Assignment Instructions: Assignment 3

Purpose

The purpose of this assignment is to apply RNNs to text and sequence data.

Directions

In this assignment, you will accomplish the following:

1. Apply RNNs to text and sequence data
2. Demonstrate how to improve performance of the network, especially when dealing with limited data
3. Determine which approaches are more suitable for prediction improvement

Consider the IMDB example from Chapter 6. Re-run the example modifying the following:

1. Cutoff reviews after 150 words
2. Restrict training samples to 100
3. Validate on 10,000 samples
4. Consider only the top 10,000 words
5. Consider both a embedding layer, and a pretrained word embedding. Which approach did better? Now try changing the number of training samples to determine at what point the embedding layer gives better performance.

Learning Outcomes

CLO 1: Understand what we mean by deep learning
CLO 2: Know the mathematical foundation and structure of neural networks
CLO 4: Understand recurrent neural networks (RNN) and its application to text analysis and sequence data

Requirements

All due dates are included in the Assignment Schedule.

General Submission Instructions

All work must be your own. Copying other people's work or from the Internet is a form of plagiarism and will be prosecuted as such.

You will upload the following to your github account.

3. Your R code (as RMD), and well-documented knitted output as html/pdf/word.
4. A summary, graph/table of your results.
5. Your final grade will be a combination of validation accuracy and your presentation of the results.

You should adhere to the following:

- Remember to use the same repository for the class that you used in Assignment 1.
- Create a new folder under that repository. Call it Assignment 3.
- Upload all files (R, html, doc, etc.) to that folder.
- Provide the link to your git repository in Blackboard Learn for the assignment. The git link should end in **.git**.