Paper: *A Fast and Accurate Dependency Parser using Neural Networks*

Mushfika Rahman

Date: 2/4/2022

Quote *...our model is able to automatically identify the most useful information for predictions, instead of hand-crafting them as indicator features.* pp.12.

Overview The authors in this paper are trying to solve the problems with the existing dependency parser. Though the dependency parsers have enormous success, they have several problems. Firstly, the parsers classify on sparse indicators. This is why they use poorly estimated feature weights. Secondly, parsers depend on the manually designed set of feature templates; these are not generally complete and require knowledge. Lastly, the authors indicate that the feature extraction process is very time-consuming. The authors of this paper state that using dense features instead of sparse indicator features can be efficient in terms of computation and accuracy. Sparsity can be reduced by giving statistical information between similar words. In this way, it can contribute a good starting point for composing features of words. Further, the authors have indicated two challenges of encoding available information and presenting higher-dimensional features on dense representation.

This paper presents a unique way of training a neural network classifier by using a greedy, transitional-based dependency parser. This paper utilizes a neural network that trains and uses 200 dense learned features. Thus it is fast and showed better accuracy in both UAS and LAS on English and Chinese datasets. They have further introduced a new activation function that can capture higher dimension interaction features.

Intellectual Merit The researchers at Stanford University explored the work of distributed word representation. Their neural network architecture is comprised of the existing word embeddings. The mapping between the input and hidden layers consists of a cube activation function, and a softmax layer is added on top of the hidden layer[Chen and Manning (2014)]. They have shown a comparison between existing parsers and their parser. Further, they have stated comparison between cube activation function and other functions. According to their experiment, their parser has performed significantly better accuracy and speed. They have also discussed different components that affect their parser, which makes their work rationale. Their work has used several existing tools and datasets such as Stanford Dependency Conversion Tool, Stanford Basic Dependency, and so on. This work has an impact on downstream NLP tasks.

Broader Impact The research has not only introduced a novel neural-based architecture but also demonstrated different model analysis. This has a further impact on various NLP-related tasks. Their work has also indicated how previous word2vec techniques have influenced new techniques, and this work has further influenced to make new dependency parsers. They have clearly stated all the parameters and hyper-parameters used in the experiment. The research involved one Ph.D. student and a professor from Stanford University.

Keywords Computing methodologies, Natural language processing, Deep learning, Computational Linguistics

Discussion Questions

- One interesting observation is that the paper discusses initialization of pre-trained word embeddings was useful in chinese language than English.
- Another surprising observation is that how the model was able to achieve comparable accuracy without the help of pre-trained word embeddings [Chen and Manning (2014)].

Table 1: Grade deductions by section

Overview	Intellectual M.	B. Impact	Keywords	Questions	Is Online?

References

D. Chen and C. D. Manning. A fast and accurate dependency parser using neural networks. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 740–750, 2014.