

Sequence models and literature

✓

Video: A conversation with Andrew Ng

1 min

✓

Video: Introduction

1 min

✓

Video: Looking into the code

57 sec

✓

Video: Training the data

2 min

✓

Video: More on training the data

1 min

✓

Reading: Check out the code!

10 min

✓

Video: Notebook for lesson 1

8 min

✓

Video: Finding what the next word should be

2 min

✓

Video: Example

1 min

✓

Video: Predicting a word

1 min

✓

Video: Poetry!

40 sec

✓

Reading: link to Laurence's poetry

10 min

✓

Video: Looking into the code

1 min

✓

Video: Laurence the poet!

1 min

✓

Reading: Check out the code!

10 min

✓

Video: Your next task

1 min

✓

Reading: Link to generating text using a character-based RNN

10 min

✓

Quiz: Week 4 Quiz

8 questions

Weekly Exercise- Using LSTMs, see if you can write Shakespeare!

Course 3 Wrap up

☰

☰

✓

Congratulations! You passed!

QUIZ

TO PASS 80% or higher

Keep Learning

GRADE

100%

Week 4 Quiz

Week 4 Quiz

LATEST SUBMISSION GRADE

100%

✓

Submit your assignment

Try again

DUE DATE Sep 7, 12:29 PM IST

ATTEMPTS 3 every 8 hours

1. What is the name of the method used to tokenize a list of sentences?

1 / 1 point

○

tokenize\_on\_text(sentences)

✓

Receive grade

○

fit\_to\_text(sentences)

○

fit\_on\_texts(sentences)

○

tokenize(sentences)

TO PASS 80% or higher

✓

Correct

Grade

100%

We keep your highest score

View Feedback

👍

👎

📄

2. If a sentence has 120 tokens in it, and a Conv1D with 128 filters with a Kernal size of 5 is passed over it, what's the output shape?

1 / 1 point

○

(None, 120, 124)

●

(None, 116, 128)

○

(None, 120, 128)

○

(None, 116, 124)

✓

Correct

3. What is the purpose of the embedding dimension?

1 / 1 point

○

It is the number of dimensions required to encode every word in the corpus

●

It is the number of dimensions for the vector representing the word encoding

○

It is the number of letters in the word, denoting the size of the encoding

○

It is the number of words to encode in the embedding

✓

Correct

4. IMDB Reviews are either positive or negative. What type of loss function should be used in this scenario?

1 / 1 point

●

Binary crossentropy

○

Binary Gradient descent

○

Adam

○

Categorical crossentropy

✓

Correct

5. If you have a number of sequences of different lengths, how do you ensure that they are understood when fed into a neural network?

1 / 1 point

○

Make sure that they are all the same length using the pad\_sequences method of the tokenizer

○

Specify the input layer of the Neural Network to expect different sizes with dynamic\_length

●

Use the pad\_sequences object from the tensorflow.keras.preprocessing.sequence namespace

○

Process them on the input layer of the Neural Network using the pad\_sequences property

✓

Correct

6. When predicting words to generate poetry, the more words predicted the more likely it will end up gibberish. Why?

1 / 1 point

○

It doesn't, the likelihood of gibberish doesn't change

●

Because the probability that each word matches an existing phrase goes down the more words you create

○

Because you are more likely to hit words not in the training set

○

Because the probability of prediction compounds, and thus increases overall