CSE 3024: Web Mining **Slot: L39 + L40**

Online Submission Deadline: 13th April 2021

Web Structure Mining

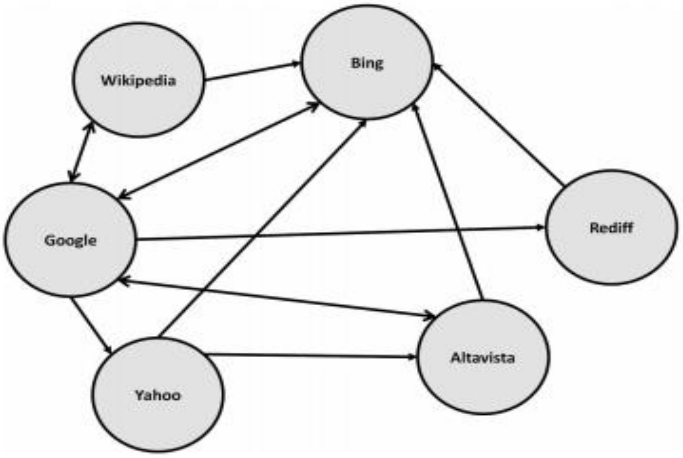
[2.5 x 4]

- **This assignment can be carried out individually or in a group of 2.**
- **In case of group assignment, please mention the name and registration number of both the members in the title page and both the members must upload the same copy.**
- **Upload your code and result as a single PDF file in VTOP [Mandatory] and MS Team Assignment [optional] on or before the deadline.**
- **No other form of submission will be acceptable.**
- **If you failed to upload in VTOP on or before the deadline, but successfully uploaded in MS Team Assignment, then 2 marks of penalty will be imposed on the secured marks.**
- **If you fail to upload your assignment in both VTOP and MS Team Assignment, then your assignment will not be evaluated and ZERO (0) mark will be awarded.**
- **File should contain**
  - **Question**
  - **Code**
  - **Result / Output screen**

---

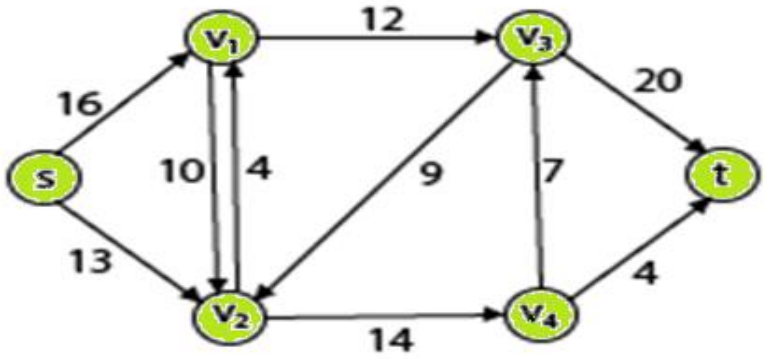
1. Write a python program to calculate the degree prestige, proximity prestige and rank prestige using a graph dataset given in the following link.
   http://snap.stanford.edu/data/wiki-Vote.txt.gz

2. Write a python program to show the implementation Page Rank algorithm for the following graph. Update and display the page rank for all the nodes after each iteration. Count and display the total number of iterations required to meet the stopping criteria. (**stopping criteria:- ε = 0.05, damping factor = 0.9)**



3. Write a python program to show the implementation HITS algorithm for the above graph and display the authority as well as hub score for all the nodes after each iteration. Count and display the total number of iterations required to meet the stopping criteria. (**stopping criteria:- ε = 0.05 for both hub and authority)**

4. Write a python program to calculate the maximum flow in the following graph using Ford-Fulkerson Max flow algorithm.



Display the augmented graph after each step. Display the value for each edge in the form of [flow / capacity]. Display the final result using the format given below using pandas.

Augmented Path	Bottleneck Capacity
S -> A -> B -> T	10