

周报 2025014~20250120

transformers复现

```
d_model = 512
n_layers = 6
n_heads = 8
ff_hidden_dim = 2048
```

均使用贪心搜索。 原论文使用 beam=4, alpha=0.6 的束搜索
原论文小模型bleu4分数为 28.4

lr	batch_size	epoch_cnt	final_loss	bleu4 score	description
0.000005	64	2	3.23	6	
0.00005	64	2	1.95	23.7	
0.00003	64	3	1.93	24.0	
0.00001	64	4	2.08	21.5	
0.0005	?	?	?	28.4	原论文

代码能力提升计划

“良好的代码能力”究竟指什么

不太懂，感觉在语境下指的是类似于项目能力的样子，能自行完成环境搭建和代码编写用来把积木搭出来。

环境搭建需要较强的信息检索能力，代码编写主要要求能理解代码运行逻辑（然后其实就可以写了吧.....）。

可能的一些建议

针对有c以及python基础，面向对象语言基础的同学，如群里的“走肖木同”，我觉得或许他甚至已经可以直接上手。

他可能可以的改进方向是学习简单的算法竞赛知识，比如前缀和、差分等，这些可以让他不至于搞出非常愚蠢的代码导致时间复杂度爆炸（。

我认为动态规划、滑动窗口、图论对深度学习的学习有辅助作用，可以适当接触。

针对没有c及python基础的同学，建议先搞懂基础语法，我没有具体的建议，因为我学这些时已经有在minecraft中敲了4年指令经验，这导致我完全速通了它们。

要记得搞懂递归等有些难度的东西，理解并熟练使用面向对象的思想。

我认为信息检索能力是最重要的，但我不了解这个如何训练QwQ。

组会上所说的方向

还是有点没太搞懂到底这是啥来着2333

目前还在发散性思考ing

以下是一些发散性思考，求批评指正QwQ:

1. 在人工智能点网页这个故事背景下，回退的情况我认为原因是网页内容与预期不同，是我们的模型在点击后才发现的情况导致它意识到这一问题而选择回退。因此，存在回退的基本条件是我们新的状态中能获得更多的信息。
2. 很难想象它是如何用在AI编程中，或许是调用一下编译器告诉一下它存在编译错误或者是解析结果反馈代码模块？
3. 它们的关系很像图，如果是稀疏图的话或许可以用图论优化些什么？（

有以下一些问题QwQ:

1. 去哪里找资料/论文比较好，如何挑选有东西的论文

下周计划

启动那个搞不懂到底叫啥的方向（

找些论文了解情况。