**COMP9444 Assignment 2 Report**

**z5089358**

**Fengting YANG**

**Design:**

Each sentence has a vector of 50 dims and the max words in review is set to 200. Batch size is 128.

In preprocess, “<br />”is replaced by space. Then, all punctuations and digits are also replaced by spaces. All words are in lower case. As an addition of filter, all words have a length less than 3 are removed. Stop words and some special cases are also removed.

In training part, one RNN layer is followed by bidirectional LSTM with attention layer. Forward sequence and backward sequence are considered. For each LSTM, it has 128 LSTM units. In RNN part, weights are initialized by truncated normal distribution with standard deviation 0.2. Biases are initialized by constant 0. Softmax function is considered as activation function. Additionally, l2 penalty is used to prevent overfitting. The penalty factor is 1e-4.

**Result:**

At iteration 3700, training set accuracy reaches 94% while evaluation set only has a 87.1% accuracy. As the training continues, it becomes overfit.