Populating the Emission Matrix

To populate the emission matrix, you have to keep track of the words associated with their parts of speech tags.

$B =$

	in	a	...
NN (noun)	0		
VB (verb)	0		
O (other)	2		

<s> in a station of the metro
<s> the apparition of these faces in the crowd :
<s> petals on a wet , black bough .

Ezra Pound – 1913

To populate the matrix, we will also use smoothing as we have previously used:

$$P(w_i \mid t_i) = \frac{C(t_i, w_i) + \epsilon}{\sum_{j=1}^{V} C(t_i, w_j) + N * \epsilon}$$
$$= \frac{C(t_i, w_i) + \epsilon}{C(t_i) + N * \epsilon}$$

Where $C(t_i, w_i)$ is the count associated with how many times the tag t_i is associated with the word w_i. The epsilon above is the smoothing parameter. In the next video, we will talk about the Viterbi algorithm and discuss how you can use the transition and emission matrix to come up with probabilities.