Homework Assignment 4

CSDS 446: Machine Learning on Graphs (Kevin S. Xu)
Case Western Reserve University

Assigned: Friday, February 23, 2024

Due: Wednesday, March 6, 2024, at 11:59pm

You must submit your homework electronically on Canvas, including any code you wrote to solve the problems. The homework assignment will be graded for relevant completion.

Problem 1

(Continuation of Problem 7 from Homework Assignment 2)

Remove the following 8 edges from Zachary's Karate Club network: (1,5), (2,4), (3,29), (6,17), (9,34), (16,33), (24,26), (25,32). Use each of the following methods to perform link prediction over all non-edges:

a) Degree-corrected stochastic block model with 2 groups.
b) Latent space model in 2 dimensions.
c) DeepWalk in 2 dimensions. I suggest you use the implementation of DeepWalk contained in the `karateclub` Python package[1].

For each method, list the top 10 predictions, sorted in decreasing order of similarity, and evaluate the improvement over random guessing. (Note: in practice, you would repeat this experiment over multiple sets of randomly selected edges to remove and average over the multiple repetitions.) ***Submit your code in a file named karate_link_prediction_models.py!***

Problem 2

Consider the following latent variable generative model for an undirected graph with no self-edges:

1. For each node i, sample a latent variable $x_i \sim \mathcal{N}(\mu, 1)$ independent of all other nodes.
2. For each pair of nodes (i, j) where both $x_i, x_j > 0$ or both $x_i, x_j \leq 0$, form an edge independent of all other pairs of nodes with probability θ.
3. For each pair of nodes (i, j) where $x_i > 0, x_j \leq 0$ or $x_i \leq 0, x_j > 0$, form an edge independent of all other pairs of nodes with probability ϕ.

Compute the probability that there is an edge between an arbitrary pair of nodes (i, j). Write the probability as a function only of μ, θ, ϕ and the cumulative distribution function of the normal distribution Φ, where $\Phi(x) = p(X \leq x)$.

[1] https://karateclub.readthedocs.io/