Review of paper – Divide and Conquer: Text Semantic Matching with Disentangled Keywords and Intents

Xuan Chen - Student #82286956
Paper link: https://arxiv.org/abs/2203.02898

Summary of the paper (3-4 sentences)

A common way for most state-of-the-art matching models to implement text semantic matching, like BERT, is to directly perform text comparison by processing each word uniformly. The paper proposes a training strategy for text semantic matching in a divide-and-conquer manner by disentangling keywords from intents, which is keyword matching and intent matching. Where keywords refer to factual information reflecting entities or actions to be strictly matched; the other elements constitute abstract intentions, which can often be expressed in a variety of ways to convey the same concept or idea.

Main contributions (2-3 bullet points)

1. We introduce a novel training regime for text matching, which disentangles keywords from intents based on different levels of matching granularity in a divide-and-conquer manner.
2. The proposed approach is simple yet effective, which can be easily combined with PLMs plus few auxiliary training parameters while not changing their original inference efficiency.
3. Experimental results on three benchmarks across two languages demonstrate the effectiveness of our approach in different aspects.

Positive and negative points (2-3 points each)

1. Positive:
   - By separating keywords from intentions, the matching process can be divided into two easier sub–problems, and they require different levels of matching granularity, different from using the same fine–tuning paradigm and treat text sequences in a uniform manner like most models.
   - With the labels of keywords and intent can model learn how to decompose keywords and intentions, and zoom out the distance of them in PLM, which allows modeling at different levels of matching.
2. Negative: The improvement of accuracy on QQP and MRPC datasets with using DC-Match training strategy is not significant.

Unclear (2-3 points)

1. The paper directly infers the matching category for a sentence pair based on the conditional probability of the original problems, why is annotation information of keywords and intents not available at the inference time?
2. How can the strategy guarantee to keep similar distribution of original issues and solution of subproblems?