

分类号 C8
UDC 311

学校代码 10590
密 级 公开

深圳大学硕士学位论文

基于 LSTM 模型的日内动量策略研究 ——以沪深 300 股指期货为例

陈 壮

| | |
|---------|----------|
| 学 位 类 别 | 应用统计专业学位 |
| 专 业 名 称 | 应用统计硕士 |
| 学院（系 所） | 经济学院 |
| 指 导 教 师 | 李猛教授 |

摘要

股票市场作为一国经济的“晴雨表”，一直以来都是热门的研究对象。沪深 300 指数反映了中国 A 股市场上股票价格的整体表现。而将沪深 300 股指期货作为交易标的的沪深 300 股指期货是中国金融期货市场上最活跃的期货合约之一。它流动性强，可以进行大额交易，既是一个重要的价格发现工具，也是一个为企业和机构投资者提供保值和风险管理工具。中国 A 股市场实行 T+1 交易制度，购买的股票次日方可卖出；而沪深 300 股指期货则实行 T+0 交易制度，即可当日购买当日卖出。因此，选择沪深 300 股指期货作为研究对象更具有实际意义。

动量效应作为少有的被公开后仍然具有活力的市场异象，为投资者提供了一种相对较为稳健的投资策略，适用于各种不同的市场和投资品种；研究动量效应可以帮助人们更好地理解市场的行为和机制。因此本文聚焦于沪深 300 股指期货日内动量效应的研究。

本研究分析了沪深 300 股指期货的日内动量效应，将一个交易日分为 8 个半小时，分别计算每个半小时的收益率并加上隔夜收益，样本期间为 2010 年 4 月 16 日至 2022 年 12 月 16 日。结果表明，隔夜收益以及倒数第二个半小时的收益率对最后一个半小时的收益率有显著正向的预测作用。对此，本文给出了可能的解释：投资者非理性行为、不频繁的资产组合调整和“后知后觉”的投资者的存在。另外，在熊市、高成交量或高波动率的市场条件下，日内动量效应更加显著。本文还检验了沪深 300 指数的日内动量效应，最终结果一致，也证明了日内动量效应的稳健性。接着分别以隔夜收益、倒数第二个半小时收益以及两者结合起来作为做多或做空信号，建立了三个简单的日内动量策略，年化收益分别为 5.115%，6.727%，5.861%，相比于简单的买入并持有年化收益有很高的提升。

最后，本文使用半小时收益、波动率、成交量和市场状况等相关指标，利用 XGBoost 和 LSTM 对每日最后半小时涨跌进行预测，进一步优化策略，最终年化收益分别为 6.336% 和 9.136%，夏普比率也达到了 1.507 和 2.164。最终策略收益和风险最佳的是基于 LSTM 的日内动量策略，年化收益 9.136%，最大回撤 4.935%，夏普比率 2.164，在高收益的同时具有更低的风险。

关键词： 沪深 300 股指期货；日内动量效应；XGBoost；LSTM

Abstract

As an "economic barometer" of a country's economy, the stock market has always been a popular research subject. The CSI 300 Index reflects the overall performance of the listed stock prices in China's A-share market. The CSI 300 Index Futures, which uses the CSI 300 Index as its trading target, is one of the most active futures contracts in China's financial futures market. It has strong liquidity and can carry out large-scale transactions, serving as an important price discovery tool and a hedging and risk management tool for enterprises and institutional investors.

China's A-share market implements a T+1 trading system, meaning that purchased stocks can only be sold on the next trading day. In contrast, the CSI 300 Index Futures adopts a T+0 trading system, meaning that it can be bought and sold on the same day. Therefore, choosing CSI 300 Index Futures as a research subject is more practical.

Momentum effect, as one of the few market anomalies that remains active after being made public, provides investors with a relatively stable investment strategy applicable to various markets and investment products. Studying momentum effect can help people better understand market behavior and mechanisms. Therefore, this article focuses on the study of intraday momentum effect of CSI 300 Index Futures.

This study analyzes the intraday momentum effect of CSI 300 Index Futures during the sample period from April 16, 2010 to December 16, 2022. The research method is to divide a trading day into eight half-hour periods, calculate the return of each half-hour period and add the overnight return. The results show that the overnight return and the return of the second-to-last half-hour period have a significant positive predictive effect on the return of the last half-hour period. This article provides possible explanations for this phenomenon, such as irrational investor behavior, infrequent asset portfolio adjustments, and the existence of "hindsight bias". In addition, the intraday momentum effect is more significant in bear markets, high trading volume or high volatility market conditions. This study also tests the intraday momentum effect of the CSI 300 Index, and the final results are consistent, which also proves the robustness of the intraday momentum effect.

Three simple intraday momentum strategies were established based on the overnight return, the return of the second-to-last half-hour period, or a combination of both as buy or sell signals, with annualized returns of 5.115%, 6.727%, and 5.861%, respectively. These strategies significantly outperformed the simple buy-and-hold strategy.

Finally, this article uses relevant indicators such as half-hour returns, volatility, trading volume, and market conditions, and uses XGBoost and LSTM to predict the rise and fall of the last half-hour period of each day, further optimizing the strategy. The final annualized returns were 6.336% and 9.136%, and the Sharpe ratios reached 1.507 and 2.164, respectively. The strategy with the best returns and risks was the intraday momentum strategy based on LSTM, with an annualized return of 9.136%, a maximum drawdown of 4.935%, and a Sharpe ratio of 2.164, which had lower risk while achieving higher returns.

Keywords: CSI 300 Index Futures Contract; Intraday Momentum Effect; XGBoost; LSTM

目 录

| | |
|------------------------|----|
| 摘要 | I |
| Abstract | II |
| 第一章 绪论 | 1 |
| 第一节 研究背景 | 1 |
| 第二节 研究意义与目的 | 2 |
| 第三节 创新点 | 3 |
| 第四节 文献综述 | 3 |
| 一、横截面动量 | 3 |
| 二、时间序列动量 | 6 |
| 三、深度学习在动量策略中的应用 | 7 |
| 第五节 本文结构与框架 | 8 |
| 第二章 相关算法理论 | 10 |
| 第一节 XGBoost 模型 | 10 |
| 第二节 长短期记忆神经网络 | 11 |
| 第三节 贝叶斯优化算法 | 13 |
| 第四节 神经网络训练相关概念 | 14 |
| 一、过拟合 | 14 |
| 二、损失函数 | 15 |
| 三、Adam 优化算法 | 16 |
| 第五节 模型评估方法 | 17 |
| 第三章 日内动量效应实证分析 | 19 |
| 第一节 数据来源 | 19 |
| 第二节 日内动量效应的存在性检验 | 19 |
| 第三节 不同指标下的日内动量效应 | 20 |
| 一、牛市与熊市下的日内动量效应 | 20 |
| 二、不同成交量下的动量效应 | 22 |
| 三、不同波动率下的日内动量效应 | 24 |

| | |
|----------------------------|----|
| 第四节 日内动量策略构建 | 25 |
| 第五节 现象解释 | 29 |
| 第六节 稳健性检验 | 31 |
| 第四章 深度学习动量策略实证分析 | 33 |
| 第一节 特征选择 | 33 |
| 第二节 基于 XGBoost 的动量策略 | 34 |
| 一、贝叶斯超参数优化 | 34 |
| 二、模型评估 | 35 |
| 第三节 基于 LSTM 的动量策略 | 37 |
| 一、数据选取与处理 | 37 |
| 二、贝叶斯超参数优化 | 37 |
| 三、模型评估 | 38 |
| 第四节 本章小结 | 39 |
| 第五章 总结与展望 | 41 |
| 第一节 研究总结 | 41 |
| 第二节 研究不足与展望 | 42 |
| 参考文献 | 43 |
| 致谢 | 47 |
| 指导教师对研究生学位论文的学术评语 | 48 |
| 学位论文答辩委员会决议书 | 49 |

第一章 绪论

第一节 研究背景

作为我国实体经济的重要支撑，中国的股票市场已经建立了 33 年。A 股市场成立于 20 世纪 90 年代初，2014 年 11 月我国股市成为全球第二大股东市，截至 2021 年 12 月 13 日中国 A 股共有 4662 家上市公司，总市值约为 92 万亿。市值总量居新兴市场之首，中国资本市场已成为全球金融市场中不可忽视的一员。股票市场反映国民经济运行状况，对其研究一直受到关注，而市场异象更是研究的重点。

市场异象一般被看作是市场定价效率不足的表现。而动量效应（Momentum Effect）也称为惯性效应，是一种被广泛研究和实践的市场异象，即过去表现良好的股票将继续表现良好。这种效应被称为动量效应或惯性效应，是市场中生命力最强的异象之一，Fama 和 French（2008）称其为“首要异象”。

在股票投资领域，策略通常是指一种特定的投资方法或计划，用于指导买入和卖出股票的决策。股票策略的目的是帮助投资者实现其投资目标，如获得高回报、降低风险或保值增值等。基于动量效应构建的套利组合能够获得显著而稳健的超额收益，这种策略被称为动量交易策略或惯性策略。自被披露以来，动量效应就一直是股票市场中的研究热点和投资实践活动中的热点策略。它是少有的被公开后仍然具有活力的市场异象之一（Schwert, 2003）。金融科技的发展促进了金融市场研究的兴起，高频交易日益普及，因此投资者开始关注日内动量效应。在美国标准普尔 500ETF 中，Gao 等（2018）首先探究了日内收益的可预测性，发现了日内每半小时收益率之间的关系。

本研究的出发点在于，动量效应作为金融市场中“首要异象”，在许多市场中普遍存在，但在中国 A 股市场中，反转效应却更加常见，而在中长期（月度、年度）内动量效应一般不存在，或仅在特定的条件下存在（田利辉等，2014）。最近的一些文献引起了对 A 股市场动量效应话题的再次关注，他们发现在 A 股市场中存在着短期动量效应（唐也然等，2020；曲荣华和刘扬，2020），而对于形成期在一日以内的动量效应的研究还较少。

其次，目前几乎所有研究都使用股票指数，来检验中国是否存在日内动量。虽然股票指数 ETF 是可交易的，但在中国由于 T+1 交易制度，T 日买入的资产不能在同一天出售。这限制了日内动量效应的实际运用。测试日内动量的经济意义需要在最后一个交易节点开始时买入（或卖出）资产，并在最后一个交易节点结束时平仓。非流通股指数

不符合这一要求，其对投资者的经济价值也不明确。股指期货是唯一符合上述所有标准的金融工具，因此本文将沪深 300 股指期货作为研究对象。

再次，中国股指期货市场的交易周期与美国相应市场的交易周期不同。2015 年 12 月 31 日之前，交易时段为 9:15 至 15:15，午休时间为 11:30 至 13:00。2016 年 1 月 1 日后交易时间为 9:30 至 15:00，11:30 至 13:00 为午休时间。这种差异可能会对日内时间序列动量产生影响，因此本文实证检验中国沪深 300 股指期货市场日内动量效应的存在性。

第二节 研究意义与目的

随着国民经济的发展，我国也进入了新阶段，资本市场将进一步加强市场化（国务院金融稳定发展委员会，2020），股票市场的有效运作需要有准确的市场信号。沪深 300 股指期货作为沪深 300 指数的衍生品，有着价格发现的功能。因此，研究股指期货市场的动量效应在理论和实践层面上具有重要意义。

理论价值方面，动量效应是资产定价中的一个重要因素，它描述了股票等金融资产价格在短期内出现的惯性趋势。尽管动量效应在股票市场中得到了广泛的应用和认可，但其在股指期货市场中的应用却饱受争议。通过对股指期货市场中的动量效应进行深入研究和分析，可以更好地理解和解释市场中的价格波动，因此，本研究的理论延伸与应用可以帮助重新审视动量效应在股指期货市场中的应用。

从投资者的角度而言，研究超短期动量效应有利于改进其资产配置方式，并协助市场参与者更好地理解股指期货市场的特征，从而更好地规避风险，实现更好的投资回报。随着金融科技的发展，投资者的交易频率日益提高，因此研究超短期动量效应更具有实际意义。根据金戈（2019）的研究，超短期动量效应可以用于构建日内套利策略，从而实现更高的投资收益。当然，要正确应用日内动量效应，需要投资者具备丰富的市场经验和较高的交易技能。

对于监管机构来说，研究动量效应具有重要意义。首先，了解和掌握市场的动量效应能够帮助监管机构更好地了解市场的行为和机制，以及市场上的投资者行为和决策过程，有助于监管机构更好地制定监管政策和规则。其次，监管机构可以利用动量效应研究成果提供投资者更为准确和全面的市场信息，促进市场的透明度和公平性。此外，研究动量效应还可以帮助监管机构发现市场中存在的潜在风险，及时采取措施防范风险，维护市场的稳定和安全。

第三节 创新点

本文的创新之处有以下三点：

第一，本文将交易时间提高到 30 分钟的频率考察动量效应。现有文献对动量效应的研究主要集中在中长期的频率（年度，月度），关于日内动量效应的研究较少，本文将一个交易日内的交易时间分成 8 个 30 分钟的超高频时段，探究沪深 300 股指期货的动量效应。

第二，发现了隔夜效应对日内动量具有显著影响。研究结果显示，在沪深 300 股指期货市场上，与 Gao 等人（2018）在美股市场中的发现不同，第一个半小时对最后半小时收益并没有显著的预测作用。相反，本文发现了隔夜效应的存在，即前一日收盘价格与当日开盘价格之差的比率与倒数第二个半小时收益对最后半小时收益具有显著的正向预测作用。这一发现对投资者具有重要的启示意义。投资者可以在进行股指期货交易时，结合隔夜效应和倒数第二个半小时收益对最后半小时收益的预测作用，优化自己的交易策略，提高投资效益。同时，本研究也提醒投资者注意前一日收盘价格与当日开盘价格之间的差异，了解市场变化的趋势，合理调整自己的投资决策。

第三，将深度学习算法 LSTM 模型引入到日内动量策略的研究中。对沪深 300 股指期货最后半小时收益的涨跌进行二分类预测，建立基于 XGBoost 和 LSTM 的日内动量策略。最终基于 LSTM 的日内动量策略取得了 9.136% 的年化收益，夏普比率 2.164，高收益的同时风险也较低。

第四节 文献综述

根据的测度方法的差异，动量策略传统上分为两类：（多元）横截面动量（Cross-Sectional Momentum）（Jegadeesh 和 Titman, 1993; Kim, 2019）和（单变量）时间序列动量（Time-Series Momentum）（Moskowitz Ooi 和 Pedersen, 2012; Baltas 和 Kosowski, 2017）。横截面动量，也称为相对动量或多资产动量，是反映赢家组合和输家组合之间收益率差异的指标（Gray 和 Vogel, 2016）。而时间序列动量效应，也被称为绝对动量或单资产动量，是根据某项资产过去的收益率来计算动量效应的，与其他资产价格或收益率无关（Antonacci, 2015; Moskowitz 等, 2012）。

一、横截面动量

横截面动量策略侧重于证券之间的相对表现，买入相对赢家和卖出相对输家。

Jegadeesh 和 Titman (1993) 通过根据过去的回报率对股票进行排名, 并将最高十分之一与最低的十分之一进行交易, 发现最近在过去 3 到 12 个月中表现优于同行的证券在下个月继续平均表现优于同行。横截面动量的表现也被证明在各种市场和资产类别上是稳定的 (Baz 等, 2015)。

在空间维度上, 动量效应在不同国家和地区的资本市场中广泛存在, 并且在发达资本市场中普遍存在。Grin 等 (2005) 发现短期动量效应不仅存在于美国股票市场, 而且也存在于其他国家和地区的股票市场, 包括欧洲、亚洲、拉丁美洲等, 并且具有一定的全球性特征。Fama 和 French (2012) 通过对北美、欧洲、日本和亚太四个地区的股票市场进行研究, 发现除了日本市场外, 动量效应在其他市场中广泛存在。Asness (2011) 和 Chaves (2012) 通过对动量策略进行不同的调整, 在日本的股票市场中也发现了显著的动量效应。Chui 等 (2000) 的研究指出, 这种动量效应在许多欧洲国家十分显著, 而在许多新兴国家, 虽然不及欧洲国家强烈, 但仍能够起到积极的预测作用。可以说, 这种效应是具有普遍性的, 并且在不同国家的市场中表现出不同程度的效力。此外, 在至少五个亚洲国家也是如此。

在时间维度上, Geezy 和 Samonov (2016) 发现, 市场上存在已超过 200 年的动量效应现象, 不仅局限于股票市场, 还出现在货币市场中 (Shleifer 和 Summers, 1990; Kho, 1996; Lebaron, 1999)。此外, 众多学者发现商品市场中也存在着动量效应, 且时间跨度较大 (Okunev 和 White, 2003; Erb 和 Harvey, 2006; Gorton 等, 2013)。Zaremba 等 (2019) 发现短期动量效应不仅存在于股票市场, 而且也存在于其他资产类别, 包括政府债券、国库券、商品和货币, 并且在过去两个多世纪中一直保持稳定。

许多研究表明, 动量效应的存在与否与其形成期有关。根据形成期的不同, 可以将动量效应划分为短期动量效应 (Short-term Momentum)、中期动量效应 (Intermediate-term Momentum) 和长期动量效应 (Long-term Momentum)。其中, 短期动量效应指的是动量效应的形成期在一个月以内。短期动量效应是由 Jegadeesh (1990) 和 Lehmann (1990) 最早发现并证明的, 他们分别使用了美国股票市场中每周和每日频率下的数据, 并且发现了短期反转效应在不同季节和不同市场状态下的差异。他们认为短期反转效应可能是由市场对于新信息过度反应而导致的, 即投资者对于最近发生的事件或消息过于敏感, 并且忽视了先验概率或样本大小等信息。Avramov 等 (2016) 认为, 机构投资者流动性对短期反转效应有很大的解释作用, 低流动性市场中的反转效应更为显著。Gray 和 Vogel (2016) 发现短期动量效应与市场状态有关, 即在牛市中, 赢家组合会持续跑赢输家组合;

而在熊市中，赢家组合会持续跑输输家组合，并且提出了一种基于市场状态调整动量策略权重的方法。中期动量效应指的是形成期在三个月到一年以内，中期动量效应是由 Jegadeesh 和 Titman（1993）最早提出的，他们在美国股票市场中，以过去三到十二个月的收益率作为排序依据，持有未来三到十二个月，发现赢家组合（Past Winners）与输家组合（Past Losers）之间存在显著且稳健的正向收益差。他们认为中期动量效应可能是由市场反应不足或过度反应导致的，也可能是由投资者行为或制度因素引起的。Li（2021）研究了产品市场竞争与中期动量效应的关系，发现在美国股票市场中，随着行业集中度的提高，收益反转模式只在竞争性行业中可观察到。当行业集中度达到最高分位数时，这些模式逐渐消失并反转。他认为中期动量效应的表现不仅受到产品市场竞争水平的影响，还受到近期经济衰退的影响。长期动量效应则反映的是资产价格在长期内的变化趋势，一般会持续数年。DeBondt 和 Thaler（1985）在美国股票市场中，以过去三到五年的收益率作为排序依据，持有未来三到五年，发现输家组合与赢家组合之间存在显著且稳健的负向收益差异。

在中国股票市场的动量效应的存在性研究中，因为中国股市的有效历史数据不足，所以在 Rouwenhorst（1999）验证动量效应的研究中，没有将中国股市纳入研究范畴。近年来，随着中国股市的发展，人们越来越关注 A 股市场中的动量效应。Cheema 和 Nartea（2017）认为中国股市可能具有与发达市场不同的特征，因此对于中国股市的动量效应的存在性存在分歧。然而，多数研究者认为在 A 股市场中存在动量效应，但其持续的期间相较于发达市场可能更短。在动量效应的验证中，策略构造起着重要的作用，因此可以按照策略构造中最关键的形成期和持有期对文献进行分类梳理。

表 1.1 不同策略期限下的动量效应研究

| 策略构造（形成期限） | A 股市场表现 | 代表性文献 |
|---------------|--------------|--|
| 超高频形成期（1 日以内） | 动量效应 | 刘清源等（2016） |
| 超短形成期（1 周以内） | 动量效应 | 唐也然等（2020） 曲荣华和刘扬（2020） 王剑波和钟源（2021） |
| 短形成期（1 月以内） | 动量效应 | 周琳杰（2002） 高秋明等（2014） 鲁臻和邹恒甫（2007） |
| 中形成期（1 年以内） | 动量效应（反转效应较弱） | 游家兴（2008） 樊家鸣（2011） 史永东等（2015） |
| 长形成期（1 年以上） | 反转效应 | Naughton 等（2008） |

从动量策略构建的角度来看，策略的形成期反映了选取历史数据的时间范围，而持

有期则反映了策略组合的再平衡频率。王剑波和钟源（2021）发现：一是我国上证 A 股市场存在排序期为 1 周，持有期为 1 至 4 周的超短期动量效应；二是信息传播速度慢的公司，动量效应明显；三是从投资者过度自信的角度可以很好地解释动量效应。尹瑞和魏晓峰（2020）通过 2010-2019 年中国 A 股市场的数据，研究了投资者理性特征对动量效应的影响，采用了异质信念、处置效应和过度自信三个指标来衡量投资者理性特征，并分别构建了相应的动量策略，发现中国股票市场存在着短期动量效应。中期动量效应在中国股市的存在性得到了多项研究的支持，包括 Kang 等（2002）、徐信忠和郑纯毅（2006）等的研究。虽然有许多学者证实了中国股市存在着中期动量效应，但仍有许多学者对中期和长期动量效应的存在性表示质疑（王永宏和赵学军，2001；马超群和张浩，2005；郭磊等，2007；田利辉等，2014）。其中，形成期为 6 个月的中期动量效应被广泛关注（Cheema 和 Nartea，2014）。然而，对于形成期在一日以内的动量效应的研究，国内较少，具有代表性的是刘清源等（2016）发现隔夜收益是沪深市场动量策略收益的主要来源。

二、时间序列动量

时间序列动量（Time Series Momentum, TSM）专注于资产本身的过去回报，建立包含所有正在考虑的证券的投资组合。这最初是由 Moskowitz 等（2012）提出的，他们描述了一种具体策略，该策略使用波动率缩放并根据过去一年的收益情况交易头寸，他们在 25 年的数据中分别展示了 58 种流动投资工具的盈利能力。从那时起，提出了许多交易规则，各种趋势估计技术和方法将应用到交易中。例如，Bruder 等（2013）记录了广泛的线性和非线性过滤器来测量趋势，并统计其重要性，尽管他们没有直接讨论用这些估计来衡量头寸的方法。Baltas 和 Kosowski（2017）采用了类似于 Moskowitz 等（2012）的方法，将过去 12 个月的对数价格与时间倒退，并使用回归系数 t 统计来确定交易头寸的方向。尽管夏普比率在两者之间具有可比性，但基于 t 统计的趋势估计导致投资组合周转率降低了 66%，从而降低了交易成本。Baz 等（2015）、Rohrbach 等（2017）提出了更复杂的交易规则，以波动率的异同移动平均线（Moving Average Convergence and Divergence, MACD）指标作为输入。

随着计算机存储技术的快速发展和金融高频数据的日益可用，日内数据引起了国内外学者在波动性研究领域的广泛关注。然而，国内外缺乏基于日内数据收益率预测贡献的相关研究。需要强调的是，Gao 等（2018）首先研究了标普 500ETF 以每半小时为采

样频率的日内收益率，并提出了日内动量策略。该策略通过统计和经济学的方 法，在交易日的第一个半小时（16:00（前一天收盘价）—10:00（当天））和倒数第二个半小时（15:01—15:30）的日内收益率上进行预测，以提高最后半小时（15:31—16:00）的收益率预测准确性，即利用日内信息可以提高预测准确性。Elaut 等（2018）发现，RUB-USD 货币市场的后半小时回报可以由前半小时回报预测。Zhang 等（2019）还通过使用中国股票市场上的上证综指证明了显著的日内动量效应。研究还提供了两种可能的经济解释：一是投资者存在信息滞后现象，即开盘前的信息会影响投资者的反应，使得前半小时和后半小时的交易方向呈现正相关；二是投资者不频繁调整资产组合策略（Investors' Infrequent Rebalancing）（Bogousslavsky, 2016）。随后，其他学者也在中国股市进行了实证检验，证实了日内动量的存在。特别是在 Gao 等（2018）的研究基础上，Zhang 等（2019）和 Li 等（2020）发现，中国股市也存在日内动量效应，即每日的第一个半小时的收益率和倒数第二个半小时（14:01—14:30）的收益率有助于预测最后半小时（14:31—15:00）的收益率。

三、深度学习在动量策略中的应用

在机器学习，尤其是神经网络的使用，为金融时间序列预测提供了新的方法。深度学习在医学、需求预测和金融等领域表现良好（Miotto 等，2017；Ke 等，2017；Heaton 等，2016）。Gu 等（2017）使用各种线性、基于树的神经网络模型证明了非线性方法的优势，深度神经网络，特别是三层多层感知器（Multi-Layer Perceptron, MLP），具有最佳的样本外预测 R^2 。然后，通过使用模型预测每月对股票进行排名，建立机器学习投资组合，最佳策略来自四层 MLP，交易前十分之一与预测十分之一。Kim（2019）表示深度学习也可以应用于趋势预测，以便将其应用于动量策略。与回溯法相比，他们使用一种称为堆叠降噪自动编码器的架构进行预测时，在收益方面取得了显著的改善。

然而，尽管 Kim（2019）使用了深度学习技术并取得了很好的效果，但没有考虑诸如递归神经网络（Recurrent Neural Networks, RNN）等更现代的体系结构。许多论文发现，RNN 能够准确预测金融时间序列，例如，在预测货币波动性（Dunis 和 Huang, 2002）和股票收益预测方面（Rather 等，2015）。尤其是最现代的 RNN 结构，如长短期记忆（Long Short Term Memory, LSTM）和门控循环单元（Gated Recurrent Units, GRU），与常规前馈人工神经网络相比表现良好。

深度神经网络通常以纯粹的数据驱动方式捕获时间动态和横截面关系。TensorFlow

(Abadi 等 2015) 和 PyTorch (Paszke 等, 2017) 等强大的开源框架也促进了深度神经网络的采用, 它们使用自动微分来计算反向传播的梯度, 避免了手动推导梯度的繁琐过程。反过来, 这种灵活性使深度神经网络能够超越标准的分类和回归模型。例如, 已观察到将传统时间序列模型与神经网络组件相结合的混合方法在任一类别(Makridakis 等, 2018)中都优于纯方法, 例如指数平滑 RNN (Smyl 等, 2018)、自回归 CNN (Binkowski 等, 2018) 和卡尔曼滤波器 (Fraccaro 等, 2017; Rangapuram 等, 2018)。此外, 这些框架能够开发用于训练神经网络的新损失函数, 例如生成对抗网络 (GAN) 中的对抗损失函数 (Goodfellow 等, 2014)。

综上所述, 在动量效应的存在性方面, 大部分学者认为动量效应广泛存在不同的地区和资本市场中, 并且在发达资本市场中更加普遍。且动量效应在不同时间维度上也有不同的表现, 在中国股票市场中, 多数研究者认为存在动量效应, 但其持续的期间相较于发达市场可能更短。深度学习等技术的发展进一步拓宽了金融时间序列预测的方法, 但是对于形成期在一日以内的动量效应的研究还较少, 且与深度学习等先进技术的结合不足, 本研究将进一步丰富深度学习方法在日内动量效应中的应用。

第五节 本文结构与框架

本文的结构如下:

第一部分: 绪论。首先介绍了动量效应的背景、意义和目的; 其次介绍了本文的创新点; 再次分三部分介绍了动量效应的研究现状, 包括横截面动量效应, 时间序列动量效应和深度学习在动量策略中的应用; 最后介绍本文的框架结构。

第二部分: 相关算法理论。本章主要介绍了会使用到的深度学习模型, 包括 XGBoost 模型、长短期记忆神经网络模型、贝叶斯优化和相关概念。

第三部分: 日内动量效应实证分析。本章主要是检验我国股指期货日内动量效应的存在性, 以及在不同情况下的日内动量效应, 包括不同成交量、波动率和牛熊市下的日内动量效应。

第四部分: 动量策略的构建对比。根据第三章的检验结果构建动量策略, 分别建立了三个简单的动量策略和两个基准策略, 然后使用 XGBoost 和 LSTM 分别建立日内动量策略。

第五部分: 总结与展望。

以下是技术路线图:

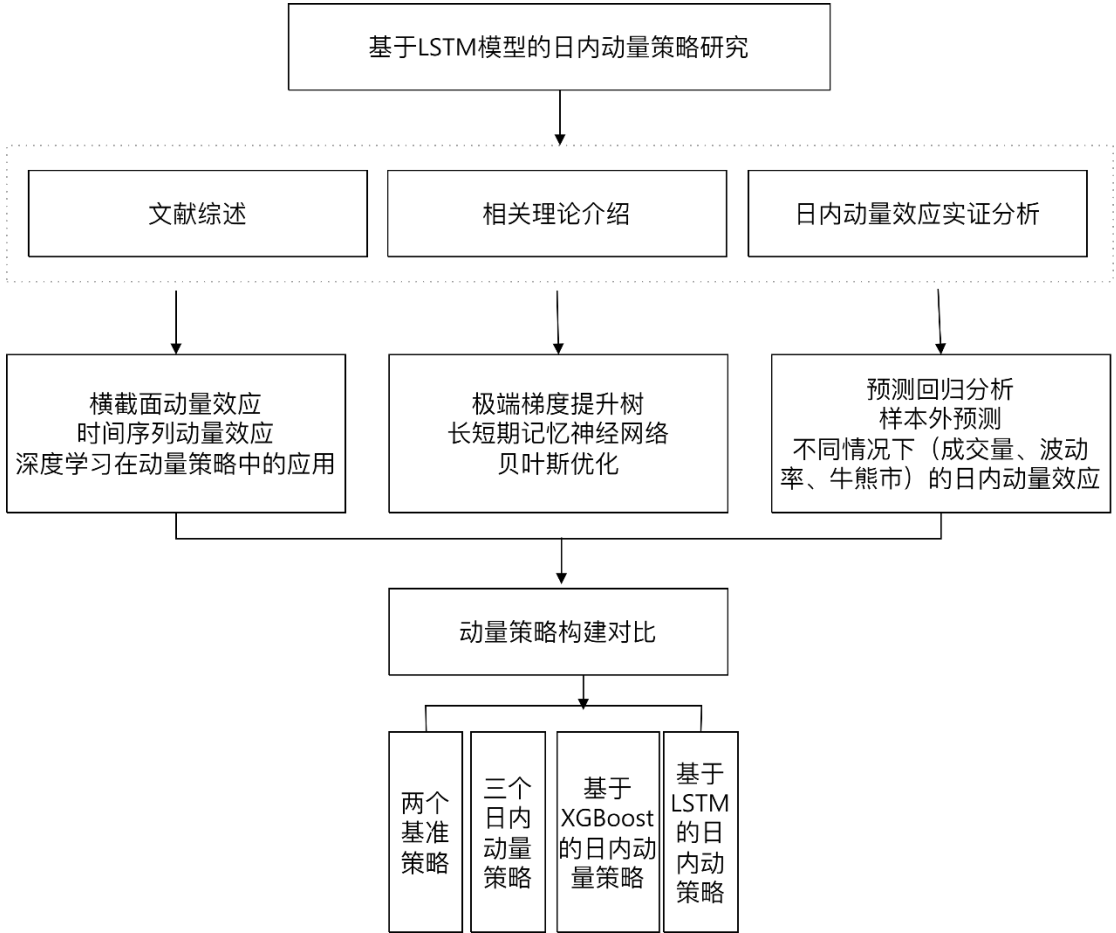


图 1.1 技术路线图

第二章 相关算法理论

第一节 XGBoost 模型

极限梯度提升 (Extreme Gradient Boosting, XGBoost) 是一种梯度提升框架, 被广泛应用于分类、回归和排序等问题。它采用了一个可扩展的、高效的树决策模型, 通过在每一轮迭代中加入新的树, 逐步提高模型的准确性。XGBoost 还包含了多种功能, 如自定义目标函数、正则化、缺失值处理和并行计算等, 使得它具备了较高的可定制性和适应性。此外, XGBoost 还利用了稀疏数据的优化方法, 减少了对存储空间的要求, 并通过直方图算法有效地处理连续型变量。在实践中, XGBoost 常常被认为是一种高度可靠和具有广泛适应性的机器学习方法, 不仅可以应用于传统的结构化数据分析, 还能够处理文本和图像等非结构化数据。因此, XGBoost 已经成为许多机器学习和数据挖掘竞赛中的重要算法, 并且被广泛应用于各种商业和科研领域。

XGBoost 算法最基本的预测模型可以表示如下:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (2.1)$$

其中, x_i 为样本特征, K 为树的个数, f_k 表示第 k 颗树对样本 x_i 的预测结果。

令损失函数为 f_{obj} , 其中, $L(\theta)$ 为误差项, $\Omega(\theta)$ 为正则化项, 得到:

$$f_{obj} = L(\theta) + \Omega(\theta) \quad (2.2)$$

$$L(\theta) = l(y_i, \hat{y}_i) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.3)$$

$$\Omega(\theta) = \sum_{k=1}^K \Omega(f_k) \quad (2.4)$$

加入一个新函数 $f(x_i)$, 使得损失函数最小化。

$$f_{obj}^{(t)} = \sum_{i=1}^n (y_i - (\hat{y}_i + f_t(x_i)))^2 + \Omega(f_t) + C \quad (2.5)$$

然后通过泰勒展开近似损失函数:

$$f_{obj}^{(t)} \approx \sum_{i=1}^n \left[l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i) \right] + \Omega(f_t) + C \quad (2.6)$$

$$h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$$

对于每一颗回归树, 其模型可以表示如下:

$$f_t(x) = w_{q(x)}, w \in R^T, q: R^d \rightarrow \{1, 2, \dots, T\} \quad (2.7)$$

其中, w 为叶子节点的权重, $q(x)$ 为叶子节点。 T 为叶子节点个数。

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (2.8)$$

为了防止模型过拟合, 加入参数 γ 和 λ , 则损失函数变为:

$$f_{obj}^{(t)} = \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \quad (2.9)$$

令 $G_j = \sum g_i$, $H_j = \sum h_i$, G_j 表示当前叶子结点包含的样本的一阶导数之和, H_j 表示当前叶子结点包含的样本的二阶导数之和

$$f_{obj}^{(t)} = \sum_{j=1}^T \left[G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2 \right] + \gamma T \quad (2.10)$$

对 w_j 求偏导, 并使偏导等于 0, 有

$$G_j + (H_j + \lambda) w_j = 0 \quad (2.11)$$

求解得

$$w_j^* = \frac{-G_j}{H_j + \lambda} \quad (2.12)$$

最后, 得到当前树结构下的最佳目标函数值:

$$f_{obj}^* = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \quad (2.13)$$

式中, f_{obj}^* 值越小, 代表对应的模型越好。

XGBoost 采用了并行算法和缓存优化技术, 因此在训练数据时速度非常快, 具有非常高的预测精度, 在许多数据挖掘比赛中都有很好的表现, 具有很好的泛化能力, 即在新数据上具有很好的预测能力, 具有很好的可解释性, 可以方便地确定特征的重要性。

第二节 长短期记忆神经网络

长短期记忆神经网络是基于循环神经网络的一种模型。循环神经网络 (Recurrent

Neural Network, RNN) 是一类用于处理序列数据的神经网络 (Rumelhart 等, 1986), 被广泛应用于自然语言处理、语音识别、机器翻译、时间序列预测等领域。相较于传统的前馈神经网络, 循环神经网络不仅可以接受其他神经元的信息, 还可以接受自身的信息, 形成有环路的网络结构。这样的结构可以实现对序列数据的记忆, 进而对未来的数据进行预测。但 RNN 的参数数量随着时间步数的增加而呈指数级增加, 容易导致过拟合, 当梯度在通过循环单元时逐渐变小或变大时, 很难传递长期依赖性, 从而影响模型的训练和泛化能力。

为了克服循环神经网络中的梯度消失问题, 引入了门控循环神经网络 (Gated RNN)。门控循环神经网络使用门控机制来有选择性地引入新信息和遗忘之前的信息。Hochreiter (1997) 等人提出了长短期记忆网络 (LSTM), 该网络与传统的循环神经网络不同, 可以很好地学习任意长度的时间序列并进行预测。此外, LSTM 网络可以随时保存存储单元的时间信息, 从而在一定程度上解决了循环神经网络的梯度消失问题。

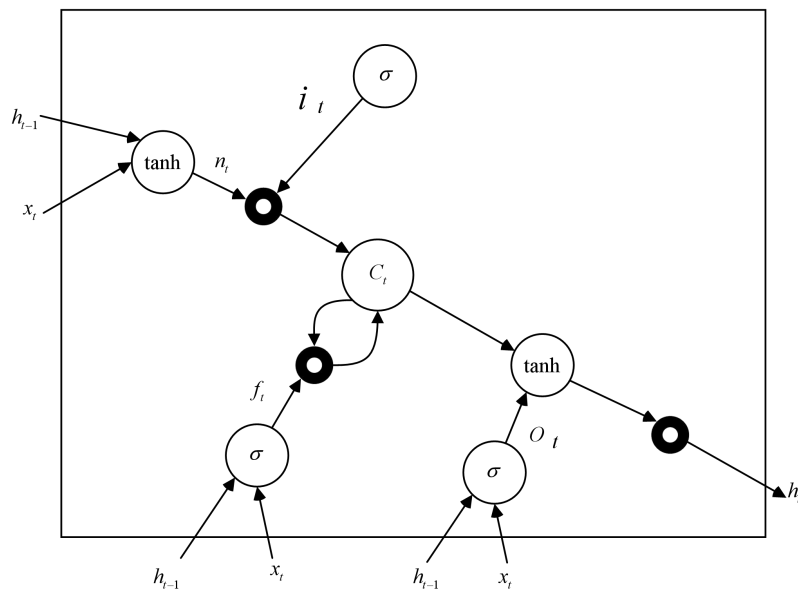


图 2.1 长短期记忆网络存储单元

具体来说, 每个 LSTM 单元由三个门和一个记忆单元组成。这些门分别是输入门 (Input Gate)、遗忘门 (Forget gate) 和输出门 (Output gate), 它们可以通过学习得到自适应的权重。输入门控制着新输入数据的流入, 遗忘门控制着之前记忆单元中的信息是否需要被遗忘, 输出门控制着当前记忆单元中的信息是否需要输出。

LSTM 单元的计算可以分为以下三个步骤:

输入门控制新数据的输入。对于输入数据 x_t 和上一时刻的隐藏状态 h_{t-1} , 输入门的

输出为一个介于 0 到 1 之间的数值向量 i_t ，表示当前时刻有多少信息需要被记忆。

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (2.14)$$

其中， W_{ix} 和 W_{ih} 是权重矩阵， b_i 是偏置向量， σ 是 sigmoid 函数。

遗忘门控制旧数据的遗忘。对于输入数据 x_t 和上一时刻的隐藏状态 h_{t-1} ，遗忘门的输出为一个介于 0 到 1 之间的数值向量 f_t ，表示当前时刻有多少信息需要被遗忘。

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (2.15)$$

更新记忆单元。对于输入数据 x_t 、上一时刻的隐藏状态 h_{t-1} 、输入门的输出 i_t 和遗忘门的输出 f_t ，可以计算当前时刻的记忆单元 c_t ，用于保存历史信息。

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \quad (2.16)$$

其中， \odot 是逐元素相乘操作， \tanh 是双曲正切函数。

输出门控制输出。对于当前记忆单元 c_t 和当前隐藏状态 h_t ，输出门的输出为一个介于 0 到 1 之间的数值向量 o_t ，表示当前时刻有多少信息需要输出。

第三节 贝叶斯优化算法

贝叶斯优化是一种基于概率论的优化方法，它利用贝叶斯定理和概率模型来对参数进行估计和优化。这种方法具有自动化和灵活性，适用于各种类型的优化问题，并且能够考虑不确定性和噪声。贝叶斯定理指出：给定一个随机变量 x ，其关于一个固定参数 θ 的后验分布可以由先验分布 $P(\theta)$ 和数据 D 通过贝叶斯公式计算得到：

$$P(\theta|D) = P(D|\theta)P(\theta)/P(D) \quad (2.17)$$

其中， $P(\theta|D)$ 表示关于 θ 的后验分布， $P(D|\theta)$ 表示给定 θ 的数据的似然分布， $P(\theta)$ 表示关于 θ 的先验分布， $P(D)$ 表示数据的先验概率。

贝叶斯优化利用这个公式，在后验分布中寻找使得目标函数最大的参数值，并用该值作为优化结果。即：

$$\theta^* = \operatorname{argmax} P(\theta|D) \quad (2.18)$$

通常， $P(\theta)$ 和 $P(D|\theta)$ 都是需要通过一些方法进行建模和评估的。在贝叶斯优化中，先验分布 $P(\theta)$ 可以是一个预先确定的分布，也可以是通过分析数据得出的分布。同时，似然分布 $P(D|\theta)$ 可以通过模拟数据生成的方法得到。

最终，贝叶斯优化通过不断迭代计算后验分布并选择最大值，来实现对参数的优化。贝叶斯优化具有很多优点：可以自动评估不确定性，贝叶斯优化可以通过评估先验分布和后验分布的不确定性来评估参数的不确定性。这对于评估参数的置信度非常有用；具

有灵活性，贝叶斯优化可以应用于各种不同类型的优化问题，并且能够适应不同类型的先验知识。适用于高维问题，贝叶斯优化对于高维问题是有效的，因为它能够自动评估不确定性和噪声。

第四节 神经网络训练相关概念

一、过拟合

过拟合是指机器学习模型在训练数据上表现良好，但在测试数据上表现较差的现象。这是因为模型过于拟合了训练数据，导致无法泛化到新的数据。在训练过程中，如果模型的复杂度过高或者训练数据量太小，就容易出现过拟合问题。下面介绍几种缓解过拟合的方法：

正则化（Regularization）是一种减少模型复杂度的方法，可以提高模型的泛化能力，从而避免过拟合的问题。通常会在损失函数中加入一个正则化项，它会惩罚模型的复杂度。常用的正则化方法包括 L1 正则化和 L2 正则化。L1 正则化在损失函数中添加模型参数的 L1 范数（参数绝对值之和）与系数 λ 的乘积作为正则化项，L1 正则化会让模型的某些参数为 0，从而实现特征选择的功能，即去除无用的特征，使得模型更加简单和易于解释。而 L2 正则化是将权重向量中所有元素的平方和添加到损失函数中，该损失函数通常称为 L2 损失函数。通过 L2 正则化，模型将被迫使权重参数的值尽量小，从而减少权重参数对模型输出的影响。这使得模型更加平滑，更容易泛化到新的数据集。

$$\begin{aligned}\theta^* &= \underset{\theta}{\operatorname{argmin}} \mathbb{R}_D^{struct}(\theta) \\ &= \underset{\theta}{\operatorname{argmin}} \mathbb{R}_D^{emp}(\theta) + \frac{1}{2} \|\theta\|^2 \\ &= \underset{\theta}{\operatorname{argmin}} \frac{1}{N} L(y^{(n)}, f(x^{(n)}; \theta)) + \frac{1}{2} \|\theta\|^2\end{aligned}\quad (2.19)$$

其中， $\|\theta\|$ 是 L2 范数的正则化项，用来减少参数空间，避免过拟合； λ 用来控制正则化的强度。

提前停止（Early Stopping）是一种常用的正则化技术，主要用于解决过拟合的问题。提前停止通过在训练过程中监控模型的性能指标，当性能指标在一段时间内没有进一步提升时，就停止训练，以避免过拟合。

具体来说，提前停止是指在模型训练过程中，将训练集分为训练集和验证集，使用

训练集进行模型训练，同时使用验证集监控模型的性能指标，如损失函数或准确率等。当模型在验证集上的性能指标连续若干轮没有进一步提升时，就停止训练，将模型参数设置为最好的一次训练的参数。提前停止的核心思想是防止模型在训练集上过拟合，通过在验证集上的性能指标判断是否过拟合。当模型在训练集上表现得越来越好，但在验证集上的性能指标开始下降时，就说明模型已经开始过拟合，需要停止训练。提前停止相比其他正则化技术，如 L1、L2 正则化等，不需要手动设置超参数，可以自适应地停止训练。

丢弃法（Dropout）由 Hinton 等（2012）提出，经过实验证明可以在深度学习任务中取得非常好的效果。丢弃法的思想是在神经网络的训练过程中，对每个神经元以一定概率进行随机舍弃（即输出值为 0），相当于将该神经元从网络中暂时去除。丢弃法随机地“丢弃”一些神经元，使得每次迭代的网络结构都不同，从而在一定程度上减轻了神经网络的过拟合现象。

具体来说，对于一个具有 n 个神经元的神经网络，在进行训练的时候，每个神经元都有 p 的概率被舍弃（即输出值为 0），剩下的神经元的输出值需要按比例进行缩放，使得输出值的期望值保持不变。在测试阶段，由于此时需要考虑所有的神经元，因此不再进行舍弃，但需要将每个神经元的输出值按照 $(1-p)$ 的比例进行缩放，以保证输出值的期望值不变。

丢弃法可以减轻神经网络的过拟合现象，因为通过随机舍弃神经元，可以减少神经元之间的共适应性，增强网络的泛化能力。此外，由于每次迭代的网络结构都不同，因此丢弃法也具有一定的正则化作用。

二、损失函数

损失函数是机器学习中用于衡量模型预测结果与实际值之间差距的函数。通常，模型的目标是最小化损失函数，以使预测结果与实际值尽可能接近。

在监督学习中，损失函数通常基于模型预测结果和实际标签之间的差异来计算。最常用的损失函数之一是均方误差（Mean Squared Error, MSE），它计算预测值和实际值之间的平均平方差。MSE 在回归问题中广泛使用，但在分类问题中可能不太适用。其公式如下：

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2.20)$$

其中 y_t 代表真实值, \hat{y}_t 代表模型预测值。该损失函数适用于回归问题。

对于分类问题, 通常使用交叉熵 (Cross Entropy) 作为损失函数。交叉熵衡量了预测结果与实际标签之间的差异, 并尝试使其最小化。对于二分类问题, 使用二元交叉熵 (Binary Cross Entropy); 对于多分类问题, 使用多元交叉熵 (Categorical Cross Entropy)。除了常见的均方误差和交叉熵外, 还有其他损失函数, 如绝对误差 (Absolute Error)、对数损失 (Logarithmic Loss) 等, 每个损失函数都有其特定的优点和适用范围。选择损失函数应该根据具体问题和模型类型进行选择。在训练模型时, 目标是最小化损失函数, 因此选择适合的损失函数可以提高模型的性能和精度。

三、Adam 优化算法

Adam (Adaptive Moment Estimation) 优化算法是一种自适应学习率的优化算法, 由 Kingma 和 Ba (2014) 提出, 其基本思想是为每个参数计算一个自适应的学习率, 不同于传统的随机梯度下降 (SGD) 采用相同的学习率来更新所有的参数。Adam 算法结合了动量法和 RMSProp 算法的优点, 既可以利用梯度的一阶矩估计 (即梯度的平均值), 又可以利用梯度的二阶矩估计 (即梯度的方差)。

Adam 算法的更新规则如下:

计算梯度的一阶矩估计 (即梯度的平均值):

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (2.21)$$

计算梯度的二阶矩估计 (即梯度的方差):

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (2.22)$$

对一阶矩估计和二阶矩估计进行偏差修正:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (2.23)$$

利用修正后的一阶矩估计和二阶矩估计来更新参数:

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t \quad (2.24)$$

其中, g_t 是当前时刻的梯度, m_t 和 v_t 分别是梯度的一阶矩估计和二阶矩估计, \hat{m}_t 和 \hat{v}_t 是对 m_t 和 v_t 的偏差修正后的估计, β_1 和 β_2 是平滑系数, t 表示当前迭代次数, w_t 和 w_{t+1} 分别表示当前和下一时刻的参数值, η 表示学习率, ϵ 是为了数值稳定性而添加的一个小数值。

Adam 算法的优点在于能够自适应地调整学习率, 且在处理大规模数据时效果显著。

然而，Adam 算法也有一些缺点，比如对于非平稳的数据（如带有周期性变化的数据）可能存在不良影响。此外，对于小批量的数据和稀疏梯度，Adam 算法的表现可能不如其他优化算法。

第五节 模型评估方法

由于本文是预测沪深 300 股指期货最后半小时的涨跌情况，因此是一个分类问题。分类模型的评估方法是用于评估分类模型的性能的技术。常见的分类模型评估方法如下：

Confusion Matrix: 混淆矩阵是一种用于评估分类模型性能的常用技术，它用于统计分类模型预测的正确率和错误率。混淆矩阵中的数据可以用于计算一些评估指标，如精确率，召回率和 F1 分数等。

表 2.1 混淆矩阵

| | | Actual | |
|-----------|----------|----------|----------|
| | | Positive | Negative |
| Predicted | Positive | TP | FP |
| | Negative | FN | TN |

其中，True Positive（TP）代表模型正确预测的正例数；False Positive（FP）代表模型错误预测的正例数；False Negative（FN）代表模型错误预测的负例数；True Negative（TN）代表模型正确预测的负例数。

Accuracy 代表准确率，是评估分类模型的常用指标，它表示模型的预测正确率。准确率的公式为： $\text{Accuracy} = (\text{True Positives} + \text{True Negatives}) / \text{Total Predictions}$ 。

Precision 代表精确率，是另一个评估分类模型的常用指标，它表示预测为正的样本中有多少是真的正样本。精确率的公式为： $\text{Precision} = \text{True Positives} / (\text{True Positives} + \text{False Positives})$ 。

Recall 代表召回率，它表示真实正样本中有多少被预测为正的。召回率的公式为： $\text{Recall} = \text{True Positives} / (\text{True Positives} + \text{False Negatives})$ 。

F1 Score 表示 F1 分数，是精确率和召回率的调和平均数，用于评估分类模型的性能。F1 分数的公式为： $\text{F1 Score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$ 。

ROC（Receiver Operating Characteristic）曲线和 **AUC**（Area Under the Curve）是评估分类模型性能的另一种方法。ROC 曲线是用于可视化真正率（TPR）和假正率（FPR）之间的关系的技术。AUC 是评估模型性能的指标，表示 ROC 曲线下方的面积。

Cross-Validation 表示交叉验证，是一种评估分类模型性能的常用技术，它通过将训

练数据分为多个小数据集，然后在每个数据集上分别训练模型，并在剩下的数据集上评估模型的性能，来评估模型的泛化能力。

本章主要介绍本文后续实验所用方法的理论知识，为下文建模打下理论基础。本章介绍了 XGBoost 的原理方法及优点，然后介绍了 LSTM 的原理优点和神经网络训练相关知识，最后介绍了分类模型的评估方法及指标。XGBoost 是一种基于决策树的集成学习模型，通过在决策树之间构建弱分类器来提高模型的准确性和泛化能力，适用于结构化数据，而本文所使用的数据为时间序列数据，LSTM 是一种基于循环神经网络的深度学习模型，特别适合于处理序列数据，如股票价格序列数据。LSTM 可以通过记录长期依赖关系和消除短期干扰来保存和提取历史信息，因此，对于时间序列相关的问题，LSTM 往往比其他模型更有效。因此理论上来说，基于 LSTM 模型的日内动量策略会比基于 XGBoost 模型的日内动量策略效果更好。

第三章 日内动量效应实证分析

第一节 数据来源

本文从米筐量化交易平台收集了 2010 年 4 月 16 日至 2022 年 12 月 16 日沪深 300 股指期货主力连续期货合约共计 3081 个交易日的交易数据。

本文主要遵循 Gao 等（2018）的观点，考察中国股指期货市场的日内动量。中国股指期货市场的一个特点是开盘和收盘时间与美国不同。具体来说，2016 年之前，交易时段从 9:15 开始，比股市早 15 分钟，在 15:15 结束，比股市晚 15 分钟。然而，2016 年后，股指期货合约的交易周期与股市相同。为了数据的一致性，本文将 2016 年以前的上午 9:15 至 10:00 作为第一个半小时，14:30 至 15:15 作为最后一个半小时。由此得出每天八个半小时的收益率。

沪深 300 股指期货于 2010 年推出，是中国第一个也是最受欢迎的股指期货，流动性高，持续时间长，已被众多研究所采用。沪深 300 股指期货与沪深 300 指数不同，沪深 300 指数采用的 T+1 日的交易机制，无法当日买入当日卖出，而股指期货可以当日买入当日卖出，因此，研究沪深 300 股指期货的日内动量效应更具有实际意义。

第二节 日内动量效应的存在性检验

为了考虑不同时间段的动量效应的影响，本文使用 $P_{i,j}$ 来表示第 i 个交易日中第 j 个 30 分钟的收盘价格。用 $V_{i,j}$ 来表示第 i 个交易日中第 j 个 k 分钟的成交量。因此，第 t 个时间段的收益率 $t_{i,j}$ 为：

$$t_{i,j} = \frac{P_{i,j}}{P_{i,j-1}} - 1 \quad (3.1)$$

隔夜收益是指由于隔夜效应而对投资者带来的收益或损失，它包含了前一个交易日收盘到当日开盘之间市场上的所有信息。第 i 个交易日隔夜收益为 $t_{i,0}$ ，表达式如下：

$$t_{i,0} = \frac{P_{i,1_open}}{P_{i-1,8_close}} - 1 \quad (3.2)$$

其中 $P_{i,1_open}$ 表示第 i 个交易日 9:30 的收盘价， $P_{i-1,8_close}$ 表示第 $i-1$ 个交易日尾盘的收盘价。

接下来按照不同的划分区间，用线性回归方程来检验前面几个区间的收益率对最后一个区间收益率的影响。以半小时为例，每个交易日可以获得 9 个收益率数据，每个时

间段都可以获得该时间段的收益率数据，见下列公式：

$$t_{i,j} = \alpha_j + \beta_j t_{i,j} + \epsilon_j \quad j = 0, 1, \dots, 8; i = 1, 2, \dots, T \quad (3.3)$$

表 3.1 隔夜收益及日内七个半小时收益对最后一个半小时的收益的影响

| | t_0 | t_1 | t_2 | t_3 | t_4 | t_5 | t_6 | t_7 |
|-----------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| α | -0.415*** (-4.7857) | -0.371*** (-4.2751) | -0.375*** (-4.3095) | -0.375*** (-4.3058) | -0.374*** (-4.3081) | -0.375*** (-4.3125) | -0.373*** (-4.2954) | -0.376*** (-4.3629) |
| β_0 | 0.0724*** (5.8744) | | | | | | | |
| β_1 | | -0.0286* (-1.7766) | | | | | | |
| β_2 | | | 0.0513** (2.5368) | | | | | |
| β_3 | | | | 0.0065 (0.3257) | | | | |
| β_4 | | | | | -0.050*** (-3.3452) | | | |
| β_5 | | | | | | 0.0301 (1.5129) | | |
| β_6 | | | | | | | 0.0515*** (2.6112) | |
| β_7 | | | | | | | | 0.1260*** (7.2671) |
| R^2 | 1.102% | 0.102% | 0.208% | 0.003% | 0.362% | 0.074% | 0.221% | 1.685% |

注：表中***，**，*分别表示在 1%，5%，10%的置信水平下显著，括号中为 T 值

如上表所示， t_0 、 t_4 、 t_7 在 1%的显著性水平显著，其中 t_0 、 t_4 的 R^2 分别为 1.102%、1.685%，根据 Kandel（2005）和 Gao（2018）等人的研究，在股票市场中，当 R^2 达到 0.5%或以上时，可以视为显著的预测回报。与 Gao 等 2018 年在标普 500ETF 发现第一个半小时和最后一个半小时能够显著地预测最后半小时的收益不同，本文未发现第一个半小时对最后一个半小时有显著影响，而是隔夜效应 t_0 和倒数第二个半小时的收益对最后半小时收益有显著正向影响。

第三节 不同指标下的日内动量效应

一、牛市与熊市下的日内动量效应

不同的市场行情下，股市会表现出不同的特点，同一种策略很有可能出现很大差异，

为了了解在不同市场行情下，日内动量效应是否有显著的变化。所谓牛市指的是，股市行情强劲，价格持续上涨，投资者的信息高涨的情况；熊市则相反，指股市行情疲弱，价格持续下跌，投资者信息低落的情况。本文将沪深 300 股指期货划分为牛熊市，按照上文的回归方法，分别对牛市和熊市进行检验。

本文对于牛熊市的划分主要基于道琼斯的判别方法，具体而言，如果市场连续半年或以上上涨或下跌超过 30%，则被定义为牛市或熊市。在本文中，牛市和熊市的定义是根据市场的最高点和最低点来判断的。

表 3.2 沪深 300 股指期货牛-熊市后划分区间

| 周期 | 时间区间 | 指数区间 | 区间差值 | 幅度 | 振幅 |
|----|----------------------|---------------|---------|---------|---------|
| 熊市 | 2010/4/16-2014/6/5 | 3450-1702.6 | -1747.4 | -50.65% | 118.43% |
| 牛市 | 2014/6/6-2015/6/15 | 1702.6-4679.4 | 2976.8 | 174.84% | 174.84% |
| 熊市 | 2014/6/6-2015/6/15 | 4679.4-2443.4 | -2236 | -47.78% | 91.51% |
| 牛市 | 2015/8/28-2018/1/23 | 2443.4-4973.2 | 2529.8 | 103.54% | 103.54% |
| 熊市 | 2018/1/24-2019/1/4 | 4973.2-3681.2 | -1292 | -25.98% | 36.48% |
| 牛市 | 2019/1/5-2021/2/18 | 3681.2-6911.6 | 3230.4 | 87.75% | 87.75% |
| 熊市 | 2021/2/19-2022/12/16 | 6911.6-5290.8 | -1620.8 | -23.45% | 42.27% |

如上表所示，本文所选取的样本区间中，从沪深 300 股指期货建立以来，共经历了 4 次熊市和 3 次牛市。将牛市和熊市分别进行回归分析，结果如下表所示：

表 3.3 熊市期间回归结果

| | t_0 | t_7 | $t_0 \& t_7$ |
|-----------|---------------------------|--------------------------|---------------------------|
| α | -0.567259*** (-5.0348) | -0.53168*** (-4.7373) | -0.553213*** (-4.9659) |
| β_0 | 0.087372*** (5.6454) | | 0.081411*** (5.3145) |
| β_7 | | 0.14328*** (6.6235) | 0.136282*** (6.3414) |
| R^2 | 1.808 | 2.418 | 4.038 |

表 3.4 牛市期间回归结果

| | t_0 | t_1 | t_4 | t_7 |
|-----------|-------------------------|-------------------------|---------------------------|-------------------------|
| α | -0.3776*** (-2.0825) | -0.2398*** (-1.3353) | -0.3182*** (-1.8311) | -0.3304*** (-1.8392) |
| β_0 | 0.0770*** (2.4576) | | | |
| β_1 | | -0.1258*** (-3.9585) | | |
| β_4 | | | -0.234164*** (-6.9526) | |
| β_7 | | | | 0.118975*** (2.7725) |
| R^2 | 1.006 | 2.576 | 7.608 | 1.292 |

注：表中***，**，*分别表示在 1%，5%，10%的置信水平下显著，括号中为 T 值

从表 3.3 中可以看出，在市场上行情不好的情况下即本文所划分的熊市期间，隔夜效应和倒数第二个半小时在 1%的显著性水平下具有统计意义，即具有显著的正相关关系，且在将两者联合起来进行预测，将一元线性回归扩展为二元线性回归的情况下，也具有统计学意义，且两种情况下斜率大致相同， R^2 大概为 t_0 、 t_7 分别回归的 R^2 之和。

在表 3.4 中，在牛市行情下， t_0 和 t_7 在 1%的显著性水平下，也具有统计学意义，而 t_1 和 t_4 在 1%的显著性水平下具有显著的反转效应，即 t_1 和 t_4 每变动一个单位， t_8 分别减少 0.1258、0.2342 个单位。

两者相比较，在熊市期间， t_0 和 t_7 的拟合优度 R^2 都大于牛市期间，且回归系数也都是熊市大于牛市，熊市的动量效应要显著强于牛市期间，这也与 Gao 等（2018）的研究结论一致。

二、不同成交量下的动量效应

成交量是指一段时间内股票交易的数量。对于股票收益，成交量有以下几种影响：第一，市场流动性。如果成交量较高，市场的流动性就较强，投资者可以轻松买卖股票，这有利于维护股票价格的稳定。第二，供求关系。成交量的多少可以反映供求关系，如果成交量较高，说明需求很强，价格可能会上涨；如果成交量较低，说明供给较多，价格可能会下跌。第三，关注度。成交量高的股票往往受到更多关注，有可能更容易引起投资者的注意，进而影响股票价格。

动量效应是指股票价格近期上涨的股票今后继续上涨，而股票价格近期下跌的股票

今后继续下跌的现象。成交量可以作为动量效应的一个指标，因为成交量的多少可以反映市场对股票的关注度和需求，如果成交量较高，说明需求很强，投资者对该股票有很高的兴趣，这可以支持动量效应的发生。反之，如果成交量较低，说明需求较弱，投资者对该股票兴趣不高，这可能会削弱动量效应。因此，成交量可以作为动量效应的一个重要指标，有助于判断股票价格的趋势。

如下图 3.1 为沪深 300 股指期货八个半小时时间段的总体成交量图，v1 代表当日开盘到 10 点的成交量，v2 代表 10 点到 10 点 30 分的成交量，以此类推。可以看出图形大体上呈现出 U 型，第一个半小时和最后一个半小时的成交量最高，都在 1.2 亿成交量以上。其余时间流动性相对较弱，在开盘半小时之前，投资者通常需要进行市场调研、分析股市走势等，而在尾盘半小时之后，许多投资者会平仓离场，避免隔夜风险。因此，这两个时间段内，许多投资者可能会集中进行交易，导致成交量最多。从图中也可以看出在 v4 即 11 点至 11 点 30 分闭盘时成交量也较高，在上午交易时段，投资者可能会根据市场行情和自身策略进行买卖操作，但由于上午只有短短的三个小时交易时间，因此很多投资者可能需要在上午闭盘前进行必要的调整，以控制风险和获得更好的收益。这可能导致上午闭盘时成交量较高。此外，一些重要的财经数据和公告通常会在上午发布，例如就业数据、GDP 数据等。这些数据可能会对市场情绪和投资者的决策产生影响，因此在这些数据公布前后，成交量也可能会有所增加。

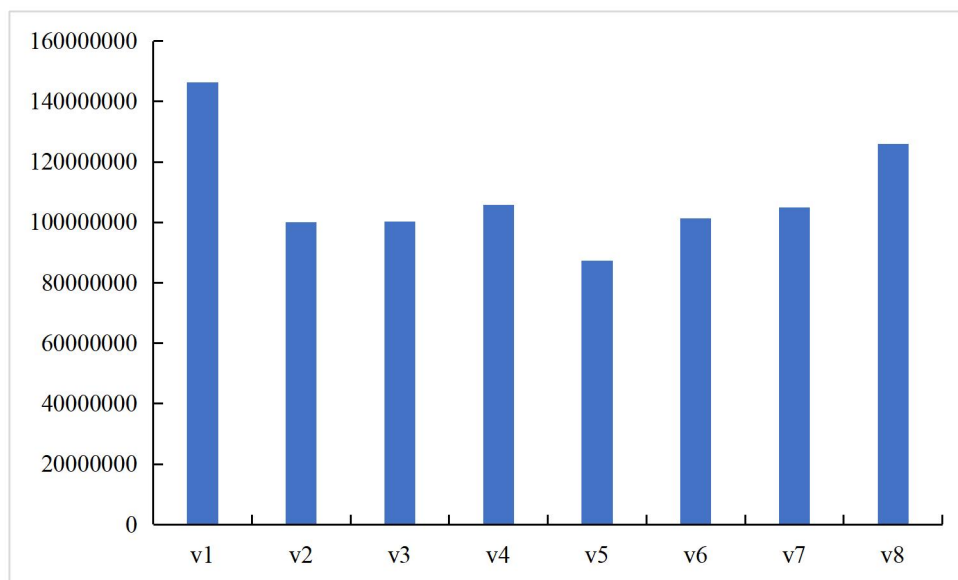


图 3.1 不同时间段沪深 300 股指期货成交量

因此本文将所有样本根据每日的成交量大小排序，并根据成交量分成两组：Low，High。再分别进行回归分析，结果如下：

表 3.5 不同成交量水平下回归结果

| | Low | | High | |
|-----------|------------------------|-------------------------|-------------------------|-------------------------|
| | t_0 | t_7 | t_0 | t_7 |
| α | -0.1579** (-2.0742) | -0.1515** (-2.0027) | -0.6582*** (-0.6582) | -0.6001*** (-3.8797) |
| β_0 | 0.0151** (1.0332) | | 0.0945*** (5.1528) | |
| β_7 | | 0.062974*** (2.6904) | | 0.1429*** 5.7693 |
| R^2 | 0.690 | 0.768 | 1.687 | 2.114 |

注：表中***，**，*分别表示在 1%，5%，10%的置信水平下显著，括号中为 T 值

从表 3.5 中可以看出，低成交量组 t_0 在 5% 的显著性水平下显著， R^2 为 0.69%， t_7 在 1% 的显著性水平下具有统计学意义， R^2 为 0.786%。高成交量组在 1% 的显著性水平下具有统计学意义，且 R^2 远高于低成交量组，分别为 1.687%、2.114%。检验结果也与上述论证相符，即成交量越高，动量效应越显著。

三、不同波动率下的日内动量效应

波动率是一个衡量股票价格波动程度的指标，它反映了股票价格在一定时间内的变动情况。波动率较高的股票价格波动较大，波动率较低的股票价格波动较小。波动率对股票收益有重要影响。首先，波动率较高的股票风险也较高，因为价格波动较大，投资者很难预测股票未来的价格走势，存在较高的价格波动风险。这可能导致投资者对该股票的信心不足，从而影响股票的收益。其次，波动率较高的股票可能难以吸引投资者的兴趣，因为这种风险很难被投资者接受。这可能导致股票的需求减少，从而影响其价格，从而影响投资者的收益。

但是，波动率也可以是一种机会。对于风险偏好的投资者，波动率较高的股票可能是一个不错的机会，因为这些股票的价格有可能在未来上涨。这可以为投资者带来较高的收益。波动率对动量效应有一定影响，如果波动率较高，股票的价格波动较大，那么动量效应可能会更明显，因为价格波动较大的股票更容易观察到动量效应。同时，波动率也可以影响投资者对动量效应的信心。如果波动率较高，那么投资者可能对动量效应产生疑问，因为价格波动较大的股票可能不适合所有投资者。因此，波动率对动量效应的影响是复杂的。

本文选取样本区间（2010 年 04 月 16 日—2022 年 12 月 16 日）内频率一分钟的数

据，计算出每日每分钟的平均价格，得到每天 240 分钟的平均价格，再计算标准差，以此作为该日的波动率，再根据波动率的大小，将样本分为三个区间：Low、Middle、High，再分别对样本进行回归分析，以此来验证低、中、高波动率下动量效应的效果。

回归结果如下：

表 3.6 不同波动率下回归结果

| | Low | | Middle | | High | |
|-----------|---------------------|-------------------------|----------------------|---------------------|-------------------------|-------------------------|
| | t_0 | t_7 | t_0 | t_7 | t_0 | t_7 |
| α | 0.0666 (0.9930) | 0.063806 (0.9567) | -0.1408 (-1.3934) | | -1.1389*** (-4.9948) | -1.0979*** (-4.8563) |
| β_0 | 0.0270* (1.6698) | | 0.0044 (0.2737) | | 0.1123*** | |
| β_7 | | -0.1029*** (-3.7018) | | 0.07539 (0.2672) | | 0.1792*** (5.9419) |
| R^2 | 0.570 | 1.316 | 0.0007 | 0.070 | 2.114 | 3.322 |

注：表中***，**，*分别表示在 1%，5%，10%的置信水平下显著，括号中为 T 值

从上表可以看出，在低波动率组中，第七个半小时收益对最后一个半小时的收益的影响呈现出显著的负相关，具有显著的反转效应。在中波动率组中，未发现其他半小时收益与最后一个半小时收益具有显著的相关关系，不具有动量效应。而在高波动率组中，在 1%的显著性水平下，具有显著的动量效应， t_0 和 t_7 ，对最后半小时的收益具有正向预测作用， R^2 分别达到了 2.114%和 3.322%。在高频交易的情况下，投资者对于股票价格的变动非常敏感，波动率越大，机会越多，投资者获利的可能性越大，因此形成了动量效应。

第四节 日内动量策略构建

由本章前两节的实证检验可知，在沪深 300 股指期货市场中存在显著的动量效应，开盘跳空的幅度即隔夜收益 t_0 和第七个半小时收益对最后一个半小时收益具有显著的正向预测作用。而在不同的市场条件、不同成交量和不同波动率条件下，动量效应显示出不同的特点。具体来说，在熊市期间，高成交量和高波动率情况下，动量效应更加显著。

如果此种动量效应是真实存在的，那么应用这种动量效应应该就能获得超额收益，因此本章就利用隔夜收益与第七个半小时收益对最后半小时构建股票择时投资策略。投资策略的标的是沪深 300 股指期货，期间为 2010 年 4 月 16 日至 2022 年 12 月 16 日，共计 3081 个工作日。根据隔夜收益 t_0 和倒数第二个半小时收益 t_7 来决定最后半小时是做

空还是做多。根据第三章的结论，本文形成了三个动量策略：第一个策略是将隔夜收益 t_0 作为沪深 300 股指期货最后一个半小时 t_8 的交易信号；第二策略是将倒数第二个半小时收益 t_7 作为一个半小时的交易信号；第三个策略是将隔夜收益 t_0 和倒数第二个半小时收益 t_7 结合起来作为最后一个半小时的交易信号。此外，还设立了两种基准策略进行对比。并对每个策略进行对比，具体策略如下：

(1) 策略 $\gamma(t_0)$ 仅将隔夜收益 t_0 的正负作为最后半小时买入或卖出的信号，若当日开盘跳空的幅度 t_0 大于 0，则做多，在最后半小时之前买入沪深 300 股指期货并在收盘前平仓；若当日开盘跳空的幅度 t_0 小于 0，则做空，在最后半小时之前卖出沪深 300 股指期货，并在收盘前平仓；若当日开盘跳空的幅度 t_0 等于 0，则不做任何操作。

$$\gamma(t_0) = \begin{cases} -t_8, & t_0 < 0 \\ 0, & t_0 = 0 \\ t_8, & t_0 > 0 \end{cases} \quad (3.4)$$

(2) 策略 $\gamma(t_7)$ 仅将倒数第二个半小时收益 t_7 的正负作为最后半小时买入或卖出的信号，若当日开盘跳空的幅度 t_7 大于 0，则做多，在最后半小时之前买入沪深 300 股指期货并在收盘前平仓；若当日开盘跳空的幅度 t_7 小于 0，则做空，在最后半小时之前卖出沪深 300 股指期货，并在收盘前平仓；若当日开盘跳空的幅度 t_0 等于 0，则不做任何操作。

$$\gamma(t_0) = \begin{cases} -t_8, & t_0 < 0 \\ 0, & t_0 = 0 \\ t_8, & t_0 > 0 \end{cases} \quad (3.5)$$

(3) 策略 $\gamma(t_0, t_7)$ 同时考虑隔夜收益 t_0 和倒数第二个半小时收益 t_7 作为最后半小时买入或卖出的信号，如果 t_0 和 t_7 都大于 0，则做多，在最后半小时之前买入沪深 300 股指期货并在收盘前平仓；若隔夜收益 t_0 和 t_7 都小于 0，则做空，在最后半小时之前卖出沪深 300 股指期货，并在收盘前平仓；其他情况下，则不做任何操作。

$$\gamma(t_0) = \begin{cases} -t_8, & t_0 < 0 \text{ and } t_7 < 0 \\ t_8, & t_0 > 0 \text{ and } t_7 > 0 \\ 0, & \text{otherwise} \end{cases} \quad (3.6)$$

(4) 基准策略 1，在最后半小时之前买入，收盘前直接平仓，即直接计算最后半小时的收益率。基准策略 $1 = t_8$ 。

(5) 基准策略 2，计算沪深 300 股指期货整个样本期间 2010 年 4 月 16 日至 2022 年 12 月 16 日的收益率。

本文五种策略主要通过年化收益、最大回撤、夏普比率和胜率四个指标来进行对比。

最大回撤（Maximum Drawdown）是指投资组合价值的最大下降幅度。它反映了投资者在投资期间所承受的最大损失。

计算最大回撤的方法如下：找出投资组合价值的历史最高价值（以下称“峰值”）；计算从峰值到下一个最低价值（以下称“谷值”）的回撤率，即（峰值-谷值）/峰值；对所有的回撤率求最大值，这就是最大回撤。

最大回撤具有重要的风险评估意义。它可以帮助投资者评估投资组合的风险水平，并为决策提供依据。投资者可以根据最大回撤的大小来评估自己的风险承受能力，并决定是否继续投资。

夏普比率（Sharp Ratio）是一种衡量投资组合投资回报与风险水平的指标。它通过将投资组合的收益率与其对应的市场风险比较，评估投资组合的效率和优势。

计算夏普比率的方法如下：

$$SR = \frac{E(R_p) - R_f}{\sigma_p} \quad (3.7)$$

其中 $E(R_p)$ 为投资组合的预期年化收益率， R_f 为年化无风险利率， σ_p 为投资组合年化报酬率的标准差。夏普比率越高，说明投资组合的收益率与风险水平的比率越高，投资组合效率越高。夏普比率为 0 或负数，说明投资组合的收益不高于无风险利率，投资组合效率较低。

胜率是指在所有有交易的次数中，产生正收益的个数所占的比率。

表 3.7 五种不同策略对比

| | 年化收益(%) | 最大回(%) | 夏普比率 | 胜率(%) |
|--------------------|---------|--------|--------|-------|
| $\gamma(t_0)$ | 5.115 | 14.902 | 0.968 | 58.61 |
| $\gamma(t_7)$ | 6.727 | 16.255 | 1.274 | 60.57 |
| $\gamma(t_0, t_7)$ | 5.861 | 7.518 | 1.549 | 75.40 |
| 基准策略 1 | -9.446 | 69.903 | -1.237 | 47.32 |
| 基准策略 2 | 4.412 | 6.151 | 0.18 | - |

从表 3.7 中可以看出，采用了动量效应的策略都产生较高的年化收益。其中第一个策略 $\gamma(t_0)$ 以 t_0 作为做多或做空的策略信号，年化收益为 5.115%，第二个策略 $\gamma(t_7)$ 以 t_7 作为做多或做空的策略信号，年化收益为 6.727%，同时考虑隔夜收益 t_0 和倒数第二个半小时收益 t_7 的策略的收益为 5.861%，而基准策略 1、基准策略 2 的年化收益分别为 -9.446%、4.412%。单从收益的角度来说，第二个策略的年化收益是最高的。基准策略一是在最后

半小时买入，在收盘前平仓，其收益非常低，达到年化-9.446%。可能有以下原因：（1）投资者的套利行为，有些投资者可能会利用期货市场与现货市场的价格差距进行套利，在最后半小时买入期货并卖出现货，这种行为可能导致期货市场价格下跌；（2）调整现货指数，有些投资者可能会在最后半小时调整现货指数，即投资者在市场收盘前对其现货指数类投资组合进行调整，以实现特定的投资目标，使得指数更接近期货市场价格，这种调整也可能导致期货市场价格下跌；（3）结算行为：在期货交易中，投资者需要在结算日前对持有的合约进行结算，如果有许多投资者需要同时结算，这可能导致期货市场价格下跌。

从最大回撤来看，最低的是基准策略 2，为 6.151%，其次是第三个策略 $\gamma(t_0, t_7)$ ，为 7.518%，最高的为基准策略 1。表明第三个动量策略和基准策略 2 的风险都是比较低的。从夏普比率来看，最高的是第三个动量策略 $\gamma(t_0, t_7)$ ，其次是第二个动量策略 $\gamma(t_7)$ ，分别为 1.549、1.274，说明这两种策略投资组合的收益率与风险水平的较高，投资组合较越高，动量策略能够获得显著的超额收益。从胜率上来看，最高的为第三种策略，由于第三种策略的交易条件相对于第一、二种策略的条件更为严苛，所以实际交易次数相对较少，获得正收益的次数更多。

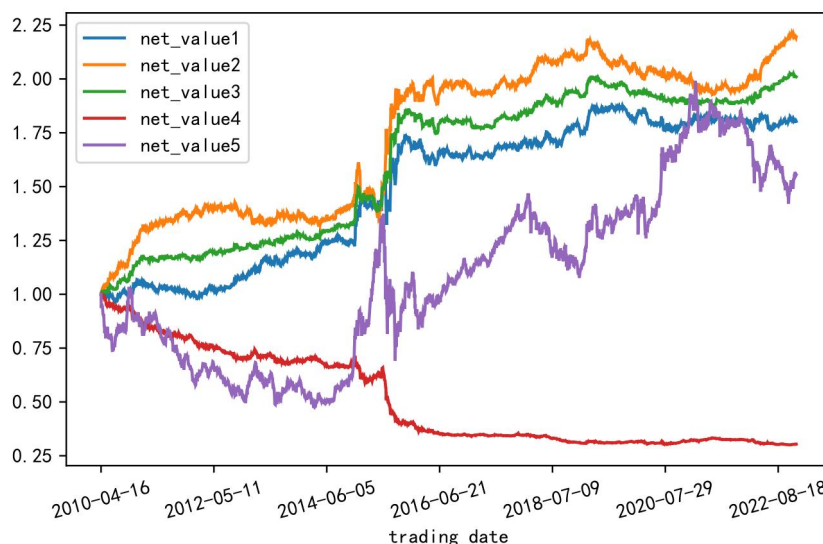


图 3.2 不同策略净值曲线图

图 3.2 为上文五种不同策略的净值曲线图，net_value1，net_value2，net_value3，为三种动量策略，net_value4 和 net_value5 为两种基准策略。可以看出三种动量策略的净值曲线是显著高于两种基准策略，且波动相较于基准策略更小。

综上所述，五种策略中，表现最好的是第三种动量策略 $\gamma(t_0, t_7)$ ，在 t_0 和 t_7 都大于 0

时，做多，在隔夜收益 t_0 和第七个半小时收益 t_7 都小于 0 时做空，其他情况下不进行交易，并在收盘前平仓。取得了五种策略中，第二高的年化收益，第二低的最大回测，最高的夏普比率和最高的胜率，表明我们能够根据动量效应获得显著的超额收益。

第五节 现象解释

从上述实证分析来看，沪深 300 股指期货样本期间内存在着日内动量效应。从实际应用的角度来看，考虑这种现象背后的经济力量也是有价值的。

从本章的第四节中可以得知，在高成交量的情况下，日内动量效应更加显著，U 形的模式反映了在第一个半小时内新信息的消化以及在最后一个半小时内交易的愿望，以便结算并避免隔夜风险。本文认为主要有以下几种原因：

基于行为金融的理论模型和假说认为，动量效应是由市场参与者的非理性行为或有限理性行为导致的，即市场参与者对新信息反应不足或过度反应，从而导致股票价格偏离其内在价值。这些非理性或有限理性行为可能包括代表性启发式（Representativeness Heuristic）、保守主义（Conservatism）、自信过度（Overconfidence）、反馈交易、羊群效应（Herding Effect）等等。这些行为偏差会影响市场参与者（Feedback Trading）对股票未来收益率的预期，从而产生动量效应。

代表性启发式是指人们根据一些特征或样本来判断一个事件或对象属于某一类别的概率而忽视了先验概率或样本大小等信息。在股票市场中，代表性启发式可能导致投资者对过去收益率较高或较低的股票形成过于乐观或悲观的预期，并且认为这种趋势会持续下去。保守主义是指人们在面对新证据时倾向于坚持原有的信念，并且对新证据反应不足。在股票市场中，保守主义可能导致投资者对新信息没有及时地更新他们对股票内在价值的估计，并且低估了新信息对未来收益率变化的影响。

Bogousslavsky(2016)强调了不频繁的投资组合再平衡在金融中的作用。Duffie(2010)则指出，缓慢移动的资本和不频繁的决策也是重要因素。同时，Murphy（2013）和Thirumalai（2016）的研究为不频繁投资组合再平衡提供了具体证据，表明机构投资者进行重复交易的情况比较普遍。Bogousslavsky（2016）的研究着眼于解释回报自相关性和季节性，他强调不频繁的再平衡在日内动量现象中扮演的重要角色。他的研究理论上证明了投资者只需将再平衡交易延迟到市场收盘而不是市场开盘，就能驱动日内动量。因此，在交易日的最后半小时内，如果交易方向与第一个半小时的交易方向相同，将导致两个回报之间的正相关性。

另一种解释认为,由于投资者信息处理的不同,有些人可能会比其他人更晚地了解市场上的好消息。这些“后知后觉”的投资者可能无法在市场上最初推高价格的前半小时做出反应,而是在较晚的时间才能追逐购买。在股指期货交易中,最后半小时往往是最活跃的时间段,因此选择在这个时间段进行交易是最佳的选择之一。如果交易方向与前半小时市场趋势相同,就有可能产生正回报和正相关性。一些共同基金也倾向于在接近收盘价的价格交易,这不仅有利于日终结算,还对基于收盘价格的实证因子模型的应用有利。共同基金投资者只能在收盘时进行交易,因此他们也可以被视为“后知后觉”者。

此外,在熊市、高成交量、高波动率的市场条件下,日内动量效应更显著,可能有以下原因:

在熊市市场条件下,股票市场整体处于下跌趋势,投资者信心下降,风险厌恶情绪加强。在这种情况下,投资者更加关注股票价格的变化,期望通过短期交易获得更多的收益。因此,日内动量效应在熊市下更加明显。此外,由于投资者普遍预期市场将继续下跌,交易活跃度和成交量较高,更多的投资者选择通过日内交易获得收益,从而加剧了日内动量效应的出现。熊市下的日内动量效应,与牛市相比,更具有持久性和普适性。

成交量与价格趋势密切相关。当成交量高时,通常意味着市场参与者的交易活跃度增加,市场流动性增强。此时,市场价格可能会出现明显的上升或下降趋势,即出现较强的趋势性。其次,成交量也与市场信息的披露和反应有关。当市场重要信息发布后,高成交量的情况下,市场参与者对于该信息的反应较为一致,市场反应更为迅速和明显,可能形成显著的价格趋势,从而出现较强的日内动量效应。最后,成交量还反映了市场的交易压力和市场参与者的信心水平。高成交量往往意味着市场参与者的交易压力较大,市场的波动性增加,也可能反映市场参与者的信心水平较高,市场热度较高。在这种情况下,市场参与者的交易决策更容易受到情绪和群体效应的影响,可能导致出现较为明显的价格趋势和日内动量效应。

高波动率表示市场价格的波动程度比较大,因此当价格出现变动时,其幅度也会比较大,这就为日内动量效应的出现提供了更多的机会。在高波动率的情况下,交易者更容易捕捉到价格变化的信号,并通过买入或卖出来获得更高的收益。此外,高波动率通常伴随着市场情绪的波动,也就是说,投资者的情绪和信念在价格形成中的作用变得更加重要。当市场情绪较为明显时,交易者更容易受到情绪的影响,进而出现过度买入或过度卖出的情况,从而导致日内动量效应的出现。

这些解释提供了强有力的统计证据支持日内动量现象的存在。其他可能的解释还需进一步进行研究。

第六节 稳健性检验

稳健性检验，也称为鲁棒性检验，是为了检验模型的稳定性和适用性。这是因为在实际的数据分析和建模过程中，数据可能存在偏差和异常值，这些值可能会对模型产生不利影响，使得模型的结果不准确。稳健性检验可以帮助我们了解模型对于异常值和偏差数据的敏感程度，并评估模型的性能如何受到这些数据的影响。通过进行稳健性检验，可以确定模型是否需要进一步的改进，以提高模型的稳健性。

为此，本文选取 2010 年 4 月 16 日至 2022 年 12 月 16 日共计 3081 个工作日沪深 300 指数 1 分钟频率的数据，再按照本章第二节中，公式 3.1 和 3.2 进行处理，将数据处理成每天八个半小时的收益率数据。然后根据下列公式进行回归分析：

$$t_{i,j} = \alpha_j + \beta_j t_{i,j} + \epsilon_j \quad j = 0, 1, \dots, 8; i = 1, 2, \dots, T \tag{3.8}$$

表 3.8 沪深 300 指数日内动量效应回归分析

| | t_0 | t_1 | t_2 | t_3 | t_4 | t_5 | t_6 | t_7 |
|-----------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| α | 0.0263*** (3.6023) | 0.0210*** (2.8740) | 0.0224*** (3.0771) | 0.0226*** (3.1048) | 0.0226*** (3.1049) | 0.0226*** (3.1039) | 0.0225*** (3.0976) | 0.0236*** (3.2991) |
| β_0 | 0.0489*** (4.3831) | | | | | | | |
| β_1 | | 0.0305** (2.2256) | | | | | | |
| β_2 | | | 0.0235 (1.2857) | | | | | |
| β_3 | | | | -0.0019 (-0.1013) | | | | |
| β_4 | | | | | 0.0018 (0.1113) | | | |
| β_5 | | | | | | 0.0134 (0.7244) | | |
| β_6 | | | | | | | 0.0947*** (4.7350) | |
| β_7 | | | | | | | | 0.1980*** (10.9130) |
| R^2 | 0.618 | 0.160 | 0.054 | 0.000 | 0.000 | 0.017 | 0.722 | 3.721 |

注：表中***，**，*分别表示在 1%，5%，10%的置信水平下显著，括号中为 T 值

如上表所示, t_0 , t_6 , t_7 在显著性水平为 1% 的情况小, 都具有统计学意义, R^2 分别为 0.618%, 0.722%, 3.721%, 与本章第二节中实证结果一致, 且倒数第二个小时具有更高的 R^2 , 说明在沪深 300 指数中也存在着显著的动量效应。

本章检验了沪深 300 股指期货市场 2010 年 4 月 16 日至 2022 年 12 月 16 日日内动量效应的存在性, 以及在不同情况下日内动量效应的存在性。本文第一个半小时对最后半小时收益具有正向预测性, 而是发现了隔夜收益和倒数第二个半小时对最后半小时具有正向预测性, 在 1% 的显著性水平下具有统计学意义。本文也给出了两个可能的解释: 投资者不频繁的资产组合调整和“后知后觉”的投资者的存在。另外还发现在熊市、高成交量组或波动率的市场条件下, 日内动量效应更加显著, 本文也给出了相应的解释。最后本文利用同样的样本区间检验了沪深 300 指数日内动量效应的存在性, 检验结果与股指期货市场一致。

第四章 深度学习动量策略实证分析

沪深 300 股指期货是一个重要的行业指数，反映了中国 A 股市场的整体表现，它也是一个有效的风险管理工具，可以用作投资多样性的工具，与其他金融工具配合使用，以增加投资组合的多样性和风险分散能力。股指期货市场的波动性往往比现货市场要更高，而且往往是非线性的。

在第四章中，本文根据隔夜收益 t_0 和第七个半小时收益 t_7 对最后半小时收益 t_8 的正向预测性，构建了三个简单的动量策略和两种基准策略。三种动量策略都取得了较高的年化收益，本章将进一步把动量效应与机器学习和深度学习的模型结合起来，看能否获得更高的超额收益。

本文从米筐量化交易平台选取了样本期间为沪深 300 股指期货从开市以来 2010 年 4 月 16 日至 2022 年 12 月 16 日共计 3081 个工作日的高频数据，其中包括每日每分钟的开盘价、收盘价、成交量等。

第一节 特征选择

特征工程是机器学习中的一个重要环节，它决定了模型的训练效果。特征工程的部分，需要考虑以下几点：数据预处理：清洗数据、缺失值处理、数据标准化等；特征选择：选择有意义的特征，去除无关或冗余的特征，以提高模型的准确率；特征抽取：从原始数据中提取新的特征，例如统计特征、图像特征等；特征变换：对原始特征进行变换，例如归一化、离散化等；特征组合：将多个特征组合成一个新的特征，以便更好地表示数据等。

本文选取的特征主要有沪深 300 股指期货相关数据共选取了 23 个特征进行预测，如下表 4.1。特征 1 到 8 是利用每半小时的开盘价与收盘价计算出前七个半小时的收益率 $t_i (i=1, \dots, 7)$ ，以及 t_0 前一日收盘价与今日开盘价差，即隔夜收益率。

本文第三章中检验了成交量的高低对于动量效应具有显著影响，成交量越高动量效应月显著，为提现这种影响，本文特征 9 到 15 是每半小时成交量相对于前五个工作日的变化率，具体计算方法如下：

$$\Delta V_{i,j} = \frac{V_{i,j}}{\overline{V_{i,j}}} - 1 (i \geq 5, j = 1, \dots, 7) \quad (4.1)$$

其中， $V_{i,j}$ 表示第 i 日第 j 个半小时的成交量， $\overline{V_{i,j}}$ 表示第 i 日的前五个工作日第 j 个

半小时的成交量的平均值，因此 $\Delta V_{i,j}$ 表示第 i 日第 j 个小时与前五个工作日第 j 个半小时的成交量的平均值的变化率。

由于波动率的大小也影响动量效应的显著性，特征 16 到 22 为每半小时的波动率 $\sigma_i (i=1, \dots, 7)$ ，具体计算方法是选取每个工作日中每半小时股指期货价格的标准差，以此作为波动率的大小。

第 23 个特征为市场状况，即牛市或者熊市，根据第三章第四节中的分类方法进行分类，牛市记为 1，熊市记为 0。

被解释变量为沪深 300 股指期货最后半小时的收益，即 t_8 ，若收益为正则记为 1，收益为负或者 0 则记为 0。因此本文策略部分即成为一个二分类预测的问题。

表 4.1 特征选取

| 特征 | 含义 |
|--|----------------------|
| t_0 | 隔夜收益率 |
| $t_i (i=1, \dots, 7)$ | 每日前 7 个半小时的收益率 |
| $\Delta V_{i,j} (i \geq 5, j = 1, \dots, 7)$ | 每半小时成交量相对于前五个工作日的变化率 |
| $\sigma_i (i=1, \dots, 7)$ | 价格波动率 |
| M | 市场状况，牛市为 1，熊市为 0 |

第二节 基于 XGBoost 的动量策略

本文第二章中详细介绍的 XGBoost 算法的原理和优缺点等。现在将该算法进行实际运用。将全部样本的前 80%作为训练集，后 20%作为测试集，由于本文计算成交量的变化率时运用了前五个工作日的数据，因此，训练集期间为 2010 年 4 月 22 日至 2020 年 6 月 8 日共计 2461 个工作日，测试集期间为 2020 年 6 月。首先用 scikit-learn 中 XGBoost 的默认参数直接进行建模，然后再进行超参数优化。

一、贝叶斯超参数优化

常见的调参方法有网格搜索、随机调参等，本文所使用的超参数优化方法是使用贝叶斯优化算法进行调参。贝叶斯调参是一种基于概率理论的模型调参方法。它利用贝叶斯定理来评估模型的不确定性，并通过更新先验概率分布来不断优化模型的参数。

在机器学习中，模型的参数往往有很多，且对于不同的数据集，参数的最佳取值可能不同。因此，调参过程需要对模型参数进行试错和优化。而贝叶斯调参方法通过使用概率模型，可以在调参过程中更加科学地评估模型参数的不确定性，从而避免随意试错

带来的误差。

首先需要选择一个合适的概率模型，并且对参数设定一个合适的先验分布。然后，根据 XGBoost 的算法原理，使用交叉验证的方法评估模型的性能，并且得出参数的后验分布。最后，根据参数的后验分布和先验分布，选择一组合适的参数，以获得最佳的模型性能。

在 XGBoost 中，常用的参数包括学习率、树的数量、树的深度等。在使用贝叶斯调参方法时，可以对这些参数分别设定一个合适的先验分布，然后利用交叉验证的方法评估模型的性能，得出这些参数的后验分布。最后，根据参数的后验分布和先验分布，选择一组合适的参数，以获得最佳的模型性能。

总的来说，使用贝叶斯调参方法对 XGBoost 进行调参是一种非常有效的方法，可以提高模型的性能，并且具有较好的泛用性。在贝叶斯调参方法中，可以明确地考虑参数的不确定性，从而更准确地评估模型的性能。此外，贝叶斯调参方法还可以通过计算参数的后验概率分布来选择最优的参数。

表 4.2 XGBoost 超参数优化

| 参数名 | 含义 | 范围 | 最佳参数值 |
|------------------|-------------------|------------|--------|
| colsample_bytree | 每棵树在随机采样的特征数量 | (0.5,0.99) | 0.934 |
| gamma | 节点分裂时对损失函数的最小改进需求 | (0,5) | 0.549 |
| learning_rate | 学习率 | (0.01,0.5) | 0.412 |
| max_depth | 树的最大深度 | (1,6) | 3 |
| min_child_weight | 决策树分裂所需的最小样本权值和 | (2,50) | 14.551 |
| n_estimators | 弱学习器的数量 | (50,200) | 110 |
| reg_alpha | L1 正则项的系数 | (0,1) | 0.553 |
| reg_lambda | L2 正则项的系数 | (0,1) | 0.625 |
| subsample | 每棵树随机采样的样本数量 | (0.7,0.9) | 0.715 |

二、模型评估

在机器学习模型中，对于模型性能的评估有一套特定的方法。本文采用的交叉验证（Cross-validation）是一种常用的评估模型性能的方法，它通过将训练数据分为若干个互不重叠的子集，并利用每一个子集分别训练和验证模型的方法来评估模型的性能。本文采用的是 10 折交叉验证。

在交叉验证中，通常采用 K-fold 交叉验证的方法。其中，K 表示将训练数据分为 K 份，每一份作为验证数据，其余 K-1 份作为训练数据。每一次训练和验证都使用不同的

验证数据，最终得到 K 次评估结果，并通过求平均值等方法得到最终的评估结果。交叉验证的优点在于，它考虑了模型的泛化能力，避免了过拟合的问题。此外，交叉验证的评估结果更具有可靠性和可比性，可以更准确地评估模型的性能。总的来说，交叉验证是一种很好的评估模型性能的方法，对于评估 XGBoost 模型的性能非常有效。通过交叉验证，可以准确地评估模型的泛化能力，避免过拟合的问题，从而提高模型的性能。

模型评价指标主要有一下几个：准确率、精确率、召回率 ROC 曲线和 AUC 等。

由于在股指期货中我们可以通过做多和做空来获取收益，因此本文更关心模型预测的准确率。本文所采用的评估指标主要有准确率和 AUC 值，训练集准确率 75.66%，测试集准确率 52.11%，AUC51.93%。此外再根据预测结果计算出策略的相关指标如下表 4.3，最终基于 XGBoost 的动量策略年化收益为 6.336%，最大回撤为 2.564%，夏普比率为 1.507，相对于本文第四章中三种动量策略拥有更好的收益和更低的风险。

表 4.3 基于 XGBoost 的日内动量策略

| | 年化收益率(%) | 最大回撤(%) | 夏普比率 | 胜率(%) |
|---------|----------|---------|-------|--------|
| XGBoost | 6.336% | 2.564 | 1.507 | 52.11% |

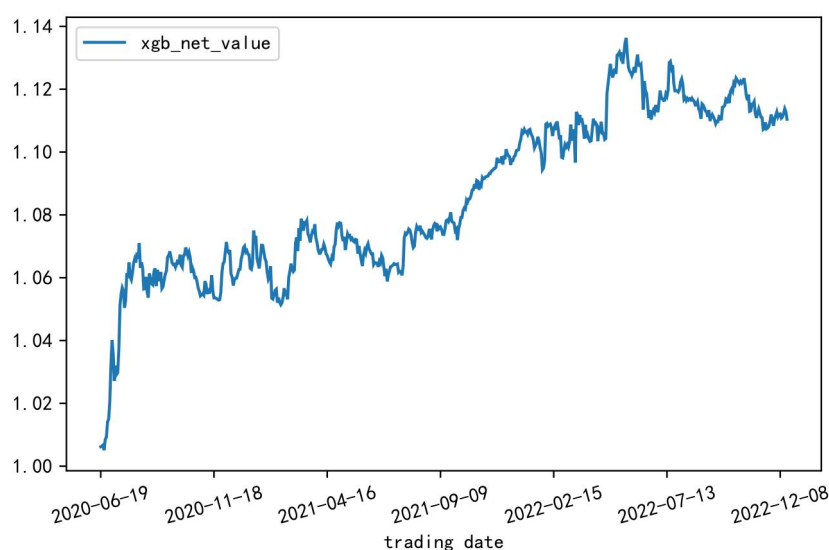


图 4.1 基于 XGBoost 的日内动量策略净值曲线图

表 4.2 和图 4.1 分别为基于 XGBoost 的日内动量策略的相关指标和净值曲线图，净值曲线的总体趋势是在增长的，但在 2020 年 7 月至 2021 年 9 月之间比较平缓，整体回撤较低。

第三节 基于 LSTM 的动量策略

首先, LSTM 是一种基于循环神经网络的深度学习模型, 特别适合于处理序列数据, 如股票价格序列数据。LSTM 可以通过记录长期依赖关系和消除短期干扰来保存和提取历史信息, 因此, 对于时间序列相关的问题, LSTM 往往比其他模型更有效。

相反, XGBoost 是一种基于决策树的集成学习模型, 特别适合于处理结构化数据。XGBoost 通过在决策树之间构建弱分类器来提高模型的准确性和泛化能力。由于 XGBoost 适用于结构化数据, 因此, 在处理非时间序列相关的问题时, 它往往比 LSTM 更有效。

因此, 在股票预测涨跌问题中, LSTM 往往更适合于处理历史价格数据, 而 XGBoost 往往更适合于处理经济、市场等因素数据。由于本文的数据主要是每日收益率、成交量和价格波动率的时间序列数据, 理论上来说基于 LSTM 的日内动量策略的效果比 XGBoost 要更好。

一、数据选取与处理

本节仍然选用第一节中的数据进行建模预测。LSTM 网络中的神经元节点是基于数值计算的, 因此输入数据的数值范围必须是统一的。如果输入数据中有一个特征的数值范围比其他特征大得多, 那么这个特征对网络的影响会远远大于其他特征, 导致模型的不稳定。因此, 在使用 LSTM 建模预测股票涨跌前, 对数据进行标准化或归一化是很重要的。本文将除了牛熊市情况以外的特征全部进行标准化处理, 以便所有特征的数值范围都是相同的, 帮助模型较好地捕捉输入数据中的信息。

二、贝叶斯超参数优化

本文在训练模型时, 使用滑动窗口预测, 使用前九天的数据以及当天前七个半小时的相关数据来对最后一个半小时的收益进行预测。本节仍然使用贝叶斯调参对 LSTM 模型的相关参数进行优化。LSTM 需要调整参数主要有: 隐藏层大小、批量大小、学习率等。具体如下表:

表 4.4 LSTM 超参数优化

| 名称 | 含义 | 范围 | 最佳参数 |
|-------------------|-----------|---------------------------|-------|
| batch_size | 每次训练的样本数量 | 通常为 2 的整数次幂 | 32 |
| epochs | 训练循环次数 | 整数，一般在 50-200 之间 | 50 |
| hidden_layer_size | 隐藏层向量的维度 | 整数，一般在 64-512 之间 | 128 |
| learning_rate | 学习率 | (0,1) | 0.001 |
| optimizer | 优化器算法 | 'adam', 'rmsprop', 'sgd'等 | adam |
| num_layers | LSTM 层数 | (1,5) | 2 |

三、 模型评估

经过模型训练及参数优化，最终测试集准确率为 56.178，AUC 为 56.22%。策略具体表现如下表 4.5，年化收益为 9.136%，为所有策略中最高的年化收益率，最大回撤为 4.935%，夏普比率为 2.164，也是所有策略中最高的夏普比率，说明基于 LSTM 的日内动量策略具有较低的风险和较高的收益。

表 4.5 基于 LSTM 的日内动量策略

| | 年化收益率(%) | 最大回撤(%) | 夏普比率 | 胜率(%) |
|------|----------|---------|-------|--------|
| LSTM | 9.136 | 4.935 | 2.164 | 56.178 |

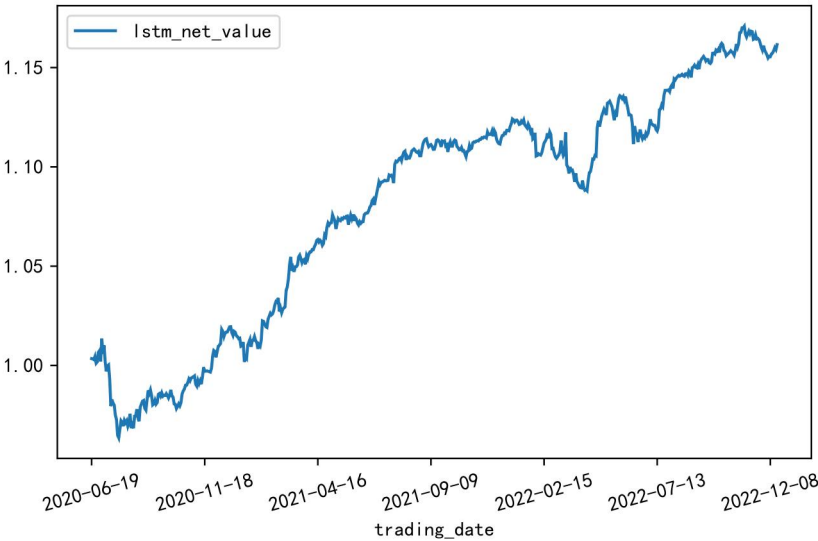


图 4.2 基于 LSTM 的日内动量策略净值曲线图

上图 4.2 为基于 LSTM 的日内动量策略的净值曲线图，可以看出净值整体上是增加，且斜率要大于基于 XGBoost 的日内动量策略，增长得更快。

第四节 本章小结

通过第三章日内动量效应的存在性检验,本文发现了隔夜收益与倒数第二个半小时对最后一个半小时收益的正向预测性,即具有动量效应。然后据此在第四章中,分别以 t_0 、 t_7 以及两者结合起来作为交易信号建立动量策略。在第四章和第五章中分别用 XGBoost 模型和 LSTM 模型进行预测,建立日内动量策略。策略对比结果如下表 4.6:

从下表中可以看出,从年化收益来说,最高的是 LSTM,其次是 $\gamma(t_7)$,然后是 XGBoost,年化收益最低的是基准策略 1,其次是基准策略 2。从最大回撤的角度来说,最低的是 XGBoost 策略,其次是 LSTM 策略,最高的是基准策略 1。从夏普比率的角度来说,最高的是 LSTM 策略,其次是 $\gamma(t_0, t_7)$ 和 XGBoost。

表 4.6 策略对比

| | 年化收益率(%) | 最大回撤(%) | 夏普比率 | 胜率(%) |
|--------------------|----------|---------|--------|-------|
| $\gamma(t_0)$ | 5.115 | 14.902 | 0.968 | 58.61 |
| $\gamma(t_7)$ | 6.727 | 16.255 | 1.274 | 60.57 |
| $\gamma(t_0, t_7)$ | 5.861 | 7.518 | 1.549 | 75.40 |
| XGBoost | 6.336 | 2.564 | 1.507 | 52.11 |
| LSTM | 9.136 | 4.935 | 2.164 | 56.18 |
| 基准策略 1 | -9.446 | 69.903 | -1.237 | 47.32 |
| 基准策略 2 | 4.412 | 6.151 | 0.18 | - |

由于前三个动量策略 $\gamma(t_0)$ 、 $\gamma(t_7)$ 和 $\gamma(t_0, t_7)$ 是使用样本内数据进行测算年化收益、最大回撤、夏普比率和胜率,而 XGBoost 和 LSTM 是用样本外的数据进行预测然后计算。因此实际上效果最好的仍然是基于 LSTM 和 XGBoost 的日内动量策略最好。需要特别说明的是,基准策略 1 仅计算了最后半小时涨跌作为收益测算,而会有很多原因导致沪深 300 股指期货市场最后半小时价格下跌,例如:投资者的套利行为,有些投资者可能会利用期货市场与现货市场的价格差距进行套利,在最后半小时买入期货并卖出现货,这种行为可能导致期货市场价格下跌;调整现货指数,有些投资者可能会在最后半小时调整现货指数,即投资者在市场收盘前对其现货指数类投资组合进行调整,以实现特定的投资目标,使得指数更接近期货市场价格,这种调整也可能导致期货市场价格下跌;结算行为,在期货交易中,投资者需要在结算日前对持有的合约进行结算,如果有许多投资者需要同时结算,这可能导致期货市场价格下跌。

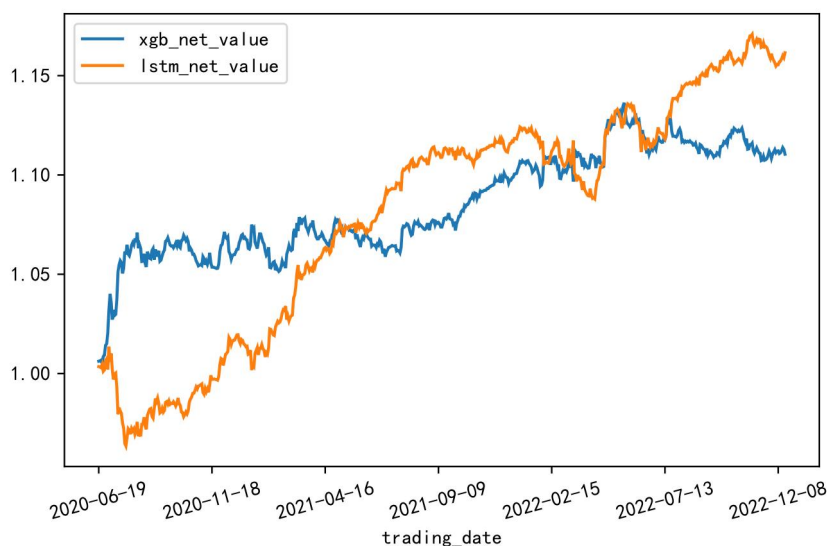


图 4.3 两种机器学习日内动量策略净值曲线对比图

从上图中可以看出，在测试集为 2020 年 6 月 19 日至 2022 年 12 月 16 日区间内，基于 XGBoost 和 LSTM 的日内动量策略都取得了不错的效果。2020 年 6 月至 2021 年 5 月之间，基于 XGBoost 的策略要好于基于 LSTM 的策略，但从总体趋势来看，基于 LSTM 的策略增长更快。

综上所述，如果以最高收益为目标，LSTM 是最好的选择，如果更关注风险控制，则 XGBoost 策略可能是更优的选择；如果需要在收益和风险之间寻找平衡，LSTM 则更胜一筹。

第五章 总结与展望

本文通过文献研究法和实证研究法对基于 LSTM 模型的日内动量策略进行了研究，本章将对全文进行总结。

第一节 研究总结

沪深 300 股指期货作为一种衍生金融工具，自 2010 年 4 月 16 日设立以来，就受到了广大投资者和学者的追捧。沪深 300 股指期货不仅能使金融资本市场实现多元化发展，变得更加丰富，而且更为关键的作用是对风险进行有效的防控，同时使金融资本市场能够更好的抵御市场风险。因此研究沪深 300 股指期货意义重大。

本文通过第一章中的文献研究发现关于动量效应的文献主要分为两类，一种是横截面动量，即买入赢家组合卖出输家组合，更侧重于选股；另一种是时间序列动量，专注于资产本身的过去回报，建立包含所有正在考虑的证券的投资组合，更侧重于择时。而本文利用沪深 300 股指期货作为研究标的，建立时间序列动量策略，属于第二种择时策略。另外，从动量效应的形成期来说，现有文献对动量效应的研究主要集中在中长期，如年、月、周等，本文则将动量效应的形成期限定在一日以内，发现了沪深 300 股指期货的日内动量效应。

本文第二章介绍了相关的理论算法，XGBoost 能够处理多类别分类、回归、排名和其他机器学习任务。优势在于其高效性和准确性，可以并行处理数据并在大规模数据集上快速训练，其算法的优化可以最小化计算资源的使用，从而实现了高效率的训练和预测。此外，XGBoost 还具有强大的特征选择功能，可以自动选择重要的特征进行训练和预测，从而提高了模型的准确性。LSTM 是一种特殊的循环神经网络，它能够处理具有时间序列结构的数据。相对于传统的循环神经网络，LSTM 具有更好的记忆能力，可以学习和保存更长时间的上下文信息，从而更好地处理时间序列数据。LSTM 的优势在于其能够处理非常复杂的序列模式，具有更好的长期记忆和序列预测能力。此外，LSTM 还具有抗噪声能力和对数据分布变化的适应性，可以适用于各种不同的时间序列预测任务。贝叶斯调参是一种基于贝叶斯理论的超参数调优方法，相比于传统的网格搜索、随机搜索等调参方法，具有高效性、自适应性、鲁棒性和可解释性等优点。

第三章检验了沪深 300 股指期货市场上日内动量效应的存在性，选取的数据是 2010 年 4 月 16 日至 2022 年 12 月 16 日沪深 300 股指期货主力连续合约的相关数据。通过检验，与其他学者不同的是，本文并未发现第一个半小时收益率对最后一个半小时收益率

的具有显著作用，而发现隔夜收益率即隔夜效应即倒数第二个半小时收益对最后半小时收益具有显著的正向预测作用，此外，还发现在熊市的情况下，日内动量效应更加显著，以及在高波动率及高成交量的情况下，日内动量效应更加显著。根据检验结果，本文建立三个日内动量策略，第一个是以隔夜收益 t_0 的正负作为做多或做空的信号，第二个是以 t_1 的正负为做多或做空的信号，第三个是以 t_0 和 t_1 联合起来作为做多或做空的信号。三种策略分别取得了 5.115%、6.727%、5.861%的年化收益。

在第四章中，本文以每日半小时收益率、成交量的变动率和价格波动率作为特征，利用 XGBoost 和 LSTM 模型对最后半小时的涨跌进行分类预测，取得了远超上述三个动量策略的年化收益率，分别是 6.336%和 9.136%，同时具有更低的回撤的和更高的夏普比率，收益高的同时风险更低。

本文的研究话题对于期货市场的动量效应具有重要意义，对于投资者也具有实际价值，能够给投资者以实际的参考作用。通过研究市场动量效应，可以对市场数据进行深入分析，对市场趋势和走势有更清晰的了解。研究使用 XGBoost 和 LSTM 算法预测市场动量效应是技术探索的一个重要方面，也可以帮助开发更加高效的金融预测算法。

第二节 研究不足与展望

本文利用机器学习和深度学习的算法构建了日内动量策略，取得了较高的收益同时风险相对较低。但仍然存在许多不足之处，具体如下：

XGBoost 和 LSTM 都是目前比较主流的算法，且本文都取得了比较好的效果，但目前有很多更新效果更好的深度学习算法，例如 Attention、强化学习等。未来可以考虑利用更加有效的算法进行预测。

影响日内动量效应的因素除了成交量、波动率和市场情况以外还有很多其他因素，如投资者的情绪：当投资者情绪悲观时，市场动量效应可能变得更强，反之亦然；股票流通性：股票的流通性对市场动量效应有影响，如果股票流通性很高，这可能会导致市场动量效应变弱；信息披露：信息披露对市场动量效应有影响，如果有重要的信息披露，这可能会引起市场动量效应的变化等；未来可以进一步将这些因素引入模型之中。

本文仅检验了沪深 300 股指期货的日内动量效应，日内动量效应可能还存在与其他股票市场或者期货市场之中。可以进一步进行探索。

虽然存在以上不足，但是本文也为广大投资者提供了新的思路，本文所取得的收益和风险控制具有实践意义。

参 考 文 献

- [1] Abadi M, Agarwal A, Barham P, et al. Tensorflow: Large-scale Machine Learning on Heterogeneous Distributed Systems[J]. arXiv preprint arXiv:1603.04467, 2016.
- [2] Antonacci G. Dual momentum investing : an Innovative Strategy for Higher Returns with Lower Risk[M]. McGraw-Hill Education, 2015.
- [3] Baltas N, Kosowski R. Demystifying Time-series Momentum Strategies: Volatility Estimators, Trading Rules and Pairwise Correlations[J]. Market Momentum: Theory and Practice", Wiley, 2020.
- [4] Baz, J, N. Granger, H. Nicolas, et al. "Dissecting Investment Strategies in the Cross Section and Time Series." SSRN, 2015.
- [5] Bengio Y, Lamblin P, Popovici D, et al. Greedy Layer-wise Training of Deep Networks[J]. Advances in Neural Information Processing Systems, 2006, 19.
- [6] Bogousslavsky V. Infrequent Rebalancing, Return Autocorrelation, and Seasonality[J]. Journal of Finance, 2016, 71(6):2967-3006.
- [7] Buehler H, Gonon L, et al. Deep hedging[J]. Quantitative Finance, 2019, 19(8): 1271-1291.
- [8] Cheema M A, Nartea G V. Momentum Returns and Information Uncertainty: Evidence from China[J]. Pacific-Basin Finance Journal, 2014, 30: 173-188.
- [9] Dunis C L, Huang X. Forecasting and Trading Currency Volatility: An Application of Recurrent Neural Regression and Model Combination[J]. Journal of Forecasting, 2002, 21(5): 317-354.
- [10] Erb C B, Harvey C R. The Strategic and Tactical Value of Commodity Futures[J]. Financial Analysts Journal, 2006, 62(2): 69-97.
- [11] Fama E F, French K R. Dissecting Anomalies[J]. The Journal of Finance, 2008, 63(4):1653-1678.
- [12] Fraccaro M, Kamronn S, Paquet U, et al. A Disentangled Recognition and Nonlinear Dynamics Model for Unsupervised Learning[J]. Advances in Neural Information Processing Systems, 2017, 30.
- [13] Gao L, Han Y, Li S Z, et al. Market Intraday Momentum[J]. Journal of Financial Economics, 2018, 129(2): 394-414.
- [14] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Advances in Neural Information Processing Systems[J]. Curran Associates, Inc, 2014, 27: 2672-2680.
- [15] Gorton G B, Hayashi F, Rouwenhorst K G. The Fundamentals of Commodity Futures Returns[J]. Review of Finance, 2013, 17(1): 35-105.

- [16] Gray W, Vogel J. Quantitative Momentum: A Practitioner's Guide to Building A Momentum-based Stock Selection System[M]. John Wiley & Sons, 2016.
- [17] Gu S, Kelly B T, Xiu D. Empirical Asset Pricing Via Machine Learning[J]. Chicago Booth Research Paper, 2019 (18-04): 2018-09.
- [18] Heaton J B, Polson N G, Witte J H. Deep Learning for Finance: Deep Portfolios[J]. Applied Stochastic Models in Business and Industry, 2017, 33(1): 3-12.
- [19] Hinton G E, Osindero S, Teh Y W. A Fast Learning Algorithm for Deep Belief Nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [20] Hochreiter S, Schmidhuber J. Long Short-Term Memory[M]. Berlin: Springer, 1997: 1735-1780.
- [21] Hong H, Lim T, Stein J C. Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies[J]. The Journal of Finance, 2000, 55(1): 265-295.
- [22] Huang N E, Shen Z, Long S R, et al. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis[J]. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 1998, 454(1971): 903-995.
- [23] Jegadeesh N, Titman S. Cross-sectional and Time-series Determinants of Momentum Returns[J]. The Review of Financial Studies, 2002, 15(1): 143-157.
- [24] Jegadeesh N, Titman S. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency[J]. Journal of Finance, 1993, 48(1): 65-91.
- [25] Kang J, Liu M, Ni S. Contrarian and Momentum Strategies in The China Stock Market: 1993-2000[J]. Pacific-Basin Finance Journal, 2002, 10(3): 243-265.
- [26] Ke J, Zheng H, Yang H, et al. Short-term Forecasting of Passenger Demand Under On-demand Ride Services: A Spatio-temporal Deep Learning Approach[J]. Transportation Research Part C: Emerging Technologies, 2017, 85: 591-608.
- [27] Kho B. Time-Varying Risk Premia, Volatility, and Technical Trading Rule Profits: Evidence from Foreign Currency Futures Markets[J]. Journal of Financial Economics, 1996, 41(2): 249-290.
- [28] Kim, S. 2019. "Enhancing the Momentum Strategy through Deep Regression." Quantitative Finance 19 (7): 1121–1133.
- [29] Kingma D P , Ba J L. Adam: A Method for Stochastic Optimization[C]. International Conference on Learning Representations, 2015.
- [30] Kolm, P. N., and G. Ritter. "Dynamic Replication and Hedging: A Reinforcement Learning Approach."

The Journal of Financial Data Science, 2019, 1 (1): 159–171.

[31] Lebaron B. Technical Trading Rule Profitability and Foreign Exchange Intervention[J]. Journal of International Economics, 1999, 49(1): 125-143.

[32] Li Y, Shen D, Wang P, et al. Does intraday Time-series Momentum Exist in Chinese Stock Index Futures Market?[J]. Finance Research Letters, 2020, 35: 101292.

[33] Makridakis S, Spiliotis E, Assimakopoulos V. The M4 Competition: Results, Findings, Conclusion and Way Forward[J]. International Journal of Forecasting, 2018, 34(4): 802-808.

[34] Miffre J, Rallis G. Momentum Strategies in Commodity Futures Markets[J]. Journal of Banking & Finance, 2007, 31(6): 1863-1886.

[35] Miotto R, Wang F, Wang S, et al. Deep Learning for Healthcare: Review, Opportunities and Challenges[J]. Briefings in Bioinformatics, 2018, 19(6): 1236-1246.

[36] Moskowitz T J, Ooi Y H, Pedersen L H. Time Series Momentum[J]. Journal of Financial Economics, 2012, 104(2): 228-250.

[37] Naughton T, Truong C, Veeraraghavan M. Momentum Strategies and Stock Returns: Chinese evidence[J]. Pacific-Basin Finance Journal, 2008, 16(4): 476-492.

[38] Okunev J, White D. Do Momentum-Based Strategies Still Work in Foreign Currency Markets? [J]. Journal of Financial & Quantitative Analysis, 2003, 38(2): 425-447.

[39] Rather A M, Agarwal A, Sastry V N. Recurrent Neural Network and a Hybrid Model for Prediction of Stock Returns[J]. Expert Systems with Applications, 2015, 42(6): 3234-3241.

[40] Schwert G W. Chapter 15: Anomalies and Market Efficiency[M] Arrow K J et al. Hand book of the Economics of Finance[M]. Elsevier, 2003.

[41] Shleifer A, Summers L H. The Noise Trader Approach to Finance[J]. Journal of Economic Perspectives, 1990, 4(2): 19-33.

[42] Silver D, Schrittwieser J, Simonyan K, et al. Mastering the Game of Go Without Human Knowledge[J]. Nature, 2017, 550(7676): 354-359.

[43] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1): 1929-1958.

[44] Wen Z, Gong X, et al. Intraday Momentum and Return Predictability: Evidence from the Crude Oil Market[J]. Economic Modelling, 2021, 95: 374-384.

[45] Werbos P. Backpropagation Through Time: What It Does and How to Do It[J]. Proceedings of the

IEEE, 1990, 78(10): 1550-1560.

[46] Zhang Y, Ma F, Zhu B. Intraday Momentum and Stock Return Predictability: Evidence from China[J]. Economic Modelling, 2019, 76: 319-329.

[47] 樊家鸣. 股权分置改革前后沪深 A 股动量和反转效应的实证研究[J]. 经济研究参考, 2011(53): 66-78+83.

[48] 高秋明, 胡聪慧, 燕翔. 中国 A 股市场动量效应的特征和形成机理研究[J]. 财经研究, 2014, 40(2): 97-107.

[49] 郭磊, 吴冲锋, 刘海龙. 基于收益分解的股票市场动量效应国际比较[J]. 系统管理学报, 2007(2): 189-193.

[50] 刘清源, 魏先华, 王霖. 动量策略收益主要来源于隔夜?——基于沪市 A 股动量策略收益分解的探讨[J]. 投资研究, 2016, 35(3): 92-104.

[51] 鲁臻, 邹恒甫. 中国股市的惯性与反转效应研究[J]. 经济研究, 2007(9): 146-156.

[52] 马超群, 张浩. 中国股市价格惯性反转与风险补偿的实证研究[J]. 管理工程学报, 2005(2): 64-69.

[53] 邱锡鹏. 神经网络与深度学习[M]. 北京: 机械工业出版社, 2019: 185-186.

[54] 曲荣华, 刘扬. 中国 A 股的隔夜-日内反转效应[J]. 经济学报, 2020, 7(4): 64-87.

[55] 史永东, 王谨乐, 胡丹. 中国股票市场个人投资者和机构投资者的过度自信差异研究[J]. 投资研究, 2015(1): 82-96.

[56] 唐也然, 龙文, 石勇. 投资者非理性行为是动量效应产生的原因吗? ——来自互联网金融论坛的证据[J]. 证券市场导报, 2020(4): 62-70.

[57] 田利辉, 王冠英, 谭德凯. 反转效应与资产定价: 历史收益率如何影响现在[J]. 金融研究, 2014(10): 177-192.

[58] 王剑波, 钟源. 我国 A 股市场动量效应探究[J]. 经济研究导刊, 2021(6): 60-63.

[59] 王永宏, 赵学军. 中国股市惯性策略和反转策略的实证分析[J]. 证券市场导报, 2001(6): 28-33.

[60] 徐信忠, 郑纯毅. 中国股票市场动量效应成因分析[J]. 经济科学, 2006(1): 85-99.

[61] 尹瑞, 魏晓峰. 市场动量效应与盈利不稳定性——基于中国市场的实证研究[J]. 经济学(季刊), 2020, 19(4): 1421-1446.

[62] 游家兴. 谁反应过度, 谁反应不足——投资者异质性与收益时间可预测性分析[J]. 金融研究, 2008(4): 161-173.