

在query理解中能ALL IN BERT吗？

原创 WePlayData WePlayData 2020-04-04

BERT模型在多个NLP任务中取得不错的效果，其在预训练的embedding基础上，利用finetune机制，上面对接不同类型的NLP任务，以分类、匹配、序列标注、回归为主。考虑到BERT的耗时，往往采用更简单的模型（LSTM / CNN）通过蒸馏的方式进行上线。

越来越多的nlper在工作中采用BERT方法，那对于query理解任务，是不是都可以用BERT重构一下呢？因为query理解的模块大都可以归类到finetune中的几种类型的任务。

序列标注：分词，实体识别，phrase识别...

分类：非必留，同义替换，意图...

回归：词权重，紧密度...

看上去是个不错的想法，以不变应万变，集中力量的打造精品的预训练embedding，然后再结合不同任务各个击破。笔者会有个疑问，倘若这样的一个流程真的如此有效，那query理解的门槛在哪里？是不是从事query理解的nlper都要失业了？不可否认bert的效果在各种各样的任务上取得了不错的效果，但all in bert是不是太绝对了？如果真的万物皆可bert，那什么才算是nlper的核心竞争力？别的团队，只要能爬数据，只要能买gpu，是不是也可以在相应任务上做出同等效果呢？可能我们会有更多数据，我们能做更好的预训练，结果往往事与愿违，可能真的能在评价指标上提升几个点，但是这些提升是否有意义，是过拟合还是真的学到了一些新的泛化能力？可能汇报指标挺好看，但反而会跟领导一种没有结合任务场景去思考的印象。

BERT在某些任务上确实比其他方法更合适，但对于All in BERT的做法不敢苟同。说BERT作为大多任务的冷启动方法或者互补方法是认同的。笔者认为query理解是个复杂的任务，需要结合任务特点+垂类特点（医疗、音乐、视频...）来做。并且实际的在线系统往往不是单一的模型：1) 准确的词典/模板覆盖头部query；2) 轻量模型来解决中长尾的query；3) 复杂模型或前沿模型更多是方法上的一个探索。

相关阅读

1. Query理解 - 搜索引擎“更懂你”
2. 从搜一搜中检“相关性排序”的排序结果说起...
3. 搜索排序 = 相关性排序？
4. 搜索引擎新的战场 - 百度、头条、微信
5. 搜索引擎的两大问题（1） - 召回
6. 搜索引擎的两大问题（2） - 相关性