COMP9444 Neural Networks and Deep Learning
Session 2, 2018
Project 2 - Recurrent Networks and Sentiment Classification

Prajwal Rao | z5176504

Overview

Started with basic model, trained with different hyper parameters and tweaked the text pre-processing.

Pre-processing

Converted a given movie review to a list of words by:

- Changing text to lower case, removing additional spaces and stop words.
- Removing HTML tags.
- Retaining only alphabetic characters. (Numbers are relevant, but could be taken out of context)
- Modifying stop words list to exclude '*not*' as it could possibly change the entire meaning of a sentence. Example: "*The movie was **not** bad.*" (Sentiment: Non-Negative) vs. "*The movie was bad.*" (Sentiment: Negative)

TF Graph

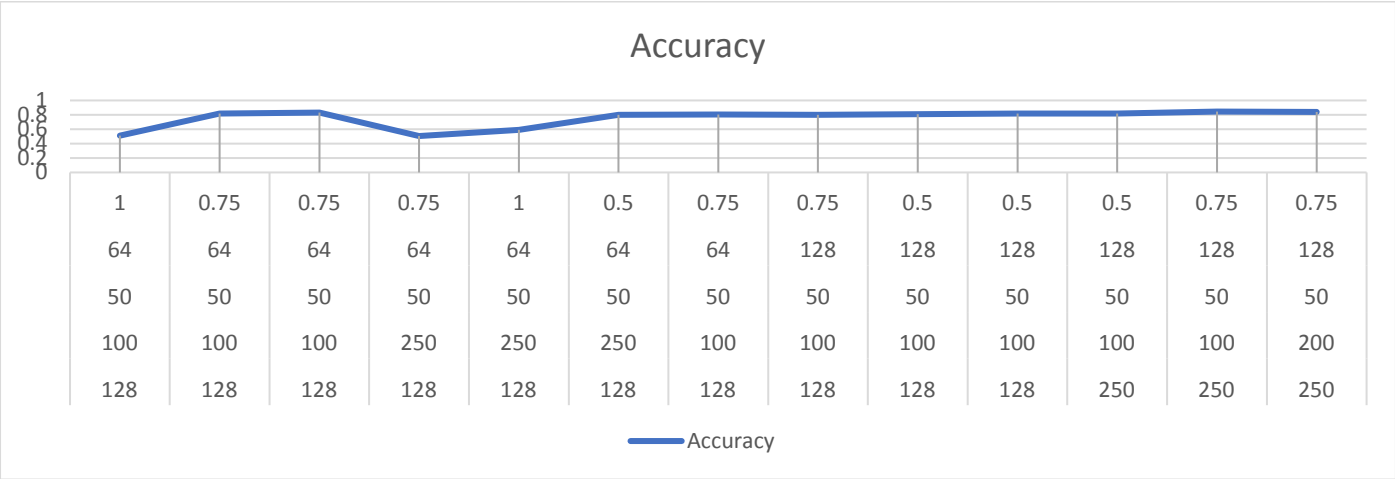
Chose basic Long Short-Term Memory (LSTM) cell with a count of 128 units in a Dynamic Recurrent Neural Network.

Passed the last vector of the output from this RNN into a dense layer (TF.layers.dense) to obtain predictions.

Finally, calculated the error with softmax function and summed it over entire batch to get loss.

Used an Adam optimizer to minimize the loss.

Hyperparameter Combination vs Accuracy



Summary

The highest validation accuracy – **0.84** was obtained with the following hyper parameter combination:

BATCH_SIZE	MAX_WORDS_IN_REVIEW	EMBEDDING_SIZE	LSTM_UNITS	DROPOUT
250	200	50	128	0.75