

10/06/2025 BE IT SEM-VII C-SCHEME IRS QP CODE: 10082376

Time: 3 hours

Max. Marks: 80

N.B

1. Q.1 is compulsory
2. Attempt any three from the remaining five questions.
3. Assume suitable data, if required and state it clearly.

Q 1 Attempt all.

(20 marks)

- a. Recall objectives of Information Retrieval Process
- b. Differentiate between Information versus Data Retrieval.
- c. Identify the various task performed by web search engine.
- d. Importance of metadata and its types.

Q 2. Attempt all.

(20 marks)

- a. Show taxonomy of IR models and recall browsing models in detail.
- b. Consider a very small collection C that consists in the following three documents:
  - d1: "beautiful garden"
  - d2: "evening garden time"
  - d3: "garden time is beautiful"

Given the following query: "garden time", calculate the rank of each document using vector space retrieval model. (Use tf-idf vector for the query, and compute the score of each document in C relative to this query, using the cosine similarity measure.)

Q3. Attempt all.

(20 marks)

- a. Explain Boolean Models in detail with example.
- b. Explain Natural language modelling issues and its solution

Q.4. Attempt all.

(20 marks)

- a. What is the purpose of using keyword based query? Briefly explain any 3 types of keyword based queries.
- b. Apply Boyer Moore algorithm to construct bad match table and find the index of the given pattern for the string below with steps

**String: STUDENTS ARE SMART I LIKE SMART STUDENTS**

**Pattern: SMART**

Q 5. Attempt all.

(20 marks)

- a. Summarize various visualization techniques with respect to user interface design.
- b. Construct a Suffixes, suffix trie and suffix tree, suffix array and supra index for the above sample text.

**Roses are red. Red roses are beautiful. Many people like red roses.**

Q 6. Write short note on

(20 marks)

- a. Different types of Information Systems
- b. Different markup languages and its applications.

\*\*\*\*\*