

## **COMP9444 Homework 2**

First a convolution layer was chosen because of its ability to try and link sentiments between words. A filter size of 32 was found to work the best while evaluating in strides of size 3. A ReLu activation function is used after this step to replace negative outputs with zero.

The output from the convolution layer is then pooled to allow further processing and to reduce overfitting.

Afterwards various dropout rates are tested to see which provides the best results in reducing overfitting. I settled on a dropout rate of 1.0 as it seemed to prove to work the best.

The result is then passed through a batch normalisation filter to further reduce overfitting of the network.

Finally, the network is passed through a LSTM node to allow the recurrent memory component of the network. After numerous testing between a GRU and LSTM unit, it was found a LSTM with 128 units worked the best.

Finally the output is passed through a dense layer to allow calculation of the loss and error.