

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

Computer Science & Information Technology, VIII-Semester

Departmental Elective CSIT-802(C) Web & Information Retrieval

Course Objective:

This course aims at introducing the area of Information Retrieval and at examining the theoretical and practical issues involved in designing, implementing and evaluating Information Retrieval systems.

Course Outcomes: After the completion of this course, the students will be able to:

1. To identify basic theories and analysis tools as they apply to information retrieval.
2. To develop understanding of problems and potentials of current IR systems.
3. To learn and appreciate different retrieval algorithms and systems.
4. To apply various indexing, matching, organizing, and evaluating methods to IR problem.
5. To become aware of current experimental and theoretical IR research.

Unit-I

Introduction: Information versus data retrieval, the retrieval process, taxonomy of Information Retrieval Models.

Unit-II

Classic Information Retrieval Techniques: Boolean Model, Vector model, Probabilistic Model, comparison of classical models. Introduction to alternative algebraic models such as Latent semantic Indexing etc.

Unit-III

Keyword based Queries, User Relevance Feedback: Query Expansion and Rewriting, Document preprocessing and clustering, Indexing and Searching: Inverted Index construction, Introduction to Pattern matching.

Unit-IV

Web Search: Crawling and Indexes, Search Engine architectures, Link Analysis and ranking algorithms such as HITS and Page Rank, Meta searches, Performance Evaluation of search engines using various measures, Introduction to search engine optimization.

Unit-V

Introduction to online IR Systems, Digital Library searches and web Personalization.

Recommended Books:

1. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, "Modern Information Retrieval" Pearson Education
2. C. Manning, P. Raghvan and H. Schutze, "Introduction to Information Retrieval", Cambridge University Press.
3. Amy N. Langville and Carl D. Meyer, "Google's PageRank and Beyond: The Science of Search Engine Rankings", Princeton University Press
4. Pierre Baldi, Paolo Frasconi and Padhraic Smythe, "Modelling the internet and the web: Probabilistic methods and Algorithms", John Wiley

List of Experiments:

1. Students must experiment on various information retrieval systems like page rank etc