

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

Program: B. Tech in Computer Engineering
Subject/Course: Web Intelligence
Date: 26/5/2023

Max. Marks: 75
Time: 10.30 AM-1.30 PM
Duration: 3 Hours

REGULAR 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)	Describe the PageRank algorithm and how it is used to rank web pages in search engine results. OR Describe the basic components of a web crawler algorithm.	[10]
Q1 (b)	Short note on Web Spamming. OR Explain the concept of web mining and how it is useful for businesses and organizations.	[5]
Q2 (a)	Describe the process of wrapper generation and how it is used to extract structured data from web pages.	[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)	Discuss some common techniques used to combine similarities in schema matching. 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 spam and how it is detected in practice?	[05]
Q5 (a)	Draw and Explain process of data collection and pre-processing in web usage mining.	[10]
Q5 (b)	Explain Collaborative Filtering Using Association Rules. OR Explain Computational Advertising in short.	[05]

All the Best!