VU1F1920014
Tejas Shinde
BE comps-A/A

ML Exp No: 7

Aim: To understand and implement graph-based clustering (Kmeans)

Theory:

K-Means clustering is an Unsupervised learning algorithm, which groups the unlabeled dataset into different clusters. A convenient way to discover the categories of groups on its own without the need for any training.

Step-1: Select the number k to decide number of cluster.

Step-2: Assign each data point to their closest centroid, which will form the predefined K clusters.

Step-3: Calculate the variance and place new centroid of each cluster

step-4: Repeat 2nd step which means reassign each datapoint to new closest centroid of each cluster.

step-5: If any reassignment occurs, then go to step 3 else go to Finish

Parameter

n_clusters :- The number of clusters to form

Cluster_centers_ :- Coordinates of cluster centers

inertia_ :- Sum of squared distance of sample to their closest cluster center

In program, algorithm has divided data into 3 clusters.
Red for Setosa
Green for Versicolor
Blue for Virginica

Conclusion :- Hence, we have successfully implemented Kmeans clustering.