

和学弟聊了一下午，关于在这个AI狂飙的时代，我们到底该怎么学

今天和学弟聊了很久，他问我平时看什么公众号，怎么把AI理解得这么深。其实说实话，我从不迷信某个特定的公众号，我的核心心法只有一个：**抛弃二手的知识，直接去读源码。**

我一直觉得，计算机世界最性感的地方就在于它的“开源精神”。其他学科的好东西可能被藏着掖着，但在这里，最顶级的代码、最前沿的思想，就那样赤裸裸地躺在GitHub上。知识获取的成本极其低廉，这本身就是我们最大的红利。所以，别再花大把时间去刷那些冗长的视频课了，那个效率太低。在这个时代，最好的学习方式就是把GitHub上的顶尖项目拉下来，不懂的地方直接一行一行问AI。像Cursor或者Traq这种工具，现在对学生党这么友好，完全可以把它们当成你的24小时私教。代码看不懂？甩给它，让它给你讲逻辑、讲数学原理，这才是“降维打击”式的学习。

聊到具体的路线，我发现很多人还在传统的机器学习（ML）里打转。虽然大三必修课会讲，但说实话，工业界早就不怎么用那一套了。对于ML，我觉得你只需要搞懂线性回归、逻辑回归、K-means和PCA这四个核心算法，会用scikit-learn，理解监督和无监督的本质就足够了，千万别恋战。

真正的重头戏在深度学习（DL）。这里我有个非常重要的建议：**重数学，轻记忆。** 比如反向传播（BP）算法，你必须把它的数学推导吃透，这是基本功。但对于PyTorch里那些繁琐的张量形状变换接口，千万别去死记硬背。那是工程层面的事，容易晕也容易忘。现在的玩法是，你直接在代码注释里写清楚你想要的输入输出形状，剩下的代码补全交给AI去做。我们要掌握的是架构的思想，比如Transformer里的注意力机制，去读《Attention is All You Need》的论文，去理解它为什么能取代RNN和LSTM，而不是纠结于某行代码具体的写法。

再往后走，就是大模型（LLM）和Agent了。别被那些高大上的名词吓住，去看看Karpathy大神的 nanoGPT，那是理解LLM最小最好的原型。想动手练手？别用自己的笔记本硬跑，去AutoDL租个GPU服务器，学会用SSH远程连接，学会用 nohup 挂后台。对了，Python的包管理也别再用Conda了，试试Rust写的 uv，速度快到飞起，记得坚持“一个项目一个虚拟环境”的原则，能省去你未来无数的麻烦。

至于现在大火的强化学习（RL），它才是让模型超越人类数据的关键。虽然数学很难，但耐着性子多啃几遍论文，结合 cleanrl 这种清爽的代码库去学，你会打开新世界的大门。特别是像DeepSeek都在用的GRPO算法，绝对是当下的显学。而如果你的志向是做那个“卖铲子的人”，那么AI Infra方向，特别是搞懂 vllm 的PagedAttention机制，绝对是中美大厂都疯抢的核心竞争力。

最后我们聊到了“世界模型”（World Model）。这可能就是继ML、DL、RL、LLM、Agent之后的第六大江山。现在的RL本质上还是暴力搜索，效率太低。而世界模型能让AI像人一样“先做梦、先规划”，在脑子里预演一遍再去行动，这种“聪明地探索”才是通往通用人工智能的必经之路。

这一路聊下来，其实千言万语汇成一句话：**在这个时代，别做“听课”的学生，要做“读源码”的工程师。** 所有的秘密都在代码里，去跑起来，去debug，去和AI结对编程，这才是最快的捷径。