

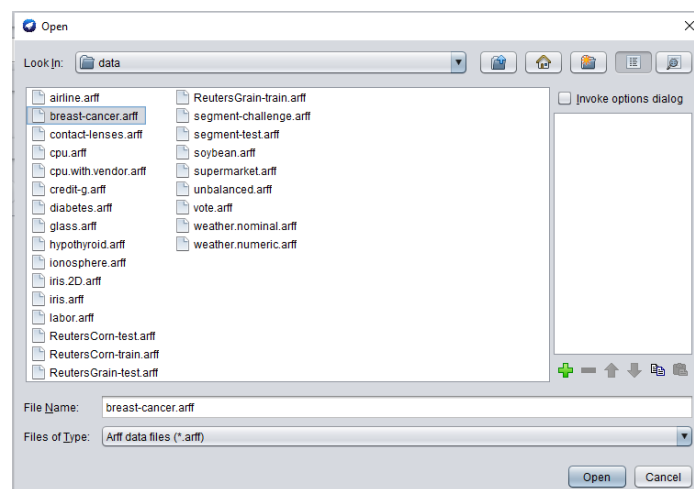
Building Naive Bayesian classifier with WEKA

The use of the Naive Bayesian classifier in Weka is demonstrated in this assignment. The “weather-nominal” data set used in this experiment is available in ARFF format. This assignment assumes that the data has been properly preprocessed.

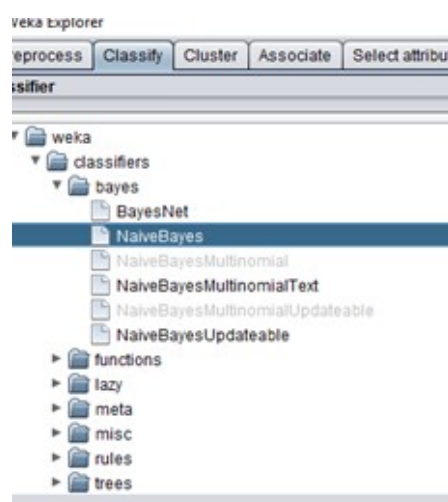
The Bayes’ Theorem is used to build a set of classification algorithms known as Naive Bayes classifiers. It is a family of algorithms that share a common concept, namely that each pair of features being classified is independent of the others.

Steps to be followed:

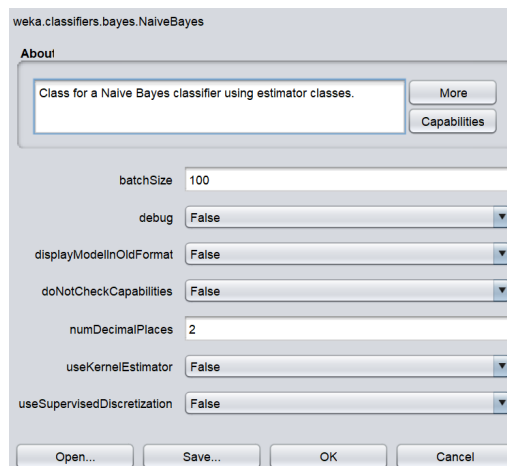
- Initially, we have to load the required dataset in the weka tool using choose file option. Here we are selecting the weather-nominal dataset to execute.



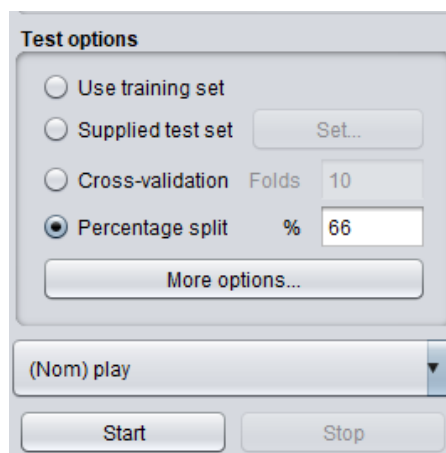
- Now we have to go to the classify tab on the top left side and click on the choose button and select the Naive Bayesian algorithm in it.



- Now to change the parameters click on the right side at the choose button, and we are accepting the default values in this example.



We choose the Percentage split as our measurement method from the “Test” choices in the main panel. Since we don’t have a separate test data collection, we’ll use the percentage split of 66 percent to get a good idea of the model’s accuracy. Our dataset contains 14 examples, with h9 being used for training and 5 being used for testing.



- To generate the model, we now click “start.” When the model is done, the evaluation statistic will appear in the right panel.

```
Time taken to build model: 0 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

=== Summary ===

Correctly Classified Instances      3          60  %
Incorrectly Classified Instances    2          40  %
Kappa statistic                     0
Mean absolute error                 0.4437
Root mean squared error             0.5023
Relative absolute error             93.8582 %
Root relative squared error        102.2471 %
Total Number of Instances          5

=== Detailed Accuracy By Class ===
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1.000	1.000	0.600	1.000	0.750	?	0.667	0.867	yes
	0.000	0.000	?	0.000	?	?	0.667	0.583	no
Weighted Avg.	0.600	0.600	?	0.600	?	?	0.667	0.753	

```

=== Confusion Matrix ===

a b  <-- classified as
3 0 | a = yes
2 0 | b = no

```