

# 组会汇报

陈钊杰  
专业:计算数学

September 5, 2023

# 目录

1

## 代码调试

- 对测试集数据进行预测,并计算预测误差
- 预测结果可视化

# 时间序列模型与自然语言模型对比实验

- ❶ 将50支股票分别选取6支并做对比实验
- ❷ 进行可视化对比语言模型和时间序列模型的对比
- ❸ 线性模型的指标对比中，对数据进行四舍五入。

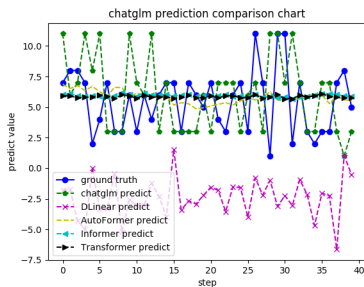
# 评估指标

- ① 准确率：准确率是模型正确分类的样本数与总样本数之比。它衡量了模型的整体性能，但在不平衡类别分布的情况下可能不是一个很好的指标。
- ② 宏平均精确度：它计算每个类别的精确度，然后取平均值。不考虑类别不平衡，对每个类别的性能平等看待。
- ③ 宏平均召回率：与宏平均精确度类似，它计算每个类别的召回率，然后取平均值。也不考虑类别不平衡，对每个类别的性能平等看待。
- ④ 宏平均F1分数：它计算每个类别的F1分数，然后取平均值。对每个类别的性能平等看待，适用于类别不平衡的问题。
- ⑤ 微平均精确度：它计算所有类别的真正例和假正例的总数，然后计算总的精确度。在类别不平衡时，它更受到正例类别的影响。
- ⑥ 微平均召回率：它计算所有类别的真正例和假负例的总数，然后计算总的召回率。也在类别不平衡时受到正例类别的影响。
- ⑦ 微平均F1分数：它计算所有类别的真正例、假正例和假负例的总数，然后计算总的F1分数。同样，它在类别不平衡时受到正例类别的影响。

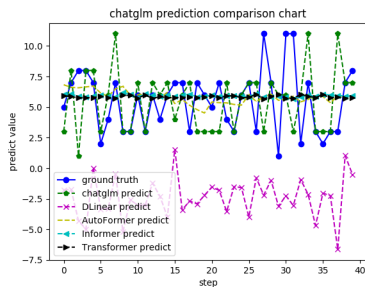
Table: 准确率结果

sss	chatglm	DLinear(one/much)	Autoformer(one/much)	Informer(one/much)	Transformer(one/much)
样例1	0.16	0.0021/0.0	0.1203/0.1266	0.1701/0.0	0.1701/0.029
样例2	0.1684	0.0021/0.0	0.1245/0.1203	0.1743/0.0	0.1743/0.029
样例3	0.8877	0.0/0.0	0.4772/0.3672	0.7573/0.0	0.187/0.0332
样例4	0.158	0.0/0.0	0.1037/0.11	0.1556/0.0	0.1701/0.029
样例5	0.1601	0.0083/0.0	0.1017/0.0975	0.0996/0.0	0.1058/0.0436
样例6	0.2869	0.0/0.0	0.0166/0.0145	0.0041/0.0	0.0041/0.0041

使用一个时间序列模型进行训练

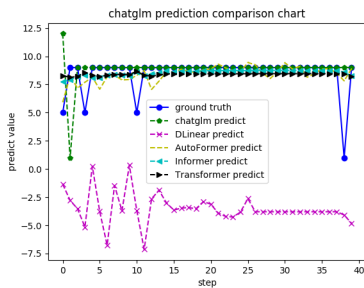


(a) 测试例1

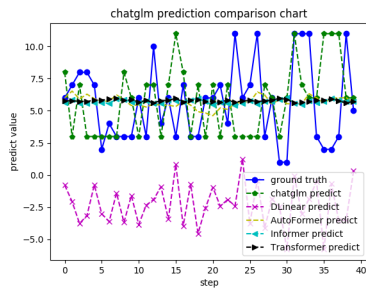


(b) 测试例2

使用一个时间序列模型进行训练

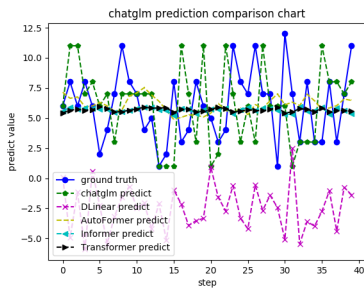


(c) 测试例3

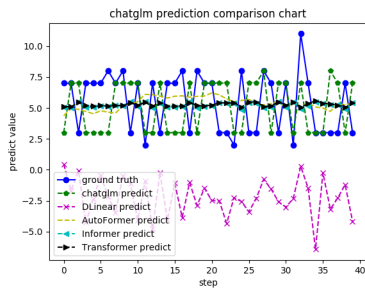


(d) 测试例4

使用一个时间序列模型进行训练



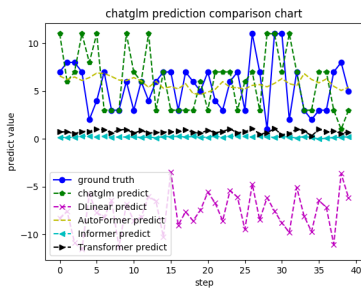
(e) 测试例5



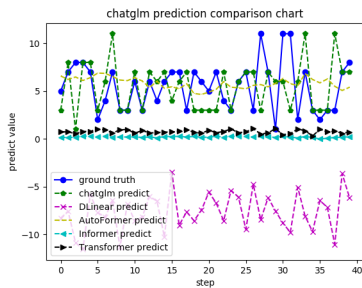
(f) 测试例6



使用多个时间序列模型进行训练

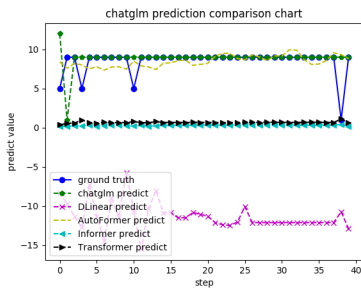


(g) 测试例1

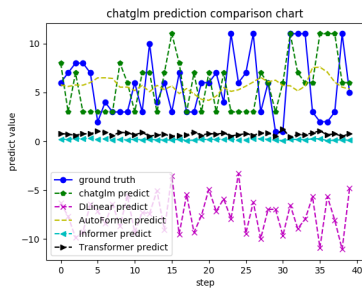


(h) 测试例2

使用多个时间序列模型进行训练

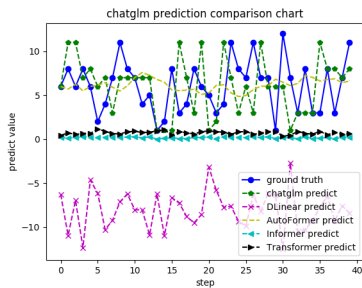


(i) 测试例3

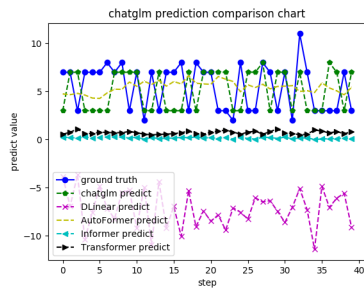


(j) 测试例4

使用多个时间序列模型进行训练



(k) 测试例5



(l) 测试例6

# k线图相关结论

- 经过多种模型测试，语言模型的准确率均是最高的
- 对于分类任务的序列预测，语言模型要优于时间序列模型，而且语言模型还具有强大的泛化能力。

# 谢谢老师和同学们的聆听!