

# Report1

Refael Shaked Greenfeld 305030868,

Danit Yshaayahu 312434269

The training succeed on both the training, dev and test datasets. The model managed to converge on the training data pretty fast. Seem with the loss it took about 6 iteration and it got 100% on dev after 6 iteration, and because that in our setup we take the model that got the best results on the dev, this is the model we ran on the test data which also got 100%.

## Training Details:

Training data: **4000**

dev data: **400**

Test data: **400**

Number of iteration: **6**

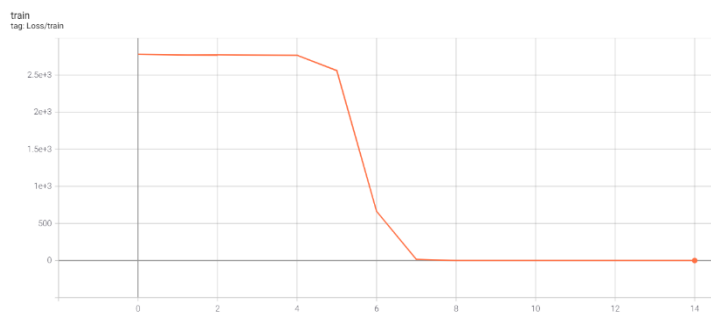
Time: **41.0** seconds (on a single GTX 1080 GPU)

## Questions

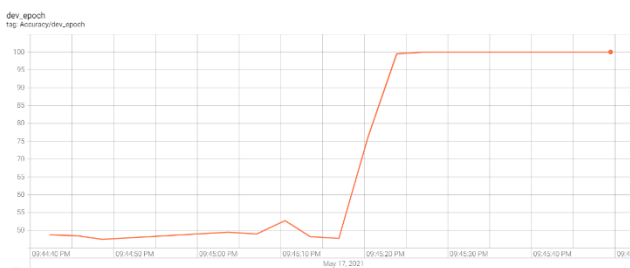
In the beginning we had an extremely long sequences not limited by length (the limitation was about 2000 characters and was an average of about 1000), when we limited the sequence to 180 (20 for every sub sequence) the model managed to learn. We didn't need to do any hyper parameter tuning and we ran with a batch of 10 lr 0.001 hidden dim in the LSTM of 100 and a vector size of 30.

We also added graph of the training Loss and the Dev accuracy both with global time view and both with epoch view.

## Train Loss



## Dev accuracy – Global time



## Dev accuracy – Epochs

