Experiment-6

Aim:- Demonstrate knowledge flow application on data sets

→ Develop a knowledge flow layout for finding strong association rules by using Apriori, FP Growth algorithms.

→ set up the knowledge flow to load an ARFF and perform a Cross validating using J48 algorithm.

→ Demonstrate plotting multiple ROC curves in the same plot window by using j48 and Random forest tree.

Input:-

Weather.nominal.arff is a predefined data set.

Analysis:-

classification algorithms - J48 algorithm,

Random forest tree.

Association rules by using Apriori.

Objective:-

The knowledge flow provides an alternative way to the explorer as a graphical front end to WEKA's algorithm. Knowledge flow is a working progress. so, some of the functionality from explorer is not yet available.

Knowledge Flow:-

The knowledge Flow presents a data-flow inspired interface to Weka. the User can select Weka components from a tool bar, place them on a layout canvas and connect them together in order to form a 'knowledge flow' for processing and analyzing data.

Arff Viewer:-

The Arff Viewer is a little tool for viewing ARFF files in a tabular format. The advantage of this kind of display over the file representation is that attribute name, type and data are directly associated in columns and not separated in definition and data part.

Association Rule:-

Association rules are mined out after frequent itemsets in a big dataset are found. These datasets are found out using mining algorithms such as Apriori and FP Growth.

Apriori Algorithm:-

Apriori is the simple algorithm, which applied for mining of repeated the patterns from the transaction dataset to find frequent itemsets and association between various item sets.

FP Growth algorithm:-

FP Growth is an algorithm for finding patterns in data and it's much more efficient than its predecessor. Weka implementation of FP Growth requires data be supplied in binary format: 0 or 1.

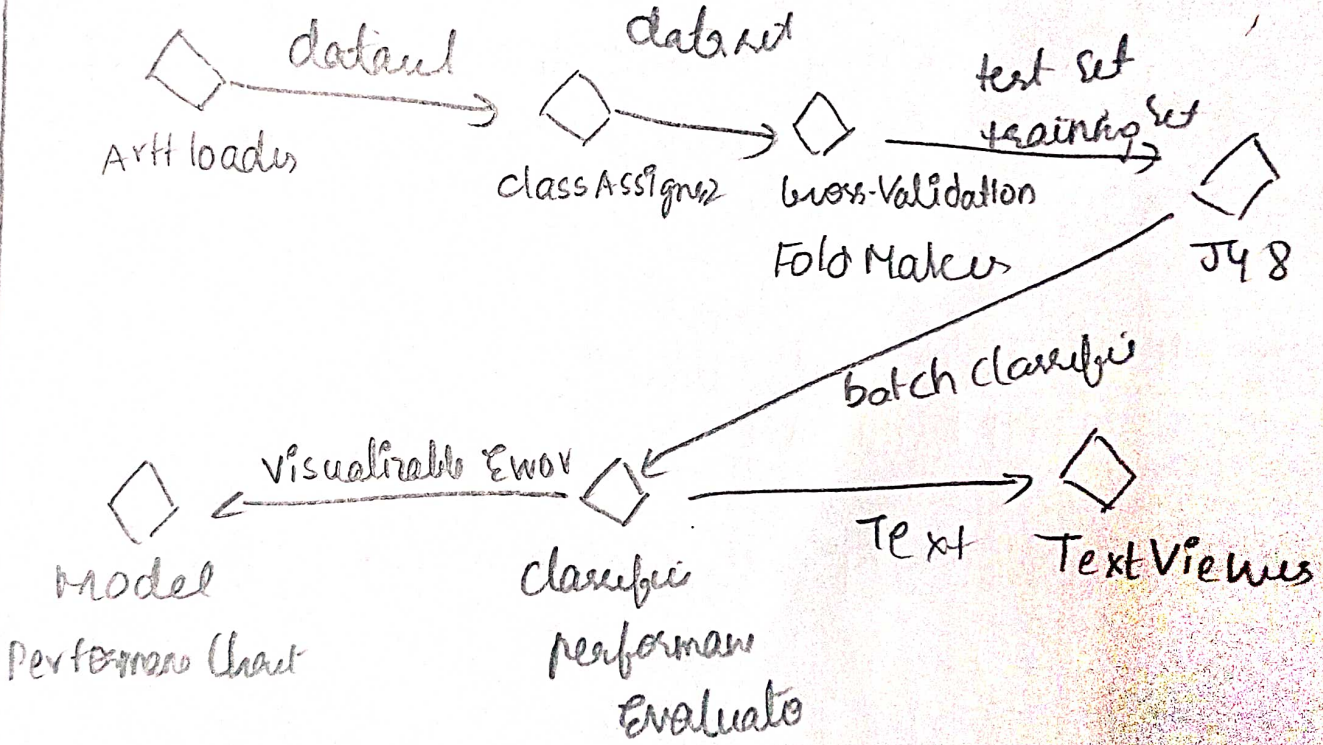
Roc Curve:-

The Weka explorer enables you to plot the ROC (Receiver operating characteristic) curve for a certain class label of dataset, run a classifier on a dataset, right-click in the result list on the result you want to display the curve for. Select Visualize threshold curve and choose.

Cross-Validation:-

Weka does stratified cross-validation by default. And with 10-fold cross validation. Weka invokes the learning algorithm.

output:-


dataset
dataset
test set
training set
ArffLoader
ClassAssigner
Cross-Validation
Fold Maker
J48
batch classifier
Visualizable Error
Text
TextViewer
model
Performance Chart
classifier
performance
evaluator