

# chatglm模型序列预测结论汇总

陈钊杰  
学号: 22135030  
July 13, 2023

## 1 序列相关的基本性质汇总(预测已经见过的任务)

- 预测长度越短, 预测结果的误差就越小.
- 对于相同的训练任务, 给定的训练数据较多时, 预测结果的误差就越小。
- 对于相同的训练任务, 给定的训练数据超过某个阈值, 会使得预测结果的误差变大。(这可能是因为过拟合导致的)
- 进行预测时,历史序列的长度越长,最终预测结果的误差就越小.
- 进行预测时,需要预测的长度越短,最终预测结果的误差就越小.
- 对于两种不同的函数序列(其中预测维度的函数相同),那么其中维度更高的函数序列,最终的预测误差会更小.
- 在多种类函数序列训练任务中,预测结果的误差,以及任务的通过率均低于单一训练任务.
- 在对于多种预测长度的训练任务中,同样可以看到历史序列的长度越长以及需要预测的长度越短,最终预测结果的误差就越小.
- 不同的预测长度的训练任务对于预测结果精确度影响不大, 不同的函数序列训练任务对精确度的影响比较大。
- 同上,通过增加训练数据的数量,可以改善预测结果的误差.

## 2 泛化性结论汇总(未见过的预测任务)

- 预测的函数序列和训练的函数序列相同,但预测序列的长度和训练的长度不同,,结果发现正确的预测长度应该是4,但是结果预测长度全部是8.从中可以看出,模型在解决预测长度不同的新问题上泛化能力较弱.
- 在预测没见过的任务上,误差明显大于见过的任务.
- 可能预测序列的维度对预测结果的误差影响不大.
- 提高训练的数据量,对于最终的通过率有明显的提升.但是预测的准确率没有明显的上升.

## 3 从相关论文中总结出的一些结论

- 在使用的10个语言模型和10个时间序列模型的比较中
  1. 大部分语言模型的性能是要高于时间序列模型。在三个测试数据集中, 大部分语言模型的MAE和MSE均超越了时间序列预测模型。
  2. 测试的几个的预训练模型中,Pretrained Model Size越大,预测效果也越好.
- 关于模型泛化能力的测试:
  1. 在泛化能力的测试上面, 语言模型的预测结果的误差小于时间序列模型。说明语言模型的泛化能力更强.
  2. 语言模型在处理泛化任务上,也能够保证一定的缺失率(通过率)。
- 关于序列预测长度的相关结论
  1. 给定同样的历史序列长度, 预测长度越短, 预测结果误差越小
  2. 预测相同的长度, 给定的历史序列越长, 预测结果误差越小
- 提示指令的相关结论
  1. 关于提问, 相互间隔有规律的提问方式(类似表格形式的提问方式)比并排(简单的陈述句)的方式在预测效果上更好。
  2. 通过热力图说明transformer的多头注意力机制是如何解读指令数据的内容.
  3. 对于语言模型而言,他能够将注意力的大部分比重都放在有效提示中(提示命令中的序列语句),在无效的提示词中有更小的比重.
  4. 语言模型确实用到了历史序列的信息, 这为语言模型进行时间序列的预测奠定基础.