Dependency Parser on StanfordNLP model

Yuan-Chia Cheng

May 2019

1 The Dependency Parser Model of StanfordNLP

We apply dependency parser model of stanfordNLP to go through our task. Furthermore, we append extra neural network feature to enhance the performance which will be expanded on later.

1.1 Standard Model Structure

StanfordNLP model (2018) is contributed to extract information from word embedding, character embedding, POS embeddingm, lemma embedding and summed UFeats embedding, and then pass these extracted vectors through a highway LSTM to have a set of scores on each head candidates per word. What's more, following the output of highway LSTM, they also take the distance of head candidates and linear order into consideration, that is the length of arcs and the direction of the arcs, by mainly applying Baye's rule and Deep-Biaffine Model to place some constraints on the distances and linear order.

1.2 Modification on Standard Model - Sentence Domain

To fuse cross domain information into the model, we apply a simple BiLSTM to the input sequence of words, considering the output vector of the BiLSTM as the domain information of the sentence, and then put it into the original highway LSTM along with the other original vectors such as word embedding. By so, we hope the model can be more sensentive to which domain the input sentence comes from.

1.3 Experiments and Results

UAS(F1)	Description
~ 0.77	Standard Stanford model w/o Char. Emb. w/ S.D.
~ 0.76	Standard Stanford model w/o Char. Emb.
~ 0.74	Standard Stanford model w/ S.D
~ 0.73	Standard Stanford model
~ 0.70	Standard Stanford model w/o distance constraint

* S.D denotes our Sentence Domain feature
* Char. Emb. denotes Character Embedding

1.4 Other sources we adpat

We adapt pretrained word embedding dataset to train our model. FastText supported by Facebook is adpated this time which is a 300-dim word embedding dataset.

References

[1] Timothy Dozat, Peng Qi, and Christopher D. Manning. Stanford's graph-based neural dependency parser at the CoNLL 2017 shared task. In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 20–30, Vancouver, Canada, August 2017. Association for Computational Linguistics.

[2] Peng Qi, Timothy Dozat, Yuhao Zhang, and Christopher D. Manning. Universal dependency parsing from scratch. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 160–170, Brussels, Belgium, October 2018. Association for Computational Linguistics.