

# Part 1

- Name: Jonathan Zeltser
- ID: 204331375

## Considerations

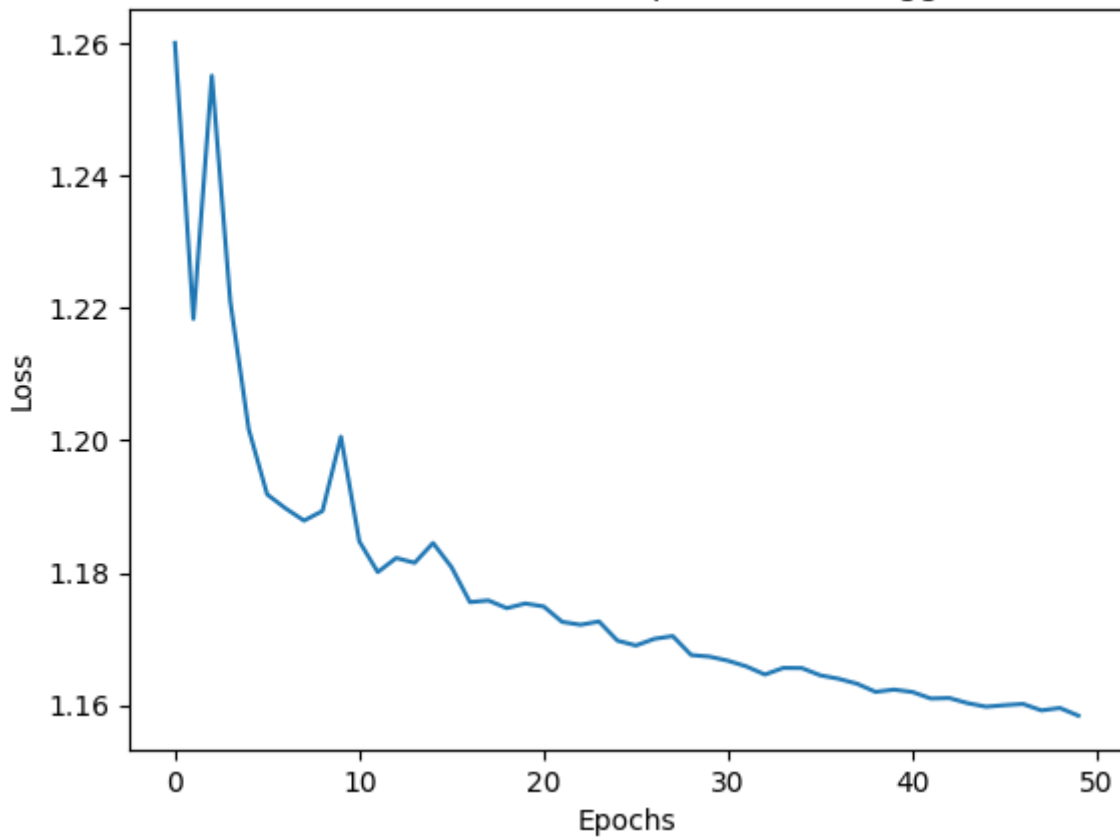
For both tasks, when a word appears in the dev dataset that is not in the train dataset, it the embedding of special `<UNK>` word, and was tagged with the real tag in the dataset. For the first words and last words, a `<PAD>` embedding was initialized and given a `PAD_TAG` .

## NER

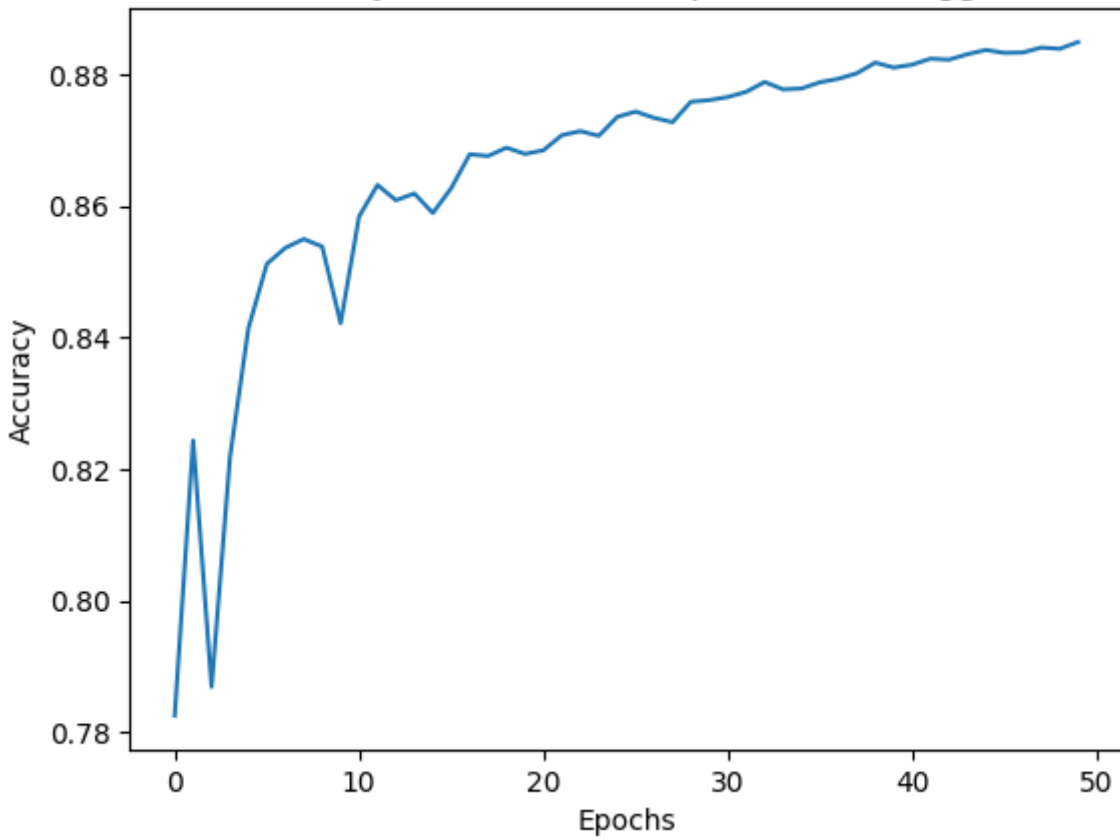
### Hyperparameters:

- Learning Rate: 0.01
- Hidden dimension size: 2000

Loss as a function of epochs - NER Tagger



Accuracy as a function of epochs - NER Tagger



**POS**

**Hyperparameters:**

- Learning rate: 0.03
- Hidden dimensions size: 2500

