Roll No. .....

Total Pages : 03

016602

May 2024

B. Tech. (CE (DS)) (Sixth Semester)

Data Acquisition Analysis and Visualization

(PCC-DS-602)

Time : 3 Hours] [*Maximum Marks : 75*

Note : It is compulsory to answer all the questions (1.5 marks each) of Part A in short. Answer any *four* questions from Part B in detail. Different sub-parts of a question are to be attempted adjacent to each other.

Part A

1. (a) How is Multilabel classification different from Multiclass ? **1.5**

(b) What does sharding technique refer to ? **1.5**

(c) What is the role of multiplexer in data acquisition system ? **1.5**

(d) What do you understand by Agile Manifesto ? **1.5**

(e) Define Data Sampling Rate. **1.5**

(f) List the features of NoSQL databases. **1.5**

(g) Specify the ways through which various types of data can be visualized. **1.5**

(h) Differentiate between Scale-up and Scale-out. **1.5**

(i) State two benefits of Data Acquisition System. **1.5**

(j) Define the term SLA. **1.5**

Part B

2. (a) Differentiate between NoSQL and traditional databases. **10**

   (b) Explain *two* applications of big data. **5**

3. (a) What are the *five* P's of Data Science ? **10**

   (b) What components constitute Data Science ? List the major steps of any Data Science process. **5**

4. (a) Explain the components of data acquisition system in detail with the help of a suitable diagram. **10**

   (b) Explain the various approaches to acquire data for applying machine learning algorithms. What are the major tools and techniques available for this data acquisition ? **5**

5. (a) How are failures handled in MapReduce ? **10**

   (b) How is the performance of a classification algorithm measured ? State the usage of each performance metric. **5**

6. (a) How does clustering process differ from classification ? Explain through a suitable example. **10**

   (b) Provide design of a NoSQL database for a university undergraduate student. How is this design different from SQL oriented database design ? **5**

7. Write short notes on the following :

   (a) Signal conditioning **5**

   (b) Data Ingestion **5**

   (c) YARN. **5**