Report1

Refael Shaked Greenfeld 305030868,

Danit Yshaayahu 312434269

100 and vector dim of 30.

Training data: 4000

dev data: 400

Test data: 400

Number of iteration: 6

Time: 41.0 seconds (on GPU GTX 1080)

It succeed on both. It manage to converge in the train pretty fast (seem with the loss it took about 4 iteration and it got 100% on dev in the 6 iteration, where it took the best model from, and ran it on test which also got 100%)

In the beginning we had an extremely long sequences not limited by length (the limitation was about 2000 characters and was an average of about 1000), when we limited the sequence to 180 (20 for every sub sequence) the model managed to learn. We didn’t need to do any hyper parameter tuning and we ran with batch of 10 lr 0.001 hidden dim of LSTM



