
CS5787: Exercises 2

FULL_CODE_FOLDER_URL

Mitchell Krieger
mak483@cornell.edu

1 Theory: Question 1 [10 pts]

TODO: Provide your answers to Question 1 here

2 Theory: Question 2 [10 pts]

TODO: Provide your answers to Question 2 here.

3 Theory: Question 3 [10 pts]

In a LSTM cell there are 4 different gates each with its own learnable parameters. For a LSTM with an input of 200 and a hidden state of 200 each gate will have an input weight matrix of 200×200 so that the output of the gate will be a vector of 200 units. The same is true for the hidden state weight matrix. Lastly, each gate has a bias vector of 200. So for one gate there will be $200 \times 200 + 200 \times 200 + 200 = 80200$ parameters. Since there are 4 gates, we can multiply this number by 4 to get the total number of $80200 \times 4 = 320800$ parameters in the LSTM cell.

4 Theory: Question 4 [20 pts]

TODO: Provide your answers to Question 4 here.

5 Practical [50 pts]

TODO: Provide a report detailing your experiments, results and discussion in this section.