

Word Embeddings

- ✓

**Video:** A conversation with Andrew Ng  
2 min
- ✓

**Video:** Introduction  
2 min
- ✓

**Video:** The IMBD dataset  
1 min
- ✓

**Reading:** IMDB reviews dataset  
10 min
- ✓

**Video:** Looking into the details  
4 min
- ✓

**Video:** How can we use vectors?  
2 min
- ✓

**Video:** More into the details  
2 min
- ✓

**Reading:** Check out the code!  
10 min
- ✓

**Video:** Notebook for lesson 1  
10 min
- ✓

**Video:** Remember the sarcasm dataset?  
1 min
- ✓

**Video:** Building a classifier for the sarcasm dataset  
1 min
- ✓

**Video:** Let's talk about the loss function  
1 min
- ✓

**Reading:** Check out the code!  
10 min
- ✓

**Video:** Pre-tokenized datasets  
43 sec
- ✓

**Reading:** TensorFlow datasets  
10 min
- ✓

**Video:** Diving into the code (part 1)  
1 min
- ✓

**Reading:** Subwords text encoder  
10 min
- ✓

**Video:** Diving into the code (part 2)  
2 min
- ✓

**Reading:** Check out the code!  
10 min
- ✓

**Video:** Notebook for lesson 3  
5 min
- ✓

**Quiz:** Week 2 Quiz  
8 questions
- 📖

**Reading:** Week 2 Wrap up  
10 min

Weekly Exercise- More on the BBC News Archive

✓

Congratulations! You passed!

TO PASS 80% or higher

Keep Learning

GRADE  
100%

Week 2 Quiz

Week 2 Quiz

LATEST SUBMISSION GRADE

100%

✓

Submit your assignment

Try again

**DUE DATE** Oct 12, 12:29 PM IST

**ATTEMPTS** 3 every 8 hours

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

1 / 1 point

- Ⓐ

TensorFlow Datasets

✓

Receive grade

Ⓑ

There is no library of common data sets, you have to use your own

TO PASS 80% or higher

Ⓒ

TensorFlow Data Libraries

Ⓓ

TensorFlow Data

Grade  
100%

View Feedback

We keep your highest score



✓

Correct

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

1 / 1 point

- Ⓐ

60,000 records, 50/50 train/test split
- Ⓑ

60,000 records, 80/20 train/test split
- Ⓒ

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

50,000 records, 80/20 train/test split

✓

Correct

3. How are the labels for the IMDB dataset encoded?

1 / 1 point

- Ⓐ

Reviews encoded as a number 0-1
- Ⓑ

Reviews encoded as a number 1-5
- Ⓒ

Reviews encoded as a number 1-10
- Ⓓ

Reviews encoded as a boolean true/false

✓

Correct

4. What is the purpose of the embedding dimension?

1 / 1 point

- Ⓐ

It is the number of dimensions required to encode every word in the corpus
- Ⓑ

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

It is the number of words to encode in the embedding
- Ⓓ

It is the number of letters in the word, denoting the size of the encoding

✓

Correct

5. When tokenizing a corpus, what does the num\_words=n parameter do?

1 / 1 point

- Ⓐ

It errors out if there are more than n distinct words in the corpus
- Ⓑ

It specifies the maximum number of words to be tokenized, and picks the first 'n' words that were tokenized
- Ⓒ

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

✓

Correct

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

1 / 1 point

- Ⓐ

tf.keras.layers.Word2Vector
- Ⓑ

tf.keras.layers.Embedding
- Ⓒ

tf.keras.layers.WordEmbedding
- Ⓓ

tf.keras.layers.Embed

✓

Correct