

# Lab 5X report

Ludvig Noring   ludno249  
Michael Sörsäter   mics0554  
Victor Tranell   victr593

## Test data

After taking the test we got 83 % accuracy.

## Word space model

We train the word space model with the two given files. Compute the similarity between the gold standard word and all the candidates. The candidate with the highest score is most likely the synonym.

## Evaluation

The performance of our two models:

File	#words	Training time	Accuracy
oanc.txt	11.4 million	1.5 minute	58.75 %
enwik9.txt	124.3 million	31 minutes	77.50 %

The first model, OANC, got an lower accuracy than a non-native english speaker. However when training on the entire english wikipedia the model performed better than the non-native speakers. We still performed better :)