검
색

- 설정
- 업데이트
- 성취도
- 도움말 센터
- 로그아웃
-
- Youngeun In

Sequence Architectures

Audio Data

nal)

- **Quiz**
  - 테스트: Sequence Models & Attention Mechanism
    10개의 질문
- **Programming Assignments**

테스트테스트 • 30 min30 minutes

Sequence Models & Attention Mechanism

과제 제출
기한년 9월 6일 오후 3:59 KST년 9월 6일 오후 3:59 KST
시도하기8 hours당 3회

다시 시도해주십시오

성적 받기
통과 점수:80% 이상
성적
100%

피드백 보기

최고 점수가 유지됩니다.

탐색 확인

이 페이지에서 나가시겠습니까?

이 페이지에 머물기 이 페이지에서 나가기

Sequence Models & Attention Mechanism
성적 평가 퀴즈 • 30 min

만료 년 9월 6일 오후 3:59 KST

축하합니다! 통과하셨습니다!
통과 점수: 80% 이상

학습 계속하기

성적
100%

Sequence Models & Attention Mechanism

최신 제출물 성적
100%
1.
질문 1

Consider using this encoder-decoder model for machine translation.

This model is a “conditional language model” in the sense that the encoder portion (shown in green) is modeling the probability of the input sentence xx.

1 / 1점

False

True

맞습니다
2.
질문 2

In beam search, if you increase the beam width BB, which of the following would you expect to be true? Check all that apply.

1 / 1점

Beam search will converge after fewer steps.

Beam search will use up more memory.

맞습니다

Beam search will run more slowly.

맞습니다

Beam search will generally find better solutions (i.e. do a better job maximizing $P(y \mid x)$P(y | x))

맞습니다
3.
질문 3

In machine translation, if we carry out beam search without using sentence normalization, the algorithm will tend to output overly short translations.

1 / 1점

True

False

맞습니다
4.
질문 4

Suppose you are building a speech recognition system, which uses an RNN model to map from audio clip xx to a text transcript yy. Your algorithm uses beam search to try to find the value of yy that maximizes $P(y \mid x)$P(y | x).

On a dev set example, given an input audio clip, your algorithm outputs the transcript $\hat{y} =$ y^ = “I’m building an A Eye system in Silly con Valley.”, whereas a human gives a much superior transcript $y^* =$ y∗ = “I’m building an AI system in Silicon Valley.”

According to your model,

$P(\hat{y} \mid x) = 1.09 * 10^-7$P(y^ | x) = 1.09 ∗ 10−7

$P(y^* \mid x) = 7.21 * 10^-8$P(y∗ | x) = 7.21 ∗ 10−8

Would you expect increasing the beam width B to help correct this example?

1 / 1점

No, because $P(y^* \mid x) \leq P(\hat{y} \mid x)$P(y∗ | x) ≤ P(y^ | x) indicates the error should be attributed to the RNN rather than to the search algorithm.

No, because $P(y^* \mid x) \leq P(\hat{y} \mid x)$P(y∗ | x) ≤ P(y^ | x) indicates the error should be attributed to the search algorithm rather than to the RNN.

Yes, because $P(y^* \mid x) \leq P(\hat{y} \mid x)$P(y∗ | x) ≤ P(y^ | x) indicates the error should be attributed to the RNN rather than to the search algorithm.

Yes, because $P(y^* \mid x) \leq P(\hat{y} \mid x)$P(y∗ | x) ≤ P(y^ | x) indicates the error should be attributed to the search algorithm rather than to the RNN.

맞습니다
5.
질문 5

Continuing the example from Q4, suppose you work on your algorithm for a few more weeks, and now find that for the vast majority of examples on which your algorithm makes a mistake, $P(y^* \mid x) > P(\hat{y} \mid x)$P(y∗ | x) > P(y^ | x). This suggests you should focus your attention on improving the search algorithm.

1 / 1점

False.

True.

맞습니다
6.
질문 6

Consider the attention model for machine translation.

Further, here is the formula for $\alpha^{<t,t'>}$ α<t,t’>.

Which of the following statements about $\alpha^{<t,t'>}$ α<t,t’> are true? Check all that apply.

1 / 1점

☐ ☐

We expect $\alpha^{<t,t'>}$ α<t,t’> to be generally larger for values of $a^{<t>}$ a<t> that are highly relevant to the value the network should output for $y^{<t'>}$ y<t’>. (Note the indices in the superscripts.)

☐ ☐

$\sum_t \alpha^{<t,t'>} = 1$∑t α<t,t’> = 1 (Note the summation is over tt.)

☑ ☐

$\sum_{t'} \alpha^{<t,t'>} = 1$∑t’ α<t,t’> = 1 (Note the summation is over t't’.)

✓

맞습니다

☑ ☐

We expect $\alpha^{<t,t'>}$ α<t,t’> to be generally larger for values of $a^{<t'>}$ a<t’> that are highly relevant to the value the network should output for $y^{<t>}$ y<t>. (Note the indices in the superscripts.)

✓

맞습니다

7.

질문 7

The network learns where to “pay attention” by learning the values $e^{<t,t'>}$ e<t,t’>, which are computed using a small neural network:

We can't replace $s^{<t-1>}$ s<t−1> with $s^{<t>}$ s<t> as an input to this neural network. This is because $s^{<t>}$ s<t> depends on $\alpha^{<t,t'>}$ α<t,t’> which in turn depends on $e^{<t,t'>}$ e<t,t’>; so at the time we need to evaluate this network, we haven’t computed $s^{<t>}$ s<t> yet.

1 / 1점

◉ ○

True

○ ○

False

✓

맞습니다

8.

질문 8

Compared to the encoder-decoder model shown in Question 1 of this quiz (which does not use an attention mechanism), we expect the attention model to have the greatest advantage when:

1 / 1점

◉ ○

The input sequence length T_xTx is large.

○ ○

The input sequence length T_xTx is small.

✓

맞습니다

9.

질문 9

Under the CTC model, identical repeated characters not separated by the “blank” character (_) are collapsed. Under the

CTC model, what does the following string collapse to?

__c_oo_o_kk___b_ooooo__oo__kkk

1 / 1점

cookbook

cook book

coookkbooooooookkk

cokbok

맞습니다

10.

질문 10

In trigger word detection, $x^{<t>}$ x<t> is:

1 / 1점

The tt-th input word, represented as either a one-hot vector or a word embedding.

Features of the audio (such as spectrogram features) at time tt.

Whether someone has just finished saying the trigger word at time tt.

Whether the trigger word is being said at time tt.

맞습니다