(Following Paper ID and Roll No. to be filled in your Answer Book)

PAPER ID : 110755

Roll No.

B. Tech.

(SEM. VII) (ODD SEM.) THEORY EXAMINATION, 2014-15

DATA MINING & DATA WAREHOUSING

Time : 3 Hours] [Total Marks : 100

Note : Attempt all questions.

Attempt any **four** parts of the following : **5×4=20**

(a) What is data mining? Define the major issues in data mining.

(b) Explain Data, Information and Knowledge.

(c) What are the different forms of data processing?

(d) Explain the data cleaning, data Integration and transformation in brief.

(e) Distinguish between dimensionality reduction and Numerosity reduction.

(f) Explain concept hierarchy generation for categorical data.

https://quescol.com | Free PYQs, Solutions, Syllabus, Notes etc

2 Attempt any **four** parts of the following : **5×4=20**

(a) What are the properties of standard deviation and give its formula ?

(b) Explain data cube approach and attribute oriented approach.

(c) What is the role of statistics in data mining?

(d) What is association rule mining? Explain the Apriori algorithm to find the frequent item sets.

(e) Explain multi-level association rules for transactional database.

(f) Write short notes on :

(i) Quartiles

(ii) Histograms

(iii) Scatter plots.

3 Attempt any **two** parts of the following : **10×2=20**

(a) Explain the density based clustering method based on connected regions with sufficiently high density (DBSCAN).

(b) (i) Explain the different data types used in cluster analysis.

(ii) Describe the role of Genetic Algorithm in data mining.

(c) Explain multilayer feed-back neural network. Differentiate between feed-forward and feedback system.

4 Attempt any **two** parts of the following : **10×2=20**

(a) Explain 3 tier architecture of data warehouse. Explain the difference between Data mart and Data warehouse.

(b) Explain all steps and guidelines for data warehouse implementation.

(c) Explain with diagram the STAR, SNOWFLAKE and FACT constellation schemas for multidimensional databases.

5 Attempt any **two** parts of the following : **10×2=20**

(a) Explain different backup and recovery models in data warehousing.

(b) Describe the basic similarities and differences among ROLAP, MOLAP and HOLAP.

(c) Describe the following in detail :

(i) Data mining interface.

(ii) Testing of data warehouse.

https://quescol.com | Free PYQs, Solutions, Syllabus, Notes etc