2 Mark Questions

UNIT I
1. What leads to information explosion?
2. List few challenges in IR systems.
3. Give a recovery technique that can be used for very large archives.
4. Which information technique is quicker? Why?

UNIT II
1. Difference between case based and rule based reasoning
2. What are the different statistical methods in clustering
3. What do you mean by market-basket analysis.
4. Write short notes on theharus based approach.
5. How is search accomplished in genetic algorithm.
6. What are the parser free techniques for NLP.

UNIT III
1. Which searching technique is used to search in sorted array implementation of inverted files? Why?
   Ans: Binary Search
2. List the advantages of fast inversion algorithm.
3. What do you mean by false drop probability?
4. What is the advantage of VBC over BC?
5. Differentiate proximity searching and range searching.
6. What are the two types of division of node in Time-split B Tree?
   Ans: Time splits and key splits
7. Use of checkpointing.
8. Lexical analyzer generator.
9. What is a negative directory?

UNIT IV
1. Break the term ‘statistics’ into unique diagrams using n-gram stemming algorithm.
2. What are signature processors.
3. Differentiate hierarchical and non-hierarchical clustering methods.
4. Write notes on parallel inverted files.
5. What is the need of clustering algorithm.

UNIT V
1. Analyse the retrieval in diet expert system.
2. How can speech data retrieved?
3. Write short notes on information retrieval in NLP.
4. Differentiate software agent and personal agent.
5. Discuss the ways of summarizing documents.
6. How can data be retrieved and translated.

16 Mark Questions

UNIT I

1. Explain the historical development of Information Systems. Discuss the sophistication in technology in detail.
2. Analyze the challenges in IR system and give your suggestion to overcome that.
3. What is the need for self recovery? Explain how it can be implemented for very large archives.
4. Explain various retrieval multimedia documents.

UNIT II

1. How can clustering classified using statistical techniques.? Describe in detail.
2. Write short notes on
   a. Rough sets and fuzzy logic
   b. Neural networks
3. Differentiate case based reasoning and rule based reasoning with suitable applications.
4. What is the need for Market-Basket analysis. Discuss its advantages and disadvantages.

UNIT III

1. In what way is the signature approach advantageous over other text retrieval methods? Classify the signature based methods.
2. How is a Patricia tree constructed? Construct a Patricia tree for 01100100010111. Discuss the algorithms on PAT tree.
3. What is a use of PAT tree? How can PAT trees be represented as arrays?
4. List the technical issues for file system. Explain WOB T with three nodes having C,D, F, G and H as records.
5. Discuss the procedure to implement a lexical analyzer.
6. What is a stop list? Give some examples of stop words. How can it be used along with lexical analyzer.

UNIT IV

1. Explain Boyer-Moore algorithm in searching
2. Discuss Shift-OR algorithm and KARP-RABIN algorithm.
3. How are Boolean expressions represented in sets? Give the different operations on sets.
5. Discuss the different types of stemming algorithms in detail.

UNIT V

1. Take diet expert system as a case study and give all the data required for storing. Explain various methods to retrieve data efficiently.
2. How are speech data stored? How can the speech data retrieved? Discuss in detail.
3. Analyze the techniques used in Translation, question answering and summarizing.
4. Compare the data storage and retrieval process in software agent and personal agent.
5. List the different implementation issues in storage architecture. What solution can be applied to solve it.