Roll No ....

IT-8003 (2) (CBGS)
B.E. VIII Semester Examination, June 2020
Choice Based Grading System (CBGS)
Data Science

Time : Three Hours

Maximum Marks : 70

Note: i)   Attempt any five questions.
      ii)   All questions carry equal marks.

1. a)   Explain big data management architecture with neat block diagram.   7

   b)   Explain the role of relational databases in big data.   7

2. a)   Classify the different data types with reference to big data.   7

   b)   Define a normal distribution and derive the mean, median, mode of a normal distribution 7

3. a)   Derive the mean and the variance of a binomial distribution.   7

   b)   Explain the role of support vector machine in data analysis.   7

4. Explain the following: 14
   i)   Naive bayes classifier
   ii)   Linear regression with nonlinear substitution
   iii)   Evolution of data management

5. a)   Explain parallel vs distributed processing.   7

   b)   Explain data warehousing architecture.   7

6. a)   Explain data preparation, model planning and model building.   7

   b)   Differentiate cluster architecture and traditional architecture.   7

7. a)   Explain core components of hadoop ecosystem with neat diagram.   7

   b)   Explain loops and conditional execution in R language.   7

8. a)   Explain data manipulation and statistical analysis with R.   7

   b)   Explain various data structures used in R.   7

******

IT-8003(2) (CBGS)