CS5542 Big Data Apps and Analytics

In Class Programming –7
8th October 2020

Submit ICP Feedback in Class. : [Lnik to Feed back Form](#)

K-means clustering:

Use a different data and use the model provided in ICP7 to perform clustering. You must try 5 different number of clusters (for example n_clusters= 5 or n_clusters=6,7,8,or 9 etc) based on elbow curve and for each cluster visualize the clustering results and report your findings in detail.

ICP Requirements:

1) Successfully executing the code and trying 5 different number of clusters based on elbow curve and visualizing those clusters using python plotting libraries (75 points)
2) Using a new and good dataset (5 points)
3) overall code quality (10 points)
4) Pdf Report quality, video explanation (10 points)

Submission Guidelines:

Same as previous ICPs.