ROYAL COLLEGE OF ARTS, SCIENCE & COMMERCE
TYBSc Computer Science Semester VI
2019-20
USCS604 Information Retrieval
Sample Questions

Marks:20

1. The _____ is a wild card that represents one or more characters   [1M]

    A. question mark
    B. asterisk
    C. exclamation mark
    D. dollar sign

2. A _____is a software program that traverses web pages, downloads them for indexing,

    and follows the hyperlinks.   [1M]

    A. Search Engine
    B. Web crawler
    C. Websites
    D. URL

3.Many pages are part of hidden web or invisible web is also called as_____.   [1M]

    A. deep web
    B. soft web
    C. Websites
    D. URL

4.There can be_____ data nodes in Hadoop.   [1M]

    A. multiple
    B. Single
    C. Two
    D. Eight

5. HDFS in hadoop provides _____.   [1M]

    A. Centralized Storage
    B. Distributed Storage.
    C. Very les storage
    D. Fixed storage

6.. _____ can be used to store their past interaction with the search service, and the

    inference mechanism can then personalize their query results.   [1M]

    A. Cookies
    B. Personal Information
    C. Google ratings of page
    D. Web ratings of page

7. _____is an Internet advertising model in which advertisements appear along with

relevant search results from the web search engine. [1M]

    A. Paid Placement
    B. Pay Per Click
    C. Paid Websites
    D. Paid Link

8.SEO stands for _____. [1M]

    A. Search Engine Order
    B. Search Engine Organizer
    C. Search Engine Option
    D. Search Engine Optimization

9. CPC stands for_____. [1M]

    A. Cost Per Click
    B. Cost Per Customer
    C. Cost Per Coverage
    D. Cost Per Content

10. A web server communicates with a client (browser) using which protocol: [1M]

    A. HTML
    B. HTTP
    C. FTP
    D. Telnet

11.Map function breaks the input into _____. [1M]

    A. Bits
    B. Key/value pair
    C. Bytes
    D. Megabytes

12.A model of information retrieval in which we can pose any query in which search terms are combined with the operators AND, OR, and NOT [1M]

    A. Ad Hoc Retrieval
    B. Ranked Retrieval Model
    C. Boolean Information Model
    D. Proximity Query Model

13._____is a way of measuring the importance of website pages. [1M]

    A. Querying
    B. Page Rank
    C. Link Analysis
    D. HITS

14.A data structure that maps terms back to the parts of a document in which they occur

is called an [1M]

    A. Postings list
    B. Incidence Matrix
    C. Dictionary
    D. Inverted Index

15._____are indexed units in the incidence matrix. [1M]

    A. Terms
    B. Collection
    C. Information
    D. Data

16. The basic operation of a web browser is to pass a request to the web server. This request is an address for a web page and is known as the _____ [1M]

       A. UAL: Universal Address Locator
       B. HTML: Hypertext Markup Language
       C. URL: Universal Resource Locator
       D. HTTP: Hypertext transfer protocol

17. Map-Reduce program executes in three stages, they are _____ [1M]

       A. map stage, store stage, and reduce stage.
       B. map stage, re-match stage, and reduce stage.
       C. map stage, shuffle stage, and reduce stage.
       D. map stage, reduce stage, and again map stage.

18. The way to avoid linearly scanning the texts for each query is to _____the documents in advance. [1M]

       A. map
       B. IR
       C. index
       D. shuffle

19. Soundex Algorithm is used in _____. [1M]

       A. Spelling Correction
       B. Boolean Retrieval
       C. Phonetic Correction
       D. Tolerant Retrieval

20. Permuterm indexes are used for solving_____ [1M]

       A. Spelling Checking
       B. Boolean queries
       C. Phrase queries
       D. Wildcard queries