***Chaitanya Bharathi Institute of Technology***

**Department of Information Technology**

Subject: Data Mining Lab

Set-1

1. Convert .arff file into .csv file using Java-WEKA API
2. Linear Regression using R-Tool

Set-2

1. Convert .csv file into .arff file using Java-WEKA API
2. K-NN in WEKA

Set-3

1. Convert .arff file into .csv file using WEKA
2. Demonstrate the classifier J48 in WEKA and Visualise Decision Tree in WEKA

Set-4

1. Perform K-NN classification of Time series data using R-Tool
2. Perform the following preprocessing operations in WEKA
   1. Attribute selection
   2. Handling missing values
   3. Discretisation
   4. Converting nominal attributes to binary attributes
   5. Standardisation

Set-5

1. Random Forest in R
2. Perform the following preprocessing operations in WEKA
3. Attribute Selection
4. Normalisation
5. Outlier detection
6. Discretisation
7. Handle missing values

Set-21

1. Convert .csv file into .arff file using WEKA
2. Data visualisation and exploration in R
3. Read the dataset into R-Dataframe
4. Check the dimensionality of the chosen dataset
5. Display Variable names or column names
6. Check the Structure of the object
7. Get the last 6 rows
8. Distribution of every dimension
9. Variance of a numeric attribute
10. Scatter plot

(Note your observations, Comment on the data distribution, try plotting commands for different kinds of dimensions, try different plotting function options: symbols, size of plotting symbol, legends, x,y-axis labels, titles of graphs, etc)

Set-20

1. DBSCAN in WEKA
2. Data visualisation and exploration in R
3. Read the dataset into R-Dataframe
4. Get second attributes of the first 10 rows
5. Covariance of two attributes
6. Pair Plot
7. Box-Whisker Plots
8. Frequency of each class type
9. Density
10. Line charts for both numeric and categorical dimensions

(Note your observations, Comment on the data distribution, try plotting commands for different kinds of dimensions, try different plotting function options: symbols, size of plotting symbol, legends, x,y-axis labels, titles of graphs, etc)

Set-19

1. Hierarchical Clustering in WEKA
2. Data visualisation and exploration in R
3. Read the dataset into R-Dataframe
4. Get the first 5 rows
5. Correlation of two dimensions
6. Histogram of an attribute
7. Cleveland Dot Charts
8. Bar Charts
9. Pie chart
10. Line charts for both numeric and categorical dimensions

(Note your observations, Comment on the data distribution, try plotting commands for different kinds of dimensions, try different plotting function options: symbols, size of plotting symbol, legends, x,y-axis labels, titles of graphs, etc)

Set-13

1. Generate Association rules using Apriori in R-Tool
2. Adaboost in R

Set-14

1. Generate Association rules using FP-Growth in R-Tool
2. K-NN in R

Set-18

1. Generate Association rules using ECLAT in R-Tool
2. Bagging in R

Set-15

1. Build a Decision Tree Classifier in R-Tool using the packages Party.
2. Hierarchical Clustering in R

Set-17

1. Build a Decision Tree Classifier in R-Tool using the packages rpart.
2. K-Means in R

Set-16

1. Build a Decision Tree Classifier in R-Tool using the packages caret.
2. Naive Bayes in R

Set-10

1. DBSCAN in R
2. Generate association rules using Apriori in WEKA

Set-12

1. Write a program in R-tool for displaying word cloud
2. Generate association rules using Apriori and FP-Growth in WEKA

Set-11

1. Write a program in R-Tool for performing sentiment analysis in twitter data
2. K-Means in WEKA

Set-7

1. Perform time series decomposition and forecasting in R
2. Adaboost in WEKA

Set-9

1. Perform hierarchical clustering on time series data in R
2. Random Forest in WEKA

Set-8

1. Classify Time series data using R-tool
2. Bagging in WEKA

Set-6

1. Analyse time series data using Dynamic Time Warping using R-Tool
2. Naive Bayes in WEKA