



المادة: Social Network Analysis
المدة: ساعة ونصف
الأستاذ: د. حسين هزيمة

المرحلة: الإجازة
السنة المنهجية: الثالثة

الاختصاص: علم البيانات - Data Science

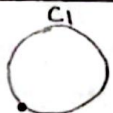
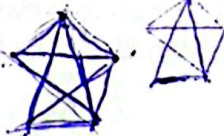
Documents are NOT authorized

Question 1: (16 pts) true-false questions:

Answer the following questions by true or false, and explain the false statements.

1. In directed graphs a node's edge is bidirectional.	T	(F)	There is directions a to b ≠ b to a
2. A graph that has loops is called a simple graph	T	(F)	complex graph
3. A bipartite graph exists has no two vertices within the same set are connected.	(T)	F	
4. A path is the same as shortest path	T	(F)	
5. Local neighbourhood overlap is a node level feature	T	(F)	Link-level feature
6. Google knowledge graph is public	T	(F)	
7. Wikidata knowledge graph is public	(T)	F	
8. Knowledge graphs represents entities of same types	T	(F)	

Question 2: (16 pts) multiple-choice questions:

1. A path from v_0 to v_0 in a graph G , can repeat a vertex or edge, is called	a. Walk b. Circuit c. Cycle d. Path
2. A path from v_0 to v_0 in a graph G , cannot repeat a vertex or edge, is called	a. Walk b. Circuit c. Cycle d. Path
3. Graph C_1 is called 	a. Complete graph b. Connected graph c. Simple graph d. Cycle graph
4. Given that the following graph is complete, what is it's the number edges? <u>Justify with calculation.</u> 	a. 20 b. 15 c. 5 d. 10 $E = \frac{n(n-1)}{2} = \frac{5(5-1)}{2} = \frac{20}{2} = 10$
5. A centrality measure that finds how many people can this person reaches directly?	a. Degree b. Closeness c. Betweenness d. Eigenvector
6. A centrality measure that finds how well is this person connected to other well connected	a. Degree b. Closeness

people?	c. Betweenness (d) Eigenvector
7. Measures how connected v 's neighboring nodes are	a. Graphlet Degree Vector (b) Clustering coefficient c. Eigenvector d. Degree centrality
8. Clustering coefficient is a	(a) Node level feature b. Graph level feature c. Link level feature

Question 3: (38 pts) centrality measures:

1. Given the following graph G:

Node	Edge(s)
a	d
b	d
c	d, e
d	a, b, c
e	b

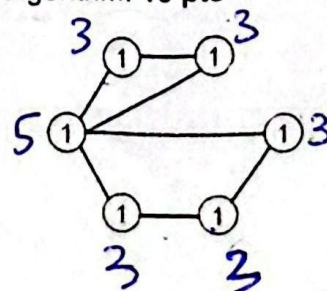
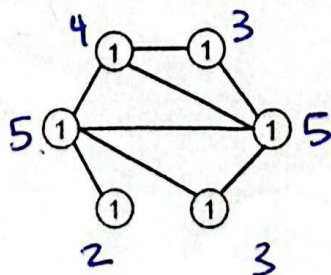
- Draw the network for the above table. **3 points**
- Find the **degree** centrality (dc) of each node in the network, and determine the node with the highest dc. **10 points**
- Calculate the **closeness** centrality (cc) of each node, and determine the node that has the highest cc. **17 points**
- Assume that we added the following node(s) edge(s) to the network, to become G' , how many components does G' have?

Node	Edge(s)
f	g

a. Can we calculate the **betweenness** centrality of graph G' ? Why? **8 points**

Question 4: (30 pts) graph features:

- Find the **clustering coefficient** of graph G in part 3). **10 pts**
- What is the possible number of **graphlets** of G ? **5 pts**
- Given the following two graphs G^1 and G^2 , find the similarity between these two graphs using **Weisfeiler-Lehman Kernel** color refinement algorithm. **15 pts**



Good Work

Appendix

$cc(v) = 2N_v / K_v(K_v - 1)$	$s = C(n, K)$ where $k = n(n-1)/2$
$C_{D_i} = \frac{\sum_{j=1}^n a_{ij}}{n-1}$	$C(x) = \frac{N-1}{\sum_y d(y, x)}$
$C_{btw}(v) = \sum_{s, t \in N} \frac{\sigma_{s, t}(v)}{\sigma_{s, t}}$	
$a \cdot b = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$	$K(G, G') = \emptyset (G)^T \emptyset (G')$