

# 组会汇报

陈钊杰  
专业:计算数学

November 7, 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
- 模型的修改思路

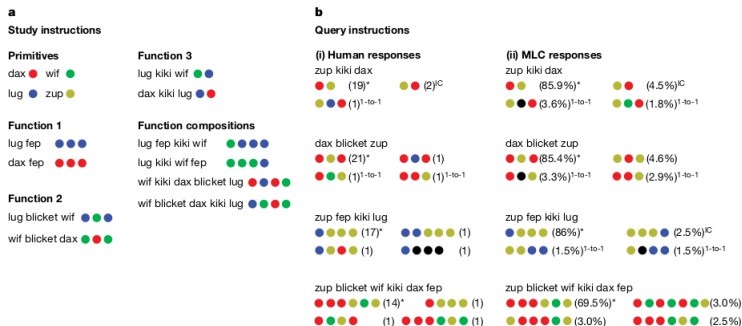
# 这周主要做的事情

- 1 阅读了一篇关于探索神经网络的系统性组合和概括能力的论文
- 2 修改LSTM模型,使其固定节点个数为25,并选择频率较高的二/三元组作为序列元素进行预测.
- 3 将时间序列预测的informer模型修改为序列分类的模型

# MLC模型的系统性组合和概括能力

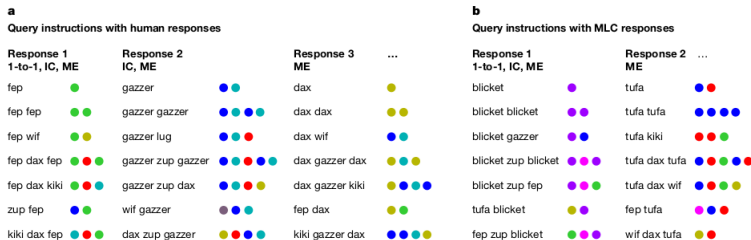
- 这篇文章的通过创建元学习神经网络来模拟类人系统泛化,简单说就是让神经网络模型具有像人一样的基本泛化能力.
- 文章通过对比人和模型在一些简单的泛化任务的表现来表明观点.
- 关于MLC模型的架构.
- 关于MLC模型的结论.

# 人类与MLC模型的对比实验1



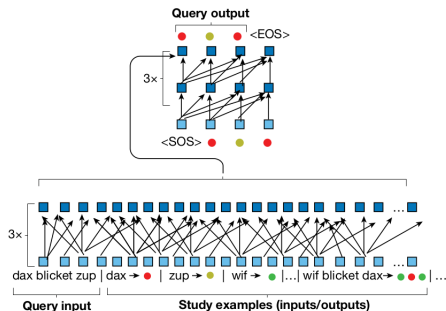
- 对人类参与者 (n = 25), MCL模型提供了一套由14个学习指令 (输入/输出对) 组成的课程, 并被要求为查询指令生成输出.

# 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.

# 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%

- 使用频率较高的节点作为基底

- ① 将k线图类别中频率最高的前200个作为基底.  
比如将若干个节点组"A,B,D,G,C,...,G,C,G"变成"AB","D","GC",...,"GCG"(个数为25).
- ② 每次选择25个节点来预测一个节点  
比如根据25个节点:"AB","D","GC",...,"GCG",去预测最后一个节点,比如预测结果是"AB",那么就把"A"作为预测的结果.

- 关于informer主要做了如下两个修改:

- 取消了原本的标准化,数据的标准化以后会导致所有值都在 $[-1,1]$ 附近不同值之间的特征变得不明显.
- 对于informer的输出部分的线性层和softmax处理器的修改,新增了一层12类别的线性层,并把损失函数从`nn.MSELoss()`改为`nn.CrossEntropyLoss()`

# 谢谢老师和同学们的聆听!