Syllabus

✿ **Unit 1: Introduction**: Concepts and terminology of information retrieval systems, Information Retrieval Vs Information Extraction; Indexing: inverted files, encoding, Zipf's Law, compression, boolean queries;

✿ **Unit 2: Fundamental IR models and performance evaluation**: Boolean, Vector Space, probabilistic, TFIDF, Okapi, language modeling, latent semantic indexing, query processing and refinement techniques; Performance Evaluation: precision, recall, F-measure;

✿ **Unit 3: Classification and Clustering**: Classification: Rocchio, Naive Bayes, k-nearest neighbors, support vector machine; Clustering: partitioning methods, k-means clustering, hierarchical;

✿ **Unit 4: Advanced topics**: Searching, relevance feedback, ranking, query expansion.