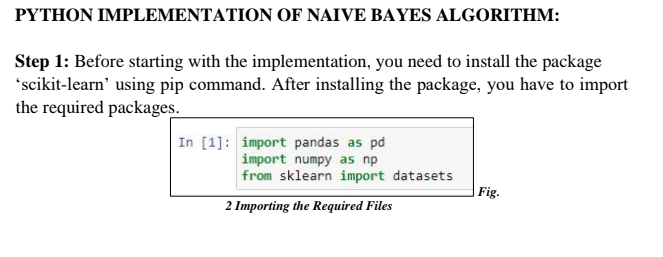
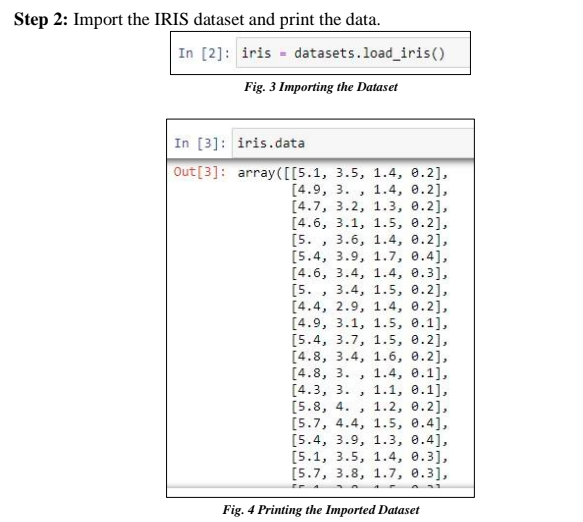
\*\*\*\*\*\*\*\*\*\*\*\*\*Exp no 6\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**AIM:** Implementation of Naive Bayes Algorithm using Python.





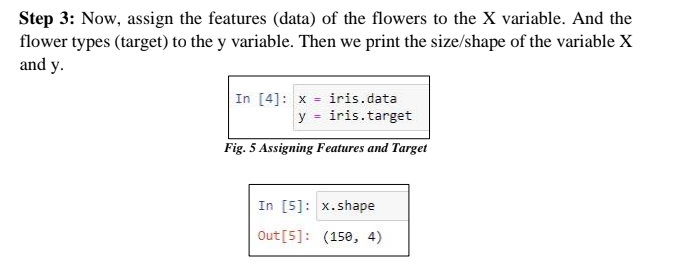
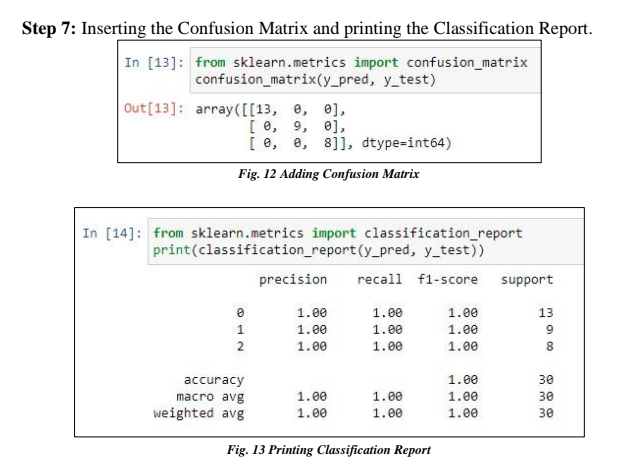
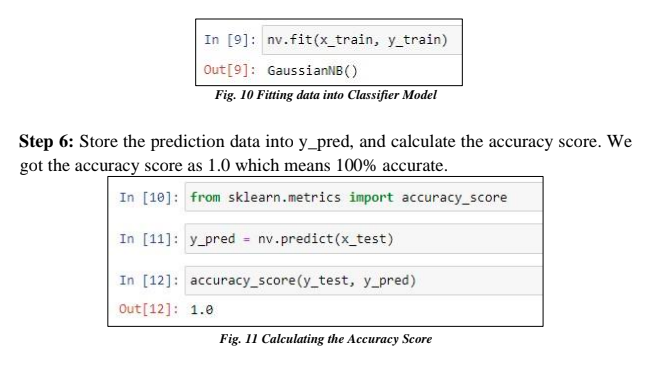
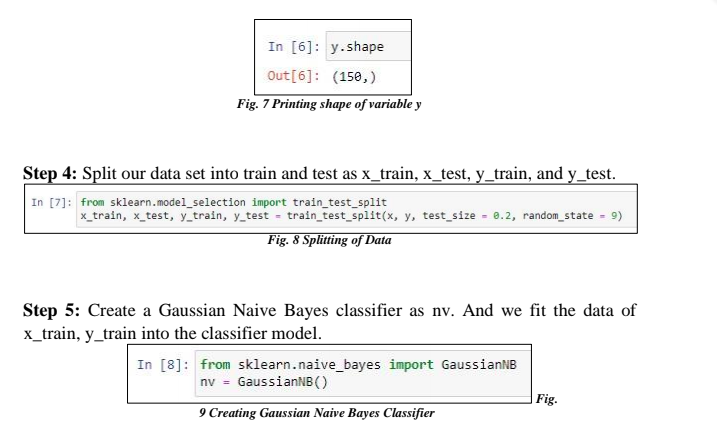


Fig. 6 Printing shape of variable x

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Exp no 7\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Aim: Implementation of K-means clustering using Python.

Theory:

K-MEANS ALGORITHM:

K-Means Clustering is an unsupervised learning algorithm that is used to solve the

clustering problems in machine learning or data science. In this topic, we will learn

what is K-means clustering algorithm, how the algorithm works, along with the

Python implementation of k-means clustering.

K-MEANS CLUSTERING :

Recall the first property of clusters – it states that the points within a cluster should

be similar to each other. So, our aim here is to minimize the distance between the

points within a cluster.

WHEN TO USE CLUSTER ANALYSIS?

This is one of the decisions we have to take while dealing with problems. Taking

Decision is not a tedious task as it solely depends upon the type of data we are using.

If we are using a labeled data we can use classification technique whereas in case

when the data is not labeled we can cluster the data based on certain feature and try

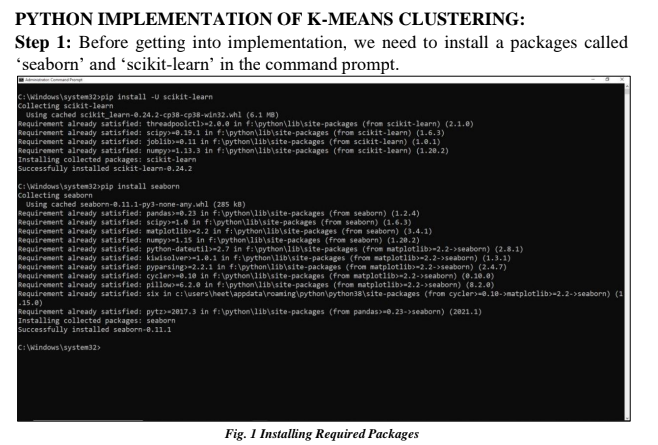
to label it on our own.

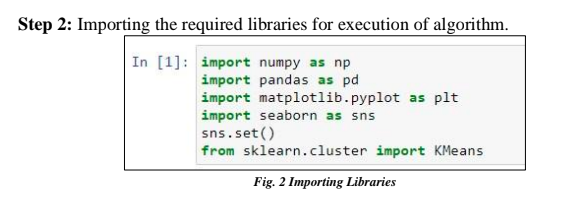
So when we use cluster analysis we don’t have labels (i.e. data is not labeled) in the

context of machine learning this is called unsupervised learning.

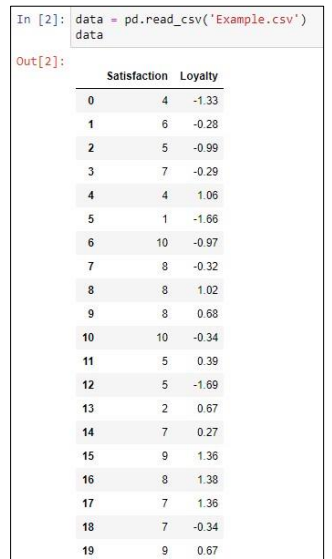
The goal of clustering is to maximize the similarity of observation within the cluster

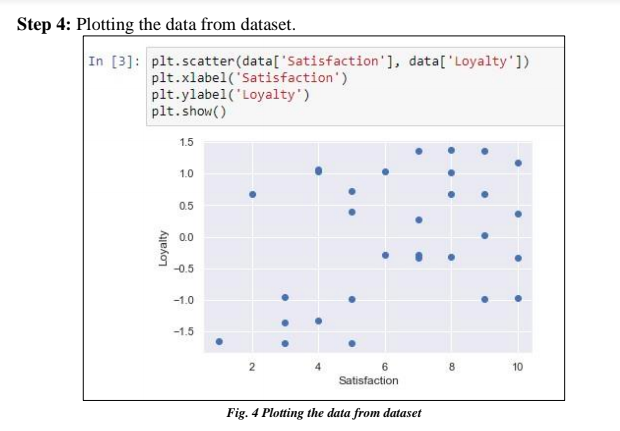
and maximize the dissimilarity between the clusters.

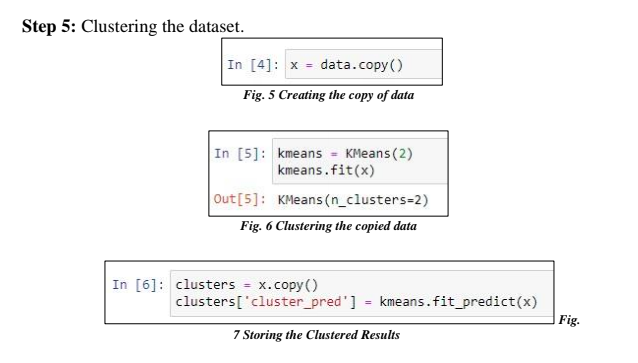


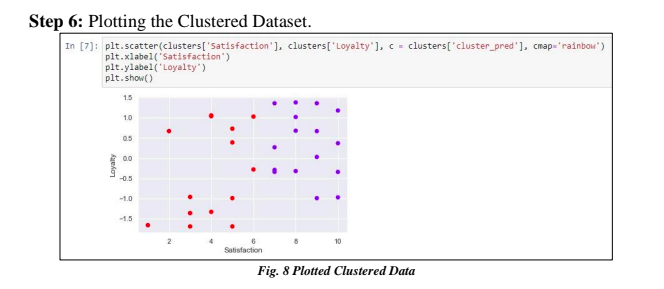


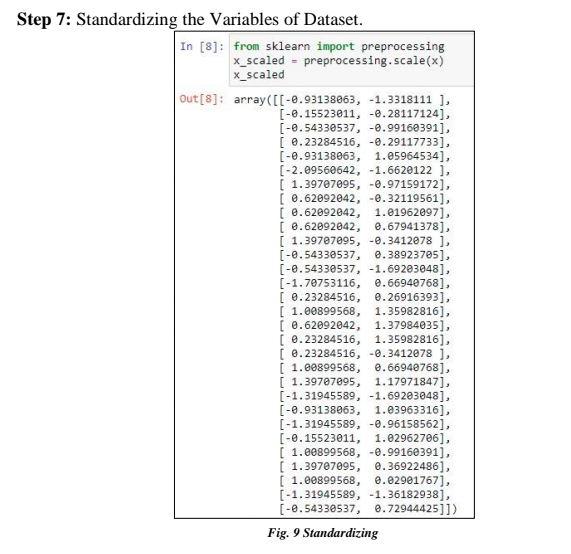
Step 3: Loading the desired dataset and reading it.

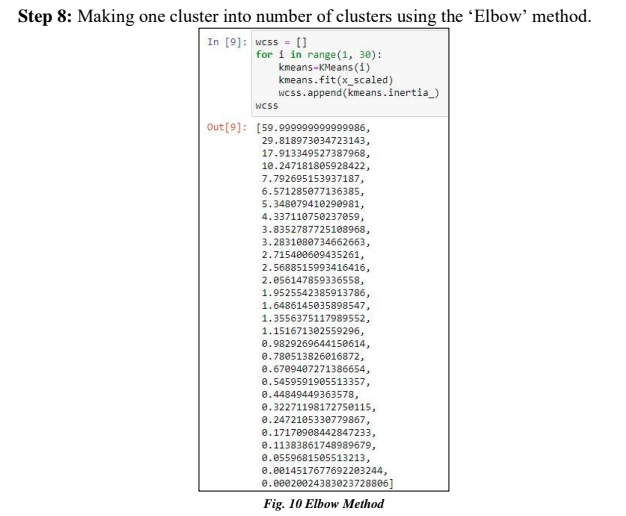


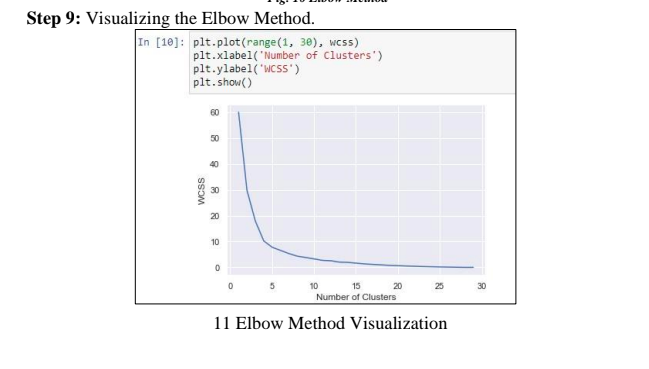






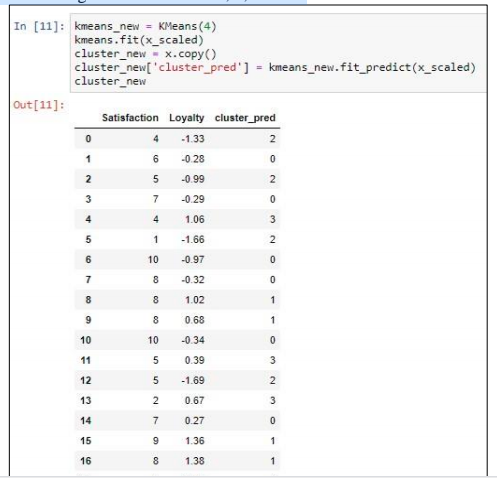


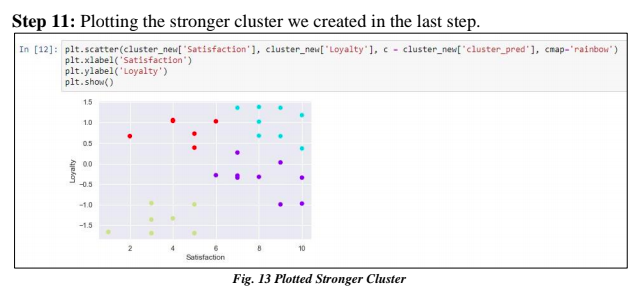




Step 10: Making a stronger cluster for the dataset. Here there are four clusters so our

whole data is categorized into either 0, 1, 2 or 3.





Step 12: Analyzing the data (Final Step in K-Means Clustering).

Through the given figure following things can be interpreted:

1. The yellow dots are the people who are less satisfied and less loyal and therefore

can be termed as alienated.

2. The red dots are people with high loyalty and less satisfaction.

3. The sky blue dots are the people with high loyalty and high satisfaction and they

are the fans.

4. The purple dots are the people who are in the midst of thing