Information Retrieval System

Sample Questions

Module 1:

1. State and explain the metrics used for evaluating performance of the IR system.
2. State and explain components of search engine and browser.

Module 2:

1. Write a formal definition of the IR model.
2. Explain the formal characteristics of IR models.
3. Explain taxonomy of IR models.
4. Explain the working of boolean model/vector-space model / probabilistic model /Extended boolean model / fuzzy set model with example
5. State Advantages and disadvantages of each type of model.
6. What are the different approaches for ranking the documents? State advantages and disadvantages of each approach.
7. How to calculate cosine similarity between query vector and document vector?
8. Explain the different types of browsing.
9. Compare boolean and extended boolean model.
10. Compare the Boolean model with the Vector model.
11. Explain structured retrieval model based on non overlapping lists .
12. Explain structured retrieval model based on proximal nodes. .

Module 3:

1. What is the purpose of keyword based query? Explain the types of keyword based queries
2. Explain the various types of patterns in IR.
3. Explain the three types of structures in structural queries.
4. Explain the different samples of hierarchical models.
5. Write a short note on Query Protocols.
6. What is query modification ? What is the purpose of query modification?
7. Explain query expansion methods.
8. Explain how term reweighting is done in a vector-space model with examples.
9. Explain term reweighing process in a probabilistic model with example.
10. What are the advantages and disadvantages of user relevance feedback techniques?
11. What do you mean by residual collection in the IRS? Why to use residual collection in the evaluation of relevance feedback strategies?
12. Explain the key challenges in multimedia information retrieval?
13. State and explain the different approaches and models used in multimedia IR
14. Explain data modeling techniques in multimedia IR models.
15. Write a short note on MULTOS data model.
16. Write SN on User relevance feedback

Module 4:

1. Explain the logical view of the document.
2. Write short notes on following
   1. SGML
   2. HyTime
   3. Multimedia data formats
   4. Types of Metadata.
   5. HTML Vs. SGML
3. Explain in detail the steps involved in document preprocessing.. Discuss any one application for the same.
4. What is thesaurus? Explain the components of thesaurus.Explain use of thesaurus in IR.
5. Explain Local Document clustering / Explain how Query Expansion is done through Local Clustering.
6. Explain Global Document clustering.

Module 5:

1. Short note on sequential searching
2. Write short note on Information retrieval in digital libraries
3. Describe Metasearchers and its merits with example
4. Describe the process of creating Inverted index with an example. How can this process be optimized using block addressing?
5. Describe sequential searching. Explain any one algorithm for the same.
6. Write short notes on Multimedia indexing approach, interface support search process
7. Describe concept of text search engine.
8. Explain inverted file indexing with suitable examples.
9. Explain Search engine architecture.
10. What are the different parameters to judge quality of information
11. SN on scalar and metrics clusters
12. What is the query specification? Describe direct manipulation and natural language in the context of boolean query formation
13. SN on Web information retrieval system
14. SN-automatic feature extraction
15. SN- A Generic Multimedia indexing approach
16. SN-statistical text compression and dictionary based text compression

**Note: For module 5 refer chapter 8 and 10 of the Text Book-modern-information-retrieval-Ricardo Baeza-Yates**

**(Soft copy available in the shared IRS folder)**

Module 6:

1. List and explain User interface requirements (good and bad interface).
2. Explain basic concepts and steps of audio information retrieval.
3. Explain basic concepts and steps of video information retrieval.