

## SI630 HW2 Mengying Zhang

### Task1, Problem 5&6

From the graphs, we can see that negative log likelihood(nll) displays decreasing trend overtime, and eventually converges to a limit. The convergence rate is pretty fast. When context window is  $[-2, -1, 1, 2]$ , the nll starts with 66,969 and ends with 49,165, converging to around 50,000. When context window is  $[-4,-3, -2, -1]$ , the nll starts with 67,365 and ends with 51,722, converging to around 51,000, which is slightly higher than the original context window. Very similar convergence pattern was found for context window  $[1,2,3,4]$ , with nll starting 67347 and ending 51,772, converging to around 52,000. The 3 resulting nll plots look like as follows:

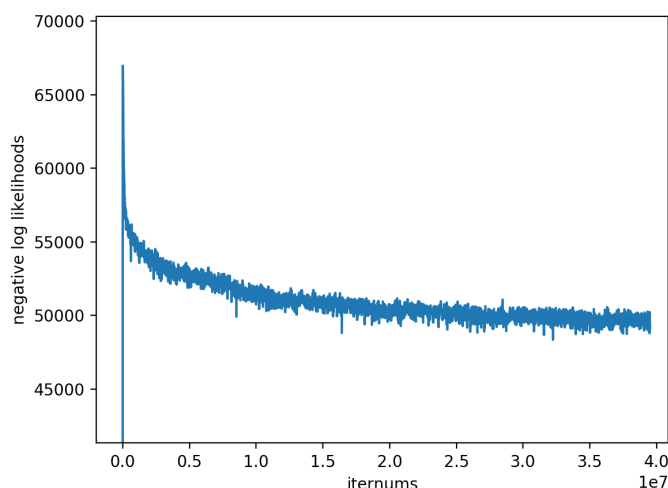


Figure 1: change in negative log likelihood vs. number of iterations using context window =  $[-2, -1, 1, 2]$

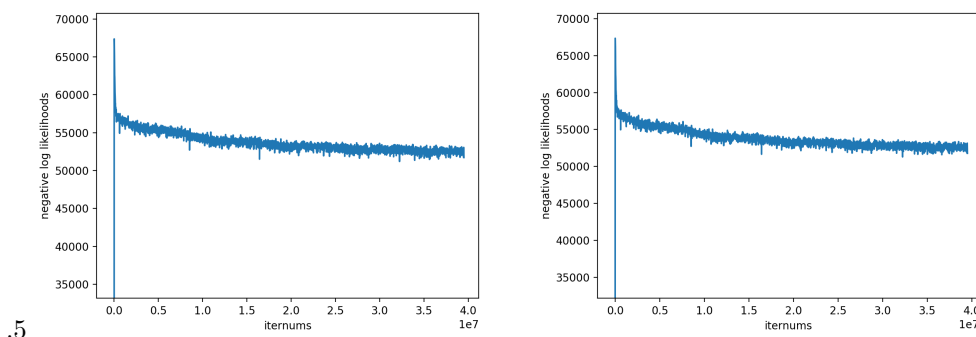


Figure 2: change in negative log likelihood vs. number of iterations using context window =  $[-4,-3,-2,-1]$ (Left),  $[1, 2, 3, 4]$ (Right)

## Task2, Part3

I have chosen the following as my word variation patterns: "-ful", "-less", "re-", "un-", "dis-". The average Precision@K looks like the following:

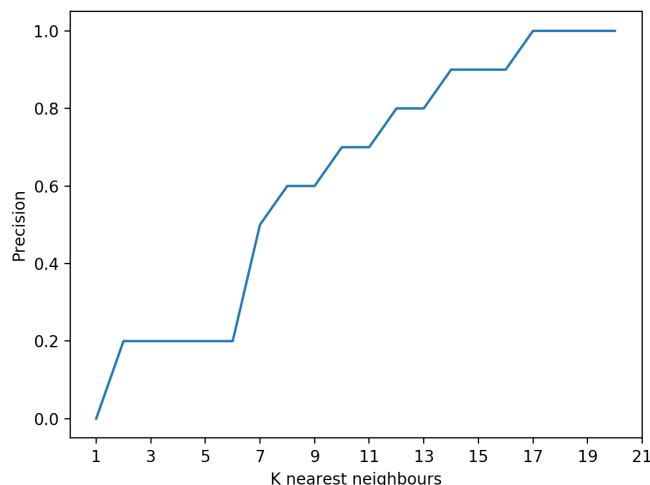


Figure 3: Precision@K using W1 trained by context window = [-2, -1, 1, 2]

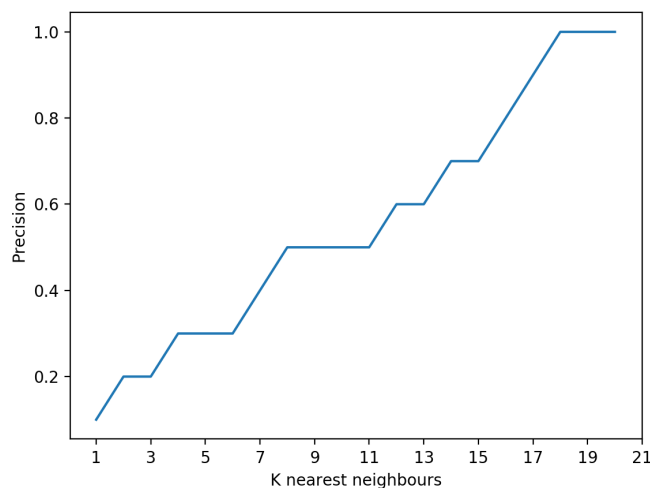


Figure 4: Precision@K using randomly generated W1(200\*20)

In contrast, I have generated the same Precision@K plot using randomly generated W1, see Figure 4. The plot shows that my trained W1 detects morphology more accurate than otherwise not trained model when  $K \geq 7$ . I am very happy that my model actually learned something!

### **Task3, Problem 10**

All three models did a quite good job in catching some of the similarity words with around .4 to .5 cosine similarity score. However, there are some words that are not semantically similar to the target words, but share same part-of-speech(eg. adj.). More interestingly, some even find the opposite word, for example "good"'s similarity words contain "bad". A few contain totally irrelevant words, for example "scary"'s similarity words contain "throw". Details can be found in p8\_output\_1.txt and p8\_output\_2.txt and p9\_output.txt.