GUJARAT TECHNOLOGICAL UNIVERSITY

POST GRADUATE DIPLOMA IN BIOINFORMATICS SEMESTER - 1 WINTER 2021 EXAMINATION

Subject Code:1310204   **Date:14 Mar 2022**
Subject Name:Cloud Computing, Data Mining & Visualization
Time:10:30 AM TO 01:00 PM   **Total Marks: 70**
Instructions:
1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**
4. **Draw neat and clean diagrams as required**

Q.1   **Write a note on following**   **(Marks-10X2=20)**

1. What are the types of cloud as far as the computation is concerned?
2. What do you mean by virtualization?
3. Which are the different layers that define cloud architecture?
4. Which platforms are used for large scale cloud computing?
5. What are the different models for deployment in cloud computing?
6. Define: data mining, data warehousing.
7. Write bayes theorem.
8. What do you mean by frequent patterns in transaction dataset?
9. Define: minimum confidence, association rule.
10. Define: supervised machine learning, unsupervised learning.

Q.2   **Answer the following (Any 2 out of 3)**   **(Marks-2X10=20)**

1. What are the benefits and disadvantages of cloud computing? Explain them in detail.
2. Discuss in detail any one case study of cloud based services in the field of bioinformatics.
3. With suitable examples, show different types of charts and graphs with their specific application that are used to visualize the data.

Q.3   **Answer the following (Any 6 out of 8)**   **(Marks-6X5=30)**

1. Write the desired features of cloud as far as computation is concerned.
2. What are the challenges and risks involved in providing cloud infrastructure and platform as service?
3. Explain in brief about any two tools that are used to build private cloud.
4. What are various virtualization technologies for cloud computing platform? Also mention its characteristics.
5. What is OLAP? With suitable examples, show any two OLAP operations on multidimensional data.
6. Write the steps to construct the decision tree for given supervised dataset.
7. What do you mean by missing values in dataset? What are the methods to cope-up with missing values?
8. Write the steps of apriori algorithm to mine frequent item-sets from given transaction dataset.

****