

Exam.Code:1046  
Sub. Code: 35460

2056

**M.E. Computer Science and Engineering (Cyber Security)**  
**Second Semester**  
**CSN-8205: Information Retrieval**  
**(Common with CS-8304)**

**Time allowed: 3 Hours**

**Max. Marks: 50**

*NOTE: Attempt five questions in all, including Question No. 1 which is compulsory and selecting two questions from each Section.*

*x-x-x*

|   |   |                                                                                                                                                                   |    |
|---|---|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1 | a | Define Information Retrieval (IR) and explain its main objectives.                                                                                                | 2  |
|   | b | Differentiate between stemming and lemmatization.                                                                                                                 | 2  |
|   | c | What are stop words? Why are they removed during indexing?                                                                                                        | 2  |
|   | d | What is Naïve Bayes classification in text processing?                                                                                                            | 2  |
|   | e | What is the significance of cosine similarity measure in the vector space model?                                                                                  | 2  |
|   |   | <b>Section A</b>                                                                                                                                                  |    |
| 2 |   | Explain the process of text preprocessing in Information Retrieval. Discuss tokenization, stop word removal, stemming, and phrase queries with suitable examples. | 10 |
| 3 |   | Describe the construction of inverted indexes. Explain sort-based indexing, dynamic indexing, and positional indexes.                                             | 10 |
| 4 |   | Explain the Vector Space Model for Information Retrieval. Discuss term weighting, cosine similarity, and document ranking.                                        | 10 |
|   |   | <b>Section B</b>                                                                                                                                                  |    |
| 5 |   | Discuss efficient scoring techniques in IR systems. Explain components of an IR system and methods for computing document scores.                                 | 10 |
| 6 |   | Explain different text classification techniques used in IR systems. Discuss k-Nearest Neighbors classifies with example.                                         | 10 |
| 7 |   | Explain Web Information Retrieval. Discuss web crawling, web indexing, near-duplicate detection, and link analysis.                                               | 10 |

*x-x-x*