



Indian Institute of Information Technology, Allahabad

Department of Information Technology

Syllabus

Name of the Course: **Information Retrieval**

Course Instructor- **Prof. U.S. Tiwary**

Under Graduate: **IINR730E**

Post Graduate: **IINR340E**

1. Objective of the course:
 - a. To provide an overview of Information Retrieval.
 - b. To introduce students about insights of the several topics of Information retrieval such as – Boolean retrieval model, Vector space model, Latent semantic indexing, XML and Image retrieval model.
 - c. To provide comprehensive details about various Evaluation methods.
 - d. To provide implementational insight about the topics covered in the course.

2. Outcome of the course:
 - a. Students will get the understanding different Information retrieval model.
 - b. Students will get to know about evaluation methods of the information retrieval model.
 - c. Students will get to know the challenges associated with each topic.

3. Course Plan: As per the below format only

Unit	Title	Topics for Coverage
Unit 1	<u>Introduction to Information retrieval</u>	Information retrieval process, Indexing, Information retrieval model, Boolean retrieval model
	<u>Dictionary and Postings</u>	Tokenization, Stop words, Stemming, Inverted index, Skip pointers, Phrase queries

Unit 2	<u>Tolerant Retrieval</u>	Wild card queries, Permuterm index, Bigram index, Spelling correction, Edit distance, Jaccard coefficient, Soundex
	<u>Term Weighting and Vector Space Model</u>	Wild card queries, Permuterm index, Bigram index, Spelling correction, Edit distance, Jaccard coefficient, Soundex
Unit 3	<u>Evaluation</u>	Precision, Recall, F-measure, E-measure, Normalized recall, Evaluation problems
	<u>Latent Semantic Indexing</u>	Eigen vectors, Singular value decomposition, Low-rank approximation, Problems with Lexical Semantics
Unit 4	<u>Query Expansion</u>	Relevance feedback, Rocchio algorithm, Probabilistic relevance feedback, Query Expansion and its types, Query drift
	<u>Probabilistic Information Retrieval</u>	Probabilistic relevance feedback, Probability ranking principle, Binary Independence Model, Bayesian network for text retrieval
Unit 5	<u>XML Indexing and Search</u>	Data vs. Text-centric XML, Text-Centric XML retrieval, Structural terms
	<u>Content Based Image Retrieval</u>	Introduction to content Based Image retrieval, Challenges in Image retrieval, Image representation, Indexing and retrieving images, Relevance feedback
Unit 6	Projects	

Books:

1. **Introduction to Information Retrieval** by Christopher D. Manning
2. **Natural Language Processing And Information Retrieval** by *Tanveer Siddiqui and U. S. Tiwary*