PRODUCT CLUSTER PART:

Fig 1: Word Cloud of type of product in each cluster.

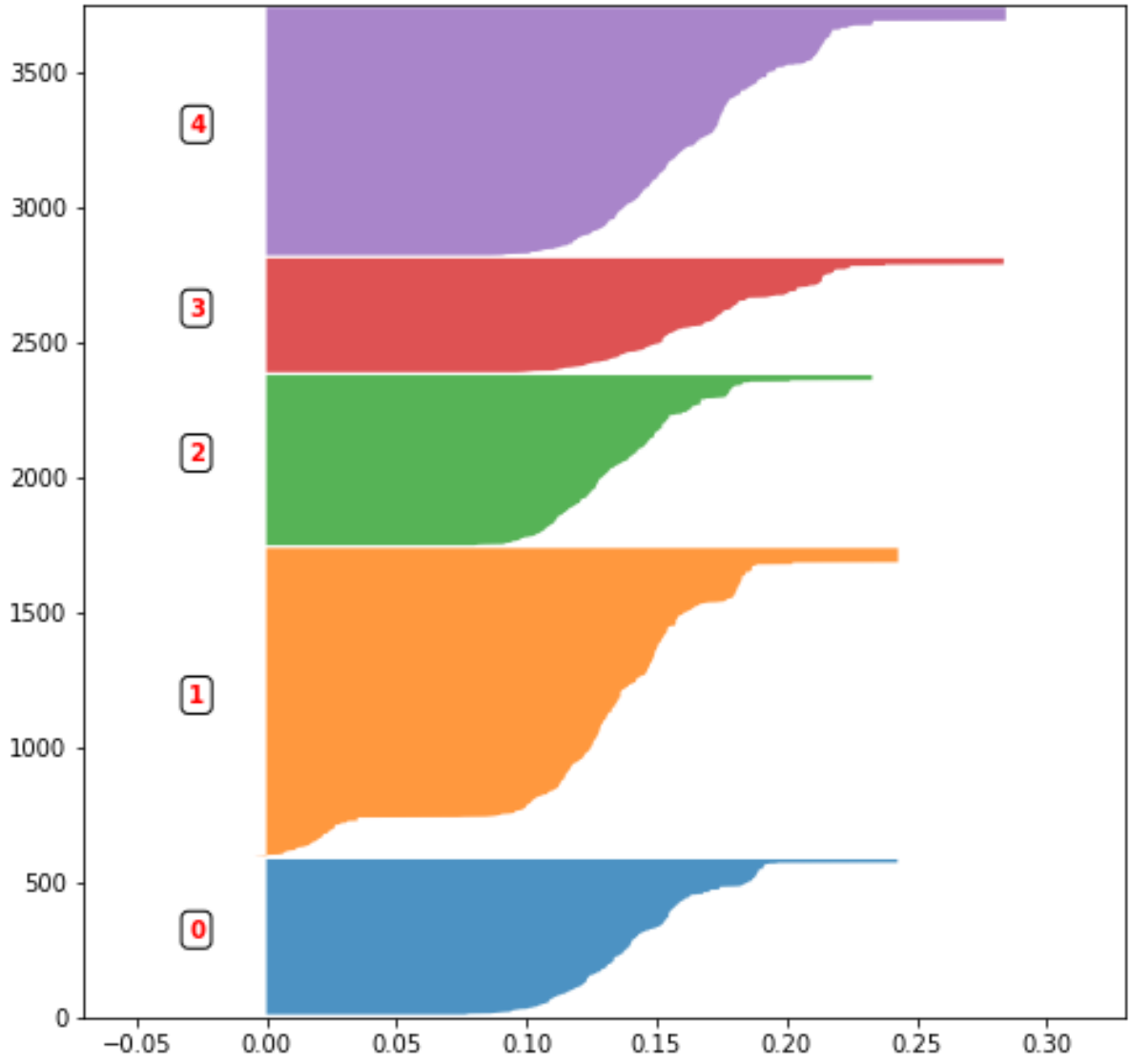


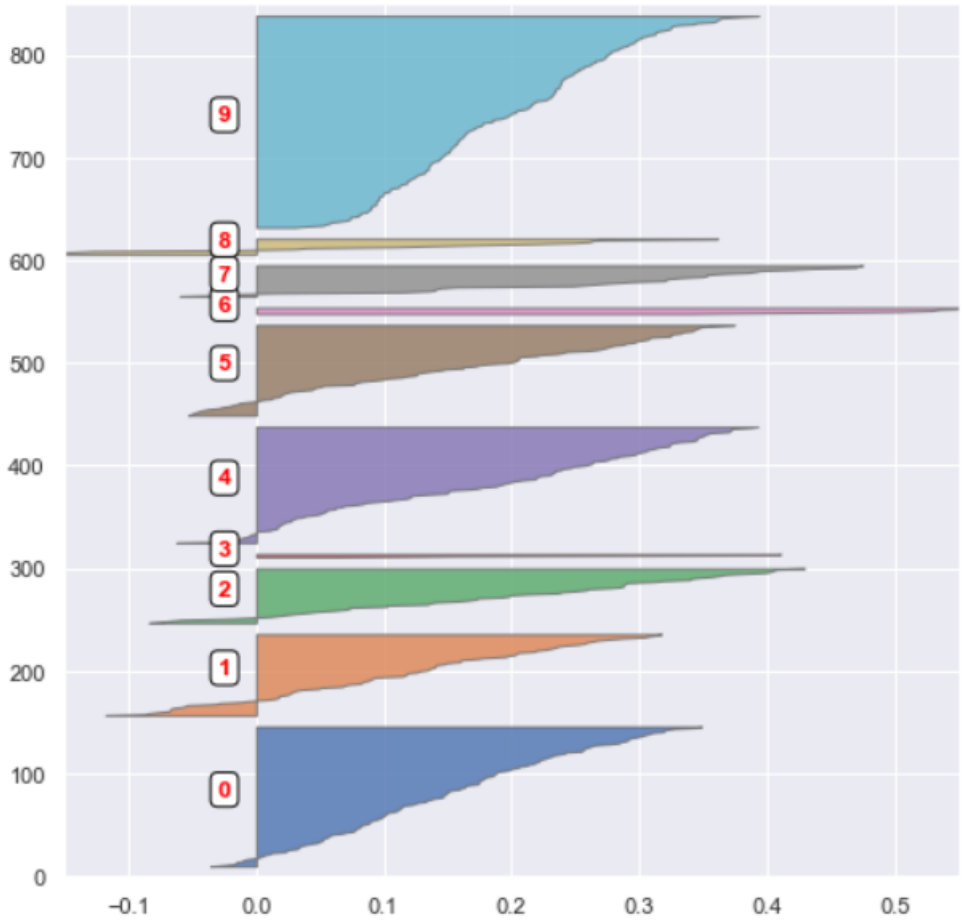
Fig 2: Silhouette scores of each element of the different product clusters.

CUSTOMER CLUSTER PART:-

```
For n_clusters :  3 The average silhouette_score is :  0.16482751977463972
For n_clusters :  4 The average silhouette_score is :  0.1781723969674145
For n_clusters :  5 The average silhouette_score is :  0.1899217010678906
For n_clusters :  6 The average silhouette_score is :  0.18515824366660452
For n_clusters :  7 The average silhouette_score is :  0.18547415126069844
For n_clusters :  8 The average silhouette_score is :  0.18098901076346333
For n_clusters :  9 The average silhouette_score is :  0.18401201900493666
For n_clusters :  10 The average silhouette_score is :  0.16957408453315828
For n_clusters :  11 The average silhouette_score is :  0.15736947638276239
For n_clusters :  12 The average silhouette_score is :  0.16213159944408528
For n_clusters :  13 The average silhouette_score is :  0.15270769812287285
For n_clusters :  14 The average silhouette_score is :  0.15331188085678152
For n_clusters :  15 The average silhouette_score is :  0.14928545363500062
For n_clusters :  16 The average silhouette_score is :  0.1504785666270165
For n_clusters :  17 The average silhouette_score is :  0.15278233438631114
For n_clusters :  18 The average silhouette_score is :  0.15550183447579233
For n_clusters :  19 The average silhouette_score is :  0.15213919311232402
For n_clusters :  20 The average silhouette_score is :  0.15572240813473684
For n_clusters :  10 The average silhouette_score is :  0.16904771993462991
Silhouette Score : 0.171
```

Fig 3: Silhouette score to define the number of clusters.

```
Number of customers :  739
```



Ye dono graph same hai jo achaa lage wo laga lio

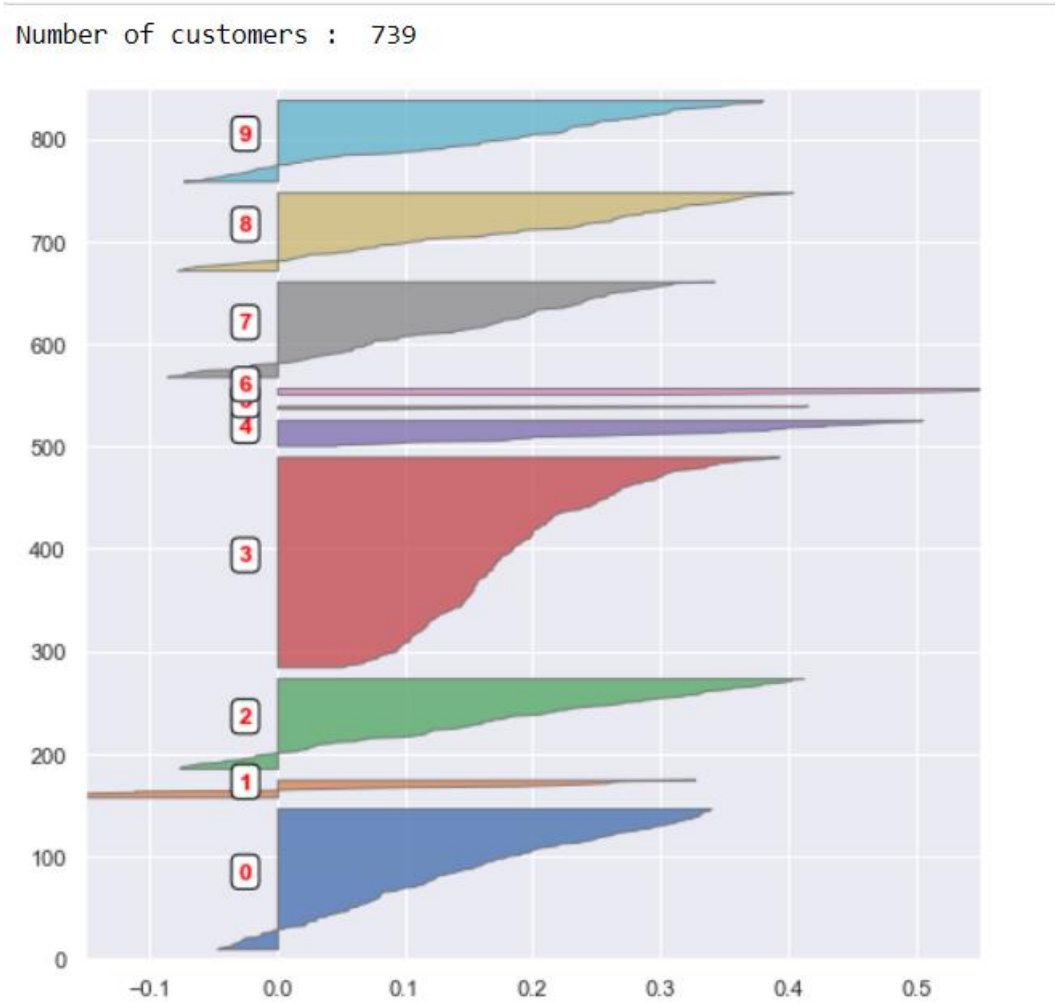


Fig 4: Silhouette scores of each element of the different customer clusters.

7]:

Unnamed: 0	CustomerID	count	min	max	mean	sum	categ_0	categ_1	categ_2	categ_3	categ_4	LastPurchase	FirstPurchase	cluster
0	12347	5	382.52	711.79	558.172	2790.86	32.408290	29.105724	11.173617	18.636191	8.676179	59	297	1
1	12359	3	547.50	1803.11	1153.310	3459.93	15.019090	9.916386	3.985052	44.655528	26.423945	119	261	4
2	12362	5	303.76	829.99	510.908	2554.54	17.343631	34.424985	5.787343	34.123560	8.320480	2	225	0
3	12380	2	607.55	626.01	616.780	1233.56	12.569312	22.325627	7.052758	51.437303	6.615000	8	115	4
4	12381	1	1227.39	1227.39	1227.390	1227.39	10.522328	23.602930	8.455340	43.776632	13.642770	49	49	7

Fig 5: Customers classified in the different client categories.

TRAINING MODELS:-

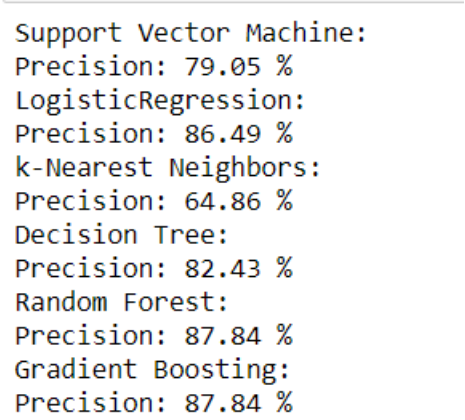


Fig 6: Precision of train data using different classifiers.

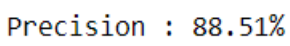


Fig 7: Precision of train data using voting classifier.

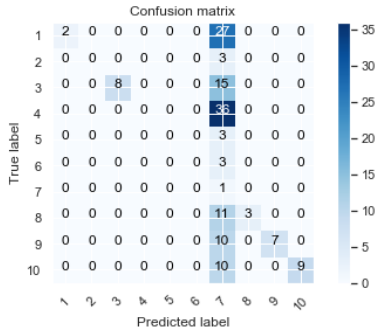


Fig 8(a): SVC

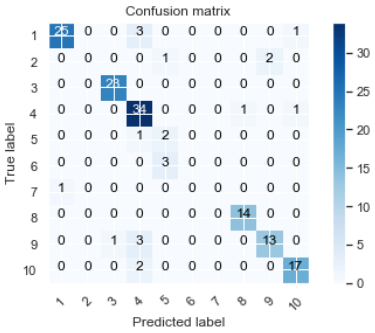


Fig 8(b): Logistic Regression

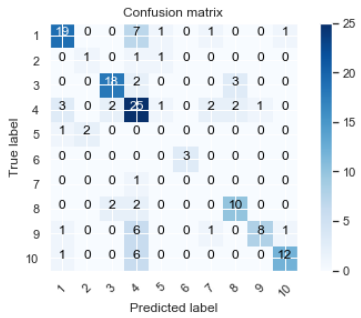


Fig 8(c): KNN

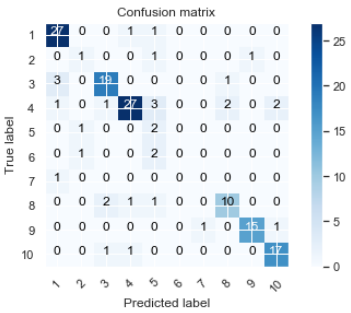


Fig 8(d): Decision Tree

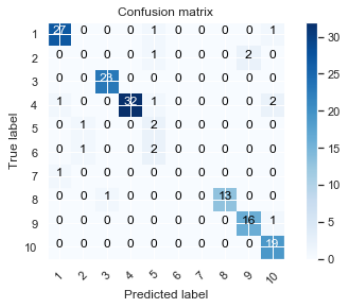


Fig 8(e): Random Forest

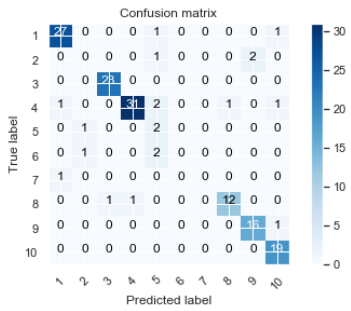


Fig 8(f): Gradient Boosting

Fig 8: Predictions and real values comparisons to the breasts of the different classes

OR

Fig 8: Confusion Matrix of different classifiers.

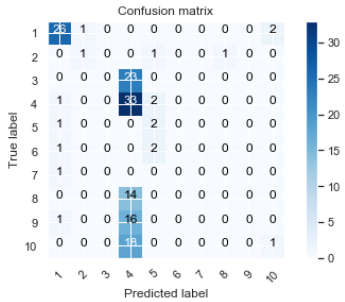
TESTING MODELS:-

```
_____________________________________
Support Vector Machine
Precision: 80.56 %

_____________________________________
Logostic Regression
Precision: 81.48 %

_____________________________________
k-Nearest Neighbors
Precision: 62.43 %

_____________________________________
Decision Tree
Precision: 73.54 %

_____________________________________
Random Forest
Precision: 80.56 %

_____________________________________
Gradient Boosting
Precision: 81.88 %
```

Fig 9: Precision of test data using different classifiers that have been trained.

```
Precision : 82.41%
```

Fig 10: Precision of test data using voting classifier.



Confusion Matrix of adda bost classifier.