Part 3: BILSTM

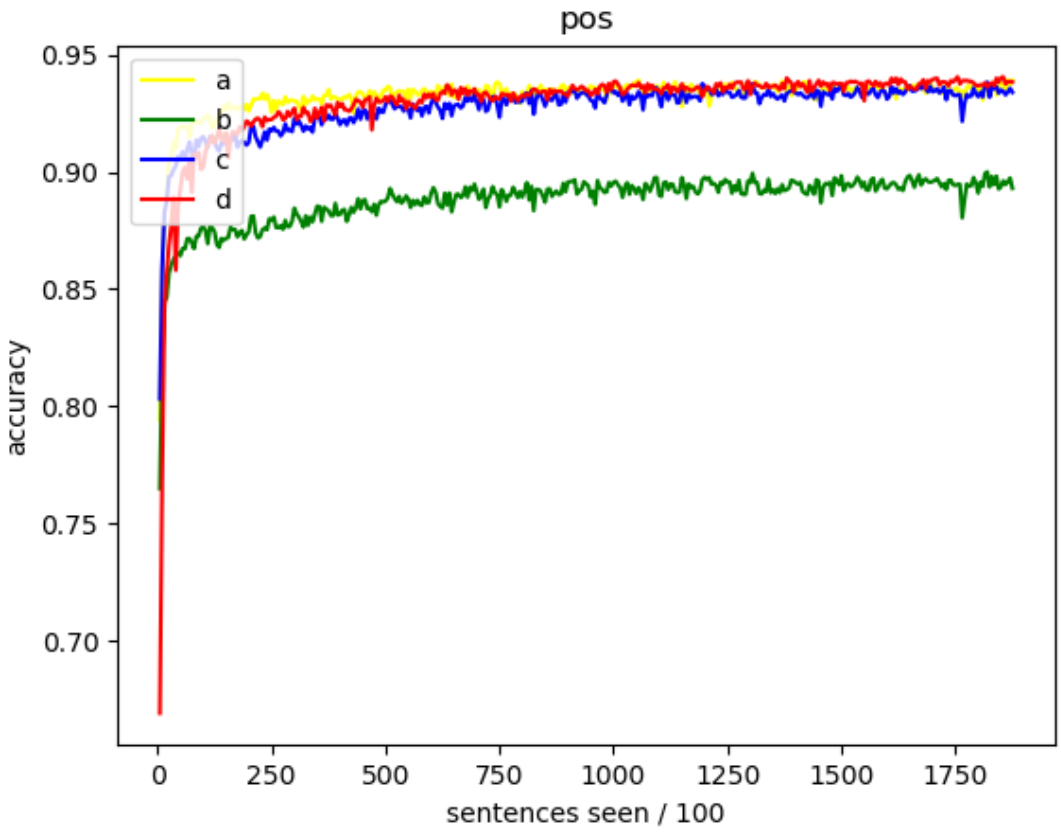
Lior Shimon 341348498
Lev Levin 342480456

For POS tagging:

type	*epochs*	*Output size of bilstm and lstm*	*eta*	*Batch size*	*optimizer*	*Initialization of embedding*	*dropout*	*regularization*
a	5	BILSTM: 100 LSTM: no lstm in this model	0.01	BLSTM: batch size of 1 Linear layer after BLSTM: number of words in sentence	Adam	1.) Uniform distribution between [-1,1] 2.) Length of embedding vector: 100	no	L2: 1e^5
b	*5*	BILSTM: 100 LSTM: 50		For character LSTM: number of words in sentence BLSTM: batch size of 1. Linear layer after BLSTM: number of words in sentence	Adam	1.) Uniform distribution between [-1,1] 2.) Length of embedding vector: 20	0.4 in bilstm	L2: 1e^5
c	*5*	BILSTM: 150 LSTM: no lstm in this model	0.001	BLSTM: batch size of 1 Linear layer after BLSTM: number of words in sentence	Adam	1.) Uniform distribution between [-1,1] 2.) Length of embedding vector: 100	no	no.
d	*5*	BILSTM: 100 LSTM: 100	0.001	For character LSTM: number of words in sentence	Adam	1.) Uniform distribution between [-1,1] 2.) Length of embedding vector: 100	0.3 at bilstm	L2: 1e^5

				Linear layer after concatenation: number of words in sentence BLSTM: batch size of 1. Linear layer after BLSTM: number of words in sentence		3.)without applying lowercasing to words.		

Graph



For NER tagging:

type	*epochs*	*Output size of bilstm and lstm*	*eta*	*Batch size*	*optimizer*	*Initialization of embedding*	*dropout*	*regularization*
a	5	BILSTM: 42 LSTM: no ltm in this model	0.001	BLSTM: batch size of 1 Linear layer after BLSTM: number of words in sentence	Adam	1.) Uniform distribution between [-1,1] 2.) Length of embedding vector: 30	no	L2: no
b	*5*	BILSTM: 100 LSTM: 100		For character LSTM: number of words in sentence BLSTM: batch size of 1. Linear layer after BLSTM: number of words in sentence	Adam	1.) Uniform distribution between [-1,1] 2.) Length of embedding vector: 10	0.4 in bilstm and 0.1 in listm	L2: 1e-4.
c	*5*	BILSTM: 60 LSTM: no lstm in this model	0.001	BLSTM: batch size of 1 Linear layer after BLSTM: number of words in sentence	Adam	1.) Uniform distribution between [-1,1] 2.) Length of embedding vector: 60	no.	1e-3.
d	*5*	BILSTM: 100 LSTM: no lstm in this model 60	0.002	For character LSTM: number of words in sentence Linear layer after concatenation:	Adam	1.) Uniform distribution between [-1,1] 2.) Length of embedding vector: 60	Dropout 0.3 in bilstm.	L2: 1e^5

				number of words in sentence <u>BLSTM:</u> batch size of 1. <u>Linear layer after BLSTM:</u> number of words in sentence				

Graph