

周报（4.30-5.7）

汪李昀虹 2023141460197

本周进展

1. 继续研读探讨论文原文以及相关论文，已明了实现原理；
2. 尝试接入deepseek模型作为扰动模型；
3. 解决了之前每次运行都要重新下载模型的问题。

现存问题

1. 现有原理存在一定缺陷，比如生僻字较多的文章识别不准（已经过实验验证：《滕王阁序》被判定为ai，而ai生成的生僻字较多的文章被其中一个模型判断为人类生成）；（解决方案：可能是该方法自带的问题，难以规避，毕竟越生僻的字/词的同义字/词越少）
2. 同一篇文章，使用不同模型，有时结果差异显著，原理不明（解决方案：有待进一步商榷）

后续工作

1. 接入deepseek进行实验；
2. 探讨更多优化可能性；
3. 反馈功能上线。