

Please check that this question paper contains 9 questions and 2 printed pages within first ten minutes.

[Total No. of Questions: 09]

[Total No. of Pages: 2]

Uni. Roll No. 1805158.....

Program: B.Tech. (Batch 2018 onward)

Semester: 8th

Name of Subject: Big Data

Subject Code: PECS-118

Paper ID: 17573

Time Allowed: 03 Hours

Max. Marks: 60

NOTE:

- 1) Parts A and B are compulsory
- 2) Part-C has Two Questions Q8 and Q9. Both are compulsory, but with internal choice
- 3) Any missing data may be assumed appropriately

Part – A

[Marks: 02 each]

Q1.

- a) Illustrate key roles for new Big Data Ecosystem.
- b) Define different modes of Hadoop.
- c) Classify the difference between Hbase and HDFS.
- d) Discuss briefly XML files required for configuring Hadoop.
- e) Explain types of NoSql databases in practice.
- f) Design Business Analytic role in Health Care Analytics.

Part – B

[Marks: 04 each]

- Q2.** Explain the working of Map Phase and Combiner Phase of Map-Reduce paradigm.
- Q3.** Describe Descriptive, Predictive and Perspective analytics.
- Q4.** Differentiate between Data at Rest vs Data in Motion. Do they require different computation power? Explain.
- Q5.** Compare and Contrast HIVE and SPARK.
- Q6.** Write the major role of Graph Databases in Big Data Analytics.
- Q7.** Analyse whether k-means clustering is good for large datasets? Support your answer with an example.

Part – C

[Marks: 12 each]

- Q8.** Explain HDFS architecture in detail. Explain the function of Name node and Data node. Discuss the need and importance of Secondary name node, Task-Tracker and Job Tracker in HDFS.

OR

Explain various building blocks in Hadoop. What is checkpointing? Specify server roles of Hadoop Cluster.

- Q9.** Classify the difference between:

- a) MongoDB and Cassandra
- b) Linear Regression and Logistic Regression

OR

Evaluate various Analytic Methods and models used in business analytics in context to sports and web analytics with suitable examples.
