

1. Are there any dimension limitations on the input of RNN models? Why?

Answer:

2. What are the main reasons to add the gate units in LSTM models?

Answer:

3. Why is GRU faster as compared to LSTM?

Answer:

4. Tell the reason for adding the attention mechanism to the RNN/ Seq2Seq model.

Answer:

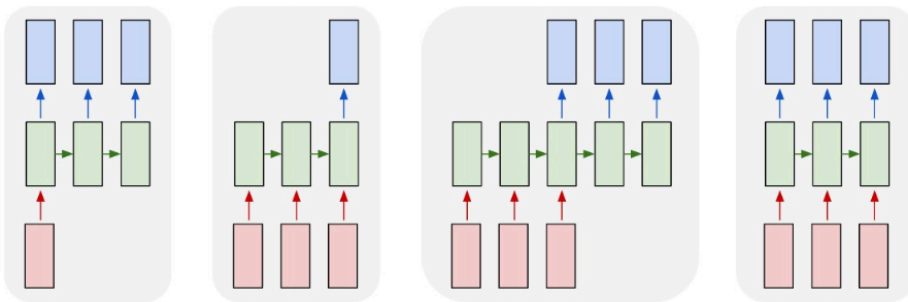
5. Tell the difference between the attention mechanism in Seq2Seq models and Transformer.

Answer:

6. How is the transformer architecture better than RNN?

Answer:

7. Which of the following figure shows the incorrect configuration of RNN structure?



- a) 1
- b) 2
- c) 3
- d) None

Answer:

8. One of the RNN's issue is 'Exploding Gradients'. What is that?

- a) When the algorithm assigns a stupidly high importance to the weights, when your dataset is too big.
- b) When the algorithm assigns a stupidly high importance to the weights, without much reason.
- c) When the algorithm assigns a stupidly high importance to the weights, because the better features.
- d) When the algorithm assigns a stupidly high importance to the weights, when your data is too small.

Answer: B

9. The other RNN's issue is called 'Vanishing Gradients'. What is that?

- a) When the values of a gradient are too big and the model stops learning or takes way too long because of that.

- b) When the values of a gradient are too big and the model joins in a loop because of that.
- c) When the values of a gradient are too small and the model joins in a loop because of that.
- d) When the values of a gradient are too small and the model stops learning or takes way too long because of that.

Answer: D

© 2023 UW STAR Lab