A NOVEL DATABASE MANAGEMENT APPROACH FOR SMART FARMING

A PROJECT REPORT

Submitted by

GOPINARESH B **(422619104013)**

HARI PRIYA D **(422619104015)**

PRABAKARAN R **(422619104032)**

In partial fulfillment for the award of the degree

Of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

UNIVERSITY COLLEGE OF ENGINEERING, PANRUTI

ANNA UNIVERSITY: CHENNAI 600-025

MAY 2023

ANNA UNIVERSITY: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this project report "**A NOVEL DATABASE MANAGEMENT APPROACH FOR SMART FARMING"** is the bonafide work of **"GOPINARESH B (422619104013), HARI PRIYA D (422619104015), PRABAKARAN R (422619104032)"** who carried out the project work under my supervision.

SIGNATURE

Dr.D. MURUGANANDAM,

M.Tech Ph.D.,

ASSISTANT PROFESSOR

HEAD OF THE DEPARTMENT

Computer Science & Engineering

University College of Engineering

Panruti – 607106.

SIGNATURE

Mr.S.SUNDARARAJAN, M.E.,

TEACHING FELLOW

SUPERVISOR

Computer Science & Engineering

University College of Engineering

Panruti – 607106.

EXAMINATION HELD ON _____

INTERNAL EXAMINER

EXTERNAL EXAMINER

DECLARATION

We hereby declare that the work entitled "**A NOVEL DATABASE MANAGEMENT APPROACH FOR SMART FARMING"** is submitted in partial fulfillment for the award of the degree in **Bachelor of Engineering in Computer Science & Engineering. University College of Engineering, Panruti** is a record of our own work carried out by us during the academic year 2022-2023. Under the supervision and guidance of **Mr.S.Sundararajan, M.E.**, Department of Computer Science and Engineering, UNIVERSITY COLLEGE OF ENGINEERING PANRUTI. The extent and source of information are derived from the existing literature and have been indicated through dissertation at the appropriate places. The matter embodied in this work is original and has not been submitted for the award of any other degree or diploma, either in this or any other university.

REGISTER NUMBER	NAME	SIGNATURE
422619104013	GOPINARESH B	
422619104015	HARI PRIYA D	
422619104032	PRABAKARAN R	

I certify that the declaration made above by the candidate is true.

SIGNATURE

Mr. S. SUNDARARAJAN, M.E.,

TEACHING FELLOW

SUPERVISOR

Computer Science & Engineering

University College of Engineering

Panruti – 607106.

ABSTRACT

Tamil Nadu is an agricultural state in southern India with a diverse range of crops grown throughout the year due to its favorable climate. The state is known for its high production of rice, sugarcane, cotton, tea, coffee, maize, black gram, coconut, groundnut and ragi. However, like other agricultural regions, Tamil Nadu faces several challenges such as water scarcity, land degradation, and low productivity. Smart farming is a rapidly growing field that utilizes technology to improve agricultural practices. To overcome the challenges, crop prediction is helping for crop selection as per the resource of land and season. A crop recommendation system is a software application that utilizes machine learning algorithms to assist farmers in selecting the most suitable crop for their agricultural field based on factors such as NPK levels, temperature, humidity, soil pH and rainfall. The system generates recommendations that help farmers maximize yield and minimize resource waste. Ultimately, the goal of the project is to help the young farmers in a better way by suggesting suitable crops for productivity. The benefits of such a system include increased efficiency, improved decision-making, and potentially higher profits for farmers. The crop will be predicted with the best accuracy using Random Forest Classifier and Decision Tree Classifier algorithms. Additionally, the system also suggests the duration of crop.

TABLE OF CONTENTS

CHAPTER 1
1. INTRODUCTION

1.1 DATA COLLECTION

Data collection is the most common approach for gathering and analyzing information from various sources. The dataset must have the following qualities to provide an approximate data set for the system. These criteria will be considered for crop recommendation: i) NPK levels ii) Temperature iii) Humidity iv) Soil pH v) Rainfall. Collecting data allows you to capture a record of past events to find recurring patterns. From those patterns, build predictive models using machine learning algorithms that looks for trends and predicts future changes.

Predictive models are only as good as the data from which they are built, so good data collection practices are crucial to developing high-performing models. The data need to be error-free (garbage in, garbage out) and contain relevant information for the task at hand.

1.2 DATA PREPROCESSING

Pre-processing involves the elimination of disturbances and outliers present in the CSV dataset. A lot of times data can go missing that directly affects the efficiency and accuracy of the final machine learning model. This needs to be addressed through various techniques such as median and mean of the whole column.

Using Sklearn, we can conveniently clean the dataset. It provides an imputer class that addresses and fixes missing values. The imputer class accepts parameters like missing values and strategies that the imputer internally uses. This strategy involves replacing missing values using the mean on-axis.

1.3 FEATURE SELECTION

Not all columns contribute majorly towards favorable prediction. Feature selection is a specific process in the pipeline where such features are selected

which contribute most towards the prediction variable. Feature selection reduces the overfitting of data. An optimized dataset means a machine learning model makes decisions against a low amount of noise. It directly increases the accuracy of the model.

1.4 TRAINING THE MODEL

Training the model is a critical step in the development of the crop recommendation system.

Random Forest Algorithm

Random Forest is one of the most popular and widely accepted supervised learning techniques. It consists of a number of decision trees for different subsets of data, rather than using the whole data as a single unit. This helps maximize the accuracy of the prediction of each variable. It makes sense to use Random Forest on a large dataset as it can provide results with maximum accuracy in a short span of time.

Decision Tree Algorithm

In a Decision tree, there are two nodes, which are the Decision Node and Leaf Node. Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches. The decisions or the test are performed on the basis of features of the given dataset.

1.5 TESTING THE MODEL

The goal is to evaluate the model's performance and ensure that it provides accurate and reliable recommendations. The accuracy of Random Forest Classifier and Decision Tree Classifier algorithm is 1.0 for both.

1.6 DEPLOYMENT

Once the model has been tested and refined, it can be deployed for use in the crop prediction system. The model should be integrated with the user interface and tested thoroughly to ensure that it provides accurate and reliable recommendations. Python Flask API is used to deploy the model into web application.

CHAPTER 2

2. LITERATURE SURVEY

2.1 TITLE: Design and development of crop database website to help farmers

AUTHOR: Ramya, Rohit, Prashanth, Pratibha, 2022

The United Nations report suggests that, by 2050, the World's population will reach around 9.7 billion people. To meet the demands of this growing population, India must step up its food production by 70%. An efficient way to achieve this is to help the farmers of our nation with proper knowledge on crops and how to grow them by also making certain smart agricultural devices available to them through a website. A database with data of 30 crops is created as an example for analyzing crop data. This aims to give information to the farmer as to which crop is suitable for the soil conditions and climatic conditions of his land along with the market value of the crops. Thus, by assisting our farmers with all the information, we can help them get a better yield and to grow the economy of the nation, in turn.

2.2 TITLE: Agricultural Crop Recommendations based on Productivity and Season

AUTHOR: S. Vaishnavi, M. Shobana, R. Sabitha and S. Karthik, 2021

Tamil Nadu being 7th largest area in India has 6th largest population. It is the leading producer of agriculture products. Agriculture is the main occupation of Tamil Nadu people. Agriculture has a sound tone in this competitive world. Cauvery is the main source of water. Cauvery delta regions are called rice bowl, Tamil Nadu. Rice is the major crop grown in Tamil Nadu. Other crops like Paddy, Sugarcane, Cotton, Coconut and groundnut are grown. Agriculture directly depends on the environmental factors such as sunlight, humidity, soil type, rainfall, Maximum and Minimum Temperature, climate, fertilizers, pesticides etc. Knowledge of proper harvesting of crops is in need to bloom in Agriculture. Due to the diversity of season and rainfall, assessment of suitable crops to cultivate is necessary. Farmers face major problems such as crop management, expected crop yield and productive yield from the crops. Farmers or cultivators need proper assistant regarding crop cultivation as now-a-days many fresh youngsters are

interested in agriculture. Impact of IT sector in assessing real world problem is moving at a faster rate. Data is increasing day by day in the field of agriculture. With the advancement in Internet of Things, there are ways to grasp huge data in field of Agriculture. There is a need for a system to have obvious analyzes of data of agriculture and extract or use useful information from the spreading data. To get insights from data, it has to be learnt.

2.3 TITLE: Machine Learning Approaches for Crop Recommendation

AUTHOR: P. Parameswari, N. Rajathi, K.J. Harshanaa, 2021

Countries' competitiveness and economic growth are fueled by innovation. Software is used in sustainable agriculture to provide farmers with data and assistance on crop rotation, harvesting details, and soil management. Sensors are used to measure the soil moisture and temperature. This research contributes to the development of a model that assists farmers by providing crop-related information or crop recommendations based on various attributes such as crop details, soil composition, weather conditions that crop can grow in, temperature, soil PH, and rainfall. This research employs machine learning algorithms such as PART, Decision table, and JRip. Experiments are run on data gathered by the Indian Chamber of Food and Agriculture from the Kaggle repository. The effects of all three algorithms are measured on a variety of scales, including Accuracy, Precision, Recall, and F-Measure. Accuracy is calculated by comparing instances that were correctly and incorrectly predicted. The results show that the PART algorithm performs well, with the highest precision of 98.33% when compared to other methods, and it consumes less time to build the model.

2.4 TITLE: Recommendation of Agricultural Crop Based on Productivity and Season Using Machine Learning

AUTHOR: Sowmya, Aliya Sultana Imam Sabh Athar, Mubeena Banu, 2022

Agriculture gave birth to civilization. India is an agrarian country and its economy largely based upon crop productivity. Thus, agriculture is that the backbone of all business in India. Now India stands in second rank in worldwide in farm production. India is an agricultural country but remains using traditional ways of recommendations for agricultural purpose. Presently, recommendations for

farmers are supported one to at least one interaction between farmers and therefore the experts and different experts have different recommendations. Agriculture directly depends on the environmental factors such as sunlight, humidity, soil type, rainfall, Maximum and Minimum Temperature, climate, fertilizers, pesticides etc. Knowledge of proper harvesting of crops is in need to bloom in Agriculture. India has seasons of 1) Winter which occurs from December to March 2) Summer season from April to June 3) Monsoon or rainy season lasting from July to September and 4) Post-monsoon or autumn season occurring from October to November. Due to the diversity of season and rainfall, assessment of suitable crops to cultivate is necessary. Farmers or cultivators need proper assistance regarding crop cultivation as now-a-days many fresh youngsters are interested in agriculture.

2.5 TITLE: A proposal for future data organization in enterprise systems—an analysis of established database approaches

AUTHOR: Benedict Bender, Clementine Bertheau, Tim Korppen, Hannah Lauppe, Norbert Gronau, 2022

The digital transformation sets new requirements to all classes of enterprise systems in companies. ERP systems in particular, which represent the dominant class of enterprise systems, are struggling to meet the new requirements at all levels of the architecture. Therefore, there is an urgent need to reconsider the overall architecture of the systems and address the root of the related issues. Given that many restrictions ERP pose on their adaptability are related to the standardization of data, the database layer of ERP systems is addressed. Since database serve as the foundation for data storage and retrieval, they limit the flexibility of enterprise systems and the chance to adapt to new requirements accordingly. So far, relational databases are widely used. Using a systematic literature approach, recent requirements for ERP systems were identified. Prominent database approaches were assessed against the 23 requirements identified. The results reveal the strengths and weaknesses of recent database approaches. To this end, the results highlight the demand to combine multiple database approaches to fulfill recent business requirements. From a conceptual point of view, this paper supports the idea of federated databases which are interoperable to fulfill future requirements and support business operation. This research forms the basis for renewal of the current generation of ERP systems and proposes to ERP vendors to use different database concepts in the future.

OVERVIEW OF LITERATURE SURVEY

S.NO	TITLE	AUTHOR	YEAR	MERITS	DEMERITS
1.	Design and Development of Crop Database Website to help Farmers	Ramya, Rohit, Prashanth, Pratibha	2020	Helps farmers to make decisions on which crop to grow in what type of soil.	The climatic changes and extreme weather conditions are posing more problems in farming.
2.	Agricultural Crop Recommendations based on Productivity and Season	S. Vaishnavi, M. Shobana, R. Sabitha and S. Karthik	2021	More personalized and relevant recommendations can be given to farmers which makes them yield a good volume of production.	Problems in crop management and crop yields.
3.	Machine Learning Approaches for Crop Recommendation	P.Parameswari, N.Rajathi, K.J.Harshanaa	2021	Considering the weather conditions, soil, and water requirements.	It is critical that the recommendation offered is exact and precise, as errors can result In capital loss.
4.	Recommendation of Agricultural Crop Based on Productivity and Season Using	Sowmya, Aliya Sultana Imam Sabh Athar, Mubeena Banu	2022	Practiced for good volume of production	Chosen of unsuitable crops based on their soil requirements.

	Machine Learning				
5.	A proposal for future data organization in enterprise systems—an analysis of established database approaches	Benedict Bender, Clementine Bertheau, Tim Korppen, Hannah Lauppe, Norbert Gronau	2022	Arrange the data in order that it can easily read and work.	Problems in database level due to new requirements of digital transformation give reason to assume that the relational concept may be outdated.

CHAPTER 3

3. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

Farmers have word-of-mouth in past decades but now it cannot be used due to climatic factors. Due to diversity of season and rainfall, assessment of suitable crops to cultivate is necessary. Farmers face major problems such as crop management, expected crop yield and productive yield from the crops. Farmers or cultivators need proper assistance regarding crop cultivation as nowadays many fresh youngsters are interested in agriculture. For the crop recommendation system, the machine learning algorithm has tremendous potential as it allows using crop details for research and improving the agricultural productivity with minimum cost. It is essential to have a certain temperature, humidity, soil pH, sunlight, and soil moisture for a plant to be grown healthy. The section went into greater detail about how the information was used to assist farmers with their farming. Accuracy is calculated by comparing instances that were correctly and incorrectly predicted. The previously collected dataset is trained for prediction.

3.2 DISADVANTAGES

- Previously collected dataset is used.
- Chosen of unsuitable crops based on their soil requirements.
- Less accuracy and low performance.
- Sometimes wrong predictions occur.

3.3 PROPOSED SYSTEM

The proposed crop prediction system can assist farmers in selecting the best crops to be grown based on the soil, climate, and other environmental factors. This can help farmers optimize their crop management practices, minimize risk, and maximize profits. Collecting data on soil properties, weather patterns, crop requirements, and yield information for various crops in a particular region. The collected data needs to be preprocessed to remove any missing or erroneous values and normalized to ensure consistency. Identifying the most important features or

variables that affect crop growth and yield. Selecting the appropriate machine learning algorithm, such as decision trees, random forest that can handle the selected features. Training the selected machine learning model using the preprocessed data and the selected algorithm. Evaluating the performance of the trained model using metrics such as accuracy. Developing a user interface that takes inputs such as soil type, weather patterns and provides recommendations for the most suitable crop to be grown. Deploying the crop recommendation system in a controlled environment and testing it with real-world data. Maintaining the system by regularly updating it with new data and improving the model performance as needed.

3.4 ADVANTAGES

- Improved crop management.
- High accuracy
- Best performance
- Right predictions
- New dataset

COMPARISON SCENARIO

S.NO	EXISTING SYSTEM	PROPOSED SYSTEM
1.	Previously collected dataset	New dataset
2.	Wrong predictions	Accurate predictions
3.	Less accuracy	High accuracy

CHAPTER 4

4. SYSTEM REQUIREMENTS

4.1 HARDWARE REQUIREMENTS

- Processor: AMD Radeon R5 Graphics 2.20 GHz
- RAM: 4 GB
- Hard disk: 160 GB
- Compact Disk: 650 Mb
- Keyboard: Standard keyboard
- Monitor: 15-inch color monitor

4.2 SOFTWARE REQUIREMENTS

- Operating system: Windows OS
- Front end: Visual Studio Code, Jupyter Notebook

CHAPTER 5

5. SOFTWARE DESCRIPTION

5.1 FRONT END: VISUAL STUDIO CODE

- Visual Studio Code, commonly known as VS Code, is a lightweight source-code editor developed by Microsoft that supports multiple programming languages and platforms. It is built on the Electron framework and is based on the open-source project Monaco Editor.
- VS Code is known for its ease of use, customizable user interface, and a wide range of features such as IntelliSense, debugging, syntax highlighting, code completion, and source control integration. It also has a built-in terminal, which allows developers to run commands within the editor.
- One of the main advantages of VS Code is its vast ecosystem of extensions, which provide additional functionality and support for a wide range of programming languages and technologies. These extensions can be easily installed from the VS Code marketplace.
- VS Code is also known for its cross-platform support, as it is available on Windows, Mac, and Linux platforms. It is also highly customizable with a wide range of settings and configurations that can be modified to suit the needs of the developer.
- Overall, VS Code is a popular and powerful code editor that provides developers with a wide range of features and extensions to make development easier and more efficient.

FEATURES OF VISUAL STUDIO CODE

- **IntelliSense:** VS Code provides intelligent code completion and suggestions based on the code being written.
- **Debugging:** The editor has built-in debugging features that allow developers to debug their code right from the editor.
- **Git integration:** VS Code has built-in Git integration, making it easy for developers to manage their source code.
- **Extensions:** There are a vast number of extensions available in the VS Code marketplace that provide additional features and support for additional programming languages.

- **Terminal integration:** VS Code has a built-in terminal, which allows developers to run commands within the editor.
- **Syntax highlighting:** The editor provides syntax highlighting for a wide range of programming languages.
- **Customizable user interface:** The editor has a customizable user interface with a wide range of settings and configurations that can be modified to suit the needs of the developer.
- **Cross-platform support:** VS Code is available on Windows, Mac, and Linux platforms.
- **Multiple cursor support:** VS Code allows developers to work with multiple cursors simultaneously, making it easier to make changes to multiple lines of code at once.
- **Snippet support:** The editor provides support for code snippets, allowing developers to quickly insert frequently used code.

USES OF VISUAL STUDIO CODE

- VS Code is primarily used for code editing, with support for a wide range of programming languages and syntax highlighting.
- The editor has built-in debugging features that allow developers to debug their code right from the editor.
- VS Code has support for real-time collaboration, allowing multiple developers to work on the same codebase simultaneously.
- Task automation: The editor has built-in support for task automation, allowing developers to automate common tasks such as building, testing, and deploying their code.
- VS Code has a wide range of extensions and features that make it a popular choice for web developers, with support for popular frameworks such as React, Angular, and Vue.js.
- Machine learning development: VS Code has support for popular machine learning frameworks such as TensorFlow and PyTorch, making it a popular choice for machine learning development.

VISUAL STUDIO CODE FILES

- **Source code files:** This includes files that contain source code for a wide range of programming languages such as Java, Python, C++, Ruby, and many others.
- **Configuration files:** This includes files that are used to configure applications, servers, and other tools, such as JSON, YAML, INI, and XML files.
- **Markup files:** This includes files that are used to create structured documents such as HTML, XML, and Markdown files.
- **Stylesheet files:** This includes files that are used to define the styles and formatting for web pages such as CSS and SCSS files.
- **Build files:** This includes files that are used to automate the building and packaging of software, such as Makefiles, Gradle files, and Maven files.
- **Git files:** This includes files that are used by Git version control system, such as .gitignore, .gitattributes, and .gitmodules files.
- **Database files:** This includes files that are used to store data in various database systems such as MySQL, PostgreSQL, MongoDB, and others.
- **Other file types:** VS Code can also work with other file types such as images, audio, video, and text files.

5.2 JUPYTER NOTEBOOK

Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. It is commonly used in data science, scientific research, and education to perform interactive data analysis, exploration, and visualization. Jupyter Notebook is a powerful and versatile tool that allows you to perform interactive data analysis and visualization in a wide range of programming languages, making it a popular choice for data scientists, researchers, and educators.

FEATURES OF JUPYTER NOTEBOOK

- **Multi-language support:** Jupyter Notebook supports over 40 programming languages, including Python, R, and Julia, among others.

- **Interactive computing:** You can write and execute code interactively, allowing you to explore data and experiment with different approaches.
- **Data visualization:** Jupyter Notebook supports a wide range of data visualization libraries, such as Matplotlib and Plotly, allowing you to create interactive plots and charts.
- **Collaboration:** Jupyter Notebook allows you to share your notebooks with others, making it easy to collaborate and work on projects together.
- **Reproducibility:** Jupyter Notebook allows you to create reproducible workflows and analyses, making it easier to share and reproduce your work.

USES OF JUPYTER NOTEBOOK

- Jupyter Notebook is widely used in data science for interactive data analysis and visualization. It allows you to write and execute code, and create visualizations and charts, making it a powerful tool for exploring and analyzing data.
- Jupyter Notebook is also commonly used in machine learning for tasks such as model training, data preprocessing, and model evaluation. It supports a wide range of machine learning libraries such as Scikit-Learn and TensorFlow.
- Jupyter Notebook is used by scientists and researchers to analyze and visualize data, and to document and share their research findings. It allows them to create reproducible workflows and analyses, making it easier to share and reproduce their work.
- Jupyter Notebook is a popular tool for teaching and learning programming and data science. It provides an interactive and engaging environment for students to learn and experiment with programming and data analysis.
- Jupyter Notebook is used by journalists and news organizations to analyze and visualize data, and to create interactive stories and reports. It allows them to present data in a compelling and interactive way, making it easier for readers to understand complex information.

JUPYTER NOTEBOOK FILES

- **Notebook files (.ipynb):** These are the main files used in Jupyter Notebook. They contain a combination of code, text, and visualizations, organized into cells. You can create new notebook files, open existing ones, and save them to your computer.
- **Kernel files (.json):** These files contain the state of the kernel associated with a notebook. They are automatically created when you run a notebook and are used to store variables and other data created during the session. You generally don't need to interact with kernel files directly, but they can be useful for troubleshooting issues with a notebook.

CHAPTER 6

6. SYSTEM DESIGN

6.1 SYSTEM ARCHITECTURE

System architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.

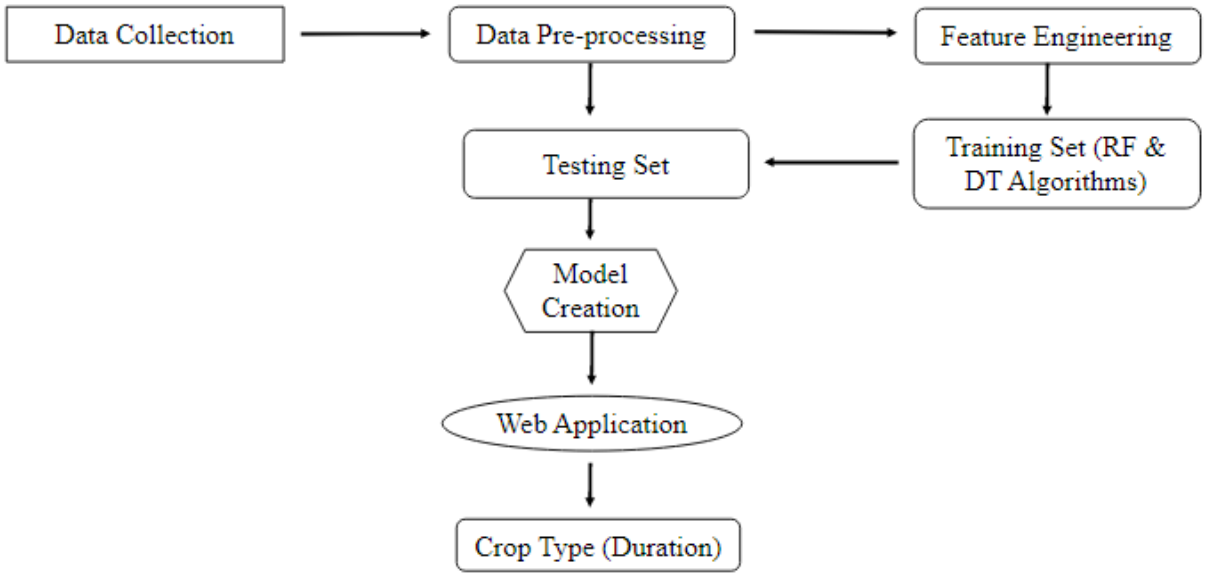


FIG 1. SYSTEM ARCHITECTURE

6.2 DATA FLOW DIAGRAM

A two-dimensional diagram explains how data is processed and transferred in a system. The graphical depiction identifies each source of data and how it interacts with other data sources to reach a common output. Individuals seeking to draft a data flow diagram must identify external inputs and outputs, determine how the inputs and outputs relate to each other, and explain with graphics how these

connections relate and what they result in. This type of diagram helps business development and design teams visualize how data is processed and identify or improve certain aspects.

Data flow Symbols:

SYMBOL	DESCRIPTION
	An entity. A source of data or a destination for data.
	A process or task that is performed by the system.
	A data flow.

FIG 2. DATA FLOW DIAGRAM

CHAPTER 7

7. SYSTEM IMPLEMENTATION

Systems implementation is the process of defining how the information system should be built (i.e., physical system design), ensuring that the information system is operational and used, ensuring that the information system meets quality standard (i.e., quality assurance).

7.1 Data collection:

Collect data on various crops, including their growing conditions, climate requirements and soil requirements. This data can be obtained from various sources, such as agricultural research centers, government reports, and farming communities.

7.2 Data preprocessing:

Clean and preprocess the collected data, such as removing duplicate entries, handling missing values, and normalizing the data.

7.3 Feature selection:

Select the most relevant features from the preprocessed data that are useful for predicting the crop yield.

7.4 Model selection:

Choose an appropriate machine learning algorithm that is suitable for the data and the problem statement. Some common algorithms used in crop recommendation systems include random forest and decision trees.

7.5 Model evaluation:

Evaluate the performance of the trained model using various metrics such as accuracy.

7.6 Deployment:

Deploy the model in a web-based application that farmers can use to get personalized crop recommendations based on their soil requirements and other factors.

CHAPTER 8

8. SYSTEM TESTING

8.1 TESTING

Testing is the process of evaluating a system or its component(s) with the intent to find whether it satisfies the specified requirements or not. In simple words, testing is executing a system in order to identify any gaps, errors, or missing requirements in contrary to the actual requirements.

8.2 TEST CASES

8.2.1 Test-Case table for Web Application Running

ID	Test cases	Input data	Steps of execution	Expected Output	Result
1.	User can access the application	Running code in terminal	Run the html file	Greets user	Pass

8.2.2 Test-Case table for Machine Learning Prediction

ID	Test cases	Input data	Steps of execution	Expected Output	Result
1.	User trains the model	-	Fits the RF algorithm	Shows the accuracy of algorithm	Pass
2.	User trains the model	-	Fits the DT algorithm	Shows the accuracy of algorithm	Pass
3.	User checks for the prediction	Gives required input	Run the Jupyter Notebook	Predicts the right crop with duration	Pass

8.2.3 Test-Case table for Web Application deployed with model

ID	Test cases	Input data	Steps of execution	Expected Output	Result
1.	User tests for the expected results in web application	User gives the required inputs in the field and click predict	Run the flask file	Predicts the right crop with duration	Pass

CHAPTER 9

9. SYSTEM STUDY

SOURCE CODE

sample.html

```
<!DOCTYPE html>

<html lang="en">

<head>

  <meta charset="UTF-8" />

    <meta name="viewport" content="width=device-width, initial-scale=1.0" />

  <title>Smart Farming</title>

  <link rel="stylesheet" href="../static/css/style2.css">

</head>

<body>

 <div class="login">

  <h3>A NOVEL DATABASE MANAGEMENT APPROACH FOR SMART FARMING</h3>

  <form action="{{ url_for('predict')}}"method="post">

                  <input   type="float"   name="Nitrogen"   placeholder="Nitrogen"
required="required" />

                <input   type="float"   name="Phosphorus"   placeholder="Phosphorus"
required="required" />

                 <input   type="float"   name="Potassium"   placeholder="Potassium"
required="required" />

              <input   type="float"   name="Temperature"   placeholder="Temperature"
required="required" />

                  <input   type="float"   name="Humidity"   placeholder="Humidity"
required="required" />

                 <input   type="float"   name="Soil   pH"   placeholder="Soil   pH"
required="required" />
```

```
                <input type="float" name="Rainfall" placeholder="Rainfall"
required="required" />

                 <button type="submit" class="btn btn-primary btn-block
btn-large">Predict</button>

</form>

<br>

<br>

<!-- {{ prediction text }} -->

<h2 class="output">{{prediction_text}}</h2>

 </div>

</body>

</html>
```

style2.css

```
* {

  padding: 0;

  margin: 0;

  box-sizing: border-box;

  font-family: Arial, Helvetica, sans-serif;

  text-transform: capitalize;

  text-decoration: none;

}

body {

  min-height: 100vh;

background:url("https://cdn.lifestyleasia.com/wpcontent/uploads/sites/6/2020/02/0
7160345/changluntwitter-1024x681.jpeg")

    no-repeat;

  background-size: cover;
```

```
background-position: center;
}
input[type="float"] {
  font-size: 16px;
  display: block;
  padding: 10px;
  border-radius: 5px;
  border: none;
}
input[type="float"]:focus {
  outline: 2px dotted #fe6244;
}
.login {
  min-height: 100vh;
  display: flex;
  flex-direction: column;
  justify-content: center;
  align-items: center;
}

form {
  display: grid;
  grid-template-columns: minmax(150px, 200px);
  gap: 15px;
  margin-top: 80px;
  background-color: rgba(255, 255, 255, 0.75);
  padding: 30px;
  border-radius: 10px;
}
```

```
.btn-primary {
  padding: 10px 10px;
  background-color: rgba(2, 2, 2, 0.775);
  border: none;
  color: #fff;
  font-weight: 500;
  font-size: 16px;
  text-transform: uppercase;
  letter-spacing: 1px;
  transition: all 0.3s;
  border-radius: 5px;
  cursor: pointer;
}
.btn-primary:hover {
  background-color: rgba(0, 0, 0, 0.899);
}
.output {
  margin-top: 10px;
  background-color: hsla(108, 95%, 43%, 0.733);
  color: #333;
  color: #f0f0f0;
  padding: 10px;
  border-radius: 5px;
  font-weight: 500;
}
h3{
  position: absolute;
  top: 30px;
  color: #ffffff;
```

```
}
```

model.py

```
import pandas as pd
from sklearn.metrics import classification_report
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn import metrics
import pickle
from sklearn.preprocessing import StandardScaler

df = pd.read_csv("../final project telegram/new/crop data.csv")
df.head()
x = df.iloc[:,df.columns!='Crop'] #data
y = df.iloc[:,df.columns=='Crop'] #outcome or Label
features = df[['Nitrogen', 'Phosphorous','Potassium','Temperature', 'Humidity',
'Soil pH', 'Rainfall']]
target = df['Crop']
labels = df['Crop']
xtrain, xtest, ytrain, ytest = train_test_split(x,y,test_size = 0.2)
acc = []
model = []
# Feature scaling
sc = StandardScaler()
X_train = sc.fit_transform(xtrain)
X_test= sc.transform(xtest)
```

```
#Random Forest Classifier
classifier = RandomForestClassifier(n_estimators=20, random_state=0)
classifier.fit(xtrain,ytrain.values.ravel())
predicted_values = classifier. Predict(xtest)
x = metrics.accuracy_score(ytest, predicted_values)
acc.append(x)
model.append('classifier')
print("RF's Accuracy is: ", x)
print(classification_report(ytest,predicted_values))

DecisionTree=
DecisionTreeClassifier(criterion="entropy",random_state=2,max_depth=5)
DecisionTree.fit(xtrain,ytrain)
predicted_values = DecisionTree.predict(xtest)

pickle.dump(classifier, open("model.pkl", "wb"))
```

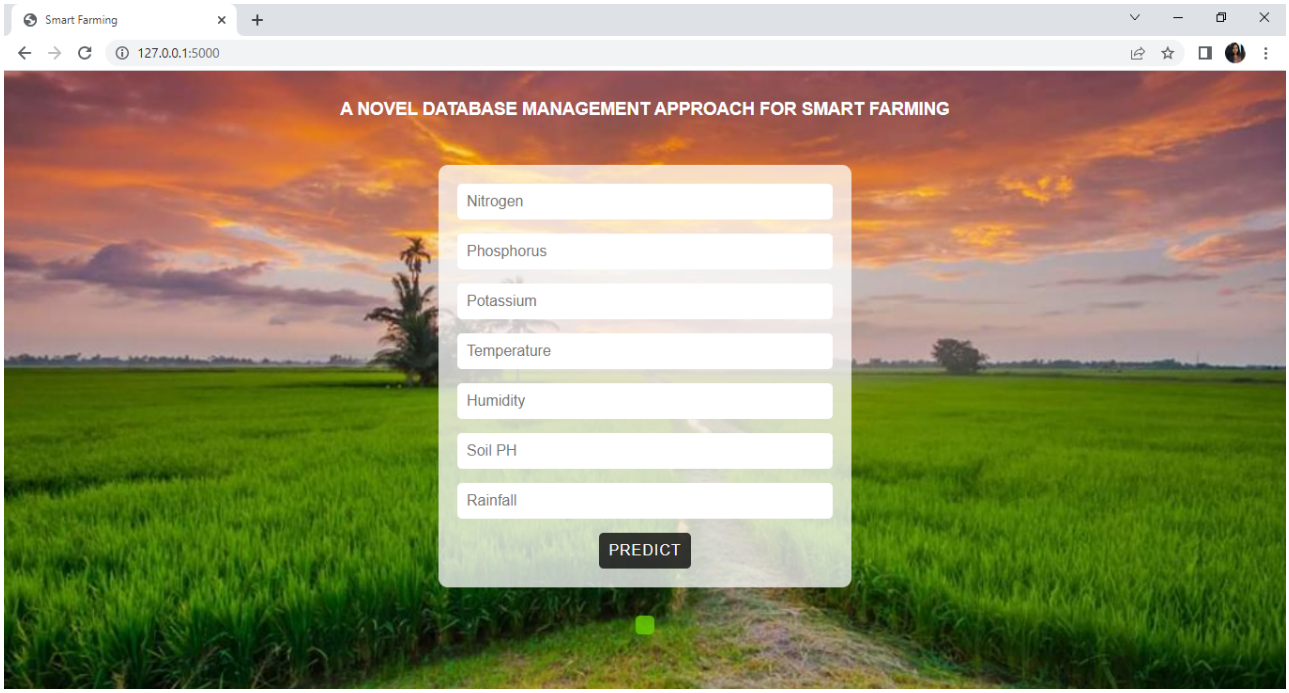
app.py

```
import numpy as np
import pandas as pd
from flask import Flask, request, render_template
import pickle
flask_app = Flask(__name__)
model = pickle.load(open("model.pkl", "rb"))

@flask_app.route('/')
def Home():
    return render_template('sample.html')
```

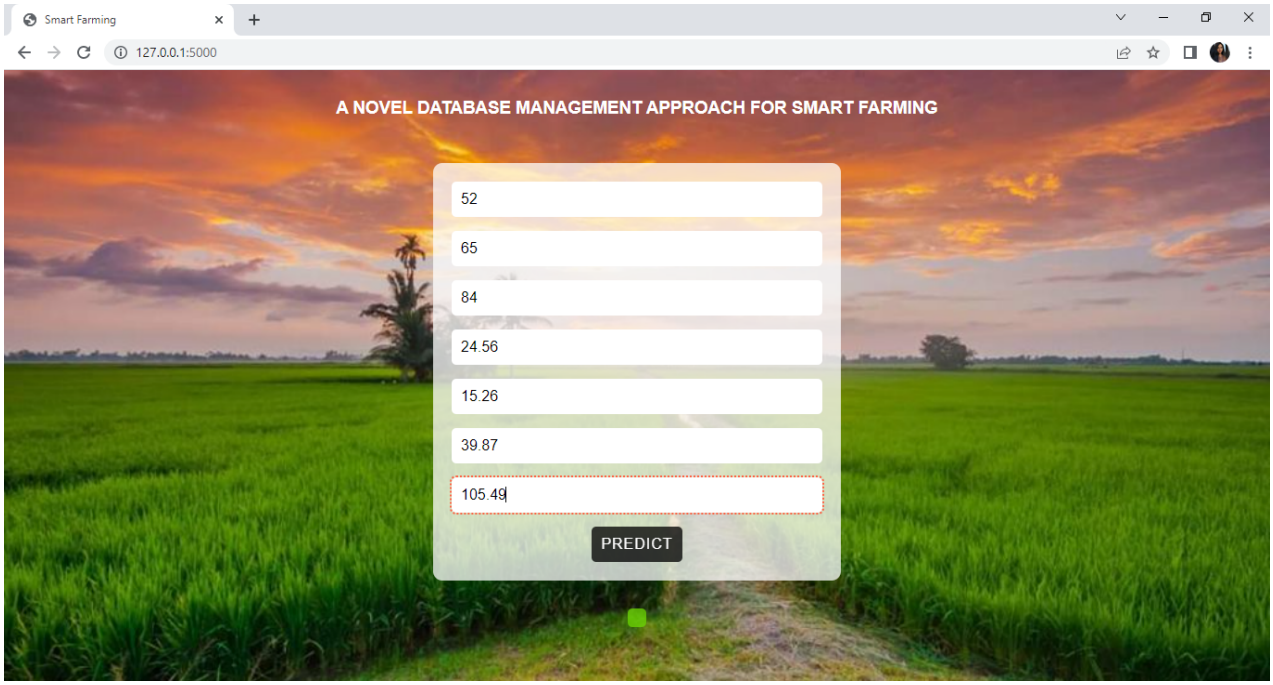
```
#prediction of model using machine learning

@flask_app.route('/predict', methods=['GET','POST'])
def predict():
        float_features = [float(x) for x in request.form.values()]
        features = [np.array(float_features)]
        prediction = model.predict(features)


      return  render_template("sample.html", prediction_text = "The recommended
      crop is{}".format(prediction))

if __name__ == "__main__":
    flask_app.run(debug=True)
```

CHAPTER 10

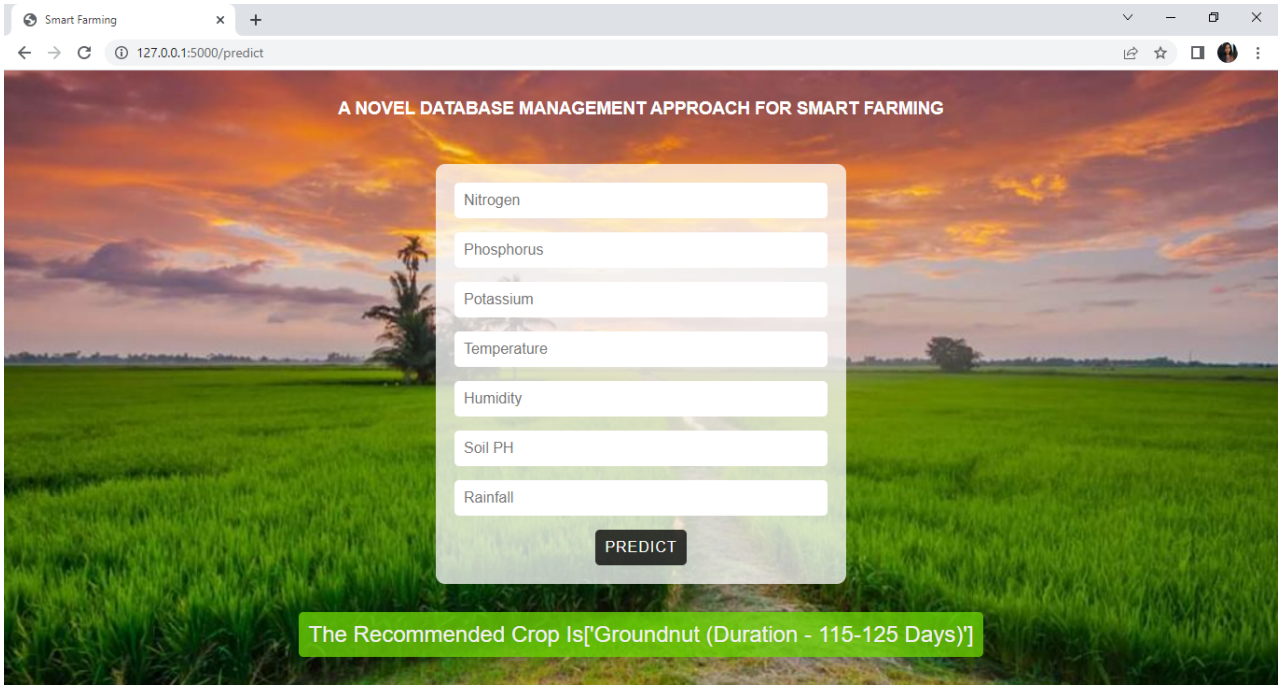
10. SCREENSHOTS



SCREENSHOT 1. DESIGN OF WEB APPLICATION



SCREENSHOT 2. INPUT IMAGE



SCREENSHOT 3. OUTPUT IMAGE

crop data.csv

final project telegram > new > crop data.csv

```
"Nitrogen","Phosphorous","Potassium","Temperature","Humidity","Soil pH","Rainfall","Crop"
"85","58","41","21.77","80.32","7.04","226.66","Rice (Duration - 100-140 days)"
"60","55","44","23.00","82.32","7.84","263.96","Rice (Duration - 100-140 days)"
"74","35","40","26.49","80.16","6.98","242.86","Rice (Duration - 100-140 days)"
"78","42","42","20.13","81.60","7.63","262.72","Rice (Duration - 100-140 days)"
"69","37","42","23.06","83.37","7.07","251.05","Rice (Duration - 100-140 days)"
"69","55","38","22.71","82.64","5.70","271.32","Rice (Duration - 100-140 days)"
"68","58","38","23.22","83.03","6.34","221.21","Rice (Duration - 100-140 days)"
"78","58","44","26.80","80.89","5.11","284.44","Rice (Duration - 100-140 days)"
"60","48","39","24.28","80.30","7.04","231.09","Rice (Duration - 100-140 days)"
"85","38","41","21.59","82.79","6.25","276.66","Rice (Duration - 100-140 days)"
"77","38","36","21.87","80.19","5.95","224.56","Rice (Duration - 100-140 days)"
"88","35","40","23.58","83.59","5.85","291.30","Rice (Duration - 100-140 days)"
"76","40","43","25.16","83.12","5.07","231.38","Rice (Duration - 100-140 days)"
"67","59","41","21.95","80.97","6.01","213.36","Rice (Duration - 100-140 days)"
"66","53","41","25.08","80.52","7.78","257.00","Rice (Duration - 100-140 days)"
"60","49","44","20.78","84.50","6.24","240.08","Rice (Duration - 100-140 days)"
"84","51","35","22.30","80.64","6.04","197.98","Rice (Duration - 100-140 days)"
"73","57","41","21.45","84.94","5.82","272.20","Rice (Duration - 100-140 days)"
"92","35","40","22.18","80.33","6.36","200.09","Rice (Duration - 100-140 days)"
"85","37","39","24.53","82.74","6.36","224.68","Rice (Duration - 100-140 days)"
"60","43","44","21.02","82.95","7.42","298.40","Rice (Duration - 100-140 days)"
"63","44","41","24.17","83.73","5.58","257.03","Rice (Duration - 100-140 days)"
"62","42","36","22.78","82.07","6.43","248.72","Rice (Duration - 100-140 days)"
"64","45","43","25.63","83.53","5.53","209.90","Rice (Duration - 100-140 days)"
"82","40","40","23.83","84.81","6.27","298.56","Rice (Duration - 100-140 days)"
"85","52","45","26.31","82.37","7.22","265.54","Rice (Duration - 100-140 days)"
"76","49","42","24.96","84.48","5.21","196.96","Rice (Duration - 100-140 days)"
"74","39","38","23.24","84.59","7.78","233.05","Rice (Duration - 100-140 days)"
"79","43","39","21.67","80.71","7.06","210.81","Rice (Duration - 100-140 days)"
"88","55","45","24.64","80.41","7.73","253.72","Rice (Duration - 100-140 days)"
"60","36","43","23.43","83.06","5.29","219.90","Rice (Duration - 100-140 days)"
"65","60","43","21.97","81.00","5.66","227.36","Rice (Duration - 100-140 days)"
```

FIG 3. CSV FILE

```
314  "89","28","33","26.44","53.84","6.99","175.37","Coffee (Duration - 3-4 years)"
315  "112","39","29","26.12","63.37","6.73","147.80","Coffee (Duration - 3-4 years)"
316  "111","28","26","27.77","64.48","6.94","192.71","Coffee (Duration - 3-4 years)"
317  "114","20","26","25.56","62.67","7.28","193.59","Coffee (Duration - 3-4 years)"
318  "117","26","30","27.92","67.97","7.08","115.23","Coffee (Duration - 3-4 years)"
319  "111","29","31","26.06","52.31","6.14","161.34","Coffee (Duration - 3-4 years)"
320  "119","30","28","26.36","64.58","6.51","163.63","Coffee (Duration - 3-4 years)"
321  "116","40","33","24.91","54.15","7.04","129.55","Coffee (Duration - 3-4 years)"
322  "95","37","35","27.31","68.42","6.35","192.43","Coffee (Duration - 3-4 years)"
323  "86","40","33","26.14","52.26","7.43","136.30","Coffee (Duration - 3-4 years)"
324  "117","37","32","23.11","67.06","6.79","162.58","Coffee (Duration - 3-4 years)"
325  "105","18","35","23.53","68.44","6.74","171.88","Coffee (Duration - 3-4 years)"
326  "109","23","25","25.12","68.48","7.01","194.88","Coffee (Duration - 3-4 years)"
327  "80","18","31","24.03","58.85","7.30","134.68","Coffee (Duration - 3-4 years)"
328  "101","31","26","26.71","69.71","6.86","158.86","Coffee (Duration - 3-4 years)"
329  "103","33","33","26.72","50.50","7.13","126.81","Coffee (Duration - 3-4 years)"
330  "93","26","27","24.59","56.47","7.29","137.70","Coffee (Duration - 3-4 years)"
331  "104","35","28","27.51","50.67","6.98","144.00","Coffee (Duration - 3-4 years)"
332  "120","72","77","17.02","16.99","7.49","88.55","Sugarcane (Duration - 12-18 months)"
333  "122","72","84","19.02","17.13","6.92","79.93","Sugarcane (Duration - 12-18 months)"
334  "132","58","85","17.89","15.41","6.00","68.55","Sugarcane (Duration - 12-18 months)"
335  "121","72","85","18.87","15.66","6.39","88.51","Sugarcane (Duration - 12-18 months)"
336  "125","67","77","18.37","19.56","7.15","79.26","Sugarcane (Duration - 12-18 months)"
337  "118","73","81","20.45","15.40","5.99","92.68","Sugarcane (Duration - 12-18 months)"
338  "122","70","84","20.65","16.61","6.23","74.66","Sugarcane (Duration - 12-18 months)"
339  "115","70","84","17.33","18.75","7.55","82.62","Sugarcane (Duration - 12-18 months)"
340  "117","62","75","18.18","18.90","7.01","81.85","Sugarcane (Duration - 12-18 months)"
341  "135","74","81","18.01","18.31","8.75","81.99","Sugarcane (Duration - 12-18 months)"
342  "123","66","79","20.99","19.33","8.72","93.55","Sugarcane (Duration - 12-18 months)"
```

FIG 3.1 CSV FILE

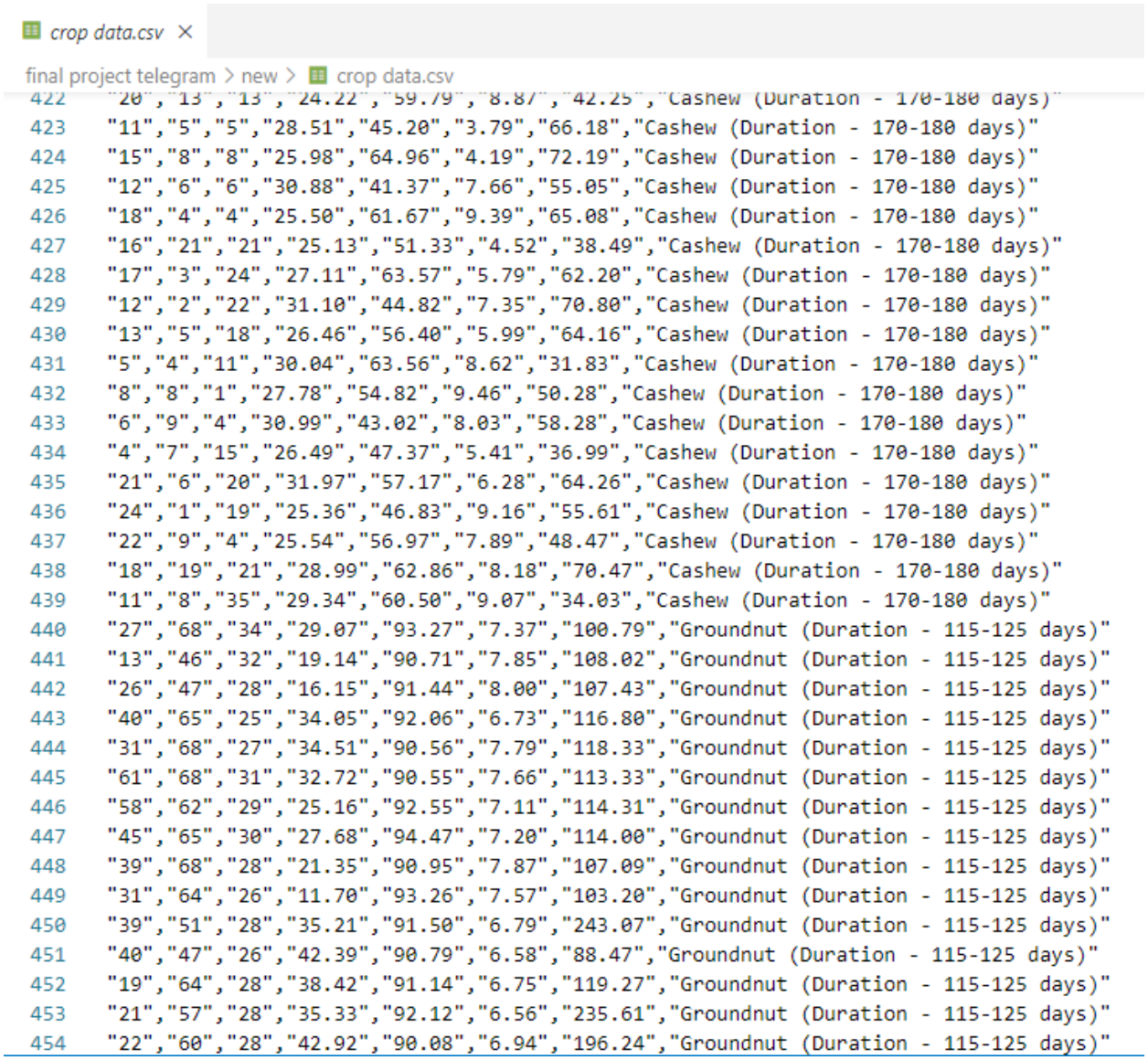


FIG 3.2 CSV FILE

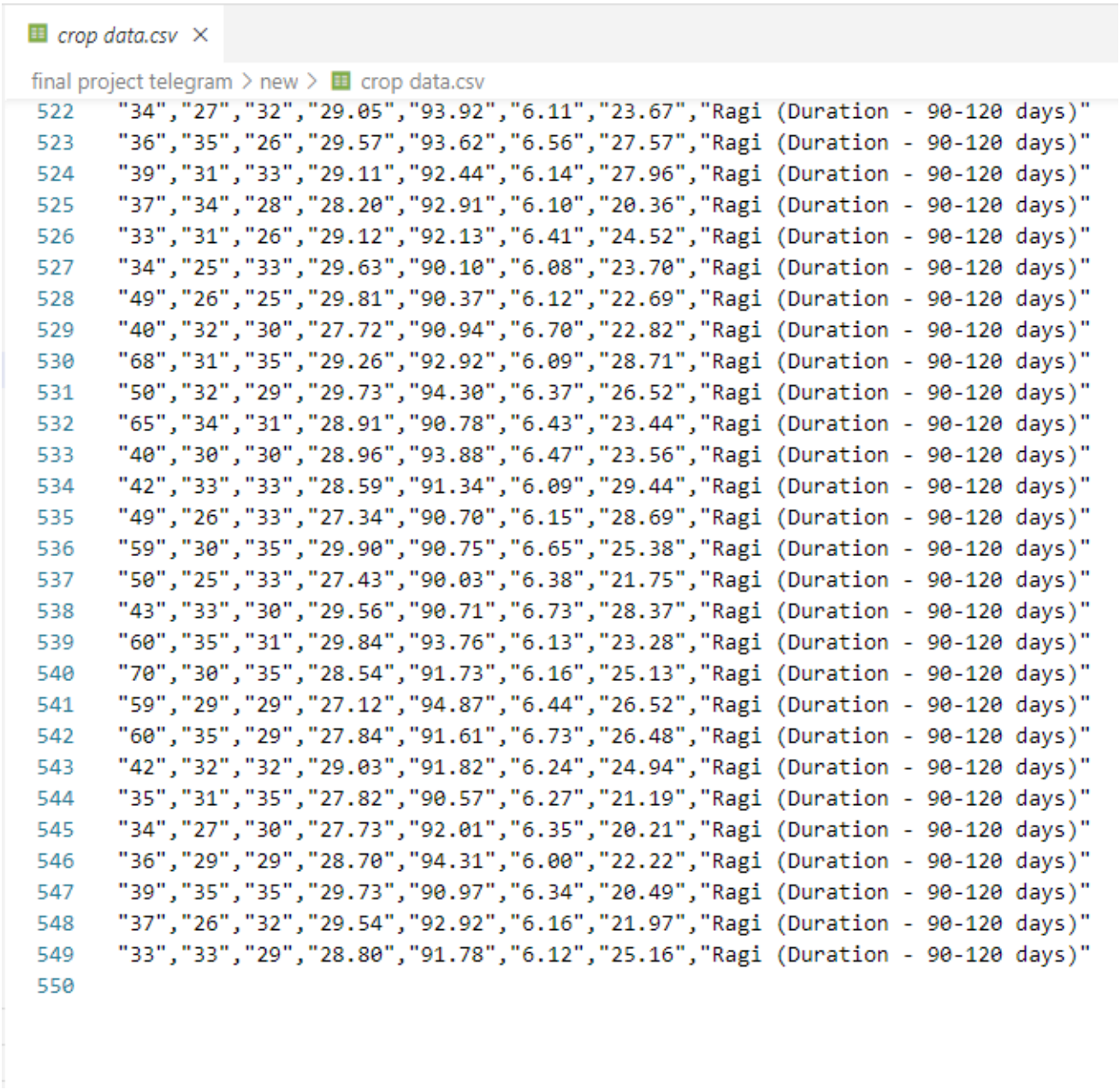


FIG 3.3 CSV FILE

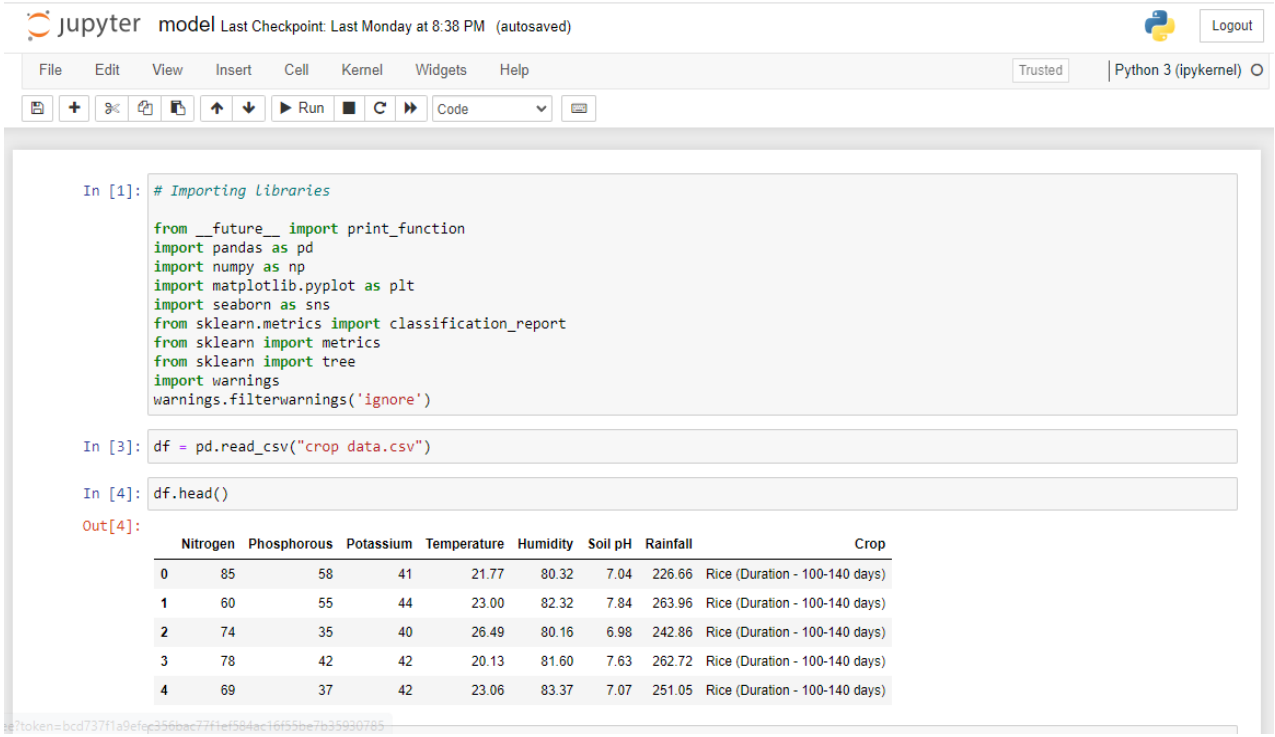


FIG 4. PREPROCESSING THE DATASET IN JUPYTER NOTEBOOK

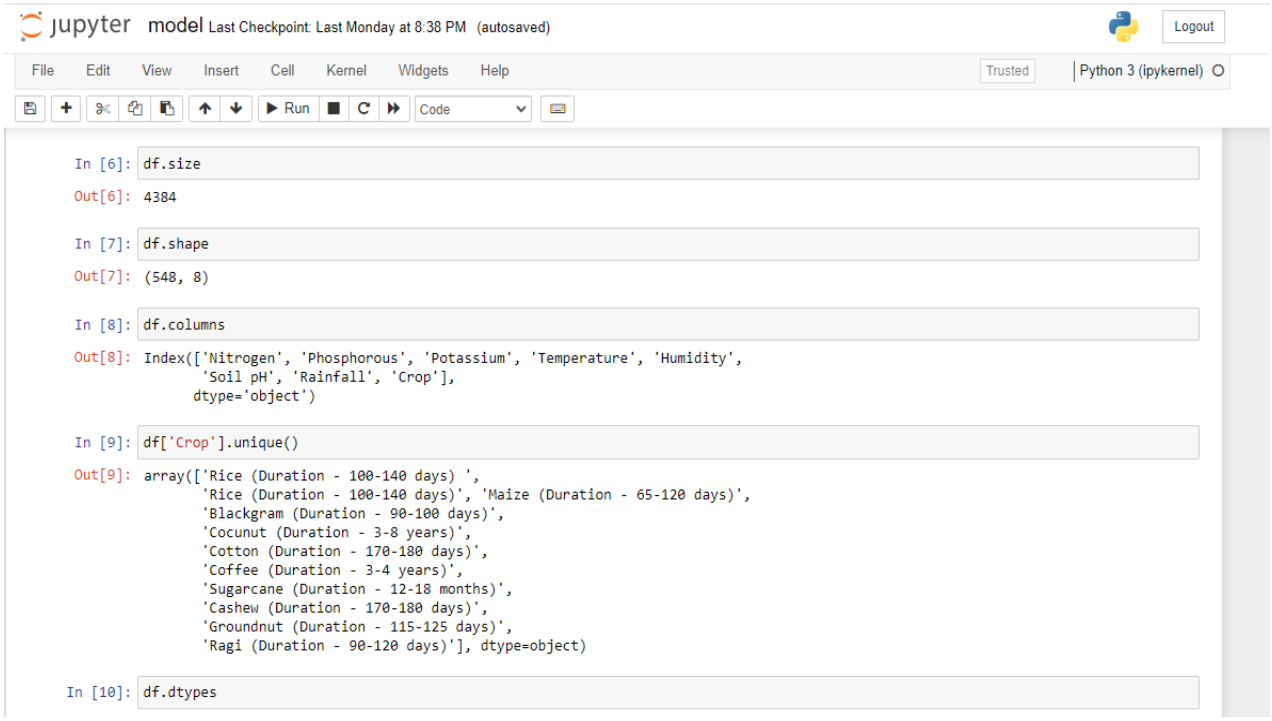


FIG 4.1 PREPROCESSING THE DATASET IN JUPYTER NOTEBOOK

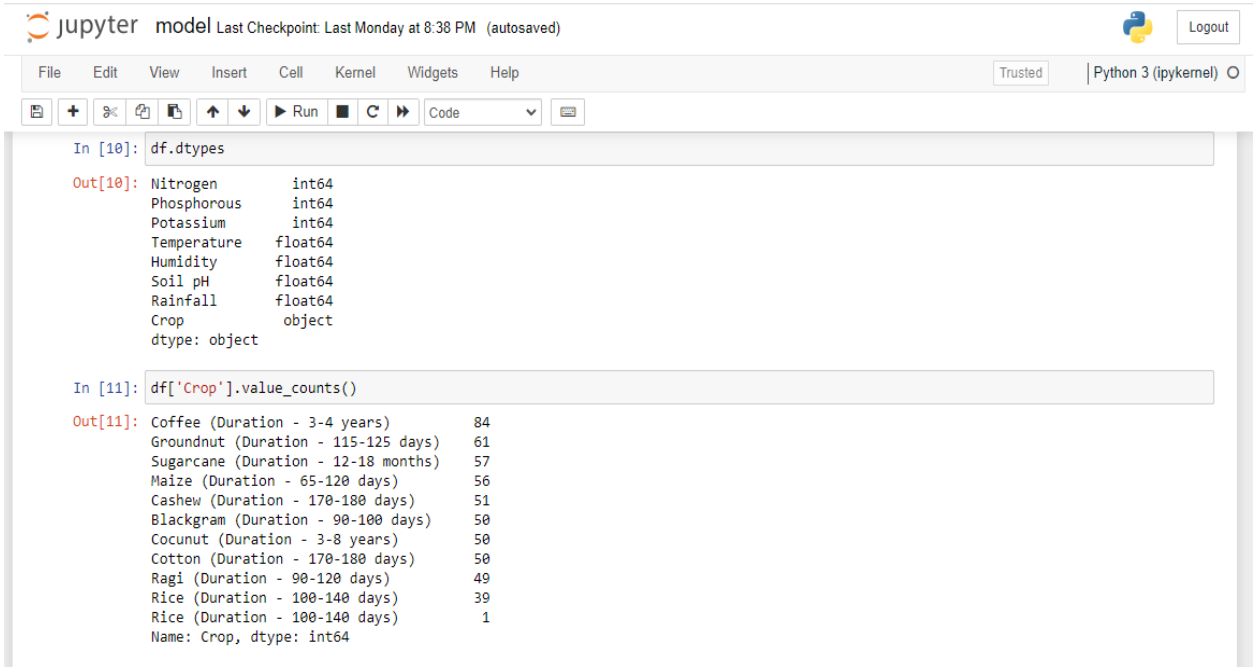


FIG 4.2 PREPROCESSING THE DATASET IN JUPYTER NOTEBOOK

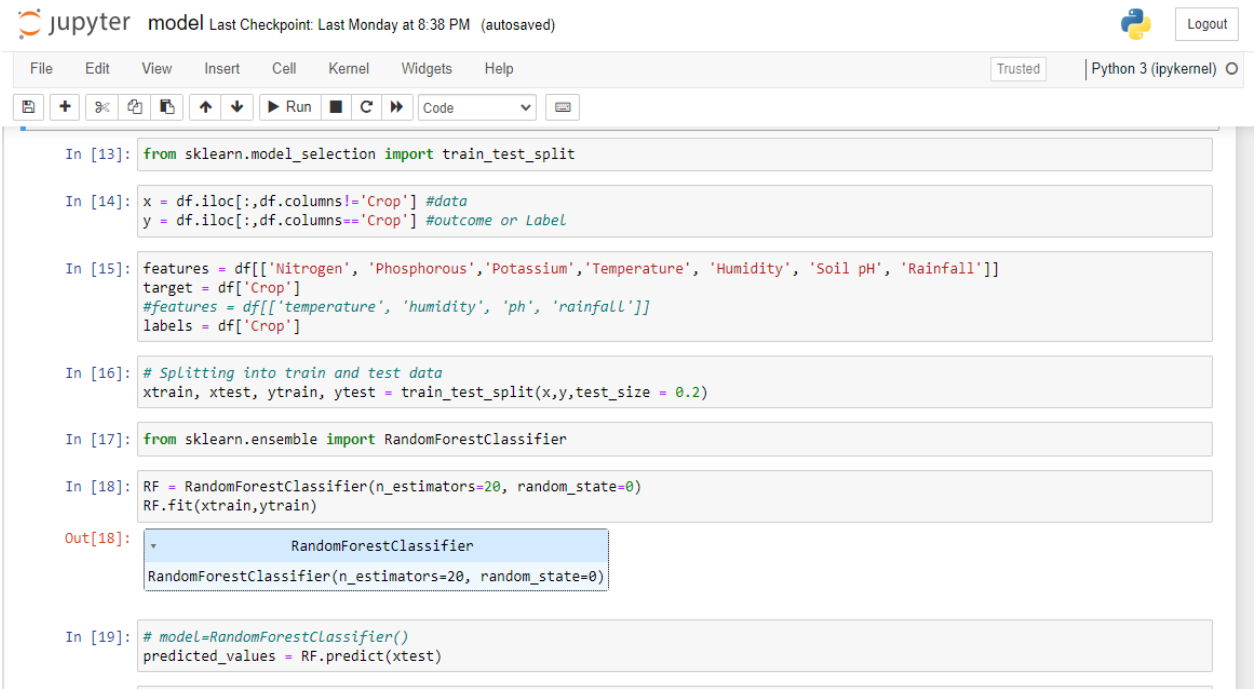


FIG 5. SPLITTING THE DATA AND TRAINING WITH RANDOM FOREST ALGORITHM

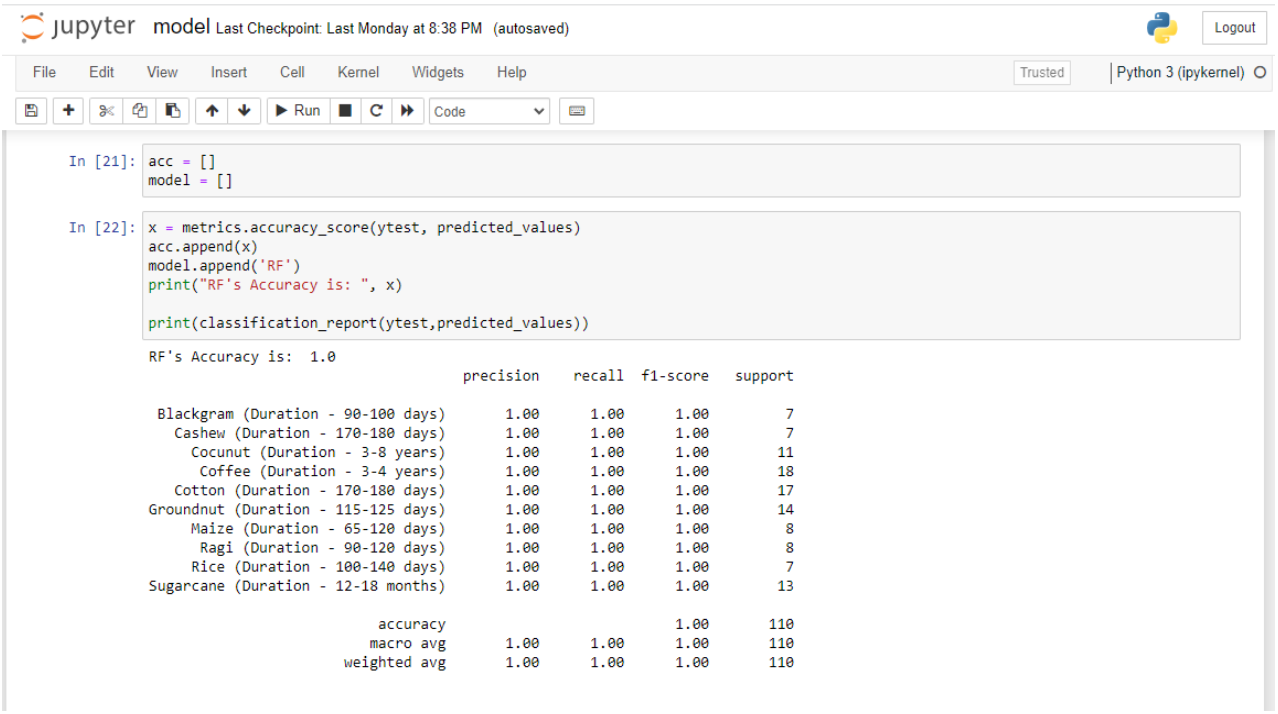


FIG 5.1 ACCURACY FOR RANDOM FOREST ALGORITHM

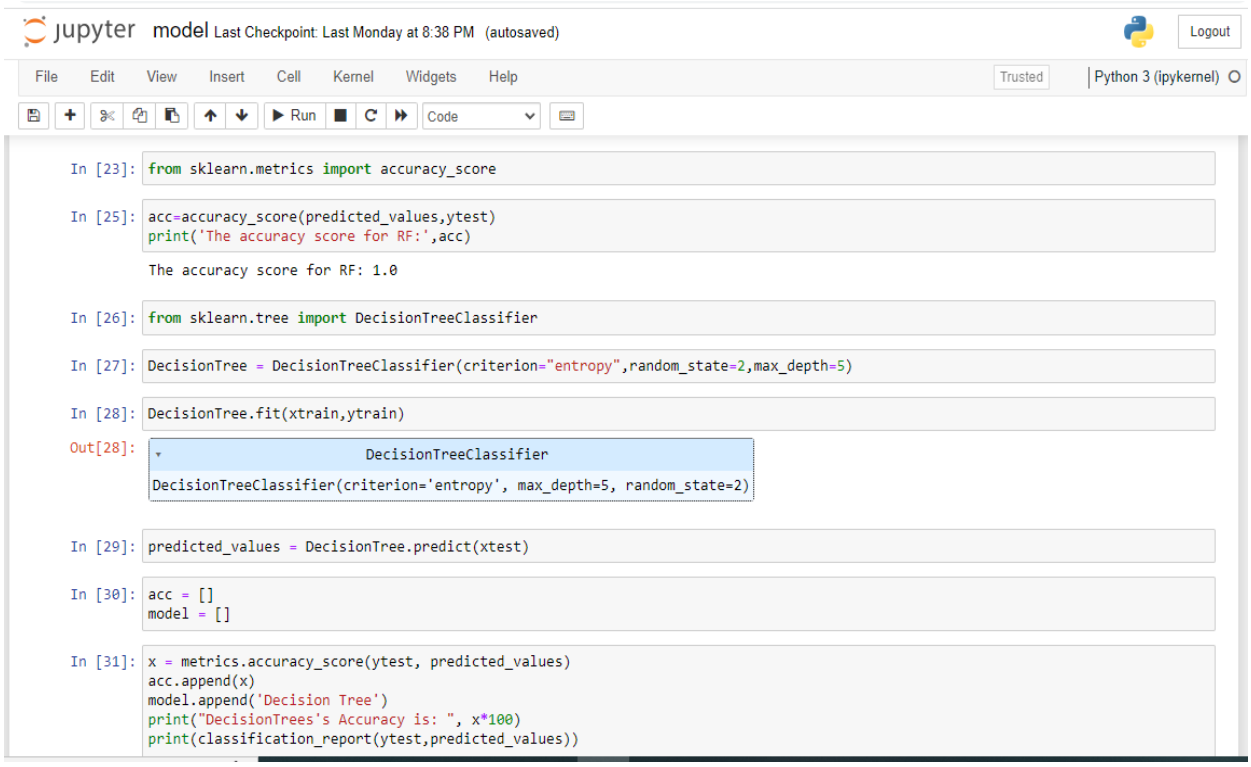


FIG 5.2 TRAINING WITH DECISION TREE ALGORITHM

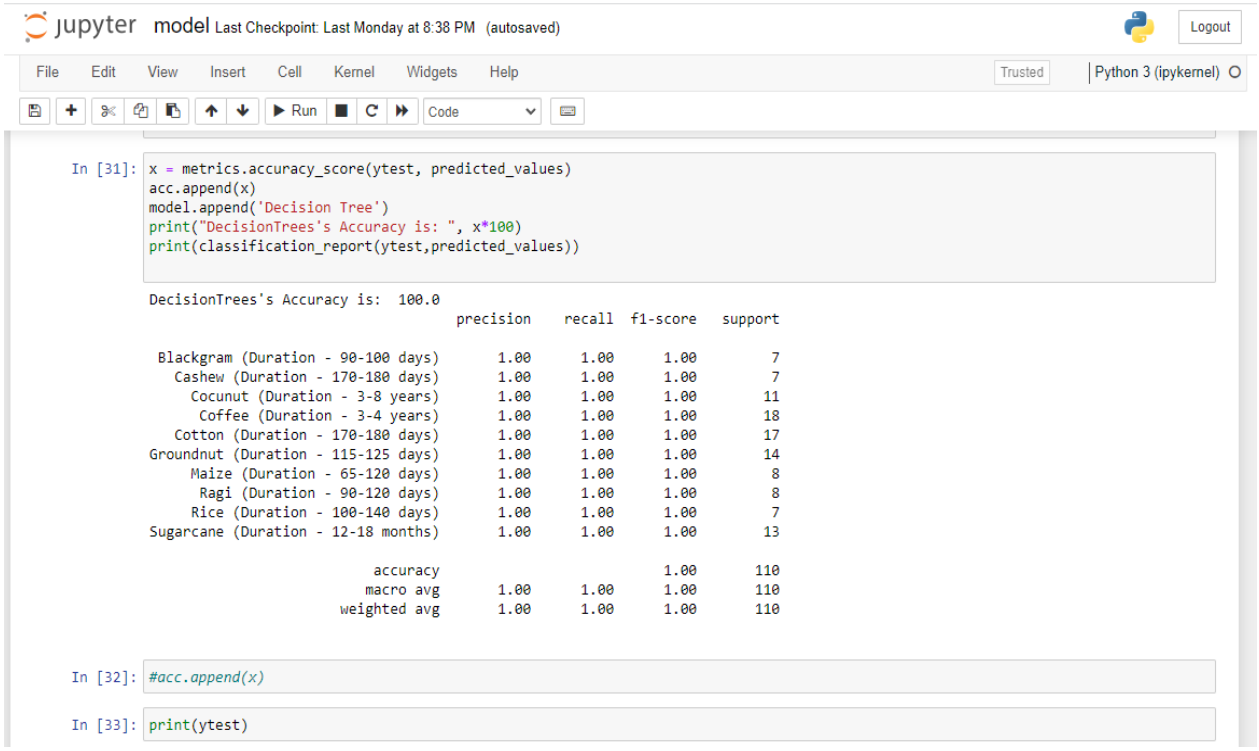


FIG 5.3 ACCURACY FOR DECISION TREE ALGORITHM

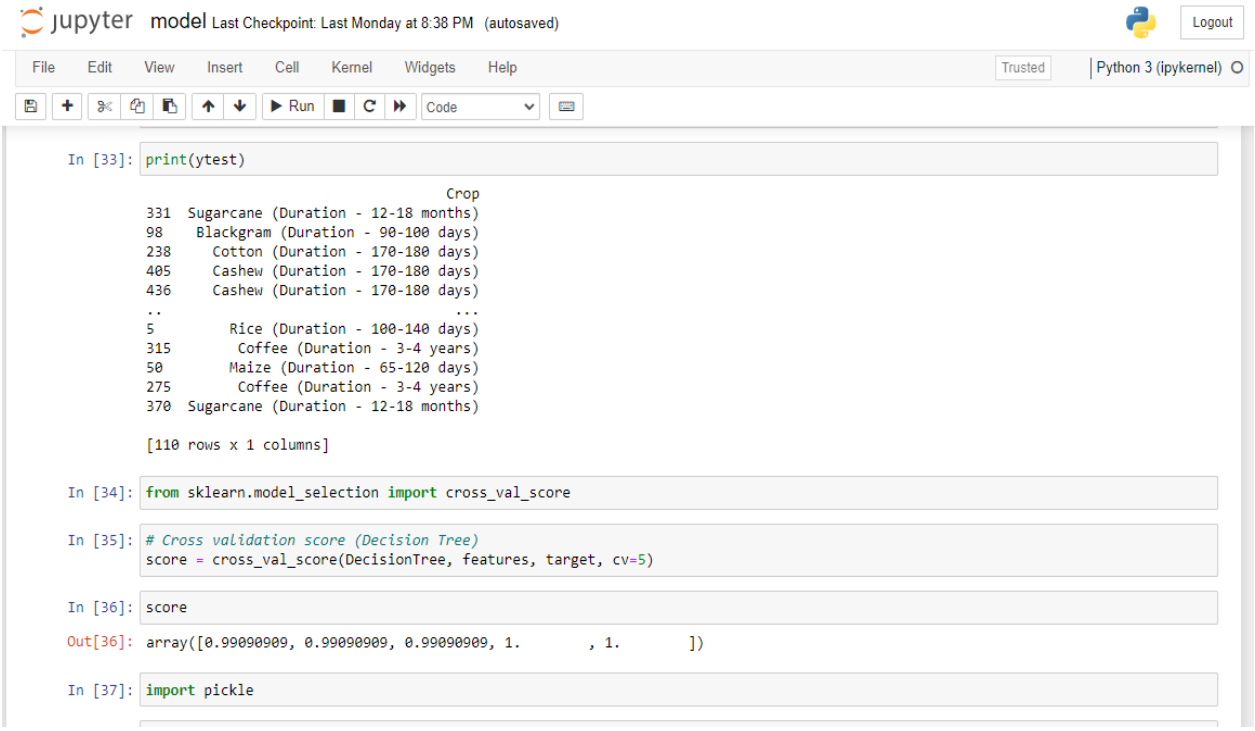


FIG 5.4 CROSS VALIDATION SCORE

FIG 6. CREATING PICKLE FILE

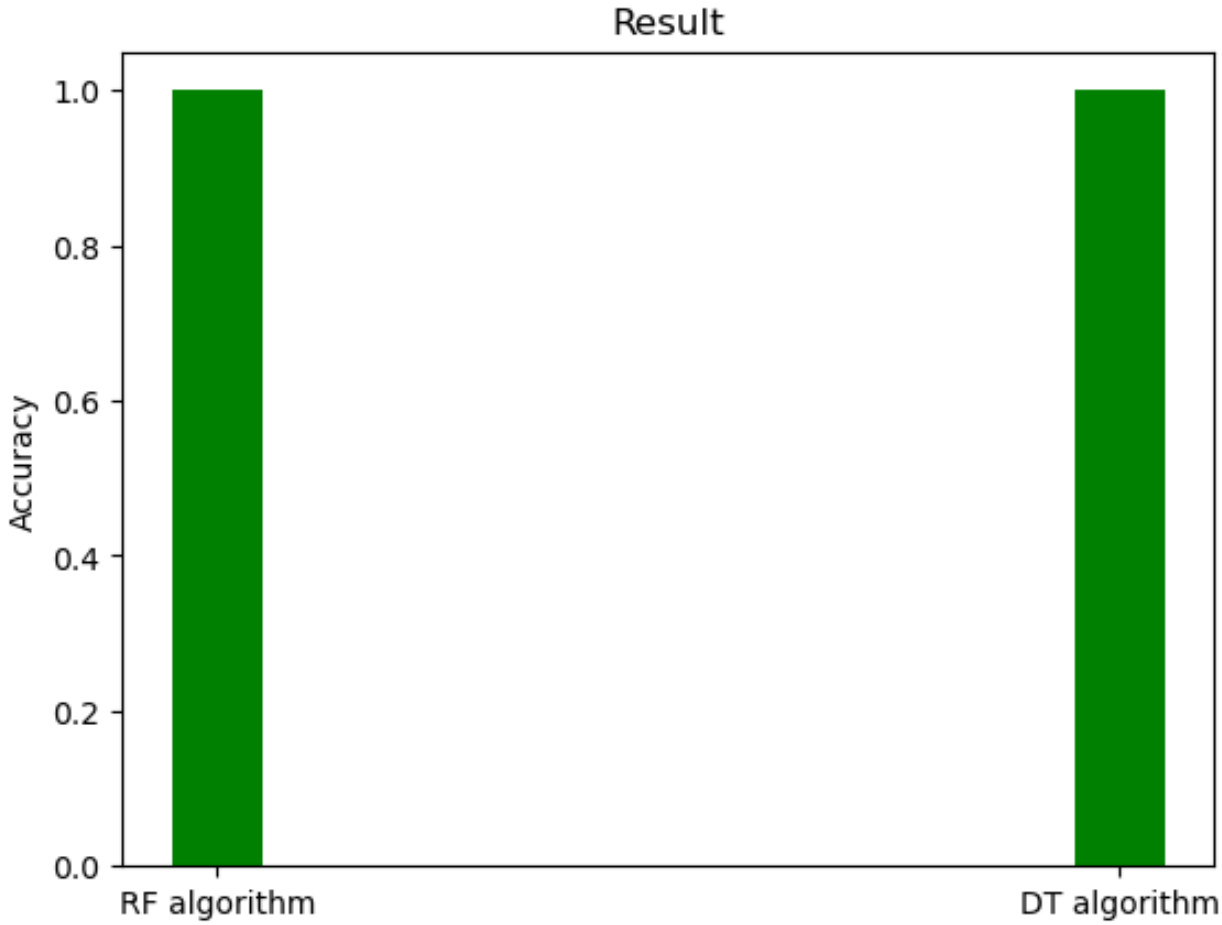


FIG 7. COMPARISON BAR GRAPH FOR ALGORITHM ACCURACY

CHAPTER 11

11. CONCLUSION AND FUTURE ENHANCEMENT

11.1 CONCLUSION

In conclusion, this crop prediction system is a valuable tool for farmers that can help them increase their crop yield and profitability. By leveraging machine learning algorithms and data on various crops, weather conditions, and soil requirements, a crop prediction system can provide personalized recommendations to farmers on the best crops to grow in their location and conditions.

The implementation of a crop recommendation system requires careful data collection, preprocessing, feature selection, model selection, and training, as well as deployment and maintenance of the system. However, with the right expertise and resources, a crop recommendation system can be a powerful tool for improving agriculture and helping farmers achieve their goals.

Overall, a crop recommendation system is a promising application of machine learning in the agricultural industry and has the potential to make a significant impact on the lives and livelihoods of farmers around the world.

11.2 FUTURE ENHANCEMENT

In the future, can try applying a data independent system. That is whatever be the format, the system should work with same accuracy. Integrating soil details to the system is an advantage, as for the selection of crops knowledge on soil is also a parameter. Proper irrigation is also a needed feature of crop cultivation. In reference to rainfall can depict whether extra water availability is needed or not. Also, we can develop the model to avoid over and under crisis of the food. When the farmers sow a particular crop, there might be some issues or diseases in the crop before harvesting. In that case, they can upload photographs of the crop and the soil report. Then the AI model can identify the problems and provide them with probable solutions.

12. REFERENCES

[1] Ramya M1, Rohit N2, Prashanth M3, Pratibha K4., (June 2020)" Design and Development of Crop Database Website to help Farmers", Dept. of Electronics and Communication, RV College of Engineering, Bangalore, Karnataka, India. Volume: 07 Issue: 06.

[2] S. Vaishnavi., M. Shobana., R. Sabitha. and S. Karthik., (March 2021) "Agricultural Crop Recommendations based on Productivity and Season", 7th International Conference on Advanced Computing and Communication Systems (ICACCS).

[3] P.Parameswari, N.Rajathi, K.J.Harshanaa (2021) "Machine Learning Approaches for Crop Recommendation", Department of IT Kumaraguru College of Technology Coimbatore,India & International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA).

[4] Sowmya, Aliya Sultana Imam Sabh Athar, Mubeena Banu, Rakshitha Jain, Divya M (2022) "Recommendation of Agricultural Crop Based on Productivity and Season Using Machine Learning" 4UG Students, Assistant Professor, Department of Information Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere, India., Volume 10 Issue VII.

[5] Benedict Bender, Clementine Bertheau, Tim Korppen, Hannah Lauppe, Norbert Gronau., (4 May 2022) "A proposal for future data organization in enterprise systems—an analysis of established database approaches", Chair of Business Informatics, Esp. Processes and Systems, University of Potsdam, Germany.