CLL 788

Assignment 4

Carbon and nitrogen emission tests of 2 different types of vehicles were done. Test results are provided in Excel sheets. Your task is to identify the two groups of vehicles from the data.

1. Plot the data (Data.xlsx) to get an idea of the data distribution. Plot Result 1 on x-axis and Result 2 on y-axis. Report your visual observations.
2. Apply K-Means clustering on the data to find out the 2 clusters. Make appropriate plots.
3. Plot the data (Data_GMM.xlsx) to get an idea of the data distribution. Plot Result 1 on x-axis and Result 2 on y-axis. Report your visual observations.
4. Apply Gaussian Mixture Model on the Data_GMM.xlsx to find out the 2 clusters. Make appropriate plots.
5. Compare the two methods used.
6. Manually perform K Means clustering on Manual_Data.xlsx. There are 10 data points given and you have to separate them into 2 clusters.

Submission Instructions

1. Submit a zip file on moodle named "EntryNumber.zip" with all the code files and a **report with all the graphs and analysis**. Only Matlab & python are allowed.
2. Use of GMM and K-means libraries are strictly not allowed.
3. Deadline for the submission is 5th July 2020.
4. For any doubts in the assignment, contact: Devansh Agrawal: devansh.agrawal1997@gmail.com