

Learning Natural Language Inference using Bidirectional LSTM model and Inner-Attention

1. Data Used

At 570K pairs, SNLI is two orders of magnitude larger than all other resources of its type. The dataset is constructed by crowdsourced efforts, each sentence written by humans. The target labels comprise three classes: Entailment, Contradiction, and Neutral

2. Methodology

In this paper, the author proposed a sentence encoding-based model for recognizing text entailment. In the author's approach, the encoding of sentence is a two-stage process. Firstly, average pooling was used over word-level bidirectional LSTM (biLSTM) to generate a first-stage sentence representation. Secondly, attention mechanism was employed to replace average pooling on the same sentence for better representations. Instead of using target sentence to attend words in source sentence, the author utilized the sentence's first-stage representation to attend words appeared in itself, which is called "Inner-Attention".

3. Model Used

The author proposed a unified deep learning framework for recognizing textual entailment which does not require any feature engineering, or external resources. The basic model is based on building bi-LSTM models on both premises and hypothesis. *The basic mean pooling encoder can roughly form a intuition about what this sentence is talking about.* Obtained this representation, the author extended this model by utilizing an Inner-Attention mechanism on both sides. This mechanism helps generate more accurate and focused sentence representations for classification. In addition, the author introduced a simple effective input strategy that get ride of same words in hypothesis and premise, which further boosts our performance. Without parameter tuning, the author improved the art-of-the-state performance of sentence encoding based model by nearly 2%.

4. Key Conclusions

In this paper, the author proposed a bidirectional LSTM-based model with Inner-Attention to solve the RTE problem. The augthor come up with an idea to utilize attention mechanism within sentence which can teach itself to attend words without the information from another one. The Inner-Attention mechanism helps produce more accurate sentence representations through attention vectors. In addition, the simple effective diversing input strategy introduced by the author's further boosts our results. And this model can be easily adapted to other sentence-matching models.