Assignment Instructions: Assignment 4

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 11 (Section 11.3, chapter11_part02_sequence-models.ipynb). 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) Before the `layers.Bidirectional` layer, consider a) an embedding layer, and b) a pretrained word embedding.

1. Which approach works better?

2. 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 strict and late submissions will be given zero credit.

General Submission Instructions

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.

Before the deadline: You will upload the following to Canvas *before the deadline*. Provide the link to your git repository in Canvas for the assignment. The git link should end in .git. No files shall be uploaded to this folder yet.
After the deadline: please upload another copy to your github account after the assignment deadline.

1. A summary report that summarizes your results. The graph or table (if any) that you produced should clearly indicate your final conclusions.
2. Your Python code and well-documented comments.
3. Number your question and answers for better clarity.
4. Your final grade will be a combination of validation accuracy and your presentation of the results.

You should adhere to the following:

- Remember to create a new repository for the class, and include your username as part of the repository. For example, Username_64061, or Username_MachineLearning. The Username is very important.
- You will use that repository for ALL assignments. Do not create a new repository for each assignment. Instead, create a subfolder for each assignment. For this assignment, call it Assignment 4.