✓ **Congratulations! You passed!**
TO PASS 80% or higher

Keep Learning

GRADE
100%

Week 3 Quiz

LATEST SUBMISSION GRADE
100%

1. Why does sequence make a large difference when determining semantics of language?

`1 / 1 point`

◉ Because the order in which words appear dictate their impact on the meaning of the sentence

○ Because the order in which words appear dictate their meaning

○ It doesn't

○ Because the order of words doesn't matter

✓ **Correct**

2. How do Recurrent Neural Networks help you understand the impact of sequence on meaning?

`1 / 1 point`

○ They shuffle the words evenly

◉ They carry meaning from one cell to the next

○ They don't

○ They look at the whole sentence at a time

✓ **Correct**

3. How does an LSTM help understand meaning when words that qualify each other aren't necessarily beside each other in a sentence?

`1 / 1 point`

◉ Values from earlier words can be carried to later ones via a cell state

○ They load all words into a cell state

○ They don't

○ They shuffle the words randomly

✓ **Correct**

4. What keras layer type allows LSTMs to look forward and backward in a sentence?

`1 / 1 point`

○ Unilateral

◉ Bidirectional

○ Bothdirection

○ Bilateral

✓ **Correct**

5. What's the output shape of a bidirectional LSTM layer with 64 units?

`1 / 1 point`

○ (128,1)

◉ (None, 128)

○ (None, 64)

○ (128,None)

✓ **Correct**

---

6. When stacking LSTMs, how do you instruct an LSTM to feed the next one in the sequence?   `1 / 1 point`

○ Do nothing, TensorFlow handles this automatically

○ Ensure that return_sequences is set to True on all units

○ Ensure that they have the same number of units

◉ Ensure that return_sequences is set to True only on units that feed to another LSTM

✓ **Correct**

---

7. If a sentence has 120 tokens in it, and a Conv1D with 128 filters with a Kernal size of 5 is passed over it, what's the output shape?   `1 / 1 point`

○ (None, 120, 124)

○ (None, 116, 124)

○ (None, 120, 128)

◉ (None, 116, 128)

✓ **Correct**

---

8. What's the best way to avoid overfitting in NLP datasets?   `1 / 1 point`

○ Use LSTMs

○ Use GRUs

○ Use Conv1D

◉ None of the above

✓ **Correct**