

Lab Report 5

We downloaded training and test sets. We discussed the conditions on the stack and queue which are that they have available elements, and that the same word hasn't been added in the preceding round.

We completed *extractFeatures* by popping an element from the stack and peeking on the next one, storing these values and then pushing the element back. In the parse method we called the oracle functions and performed the corresponding actions.

We then created Weka models. Probably because of low RAM this took 5 minutes for the unlabeled graph around around 20 for the labeled one. Here are the results:

UNLABELED 4 ARGUMENTS

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	308449	91.5417 %
Incorrectly Classified Instances	28500	8.4583 %
Kappa statistic	0.8855	
Mean absolute error	0.0636	
Root mean squared error	0.1821	
Relative absolute error	17.2135 %	
Root relative squared error	42.3602 %	
Coverage of cases (0.95 level)	98.392 %	
Mean rel. region size (0.95 level)	35.1262 %	
Total Number of Instances	336949	

LABELED 4 ARGUMENTS

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	283038	84.0002 %
Incorrectly Classified Instances	53911	15.9998 %
Kappa statistic	0.8112	
Mean absolute error	0.0039	
Root mean squared error	0.0456	
Relative absolute error	26.1062 %	
Root relative squared error	52.4571 %	
Coverage of cases (0.95 level)	96.2888 %	

Mean rel. region size (0.95 level)	1.9868 %
Total Number of Instances	336949

UNLABELED 6 ARGUMENTS

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	308840	91.6578 %
Incorrectly Classified Instances	28109	8.3422 %
Kappa statistic	0.8871	
Mean absolute error	0.0627	
Root mean squared error	0.1818	
Relative absolute error	16.9662 %	
Root relative squared error	42.2875 %	
Coverage of cases (0.95 level)	98.2591 %	
Mean rel. region size (0.95 level)	34.8911 %	
Total Number of Instances	336949	

LABELED 6 ARGUMENTS

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	283722	84.2032 %
Incorrectly Classified Instances	53227	15.7968 %
Kappa statistic	0.8137	
Mean absolute error	0.0039	
Root mean squared error	0.0456	
Relative absolute error	25.6928 %	
Root relative squared error	52.5388 %	
Coverage of cases (0.95 level)	95.8534 %	
Mean rel. region size (0.95 level)	1.9396 %	
Total Number of Instances	336949	