SVKM'S
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
Academic Year (2022-23)
YEAR IV / Semester VIII

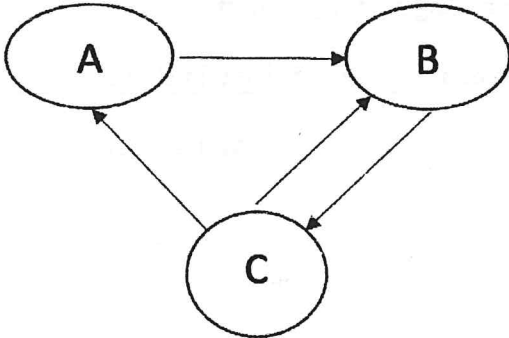Program: B. Tech in Computer Engineering   Max. Marks: 75
Subject/Course: Web Intelligence   Time: 9:00 a.m. to 12.00 p.m.
Date: 26/07/2023   Duration: 3 Hours

RE-EXAMINATION

Instructions: Candidates should read carefully the instructions printed on the question paper and on the cover page of the Answer Book, which is provided for their use.
(1) This question paper contains 02 page.
(2) All Questions are Compulsory.
(3) Answer to each new question is to be started on a fresh page.
(4) Figures in the brackets on the right indicate full marks.
(5) Assume suitable data wherever required, but justify it.
(6) Support your answers with neat-labelled diagrams, wherever necessary.

Question No.		Max. Marks
Q1 (a)	Compute Efficient PageRank with the damping factor d=0.85 for web. Initial Page Rank=1 Iteration=2 **OR** Explain Crawling with its types.	[10]
Q1 (b)	Short note on Web Spamming. **OR** Explain Inverted Index and Its Compression	[5]
Q2 (a)	Describe the process of wrapper induction	[10]
Q2 (b)	Explain how DOM trees are used in structured data extraction. **OR** Explain Instance-Based Wrapper Learning.	[05]

Q3 (a)	What are some pre-processing techniques used in schema matching? How are they used to improve the accuracy of schema matching?	[10]
Q3 (b)	Explain Integration of Web Query Interfaces. **OR** Explain Domain and Instance-Level Matching in short.	[05]
Q4 (a)	Explain Document Sentiment Classification in detail.	[10]
Q4 (b)	What are the Problems in Opinion Mining? **OR** What is opinion lexicon expansion and how is it used to improve the accuracy of opinion mining?	[05]
Q5 (a)	Describe the process of data collection and pre-processing in web usage mining.	[10]
Q5 (b)	Explain Query Log Mining in short. **OR** Explain Computational Advertising in short.	[05]

All the Best!