

CAP 6776 Information Retrieval

Credits: 3 credits

Textbook, Title, Author, and Year: Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze: Introduction to Information Retrieval, Cambridge University Press, July, 2008. ISBN: 9780521865715 Handouts and notes.

Reference Materials: Bruce Croft, Donald Metzler, Trevor Strohman: Search Engines: Information Retrieval in Practice. ISBN-10: 0136072240, ISBN-13: 9780136072249

Specific Course Information:

Catalog Description: This course teaches concepts, techniques, and popular tools and applications in information retrieval (IR), which aims to obtain relevant information from a collection of resources. The class will cover efficient text indexing, text processing, web search, and text mining. New applications will also be introduced.

Prerequisites: COP3530 Data Structures and Algorithm Analysis, or permission from the instructor.

Specific Goals for The Course: This course will provide students with both theory and applications of Information Retrieval. Students will gain basic to advanced knowledge and hands-on experience. At the end of the class, students should be able to master latest techniques of text indexing, web search, text mining and system evaluation including building index, calculating term weights and ranking scores, etc. Students will form teams and apply these techniques on real-world web data using IR tools.

Brief List of Topics to Be Covered:

Topics	Approx. 1.5 hr. Lecture
Indexing, term weighting, vector space model	4
Scoring and ranking in a search system	2
Useful text processing tools	2
System evaluation	2
Text clustering	4
Text classification	2
Text summarization	4
Tools and Applications	4
Other IR topics	4