

Machine Learning Assignment

You are given a dataset containing text from Mueller Report (dataset [here](#)). You have to cluster sentences from this report to get optimal number of clusters. Do mention how you can optimize number of clusters.

Do note that you have to cluster sentences not lines.

Deadline: 60hrs from when assignment problem statement is given

Marking Scheme:

- | | |
|-----------------------|-----|
| 1. Preprocessing | - 5 |
| 2. Sentence Embedding | - 8 |
| 3. Clustering | - 5 |
| 4. Optimizing | - 2 |

You have to submit a **python file** directly to us. Do not share your code on any public code sharing platform like GitHub, Kaggle, etc. Do not share your code with anyone privately even.

For any query and submission mail us at: grover.puneet1995@gmail.com and (cc) hr@myways.in

Good Luck!,
Team MyWays