

Part1 –

Ben cohen 322693722

Asaf Rosemarin : 211544846

Our model includes batch normalization – 100 inputs on each batch .

Regularization : $5e-5$

Optimizer : Adam (pytorch)

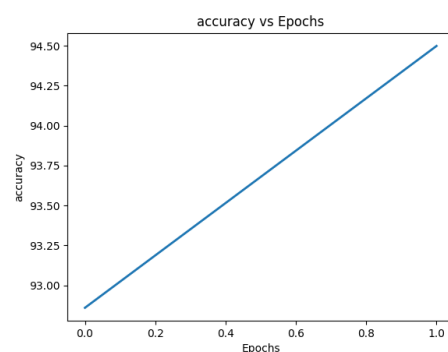
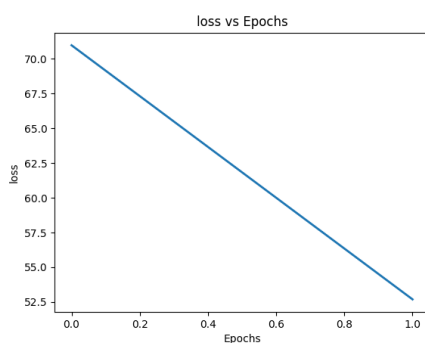
The size of the hidden layer is 100.

Our solution for a words appears in the train set and not in the dev set is to use lines in the embedding matrix for common suffix and prefix and also for unknown words, when word doesn't appear in the train we refer to its suffix / prefix and if it doesn't have common suf/pref we are relating to it as unknown word . We also added lines for padding sentences <s>(start of sentence) </s> (end of sentence)

Learning Rate : e^{-3}

Our model results :

POS : accuracy – 94.637 2 epoch



NER : accuracy – 75.742 6 epoch

