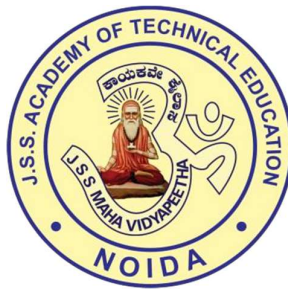# MINI PROJECT REPORT

ON

Profitability and Yield Prediction on Agricultural Crops of India

B. TECH COMPUTER SCIENCE & ENGINEERING

Submitted by

Samridh Pathela	2000910100155
Avneet Singh	2000910100043
Chaitanya Jindal	2000910100050
Gaurav Shukla	2000910100067

GROUP NO: 39

Department of Computer Science and Engineering

JSS Academy of Technical Education

Noida

ODD SEM 2022

TABLE OF CONTENTS

1.   INTRODUCTION

Agriculture is a vital sector for the economic development and food security of India, with over 50% of the population depending on agriculture for their livelihoods. However, agriculture in India faces numerous challenges, including variability in crop yields, fluctuating market prices, and limited access to resources and technology. As a result, there is a need for effective tools and strategies to improve the efficiency and profitability of agriculture in India.

Profitability and Yield Prediction on Agricultural Crops of India using python is a project that aims to address this need by using machine learning techniques to predict crop yield and profitability. By collecting and analyzing data on numerous factors that affect crop yield and profitability, such as weather conditions, soil quality, and market prices, and using machine learning algorithms to make predictions, this project can help farmers make informed decisions about what and when to plant, and how to allocate their resources. The predictive model developed in this project can also be used by agricultural organizations and government agencies to plan and allocate resources more efficiently, and to enhance food security and sustainability in India.

- The project involves collecting and analyzing data from a variety of sources, including agricultural statistics, weather data, and market data.
- The predictive model developed in this project is based on machine learning techniques, which can learn and adapt to new data and make accurate predictions based on past patterns.
- The model can be accessed through a user-friendly interface, allowing farmers to easily input data and receive predictions about crop yield and profitability.
- The project aims to provide a valuable tool for farmers and agricultural organizations to optimize their resources and improve the efficiency and profitability of their operations.
- The project also aims to contribute to the overall development and prosperity of the agricultural sector in India, by enhancing food security and sustainability and promoting sustainable and environmentally friendly practices.

2. MOTIVATION

Agriculture is the most important sector that influences the economy of India. It contributes to 18% of India 's Gross Domestic Product (GDP) and gives employment to 50% of the population of India. People of India have been practicing Agriculture for years, but the results are never satisfying due to numerous factors that affect the crop yield [3]. To fulfill the needs of around 1.2 billion people, it is especially important to have a good yield of crops.

Due to factors like soil type, precipitation, seed quality, lack of technical facilities etc. the crop yield is directly influenced. We focus on implementing crop yield prediction systems by using Machine learning techniques by doing analysis on agriculture dataset. For evaluating performance Accuracy is used as one of the factors. The classifiers are further compared with the values of Precision, Recall and F1 Score. Lesser the value of error, more accurate the algorithm will work. The result is based on comparison among the classifiers.

In agriculture, Machine Learning is considered as a novel field, as a variety of work has been done with the help of machine learning in the field of agriculture. There are different philosophies made and evaluated by the researchers all through the world in the field of agriculture and related sciences.
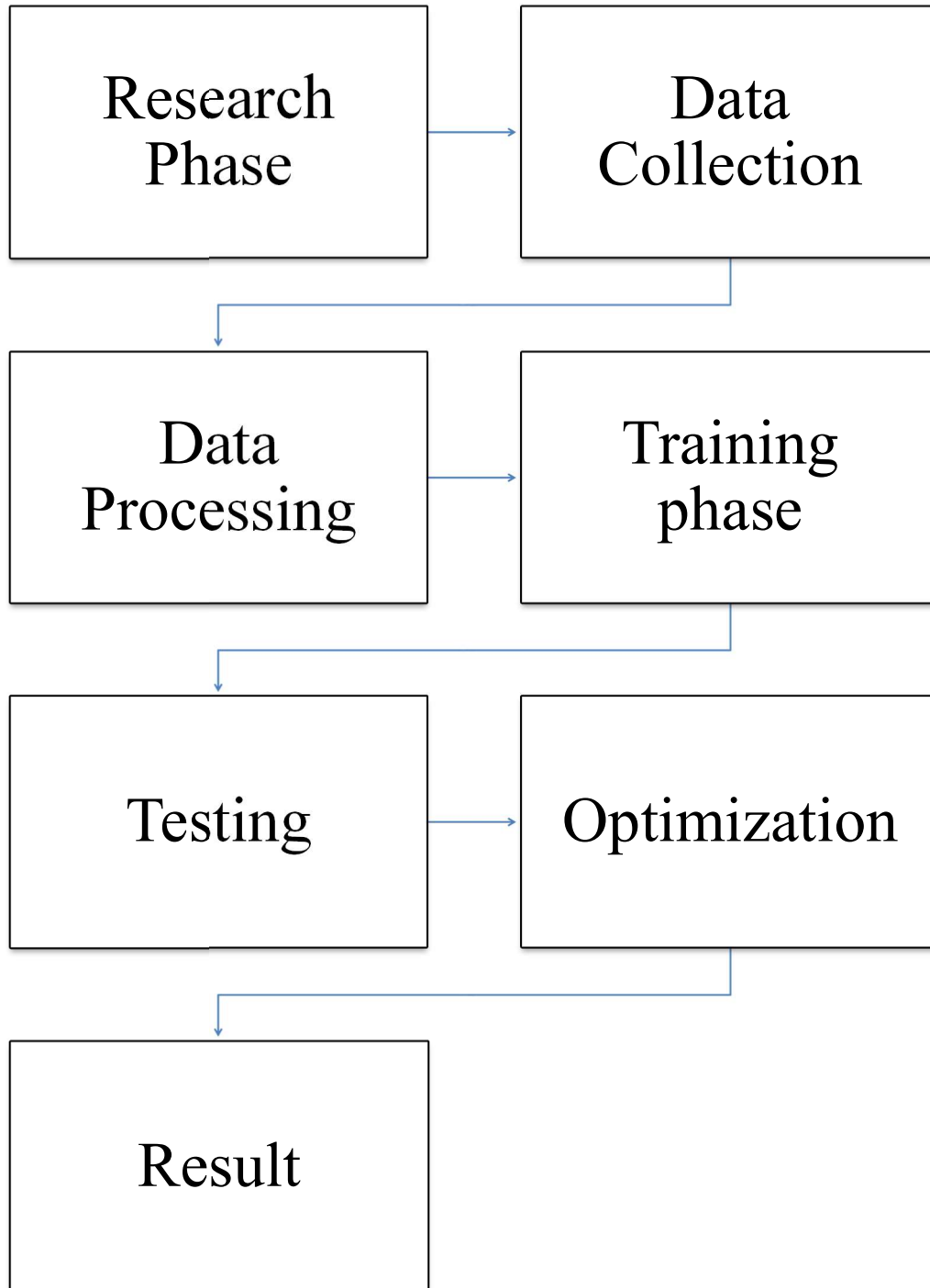
CH. Vishnu Vardhan Chowdary, Dr. K. Venkataramana [5], developed id3 algorithm for getting improved and excellent quality of crop yield of Tomato and is executed in PHP platform and datasets are used as csv. Temperature, area, humidity and the production of tomato crop are the different parameters used in this study. R. Sujatha and P. Isakki [6], utilizes data mining techniques for prediction. This model worked on different parameters such as crop name, land area, soil type, pH value, seed type, water and foreseen the boom and diseases of plants and in this way empowered to choose the descent crop dependent on climatic data and required parameters. N. Gandhi, L. J. Armstrong, O. Petkar and A. K. Tripathy [7], proposed the SVM for crop yield prediction of rice. In this method, the dataset used consists of different parameters such as place, temperature, precipitation and manufacturing. On this dataset, the implemented classifier is sequential minimal optimization.

3.   OBJECTIVES

The Proposed Work aims to meet the following objectives (s):

1.   To develop a predictive model that can accurately predict the yield and profitability of agricultural crops in India.

2.   To collect and analyse data on various factors that affect crop yield and profitability, including weather conditions, soil quality, and market prices.

3.   To use python programming language and machine learning algorithms such as linear regression and decision trees to analyse the data and make predictions.

4.   To improve the efficiency and profitability of agriculture in India by providing farmers with accurate and reliable predictions about crop yield and profitability.

5.   To help farmers make informed decisions about what and when to plant, and how to allocate their resources.

6.   To identify the most important factors that affect crop yield and profitability in India, and incorporate these factors into the predictive model.

7.   To develop a user-friendly interface for the predictive model, allowing farmers to easily access and use the model.

8.   To validate the predictive model using real-world data, and compare the model's performance with other existing models.

9.   To use the predictive model to generate forecasts for future crop yield and profitability, and provide this information to farmers in a timely manner.

10.   To work with farmers and agricultural organizations to identify areas where the predictive model can be most useful, and develop strategies for implementing and promoting the model.

4.   METHODOLOGY / PROCESS FLOW

Research Phase	→	Data Collection

Data Processing	→	Training phase

Testing	→	Optimization

Result

4.1 Data Collection:

The first step in this process is to collect data on various factors that affect crop yield and profitability, including weather conditions, soil quality, and market prices. The data is collected from various sources such as government agencies, research institutes, and online databases.

We have considered 2 datasets. One finds out the profit and classifies it if there is profit or loss, The second dataset predicts the production.

4.1.1 Dataset 1:

datainput - DataFrame

Index	Crop	State	ultivation (`/Hecta	Cultivation (`/Hec	Production (`/Qui	d (Quintal/ Hecta	Support price
0	ARHAR	Uttar Pradesh	9794.05	23076.7	1941.55	9.83	6000
1	ARHAR	Karnataka	10593.1	16528.7	2172.46	7.47	6000
2	ARHAR	Gujarat	13468.8	19551.9	1898.3	9.59	6000
3	ARHAR	Andhra Pradesh	17051.7	24171.7	3670.54	6.42	6000
4	ARHAR	Maharashtra	17130.5	25270.3	2775.8	8.72	6000
5	COTTON	Maharashtra	23711.4	33116.8	2539.47	12.69	5515
6	COTTON	Punjab	29047.1	50828.8	2003.76	24.39	5515
7	COTTON	Andhra Pradesh	29140.8	44756.7	2509.99	17.83	5515
8	COTTON	Gujarat	29616.1	42070.4	2179.26	19.05	5515
9	COTTON	Haryana	29919	44018.2	2127.35	19.9	5515

We combined data from different sources [1].The data contains columns:
- Crops
- State
- Cost of Cultivation (`/Hectare) A2+FL
- Cost of Cultivation (`/Hectare) C2
- Cost of Production (`/Quintal) C2
- Yield produced.

The profit for each row was calculated using the formula

C1 -> Cost of cultivation(`/Hectare) A2+FLC2 -> Cost of Cultivation (`/Hectare) C2
Cp -> Cost of Production (`/Quintal)
Profit = (Yield *Support Price) - (C1 + C2 + (Yield*Cp))

The govt. fixes support prices [2] per Quintal for various commodities, forexample various Kharif and Rabi crops.

If the yield produced will result in profit based on support prices declared bythe government, class 1 was allotted; else it was classified as class 0.
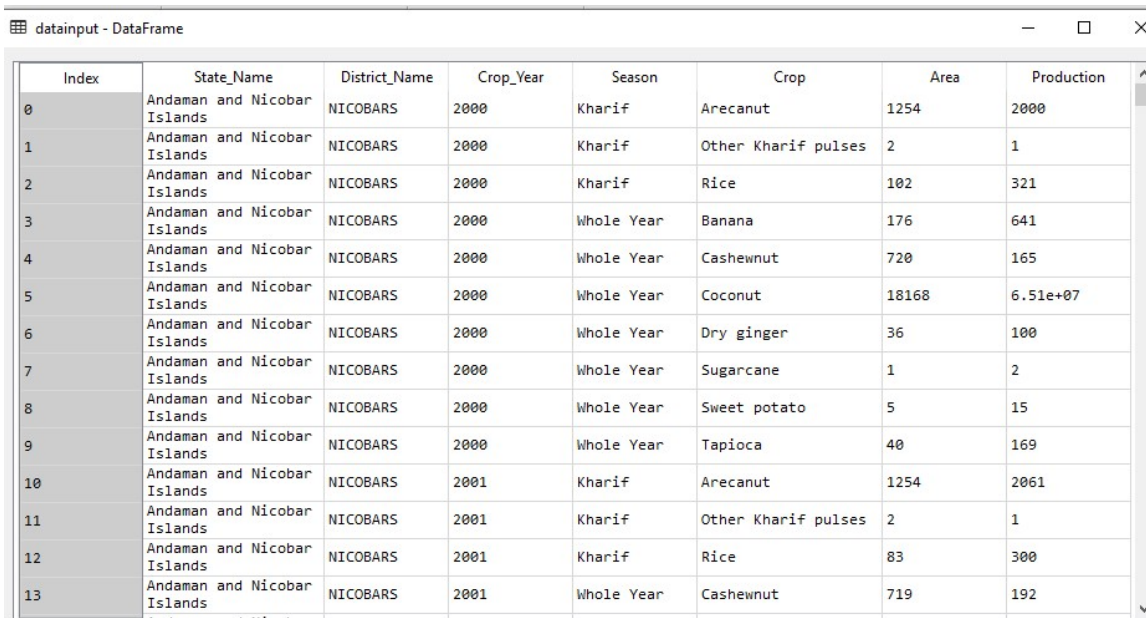
Advantages

- This dataset is compiled by using data from an official governmentsite which proves its authenticity.
- Farmers can directly find out if the crop they are about to sow willresult in profit after cultivation

Disadvantage

- Does not have many instances

4.1.2 Dataset 2:

Index	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production
0	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Arecanut	1254	2000
1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	2	1
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	102	321
3	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	176	641
4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	720	165
5	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Coconut	18168	6.51e+07
6	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Dry ginger	36	100
7	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sugarcane	1	2
8	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sweet potato	5	15
9	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Tapioca	40	169
10	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Arecanut	1254	2061
11	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Other Kharif pulses	2	1
12	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Rice	83	300
13	Andaman and Nicobar Islands	NICOBARS	2001	Whole Year	Cashewnut	719	192

In the second dataset we have the following columns:

- State_Name
- District_Name
- Crop_Year
- Season
- Crop
- Area
- Production.

We will be predicting the production of the crops using regressors.

Advantages:

- Huge dataset about 2 lakh entries
- Takes season as well as crop year into consideration

Disadvantages:

- Many missing values
- Rainfall and temperature are not considered.
- Numerous Categories: The categorical variables have many values foreach attribute.
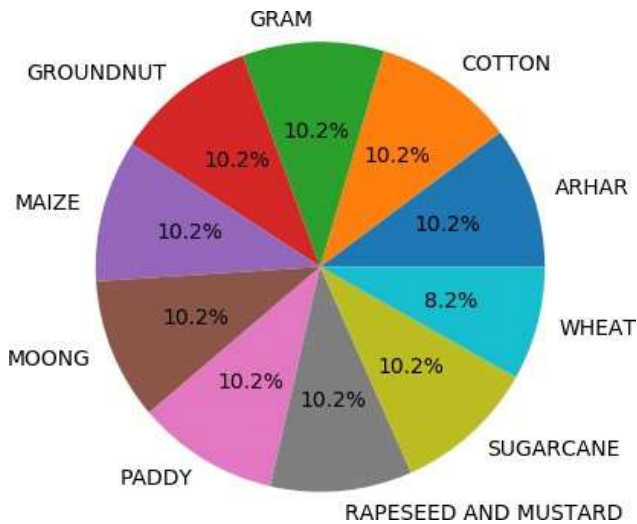
Figure 1 Crop-wise distribution in percentage.
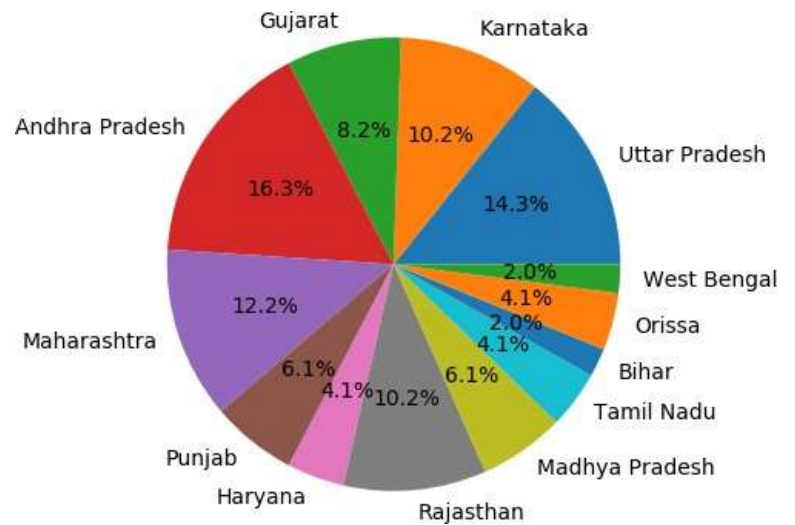
Figure 2 State wise distribution in dataset.
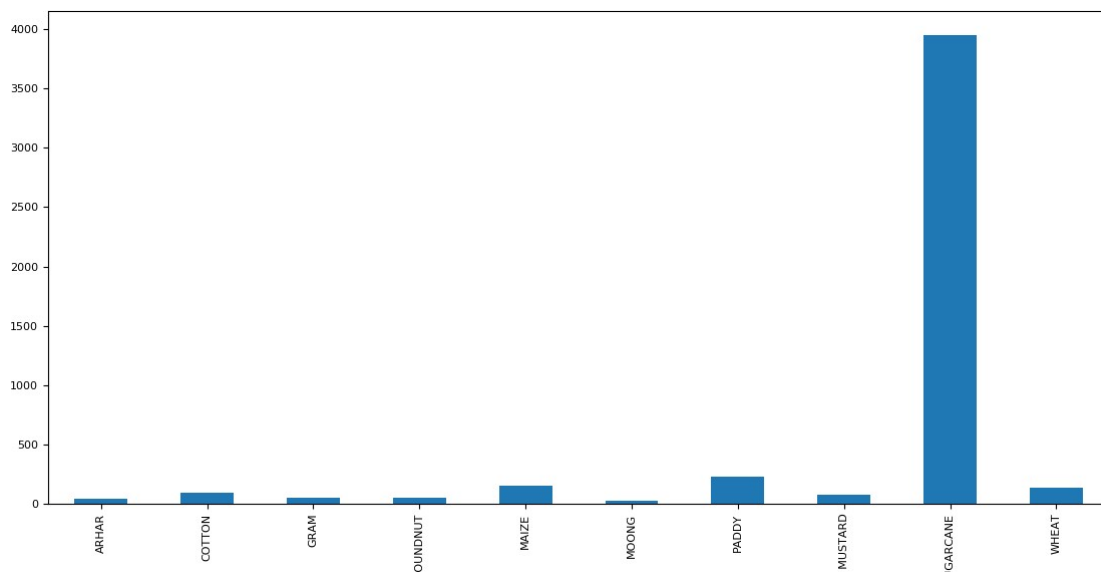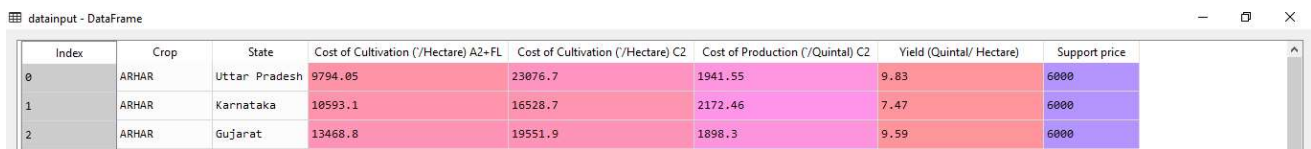
Figure 3 Total Yield crop wise.

4.2 Data Preprocessing:

Once the data is collected, it is pre-processed to ensure that it is in a usable format. This involves cleaning the data to remove any missing or incorrect values, and standardizing the data to ensure that all the variables are on the same scale.

After adding the support price column and profit in our dataset and labellingthem as 0 and 1, preprocessing techniques were applied such as missing values. The crops and state columns were encoded using labels and the onehot encoder was applied so as to avoid ranking.

Index	Crop	State	Cost of Cultivation ('/Hectare) A2+FL	Cost of Cultivation ('/Hectare) C2	Cost of Production ('/Quintal) C2	Yield (Quintal/ Hectare)	Support price
0	ARHAR	Uttar Pradesh	9794.05	23076.7	1941.55	9.83	6000
1	ARHAR	Karnataka	10593.1	16528.7	2172.46	7.47	6000
2	ARHAR	Gujarat	13468.8	19551.9	1898.3	9.59	6000

A snapshot of Data frame

Dataset 2 contains many missing values. During the preprocessing step these rows are dropped since the number of instances is very large.

4.3 Model Development:

After the data is pre-processed, machine learning algorithms such as linear regression and decision trees are used to analyse the data and make predictions. The algorithms are trained on a subset of the data, and their performance is tested on a separate validation set to ensure that the models are accurate and reliable.

In the system, we propose tests of many algorithms and by studying the classification report we compare the algorithms and choose the best one. It must find accuracy of the training dataset, accuracy of the testing dataset, specification, False Positive rate, precision and recall by comparing algorithms using python code.

The following Involvement steps are:

1. Define a problem
2. Preparing data
3. Evaluating algorithms
4. Improving results
5. Predicting results

We will be applying classification algorithms on dataset 1 and regression for prediction of Production on dataset 2.

Algorithms applied:

1. Classification
   o Decision Tree
   o Logistic Regression
   o K nearest neighbor
   o Random forest Classifier
2. Clustering
3. Regression
   o Decision Tree
   o Random forest

4.4 Model Evaluation:

The developed models are evaluated using various performance metrics such as mean absolute error, root mean squared error, and accuracy. The models are also compared with other existing models to determine their effectiveness.

Let us first briefly understand some of the performance evaluation metrics:

4.4.1 General Definitions:

1. True Positive (TP) depicts the number of instances where the system detects for a condition when it is really present.
2. True Negative (TN) depicts the number of instances where the system does not detect a condition when it is absent.

Observations to the total predicted positive observations. Low false positive rate means high precision. In this research the precision 0.788 is obtained which is pretty good.

4.4.2 Recall:

Positive observed values proportion is correctly predicted. (Actual defaulter's model will correctly predict the proportion)

Recall = TP / (TP + FN)

Recall (Sensitivity) - Recall is the ratio of correctly predicted positive observations to the all observations in actual class -yes.

4.4.3 F1 Score:

F1 score is the process of finding the calculated weighted average of Precision and Recall. The score is considered for both false positives and false negatives. Intuitively it is not easy to understand accuracy, but F1 is usually more useful than accuracy, especially if uneven classdistribution is considered. Accuracy is the best way, if false positives and false negatives have similar cost. To better look at the precision and recall, the cost of false positives and false negatives should be very different.
General Formula:
F- Measure = 2TP / (2TP + FP + FN)
F1-Score Formula:
F1 Score = 2*(Recall * Precision) / (Recall + Precision)

4.4.4 Precision:

When the model predicts positive, how often is it correct? Precision = TP/TP+FP Precision helps when the costs of false positives are high. So let's assume the problem involves the detection of skin cancer. If we have a model that has very low precision, then many patients will be told that they have melanoma, and that will include some misdiagnoses. Lots of extra tests and stress are at stake. When false positives are too high, those who monitor the results will learn to ignore them after being bombarded with false alarms.

4.4.4.1 Decision Tree:

- **Confusion Matrix:**

- **Classification Report:**

Class	Precision	Recall	F1 Score	Support (num ofex-amples)
0	0.78	1.00	0.88	7
1	1.00	0.75	0.86	8
Accuracy	0.87			15

- **R2 score (Dataset 2):**

The R2 score comes out to be 0.84 using the Decision tree regressor. The mean absolute error comes out to be 167163.3086041714.

4.4.4.2 Logistic Regression:

- **Confusion Matrix:**
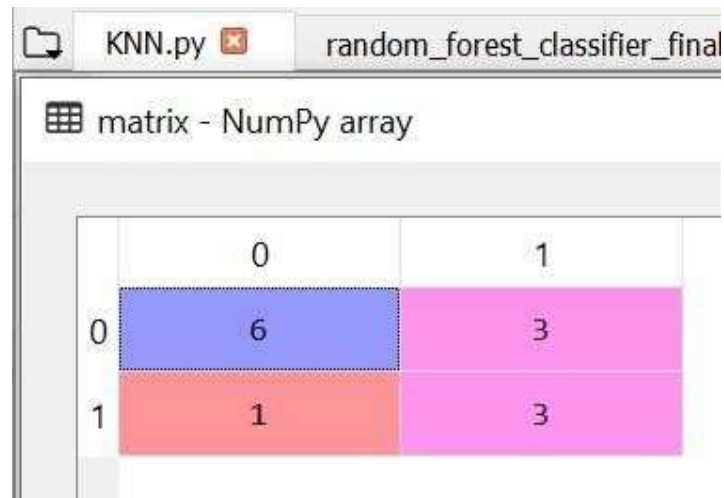
- **Classification Report:**

Class	Precision	Recall	F1 Score	Support (num ofex-amples)
0	1.00	0.86	0.92	7
1	0.89	1.00	0.94	8
Accuracy	0.93			15

4.4.4.3   K-Nearest Neighbor:

- **Confusion Matrix:**
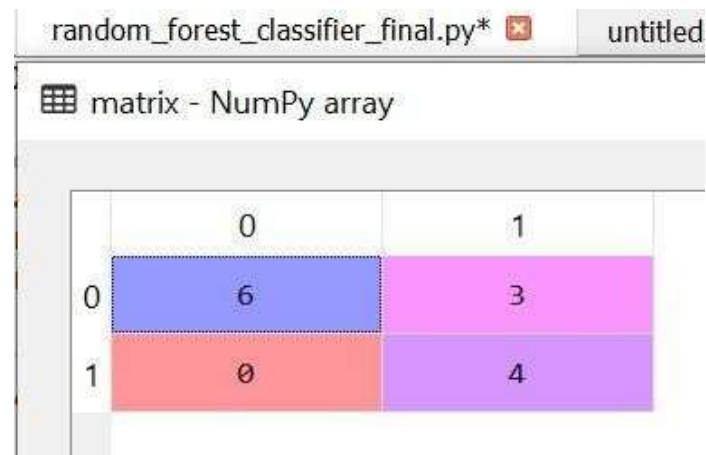
- **Classification Report:**

Class	Precision	Recall	F1 Score	Support (num ofex-amples)
0	0.50	0.75	0.60	4
1	0.86	0.67	0.75	9
Accuracy	0.69			13

4.4.4.4   Random Forest Classifier:

- **Confusion Matrix:**

- **Classification Report:**

Class	Precision	Recall	F1 Score	Support (num of examples)
0	0.57	1.00	0.73	4
1	1.00	0.67	0.80	9
Accuracy	0.77			13

- **Random Forest Regressor on Dataset 2:**

The mean absolute error comes out to be 155503.99436675265. The R2 score is 0.91.

4.5 Clustering:

After applying clustering we plotted the elbow graph to check how many clusters gave optimal results. An ideal way to figure out the right number of clusters would be to calculate the Within-Cluster-Sum-of-Squares (WCSS). WCSS is the sum of squares of the distances of each data point in all clusters to their respective centroids. The idea is to minimise the sum.
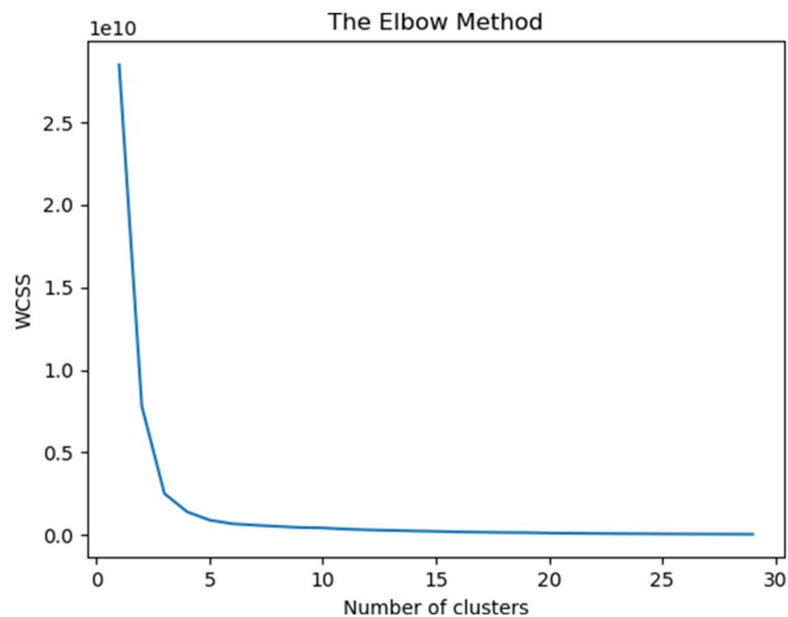
This graph shows that 3 clusters are best suited for the dataset.

4.6 Result & Discussion:

Algorithm	Precision		Recall		F1 Score		Accuracy
	Class 0	Class 1	Class 0	Class 1	Class 0	Class 1	
Logistic Regression	1.0	0.89	0.86	1.0	0.92	0.94	0.93
DecisionTree	0.78	1.0	1.0	0.75	0.88	0.86	0.87
Randomforest	0.86	0.75	0.75	0.86	0.80	0.80	0.80
K nearest	0.50	0.86	0.75	0.67	0.60	0.75	0.69

Algorithm	R2 score	Mean absolute error
Decision Tree	0.84	167163.3086041714
Random Forest	0.91	155503.99436675265

For classification algorithms, Logistic regression performed the best for predicting the profit on a given crop, state, costs of cultivation (C1,C2), cost of production (Cp), and support prices provided by the government for the year 2020-21.

As we saw the second dataset did not perform so well. Additional columns like rainfall and temperature need to be added to improve the accuracy of the models.

4.7 Model Deployment:

If the developed models meet the required accuracy threshold, they are deployed for use by farmers. The model can be accessed through a user-friendly interface, allowing farmers to easily input data and receive predictions about crop yield and profitability.

4.8 Model Maintenance:

The developed models are continuously monitored and updated as needed to ensure their accuracy and effectiveness. Any changes or updates to the model are carefully tested to ensure that they do not negatively impact its performance.

5. HARDWARE & SOFTWARE REQUIREMENTS
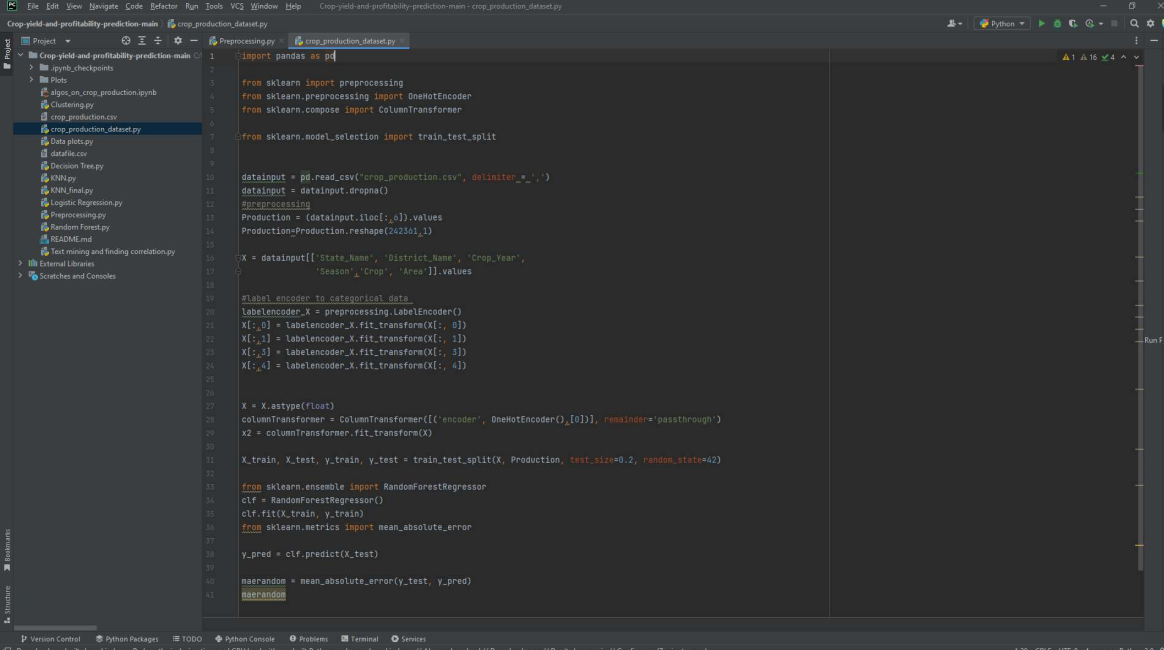
Hardware Requirements:

- A computer or server with sufficient processing power and memory to handle the data analysis and machine learning tasks involved in the project.

- A reliable internet connection for accessing data sources and deploying the predictive model.

- Storage: A large amount of data may be collected and analysed as part of this project, so it is important to have sufficient storage capacity to store the data and the machine learning models.

- Graphics processing unit (GPU): A GPU may be required if the project involves training machine learning models that are computationally intensive, such as deep learning models.

Software Requirements:
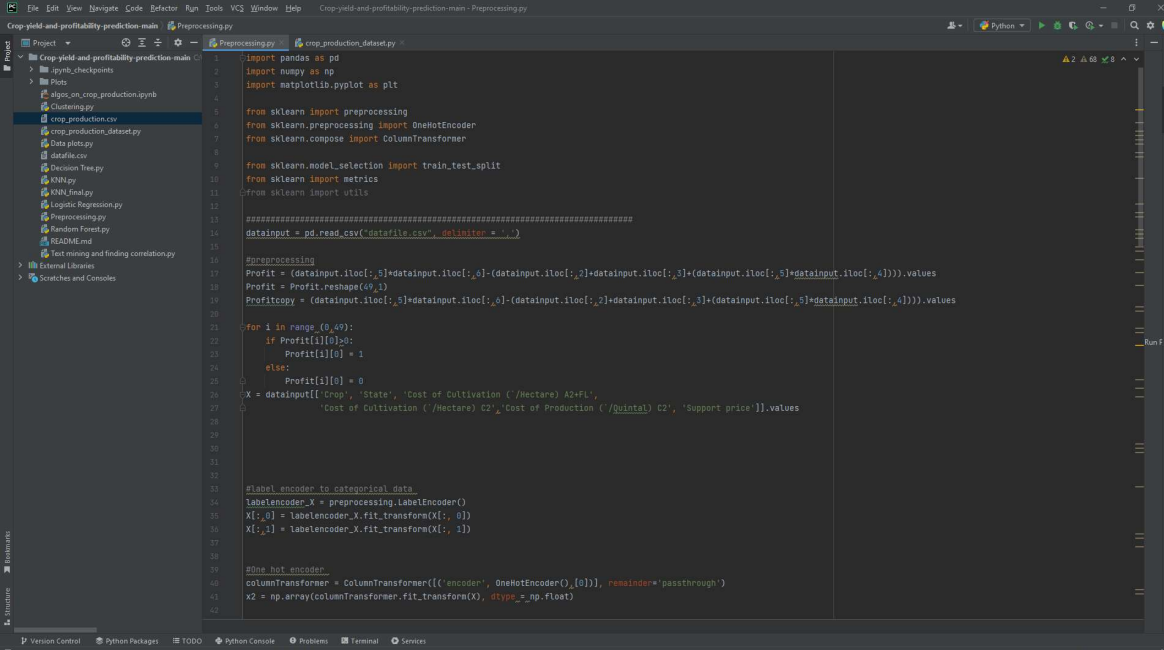
- Python programming language: Python is the primary programming language used for data analysis and machine learning tasks in this project.

- Data manipulation libraries: Libraries such as Pandas and NumPy used to manipulate and process the data.

- Machine learning libraries: Python has numerous libraries for machine learning tasks such as scikit-learn, TensorFlow, and PyTorch. These libraries can be used to develop and train machine learning models.

- Data visualization libraries: Python libraries such as Matplotlib and Seaborn can be used to visualize and analyse the data.

- Cloud computing platforms: Cloud computing platforms such as AWS or Google Cloud may be used to store and analyse the data, and to deploy the predictive model.

- Data management tools: Tools such as MySQL or MongoDB may be used to store and manage the data.

- Collaboration and project management tools: Tools such as Trello or Asana can be used to collaborate and manage the project.

6. SNAPSHOTS OF PROJECT

Data Collection

Data Preprocessing

Decision Tree

Logistic Regression

K-Nearest Neighbor

Random Forest

Clustering

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

from sklearn import preprocessing
from sklearn.preprocessing import OneHotEncoder
from sklearn.compose import ColumnTransformer

from sklearn.model_selection import train_test_split

###################################################################
datainput = pd.read_csv("datafile.csv", delimiter = ',')

#preprocessing
Profit = (datainput.iloc[:,5]*datainput.iloc[:,6]-(datainput.iloc[:,2]+datainput.iloc[:,3]+(datainput.iloc[:,5]*datainput.iloc[:,4]))).values
Profit = Profit.reshape(49,1)
Profitcopy = (datainput.iloc[:,5]*datainput.iloc[:,6]-(datainput.iloc[:,2]+datainput.iloc[:,3]+(datainput.iloc[:,5]*datainput.iloc[:,4]))).values

for i in range (0,49):
    if Profit[i][0]>0:
        Profit[i][0] = 1
    else:
        Profit[i][0] = 0
X = datainput[['Crop', 'State', 'Cost of Cultivation (`/Hectare) A2+FL',
               'Cost of Cultivation (`/Hectare) C2','Cost of Production (`/Quintal) C2', 'Support price']].values

#label encoder to categorical data
labelencoder_X = preprocessing.LabelEncoder()
X[:,0] = labelencoder_X.fit_transform(X[:, 0])
X[:,1] = labelencoder_X.fit_transform(X[:, 1])


#One hot encoder
columnTransformer = ColumnTransformer([('encoder', OneHotEncoder(), [0])], remainder='passthrough')
x2 = np.array(columnTransformer.fit_transform(X), dtype = np.float)

columnTransformer = ColumnTransformer([('encoder', OneHotEncoder(), [10])], remainder='passthrough')
x3 = np.array(columnTransformer.fit_transform(x2), dtype = np.float)

#output col in y
```

Plotting

```python
import pandas as pd
import matplotlib.pyplot as plt

###################################################################

###################################################################
datainput = pd.read_csv("datafile.csv", delimiter = ',')

###################################################################
#Plot 1. Crops in percentage
Crops = datainput.iloc[:,0]
CropsCount = {}
for crop in Crops:
    CropsCount[crop] = CropsCount.get(crop, 0)+1

#extract values and keys of dict:CropsCount
labels = list(CropsCount.keys())
values = list(CropsCount.values())

plt.pie(values, labels=labels, autopct='%1.1f%%')
plt.show()

###################################################################
#Plot 2. State distribution
States = datainput.iloc[:,1]
StatesCount = {}
for state in States:
    StatesCount[state] = StatesCount.get(state, 0)+1

#extract values and keys of dict:CropsCount
labels = list(StatesCount.keys())
values = list(StatesCount.values())

plt.pie(values, labels=labels, autopct='%1.1f%%')
plt.show()

###################################################################
#Plot 3. Yield per Crop
plt.tick_params(labelsize=8)
datainput.groupby("Crop")["Yield (Quintal/ Hectare) "].agg("sum").plot.bar()

###################################################################
#Plot 4.
```

7.   APPLICATION OF PROJECT

The application of Profitability and Yield Prediction on Agricultural Crops of India using python has the potential to greatly improve the efficiency and profitability of agriculture in India. Some potential applications of this project include:

1. Providing farmers with accurate and reliable predictions about crop yield and profitability: The predictive model developed in this project can be used by farmers to make informed decisions about what and when to plant, and how to allocate their resources. This can help farmers maximize their profits and minimize their risks.

2. Improving crop planning and resource allocation: The predictive model can be used by agricultural organizations and government agencies to plan and allocate resources more efficiently. For example, the model can be used to identify areas where certain crops are most likely to be profitable, and to allocate resources accordingly.

3. Reducing waste and improving resource utilization: By providing accurate predictions about crop yield and profitability, the predictive model can help farmers avoid planting crops that are likely to be unprofitable, reducing waste and improving resource utilization.

4. Enhancing food security and sustainability: By improving the efficiency and profitability of agriculture, the predictive model can help to enhance food security and sustainability in India. This can be achieved by reducing food waste, improving resource utilization, and minimizing the environmental impact of agriculture.

5. Overall, the application of Profitability and Yield Prediction on Agricultural Crops of India using python has the potential to greatly benefit the agricultural sector in India and contribute to the overall development and prosperity of the country.

8. CONCLUSION

In conclusion, Profitability and Yield Prediction on Agricultural Crops of India using python is a valuable project that has the potential to greatly improve the efficiency and profitability of agriculture in India. By collecting and analyzing data on various factors that affect crop yield and profitability, and using machine learning algorithms to make predictions, this project can help farmers make informed decisions about what and when to plant, and how to allocate their resources. The predictive model developed in this project can also be used by agricultural organizations and government agencies to plan and allocate resources more efficiently, and to enhance food security and sustainability in India.

Overall, the application of Profitability and Yield Prediction on Agricultural Crops of India using python has the potential to greatly benefit the agricultural sector in India, and contribute to the overall development and prosperity of the country.

- The predictive model developed in this project can be accessed through a user-friendly interface, allowing farmers to easily input data and receive predictions about crop yield and profitability.

- The model can be continuously monitored and updated as needed to ensure its accuracy and effectiveness.

- The project can work with farmers and agricultural organizations to identify areas where the predictive model can be most useful, and develop strategies for implementing and promoting the model.

- The project can also explore the potential for using the predictive model to optimize supply chain management, facilitate market analysis and decision-making, and promote sustainable and environmentally friendly practices in the agricultural sector.

- Overall, the Profitability and Yield Prediction on Agricultural Crops of India using python project has the potential to greatly benefit the agricultural sector in India, and contribute to the overall development and prosperity of the country.

9.   FUTURE SCOPE

There are many potential areas for future research and development in the field of Profitability and Yield Prediction on Agricultural Crops of India using python. Some potential areas for future work include:

• Improving the accuracy and reliability of the predictive model: The accuracy of the model can be improved by collecting and analyzing more data, and by developing and testing more advanced machine learning algorithms.

• Expanding the scope of the model to cover additional crops and regions: The model can be expanded to cover a wider range of crops and regions, in order to provide more comprehensive predictions about crop yield and profitability.

• Developing more advanced predictive models: The project can explore the use of more advanced machine learning techniques such as deep learning, in order to improve the accuracy and reliability of the model.

• Enhancing the user-friendliness and accessibility of the model: The project can work to make the model more user-friendly and accessible for farmers, for example by developing a mobile app or integrating the model with existing agricultural tools and platforms.

• Investigating the potential for using the model to optimize resource allocation and improve sustainability: The project can explore the potential for using the model to optimize resource allocation and promote sustainable and environmentally friendly practices in the agricultural sector.

Overall, there are many potential areas for future research and development in the field of Profitability and Yield Prediction on Agricultural Crops of India using python, with the potential to greatly benefit the agricultural sector and contribute to the overall development and prosperity of the country.

10.   REFERENCES

[1] https://data.gov.in/

[2] https://farmer.gov.in/mspstatements.aspx

[3] N. P. Sastra and D. M. Wiharta, ―Environmental monitoring as an IoT application in building smart campus of UniversitasUdayana,‖ in Proc. Int. Conf. Smart Green Technol. Elect. Inf. Syst. (ICSGTEIS), Oct. 2016, pp. 85–88.

[4] M. Suganya., Dayana R and Revathi.R, Crop Yield Prediction Using Supervised Learning Techniques, International Journal of Computer Engineering and Technology, 11(2), 2020, pp. 9-20

[5] CH. Vishnu Vardhanchowdary, Dr.K.Venkataramana, Tomato Crop Yield Prediction using ID3, March 2018,IJIRT Volume 4 Issue 10 pp,663-62.

[6] R. Sujatha and P. Isakki, A study on crop yield forecasting using classification techniques 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), Kovilpatti, 2016, pp. 1-4.

[7] N. Gandhi, L. J. Armstrong, O. Petkar and A. K. Tripathy, Rice crop yield prediction in India using support vector machines 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), KhonKaen, 2016, pp. 1-5.