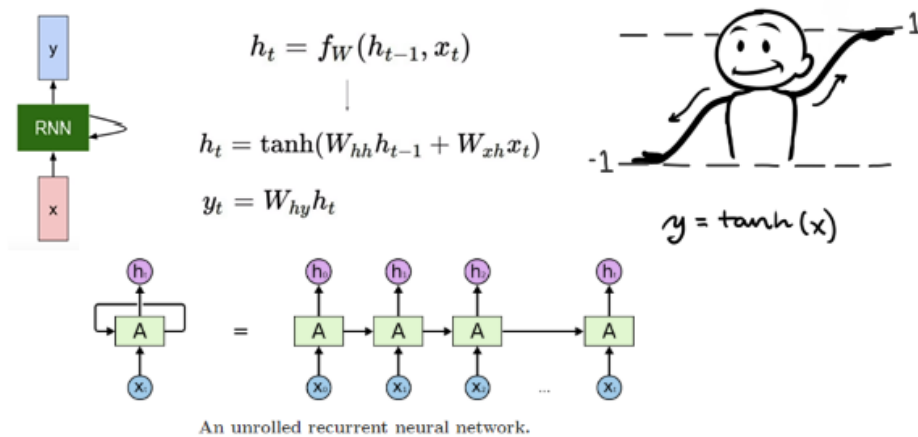
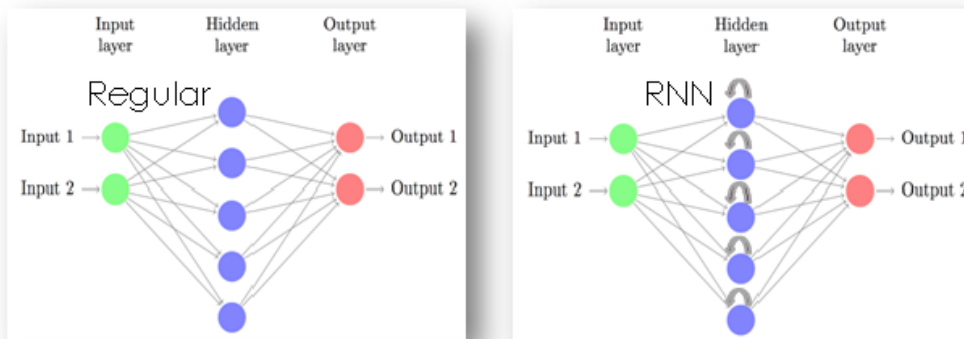


Dr-Suess Machine Learning LSTM

This project focuses on using LSTM machine learning models to create Dr Seuss like stories from a user input seed.

Understanding LSTM

RNN and LSTM



Good for predicting something like : "the clouds are in the **'sky'**"
 (If its seen a sentence like that before)

Bad for predicting something like: "I grew up in France.
 Although I live in America, I can still speak fluent **'french'**"
 (because its further back)

Long Short Term Memory (LSTM)

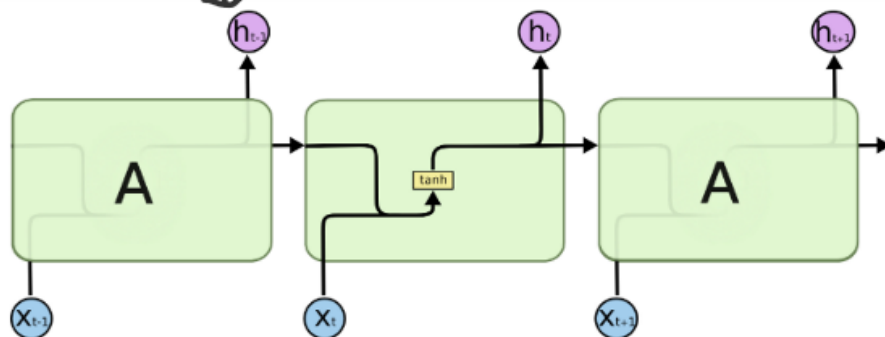
Vanilla RNN

LSTM

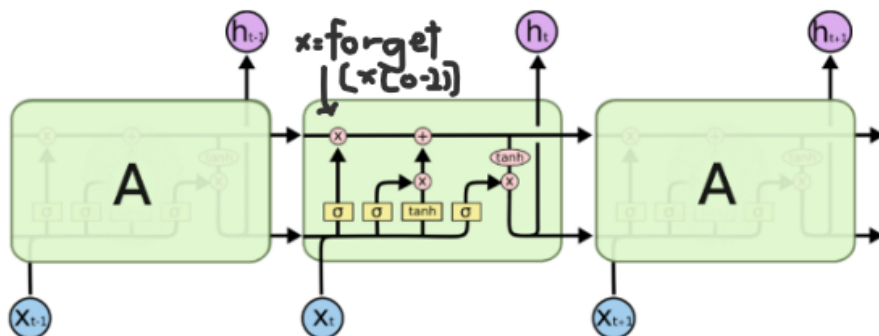
$$h_t = \tanh \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \right)$$

output of a cell
 $h =$

$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$
$$c_t = f \odot c_{t-1} + i \odot g$$
$$h_t = o \odot \tanh(c_t)$$



The repeating module in a standard RNN contains a single layer.



The repeating module in an LSTM contains four interacting layers.

Hochreiter and Schmidhuber, "Long Short Term Memory", Neural Computation

RNN vs LSTM cell representation, source: stanford

Getting Started

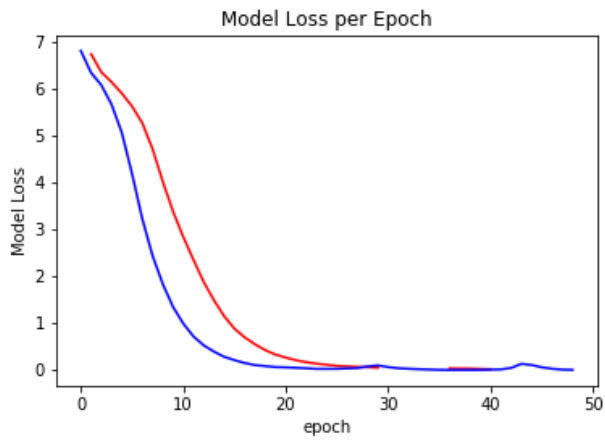
word_based.ipnb is our experimentation with word based LSTM models. This has the advantage of outputting only words that it has seen, and will therefore be more readable. However this makes for less training data availability than the character based model. Our character based model can be found in seuss.ipnb or seuss.py

Prerequisites

The model is very big, and may require a dedicated GPU to run. AWS provides GPU's and can be experimented on with a free trial. You'll need: npm, flask, and sufficient hardware for model utilization, or training should you want to play with the model.

Running the tests

The parameters of this model creation code are easily manipulatable for experimentation. When the code runs it will save model information, as well as training loss, accuracy and model output data in a json file for visibility. You can see previous output files in AlldataX.json



Built With

- [Dedicated GPU] (<https://aws.amazon.com/ec2/instance-types/p3/>) - Or similar

Authors

- Jennifer Shtaway - *Model Training, web developement, graphical representations of results*

Acknowledgments

- Hat tip to my team, Vish, Greyson, and Alper who collaborated with me on this.