

Sentence-BERT主要是解决Bert语义相似度检索的巨大时间开销和其句子表征不适用于非监督任务如聚类，句子相似度计算等而提出的。Sentence-BERT使用鉴孪生网络结构，获取句子对的向量表示，然后进行相似度模型的预训练即为sentence-BERT。其预训练过程主要包括如下步骤：

（1）孪生网络获取句向量表示

使用孪生网络结构，使用Bert进行finetune来进行句子的语义相似度模型的预训练，其具体做法是：将句子对输入到参数共享的两个bert模型中，将Bert输出句子的所有字向量进行平均池化（既是在句子长度这个维度上对所有字向量求均值）获取到每个句子的句向量表示。

（2）分类器特征拼接

然后将两向量的元素级的差值向量与这两个句向量进行拼接，最后接softmax分类器来训练句子对分类任务，训练完成后就得到了sentence-Bert语义相似度预训练模型。

这样做的目的是：减小Bert语义检索的巨大时间开销，并使其适用于句子相似度计算，文本聚类等非监督任务。

实验结果也正是如此，对于同样的10000个句子，我们想要找出最相似的句子对，只需要计算10000次，需要大约5秒就可计算完全。从65小时到5秒钟，检索速度天壤之别。