



NLP Assignment 2: Answers

Name: Alireza Dastmalchi Saei

Stu No.: 993613026

POS Tagging

0.1 Probability Table

First, the count matrix must be calculated to:

Table 1: Count Table

Words	Count	Noun	Modal	Verb
Mark	1	0	2	
Can	0	3	0	
Watch	2	0	2	
Tom	2	0	0	
Will	1	1	0	

Now that we have the count vector, the probability table of each word being Noun, Model, or Verb will be as follows:

Table 2: Probability Table

Words	Probs	Noun	Modal	Verb
Mark	$\frac{1}{6}$	0	$\frac{1}{2}$	
Can	0	$\frac{3}{4}$	0	
Watch	$\frac{2}{6}$	0	$\frac{1}{2}$	
Tom	$\frac{2}{6}$	0	0	
Will	$\frac{1}{6}$	$\frac{1}{4}$	0	

0.2 Sentences with <S>and <E>

Sentences are with Start and End Sentence Tags are:

- <S>Mark can watch <E>(S N M V E)
- <S>Will can mark watch <E>(S N M V N E)
- <S>Can Tom watch? <E>(S M N V E)
- <S>Tom will mark watch <E>(S N M V N E)

Now with these new sentences the below count table can be created:

Table 3: Occurrence Table

	Noun	Modal	Verb	<E>
<S>	3	1	0	0
Noun	0	3	1	2
Modal	1	0	3	0
Verb	2	0	0	2

Now by dividing each number by sum of row numbers, we will get probabilities happening specific POS sequences:

Table 4: Prob. Occurrence Table

	Noun	Modal	Verb	<E>
<S>	$\frac{3}{4}$	$\frac{1}{4}$	0	0
Noun	0	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$
Modal	$\frac{1}{4}$	0	$\frac{3}{4}$	0
Verb	$\frac{1}{2}$	0	0	$\frac{1}{2}$

0.3 Applying POS Tagging

The sentence is "Can Tom mark watch" and the POS tagging must result:
<S>Modal Noun Verb Noun <E>

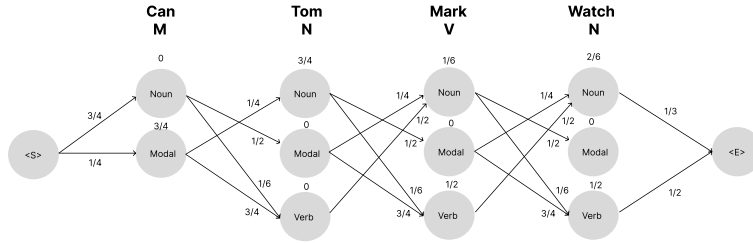


Figure 1: Hidden Markov Model for the given Sentence

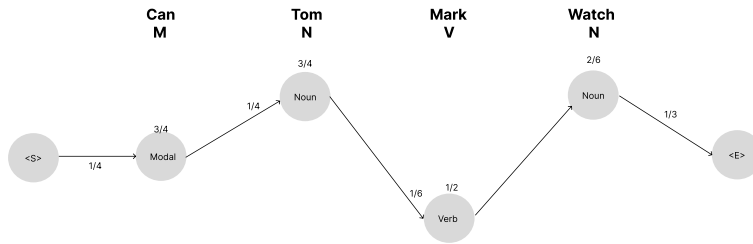


Figure 2: Simplified Hidden Markov Model

$$\begin{aligned}
 P(\text{Sentence}) &= \frac{1}{4} \times \frac{3}{4} \times \frac{1}{4} \times \frac{3}{4} \times \frac{1}{6} \times \frac{1}{2} \times \frac{1}{2} \times \frac{2}{6} \times \frac{1}{3} \\
 &= 0.0001627604
 \end{aligned}$$