**Code No: MC1343/R13**

### MCA IV Semester Regular/ Supplementary Examinations, June-2016
## DATA WAREHOUSING AND MINING

**Time: 3 Hours**                                                                 **Max. Marks: 60**

*Answer Any* **FIVE** *Questions*
*All Questions Carry Equal Marks*

| | | | |
|---|---|---|---|
| 1. | a | What is data preprocessing? Explain in detail. | 6 |
| | b | Given two objects represented by the tuples: (22,1,42,10) and (20,0,36,8), compute the Euclidean distance, Manhanttan distance, Minkowski distance between the two objects using q=3. | 6 |
| 2. | a | Describe schemes for multidimensional databases. | 6 |
| | b | What are summery statistics? Explain | 6 |
| 3. | a | Discuss about problem definition of association analysis. | 6 |
| | b | Explain about evaluating the performance of a classifier. | 6 |
| 4. | | Describe Naïve bayes classifier with examples. | 12 |
| 5. | | Apply Apriori algorithm for the following transaction database and identify all frequent item sets systematically with minsup = 0.3 | 12 |

| Row | Transaction |
|---|---|
| 1 | {a, b} |
| 2 | {b, c} |
| 3 | {a, b, c} |
| 4 | {d, e, f} |
| 5 | {a, b, c} |
| 6 | {d, f} |
| 7 | {c, d, e, f} |
| 8 | {a, b, c, d, e} |

| | | | |
|---|---|---|---|
| 6. | a | Write the basic K-means algorithm and explain. | 6 |
| | b | Describe strengths and weaknesses of K-means. | 6 |
| 7. | a | Discuss in detail about agglomerative hierarchical clustering | 7 |
| | b | What are the key issues in hierarchical clustering? | 5 |
| 8. | a | Explain Functionalities of search engines. | 6 |
| | b | Discuss about mining the web link structure to identify authoritative web pages. | 6 |

*****