**Week 1 Quiz**

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

CharacterTokenizer

WordTokenizer

Tokenizer

TextTokenizer

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

fit\_on\_texts(sentences)

tokenize(sentences)

fit\_to\_text(sentences)

tokenize\_on\_text(sentences)

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

text\_to\_sequences(sentences)

texts\_to\_sequences(sentences)

text\_to\_tokens(sentences)

texts\_to\_tokens(sentences)

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

out\_of\_vocab=<Token>

oov\_token=<Token>

unknown\_token=<Token>

unknown\_word=<Token>

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

The word isn’t encoded, and the sequencing ends

The word isn’t encoded, and is replaced by a zero in the sequence

The word is replaced by the most common token

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

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

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

7. If you have a number of sequences of different length, and call pad\_sequences on them, what’s the default result?

Nothing, they’ll remain unchanged

They’ll get cropped to the length of the shortest sequence

They’ll get padded to the length of the longest sequence by adding zeros to the end of shorter ones

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

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 ‘post’

Call the padding method of the pad\_sequences object, passing it ‘after’

Pass padding=’after’ to pad\_sequences when initializing it