



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

Keep Learning

GRADE  
100%

### You have new skill scores!

Great job! Keep learning and making progress in your courses to increase your skill scores.

Artificial Neural Networks

390



Computer Programming

153



Deep Learning

249



Machine Learning

317



Python Programming

234



## Week 3 Quiz

LATEST SUBMISSION GRADE

100%

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

1 / 1 point

- ☐ It doesn't
- ☐ Because the order of words doesn't matter
- ☐ Because the order in which words appear dictate their meaning
- ☒ Because the order in which words appear dictate their impact on the meaning of the sentence

✓ Correct

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

1 / 1 point

- ☐ They look at the whole sentence at a time
- ☐ They don't
- ☒ They carry meaning from one cell to the next
- ☐ They shuffle the words evenly

✓ 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

- ☐ They don't
- ☐ They shuffle the words randomly
- ☒ Values from earlier words can be carried to later ones via a cell state
- ☐ They load all words into a cell state

✓ Correct

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

1 / 1 point

- ☐ Bothdirection
- ☐ Bilateral
- ☒ Bidirectional
- ☐ Unilateral

✓ Correct

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

1 / 1 point

- ☐ (None, 64)
- ☐ (128, None)
- ☒ (None, 128)

 (128,1)

☐ (128,1)

 Correct

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

1 / 1 point

- ☐ Ensure that return\_sequences is set to True on all units
- ☐ Do nothing, TensorFlow handles this automatically
- ☐ 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 Kernel size of 5 is passed over it, what's the output shape?

1 / 1 point

- ☐ (None, 116, 124)
- ☐ (None, 120, 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