

CS6957 NLP with Neural Networks
Mini Project 2

R. Rishanth (u1419542)

1.
Done

2.

Embedding	Mean		Concatenate	
	UAS	LAS	UAS	LAS
Glove 6B 50d	0.34	0.29	0.70	0.64
Glove 6B 300d	0.41	0.36	0.71	0.65
Glove 42B 300d	0.43	0.37	0.72	0.66
Glove 840B 300d	0.44	0.39	0.74	0.68

3.
Concatenating vectors works better which makes intuitive sense as it avoids the loss of information on averaging.
Larger Glove embeddings lead to small yet consistent improvements in performance across strategies.

4.
(a) Mary had a little lamb . (POS tags: PROPN AUX DET ADJ NOUN PUNCT)
Actions:
SHIFT SHIFT SHIFT SHIFT SHIFT REDUCE_L_amod REDUCE_L_det REDUCE_L_aux
REDUCE_R_nmod SHIFT REDUCE_R_punct

(b) I ate the fish raw . (POS tags: PRON VERB DET NOUN ADJ PUNCT)
Actions:
SHIFT SHIFT REDUCE_L_nsubj SHIFT SHIFT REDUCE_L_det SHIFT REDUCE_R_conj
REDUCE_R_obj SHIFT REDUCE_R_punct

(c) With neural networks , I love solving problems . (POS tags: ADP ADJ NOUN PUNCT PRON
VERB VERB NOUN PUNCT)
Actions:
SHIFT SHIFT SHIFT REDUCE_L_amod REDUCE_L_case SHIFT REDUCE_R_punct SHIFT
SHIFT REDUCE_L_nsubj REDUCE_L_nsubj SHIFT SHIFT REDUCE_R_obj REDUCE_R_advcl
SHIFT REDUCE_R_punct

5.

Chen and Manning include the labels too in the input (as in the Extra credit setup) besides words in the stack and buffer and their POS tags.

The words they consider include not just top-k words on the stack and buffer but also the first and second leftmost and rightmost children of the top-2 words on the stack and the leftmost and rightmost children of the rightmost children of the top 2 words of the stack.

One of the main points they make is that such higher-order features are very important for the task of dependency parsing.

Not just the parse state but their non-linearity is also different. They use the cube function as the activation function instead of ReLU as they hypothesize that the cube function would help capture relationship between every pair of input features (words, POS, labels) and also of the triple as a whole.

Extra Credit

All results are on test set

Embedding	Mean		Concatenate	
	UAS	LAS	UAS	LAS
Glove 6B 50d	0.25	0.15	0.81	0.79
Glove 6B 300d	0.29	0.21	0.82	0.80
Glove 42B 300d	0.31	0.23	0.83	0.80
Glove 840B 300d	0.32	0.24	0.84	0.82

It is interesting to note that including labels deteriorates performance when using mean strategy and improves performance when using concatenate strategy compared to when not including them.
