Regression Model Tuning

1. Baseline Performance of Each Algorithm using Default Settings:

Baseline Performance (Root Relative Squared Error) on Test Set using All Default Settings

	LWL with Linear Regression Base Classifier	SMOreg with PolyKernel
Test Set – Fold 1	33.6934 %	45.3568 %
Test Set – Fold 2	32.5823 %	13.2175 %
Test Set – Fold 3	29.3881 %	16.2305 %
Test Set – Fold 4	22.4736 %	25.3912 %
Test Set – Fold 5	44.6414 %	52.8507 %
Test Set Average	**32.5558 %**	**30.6093 %**

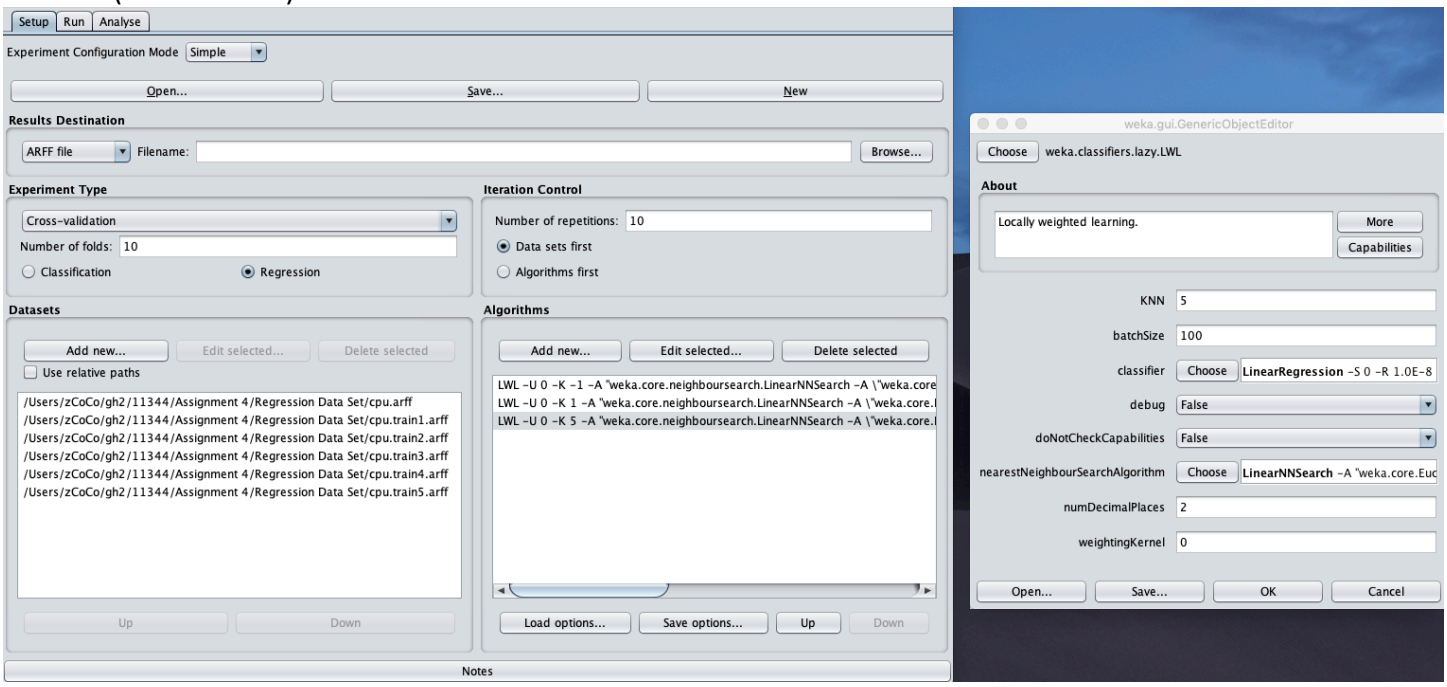
- For each fold, the fold's training set was loaded in the Preprocess tab of Weka Explorer, then the model's performance was assessed by classifying the fold's test set as the supplied test set in the Classify tab.
- Note: The average performance values determined here could have also been obtained by just using 5-fold cross-validation directly since no choices regarding model settings were made on each fold. I chose to do it this way to get more experience with the method.

2.

<u>For Locally Weighted Learning:</u>

Stage 1: Chosen parameter for tuning is the k of the K Nearest Neighbors used by LWL to choose the subset of the space to apply the base classifier over. This was chosen because with a small k the prediction will be heavily influenced by noise, with a moderate k it will be less susceptible to noise, but with a large k (all values) the model is essentially just applying its base classifier over the entire space.

Chosen setting values to tune over were the default value k=-1 (all neighbors), k=1 (just the nearest neighbor), and k=5 (used in class).



Stage 3:

First, the performance of each algorithm configuration was tested for each fold of the training set using the Weka Experiment Environment. For each fold of the training set, the setting which created the configuration which gave the best (lowest root relative squared) performance was chosen as the optimal setting. The model was then configured with this setting in Weka Explorer, trained against the relevant fold of the training data, and its performance was tested against the relevant fold of the test data.

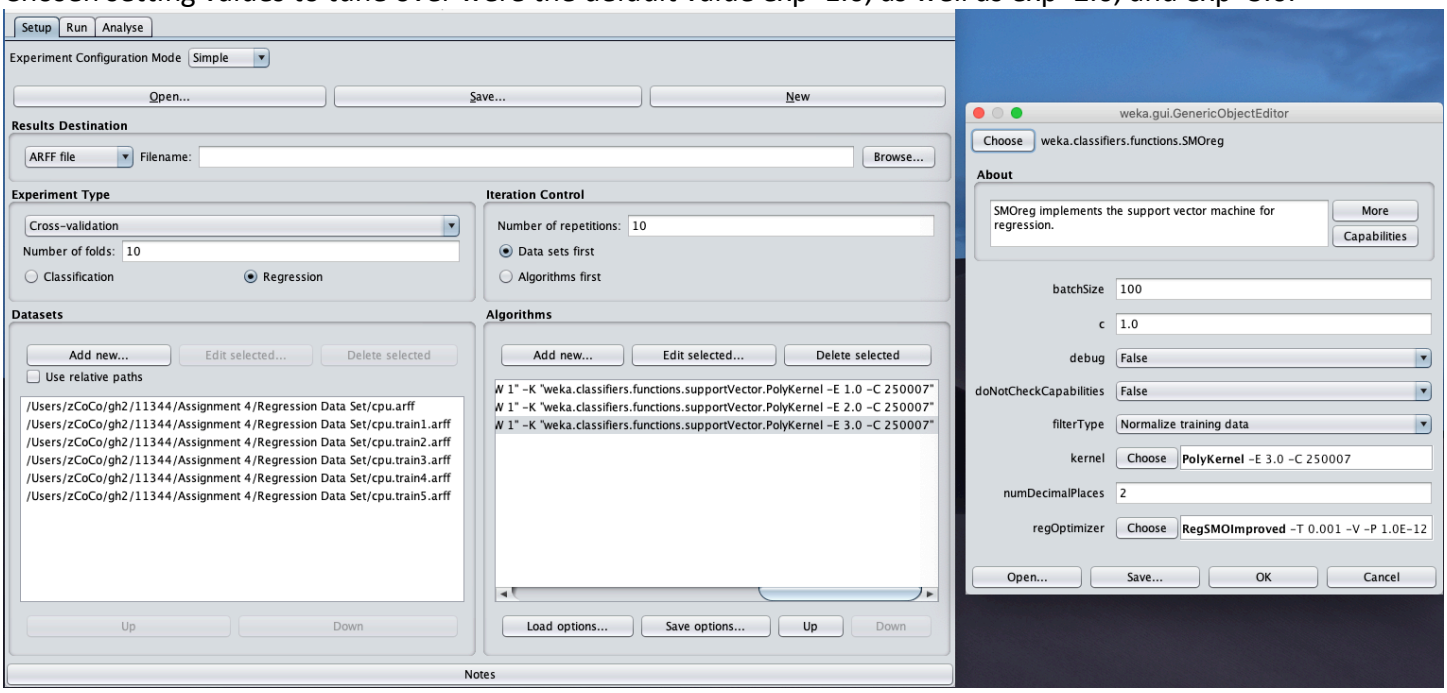
The outcomes of these tests are presented in the following table.

	Training Set Performance (RRSE) for k=-1	Training Set Performance (RRSE) for k=1	Training Set Performance (RRSE) for k=5	Optimal Setting	Test Set Performance using Optimal Setting
Fold 1	24.58	22.37	21.69	k=5	63.2805 %
Fold 2	25.90	22.46	25.07	k=1	13.9935 %
Fold 3	26.52	25.15	26.46	k=1	15.7612 %
Fold 4	26.65	25.28	25.11	k=5	189.9755 %
Fold 5	22.71	23.04	25.39	k=-1	44.6414 %
				Average:	**65.5304 %**

For Support Vector Machine for Regression (SMOreg):

Stage 1: Chosen parameter for tuning is the exponent of the polynomial kernel (PolyKernel) used by the SVM. This was chosen because changing the polynomial exponent will change the SVM's mapping between the feature space and the space in which the SVM's marginal hyperplane is drawn, thus allowing it to fit different "shapes" (polynomial orders) within the feature space.

Chosen setting values to tune over were the default value exp=1.0, as well as exp=2.0, and exp=3.0.

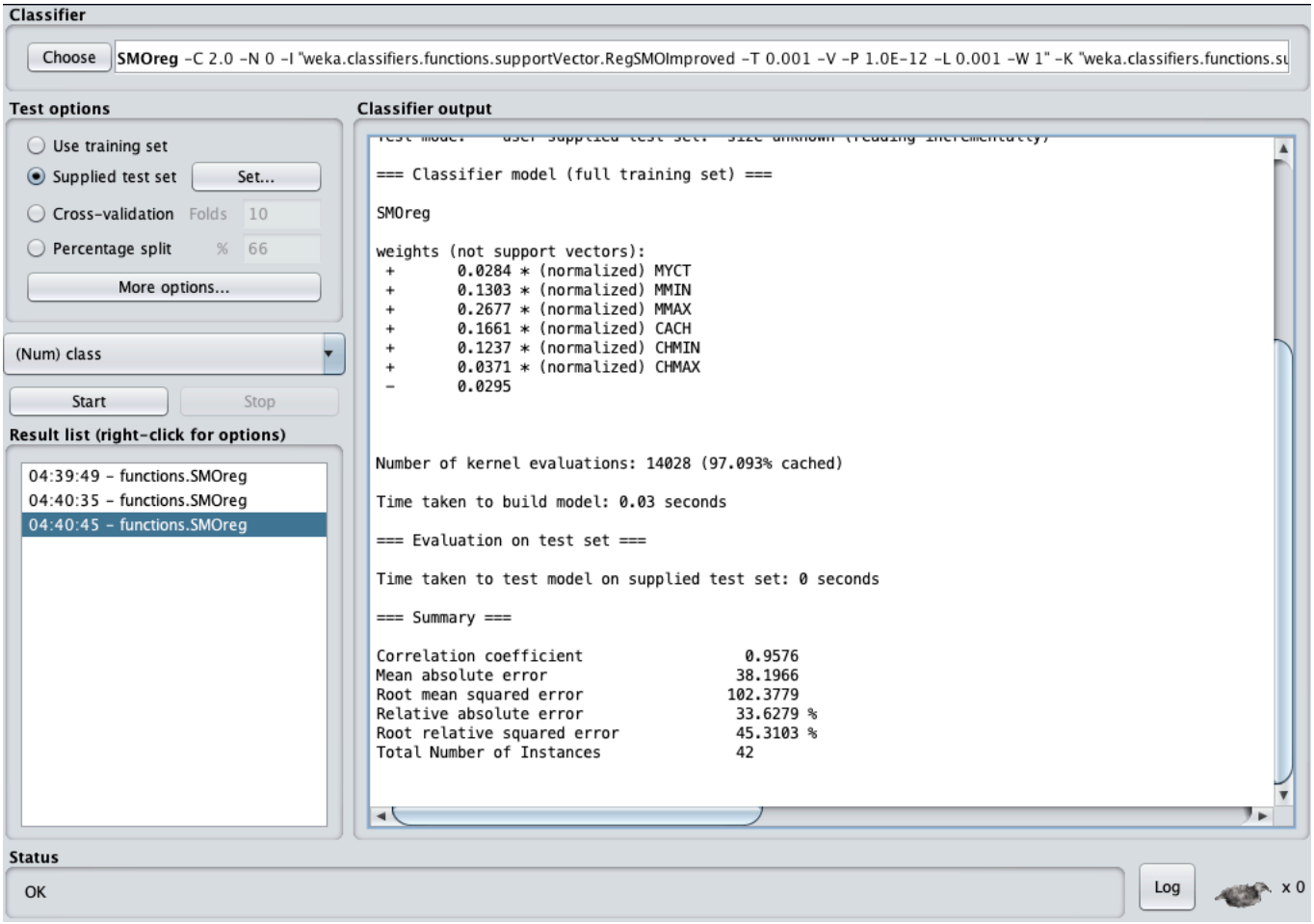


Stage 3:

First, the performance of each algorithm configuration was tested for each fold of the training set using the Weka Experiment Environment. For each fold of the training set, the setting which created the configuration which gave the best (lowest root relative squared) performance was chosen as the optimal setting. The model was then configured with this setting in Weka Explorer, trained against the relevant fold of the training data, and its performance was tested against the relevant fold of the test data.

The outcomes of these tests are presented in the following table.

	Training Set Performance (RRSE) for exp=1	Training Set Performance (RRSE) for exp=2	Training Set Performance (RRSE) for exp=3	Optimal Setting	Test Set Performance using Optimal Setting
Fold 1	29.84	9.28	28.57	exp=2	45.3103 %
Fold 2	28.17	12.47	58.61	exp=2	13.4065 %
Fold 3	28.98	13.16	60.06	exp=2	16.1484 %
Fold 4	26.81	11.70	55.54	exp=2	25.2143 %
Fold 5	23.15	16.59	72.44	exp=2	53.4160 %
				Average:	**30.6991 %**



3. For the LWL model, it was not worth it to perform tuning since the tuned data-set exhibited significantly worse performance on the test dataset than the model with the default parameters (relative root squared error of 65.5304% tuned vs 32.5558% default). Alternatively, tuning could have been performed better by choosing different, more, or a wider range of setting values to test.
For the SVMreg model, it was not worth it to performing tuning since, although the tuned data-set exhibited marginally better performance on the test dataset than the model with the default parameters (RRSE of 30.6093% tuned vs 30.6991% default), the difference was only 0.29%.

Classification Model Tuning

The algorithm chosen to tune was Locally Weighted Learning with Naïve Bayes as the Base Classifier.

1. Baseline Performance using Default Settings:

Baseline Performance (Percent Correct) on Test Set using All Default Settings

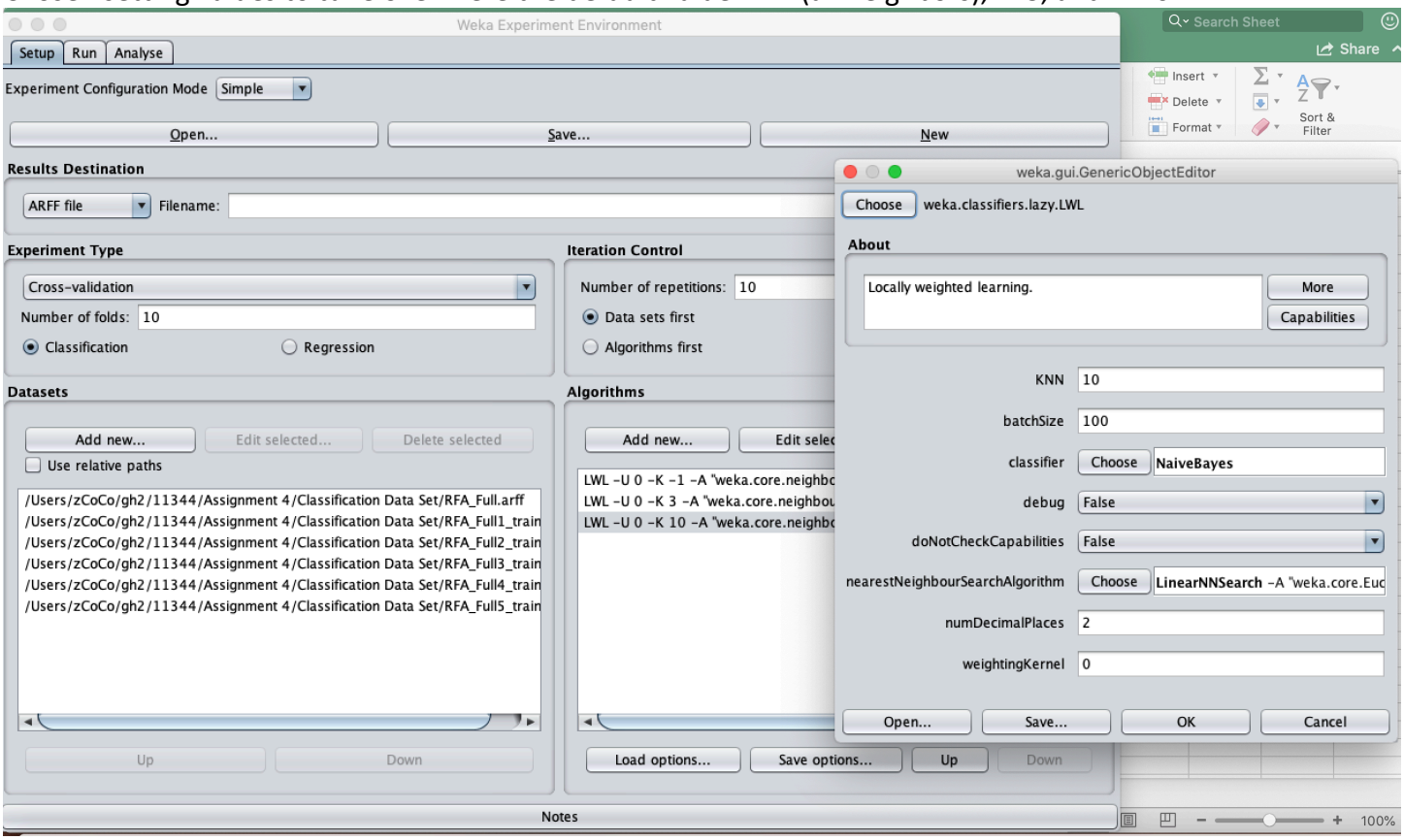
Test Set – Fold 1	68.5714 %
Test Set – Fold 2	71.4286 %
Test Set – Fold 3	62.3188 %
Test Set – Fold 4	69.5652 %
Test Set – Fold 5	62.8571 %
Test Set Average	**66.9482 %**

- For each fold, the fold's training set was loaded in the Preprocess tab of Weka Explorer, then the model's performance was assessed by classifying the fold's test set as the supplied test set in the Classify tab.
- Note: The average performance values determined here could have also been obtained by just using 5-fold cross-validation directly since no choices regarding model settings were made on each fold. I chose to do it this way to get more experience with the method.

2.

Stage 1: Chosen parameter for tuning is the k of the K Nearest Neighbors used by LWL to choose the subset of the space to apply the base classifier over. This was chosen because with a small k the prediction will be heavily influenced by noise, with a moderate k it will be less susceptible to noise, but with a large k (all values) the model is essentially just applying its base classifier over the entire space.

Chosen setting values to tune over were the default value k=-1 (all neighbors), k=3, and k=10.

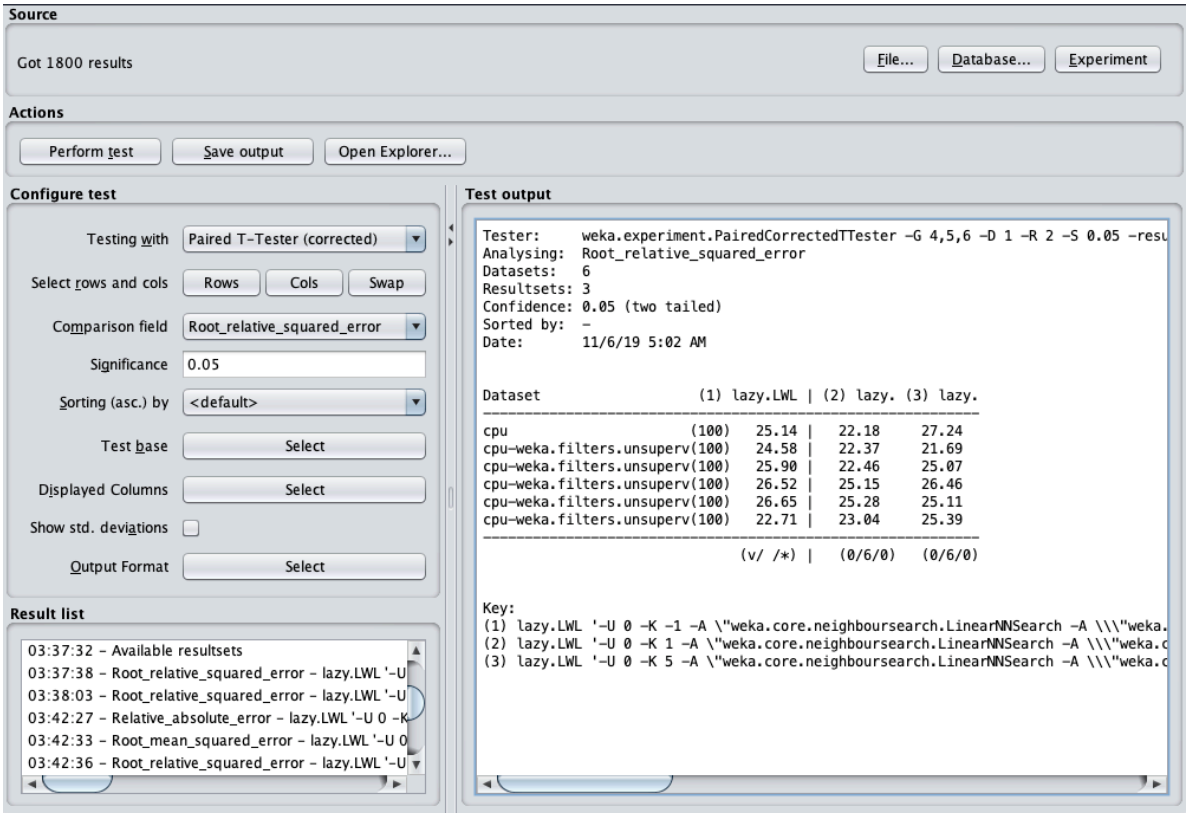


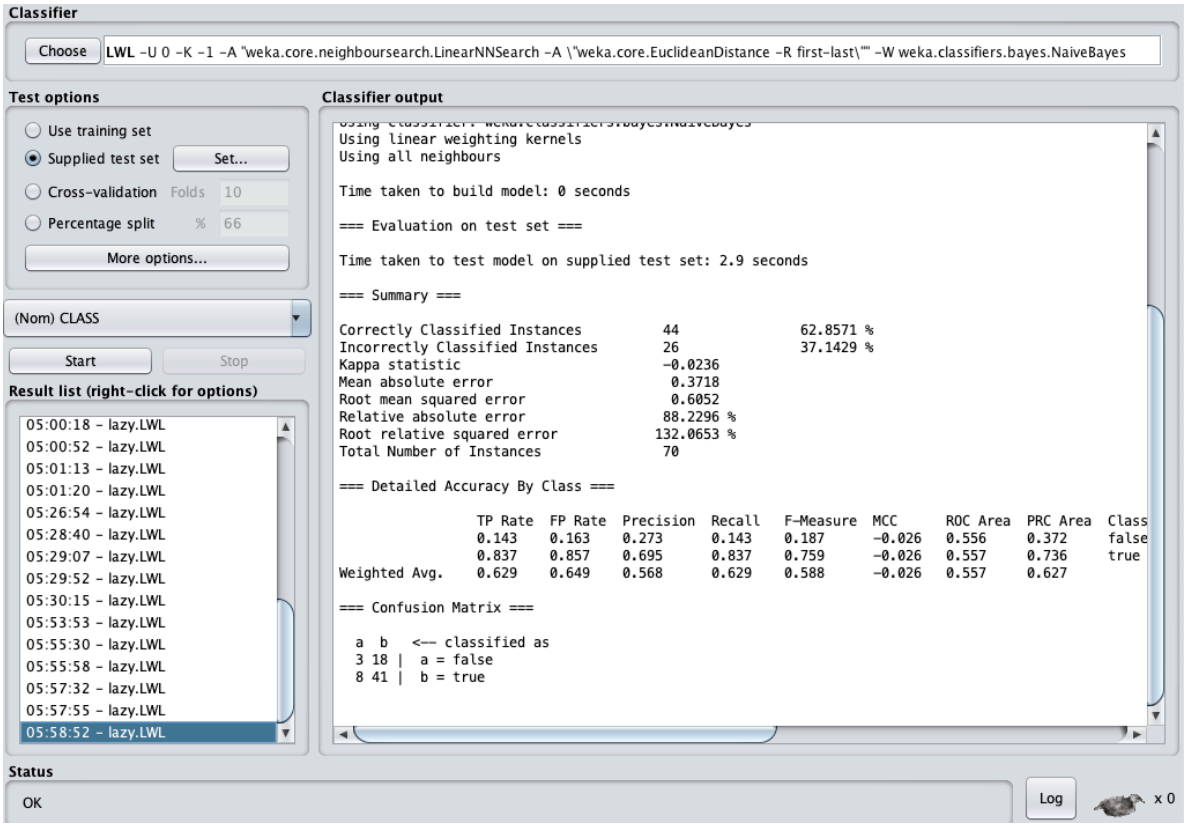
Stage 3:

First, the performance of each algorithm configuration was tested for each fold of the training set using the Weka Experiment Environment. For each fold of the training set, the setting which created the configuration which gave the best (lowest root relative squared) performance was chosen as the optimal setting. The model was then configured with this setting in Weka Explorer, trained against the relevant fold of the training data, and its performance was tested against the relevant fold of the test data.

The outcomes of these tests are presented in the following table.

	Training Set Performance (Percent Correct) for k=-1	Training Set Performance (Percent Correct) for k=3	Training Set Performance (Percent Correct) for k=10	Optimal Setting	Test Set Performance using Optimal Setting
Fold 1	24.58	22.37	21.69	k=10	74.2857 %
Fold 2	25.90	22.46	25.07	k=3	54.2857 %
Fold 3	26.52	25.15	26.46	k=3	47.8261 %
Fold 4	26.65	25.28	25.11	k=10	56.5217 %
Fold 5	22.71	23.04	25.39	k=-1	62.8571 %
				Average:	**59.1553 %**





3.

It was not worth it to perform tuning since the tuned data-set exhibited significantly worse performance on the test dataset than the model with the default parameters (average percent correct of 59.1553% tuned vs 66.9482% default). Alternatively, tuning could have been performed better by choosing different, more, or a wider range of setting values to test.