

Oren Cohen  
305164295  
DL ass2 part3

number of epochs = 6

Learning rate= 0.01

cross-entropy loss (as required)

additionally I add ADAM optimizer (worked better then SGD)

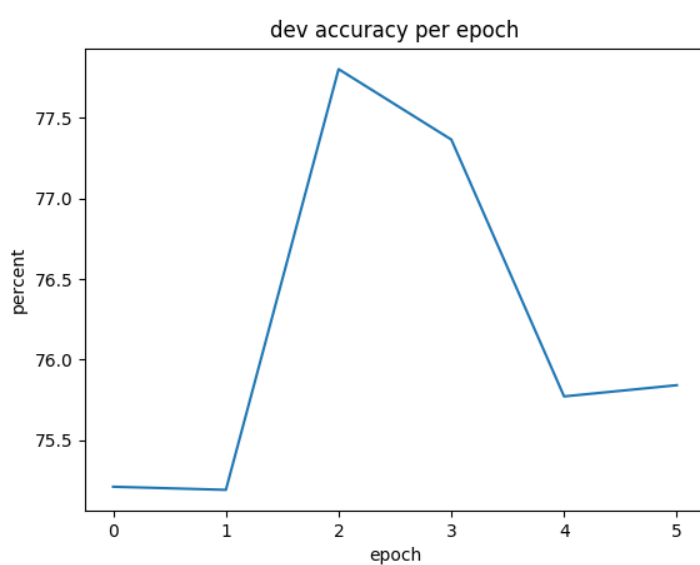
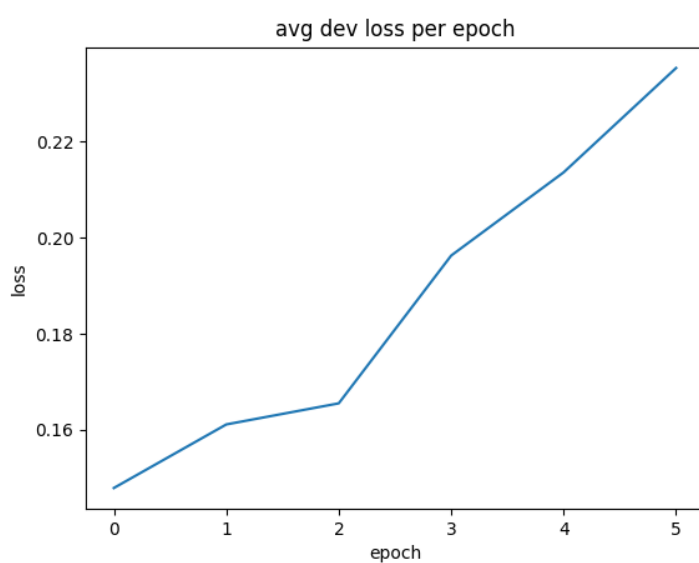
(in the training I also used batching, with size of 1024)

NER model:

after 6 epochs:

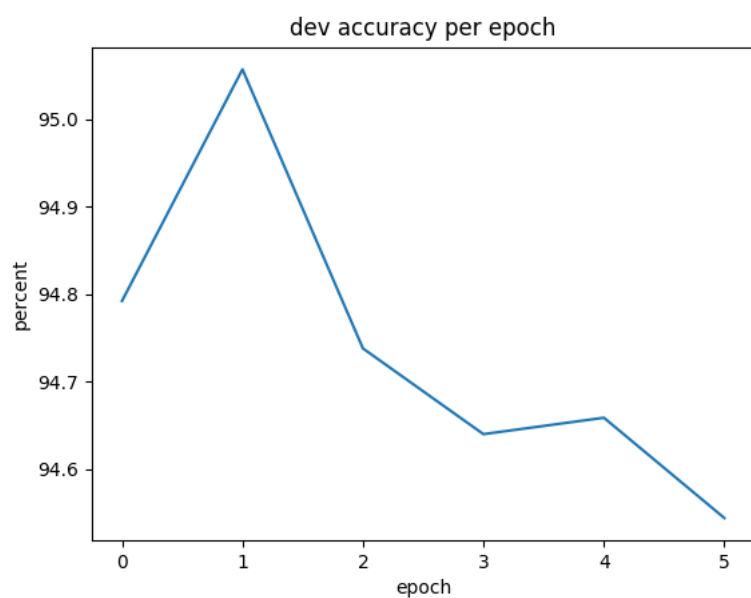
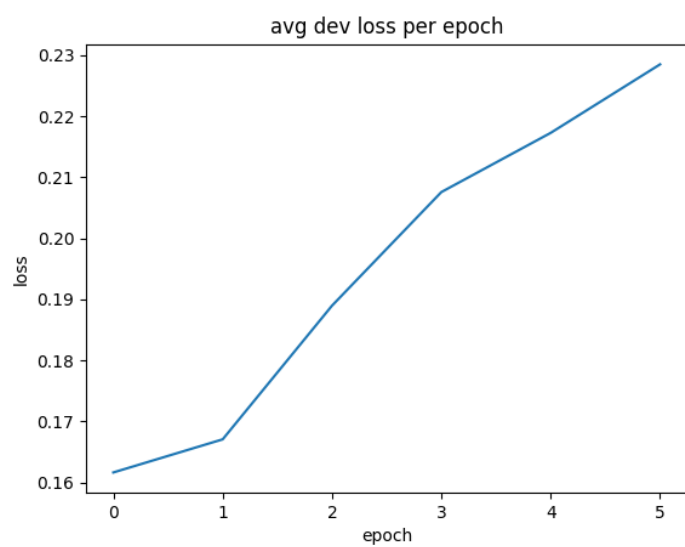
average loss: ~0.2

accuracy: ~77%



Oren Cohen  
305164295  
DL ass2 part3

POS model:  
after 6 epochs:  
average loss: ~0.2  
accuracy: ~94%



Oren Cohen  
305164295  
DL ass2 part3

**conclusion:**

the POS model was much more accurate!

Considerations:

as before – each word that does not appears in the dev/test set – marked as UNK.

Here we also changed the data input, to lowercase input, so the pre-trained vector will work with out data sets.

**Its seems that the pre-trained vector was a little bit helpful – but no more then that.  
The accuracy o NER and POS increase by 2-4 persents.**