EEL 6935

Programming Assign 2 – NLP

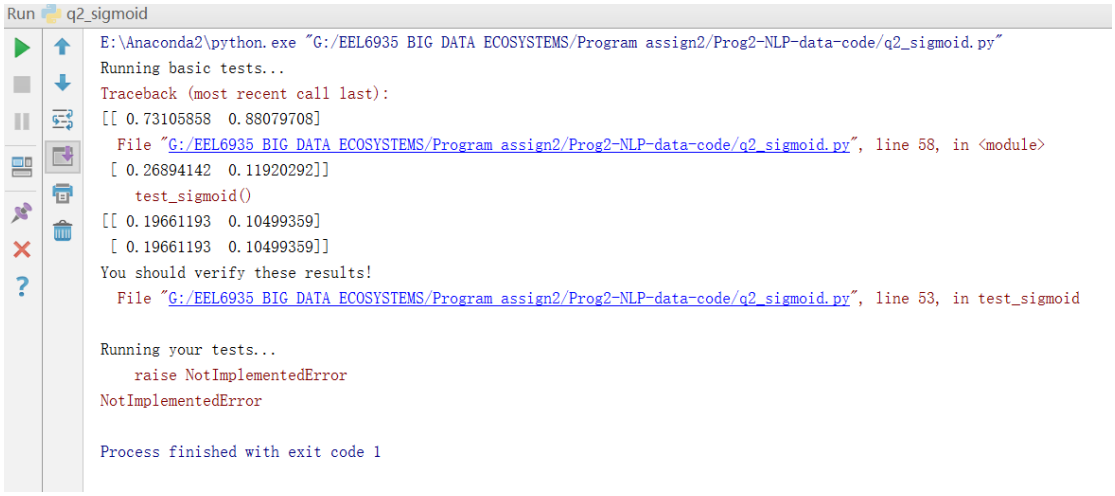
Github link: https://github.com/jeness/EEL6935BIGDATA-project2
Google Doc link: https://docs.google.com/document/d/1IzXEzADv-hvVY_Fj_XyC6DAR5aBYKrIBL3GzovdNorw/edit?usp=sharing

Haoran Yu
UFID:18106994

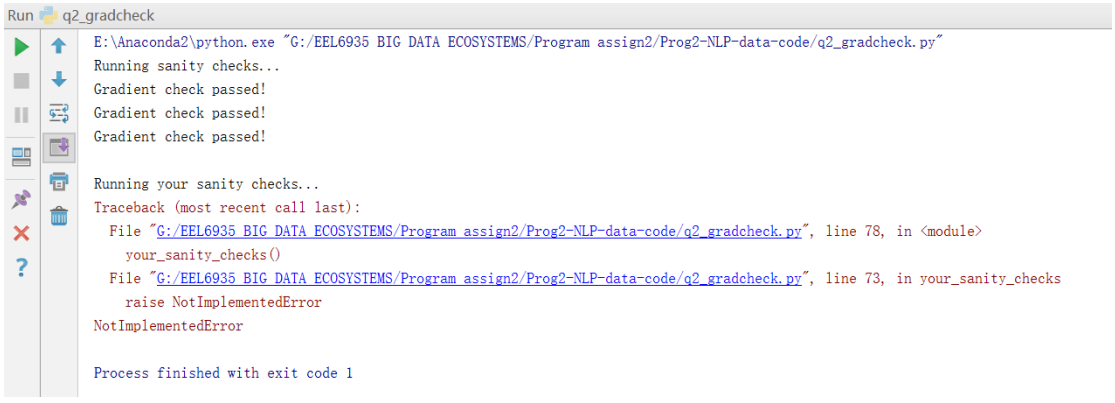
3.1 Softmax

3.2 Neural Network Basics

q2_sigmoid.py



q2_gradcheck.py



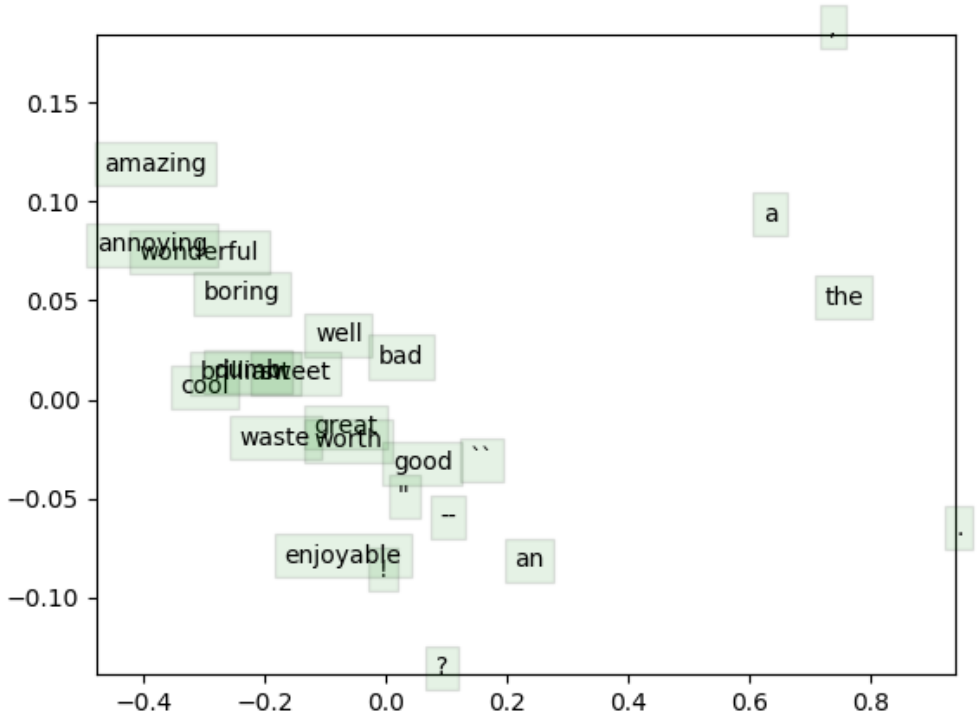
q2_neural.py

```
Run q2_neural
E:\Anaconda2\python.exe "G:/EEL6935 BIG DATA ECOSYSTEMS/Program assign2/Prog2-NLP-data-code/q2_neural.py"
Running sanity check...
Gradient check passed!
Traceback (most recent call last):
Running your sanity checks...
  File "G:/EEL6935 BIG DATA ECOSYSTEMS/Program assign2/Prog2-NLP-data-code/q2_neural.py", line 89, in <module>
    your_sanity_checks()
  File "G:/EEL6935 BIG DATA ECOSYSTEMS/Program assign2/Prog2-NLP-data-code/q2_neural.py", line 84, in your_sanity_checks
    raise NotImplementedError
NotImplementedError

Process finished with exit code 1
```

3.3 Word2Vec

q3_word_vectors.png



3.4 Sentiment Analysis

I mainly try two ways of regularization.

First one is,

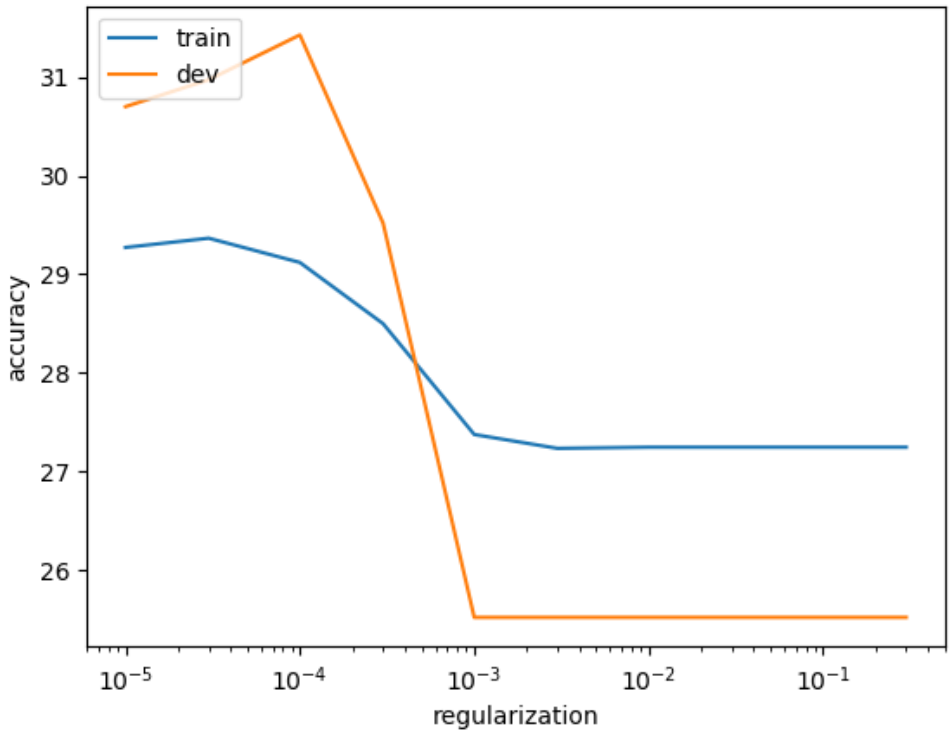
REGULARIZATION = [0.1,0.3,0.01,0.03,0.001,0.003,0.0001,0.0003,0.00001,0.00003]

Console output:

```
=== Recap ===
Reg      Train       Dev
1.000000E-05    29.272004    30.699364
3.000000E-05    29.365637    30.971844
1.000000E-04    29.119850    31.425976
3.000000E-04    28.499532    29.518619
1.000000E-03    27.375936    25.522252
3.000000E-03    27.235487    25.522252
1.000000E-02    27.247191    25.522252
3.000000E-02    27.247191    25.522252
1.000000E-01    27.247191    25.522252
3.000000E-01    27.247191    25.522252

Best regularization value: 1.000000E-04
Test accuracy (%): 27.556561

Process finished with exit code 0
```



Second one is,

REGULARIZATION = np.logspace(-5,1,20)

Console output:

Run q4_sentiment

```
Dev accuracy (%): 12.806540

=== Recap ===
Reg      Train       Dev
1.000000E-05    29.272004   30.699364
2.069138E-05    29.353933   31.062670
4.281332E-05    29.201779   31.153497
8.858668E-05    29.236891   31.244323
1.832981E-04    29.014513   30.154405
3.792690E-04    28.288858   27.792916
7.847600E-04    27.621723   26.067212
1.623777E-03    27.258895   25.431426
3.359818E-03    27.235487   25.522252
6.951928E-03    27.247191   25.522252
1.438450E-02    27.247191   25.522252
2.976351E-02    27.247191   25.522252
6.158482E-02    27.247191   25.522252
1.274275E-01    27.247191   25.522252
2.636651E-01    27.247191   25.522252
5.455595E-01    25.971442   26.339691
1.128838E+00    12.816011   12.806540
2.335721E+00    12.816011   12.806540
4.832930E+00    12.816011   12.806540
1.000000E+01    12.816011   12.806540

Best regularization value: 8.858668E-05
Test accuracy (%): 27.375566

Process finished with exit code 0
```

4: Run 6: TODO Terminal Python Console