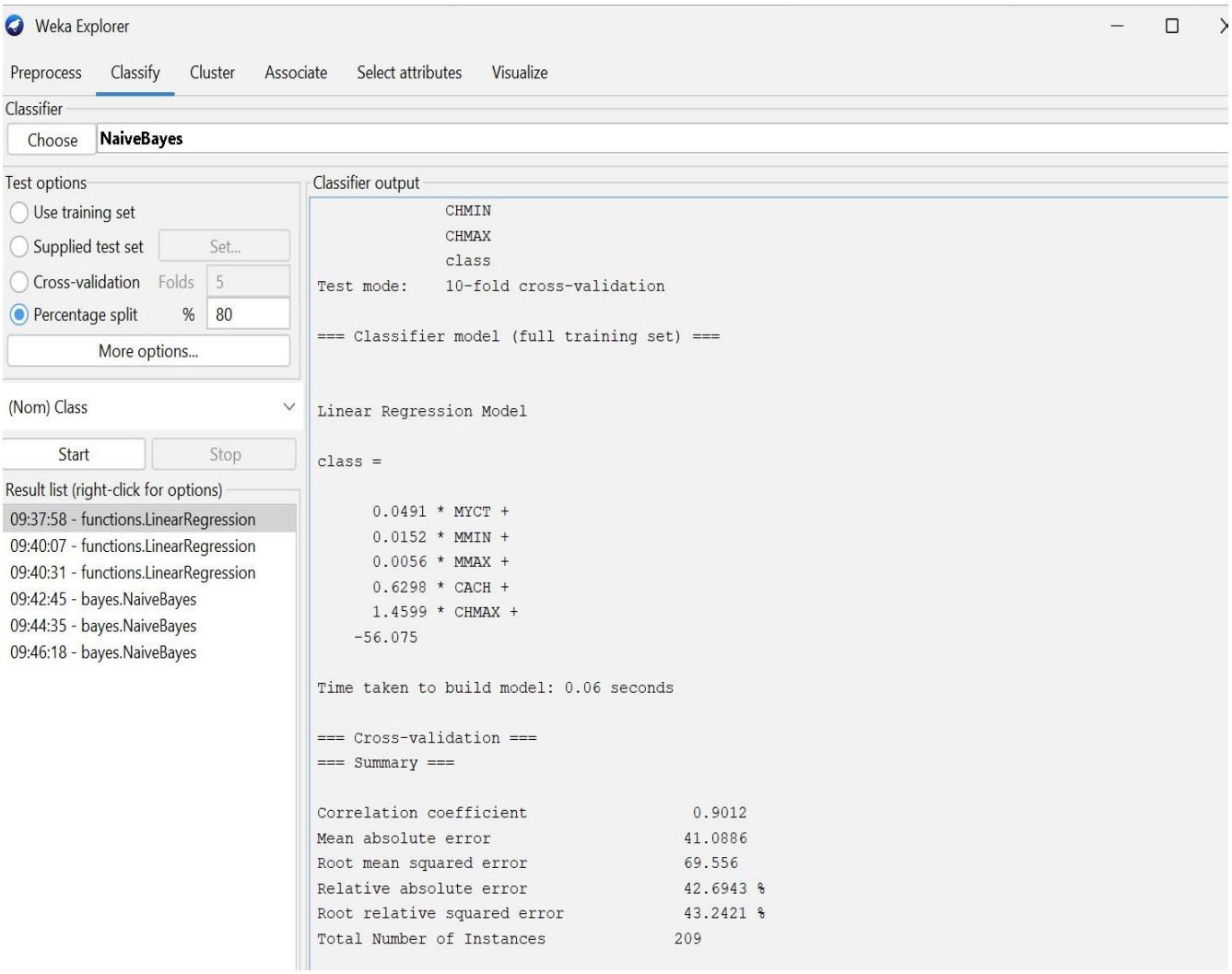
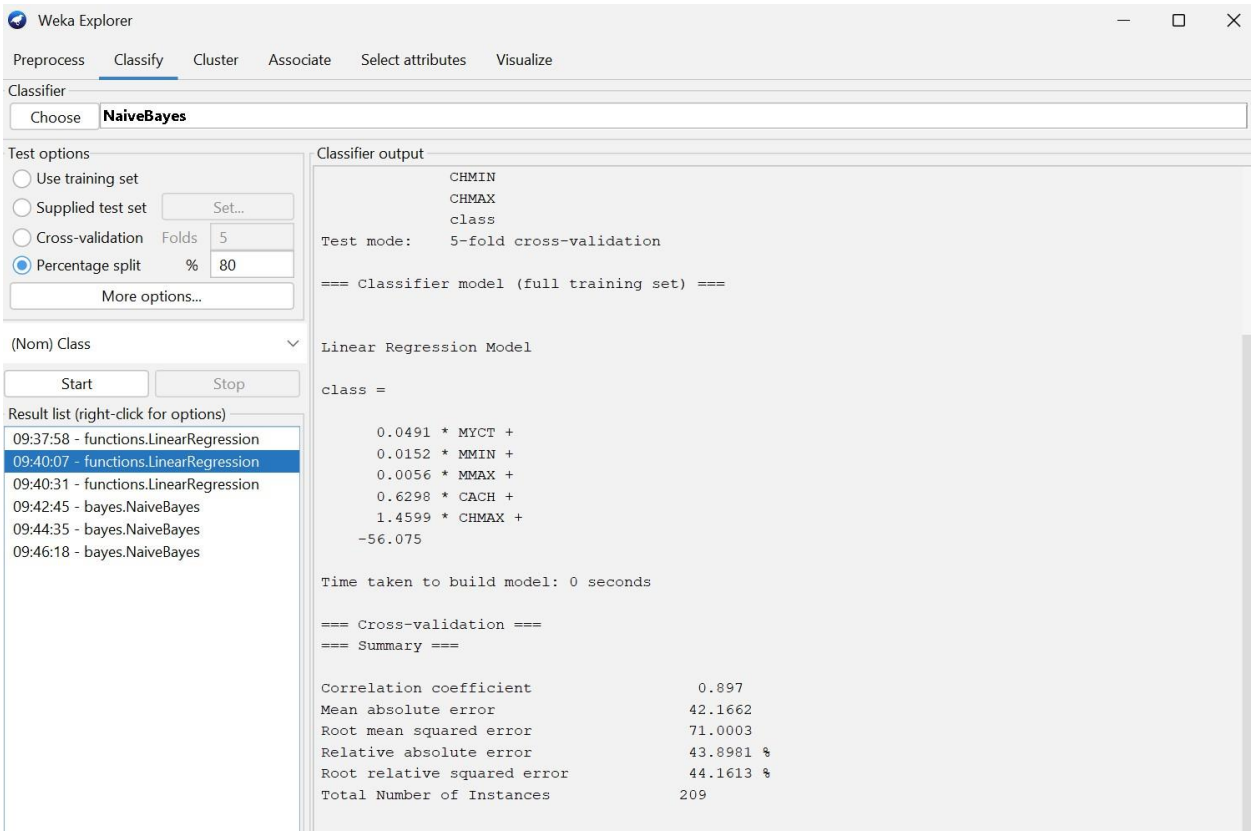
EXPERIMENT-1 **WEKA**

1. **Linear Regression-**Linear regression aims to fit a linear equation to observed data given by:

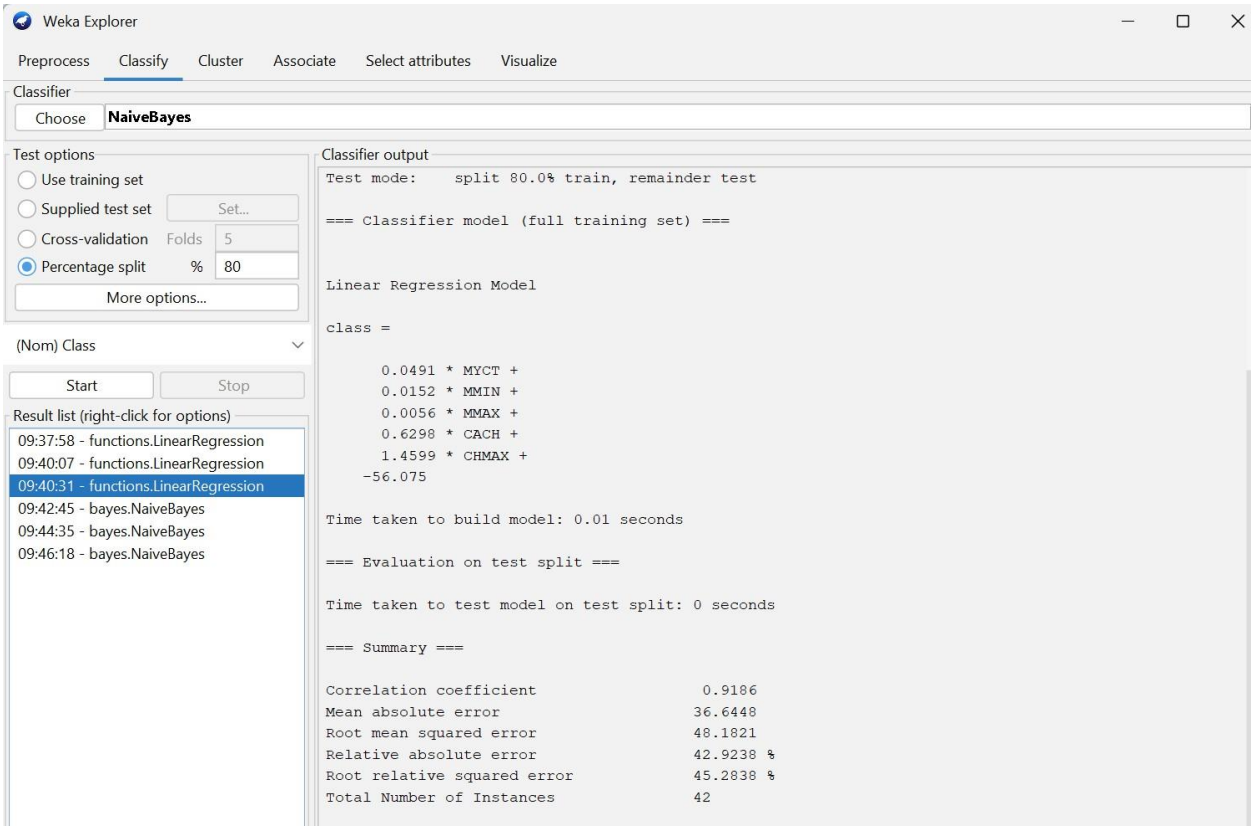
Where: y and x are the dependent and independent variables



2.

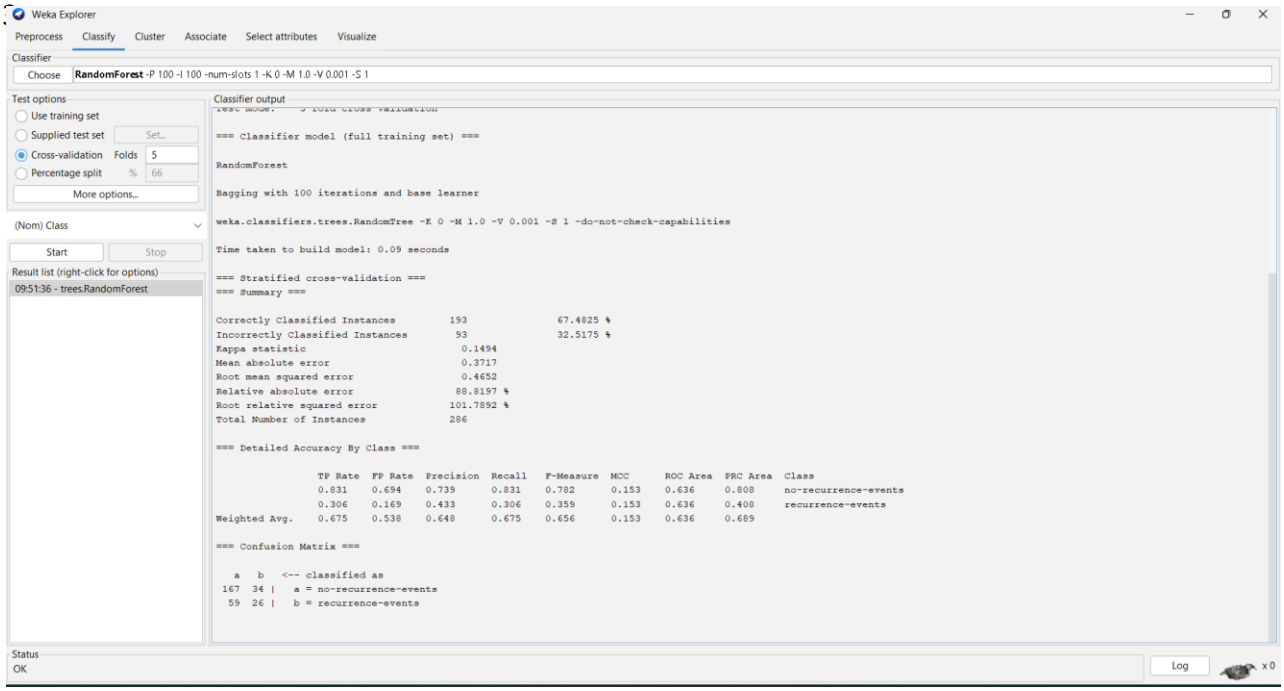

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
NaiveBayes
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
5
Percentage split
%
80
More options...
(Nom) Class
Start
Stop
Result list (right-click for options)
09:37:58 - functions.LinearRegression
09:40:07 - functions.LinearRegression
09:40:31 - functions.LinearRegression
09:42:45 - bayes.NaiveBayes
09:44:35 - bayes.NaiveBayes
09:46:18 - bayes.NaiveBayes
Classifier output
CHMIN
CHMAX
class
Test mode: 5-fold cross-validation
=== Classifier model (full training set) ===
Linear Regression Model
class =
0.0491 * MYCT +
0.0152 * MMIN +
0.0056 * MMAX +
0.6298 * CACH +
1.4599 * CHMAX +
-56.075
Time taken to build model: 0 seconds
=== Cross-validation ===
=== Summary ===
Correlation coefficient 0.897
Mean absolute error 42.1662
Root mean squared error 71.0003
Relative absolute error 43.8981 %
Root relative squared error 44.1613 %
Total Number of Instances 209

3.

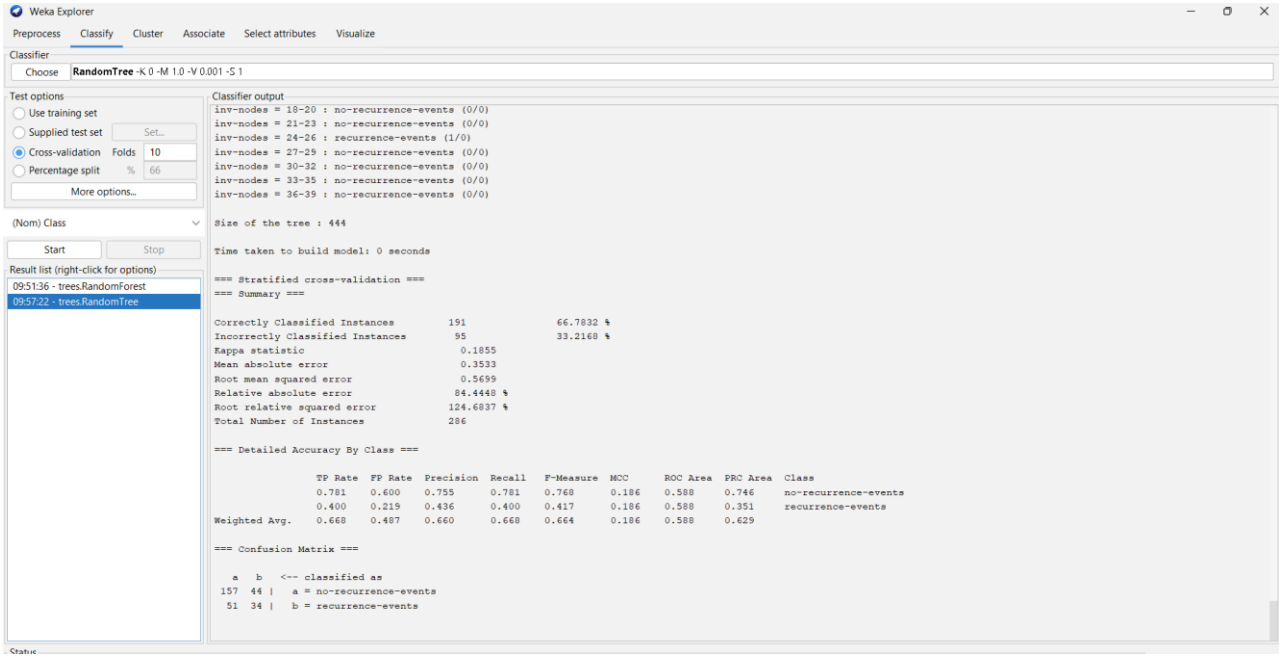


4. The Naïve Bayes classifier is a supervised machine learning algorithm that is used for classification tasks such as text classification. They use principles of probability to perform

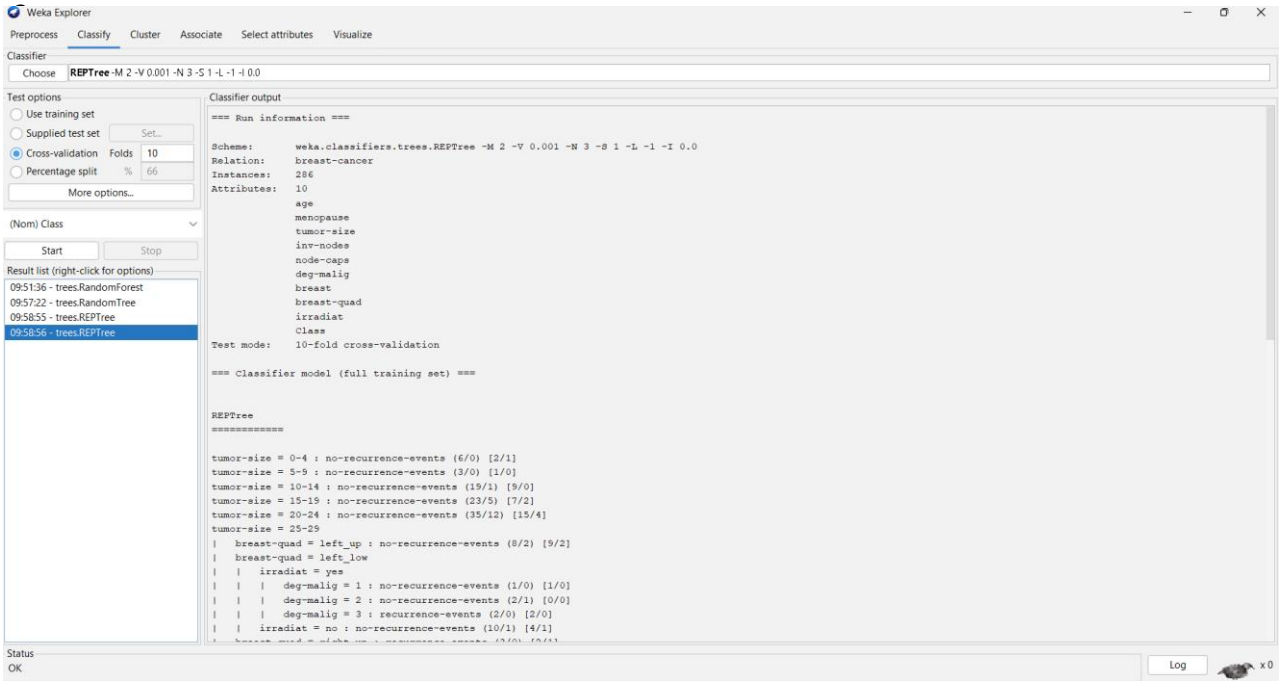
Classifiers-



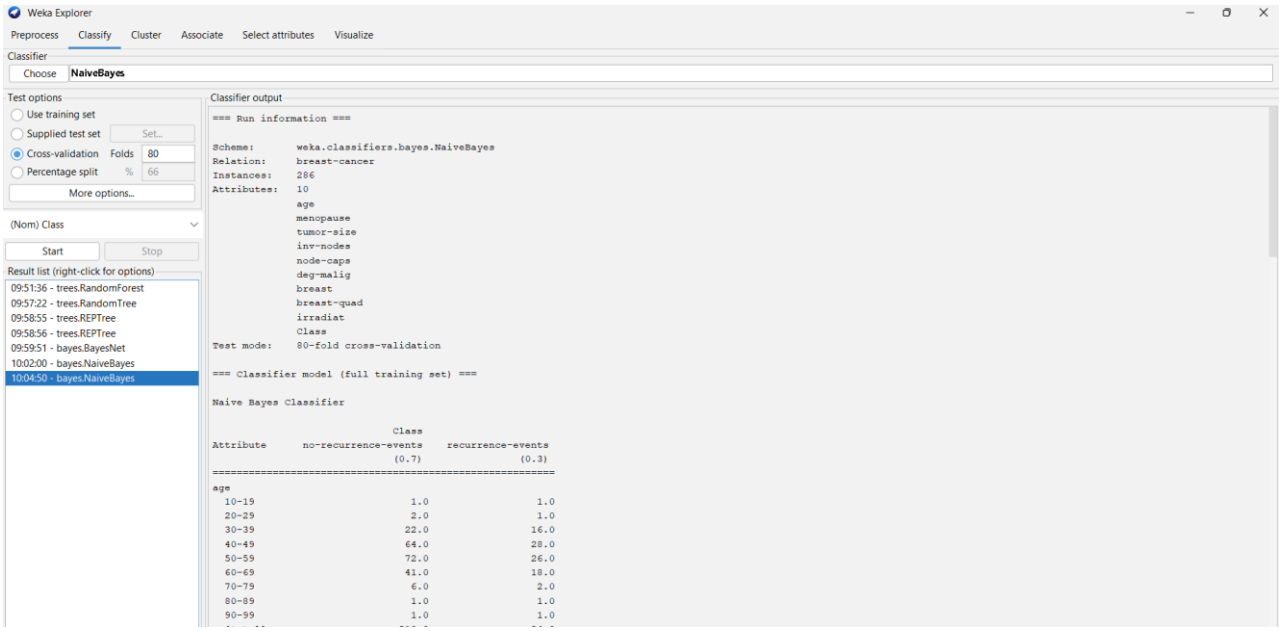
Tree.random forest
Fold – 5



Tree.random.tree
Fold – 10

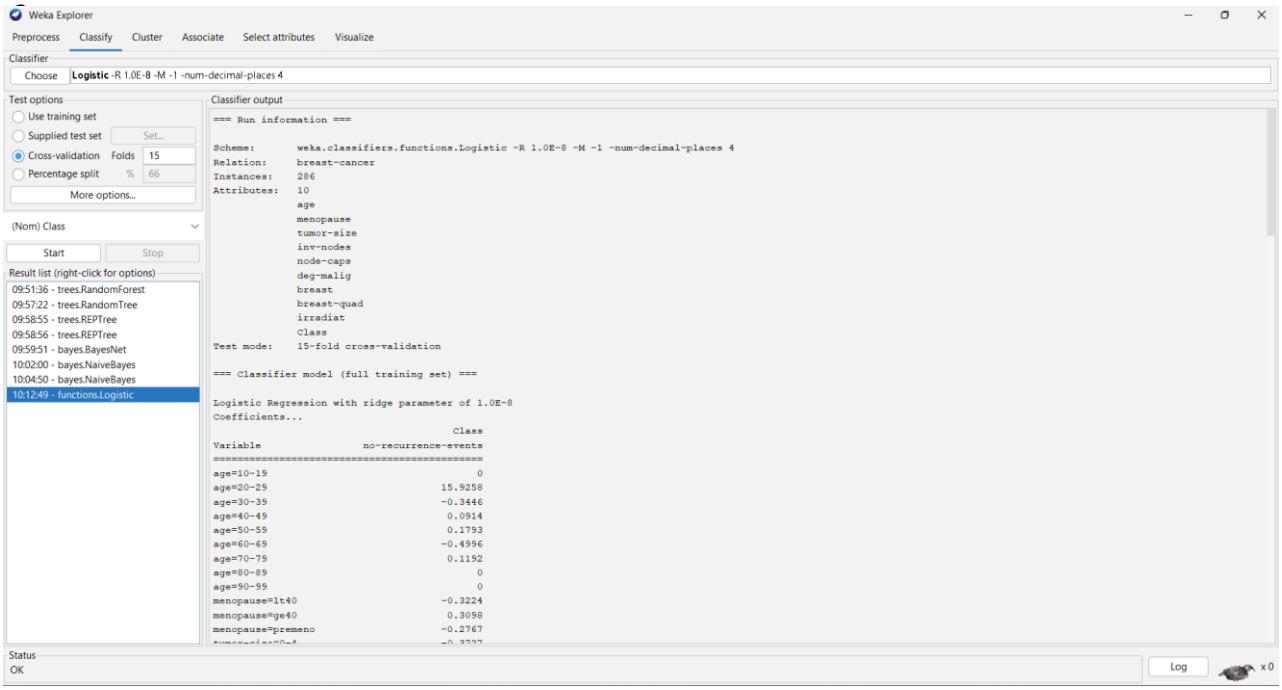


Tree_reptTree
Fold-10



bayers.NaiveBayers

Fold - 80



Function.logistic
Fold- 15

TP Rate: rate of true positives (instances correctly classified as a given class)

FP Rate: rate of false positives (instances falsely classified as a given class)

Precision: proportion of instances that are truly of a class divided by the total instances classified as that class

Recall: proportion of instances classified as a given class divided by the actual total in that class (equivalent to TP rate)

F-Measure: A combined measure for precision and recall calculated as 2 * Precision * Recall / (Precision + Recall) MCC is used in machine learning to measure the quality of binary (two-class) classifications. It considers true and false positives and negatives and is generally regarded as a balanced measure which can be used even if the classes are of
very different sizes

ROC(Receiver Operating Characteristics) area measurement: One of the most important values output by Weka. They give you an idea of how the classifiers are performing in general