Student Answer Sheet

July 5, 2025

1

1 0 0 1 1 1 0 1 0 0 0 1

Answer: 1) a) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ rows = vertices (v_1, v_2, v_3, v_4)

columns = edge (E_1, E_2, E_3)

E_1: Connect v_1 and v_2 (Since $a_{i,j}$ in rows 1 and 2)

E_2: Connect v_2 and v_3

Adjacency matrix $A[i][j] = 1$

v_1 $\begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}$

v_2 $\begin{bmatrix} 1 & 0 & 1 & 1 \end{bmatrix}$

v_3 $\begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}$

v_4 $\begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}$

b) → (B) Erdos- Renyi (Random Network) model which assume that each edge between a pair of nodes are formed with the same (uniform) probability independent of all edges

c) → (C) Nash-Equilibrium - where no. player can improve their outcome by unilaterally changing their strategy of other is called Nash Equilibrium

d) → (B) Assertive Mixing is the tendency for individuals in social network to associate and bond with similar to others.

e) → (D) Because it qualifies how often a node lies on the shortest path between other nodes, it is become of betweenness and centrality.

f) → (C) The presence of many nodes with very high degree (hubs) that maintain connectivity

g) → (A) The no. of community intra-community edges is significantly higher than expected in random sequence.

h) → (B) For nodes X (neighbor A,B,C,D) and Y (neighbor C,D,E) the Jaccard Co-efficient is

Intersection $\{C, D\}$ size 2

Union $\{A, B, C, D, E\}$ size 5

$= 2/5$. Ans

i) → (A) ICM - Independently; LTM uses a weighted sum of active neighbors compared to a node threshold.

(j) → (B) Because aggregating feature from dissimilar neighbor can blur the node's own representative feature, making classification harder:

2) To minimize total infection when only 5% of the population can be vaccinated, combine these two network

1. Degree centrality: which will mean how many direct connection (edges) a node (Person) has in the network

Individual with the highest degree centrality are the most connected and thus have greatest potential to spread the infection to many others. Vaccinating those "hubs" can significantly reduce the no. of transition pathways and slow down the epidemic. → Targeting high degree node is effective because of removing them from the susceptible pool and disrupt largest no. of possible transmission route.

2) Betweenness and Centrality: Which will act as a bridge between different groups and clusters

Rigon betweenness Centrality connect otherwise separate part of the network. Vaccinating these people can prevent the disease from jumping between community or clusters.

By Immunizing these nodes, "bridge" nodes, we limit the epidemic's ability to spread local outbreak containing with smaller group and preventing City wide transmission.

3) Combined Approach Step 1: Calculate degree of Centrality and betweenness centrality. Step 2: Select the top 5% of individual to vaccinate prioritizing those who are highly ranked in either or both the measures.

This dual approach can disrupt both most prolific spreaders and the key connectors, maximizing impact of limited vaccination resources, and minimizing the total infection.

3) Step-1 Node Embedding with node to vector: learns researcher from co-authorship/citation network

Step-2 Link - Prediction - which will predict collaboration using embedding based feature

Step-3 Homiphily → be commendation focus Similar field.

Step -4 Cross disciplinary - Adds diversity - awars scoring to Promotion surface interdiscipline Suggestions.

Explanation Node 2 Vec Embedding: It learns low dimensional vector representation for each researcher global network structure such as co-authorship link or citation pattern. These embedding include similarity and proximity in academic network

Link prediction: Using node embedding we can train ML models to predict the livelihood of future collaboration to link between any two researcher. the model can use feature such as the concatenation between embedding of two node

Role of homophily It refers to the tendency of researchers to collaborate within the similar field, or with those who have similar research interest. It naturally captured by the embedding as researcher in same domain with frequently collaborator will have similar vector

Promoting Cross disciplinary collaboration

Diversity Recommenders: use topic modeling to identify researcher in related but different field and the ranking occasion.

Hybrid scoring: It combine the link prediction score with a diversity score that favors cross-disciplinary connection and balancing relevance novelty.

4) a) Core Idea behind the Girvan-Newman Algorithm

The Girvan-Newman Algorithm detects communities in the network by progressively removing edges with the highest edge between centrality, which are most likely connect different communities. As these "bridge" edges are removed the network splits into smaller more densely connected groups dividing community structure.

b) Use of Edge betweenness Centrality 1. Compute edge betweenness centrality for all edges (i.e.: Count of shortest paths through each edge)

2. Remove the edge with highest betweenness then recompute betweenness on the pruned graph, repeat until desired community emerged.

c) Major Computational Limitation

The edge edge betweenness centrality must be recalculated after every edge removal, making the algorithm computationally expensive and unsuitable for large network and computationally expensive

(d) Louvain Method as Scalable alternative: If scales roughly as O(m*n) or (O(m*2) for sparse graph) making impractical large graph network

(e) Louvain method's Scalable Modularity Optimization!

It offers greater scalability by using greedy optimization approach to maximize the modularity and then aggregating communities into super nodes, repeating process recursively which is faster and more efficient network.

5) a) Intuition behind Pagerank

Pagerank models a "random surfer" navigating the web * At each step the surfer either the surfer follows an outgoing link from the current page, or teleport the random page.

* Node importance is the steady state probability of finding the surfer at that node. * Page linked by many pages accrue higher probability capturing a recursive notion of authority.

b) Role of damping factor

* With probability d, the surfer follows random outgoing link of the current node.

* With probability (1-d), the surfer teleports to a uniformly chosen node.

Purpose: 1. It breaks cycles and prevents rank sink in strongly connected components

2. Ensure the transition matrix is irreductable and aperiodic, guarantees a unique stationary distribution.

(c) Dangling nodes - more on this

(e) - Dangling Nodes: Problem and solution This node does not have any outgoing links. So a random surfer could get reaching such a node disrupt the calculation and prevent convergence

To handle this, algorithm redistributes the rank from the dangling node uniformly to all nodes (as if the surfer teleports to a random node) with link to every node. This adjustments ensure the Markov process remains valid and the page rank calculation converges.

Problems: Nodes with Zero-Outlink

absorb probability (rank leak) and stall convergence

Replace each dangling node column in the transition matrix V/n (Uniform links to all nodes) or equivalents feed its rank into the teleportation step each iteration. This restores a fully stochastic matrix greater convergence.

6) Pure strategy for Nash Equilibrium (U,A) (L,B) a)

	Strategy A	Strategy B
Strategy U	(3,2)	(0,1)
Strategy L	(2,0)	(2,3)

A pure strategy Nash equilibrium occurs when neither player can improve their payoff by unilitely changing Strategies

* At (U,A): Player 1 gets 3, Player 2 gets 2. If Player 2 switches to B their payoff drops to 1. If Player switches to L, their payoff drops 2. So, (U,A) is a Nash equilibrium.

* At (L,B): Player 1 gets 2, Player 2 gets 3. If Player 2 switches to A their payoff dropsto 0. If Player 1 switches to U, their payoff drops to 0, so (L,B) is a nash equilibrium

(b) Expected payoff for player 2

Let player 1 play U with probability P 1-P If Player plays B:

Expected payoff $= 2 \cdot P + 0 \cdot (1 - P) = 2P$

If Player 2 plays B Expected payoff $= 1 \cdot P + 3(1 - P) = 2 \cdot P + 3(1 - P) = 3 - 2P$

(c) Expected outcome if P=0.7 * If Player 2 plays A: Expected payoff = 2 x 0.7 = 1.4 * If Player 2 plays B

Expected payoff = 3 - 2 x 0.8 = 3 - 1.4 = 1.6

So if player 1 chooses U with, Probability 0.7. Player 1 Expected payoff 1.4. Player 2 should play in B in this case.

7) Given $h_A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $h_C = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$, $h_D = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

$W = \begin{bmatrix} 0.5 & 0 \\ 0.1 & 0.2 \end{bmatrix}$

Aggregate neighbor

$h_M = \frac{1}{3}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \end{bmatrix}\right)$

$= \frac{1}{3}\begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Linear Transformation $h_B(B) = \frac{1}{3}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \end{bmatrix}\right)$

$W \cdot h_A(B) = \begin{bmatrix} 0.5 & 0 \\ 0.1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.1 + 0.4 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$

ReLU Activation

ReLU $\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$

$h_B^{(1)} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$