



**LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034**

**M.Sc. DEGREE EXAMINATION – STATISTICS**

**THIRD SEMESTER – NOVEMBER 2015**

**ST 3957 - DATA WAREHOUSING AND DATA MINING**

Date : 11/11/2015  
Time : 09:00-12:00

Dept. No.

Max. : 100 Marks

**PART – A**

**Answer all the questions**

**(10x2=20 Marks)**

1. Define Data mining
2. State any two applications of Data warehousing
3. Provide the goodness of split for i) classification tree ii) Regression tree
4. Define Support and Confidence in Association rule mining
5. Explain transformation step in ETL process
6. State the use of Buzz and Sentiment Analysis
7. Define ROC and AUC
8. Define classifier and state any two non linear classifier
9. Define Data mart and Meta data
10. Define Fact table and Dimension table

**PART – B**

**Answer any FIVE questions**

**(5x8=40 Marks)**

11. Distinguish between OLTP and OLAP
12. Explain Bayesian Network with an example
13. Discuss the steps involved in Chi-square Automated Interaction Detector
14. Discuss the types of OLAP schemas
15. Explain Support Vector Machine
16. a. Define  
i) Association rule                      ii) Support  
iii) Confidence                          iv) Frequent itemset  
b. Explain the Apriori algorithm for obtaining Association rules
17. Explain i) ROC and AUC ii) Classification table ,Sensitivity and Specificity  
iii) Gains chart iv) Optimal cut value for a binary classification problem
18. Explain the steps involved in Adaptive Boosting method

**PART – C**

**Answer any TWO questions**

**(2x20=40 Marks)**

19. Explain ANN structure and Back propagation algorithm with an example
20. i) Explain CRISP-DM Architecture  
ii) Explain ETL process in detail
21. i) Explain KNN Method for regression and classification  
ii) Explain Random Forest method for classification and regression
22. i) Discuss the steps involved in construction of CART **(15)**  
ii) Discuss the steps involved in the construction of Gains chart **(5)**

\*\*\*\*\*