The False Promise of Imitating Properties of LLMs Summary

沙之洲 2020012408

本篇工作的核心论点是，认为通过 Imitation Learning 的捷径，实现对 LLM 能力的复制是不可行的，一个更正确的方法是增加 base LM 的 capability，可以通过更好的模型结构，更大的模型参数和更好的预训练数据集和预训练方法。

本文的优点在于，证明了 Imitation Learning 得到的模型，在除了 Imitation dataset 之外的任务上表现的非常差。具体来说，作者用了 ShareMix 这个 Broad Coverage 的 Imitation dataset finetune base model，甚至导致了模型在 NQ 数据集上的能力下降。这个实验结果很好地证明了作者上述的观点。除此之外，本文还证明了人类评价员已经不能很好地作为一个评价 imitation model 的指标了。因为 imitation model 在模仿的过程中会学习 LLM 的语言风格。而人类评价员往往会被这种外表华丽的语言风格所迷惑，给 imitation model 打很高的分，但是实际上 imitation model 输出的内容有很多事实性错误，而这一部分被人类所忽略了。

这篇工作的不足之处在于，它只证明了 Broad Coverage Imitation dataset 会使得模型的能力下降，但是本质上想让一个小模型学会 LLM 的所有能力显然是不可能的事情，而过于 diverse 的 imitation dataset 的分布根本不是模型能够拟合的，所以会导致能力的下降。但是一种更加可行的方案是，让小模型学会 LLM 的少数几个能力，比如说两个特定的能力，那么直观理解小模型在能力上是具备的。而本文缺少了这一部分的实验。

其次，本文的作者背景大多是 RL ，而且在本文中大量强调了 base model 的 capability 作为商业护城河的事情，这不由得让人们怀疑这篇工作是否别有用心。作者在文章中提到一点，只有 model capability 能够作为商业的护城河(moat)，那些靠 imitation dataset 作为护城河的公司，如果在竞争对手有相同的 capability 的model 的前提下，是很容易被超越的。但是作者并没有做任何证实这个观点的实验。

本篇工作未来可能的发展方向，以及课上的一些有趣的讨论。这篇工作生成 local imitation dataset 的方法是给 LLM 一堆 QA pair，让 LLM 生成更多的这样的 pair，然后用这些数据去 finetune 小模型。一种理解思路是认为这是一种数据增强的方法。而另一种理解思路是这是一种新颖的蒸馏 LLM 的思路，因为以往我们的思路是，相同的问题喂给小模型和 LLM，让后以 LLM 的回答作为 ground truth，来实现 finetune。这篇工作给我们提供了一种新的蒸馏 LLM 的方式，也即通过 dataset generation 的方式。但是这种方式会面临几个问题，首先可能 LLM 自己生成出来的答案有问题，所以这种事实性的错误也会传递给小模型，第二是 LLM 生成的 pair 在某种程度上会有同质化，可能会让小模型过拟合，实际效果变差。而这实际上暗示着我们，未来的研究可能会重点关注如何做更好的数据增强上，以及对 LLM 的蒸馏方法上。实际上，LLM 性能虽然很强，但是无论是训练一个 LLM，还是持续改进一个 LLM 的成本都很高昂，一种 alternative 的解决方案是训练出若干在每个任务上能力能和 LLM 媲美的小模型，相当于从 general，又回到了 task specific，不过，这一次的路径和人们更早时期的路径不同，因为这次是 from distilling LLM