Roll No ....

CS-7005(2)-CBGS
B.E. VII Semester
Examination, December 2020
Choice Based Grading System (CBGS)
Data Science and Big Data
Time : Three Hours

Maximum Marks : 70

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

1.   a)   What is Data Wrangling? Define exploratory data analysis. Why EDA is required in Data Analysis? Discuss about some tools and techniques used for EDA.

    b)   Define and differentiate data cleaning and data transformation. What kind of issues affect the quality of data? How can I detect and fix these issues?

2.   a)   What is Feature Extraction? Discuss advantages of the Feature Extraction process in data analytics.

    b)   What do you understand by Data Visualization? Discuss some Python's data visualization tools such as Box plots, Pie Charts and Bar Charts in brief.

3.   a)   What is reasoning under uncertainty? What is probability reasoning? How logical reasoning is differing with probability reasoning?

    b)   Explain and prove the Bayes Theorem. What is meant by conditional probability? Explain the method of performing exact inference in Bayesian Networks.

4. a) What do you mean by Clustering? Explain working of any clustering algorithm.
   b) What do you understand by dimensionality reduction? Explain working of the PCA.

5. a) Explain supervised and unsupervised learning with suitable example.
   b) Explain any one clustering algorithm in detail.

6. a) Explain training and testing of data with example.
   b) Explain vector model of information retrieval.

7. a) Discuss about information retrieval. Discuss about various models used in information retrieval in detail.
   b) Explain text mining. What kind of problem can be addressed using text mining? What kind of sentiments can be found in the text? Why is text mining useful in the age of social media?

8. a) Explain the process of Data storage in Hadoop Distributed File System (HDFS) with help of a suitable example.
   b) What is Map Reduce programming model? Explain its Execution process and working with help a suitable flow diagram and example.

******

CS-7005(2)-CBGS