

Recurrent Neural Networks

Nikhil Sardana

April 26, 2017

1 Long Short-Term Memory Problem

Recall that LSTMs are used to remove the vanishing gradient problem that plagues standard Recurrent Neural Networks. The LSTM structure is shown below.

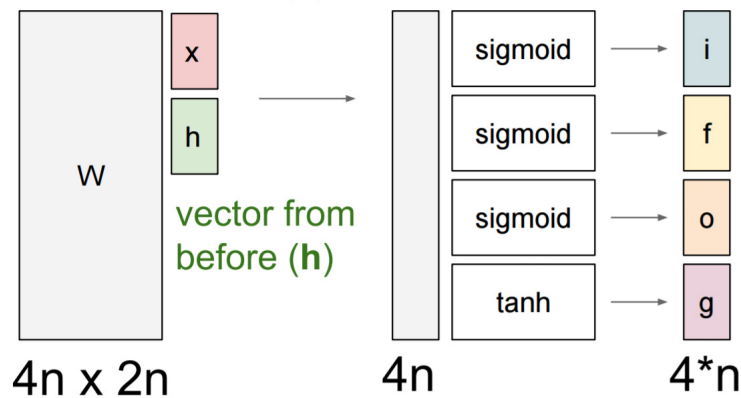


Figure 1: LSTM

$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \text{sigm} \\ \text{sigm} \\ \text{sigm} \\ \text{tanh} \end{pmatrix} W \begin{pmatrix} x_t \\ h_{t-1} \end{pmatrix}$$

$$c_t = f \odot c_{t-1} + i \odot g$$

$$h_t = o \odot \tanh(c_t)$$

Given some input vector x_t , previous hidden vector h_{t-1} , and previous cell state c_{t-1} , write the python code to calculate the new cell state, c_t , and the new hidden vector h_t .

Use the numpy library, but do not use any machine learning libraries (Tensorflow, Scikit-Learn, etc.).