

PARUL UNIVERSITY
FACULTY OF IT & COMPUTER SCIENCE
BCA/B.Sc.(IT)/IMCA(AY_IV) Summer 2022-23 Examination

Semester: 4**Date: 21/03/2023****Subject Code: (05101266/05102252/05302252)****Time: 10.30am to 1.00pm****Subject Name: Big Data Analytics –I****Total Marks: 60****Instructions:**

1. All questions are compulsory.
2. Figures to the right indicate full marks.
3. Make suitable assumptions wherever necessary.
4. Start new question on new page.

Q.1 Answer the followings.**A. Answer / Define the following in short. (05)**

1. State the term Big Data.
2. What is structured data?
3. Define: Business Intelligence.
4. Give one disadvantage of Distributed Computing with respect to Big Data.
5. What is the main use of HDFS?

B. Multiple choice type questions/ Give the sentence true or false. (Each of 01 marks) (10)

1. Which of the following is not a characteristics of Big Data?
 - A. Volume
 - B. Variability
 - C. Variety
 - D. Velocity
2. Who among the following do you think would be able to deal with growing number of data source efficiently?
 - A. Business developer
 - B. Data Scientist
 - C. Sales executive
 - D. Web Designer
3. Which application of social network data analysis is used by a customer retention manager?
 - A. Business intelligence
 - B. Marketing
 - C. Product design and development
 - D. Insurance fraud
4. Identify the type of fraud that typically impact online retailers.
 - A. Credit card fraud
 - B. Forward fraud
 - C. Corporate fraud
 - D. Insurance fraud
5. In distributing computing, _____.
 - A. The computing task is divided among
 - B. Additional high-capacity disks are added
 - C. The results are shared among several users of the network
 - D. The computing task is moved to the cloud
6. Why are Big Data applications susceptible to latency?
 - A. Big data may reside in a different location from use in-memory
 - B. The volume of big data is too large to be analyzed rapidly
 - C. Big data cannot use in-memory computing
 - D. Big data applications are still in the early stage of development

7. Which of the following is used to denote the small subsets of a large file created by HDFS?
 - A. NameNode
 - B. DataNode
 - C. Blocks
 - D. Namespace
8. What message is generated by a DataNode to indicate its connectivity with NameNode?
 - A. Beep
 - B. Heartbeat
 - C. Analog pulse
 - D. Map
9. Which of the following options most aptly explains the reason behind the creation of Map Reduce?
 - A. Need to increase the processing power of new hardware
 - B. Need to perform complex analysis of structured data
 - C. Need to increase the number of web users
 - D. Need to spared distributed computing
10. In designing the Map Reduce framework, which of the following needs did the engineers consider?
 - A. It should be cheap and distributed free of cost
 - B. Processing should expand and contract automatically
 - C. Processing should be stopped in the case of network failure
 - D. Developers should be able to create new language

Q.2 Answer the followings. (3 Marks Questions.) (Any Five) (15)

1. Explain in brief about elements of Big Data.
2. Write any three point of difference between Internal Source and External source of Data.
3. Give brief about the use of social networking in Big Data with example.
4. Explain how distributed computing work with appropriate diagram.
5. Discuss in details the Hadoop Ecosystem.
6. Describe the MapReduce framework with example.

Q.3 Answer the following. (5 Marks Questions)(Any three) (15)

1. Explain in detail about evaluation of Big Data with diagram.
2. Discuss with various example, the use of Big Data in preventing Fraudulent activities.
- 3 Describe the Hadoop multi-nodes cluster architecture.
4. Give main significant of NameNode. Explain HDFS architecture in details.

Q.4 Answer the following in detail.

- A.** Describe various areas in which decision-making process are influenced by social network data. **(05)**
- B.(1)** Discuss in details Use of Big Data in Retail Industry. **(05)**
- B.(2)** Enlist Technologies for Handling Big Data. Give brief about each of them. **(05)**
- OR
- B.(1)** Write steps to setting up and installing Hadoop in its two operating modes. **(05)**
- B.(2)** Describe with example how to implement file management tasks in Hadoop. **(05)**