

Assignment 3 - RNN Acceptors and BiRNN Transducers

Itay Etelis, 209041474,
Matan Avitan, 318287455

March 2023

1 Part 3: Transducer Capabilities

In the third part we use the Bi-LSTM as a transducer to solve two tasks, the first is the Part Of Speech (POS) problem, the second is the Named Entity Recognition (NER) problem.

In this task we were limited to only five epochs with each approach (a, b, c, d). This resulted us trying several sets of hyper-parameters to find the most appropriate once for the mission.

1.1 Model Architecture

The Bi-LSTM model used is as described in the assignment itself, the hyperparameters used are summarized in Table 4.

Parameter	Value
Batch Size	32
Learning Rate	0.001
Embedding Dimension	10
LSTM Hidden Units	64
Fully Connected Hidden Units	32
Training Epochs	5

Table 1: Model parameters for approach a

Parameter	Value
Batch Size	32
Learning Rate	0.001
Embedding Dimension	10
LSTM Hidden Units	64
Fully Connected Hidden Units	32
Training Epochs	5

Table 2: Model parameters for approach b

Parameter	Value
Batch Size	32
Learning Rate	0.001
Embedding Dimension	10
LSTM Hidden Units	64
Fully Connected Hidden Units	32
Training Epochs	5

Table 3: Model parameters for approach c

Parameter	Value
Batch Size	32
Learning Rate	0.001
Embedding Dimension	10
LSTM Hidden Units	64
Fully Connected Hidden Units	32
Training Epochs	5

Table 4: Model parameters for approach d

1.2 Training Results

Our models performs well on both tasks, this is a strong indication ...

Problem	Training Acc
a approach	92.96%
b approach	93.03%
c approach	91.04%
d approach	92.70%

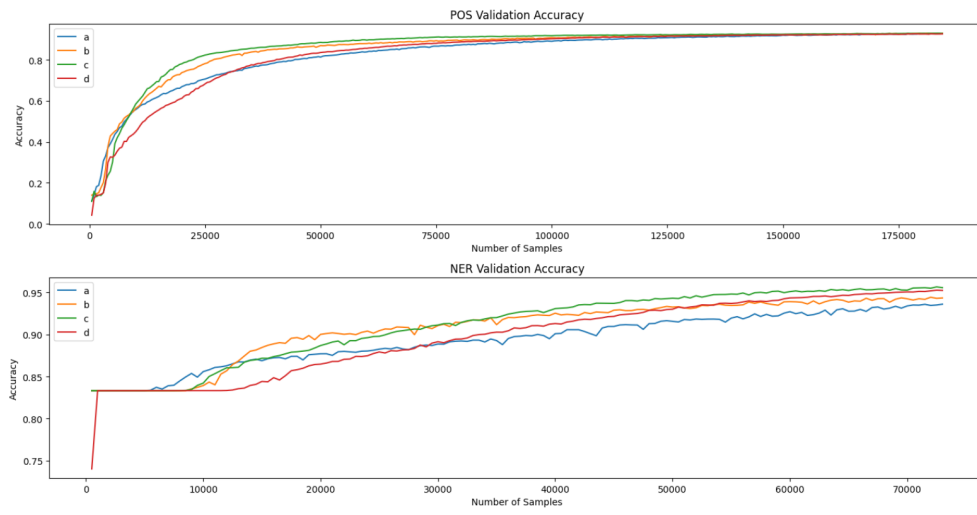
Table 5: Training results for Part 3 - Part of Speech

Problem	Training Acc
a approach	93.61%
b approach	94.43%
c approach	95.66%
d approach	92.27%

Table 6: Training results for Part 3 - Named Entity Recognition

1.3 Discussion

These results underscore the challenges associated with applying LSTM models to context-sensitive languages. Such languages require more than simple counting abilities, pushing the limits of what LSTMs can achieve.



POS				
	a	b	c	d
Max Accuracy	0.9296	0.9303	0.9304	0.927

NER				
	a	b	c	d
Max Accuracy	0.9361	0.9443	0.9566	0.9527

Figure 1: Training Testing with a,b,c,d approaches