KG REDDY
College of Engineering
& Technology
AN AUTONOMOUS INSTITUTION

Crop analysis and pesticides Recommendation system

Mr. Devdas - Guide

MINI PROJECT GROUP 7

01 **B. SAI BHARGAV**

20QM1A6608
3RD CSM
KGRCET

02 **CH. LIKITHA**

20QM1A6609
3RD CSM
KGRCET

03 **P. VENILLA**

20QM1A6639
3RD CSM
KGRCET

04 **SHAHNAWAAZ**

20QM1A6630
3RD CSM
KGRCET

DATASETS :

1) Crop Prediction :

Crop Prediction Data, CSV file format

Dist Code	Year	State Code	State Nam	Dist Name	RICE AREA	RICE PROD	RICE YIELD	WHEAT AR	WHEAT PR	WHEAT YII	KHARIF SO	KHARIF SO	KHARIF SO	RABI SORG	RABI SORG	RABI SORG	SORGHUM	SORGHUM	SORGHUM	PEARL MIL	PEARL MIL	PEARL MIL
1	1966	14	Chhattisga	Durg	548	185	337.59	44	20	454.55	0.6	0.4	666.67	0	0	0	0.6	0.4	666.67	0	0	0
1	1967	14	Chhattisga	Durg	547	409	747.71	50	26	520	1.1	0.9	818.18	0	0	0	1.1	0.9	818.18	0	0	0
1	1968	14	Chhattisga	Durg	556.3	468	841.27	53.7	30	558.66	0.5	0.4	800	0	0	0	0.5	0.4	800	0	0	0
1	1969	14	Chhattisga	Durg	563.4	400.8	711.4	49.4	26.5	536.44	0.8	0.6	750	0	0	0	0.8	0.6	750	0	0	0
1	1970	14	Chhattisga	Durg	571.6	473.6	828.55	44.2	29	656.11	0.9	0.6	666.67	0	0	0	0.9	0.6	666.67	0	0	0
1	1971	14	Chhattisga	Durg	581.8	412.9	709.69	44.4	25.8	581.08	0.3	0.2	666.67	0	0	0	0.3	0.2	666.67	0	0	0
1	1972	14	Chhattisga	Durg	582.2	381	654.41	39.6	20.6	520.2	0.3	0.3	1000	0	0	0	0.3	0.3	1000	0	0	0
1	1973	14	Chhattisga	Durg	600	471.9	786.5	37.3	18.6	498.66	0.2	0.2	1000	0	0	0	0.2	0.2	1000	0	0	0
1	1974	14	Chhattisga	Durg	587.4	219	372.83	36.5	22.4	613.7	0.5	0.4	800	0	0	0	0.5	0.4	800	0	0	0
1	1975	14	Chhattisga	Durg	598.3	454	758.82	49.2	27.8	565.04	0.2	0.2	1000	0	0	0	0.2	0.2	1000	0	0	0
1	1976	14	Chhattisga	Durg	593.6	327.1	551.04	46.9	10	213.22	0.1	0.1	1000	0	0	0	0.1	0.1	1000	0	0	0
1	1977	14	Chhattisga	Durg	600.7	572.4	952.89	53.1	27.1	510.36	0.1	0.1	1000	0	0	0	0.1	0.1	1000	0	0	0
1	1978	14	Chhattisga	Durg	612.5	362.2	591.35	48.7	25.6	525.67	0.1	0.1	1000	0	0	0	0.1	0.1	1000	0	0	0
1	1979	14	Chhattisga	Durg	616.8	330.6	535.99	44.6	17.8	399.1	0.5	0.5	1000	0	0	0	0.5	0.5	1000	0	0	0
1	1980	14	Chhattisga	Durg	634.9	515.6	812.1	44.1	33.6	761.9	0.2	0.2	1000	0	0	0	0.2	0.2	1000	0	0	0
1	1981	14	Chhattisga	Durg	630	506.9	804.6	41.5	23.6	568.67	0.2	0.2	1000	0	0	0	0.2	0.2	1000	0	0	0
1	1982	14	Chhattisga	Durg	627.9	513.3	817.49	41.1	23.9	581.51	0.1	0.1	1000	0	0	0	0.1	0.1	1000	0	0	0
1	1983	14	Chhattisga	Durg	626.7	711	1134.51	39.9	20.6	516.29	0.1	0.1	1000	0	0	0	0.1	0.1	1000	0	0	0
1	1984	14	Chhattisga	Durg	632.2	563.8	891.81	40.5	19.9	491.36	0.3	0.3	1000	0	0	0	0.3	0.3	1000	0	0	0
1	1985	14	Chhattisga	Durg	630.8	699.8	1109.38	39.4	21	532.99	0.3	0.2	666.67	0	0	0	0.3	0.2	666.67	0	0	0
1	1986	14	Chhattisga	Durg	643	525	816.49	37	24	648.65	0.2	0.1	500	0	0	0	0.2	0.1	500	0	0	0
1	1987	14	Chhattisga	Durg	648	523	807.1	43	23	534.88	0.3	0.2	666.67	0	0	0	0.3	0.2	666.67	0	0	0
1	1988	14	Chhattisga	Durg	652.7	549.7	842.19	43.7	20.2	462.24	0.1	0.1	1000	0	0	0	0.1	0.1	1000	0	0	0
1	1989	14	Chhattisga	Durg	660.2	457.3	692.67	43.8	22.7	518.26	0.1	0.1	1000	0	0	0	0.1	0.1	1000	0	0	0
1	1990	14	Chhattisga	Durg	682.9	806.8	1181.43	36.2	24.8	685.08	0.5	0.4	800	0	0	0	0.5	0.4	800	0	0	0
1	1991	14	Chhattisga	Durg	680.8	773.6	1136.31	34.8	21.6	620.69	2.3	1.8	782.61	0	0	0	2.3	1.8	782.61	0	0	0

Crop Prediction Data

DATASETS :

2) Crop Analysis:

Crop Analysis Data, CSV file format

	A	B	C	D	E	F	G	H
1	N	P	K	temperature	humidity	ph	rainfall	label
2	90	42	43	20.87974371	82.00274423	6.502985292	202.9355362	rice
3	85	58	41	21.77046169	80.31964408	7.038096361	226.6555374	rice
4	60	55	44	23.00445915	82.3207629	7.840207144	263.9642476	rice
5	74	35	40	26.49109635	80.15836264	6.980400905	242.8640342	rice
6	78	42	42	20.13017482	81.60487287	7.628472891	262.7173405	rice
7	69	37	42	23.05804872	83.37011772	7.073453503	251.0549998	rice
8	69	55	38	22.70883798	82.63941394	5.70080568	271.3248604	rice
9	94	53	40	20.27774362	82.89408619	5.718627178	241.9741949	rice
10	89	54	38	24.51588066	83.5352163	6.685346424	230.4462359	rice
11	68	58	38	23.22397386	83.03322691	6.336253525	221.2091958	rice
12	91	53	40	26.52723513	81.41753846	5.386167788	264.6148697	rice
13	90	46	42	23.97898217	81.45061596	7.50283396	250.0832336	rice
14	78	58	44	26.80079604	80.88684822	5.108681786	284.4364567	rice
15	93	56	36	24.01497622	82.05687182	6.98435366	185.2773389	rice
16	94	50	37	25.66585205	80.66385045	6.94801983	209.5869708	rice
17	60	48	39	24.28209415	80.30025587	7.042299069	231.0863347	rice
18	85	38	41	21.58711777	82.7883708	6.249050656	276.6552459	rice
19	91	35	39	23.79391957	80.41817957	6.970859754	206.2611855	rice
20	77	38	36	21.8652524	80.1923008	5.953933276	224.5550169	rice
21	88	35	40	23.57943626	83.58760316	5.85393208	291.2986618	rice
22	89	45	36	21.32504158	80.47476396	6.442475375	185.4974732	rice
23	76	40	43	25.15745531	83.11713476	5.070175667	231.3843163	rice
24	67	59	41	21.94766735	80.97384195	6.012632591	213.3560921	rice
25	83	41	43	21.0525355	82.67839517	6.254028451	233.1075816	rice
26	98	47	37	23.48381344	81.33265073	7.375482851	224.0581164	rice
27	66	53	41	25.0756354	80.52389148	7.778915154	257.0038865	rice

DATASETS :

3) Crop Clustering :

In crop clustering, we use data that we have extracted from the crop analysis.

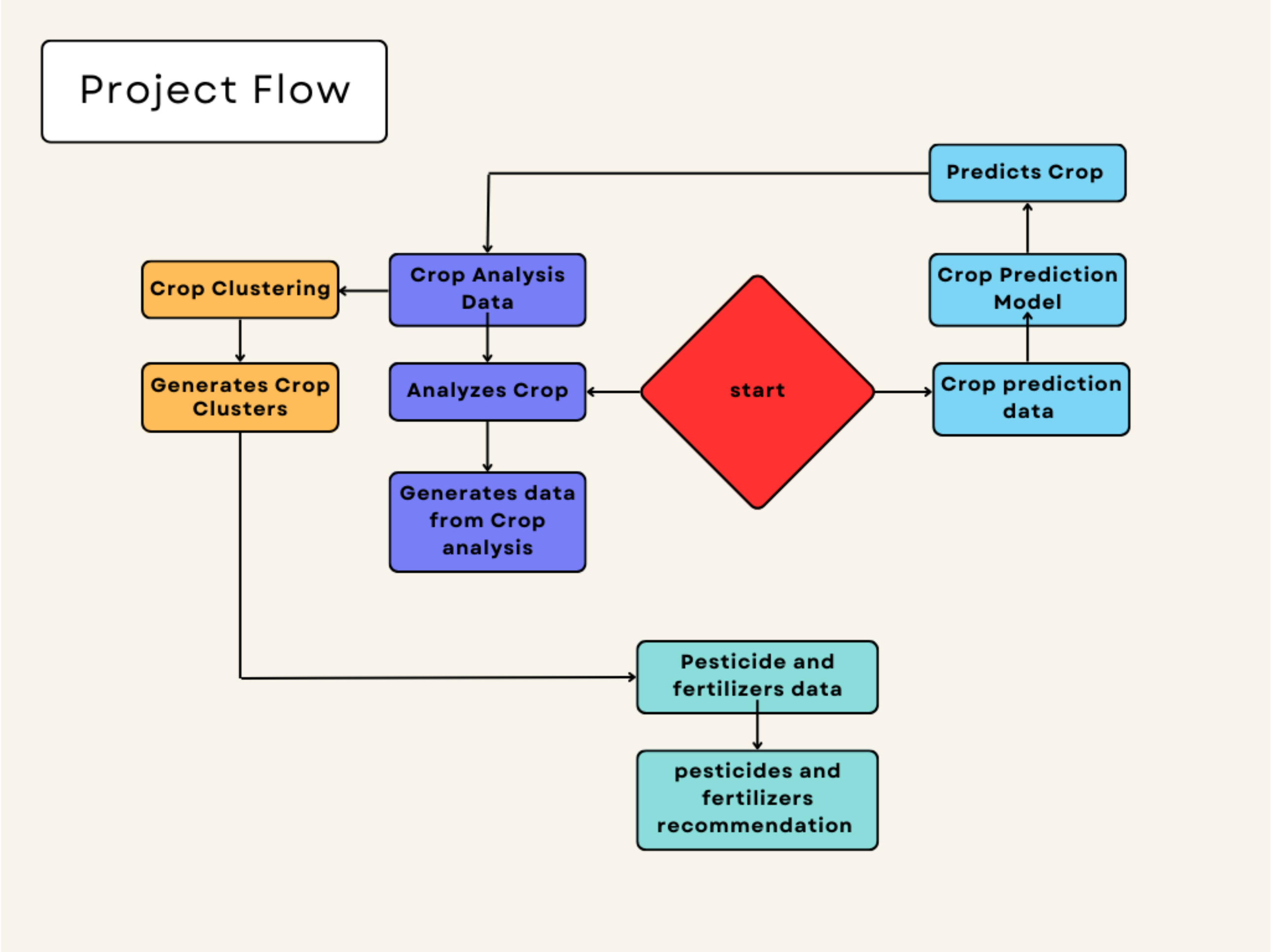
DATASETS :

4) Pesticides and Fertilizers Recommendation:

P&F Data, CSV file format

1	Entity	Code	Year	Nutrient
2	Afghanistan	AFG	1961	0.13
3	Afghanistan	AFG	1962	0.13
4	Afghanistan	AFG	1963	0.13
5	Afghanistan	AFG	1964	0.13
6	Afghanistan	AFG	1965	0.13
7	Afghanistan	AFG	1966	0.13
8	Afghanistan	AFG	1967	1.13
9	Afghanistan	AFG	1968	1.75
10	Afghanistan	AFG	1969	1.88
11	Afghanistan	AFG	1970	2
12	Afghanistan	AFG	1971	1.84
13	Afghanistan	AFG	1972	1.99
14	Afghanistan	AFG	1973	2.2
15	Afghanistan	AFG	1974	3.01
16	Afghanistan	AFG	1975	3.44
17	Afghanistan	AFG	1976	3.81
18	Afghanistan	AFG	1977	4.6
19	Afghanistan	AFG	1978	4.41
20	Afghanistan	AFG	1979	4.13

1	Entity	Code	Year	Pesticides (to
2	Africa (FAO)		1990	65944.26
3	Africa (FAO)		1991	62723.2
4	Africa (FAO)		1992	54428.07
5	Africa (FAO)		1993	49620.53
6	Africa (FAO)		1994	51095.37
7	Africa (FAO)		1995	56939.78
8	Africa (FAO)		1996	58563.56
9	Africa (FAO)		1997	59473.66
10	Africa (FAO)		1998	60770.25
11	Africa (FAO)		1999	63027.14
12	Africa (FAO)		2000	63894.15
13	Africa (FAO)		2001	64523.1
14	Africa (FAO)		2002	67516.16
15	Africa (FAO)		2003	69697.31
16	Africa (FAO)		2004	73268.59
17	Africa (FAO)		2005	71312.94
18	Africa (FAO)		2006	77401.08
19	Africa (FAO)		2007	75657.45
20	Africa (FAO)		2008	79274.89


Project Flow
Predicts Crop
Crop Clustering
Crop Analysis Data
Crop Prediction Model
Generates Crop Clusters
Analyzes Crop
start
Crop prediction data
Generates data from Crop analysis
Pesticide and fertilizers data
pesticides and fertilizers recommendation

MODULES:

1) Crop Prediction :

1) Decision Tree

• **Step-1:** Begin the tree with the root node, says S, which contains the complete dataset.

•**Step-2:** Find the best attribute in the dataset using **Attribute Selection Measure (ASM).**

•**Step-3:** Divide the S into subsets that contains possible values for the best attributes.

•**Step-4:** Generate the decision tree node, which contains the best attribute.

•**Step-5:** Recursively make new decision trees using the subsets of the dataset created in step -3. Continue this process until a stage is reached where you cannot further classify the nodes and called the final node as a leaf node.

Example

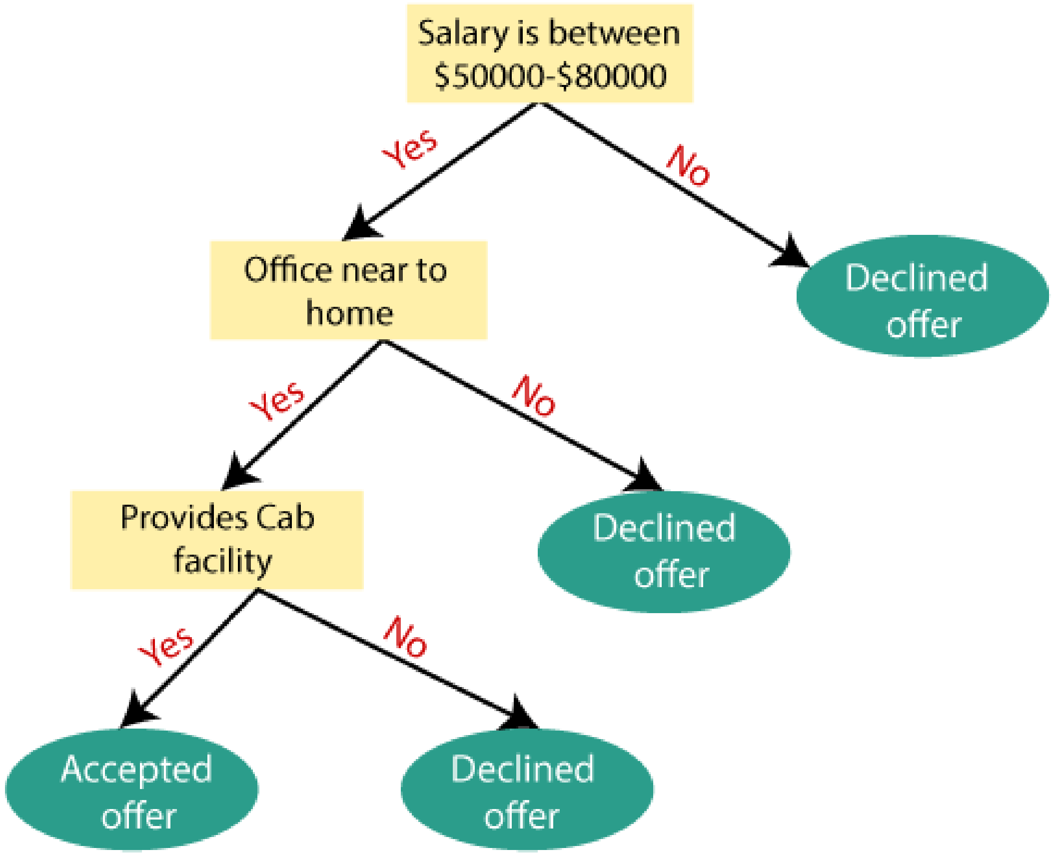
1) Suppose there is a candidate who has a job offer and wants to decide whether he should accept the offer or Not.

2) So, to solve this problem, the decision tree starts with the root node (Salary attribute by ASM).

3) The root node splits further into the next decision node (distance from the office) and one leaf node based on the corresponding labels.

4) The next decision node further gets split into one decision node (Cab facility) and one leaf node.

5) Finally, the decision node splits into two leaf nodes (Accepted offers and Declined offer).


Salary is between $50000-$80000
Yes
No
Office near to home
Declined offer
Yes
No
Provides Cab facility
Declined offer
Yes
No
Accepted offer
Declined offer

MODULES:

1) Crop Prediction :

2) Random Forest :

Step-1: Select random K data points from the training set.

Step-2: Build the decision trees associated with the selected data points (Subsets).

Step-3: Choose the number N for decision trees that you want to build.

Step-4: Repeat Step 1 & 2.

Step-5: For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority votes.

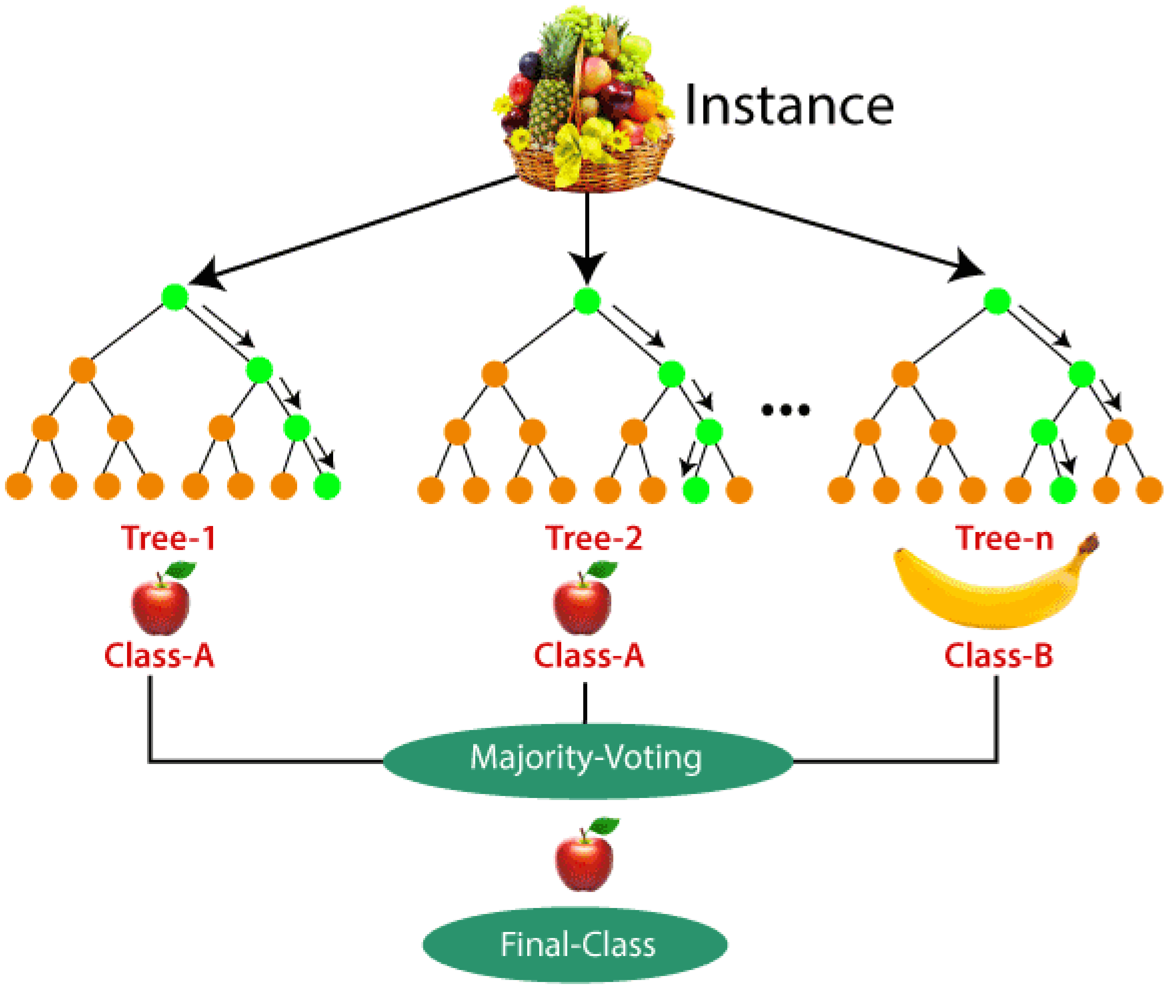
Example:

Suppose there is a dataset that contains multiple fruit images.

So, this dataset is given to the Random forest classifier. The dataset is divided into subsets and given to each decision tree.

During the training phase, each decision tree produces a prediction result

when a new data point occurs, then based on the majority of results, the Random Forest classifier predicts the final decision


Instance
Tree-1
Class-A
Tree-2
Class-A
Tree-n
Class-B
Majority-Voting
Final-Class

MODULES:

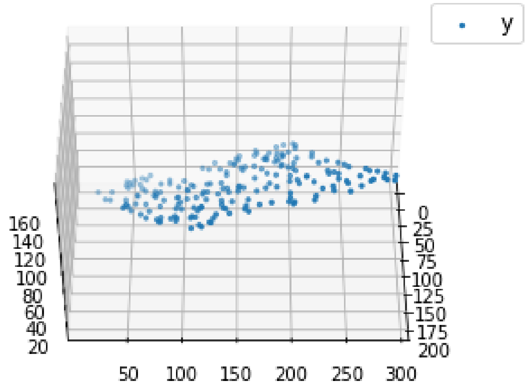
2) Crop Analysis :

1) Multiple Linear Regression :

Step 1: Encoding the Categorical Data.
Step 2: Splitting the Data set into Training Set and Test Set.
Step 3: Fitting Multiple Linear Regression to the Training set
Step 4: Predict the Test set results.

Example:

GENDER	MALE	FEMALE
Male	1	0
Male	1	0
Female	0	1
Female	0	1
Male	1	0
Female	0	1
Male	1	0



MODULES:

2) Crop Analysis :

2)Principal Component Analysis:

Step 1: Getting the dataset
Step 2: Representing data into a structure
Step 3: Standardizing the data
Step 4: Calculating the Covariance of Z
Step 5: Calculating the Eigen Values and Eigen Vectors
Step 6: Sorting the Eigen Vectors
Step 7: Calculating the new features Or Principal Components
Step 8: Remove less or unimportant features from the new dataset.

Example:

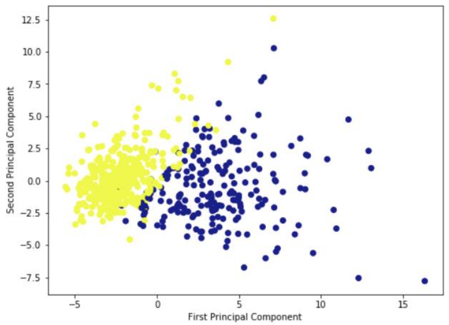
	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	mean fractal dimension	...	worst radius	worst texture	worst perimeter	worst area	worst smoothness
0	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.3001	0.14710	0.2419	0.07871	...	25.38	17.33	184.60	2019.0	0.1622
1	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.0869	0.07017	0.1812	0.05667	...	24.99	23.41	158.80	1956.0	0.1238
2	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.1974	0.12790	0.2069	0.05999	...	23.57	25.53	152.50	1709.0	0.1444
3	11.42	20.38	77.58	386.1	0.14250	0.28390	0.2414	0.10520	0.2597	0.09744	...	14.91	26.50	98.87	567.7	0.2098
4	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.1980	0.10430	0.1809	0.05883	...	22.54	16.67	152.20	1575.0	0.1374

5 rows × 30 columns

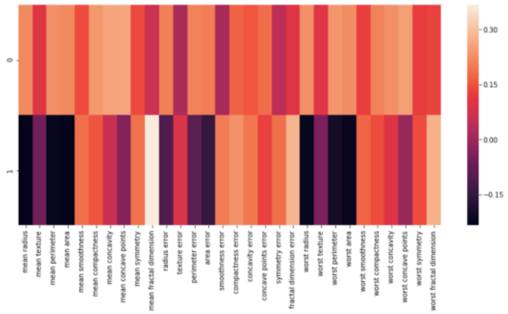
```
pca.components_
```

```
array([[ 0.21890244,  0.10372458,  0.22753729,  0.22099499,  0.14258969,
         0.23928535,  0.25840048,  0.26085376,  0.13816696,  0.06436335,
         0.20597878,  0.01742803,  0.21132592,  0.20286964,  0.01453145,
         0.17039345,  0.15358979,  0.1834174 ,  0.04249842,  0.10256832,
         0.22799663,  0.10446933,  0.23663968,  0.22487053,  0.12795256,
         0.21009588,  0.22876753,  0.25088597,  0.12290456,  0.13178394],
       [-0.23385713, -0.05970609, -0.21518136, -0.23107671,  0.18611302,
         0.15189161,  0.06016536, -0.0347675 ,  0.19034877,  0.36657547,
        -0.10555215,  0.08997968, -0.08945723, -0.15229263,  0.20443045,
         0.2327159 ,  0.19720728,  0.13032156,  0.183848  ,  0.28009203,
        -0.21986638, -0.0454673 , -0.19987843, -0.21935186,  0.17230435,
         0.14359317,  0.09796411, -0.00825724,  0.14188335,  0.27533947]])
```

Text(0,0.5,'Second Principal Component')



<matplotlib.axes._subplots.AxesSubplot at 0x1a1fa171d0>



MODULES:

3) Crop Clustering:

1)K-Means Clustering :

Step-1: Select the number K to decide the number of clusters.

Step-2: Select random K points or centroids. (It can be other from the input dataset).

Step-3: Assign each data point to their closest centroid, which will form the predefined K clusters.

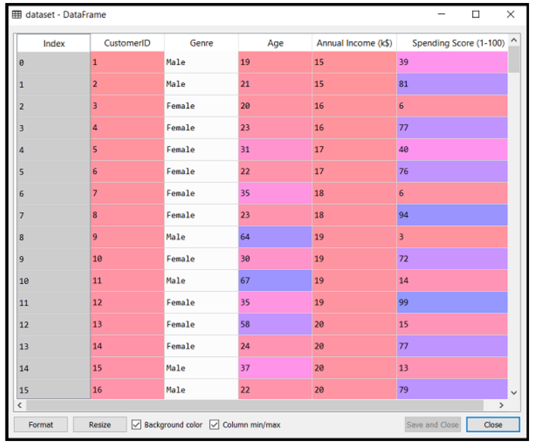
Step-4: Calculate the variance and place a new centroid of each cluster.

Step-5: Repeat the third steps, which means reassign each datapoint to the new closest centroid of each cluster.

Step-6: If any reassignment occurs, then go to step-4 else go to FINISH.

Step-7: The model is ready.

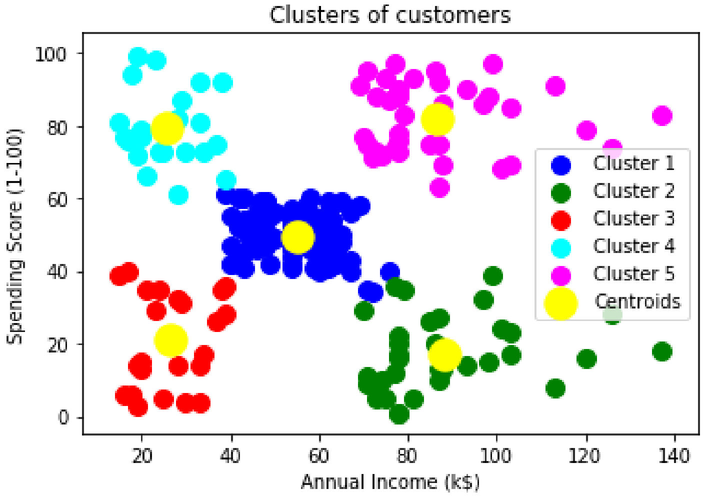
Example:



dataset - DataFrame

Index	CustomerID	Genre	Age	Annual Income (k$)	Spending Score (1-100)
0	1	Male	19	15	39
1	2	Male	21	15	81
2	3	Female	20	16	6
3	4	Female	23	16	77
4	5	Female	31	17	40
5	6	Female	22	17	76
6	7	Female	35	18	6
7	8	Female	23	18	94
8	9	Male	64	19	3
9	10	Female	30	19	72
10	11	Male	67	19	14
11	12	Female	35	19	99
12	13	Female	58	20	15
13	14	Female	24	20	77
14	15	Male	37	20	13
15	16	Male	22	20	79

Format Resize Background color Column min/max Save and Close Close



MODULES:

3) Crop Clustering:

2) Fuzzy Clustering :

Step 1: Initialize the data points into the desired number of clusters randomly.
Step 2: Find out the centroid.
Step 3: Find out the distance of each point from the centroid.
Step 4: Updating membership values.
Step 5: Repeat the steps(2-4) until the constant values are obtained for the membership values or the difference is less than the tolerance value
Step 6: Defuzzify the obtained membership values.

Example:

Cluster	(1, 3)	(2, 5)	(4, 8)	(7, 9)
1)	0.8	0.7	0.2	0.1
2)	0.2	0.3	0.8	0.9

Cluster Centers:
[[0.42363557 0.68304616]
[0.52768166 0.38180987]
[0.39967863 0.31042639]]

Cluster Membership:
[1 0 1 2 1 0 0 2 0 2 1 0 0 0 1 2 1 1 0 2 1 0 0 2 1 1 2 0 1 0
0 2 0 2 1 1 0
1 0 2 0 0 2 0 0 1 1 1 1 2 2 0 1 1 0 1 0 0 2 2 2 0 1 1 2 0 0
0 2 1 0 1 0 0
1 2 0 0 2 2 1 1 0 2 1 0 2 2 1 1 0 2 1 0 2 1 0 2 2 2 0 1 0 2]

MODULES:

4) Pesticide and Fertilizer recommendation

1) Collabarative Based filtering:

Step 1: Finding the similarity of users to the target user U.

Step 2: Prediction of missing rating of an item

Example: Consider a matrix that shows four users Alice, U1, U2 and U3 rating on different news apps. The rating range is from 1 to 5 on the basis of users' likability of the news app.
Step 1: Calculating the similarity between Alice and all the other users At first we calculate the averages of the ratings of all the user excluding I5 as it is not rated by Alice.
Step 2: Predicting the rating of the app not rated by Alice Now, we predict Alice's rating for BBC News App.

Example :

Name	Inshorts(I1)	HT(I2)	NYT(I3)	TOI(I4)	BBC(I5)
Alice	5	4	1	4	?
U1	3	1	2	3	3
U2	4	3	4	3	5
U3	3	3	1	5	4

Step 1:

Name	Inshorts(I1)	HT(I2)	NYT(I3)	TOI(I4)
Alice	1.5	0.5	-2.5	0.5
U1	0.75	-1.25	-0.25	0.75
U2	0.5	-0.5	0.5	-0.5
U3	0	0	-2	2

Step 2:

[Tex]r_{(Alice,I5)}=3.5 + \frac {(0.301*0.75)+(-0.33*1.5)+(0.707*1)}{|0.301|+|-0.33|+|0.707|}=3.83\newline

MODULES:

4) Pesticide and Fertilizer recommendation

2) Content Based filtering:

Content-based filtering makes recommendations by using keywords and attributes assigned to objects in a database (e.g., items in an online marketplace) and matching them to a user profile. The user profile is created based on data derived from a user's actions, such as purchases, ratings (likes and dislikes), downloads, items searched for on a website and/or placed in a cart, and clicks on product links.

Example: suppose you're recommending accessories to a user that just purchased a smartphone from your website and has previously bought smartphone accessories. Aside from keywords such as the smartphone manufacturer, make, and model, the user profile indicates prior purchases include phone holders with sleeves for credit cards. Based on this information, the recommender system may suggest similar phone holders for the new phone with attributes such as an RFID blocking fabric layer to help prevent unauthorized credit card scanning. In this example, the user would expect recommendations for similar phone holders, but the RFID blocking feature may be something they didn't expect yet appreciate.

KG REDDY
College of Engineering
& Technology
AN AUTONOMOUS INSTITUTION

Thank You