Assignment\_2 Report

## Ziyang Lin

1. How did you preprocess your dataset and why?

Convert each original headline into a normal sentence (remove ”<” and “/>”).

Get the edited version of headlines by doing word substitution using regular expression.

Pair the original headline and the corresponding edited one and put them into a list.

Do tokenization and lowercasing for each headline pair in the list.

Reason:

To improve the generalisability of the model.

1. Briefly describe your model. Why did you pick up this model? What are the final set hyperparameters? How did you tune them?

My model is a two inputs’ feed forward neural network in which two input matrices representing all the original headlines and their corresponding edited headlines respectively are passed simultaneously to the first so called embedding layer of the model to get the word embedding of the fixed dimension for each word in the headline. Following above the model will do averaging for each headline to get the ‘document representation (vector)’ for each headline. Then these headlines’ vector representations are passed to a combination of three concatenated fully connected layers where the information about “how humour they are” are encoded. The Relu activation is applied after output from each of the first two hidden layers to prevent gradient vanishing and gradient explording. Finally all weighted sums or all vector products between the n-th row of the original matrix and the n-th row of the edited matrix are computed such that a vector with the size (origin\_headlines\_num, 1) is returned. The loss function of my model is RMSE, which is broadly applied for the regression tasks.

Hyperparameters:

EPOCHS = 100

LRATE = 0.145

EMBEDDING\_DIM = 300

HIDDEN\_DIM\_1 = 100

HIDDEN\_DIM\_2 = 50

HIDDEN\_DIM\_3 = 10

Learning rate too high will make the loss function keep oscillating strongly. The number of Epochs and the number of embedding dimensions are setted in terms of the size of the corpus and vocabulary.

The learning rate is tuned (reduced) automatically throughout the epoches using the scheduler.

1. How did you evaluate the performance of your model and why did you decide to do so?

Using the Root Mean Square Error (RMSE) to evaluate the model. Because the model is designed for the regression task such that the model seeks to predict the continuous values in which we just compare the predicted output value with the golden truth and using the difference between these two values as the metric to look for a function that is best fitted in with these ground true data.