USN ☐☐☐☐☐☐☐☐☐☐   **16/17SCN24**

Second Semester M.Tech. Degree Examination, June/July 2018
Managing Big Data

Time: 3 hrs.   Max. Marks: 80

Note: Answer FIVE full questions, choosing one full question from each module.

Module-1

1 a. Write a brief note on the following:
i) Big data   ii) Big data Analysis   iii) Unstructured data   iv) Volume and variety of big data.   **(10 Marks)**
   b. Explain how big data analysis can be used for the following areas:
i)   Understanding and targeting customers
ii) Improving and optimizing cities and countries.   **(06 Marks)**

OR

2 a. What are 2 critical components of hadoop and give brief explanation of each.   **(08 Marks)**
   b. What is open source technology and explain how it is contributing to big data analytics.
                                                                             (04 Marks)
   c. What is crowd sourcing analytics and how it is helpful for big data analysis?   **(04 Marks)**

Module-2

3 a. What is NOSQL and explain at least four reasons for using NOSQL data base for big data analysis?   **(10 Marks)**
   b. Compare and contrast SQL and NOSQL taking at least six points.   **(06 Marks)**

OR

4 a. What is data aggregation? Explain different data models used in NOSQL with examples.
                                                                             (10 Marks)
   b. What is NOSQL polygot persistent explain with a neat diagram?   **(06 Marks)**

Module-3

5 a. Explain different types of data input format and output format supported by Hadoop with an example.   **(10 Marks)**
   b. What is Hadoop pipes give a brief explanation with an example.   **(06 Marks)**

OR

6 a. Explain data integrity mechanism supported by HDFS and how data nodes and clients do integrity check.   **(08 Marks)**
   b. Explain combiner functions and Hadoop streaming in brief.   **(08 Marks)**

Module-4

7 a. What is map reduce workflow and explain with an example how a problem can be translated into map reduce work flow with an example.   **(10 Marks)**
   b. Explain with a neat diagram how Hadoop runs a Map Reduce job using YARN.   **(06 Marks)**

OR

8 a. Explain how Hadoop runs a Map Reduce job using classic framework. **(10 Marks)**

b. Which are the failures to be considered while running Map Reduce on YARN give brief explanation? **(06 Marks)**

Module-5

9 a. Explain data model of HBASE. **(06 Marks)**

b. Compare and contrast HBASE and RDBMS. **(04 Marks)**

c. List characteristics of HBASE with brief explanation. **(06 Marks)**

OR

10 a. Briefly explain the execution types of PIG. **(06 Marks)**

b. Compare and contrast PIG and data bases. **(04 Marks)**

c. Explain HiVE services. **(06 Marks)**

* * * * *