

组会汇报

陈钊杰
专业:计算数学

November 14, 2023

目录

1 相关文献阅读-MLC模型

2 代码调试

- 当前序列分类任务汇总结果
- 增加k线图特征每支股票误差可视化结果

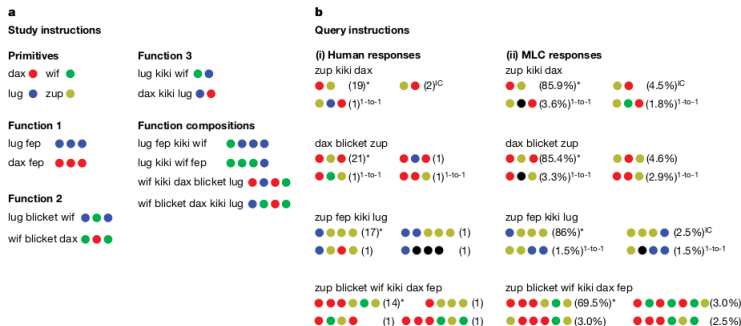
这周主要做的事情

- 1 组合性元学习(MLC)模型是如何学得组合技能的?
- 2 为股票数据增加一个k线图特征维度,并使用informer模型进行对比效果.
- 3 修改了LSTM模型对于高频节点的选择以及预测.

MLC模型的系统性组合和概括能力

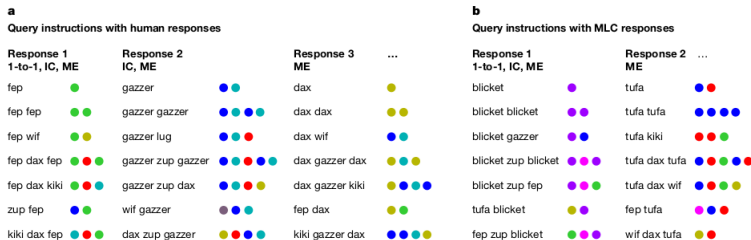
- 这篇文章的通过创建元学习神经网络来模拟类人系统泛化,简单说就是让神经网络模型具有像人一样的基本泛化能力.
- 文章通过对比人和模型在一些简单的泛化任务的表现来表明观点.
- 关于MLC模型的架构.
- 关于MLC模型的结论.

人类与MLC模型的对比实验1



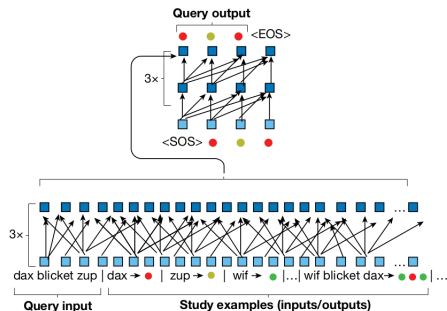
- 对人类参与者 (n = 25), MCL模型提供了一套由14个学习指令 (输入/输出对) 组成的课程, 并被要求为查询指令生成输出.

人类与MLC模型的对比实验2



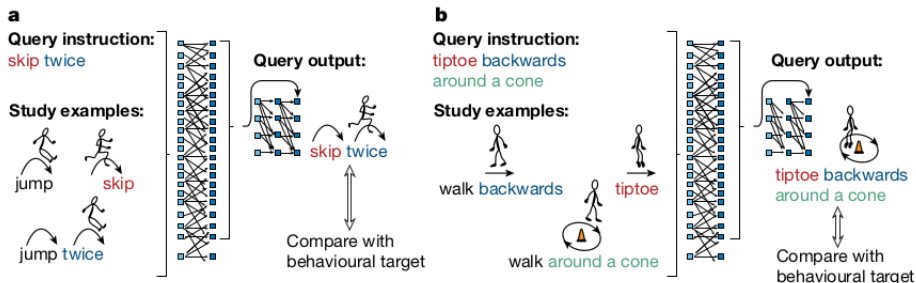
- 参与者在未看到任何学习示例的情况下对查询(语言字符串)产生了响应(彩色圆圈的序列).其中1-to-1,IC,ME分别表示以一对一,从左到右的方式将查询翻译为图标串联,相互排斥的行为方式.
- 人类的回应非常有结构,且遵循上述的行为方式,MLC模型也得到了有结构性的回答,且符合上述的行为方式.

MLC模型的架构



- MLC模型使用了标准的transformer架构,MCL模型学习一种例子的方式就相当于元学习，在训练阶段,优化是在动态变化的情况下进行的（每种情况都有新的学习和问题示例），而不是在静态数据集上进行的.在测试情节中，模型的权重被冻结，不提供任务特定的参数,随机的将每个问题与输出序列进行比对.

MLC模型如何获得组合技能？



- MLC模型:通过动态的组合任务来获得组合技能.
- a图中的输入中把复杂任务jump twice通过jump这个基本任务进行组合得到,因此对于skip也可以通过类似的组合得到skip twice.

MLC模型的学习本质

- 在传统的机器学习中，模型通常在大量的训练数据上进行学习，然后在测试数据上进行泛化。然而，这种学习方式不一定能够使模型具备良好的泛化能力，特别是在面对新颖、复杂的任务。
- MLC的目标是通过训练模型在一系列少量的组合任务上进行学习，以鼓励模型学习到更加通用和抽象的表示。这些组合任务通常是由多个基本任务组成，每个基本任务都涉及到不同的输入和输出组合。通过在这些组合任务上进行学习，模型可以学会将不同的基本任务组合起来解决更复杂的问题，从而具备了更好的泛化能力。
- MLC的关键思想是通过训练模型在多个任务上进行学习，使其能够学习到任务之间的共享结构和模式，从而能够在新任务上进行快速、准确的泛化。这种学习方式可以帮助模型发现任务之间的相似性和关联性，并从中抽取出通用的知识和模式，以便在未见过的任务上进行应用。

文章的结论

- MLC通过元学习获得其能力,系统性泛化和概括能力不是神经网络架构的固有属性,而是从数据中诱导出来的.
- 展示了MLC如何使标准的神经网络经过优化以实现其组合技能,在比较中,以模仿或超越人类的系统性泛化.MLC的系统性远远强于以标准方式训练的神经网络.
- MLC也有一些不足之处,比如MLC无法处理更长的输出序列以及更复杂的句子结构,这是将来要解决的问题。
- 由于这篇文章的源代码还没有,所以只能知道大致的框架,无法得知其具体的做法.

Table: 准确率结果(均用共享型模型进行测试)

| 模型 | 平均准确率 |
|--------------------------|--------|
| Word2Vec方法, 窗口大小为1 | 29.49% |
| Word2Vec方法,窗口大小为2 | 28.90% |
| Word2Vec方法,窗口大小为3 | 27.06% |
| TCN模型 | 34.52% |
| InceptionTime模型 | 34.80% |
| TSiT模型 | 34.71% |
| transformer模型 | 34.55% |
| chatglm模型 | 32.34% |
| informer模型使用交叉熵损失函数 | 23.46% |
| Word2Vec方法,选择频率较高的节点 | 34.33% |
| Word2Vec方法,选择频率较高的节点(更新) | 34.46% |

- 使用word2Vec方法编码并选择频率更高的节点

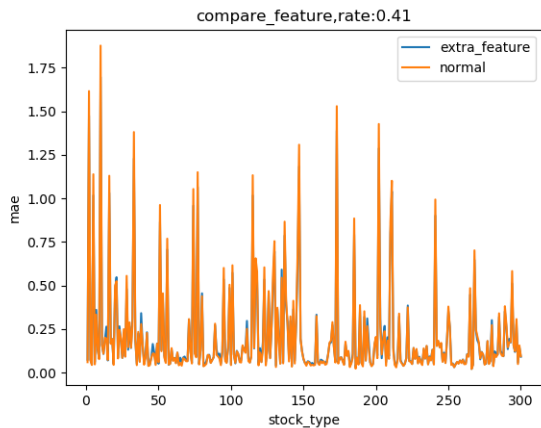
- ① 例如有一个历史序列节点为"A,B,C,D,E,F,G,H",预测值为"I"
- ② 通过选择高频率历史结点"A,AB,BC,CD,E,EF,FG,FGH",预测值为"C"或"BI" 或"GCH"
- ③ 那么我只选择预测值中的最后一个节点,比如"C"的话就选择C,"BI" 就选择I,"GCH"就选择H,

- 增加的特征维度具体操作

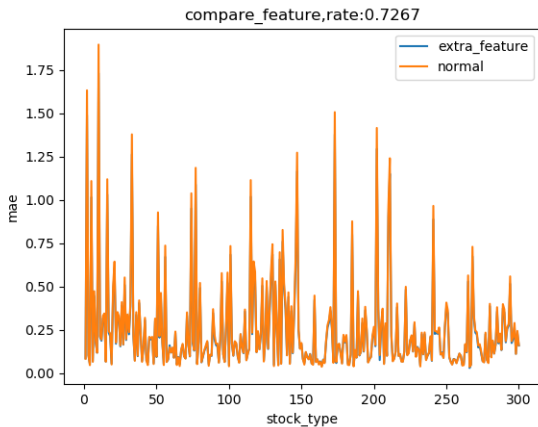
- ① 使用数字来表示类别标签.
- ②

```
1 date,open,close,high,low,volume,turnover,amplitude,percentage change,price change,turnover rate,feature
2 2020-06-26,7.29,7.24,8.29,7.1,37289540,591168816.0,15.28,-7.06,-0.55,7.78,11
3 2020-06-29,7.68,7.36,7.8,7.29,7374005,116415146.0,7.04,1.66,0.12,1.54,3
4 2020-06-30,7.36,7.56,7.69,7.24,5047192,79090061.0,6.11,2.72,0.2,1.05,7
5 2020-07-01,7.58,7.55,7.68,7.34,3415231,53630063.0,4.5,-0.13,-0.01,0.71,11
6 2020-07-02,7.61,7.6,7.68,7.39,7213077,113699701.0,3.84,0.66,0.05,1.51,11
7 2020-07-06,7.57,7.31,7.57,7.17,3460806,53791957.0,5.26,-3.82,-0.29,0.72,2
8 2020-07-07,7.19,6.97,7.23,6.89,4228006,64412773.0,4.65,-4.65,-0.34,0.88,2
9 2020-07-08,6.9,6.95,7.1,6.83,3691465,56029264.0,3.87,-0.29,-0.02,0.77,11
10 2020-07-09,6.9,6.54,6.98,6.52,5107052,76136533.0,6.62,-5.9,-0.41,1.07,4
11 2020-07-10,6.56,6.39,6.65,5.9,6979252,100685702.0,11.47,-2.29,-0.15,1.46,3
12 2020-07-13,6.39,5.79,6.39,5.72,4380110,62046599.0,10.49,-9.39,-0.6,0.91,2
13 2020-07-14,5.71,6.38,6.51,5.51,4127624,58816955.0,17.27,10.19,0.59,0.86,7
```

对比差异可视化:多输入多输出结果



对比差异可视化:多输入单输出结果



谢谢老师和同学们的聆听!