5092

B.Tech. Examination, 2017

(Seventh Semester)

(C.S. & I.T. Branch)

DATA MINING & DATA WARE HOUSING

Paper - III

Time Allowed : Three Hours

Maximum Marks : 100

Note : Attempt any five questions. All questions carry equal marks.

Q. 1. What is data mining ? In your answer, address the following :

(a) Explain how the evolution of data base technology led to data mining.

(b) Describe the step involved in data mining when view as a process of knowledge discovery.

Q. 2. (a) What is difference between discrimination and classification ? Between characterization and clustering ? Between classification and prediction ? For each of these pair of task how are they similar.

(b) Use the two method below to normalize the following group of data :

200, 300, 400, 600, 1000

(i) Min-max normalization by setting min = 0 and max = 1

(ii) Z-Score normalization

3. (a) Explain Architecture of data mining system with neat diagram.

(b) In real-world data, tuples with missing value for some attribute are a common occurrence Describe various method for handling this problem.

4. (a) Explain various component of Three-Tier Data Warehouse Architecture with suitable diagram.

(b) Explain in detail about classification by Backpropagation.

5. (a) Briefly compare the following concept. [You may use an example to explain your point.]

(i) Star schema

(ii) Snowflake schema

(iii) Fact constellation schema

(b) Define following normalization method :

(i) Min-max Normalisation

(ii) Z-score Normalization

(iii) Normalization by decimal scaling

Q. 6 (a) Suppose that the data for analysis include the attribute age. The age value for the data tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33,33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70.

Using this data for age answer the following :

(i) Use Min-max normalization to transform the value 35 for age, onto the range [0.0, 1.0].

(ii) Use Z-score normalization to transform the value 35 for age, where the standard deviation of age is 12.94 year.

(b) List and explain comparison between OLTP and OLAP system.

Q. 7. (a) What is association rule mining ? Explain Apriori Algorithm to find frequent item sets.

(b) Illustrate the association rule and its purposes. Discuss how market basket analysis form the

association rules with the help of suitable example.

Write short notes on following :

(a) HOLAP and MOLAP

(b) Clustering Analysis

(c) Genetic Algorithm

(d) Hetrogeneous database and legacy database

(e) Challenges of data mining

1092

B.Tech. Examination, 2016

(Seventh Semester)

(C.S. & I.T. Branch)

DATA MINING & DATA WARE HOUSING

Paper - III

Time Allowed : Three Hours

Maximum Marks : 100

Note : Attempt any five questions. All questions carry equal marks.

Q. 1. (a) Why multidimensional views of data and data cubes are used ? With a neat diagram explain data cube implementations.

(b) What are data cube operations ? Explain.

Q. 2. (a) What is data mining ? Explain various data mining tasks.

(b) Why data preprocessing required in data mining ? Explain various steps involved in data preprocessing.

Q. 3. (a) Discuss about FP-Tree growth algorithm.

(b) Explain various categories of Association rules.

Q. 4. (a) Describe the steps involved in knowledge discovery in database.

Describe the different methods for data cleaning.

Q. 5. (a) Compare and contrast data warehouse and data mart. Also explain the reasons for creating the data mart.

(b) Discuss about the typical OLTP operations on multidimensional data with example.

Q. 6. (a) What is web mining ? Explain the techniques in web mining.

(b) Explain K-means clustering algorithm. What are its limitations ?

Q. 7. Write short note on :

(i) Genetic algorithm

(ii) ROLAP, MOLAP and HOLAP

2257

B.Tech. Examination, 2015

(Seventh Semester)

(C.S. & I.T. Branch)

DATA MINING & DATA WARE HOUSING

Paper-III

Time Allowed : Three Hours

Maximum Marks : 100

Note : Attempt any five questions. All questions carry equal marks.

Q. 1. (a) Explain basic data mining tasks with an example. **10**

(b) Give details on data mining versus knowledge discovery in databases. **10**

Q. 2. (a) What do you mean by hypothesis testing ? Give an example. **10**

(b) What is metadata and why is it important ? Discuss the multidimensional data. **10**

Q. 3. (a) Explain with an example Bayesian classification. **10**

P.T.O.

(b) Explain decision tree-based algorithms. 10

Q. 4. (a) Discuss about the typical OLAP operations on multidimensional data with an example. 10

(b) How is web usage mining different from web structure mining and web content mining ? 10

Q. 5. Explain the following : 20

(i) Nearest Neighbor algorithm

(ii) PAM algorithm

Q. 6. (a) With a neat diagram, give a brief explanation of the various components of three-tier data warehouse architecture. 10

(b) Draw a snowflake schema diagram for the data warehouse ? 10

Q. 7. Write short notes on any four of the following : 20

(a) Data mining metrics

(b) Decision Support Systems

(c) Neural Network

(d) Hierarchical Clustering

(e) HOLAP

(f) Tuning Data Warehouse.

396

B.Tech. Examination, 2014

(Seventh Semester)

(C.S. & I.T. Branch)

DATA MINING & DATA WARE HOUSING

Paper - III

Time Allowed : Three Hours

Maximum Marks : 100

Note : Attempt any five questions. All questions carry equal marks.

Q. 1. (a) How is a Data Warehouse different from a database ? How are they similar to each other ? Explain. 10

(b) How are Data Marts different from data warehouse ? Explain different types of data marts. 10

Q. 2. (a) What are the various components of data warehouse ? List them, and explain briefly. 10

(b) Discuss different approaches used to build a data warehousing. Which approach is generally used for building a data warehouse ? 10

Q. 3. (a) In Association Rule Mining, what is the anti-monotone property ? How is this property utilized in the Apriori Algorithm ? 10

(b) What do you mean by Web Mining ? How Web Content Mining is different from Web Structure Mining ? 10

Q. 4. (a) What is the role of OLAP in data Warehouse ? Explain any one OLAP model. 10

(b) Explain Rool-up display and Drill down operation. 10

Q. 5. (a) Differentiate between classification and clustering. 10

(b) Briefly explain the following : 10

(i) Spatial Datamining

(ii) Text Mining

(iii) Decision Tree

Q. 6. Define and differentiate MOLAP and ROLAP. Give arguments in favour of MOLAP and ROLAP. State rules for evaluating OLAP products developed by E.F Codd 20

Q. 7. Write short notes on any four of the following : 20

(a) Data cleaning and Data Integration

(b) Classification by Back propagation

(c) Lazy Learners

(d) Priodicity Analysis

(e) DBSCAN and OPTICS

(f) Principles of Dimensional Modeling.