



Week 1 Quiz

TOTAL POINTS 8

1.Question 1

What is the name of the object used to tokenize sentences?

Tokenizer

2.Question 2

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

`fit_on_texts(sentences)`

3.Question 3

Once you have the corpus tokenized, what's the method used to encode a list of sentences to use those tokens?

`texts_to_sequences(sentences)`

4.Question 4

When initializing the tokenizer, how to you specify a token to use for unknown words?

`oov_token=<Token>`

5.Question 5

If you don't use a token for out of vocabulary words, what happens at encoding?

The word isn't encoded, and is skipped in the sequence

6.Question 6

If you have a number of sequences of different lengths, how do you ensure that

they are understood when fed into a neural network?

Use the `pad_sequences` object from the `tensorflow.keras.preprocessing.sequence` namespace

7.Question 7

If you have a number of sequences of different length, and call `pad_sequences` on them, what's the default result?

Radio Checked

They'll get padded to the length of the longest sequence by adding zeros to the beginning of shorter ones

8.Question 8

When padding sequences, if you want the padding to be at the end of the sequence, how do you do it?

Pass `padding='post'` to `pad_sequences` when initializing it

END