PHYS/4036 Workshop 5

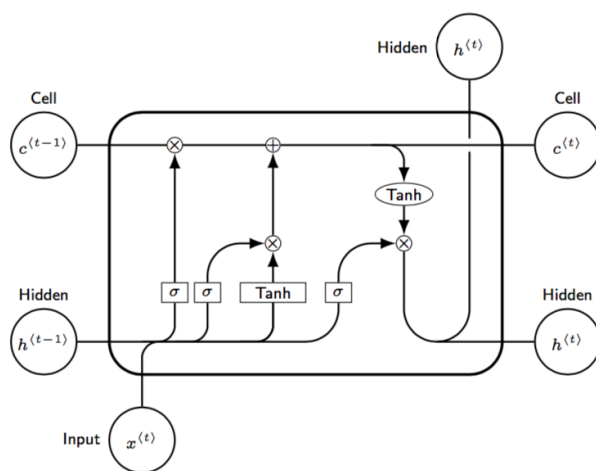Adam Moss and Maggie Lieu

March 2021

Question 1:

Figure 1: LSTM Cell.

The basic LSTM cell as discussed in lectures is shown in Fig. 1. During back-propagation, suppose the error

$$\frac{\partial J}{\partial \boldsymbol{h}^{(t)}} \tag{1}$$

is known, and that

$$\frac{\partial J}{\partial \boldsymbol{c}^{(t)}} = 0 \,. \tag{2}$$

Calculate the resulting error

$$\frac{\partial J}{\partial \boldsymbol{h}^{(t-1)}} \,. \tag{3}$$

1

Question 2:

In lectures we hand-coded the BPTT algorithm to train an RNN language model to predict the next word in a sentence. Using the same training corpus, train a many-to-many LSTM model using TF2 to perform the same task, and compare your results against a vanilla RNN.

In this example we concatenate all the sentences into a single vector to make it easier to feed into the TF2 dataset API. We therefore only have a single stop word between sentences. The downside of this approach is that sentences in different reviews will have different context, so ideally we would treat different reviews separately and pad inputs where necessary.

Tensorflow has a similar character level RNN (not at the word level) here to help you: `https://www.tensorflow.org/tutorials/text/text_generation`

We have written the initial data-processing in the worksheet below, so your task is to build and train the model, and assess its performance.

```
   https://colab.research.google.com/github/adammoss/MLiS2/
blob/master/workshops/workshop5/rnn.ipynb
```