

Quiz for Natural Language Processing

Tuesday, June 25, 2019

11:13 AM

[Back](#)

Week 1 Quiz

Quiz, 8 questions

Question 1

1
point

1. Question 1

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



Tokenizer



CharacterTokenizer



TextTokenizer



WordTokenizer

Question 2

1
point

2. Question 2

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



tokenize_on_text(sentences)



fit_to_text(sentences)



tokenize(sentences)



fit_on_texts(sentences)

Question 3

1
point

3. Question 3

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



text_to_tokens(sentences)



text_to_sequences(sentences)

- ☐ text_to_sequences(sentences)
- ☒ texts_to_sequences(sentences)
- ☐ texts_to_tokens(sentences)

Question 4

1
point

4. Question 4

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

- ☒ oov_token=<Token>
- ☐ out_of_vocab=<Token>
- ☐ unknown_word=<Token>
- ☐ unknown_token=<Token>

Question 5

1
point

5. Question 5

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

- ☐ The word is replaced by the most common token
- ☐ The word isn't encoded, and the sequencing ends
- ☒ The word isn't encoded, and is skipped in the sequence
- ☐ The word isn't encoded, and is replaced by a zero in the sequence

Question 6

1
point

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?

- ☐ Specify the input layer of the Neural Network to expect different sizes with dynamic_length
- ☐ Make sure that they are all the same length using the pad_sequences method of the tokenizer
- ☐ Process them on the input layer of the Neural Network using the pad_sequences property
- ☒ Use the pad_sequences object from the tensorflow.keras.preprocessing.sequence namespace

Question 7

1

point

7. Question 7

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

- ☒ They'll get padded to the length of the longest sequence by adding zeros to the beginning of shorter ones
- ☐ They'll get padded to the length of the longest sequence by adding zeros to the end of shorter ones
- ☐ They'll get cropped to the length of the shortest sequence
- ☐ Nothing, they'll remain unchanged

Question 8

1

point

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
- ☐ Call the padding method of the `pad_sequences` object, passing it 'after'
- ☐ Pass `padding='after'` to `pad_sequences` when initializing it
- ☐ Call the padding method of the `pad_sequences` object, passing it 'post'