

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

**VASAVI COLLEGE OF ENGINEERING (Autonomous), HYDERABAD**  
**M.C.A. II Year II-Semester Main & Backlog Examinations, May-2017**

**Data Mining**

Time: 3 hours

Max. Marks: 70

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

**Part-A (10 X 2=20 Marks)**

1. What is the difference between Classification and clustering?
2. What are the data smoothing methods?
3. Define data mart and virtual data warehouse.
4. Define different types of OLAP Servers.
5. What is confidence?
6. Define multilevel association rule.
7. What is prediction?
8. Define Tree pruning.
9. Write briefly the density-based method for Clustering.
10. List the different types of Outliers.

**Part-B (5 × 10=50 Marks)**

11. a) Given age attribute values for the data tuples are 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52 and 70. Use smoothing by bin means to smooth the data, using a bin depth of 3. [5]
- b) List and describe the five primitives for specifying a data mining task. [5]
12. a) Explain OLAP operations in multidimensional data model. [5]
- b) Discuss Data warehouse schemas. [5]

13. a)

TID	Items bought
T100	{M, O, N, K, E, Y}
T200	{D, O, N, K, E, Y}
T300	{M, A, K, E}
T400	{M, U, C, K, Y}
T500	{C, O, O, K, I, E}

- Find all frequent itemsets using FP-Growth tree algorithm. [6]
- b) What is the need of correlation analysis to association analysis? Discuss different correlation measurements. [4]
14. a) Discuss Back Propagation algorithm. [6]
- b) Why is Naïve Bayesian Classification called “naïve”? Briefly outline the major ideas of naïve Bayesian classification. [4]

- 15. a) Describe K-Means Algorithm. [5]
- b) Describe the STING method for Clustering. [5]
- 16. a) Explain Curse of dimensionality. [4]
- b) Discuss in detail data reduction techniques. [6]
- 17. Answer any two of the following:
  - a) Outline multilevel association rules. [5]
  - b) Explain Associative Classification. [5]
  - c) List and briefly discuss the Applications of Data Mining [5]

\*\*\*\*\*

1300	(M, A, K, E, Y)
1300	(M, A, K, E)
1300	(D, O, N, K, E, Y)
1300	(M, O, N, K, E, Y)
1300	(M, A, K, E, Y)