Reg. No. :

Duration: Three hours   Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 1 = 10 Marks)

1. Choose the best answer. Big data analytics technologies are necessary to:

    (a) Formulate eye-catching charts and graphs
    (b) Extract valuable insights from the data
    (c) Integrate data from internal and external sources
    (d) Perform RDBMS query process and display the result

2. Which one from the following is a not analytics method?
    (a) Regression   (b) Monte Carlo   (c) Spatial   (d) Classification

3. Most NoSQL databases support automatic _____, meaning that you get high

    availability and disaster recovery

    (a) Processing   (b) Scalability   (c) Replication   (d) All of the mentioned

4. Pipes use _____ as the channel for communication.

    (a) bus   (b) sockets   (c) data   (d) control

5. HDFS provides a command line interface called _____ used to interact with
    HDFS.
    (a) "HDFS Shell"   (b) "FS Shell"   (c) "DFS Shell"   (d) "HDF Shell"

6. The High availability of Namenode is achieved in HDFS2.x using _____

   (a) Polled Edit Logs   (b) Synchronized Edit Logs

   (c) Shared Edit Logs   (d) Edit Logs Replacement

7. When You are developing a combiner that takes as input Text keys, IntWritable values, and emits Text keys, IntWritable values. Which interface should your class implement?

   (a) Combiner <Text, IntWritable,Text, IntWritable>

   (b) Reducer <Text, IntWritable,Text, IntWritable>

   (c) Combiner <Text,Text, IntWritable, IntWritable>

   (d) Combiner <Text, Text, IntWritable, IntWritable>

8. Which of the following is web-based client software for MongoDB ?

   (a) BI Studio   (b) Database Master   (c) Fang of Mongo   (d) Mongo3

9. Which of the following is a wide-column store?

   (a) Cassandra   (b) Riak   (c) MongoDB   (d) Redis

10. _____ provides statistics on the per-collection level.

   (a) Mongosniff   (b) Mongotop   (c) Mongooplog   (d) Mongofiles

PART - B (5 x 2 = 10 Marks)

11. Draw the Credit Risk Framework.

12. What is NoSQL?

13. Why Key Type need be both Writable and Comparable in Map-Reduce Programs.**?**

14. What is Hadoop streaming?

15. What is ZOO keeper?

PART - C (5 x 16 = 80 Marks)

16. (a) What is Inter-Firewall and Trans-Firewall Analytics? Explain the Value Chain for Inter-Firewall and Trans-Firewall Analytics.   (16)

   Or

   (b) Explain with illustrative examples Big data in healthcare..   (16)

49809

17. (a) Write a short notes on following terminology in NoSQL with suitable application
    (i) Aggregates.   (8)
    (ii) Application Side Joins.   (2)
    (iii) Dimensionality Reduction.   (4)
    (iv) Composite Key Index.   (2)

                                    Or

  (b) Analyze with an example How Hadoop uses Scale out feature to improve the
      Performance   (16)

18. (a) Elucidate the following terms with respect to the basic HADOOP technology.   (16)
     (i) Map reduce.
     (ii) Hadoop Scale out.
     (iii) HAdoop Streaming..
     (iv) Hadoop pipes.

                                    Or

  (b) Analyze Hadoop streaming in detail with suitable examples   (16)

19. (a) How does map reduce on YARN technology work? Details the failures in YARN
        over map reduce.   (16)

                                    Or

  (b) Draw the architecture of Map Reduce process and explain the mapper and
      master in detail.   (16)

20. (a) (i) Write short notes on primary and complex data types in Hive data types.   (8)

       (ii) What are the CRUD operations in HBase? Among them explain Write and Read
            operations with supportive commands.   (8)

                                    Or

  (b) What are the good rules and goals for creating data model in Cassandra?   (16)

49809