Hate Speech Detection - TiMBL

Ravi Regulagedda, Jieli Liu, Ugonna Virginia Ahumibe

Initial Accuracy with no hyper-parameter tuning: 70.5451%

Figure 1: Result showing the baseline accuracy

After many tests the best results were achieved with `-a 4`. Every other parameter had no affect on the data, or produced worse results. The final accuracy was an improvement of 0.1% bringing the final accuracy upto 70.6571%.

In some cases, setting parameters like `-k 7` etc., improved the performance but they never gave a better performance than the above parameter. Another interesting observation was the test set and dev set had different optimal parameters. The dev set had greatest accuracy at `-a 3 -k 7 -b 9`. For these parameters, the test set had a lower accuracy than the baseline at 70.3%

The output showing the accuracy for the test set is below.

Figure 2: Result showing the baseline accuracy