# NLP Assignment 1

Mrinal Khubchandani
2018101053

As the values of perplexity were large in some cases, I have calculated and reported log(perplexity values)

The log perplexity scores for different LMs were:

1) LM1
    1) LM1-train: 1.2836997877885
    2) LM1-test: 4.801202830329139
2) LM2
    1) LM2-train: 0.8545999391030258
    2) LM2-test: 5.687357037684344
3) LM3
    1) LM3-train: 1.1075002710771014
    2) LM3-test: 5.097865813945379
4) LM4
    1) LM4-train: 0.6750257390612044
    2) LM4-test: 6.2956088768537395

As we can see, the perplexity scores for test sets are much higher than the perplexity scores for train set. This can be attributed to the fact that we constructed n-grams from the train set.

By adjusting the values of n in our program, we can see that Kneyser-Ney smoothing performs better when value of n is larger and Witten Bell smoothing performs better when value of n is small. This is because Kneyser-Ney is an absolute discounting method which gives more weightage to higher order n-gram.