

CS4412 Class Assignment #2

Due: 03/24/2020 at 11:59 pm

Instructions:

1. Please include your name, class and assignment number in the comments at the start of the program.
2. Submit your assignment (python code) as Zip file through D2L using the appropriate assignment link.
3. This is a class assignment for dates 3/19, 3/24.

Objective:

The goal of Class Assignment #2 is to reinforce the concept of clustering and its implementation.

Assignment Problems:

1. Implement K-Means and Hierarchical Agglomerative Clustering (HAC) in Python using the dataset (100 points)
<https://archive.ics.uci.edu/ml/datasets/Facebook+Live+Sellers+in+Thailand>
 - a) Consider only the following attributes for clustering
status_type,num_reactions,num_comments,num_shares,num_likes,num_loves,num_wows,num_hahas,num_sads,num_angrys
 - b) Use the following as reference
<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>
<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.AgglomerativeClustering.html>
 - c) Try different values of 'n_clusters' parameter in K-means and 'linkage' parameter in HAC (though you can as well as use other parameters also).
 - d) Use scatter plot to visualize the clusters.