



BRAINWARE UNIVERSITY

Term End Examination 2023
Programme – MCA-2020/MCA-2021
Course Name – Big Data Analysis
Course Code - MCA305A
(Semester III)

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) Predict the reason to use an SQL database?
 - a) It's ACID-compliant.
 - b) It can easily store unstructured data. It's ACID-compliant.
 - c) It can enable development in the cloud
 - d) None of above
- (ii) Validate the correct option for the given blank: _____ is a online NoSQL developed by Cloudera.
 - a) Oozi
 - b) Hbase
 - c) Imphala
 - d) Hcatalog
- (iii) Select the correct option which is general-purpose computing model and runtime system for distributed data analytics.
 - a) Oozie
 - b) Mapreduce
 - c) Drill
 - d) None of the mentioned
- (iv) Match the suitable option for the blank: _____ part of the MapReduce is responsible for processing one or more chunks of data and producing the output results.
 - a) Maptask
 - b) Mapper
 - c) Task execution
 - d) All of the mentioned
- (v) Name the type of analytics that answers the question "What will happen?".
 - a) Descriptive analytics
 - b) Predictive analytics
 - c) Prescriptive analytics
 - d) None
- (vi) Select the suitable option for the blank: _____ function is responsible for consolidating the results produced by each of the Map() functions/tasks.
 - a) Reduce
 - b) Map
 - c) Reducer
 - d) All of the mentioned
- (vii) Report the number of iterations in apriori _____

- a) Increases with the size of the data
b) Decreases with the increase in size of the data
c) Increases with the size of the maximum frequent set
d) Decreases with increase in size of the maximum frequent set
- (viii) Express the suitable option for the blank: HDFS works in a _____ fashion
a) master-worker
b) master-slave
c) worker/slave
d) all of the mentioned
- (ix) Explain a dendrogram .
a) A hierarchical structure
b) A diagram structure
c) A graph structure
d) None of these
- (x) Choose the correct option: clustering is a
a) Supervised learning
b) Unsupervised learning
c) Reinforcement learning
d) None of these
- (xi) Determine the data type that we cannot perform for cluster analysis?
a) Multimedia data
b) Time series data
c) Text data
d) None of these
- (xii) Focus the application of Big data analysis?
a) Stock market prediction
b) Sentiment analysis
c) Real time analysis
d) All of these
- (xiii) Analyze the following statements that is true about data streaming?
a) Stream data is always unstructured data.
b) Stream data often has a high velocity.
c) Stream elements cannot be stored on disk.
d) Stream data is always structured data.
- (xiv) Find errors from following options: Which is the type of a NoSQL database?
a) Cassandra
b) MongoDB
c) SQL Server
d) None of the above
- (xv) Select the correct statement.
a) HBase is a distributed column-oriented database
b) Hbase is not open source
c) Hbase is horizontally scalable.
d) Both A and C

Group-B

(Short Answer Type Questions)

3 x 5=15

2. Report the types of mechanism works in NoSQL and write down their name. (3)
3. Write down some of the pros and cons of K-means clustering algorithm. (3)
4. Report how will you solve a classification problem using Decision Tree. (3)
5. Explain metadata and which node stores metadata in HDFS. (3)
6. Identify the task of map task and reduce task (3)

OR

Illustrate the three types of data analytics . (3)

Group-C

(Long Answer Type Questions)

5 x 6=30

7. Discuss about MapReduce with a example. (5)
8. Illustrate Naive Bayes Classification in detail . (5)
9. Compare the functions of name node, data node and secondary node. (5)
10. Illustrate collaborative filtering in detail. (5)
11. Write down the advantages and disadvantages of NoSQL. (5)
12. Describe Big data and the importance of big data. (5)

OR

Define the characteristics of Big Data.

(5)

Library
Brainware University
398, Ramkrishnapur Road, Barasat
Kolkata, West Bengal-700125