Lea Setruk
Yoel Benabou

Deep Learning - 89687

Ass2 – Part 1

POS

We are using the PyTorch package to build our MLP model.

The parameters we use are:

- Adam Optimizer which gave us better results than SGD.
- One hidden layer, as required, with 200 neurons (we tried different number of neurons, and this number gave us the best results).
- Learning rate = 0.01
- Number of epochs = 25 (we stop when it converges).
- Batch size = 1024 (to be faster)
- Loss = Cross Entropy (as required)
- Activation function = tanh (as required)

Accuracy = 92%

NER

We are using the PyTorch package to build our MLP model.

The parameters we use are:

- Adam Optimizer.
- One hidden layer, as required, with 100 neurons.
- Learning rate = 0.01
- Number of epochs = 30 (we stop when it converges).
- Batch size = 1024
- Loss = Cross Entropy (as required)
- Activation function = tanh (as required)

Accuracy = 72.7%

Considerations

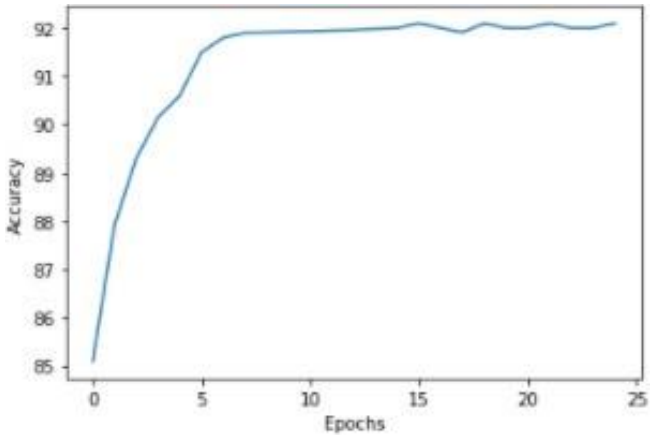
- If there is a word that appears in the train set and not in the dev set, we embed it by 'UUUNKKK'. This 'word' is added during the training.

- We want a window of 5 words. Which means that for the first word, we'll add before two '<s>', '<s>'. And same for the last word, at the end, we'll add to the vector '</s>', '</s>'.

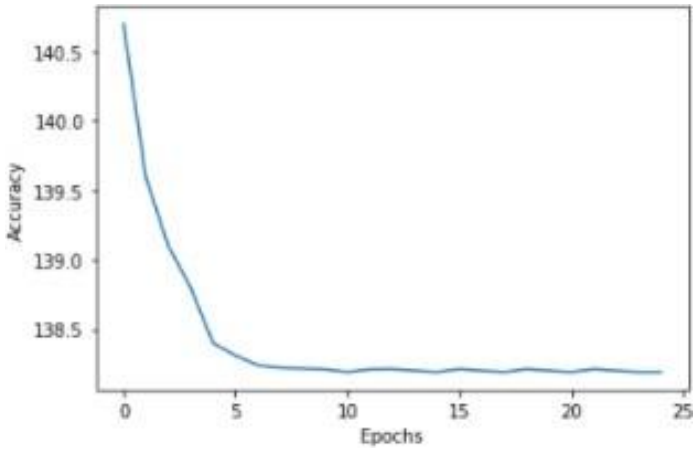
GRAPHS:

POS accuracy:

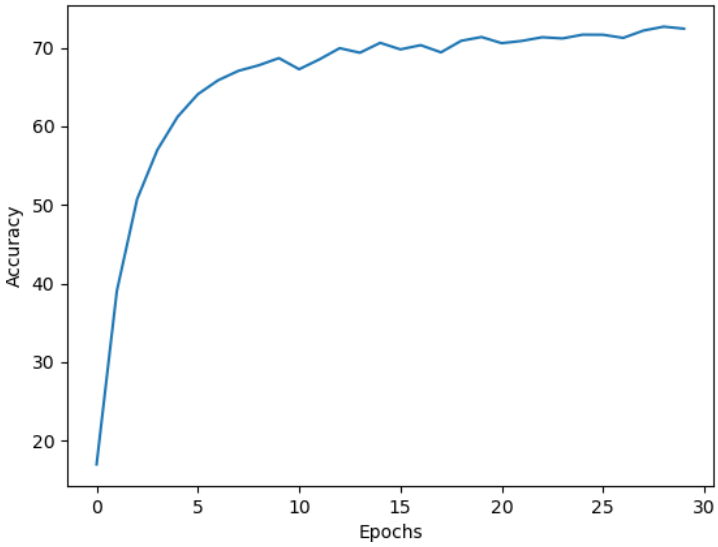
Lea Setruk
Yoel Benabou



POS loss:



NER accuracy:



NER loss:

Lea Setruk
Yoel Benabou