

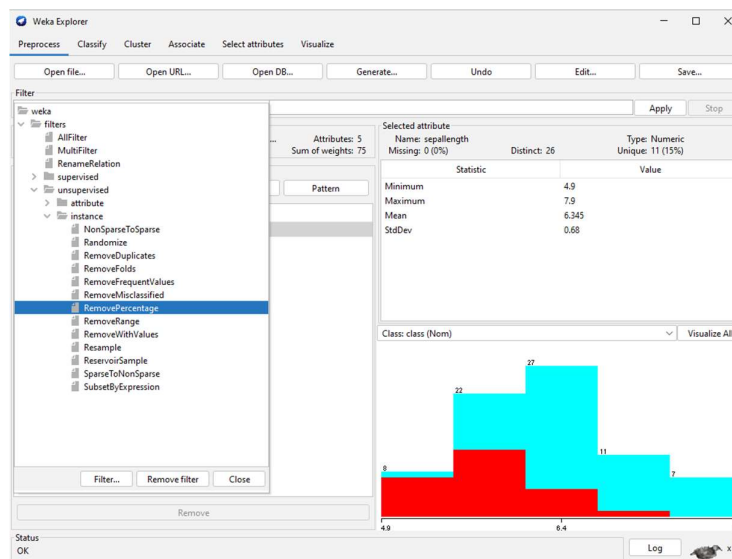
## Practical No 5

16/03/23

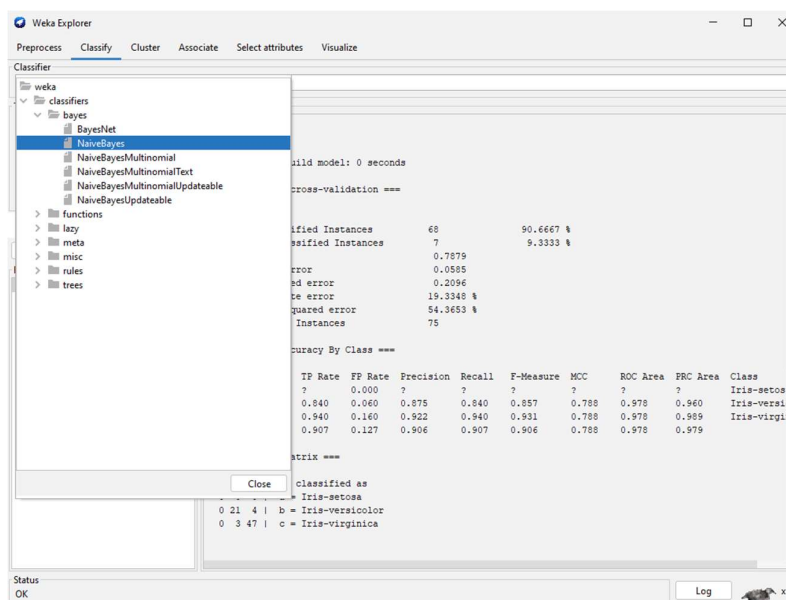
**Aim:** Pre-process the given dataset and hence classify the resultant data set using statistical based classifiers. Interpret the result

Steps:

1. Open **Weka** > click **Explorer** > click **open file** > select file **iris.arff** .
2. Click on choose > filter > unsupervised > attribute > RemovePercentage and on apply



3. Click on **Classify** tab, click on **choose** > **bayes** > select **NaiveBayes**.



#### 4. Click on **Start**.

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The 'Classifier' dropdown is set to 'NaiveBayes'. Under 'Test options', 'Cross-validation' is selected with 'Folds' set to 10. The 'Start' button is visible. The 'Classifier output' pane displays the following information:

```
=== Run information ===
Scheme:      weka.classifiers.bayes.NaiveBayes
Relation:    iris-veka.filters.unsupervised.instance.RemovePercentage-P50.0
Instances:   75
Attributes:  5
  sepallength
  sepalwidth
  petallength
  petalwidth
  class
Test mode:   10-fold cross-validation

=== Classifier model (full training set) ===

Naive Bayes Classifier

      Class
Attribute  Iris-setosa Iris-versicolor Iris-virginica
              (0.01)      (0.33)      (0.65)
=====
sepallength
mean       0          5.8752      6.5832
std. dev.  0.02       0.4721      0.6336
weight sum 0          25         50
precision  0.12       0.12       0.12

sepalwidth
mean       0          2.7657      2.9806
std. dev.  0.019     0.2646      0.3124
weight sum 0          25         50
precision  0.1143    0.1143     0.1143
```

The status bar at the bottom shows 'Status OK' and a 'Log' button.

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The 'Classifier' dropdown is set to 'NaiveBayes'. Under 'Test options', 'Cross-validation' is selected with 'Folds' set to 10. The 'Start' button is visible. The 'Classifier output' pane displays the following information:

```
Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      68          90.6667 %
Incorrectly Classified Instances     7           9.3333 %
Kappa statistic                    0.7879
Mean absolute error                  0.0585
Root mean squared error              0.2096
Relative absolute error             19.3348 %
Root relative squared error         54.3653 %
Total Number of Instances          75

=== Detailed Accuracy By Class ===
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      ?         ?         ?         ?         ?         ?         ?         ?         Iris-setosa
0.840  0.060  0.875  0.840  0.857  0.788  0.978  0.960  Iris-versi.
0.940  0.160  0.922  0.940  0.931  0.788  0.978  0.989  Iris-virgi.
Weighted Avg.  0.907  0.127  0.906  0.907  0.906  0.788  0.978  0.979

=== Confusion Matrix ===
  a  b  c  <-- classified as
  0  0  0 | a = Iris-setosa
  0 21  4 | b = Iris-versicolor
  0  3 47 | c = Iris-virginica
```

The status bar at the bottom shows 'Status OK' and a 'Log' button.