Week 2 Quiz

CALIFICACIÓN DEL ÚLTIMO ENVÍO

100%

1. What is the name of the TensorFlow library containing common data that you can use to train and test neural networks? **1 / 1 punto**

- ○ TensorFlow Data
- ◉ TensorFlow Datasets
- ○ There is no library of common data sets, you have to use your own
- ○ TensorFlow Data Libraries

✓ **Correcto**

2. How many reviews are there in the IMDB dataset and how are they split? **1 / 1 punto**

- ○ 50,000 records, 80/20 train/test split
- ◉ 50,000 records, 50/50 train/test split
- ○ 60,000 records, 80/20 train/test split
- ○ 60,000 records, 50/50 train/test split

✓ **Correcto**

3. How are the labels for the IMDB dataset encoded? **1 / 1 punto**

- ○ Reviews encoded as a boolean true/false
- ○ Reviews encoded as a number 1-5
- ○ Reviews encoded as a number 1-10
- ◉ Reviews encoded as a number 0-1

Correcto

4. What is the purpose of the embedding dimension? 1 / 1 punto

- (●) 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
- ( ) It is the number of dimensions required to encode every word in the corpus

Correcto

5. When tokenizing a corpus, what does the num_words=n parameter do? 1 / 1 punto

- ( ) It specifies the maximum number of words to be tokenized, and stops tokenizing when it reaches n
- (●) It specifies the maximum number of words to be tokenized, and picks the most common ‘n’ words
- ( ) It specifies the maximum number of words to be tokenized, and picks the first ‘n’ words that were tokenized
- ( ) It errors out if there are more than n distinct words in the corpus

Correcto

6. To use word embeddings in TensorFlow, in a sequential layer, what is the name of the class? 1 / 1 punto

- (●) tf.keras.layers.Embedding
- ( ) tf.keras.layers.Embed
- ( ) tf.keras.layers.WordEmbedding
- ( ) tf.keras.layers.Word2Vector

Correcto

7. IMDB Reviews are either positive or negative. What type of loss function should be used in this scenario? 1 / 1 punto

- ○ Binary Gradient descent
- ○ Adam
- ◉ Binary crossentropy
- ○ Categorical crossentropy

Correcto

8. When using IMDB Sub Words dataset, our results in classification were poor. Why? 1 / 1 punto

- ○ We didn't train long enough
- ○ Our neural network didn't have enough layers
- ○ The sub words make no sense, so can't be classified
- ◉ Sequence becomes much more important when dealing with subwords, but we're ignoring word positions

Correcto