Mini Project 3 – Network Analysis

Congratulations! You reached the last milestone of the course!

Data scientists not only process and utilize data but also often propose project ideas, plan them carefully, implement them and finally report their findings. This is the goal of mini project 3. This project will take 2 weeks. The goal of this project is to use data of your choice to answer an interesting question of your own using network analysis. This data can be crawled from Twitter, Reddit, Facebook, Linkedin etc. if you want or you can use an already existing dataset. Some examples of place where you can find datasets:

- https://snap.stanford.edu/data/
- https://networkrepository.com/
- https://www.kaggle.com/datasets
- https://datasetsearch.research.google.com/

Please follow the steps below:

Week I

- Setup the necessary tools and get the permissions (if needed) for accessing the data.
- Think of an interesting question you want to answer using the data.
- Send a one-page description including the following parts through the link in the Moodle page:

  - Problem statement
  - Problem importance
  - Your proposed approach
  - The expected outcome

I will either approve or suggest improvements on the problem description within 48 hours. If your proposal gets accepted you can start working on it immediately; if not, you have 24 hours to apply the comments and send the description back to me.

Week II

- Implement the project
- Write a scientific report for the project
- Use visualization to explain better the results and the data
- Submit the code and the report to Moodle