

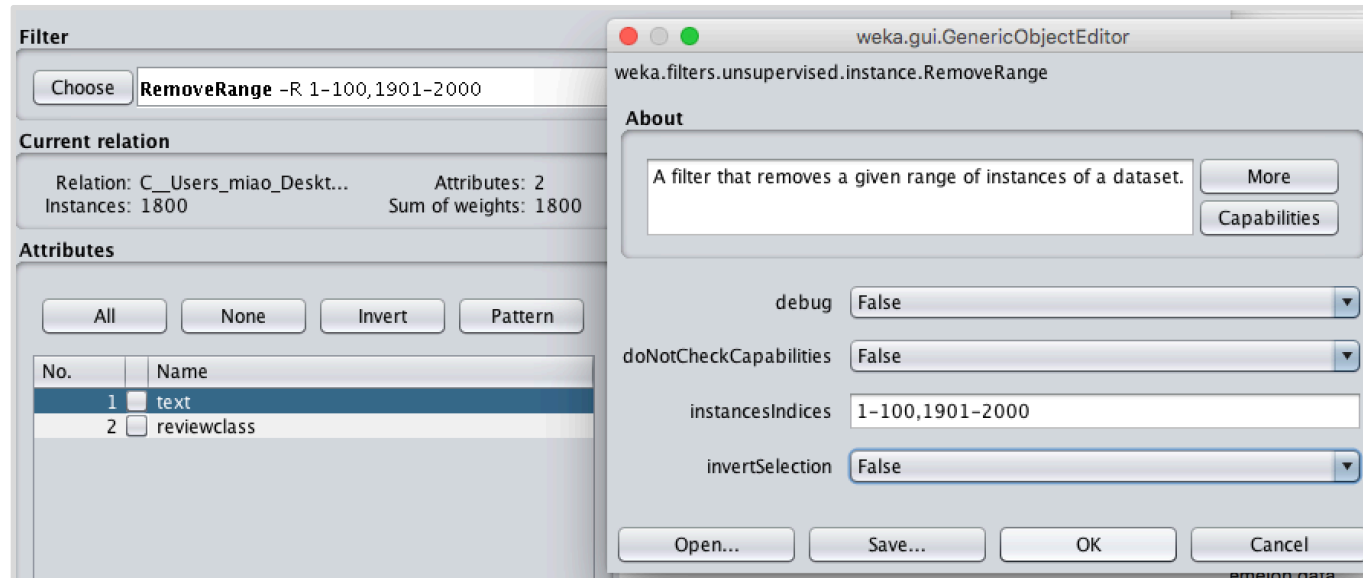


Weka Demo: Training Error and Test Error

School of Information Studies
Syracuse University

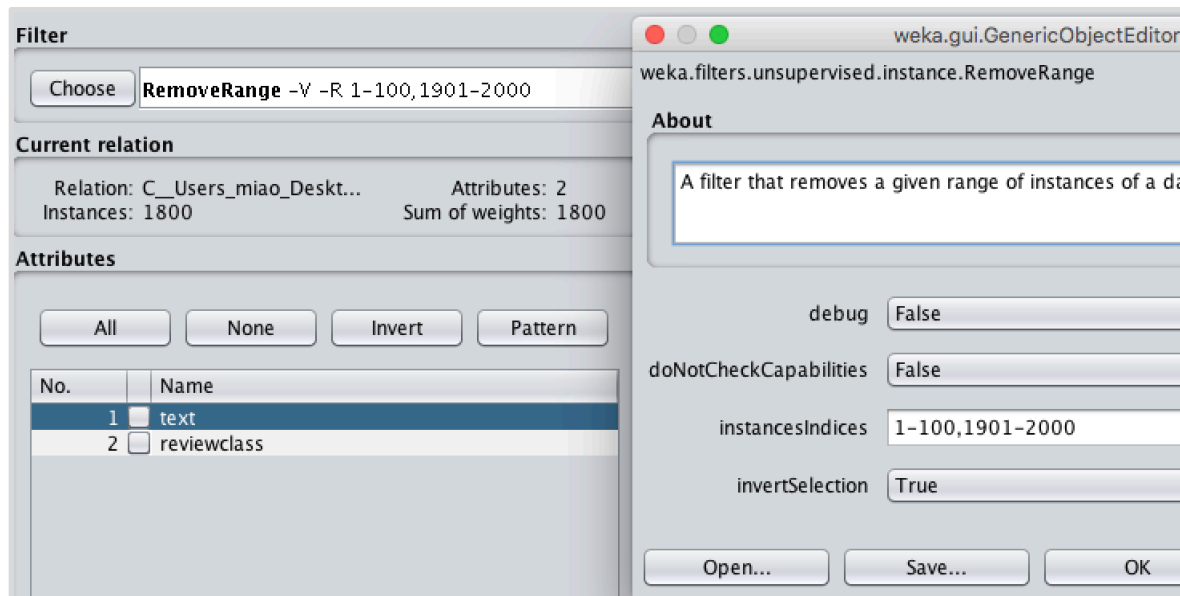
Exercise 1: Training vs. Test Error

First, create a training data set that includes 1,800 examples by removing the first 100 and the last 100 examples using Weka's RemoveRange filter (under filters -> unsupervised -> instance). Save to a new data file "training1800.arff."



Exercise 1: Training vs. Test Error

Second, undo to go back to the original data set of 200 examples. Create a test data set that includes the first 100 and the last 100 examples by turning the “invertSelection” option to “TRUE” in the RemoveRange filter. Save to a new file “test200.arff.”



Exercise 1: Training vs. Test Error

Then, load in “train1800.arff,” and then build and test the MNB model by choosing “use training set” option. Report your training error.

Then, change the test option to “supplied test set” and choose “test200.arff” as test data. Report your test error.

How much difference between training and test error for the same model built on the 1,800 training examples?