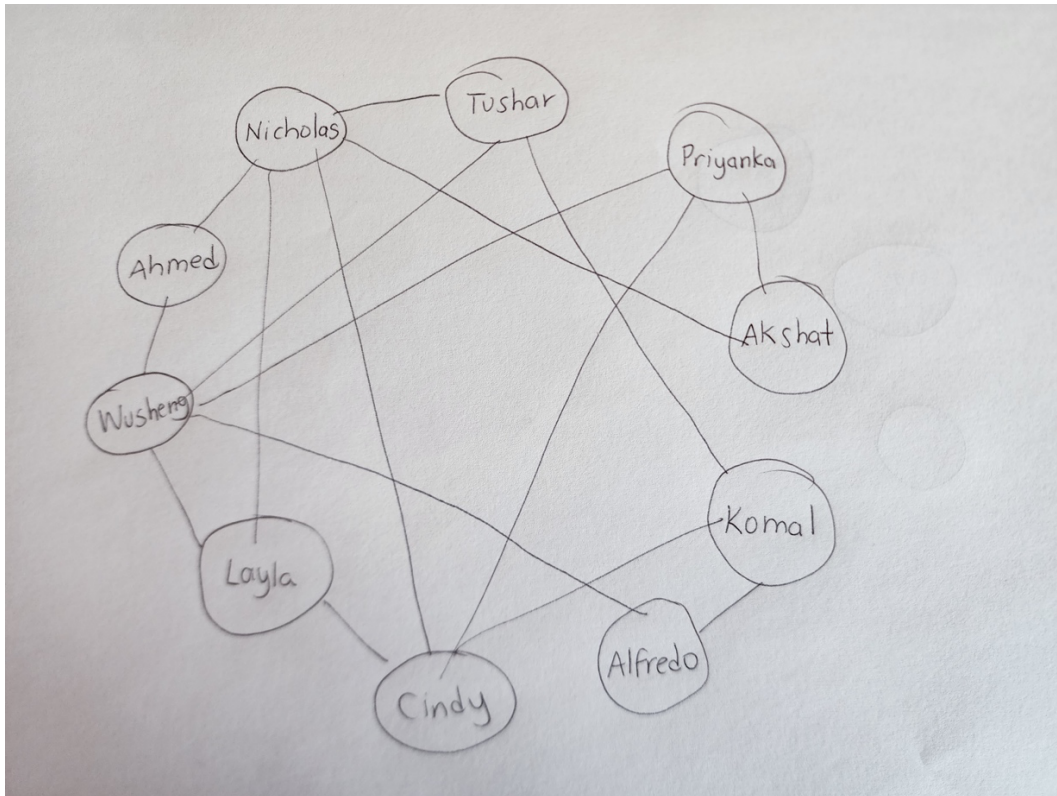


CS 579 Homework 1 due 11:59pm Saturday 3/2. No late submissions after 8am (Chicago time) Monday 3/4.

- Given the friendship graph below, calculate and plot the degree distribution of the graph. Be sure to label the plot axes.



- Draw the graph specified in the adjacency matrix below. Is this graph connected? If yes, is it weakly connected or strongly connected?

$$\begin{bmatrix} 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 7 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

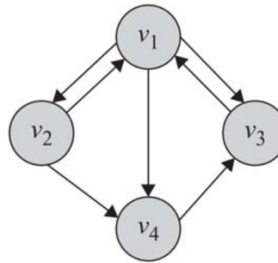
- Use Dijkstra's or Prim's Algorithm to create a shortest path table for the friend graph from problem 1. What is the diameter of this graph? Show a minimum spanning tree of this graph as an adjacency matrix.

- Given the following flow network from  $v_0$  (source) to  $v_7$  (sink), use the Ford-Fulkerson algorithm to determine the maximum flow. Provide the resulting flow, and draw and label the flow network and residual network.

Edge list:

$\{(v_0, v_1, 1), (v_0, v_2, 3), (v_0, v_3, 3), (v_1, v_6, 4), (v_2, v_1, 2), (v_2, v_5, 1), (v_3, v_2, 3), (v_3, v_4, 2), (v_4, v_7, 4), (v_5, v_1, 5), (v_5, v_4, 3), (v_5, v_7, 3), (v_6, v_2, 3), (v_6, v_5, 2), (v_6, v_7, 3)\}$

- For the friendship graph in problem 1, calculate the degree centrality, betweenness centrality and closeness centrality for each node. Provide a table showing the rank of each node for each measure.
- For the friendship graph in problem 1, what is the clustering coefficient?
- For the friendship graph in problem 1, assume that Cindy is a foe of Nicholas and Priyanka, Tushar is a foe of Wusheng and Komal and all the other edges in the graph represent friendship. According to social balance theory, is this new friend/foe graph balanced?
- Calculate PageRank values for the graph below when



- $\alpha=1, \beta=0$
- $\alpha=\beta=0.3$
- $\alpha=0.85, \beta=1$
- $\alpha=0, \beta=1$

Discuss the effects of different values of  $\alpha$  and  $\beta$  for this particular problem.

- You have been tasked to design a classifier that decides whether students will be admitted to a CS graduate program. Applications to the program are received from students all around the world. Applications contain student name, address, mobile phone number, final grade point average in undergraduate program and transcript. Describe what information from the application you would use as input to your classifier. For each piece of information, what (if anything) would need to be done to clean or transform the information into input data. For each of the transformed data input, identify the type (nominal, ordinal, interval or ratio). Give at least 3 other pieces of information that would be helpful and describe why you think they would help.

10. You are given the following set of data

Name	City	Likes Beyoncé	In a relationship	Age	Number of concerts per year	Bought ticket to see Taylor Swift
Kate	Chicago	Yes	No	23	8	Yes
Joe	New York	No	No	36	4	Yes
Mena	New York	Yes	Yes	43	20	No
Pat	Chicago	No	Yes	19	2	No
Tim	Chicago	Yes	No	20	14	Yes
Tina	Chicago	Yes	Yes	54	7	Yes

Using entropy as a measure of purity, design a decision tree to predict whether someone bought a ticket to see Taylor Swift or not. Show how each decision node was selected.