PREPAREED BY TEAM 6:

1. SHERRANETTE TINAPUNAN
2. YOHANNES DEBOCH

6th May 2019

PROJECT PROPOSAL

The main purpose of this project will be to carry out a network analysis on social media data, which will entail a detailed analysis of social structures using graphical theorems in R studio. The dataset in consideration is the social circle Facebook dataset, obtained from <https://snap.stanford.edu/data/#socnets>. It does consist of different circles of friends, all of which are from Facebook. The data was collected by using a survey statistic of all individuals using the Facebook application. It does include all the node features, that is each profile of an individual, the circles (friends), as well as the ego networks. A characteristic that the data has is the anonymity factor where each user id has been anonymized and respective feature vectors provided.

The project proposal therefore seeks to carry out a network analysis on the circles of each profile id, and thus we propose to focus mainly on the edges file in the provided dataset. By the use of the *igraph* package on R studio, we will analyze the data to find out the most susceptible community in the dataset that is likely to be affected by social network analysis. After a clear description, we will then proceed and seek to define the shortest network path that one can use to identify a unique profile in the data.