Lea Setruk
Yoel Benabou

Deep Learning - 89687

Ass2 – Part 4

POS

The parameters we use are:

- Adam Optimizer
- One hidden layer, as required, with 500 neurons
- 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 = 93%.

NER

The parameters we use are:

- Adam Optimizer.
- One hidden layer, as required, with 100 neurons.
- Learning rate = 0.01
- Number of epochs = 25(we stop when it converges).
- Batch size = 1024
- Loss = Cross Entropy (as required)
- Activation function = tanh (as required)

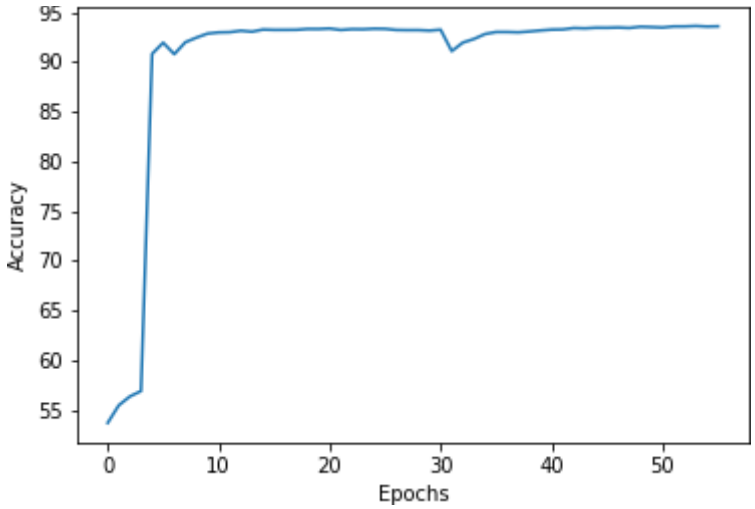
Accuracy = 75.5%

Using pre-trained word vectors gives us great results in accuracy of POS and NER. Our training produces word vectors according to what we want and what we need (the data, the model…). When we use sub-words, we also have more information. Good accuracy is being seen after a very small number of epochs. Common prefixes and suffixes give us more information.

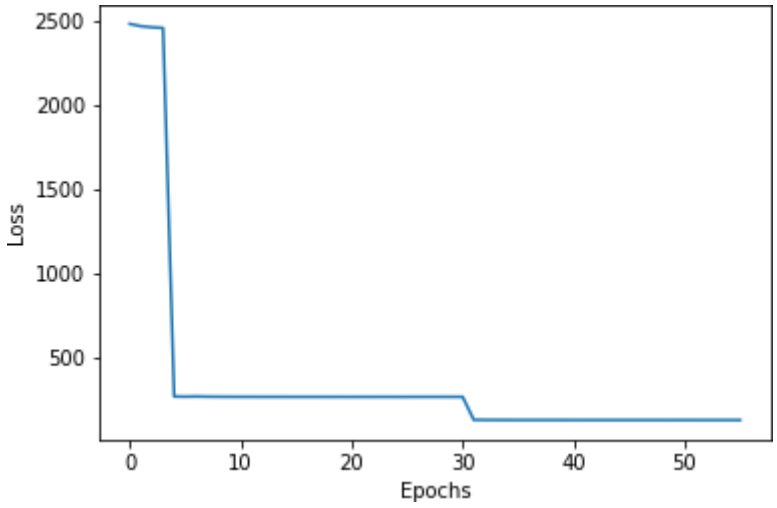
GRAPHS

POS Accuracy:

Lea Setruk
Yoel Benabou

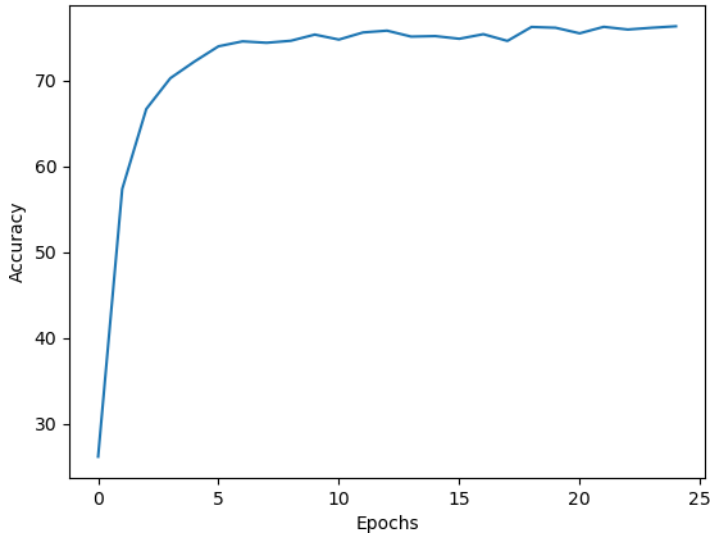


POS Loss:



NER Accuracy:

Lea Setruk
Yoel Benabou



NER Loss: