Assignment 3

Advanced Topics in High-Performance Computing (MCSC 6230G/7230G)

Faisal Qureshi

http://faculty.uoit.ca/qureshi

Computer Science, Faculty of Science

University of Ontario Institute of Technology

Due Back Nov. 27, 11:59 pm

Introduction

The goal of this assignment is for you to experiment with a Long Short-Term Memory (LSTM) neural network for modeling the English language. This task is similar to the LSTM code that we discussed in the class. That code, which is available from the Andrej Karpathy's website (http://karpathy.github.io/2015/05/21/rnn-effectiveness/), models individual characters. You are however asked to model words, i.e., each word is a unique feature. Treating language as such, your network will be able to create coherent English sentences.

I recommend that you use Keras for this assignment. Keras supports all kinds of recurrent layers. For more documentation check the following link https://keras.io/layers/recurrent/.

For training data you should download your favorite pieces of literature (in English) from Project Gutenberg.

Goal of this network

The simplest goal of this network would be to create a sentence given a single "intial" word. However, in order to get the full marks, you will also need to support situations where we can feed this network a sentence (or a paragraph) and the network spits out subsequent sentences, paragraphs.

Submission

Submit a 2 to 4 page report outlining your findings. Please structure your report as if you are submitting a conference/journal paper. The paper should describe your network structure and how you went about constructing feature vectors from words. Again you will be marked on the novelty of your solution, the depth of experiments and the thougtfullness of your conclusions.

Submit via Blackboard.