SRN						



## **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 numeric	cal problems.					
For all other questions, write brief answers in maximum 5 sentences.							

1	a)	Briefly explain the following:  (a) Closeness centrality vs. Harmonic centrality  (b) Relation (direct or inverse?) between density and brokerage opportunity of an ego in ego's network  (c) Eigen Vector centrality vs. Katz centrality											
	b)	Provide your justifications on the types of grap	our justifications on the types of graphs as indicated below:										
			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.					2*3						
		Wikipedia Knowledge Graph: consists of entities represented through head-relation-tail (h-r-t) tuples											
	c)	v1 v2 v6 v6	V7 V8	v10 v10			1+3						
		For the graph shown, answer using the principles of centrality:  (a) Which are the most popular nodes?  (b) Which nodes can most easily obtain information from other nodes?											

SRN						

	d)	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:  a) Foci closure vs. membership closure  b) Community detection vs. aggregation in Louvain method of community detection  c) 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, calcluate 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)	<ul> <li>Briefly explain the following:</li> <li>The sources of randomness in G(n,p) and G(n,m) Random Graph Model are different.</li> <li>Identify the specific findings of Milgram's experiment that are addressed by Small World model and Decentralized search.</li> <li>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.</li> </ul>	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

SRN						

		<ul> <li>Bass model of diffusion has two parameters p and q. For p &gt; q and p<q,, be="" by="" difference="" in="" li="" model.<="" product="" represented="" scenarios="" the="" there="" will=""> <li>Pluralistic ignorance can turn a "collective action event" such as organizing a protest into a failure due to integration of network effect in both population level and local level</li> </q,,></li></ul>	
	c)	Using SIR epidemic model, you are modelling Covid 19 epidemic for your city about which all the statistics are known. All usual assumptions are made while $\beta$ is the effective infection rate and $\gamma$ is the effective recovery rate. <b>Given</b> : $S(t=0) = 200$ , $000$ ; $I(t=0) = 200$ ; $R(t=0) = 10$ ; $\beta = .0001$ ; $\gamma = 1/25$ .  What will be the number of susceptible, infected and recovered <b>at t=1</b> ?	4
	d)	In the network shown, what is the density of the clusters {1,2,3,4}, {5,6,7,8} and {9,10,11,12}?	4
4	a)	<ul> <li>Answer briefly the following questions:</li> <li>Why is the Adamic Adar method more effective than Jaccard similarity in local neighborhood similarity-based link prediction in a social graph set up?</li> <li>What kind of downstream graph ML task is "drug discovery"?</li> <li>In a downstream classification task in graph ML, when is precision more useful as a metric compared to recall?</li> </ul>	2*3
	b)	<ul> <li>Answer briefly the following questions:</li> <li>What is spectrum in spectral graph theory?</li> <li>What will be an obvious Eigenpair of a k-regular graph?</li> <li>How is eigengap used to select the most stable cluster in k-way spectral clustering?</li> </ul>	2*3
	c)	In the attached graph of 7 nodes (1 to 7), labels are provided for only two nodes i.e. node 1 is of class A and node 7 is of class B. Rest of the nodes do not have any label. Rest of the nodes do not have any label. Using Label Propagation Algorithm (LPA) for one iteration find out which nodes will be having label A or B.	4
	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:  (a) Out of the 500 relevant items, 300 are not recommended  (b) Out of the other 500 non-relevant items, 200 are not recommended.  The previous generation pure collaborative has a F1 score of 0.71. Has the new recommender improved things?	4

SRN						

		You are planning to use Node2	•	E to derive <b>node</b>	<b>embedding</b> fo	r a graph dataset.						
5	a)	Compare the two approaches as i	ndicated below:		N - 1 - 2 \( \tau \)	01-04-05						
	,	La divetti e ve Transadvativa la anci		Node2Vec	GraphSAGE	2*2						
		Inductive vs. Transductive learni	<u> </u>	ian laannina mada			2*3					
		Sharing of learning parameter embedding	s between nodes i	or learning node								
		Incorporation of node features a	and relations hetwe	en nodes								
		mediporation of flode reatures to	and relations between	en nodes								
	b)	Compare the Graph Neural Netv	vork architectures G	raphSAGE and GAT	as indicated	pelow :						
	-,	Compare the Graph Neural Network architectures GraphSAGE and GAT as indicated below:  GraphSAGE GAT										
		Set of learnable parameters	•				2*3					
		Aggregation mechanisms										
		Update mechanisms										
				1								
	c)	You are using Graph Neural Network based approaches for the learning problems mentioned below.										
		Provide your brief justification ab	out how you will set	t up the learning pr	oblem.							
		<ul> <li>A security agency is taske</li> </ul>		= :		· · · · · · · · · · · · · · · · · · ·						
		to decipher the network	•									
		media platforms. The sec	urity agency has a p	rofile of possible su	spects obtaine	d by analysing their	2+2					
		foot prints on multiple so	cial media platform	S.								
		<ul> <li>On a social media websit</li> </ul>	te where users shar	e creative content	around their i	nterests, users can						
		follow each other and use	hashtags to classify	their genre or inte	rests. Any part	icular user can use						
		multiple such hashtags in	his or her posts. Y	ou are asked to cor	me up with a G	NN based solution						
		to recommend followersh	· · · · · · · · · · · · · · · · · · ·		·							
			•									
	d)	You are given a list of 4 scenarios	and a list of GNN m	odelling technique	S.							
		Characteristics (a) Tuning test and										
		<b>Given scenarios</b> : (a) Train -test sp shallow GNN (c) Graph being co				•						
		limited number of GNN layers is r		ise (a) The learning	g problem req	ulles deep Giviv as						
		inniced ridinger of Giviv layers is t	iot abic to icarri									
		Applicable techniques: (a) Replace	ce adjacency matrix	A with $A+A^2$ (b)	Skip connection	ns (c) Global Max						
		pooling (d) Transductive split of t	the dataset (e) Indu	uctive Split of the d	ataset (f) Usir	ng MLP (multi-layer	1*4					
		perceptron) based pre-processing layer before GNN layer										
		Identify applicable techniques for	the given scenarios	in the table below	:							
		Scenario	Aj	oplicable GNN mod	elling techniqu	ie						