Columbia University
IEOR4742 – Deep Learning for OR & FE (Hirsa)
Assignment 4 – Due 11:40 am on Tuesday April 21, 2020

Problem 1 (Long Short-Term Memory - LSTM): In this assignment our goal is to assess `LSTM` capabilities and shortcomings. Use the sample code *example_LSTM.jpynb* to train LSTM separately for each short story, *simpleStory1.txt* & *simpleStory2.txt*, for the following parameters:

- learning_rate = 0.001

- n_input= 3, 4

- n_out= 8

- training_iters = 2000, 5000, 10000

- LSTM layer = 1, 2

- n_hidden in each LSTM = 2, 3, 4

There are total of 36 different training for each story. In testing, for each of these cases, if you insert exactly the same words as were used originally in the short stories the trained model would probably print out the rest of story (up to n_out) verbatim (e.g. for n_input = 3 if you insert Sarah had dinner you would probably get with Steve . Steve had breakfast with David). What you should do it to test it for cases that are not similar to the story (e.g. Sarah had breakfast) and assess the trained model.