

组会汇报

陈钊杰
专业:计算数学

December 19, 2023

目录

1

代码调试

- MLC模型学习基本加减乘法来进行任意数值的计算

模型部署说明

- 模型实现功能:只学习一种规则来对数字完成基本的加减乘以及混合运算.
- 现有模型特点:传统语言模型一般在数字理解能力上比较弱, 比如像chatglm类似的语言模型, 一般在数字计算上都是通过大量的预训练数据去训练, 比如计算 $3+4=?$ 的时候, 现在常见的语言模型都是去回忆以前有没有做过类似的题, 而不是理解‘3’, ‘+’, ‘4’分别指代的意义。因此在数字计算上的能力取决于训练数据, 但不管多少训练数据, 数字计算是无穷无尽的, 模型如果不理解数字就难以拥有具有强大的计算能力。
- 我们的模型希望从数学角度上对数字进行理解。因此通过给定各个数字的运算规则, 并给少量运算样例, 实现数字的计算。
- 我们模型主要能够实现以下的功能:
模型自行理解数字, 取余以及整除算法, 来进行任意数字之间加减乘计算。

加减乘的设计思路

- 首先让模型进行训练学习最基本的乘法，加法表，以及什么是整除，什么是取余，什么是进位。
- 对于任意的两个整数使用模型所理解的加法，进位，进行数字之间的运算。
- 训练数据集示例

```
1 *SUPPORT*
2 IN: 8 sum 4 OUT: 1 2
3 IN: 9 sum 2 OUT: 1 1
4 IN: 4 sum 6 OUT: 1 0
5 IN: 9 sum 6 OUT: 1 5
6 IN: 7 sum 7 OUT: 1 4
7 IN: 2 prod 4 OUT: 8
8 IN: 7 prod 7 OUT: 4 9
9 IN: 9 prod 6 OUT: 5 4
10 IN: 1 prod 6 OUT: 6
11 IN: 4 prod 8 OUT: 3 2
12 IN: 3 sub 3 OUT: 0
13 IN: 0 sub 6 OUT: -6
14 IN: 6 sub 7 OUT: -1
15 IN: 1 sub 5 OUT: -4
16 IN: 17 mod 10 OUT: 7
17 IN: 16 mod 10 OUT: 6
18 IN: 13 div 10 OUT: 1
19 IN: 15 div 10 OUT: 1
```

Table: MLC模型评估参数

相关参数	数值
最大评估长度	20
模型参数量	1395344
编码器,解码器层数	3
多头注意力机制中的头数	8
dropout	0.1
前馈神经网络的激活函数	GELU

Table: 准确率结果(模型)

四则算法种类	数据量	训练次数	准确率
随机5位数加法计算	5k	100	100%
随机6位数加法计算	5k	100	100%
随机7位数加法计算	5k	100	100%

- ① MLC模型在通过训练学习到的数字能力，进行高位数计算，不论数字的位数均可以进行准确的计算得到。

下一步计划

① 将任务在数值上更加复杂,探究更复杂的运算能力.

① 解决乘法问题。

① 乘法的交叉相乘，步骤拆分的时候,步骤比较多，模型进行计算时候步骤比较繁琐。

$$\begin{array}{r} 3 \ 4 \\ 5 \ 9 \end{array}$$

② 尝试训练模型在现在基础上，进行结合律，交换律等规则的学习，来提高数字计算效率。

谢谢老师和同学们的聆听!