

# Seniment Analysis

---

Changmao Li

## Model Architecture

The main architecture is mainly made of trainable embedding layer, convolutional layer, attention layer. The following is details:

Layer (type) # Comment	Connected to
input_1 (InputLayer)	
embedding_1 (Embedding) # trainable embedding layer	input_1
bidirectional_1 (Bidirectional) # bidirectional lstm	embedding_1
conv1d_1 (Conv1D) # convolution layer	bidirectional_1
lambda_1 (Lambda) # attention layer	conv1d_1
global_max_pooling1d_1 (GlobalM) # max pooling layer	conv1d_1
concatenate_1 (Concatenate) # merge pooling and attention	lambda_1, global_max_pooling1d_1
dense_1 (Dense) # output layer	concatenate_1

### - Trainable embedding layer

- Build word dict  
I use all training, dev and test data to build word dict, so when you run it, **don't remove** any of them.
- Build embedding matrix
- Make the input turn to sequence of id in word dict

### - Bidirectional LSTM

- Use a bidirectional lstm to collect order information.

### - Convolution layer

- Convolution kernel that is convolved with the layer input over a single spatial (or temporal)

dimension

- Capture n-gram feature

## - Attention layer

- $Attn = sentence^T * normalize(tanh(sentence * sentence^T))$

## Tuning

- Three important hyperparameters: filters and kernels, hidden unit of lstm
- I run full search by talos and get the full result of hidden unit{100,200,300}, filters{64,128,256}, kernels {2,3,4} which in the "hw2\_007.csv" file. The highest six results are following which are all above 0.46:

filters	kernel_size	Hidden unit	val_acc	val_loss
256	3	200	0.4668483200341761	0.7055351275766129
128	3	100	0.4668483200341761	0.752702211688701
128	2	300	0.4650317896072914	0.7158156024417102
128	4	300	0.46049046354007983	0.6899680490585356
128	4	200	0.46049046354007983	0.7111503604159505
64	4	300	0.46049046354007983	0.7178201186934767

And I got that the two best choice of the hyperparameters are 256,3,200 and 128,3,100. And they have same highest dev score, and since the talos cannot save model, so I reran the program third times for each hyperparameter and pick the highest dev score model.