CST 463 - Advanced Machine Learning

Dr. Glenn Bruns

# Lab: RNN model

The goal of this lab is for you to understand a simple RNN written in pure Python (no TensorFlow). It is helpful to understand an RNN that is written from scratch. Work closely with your team. Talk through the code.

1. Get Andrej Karpathy's simple character-level RNN from his github:

<https://gist.github.com/karpathy/d4dee566867f8291f086>

The code reads text from an input file and creates a "character-level" RNN. This means that the input sequence is composed of characters.

Read the code, and try to answer the following questions: (don't bother looking at the 'backward pass' part of function loss\_fun() for now)

1. How is vocab\_size computed and what does it mean?
2. What kind of variable is char\_to\_ix and what is it for?
3. How many neurons in a layer?
4. What is the input sequence length?
5. What is the output sequence length?
6. What is the size of each element of the input sequence?
7. How are inputs encoded? (for example, as characters, integers, one-hot encoded integers?)
8. Where is the main execution loop in the program?
9. What is the size of a mini-batch?
10. Are the state and the output the same, as we saw in the first simple RNN in our textbook?
11. Why is something called 'smooth\_loss' printed, instead of simply the 'loss' value?
12. In this code, is the RNN unrolled?