ISR and LP III Oral sample question bank for theory and practical assignments

Set 1

Theory

1. Define Information Retrieval. How does it differ from Data Retrieval?

2. Explain Luhn's ideas in Automatic Text Analysis.

3. What is the purpose of Index Term Weighting?

4. Describe the structure and function of an IR System block diagram.

5. What is the Cluster Hypothesis?

Practical

1. Implement the **Conflation Algorithm** and explain its output.

2. Demonstrate retrieval of documents using **inverted file indexing**.

Set 2

Theory

1. Explain the relationship between Text Mining and Information Retrieval.

2. What is Probabilistic Indexing?

3. Compare Rocchio's Algorithm and Single Pass Algorithm in clustering.

4. Explain different matching coefficients used in IR.

5. Describe Boolean Search and Vector Space Model.

Practical

1. Implement **Single-pass algorithm** for clustering of 4–5 text files.

2. Write a Python program to calculate **Precision and Recall** for a given query.

Set 3

Theory

1. What are the main components of the IR system architecture?

2. Explain Signature files and Scatter Storage methods for indexing.

3. Define Query Languages and Types of Queries.

4. Discuss the **Probabilistic Model** of IR.

5. What are the advantages of using cluster-based retrieval?

Practical

1. Demonstrate document retrieval using **inverted files**.

2. Compute **Harmonic Mean (F-measure)** and **E-measure** for given data.

Set 4

Theory

1. Define and differentiate Precision, Recall, and F-Score.

2. What is **Mean Reciprocal Rank (MRR)** and **NDCG**?

3. Explain visualization in Information Systems.

4. Discuss **User Relevance Judgment** and **Interface Support** in IR.

5. What are user-oriented measures in IR evaluation?

Practical

1. Write a Python program to compute **Precision**, **Recall**, and **F-measure**.

2. Explain the role of visualization in query processing (theoretical explanation).

Set 5

Theory

1. Define Distributed Information Retrieval and its importance.

2. Explain Collection Partitioning and Source Selection.

3. What is Multimedia IR?

4. Discuss One-Dimensional Time Series and Two-Dimensional Color Image IR.

5. Explain Automatic Feature Extraction techniques.

Practical

1. Implement **feature extraction** for a 2D color image and plot its **color histogram**.

2. Discuss trends and research issues in Multimedia IR.

Set 6

Theory

1. What are the main challenges in Web Searching?

2. Explain the working of a **Search Engine** with Centralized Architecture.

3. What is Web Crawling?

4. Define Meta-searchers and Hyperlink-based searching.

5. Discuss Python tools used for **Web Scraping**.

Practical

1. Build a simple **Web Crawler** to fetch product information and links.

2. Write a program to fetch the **live weather report** of a given city.

Set 7

Theory

1. Explain the concept and structure of XML documents.

2. What are the challenges in XML retrieval?

3. Discuss **Vector Space Model** for XML Retrieval.

4. Compare Text-Centric and Data-Centric XML Retrieval.

5. Explain the importance of evaluation in XML Retrieval.

Practical

1. Write a case study on a **Recommender System** (Product, Doctor, or Music).

2. Demonstrate **Collaborative Filtering** and **Content-Based Filtering** approaches.

Set 8

Theory

1. Explain the concept of a **Recommendation System** and its applications.

2. Differentiate between Collaborative and Content-Based Recommendations.

3. What is the **Semantic Web**?

4. Discuss trends and research issues in modern IR systems.

5. How does Agent-based IR differ from traditional models?

Practical

1. Implement a simple **Recommendation System** for product or movie data.

2. Demonstrate data extraction using **BeautifulSoup** in Python.