

# 谈谈怎样提高炼丹手速

原创 夕小瑶 夕小瑶的卖萌屋 2021-03-16 22:20



文 | 夕小瑶

最近搞定几件焦头烂额的大事后，终于有了一丢丢的时间来写写文章，并且偶尔思考下算法工程师的核心竞争力是什么。

前不久一时兴起写了篇标题党文章《[惊了！掌握了这个炼丹技巧的我开始突飞猛进](#)》，简单描述了一下我的升级打怪路线图。后来想了想，发现还有一点极其重要的基本功常常被大家忽略，但确是初级&中级算法工程师之间拉开差距的重要因素，那就是：**手速**。

## 手速有什么用呢？

假如一件事情工程量较大（大型实验 or 策略pipeline复杂 or 从0到1单纯的代码量大），那么别人整整一周才能完成的工作你两天就做完了，于是你就多出来了5天的时间去刷刷paper、或者冥（fa）想（dai），于是你就成了一个有灵魂的打工人（有时间自我提升了！还能储备套路和灵感！）。

换个视角，假如你手速不行，大家预期你两周能完成的工作，实际你做了三周，这就会导致你陷入到了“不断赶ddl”的节奏怪圈和产出焦虑中，不仅失去了成长和灵魂，还可能失去生活和女朋友！哪怕你曾经是个聪明的小机灵，也会由于疲于赶ddl而失去思考的机会。



再比如，假如一件事情风险很大（成功后收益当然也更大），对正常人来说是不敢尝试的，毕竟谁都不想怒肝五天后发现是个代码全废的失败idea，于是周报上：“本周无进展，只有个失败的尝试”。

于是，大部分人只去尝试一些确定性强的有收益的事情，当然这类事情也很难带来大的惊喜和差异化，而手速快的你就有资本去尝试这类风险较高的操作（毕竟失败了你也没耽误主线进度），于是你就相比其他人有了命中高收益事件的可能性。



短时间内，这个“灵魂收益”+“高收益事件命中收益”可能看不出太大意义，但长线下去就会引起质的差异。毕竟“文能创新写paper，武能光速肝产出”的小可爱在职场中是相当稀少的，绝大部分人仅有一，而当两者兼备时便能产生奇妙的化学反应，取个五毛钱的名字，叫做“核心技术差异化”。

get了手速的重要性，接下来就谈谈怎么提高炼丹手速。

## 提高炼丹手速

最重要的是工具层面的基本功：

1. 熟练掌握主力编程语言（比如python）。不是单纯的记住语法就叫熟练了，还包括常见的写法优化、代码效率优化、高频代码段的光速实现、光速修报错（常见python层报错还靠Google的话建议面壁）等。
2. 熟练掌握shell，积累替代python代码段的奇技淫巧。带过几个实习生，发现实习生手速慢的其中一个重要原因是普遍不熟悉shell这种high-level脚本语言，经常管道黏几个命令就能搞定的事情，实习生往往要花几十分钟写一坨python代码，还效率极差，数据规模一大就尴尬。
3. 熟练掌握主力深度学习框架。这里不是说必须要去深究框架层源码细节，而是说要对框架底层运行机制和常用API的行为有非常清晰的把控，如果还能做到框架能力覆盖面（哪些功能/操作这个框架实现不了，或者实现代价高，甚至有bug）方面的把控和大版本差异化点方面的了解，那就更好了。这会在你尝试一些比较有脑洞的想法，或者快速验证一些fancy但闭源的论文idea时有很大的帮助。如果你想实现一个操作却不知道这个框架有没有API支持，那就说明对这个框架的了解还不够。

这种工具层面的基本功怎么提升呢？没有好的捷径，那就是多写代码，尤其是主导研发万行代码以上的项目，会非常锻炼你的语言/框架基本功。如果没有这种机会，也可以偶尔抽出些精力去重构实验项目代码，提升其易用性和灵活性。

## 再说说积累和习惯问题

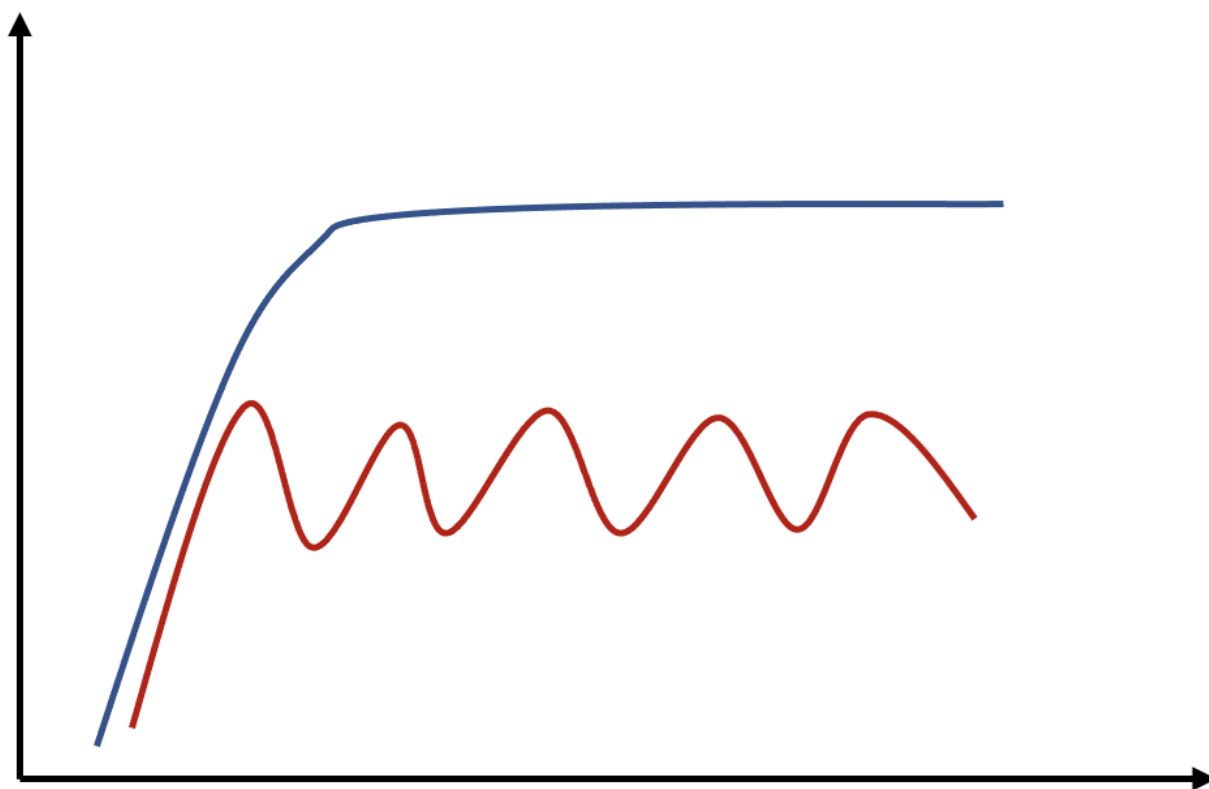
为每种经典算法任务都积累一个重构的不错的完整、易用、灵活的project，这种project就是将来你完成其他相近project的热启动project。由于这个project是经过你的手重构的，所以二次上手成本和魔改成本会远低于你去找个外部的github repo。

以NLP为例，基于传统cnn/lstm的文本分类、文本匹配、序列标注、自回归生成、阅读理解基线project，基于BERT的分类匹配标注阅读理解MLM project，离散检索&向量检索project，离散聚类&向量聚类project，serving project，知识蒸馏project，多feature ensemble，多任务学习project等都可以根据自己实际需要抽时间折腾重构&沉淀，亲测上手新任务时还是非常省心省力省时间的。建议大家业余抽点精力将自己工作内容相关的project时不时的重构、打磨一下，一年后做类似的project时你会非常感激自己当年的勤劳的！

最后说一些不建议频繁尝试的提高手速的奇技淫巧：

再次警告下，不建议频繁尝试！否则后果自负！

1. 咖啡。在非常困和非常精神的时候就别考虑了，在微困和微乏时就是效率神器了。强烈建议不要喝太浓的、太劣质的和太甜的。如果发现出现眩晕、头皮发麻、焦虑、注意力不集中或易出汗等现象，就说明这杯咖啡对你来说太浓了，或者太劣质了；如果喝完没有任何感觉，反而更想睡觉了，说明糖加多了或者太劣质了。另外，熬夜后第二天千万不要喝咖啡，强行提神很容易造成免疫系统异常（生病、过敏甚至影响很长一段时间的精神状态）。
2. 换个环境。当前所处环境干扰较多（尤其是非背景白噪声的干扰，比如时不时有人说话、手机提醒音、有人来回走动甚至戳你聊会天）时，找个没人的会议室/角落可以大大提升注意力。长期未被打断的沉浸式做一件事情时，人的注意力状态/效率大概是蓝色线，长期处于敏感干扰较多的环境中时，人的效率大概是红色线。当然，一人占用会议室是浪费公共资源，且乱换环境也有被怀疑摸鱼的风险，所以虽然可以大大提高手速，但是尝试要慎重。



3. 扔掉手机，屏蔽聊天工具通知甚至直接退出（后果自负。。。)

4. 隐藏mac docker栏/windows 通知栏，营造沉浸式开发环境（后果自负)

若想再进一步，就是研究一些高阶的时间管理技巧了，统筹规划下自己的时间碎片和事件列表，尤其当N多条并行推进时，若缺乏必要的时间管理能力和意识，很容易丢三落四、将事情搞得一团糟并且导致自己无暇炼丹，更不必说手速了。这方面又能写一篇文章了，大家感兴趣的可以自行去知乎做做功课，这也是打工人进阶的必备能力！

好啦，本期分享就到这里了，大家有什么困惑或者想听的经验分享也可以在评论区留言，下期分享的锦鲤主题可能就是你！没抽中的小伙伴也不要难过，下次一定！

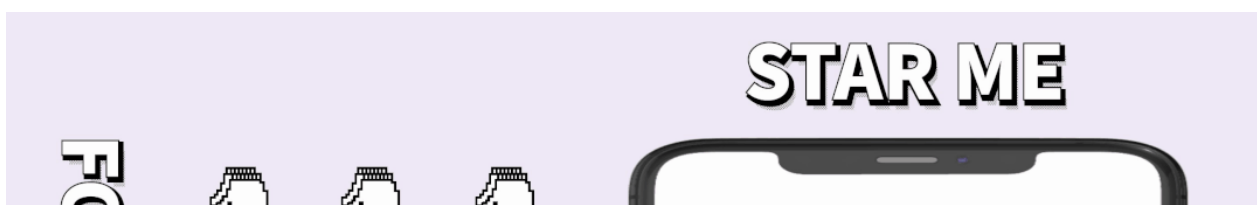


后台回复关键词【**入群**】

加入卖萌屋NLP/IR/Rec与求职讨论群

后台回复关键词【**顶会**】

获取ACL、CIKM等各大顶会论文集！





喜欢此内容的人还喜欢

若被制裁，中国AI会雪崩吗？

夕小瑶的卖萌屋