CST 463 - Advanced Machine Learning

Dr. Glenn Bruns

# Lab: Attention

**Please work with your team**.

It is very helpful to try to run experiments to test claims that you hear in lecture, read in books, or hear online.

I wanted to test the idea that RNNs have problems learning long-term dependencies, and see if LSTMs do much better with them.

Please copy [this code](https://drive.google.com/file/d/1tv_QP0xx9_nkyka18KKk8FZT2coWbP0S/view?usp=sharing), read it, and run it. The code modifies the IMDB data by selecting two words - one to be used to model positive sentiment, and one to be used to model negative sentiment. For all the IMDB reviews, either the positive word or the negative word replaces a randomly-located word in the original review. If the positive word was used, the target value for that review is set to 1, else the target value is set to 0.

After training, predictions were made on the test data, which was also "injected" with the same positive and negative words.

I wanted to know, would the predictions be better when the random location of the positive or negative word was closer to the end of the review?

Run the code to see my results, then make the following changes:

1. See if the result is different if shorter or longer reviews are used.
2. See if the result is different if an LSTM cell is used instead of an RNN cell.
3. If you still have time, try a different scheme for injecting words and modifying the target values. For example, maybe pairs of words get injected, and the order of the two words controls the target value. Then see if the distance between the words influences the performance of the classifier.