SCHEME OF COURSE WORK

Course Details:

Course Title	Information Retrieval Systems (Professional Elective-VI)		
Course Code	15CT1139	L T P C	:3003
Program:	B.Tech.		
Specialization:	Information Technology		
Semester	VII		
Prerequisites	Database Management Systems		
Courses to which it is a prerequisite	None		

Course Outcomes (COs):

1	Design pre-processing methods for effective information retrieval
2	Build tolerant information retrieval
3	Implement index compression process
4	Formulate textual information into vectors
5	Analyze ranked and unranked search results

Program Outcomes (POs):

A graduate of Information Technology will be able to

1	Apply the knowledge of mathematics, science, engineering fundamentals and principles of Information Technology to solve problems in different domains.
2	Analyze a problem, identify and formulate the computing requirements appropriate to its solution.
3	Design & develop software applications that meet the desired specifications within the realistic constraints to serve the needs of the society.
4	Design and conduct experiments, as well as to analyze and interpret data
5	Use appropriate techniques & tools to solve engineering problems.
6	Understand the impact of information technology on environment and the evolution and importance of green computing
7	Analyze the local and global impact of computing on individual as well as on society and incorporate the results in to engineering practice.
8	Demonstrate professional ethical practices and social responsibilities in global and societal contexts.
9	Function effectively as an individual, and as a member or leader in diverse and multidisciplinary teams.
10	Communicate effectively with the engineering community and with society at large.
11	Understand engineering and management principles and apply these to one's Own work, as a member and leader in a team, to manage projects.
12	Recognize the need for updating the knowledge in the chosen field and imbibing learning to learn skills.

Course Outcome Versus **Program Outcomes:**

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO-1			2		3								2		
CO-2			3	3									2		
CO-3		3											2		
CO-4				3									2		
CO-5			2		2								2		

S - Strongly correlated, *M* - *Moderately* correlated, *Blank* - *No correlation*

Assessment Methods:	Seminar / Mid-Test / End Exam

Teaching-Learning and Evaluation

Week	TOPIC / CONTENTS	Course Outcomes	Sample questions	TEACHING-LEARNING STRATEGY	Assessment Method & Schedule
1	**UNIT-I: Boolean Retrieval:** An example information retrieval problem, A first take at building an inverted index,	CO1	Define the term Boolean Query.	▫ Lecture	Mid –Test 1 Assignment-1 Quiz-1
2	Processing Boolean queries, The extended Boolean model versus ranked retrieval.	CO1	What are the advantages of extended Boolean model over traditional model?	▫ Lecture	Mid –Test 1 Assignment-1 Quiz-1
3	**The Term vocabulary and postings lists :** Document delineation and character sequence decoding, Obtaining the character sequence in a document, Choosing a document unit, Determining the vocabulary of terms.	CO1	How to define a document unit?	▫ Lecture	Mid –Test 1 Assignment-1 Quiz-1
5	Stemming and lemmatization, Faster postings list intersection via skip pointers,	CO1	How stemming is helpful in information retrieval?	▫ Lecture / ▫ Demonstration	Mid –Test 1 Assignment-1 Quiz-1
6	Positional postings and phrase queries , Biword indexes , Positional indexes , Combination schemes	CO1	What is meant by Biword indexes?	▫ Lecture	Mid –Test 1 Assignment-1 Quiz-1
7	**UNIT II: Dictionaries and tolerant retrieval :** Search structures for dictionaries, Wildcard queries, General wildcard queries, k-gram indexes for wildcard queries, Spelling correction.	CO2	What are the different search statements for dictionaries	▫ Lecture	Mid –Test 1 Assignment-1 Quiz-1
8	Implementing spelling correction, Forms of spelling correction , Edit distance , k-gram indexes for spelling correction, Context sensitive spelling correction , Phonetic correction.	CO2	What is isolated-term correction and context-sensitive correction?	▫ Lecture	Mid –Test 1 Assignment-1 Quiz-1
9	**Index construction :** Hardware basics , Blocked sort-based indexing, Single-pass in-memory indexing , Distributed indexing , Dynamic indexing , Other types of indexes	CO2	What are the hardware issues associated with indexing?	▫ Lecture / ▫ Demonstration	Mid –Test 2 Assignment-2 Quiz-2
10	**Mid-Test 1**	CO1 & CO2			Mid-Test 1 (Week 9)
11	**UNIT –III: Index compression:** Statistical properties of terms in information retrieval, Heaps' law: Estimating the number of terms , Zipf's law:Modeling the distribution of terms ,	CO3	What is Zipf's law?	▫ Lecture ▫ Problem solving	Mid –Test 2 Assignment-2 Quiz-2
12	Dictionary compression, Dictionary as a string , Blocked storage , Postings file compression, Variable byte codes , ã codes	CO3	What is the significance of variable bye encoding?	▫ Lecture	Mid –Test 2 Assignment-2 Quiz-2

13	**Scoring, term weighting :** Parametric and zone indexes, Weighted zone scoring, Learning weights, The optimal weight g, Term frequency and weighting , Inverse document frequency, Tf-idf weighting.	CO3	Write briefly about inverse document frequency	▫ Lecture / ▫ Demonstration	Mid –Test 2 Assignment-2 Quiz-2
14	**UNIT –IV: The vector space model:** The vector space model for scoring, Dot products , Queries as vectors , Computing vector scores, Variant tf-idf functions ,	CO4	What are the weighting functions/schemes used in variant idf functions	▫ Lecture	Mid –Test 2 Assignment-2 Quiz-2
15	Sublinear tf scaling, Maximum tf normalization, Document and query weighting schemes, Pivoted normalized document length.	CO4	List some of the query weighting schemes.	▫ Lecture	Mid –Test 2 Assignment-2 Quiz-2
16	**UNIT –V : Evaluation in information retrieval :** Information retrieval system evaluation, Standard test collections , Evaluation of unranked retrieval sets , Evaluation of ranked retrieval results,	CO5	Discuss about Evaluation of unranked retrieval sets.	▫ Lecture	Mid –Test 2 Assignment-2 Quiz-2
17	Assessing relevance , Critiques and justifications of the concept of Relevance, A broader perspective: System quality and user utility, System issues, User utility, Refining a deployed system, Results snippets	CO5	How do you measure the relevance of retrieved results? What is a snippet?	▫ Lecture / ▫ Demonstration	Mid –Test 2 Assignment-2 Quiz-2
18	Mid-Test 2	CO3, CO4, CO5			Mid-Test 2
19/20	END EXAM				