INFORMATION RETRIEVAL SYSTEMS

OBJECTIVES:
- To provide the foundation knowledge in information retrieval.
- To equip students with sound skills to solve computational search problems.
- To appreciate how to evaluate search engines.
- To appreciate the different applications of information retrieval techniques in the Internet or Web environment.
- To provide hands-on experience in building search engines and/or hands-on experience in evaluating search engines.

UNIT-I:
Introduction to Information Storage and Retrieval System: Introduction, Domain Analysis of IR systems and other types of Information Systems, IR System Evaluation.
Introduction to Data Structures and Algorithms related to Information Retrieval: Basic Concepts, Data structures, Algorithms

UNIT-II:
Inverted files: Introduction, Structures used in Inverted Files, Building Inverted file using a sorted array, Modifications to Basic Techniques.

UNIT-III:
Signature Files: Introduction, Concepts of Signature Files, Compression, Vertical Partitioning, Horizontal Partitioning.

UNIT-IV:
New Indices for Text: PAT Trees and PAT Arrays: Introduction, PAT Tree structure, algorithms on the PAT Trees, Building PAT trees as PATRICA Trees, PAT representation as arrays.

UNIT-V:
Stemming Algorithms: Introduction, Types of Stemming Algorithms, Experimental Evaluations of Stemming to Compress Inverted Files

UNIT-VI:
Thesaurus Construction: Introduction, Features of Thesauri, Thesaurus Construction, Thesaurus construction from Texts, Merging existing Thesauri

OUTCOMES:

- Identify basic theories in information retrieval systems
- Identify the analysis tools as they apply to information retrieval systems
- Understands the problems solved in current IR systems
- Describes the advantages of current IR systems
- Understand the difficulty of representing and retrieving documents.
- Understand the latest technologies for linking, describing and searching the web.

TEXT BOOKS:
1. Frakes, W.B., Ricardo Baeza-Yates: Information Retrieval Data Structures and Algorithms, Prentice Hall, 1992.
2. Modern Information Retrieval by Yates Pearson Education.
3. Information Storage & Retrieval by Robert Korfhage – John Wiley & Sons.

REFERENCE BOOKS:
1. Kowalski, Gerald, Mark T Maybury: Information Retrieval Systems: Theory and Implementation, Kluwer Academic Press, 1997.
2. Information retrieval Algorithms and Heuristics, 2ed, Springer