

Course Code	Course Title	Credits	Lectures /Week
USCS6032	Information Retrieval	2	3
<p>About the Course: This course introduces the principles, techniques, and technologies underlying information retrieval (IR) systems. Students will explore the fundamental concepts of document indexing, storage, and retrieval, as well as advanced topics such as retrieval models, text categorization, web information retrieval, and evaluation techniques. Through a combination of theoretical study, practical exercises, and reference to industry-standard books, students will gain a solid foundation in the field of information retrieval.</p>			
<p>Course Objectives:</p> <ul style="list-style-type: none"> • To understand the fundamental principles and components of information retrieval systems. • To explore various techniques for document indexing, storage, and retrieval. • To analyze and compare different retrieval models and understand their strengths and limitations. • To gain practical experience in implementing and evaluating information retrieval systems. • To explore advanced topics in information retrieval, such as web information retrieval and machine learning techniques. 			
<p>Learning Outcomes:</p> <p>After successful completion of this course, students would be able to</p> <ul style="list-style-type: none"> • Explain the key components and principles of information retrieval systems. • Apply indexing, storage, and retrieval techniques to efficiently retrieve relevant documents. • Compare and contrast different retrieval models and select appropriate models for specific search scenarios. • Develop practical skills in implementing and evaluating information retrieval systems. • Demonstrate an understanding of advanced topics in information retrieval, including web search and machine learning techniques. 			
Unit	Topics	No of Lectures	
I	<p>Foundations of Information Retrieval</p> <p>Introduction to Information Retrieval (IR) systems: Definition and goals of information retrieval, Components of an IR system, Challenges and applications of IR</p> <p>Document Indexing, Storage, and Compression: Inverted index construction and compression techniques, Document representation and term weighting, Storage and retrieval of indexed documents,</p> <p>Retrieval Models: Boolean model: Boolean operators, query processing, Vector space model: TF-IDF, cosine similarity, query-document matching, Probabilistic model: Bayesian retrieval, relevance feedback</p> <p>Spelling Correction in IR Systems: Challenges of spelling errors in queries and documents, Edit distance and string similarity measures, Techniques for</p>	15	

	<p>spelling correction in IR systems</p> <p>Performance Evaluation: Evaluation metrics: precision, recall, F-measure, average precision, Test collections and relevance judgments, Experimental design and significance testing</p>	
II	<p>Advanced Topics in Information Retrieval</p> <p>Text Categorization and Filtering: Text classification algorithms: Naive Bayes, Support Vector Machines, Feature selection and dimensionality reduction, Applications of text categorization and filtering</p> <p>Text Clustering for Information Retrieval: Clustering techniques: K-means, hierarchical clustering, Evaluation of clustering results, Clustering for query expansion and result grouping</p> <p>Web Information Retrieval: Web search architecture and challenges, Crawling and indexing web pages, Link analysis and PageRank algorithm</p> <p>Learning to Rank: Algorithms and Techniques, Supervised learning for ranking: RankSVM, RankBoost, Pairwise and listwise learning to rank approaches Evaluation metrics for learning to rank</p> <p>Link Analysis and its Role in IR Systems: Web graph representation and link analysis algorithms, HITS and PageRank algorithms, Applications of link analysis in IR systems</p>	15
III	<p>Advanced Topics in Information Retrieval</p> <p>Crawling and Near-Duplicate Page Detection: Web page crawling techniques: breadth-first, depth-first, focused crawling, Near-duplicate page detection algorithms, Handling dynamic web content during crawling</p> <p>Advanced Topics in IR: Text Summarization: extractive and abstractive methods, Question Answering: approaches for finding precise answers, Recommender Systems: collaborative filtering, content-based filtering</p> <p>Cross-Lingual and Multilingual Retrieval: Challenges and techniques for cross-lingual retrieval, Machine translation for IR, Multilingual document representations and query translation, Evaluation Techniques for IR Systems</p> <p>User-based evaluation: user studies, surveys, Test collections and benchmarking, Online evaluation methods: A/B testing, interleaving experiments</p>	15
<p>Textbook(s):</p> <ol style="list-style-type: none"> 1. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, —Modern Information Retrieval: The Concepts and Technology behind Search, Second Edition, ACM Press Books 2. C. Manning, P. Raghavan, and H. Schütze, —Introduction to Information Retrieval, Cambridge University Press <p>Additional Reference(s):</p> <ol style="list-style-type: none"> 1. Ricci, F, Rokach, L. Shapira, B. Kantor, —Recommender Systems Handbook, First Edition. 2. Bruce Croft, Donald Metzler, and Trevor Strohman, Search Engines: Information Retrieval in Practice, Pearson Education. 3. Stefan Buttcher, Charlie Clarke, Gordon Cormack, Information Retrieval: Implementing and Evaluating Search Engines, MIT Press. 		