

Hide menu

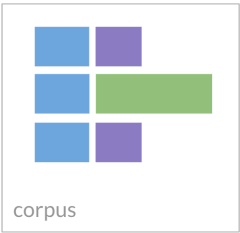
Lecture: Part of Speech Tagging

- Video:** Week Introduction
1 min
- Video:** Part of Speech Tagging
2 min
- Reading:** Part of Speech Tagging
4 min
- Lab:** Lecture Notebook - Working with text files
20 min
- Video:** Markov Chains
3 min
- Reading:** Markov Chains
3 min
- Video:** Markov Chains and POS Tags
4 min
- Reading:** Markov Chains and POS Tags
6 min
- Video:** Hidden Markov Models
3 min
- Reading:** Hidden Markov Models
6 min
- Video:** Calculating Probabilities
3 min
- Reading:** Calculating Probabilities
5 min
- Video:** Populating the Transition Matrix
4 min
- Reading:** Populating the Transition Matrix
6 min
- Video:** Populating the Emission Matrix
2 min
- Reading:** Populating the Emission Matrix
5 min
- Lab:** Lecture Notebook - Working with tags and Numpy
20 min
- Video:** The Viterbi Algorithm
4 min
- Reading:** The Viterbi Algorithm
5 min
- Video:** Viterbi: Initialization
2 min
- Reading:** Viterbi Initialization
5 min
- Video:** Viterbi: Forward Pass
2 min
- Reading:** Viterbi: Forward Pass
10 min
- Video:** Viterbi: Backward Pass
5 min
- Reading:** Viterbi: Backward Pass
10 min

Week 2 > Calculating Probabilities

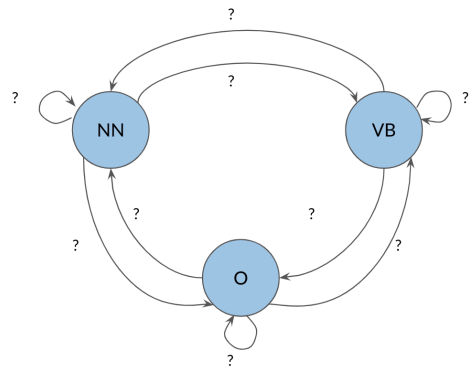
Calculating Probabilities

Here is a visual representation on how to calculate the probabilities:



transition probability: + = 2/3

The number of times that blue is followed by purple is 2 out of 3. We will use the same logic to populate our transition and emission matrices. In the transition matrix we will count the number of times tag $t(i-1), t(i)$ show up near each other and divide by the total number of times $t(i-1)$ shows up (which is the same as the number of times it shows up followed by anything else).



1. Count occurrences of tag pairs
 $C(t_{i-1}, t_i)$
2. Calculate probabilities using the counts

$P(t_i|t_{i-1}) = \frac{C(t_{i-1}, t_i)}{\sum_{j=1}^N C(t_{i-1}, t_j)}$

$C(t(i-1), t(i))$ is the count of times *tag (i-1)* shows up before *tag i*. From this you can compute the probability that a tag shows up after another tag.

Mark as completed

Like Dislike Report an issue

