TE - sem-VI - old-computers

DWM

14/12/15

QP Code : 2055

TIME - 3 Hrs **Marks – 100**

Note:

1. **Question 1 is compulsory.**
2. **Answer any 4 out of the remaining questions.**

Q 1.

A bank wants to develop a data warehouse for effective decision-making about their loan schemes. The bank provides loans to customers for various purposes like house building loan, car loan, educational loan, etc. The whole country is categorized in to regions, namely north, south, east and west. Each region consists of a set of states. Loan is disbursed to customers at interest rates that change from time to time. Different types of loans have different rates. The data warehouse should record an entry for each disbursement of loan to customer. For this system:

(a) Draw an information package diagram **[10]**

(b) Draw star schema diagram. **[10]**

Q 2

(a) Consider a data warehouse for a hospital, where there are three dimensions: Doctor, Patient Time and two measures i.e. count and charge, where charge is the fee that the doctor charges a patient for a visit. Using these describe the following operations: Slice, Dice, Rollup, Drill down and Pivot. **[10]**

(b) Describe ETL cycle in data ware house. **[10]**

Q 3

(a) Consider the following transaction data base and apply the Apriori algorithm with minimum support of 30% and minimum confidence of 70% and find all the association rules in the data set. **[10]**

TID	Items
1.	A, B, C, D
2.	A, B, C, D, E, G
3.	A, C, G, H, K
4.	B, C, D, E, K
5.	D, E, F, H, L
6.	A, B, C, D, L
7.	B, I, E, K, L
8.	A, B, D, E, K
9.	A, E, F, H, L
10.	B, C, D, F

TURN OVER

QP-Con. 11358-15.

(b) Using an example dataset illustrate any one classification technique. [10]

Q 4

(a) Suppose you are employed as a data mining consultant for an Internet search engine company. Describe how data mining techniques can be used to help the company by giving specific examples with respect to clustering, classification and association rule mining techniques. [10]

(b) What is Clustering Technique? Discuss hierarchical clustering techniques. [10]

Q5

(a) What is K-Means Clustering? Perform the K-Means algorithm with the following data for two clusters. Data Set {10, 4, 2, 12, 3, 20, 30, 11, 25, 31} [10]

(b) Discuss Multidimensional and Multilevel association mining. [10]

Q 6

(a) What are the primitives that define data mining task? Show how DMQL can be used to define task primitives. [10]

(b) Describe the three tier architecture of a data warehouse [10]

Q 7 Write short notes on any two: [20]

(a) Application of Data Mining in ecommerce

(b) Web Mining