

组会汇报

陈钊杰
专业:计算数学

December 12, 2023

目录

1

代码调试

- MLC模型在解决大数值上加减乘计算能力

模型部署说明

- 模型实现功能:能够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:
我们的模型能够对数字进行加减乘及混合运算的规则进行学习, 然后进行其他相关数值的计算。
比如我们的模型学习 $3+5-4$ 的学习, 这种混合运算的模型, 那么对于所有 $x+y-z$ 都能以极高的准确率进行问题的回答。

100以内加减的设计思路

- 使用10进制进行计算，无需增加新的规则，只需理解1到9即可，比如34,可以理解看成10 10 10 1 1 1 1。
- 同理类推，对于大数都可以用类比思路，这样的思路之下，只需要教模型理解数字1到9即可，所有数字都可以进行类型的加减了。

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
10以内加减混运算	5k	200	100%
10以内加减乘混运算	5k	200	96.83%
10以内加减混运算	100k	21	76.5%
10以内加减乘混运算	100k	21	55.67%
100以内的加减混运算	100k	20	75%
100以内的加减混运算	25k	11	87.5%

- ① 可以使用较少的数据量在训练次数上足够多，就可以以高准确率来完成计算任务。
- ② 100以内的加减计算，使用较少的训练次数就可以达到比较高的准确率了

训练结果计算图展示(100k,20次)

- 关于100以内数值加减计算(测试结果)

$$+13-5 = 8$$

$$+5+21 = 26$$

$$-21+13 = -8$$

$$-2+4-61 = -59$$

- 关于100以内数值加减计算(泛化结果)

$$+26-3 = 23$$

$$+35-2 = 37 \text{ (** target: 33)}$$

$$-36+42 = 6$$

$$-52+32 = -20$$

$$+5+43 = 48$$

$$+2+62 = 64$$

$$-3+5-12 = -10$$

$$-2+3-12 = -11 \text{ (** target: -12)}$$

训练结果计算图展示(25k,11次)

- 关于100以内数值加减计算(测试结果)

$$+13-5 = 8$$

$$+5+21 = 26$$

$$-21+13 = -8$$

$$-2+4-61 = -59$$

- 关于100以内数值加减计算(泛化结果)

$$+26-3 = 26$$

$$+35-2 = 35$$

$$-36+42 = 42$$

$$-52+32 = 32$$

$$+5+43 = 48$$

$$+2+62 = 64$$

$$-3+5-12 = 5$$

$$-2+3-12 = 3 \text{ (** target: -12)}$$

下一步计划

- ① 将任务在数值上更加复杂,探究其泛化能力.
 - ① 将数值进一步的提高,理论上可以进行任何数值的加减计算。
 - ② 开始准备高位数的乘法运算模型搭建。

谢谢老师和同学们的聆听!