User manual

Weiheng Li

I build this program on the basis of SemEval 2017 task4 champion team Datastories. They used a very good twitter sentence preprocess tool and trained a good set of word2vector database. It is impossible for me to write them on my own in a few weeks and my program is targetting at find the best architecture of Neural network which can be used in sentiment analysis task.

The discription of my dir:

datarecord.docx: to record my experiment output

Finalreport.docx: final report

final.py: my final program including load data, preprocess and code for NN.

generate\_figure.py: to generate image for final report

taskimage.png: current running neural network structure.

dataset: data for taskA and the respondingly loading script

others: other code used in preprocessing

To run the final.py, you need to download

<https://mega.nz/#!u4hFAJpK!UeZ5ERYod-SwrekW-qsPSsl-GYwLFQkh06lPTR7K93I>

And put it in the root.

In the final.py, There are two target cnn named target\_RNN and target\_RNN2, target\_RNN2 is another implementation of attention layer which has been proved to be an mistake. In the target\_RNN, I have given discription which structure is every episode coding for. You can also look up the structure of NN in the taskimage.png after your training begin.

The code episode with tag #bilstm+bilstm+attention merge bilstm+attention and #bilstm+bilstm+attention is N1 and N2 as mentioned in the report. The origin in myode means N2 structure.

This code can generate result that rank 1st in the 2017 SemEval task4. But my hardware limited my result, I can’t fully test all my structure and run it for more epoch. And I also want to try more structure. I would try that in the future.