

Network Analysis Web Application

This project helps in better understand a social media network by detecting and analysing of various key parameter of the network formed using graph algorithms.



WEB TECHNOLOGIES LIST



Backend

- Flask
- MongoDB
- Instaloader
- numpy
- matplotlib
- plotly

Front End

- HTML
- CSS
- JS
- Dash
- Dash-html-components

Functional requirements of project

1. Network Visualization
2. Cluster Analysis
3. Finding Centroidal Node
4. Metric Analysis of Network
5. Data processing and scraping from social media.
6. Develop Backend API to serve scraped data and visualisations
7. Develop Front End Independent Components for each Graph and Feature
8. Unit testing and integration testing on final Feature merge.
9. Set Up github actions and CI/CD for Deployment on Heroku.

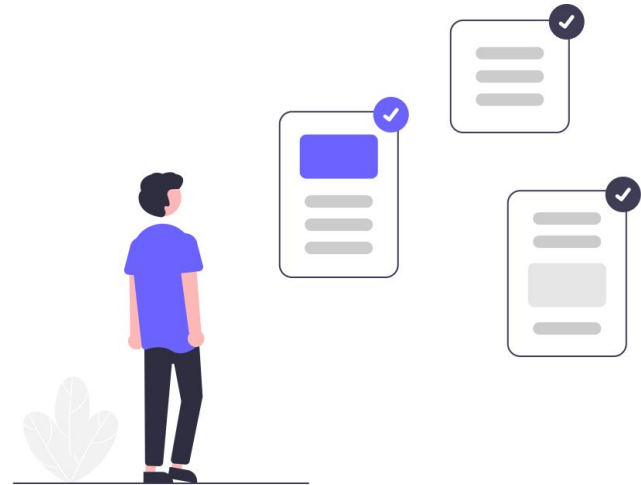


Backend Tasks completed

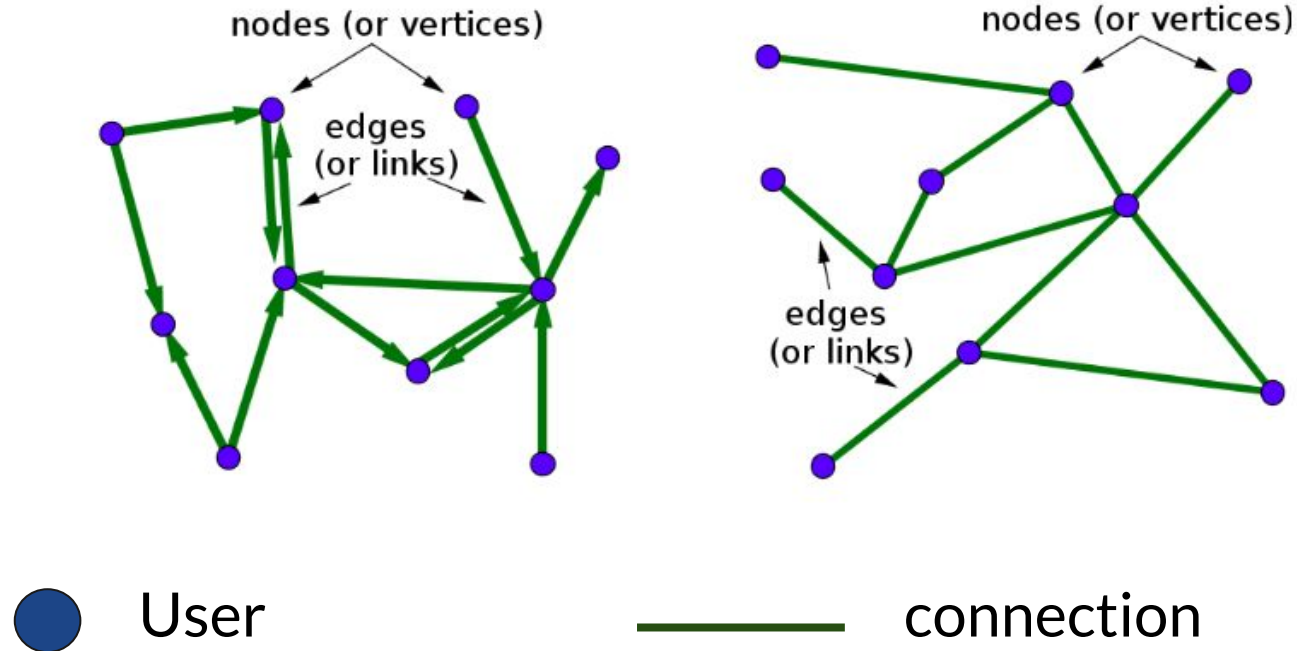
1. Data processing and Scraping from social media
2. Backend API to serve data and visualisations.
3. Using Matplotlib and plotly for visualization and Analysis

Frontend Tasks completed

1. API sync and adding service to fetch data from Backend.
2. Modularization & Components Design visualisations

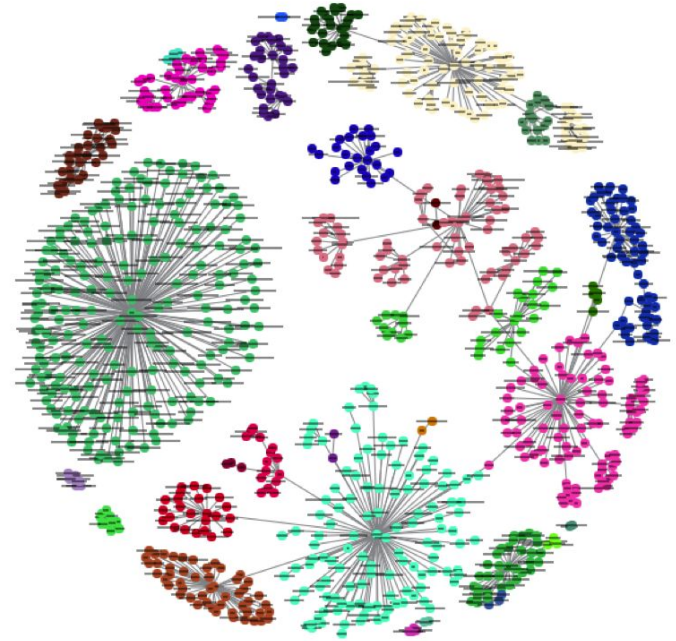


Definitions and Overview



CLUSTER ANALYSIS OF NETWORK

Hierarchical k-means clustering is used to form clusters. Each user is initialised to be an independent cluster, then with each iteration in the **NETWORK GRAPH**, clusters that are close to each other are merged, and clusters having at least one pair of users with distance higher than some threshold are split into two.

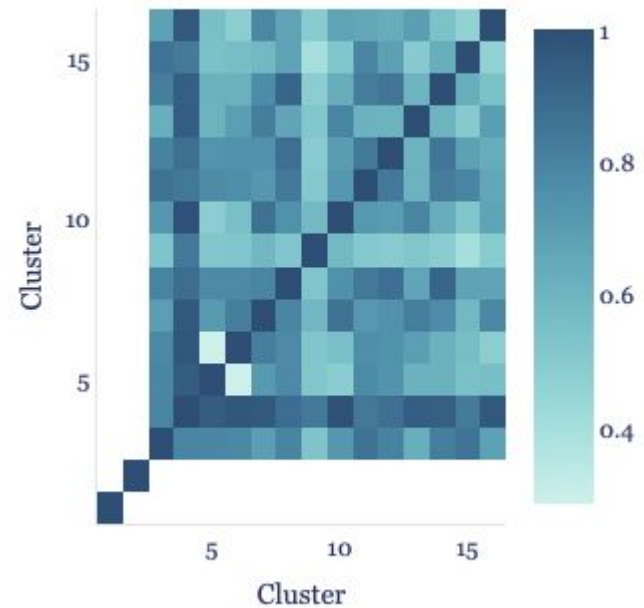


EVALUATING CENTROIDAL NODE

User in a cluster is ranked by a combination of two factors -

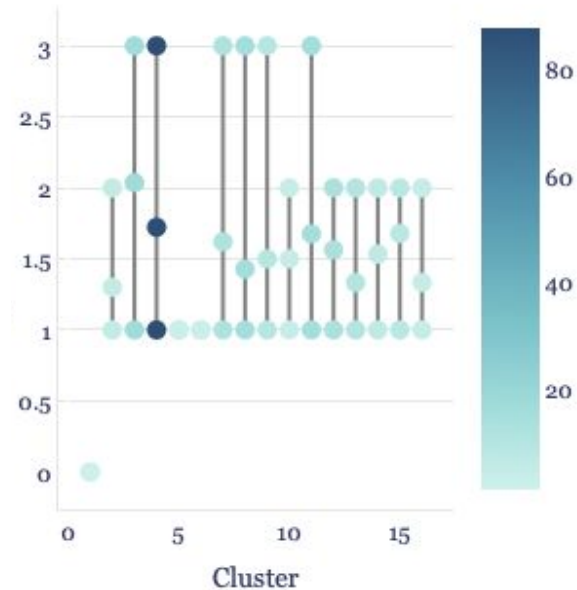
- Mean distance to reach other users in the same cluster.
- Minimum distance to reach any one of the users in the cluster

The user with the lowest score will be the centroid of the cluster.



METRIC ANALYSIS USING COHESION

Performing a quantifiable measure of the interconnection and closeness of the nodes of the **GRAPH** using the Distance between two nodes, which is measured using the shortest path between users, i.e. the minimum number of profiles a user has to visit before the target user's profile can be reached.



ADARSH AGARWAL
2K17/MC/008

SACHIN DUHAN
2K17/MC/087

TANAY SINGHANIA
2K17/MC/107



Thank You!