LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – **COMPUTER SCIENCE**

FIRST SEMESTER – **NOVEMBER 2018**

17/18PCS1MC01– DATA MINING

Date: 25-10-2018 Dept. No. Max. : 100 Marks

Time: 01:00-04:00

SECTION- A

Answer All the questions **(10x2=20)**

1. How will you Import from and Export to .CSV Files?
2. How will you save a chart as a .pdf file.
3. What is decision tree?
4. What is regression? Write the expression for linear regression
5. What are outliers?
6. What is the difference between ***pam()*** and ***pamk()?***
7. What are the dissimilarity measures of time series clustering?
8. What are the feature extraction Techniques?
9. Define confidence
10. What is Apriori algorithm?

SECTION – B

Answer All the questions **(5x8=40)**

11. a) How will you explore multiple variables in R and plot them as a chart?

    (OR)

    b) How will import data to R from various file types of Excel files.

12. a)Write down the steps to build random forest for Iris data where Species is the target variable. Plot the tree.

    (OR)

    b)Explain the steps to build decision tree using the packages ***party*** and ***rpart.***

13. a) Explain K- means clustering on Bodymass dataset with three clusters and plot them.

    (OR)

    b) Write down the distance measures used for clustering.

14. a) Explain time series forecasting with Airpassenger dataset.

    (OR)

    b) What is Time series classification?

15. a) How will you remove redundant rules form association rules formed by Apriori?

    (OR)

    b) How will you perform association rule mining in R? Apply with Titanic dataset.

SECTION –C

Answer Any TWO questions **(2x20=40)**

16 i) Explain the data visualization features of R.
ii) Explain linear regression and non linear regression for prediction with a sample data.

17 i) Explain K- Medoid clustering and hierarchical clustering with example.
ii) Explain Time series clustering and time series classification.

18. i) How will you represent the association rules through balloon plot and graphs.
ii) Explain outlier detection through LOF and clustering with example.

****