

# 基于多任务学习的金融时间序列预测研究

胡志博<sup>1</sup>, 叶正<sup>1\*</sup>, 葛君<sup>2</sup>

(1 中南民族大学 a. 计算机学院; b. 信息物理融合智能计算国家民委重点实验室, 武汉 430074;  
2 武汉纺织大学外经贸学院, 武汉 437100)

**摘要** 在量化投资领域, 实现连续交易期内投资收益最大化与风险最小化, 是至关重要的任务. 此过程中, 金融时间序列的准确预测扮演着核心角色. 文中创新性地引入多任务学习理念, 并融合信息反馈机制, 极大地丰富了模型功能. 这不仅仅使得模型能够进行多维度的预测, 同时也实现了在提升量化交易收益的同时降低投资风险的双重目标. 研究首先利用实际股票数据, 通过预测一周后的涨跌幅, 将其作为评价分数纳入股票回测评分系统中. 这种方法既能提高对股票走势预测的精确度以提升投资收益, 也有效降低了风险. 进一步地, 通过使用模型预测的未来一周涨跌幅来替代历史数据, 优化了股票评分体系, 从而提高了股票走势的预测准确率. 实验结果表明: 与仅预测单日涨跌幅相比, 应用Transformer、LSTM及IGMTF模型能显著提高预测精度, 并在单位风险下实现收益的显著提升.

**关键词** 时间序列预测; 机器学习; 多任务学习; 股票预测; 特征融合

中图分类号 TP389.1 文献标志码 A 文章编号 1672-4321(2026)03-0392-09

doi: 10.20056/j.cnki.ZNMDZK.20250849

## Research on financial time series prediction based on multi-task learning

HU Zhibo<sup>1</sup>, YE Zheng<sup>1\*</sup>, GE Jun<sup>2</sup>

(1 South-Central Minzu University, a. College of Computer Science; b. Information Physics Fusion Intelligent Computing Key Laboratory of the National Ethnic Affairs Commission, Wuhan 430074, China; 2 College of International Business and Economics, WTU, Wuhan 437100, China)

**Abstract** In the domain of quantitative investment, achieving maximized returns and minimized risks during continuous trading periods is a pivotal task. Accurate forecasting of financial time series plays a crucial role in this process. The concept of multi-task learning is introduced innovatively and an information feedback mechanism is incorporated, significantly enhancing the model's capabilities. The advancement allows the model to not only perform multi-dimensional predictions but also simultaneously increase quantitative trading profits while reducing investment risks. The research initially utilizes actual stock data to predict the price changes a week ahead, incorporating these predictions as a critical factor in the stock backtesting scoring system. The method can improve the accuracy of predicting stock trends to enhance investment returns and effectively reduce risks. Further, by replacing historical data with model-predicted future price changes for a week ahead, the stock scoring system is optimized, thereby improving the accuracy of stock trend predictions. Experimental results demonstrate that compared to predictions based solely on daily price changes, the application of Transformer, LSTM, and IGMTF models significantly enhances prediction accuracy and achieves a noticeable increase in returns per unit of risk.

**Keywords** time series prediction; machine learning; multi-task learning; stock forecast; feature fusion

近年来, 随着机器学习技术的持续成熟与发展, 其在量化投资领域的应用引起了广泛的关注.

收稿日期 2024-03-11

\* 通信作者 叶正(1981-), 男, 教授, 博士, 研究方向: 时间序列预测、自然语言处理, E-mail: yezheng@scuec.edu.cn

基金项目 教育部产学合作协同育人资助项目(202102191002); 中南民族大学引进人才资助项目(YZZ20001); 中央高校科研业务费专项资金资助项目(CZZ24009)

研究界开始深入探讨人工智能在量化投资中的潜能.事实上,人工智能技术已在实际投资策略中得到广泛应用<sup>[1]</sup>.这得益于人工智能在通过模型预测股票市场走势,为投资者提供决策依据方面的能力.

当前,众多机器学习模型如 Transformer<sup>[2]</sup>、LightGBM<sup>[3]</sup>、LSTM<sup>[4-5]</sup>及 IGMTF<sup>[6]</sup>等已被应用于股票市场走势的预测.这些模型各具特色,能够从不同维度辅助投资者进行股票市场分析.然而,在实际股票交易中,最终收益往往依赖于投资者持有的股票组合所产生的对冲效果及损失.尽管机器学习模型能够提供高度精确的市场预测,以构建最大化收益的交易策略,但目前的模型在股票市场趋势预测方面尚存局限.尽管存在这些局限,利用机器学习模型进行量化交易仍被视为一种具有潜力的策略.为了进一步提升机器学习模型在量化交易中的收益效率,本研究旨在探索如何提高机器学习模型在金融时间序列预测上的准确度.

本文旨在不牺牲模型预测性能的前提下,在单一模型框架中实现量化交易收益的最大化与风险的最小化.这一目标提出了一个具有挑战性的问题:在保持模型性能稳定的同时,找到优化时间序列预测结果的方法,以提高投资收益并降低风险.

为应对此挑战,本文引入了多任务学习的策略<sup>[7-9]</sup>,并将其应用于金融时间序列的预测中.通过预测未来一周的股票走势,并将预测的涨跌幅作为影响股票回测结果的关键因素,本研究提出了一种新的方法论.与传统仅依据历史价格走势进行预测的方法相比,本研究的方法不仅考虑即时的市场动态,还将长期趋势变化纳入分析框架.

在实施股票回测时,根据本研究方法的预测结果来制定买入与卖出策略.实验结果显示,此方法不仅能提高对股票走势预测的精确度以提升投资收益,同时也有助于降低投资风险.这为量化投资中的时间序列预测问题提供了一种全新的视角和解决方案.

## 1 相关工作

近年来,Transformer模型凭借其Attention机制在捕捉时间序列数据的点对点关系方面显示出了显著优势,从而在时序预测任务中取得了良好的成效.然而,该模型在时间复杂度和预测精度方面尚

存改进空间.与此同时,LSTM模型成功解决了传统RNN<sup>[10]</sup>模型由于梯度消失问题而只能实现短期记忆的局限,对于股票交易等领域中长期记忆信息的重要性提供了有效的解决方案.在实际的股票交易应用中,LSTM的性能超过了Transformer模型.LightGBM作为GBDT<sup>[11]</sup>模型的优化版本,解决了后者在处理大规模数据时训练速度下降的问题,因而在训练效率和预测准确性方面均表现出色,在股票预测问题上也展现了其优势.

对于Transformer模型的改进并未停止,Informer<sup>[12]</sup>、Autoformer<sup>[13]</sup>、FEDformer<sup>[14]</sup>等模型都在Transformer的基础上针对时间序列问题进行了不同的改进.Informer提出了ProbSparse Self-Attention在论文中替代了传统的Self-Attention,在预测部分,只需一步就可以生成序列的输出,从而避免了误差的传播.Autoformer将Transformer的架构改为深度分解架构,包括内部的序列分解单元、自相关机制以及对应的encoder-decoder.由于预测问题中未来的不可知性,通常先对过去序列进行分解,再分别预测.但这会导致预测结果受限于分解效果,并且忽视了未来各个组分之间的相互作用.Autoformer提出深度分解架构,将序列分解作为Autoformer的一个内部单元,嵌入到encoder-decoder中.在预测过程中,模型交替进行预测结果优化和序列分解,即从隐变量中逐步分离趋势项与周期项,实现渐进式分解.FEDformer将Transformer与季节趋势分解方法相结合,开发了一种频率增强的Transformer.针对长时间序列预测问题,大幅提高了预测精度和模型运行效率.无论是哪种模型,对股票市场的预测能力都是有限的,如果能够使用更好的策略来充分利用模型的预测结果,那么也是可以达到提高最终收益的目的.

在量化交易策略研究领域,最近提出了一种基于深度确定性策略梯度的强化学习框架的新颖投资组合管理策略<sup>[15]</sup>.该策略通过构建两个长短期记忆网络和两个全连接网络,展示了引入深度确定性策略梯度理念的有效性,在不牺牲模型预测性能的前提下,实现了收益最大化和风险最小化的目标.这为量化交易策略的研究提供了新的视角和方法论,对于那些旨在量化投资中获得更高收益和更低风险的投资者与研究人员而言,具有重要的吸引力.本文工作在保持模型预测性能不变的基础上,通过引入多任务学习方法,旨在提升时间序列预测模型在金融时间序列预测上的表现.

## 2 方法

本研究采用以 Transformer 模型为代表的方法流程,如图 1 所示,首先对复权后的历史股票数据进行归一化处理,对于股票在部分交易日没有数据的问题,使用前一个有数据的交易日进行补充.利用 Transformer 模型预测股票未来一天及未来一周后的收盘价,该模型的编码器和解码器均使用六个多头自注意力机制模块(Multi-Head)和前馈网络(Feed-

Forward).通过解码器、线性层(Linear)和 softmax 层得到模型的输出.基于模型的预测结果,通过计算股票未来一天及一周后收盘价的涨跌幅来评估股票的表现.这里的评分标准基于 Qlib<sup>[15]</sup>平台的方法计算.接着,根据一天后和一周后的涨跌幅评分,采用不同的权重对当前股票进行重新评分.最终,依据模拟交易策略,即持有评分最高的 a 支股票和卖出评分最低的 b 支股票,来执行模拟交易并获取交易结果.

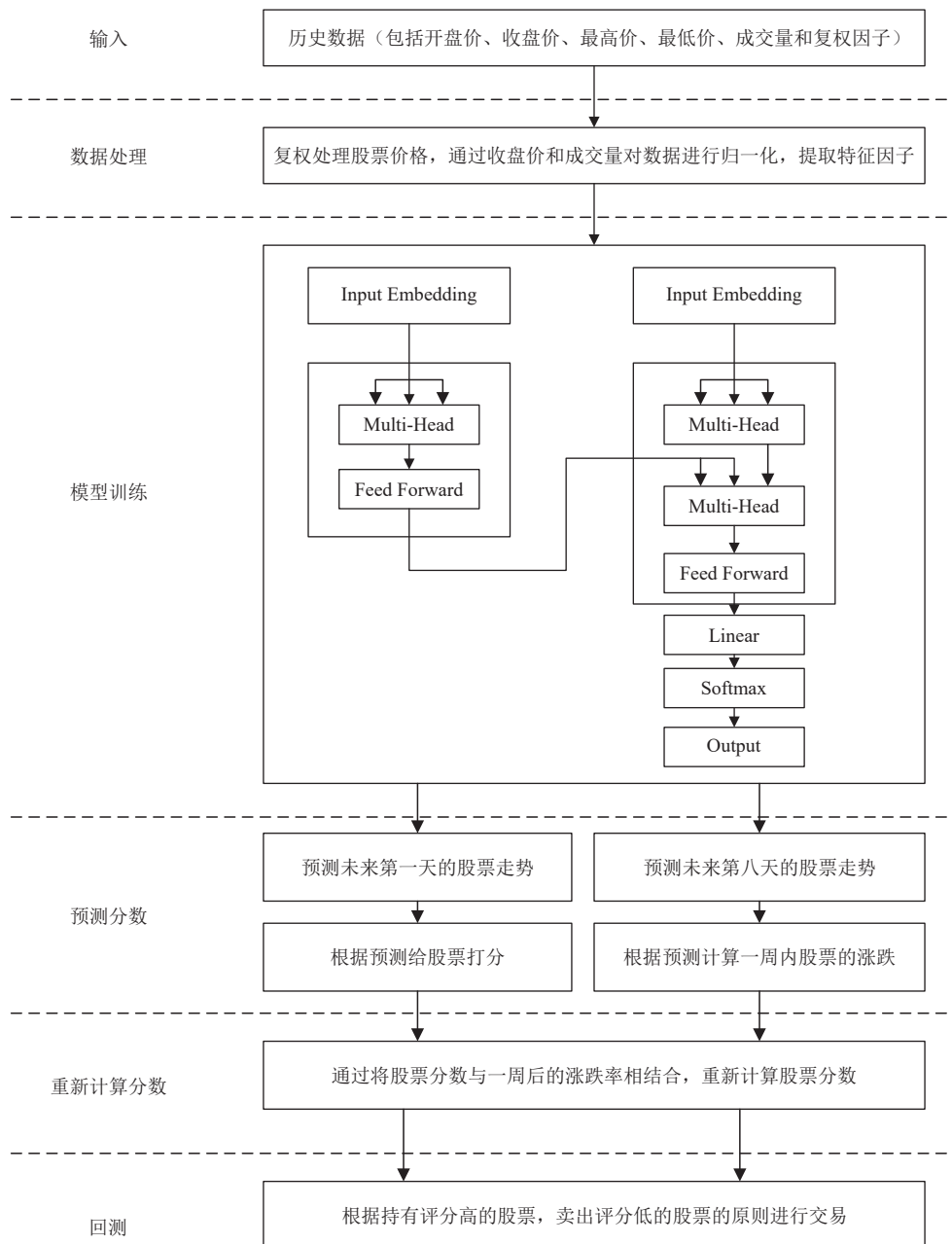


图1 融合多任务学习的量化交易评估示意图

Fig. 1 Schematic diagram of quantitative transaction evaluation integrating multi task learning

### 2.1 多任务学习的股票预测

在图 1 展示的模型基础上,本文也考虑将

Transformer 模型替换为其他模型,在第 3 节的实验中探索不同模型对股票收盘价预测的效果.本研究

引入多任务学习的方法,运用同一模型同时对股票在不同时间点的收盘价进行预测.通过结合一天后和一周后的涨跌幅,重新计算股票评分,旨在提高最终的投资收益.

此方法详细介绍了使用Transformer模型在量化交易中进行股票预测的流程,同时强调了多任务学习在提升预测精度和最终收益中的应用潜力.首先对预处理后的数据进行模型训练,在预测分数部分使用模型预测一天后股票走势并对股票打分,然后预测一周后的涨跌幅,最后重新计算股票分数,在回测中持有评分较高的股票,卖出评分较低的股票.通过这种方法,可以更有效地利用历史数据,提升模型在短期和中期股票收盘价预测上的表现,从而为量化交易策略的制定和优化提供支持.

## 2.2 评价分数的调整

$Week_{score}$ 是根据模型预测一周后的收盘价计算得到的股票涨跌幅,是新增的重要参数. $Week_{score}$ 的计算见公式(1):

$$Week_{score} = \frac{Close_7}{Close_1} - 1. \quad (1)$$

其中, $Close_1$ 是模型预测的未来一天的收盘价, $Close_7$ 是模型预测的未来第七天的收盘价.根据公式(1)可以得到股票一周后的涨跌幅.

本文将模型通过对股票未来一天的预测结果得到的评价分数记为Signal,Signal数据类型是浮点型,取值范围为 $[-1.0, 1.0]$ ,用来评价股票是否值得交易.在得到Signal之后,加入股票在一周后的涨跌幅,得到最终的预测分数,具体的计算见式(2):

$$Pred_{score} = n \times Signal + m \times Week_{score} \quad (n + m = 1, n > 0, m > 0). \quad (2)$$

其中, $Pred_{score}$ 代表股票最终的评分,根据此分数来决定股票的买入卖出.Signal是根据模型进行单任务预测得到的第二天的股票评分, $Week_{score}$ 是根据模型预测一周后的结果,计算得到的股票涨跌幅. $n$ 和 $m$ 是比例系数.通过修改 $n$ 和 $m$ 可以将Signal和 $Week_{score}$ 按照不同的比例进行融合,在本研究的实验设计中,文中细致地选定了1:1、2:1及1:2,三种比例作为Signal和 $Week_{score}$ 融合的具体权重配置.此项配置旨在全面探索两者在综合评分体系中的相对重要性,考虑到不同模型对金融时间序列的短期(未来一天)与中期(未来一周)预测能力存在显著差异.理论上,赋予更高精确度预测结果较大的权重应能促进最终评分的优化,从而提高模型整体的预

测性能.这种权重分配方法不仅体现了对模型预测能力差异的认识,也为精细化调整预测框架提供了一种策略,以期达到更为准确的股票市场预测结果.

选择未来一周而非未来更长的时间点作为预测目标,是基于股票市场的高度动态性、短期预测在交易策略中的适应性高、模型预测准确率随预测期增长而降低、以及市场反馈及时性和风险管理的有效性等因素.短期预测如一周内的收盘价能够更准确地反映近期市场变化的影响,同时为量化交易策略提供灵活性和高适应性,允许投资者快速响应市场波动.此外,这种预测周期还有助于提高模型性能的迭代速度并优化风险管理策略,从而在保持预测质量的同时,增加投资回报率.

## 3 实验设计与分析

本研究的实验数据源自Qlib库提供的雅虎财经关于中国A股市场的日频数据.所选数据涵盖了从2008年1月2日至2020年7月22日间,曾经被纳入沪深300指数的股票,包括每支股票的开盘价、收盘价、最高价、最低价及交易量等信息.本实验旨在验证所提出方法的可行性与有效性.

### 3.1 实证数据验证

#### 3.1.1 数据集划分

实验使用的数据集为:选取2008年1月2日到2014年12月31日的日频数据作为训练集,2015年1月1日到2016年12月31日的日频数据作为验证集,2017年1月1日到2020年7月22日的日频数据作为测试集.作为性能基准,本研究采用沪深300指数(SH000300).对于非交易日,实验采用前一交易日的数据进行填补,以保证数据连续性.

#### 3.1.2 消融实验

(1) 利用Qlib在训练及验证集上训练Transformer模型;

(2) 应用策略,持有评分最高的50支股票、卖出评分最低的5支股票,并进行模型回测,记录结果;

(3) 在已有测试集上,依据公式(1)计算实际一周后的股票涨跌幅;

(4) 应用公式(2),设置 $n:m$ 为1:1比例,重新计算股票评分,影响交易决策,并记录回测结果.

#### 3.1.3 实验参数

Transformer模型的批处理大小设置为2048,特

征维度为20,模型维度为64,多头注意力机制的头数为2,模型层数为2,epoch为100,学习率为0.0001,使用Adam优化器,迭代没有提升提前停止的条件为5,损失函数的计算为MSE.

### 3.1.4 实验结果与分析

本文图2、图3、图4中的图例所示的 cum bench 是基准(SH000300)的累计值, cum return wo cost 是不考虑手续费的累计收益, cum return w cost 是考虑手续费的累计收益, return wo mdd 是不考虑手续费时的最大回撤, return w cost mdd 是考虑手续费时的最大回撤, cum ex return wo cost 是不考虑手续费时的超额收益, cum ex return w cost 是考虑手续费时的超额收益, turnover 是换手率, cum ex return w cost mdd 是考虑手续费时超额收益最大回撤, cum ex return wo cost mdd 是不考虑手续费时超额收益最大回撤.

图2展示了仅使用Transformer模型预测结果进行回测的收益变化.具体来说,在2018年1月29日至2018年10月12日这段时间,在Qlib量化交易平台上的股票回测的累计收益在下降,最大回撤在增大.在2019年1月2日至2020年6月30日这段时间,超额收益在下降.

图3通过加入实际数据一周后的涨跌幅进行的回测结果表明,收益呈现稳定上升趋势,显示出加入未来实际数据的有效性.因为在评分部分加入了未来实际数据进行计算涨跌幅,所以这里的收益是保持上升的.这个结果也说明通过加入实际的涨跌幅可以提高股票回测中的收益,同时降低风险.

图4是使用Transformer模型,设置股票价格评

分和股票一周后涨跌幅的评分为1:1的实验结果.从图中可以看出,在2017年4月21日至2018年1月29日这段时间,股票回测的超额收益在下降.在2018年1月17日至2018年10月12日这段时间,股票回测的累计收益在下降,最大回撤在增大.

虽然在修改评价分数后,依旧会有收益下降,风险上升的时间段,但是对比图2和图4的结果可以发现,修改后的收益下降和风险上升的总时长减少了约24%.这一结果说明对于同一模型,预测性能不变的情况下,加入股票一周后涨跌幅的评价分数,可以延长股票回测中收益上升的时间段,通过降低风险上升的时间段,也可以有效降低量化交易中的风险.

### 3.1.5 结论

实验结果验证了将一周后的涨跌幅作为评价因子对提高策略收益的正面影响.首先,通过Transformer模型预测未来一天股票价格的评分进行股票回测,和结合实际数据计算出的涨跌幅评分进行股票回测.图3的结果表明,实际数据的涨跌幅加入后,收益提高较为显著,说明加入股票一周后涨跌幅的方法是可行的.其次,设置股票价格评分和股票一周后涨跌幅的评分为1:1,对比图2中的结果,收益下降和风险上升的总时长减少了约24%.通过两组对比可以看出,通过加入一周后的涨跌幅作为评价因子,对于提高股票回测中的收益是有用的.

后续实验将进一步探讨使用模型预测的一周后涨跌幅的有效性,以加强模型预测能力和交易策略的优化.

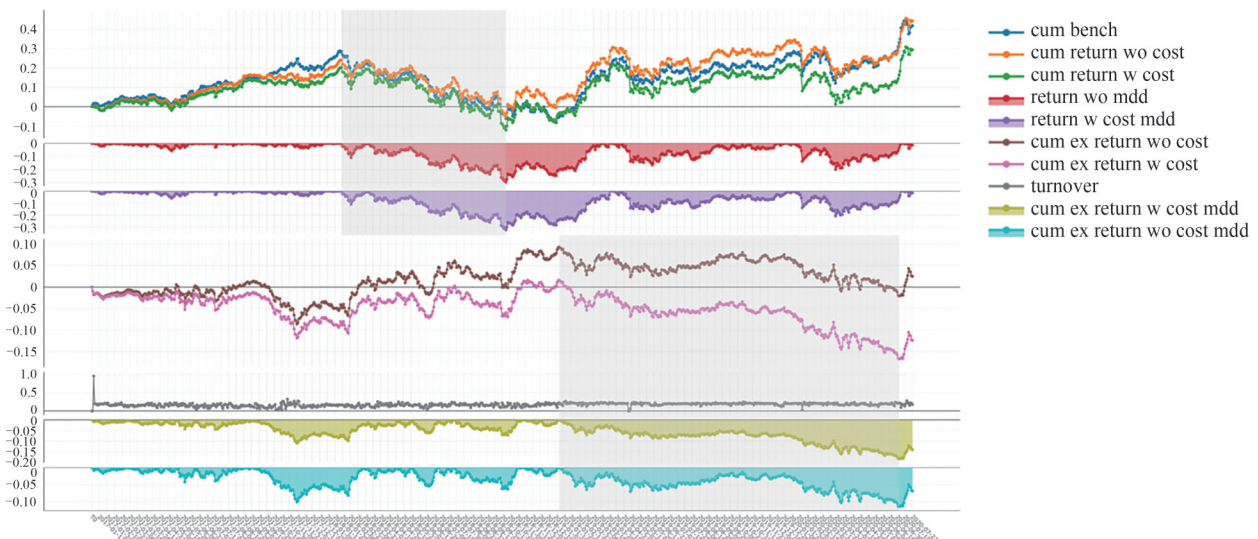


图2 Transformer模型的实验结果  
Fig. 2 Experimental results of Transformer model

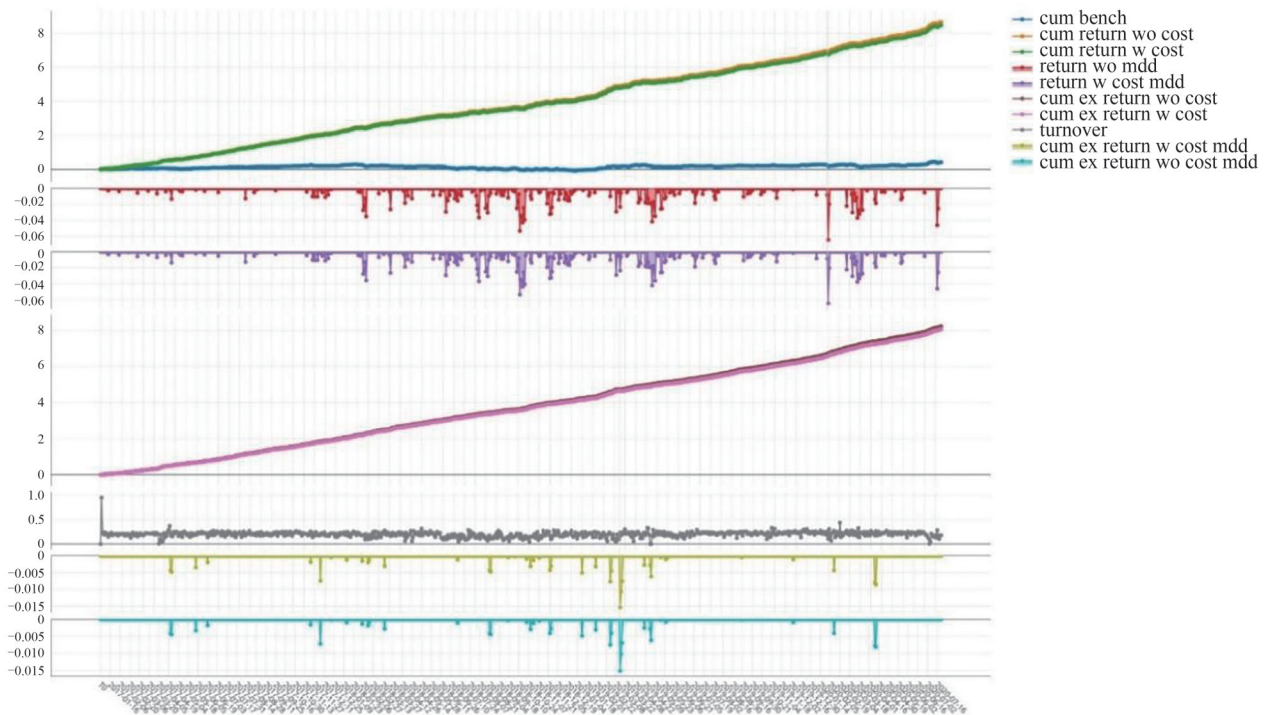


图 3 Transformer 模型设置  $n:m=1:1$  的实验结果(使用实际数据计算一周后涨跌幅)

Fig. 3 Experimental results of Transformer model with  $n:m=1:1$  (using actual data to calculate the rise and fall in a week)

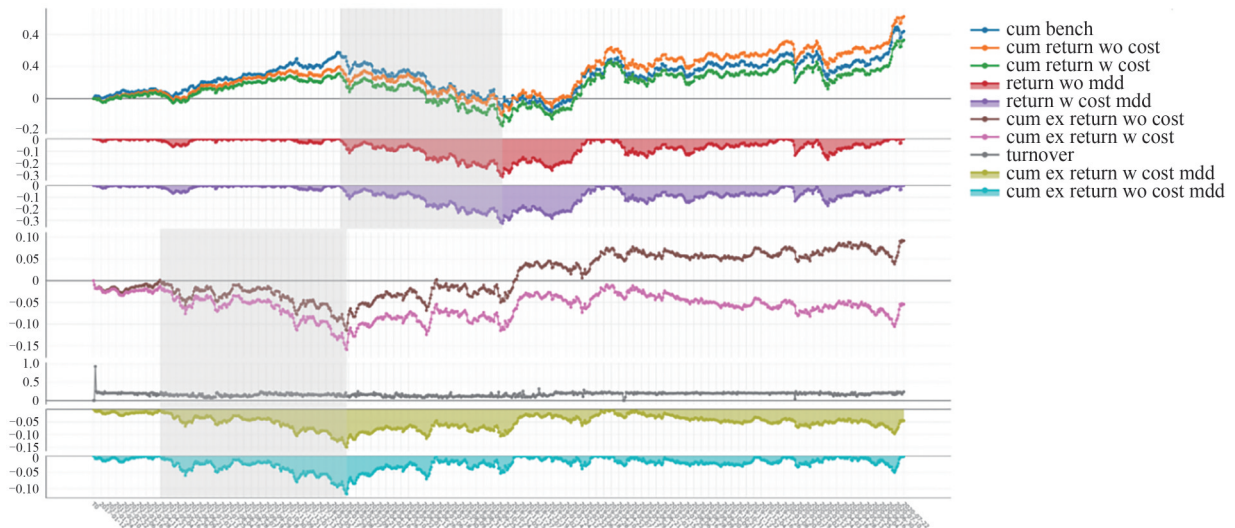


图 4 Transformer 模型设置  $n:m=1:1$  的实验结果(使用模型预测数据计算一周后涨跌幅)

Fig. 4 Experimental results of Transformer model with  $n:m=1:1$  (using model data to calculate the rise and fall in a week)

### 3.2 模型比较及评价指标融合的影响分析

本部分旨在深入探索不同模型在量化投资策略中的应用,并分析综合使用未来一天评分与一周后涨跌幅两种评价指标对投资策略效果的影响.实验数据沿用 3.1 部分设置,包括 2008 年 1 月 2 日至 2020 年 7 月 22 日的中国 A 股日频数据,并将沪深 300 指数(SH000300)作为性能基准.为确保数据连续性,非交易日数据以前一交易日数据进行填补.

#### 3.2.1 实验设计与方法

相较于 3.1 部分,本节实验设计引入 LSTM 和 IGMF 模型,与 Transformer 模型并行分析,以覆盖不同的预测技术及其对交易策略的潜在影响.通过设计以下消融实验,本研究旨在系统评估一周后涨跌幅这一评价指标的引入及其与未来一天评分的融合对模型性能的影响:

- (1) 仅使用未来一天的预测评分进行交易决策,不考虑一周后涨跌幅;

(2) 仅基于一周后涨跌幅进行交易决策,不使用未来一天的评分;

(3) 将未来一天的评分与一周后涨跌幅按照不同比例融合,进行交易决策.

在每个实验中,分别采用Transformer、LSTM、和IGMTF模型进行预测,并应用持有评分最高的50支股票、卖出评分最低的5支股票的策略进行模型回测.

### 3.2.2 实验参数

Transformer模型的参数设置同3.1.3节.LSTM模型的批处理大小为2000,特征维度为6,每个隐藏层的单元特征数量设置为64,隐藏层设置为2,epoch设置为200,使用Adam优化器,初始学习率设置为0.001,迭代没有提升提前停止的条件为20,损失函数的计算为MSE.IGMTF模型的特征维度为6,每个隐藏层的单元特征数量设置为64,隐藏层设置

为4,epoch设置为200,使用Adam优化器,初始学习率设置为0.0005,迭代没有提升提前停止的条件为20,损失函数的计算为MSE.

### 3.2.3 实验结果与分析

图2、图4和表1、表2、展示了详细的Transformer模型的回测结果和Transformer设置 $n:m=1:1$ 的回测结果.在加入了一周后股票的涨跌幅之后,同样的模型和策略进行回测,最终的账户资金也提高了8.2%,收益明显增加,从图中也能看到无论是否考虑手续费,最大回撤也至少有34.9%的降低.本文将表3、表4、表5的最佳数据做了加粗标记,把拥有最多标记的结果作为实验的最佳结果.在有成本的情况下,LSTM、IGMTF的最佳结果中信息比率分别提高了12.6%和101.2%;在无成本的情况下,LSTM、IGMTF的最佳结果中年化收益分别提高了10.7%和77.2%.

表1 Transformer模型的实验结果

Tab. 1 Experimental results of Transformer model

日期	账户资金	收益率	成交额	换手率	持仓市值	现金	基准收益率
2017-01-03	1.000000e+08	0.000000	0.000000e+00	0.000000	0.000000e+00	1.000000e+08	0.009713
2017-01-04	9.995250e+07	0.000000	9.500000e+07	0.950000	9.500000e+07	4.952500e+06	0.007803
2017-01-05	1.000167e+08	0.000816	1.144571e+08	0.194663	9.939706e+07	6.196524e+05	-0.000154
...	...	...	...	...	...	...	...
2020-07-20	1.260039e+08	0.028097	1.592207e+10	0.220553	1.225623e+08	3.441595e+05	0.029837
2020-07-21	1.251647e+08	-0.006458	1.594883e+10	0.212387	1.243981e+08	7.666256e+05	0.002295
2020-07-22	1.256161e+08	0.003780	1.597071e+10	0.174796	1.250414e+08	5.747088e+05	0.004990

表2 Transformer模型设置 $n:m=1:1$ 的实验结果

Tab. 2 Experimental results of Transformer model setting  $n:m=1:1$

日期	账户资金	收益率	成交额	换手率	持仓市值	现金	基准收益率
2017-01-03	1.000000e+08	0.000000	0.000000e+00	0.000000	0.000000e+00	1.000000e+08	0.009713
2017-01-04	9.995345e+07	0.000000	9.310000e+07	0.931000	9.310000e+07	6.853450e+06	0.007803
2017-01-05	9.993686e+07	0.000053	1.180003e+08	0.249119	9.913076e+07	8.061040e+05	-0.000154
...	...	...	...	...	...	...	...
2020-07-20	1.351345e+08	0.035119	1.567773e+10	0.172752	1.345444e+08	5.901131e+05	0.029837
2020-07-21	1.351459e+08	0.000309	1.570794e+10	0.223594	1.343641e+08	7.817966e+05	0.002295
2020-07-22	1.359088e+08	0.005882	1.573996e+10	0.236930	1.350763e+08	8.324747e+05	0.004990

通过比较不同实验设置下的模型表现,本研究揭示了评价指标融合对于提升量化交易策略性能的重要性.实验结果显示,在引入一周后涨跌幅作为新的评价因子后,模型在相同策略下实现了显著的收益提升和最大回撤的降低,证明了综合使用两种评价指标的有效性.

### 3.2.4 实验结论

实验数据和回测结果表明,在基于单一未来一天

股票评分的预测基础上引入一周后股票涨跌幅这一评价指标,能够在不改变原有模型架构和回测策略的前提下,实现收益的提升和风险的降低.这一发现突显了多任务学习思想在量化交易领域的应用价值,证明了通过融入额外的时间序列预测指标,可以显著增强模型的预测能力和决策的有效性.

进一步分析显示,依赖单一评价指标(未来一天评分或一周后涨跌幅)作出的交易决策的表现,均

表 3 Transformer模型及加入涨跌幅后的实验结果

Tab. 3 Transformer model and experimental results after adding fluctuation

Transformer	评价标准	不加入一周后的涨跌幅	只预测一周后的涨跌幅	$n:m=1:1$	$n:m=2:1$	$n:m=1:2$
无成本的超额回报	标准偏差	<b>0.005311</b>	0.005506	0.005476	0.005401	0.005528
	年化收益	0.006855	0.070178	0.066341	<b>0.077908</b>	0.066234
	信息比率	0.083659	0.826130	0.785258	<b>0.934979</b>	0.776606
	最大撤回	-0.113542	-0.092612	-0.080462	<b>-0.073823</b>	-0.094325
有成本的超额回报	标准偏差	<b>0.005312</b>	0.005508	0.005476	0.005402	0.005530
	年化收益	0.034174	0.039700	0.030562	<b>0.040862</b>	0.032145
	信息比率	0.417047	0.467171	0.361739	<b>0.490314</b>	0.376799
	最大撤回	-0.183292	-0.096776	-0.088425	<b>-0.087957</b>	-0.098191

表 4 LSTM模型及加入涨跌幅后的实验结果

Tab. 4 LSTM model and experimental results after adding fluctuation range

LSTM	评价标准	不加入一周后的涨跌幅	只预测一周后的涨跌幅	$n:m=1:1$	$n:m=2:1$	$n:m=1:2$
无成本的超额回报	标准偏差	0.004503	0.004160	0.004189	<b>0.004794</b>	0.004045
	年化收益	0.124200	0.100854	0.079926	0.123266	<b>0.124858</b>
	信息比率	1.797325	1.571618	1.236709	1.666688	<b>1.990159</b>
	最大撤回	-0.058947	-0.061564	<b>-0.055532</b>	-0.067166	-0.05724
有成本的超额回报	标准偏差	0.004499	0.004162	0.004192	<b>0.004796</b>	0.004045
	年化收益	0.082169	0.063402	0.042666	<b>0.086800</b>	0.083156
	信息比率	1.183766	0.987376	0.659692	1.173113	<b>1.332557</b>
	最大撤回	-0.067378	-0.075772	-0.082117	-0.071659	<b>-0.068212</b>

表 5 IGMTF模型及加入涨跌幅后的实验结果

Tab. 5 IGMTF model and experimental results after adding fluctuation

IGMTF	评价标准	不加入一周后的涨跌幅	只预测一周后的涨跌幅	$n:m=1:1$	$n:m=2:1$	$n:m=1:2$
无成本的超额回报	标准偏差	0.004996	0.004063	0.004103	0.004626	<b>0.004042</b>
	年化收益	0.108387	0.137952	0.076920	0.112498	<b>0.156226</b>
	信息比率	1.406396	2.212310	1.215190	1.576251	<b>2.492189</b>
	最大撤回	-0.066952	<b>-0.042195</b>	-0.056012	-0.047031	-0.050141
有成本的超额回报	标准偏差	0.004997	0.004065	0.004104	0.004624	<b>0.004045</b>
	年化收益	0.070142	0.102181	0.042394	0.076705	<b>0.114866</b>
	信息比率	0.909948	1.637494	0.669664	1.075239	<b>1.831446</b>
	最大撤回	-0.075482	-0.053453	-0.071603	-0.062715	<b>-0.056514</b>

不如两者综合应用时的结果.这一结论强调了综合应用多个影响因子在提高交易策略效率方面的重要性,揭示了实验性能提升并非仅由单一影响因子驱动,而是多指标相互作用的结果.

在对比不同模型性能时,通过调整未来一天评分与一周后涨跌幅影响因子的权重比例( $n:m$ ),实验结果表明可以实现不同程度的性能优化.本研究选取了 $1:1$ 、 $2:1$ 、 $1:2$ 三种权重比例进行测试,发现对于Transformer模型而言, $2:1$ 的配比能够最大化收益率;对于LSTM和IGMTF模型, $1:2$ 的配比展现了最佳的收益率.这一发现指出,为了充分发挥模型在特定量化交易环境中的潜力,对评价指标权重的细致调整是必要的.不同模型根据其数据的处

理能力和特征提取的差异,对评价指标的权重比例有着不同的最优需求.

## 4 总结

本研究旨在探索不改变模型预测性能的前提下,如何实现交易收益的最大化与风险的最小化.研究提出了一种创新策略,即在股票走势预测中引入一周后涨跌幅,以调整当前对股票的评分.该策略未涉及重新训练模型,但在量化交易实验中显著提升了结果.通过在Qlib平台上训练的Transformer、LSTM、IGMTF等模型,本研究发现,执行持有评分最高的50支股票及卖出评分最低的5支股票策略时,

引入一周后涨跌幅能够有效提升单位风险下的收益.

进一步的消融实验表明,仅通过预测一周后的收盘价并不能达到最优收益.然而,当将未来一天的评分与一周后的涨跌幅按照 1:1、2:1、1:2 的比例融合时,不同模型在各比例下的表现各异,但最优结果的信息比率均实现了显著提升.这一发现证实,基于多任务学习方法的金融时间序列预测能够在不改变模型结构的基础上,有效地提升交易策略的终极收益.

本研究强调,即便在模型和策略方法遇到性能瓶颈时,也可以通过融入多任务学习思想和引入不同的评价因子来提高最终收益.在所有时间序列预测模型中都存在预测误差,但本研究展示了在相同预测精度下,通过整合未来不同时间节点的影响因子,可以优化预测结果的经济收益.实验结果验证了该方法在金融时间序列预测中的有效性.

未来工作可以进一步探讨预测股票未来一天走势与未来一周变化之间更深入的关联性,以及探索未来一周是否为最优预测时间节点,或针对不同股票市场寻找更合适的影响因子,这些都是值得深入研究挑战.这不仅有助于深化对股票市场动态变化的理解,也能为投资者提供更高收益的投资策略.

#### 参 考 文 献

- [1] CHEN X, GUO H. A futures quantitative trading strategy based on a deep reinforcement learning algorithm [C]// 2023 IEEE 8th International Conference on Big Data Analytics (ICBDA). Harbin: IEEE, 2023: 175-179.
- [2] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems. Long Beach: NIPS, 2017: 5998-6008.
- [3] KE G, MENG Q, FINLEY T, et al. LightGBM: A highly efficient gradient boosting decision tree [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: ACM, 2017: 3149-3157.
- [4] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural computation*, 1997, 9(8): 1735-1780.
- [5] LI J, QI Z, LI Z, et al. Quantitative trading decision model based on LSTM neural network [J]. *Highlights in Business, Economics and Management*, 2023, 5: 583-591.
- [6] XU W, LIU W, BIAN J, et al. Instance-wise graph-based framework for multivariate time series forecasting[EB/OL]. 2021: 2109.06489.<http://arxiv.org/abs/2109.06489v1>.
- [7] CARUANA R. Multitask learning[J]. *Machine Learning*, 1997, 28(1): 41-75.
- [8] ZHANG Y, YANG Q. An overview of multi-task learning[J]. *National Science Review*, 2018, 5(1): 30-43.
- [9] VANDENHENDE S, GEORGIOULIS S, VAN GANSBEKE W, et al. Multi-task learning for dense prediction tasks: A survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(7): 3614-3633.
- [10] WILLIAMS R J, ZIPSER D. A learning algorithm for continually running fully recurrent neural networks [J]. *Neural Computation*, 1989, 1(2): 270-280.
- [11] FRIEDMAN J H. Greedy function approximation: A gradient boosting machine[J]. *The Annals of Statistics*, 2001, 29(5): 1189-1232.
- [12] ZHOU H, ZHANG S, PENG J, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(12): 11106-11115.
- [13] WU H, XU J, WANG J, et al. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting[EB/OL]. 2021: 2106.13008.<http://arxiv.org/abs/2106.13008v5>.
- [14] ZHOU T, MA Z, WEN Q, et al. FEDformer: Frequency enhanced decomposed transformer for long-term series forecasting[EB/OL]. 2022: 2201.12740.<http://arxiv.org/abs/2201.12740v3>.
- [15] ZHANG H, JIANG Z, SU J. A deep deterministic policy gradient-based strategy for stocks portfolio management [C]// 2021 IEEE 6th International Conference on Big Data Analytics (ICBDA). Xiamen: IEEE, 2021: 230-238.
- [16] YANG X, LIU W, ZHOU D, et al. Qlib: An AI-oriented quantitative investment platform [EB/OL]. 2020: 2009.11189.<http://arxiv.org/abs/2009.11189v1>.

(责编&校对 雷建云)