



BIG DATA ANALYTICS
(Computer Science and Engineering)

Maximum Marks: 70

Date:02.07.2022 Duration: 3 hours

- Note: 1. This question paper contains two parts A and B.
2. Part A is compulsory which carries 20 marks. Answer all questions in Part A.
3. Part B consists of 5 Units. Answer any one full question from each unit which carries 10M.
4. Each question carries 10 marks and may have a, b, c, d as sub questions.

Part-A

All the following questions carry equal marks

(10x2M=20 Marks)

- 1 Write the elements of data architecture
- 2 What is Data Preprocessing
- 3 Why Hadoop is used in Big data Analytics
- 4 What are standard reporting templates
- 5 How Outlier can be eliminated in Big Data Analytics?
- 6 What is Key-value data store?
- 7 Interpret hypothesis testing in machine learning algorithms
- 8 How do you prepare the input data for an algorithm?
- 9 List Quick Visual Options in Tableau
- 10 How Data Visualization is used in getting draw insights out of visualization tool?

Part-B

Answer All the following questions.

(10MX 5=50Marks)

- 11 List the guidelines for identifying and reporting an accident or emergency in detail. 10 M
- OR
- 12 Explain the process of exporting data to Cloud (AWS) 10 M
- 13 What is knowledge management? Explain about model-based techniques. 10 M
- OR
- 14 Explain in detail about Data ETL process 10 M
- 15 Describe various types of Big Data Analytics? Explain about Predictive Analytics 10 M
- OR
- 16 Explain Outlier detection and elimination process in detail? 10 M
- 17 Describe Train Model on 2/3 sample data using Machine Learning Algorithm and Test model on 1/3 sample for prediction 10 M

OR

- 18 Explain in detail about Hypothesis testing 10 M
- 19 A. What are the steps involved in production implementation 5 M
B. Draw insights out of any one Visualization Tool 5 M
- OR
- 20 Explain in detail about data visualization in Tableau. 10 M