BRAINWARE UNIVERSITY

Term End Examination 2023-2024
Programme – B.Tech.(CSE)-DS-2021
Course Name – Intelligent Database System
Course Code - PEC-CSD602A
(Semester VI)

Full Marks : 60 **Time : 2:30 Hours**

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

Group-A

(Multiple Choice Type Question) 1 x 15=15

1. *Choose the correct alternative from the following :*

(i) Select the technology that is NOT typically associated with an Intelligent Database System

   a) Machine Learning   b) Natural Language Processing
   c) Blockchain   d) Predictive Analytics

(ii) State the purpose of Natural Language Processing (NLP) in an Intelligent Database System

   a) Data encryption   b) Query optimization
   c) Understanding and responding to human language   d) Backup and recovery

(iii) State the way Machine Learning contribute to an Intelligent Database System

   a) Improves data storage efficiency   b) Enables real-time data analysis
   c) Allows the system to learn from data and make predictions   d) Enhances data encryption techniques

(iv) Choose the purpose of document delineation in information retrieval systems

   a) To determine the relevance of documents   b) To identify the structure of documents
   c) To rank documents based on relevance   d) To analyze the user query

(v) Choose the action word that is associated with character sequence decoding

   a) Evaluate   b) Retrieve
   c) Decode   d) Remember

(vi) List the significance of determining the vocabulary of terms in information retrieval

   a) It helps in analyzing user behavior   b) It facilitates faster search operations
   c) It improves relevance ranking   d) It enhances document delineation

(vii) State the primary function of character sequence decoding in information retrieval systems

   a) To identify relevant documents   b) To interpret user queries
   c) To organize document collections d) To process textual data

164- 153

(viii) Choose the primary objective of faster postings list intersection using skip pointers

    a) To increase the size of the index

    b) To decrease the efficiency of search operations

    c) To reduce the time complexity of intersection

    d) To prioritize certain documents over others

(ix) State the role do relations and correlations of statistical properties play in information retrieval

    a) Identifying document trends

    b) Enhancing search relevance

    c) Predicting user behavior

    d) Analyzing historical data

(x) Select which of the following is NOT a statistical property commonly analyzed in information retrieval?

    a) Term frequency

    b) Document length

    c) User location

    d) Inverse document frequency

(xi) Choose from the following the significance of term frequency in information retrieval

    a) Indicates document relevance

    b) Measures how often a term appears in a document

    c) Predicts user preferences

    d) Analyzes historical trends

(xii) Choose the type of indexing technique organizes documents based on predefined parameters such as author, date, or category

    a) Parametric indexing

    b) Zone indexing

    c) Inverted indexing

    d) Sequential indexing

(xiii) Choose the primary purpose of zone indexing in document retrieval systems

    a) To improve query processing speed

    b) To organize documents based on content similarity

    c) To partition documents into distinct zones for efficient indexing

    d) To facilitate user-defined indexing parameters

(xiv) Select which of the following characteristics is typically associated with the structure of the World Wide Web

    a) Centralized organization

    b) Linear hierarchy

    c) Decentralized nature

    d) Fixed content distribution

(xv) Choose from the following which factors affects the size of a search engine's index

    a) Internet connection speed

    b) Average web page size

    c) Number of search queries per day

    d) Web server hardware configuration

Group-B
(Short Answer Type Questions)
          3 x 5=15

2. Define the format of Intelligent Document   (3)
3. Explain the correlation between term frequency (TF) and inverse document frequency (IDF). (3)
4. Describe in brief the term Zone Indexes   (3)
5. Describe the primary goal of Information Retrieval (IR)   (3)
6. Analyze the primary purpose of advertising in the economic model   (3)

OR

Analyze the importance of index distribution in search engine architecture   (3)

Group-C
(Long Answer Type Questions)
          5 x 6=30

7. Illustrate the concepts of query processing, relevance ranking in IR system   (5)
8. Discuss various techniques used for dictionary compression in information retrieval systems, highlighting their advantages and limitations.   (5)

9. Explain the importance of postings compression in reducing the storage requirements of inverted indexes. Provide examples of compression techniques used in practice. (5)
10. Analyze how Faster Postings List Intersection Via Skip Pointers technique works and how it improves search efficiency: (5)
11. Describe the various types of Information Retrieval models employed in intelligent databases (5)
12. Analyze how do search engines estimate the relevance of indexed documents to user queries and discuss the key factors and techniques involved in relevance estimation. (5)

OR

Evaluate some challenges associated with maintaining the accuracy and freshness of search engine indexes and analyze how do search engines address these challenges (5)

****