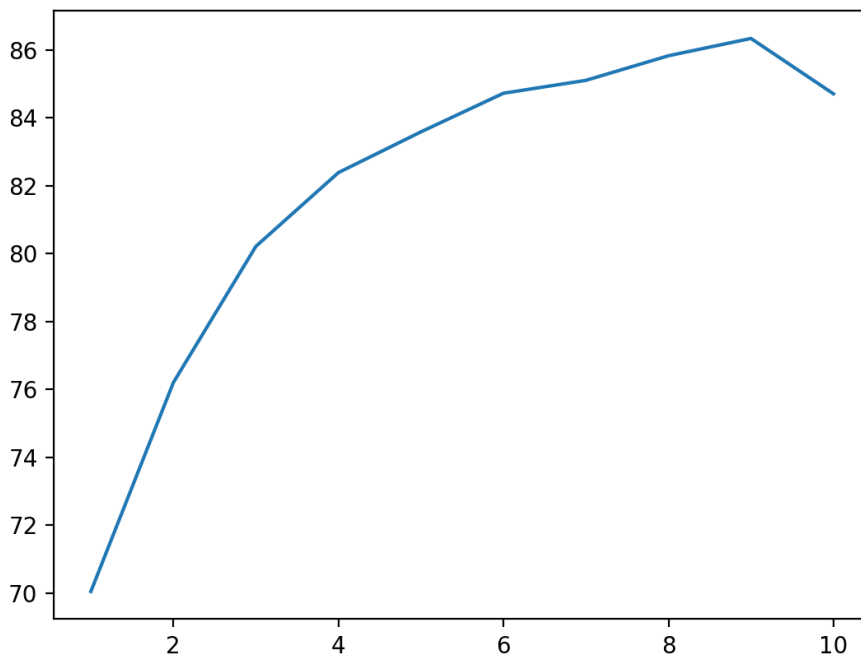# SI 630 Homework 4
Xiaoxue Xin

Task 1
See in python files

Task 2
Problem 2.1

Problem 2.2
From the plot, we can see that there is an increasing trend in the
plot. It means the training UAS score gets better and better. But
there is a fluctuation at the end point.
When the line tends to flat, which means it converge, we would want to
stop the training process.

Problem 2.3
Validation acc: 0.344
— test UAS: 71.01

Task 3
Problem 3.1 & Problem 3.2
See in python files

Problem 3.3
When the line tend to flat, which means the algorithm may converge, we
would like to turn off the training process. If we allow the weights
to change, then the contribution of gensim model seems like give it an
initial situation. And the weights will change after each iteration.
Then we may get similar result as previous one.

Problem 3.4
The plot of twitter score is shown below.
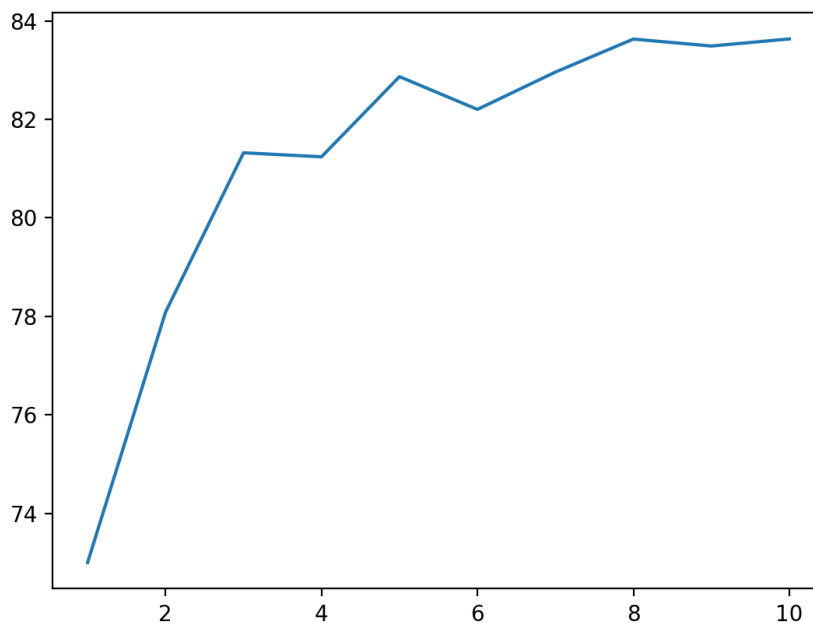
From the plot, we can see there is an increasing trend, which is
similar with previous plot. There is also fluctuations in the plot.
The range of the score is also similar with previous result. The score
has such trend, since result is better after iterations.
From the consequence, we can get a conclusion that our pre-training
works well compared with the already trained twitter dataset. Since we
didn't update the gradient in each iteration for the twitter data,
similar result means our pre-trained process makes sense.
I am not surprised with this result. The test result is shown below.
Validation acc: 0.339
— test UAS: 74.30

The plot for the wiki score is shown below.
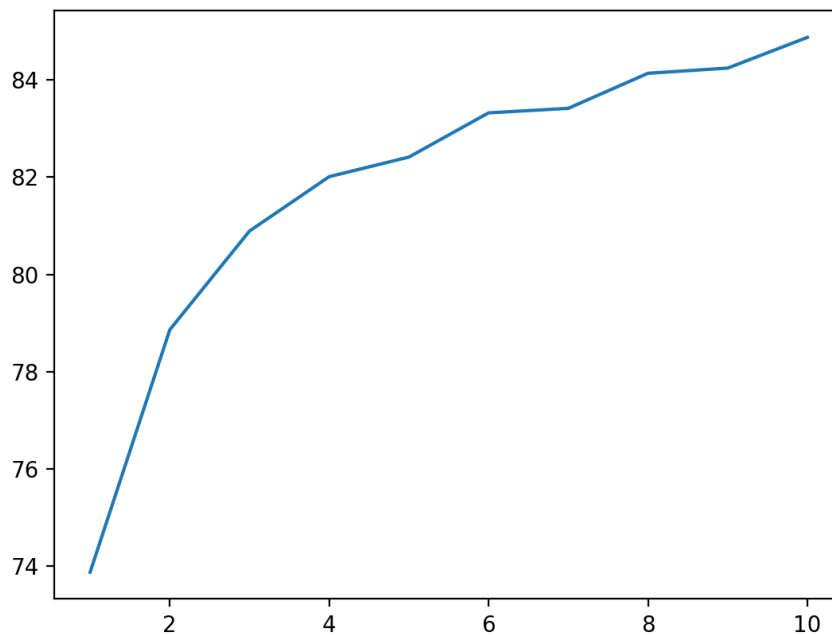
From the plot, we can see that there is an increasing trend in this plot. It looks similar with the previous plots. And there is fluctuation in the plot. Since the we still cannot see the flat tendency, maybe we could get better result when we run more epochs. The reason may be that the data in wiki file is structured. Thus the training result should be better than the twitter file. I am not surprised about the result.

Similarly with the twitter reason, without update gradient each iteration, we can compare it with the performance of our model. It seems that there is no big difference between well defined existed result and our trained model. The test consequence is shown below.

Validation acc: 0.336

— test UAS: 74.48