Experiment 8

The aim of the experiment is to understand the use of dendrogram in K-Means and Hierarchical Clustering algorithms in python.

1. Read the file 'facebooklive.csv' using read_csv() in pandas.

2. Extract two features 'num_shares' and num_loves' and scale them using StandardScalar() in sklearn.

3. Generate a linkage matrix for the data and use it to build dendrogram using scipy python package.

4. Analyze dendrogram to identify required number of clusters. Use this value to implement and visualize K-Means clustering.

5. In the dendrogram, vary the heights (h=[100,60,20]) and observe the change in number of clusters obtained.

6. Visualize the clusters obtained at each height in step 5.

7. Implement Agglomerative Clustering using sklearn package for various metrics ('euclidean', 'manhattan') and linkages ('ward','single','complete', 'average'). Visualize the clusters obtained in each case. Use step 5 and 6 to identify appropriate number of clusters.