

# DL for NLP - Ass. 3 - bilstm

Itamar Trainin, Guy Gispan

January 2020

## Parameters

Number of BiLSTM layers: 2  
Prefix/Suffix Embedding Size: 10  
Char Embedding Size: 10  
Hidden Layer: 20  
Optimization Methods: SGD  
Dropout: None  
Batch Size:  
    a, c: 8  
    b, d: 1  
Learning Rate:  
    a, c: 1  
    b, d: 0.1  
Number of Epochs:      a, c: 15  
    b, d: 20

## Graphs

Figure 1,2 show the results for POS and NER. Note that figure 3 shows the results on the NER task where the O O pairs are not dis-regarded from the loss.

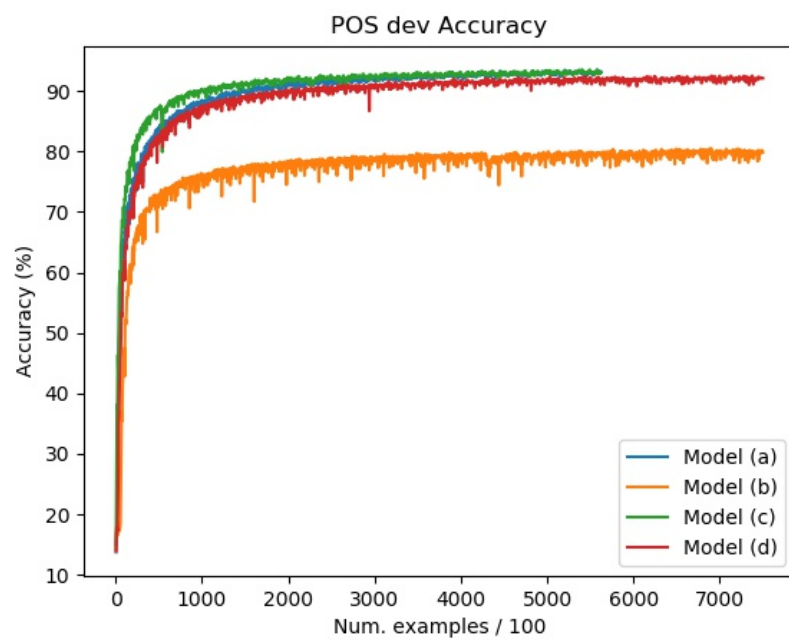


Figure 1: POS accuracy on dev

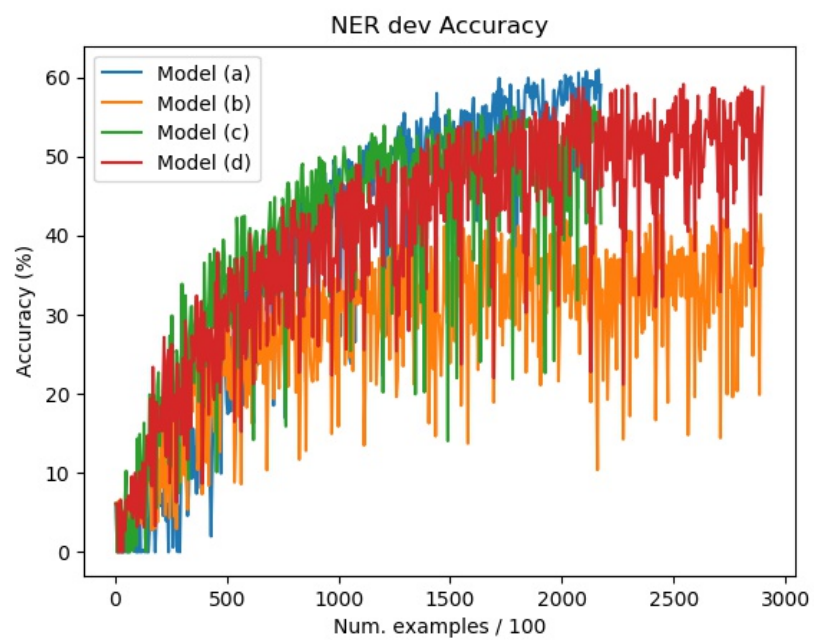


Figure 2: NER accuracy on dev

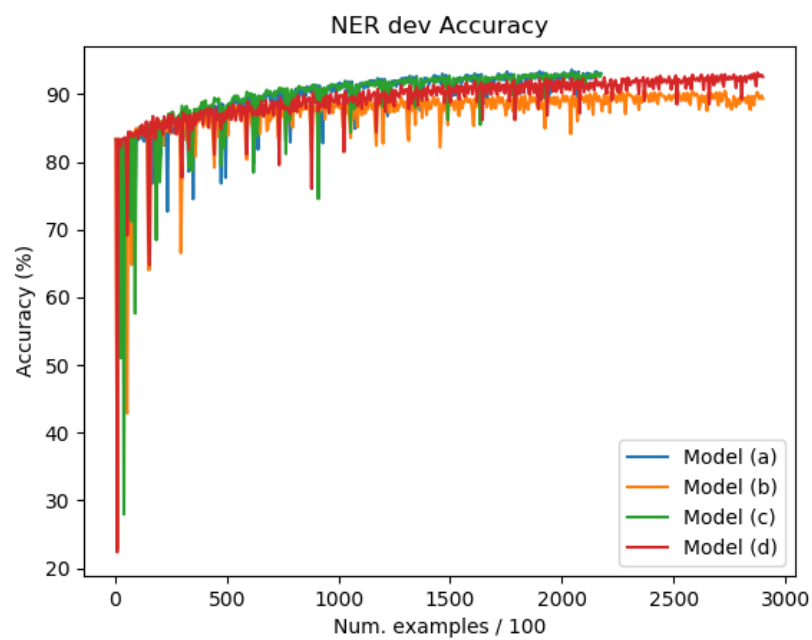


Figure 3: NER accuracy on dev