

Homework 8

N-gram exercise

Exercise

Given the Corpus:

- He is a nice guy
- It's nice to get your letter
- I am going to meet you
- I would like a meat sandwich

Estimate the probability of the following two sentences (with smoothing):

- Nice to meet you
- Nice to meat you

Hint

Add Start & End Tag:

- <s> He is a nice guy <e>
- <s> It's nice to get your letter <e>
- ...

The bigrams:

- <s> He, He is, is a, a nice, nice guy, guy <e>
- <s> It's, It's nice, nice to, to get, get your, your letter, letter <e>
- ...

Then the probabilities (the same as Exercise 1 in tutorial hour) of the 2 sentences both become 0 → Apply smoothing

