1. **Title**

Analysis of Community Structures and User Interactions in the Facebook Social Graph

1. **Key idea of project**

*2.1 What is the goal of the project?*

The primary goal of this project is to uncover and understand the structures of communities within Facebook and the nature of interactions among users. By delving into these community networks, we aim to identify how users form clusters based on shared interests and how they interact within these groups.

*2.2 Why is it worth doing?*

Investigating the organization of communities and user interactions on Facebook can help us understand the dynamics that drive group formations and social cohesion. Insights from this research are crucial for designing better community management tools, developing targeted content strategies, and enhancing user engagement tactics.

*2.3 Who did something similar before?*

Prior research has often focused on user behavior and community dynamics within specific Facebook groups or events, examining aspects like network effects and interaction patterns. However, comprehensive analyses that synthesize community detection with detailed user interaction studies are scarce, indicating a significant opportunity for deeper exploration.

*2.4 How are we going to reach the goal?*

Our approach involves deploying advanced network analysis techniques to dissect the Facebook social graph, focusing on community detection and the analysis of user interactions. We will use clustering algorithms to segment the network into communities and analyze the types of interactions that are prevalent within and across these groups.

1. **Dataset required to do the project**

*3.1 Description & link*

We will utilize the "Stanford Open Facebook Social Graph Dataset," which offers a detailed view of user interactions and community structures on Facebook. This dataset includes extensive data on user relationships and group affiliations.

**Link:** *https://snap.stanford.edu/data/ego-Facebook.html*

*3.2 Limitations*

This dataset, while comprehensive, represents only a snapshot in time and may not capture the evolving dynamics of user interactions or the formation of new community structures over longer periods. The anonymization of data also restricts the tracing of individual user trajectories across different community contexts.

1. **Anticipated SNA Methods**

*4.1 Network Construction*

A graph will be constructed with nodes representing users and edges denoting various types of interactions such as friendships, comments, or likes. This will help us visualize and analyze the structural properties of communities.

*4.2 Community Detection*

We will apply state-of-the-art community detection algorithms, such as the Louvain and Girvan-Newman methods, to identify and delineate community clusters that showcase similar interaction patterns or shared characteristics.

*4.3 Centrality Analysis*

Identifying key influencers by calculating degree, betweenness, closeness and eigenvector centrality. *Degree centrality* will be used to find active users who interact with many others; *betweenness centrality* will be used to identify users who act as bridges between different user groups; *closeness centrality* will be used to spot users who can quickly spread information across the network and, finally, *eigenvector centrality* will be used to determine users who are influential sue to the connections with the other influential groups (not their activity)

*4.4 Interaction Analysis*

By examining interaction metrics such as frequency, type, and reciprocity of user activities within communities, we can gain insights into the roles and influence of different users in these networks.

*4.5 Visualization*

Employing tools like Gephi or NetworkX in Python to visualize the network and its various characteristics.