Israel Cohen - 205812290 Hananel Hadad – 313369183

Part 1:

Considerations:

In train, we saved all labels seen, then in dev we skipped words with labels we did not see in train. It may cause some problem because now contexts are different without skipped words, but including those words cause other problems and serious headache.

For words in dev and test we did not see in train, we add to our vocabulary 'DEFAULT' and put it instead.

We also put words 'PAD_BEGIN' and 'PAD_END' for padding (two of each, at the edges of each sentence), hopping the model will come to utilize it.

The embedding is initialized uniformly, size of number_of_words X 50

The Neural network is composed of, after the embedding: dropout, tanh, hidden layer, and another dropout.

Hyper parameters:

hidden_layer_dim = 128

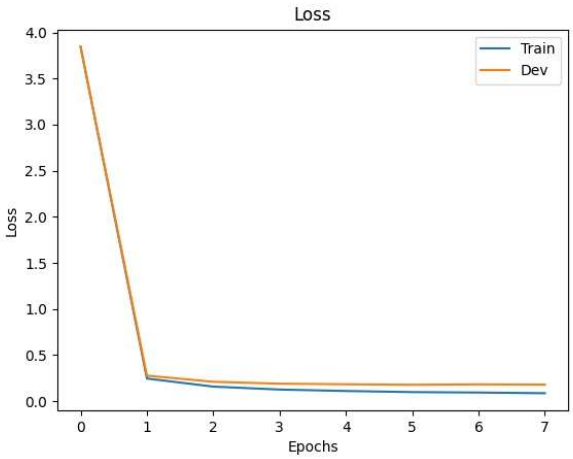
lr = 0.02

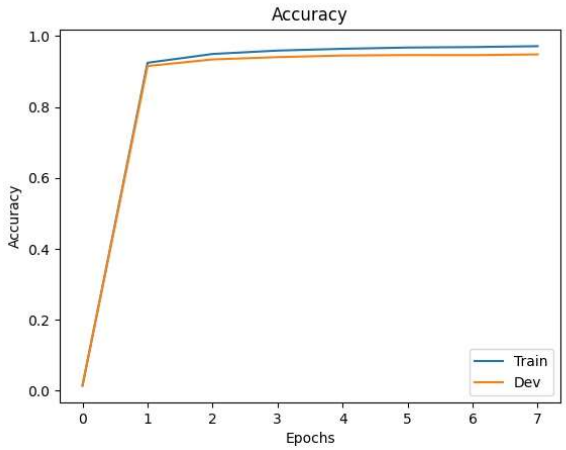
batch_size = 1000

epochs = 8

Result graphs:

Pos: 95% accuracy on dev





Ner: 96% accuracy on dev