

Name	
ASU ID Number	

## CSE 472: Social Media Mining

### Homework IV - Influence and Homophily, Recommender System

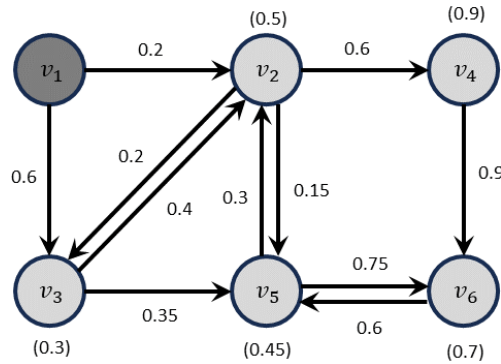
Prof. Huan Liu

Due at 2023, November 22<sup>nd</sup>, 11:59 PM

This is an *individual* homework assignment. Please submit a digital copy of this homework to **Grade-scope**. This is a fillable PDF and you are able to type into answer boxes provided for each question.

#### 1. [Influence and Homophily]

- (a) Apply the procedures of the Linear Threshold Model (LTM) to the given graph until convergence. During each step, identify the activated node(s) and explain the reasons behind their activation. Assume that node  $v_1$  is initially activated at time 0. Traverse the nodes in numerical order and explain all steps.



- (b) The graphs provided illustrate two snapshots: the graph on the left represents the state at time  $t_1$ , and the graph on the right represents the state at time  $t_2$ . Each node  $v_i$  is associated with an ordinal attribute value denoted by  $x_i$ , which can be found in Table 1. Calculate the homophily index of the graph.

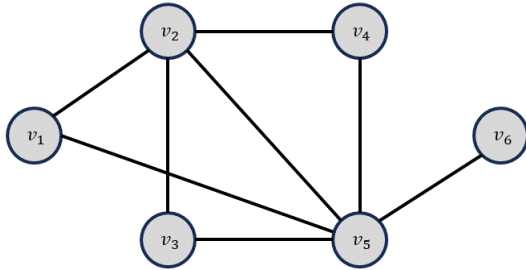


Figure 1: Snapshot at  $t = 1$

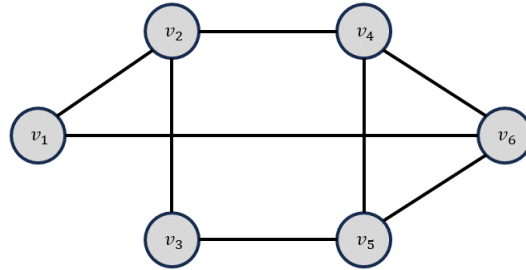


Figure 2: Snapshot at  $t = 2$

Vertex	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$
Value	19	13	21	18	22	14

Table 1: Ordinal Values of Vertices in the Graph



## 2. [Recommender System]

- (a) Compute the missing rating of these sports using **item-based** collaborative filtering (CF) where the rows denote the athletes and the columns denote the sports. When finding nearest neighbors, use Cosine similarity as your similarity measure (*Hint:  $\text{sim}(u, v) = \frac{u \cdot v}{\|u\| \cdot \|v\|}$* ).

	Basketball	Soccer	Tennis	Swimming
Ron	3	2	5	4
Ginny	2	?	4	3
George	1	4	2	3
Fred	3	4	3	5
$\bar{r}_i$		3.33		

Predict the missing rating by completing the following tasks (Rounded your answers to two decimal digits.):

- Calculate  $\bar{r}_i$  values for every sport (Soccer is given). Write down the final values in the table (last row).
- Calculate the similarity value between Soccer and Basketball (all others are provided in the following table).

	Basketball	Tennis	Swimming
Soccer		0.676	0.943

Write down your calculation in the following box.

- Identify the two nearest neighbors of Soccer.

1.	2.
----	----

- Calculate the Soccer rating for “Ginny”. *Hint:*  $r_{u,i} = \bar{r}_i + \frac{\sum_{j \in N(i)} \text{sim}(i,j)(r_{u,j} - \bar{r}_j)}{\sum_{j \in N(i)} \text{sim}(i,j)}$ . Assume that we consider 2 nearest neighbors for computing the rating.

- (b) Consider a set of six items  $I = i_1, i_2, i_3, i_4, i_5, i_6$  for which the predicted and true rankings are as follow:

	Predicted Rank	True Rank
$i_1$	1	1
$i_2$	2	4
$i_3$	3	3
$i_4$	4	6
$i_5$	5	2
$i_6$	6	5

Compute Kendall’s Tau:

Good Luck