



CS 564

INFORMATION RETRIEVAL SYSTEMS

UNIT-1

Introduction

Retrieval Strategies: Vector Space Model, Probabilistic Retrieval Strategies

Language Models: Simple Term Weights. Non Binary Independence Model

UNIT-II

Retrieval Utilities: Relevance Feedback, Clustering. N-grams, Regression Analysis. Thesauri

UNIT-III

Retrieval Utilities: Semantic Networks, Parsing

Cross-language Information Retrieval: Introduction, Crossing the Language Barrier

UNIT-IV

Efficiency : Inverted Index, Query Processing, Signature Files. Duplicate Document Detection

UNIT V

Integrating Structured Data and Text: A Historical Progression, Information Retrieval as a Relational Application. Semi-Structured Search using a Relational Schema.

Distributed Information Retrieval: A Theoretical Model of Distributed Retrieval, Web Search

Suggested Reading:

1. David A. Grossman, Ophir Frieder. "Information Retrieval - Algorithms and Heuristics" Springer, 2nd Edition (Distributed by Universities Press), 2004.
2. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze, "An Introduction to Information Retrieval", Cambridge University Press, Cambridge, England, 2009.
3. Gerald J Kowalski, Mark T Maybury, "Information Storage and Retrieval Systems", Springer, 2000