



PES University, Bengaluru-85
(Established under Karnataka Act No. 16 of 2013)

UE19CS345

May 2022: END SEMESTER ASSESSMENT, CSE, VI SEMESTER

UE19CS345 – NETWORK ANALYSIS AND MINING

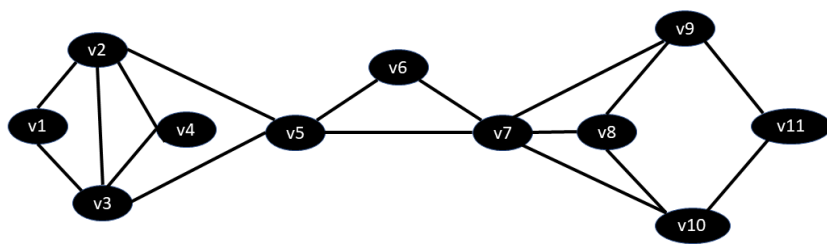
Time: 3 Hrs.

Answer All Questions

Max Marks: 100

Provide full calculation for the numerical problems.

For all other questions, write brief answers in maximum 5 sentences.

1	<p>a)</p> <p>Briefly explain the following :</p> <p>(a) Closeness centrality vs. Harmonic centrality</p> <p>(b) Relation (direct or inverse?) between density and brokerage opportunity of an ego in ego's network</p> <p>(c) Eigen Vector centrality vs. Katz centrality</p>	2*3												
	<p>b)</p> <p>Provide your justifications on the types of graphs as indicated below:</p> <table><tr><td></td><td>Simple or multi graph</td><td>Directed or undirected</td><td>Homogenous or heterogeneous</td></tr><tr><td>Recommender system dataset represented as graph: user item dataset graph when users recommend items. Also, users follow users.</td><td></td><td></td><td></td></tr><tr><td>Wikipedia Knowledge Graph: consists of entities represented through head-relation-tail (h-r-t) tuples</td><td></td><td></td><td></td></tr></table>		Simple or multi graph	Directed or undirected	Homogenous or heterogeneous	Recommender system dataset represented as graph: user item dataset graph when users recommend items. Also, users follow users.				Wikipedia Knowledge Graph: consists of entities represented through head-relation-tail (h-r-t) tuples				2*3
	Simple or multi graph	Directed or undirected	Homogenous or heterogeneous											
Recommender system dataset represented as graph: user item dataset graph when users recommend items. Also, users follow users.														
Wikipedia Knowledge Graph: consists of entities represented through head-relation-tail (h-r-t) tuples														
	<p>c)</p> <div></div> <p>For the graph shown, answer using the principles of centrality :</p> <p>(a) Which are the most popular nodes ?</p> <p>(b) Which nodes can most easily obtain information from other nodes ?</p>	1+3												

For the given graph, assume initial PageRank of $1/n$ where n is the no of nodes and a damping factor of 0.5. Calculate to find out the most central node using PageRank after one iteration.

4

2

a)

Briefly explain the following :

- Foci closure vs. membership closure
- Community detection vs. aggregation in Louvain method of community detection
- A strongly connected component of a directed graph is also a weakly connected component but the reverse may not be always true.

 2^*3

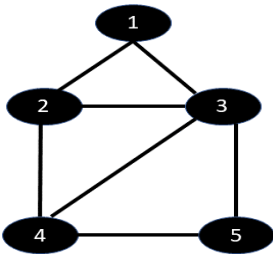
b)

Indicate how the definition of clique as a measure of cohesiveness is relaxed in

- (a) N-clique or K-Clique
- (b) K-clan
- (c) K-plex

 2^*3

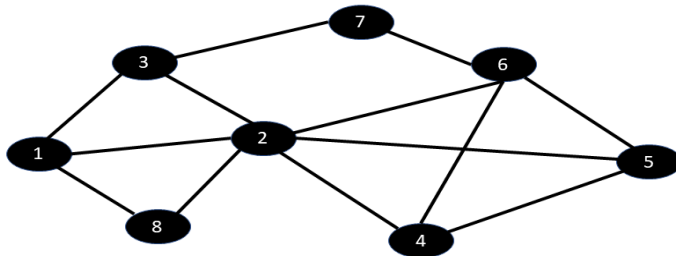
c)



For the graph given, calculate the Global Clustering coefficient using Watts and Strogatz definition.

4

d)



For the given graph, find communities of size greater than or equal to 3 using Clique Percolation method (CPM). Use matrix method. Are they overlapping ?

3+1

3

a)

Briefly explain the following :

- The sources of randomness in $G(n,p)$ and $G(n,m)$ Random Graph Model are different.
- Identify the specific findings of Milgram's experiment that are addressed by Small World model and Decentralized search.
- The random graph model differs from a scale free network with the same number of nodes both in the head and the tail of the degree distribution.

 $2*3$

b)

Briefly explain the following :

- One model out of SI, SIR, SIS and SIRS should be used for an epidemic that keeps coming after an interval of few months.

 $2*3$

<p>d) The social recommendation strategy has been implemented by appropriately modifying the pure collaborative filtering-based implementation used currently by your organization. On the evaluation dataset, you are obtaining the following:</p> <ul style="list-style-type: none"> (a) Out of the 500 relevant items, 300 are not recommended (b) Out of the other 500 non-relevant items, 200 are not recommended. <p>The previous generation pure collaborative has a F1 score of 0.71 . Has the new recommender improved things?</p>	4
---	---

5	a)	You are planning to use Node2Vec and GraphSAGE to derive node embedding for a graph dataset. Compare the two approaches as indicated below:		2*3	
		Node2Vec	GraphSAGE		
	Inductive vs. Transductive learning set up				
	Sharing of learning parameters between nodes for learning node embedding				
	Incorporation of node features and relations between nodes				
b)	Compare the Graph Neural Network architectures GraphSAGE and GAT as indicated below :			2*3	
	GraphSAGE	GAT			
Set of learnable parameters					
Aggregation mechanisms					
Update mechanisms					
c)	You are using Graph Neural Network based approaches for the learning problems mentioned below. Provide your brief justification about how you will set up the learning problem. <ul style="list-style-type: none">A security agency is tasked with identifying the hidden graph of terrorists. However, it is very hard to decipher the network of terrorists as they seldom communicate with each other on the social media platforms. The security agency has a profile of possible suspects obtained by analysing their foot prints on multiple social media platforms.On a social media website where users share creative content around their interests, users can follow each other and use hashtags to classify their genre or interests. Any particular user can use multiple such hashtags in his or her posts. You are asked to come up with a GNN based solution to recommend followership to the users.			2+2	
d)	You are given a list of 4 scenarios and a list of GNN modelling techniques . Given scenarios: (a) Train -test split in a graph level prediction task (b) Increasing expressive power of a shallow GNN (c) Graph being considered is too sparse (d) The learning problem requires deep GNN as limited number of GNN layers is not able to learn Applicable techniques: (a) Replace adjacency matrix A with $A + A^2$ (b) Skip connections (c) Global Max pooling (d) Transductive split of the dataset (e) Inductive Split of the dataset (f) Using MLP (multi-layer perceptron) based pre-processing layer before GNN layer Identify applicable techniques for the given scenarios in the table below :				1*4
	Scenario	Applicable GNN modelling technique			