EECS 16ML Introduction to RNNs
Fall 2020 Chirag Sharma & Varun Jadia Quiz

RNN Quiz

1. For which of these tasks would you use an RNN, as opposed to a feed forward network?
   (a) Sentiment Classification for Text
   (b) Image Classification, i.e. classify an image as cat or not cat
   (c) Audio transcription, i.e. take an audio file and convert it to text
   (d) Predicting changes in the business cycle (i.e. boom or bust phases)

2. You are given a many to many RNN, that has output $y^{\langle t \rangle}$ at time step t. At the t-th time step, what is the RNN estimating from a probabilistic point of view?

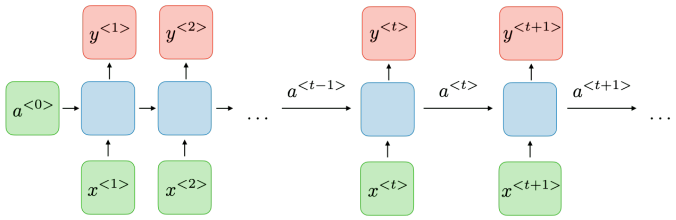

Figure 1: Traditional RNN architecture

3. You have seen how the conventional RNN architecture is prone to the vanishing/exploding gradient problem due to backpropagation, and how GRUs prevent this from happening. Intuitively speaking, is there an activation function from the ones you know of (sigmoid, tanh, ReLU, softmax) that does the same, and why?

4. Imagine you are given a sequence such that $a^{\langle t \rangle} = a^{\langle t+1 \rangle} \cdot k + b$, that is, future terms determine prior terms. Explain how would use an RNN for predicting terms in the sequence without using the bidirectional architecture.

5. How can adding an $L2$ norm regularization term to layer weights affect an RNN? If it helps, consider extreme examples to make your case.

6. Can our Bi-GRU based stock price classifier be used for real time stock predictions?

7. Explain how you would unroll an RNN through time. What neural network does the unrolled RNN look like? Provide a simple example to support your answer.

8. What is the primary difference between conventional backpropagation and backpropagation through time for RNNs?

9. How would you deal with an exploding gradient problem in an RNN without changing the architecture of the network?

10. Explain why the LSTM performed better than the simple RNN in the first section of the project notebook.