



INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

Dundigal - 500 043, Hyderabad, Telangana

COURSE CONTENT

INFORMATION RETRIEVAL SYSTEM								
Professional Elective: CSE (AI&ML)								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
ACAD08	Elective	L	T	P	C	CIA	SEE	Total
		3	0	0	3	40	60	100
Contact Classes: 48	Tutorial Classes: NIL	Practical Classes: Nil			Total Classes: 48			
Prerequisite: Linear Algebra and Calculus, Probability and Statistics								

I. COURSE OVERVIEW:

The main objective of this course is to present the scientific support in the field of information search and retrieval. This course explores the fundamental relationship between information retrieval, hypermedia architectures, and semantic models, thus deploying and testing several important retrieval models such as vector space, Boolean and query expansion. It discusses implementation and evaluation issues of new algorithms like clustering, pattern searching, and stemming with advanced data/file structures, indirectly facilitating a platform to implement comprehensive catalogue of information search tools while designing an e-commerce web site.

II. COURSES OBJECTIVES:

The students will try to learn:

- I. Methods for creating indexes to organize and speed up the retrieval process.
- II. Techniques for organizing and categorizing documents through clustering and classification methods
- III. The data/file structures necessary to design and implement information retrieval (IR) systems.

III. COURSE OUTCOMES:

At the end of the course, students should be able to:

- CO1 Outline genesis and diversity of information retrieval situations for text and hyper media.
- CO2 Interpret hands-on experience store, and retrieve information from World Wide Web using semantic approaches.
- CO3 Demonstrate the usage of different data/file structures in building computational Search engines.
- CO4 Experiment with ranked retrieval of a very large number of documents with hyperlinks between them for handling large datasets.
- CO5 Classify information visualization technologies like cognitive and perception in the Internet or Web search engine for personalized search features.
- CO6 Make use of the performance related to information retrieval using advanced Techniques and filtering over multimedia.

IV. COURSE CONTENT:

MODULE – I: INTRODUCTION TO INFORMATION RETRIEVAL SYSTEMS (09)

Definition and models of information retrieval, Probabilistic retrieval strategies: Simple term weights, Non binary independence model, Language models, Relationship to Database Management Systems. Information Retrieval System Capabilities: Search Capabilities, Browse Capabilities

MODULE – II: LEARNING CATALOGUING AND INDEXING (09)

Introduction of Indexing, Indexing Process, Automatic Indexing, Information Extraction. Retrieval Utilities: Relevance feedback, clustering, N-grams, Regression analysis, Thesauri.

MODULE – III: AUTOMATIC INDEXING (09)

Classes of Automatic Indexing, Statistical Indexing, Natural Language, Ranking algorithms, Hypertext Linkages.

Document and Term Clustering: Introduction to Clustering, Thesaurus Generation, Item Clustering, Hierarchy of Clusters.

MODULE – IV: USER SEARCH TECHNIQUES (09)

Search Statements and Binding, Similarity Measures and Ranking, Relevance Feedback, Weighted Searches of Boolean Systems, Searching the INTERNET and Hypertext. Efficiency: Inverted Index, Query processing, Signature files, Duplicate document detection.

MODULE – V: TEXT SEARCH ALGORITHMS (09)

Introduction to Text Search Techniques, Information retrieval as relational application, Semi Structured search using a relational schema.

Multimedia Information Retrieval: Spoken Language Audio Retrieval, Video Retrieval, Imagery Retrieval. Distributed Information Retrieval: A theoretical Model of Distributed retrieval, web search

V. TEXT BOOKS:

1. Gerald J. Kowalski, Mark T. Maybury, Springer, “Information Storage and Retrieval Systems – Theory and Implementation”, 2nd Edition, 2004.
2. David A. Grossman, Ophir Frieder, “Information Retrieval – Algorithms and Heuristics”, Springer, 2nd Edition (Distributed by Universal Press), 2004

VI. REFERENCE BOOKS:

1. Bharathi Kumari, “Engineering Chemistry”, VGS Book Links, 10th Edition, 2018.
2. B.Siva Shankar, “Engineering Chemistry”, Tata McGraw Hill Publishing Limited, 3rd Edition, 2015.
3. S.S.Dara, Mukkanti, “Text of Engineering Chemistry”, S.Chand & Co, New Delhi, 12th Edition, 2006

VII. ELECTRONICS RESOURCES:

1. <https://docs.oracle.com/tutorial/nlp/concepts/>
2. <https://www.w3schools.com/nlp/>
3. <https://www.edx.org/learn/sequence-models/>
4. <https://www.geeksforgeeks.org/nlp>

VIII. MATERIALS ONLINE

1. Course template
2. Tutorial question bank
3. Tech talk topics
4. Open-ended experiments
5. Definitions and terminology
6. Assignments

7. Model question paper – I
8. Model question paper – II
9. Lecture notes
10. PowerPoint presentation
11. E-Learning Readiness Videos (ELRV)