

Group 4 - Meeting

Date: 19/03/2023

Topic: Categorization of Menopause data from social media.

Ambassador Attended: Pallavit Aggarwal

Ideas discussed:

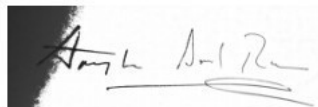
1. The discussion centered around various algorithms and techniques used in text clustering, such as K-means, DBSCAN, and hierarchical clustering. We discussed about the importance of selecting the right clustering algorithm based on the specific problem, dataset, and desired outcome, as well as understanding the strengths and weaknesses of each method.
2. The conversation also focused on evaluating and validating the results of text clustering. We discussed about different evaluation metrics, such as the silhouette score, adjusted Rand index, and mutual information. Additionally, they discussed the importance of using domain knowledge and visualization techniques (e.g., t-SNE or UMAP) to validate and interpret clustering results, as well as the need to iteratively refine the clustering process based on these insights.

Signatures To be added

Chirag Saxena,



Amogh Rao,



Vaibhav Srivastava,



Junwei Yu

