Homework Set 6, CPSC 6430/4430

LastName, FirstName

Due 05/03/2024, 11:59PM EST

DNN Implementation and Application

Please refer to Jupyter Notebook.

CNN Implementation and Application

Please refer to Jupyter Notebook.

Understanding RNN

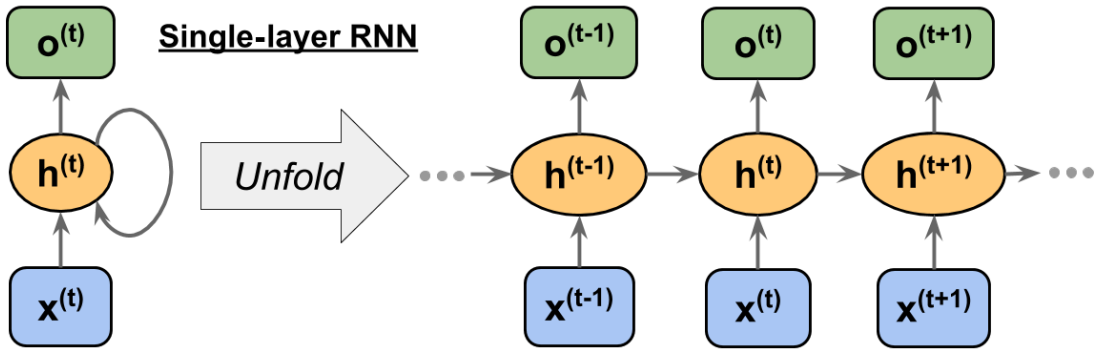

Figure 1: Assume $x^{t-1} = [1, 2, 3, 4, 5], x^t = [0, 2, 4, 6, 8], x^{t+1} = [0, 1, 2, 3, 5]$.

Assume there is no bias term and $W_{xh} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 & 9 \end{bmatrix}, W_{hh} = \begin{bmatrix} 2 & 3 \\ 3 & 7 \end{bmatrix}$. Please determine h^{t-1}, h^t, h^{t+1} where there is no activation function in each step or output.

Backpropogation

Consider the MLP structure as below and the layer h to z is connected with Sigmoid function ($z_i = \frac{1}{1+exp(-h_i)}$) while both x to h and z to y are fully connected layers. The initial weight for x_i and h_j is $i * j$ while z_k and y_l is $(k+1) * (l+1)$. Assume the groundtruth of y_1, y_2 is $5, 10$ and the loss function is defined by *Mean Squared Error (MSE)*. Given the input x_1, x_2, x_3 being $1, 2, 4$ and assume we are using gradient descent method to update the parameters with $stepsize = 0.01$ and there is no bias term.

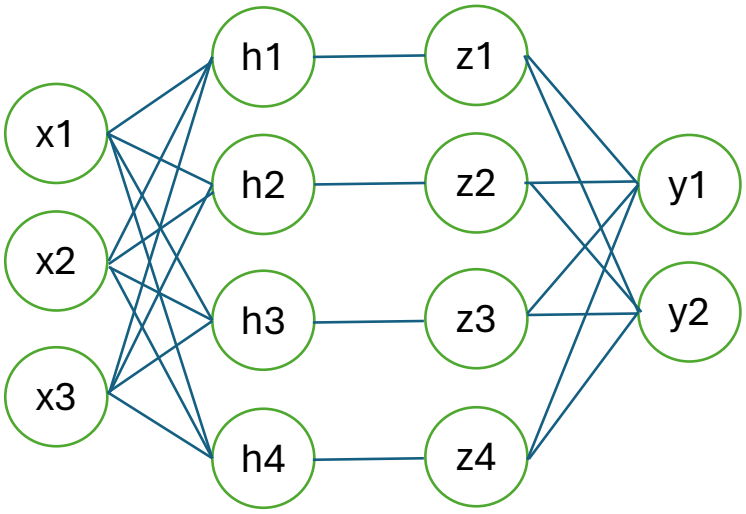

Figure 2:

1. What are the predicted y_1, y_2?
2. How many parameters do we have to learn for this network?
3. What is the weight to connect x_1 with h_1 and the weight to connect z_2 with y_1 after the first update, respectively?