CROP RECOMMENDATION SYSTEM

Team By:

S190295,
S190605

CONTENTS:

Introduction

In India,farming is not considered as a business but also has a huge impact on the social life of people which are associated to it. There are many festivals and social gatherings are celebrated in accordance with the different seasons and practices involved in farming. Hence large part of population is dependent on agriculture field directly or indirectly.Most of the times it is observed that farmers tend to saw the crop according to its market value and possible financial profits rather than taking factors like soil conditions,weather conditions etc., into the account.This may lead to undesirable results for farmers and for the nature of soil too.

In today’s time,technologies like machine learning and deep learning can become the game changers in such fields if they are used in proper manner.This project aim is to recommend the most suitable crop based on input parameters like Nitrogen(N),Phosphorus(P),Potassium(K),pHvalue of the soil,Humidity,Temperature and Rainfall using various supervised machine learning approches and recommends the most suitable crop.

GOAL OF THE PROJECT

Agriculture plays a crucial role in Indian economy and employment. Half of the country's population is still employed on agriculture sector. India is one of the largest producers of agricultural products, but still it has less farm production. The most common problem faced by the farmers is that they do not opt the crop based on the necessity of soil. This problem can be solved by using crop recommendation system.A crop recommendation system project aim is to recommend the most suitable crop based on input parameters like Nitrogen(N),Phosphorus(P),Potassium(K),pH value of the soil,Humidity,Temperature and Rainfall.This project predicts the accuracy of the future prodution of nine different crops such as pomegranate,banana,mango,grapes,watermelon,muskmelon,apple,orange,papaya crops using various supervised machine learning approches and recommends the most suitable crop.

DATASET

- *The dataset consists of parameters like Nitrogen(N), Phosphorus(P), Potassium(K),pH value of the soil,Humidity, Temperature and Rainfall. The dataset have been obtained from the Kaggle website.*
- *The dataset has 2708 instance or data that have taken from the past historic data. This dataset includes nine different cropssuch as pomegranate,banana,mango,grapes,watermelon,muskmelon,apple,orange and papaya.*

FEATURES OF DATASET

- N: ratio of Nitrogen content in soil
- P: ratio of Phosphorus content in soil
- K: ratio of Potassium content in soil
- Temperature: temperature in degree celsius
- Humidity: relative humidity
- pH: pH value of the soil
- rainfall: rainfall in mm

PREDICTION TECHNIQUES

K-Nearest Neighbour:

K-Nearest Neighbour is one of the simplest machine learning algorithms based on supervised learning technique.K-NN algorithm assumes the similarity between the new case and available cases and puts the new case into category that is most similar to available categories.K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems.It is also called as a lazy learner algorithm because it does not learn from the training set immediately instead it stores the dataset and at the time of classification,it performs an action on the dataset.

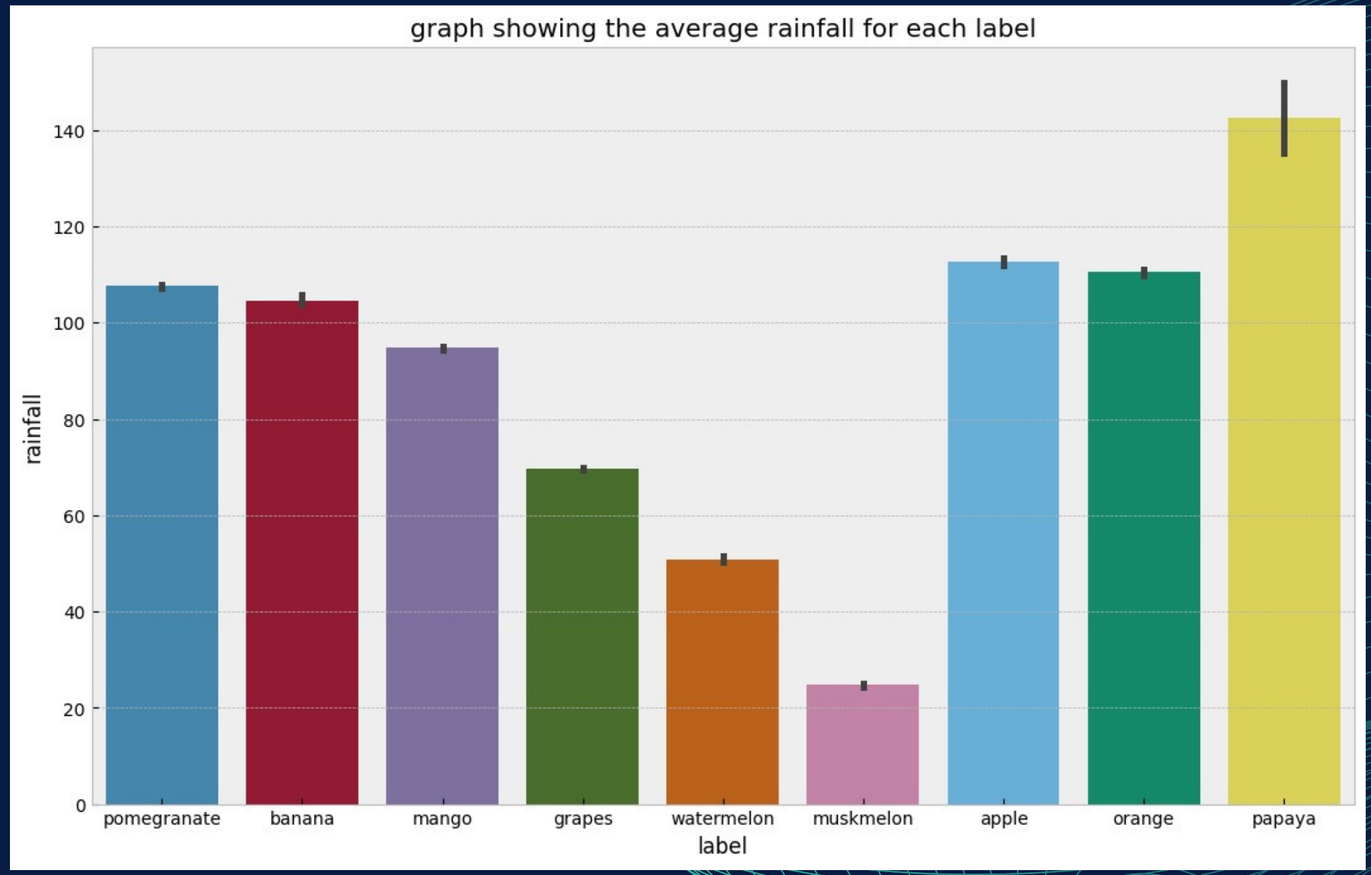
Support Vector Machines:

Support Vector Machine(SVM) is one of the most popular Supervised Learning algorithms,which is used for Classification as well as Regression problems.It is used mostly for Classification problems in Machine Learning.The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future

VISUALISATION

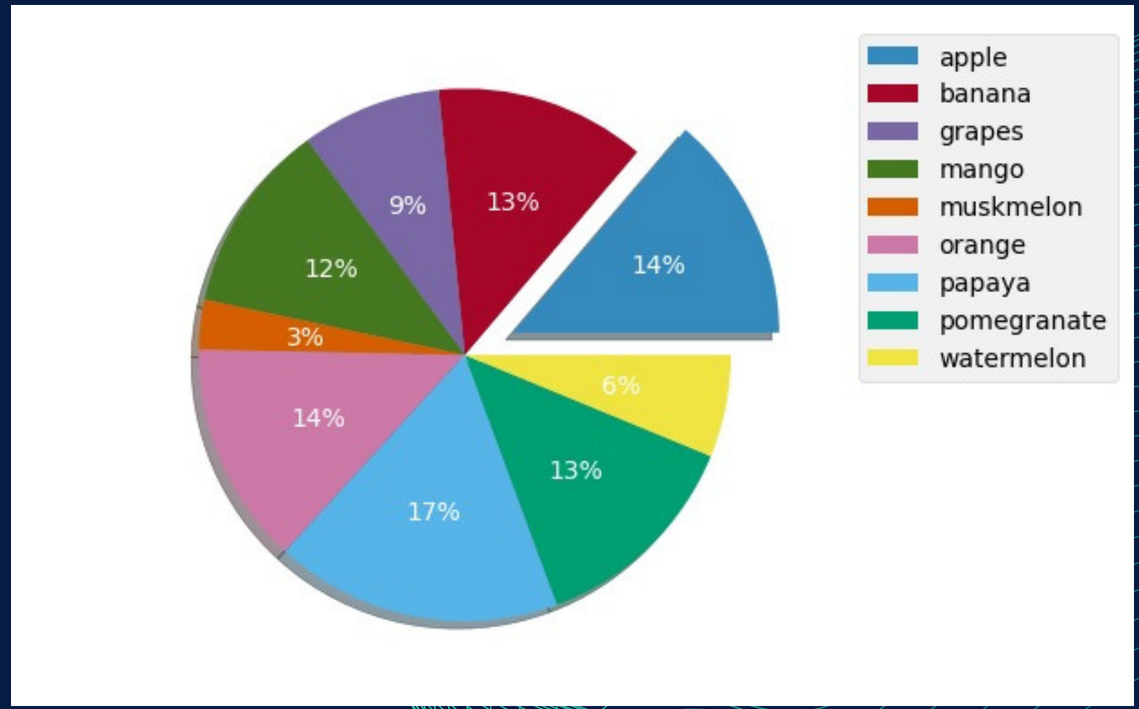
- Bar Graph represehts the average rainfall of each crop and make decisions based on the information.Farmers understand which crop is more productive.

BAR GRAPH



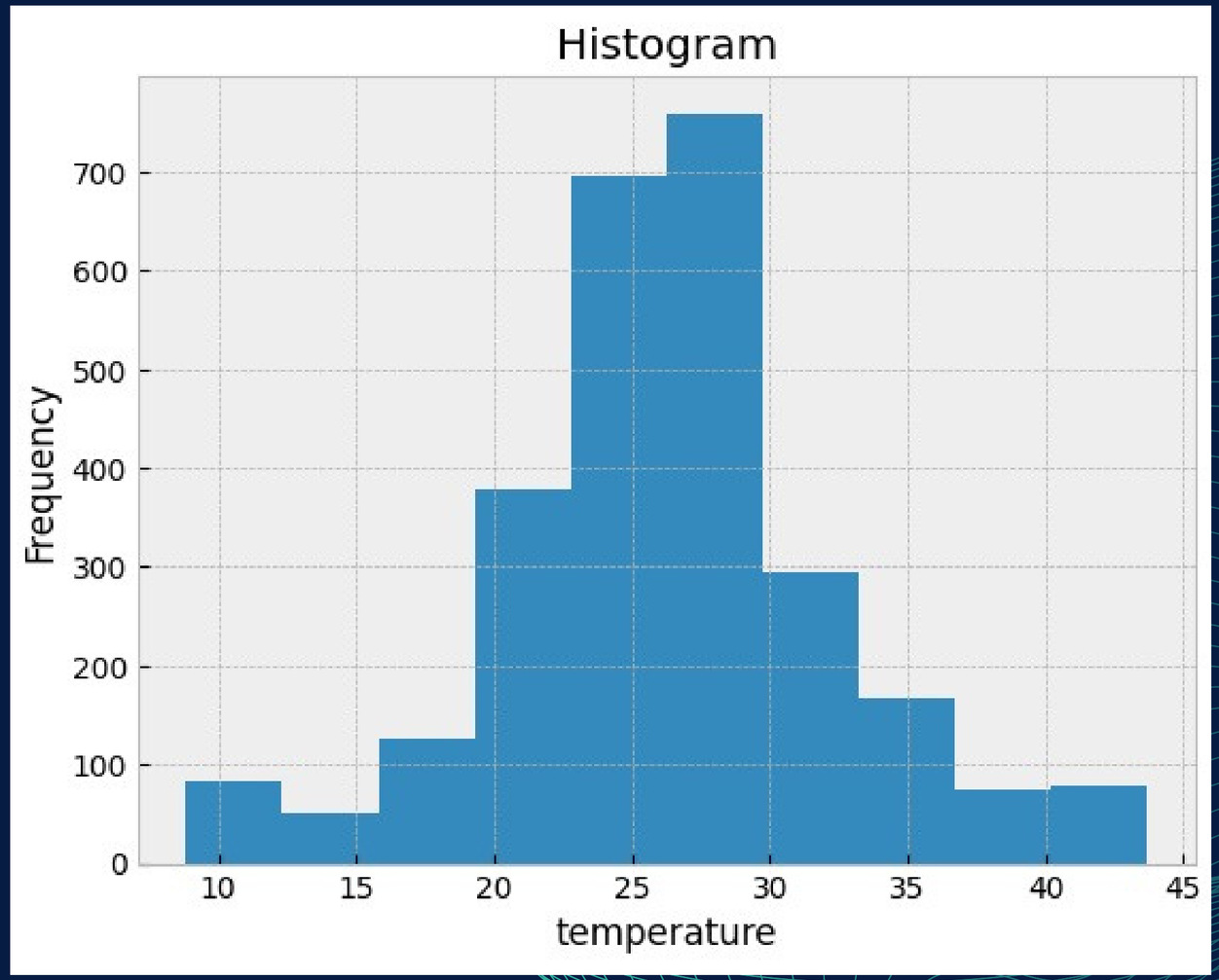
- The pie chart can show the percentage or proportion of different crops grown in area.Each crop is represented by slice and size of the slice indicates how much land is dedicated to that crop.Farmers can understand which crops are most commenly grown in a particular region

PIE CHART



HISTOGRAM

- Histogram display how temperatures are distriburted in particular area.It shows different temperature ranges on the horizontal axes such as cool,moderate,and hot.This helps farmers understand the frequency of different temperature levels in their region.



PREPROCESSING

Importing Necessary Libraries:

```
import pandas as pd
import matplotlib.pyplot as plt
import sklearn
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
from sklearn.svm import SVC
import seaborn as sns
```

Data Understanding:

```
df=pd.read_csv("cropdata.csv")
df
```

	N	P	K	temperature	humidity	ph	rainfall	label
0	NaN	NaN	NaN	34.567890	57.402477	NaN	NaN	pomegranate
1	NaN	25.445813	NaN	NaN	57.402477	NaN	NaN	pomegranate
2	NaN	NaN	6.497367	NaN	57.402477	NaN	NaN	pomegranate
3	NaN	NaN	NaN	NaN	57.402477	NaN	109.416919	pomegranate
4	2.0	24.000000	38.000000	24.559816	91.635362	5.922936	111.968462	pomegranate
...	...	...	...	...	...	...	...	...
2703	42.0	59.000000	55.000000	40.102077	94.351102	6.979102	149.119999	papaya
2704	43.0	64.000000	47.000000	38.589545	91.580765	6.825665	102.270823	papaya
2705	35.0	67.000000	49.000000	41.313301	91.150880	6.617067	239.742755	papaya
2706	56.0	59.000000	55.000000	37.035519	91.794302	6.551893	188.518142	papaya
2707	39.0	64.000000	53.000000	23.012402	91.073555	6.598860	208.335798	papaya

2708 rows × 8 columns

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2708 entries, 0 to 2707
Data columns (total 8 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   N            2700 non-null   float64
 1   P            2702 non-null   float64
 2   K            2702 non-null   float64
 3   temperature  2702 non-null   float64
 4   humidity     2708 non-null   float64
 5   ph           2700 non-null   float64
 6   rainfall     2702 non-null   float64
 7   label        2708 non-null   object 
dtypes: float64(7), object(1)
memory usage: 169.4+ KB
```

df.head()

It returns first five values

:

	N	P	K	temperature	humidity	ph	rainfall	label
0	NaN	NaN	NaN	34.567890	57.402477	NaN	NaN	pomegranate
1	NaN	25.445813	NaN	NaN	57.402477	NaN	NaN	pomegranate
2	NaN	NaN	6.497367	NaN	57.402477	NaN	NaN	pomegranate
3	NaN	NaN	NaN	NaN	57.402477	NaN	109.416919	pomegranate
4	2.0	24.000000	38.000000	24.559816	91.635362	5.922936	111.968462	pomegranate

Replacing the null values:

```
x=df["N"].mean()
df["N"].fillna(x,inplace=True)
x=df["P"].mean()
df["P"].fillna(x,inplace=True)
x=df["K"].mean()
df["K"].fillna(x,inplace=True)
x=df["temperature"].mean()
df["temperature"].fillna(x,inplace=True)
x=df["humidity"].mean()
df["humidity"].fillna(x,inplace=True)
x=df["ph"].mean()
df["ph"].fillna(x,inplace=True)
x=df["rainfall"].mean()
df["rainfall"].fillna(x,inplace=True)
print(df.tostring())
```

	N	P	K	temperature	humidity	ph	rainfall	label
0	50.261111	56.089523	75.563285	34.567890	57.402477	6.305336	90.869908	pomegranate
1	50.261111	25.445813	75.563285	26.411413	57.402477	6.305336	90.869908	pomegranate
2	50.261111	56.089523	6.497367	26.411413	57.402477	6.305336	90.869908	pomegranate
3	50.261111	56.089523	75.563285	26.411413	57.402477	6.305336	109.416919	pomegranate
4	2.000000	24.000000	38.000000	24.559816	91.635362	5.922936	111.968462	pomegranate
...	...	...	...	...	...	...	...	...
2703	42.000000	59.000000	55.000000	40.102077	94.351102	6.979102	149.119999	papaya
2704	43.000000	64.000000	47.000000	38.589545	91.580765	6.825665	102.270823	papaya
2705	35.000000	67.000000	49.000000	41.313301	91.150880	6.617067	239.742755	papaya
2706	56.000000	59.000000	55.000000	37.035519	91.794302	6.551893	188.518142	papaya
2707	39.000000	64.000000	53.000000	23.012402	91.073555	6.598860	208.335798	papaya

2708 rows × 8 columns

```
x=df["label"].nunique()
print("No.ofcrops:",x)
```

Output:

No.of crops: 9

```
print("Crop names :")
for crop in df["label"].unique():
    print(crop)
```

```
Crop names:
pomegranate
banana
mango
grapes
watermelon
muskmelon
apple
orange
papaya
```

Evolution measures:

Using Support Vector Machine(SVM) Algorithm:

```
X = data . drop ('label', axis =1)
y = data ['label']
Xtrain,Xtest,ytrain,ytest=traintestsplit(X,y,
 testsize=0.3 ,randomstate=42)
svm = SVC( )
svm . fit(Xtrain,ytrain)
 ypred=svm . predict(Xtest)
accuracy = accuracy score ( y test , y pred )
print ( " Accuracy : ", accuracy )
```

Output:

Accuracy: 0.997399753997754

Using K-Nearest Neighbour(KNN) Algorithm:

```
X = data . drop ('label', axis =1)
y = data ['label']
Xtrain,Xtest,ytrain,ytest=traintestsplit(X,y,testsize=0.3,
randomstate=42)
k = 19
knn = KNeighborsClassifier(nneighbors=k)
knn.fit(Xtrain,ytrain)
ypred=knn.predict(Xtest)
accuracy = accuracy score ( y test , y pred )
print ("Accuracy :", accuracy)
```

Output:

Accuracy: 0.997539975399754

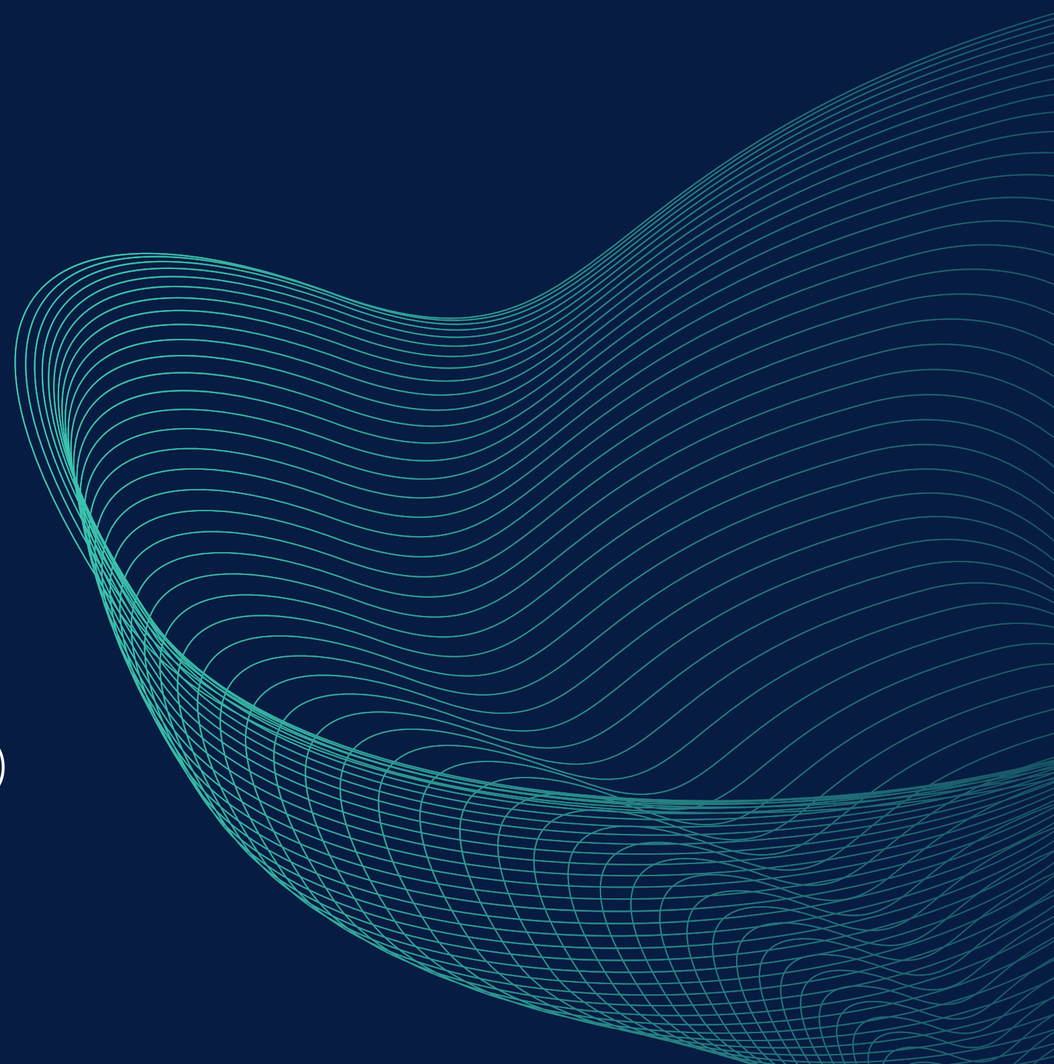
PREDICTIONS BY THE MODEL:

Using KNN Algorithm:

```
new data = {
'N': 50,
'P': 30,
'K': 20,
'temperature': 28,
'humidity': 70,
'ph': 6.5,
23'rainfall': 100
}
new input = pd.DataFrame(new data, index=[0])
predictedlabel= knn.predict(new input)
print("Predictedlabel:",predictedlabel[0])
```

Output:

Predicted label: mango



Using SVM Algorithm:

```
new data = {
'N': 50,
'P': 30,
'K': 20,
'temperature': 28,
'humidity': 70,
'ph': 6.5,
'rainfall': 100
}
new input = pd.DataFrame(new data, index=[0])
predictedlabel = svm.predict(new input)
print("Predictedlabel:",predictedlabel[0])
```

Output:

Predicted label: mango

CONCLUSION

The new data includes nitrogen(N)=75,phosphorus(P)=69,potassium(K)=47, temperature=25, humidity=80,pH=6.3 and rainfall=105,and the system predicts the label to be ”banana”.This prediction suggests that these parameter values check with the characteristics given the growth and cultivation of bananas.

In crop recommendation project,we built a system that can predict the type of crop based on certain parameters such as amount of nitrogen,phosphorus,potassium,temperature, humidity,pH,and rainfall.We used two different machine learning algorithms namely like k-Nearest Neighbors(KNN) and Support Vector Machines(SVM), to make these predictions.Our system performed well,by achieving an accuracy of 99.75%. This means the syatem correctly predict the type of crop in almost all cases. To check the performance of our models,we also created confusion matrix.

In conclusion,our project has successfully developed a crop prediction system using machine learning algorithms.By considering the given parameters,our system predicts the type of crop accurately.It helps farmers in making better decisions about crop selection.


Thank you!