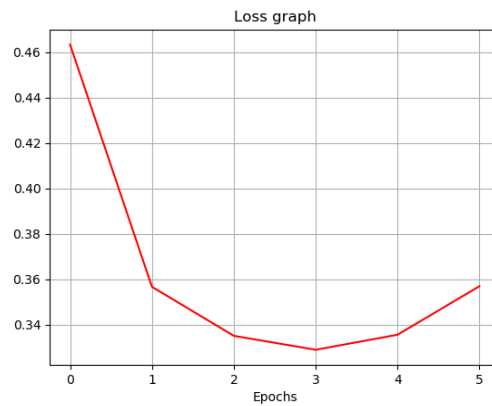
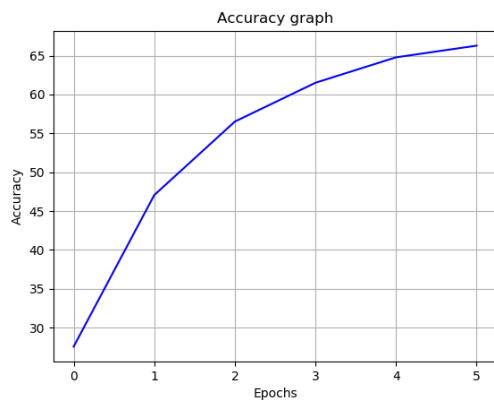


Part 4– Adding sub-word units

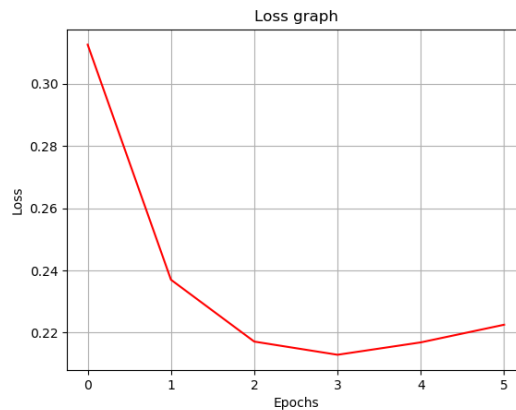
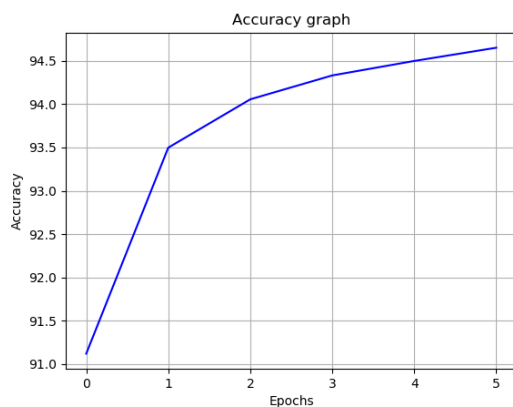
Ner parameters + pre-trained word embeddings + sub-word units:

- Learning rate: 0.001
- Epochs: 6
- Hidden dim: 110
- Batch size: 1024
- Accuracy: 66.354%
- Loss: 0.329



Pos parameters + pre-trained word embeddings + sub-word units:

- Learning rate: 0.001
- Epochs: 6
- Hidden dim: 70
- Batch size: 1024
- Accuracy: 94.650%
- Loss: 0.213

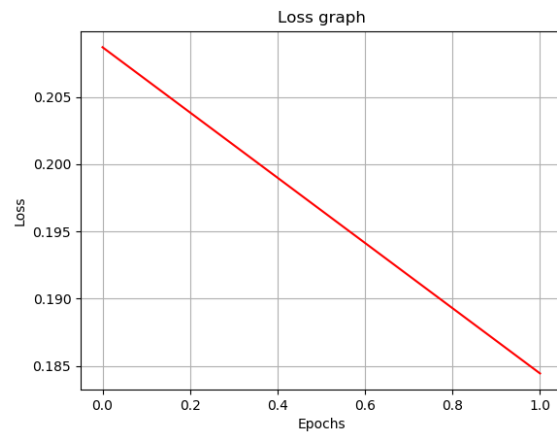
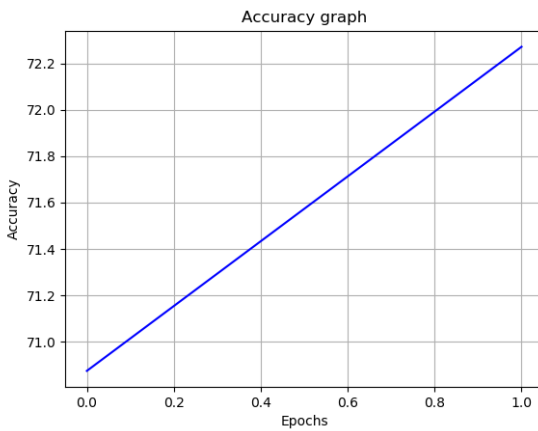


Assignment 2 – Window-based Tagging

Omer Zucker 200876548
Omer Wolf 307965988

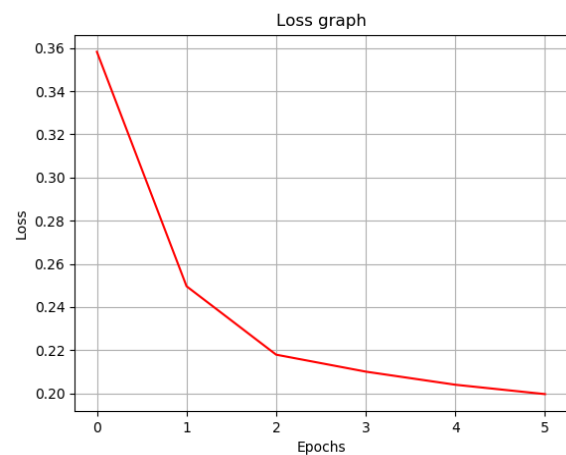
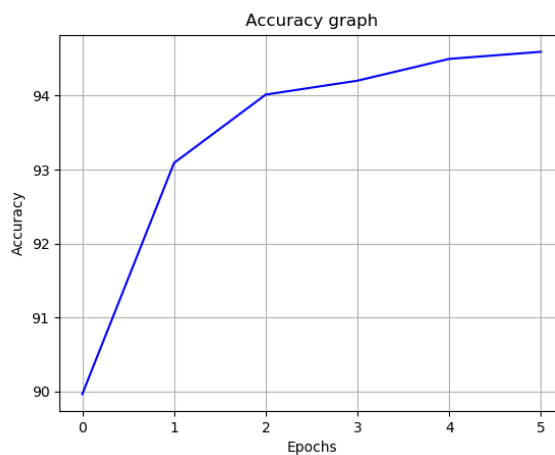
Ner parameters + no pre-trained word embeddings + sub-word units:

- Learning rate: 0.05
- Epochs: 2
- Hidden dim: 60
- Batch size: 1024
- Accuracy: 72.334%
- Loss: 0.184



Pos parameters + no pre-trained word embeddings + sub-word units:

- Learning rate: 0.001
- Epochs: 6
- Hidden dim: 60
- Batch size: 1024
- Accuracy: 94.594%
- Loss: 0.200



Analysis

Ner - We can observed that running the model with no pre-trained embedding returned higher accuracy than running it by adding the sub-words units, so the embeddings did not improve accuracy of the prediction process probably due to the low quantity of labels (the tags). We also saw that their contributions are complementary so we got changeable accuracy by running both pre-trained and adding the sub-words units together (it was changed from time to time). The most precise accuracy derived from task number 1.

Pos – Running the model by only adding sub-words units returned higher accuracy than running it with both pre-trained embeddings and sub-words unit, probably also here due to the quantity of the labels. Due to the fact there is much more tags so adding sub-words units improved the accuracy of the prediction process. We also saw that their contributions are complementary so we got higher accuracy by running both pre-trained and adding the sub-words units together. This, provided us the most precise accuracy result.