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

- Attention layer

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

Tuning

- Two important hyperparameter: filers and kernels,
- I run full scan by talos and get the full result of filers{64,128,256} kernels {2,3,4,5,6,7} which is following

filters	kernel_size	val_acc
128	4	0.4750408719616733
128	3	0.4686830157111927
256	3	0.4759491374728682
128	5	0.4732243415347887
64	3	0.4786830157111927
64	6	0.4532425071368100
128	2	0.4732243415889255
256	4	0.4650499546002735
64	4	0.4695912806810194
256	5	0.4750408722052891
64	5	0.4596003634414274
128	6	0.4704995461922142
256	6	0.4577838330145428
128	7	0.4632334242951967
64	7	0.4623251589734808
256	2	0.4596003636309063
64	2	0.4441598547046342
256	7	0.4623251589734808

And I got that the best choice of the hyperparameters are 256 and 3. And its highest dev score: 0.4759491374728682.