

摘要

近些年来，随着中国资本市场不断成熟以及机构投资者的专业程度不断提升，国内的管理期货基金（CTA）越来越受到投资者的青睐。特别是在 2020 年新冠疫情爆发之后，在其他策略基金净值受到市场的剧烈波动而表现不尽人意时，做多波动率策略的 CTA 基金表现却十分亮眼。伴随着国内金融衍生品市场的不断完善以及投资者观念进步等因素，国内 CTA 市场还将有广阔的发展空间。然而，对于 CTA 基准的评估不像股票基金那么容易，目前市场上仅有以基金管理人业绩构建的指数作为 CTA 市场的基准，还没有出现以标的资产构建的基准指数。而国内私募市场仍处于较不成熟的发展阶段，国内学者对于私募基金特别是 CTA 基金相关的研究较少，在市场前景向好的基础上，丰富国内 CTA 基金相关的研究有助于提升国内 CTA 基金市场的定价效率，营造良好的金融衍生品投资环境。

国内绝大多数 CTA 基金使用的投资策略都是趋势跟踪策略，由于 CTA 策略的特殊性，在 CTA 绩效的相关学术研究中往往使用 Moskowitz 等人（2012）与 Baltas 和 Kosowski（2013）构建的时间序列动量策略来度量 CTA 基金的时间序列动量效应。虽然时间序列动量策略在捕捉时间序列动量上是有效的，但因为模型的一些缺点，如没有对模型进行优化和过于简单的趋势和仓位大小估计等问题，招致了国内外学者的许多批评。

因此本文在 Lim 等人（2019）的研究基础上，将基于深度学习的交易规则与时间序列动量的波动率缩放框架相结合，构造深度动量网络来对国内期货市场的时间序列动量效应进行度量，进而构造 CTA 市场基于资产的基准，对个体 CTA 基金进行绩效的测度和归因研究。

本文选取 2012-2021 共 10 年的国内 70 只商品期货连续价格序列作为样本，在传统的时间序列动量模型的基础上，针对传统模型的缺点，利用深度学习的框架进行改进，构建深度动量网络，并对该模型在盈利能力、风险控制和性能比三个维度进行评估。然后将深度动量网络模型与目前市场上基于基金

管理人构建的 CTA 指数进行市场解释力研究。最后，选择国内净值披露密度较高的 CTA 私募基金作为本文个体 CTA 基金样本，利用深度动量网络作为市场基准对 CTA 绩效进行分时期分市场的测度与归因研究。

本文得出的结论主要包括：（1）基于夏普比率优化的长短期记忆神经架构的深度动量网络表现最好，能够在控制风险的同时把握时间序列动量的超额收益，波动率归一化的日回报率以及波动归一化移动平均线（8，24）对于深度动量网络训练的贡献程度较大。（2）基于资产分类的深度动量网络能够提高对管理人业绩基准的解释力，工业品和金属类别的深度动量网络对于管理人业绩基准的解释力较好，而农产品和金融衍生品类别的解释能力较差，管理人业绩基准筛选的个体 CTA 基金可能更偏好于工业品类和金属类期货。

（3）本文构建的深度动量网络对于个体 CTA 基金的解释能力较管理人业绩构建的基准要好。（4）在牛市期间，个体 CTA 基金的 Beta 和 Alpha 评估效果较好，但在熊市或者中性市场则表现一般。（5）国内 CTA 基金的业绩具有持续性，越大规模的基金公司和存续期越长的 CTA 产品能够获得更好的业绩，在牛市期间 CTA 业绩的持续性会增强。（6）在牛市配置工业品期货能给 CTA 基金带来超额收益，在熊市则是农产品的表现更好，而在全样本期间，金属类期货和工业品期货对于 CTA 基金业绩的贡献较突出。

本文的研究创新点主要有以下三个方面：（1）在由 Moskowitz 等人（2012）提出的传统的时间序列动量模型上，通过自定义的损失函数以及深度学习网络框架构建深度动量网络模型，改进了传统时间序列动量模型的缺点并融合了深度学习在时间序列预测上的优点，完善了时间序列动量在国内期货市场上的测度方法。（2）利用本文测度的时间序列动量单因子模型来对个体 CTA 基金进行绩效测度和归因研究，为国内 CTA 市场构建了基于资产的基准，提高了个体 CTA 基金的定价效率，为国内 CTA 基金绩效研究提供了新的研究思路。（3）在 CTA 基金的绩效测度和归因研究方面，将样本根据期货市场和 CTA 行业的发展情况分为牛市、熊市和中性市进行比较分析，更真实地反映了各个时期内 CTA 基金的特点，丰富了国内相关领域的研究。

由于研究水平有限，本文还存在一些不足之处：（1）目前学界对 CTA 基金基于资产的基准并无统一共识的评价标准，本文通过选择以往学者较有代表性的结果进行研究，但由于研究对象的差异性可能导致本文研究结果缺乏

一定的严谨性。(2) 由于 CTA 基金为私募基金，国内的私募基金相关数据并不需要强制披露，虽然本文在处理 CTA 基金的相关数据时尽量通过各种渠道对数据进行交叉验证，但由于各个 CTA 基金披露的数据口径存在差异，可能导致本文研究结果存在一定的片面性。

关键词：CTA 基金绩效；时间序列动量；期货市场；深度学习；CTA 基准

Abstract

In recent years, as China's capital markets continue to mature and institutional investors become more professional, managed futures funds (CTAs) in China have become increasingly popular among investors. In particular, following the outbreak of the COVID-19 in 2020, CTA funds, which are long volatility strategies, performed well when the net value of other strategy funds was underperformed by the violent market fluctuations. With the continuous improvement of the domestic financial derivatives market and the progress of investors' perceptions, the domestic CTA market will have a wide scope for future development. However, the assessment of CTA benchmarks is not as easy as that of equity funds. Currently, there are only indices constructed on the basis of fund managers' performance as benchmarks for the CTA market, and benchmark indices constructed on the basis of underlying assets have not emerged. At present, the domestic private equity market is still at a relatively immature stage of development, and domestic scholars have conducted relatively little research related to private equity funds, especially CTA funds. On the basis of the promising market outlook, enriching the research related to domestic CTA funds will help enhance the pricing efficiency of the domestic CTA fund market and create a good environment for investment in financial derivatives.

Most of investment strategies used by domestic CTA funds are trend-following strategies, and due to the specificity of CTA strategies, the time series momentum strategy constructed by Moskowitz et al. (2012) and Baltas and Kosowski (2013) is often used in academic studies related to CTA performance to measure the time series momentum effect of CTA funds. Although the time series momentum strategy is effective in capturing time series momentum, it has attracted much criticism from domestic and international scholars because of some shortcomings of the model, such as the lack of optimization of the model and oversimplified trend and position size estimation.

This paper therefore builds on the study of Lim et al. (2019) by combining deep learning-based trading rules with the volatility scaling framework of time series momentum to construct deep momentum networks to measure the time series momentum effect in the domestic futures market, and then construct asset-based benchmarks for the CTA market to measure and attribute performance to individual

CTA funds.

In this paper, 70 domestic commodity futures continuous price series for a total of 10 years from 2012-2021 are selected to construct a deep momentum network based on the traditional time series momentum model, which is improved using a deep learning framework to address the shortcomings of the traditional model, and the model is evaluated in three dimensions: profitability, risk control and performance ratio. The deep momentum network model is then studied for its market explanatory power against the CTA indices currently available in the market based on fund managers' construction. Finally, CTA private equity funds with a high density of net worth disclosure in China are selected as a sample of individual CTA funds for this paper, and the deep momentum network is used as a market benchmark to measure and attribute CTA performance over time and by market.

The main conclusions drawn in this paper include: (1) Deep momentum network based on Sharpe ratio optimized long short-term memory neural architectures performs best in capturing excess returns on time series momentum while controlling risk, with volatility normalized daily returns and volatility normalized moving averages (8, 24) contributing to a greater extent to deep momentum network training. (2) The deep momentum network based on asset classification can improve the explanatory power of the manager's performance benchmark. The explanatory power of the deep momentum network for the industrial goods and metals categories is better for the manager's performance benchmark, while the explanatory power of the agricultural products and financial derivatives categories is poor, and the individual CTA funds screened by the manager's performance benchmark may prefer the industrial goods and metals futures. (3) The explanatory power of the deep momentum network constructed in this paper is better for individual CTA funds than the benchmark constructed by the manager's performance. (4) Beta and Alpha assessments of individual CTA funds are better during bull markets, but perform generally in bear or neutral markets. (5) The performance of domestic CTA funds is persistent, and the larger the fund company and the longer the duration of the CTA product, the better the performance, and the persistence of CTA performance is enhanced during bull markets. (6) Allocating to industrial futures in bull markets gives CTA funds excess returns, while in bear markets it is agricultural products that perform better, while metal

futures and industrial futures contribute more to CTA fund performance over the full sample period.

The research innovations in this paper are mainly in the following three aspects: (1) On the traditional time series momentum model proposed by Moskowitz et al. (2012), we construct a deep momentum network model by using a custom loss function and a deep learning network framework, which improves the disadvantages of the traditional time series momentum model and integrates the advantages of deep learning in time series prediction. The measure method of time series momentum in domestic futures market is improved. (2) The time series momentum single factor model measured in this paper is used to conduct performance measurement and attribution research on individual CTA funds, which establishes an asset-based benchmark for the domestic CTA market, improves the pricing efficiency of individual CTA funds, and provides a new research idea for the performance research of domestic CTA funds. (3) In terms of CTA fund performance measurement and attribution research, the sample is divided into bull, bear and neutral markets for comparative analysis based on the development of the futures market and the CTA industry, which more realistically reflects the characteristics of CTA funds in each period and enriches domestic research in related fields.

Due to the limited research level, this paper still has some shortcomings: (1) At present, there is no unified consensus on the evaluation criteria of the asset-based benchmark of CTA fund in the academic circle. This paper selects the representative results of previous scholars for research, but the differences in the research objects may lead to the lack of a certain rigor in the research results. (2) Since CTA funds are private funds, data related to domestic private funds do not need to be mandatory disclosed. Although this paper tries to cross-verify data through various channels when processing data related to CTA funds, the data aperture disclosed by various CTA funds may be different, leading to the one-sided results of this paper.

Keywords: CTA fund performance; Time series momentum; Futures market; Deep learning; CTA Benchmark

目 录

1. 引言	1
1.1 研究背景	1
1.2 研究意义	4
1.3 研究方法	6
1.4 研究内容	7
1.5 创新之处	9
2. 文献综述	11
2.1 CTA 基金的基准研究	11
2.2 动量效应	12
2.3 时间序列动量效应	14
2.4 金融领域的机器学习	16
2.5 简要评述	19
3. 研究设计	21
3.1 样本数据	21
3.1.1 期货数据	21
3.1.2 CTA 基准指数	21
3.1.3 CTA 基金数据	22
3.1.4 无风险利率	22
3.2 经典时间序列动量策略	23
3.3 深度动量网络的构建	24
3.3.1 深度学习的神经网络架构	25
3.3.2 基金绩效评估指标定义的损失函数	29
3.3.3 超参数选择	30
3.4 管理人业绩基准的时间序列动量分解	31

3.5	个体 CTA 基金绩效研究模型	33
3.5.1	期货市场总览及市场时期划分	33
3.5.2	CTA 绩效测度模型	36
3.5.3	CTA 绩效归因模型	37
3.6	本章小结	38
4.	实证分析	40
4.1	深度动量网络训练详情及评估	40
4.1.1	训练详情	40
4.1.2	训练结果评估	41
4.2	CTA 管理人基准的时间序列动量解释	43
4.2.1	现有行业基准指数表现	43
4.2.2	深度动量网络对于管理人基准的解释力研究	44
4.2.3	分资产类别的深度动量网络解释力研究	45
4.3	CTA 基金绩效 Beta 测度	46
4.3.1	个体 CTA 基金描述性统计	46
4.3.2	不同基准的个体 CTA 基金解释力比较	47
4.3.3	分时期的 CTA 基金绩效 Beta 测度	48
4.3.4	分资产类别分时期的 CTA 基金绩效 Beta 测度	49
4.4	CTA 基金绩效 Alpha 测度	50
4.4.1	分时期的 CTA 绩效 Alpha 测度	50
4.4.2	分资产类别分时期的 Alpha 测度	51
4.5	CTA 基金绩效归因研究	52
4.5.1	全样本的 CTA 绩效归因研究	52
4.5.2	分时期的 CTA 基金绩效归因研究	54
4.6	本章小结	57
5.	结论与启示	59
5.1	论文结论	59
5.2	研究启示	60

5.3 研究展望	62
参考文献	64
附录	70
致谢	72

1. 引言

1.1 研究背景

CTA 基金的相关概念是从美国金融市场引申过来的，在美国，CTA 基金叫做商品交易顾问，而在国内则被称作管理期货基金。2015 年，随着股灾的发生导致监管趋严，许多股票量化交易受限，一些私募资金开始转战 CTA 市场，此时 CTA 基金才算在国内站稳脚跟，慢慢为投资者所熟知。

从传统意义上来说，CTA 基金仅限投资于商品期货，但随着全球金融市场的发展和相关制度的完善，CTA 基金的投资领域也逐渐扩大，延伸至几乎所有的期货品种，包括股指期货和外汇期货等。CTA 产品在国内主要通过私募基金、公募专户、券商资产管理计划、期货资产管理计划等途径发行。国内 CTA 产品可以根据不同的投资风格，将国内 CTA 产品分类标签化。从交易标的上，有些产品侧重于股指期货，有些则侧重于商品期货，也有做全品种的配置的；从交易频率上，有些侧重于高频，有些则侧重于中低频；从策略类型上，有些侧重于趋势跟踪，有些则侧重于统计套利。

从交易方式来看，我国的 CTA 产品主要有主观 CTA 与量化 CTA 之分，主观 CTA 指的是根据主观经验和判断进行交易，对宏观基本面或技术面数据进行分析 and 判断而做出交易决策；量化 CTA 主要按照事先制定的规则进行程序化交易，利用数量规则对历史价格进行处理，制定并严格按照相关交易规则执行。从交易品种来看，有相对单一化 CTA 与多元化 CTA 之分，相对单一化 CTA 指的是针对股指期货或者某类商品的 CTA 基金，交易的品种相对单一；多元化 CTA 指的是交易品种多样化，尽可能做到分散投资。从交易频率或者周期来看，有高频或日内、中低频、宏观对冲之分。高频交易主要有做市、套利或短期预测等交易方式，对交易系统要求较为严格，需要低延迟、FPGA、C++、Co-location、行情专线、独立交易单元等软件及硬件支持，同时还受到市场每日交易量的限制，其容量相对有限；中低频策略对交易系统要求较为

宽松，容量也较大；宏观对冲类策略一般对速度要求比较低，结合股、债等方面配置，容量也相对较大。

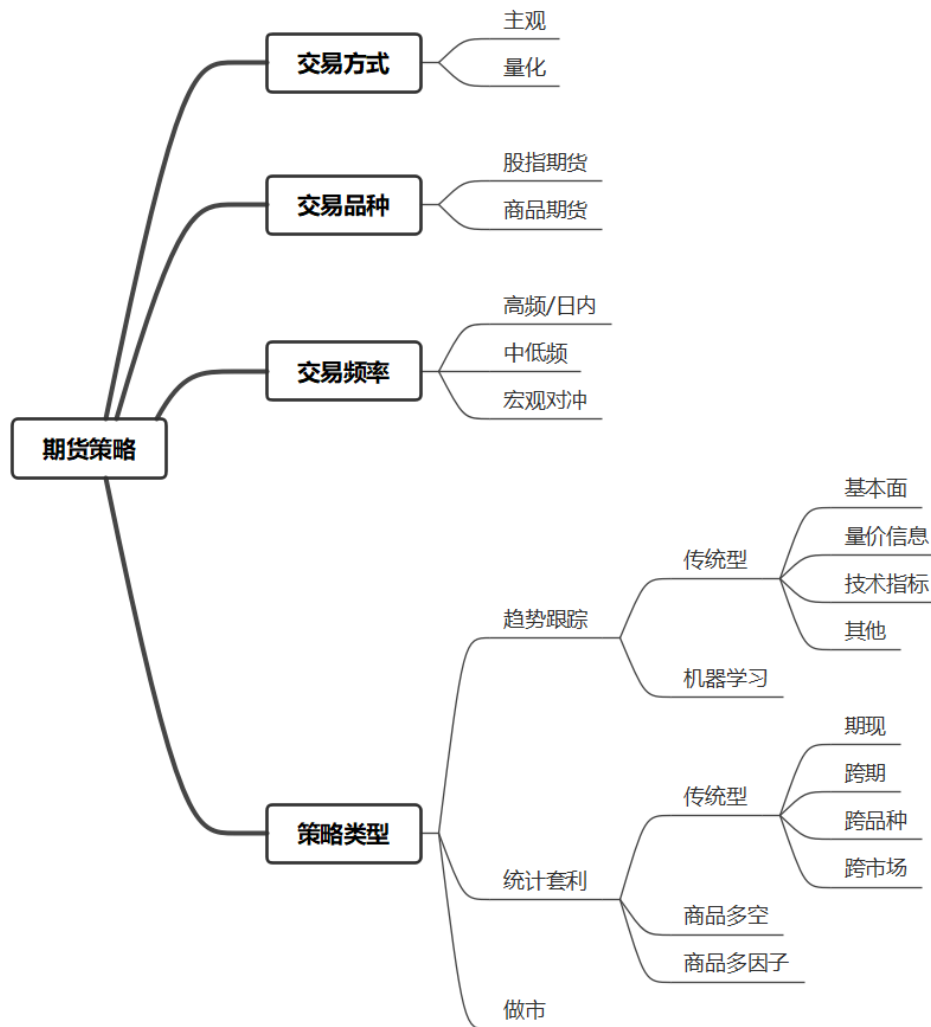


图 1-1 CTA 产品的主要分类

如果从策略类型来分，国内 CTA 产品主要使用的策略为趋势跟踪、统计套利等。趋势跟踪策略历史悠久，可以说是 CTA 基金的最主流策略。CTA 趋势跟踪类策略也经历过多轮迭代，从 80 年代最开始的简单量价类策略发展到用机器学习来做预测，意图抓住不同周期上的趋势。Hurst 等人（2012）将趋势分为三个阶段来说明趋势是如何形成、延续以及结束。趋势形成最开始都需要某个催化剂，如某个利好或者利空消息，供给侧改革、需求端偏移等。趋势开始形成之后，市场对于新的定价还不够充分，主要受到锚定效应、信息传导不充分、处置效应或者摩擦成本等因素影响。最后一个阶段指的是由于羊

群效应、确定偏差或者代表性偏差以及资产滞后配置等，市场过度反应导致趋势延续。发展到最后就是趋势结束，行情反转。

期货价差套利策略根据所选择期货合约的不同，可分为跨期套利、跨品种套利和跨市套利。传统套利策略是通过修复错误定价，做多均值回归来赚取收益。目前市场上也有 CTA 基金借鉴股票多因子的构造方法，通过构造期货市场的多因子模型进行多空组合交易，本质上也是统计套利，但由于目前大宗商品期货标的少，具有流动性及波动性的商品标的目前就 50 余个，因此在横截面上显著性值得怀疑。统计套利类策略一般胜率高，但是盈亏比较低，但会承受双边暴露风险以及定价失效的情况出现，比如 2022 年出现菜粕期货比豆粕期货的价格高、棕榈油期货价格高于豆油期货的情况，直接导致套利失败，因此套利策略的有效性目前还值得商榷，国内主流的 CTA 基金依旧是以趋势跟踪策略为主。

接下来从国内 CTA 产品运作模式角度来进行简要介绍，CTA 管理人从产品净值以及客户预期管理的角度考虑，为了保证净值曲线的相对平稳以及公司运营的需要，国内的 CTA 产品特别是有了一定规模之后的产品往往不会采用单一的趋势跟踪策略，而是采用复合策略的模式。复合策略指的是在交易过程中不仅仅采用一种交易策略，可能既有趋势跟踪策略，也有套利策略，又或者是在交易频率上存在差异。之所以国内 CTA 产品普遍采用复合策略是因为国内期货的单边趋势行情相对较少，大部分交易时间内都是以震荡为主，如果仅仅采用趋势跟踪策略，在震荡行情下产品净值回撤较大，保证回撤处在合理的范围内是产品以及投资者预期管理的重要问题。市场上某些 CTA 产品是通过“多品种、多参数、多周期、多策略”来降低回撤，也有产品是通过加入统计套利类策略来控制回撤。由于期货自带杠杆，所以一个产品大部分资金用来做增厚收益的投资，如通过逆回购或者投资低风险债券等方式，放在期货保证金账户里面的资金一部分做趋势跟踪策略，小部分资金做套利类策略以应对震荡带来的净值回撤，通过复合策略使得产品净值曲线走势相对平稳。以趋势跟踪策略为主的 CTA 产品净值曲线多呈现阶梯跳跃状，通过一波趋势行情净值创新高，震荡行情进入回撤横盘状态，等待下一波趋势行情再创新高。如果说抓取上涨或者下跌趋势是市场给予的 Beta 机会，那么如何

在震荡行情下控制好回撤幅度以及最大回撤周期就是考验每个投资经理或者产品的 Alpha 能力。

随着中国资本市场的成熟和投资机构专业水平的提高，越来越多的 CTA 基金公司成立并发展壮大，从早期的千象投资、黑翼投资，到之后的思懿投资和涵德投资等，都是具有一定管理规模的 CTA 策略公司。根据巴克莱对冲基金指数，CTA 策略在传统意义上与基于股票的策略，例如股票多空、市场中性等相关性很小。由于这种低相关性，CTA 策略正逐步被添加到越来越多的投资机构的投资组合中，如商业银行、证券公司以及期货公司等机构，以此希望能够将投资组合可能存在的市场波动风险进行对冲。此外，由于 CTA 策略中的趋势追踪策略具有较高的波动偏好，这也是 CTA 策略组合中最重要的收益贡献来源，因此 CTA 策略，尤其是 CTA 策略组合中的趋势追踪策略，在一些交易比较活跃的二级市场上能够跑赢其他策略，尤其是在国际金融市场出现严重波动时，越来越多的机构正在关注和配置 CTA 策略，将其作为应对意外或者尾部风险的缓冲。自 2020 年初，在新冠病毒开始广泛传播并对全球实体经济产生重大负面影响的背景下，CTA 策略已经成为一支不可忽视的重要力量，大多数国内 CTA 基金在 2020 年取得了显著的绝对收益。

根据巴克莱对冲基金数据库统计，全球 CTA 基金规模从 1980 年 3.1 亿美元发展到 2021 年底超过 3300 亿美元，规模显著扩大。然而截至 2021 年 6 月，国内 CTA 策略私募基金规模仅有 1500 亿元左右，存在巨大的发展前景。国内 CTA 基金主要使用的是趋势跟踪策略，时间序列动量策略所包含的动量效应或许能够很好的解释 CTA 基金收益的驱动。而在人工智能和计算机科学高度发展的今天，深度学习在金融时间序列上的预测水平也发挥了重要的作用。

而在如今国内私募市场上，目前并未有成熟的基于资产构建的 CTA 基准指数用来反映 CTA 基金行业整体的发展状况。建立基于资产的 CTA 基准指数，投资者可以通过该指数进行参考来合理安排自己的投资计划，CTA 基金的管理者则可以将该指数作为行业的比较基准，而监管者可以通过该指数来评估市场上的整体风险的情况，通过合理的分析来应对可能出现的系统风险，并提前做好相应的监管准备措施。

1.2 研究意义

随着国内金融市场环境日趋复杂、参与机构与投资策略日趋丰富，CTA 的市场规模占有量也越来越大。然而直到最近，整个 CTA 行业依旧还没有基于资产构建的基准，从业人员及学者仅仅将 CTA 基金的表现直接与基于管理人业绩构建的指数进行比较。之所以依赖基于管理人构建的指数是由于 CTA 基金主要交易策略的特殊性——趋势跟踪策略很难直接通过市场的资产直接进行构建，需要借助其它模型才能进行很好的评估。直接利用基于管理人业绩构建的 CTA 指数进行基准测试存在着以下几个缺陷，首先，基于管理人业绩构建的 CTA 指数本质上是市场上的管理人的 Alpha 与 Beta 的组合。根据 Fung 和 Hsieh (2004) 的观点，对冲基金指数可能会继承构建该指数的对冲基金一些偏见，也就是说，基于管理人业绩构建的 CTA 指数很难对其它的业内 CTA 基金进行有效的 Alpha 评估，因为很难反应业内 CTA 基金的管理技能。更重要的是，基于管理人业绩构建的 CTA 指数几乎不能反映基金的风格漂移、行业的风险敞口或者其它方面的因素，而这些因素往往是市场投资者十分在意的。随着 CTA 策略在机构投资者的投资组合中的价值越来越显著，面对市场上存在大量良莠不齐的个体 CTA 基金，如何更准确地找到适合自己风险偏好和投资组合特点的 CTA 基金以及如何把握 CTA 策略的投资时机，已经成为投资机构在实际投资时必须面对的问题。因此对于与同行业绩构建的相对基准相比，直接对基金管理人进行透明、完全基于资产构建趋势跟踪策略的基准将更有价值。

综上，从理论意义上来看，主要有以下几个方面：

(1) 利用国内期货市场的数 据，将深度学习和时间序列动量策略的内核结合起来，构成深度动量网络模型。深度动量网络通过选择的常用投资组合评估指标，如预期回报、索提诺比率、夏普比率等来优化对各个资产的敞口，利用滞后的资产表现和波动率指标形成预测。它与经典监督机器学习方法的区别在于，深度动量网络不针对明确的目标进行优化，而是针对投资组合绩效指标本身进行优化。深度动量网络与其他时间序列动量模型的不同之处在于它能够利用滞后资产数据中的非线性关系，为时间序列动量建模提供了新的方向。

(2) 丰富了国内对于 CTA 基金相关领域的研究。由于国内私募数据披

露较少及获取难度较大，国内对于私募基金领域的研究本就不多见，更遑论 CTA 基金领域的研究。本文通过改进传统的时间序列动量模型，构造出 CTA 行业基于资产的基准，通过该基准对于国内个体的 CTA 基金进行了测度及归因等相关方面的研究。

从现实意义上来看，主要有以下几个方面：

(1) 根据国内朝阳永续数据库的数据，国内 CTA 策略主要分为趋势跟踪策略、趋势反转策略和价差套利策略三种，其中趋势跟踪策略是市场的主流，约占 70% 的市场份额。趋势跟踪策略是指跟踪相关品种不同周期的趋势，买入正在上涨的资产，卖出正在下跌的资产，这正与时间序列动量策略的核心逻辑一致，因此我们可以推测国内 CTA 基金的收益可以很好地由该策略收益解释，因此恰当地对国内时间序列动量进行建模可以帮助填补国内的相关研究的空白，合理地对市场上的相关产品进行评估，帮助 CTA 市场上的产品合理定价。

(2) 国内私募基金市场发展存在着许多不透明，信息不对称的问题，CTA 基金产品由于其投资策略的特殊性，没有办法像股票基金一样能够利用如沪深 300 指数作为大盘的参考基准，投资者在对于市场产品的选择上往往会比较困难，本文构建了直接由底层资产组成的市场基准，与国内市场存在的由基金管理人业绩构建的相对比较基准组合评估，为减少私募市场信息不对称以及投资者权益保护提供了坚实的基础。

(3) 本文将深度学习在时间序列预测上的应用拓展到国内期货市场全品种当中，对期货市场最重要的时间序列动量因子进行了评估，更契合大数据背景下的金融市场属性，为量化投资在实践层面上提供了新的思路。

1.3 研究方法

CTA 策略在股票和债券市场相对单一的投资方向上进行了拓展，在金融市场出现较大的系统性风险时，为市场投资者提供了多元化的配置方向，实现了以较低的相关性平滑投资组合波动的效果。近年来市场波动越来越大，伴随着国内金融衍生品市场的发展，CTA 策略逐渐得到投资者的青睐，出现了越来越多的 CTA 策略管理团队。然而，由于 CTA 策略的复杂性，市场上仍

然没有以资产直接构建的 CTA 基准, 仅仅存在以基金管理人业绩构建的基准, 而该基准在市场的投资者对于基金进行评估的时候反映的信息有限, 因此本文结合以往学者的研究, 将深度学习的方法与传统的时间序列模型结合起来, 完成对国内 CTA 基金的绩效研究, 具体研究方法如下:

(1) 文献研究法。首先, 本文先对期货市场上的动量效应进行了理论上的详细梳理, 探讨了国内外相关学者对于不同种类动量模型优缺点的研究和实际市场表现情况。接着, 再详尽介绍了使用的深度学习网络结构, 对多层感知机、长短期记忆模型的框架搭建等细节进行了解释, 然后对本文创新的自定义损失模型进行了详尽解释。此外, 本文通过借鉴前人的研究发现, 利用回归等方法拟合个体 CTA 基金的 Alpha、Beta 对于基金绩效研究有帮助, 因此将其作为本文评估 CTA 基金绩效的模型。

(2) 多学科方法相结合。本文充分利用并结合了金融计量学、金融风险管理和统计学等诸多学科的相关研究方法和理论, 构建了本文实证研究的相关统计指标和评估参数, 使其在数学逻辑上合理的同时也存在着经济现实意义。

(3) 回归分析法。在实证研究方面, 本文首先利用 Excel 和 Python 进行样本数据的选择和处理, 接着利用本文介绍的构建深度动量网络方法, 利用 2 种神经网络架构(多层感知机、长短期记忆神经网络)以及自定义的损失函数将传统的时间序列动量模型进行了改进。本文使用了 2012 年至 2021 年底共 10 年的国内期货市场的的数据, 将改进后的模型在盈利能力、风险控制和性能比三个维度进行了比较, 从中选择了表现最好的深度动量网络模型作为 CTA 市场基于资产构建的基准, 并将该基准的训练结果进行了多方面的评估。

为了研究国内市场个体 CTA 基金的绩效, 本文收集并筛选了国内 CTA 基金的净值及相关特征数据, 结合上文构造的基于资产的 CTA 基准, 针对单个 CTA 基金的 Alpha、Beta 及相关特征开展分资产类别分时期相关的研究, 并进行 CTA 的绩效归因研究。本文在实证过程中利用表格和图片展示了研究成果, 更能直接体现研究结果以及过程。

1.4 研究内容

本文主要改进了传统的时间序列动量模型，利用深度学习的网络框架，使用基金绩效评估指标自定义了损失函数，构建了新的深度动量网络，并将深度动量网络作为 CTA 市场基于资产构建的基准，进行个体 CTA 基金的分资产类别分时期的绩效测度以及归因研究。本文主要包括五个主要章节，第一章为引言，第二章为文献综述，第三章为研究设计，第四章为实证分析，第五章为本文的结论以及启示，具体的研究框架见图 1-2 所示。

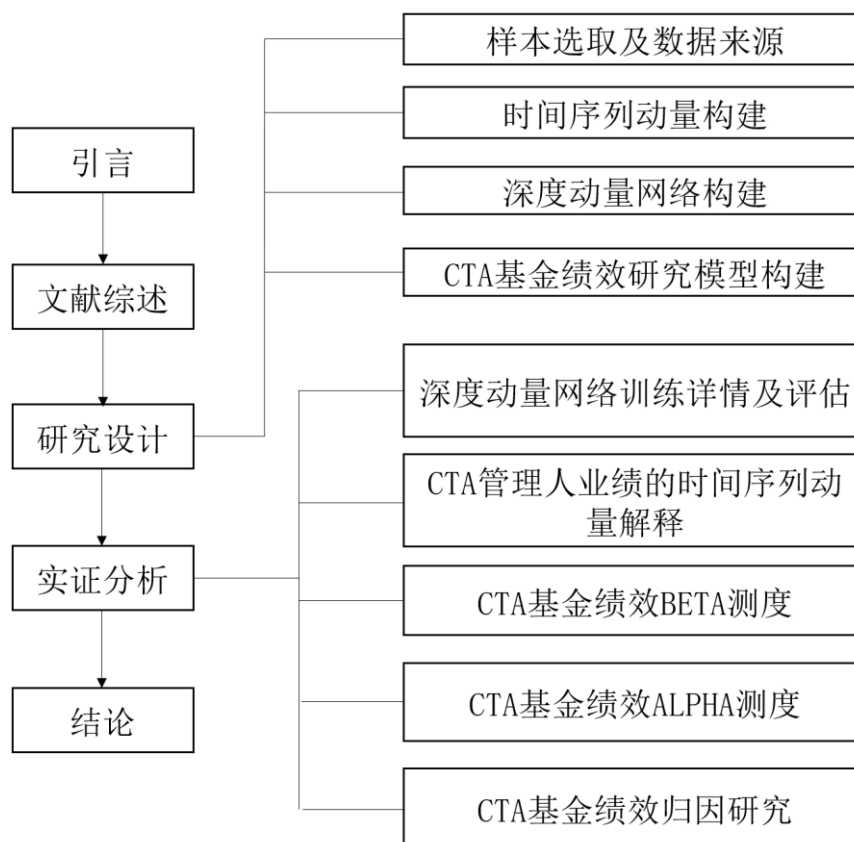


图 1-2 论文研究框架

第一章是引言。首先对本文的研究背景和研究意义进行了详细阐述，我国金融衍生品市场仍处于快速发展的时期，CTA 策略具有显著的危机 Alpha 特点，因此对 CTA 基金绩效进行有效评估对国内金融市场发展具有深刻意义。接着展示了本文所采用的研究方法、研究内容，主要是利用深度学习的框架改进了传统的时间序列动量模型，从而构建 CTA 市场基于资产的基准，然后进行 CTA 基金的横截面及面板回归对基金的绩效进行测度和归因研究。最后介绍了本文在研究中的创新之处。

第二章是文献综述。本章分别从 CTA 基金的基准研究、动量效应、经典时间序列动量模型以及机器学习在时间序列上的应用等方面，依次论述国内外学者的研究成果，最后对当前研究进行了简要评述。

第三章是研究设计。首先对本文使用的相关期货样本、市场上存在的基于管理人业绩的基准以及个体 CTA 基金样本等数据进行定义和来源介绍。接着介绍了传统的时间序列动量模型并对其核心思想进行提炼，针对传统的模型的不足，通过搭建深度学习的网络框架以及利用基金绩效评估指标构建深度动量网络，并对相关的训练细节进行介绍。最后对本文研究样本时期的划分依据进行了阐述，并对 CTA 基金绩效研究所使用的回归模型和思路进行了简要梳理。

第四章是实证分析。本章首先利用第三章介绍的期货样本数据以及模型架构训练深度动量网络模型，将所得到的模型在盈利能力、风险控制和性能比三个维度进行评估，发现训练效果最好的是基于夏普比率优化并利用 LSTM 网络框架构建的深度动量网络，并对该模型进行可解释性评估。其次，将深度动量网络作为 CTA 市场的基于资产构建的基准，与市场上存在的基于管理人业绩构建的基准进行比较研究，分析不同基准的特点。最后利用深度动量网络作为市场基准，对个体 CTA 基金进行分时期和分资产类别的绩效测度与归因研究。

第五章是结论与启示。本章在前四章的理论以及实证研究的基础上，总结了本文相关的研究成果，并结合当前中国期货以及私募市场存在的现象提出相应的政策建议，最后对本文研究过程的不足之处进行了总结，并提出了相应的研究展望。

1.5 创新之处

目前学术界关于时间序列动量效应以及私募基金绩效的相关研究已经取得了丰厚的成果，但是却鲜有类似本文利用深度学习框架构造国内 CTA 市场基于资产基准的视角，并同时利用该基准对个体 CTA 基金进行绩效的评估和归因研究。因此结合目前国内研究现状，本文主要有以下几点创新点：

- (1) 利用深度学习框架改进了国内期货市场上时间序列动量效应的度量

方式。在国内外关于时间序列动量效应的度量研究中，自从 Moskowitz 等人（2012）提出时间序列动量模型对时间序列动量效应进行了建模过后，Baz 等人（2015）利用技术指标改进了模型对于趋势的估计效果，Gert 和 Peter（2017）则利用聚合的时间序列动量模型提高了仓位估计的效率。但总体来看，以往研究中的时间序列动量模型在构建本身不仅存在着较强的主观性，并且还缺少对于模型上的优化。因此本文结合国内外学者的研究，结合深度学习的优势，将神经网络的架构引入经典时间序列动量模型的建模，即使用深度动量网络来对国内期货市场的时间序列动量效应进行评估，完善了时间序列动量在国内期货市场上的测度研究。

（2）使用基于资产构建的市场基准对个体 CTA 基金的绩效进行评估研究。国内外关于基金的绩效研究主要可以分为两大分支，一类是利用如使用 Sharpe 比率等单一评价指标来衡量基金业绩，另一类则是利用多因子业绩归因模型如 Fama-French 三因子模型等来对基金业绩进行研究。由于国内 CTA 策略的特殊性，现有多因子模型很难有效解释 CTA 基金的收益。因此本文通过国内外学者的相关研究，利用本文构建的深度动量网络构建市场基准，利用本文测度的时间序列动量单因子模型来对个体 CTA 市场进行绩效测度和归因研究，为国内 CTA 基金绩效研究提供了新的研究思路。

（3）在个体 CTA 基金绩效的测度和归因研究方面，本文通过梳理总结期货市场和 CTA 基金行业的发展历程，利用国内期货和 CTA 市场的繁荣程度和波动率为参考，将 2014 年到 2021 年底的样本区间划分为牛市、熊市和中性市进行研究，对国内个体 CTA 基金绩效测度和归因进行了不同样本期内的比较研究，更符合国内市场的实际情况，丰富了国内个体 CTA 基金绩效评估的相关研究。

2. 文献综述

传统的 CTA 策略在金融市场出现极端波动的市场环境中历来表现良好，而这种情况经常发生在系统性风险出现的时候，因此投资者认为 CTA 策略的这一属性具有危机超额收益。CTA 策略的波动性偏好主要是由趋势跟踪策略决定的，它贡献了大部分的 CTA 投资组合的风险回报。根据国内朝阳永续数据库的数据库的统计数据，国内 CTA 基金 70% 主要使用的是趋势跟踪策略，时间动量策略所包含的动量效应或许能够很好的解释 CTA 基金收益的驱动。而在人工智能和计算机科学高度发展的今天，深度学习在金融时间序列上的预测水平也发挥了重要的作用。因此，本章将从 CTA 基金的基准绩效研究、动量效应与时间序列动量和金融领域的深度学习的角度出发，对国内外相关文献进行综述，对比国内外相关理论的发展状况，为本文 CTA 基金收益驱动的研究打下坚实基础。

2.1 CTA 基金的基准研究

国内外学者认为，对于 CTA 基金的基准评估并不像普通共同基金或者对冲基金的评估那么容易。共同基金通常是由市场指数和一系列风险因子来确定基准的，如 Fama 和 French (1993) 和 Carhart (1997) 提出的因子模型。同样地，大多数对冲基金都是使用 Fung 和 Hsieh (2004) 的七因子模型进行评估的。虽然这些因子模型在解释共同基金和大多数对冲基金类别的回报方面表现良好，但它们在解释 CTA 回报变化方面的表现有限。在实践中，CTA 基金的评估仍然很大程度上依赖于基于基金管理者表现的指数。

尽管对于 CTA 基金收益建模存在着明显的障碍，国内外依然有学者试图复制他们的回报。Fung 和 Hsieh (2001) 是最早关注复制趋势跟踪对冲基金回报的人之一，他们使用基于回溯跨式期权的趋势跟踪策略 (PTF) 以捕捉市场的回报，虽然作者论证了在实践中是可以实现的，但 Harvey 等人 (2016) 认

为实践成本过高，实践价值比较低。

在 2012 年，Moskowitz 等人（2012）构建的时间序列动量模型在该领域进行了突破。他们第一次以系统的方式证明了期货和远期合约存在时间序列上的趋势效应。他们通过建立时间序列动量模型（TSMOM），该模型完全依赖于资产价格方向的延续，也就是通常意义的趋势。Moskowitz 等人（2012）研究结果表明，时间序列动量策略组合在不同的资产类别均能持续提供显著的超额回报。

在对传统的时间序列动量策略模型的方面，Baltas 和 Kosowski（2013）在 Moskowitz 等人（2012）的基础上，通过美国期货市场的数提出了一套基于技术指标构建的趋势跟踪策略，并使用了周度和日度的数据拓展了现有的时间序列动量的研究，证明了美国 CTA 基金使用时间序列中的动量效应来获得收益。值得一提的是，他们提出的时间序列动量因子显著提高了应用于 CTA 的多因子模型的解释力，并优于 Fung 和 Hsieh（2001）提出的因子。

Gert 和 Peter(2017)借鉴并拓展了 Baltas 和 Kosowski(2013)和 Moskowitz 等人（2012）提出的方法，提出了一种聚合时间序列动量模型，在投资标的趋势扩大的情况下增加了投资于该标的的头寸，并在投资标的趋势减弱时下减少了投资于该标的的头寸，并且实证研究上复制 CTA 基金回报的表现情况很好，对 CTA 基金的 Alpha 进行了有意义的评估。Greyserman 和 Kaminski(2014)则从 CTA 基金规模的角度，对趋势跟踪策略的可实施性进行了详细的研究。

自从 2010 年 CTA 投资策略在我国落地，经历了十年的高速发展为我国的投资者提供了新的投资方式，但目前我国的 CTA 基金依然处于初步的发展阶段，相关的研究依然相当匮乏，但依然有学者对 CTA 基金相关研究进行了补充。杨子江（2020）利用国内期货市场的数，发现期货市场上的市场因子和动量因子具有显著的强相关性，并利用动量因子显著优化了 CTA 策略。刘佳（2020）研究了国内期货市场 110 种 CTA 量化投资策略和 13 个期货品种。对 CTA 量化投资的策略组合优化方案和仓位调整时机进行了分析研究，发现使用风险平价模型的收益效果最好，利用周度的数据更新频率对量化策略的权重分配能使收益曲线更加平滑。

2.2 动量效应

在物理学中,动量是经典力学中的一个物理量。Jegadeesh 和 Titman(1993)在研究中发现股票的回报率倾向于以原有的趋势继续发展。这意味着,过去收益较好的股票在未来将继续获得比过去收益较低的股票更高的收益。动量效应通常被称为“惯性效应”,因为其属性与物理学中的动量相当。众多研究都证明了利用动量效应在股票、期货、债券、外汇和其他各种的交易市场,包括发展中国家和发达市场的金融市场中都能产生显著的超额回报。

例如在商品市场与股票市场的研究中,为了说明动量效应对商品市场的适用性,Narayan(2015)使用移动平均线交易规则,将样本内的商品期货品种利用规则进行排序,在投资策略中定义表现最好的品种持有多头头寸,在表现最差的品种持有空头头寸,在后续的研究中发现可以利用这种方法在商品市场中取得超额收益。Ni 等人(2015)研究了投资者是否从随机震荡指标(SOI)的股票交易中获利,研究结果表明,投资者在交易上证指数时更倾向于采用动量策略,这可能可以用中国投资者的羊群行为解释。根据 Chaves 和 Viswanathan(2016)的研究,利用均值回归策略能够在现货市场获得超额收益,但在期货市场表现不佳,而动量策略在现货和期货市场都表现出色。Kim 等人(2016)发现,如果时间序列动量和买入并持有策略不将波动率进行扩展,它们的累积收益与 Moskowitz 等人(2012)的研究完全相同,并且 Alpha 的评估将不会存在显著差异。刘俊琳(2020)发现我国 A 股市场的短期动量效应,并且发现在牛市期间动量效应显著,在熊市期间动量策略难以获得显著为正的超额收益。

此外,Zaremba(2016)研究了商品市场的金融化如何影响基于动量和期限结构的策略的盈利能力,并发现这两种策略在商品市场上表现良好,其中期限结构策略在非金融化市场中产生的回报明显较高,这些研究表明,动量策略的实施者应该将市场参与者的结构考虑在内。基于中国商品期货市场的角度,Yang 等人(2018)分析了一些不同交易频率的动量和反转交易策略,并证明这些策略可能在一段时间内产生可观的稳定回报,但存在最大回撤风险。根据 Ross 等人(2017)的研究,即使考虑到费用、预期交易成本、税收和其他与现实组合相关的摩擦因素,使用动量组合也会产生动量溢价。Wu 和 Chung(2019)利用 2010 年 1 月至 2015 年 3 月的样本数据对四种动量策略进

行了回测，以检验台湾证券交易所的资本化加权股指期货（TAIEX 期货）的市场有效性。四种策略的相关评价指标都表明台湾股票市场的有效性较弱，并证明了台湾期货市场的动量效应。

综上所述，国内外的相关学者专家在大量研究基础上认为，动量效应很适合作为投资策略组合的一个重要部分，这一观点在众多地区的金融市场上得到了证实。Miffre 和 Rallis（2007）的研究中发现传统资产投资方式的回报与商品动量策略的回报具有较低的相关性，因此可以在投资组合中加入商品动量策略来进行风险的对冲。Kang 和 Kwon（2017）对中国、美国等国家的商品市场展开了研究，发现商品期货的动量具有显著的可预测性，而且发现动量因子不能够被传统的宏观经济因子等解释。林树和陈浩（2014）发现我国期货市场存在动量效应，而且赢家组合的动量效应强于输家组合。

在以往的学术文献中，动量作为一种传统的金融风险溢价因子得到了广泛的研究，动量交易策略通常是以单个股票之间的横截面交易策略来实现的。然而，本文主要研究时间序列中的动量，Gupta 和 Kelly（2018）将其定义为利用资产自身近期收益预测其未来收益的一种业绩持久性现象，这种方法提供了比截面方法更准确的预期因子收益衡量方法。

2.3 时间序列动量效应

经典的动量策略分为两类：多变量的横截面动量和单变量的时间序列动量。横截面动量策略侧重于证券之间的相对表现，买入表现相对较好的证券，卖出表现相对不好的证券。横截面动量的表现被证明在时间上（Jegadeesh 和 Titman, 2001）与不同的市场和资产类别上是稳定的（Baz 等人, 2015）。国内学者对于横截面动量的研究也比较完善。冯珊（2020）的研究表明，赢家的投资组合比输家的投资组合有更稳定的回报，她计算了各个样本时期内的股票回报，然后将收益率按降序排列，买入排名靠前的资产，也就是赢家组合，卖出排名靠后的股票，即输家组合，通过分组的方式证明了市场上的时间序列动量效应。谢振文等人（2018）利用国内股票市场的数据证实投资者可以从时间序列动量策略中获利。

时间序列动量在横截面动量的基础上进行了拓展，专注于资产自身过去

的回报并建立包括所有证券的投资组合。Bird 等人（2017）在 24 个不同的金融市场进行了研究，比较了时间序列和截面动量策略的表现。他们发现，时间序列动量能够捕捉收益率的尾部风险，因此时间序列动量相对于横截面动量的表现更好。Han 等人（2016）分析了股票中来自短期、中期和长期的移动平均信号的组合收益，发现基于多种价格趋势投资组合的表现优于基于单一价格趋势组合的表现。而在 Niemi 等人（2017）的研究中，发现横截面或者时间序列动量在样本内都出现了显著的收益，动量策略的下行风险低于 GSCI、其子行业和其它单个商品，而且没有持续持有未来收益率曲线陡峭、换手率高的商品头寸，这将减少换手率因子的显著性。而对于具体商品样本的研究发现，自 2008 年 9 月开始，工业金属、能源商品和股票之间的相关性明显增加，到 2012 年底达到了一个较高的水平，因此股票市场的不确定性会对商品市场产生一定的影响。姚远等人（2021）则在噪声交易理论的基础上，引入了动量效应，利用上证 180 指数成分股数据构建了一种新的基于噪声交易的动量策略，并发现对于噪声交易水平较高的股票，投资者应选择动量策略来提高收益率。

Ham 等人（2019）则对中国商品期货市场的时间序列动量效应开展了研究，发现时间序列动量策略在选择一个月的回望期以及一个月的持有期中表现优异。此外，在夏普比率、风险调整后的超额收益和累积收益方面，该方法在中国期货市场上战胜了被动多头和横截面的动量策略。然而，许多投机性投资者造成的不稳定的波动市场限制了时间序列动量的持续时间。他们研究结果证实了这样一个观点：在包括中国商品期货在内的国际资产市场上，投机者从时间序列动量技术中获利，而市场上的套期保值者需要为此买单。高长春（2013）则利用中国股指期货市场的数据作为样本，发现在市场极端情况下，时间序列动量组合存在显著的收益率。

而在时间序列动量因子的定量优化方面，Moskowitz 等人（2012）建立了一种使用波动率缩放并根据过去一年的回报趋势来估计交易头寸的策略，他们在过去 25 年的数据中分别展示了 58 种流动性资产的盈利能力。自 2012 年开始，就陆续出现了各种趋势估计的技术和方法，并反映到交易头寸的设置上。例如，Bruder 等人（2013）记录了大量的线性和非线性模型来衡量趋势，

同时估计趋势的显著性，然而他们没有直接利用这些来调整仓位。Baltas 和 Kosowski (2017) 采用了类似于 Moskowitz 等人 (2012) 的方法，将过去 12 个月的对数价格对时间进行回归，并使用回归系数的 t 值来确定交易头寸的方向，基于 t 统计的趋势估计使投资组合周转率降低了 66%，从而降低了交易成本。Baz 等人 (2015) 提出了更复杂的交易规则，以波动率归一化移动平均线 (MACD) 指标作为输入。

而在历史上对于时间序列动量的研究都是对交易规则进行变化，但很少有人对交易规则本身进行比较，也就说几乎没有提供明确的证据或直观的推理来支持一个规则而不是其它规则。本文尝试使用深度神经网络直接生成这些规则，从而避免需要的显式规范，也就说，本文利用深度神经网络根据风险调整后的性能指标对其进行训练，从而直接从数据本身学习最佳训练规则。

2.4 金融领域的机器学习

机器学习是人工智能领域中最重要核心学科之一，机器学习算法利用其挖掘非线性、非参数化特征的能力来研究金融数据之间的关联性，改进了传统的计量经济学只能评估线性和参数化数据以及无法有效匹配数据波动的模式。

对于机器学习在金融领域的应用，国内外学者进行了大量的研究。Huang 等人 (2005) 为了对金融市场的预测性进行研究，使用 SVM 模型来预测日经 225 指数的每周运动方向，在与线性判别分析、二次判别分析和 Elman 反向传播神经网络的比较中，SVM 比其他分类技术表现得更好。Gu 等人 (2020) 使用随机森林模型与传统的线性回归预测方法相比较，通过相关的评估指标发现机器学习模型预测效果更好。Yu 等人 (2009) 使用演化的最小二乘法支持向量机来分析三个主要股票指数的走势：标准普尔 500 指数、道琼斯工业平均指数和纽约证券交易所指数。为了提高 SVR 的预测精度，Huang 等人 (2009) 将台湾指数期货作为研究样本，使用 SVR 与自组织特征图 (SOFM) 技术并优化滤波器的特征选择，提高了模型的预测精度。

在对股票及相关金融产品的预测方面，国内外相关学者将相关的机器学习模型进了一些改进。为了从技术和宏观角度预测股市拐点，李海燕 (2015)

建立了一个多因子 SVM 模型。Dhar 等人（2010 年）为了对印度市场的股票指数相关数据进行预测，使用了具有最优参数选择的 MLP 模型，并发现了该模型在印度市场的适用性。韩山杰和谈世哲（2018）通过使用深度学习框架开发 MLP 神经网络模型来预测苹果的每日收盘股价，证明了 MLP 模型出色的预测准确性。Haq 等人（2021）发现基于多因子的预测模型优于其他所有模型，他们使用 44 个因子的数据训练了 SVM 模型和随机森林模型，并根据每个成分的相关程度选择了最佳参数。张琬琼（2020）建立了一种基于支持向量机与核主成分的 KPCA-SVM 模型，选取上证指数、恒生指数等国内外金融市场的数据计算 MACD、KDJ 等技术指标和开盘价作为预测指标，发现 KPCA-SVM 模型在回归预测方面的性能指标均高于 ARIMA 模型。

随着卷积神经网络（CNN）和递归神经网络（RNN）等现代架构（Goodfellow 等人，2016）的发展，深度学习模型因其构建给定数据集的表示能力而受到青睐（Bengio 等人，2013），并且能以纯粹的数据驱动方式捕获时间序列以及横截面的关系。强大的开源框架也促进了深度神经网络的应用，例如 TensorFlow 和 PyTorch，它们使用自动微分来计算反向传播的梯度，而不必事先显式推导它们。也就是说这种灵活性使得深度神经网络能够超越标准分类和回归模型。例如，Makridakis 等人（2018）发现将传统时间序列模型与神经网络组件相结合的混合方法的表现优于一般的传统法，如指数平滑 RNN（Smyl 等人，2018）、自回归 CNN（Binkowski 等人，2018）和 Kalman 过滤器变体（Fraccaro 等人，2017），同时也使输出更容易被从业者解释。此外，Goodfellow 等人（2014）发现这些框架还能够开发用于训练神经网络的新损失函数，如生成对抗网络（GAN）中的对抗性损失函数。

更进一步地，关于深度学习在金融领域的应用方面，国内外学者也对该领域进行了详细的研究。为了预测国内锌期货价格的涨跌，林杰和龚正（2017）采用了 BP 神经网络和卷积神经网络模型（CNN），并取得了较高的预测精度。为了准确地预测股票价格，于卓熙等人（2018）利用广义回归神经网络（GRNN）进行建模并对结果进行评估，发现模型的评估指标如均方误差等都优于 ARIMA 模型，并利用主成分分析降低了影响股票价格的指标的维度。

而长短期记忆神经网络（LSTM）的选择性记忆功能使其能够对金融价格

时间序列进行更深入的分析。Fischer 和 Krauss（2018）在预测标普 500 指数的研究中发现，相对于以往研究中的无记忆分类模型架构，LSTM 在金融时间序列预测上的表现更好。Persio（2017）使用 RNNs、LSTM 和 GRU 对谷歌股票价格进行了比较研究，也发现 LSTM 在时间序列预测的相关研究中表现出色，并表现出模型良好的解释能力。

从预测精度角度，LSTM 模型也表现出了良好的特质。谷丽琼等人（2020）使用了基于 GRU 模型的注意力机制预测模型，并使用科大讯飞的股票数据作为样本，研究结果表明，在预测结果评价指标时，基于注意力机制的 GRU 大大优于其他模型。Hosker 等人（2018）预测了芝加哥期权交易所波动率指数的未来市场波动，发现与线性回归、主成分分析和 ARIMA 方法相比，RNN 和 LSTM 模型的平均平方误差、平均绝对误差更低，以及解释方差和相关性更高，表明它们在进行预测时更加准确。Minh 等人（2018）利用简化的 LSTM 模型，即 GRU 模型，发现该模型在金融时间序列方面的预测也出现优秀的效果。黄颖和杨会杰（2021）利用 XGBoost 模型进行特征优化并预测黄金价格涨跌趋势，并于 LSTM 模型进行比较，发现 LSTM 的模型预测效果优于 XGBoost。为了估计未来 10 天的股票收盘价，黎镭等人（2018）采用了 GRU 神经网络模型，展示了该模型的卓越的学习和概括能力。Zhang 等人（2019）利用 LSTM 和 MLP 模型，建立了对抗性网络来预测股市收盘价，发现模型的预测精度有了显著的改良。

深度学习回归方法同时也被应用在横截面策略中。Gu 等人（2017）使用各种线性的、基于树和神经网络模型，展示了非线性方法的卓越性能，也就是熟知的深度神经网络，特别是其中的三层多层感知器（MLP），具有最佳的样本外预测拟合优度。作者每月通过使用模型预测对股票进行排名来构建机器学习投资组合，其中表现最好的策略是四层 MLP，该 MLP 将前十分位数与预测的前十分位数进行交易。在其他文献中，Kim（2019）采用了类似的方法，他使用了自动编码器和去噪自动编码器架构，并将波动性缩放也纳入该模型中。虽然基础深度神经网络的结果是有一定意义的，但它们并没有考虑更现代的时间序列预测架构，例如长短期记忆神经网络，因此本文在为深度动量网络评估的架构中增加了这个模型。

由于 DeepMind 的 AlphaGo Zero 技术的成功（Silver 等人，2017），强化学习（RL）在最近也引起了许多关注。Kolm 和 Ritter（2019）发现强化学习因其在动态环境中路径相关操作的优秀能力而受到重视，在最佳执行和自动对冲的背景下的表现尤为出色。然而，强化学习方法通常需要一个逼真的模拟环境（用于 Q 学习或策略梯度方法）或世界模型（对于基于模型的深度强化学习）在训练期间向代理提供反馈，但这两者在实践中都很难获得，因此本文暂不考虑强化学习在时间序列动量的应用。

2.5 简要评述

本章主要对时间序列动量、金融领域的机器学习以及 CTA 基金的基准相关研究进行了系统性的梳理，这有利于对 CTA 基金市场的时间序列动量效应有一个更为深刻的认识，并做出自己独立的思考。通过对国内外文献的梳理，发现时间序列动量以及 CTA 基金相关研究都在迅速完善，并且随着人工智能相关技术的发展，研究越来越趋于细化并包含了多角度多方面的考虑。

（1）国内外对于基金绩效实证研究大部分的研究理论都是基于股票市场，少有主要基于期货市场的量化投资研究，更鲜有以 CTA 策略为主的相关研究。由于 CTA 基金的投资方式与普通共同基金不同，共同基金可以利用如股票市场大盘的指数等方式来很好地衡量基准，然而 CTA 基金的基准在以往的研究中更加依靠管理者管理能力指数来度量，根据 Fung 和 Hsieh（2004），如果仅依靠基金管理人业绩构建的基准对 CTA 基金收益进行评估，几乎不能反映基金的风格漂移和行业的风险敞口等信息。因此，结合我国期货市场的实际情况，选用国内期货市场数据对 CTA 市场利用资产对基准进行的建模进行研究探讨具有很大的必要性。

（2）在动量效应的发展和演变方面，传统的横截面动量相关的研究已经比较充分，国内外学者在全球主要的金融市场上都对横截面动量的重要性进行评估，但对于本文的研究对象也就是国内的个体 CTA 基金而言，Ham 等人（2019）发现中国期货市场相较于横截面动量因子，时间序列动量因子对于国内市场的影响较为显著。因此，本文结合实际情况，利用国内期货市场数据对时间序列动量进行建模，以此构建 CTA 行业基于市场的基准，对个体 CTA

基金利用时间序列动量单因子模型进行绩效测度和归因研究。

(3) 在深度学习在金融市场的应用方面, Moskowitz 等人 (2012) 提出时间序列动量模型对时间序列动量效应进行了建模, 建模后, Baz 等人 (2015) 以及 Gert 和 Peter (2017) 等学者通过对于趋势和仓位估计的定义变化来对模型进行了改进, 但并没有合理的研究来支撑这些手工修改的变量定义过程, 并缺少对模型的优化过程。而深度学习在金融时间序列的估计的表现越来越好, 因此本文尝试利用深度神经网络根据风险调整后的性能指标作为自定义的损失函数对其进行训练, 即直接从数据本身学习最佳训练规则, 规避了人工手动定义的过程, 保证了一定的客观性, 这也是本文与以往研究的不同和创新之处。

3. 研究设计

3.1 样本数据

3.1.1 期货数据

由于国内 CTA 基金的主要投资标的为商品期货，因此本文选取了 2012-2021 年共 10 年国内的 70 只商品期货主力连续合约的日度价格序列作为样本数据。本文所采用的期货数据主要来源于 Wind 数据库。具体数据见附表 A.1-1。

3.1.2 CTA 基准指数

由于 CTA 基金使用趋势跟踪策略的特殊性，因此不像股票基金可以直接使用如沪深 300，上证 50 等基于资产的指数作为基准，目前国内仅存在着基于基金业绩构成的基准指数。本文收集市场上各大 CTA 基准指数进行研究，综合考虑数据的易得性以及本文实际研究目的，选用以下 CTA 指数来进行后文的实证研究：

(1) 万得管理期货私募基金指数：主要使用的基金是活跃的，剔除已终止的产品，剔除分级基金，剔除不为母子基金的子基金。剔除投顾不是私募基金公司的产品，剔除私募基金管理人出现重大经营违规事件（重大遗漏、异常机构、虚假填报、相关主体存在不良诚信记录、违反八条底线、失联机构）的产品，产品成立时间大于等于六个月，产品到期时间大于等于两年。私募基金管理人管理规模大于等于 1 亿。最新净值近一个月有更新。产品净值按周公布，周频率的基金净值更新完整度占全年的比率大于等于 90%。数据来源于 Wind 数据库。

(2) 通联管理期货指数：使用连续 4 周每周都有披露净值且不都为 1、成立满 1 年、近 1 个月净值发布密度需大于 80%的非分级基金，剔除清盘状

态、发生重大违规违法事件、周回报标准差超过 3 个标准差的净值异常基金。数据来源于通联期货数据库。

(3) 道口 CTA 型私募基金指数：使用成立超过 6 个月、非分级基金，选取商品型基金和宏观对冲型基金中以商品期货为主要标的的私募基金，剔除清盘状态、暂停公布净值、合同变更以及基金公司发生重大违规违法事件的基金。数据来源于清华五道口金融学院。

(4) 朝阳永续五亿私募 CTA 指数：使用管理规模不低于 5 亿元、业绩公开透明、主要采用 CTA 策略并且可投的 31 家头部私募基金。数据来源于朝阳永续数据库。

3.1.3 CTA 基金数据

CTA 基金自从 2010 年就被引入到了国内市场，但直到 2013 年，国内私募基金行业得到了相关法律法规的认可之后，CTA 基金方才进入广大投资者的视野，因此本文选取了 2014 年 1 月 1 日至 2021 年 6 月 1 日成立的 CTA 基金，样本基金使用的交易策略主要为趋势跟踪策略。同时为了保证后续实证结果的稳健性，使用成立以来披露净值密度超过 50% 的基金，剔除了明显的异常值，只使用按照日度或者周度披露净值数据的 CTA 基金，最终得到个 384 有效 CTA 基金样本，所有数据均来源 Wind 数据库、通联数据库以及私募排排网。

3.1.4 无风险利率

无风险收益率是指在无风险投资中获得的收益率。它是无套利均衡定价中远期和期权定价的一个关键参数，也是本文构建的深度动量网络的一个重要评估指标。现有的无风险利率主要分为两类：定性分析方法是指利用金融市场上存在的某种利率作为无风险利率；定量分析技术是指使用相关的研究模型来对无风险利率进行建模得到的利率。

通过整理以往学者的相关研究结果后发现，定性分析方法在学术研究中主要用银行的存款或者贷款利率、银行间回购利率和国债利率等来近似替代，这也是国内外学者使用得最广泛的方法。当然，由于这些都不是理论上的无

风险资产，只是对无风险资产的一种近似替代，因此可能会出现一定的偏差性。另一种方法为定量分析，是指利用相关的模型进行建模研究从而计算出准确的无风险收益率，但这些方法还没有在现实中得到证实，因此不是很实用。

结合国内外相关学者研究，考虑到简便性与数据易得性，本文使用中国人民银行 3 个月整存整取利率作为本文的无风险收益率。

3.2 经典时间序列动量策略

动量策略传统上被分为两类：截面动量和时间序列动量。截面动量主要专注于证券于彼此之间的相对表现，买入相对赢家而卖出相对输家。将股票按照其过去几个月的收益排序，研究表明过去表现更好的股票在未来一个月也将有更好的表现。截面动量的表现在不同时间、不同市场以及不同资产类别上都是稳定的。

时间序列动量将这个想法扩展到了资产自身的历史收益上，根据资产自身的历史表现构建投资组合。该想法最初由 Moskowitz 等人（2012）提出，他们利用波动率放缩描述了具体的策略并根据过去一年的收益决定交易头寸。该策略的具体表达式如下：

$$r_{t,t+1}^{TSMOM} = SIGN(r_{t-12,t}^s) \frac{\sigma_{tgt}}{\sigma_t^s} r_{t,t+1}^s \quad (3.1)$$

$$SIGN(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{其它} \end{cases} \quad (3.2)$$

其中 s 表示正在交易的资产或合约， t 是交易时间段， σ_t 是资产的事后波动率的估计， σ_{tgt} 为目标波动率。简单来说，如果资产在过去 12 个月内是正收益率，那么则做多，否则就做空。

在这个规则下，由于已经将资产收益率用波动率进行归一化处理，因此就可以通过组合风险加权资产收益率和将风险加权资产的杠杆转换为期望的目标波动率来构建跨资产的时间序列动量投资组合。也就是如下公式：

$$r_{t,t+1}^{TSMOM} = \frac{1}{N} \sum_{s=1}^N SIGN(r_{t-12,t}^s) \frac{\sigma_{tgt}}{\sigma_t^s} r_{t,t+1}^s \quad (3.3)$$

其中 N 为资产的数量。

为了将以上公式表达得更简练，Lim 等人（2019）认为任何种类的时间序列动量策略都有两个组成成分。首先，预测趋势部分，也就是说这部分将表示下一个交易周期的交易方向以及时间跨度。其次是仓位大小，也就是说将表示下个交易周期的持仓量。根据以上定义，经典的时间序列动量策略可以用以下两个公式表示：

$$Y = r_{t-12,t} \quad (3.4)$$

$$X = \text{SIGN}(Y) \quad (3.5)$$

虽然时间序列动量策略在捕捉时间序列动量上是有效的，但因为模型的一些缺点，招致了国内外学者的许多批评：

（1）滞后的仓位估计：在该经典模型中，投资该资产所需要的仓位大小完全由波动率的估计决定，然而波动率的估计是事后波动率的估计，如果模型具有预测能力，那么应该能够及时地进行更优化的调整，在预期回报更大的时候承担更大的风险，即选择更大的仓位，经典的时间序列动量没有做到这一点。

（2）过于简单的趋势估计：经典的时间序列动量模型仅仅用了过去一年的收益率进行未来的趋势估计，然而在现实情况中，这种估计过于随意简单，所以应该构建一个有着更丰富特征空间的统计模型来进行趋势估计。

（3）模型无优化：经典的时间序列动量模型没有针对任何指标进行优化。根据通常的投资偏好，投资者可能希望投资策略能够根据如最大化收益率、最小化回撤等指标进行优化，通过这些优化指标，就可以去定义符合投资者偏好的时间序列动量策略。然而经典的时间序列动量策略却没有做到。

3.3 深度动量网络的构建

机器学习，就其最广泛的定义而言，是一项能够完成无法直接通过编程进行的活动的技术。机器学习现在被广泛用于图像识别和趋势预测等领域，它是一种模仿或模拟人类认知过程和习惯的手段，通过利用机器学习和训练大量相关数据来产生解决特定问题的模型。

机器学习根据数据是否有标签可以分为监督学习和无监督学习。监督学习是指计算机从一组已经标记的训练数据中学习，这些训练数据通常是由人

工进行标记，而标签表明了数据所属的类别，计算机将从这些标记的训练数据中构建模型，并在学习的过程中不断地对模型进行优化，直到生成最后的完整模型。

无监督学习是一个相对的概念，它反映了从无标签的训练集进行训练的过程。机器学习算法将未标记的数据归类，在每个类别中寻找共同之处，探索其中的联系，从而利用潜在的相关性。决策树、支持向量机、随机森林、马尔科夫、多层感知器和其他机器学习技术被普遍采用。

而深度学习这个概念由 Hinton 等人（2006）首次提出的，它是一个基于人工智能神经网络的前沿机器学习概念，使用多层神经网络来模拟人类大脑，建立无监督的机器学习模型。

深度学习的深度指的是人工神经网络的层数。在检索样本特征方面，深度学习优于人工神经网络学习。深度学习与人工神经网络的不同之处在于其训练方法。深度学习模型一般而言是由相当数量的神经元组成，它们通过分层学习进行无监督训练，将低层神经元利用输入和输出之间的线性关系转化为高层神经元的抽象表示，这一特点解决了传统人工神经网络费时费力的训练问题。而根据梯度下降法，由于权重和偏置被设置为任意初始值，传统的人工神经网络各层的权重和偏置参数是通过模型输出值之间的反向传播误差而改变。因此，模型训练过程非常有难度，因为梯度转移过程可能非常容易出现梯度消失和梯度爆炸的现象。

本文通过 Lim 等人（2019）提出的深度动量网络（DMN）解决了经典时间序列动量的缺点，利用构造适合的损失函数以及神经网络架构来进行时间序列动量的预测。

3.3.1 深度学习的神经网络架构

本节将介绍在深度动量网络模型中使用的 MLP、LSTM 神经网络架构，本文使用的网络架构通过 Tensorflow 中的 Keras API 建立。

多层感知机（Multi-Layer Perceptron, MLP）由感知机（PLA: Perceptron Learning Algorithm）推广而来，它最主要的特点是有多个神经元层。感知机是单个神经元模型，是神经网络的前身。神经网络的强大之处在于它们能够学

习训练数据中的表示，以及如何将其与想要预测的输出变量联系起来。从数学上讲，它们能够学习任意的映射函数，并且已经被证明是一种通用的近似算法。

神经网络的预测能力来自网络的分层或多层结构，而多层感知机是指具有至少三层节点，输入层，一些中间层和输出层的神经网络。给定层中的每个节点都连接到相邻层中的每个节点。输入层接收数据，中间层计算数据，输出层输出结果。

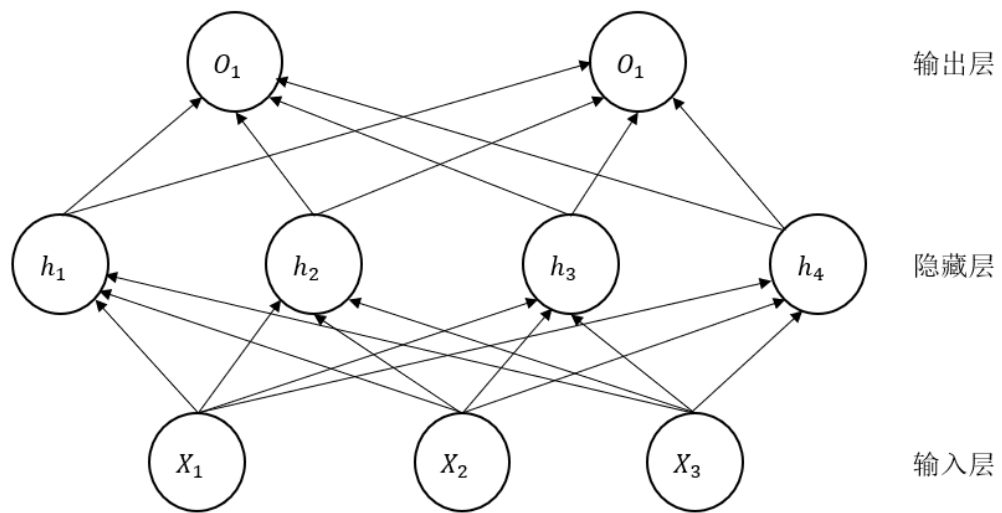


图 3-1 多层感知机网络模型

本文使用两层神经网络合并非线性效应，可以用如下公式表示：

$$h_t^{(i)} = \tanh(W_h u_{t-\tau it}^{(i)} + b_h) \quad (3.6)$$

$$z_t^{(i)} = g(W_z h_t^{(i)} + b_z) \quad (3.7)$$

其中 $h_t^{(i)}$ 是隐藏层，通过 \tanh 函数来激活， W 和 b 分别是层权重和偏差项， Z 通过预测任务灵活可变。而 \tanh （双曲正切）函数能够将元素的值变换到 $(-1, 1)$ ，方便生成本文研究中仓位的预测大小：

$$\tanh(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)} \quad (3.8)$$

循环神经网络（Recurrent Neural network, RNN）是指一类具有短期记忆能力的神经网络，适合用于处理视频、语音、文本等与时间序列相关的问题。在循环神经网络中，神经元不但可以接收其他神经元的信息，还可以接收自

身的信息，形成具有环路的网络结构。其基本结构如下图所示：

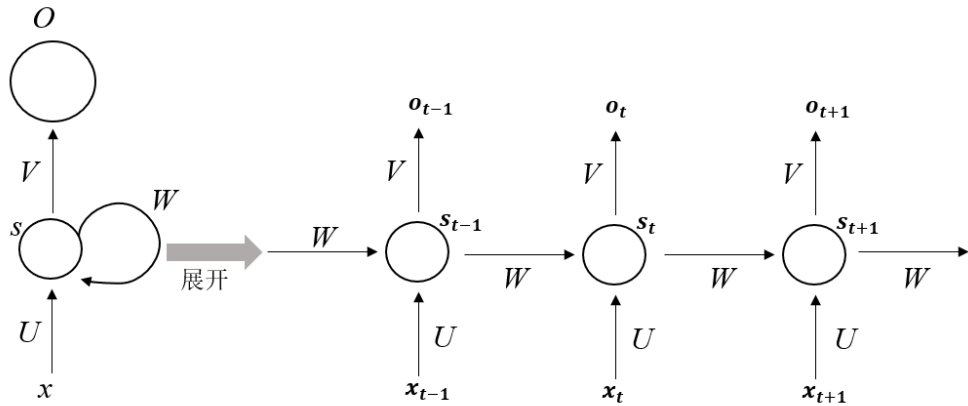


图 3-2 循环神经网络模型

在图 3-2 中， U 、 V 、 W 分别代表着输入层和隐含层、隐含层和输出层以及隐含层之间的连接权值矩阵， x 、 o 、和 s 分别代表着输入、输出和隐藏状态。 x_t 是 t 时刻的输入； s_t 是 t 时刻的隐含层状态， $s_t = f(Ux_t + Ws_{t-1})$ ，函数 f 一般选择 relu 或者 tanh 函数； s_{t-1} 为上一时刻的隐含层状态，初始值通常设为 0； o_t 是 t 时刻的输出 $o_t = \text{soft max}(Vs_t)$ 。

根据图 3-2，连接不仅存在于相邻的层与层之间（比如输入层-隐藏层），还存在于时间维度上的隐藏层与隐藏层之间，也就是说，RNN 模型可以实现记忆的功能。更进一步地，RNN 模型是一个关于时间序列上的深度学习模型。然而，由于梯度弥散，导致模型在时间序列长度过长时，无法根据后续序列来修改前向序列的参数，使得前向序列无法很好地做特征提取。因此 RNN 模型在面对长序列数据时，很容易便遭遇梯度弥散，使得 RNN 只具备短期记忆。即 RNN 面对长序列数据，仅可获取较近的序列的信息，而对较早期的序列不具备记忆功能，从而丢失信息。

根据上述 RNN 的特点，长短期记忆模型(Long Short-Term Memory, LSTM)对 RNN 模型进行了改进，它是一种特殊的 RNN。它可以避免 RNN 模型的梯度弥散，在时间序列预测上得到了广泛应用。

LSTM 模型与其他深度学习模型的不同在于它与众不同的细胞状态，在长期状态下可以学习和遗忘序列数据信息。细胞内部由遗忘门、输入门、输出门三大门限单元构成。

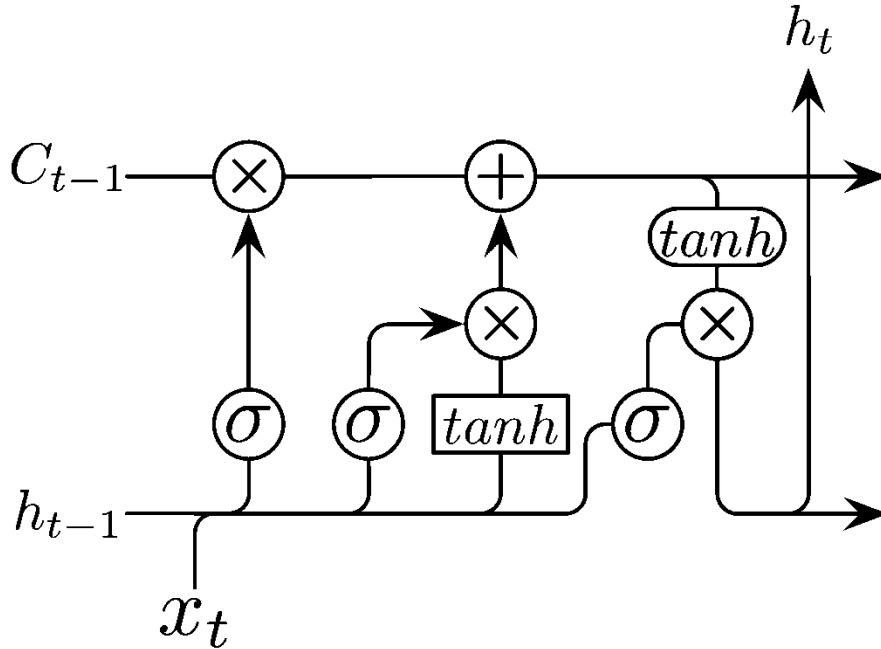


图 3-3 长短期记忆网络模型

遗忘门 (forget gate)，顾名思义，是用来控制是否对数据进行遗忘，在 LSTM 中即以一定的概率控制是否遗忘上一层的隐藏细胞状态，公式如下：

$$f_t^{(i)} = \sigma(W_f u_t^{(i)} + V_f h_{t-1}^{(i)} + b_f) \quad (3.9)$$

其中， h_{t-1} 是 t 时刻上一时刻的输出， u_t 是 t 时刻的输入， W_f 是各个变量的权重， b_f 是偏置项，激活函数 σ 是 sigmoid 函数， f_t 代表遗忘门，介于 0 到 1 之间，表示输出给每个在细胞状态中的值，0 表示全部舍弃，1 表示全部通过。

输入门 (input gate) 负责处理当前序列位置的输入，实现此过程包含两个步骤：首先，由 sigmoid 函数作为全连接层，得到一个 0 到 1 之间的 i ，确定当前细胞状态信息的更新程度；其次，选择 tanh 函数作为激活函数的全连接层，再利用输入门加入更新后的信息，储存到长期状态中，最终得到更新后的细胞状态值 $c_t^{(i)}$ ，如公式所示：

$$i_t^{(i)} = \sigma(W_i u_t^{(i)} + V_i h_{t-1}^{(i)} + b_i) \quad (3.10)$$

$$c_t^{(i)} = f_t^{(i)} \odot c_{t-1}^{(i)} + i_t^{(i)} \odot \tanh(W_c u_t^{(i)} + V_c h_{t-1}^{(i)} + b_c) \quad (3.11)$$

其中 \odot 是 Hadamard 乘积， $i_t^{(i)}$ 代表输入门，与上述遗忘门具有相同机制， W_i 与 W_c 代表相应权重， b_i 与 b_c 代表相应的偏置， $c_{t-1}^{(i)}$ 表示上一时刻的细胞状态值。

输出门（output gate）决定最终得出确定输出的序列信息，长短期记忆神经网络结构依据细胞状态决定序列信息的输出程度，最初利用 sigmoid 函数判断细胞状态的输出部分，进而将细胞状态用 tanh 函数计算后再与之前的输出部分相乘，用公式表示为：

$$o_t^{(i)} = \sigma(W_o u_t^{(i)} + V_o h_{t-1}^{(i)} + b_o) \quad (3.12)$$

$$h_t^{(i)} = o_t^{(i)} \odot \tanh(c_t^{(i)}) \quad (3.13)$$

其中 $o_t^{(i)}$ 为输出门， W_o 与 V_o 表示输出门的权重， b_o 为偏置， $h_t^{(i)}$ 为当前单元的输出值。

综上所述，LSTM 的运作过程总结如下：当模型的细胞状态 $c_{t-1}^{(i)}$ 经过 LSTM 结构时，首先经过遗忘门过滤序列数据的部分信息，然后输入门获得新的数据，并直接输出长期状态 $c_t^{(i)}$ 。最后，输出门决定能够在长期状态下被保留的信息和被过滤的信息，进而形成短期状态 h 。LSTM 利用长短期序列数据的学习，保留了长期状态中的重要信息和信息量，使得模型达到被优化的目的。

3.3.2 基金绩效评估指标定义的损失函数

深度动量网络使用传统的神经网络架构将滞后的时间序列特征映射到下一个周期趋势或仓位大小预测。在传统的监督学习问题上，我们通常存在一个标记的数据集，即明确定义的训练目标，在本文介绍的时间序列模型上指最佳的趋势估计量 Y ，接着通过如均方误差或者绝对误差来计算出最佳的模型进行优化度量。然而在实际情况并不会存在这样一个带标记的数据集，因为本文的模型就是来确定趋势估计量 Y 的，如果能够定义好最佳的趋势估计量 Y ，那么该模型就没有意义了。因此，深度动量网络就放弃优化已知目标（ Y ）以及传统的损失函数，而是通过自定义损失函数优化其对金融交易指标的预测。

本文探索了两个损失函数，分别通过优化夏普比率和平均回报率构建：

$$\mathcal{L}_{Sharpe} = - \frac{\sum R_{s,t}}{\sqrt{\sum R_{s,t}^2 - (\sum R_{s,t})^2}} \quad (3.14)$$

$$\mathcal{L}_{returns}(\theta) = -\frac{1}{N} \sum R_{s,t} = -\frac{1}{N} \sum X_{s,t} \frac{\sigma_{tgt}}{\sigma_{s,t}} r_{t,t+1}^s \quad (3.15)$$

其中 s 表示资产或者合约， t 是时间段， r 是收益率， $\sigma_{s,t}$ 是资产的事后波动率估计， σ_{tgt} 是目标波动率。在损失函数中，由于深度动量网络头寸估计量 X 必须要处于 $[-1,1]$ 的范围内，也就是说 X 为正，那么为多头仓位，如果为负，那么则为空头仓位，通过预测和优化仓位大小 X ，目的是让深度动量网络模型能够选择恰当的时机来获得更高的收益率。

3.3.3 超参数选择

深度学习模型在训练的过程中，超参数的选择尤为重要，超参数主要指的是在模型训练之前设置好的参数。本文构造深度动量网络模型主要使用的超参数主要包括：批大小（batch size）、隐藏层、模型迭代次数（number of epoch）、每层神经元个数、学习率以及丢弃比例（dropout）等。

批大小是指在训练机器学习模型时，每次迭代时使用的训练样本数量。一般来说，训练样本越多，训练的精度就越高，但是如果批大小太大，训练时间就会变长，可能会导致训练效率降低。因此，在训练机器学习模型时，需要根据实际情况来选择合适的批大小，以获得最佳的训练效果。

隐藏层是指神经网络中位于输入层和输出层之间的其他各层的总称，不参与输入和输出，但可以捕捉输入和输出之间的复杂关系。隐藏层可以接收来自输入层的信号，并将其转换为更高级别的特征，然后将这些特征传递给输出层，以便输出层可以将其转换为有用的结果，而隐藏层的神经元个数在模型训练中的预测结果来说十分重要，如果隐藏层的神经元数量太少，模型将无法从数据中提取足够的特征，这可能导致训练结果不佳。如果神经元数量太多，模型在训练集上的表现会偏好，而在测试集上的表现则会偏差。

学习率是机器学习中一个重要的参数，它控制了每次迭代中权重更新的大小。学习率越大，每次更新的步长就越大，反之，学习率越小，每次更新的步长就越小。一般来说，如果学习率过大，模型可能会发散，而如果学习率过小，模型可能会收敛到局部最优解，因此，一个合适的学习率是模型训练过程中的关键。

丢弃比例是深度学习中一种常用的正则化技术，它可以防止神经网络过

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/568051006024006027>