Session 2 Quiz

Due No due date   **Points** 100   **Questions** 13   **Time Limit** 30 Minutes

This quiz is currently locked.

Attempt History

	Attempt	Time	Score
LATEST	[**Attempt 1**](#)	18 minutes	59.17 out of 100

Score for this quiz: **59.17** out of 100
Submitted Jul 25, 2019 at 11:53pm
This attempt took 18 minutes.

Question 1	0 / 10 pts

X is the input (100 Units), O is the last output (200 Units), C is the Cell state (500 units).

For the first forget vector in LSTM, what is the dimension of the Hidden Layer (which will feed into the Sigmoid function, result of which we will multiply with the Previous Context Vector).

(Mention answer exactly in this suggestive format 122*989)

You Answered

600*200

Correct Answers

300*500

300 * 500

300x500

300 x 500

Question 2

10 / 10 pts

When we apply dropout after LSTM, what exactly is getting dropped?

○ Hidden Vector Units

Correct!

◉ Output Vector Units

○ Context Vector Units

○ Input Vector Units

Question 3

10 / 10 pts

When we apply dropout before LSTM, what exactly is getting dropped?

○ Output Vector Units

○ Hidden Vector Units

Correct! ● Input Vector Units

○ Context Vector Units

Question 4 0 / 5 pts

We are feeding in 10 words to a RNN with hidden layer of 4000 parameters. In 1 back prop step, how many total updates would happen?

Correct Answer ○ 40000

You Answered ● 4000

Question 5 5 / 5 pts

We are feeding in 10 words to a RNN with hidden layer of 4000 parameters. During forward and backward prop, how many parameters are kept in the memory?

○ 8000

○ 4000

Correct!

◉ 40000

Question 6   0 / 5 pts

We know that LSTM is better than RNN w.r.t. long term "memory". Which is better for short term memory?

You Answered

◉ RNN

Correct Answer

○ LSTM

○ Both are same

Question 7   10 / 10 pts

What is the role of the tanh function in LSTM?

Correct!   ☑ Add new information

Correct!   ☑ Act like a "down-counter"

Correct!   ☑ Act like an "up-counter"

Correct!   ☑ Remove information (remove is not equal to forget)

Question 8   0 / 5 pts

Why do we need to use Sigmoid to create "input gate", output of which we multiply with an output of tanh gate and then add to the Context Vector?

You Answered

○ This helps us crease a "refresh vector" which helps us remove some information again which forget gate wanted, but refresh not-forgotten units with new information.

Correct Answer

○ This helps us crease a "refresh vector" which helps us remove some information again which forget gate wanted, but refresh not-forgotten as well as forgotten units with new information.

○ This helps us add information only to those units which need to be updated.

○   This helps us remove information from the "new information vector" which our forget vector already wanted

Question 9

6.67 / 10 pts

Why ReLU is not such a good idea for RNNs/LSTMs (when compared to Tanh function)?

Correct!

☑   We need negative activations as well, which ReLU does not provide while TanH does.

Correct!

☑   ReLU does not have a Ceiling for positive numbers, so our activation could grow to a very large number.

☐   ReLU is discontinuous function, we need continuous function for DNNs

Correct Answer

☐   ReLU does not provide an axis of symmetry which is required to create balance between +ve and -ve contexts

Question 10

0 / 5 pts

In LSTM we have OutPut Vector, Input Vector and Cell State Vector. When we use a seed, it is fed to?

You Answered

☑   Cell State and Input Vector

You Answered   ☑ Cell State Vector

Correct Answer   ☐ OutPut Vector

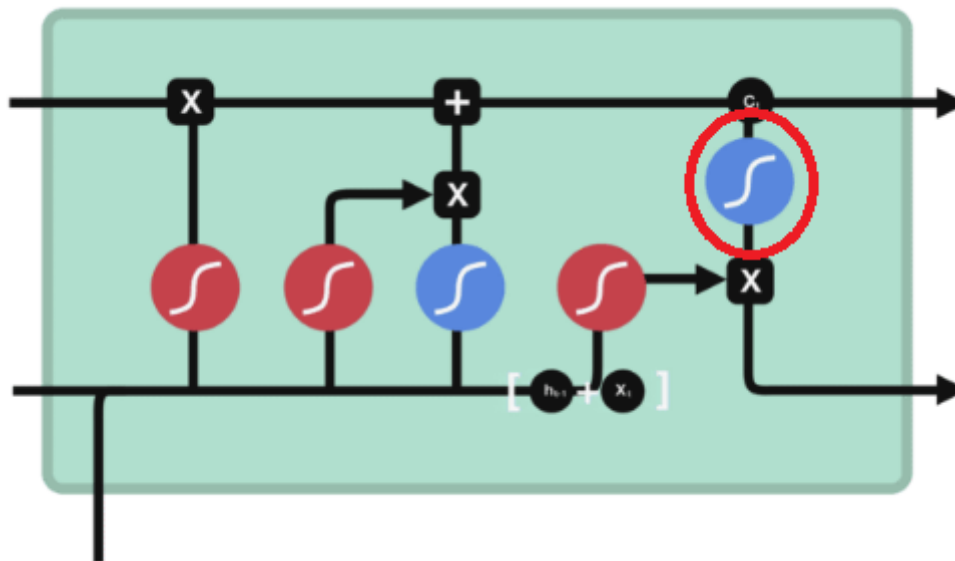You Answered   ☑ Input Vector

---

Question 11   5 / 5 pts

X is the input (100 Units), O is the last output (200 Units), C is the Cell state (500 units).

For the circled tanH gate, we have a FC preceeding it. What is the dimension of this vector?

(Mention answer exactly in this suggestive format: 122*989)

Correct!

> 500*200

Correct Answers

500 x 200

500 * 200

500*200

500x200

Question 12

7.5 / 10 pts

Select which all are true:

☐ It is possible to extract the information fed to a RNN in the first step after 5th iteration

Correct!

☑ Use can theoretically use a FC network in a loop to work (train and predict) like an RNN

Correct!

☑ Dimension of the seeded information to an LSTM must be equal to the dimension of the output needed

Correct!

☑ LSTMs are "conceptually" similar to ResNets

Correct Answer

☐ Output of LSTM can be fed to another DNN for some classification/prediction

Question 13   5 / 10 pts

You Answered

Concatenate Hidden State & Current Input and create	Forget Vector ▼

Correct Answer   Combine Vector

Correct!

What is used to Create Forget Vector	Combine Vector ▼

Correct!

Forget Vector uses which activation functioin	Sigmoid ▼

You Answered

Input vector need which gate	Candidate Context Vector ▼

Correct Answer   Sigmoid and TanH

Other Incorrect Match Options:

- Candidate Context Vector
- Forget Vector
- TanH

Quiz Score: **59.17** out of 100