

SVKM
D. J. Sanghvi College of Engineering
Programme: B.Tech in Computer Engineering
Year: IV/Semester VIII (Exam Year: 2023-2024)

Subject: Web Intelligence

Date: 25 May 2024

Time: 09:00 am - 12:00 pm (03:00 Hrs.)

Max Marks: 75

FINAL EXAMINATION(2023-2024)

Instructions:

1. This question paper contains 2 pages
2. Answer to each new question to be started on a fresh page.
3. Figure in right hand side indicates full marks
4. All the Questions are Compulsory.

1. 15

a. 10

i. Distinguish between universal crawlers, focused crawlers, and topical crawlers in the context of web crawling strategies, and how do they differ in their approach to collecting web data? 10

----- OR -----

ii. What role do authority and hub nodes play in the context of the HITS algorithm, and how are they identified within a network of web pages? 10

b. 5

i. Explain the concept of web mining and how it is useful for businesses and organizations. 5

----- OR -----

ii. Write a Short note on Web Spamming. 5

2. 15

- a. 10
- i. Explain automating the process of extracting data from web pages using wrapper generation. 10

----- OR -----

- ii. Explain Instance-Based Wrapper Learning with advantages of using instance-based wrapper learning over manual rule-based approaches. 10
- b. Explain briefly Crawler Ethics. 5
3. 15
- a. What is linguistic schema matching in the context of data integration? 10
- b. What is a unified global query interface? 5
4. 15
- a. Explain Document Sentiment Classification in detail. 10
- b. 5
- i. Explain Opinion spam detection. 5

----- OR -----

- ii. How does opinion search differ from traditional information retrieval? 5
5. 15
- a. Explain the iterative calculation process of PageRank in detail? 10
- b. 5
- i. Explain the concept of sessionization in web usage mining. 5

----- OR -----

- ii. Write a short note on Recommender system. 5