

(Autonomous Institution-UGC, Govt. of India)

(Affiliated to JNTUH, Hyderabad, Approved by AICTE, NBA&NAAC with 'A' Grade)

Maisammaguda, Kompally, Dhulapally, Secunderabad – 500100

Website: www.mrcet.ac.in

CERTIFICATE

This is to certify that this is the Bonafide record of the project entitled "AGRI AI:REVOLUTIONISING FARMING PRACTICES THROUGH MACHINE LEARNING ENABLED CROP GUIDANCE AND YIELD ESTIMATION", submitted **byM.KARTHIK KUMAR (20N31A05F9) ,M.SUKANYA (20N31A05G0) , MD ARSHAD BASHA (20N31A05E7) ,P.VINAY REDDY (20N31A05H4)** of B.Tech in the partial fulfillment of the requirements for the degree of Bachelor of Technology in Computer Science and Engineering, Department of CSE, during the year 2023-2024. The results embodied in this project report have not been submitted to any other University or Institute for the award of any degree or diploma.

Internal Guide	**Head of the Department**
Mr.SANDEEP AGARWALLA	**Dr. S.SHANTHI**
Assistant Professor	**Professor**

External Examiner

ACKNOWLEDGMENT

We feel honored to place our warm salutation to our college Malla Reddy College of Engineering and Technology (UGC-Autonomous) for allowing us to do this Project as part of our B. Tech Program. We are ever grateful to our Director **Dr. V. S. K Reddy** and Principal **Dr. S. Srinivasa Rao** who enabled us to have experience in engineering and gain profound technical knowledge.

We express our heartiest thanks to our HOD, **Dr. S. Shanthi** for encouraging us in every aspect of our course and helping us realize our full potential.

We would like to thank our Project Guide **Mr.SANDEEP AGARWALLA** for his regular guidance, suggestions, constant encouragement, continuous monitoring, and unflinching cooperation throughout project work.

We would like to thank our Project Coordinator **Mrs. T.PADMAJA** for her regular guidance, suggestions, constant encouragement, continuous monitoring, and unflinching cooperation throughout project work

We would like to thank our Class Incharge **Mr.G.RAVI** who despite being busy with his academic duties took time to guide and keep us on the correct path.

We would also like to thank all the faculty members and supporting staff of the Department of CSE and all other departments who have helped directly or indirectly in making our project a success.

We are extremely grateful to our parents for their blessings and prayers for the completion of our project which gave us the strength to do our project.

With regards and gratitude

M.KARTHIK KUMAR (20N31A05F9)

M.SUKANYA (20N31A05G0)

MD ARSHAD BASHA (20N31A05E7)

P.VINAY REDDY (20N31A05H4)

ABSTRACT

Agriculture serves as the backbone of many economies worldwide, with millions relying on its outputs for sustenance and livelihoods. In the face of climate change and environmental fluctuations, ensuring optimal crop selection and yield prediction becomes paramount for sustainable agricultural practices. Leveraging advancements in machine learning (ML), this project aims to develop a robust system for crop recommendation and yield prediction.The project integrates diverse datasets encompassing soil composition, climate conditions, historical yield records, and agronomic practices. Through data preprocessing and feature engineering, relevant insights are extracted to fuel ML algorithms. For crop recommendation, techniques such as decision trees, random forests, and support vector machines are employed to classify regions and recommend suitable crops based on environmental factors. Concurrently, yield prediction models utilize algorithms like linear regression, gradient boosting machines, and neural networks to forecast crop yields based on historical data and environmental variables. The system features an intuitive user interface accessible via web or mobile platforms, allowing stakeholders to input location-specific information and receive personalized crop recommendations and yield predictions. Continuous monitoring of model performance, coupled with updates based on new data and user feedback, ensures the system's effectiveness and adaptability over time.

By harnessing the power of ML algorithms and data-driven insights, this project endeavors to empower agricultural stakeholders with actionable information for optimizing crop selection, management practices, and productivity, thereby contributing to the advancement of sustainable agriculture and food security.

Keywords:Agriculture, Machine Learning, Crop Recommendation, Yield Prediction, Sustainable Agriculture, Climate Change, Environmental Fluctuations, Decision Trees.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In recent years, agriculture has undergone a significant transformation propelled by advancements in technology and data-driven approaches. One such innovation that has gained momentum is the application of machine learning techniques in crop recommendation and yield prediction systems. With the world's population steadily increasing and agricultural land facing various challenges such as climate change and resource constraints, there is an urgent need for more efficient and sustainable agricultural practices.

Machine learning, a subset of artificial intelligence, offers a promising avenue to address these challenges by leveraging vast amounts of agricultural data to make informed decisions. By analyzing historical crop yields, soil characteristics, weather patterns, and other relevant factors, machine learning models can provide valuable insights into which crops are best suited for specific regions and predict their potential yields with high accuracy.

This project focuses on developing a robust crop recommendation and yield prediction system using machine learning algorithms. By harnessing the power of data analytics and predictive modeling, this system aims to assist farmers, agronomists, and policymakers in making informed decisions to optimize agricultural productivity while minimizing resource inputs and environmental impact.

1.1 PURPOSE AND OBJECTIVES

Purpose:

The purpose of this project is to leverage machine learning techniques to enhance agricultural practices by providing accurate crop recommendations and yield predictions. By analyzing diverse datasets encompassing factors such as soil composition, climate conditions, and historical yield data, the project aims to empower farmers, agronomists, and policymakers with actionable insights to optimize crop selection and improve productivity.

Objectives:

1. Optimizing Crop Selection: Develop machine learning algorithms to analyze geographical and environmental data to recommend the most suitable crops for specific regions. This objective aims to assist farmers in making informed decisions regarding crop selection based on factors like soil type, climate conditions, and market demand.

2. Improving Yield Prediction: Create predictive models using machine learning techniques to forecast crop yields accurately. By incorporating variables such as soil health, weather patterns, and historical yield data, the objective is to enable farmers to anticipate and plan for potential crop yields, facilitating better resource allocation and risk management.

3. Enhancing Decision-Making: Design an intuitive user interface that allows stakeholders to interact with the system effectively. The interface should enable users to input their specific requirements and constraints, visualize recommendations and predictions, and access actionable insights to support decision-making processes related to crop selection and management.

4. Evaluating System Performance: Conduct comprehensive evaluations to assess the accuracy, reliability, and scalability of the crop recommendation and yield prediction system. Performance metrics such as precision, recall, and mean absolute error will be used to measure the effectiveness of the models and identify areas for improvement, ensuring that the system meets the needs of its users effectively.

1.2 EXISTING AND PROPOSED SYSTEM

Existing System

Before the implementation of machine learning-based crop recommendation and yield prediction systems, traditional agricultural practices relied heavily on manual observation, historical knowledge, and basic statistical analysis. While these methods have served agriculture for centuries, they often lack the precision, scalability, and predictive capabilities offered by modern technologies. Here's a breakdown of the existing system:

1. Manual Observation and Expertise: Farmers traditionally relied on their own observations and experience, as well as input from local agricultural experts, to make decisions about crop selection and management. This approach often lacks systematic analysis and may be limited by the knowledge and resources available to individual farmers.

2. Basic Statistical Analysis: Some agricultural decision-making processes incorporate basic statistical analysis of historical data, such as average yields over time. However, these analyses are typically rudimentary and may not capture the complex interactions between various environmental and agronomic factors influencing crop productivity.

3. Rule-Based Systems: In certain cases, rule-based systems have been employed to provide crop recommendations based on predefined rules and thresholds. These systems often lack adaptability and may not account for nuanced variations in environmental conditions or individual farm characteristics.

4. Crop Simulation Models: Agricultural researchers and extension services have developed crop simulation models to estimate crop yields based on mathematical equations and empirical data. While these models offer some predictive capabilities, they are often limited by simplifying assumptions and may not capture the full complexity of real-world agricultural systems.

5. Remote Sensing and GIS Technologies: Remote sensing technologies, including satellite imagery and geographic information systems (GIS), have been used to monitor crop conditions and assess environmental factors affecting agricultural productivity. While these technologies provide valuable insights, they may not always integrate seamlessly with on-the-ground decision-making processes.

Proposed System

The proposed system for crop recommendation and yield prediction using machine learning aims to revolutionize agricultural decision-making by harnessing the power of data-driven approaches. Here's an overview of the proposed system:

1.Data Collection and Preprocessing: The system will gather diverse datasets, including soil composition, climate data, historical yield records, satellite imagery, and agronomic practices. These datasets will be preprocessed to clean, normalize, and standardize the data, ensuring compatibility for machine learning algorithms.

2. Machine Learning Models:

- Crop Recommendation Model: A machine learning model will be developed to recommend suitable crops for specific regions based on input parameters such as soil type, climate conditions, and market demand. This model will employ techniques such as classification algorithms to categorize regions into suitable crop types.

- Yield Prediction Model: Another machine learning model will be constructed to predict crop yields based on various factors, including soil health, weather patterns, crop type, and management practices. Regression algorithms will be used to estimate yield outcomes, providing valuable insights for farmers to optimize resource allocation and management.

3. Model Training and Validation: The machine learning models will be trained using historical data and validated using techniques such as cross-validation to ensure robustness and generalizability. The models will be fine-tuned iteratively to improve performance and accuracy.

4. Integration with User Interface: The system will feature an intuitive user interface accessible via web or mobile platforms. Users, including farmers, agronomists, and policymakers, will interact with the interface to input their location, preferences, and constraints. The interface will display personalized crop recommendations and yield predictions generated by the machine learning models in a user-friendly manner.

5. Performance Monitoring and Updates: The system will include mechanisms for monitoring performance metrics such as prediction accuracy, model drift, and user

feedback. Regular updates and improvements will be implemented based on new data, emerging trends, and user needs to ensure the system remains effective and relevant over time.

6. Scalability and Deployment: The proposed system will be designed for scalability, allowing it to accommodate increasing volumes of data and users. It will be deployable in diverse agricultural settings, ranging from smallholder farms to large-scale commercial operations, with flexibility to adapt to local contexts and requirements.

By integrating advanced machine learning techniques with user-friendly interfaces, the proposed system aims to empower stakeholders across the agricultural value chain with actionable insights to enhance productivity, sustainability, and resilience in agriculture.

1.3 SCOPE OF THE PROJECT

- Data Collection: Gather diverse datasets related to soil, climate, historical yields, and agronomic practices.
- Data Preprocessing: Clean, preprocess, and standardize collected data for machine learning model compatibility.
- Machine Learning Model Development: Develop models for crop recommendation and yield prediction.
- Model Training and Validation: Train and validate models using historical data to ensure accuracy and reliability.
- User Interface Design: Create an intuitive interface for users to input preferences and visualize recommendations and predictions.
- Performance Monitoring: Implement mechanisms to monitor system performance and incorporate regular updates.
- Scalability and Deployment: Design the system for scalability and deployability across various agricultural settings.
- Evaluation and Validation: Conduct thorough evaluations to assess system effectiveness and reliability.

- Documentation and Knowledge Transfer: Document methodologies and findings for knowledge transfer and provide user training and support.
- 10. Ethical and Legal Considerations: Ensure compliance with ethical guidelines and regulations regarding data privacy and responsible use of machine learning algorithms.

CHAPTER 2

LITERATURE SURVEY

[1] Title: **Crop Yield Prediction using Machine Learning Algorithm**

Authors :Mr. D.Jayanarayana Reddy; Mr. M. Rudra Kumar

Agribusiness is the foundation of India's economy, with in excess of 50%of the populace occupied with cultivating. prediction (CYP) independent direction.

[2] Title: **Crop Yield Prediction in Precision Agriculture**

Author: Prof. Dr. Miklós Neményi

Environmental change, environmental change and other ecological variables altogether affect farming well being. Machine learning (ML) assumes a significant part as it is an apparatus for The prediction of crop yield is based on soil, meteorological, environmental, and crop parameters. Decision support models are broadly used to extract significant crop features for prediction. Precision agriculture focuses on monitoring (sensing technologies), management systems, variable rate technologies, and responses to inter- and inter visibility cropping systems.

[3]Title: **Crop Recommender System Using Machine LearningApproach**

Authors: Mr. Shilpa Mangesh Pande; Mr. Prem Kumar Ramesh; Mr.Anmol; Miss. b. Raishwarya; Miss. karuna rohilla

There is no question that farming and enterprises are connected to the jobs of country Indians. This is one of the principle justifications for why peripheral ranchers in India end it all

[4] Title: **Agro Consultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms**

Authors: Mr.Zeel Doshi; Mr.Subhash Nadkarni; Mr.Rashi Agrawal

Horticulture is a significant commitment to the Indian economy. The huge number of individuals living in India relies upon how they live in12 horticulture. Numerous Indian ranchers accept that they can pick plants to plant at a given time.

[5] Title:A Review on Data Mining Techniques for Fertilizer Recommendation,2018

Authors: Miss. Jignasha Miss. M. Jethva, Mr. Nikhil Gondaliya,
Miss.Vinita Shah

At the point when the dirt is insufficient in supplements, add compost to decrease it. A typical issue in farming is excrement determination and fertilizer expansion. Extreme development or absence of manure can harm vegetation and diminish efficiency. This record sums up thedifferent techniques for removing information used to develop a bunch of modern manure soils.

[6] Title: A Survey on Data Mining Techniques in Agriculture, 2015

Author: Miss.M.C.S.Geetha

Horticulture is important for the economies of emerging nations, particularly India. Mining assumes a significant part in decision-production in numerous spaces of farming. It inspects the job of data mining in the farming area and work corresponding to a couple of creators in the rural area. It likewise checks out various methods of acquiring data to resolve numerous agrarian issues. This paper integrates the work of sever all authors in a single place so it is valuable for specialists to get data of current situation of data mining systems and applications in
context to farming field.

[7] Title: AgroNutri Android Application, 2016

Authors: Miss.S.Srija, Miss.R. Geetha Chanda ,Miss.S.Lavanya, Dr. M. Kalpana Ph.D

This paper communicates the idea regarding the making of AgroNutri an android application that helps in conveying the harvest particular fertilizer amount to be applied. The future scope of the AgroNutri is that GPRS can be included so that according to location nutrients are suggested. Further this application would be incorporated as a piece of the accurcy agriculture wherein sensors can be utilized to discover the measure of NPK present in the dirt and that sum can be deducted from the suggestion and giving us the exact measure of supplements to be added.

CHAPTER 3

SYSTEM ANALYSIS

System requirements are the functionality that is needed by a system in order to satisfy the customer's requirements. System requirements are abroad and a narrow subject that could be implemented to many items. The requirements document allows the project team to have a clear picture of what the software solution must do before selecting a vendor. Without an optimized set of future state requirements, the project team has no effective basis to choose The best system for your organization.

3.1 HARDWARE AND SOFTWARE REQUIREMENTS

- Processor
- Intel i3 and above
- RAM : 4GB
- Higher Hard Disk : 500GB: Minimum

3.1.1 SOFTWARE REQUIREMENTS

- Programming Language
- Python IDE : PyCharm
- UML Design : Start UML
- Tools : PIP
- WebServer : Tomcat
- Framework : DJango,
- Google Maps API Database : SQLite

3.2 SOFTWARE REQUIREMENTS SPECIFICATION

3.2.1 FUNCTIONAL REQUIREMENTS

Functional requirements for the crop recommendation and yield prediction project can be categorized into several key areas. Here are the functional requirements:

1. Data Collection and Integration:

The system shall collect diverse datasets related to soil composition, climate conditions, historical yield records, satellite imagery, and agronomic practices from various reliable sources.

The system shall integrate the collected data into a unified database for further analysis.

2. Data Preprocessing:

The system shall preprocess the collected data to clean, normalize, and standardize it for compatibility with machine learning algorithms.

The system shall handle missing values, outliers, and inconsistencies in the data effectively.

3. Crop Recommendation:

The system shall develop machine learning algorithms for crop recommendation based on factors such as soil type, climate conditions, market demand, and historical data.

The system shall classify regions into suitable crop types and recommend crops based on environmental factors and user preferences.

4. Yield Prediction:

The system shall construct machine learning models for yield prediction using variables such as soil health, weather patterns, crop type, and management practices.

The system shall forecast crop yields accurately and reliably based on historical data and environmental factors.

5. User Interface:

The system shall provide an intuitive and user-friendly interface accessible via web or mobile platforms.

The system shall allow users to input location-specific information, preferences, and constraints.

The system shall display personalized crop recommendations and yield predictions in a visually appealing manner.

6. Performance Monitoring and Updates:

The system shall monitor the performance of machine learning models in real-time, including prediction accuracy and stability.

The system shall incorporate regular updates and improvements based on new data, emerging trends, and user feedback.

7. Scalability and Deployment:

The system shall be designed for scalability to accommodate increasing volumes of data and user traffic.

The system shall be deployable across various agricultural settings, ranging from smallholder farms to large-scale commercial operations.

8. Evaluation and Validation:

The system shall conduct thorough evaluations to assess the effectiveness and reliability of crop recommendation and yield prediction models.

The system shall measure performance metrics such as precision, recall, mean absolute error, and user satisfaction.

9. Documentation and Knowledge Transfer:

The system shall document methodologies, algorithms, and findings for knowledge transfer to stakeholders.

The system shall provide training and support for users to effectively utilize the system in agricultural decision-making processes.

10. Ethical and Legal Considerations:

The system shall ensure compliance with ethical guidelines and regulations regarding data privacy, security, and responsible use of machine learning algorithms in agriculture.

The system shall implement measures to protect sensitive information and obtain necessary consent from users for data collection and processing.

These functional requirements outline the core capabilities and features of the crop recommendation and yield prediction system, ensuring its effectiveness, reliability, and usability in agricultural decision-making.

3.2.2 NON FUNCTIONAL REQUIREMENTS

Describe user-visible aspects of the system that are not directly related to the functional behavior of the system. Non-Functional requirements include quantitative constraints, such as response time or accuracy.

- Performance
- Reliability
- Security
- Maintainability
- Compatability
- Usability
- Documentation and Maintenance

CHAPTER 4

SYSTEM DESGIN

4.1 DESCRIPTION

System design is the process of creating a system's architecture, parts, and interfaces to ensure that it satisfies the needs of its users.

Thus, in order to examine the design of this project, we first go through the specifics of establishing the concept of drone detection through a few fundamental modules that would clearly describe the workings of the system that would come from the development.

MACHINE LEARNING

Machine learning (ML) is an area of artificial intelligence (AI) that enables computers to "learn" for themselves over time from training data and develop without explicit programming. Data patterns can be found by machine learning algorithms, which can then use this information to learn and develop their own predictions. Algorithms and models used for machine learning, in essence, gain experience.

In contrast, machine learning is a process that is automated and gives computers the ability to solve issues with little to no human involvement and make decisions based on prior experiences.

While machine learning and artificial intelligence are frequently used synonymously, they are actually two distinct ideas. Machine learning, a subset of AI that enables intelligent systems to autonomously learn new things from data, is what allows intelligent systems to make decisions, gain new abilities, and solve problems in a way that is similar to people. AI is the more general notion..

4.2 SYSTEM ARCHITECTURE

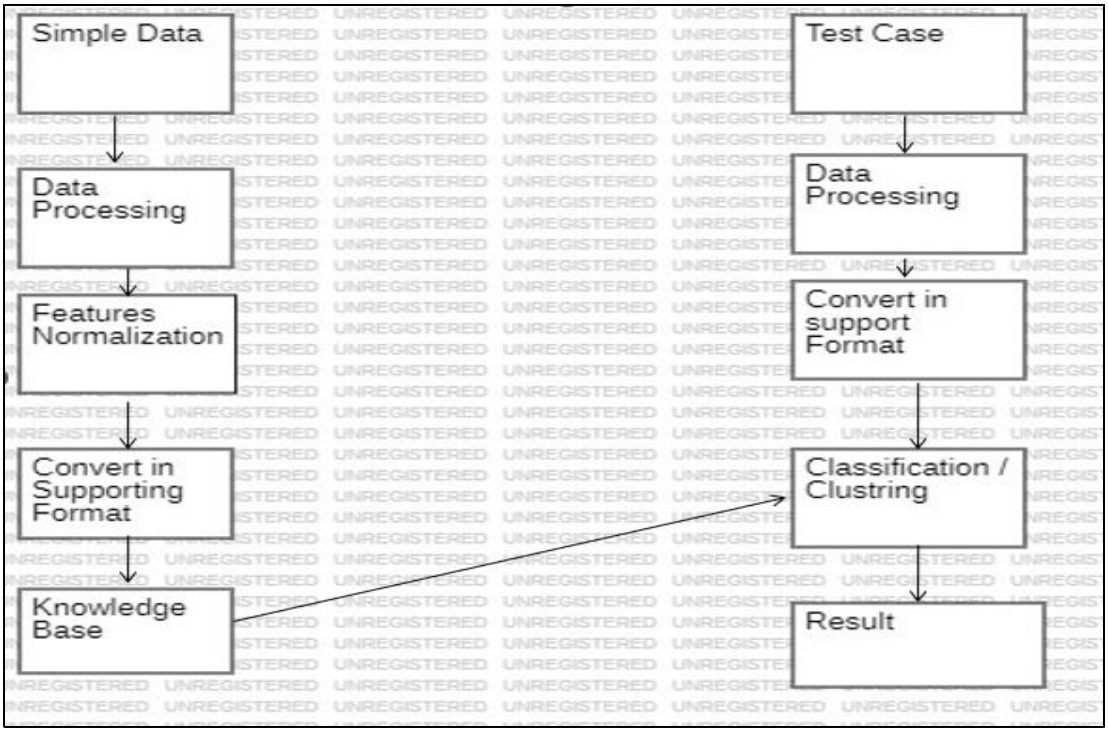

Fig 4.2.1: Architecture

Data Processing is a task of converting data from a given form to a much more usable and desired form i.e. making it more meaningful and informative. Using Machine Learning algorithms, mathematical modeling and statistical knowledge, this entire process can be automated. The output of this complete process can be in any desired form like graphs, videos, charts, tables, images and many more, depending on the task we are performing and the requirements of the machine. This might seem to be simple but when it comes to really big organizations like Twitter, Facebook, Administrative bodies like Parliament, UNESCO and health sector organizations, this entire process needs to be performed in a very structured manner.

Collection:

The most crucial step when starting with ML is to have data of good quality and accuracy. Data can be collected from any authenticated source like data.gov.in, Kaggle or UCI dataset repository. For example, while preparing for a competitive exam, students study from the best study material that they can access so that they learn the best to obtain the best results. In the same way, high-quality and accurate data will make the learning process of the model easier and better and at the time of testing, the model would yield state of the art results.

A huge amount of capital, time and resources are consumed in collecting data. Organizations or researchers have to decide what kind of data they need to execute their tasks or research.

Example: Working on the Facial Expression Recognizer, needs a large number of images having a variety of human expressions. Good data ensures that the results of the model are valid and can be trusted upon.

Preparation:

The collected data can be in a raw form which can't be directly fed to the machine. So, this is a process of collecting datasets from different sources, analyzing these datasets and then constructing a new dataset for further processing and exploration. This preparation can be performed either manually or from the automatic approach. Data can also be prepared in numeric forms also which would fasten the model's learning.

Example: An image can be converted to a matrix of N X N dimensions, the value of each cell will indicate image pixel.

Processing:

This is the stage where algorithms and ML techniques are required toper form the instructions provided over a large volume of data with accuracy and optimal computation.

Output:

In this stage, results are procured by the machine in a meaningful manner which can be inferred easily by the user. Output can be in the form of reports, graphs, videos, etc

Data Preprocessing for Machine learning in Python

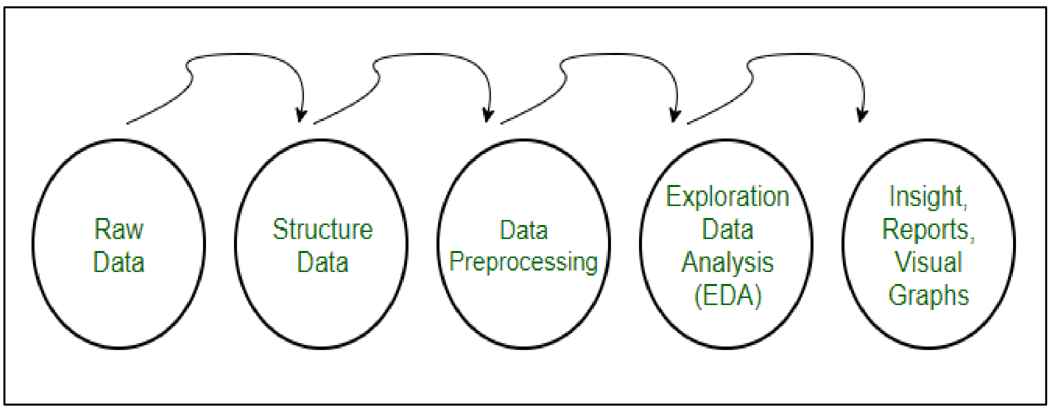

Fig 4.2.2:data processing in machine learning

• Pre-processing refers to the transformations applied to our data before feeding it to the algorithm.

• Data Preprocessing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the
data is gathered from different sources it is collected in raw format which is not feasible for the analysis.

Need of Data Preprocessing

• For achieving better results from the applied model in Machine Learning projects the format of the data has to be in a proper manner. Some specified Machine Learning model needs information in a specified format, for example, Random Forest algorithm does not support null values ,therefore to execute random forest algorithm null values have to be managed from the original raw data set.

• Another aspect is that data set should be formatted in such a way that more than one Machine Learning and Deep Learning algorithms are executed in one data set, and best out of them is chosen. Rescale Data

- When our data is comprised of attributes with varying scales ,many machine learning algorithms can benefit from rescaling the attributes to all have the same scale.
- This is useful for optimization algorithms in used in the core of machine learning algorithms like gradient descent.
- It is also useful for algorithms that weight inputs like
- regression and neural networks and algorithms that use
- distance measures like K-Nearest Neighbors.
- We can rescale your data using scikit-learn using the
- Min Max Scaler class.

Binarize Data (Make Binary)

- We can transform our data using a binary threshold. All values above the threshold are marked 1 and all equal to or
- below are marked as 0.
- This is called binarizing your data or threshold your data. It can be useful when you have probabilities that you want to
- make crisp values. It is also useful when feature engineering
- and you want to add new features that indicate something
- meaningful.
- We can create new binary attributes in Python using scikit-learn
- with the Binarizer class.

Standardize Data

- Standardization is a useful technique to transform attributes with a Gaussian distribution and differing means and standard deviations to a standard Gaussian distribution with a mean of 0and a standard deviation of 1.
- We can standardize data using scikit-learn with the Standard Scaler class.

Data Cleansing

Introduction:

Data cleaning is one of the important parts of machine learning. It plays a significant part in building a model. Data Cleaning is one of those things that everyone does but no one really talks about. It surely isn't the fanciest part of machine learning and at the same time, there aren't any hidden tricks or secrets to uncover. However, proper data cleaning can make or break your project. Professional data scientists usually spend a very large portion of their time on this step. Because of the belief that, ―Better data beats fancier algorithms .If we have a well-cleaned dataset, we can get desired results even with avery simple algorithm, which can prove very beneficial at times.

Obviously, different types of data will require different types of cleaning. However, this systematic approach can always serve as a good starting point. Steps involved in Data Cleaning

1.Removal of unwanted observations

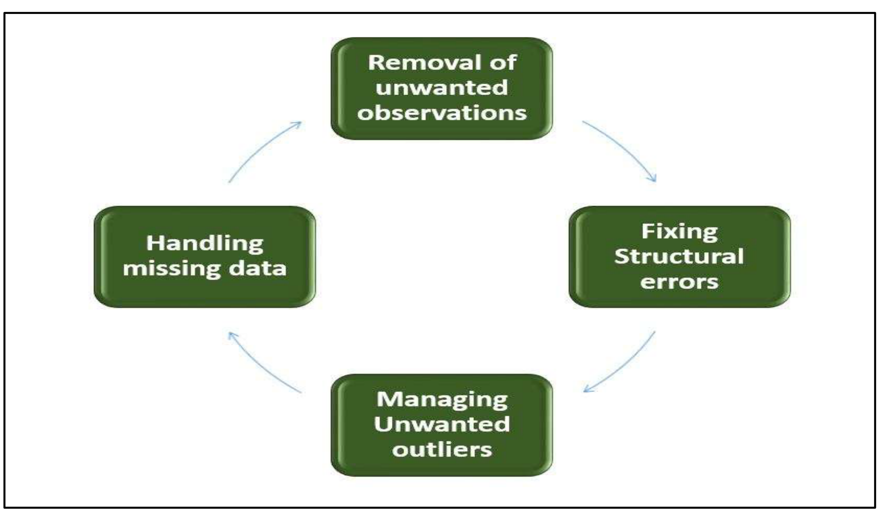

Fig 4.2.3 : process of removal of unwanted observations

This includes deleting duplicate/ redundant or irrelevant values from your dataset. Duplicate observations most frequently arise during data collection and irrelevant observations are those that don't actually fit the specific problem that you're trying to solve.

Redundant observations alter the efficiency by a great extent as the data repeats and may add towards the correct side or towards the incorrect side, thereby producing unfaithful results.

Irrelevant observations are any type of data that is of no use to can be removed directly.

2. **Fixing Structural errors**

The errors that arise during measurement transfer of data or other similar situations are called structural errors. Structural errors include typos in the name of features, same attribute with different name, mislabeled classes, i.e. separate classes that should really be the same or inconsistent capitalization.

- For example, the model will treat America and america as different classes or values, though they represent the same

value or red, yellow and red-yellow as different classes or attributes, though one class can be included in other two classes. So, these are some structural errors that make our model inefficient and gives poor quality results.

3. **Managing Unwanted outliers**

Outliers can cause problems with certain types of models. For example,

linear regression models are less robust to outliers than decision tree models. Generally, we should not remove outliers until we have a legitimate reason to remove them. Sometimes, removing them improves performance, sometimes not. So, one must have a good reason to remove the outlier, such as suspicious measurements that are unlikely to be the part of real data.

4. **Handling missing data**

Missing data is a deceptively tricky issue in machine learning. We cannot just ignore or remove the missing observation. They must be handled carefully as they can be an indication of something important. The two most common ways to deal with missing data are:

1. Dropping observations with missing values.

Dropping missing values is sub-optimal because when you drop observations, you drop information.

The fact that the value was missing may be informative in itself.

Plus, in the real world, you often need to make predictions on new data even if some of the features are missing!

2. Imputing the missing values from past observations.

Imputing missing values is sub-optimal because the value was originally missing but you filled it in, which always leads to a loss in information, no matter how sophisticated your imputation method is.

- Again, ―missingness‖ is almost always informative in itself, and you should tell your algorithm if a value was missing.

Even if you build a model to impute your values, you're not addingany real information. You're just reinforcing the patterns already provided by other features.

Both of these approaches are sub-optimal because dropping an observation means dropping information, thereby reducing data and imputing values also is sub-optimal as we fill the values that were not present in the actual dataset, which leads to a loss of information.

Missing data is like missing a puzzle piece. If you drop it, that's like pretending the puzzle slot isn't there. If you impute it, that's like trying to squeeze in a piece from somewhere else in the puzzle.

So, missing data is always informative and indication of something important. And we must aware our algorithm of missing data by flagging it. By using this technique of flagging and filling, you are essentially allowing the algorithm to estimate the optimal constant for missingness, instead of just filling it in with the mean.

4.3 UML DIAGRAMS

UML Diagrams are classified into different types such as

1. USECASE Diagram
2. SEQUENCE Diagram
3. ACTIVITY Diagram
4. CLASS Diagram

4.3.1 USECASE Diagram

Use case diagrams are the main building block of any object-oriented solution. It displays a system's classes, along with each class's properties, operations, and relationships to other classes. Most modelling tools include three elements to a class. Name is at the top, followed by attributes, then operations or methods, and finally, methods. Classes are linked together to generate class diagrams in a complex system with numerous related classes. Various sorts of arrows represent different relationships between classes.

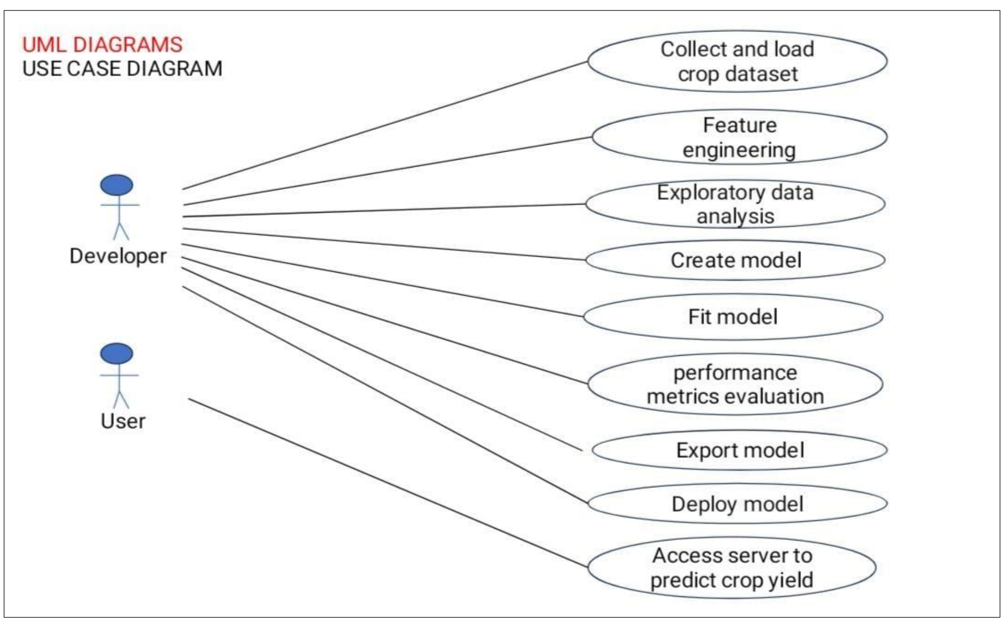

Fig. 4.3.1 usecase Diagram

4.3.2 Sequence diagram

The term "state machine" refers to a device that distinguishes between various states of an object, each of which is governed by either internal or external events. As described in the following chapter, an activity diagram is a specific type of state chart diagram. It is used to model an object's lifetime because state chart diagram defines the states.

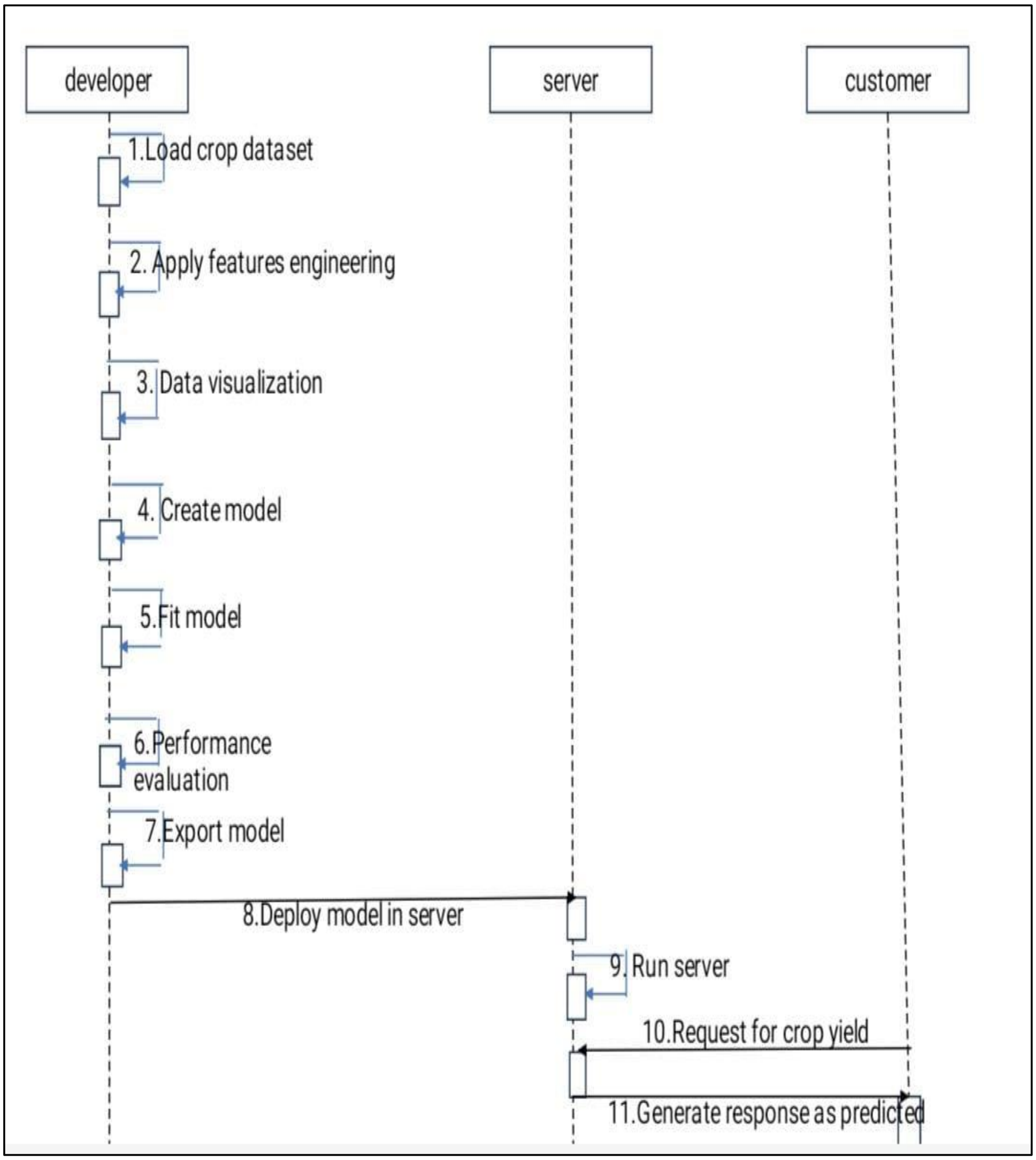

Fig. 4.3.2 Sequence Diagram

4.3.3.1Activity diagram

To depict a system's dynamic behavior, use case diagrams are often employed. Using use cases, actors, and their interactions, it captures the functionality of the system. A system or subsystem of an application's necessary duties, services, and operations are modelled. It shows a system's high- level functionality as well as how a user interacts with that system. Figure for Case

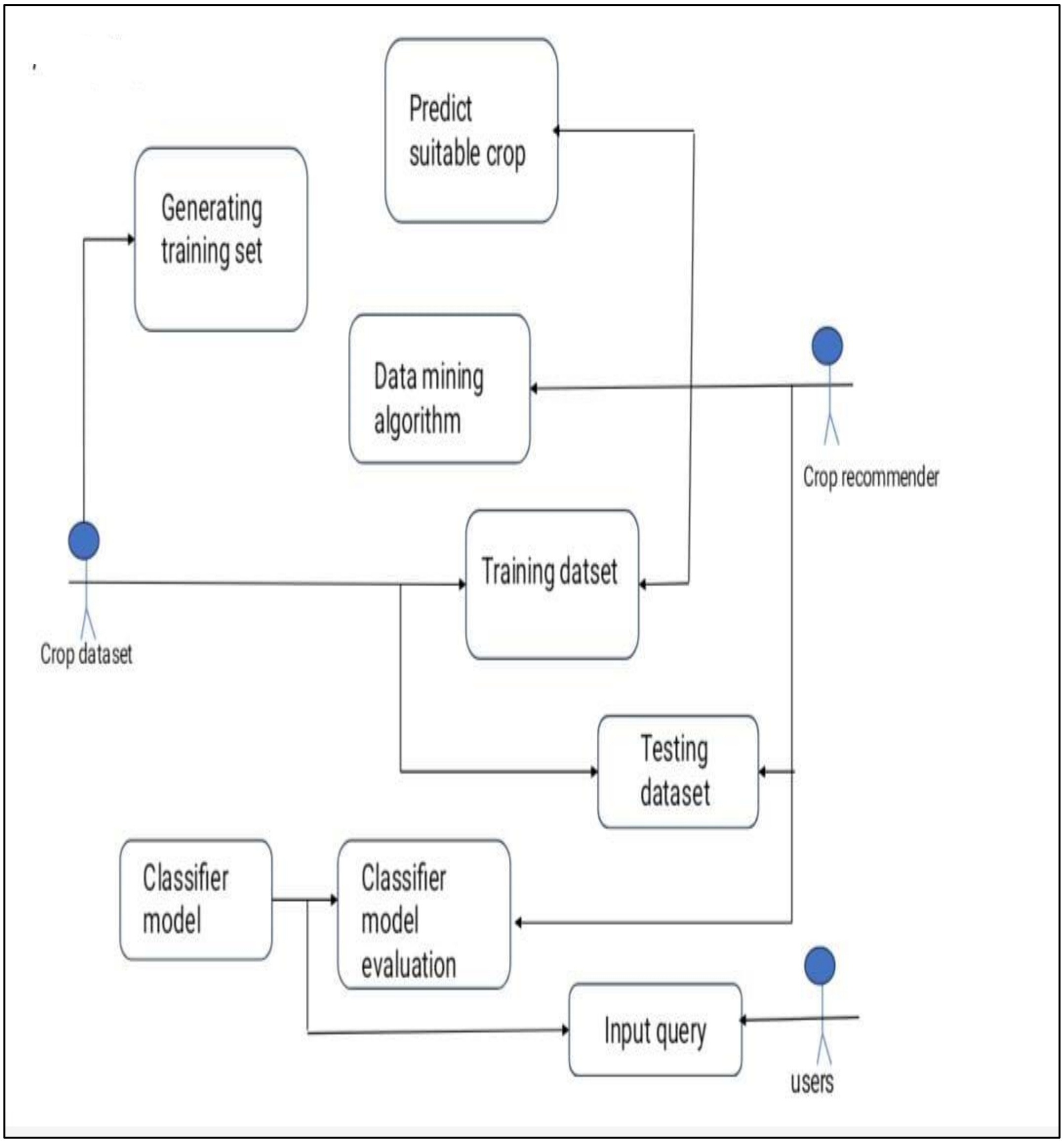

Fig. 4.3.3 Activity Diagram

4.3.3 CLASS Diagram

A data-flow diagram is a visual representation of how data moves through a system or a process (usually an information system). The data flow diagram also shows the inputs and outputs of each entity as well as the process itself. A data-flow diagram lacks control flow, loops, and decision- making processes. With a flowchart, certain operations based on the data can be depicted.

The flowchart can be used to understand how the data flows in this project. A video clip from a PTZ camera is used as the input in this case, and the number of frames on which the algorithm operates will later be converted. The result will be a picture in which a drone is recognized and shown using a rectangle frame around it.

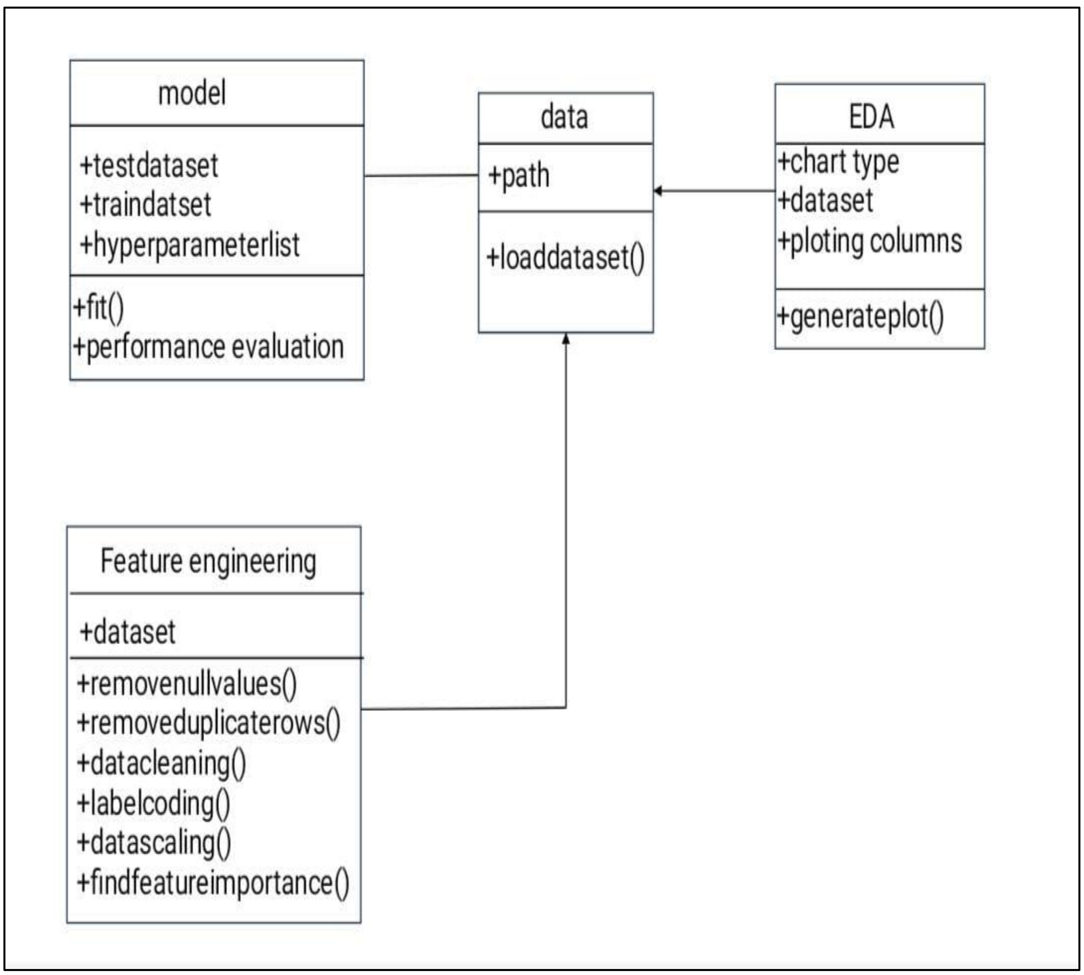

Fig. 4.3.4: CLASS Diagram

CHAPTER 5

METHODOLOGY

5.1 TECHNOLOGIES USED

The project based on crop recommendation and yield prediction using machine learning involves a variety of technologies to gather, process, analyze, and present data effectively. Here are the key technologies typically used in such a project:

1. Programming Languages:

- Python: Widely used for machine learning tasks due to its extensive libraries such as scikit-learn, TensorFlow, and Keras.
- R: Another popular language for statistical computing and data analysis, often used for exploratory data analysis and modeling.

2. Machine Learning Libraries:

- scikit-learn: Offers a wide range of machine learning algorithms and tools for tasks such as classification, regression, clustering, and dimensionality reduction.
- TensorFlow: An open-source machine learning framework developed by Google for building and training deep learning models.
- Keras: High-level neural networks API, built on top of TensorFlow, designed for fast experimentation and prototyping of deep learning models.

3. Data Processing and Analysis:

- Pandas: Python library for data manipulation and analysis, particularly useful for handling structured data such as CSV files and databases.
- NumPy: Fundamental package for numerical computing in Python, providing support for large multi-dimensional arrays and matrices.
- Matplotlib and Seaborn: Python libraries for creating static, animated, and interactive visualizations to explore and present data.

4. Data Storage and Management:

- SQL and NoSQL Databases: Utilized for storing and retrieving structured and unstructured data efficiently. Examples include MySQL, PostgreSQL (SQL), MongoDB, and Cassandra (NoSQl)
- data Lakes: Storage repositories that hold vast amounts of raw data in its native format until needed. Technologies like Amazon S3 and Apache Hadoop are commonly used for this purpose.

5. Web Development (User Interface):

- HTML, CSS, JavaScript: Standard web technologies for building user interfaces and adding interactivity to web applications.
- Flask or Django (Python) / Node.js (JavaScript): Web frameworks used for developing server-side logic and APIs to interact with machine learning models.
- React, Angular, or Vue.js: Front-end frameworks/libraries for building dynamic and responsive user interfaces.

6. Cloud Computing Platforms:

- Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP):** Provide scalable and cost-effective infrastructure for deploying machine learning models, storing data, and running web applications.

7. Version Control:

- Git: Distributed version control system used for tracking changes in code, collaborating with team members, and maintaining project history.

8. Integrated Development Environments (IDEs):

- Jupyter Notebook, PyCharm, VSCode: Popular IDEs for developing and testing machine learning models, analyzing data, and writing code efficiently.

By leveraging these technologies effectively, the project can efficiently handle data, develop accurate machine learning models, and deliver user-friendly interfaces for crop recommendation and yield prediction.

Pandas: Pandas is an open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures. The name Pandas is derived from the word Panel Data – an Econometrics from Multidimensional data. In 2008, developer Wes McKinney started developing pandas when in need of high performance, flexible tool for analysis of data. Prior to Pandas, Python was majorly used for data munging and preparation. It had very little contribution towards data analysis. Pandas solved this problem. Using Pandas, we can accomplish five typical steps in the processing and analysis of data, regardless of the origin of data — load, prepare, manipulate, model, and analyze. Python with Pandas is used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc..

5.2 MODULE DESCRIPTION

MODULES

- Data Collection
- Preprocessing Step
- Feature Extraction
- Data Prediction

Data Collection

Data is composed from a different source and optimized for data sets.And the data is used to evaluate descriptively. Several abstract online outlets, like Kaggle, Google weather forestation and data government,provide the data for up to 10years in series. The data sets such as soil nature, climatic conditions and seed data are used for the crop prediction and better crop yields.

Preprocessing Step

Preprocessing the data is considered as a significant step machine learning phase. Preprocessing involves adding the missing values, the correct set of data, and extracting the functionality. Data set form is
important to the process of analysis. The data collected in this step will induced in Google Colab platform in the form of python programming in order to get the desired output.

Feature Extraction

Extraction of the features would reduce the data size involved to characterize a wide collection of data. The characteristics of soil, crop and weather collected from the pretreatment process establish the final
training data collection. This approach selects the features based on the correlation matrix i.e. the features that has more correlation value is selected as an important predictive function for yield.

Data Prediction

In advance to this step there need to split the data into train dataset and test dataset. By applying the random forest the data is trained with available input and output data. In the test phase, the data are tested ifthe accuracy of e model is satisfied. Then the new data is predicted by machine learning module.

UML DIAGRAMS

UML is simply anther graphical representation of a common semantic model. UML provides a comprehensive notation for the full lifecycle of object-oriented development.

ADVANTAGES

To represent complete systems (instead of only the software portion)using object oriented concepts

To establish an explicit coupling between concepts and executable code

To take into account the scaling factors that are inherent to complex and critical systems

To creating a modeling language usable by both humans and machines UML defines several models for representing systems

The class model captures the static structure

The state model expresses the dynamic behavior of objects

The use case model describes the requirements of the user

The interaction model represents the scenarios and messages flow

5.3 ALGLORITHM

Decision Tree Classifier

Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, , but mostly it is preferred for solving Classification problems. The decision tree classifier creates the classification model by building a decision tree. Each node in the tree specifies a test on an attribute, each branch descending from that node corresponds to one of the possible values for that attribute.

Linear Regression

Linear Regression is the process of finding a line that best fits the datapoints available on the plot, so that we can use it to predict output values for inputs that are not present in the data set we have, with the belief that those outputs would fall on the line. The Statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variable. Linear regression algorithm shows a linear relationship between a dependent (y) and one or more independent (y) variables, hence called as linear regression.

1. Crop Recommendation:
   - Decision Trees
   - Random Forest
   - Support Vector Machines (SVM)
   - Neural Networks
2. Yield Prediction:
   - Linear Regression
   - Gradient Boosting Machines (GBM)
   - Long Short-Term Memory (LSTM)
   - Random Forest Regression
3. Data Mining and Feature Selection:
   - Principal Component Analysis (PCA)
   - Correlation Analysis

CHAPTER 6

IMPLEMENTATION

6.1SAMPLE OUTPUT

Libraries used

1. scikit-learn (sklearn)

2. TensorFlow

3. Keras

4. Pandas

5. NumPy

6. Matplotlib

7. Seaborn

8. Flask or Django (for web development)

Class Core:

CROP PREDICTION

```
<?php
include ('fsession.php');
ini_set('memory_limit', '-1');
if(!isset($_SESSION['farmer_login_user'])){
header("location: ../index.php");} // Redirecting To Home Page
$query4 = "SELECT * from farmerlogin where email='$user_check'";
$ses_sq4 = mysqli_query($conn, $query4);
$row4 = mysqli_fetch_assoc($ses_sq4);
$para1 = $row4['farmer_id'];
$para2 = $row4['farmer_name'];
?><!DOCTYPE html>
<html>
<?php include ('fheader.php');  ?>
```

```
<body class="bg-white" id="top">
<?php include ('fnav.php');  ?
<section class="section section-shaped section-lg">
<div class="shape shape-style-1 shape-primary">
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
</div>
<!--
<div class="container ">
<div class="row">
<div class="col-md-8 mx-auto text-center">
<span class="badge badge-danger badge-pill mb-3">Prediction</span>
</div>
</div>
<div class="row row-content">
<div class="col-md-12 mb-3">
<div class="card text-white bg-gradient-success mb-3">
<div class="card-header">
<span class=" text-success display-4" > Crop Prediction  </span>
</div>
<div class="card-body text-dark">
<form role="form" action="#" method="post" >
<table class="table table-striped table-hover table-bordered bg-gradient-white text-center
display" id="myTable">
<thead>
<tr class="font-weight-bold text-default">
```

```
<th><center> State</center></th>
<th><center>District</center></th>
<th><center>Season</center></th>
<th><center>Prediction</center></th>
</tr>
</thead>
<tbody>
<tr class="text-center">
<td>
<div class="form-group">
<select    onchange="print_city('state',    this.selectedIndex);"    id="sts"    name    ="stt"
class="form-control" required></select>
<script language="javascript">print_state("sts");</script>
</div>
</td>
<td>
<div class="form-group ">
<select id ="state" name="district" class="form-control" required>
<option value="">Select District</option>
</select>
<script language="javascript">print_state("sts");</script>
</div>
</td>
<td>
<div class="form-group ">

<select name="Season" class="form-control">
<option value="">Select Season ...</option>
<option name="Kharif" value="Kharif">Kharif</option>
<option name="Whole Year" value="Whole Year">Whole Year</option>
<option name="Autumn" value="Autumn">Autumn</option>
<option name="Rabi" value="Rabi">Rabi</option>
<option name="Summer" value="Summer">Summer</option>
<option name="Winter" value="Winter">Winter</option>
```

```
</select>
</div>
</td>
<td>
<center>
<div class="form-group ">
<button type="submit" name="Crop_Predict" class="btn btn-success btn-submit">Predict</button>
</div>
</center>
</td>
</tr>
</tbody>
</table>
</form>
</div>
</div>
<div class="card text-white bg-gradient-success mb-3">
<div class="card-header">
<span class=" text-success display-4" > Result  </span>
</div>
<h4>
<?php
if(isset($_POST['Crop_Predict'])){
$state=trim($_POST['stt']);
$district=trim($_POST['district']);
$season=trim($_POST['Season']);
echo "Crops grown in ".$district." during the ".$season." season are :- ";
$JsonState=json_encode($state);
$JsonDistrict=json_encode($district);
$JsonSeason=json_encode($season);

$command=escapeshellcmd("pythonML/crop_prediction/ZDecision_Tree_Model_Call.py $JsonState $JsonDistrict $JsonSeason");
```

```
$output = passthru($command);
echo $output;
}
?>
</h4>
</div>
</div>
</div>
</div>
</section>
<?php require("footer.php");?>
</body>
</html>
```

YIELD PREDICTION

```
<?php
include ('fsession.php');
ini_set('memory_limit', '-1');
if(!isset($_SESSION['farmer_login_user'])){
header("location: ../index.php");} // Redirecting To Home Page
$query4 = "SELECT * from farmerlogin where email='$user_check'";
$ses_sq4 = mysqli_query($conn, $query4);
$row4 = mysqli_fetch_assoc($ses_sq4);
$para1 = $row4['farmer_id'];
$para2 = $row4['farmer_name'];
?>
<!DOCTYPE html>
<html>
<?php include ('fheader.php');  ?>
<body class="bg-white" id="top">
<?php include ('fnav.php');  ?>
<section class="section section-shaped section-lg">
<div class="shape shape-style-1 shape-primary">
<span></span>
<span></span>
```

```
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
</div>
<!—
<div class="container ">
<div class="row">
<div class="col-md-8 mx-auto text-center">
<span class="badge badge-danger badge-pill mb-3">Prediction</span>
</div>
</div>
<div class="row row-content">
<div class="col-md-12 mb-3">
<div class="card text-white bg-gradient-success mb-3">
<form role="form" action="#" method="post" >
<div class="card-header">
<span class=" text-info display-4" > Yield Prediction  </span>
</div>
<div class="card-body text-dark">
<table class="table table-striped table-hover table-bordered bg-gradient-white text-center
display" id="myTable">
<thead>
<tr class="font-weight-bold text-default">
<th><center> State</center></th>
<th><center>District</center></th>
<th><center>Season</center></th>
<th><center>Crop</center></th>
<th><center>Area</center></th>
<th><center>Prediction</center></th>
```

```html
</tr>
</thead>
<tbody>
<tr class="text-center">

<td>
<div class="form-group">
<select  name="state" class="form-control" required>
<option value="Karnataka">Karnataka</option>
</select>
</div>
</td>
<td>
<div class="form-group ">
<select id="district" name="district" class="form-control" required>
<option value="">Select a district</option>
<option value="BAGALKOT">Bagalkot</option>
<option value="BANGALORE_RURAL">Bangalore Rural</option>
<option value="BELGAUM">Belgaum</option>
<option value="BELLARY">Bellary</option>
<option value="BENGALURU_URBAN">Bengaluru Urban</option>
<option value="BIDAR">Bidar</option>
<option value="BIJAPUR">Bijapur</option>
<option value="CHAMARAJANAGAR">Chamarajanagar</option>
<option value="CHIKBALLAPUR">Chikballapur</option>
<option value="CHIKMAGALUR">Chikmagalur</option>
<option value="CHITRADURGA">Chitradurga</option>
<option value="DAKSHIN_KANNAD">Dakshin Kannad</option>
<option value="DAVANGERE">Davangere</option>
<option value="DHARWAD">Dharwad</option>
<option value="GADAG">Gadag</option>
<option value="GULBARGA">Gulbarga</option>
<option value="HAVERI">Haveri</option>
<option value="HASSAN">Hassan</option>
```

```
<option value="KODAGU">Kodagu</option>
<option value="KOLAR">Kolar</option>
<option value="KOPPAL">Koppal</option>
<option value="MANDYA">Mandya</option>
<option value="MYSORE">Mysore</option>
<option value="RAMANAGARA">Ramanagara</option>
<option value="RAICHUR">Raichur</option>
<option value="SHIMOGA">Shimoga</option>
<option value="TUMKUR">Tumkur</option>
<option value="UDUPI">Udupi</option>
<option value="UTTAR_KANNAD">Uttar Kannad</option>
<option value="YADGIR">Yadgir</option>
</select>
</div>
</td>
<td>
<div class="form-group ">
<select name="Season" class="form-control" id="season" required>
<option value="">Select Season ...</option>
<option name="Kharif" value="Kharif">Kharif</option>
<option name="Rabi" value="Rabi">Rabi</option>
<option name="Summer" value="Summer">Summer</option>
<option name="WholeYear" value="WholeYear">Whole Year</option>
</select>
</div>
</td>
<td>
<div class="form-group" >
<select id="crop" class="form-control" name="crops" required>
<option value="">Select crop</option>
</select>
</div>
</td>
<script>
```

```
document.getElementById("season").addEventListener("change", function() {
const districtDropdown = document.getElementById('district');
const seasonDropdown = document.getElementById('season');
const cropDropdown = document.getElementById('crop');
console.log(districtDropdown);
console.log(seasonDropdown);
console.log(cropDropdown);
const selectedDistrict = districtDropdown.value;
const selectedSeason = seasonDropdown.value;
// Clear the current crop options
cropDropdown.innerHTML = '<option value="">Select crop</option>';
// If both district and season are selected, add the corresponding crop options to the
dropdown
if (selectedDistrict && selectedSeason) {
const options = cropOptions[selectedDistrict][selectedSeason];
for (const option of options) {
const optionElement = document.createElement('option');
optionElement.value = option; // Set the value to the option text
optionElement.text = option;
cropDropdown.appendChild(optionElement);
}
}
});
</script>
<td>
<div class="form-group">
<input type = "number" step=0.01 name="area" placeholder="Area in Hectares" required
class="form-control">
</div>
</td>
<td>
<center>
<div class="form-group ">
```

```
<button type="submit" value="Yield" name="Yield_Predict" class="btn btn-success btn-
submit">Predict</button>
</div>
</center>
</td>
</tr>
</tbody>
</table>
</div>
</form>
</div>
<div class="card text-white bg-gradient-success mb-3">
<div class="card-header">
<span class=" text-success display-4" > Result  </span>
</div>
<h4>
<?php
if(isset($_POST['Yield_Predict'])){
$state=trim($_POST['state']);
$district=trim($_POST['district']);
$season=trim($_POST['Season']);
$crops=trim($_POST['crops']);
$area=trim($_POST['area']);
echo "Predicted crop yield (in Quintal) is: ";
$Jstate=json_encode($state);
$Jdistrict=json_encode($district);
$Jseason=json_encode($season);
$Jcrops=json_encode($crops);
$Jarea=json_encode($area);
$command = escapeshellcmd("python ML/yield_prediction/yield_prediction.py $Jstate
$Jdistrict $Jseason $Jcrops $Jarea ");
$output = passthru($command);
echo $output;
}
```

```
?>
</h4>
</div>
</div>
</div>
</div>
</section>
<?php require("footer.php");?>
</body>
</html>
```

CROP RECOMMENDATION

```
<?php
include ('fsession.php');
ini_set('memory_limit', '-1');
if(!isset($_SESSION['farmer_login_user'])){
header("location: ../index.php");} // Redirecting To Home Page
$query4 = "SELECT * from farmerlogin where email='$user_check'";
$ses_sq4 = mysqli_query($conn, $query4);
$row4 = mysqli_fetch_assoc($ses_sq4);
$para1 = $row4['farmer_id'];
$para2 = $row4['farmer_name'];
?>
<!DOCTYPE html>
<html>
<?php include ('fheader.php');  ?>
<body class="bg-white" id="top">
<?php include ('fnav.php');  ?>
<section class="section section-shaped section-lg">
<div class="shape shape-style-1 shape-primary">
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
```

```
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
</div>
<!--
<div class="container-fluid ">
<div class="row">
<div class="col-md-8 mx-auto text-center">
<span class="badge badge-danger badge-pill mb-3">Recommendation</span>
</div>
</div>
<div class="row row-content">
<div class="col-md-12 mb-3">
<div class="card text-white bg-gradient-success mb-3">
<form role="form" action="#" method="post" >
<div class="card-header">
<span class=" text-info display-4" > Crop Recommendation  </span>
<span class="pull-right">
<button type="submit" value="Recommend" name="Crop_Recommend" class="btn btn-
warning btn-submit">SUBMIT</button>
</span>
</div>
<div class="card-body text-dark">
<form role="form" action="#" method="post" >
<table class="table table-striped table-hover table-bordered bg-gradient-white text-center
display" id="myTable">

<thead>
<tr class="font-weight-bold text-default">
<th><center> Nitrogen</center></th>
<th><center>Phosporous</center></th>
<th><center>Potasioum</center></th>
```

```
<th><center>Temparature</center></th>
<th><center>Humidity</center></th>
<th><center>PH</center></th>
<th><center>Rainfall</center></th>
</tr>
</thead>
<tbody>
<tr class="text-center">
<td>
<div class="form-group">
<input type = 'number' name = 'n' placeholder="Nitrogen Eg:90" required class="form-
control">
</div>
</td>
<td>
<div class="form-group">
<input type = 'number' name = 'p' placeholder="Phosphorus Eg:42" required class="form-
control">
</div>
</td>
<td>
<div class="form-group">
<input type = 'number' name = 'k' placeholder="Pottasium Eg:43" required class="form-
control">
</div>
</td>
<td>
<div class="form-group">
<input type = 'number' name = 't' step =0.01 placeholder="Temperature Eg:21" required
class="form-control">
</div>
</td>
<td>
<div class="form-group">
```

```
<input type = 'number' name = 'h' step =0.01 placeholder="Humidity Eg:82" required
class="form-control">
</div>
</td>

<td>
<div class="form-group">
<input type = 'number' name = 'ph' step =0.01 placeholder="PH Eg:6.5" required
class="form-control">
</div>
</td>
<td>
<div class="form-group">
<input type = 'number' name = 'r' step =0.01 placeholder="Rainfall Eg:203" required
class="form-control">
</div>
</td>
</tr>
</tbody>
</table>
</form>
</div>
</div>
<div class="card text-white bg-gradient-success mb-3">
<div class="card-header">
<span class=" text-success display-4" > Result  </span>
</div>
<h4>
<?php
if(isset($_POST['Crop_Recommend'])){
$n=trim($_POST['n']);
$p=trim($_POST['p']);
$k=trim($_POST['k']);
$t=trim($_POST['t']);
```

```
$h=trim($_POST['h']);
$ph=trim($_POST['ph']);
$r=trim($_POST['r']);
echo "Recommended Crop is : ";
$Jsonn=json_encode($n);
$Jsonp=json_encode($p);
$Jsonk=json_encode($k);
$Jsont=json_encode($t);
$Jsonh=json_encode($h);
$Jsonph=json_encode($ph);
$Jsonr=json_encode($r);
$command = escapeshellcmd("python ML/crop_recommendation/recommend.py $Jsonn
$Jsonp $Jsonk $Jsont $Jsonh $Jsonph $Jsonr ");
$output = passthru($command);
echo $output;
}
?>
</h4>
</div>
</div>
</div>
</div>
</section>
<?php require("footer.php");?>
</body>
</html>
```

FERTILIZER RECOMMENDATION

```
?php
include ('fsession.php');
ini_set('memory_limit', '-1');
if(!isset($_SESSION['farmer_login_user'])){
header("location: ../index.php");} // Redirecting To Home Page
$query4 = "SELECT * from farmer login where email='$user_check'";
$ses_sq4 = mysqli_query($conn, $query4);
```

```
$row4 = mysqli_fetch_assoc($ses_sq4);
$para1 = $row4['farmer_id'];
$para2 = $row4['farmer_name'];
?>
<!DOCTYPE html>
<html>
<?php include ('fheader.php');  ?>
<body class="bg-white" id="top">
<?php include ('fnav.php');  ?>
<section class="section section-shaped section-lg">
<div class="shape shape-style-1 shape-primary">
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
<span></span>
</div>
<!--
<div class="container-fluid "
<div class="row">
<div class="col-md-8 mx-auto text-center">
<span class="badge badge-danger badge-pill mb-3">Recommendation</span>
</div>
</div>
<div class="row row-content">
<div class="col-md-12 mb-3">
<div class="card text-white bg-gradient-success mb-3">
<form role="form" action="#" method="post" >
<div class="card-header">
```

```
<span class=" text-info display-4" > Fertilizer Recommendation  </span>
<span class="pull-right">
<button type="submit" value="Recommend" name="Fert_Recommend" class="btn btn-
warning btn-submit">SUBMIT</button>
</span>
</div>
<div class="card-body text-dark">
<table class="table table-striped table-hover table-bordered bg-gradient-white text-center
display" id="myTable">
<thead>
<tr class="font-weight-bold text-default">
<th><center> Nitrogen</center></th>
<th><center>Phosporous</center></th>
<th><center>Potasioum</center></th>
<th><center>Temparature</center></th>
<th><center>Humidity</center></th>
<th><center>Soil Moisture</center></th>
<th><center>Soil Type</center></th>
<th><center>Crop</center></th>
</tr>
</thead>
<tbody>
<tr class="text-center">
<td>
<div class="form-group">
<input type = 'number' name = 'n' placeholder="Nitrogen Eg:37" required class="form-
control">
</div>
</td>
<td>
<div class="form-group">
<input type = 'number' name = 'p' placeholder="Phosphorus Eg:0" required class="form-
control">
</div>
```

```
</td>
<td>
<div class="form-group">
<input type = 'number' name = 'k' placeholder="Pottasium Eg:0" required class="form-
control">
</div>
</td>
<td>
<div class="form-group">
<input type = 'number' name = 't'  placeholder="Temperature Eg:26" required class="form-
control">
</div>
</td>
<td>
<div class="form-group">
<input type = 'number' name = 'h'  placeholder="Humidity Eg:52" required class="form-
control">
</div>
</td>
<td>
<div class="form-group">
<input type = 'number' name='soilMoisture' placeholder="Soil Moisture Eg:38" required
class="form-control">
</div>
</td>
<td>
<div class="form-group ">
<select name="soil" class="form-control">
<option  value="">Select Soil Type</option>
<option  value="Sandy">Sandy</option>
<option  value="Loamy">Loamy</option>
<option  value="Black">Black</option>
<option  value="Red">Red</option>
<option  value="Clayey">Clayey</option>
```

```
</select>
</div>
</td>
<td>
<div class="form-group ">
<select name="crop" class="form-control">
<option  value="">Select Crop</option>
<option  value="Maize">Maize</option>
<option  value="Sugarcane">Sugarcane</option>
<option  value="Cotton">Cotton</option>
<option  value="Tobacco">Tobacco</option>
<option  value="Paddy">Paddy</option>
<option  value="Barley">Barley</option>
<option  value="Wheat">Wheat</option>
<option  value="Millets">Millets</option>
<option  value="Oil seeds">Oil seeds</option>
<option  value="Pulses">Pulses</option>
<option  value="Ground Nuts">Ground Nuts</option>
</select>
</div>
</td>
</tr>
</tbody>
</table>
</div>
</form>
</div>
<div class="card text-white bg-gradient-success mb-3">
<div class="card-header">
<span class=" text-success display-4" > Result  </span>
</div>
<h4>
<?php
if(isset($_POST['Fert_Recommend'])){
```

```
$n=trim($_POST['n']);
$p=trim($_POST['p']);
$k=trim($_POST['k']);
$t=trim($_POST['t']);
$h=trim($_POST['h']);
$sm=trim($_POST['soilMoisture']);
$soil=trim($_POST['soil']);
$crop=trim($_POST['crop']);
echo "Recommended Fertilizer is : ";
$Jsonn=json_encode($n);
$Jsonp=json_encode($p);
$Jsonk=json_encode($k);
$Jsont=json_encode($t);
$Jsonh=json_encode($h);
$Jsonsm=json_encode($sm);
$Jsonsoil=json_encode($soil);
$Jsoncrop=json_encode($crop);
$command = escapeshellcmd("python
ML/fertilizer_recommendation/fertilizer_recommendation.py $Jsonn $Jsonp $Jsonk $Jsont
$Jsonh $Jsonsm $Jsonsoil $Jsoncrop ");
$output = passthru($command);
echo $output;
}
?>
</h4>
</div>
</div>
</div>
</div>
</section>
<?php require("footer.php");?></body></html>
```

6.2 OUTPUT SCREENS

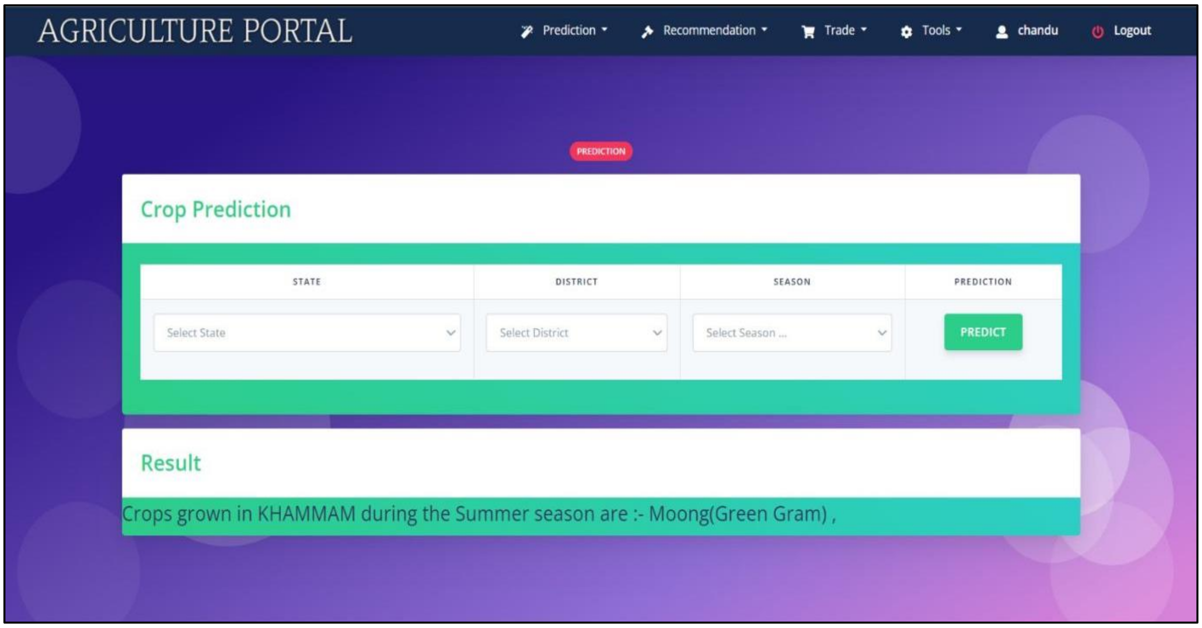

Fig6.2.1: Crop Prediction

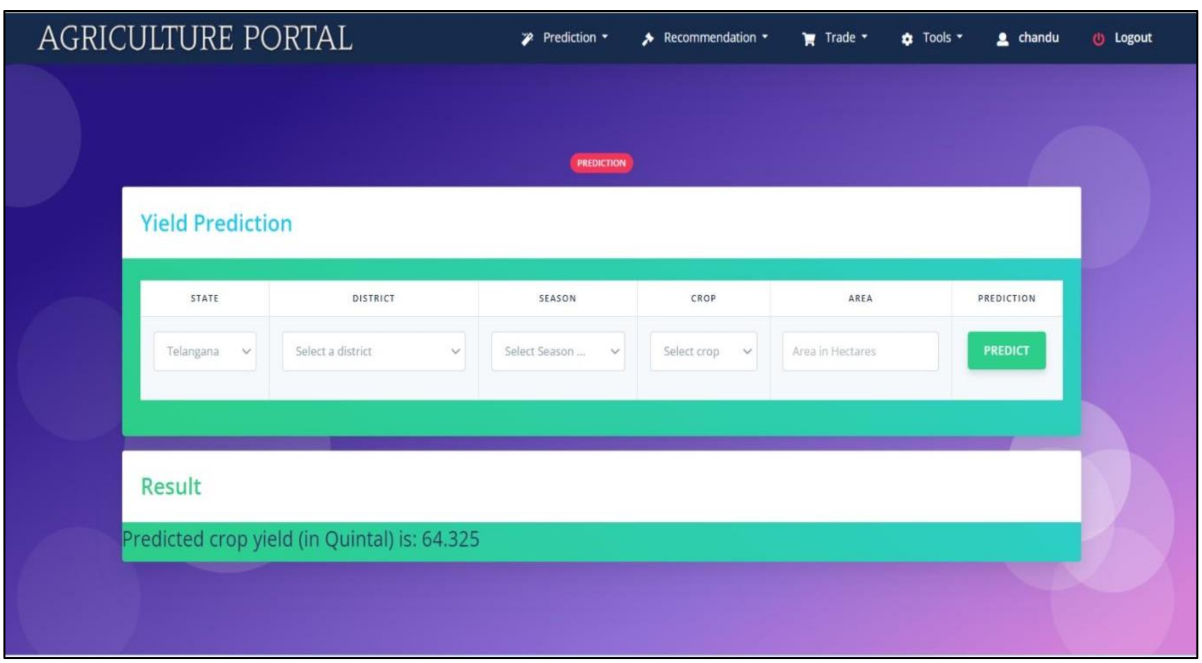

Fig6.2.2: Yield Prediction

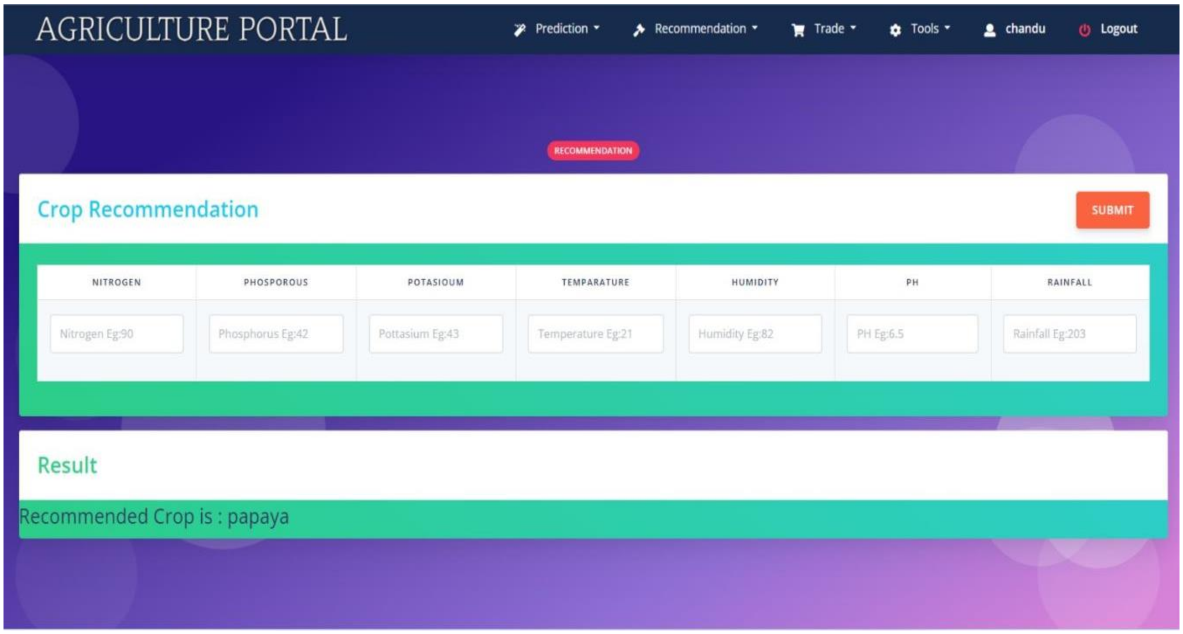

Fig6.2.3: Crop Recommendation

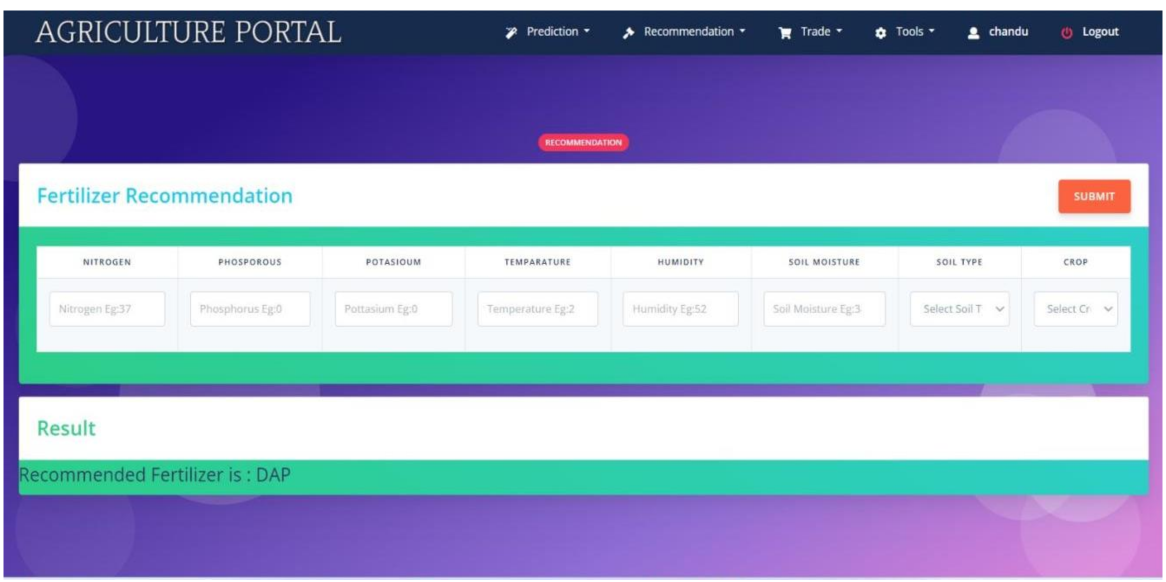

Fig6.2.4: Fertilizer Recommendation

CHAPTER 7

CONCLUSION

With this model, we have created a rudimentary model that is able to forecast to a certain extent. Even though the model is not perfect, we have one that can approximate to the past data pretty well. But for new data, we would require more parameter tuning. There exists great potential to improve sales forecasting accuracy in the Ecommerce domain. One good opportunity is to utilize the correlated and similar sales patterns available in a product portfolio.

In this paper, we have introduced a novel demand forecasting framework based on LSTMs that exploits non-linear relationships that exist in the E-commerce business. We have used the proposed approach to forecast the sales demand by training a global model across the items available in a product assortment hierarchy. Our developments also present several systematic grouping strategies to our base model, which are in particular useful in situations where product sales are sparse. Our methodology has been evaluated on a real-world E-commerce database from Walmart.com. To demonstrate the robustness of our framework, we have evaluated our methods on both category level and super-department level datasets.

The results have shown that our methods have outperformed the state-of-the art uni variate forecasting techniques. Furthermore, the results indicate that E-commerce product hierarchies contain various cross-product demand patterns and correlations are available, and approaches to exploit this information are necessary to improve the sales forecasting accuracy in this domain. So, we have discussed four different steps in data cleaning to make the data more reliable and to produce good results. After properly completing the Data Cleaning steps, we'll have a robust dataset that avoids many of the most common pitfalls. This step should not be rushed as it proves very beneficial in the further process.

Feature Scaling is a technique to standardize the independent features present in the data in a fixed range. It is performed during the data preprocessing.

BIBLIOGRAPHY

[1] Manpreet Kaur, Heena Gulati, Harish Kundra,"Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications",International Journal of Computer Applications, Volume 99– No.12, August 2014.

[2] J. Meng, "Research on the cost of agricultural products circulation and its control under the new normal economic development," Commercial Times, no. 23, pp. 145147, 2016.

[3] A. Kaloxylos etal., "Farm management systems and the future Internet era," Comput. Electron. Agricult., vol. 89, pp. 130–144, Nov.2012.

[4] N. N. Li, T. S. Li, Z. S. Yu, Y. Rui, Y. Y. Miao, and Y. S. Li, "Factors influencing farmers' adoption of new technology based on Logistic-ISM model-a case study of potato planting technology in Dingxi City, Gansu Province," Progress in Geography, vol. 33, no. 4, pp. 542-551, 2014.

[5] Y. Wang, "A neural network adaptive control based on rapid learning method and its application,"Advances In Modeling and Analysis, Vol. 46(3),pp. 27-34,199

[6] J Tang,Xia Hu,Huiji Gao,Huan Liu,"Unsupervised Feature Selection for View Data in Social Media",Proceedings of the ACM international conference on Information and Knowledge management (2015),pp.1673-1678.

[7] Haytham Elzhazel and Alex Auseem,"Unsupervised Feature Selection with ensemble learning ", Machine Learning ,Springer (2013).

[8] Mohak Shah, Mario Marchand, and Jacques Corbeil, " Feature Selection with Conjunctions of Decision Stumps and Learning ", IEEETRANSACTIONS ON PATTERN ANALYSISAND MACHINE INTELLIGENCE, VOL. 34, NO.1,(2012) pp.174-186.

[9] Qingzheng XU, Lei WANG, Baomin HE, NaWANG," Modified Opposition-Based Differential Evolution for Function Optimization", Journal of Computational Information Systems 7:5 (2011)pp.1582-1591.

[10] Hyndman, R. eTal., 2008. Forecasting with Exponential Smoothing: The State Space Approach, Springer Science & Business Media.