

SUBJECT	: DATA MINING - PRACTICAL				
SEMESTER	: VI			L T P C	
SUBJECT CODE: 17CSU612A		CLASS	: III B.Sc.CS A & B	4 0 0 2	

LIST OF PRACTICALS

1. Use the following learning schemes, with the default settings to analyze the weather data (in weather.arff). for test options, first choose “Use training set”, then choose “Percentage split” using default 66% percentage split. Report model percent error rate.
2. Use iris dataset preprocess and classify it with j4.8 and Naive Bayes classifier. Examine the tree in the classifier output panel.
3. Using the dataset ReutersCorn – Train and ReutersGrain – Train. Classify articles using binary attributes and word count attributes.
4. Apply any two association rule based algorithm for the supermarket analysis.
5. Using weka experimenter perform comparison analysis of j4.8, oneR and ID3 for vote dataset.
6. Using weka experimenter perform comparison analysis of Naive Bayes with different datasets.
7. Apply ZeroR, OneR and j4.8, to classify the iris data in an experiment using 10 train and test runs, with 66% of the data used for 34% used for testing.
8. Using Weka Knowledge flow set up a flow to load an ARFF file (batch mode) and perform a cross-validation using j4.8 (WEKS’s C4.5 implementation).
9. Draw multiple ROC curves in the same plot window, using j4.8 and RandomForest as classifiers.
10. Use any three clustering algorithm on Vehicle data set and find best among them.