

1 Considerations for the structure of lecture

1. I want to show you how you might develop a language model that suits a language-analysis problem you face.
2. The fewer hidden parameters in a model, the “easier” the problem of inferring those values from data.

2 Motivating example: modeling small-talk vs. non-small talk

Let’s consider a generative story like the following:

1. Pick a sentence length ℓ .
2. Pick a sequence of ℓ states: where the two possible state types are `st` for small talk, `nst` for not small-talk
3. For each state, pick a word according to that state’s distribution over single words.

Example; we might decide we’re going to say a five-word sentence, where the first word and the 4th and 5th words are going to be small-talk words.

2.1 Ideas for further refinement

- `st` might have a higher probability of ...
- `st` might have a higher probability of ...
- `st` might have a higher probability of

2.2 Sample data

Written “vertically” instead of “horizontally” to leave room to write on the sides.

Two sentences:

hi

i

agree

thanks

bye

hi

sell

hi [some stock ticker symbol]

now

thanks