

Week 2 Quiz

LATEST SUBMISSION GRADE

87.5%

1.Question 1

What is the name of the TensorFlow library containing common data that you can use to train and test neural networks?

TensorFlow Datasets

2.Question 2

How many reviews are there in the IMDB dataset and how are they split?

50,000 records, 50/50 train/test split

3.Question 3

How are the labels for the IMDB dataset encoded?

Reviews encoded as a number 0-1

4.Question 4

What is the purpose of the embedding dimension?

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

5.Question 5

When tokenizing a corpus, what does the num_words=n parameter do?

It specifies the maximum number of words to be tokenized, and picks the most common 'n' words

6.Question 6

To use word embeddings in TensorFlow, in a sequential layer, what is the name of the class?

`tf.keras.layers.Embedding`

7.Question 7

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

Binary crossentropy

8.Question 8

When using IMDB Sub Words dataset, our results in classification were poor. Why?

Sequence becomes much more important when dealing with subwords, but we're ignoring word positions

END