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

- ○ It doesn't
- ○ Because the order of words doesn't matter
- ◉ 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

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

- ○ 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

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

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

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

- ○ Unilateral
- ○ Bilateral
- ◉ Bidirectional
- ○ Bothdirection

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

- ◉ (None, 128)
- ○ (128,None)
- ○ (128,1)
- ○ (None, 64)

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

- ○ 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
- ○ Do nothing, TensorFlow handles this automatically

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?

- (None, 116, 128)
- (None, 116, 124)
- (None, 120, 128)
- (None, 120, 124)

8. What's the best way to avoid overfitting in NLP datasets?

- Use LSTMs
- Use GRUs
- Use Conv1D
- None of the above