B. C. A. (Sem. VI) (CBCS) Examination

March / April - 2018

CS-32 : Data Warehousing & Data Mining

Faculty Code : 003
Subject Code : 003621

Time : $2\frac{1}{2}$ Hours]　　　　　　　　　　　　[Total Marks : **70**

1　Answer the following questions :　　　　　　　　**20**

(1)　Data can be updated in _____ environment.

(2)　The star schema is composed of _____ fact table.

(3)　_____ describes the data contained in the data warehouse.

(4)　The core of the multidimensional model is the _____ which consists of large sets of facts and a number of dimensions.

(5)　Classification is an unsupervised learning. (True / False)

(6)　WEKA stands for _____

(7)　The dimension tables describes the _____

(8)　Which type of data generally data Mart can contain ?

(9)　OLAP stands for _____

(10)　Which algorithm is a collection of nodes (neurons) and edges ?

(11)　GA stands for _____

(12)　Which method maps data into the predefined groups ?

(13)　The human brain consists of a network of _____

(14)　Which is the well known association rule algorithm and is used in the most commercial products?

(15)　Which clustering method is used for iterative clustering?

(16)　MFCS stands for _____

(17) Prediction can be viewed as forecasting a _____ value.

(18) Data warehouse contains _____ data that is never found in the operational environment.

(19) Which clustering patterns/ techniques start with all records in one cluster and then try to split that cluster into small pieces ?

(20) Converting data from different sources into the common format for processing is called as _____

2 (A) Attempt any **three** :   6

   (1) Write a note on DSS system.

   (2) Differentiate: Fact table v/s Dimension table.

   (3) Differentiate: informational system v/s operational system.

   (4) Explain Bayes theorem.

   (5) Write a note on Granularity.

   (6) Explain point estimation method.

   (B) Attempt any **three** :   9

   (1) Differentiate: star schema v/s snowflake schema.

   (2) Write a note on Data Mart.

   (3) What is K-Means algorithm? Explain.

   (4) What is Data Warehouse? Explain its characteristics.

   (5) What is ETL? Why we use it?

   (6) Write a note on partitioning of data.

   (C) Attempt any **two** :   10

   (1) Explain data warehouse architecture in detail.

   (2) What is OLAP? Explain its various types.

   (3) Differentiate: OLTP v/s OLAP.

   (4) Write comparison and contradiction of any two ETL tools.

   (5) What is Data Mining? Explain its various techniques.

3 (A) Attempt any **three** : **6**

    (1) List out various application areas of data mining.

    (2) Write a note on association rule.

    (3) Explain Nearest neighbor algorithm.

    (4) What is Weka? Explain its various features.

    (5) Write a note on Machine learning technique.

    (6) List out various methods of association rule algorithm.

  (B) Attempt any **three** : **9**

    (1) Differentiate: Data Warehouse v/s Data Mart.

    (2) What is Genetic algorithm? Explain any two.

    (3) What is clustering? What is the use of it? Explain.

    (4) What is KDD? Explain.

    (5) What is statistics? Explain its usage in data mining.

    (6) Explain the term: agglomerative clustering, divisive clustering.

  (C) Attempt any **two** : **10**

    (1) Explain Apriori algorithm in detail.

    (2) What is the role of data mining in government field? Explain.

    (3) What is Neural Network? Explain in detail.

    (4) Write a note on different data mining processes in detail.

    (5) Write a step to develop the decision tree in Weka.

———————