MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY

P.O. Box 972-60200 – Meru-Kenya
Tel: +254(0) 799 529 958, +254(0) 799 529 959, + 254 (0) 712 524 293,
Website: info@must.ac.ke Email: info@must.ac.ke

University Examinations 2024/2025

FOURTH YEAR FIRST SEMESTER FOR THE DEGREE OF BACHELOR OF SCIENCE IN DATA SCIENCE

CDS 3402: SOCIAL NETWORK ANALYSIS

DATE: JANUARY 2025   **TIME: 2 HOURS**

INSTRUCTIONS: *Answer question **ONE** (Compulsory) and any other **TWO** questions*

QUESTION ONE (30 MARKS)

a) Define each of the following terms as used in Social Networks

    i.   Social Network Analysis (SNA)   (2 Marks)

    ii.   Node   (1Mark).

    iii.   Edge   (1Mark)

    iv.   Clustering Coefficient   (1 mark)

b) Using a real world example, explain the concept of 'Small-world Networks' (4 Marks)

c) Distinguish between each of the following SNA concepts

    i.   Ego Networks and Global Networks   (4 Marks)

    ii.   Degree Centrality and Betweeness Centrality   (2 Marks)

    iii.   Closeness Centrality and Eigenvector Centrality   (2 Marks)

d) Describe each of the following algorithms as used for community detection in Social networks.

    i.   Girvan-Newman Algorithm   (2 Marks)

    ii.   Louvain Method   (2 marks)

e) Clustering in Social Networks entails assembling nodes or subgraphs into clusters so that nodes inside the same cluster are more similar to one another than to those in other clusters. Briefly, explain how you would apply each of the following methods in your Facebook social network;

    i.   Modularity Optimization   (3 Marks)

    ii.   Spectral Clustering   (3 Marks)

    iii.   Hierarchical Clustering   (3 Marks)

QUESTION TWO (20 MARKS)

a) Smart Brains Ltd. has hired you as the lead scientist in their team. Your first task involves performing sentiment Analysis and Opinion Mining in Azure. The task involves a hotel chain that wishes to analyze customer reviews from their various social media platforms.

Required;

   i.   Explain the concept of Sentiment Analysis and Opinion mining with regard to the hotel chain   (4 Marks)

  ii.   Describe a Scientific data science approach that would employ in your task (4 Marks)

 iii.   Outline at least three key features of Azure's sentiment analysis tools that you Would utilize in your work   (6 Marks)

b) Social networks offer a rich field for data science applications, from analyzing public opinion to studying the spread of information and behaviors. Using examples, explain each of the following SNA features;

    i.   Privacy and Ethical Considerations   (2 Marks)

    ii.   Temporal Dynamics   (2 Marks)

    iii.   Influence Maximization   (2 Marks)

QUESTION THREE (20 MARKS)

a) In the context of node similarity in social networks, derive the mathematical formulations for each of the following metrics;

    i.   Cosine Similarity   (2Marks)

    ii.   Jaccard Index   **(2 Marks)**

    iii.   Structural Equivalence   **(2 Marks)**

    iv.   Discuss how each of the above measures are influenced by the topological structure of the graph and their implications on network clustering **(4 Marks)**

b) Dynamic social networks are networks where nodes (representing individuals or entities) and edges (representing relationships or interactions) evolve over time. Unlike static networks, where connections are fixed, dynamic networks capture temporal changes in interactions, reflecting real-world phenomena.

    i.   Describe any two real world applications that exhibit the temporal networks **(4 Marks)**

    ii.   Discuss three methods that can be applied by a user when conducting link prediction in dynamic networks **(6 marks)**

QUESTION FOUR (20 MARKS)

a) Strategic network formation focuses on how individuals or entities form links or connections within a network, guided by their own preferences or objectives, while accounting for the actions of others. Game-theoretic models are used to study these strategic behaviors, particularly in social networks, where users' decisions to form connections are often influenced by their expectations about how others will act. With relevant examples, discuss each of the following game-theoretic models;

    i.   Non-Cooperative Games   **(3Marks)**

    ii.   Cooperative Games   **(3Marks)**

    iii.   Bayesian Games   **(3Marks)**

    iv.   Dynamic Games   **(3 Marks)**

b) SNAP is a general-purpose, high-performance system designed for analyzing and manipulating large-scale networks. It is widely used in research and industry to process graphs representing social networks, web graphs, and other complex data structures. Outline any four features of SNAP that you would embrace in resolving a Data Science projects **(8 Marks)**

QUESTION FIVE (20 MARKS)

a) You are given a social network represented by the following adjacency list:

```
A: [B, C]
B: [A, C, D]
C: [A, B, D]
D: [B, C, E]
E: [D, F]
F: [E]
```

Model the spread of information using the Independent Cascade Model, assuming the probability of diffusion along an edge is 0.4

    i.   Starting from node A, simulate the first two rounds of information spread(6 Marks)

    ii.   Write Python Code to visualize the diffusion process using a graph tool. Plot the network and show which nodes are influenced after each round   (7 Marks)

b) Write a Python script to compare the spread of information using both models on the following network: (7 Marks)

```
X: [Y, Z]
Y: [X, Z, W]
Z: [X, Y, W]
W: [Y, Z]
```