

B.TECH. SEM -VII (COMPUTER) 2014 COURSE (CBCS) :

SUMMER - 2018

SUBJECT: BIG DATA ANALYTICS AND ARCHITECTURE

Day : **Wednesday**
Date : **23/05/2018**

S-2018-2483

Time : **02.30 PM TO 05.30 PM**
Max Marks : **60**

N.B. :

- 1) All Questions are **COMPULSORY**.
- 2) Figures to the **RIGHT** indicate full marks.
- 3) Assume suitable data **WHEREVER** necessary.

- Q1.** a) What is data? Describe the categories of data and its sources. (05)
b) Explain the role of RDBMS in Big data ecosystem. (05)

OR

- a) What is Big Data? Describe the key characteristics of Big data. (05)
b) Explain the role of CMS in Big data management. (05)

- Q2.** a) Explain Bayes rule and Bayesian Inference. (05)
b) Consider the following set of Points: $\{(-2,-1),(1,1),(3,2)\}$. Find the least square regression line for the given data points. (05)

OR

- a) What is Text analytics? Explain any two tools or techniques for text analytics. (05)
b) Suppose everyone who visits a retail website gets one promotional offer or no promotional offer at all, we want to see if making a promotional offer makes a difference. What statistical method you recommend for this analysis. Explain. (05)

- Q3.** What kind of tools would be used in the Data analytics lifecycle phases and for which kinds of use scenarios? Explain. (10)

OR

Describe in detail with an example, the workflow for building a predictive model using R. (10)

- Q4** What is Virtualization? Explain the importance of various types of virtualization in Big data. (10)

OR

What is NoSQL? How NoSQL is different from traditional RDBMS? Elaborate on the design goals for MongoDB. (10)

- Q5.** a) What is Map reduce? Describe how Map reduce works? (05)
b) What is Hadoop? Explain the pros and cons of Hadoop. (05)

OR

- a) Explain the HDFS Architecture with suitable diagram. (05)
b) Describe the role of Pig and Hive in Hadoop ecosystem. (05)

- Q6.** a) Describe the Privacy landscape in Big data environment. (05)
b) What is credit risk management? Describe the role of big data aspects in credit risk management. (05)

OR

- a) Define User Anonymization and Data Anonymization. Explain how they impact security and privacy in big data environment? (05)
b) What is recommendation Engine? Illustrate with an example how the collaborative filtering base recommendation system works. (05)

* * * *