Social Media Analytics Assignment 2

Find two relevant datasets related to community detection, information diffusion and recommender systems from any of the open-source databases such as Kaggle, GitHub etc. and implement above mentioned algorithms on both the datasets.

Marks:

1. Dataset1 implementation 4 marks

2. Dataset2 implementation 4 marks

3. Comparison of two datasets implementation 2 marks

Follow the below mentioned steps for implementation of the algorithms such as community detection, information diffusion and recommender systems.

1. Follow the steps in python to visualize the result for Community Detection algorithms.
    a. Load and Preprocess the Data 0.5 Marks
    b. Graph Construction and Representation 1 Mark
    c. Implement Community detection algorithms 1 Mark
    d. Evaluate the result with respective to Number of communities, Size Distribution (i.e., nodes distribution in all communities) 1 Mark
    e. Validation and Visualization of the results 0.5 Marks

2. Follow the steps in python to visualize the result for Information Diffusion algorithms.
    a. Load Preprocess the data, if necessary, assign attributes to nodes or edges 0.5 Marks
    b. Graph Construction and Representation 1 Mark
    c. Apply Information Diffusion Model 1 Mark
    d. Evaluate the results 1 Mark
    e. Visualize the graph to show the network and highlight the spread of information over time. 0.5 Marks

3. Follow the steps in python to visualize the result for Recommender Systems.
    a. Load, clean the data and convert categorical variables into numerical values if necessary 0.5 Marks
    b. Exploratory Data Analysis (EDA) by analyzing the distribution of users, items and interactions then create visualizations using histograms 1 Mark
    c. Split Data into Training and Testing Sets 0.5 Marks
    d. Apply the Recommender Systems algorithm 1 Mark
    e. Make Predictions (generate recommendations, rank items etc.) 0.5 Marks
    f. Evaluate the Model using metrics such as accuracy, precision, recall, F1-score 0.5 Marks

Set 1: a. Brute Force Community Detection Algorithm

b. Page-Rank-Based Model Information Diffusion Algorithm

Set 2: a. Louvain Community Detection Algorithm

b. Threshold Model for Information Spread Information Diffusion Algorithm

Set 3: a. Girvan-Newman Community Detection Algorithm

b. Rumor Spreading Model Information Diffusion Algorithm

Set 4: a. Label Propagation Community Detection Algorithm

b. Page Rank Based Model Information diffusion algorithm

Set 5: a. Spectral Clustering Community Detection Algorithm

b. User-Based Collaborative Filtering Algorithm of Recommender Systems

Set 6: a. Clique Percolation Method (CPM) Community Detection Algorithm

b. User-Based Collaborative Filtering Algorithm of Recommender Systems

Set 7: a. Girvan-Newman Community Detection Algorithm

b. Item-Based Collaborative Filtering Algorithm of Recommender Systems

Set 1: 1 8 15 22 29 36 43 50 57 (Group Nos.)

Set 2: 2 9 16 23 30 37 44 51 58

Set 3: 3 10 17 24 31 38 45 52 59

Set 4: 4 11 18 25 32 39 46 53 60

Set 5: 5 12 19 26 33 40 47 54

Set 6: 6 13 20 27 34 41 48 55

Set 7: 7 14 21 28 35 42 49 56