Review on Infusing Fine Tuning with Semantic Dependencies

Zhaofeng Wu, Hao Peng, Noah A. Smith

Previous models were application specific which pre-trained on large corpus but it is still unknown whether they actually capture the meaning. Some state-of-the-art models built for end-to-end tasks such as Sentiment Analysis, Question Answering are trained without understanding what relevant features are used. Therefore authors apply probing techniques to these language models by understanding semantic dependency through predicate-argument substructure. Many works have been done on syntactic parsing but not much on semantic parsing. But unlike syntax, semantics requires understanding of concepts and is very task-specific. Therefore Authors use convolutional graph encoders to encode semantic parsing into fine-tuning tasks increasing accuracy in GLUE benchmark. Authors through their experiments show that pretraining BERT highlights linguistic patterns but not FrameNet and neither does Roberta capture information about predicate-argument structure which Wu et al wants to rectify.

Authors propose Semantics-Infused Fine Tuning (SIFT) in which the input was passed through semantic dependency parser specifically DELPH-IN MRS (DM) which is combined with the pretrained model with a relational graph convolutional network (RGCN). To probe RoBERTa, authors after training the model, used a linear probe classifier to train a supervised probe model by freezing the pretrained representations and comparing the held-out data.

I liked that the authors wanted to understand what these enormous models are learning, performed ablation study and did not focus on achieving sota results. . One of the drawbacks of the paper is that it is not at all easy to read. One has to look at many complex terms which could be explained briefly. I failed to understand why the authors chose to understand only predicate-argument or why they used convolutional networks. It would be better if they explained the reason to pick these specifics. In the future, work has a lot of potential to extend to other types of semantic dependencies which can be explored so as to better design the models.

In conclusion, authors proposed SIFT which incorporated semantic dependency in RoBERTa using predicate-argument substructure and relational graph convolutional network.