Birla Institute of Technology & Science, Pilani
First Semester 2022-2023
Graph Mining (CS F426)
Comprehensive Exam 2022
PART A (Open Book)

Date : Dec 22, 2022 Duration: 1hr 45min
Total Marks : 45

Instructions:

1. There are a total of 08 questions. All questions are compulsory.
2. Write important intermediate steps in numerical. Directly writing the final correct answer is not sufficient to obtain full marks.

Q1 [2 marks]. The unnormalized Laplacian matrix L of a simple graph holds the following properties? True/False

a) L is symmetric and positive definite.
b) The number of (linearly independent) eigenvectors with zero eigenvalues for the Laplacian matrix L is equal to the number of connected components in the underlying graph.
c) The second smallest eigenvector of Laplacian can be used to define optimal clustering of nodes into k clusters.

Q2 [1 mark]. In order to design a GNN framework for graph level tasks, which of the following layers combinations graph filtering- activation-pooling is/are possible?

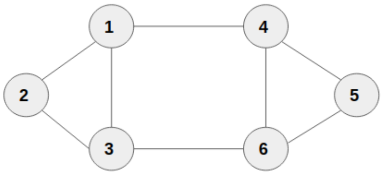
a) 1-2-1
b) 2-2-2
c) 2-2-1
d)None of these

Q3 [2 marks]. At graph level learning, which of the following is/are true about flat pooling?

a) It directly generates a graph representation from node representations.
b) It generates a new graph at each step until a single node graph is obtained.
c) It can be performed by adding a new node to the graph which is connected to all the nodes of the graph.

Q4 [4+2=6 marks]. For a RESCAL decoder, where n and m are the number of entities and relations respectively and $R_r \in R^{d\ X\ d}$ is a trainable matrix for each relation $r \in \{1, 2, \cdots, m\}$. How many total parameters are required to learn? Also explain why RESCAL is known as a 3-way interaction model.

Q5 [4+6+2=12 marks]. For the given graph with 6 nodes, shown below



a) Compute adjacency, degree, and Laplacian matrices.
b) Let us assume that eigenvalues of the Laplacian matrix are 0, 3, 1, 3, 4 and 5. Compute the eigenvectors corresponding to the first two smallest eigenvalues.
c) Also, suggest the possible partitions of the graph using the second smallest eigenvector.

Q6 [6 marks]. Suppose you have a multi-relational knowledge graph with 1000 nodes and 200 relation types. You come up with a RGCN model to learn the embedding of nodes with two hidden layers having 8 and 16 neurons. For each layer, calculate the number of parameters to be learned, and the size of the associated feature maps assuming the effect of the self-node and its neighbors on the final embeddings differently.

Q7 [4+6=10 marks]. Suppose the graph is a chain of n nodes as shown in Figure a.
a) Assume that the initial h is a column vector of ones. Compute the final hub and authorities vectors.
b) If a self-loop is added at the first node as shown in Figure b, compute hub and authority vectors as a function of k, where k is the number of iterations. At every step, normalize hub and authority vectors such that the maximum component of the vector is 1.

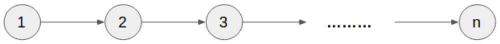


Figure a

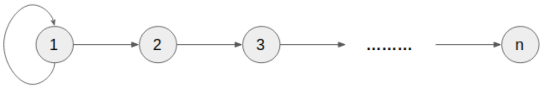


Figure b

Q8 [3+2+1=6 marks]. In the TransE model, a triplet (h, r, t) holds such that $h + r \approx t$.
a) Identify whether the given loss function will be optimized so that the valid triples are ranked above the corrupt triples or not? Justify your answer.

$$L = \sum_{(h,r,t)\,\in\, S}\ \sum_{(h',r,t')\,\in\, S'} \big(\lambda + d(h + r, t)\big) + d(h' + r, t')$$

where S and S'represent the set of true and corrupt triplets respectively, $d(.)$ is a distance function and λ is a margin constant.
b) The TransE model works well for one to one and one to many relations. True/False. Prove your answer.
c) TransE is a two-way interaction model. True/False?

*******************************END******************************

Birla Institute of Technology & Science, Pilani
First Semester 2022-2023
Graph Mining (CS F426)
Comprehensive Exam 2022
PART B (Closed Book)

Date : Dec 22, 2022 Duration: 1hr 15min
Total Marks : 30
Instructions:

1. There are a total of 06 questions. All questions are compulsory.
2. Write important intermediate steps in numerical. Directly writing the final correct answer is not sufficient to obtain full marks.

Q9 [4 marks]. Write down TransE and TransH decoder equations and compare their representational abilities.

Q10 [6 marks] Discuss three advantages of using graph neural network models over shallow embedding approaches to generate node embedding for solving downstream tasks such as classification, link prediction etc.

Q11 [2 marks]. Graph based learning is often referred to as semi-supervised learning. Why?

Q12 [5 marks]. The Jarvis-Patrick algorithm, unlike k-means, automatically determines how many clusters there are; it is still dependent on different input parameters. Explains? Discuss how the Jarvis-Patrick algorithm is similar to k-means, in the sense that results of the clustering are dependent on the parameters? Also, is it possible to partition both directed and undirected graphs using the Jarvis-Patrick clustering? Justify your answer.

Q13 [6+2=8 marks]. Let's assume there are a total of 3204 articles from the New York times newspaper belonging to six different classes: entertainment, economy, international, national, horoscope and sports. The k-means clustering is applied and grouped these articles into 3 clusters as shown in the below table. The first column of the table indicates the cluster and the next six columns together form a confusion matrix i.e. how the articles from each category are distributed in clusters.

Cluster	**Entertainment**	**Economy**	**international**	**national**	**Horoscope**	**sports**	**Total**
#1	1	1	0	11	4	676	693
#2	27	89	333	827	253	33	1562
#3	326	465	8	105	16	29	949
Total	354	555	341	943	273	738	3204

Compute the purity and entropy of each cluster (using class information available in the confusion matrix) and determine which is/are the best cluster(s) in terms of entropy and purity both. ***Note: Use log base 2 for calculation of entropy.***

Q14. [5 marks] You are asked to design a Graph Convolution Network architecture with an input layer, two hidden layers H_1 and H_2 and one output layer. Write down the graph level equation for node representation using UPDATE and AGGREGATION function at H_1 and H_2. X and A represent the input node feature and adjacency matrix respectively. Also explain the strategy to merge UPDATE and AGGREGATION steps together and limitations if any?

*******************************END******************************