

LLM 顶会"水"论文的思考

-by 韩非江

转发自小红书之前总结的热帖

下面总结两个我喜欢研究的 AI 方向，适合 Independent Researcher 或者学术界资源有限的来做。我做的这两个方向都中稿了。

声明：不是为了水论文而水，不然就是学术污点得不偿失了，除非你之后不搞学术。我一直坚持认为论文的 claim 要严谨，即便只是一个简单的问题。所以下面我仅从我个人的经验分享做什么。因为我发现太多朋友和我一样在外面很“孤独”，connection 上限低，又没钱，又没卡，但是仍然需要面对毕业或者申请的压力。积累一些原始资本后，我们才有更大的可能可以获得这些资源，进而做更有影响力的工作。（我会在几年后考古和更新这个帖子）

正文：首先，我认为[应用]方向的论文是最容易入门的。

某 Stanford PhD 曾言：“每当我对理论方向科研丧失信心时，发一篇应用论文就能给我续一口命...”

$$\text{AI} = \text{数据} + \text{模型} + \text{训练}$$

选择一个应用方向，从上面三个元素选择一个视角进行优化，就能做一篇很好的研究。

如果你选择了一个很火的应用方向，企业学术界都在做。模型架构修改入门难度高，你一个人做研究也很难比过企业，训练资源也不如他们。

这个时候就可以从数据入手，构建：

1. 更高质量的数据
2. 之前稀缺的数据
3. 更高效准确的标注 pipeline

你可以模仿李飞飞做一个新的 benchmark，或者提出一种新的数据标注范式（eg 之前需要手标，你提出一种方法自动获得 label）。缺点是这个过程需要大量时间成本。

如果你做 llm 或者 vlm，选择 training free/inference time adaptation 或许也不错。ICL/Steering/Calibration/Tuning attention/representation/logits.....

但是此时的重点是能否从原理上自洽的解释你的方法为什么 work。你的贡献不仅仅是一个 simple yet effective 的方法，还是许多 insights。这种论文首先分析一种现象，然后基于对这种现象的深入分析提出一种新方法。（不过我个人始终认为，在资源充足的条件下还是搞 training 吧，training 终将激活一切 training）

free 能带来的收益。不过后者仍然可能在可解释性上占优)

最后，从模型的视角，我感觉以 Efficiency 作为模型优化卖点也不错。例如：参数/token 剪枝，decoding 优化，模型蒸馏，skip layers..... (大量论文发现模型是冗余的，输入 token 是冗余的。架构不变，只能搞这些技术找补一下。5-10 年后模型架构更换以后这些方法还需要吗？我无法预测)

我感觉下面的公式容易讲故事：引入一个反直觉的现象。我们基于这个现象做了深入的分析，发现了一些 insights。我们发现这些 insights 反过来可以被利用来优化模型。这样研究，有 novelty (因为反直觉)，审稿人阅读起来容易看到之前没了解过的内容，读起来有趣。

有许多类似论文的卖点是“几行代码实现多个 benchmark 涨点”，我很喜欢这种方法优雅且有理论支撑和深入解释的工作。

更新：再补充一个感觉容易被 AI 审稿人给 positive 意见的故事线。

首先选择一个小众但是重要的方向，小众=参考文献很少，重要=满足顶会影响力的 bar。你可以定义一个新的问题，应用新的模型，收集新的数据。但是论文里要体现“We are the first....”，也就是说你是这个冷门方向的冷门问题的第一人。

这种论文的核心在于如何论述虽然你的方向冷门，但是重要。“第一人”效应在 AI 审稿时代会获得很高的 credits，只要 ai 被你说服，而且因为冷门搜不到资料，所以你的 novelty 就是充足的

何为冷门：数据收集困难；方向入门门槛高；偏工业届实践，所以学术界的不好想到；需要独特经历或背景才能想到的.... 想到别人想不到的问题本来就是 novelty

何为高影响力：如果这个方向能够被优化，则下游 ABCD 的应用都能被优化；提供的经验能被迁移到其他领域；之前其他类似领域的办法不能直接拿到你的方向用，你的方向具有特殊性，所以必须用你的研究才能解决.....

最后：

做热门方向的增量式研究容易被喷没有 novelty

做冷门开创性研究容易被喷毫无意义没影响力

最好的策略是：做热门方向中的冷门子方向，兼得影响力和 novelty。至于如何识别哪个方向好，就得看你的导师，还有你自己的积累和兴趣了... 祝大家都能早日发够论文脱离苦海！