IT702 Information Retrieval Systems [3 - 2 4]

Syllabus:

Overview: Overview of IR Systems and Architecture of information retrieval systems.

Document Representation: Statistical Characteristics of Text, Basic Query Processing. Data Structure and File Organization for IR.

Retrieval Models: Similarity Measures and Ranking, Boolean Matching, Vector Space Models, Probabilistic Models, Automatic Indexing and Indexing Models

Search and Filtering Techniques Relevance Feedback, User Profiles, Collaborative Filtering Automatic classification, Document and Term Clustering, Document Categorization Heuristic classification

Machine Learning and other Techniques in IR

Naive Bayes Methods, Support Vector Machines, Neural Networks, Genetic Algorithms, Symbolic Learning Indexing and storage issues, Information visualization and usage pattern analysis IR Systems and the WWW, PageRank and Hyperlink Analysis, Search Personalization N-Grams in Information Retrieval, Agent-based Information Retrieval

Multimedia and Multilingual IR

Models and Languages Cross-Language and Multilingual Information Retrieval, Retrieval from noisy documents Performance Evaluation of Information Retrieval Systems

Self-Study:

The self-study contents will be declared at the commencement of semester. Around 10% of the questions will be asked from self-study contents.

Laboratory Work:

Laboratory work will be based on above syllabus with minimum 10 experiments to be incorporated.

References:

1. D.A. Grossman, O. Frieder, Information Retrieval: Algorithms and Heuristics, Springer
2. W.B. Croft, J. Lafferty, Language Modeling for Information Retrieval, Springer
3. G. Kowalski, M.T. Maybury, Information Storage and Retrieval Systems, Springer
4. Grigoris Antoniou and Frank van Harmelen, A Semantic Web Primer, The MIT Press
5. B. Croft, D. Metzler, T. Strohman, Information Retrieval in Practice, Pearson Education