

Deep Learning / Fall 2022

Homework 3

Please upload your assignments on or before 5pm ET on October 28, 2022.

- You are encouraged to discuss ideas with each other. But you **must acknowledge** your collaborator, and you **must compose your own** writeup and code independently.
- We **require** answers to theory questions to be written in LaTeX. (Figures can be hand-drawn, but any text or equations must be typeset.) Handwritten homework submissions will not be graded.
- We **require** answers to coding questions in the form of a Jupyter notebook. It is **important** to include brief, coherent explanations of both your code and your results to show us your understanding. Use the text block feature of Jupyter notebooks to include explanations.
- Upload both your theory and coding answers in the form of a **single PDF** on Gradescope.

-
1. **(3 points)** *CNNs vs RNNs*. Until now we have seen examples on how to perform image classification using both feedback convolutional (CNN) architectures as well as recurrent (RNN) architectures.
 - a. Give two benefits of CNN models over RNN models for image classification.
 - b. Now, give two benefits of RNN models over CNN models.
 2. **(4 points)** *Recurrences using RNNs*. Consider the recurrent network architecture below in Figure 1. All inputs are integers, hidden states are scalars, all biases are zero, and all weights are indicated by the numbers on the edges. The output unit performs binary classification. Assume that the input sequence is of **even** length. What is computed by the output unit at the final time step? Be precise in your answer. It may help to write out the recurrence clearly.
 3. **(3 points)** *Attention! My code takes too long*. In class, we showed that a computing a regular self-attention layer takes $O(T^2)$ running time for an input with T tokens. Propose two different ways to reduce this running time to $O(T)$, and comment on their possible pros vs cons.
 4. **(6 points)** *Sentiment analysis using Transformer models*. Open the (incomplete) Jupyter notebook [here](#) in Google Colab (or other cloud service of your choice) and complete the missing items. Save your finished notebook in PDF format and upload along with your answers to the above theory questions in a single PDF.

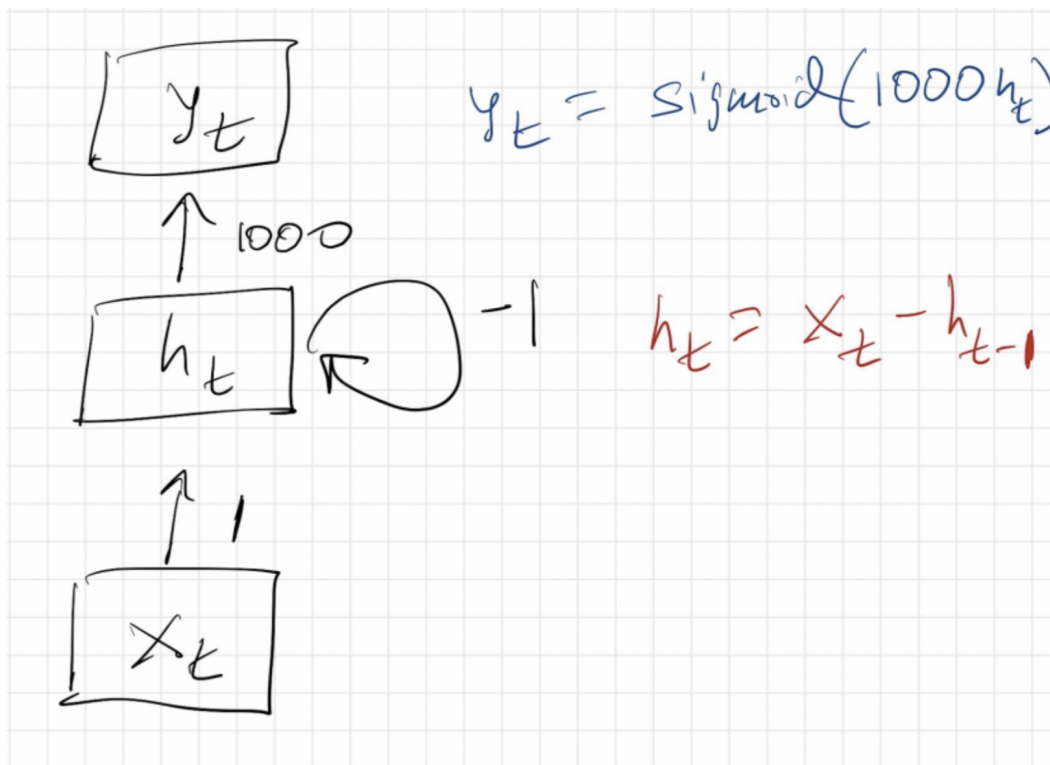


Figure 1: RNNs