Assignment 2 – Window-based Tagging

Daniel Bazar 314708181

Peleg shefi 316523638

Part 1

Architecture

- Both tasks implementing the same Network: a MLP with one hidden layer and a tanh activation function
- The network trained with a cross-entropy loss
- We Experimented with several network configurations and chose the best configuration based on the DEV accuracy

Best parameters

- NER:
  - Hidden layer size: 170
  - Dropout probability: 0.4
  - Batch size: 2048
  - Optimizer: Adam (Learning rate: 1e-3)
- POS:
  - Hidden layer size: 170
  - Dropout probability: 0.4
  - Batch size: 2048
  - Optimizer: Adam (Learning rate: 1e-3)

Considerations

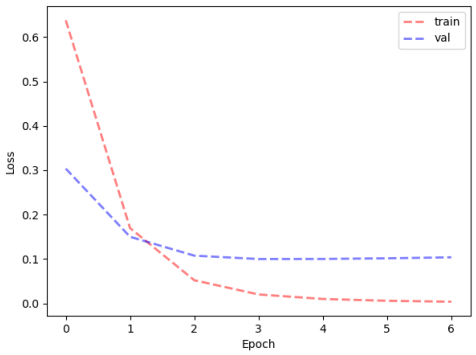
- We handle words that appear in the train set and not in the dev set by assigning them the UNK token
- We padded the sentences with SOS (start of string) and EOS (end of string) at the beginning and end of the sentence

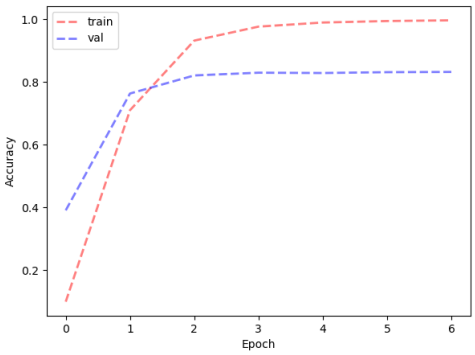
Results

- NER:
  - Loss validation: 0.104
  - Accuracy validation: 83.18%
- POS:
  - Loss validation: 0.132
  - Accuracy: 95.8%

Graphs

NER





POS