

Code No: I0505/R16

M. Tech. I Semester Regular/Supple Examinations, Jan/Feb-2018

**DATA WAREHOUSING AND DATA MINING/
DATA WARE HOUSING AND MINING****Common to Computer Science (05), Computer Science & Engineering (58)
And Neural Networks (69)****Time: 3 Hours****Max. Marks: 60**

*Answer any FIVE Questions
All Questions Carry Equal Marks*

- | | | |
|------|--|----|
| 1. a | What is the need to build a data warehouse? Write and explain various steps to build the data warehouse | 6M |
| b | Specify and explain the reasons for mapping data warehouse system with multiprocessor architecture. | 6M |
| 2. a | How to integrate data mining system with data base and data warehouse system? Explain. | 6M |
| b | Differentiate multi relational OLAP with multi dimensional OLAP. | 6M |
| 3. a | What is OLAP? How to build it? Give the guidelines.. | 6M |
| b | With an example data warehouse explain various schemas used to represent multi dimensional data. | 6M |
| 4. a | What motivated data mining? Why is it important? Explain the components of it | 6M |
| b | Explain the following major issues in data mining | 6M |
| | i) Mining methodology and user interaction | |
| | ii) Performance issues. | |
| 5. a | "Data preprocessing techniques can improve the quality of data"-Justify this statement. | 6M |
| b | Compute Euclidean, Manhattan, Minkowski distance($q=3$) between 2 objects for the given 2 objects represented by tuples (22,1,42,10) and (20,0,36,8) | 6M |

Code No: I0505/R16

6. List all frequent item sets and strong association rules with support 's' and confidence 'c' for the following transaction database 12M
I1:{T1,T4,T5,T7,T8,T9}, I2= {T1,T2,T3,T4,T6,T8,T9}
I3={T3,T5,T6,T7,T8,T9} I4={T2,T4} I5={T1,T8}
7. a Write the process of classification by support vector machines when data is both linearly separable and inseparable. 6M
b Explain the following regressions 6M
i) Linear regression
ii) Multiple linear regression
iii) Nonlinear regression.
8. a "One person's noise could be another person's signal"-Justify this statement with various computer based outlier analysis methods.. 6M
b Explain how partitioning clustering works with k-means and k-medoids algorithm. 6M
