

Hall Ticket Number:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Code No. : 32512

VASAVI COLLEGE OF ENGINEERING (Autonomous), HYDERABAD
B.E. (I.T.) III Year II-Semester Main Examinations, May-2017

Data Warehousing and Data Mining

Time: 3 hours

Max. Marks: 70

Note: Answer ALL questions in Part-A and any FIVE from Part-B

Part-A (10 × 2 = 20 Marks)

1. List the steps in Knowledge discovery process.
2. Define Data cleaning in Data mining.
3. What is multidimensional data model?
4. Define Support and Confidence for frequent Item sets.
5. Define Classification and Prediction.
6. Describe Polynomial regression.
7. List the applications of Cluster analysis.
8. Differentiate between agglomerative vs divisive approaches in hierarchical clustering method.
9. Define Spatial Data Mining.
10. List out any 3 tasks of World Wide Web mining.

Part-B (5 × 10=50 Marks)

11. a) List and explain Data Mining Task Primitives. [5]
b) Describe the Strategies in Data Reduction technique. [5]
12. a) Illustrate a three-tier data warehouse architecture with a neat diagram. [6]
b) Explain the following OLAP operations with relevant examples. [4]
i) Roll-up ii) Drill-Down iii) Slice and Dice iv) Pivot
13. a) Explain the classification by back propagation algorithm with the help of multilayer feed forward network. [5]
b) Define Bayes theorem and explain Bayesian classification with an example. [5]
14. a) Describe the k-means algorithm for partitioning with suitable example. [6]
b) Given the following measurements for the variable age: [4]
18, 22, 25, 42, 28, 43, 33, 35, 56, 28, standardize the variable by the following:
i) Compute the mean absolute deviation of age.
ii) Compute the z-score for the first four measurements.
15. a) What are the various approaches for Similarity Search in Multimedia Data? [5]
b) Describe the Concepts and Primitives of Sequential Pattern Mining. [5]
16. a) In real-world data, tuples with missing values for some attributes are a common occurrence. Describe various methods for handling this problem. [5]
b) Explain the basic concept of association rule mining and a road map of it. [5]
17. Answer any *two* of the following:
a) Describe the basic algorithm for inducing a decision tree from training tuples. [5]
b) Explain the Necessity for density-based local outlier detection in detail. [5]
c) Discuss in detail about the Basic Measures for Text Retrieval in Text Mining. [5]