

组会汇报

陈钊杰
专业:计算数学

January 2, 2024

目录

- 1 模型构建
 - 相关构建方法
 - MathGPT模型构建思路
 - 文献阅读

模型构建

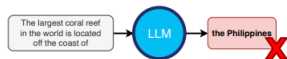
- 1 在原本的MLC模型基础上进行修改不太方便，有很多报错，所以我使用一个最简单的transformer模型，如此一来，参数量，注意力机制头数以及各种组合方式都可以自己控制，同时增加多个transformer模型合并也更好控制。目前这周就只完成了transformer模型的基本搭建,并进行了基本模型测试，还未把我们数据进行测试。

MathGPT模型

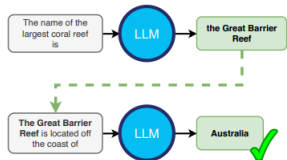
1 MathGPT测试特征

- 1 目前开源的代码那边测试的只有5位数以内的加减乘除，并没有进行高位数的测试,训练的也是五位数以内的数字。

纠正Transformer模型中推理过程中的多跳推理失败



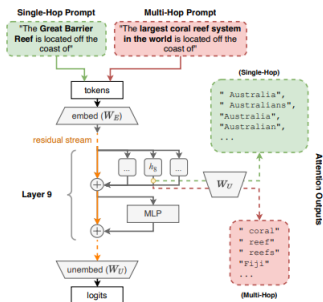
(a) Multi-hop prompt.



(b) Multi-hop prompt broken into 2 single-hop prompts.

- ❶ 直接无法得到答案，需要逐步执行获得最终结果。
- ❷ 这一步的输出可以作为下一步的输入操作。

纠正Transformer模型中推理过程中的多跳推理失败

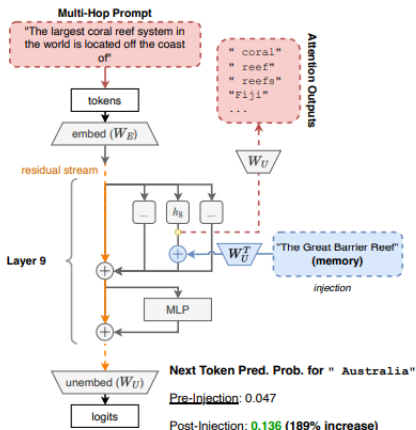


1

2

这个模型的核心思想是通过在多头注意力层中注入相关记忆来提高预测的准确率。

纠正Transformer模型中推理过程中的多跳推理失败



1

2

这个模型的核心思想是通过在多头注意力层中注入相关记忆来提高预测的准确率。

下一步计划

- 1 在基础的transformer模型上进行数据测试。

谢谢老师和同学们的聆听!