Hope Foundation's
Finolex Academy of Management and Technology
P60, P60-1, MIDC, Mirjole, Ratnagiri, Maharshtra, Pin 415639
Information Technology Department

Big Data Analytics

Date : 25th March 2020

Assignment II- Open Book Test

YEAR/SEM: *BE* VIII

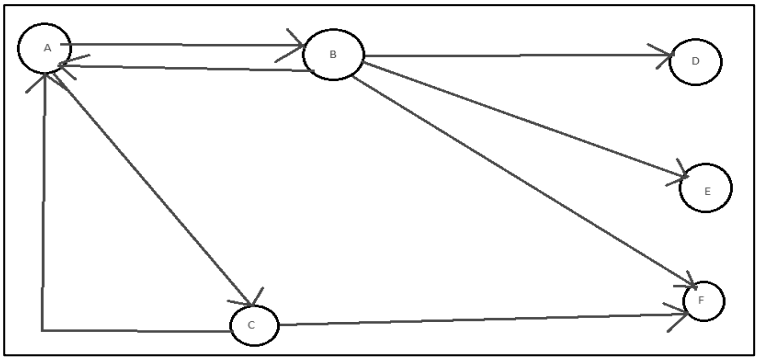
Name: Jawwad A Rahim Kazi **Roll No: 28**

Part I -Answer Following Questions.

1. What is a Data Stream ? Explain with example. (R)
2. What is Data Stream Mining ? Give four examples of Data Stream Mining. (R)
3. What is a Colloborative filtering system in view of Recommendation Systems? (R,A)
4. What is CQL , where it is used ?(R,A)
5. What are Hash Functions? Give their application and one example (R)
6. Define (a) Tendrils and (b) Tubes (c) SCC (d) Simrank (e) Pagerank (R)
7. Where do we see recommendation? List any Four. (R)
8. Draw Block diagram of typical DSMS. (R)
9. What is Random Surfer Concept? (R)
10. Define (a) Hubs (b) Authority pages (R)
11. What is Link Farm ?(R)
12. What is the use of Recommendation Systems? (R,A)
13. What is HITS algorithm? Explain differences between HITS and Pagerank. (R,A)

Part II - Answer Following Questions, Five (05) Marks per Question:

1. Explain Bloom's Filter with example. A bloom filter with m=1000 cells is used to store information about n=100 items, using k=4 hash functions. Calculate the false positive probability of this instance. Will the performance improve by increasing the number of hash functions from 4 to 5. Explain your answer. (A,AN,E)
2. Consider a Web Graph as given Below: (A,AN,E)
   Assume that Pagerank values for any page 'm' at iteration 0 is PR(m)=1 and teleportation factor for iteration is β =0.85. Perform the PageRank algorithm and determine the rank of every page at iteration 2.



Priyanka Bandagale

Asst Professor, IT