COMP9444 (18S2) Assignment 2
Submission Deadline: Sunday, 23th September 2018 at 23:59:59
z5136414 Andrew Wirjaputra

Preprocessing
The preprocessing stage is done in a fairly simple way. First, all reviews will be converted into lowercase to match the GloVe embeddings. Then, they will be stripped of punctuations first to make it easier to remove stop words and contractions. The set of stop words provided in implementation.py has been expanded using common English contractions.

Network Structure
Considering the comparative study of CNN and RNN for Natural Language Processing, a 1-layer GRU network is chosen since GRU network tends to perform better in sentiment classification. Then again, tuning hyper parameters are often more important than picking the ideal architecture.

Batch size selection is done in order to balance the algorithm's capability to jump out of a local minimum while making sure that it eventually converges instead of bouncing around. After experimenting on the batch size, I settle at BATCH_SIZE = 200, as using higher or less result in a lower validation accuracy.

We want to retain as much information from each training review. After experimenting on the maximum number of words in review, I settle at MAX_WORDS_IN_REVIEW = 200, as further increasing it no longer affect validation accuracy while using less result in a lower validation accuracy.
Note: based on the training reviews, there are ~230 words per review on average

The optimal number of hidden units depends on the complexity of the learning task. After experimenting on the number of hidden units, I settle at NUM_UNITS = 100, as further increasing it no longer affect validation accuracy while using less result in a lower validation accuracy.