



IV Year-I Semester		T	P	C
		4	0	3
INFORMATION RETRIEVAL SYSTEMS (RT41056)				

Prerequisite Course:

Data Structures and Algorithms

Course Description and 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.

Course Outcomes:

Upon completion of the course, the student will be able to achieve the following outcomes.

Cos	Course Outcomes	POs
1	Identify basic theories in information retrieval system	5
2	Identify the analysis tools as they apply to information retrieval systems	6
3	Understand the problems in current IR systems	3
4	Analyze complex problems in groups and develop group work.	6
5	Understanding difficulty of representing and retrieving documents	8
6	Understanding the latest technologies like linking, describing and searching the web	10

Syllabus:

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

TEXT BOOK:

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.

REFERENCES:

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