Università degli Studi di Padova

Computer Engeneering

Learning From Networks
Project Proposal

Alessio Cocco - 2087635
Andrea Valentinuzzi - 2090451
Giocanni Brejc - 2096046

Accademic Year 2023-2024

Motivation

Datasets:

- https://networkrepository.com/bio-mouse-gene.php
- https://networkrepository.com/bio-human-gene1.php

The two datasets contain respectively 43k, 21k Nodes and 14.5M, 12.3M Edges.

We are interested in finding similarities and differences between human and mouse genes regulatory networks.

It could be relevant to medicine in order to find parallelisms between the two organisms.

Method

Problem: Node-level analysis (closeness centrality, betweenness centrality, clustering coefficient, graphlet)

Algorithm: we are going to use exact algorithms seen in class. If the problem becomes too complex, we will use approximations.

Intended experiments:

Machine for experiments: we have 2 PCs with 24 cores and 32gb RAM each, so we plan to compute in parallel the two analysis.

We will use networkx or networKit, alternatively we will implement the algorithm ourselves if we deem it necessary.

We will compare the feature extracted for each graph looking for particular similarities in node importance or structure.