1. Markov Models
Markov Models

14:58 / 14:58 1.0x

is simply the corresponding count normalized such that

across different choices of W prime,

the entries in the table sum to 1.

So with sum over possible values of the word that

can come after W.

And this is the maximum likelihood estimate

that you could derive as the solution,

as the automizing solution to the previous criteria.

End of transcript. Skip to the start.

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Markov Symbols

1/1 point (graded)
To specify a Markov language model, what additional symbols do we need to add to our set of possible symbols?
(Choose all that apply.)

- [x] a start symbol ✔
- [x] an end symbol ✔
- [x] a symbol for unknown words ✔
- [ ] a symbol for complicated words

✔

Solution:

As shown in the lecture video, we need start and end symbols in order to specify the bounds of our sentence. We also need a symbol for unknown words as there might be words in the sentence not represented in our Markov Model. Lastly, our Markov model treats words the same regardless of complexity. We're simply representing a table of transition probabilities, so there's no need to create extra symbols for complex words.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Transition Probabilities

1/2 points (graded)

w_{i-1} \ w_i	ML	course	is	UNK	<end>
<beg>	0.7	0.1	0.1	0.1	0.0
ML	0.1	0.5	0.2	0.1	0.1
course	0.0	0.0	0.7	0.1	0.2
is	0.1	0.3	0.0	0.6	0.0
UNK	0.1	0.2	0.2	0.3	0.2

Using a first order Markov model specified above, what is the probability of generating the following sentence <beg> ML course UNK <end>?

- 0.007 ✔
- 0.01
- 0.003
- 0.005

Which of the following sentences are possible to generate?
(Choose all those apply.)

- ☑ <beg> course ML is UNK <end>
- ☐ <beg> <end>
- ☐ course is ML <end>
- ☑ <beg> ML course <end> ✔

✖

Solution:

For the first question,

$$P(ML|<beg>) \times P(course|ML) \times P(UNK|course) \times P(<end>|UNK) = 0.7 \times 0.5 \times 0.1 \times 0.2 = 0.007$$

. For the second question, all valid sentences must start with <beg>. The probability of the second word being <end> is zero. In addition, the probability of ML occuring after course is also zero. The correct answer is the only choice with nonzero transition probabilities at each step.

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Maximum Likelihood

1/1 point (graded)

Suppose our training examples are the following three sentences.

ML courses are cool.

Humanities courses are cool.

But some courses are boring.

Using a bigram model, what is the maximum likelihood estimate for the probability that the next word is 'cool', given that the previous word is 'are'?

- ◉ $\frac{2}{3}$ ✔
- ○ 1
- ○ $\frac{1}{3}$
- ○ $\frac{1}{4}$

Solution:

"are" occurs three times in the training corpus, and is followed by "cool" two out of those three times.

Submit You have used 1 of 2 attempts

ℹ Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 3 Neural networks (2.5 weeks):Lecture 11. Recurrent Neural Networks 2 / 1. Markov Models