



北京大学

# 硕士研究生学位论文

题目：

基于神经网络模型的弱有效  
市场假说检验

姓 名：夏心雨

学 号：2401212301

院 系：汇丰商学院

专 业：经济学

研究方向：

☒ 学术学位      ☐ 专业学位

二〇二四 年 四 月

## 版权声明

任何收存和保管本论文各种版本的单位和个人，未经本论文作者同意，不得将本论文转借他人，亦不得随意复制、抄录、拍照或以任何方式传播。否则，引起有碍作者著作权之问题，将可能承担法律责任。

## 摘要

本篇论文通过检验神经网络模型产生的预测是否能预测股票未来的横截面收益,来检验弱有效市场假说。如果神经网络模型预测能够带来稳定的截面超额收益,则表明技术分析仍有应用价值,与有效市场假说的弱形式相违背。经过选择和分析,文章选择网络结构为 CNN-GRU、损失函数为信息系数、以及因变量为对收益率进行截面排序后映射到标准正态累积分布反函数上的值 ( $r_{\text{norm}}$ ) 的神经网络模型来检验有效市场假说。基于模型预测值构造投资组合,对投资组合的表现进行了风险调整、技术信号中性化等操作的分析,证明了神经网络模型对股票截面有预测效果,表明中国 A 股市场尚未达到弱有效状态。最后,基于模型预测与私募量化基金 2024 年年初的大幅回撤行情,本篇论文也给出了主要针对中小盘和微盘股的分析。

关键词: 弱有效市场假说, 神经网络模型, 私募量化基金

## 目录

摘要.....	1
第一章 引言.....	1
第二章 文献综述与背景.....	3
2.1 国外股票市场有效性的研究.....	3
2.2 中国股票市场有效性的研究.....	5
2.3 神经网络模型的发展与应用.....	6
2.4 国内量化私募基金的发展与表现.....	7
第三章 数据与样本构造.....	10
3.1 样本的构造.....	10
3.2 样本的统计性描述.....	11
第四章 神经网络模型训练与预测.....	13
4.1 模型选择期.....	13
4.1.1 模型构建.....	13
① 卷积神经网络 (Convolutional Neural Network, CNN) .....	14
② 长短期记忆网络 (Long-Short Term Memory, LSTM) .....	15
③ 门控循环单元 (Gated Recurrent Unit, GRU) .....	16
④ CNN-LSTM.....	17
⑤ CNN-GRU .....	17
4.1.2 模型训练.....	18
① 损失函数的选择.....	19
② 损失函数中样本权重的选择.....	19
③ 目标值的选择.....	20
4.1.3 模型选择.....	20
4.2 主模型预测期 .....	22
第五章 基于模型预测进行截面选股的表现分析.....	25
5.1 基于 NNER 构造五分位投资组合.....	25
5.2 子周期内的预测能力 .....	26
5.3 风险调整后 NNER 投资组合的表现.....	27
5.3.1 Carhart 四因子模型回归.....	29
5.3.2 Fama-French 五因子模型回归.....	29

5.3.3 Hou-Xue-Zhang 四因子模型回归.....	30
5.4 技术指标中性化后 NNER 投资组合的表现.....	30
第六章 基于模型预测结果分析私募量化基金表现.....	34
6.1 市场指数的表现 .....	35
6.2 基于模型预测结果的分析.....	36
6.2.1 模型选股中的中小盘股票数量占比.....	36
6.2.2 NNER-5 组合收益率序列的回归分析.....	38
6.3 量化私募基金下跌的原因.....	38
6.3.1 雪球产品的敲入.....	38
6.3.2 量化中性策略的下跌.....	39
第七章 结论及展望.....	41
参考文献.....	42
北京大学学位论文原创性声明和使用授权说明.....	45



## 第一章 引言

在本篇论文中，我们通过神经网络算法来检验中国 A 股市场的有效性。我们将借鉴 Murray 等人（2024）的方法，通过训练和优化神经网络模型，基于模型产生的预测是否能预测股票未来的横截面收益，来检验弱有效市场假说。如果神经网络模型预测能够带来稳定的截面超额收益，则表明技术分析仍有应用价值，与有效市场假说的弱形式相违背。其中，本文采用的常用技术指标仅基于过去 12 个月股票的累计收益率，即能够从历史价格图中容易看出的数据。经过最终的模型预测和检验，我们发现基于深度学习模型的预测在测试期间（2016 年 1 月-2024 年 3 月）具有不错的截面选股效果和统计显著的预测能力，并且这种预测能力没有被已知的 HXZ 四因子、Carhart 四因子和 FF 五因子等所完全解释。

对于模型训练与预测，为了减轻与数据挖掘（Harvey, Liu, and Zhu（2016））或样本外预测能力（McLean and Pontti（2016），Green, Hand, and Zhang（2017））等相关的问题，我们使用 2000 年 1 月-2015 年 12 月的数据作为模型的初始训练集，使用 2016 年 1 月-2024 年 3 月的数据作为测试集。我们将整个时间区间分为模型选择期和主模型预测期。在模型选择期，我们考虑了多种神经网络架构，包括卷积神经网络（CNN）、长短期记忆网络（LSTM）、门控循环单元网络（GRU）、结合了 CNN 和 LSTM 或 GRU 的混合网络。同时，我们还尝试了不同的损失函数（如均方误差和信息系数）、赋予损失函数中每条样本不同的权重（所有样本等权、每个月内所有样本等权、每个月内按流通市值加权）、不同的目标值计算方式等，共提出了 80 种模型；在初始训练集上进行测试后，计算每个模型的预测值与股票月收益率的截面 Spearman 相关系数的时序平均值，最终按照相关性高低选择了网络结构为 CNN-GRU、损失函数为信息系数(IC)、以及因变量为对收益率进行截面排序后映射到标准正态累积分布反函数上的值( $r_{\text{norm}}$ )的模型作为论文后续检验和分析的主模型。在主模型的训练和预测期间，训练集和测试集的时间窗口是不断拓展的，即对于不同的测试集时间区间，所使用的训练集数据的时间区间可能是不同的。基于主模型，我们将生成对 201601-202403 期间的股票收益率预测。

基于模型预测，本篇论文将考察对股票进行截面排序选股而形成的投资组合的表现。除了考察投资组合在全样本区间和在子区间内的收益能力，我们还将从风险调整和技术信号中性化两个角度综合考察神经网络模型预测的选股能力。风险调整方面，对经典的 HXZ 四因子、Carhart 四因子和 FF 五因子模型进行回归，从全样本周期和日频子周期两个角度，研究了风险调整后基于神经网络预测排序选择的投资组合的表现。我

们也基于 Neely、Rapach、Tu 和 Zhou (2014) 研究的 14 个技术信号 (记为 NRTZ 信号)，通过对每一个技术信号中性化，来构造基于神经网络模型预测排序的投资组合。经过以上的检验，我们发现，基于神经网络模型预测排序构造的投资组合，经过风险调整和技术信号中性化后仍有显著的正收益，说明神经网络模型能够基于简单常用的技术指标构造出有利可图的组合，中国股票市场尚未达到弱有效状态。

最后，我们也基于神经网络模型的选股结果对 2024 年年初国内量化对冲基金的下跌行情进行了分析。我们的主要结论是小市值板块的流动性挤兑引发了本次量化私募基金产品的 beta 和 alpha 的双暴跌。

本篇论文组织如下：第二章进行了相关背景的梳理和文献综述；第三章阐述了样本的构造、组成及其统计性描述；第四章描述了我们如何训练和优化神经网络模型，并且基于选中的神经网络模型进行预测；第五章给出了我们基于模型预测，证明基于神经网络模型可以成功预测未来股票收益的横截面；第六章分析了今年国内量化对冲基金的衰落。第七章进行了总结。



## 第二章 文献综述与背景

Paul Samuelson 在 1965 年提出的有效市场假说 (Efficient Market Hypothesis, EMH) 构成了金融学理论的基石。该理论的核心观点是, 在一个有效的市场中, 所有可用信息会迅速并准确地反映在证券价格上, 意味着利用这些已知信息进行交易的投资者无法实现超过正常市场回报的收益。Fama 对有效市场假说的推进在于他对该理念进行了系统化阐述, 并发展出一个全面的理论体系。依据 Robert 和 Fama 在 1970 年的划分, 市场有效性可以根据信息反映到证券价格中的速度与程度划分为以下三个层次:

1. **弱式有效市场假说:** 在弱型有效市场上, 证券的价格已经充分反映了所有的历史价格信息 (包括成交量、成交量、融资金额等。如果弱有效市场假说成立, 则股票价格的技术分析失去作用, 基本分析还可能帮助投资者获得超额利润。
2. **半强式有效市场假说:** 在半强型有效市场上, 证券的价格已充分反映出所有可以公开获取的信息, 除了历史价格信息, 还包括公司的财务报告、新闻报道、经济数据等。假如投资者能迅速获得这些信息, 股价应该迅速作出反应。如果半强式度有效市场假说成立, 则市场中利用技术分析和基本分析都失去作用, 只有内幕消息可能获得超额利润。
3. **强式有效市场假说:** 这是有效市场假说中最强的形式, 它认为所有信息, 包括公开信息和非公开信息 (如内幕信息), 都已经完全反映在了证券价格中。在这种市场中, 即使是拥有内幕信息的投资者也无法获得超额回报, 因为市场价格已经包含了所有可能影响价格的信息。

### 2.1 国外股票市场有效性的研究

有效市场假说已经有较长时间的发展历史, 市场有效性也是所有市场参与者和学者长期关注的重要话题, 因此, 已经有了大量研究不同股票市场有效性的文献。已有的研究文献从不同的研究方法、不同的数据来源和时间跨度、不同的市场环境和不同的频率周期等多个维度出发, 得出的结论也不尽相同。

Coronel Brizio 等研究者 (2007) 通过将墨西哥股市指数与道琼斯工业平均指数的数据结合起来进行分析, 选择了 1978 至 2006 年的样本数据进行实证研究, 利用收益自相关函数和波动的非趋势性进行比较分析, 最终发现墨西哥股市的有效性相较于过去有所增强; Konak 和 Seker (2014) 研究了 FTSE 100 指数的演变方式, 以及它的演变是否支持有效市场假说, 他们发现 2001 年至 2009 年间, FTSE 100 指数遵循随机漫步理论, 即符合弱形式的有效市场假说; Aktan 等人 (2018) 利用 2006 年 6 月至 2017 年 6 月的月度数据检验了欧洲 32 个股票市场的弱形式市场有效性, 他们将欧洲市场分

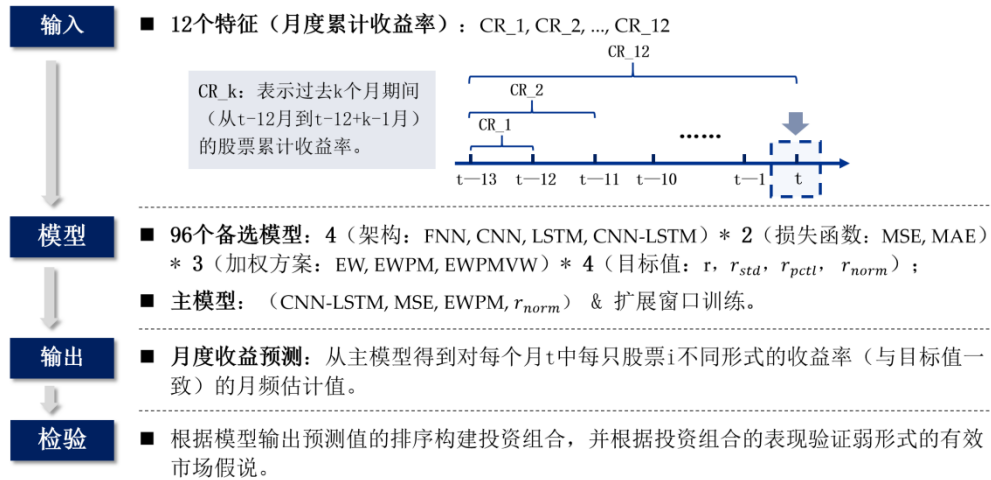
为三类:前沿市场、新兴市场和发达市场。通过六种不同的新开发的非线性面板根检验表明,不同经济发展水平与弱形式市场有效性之间存在一定的关系,且新兴市场和发达市场满足弱有效状态,而前沿市场不满足。

Lee 等人(2010)研究了 1999 年 1 月至 2007 年 5 月期间 32 个发达国家和 26 个发展中国家实际股票价格的平稳性,得出其股票市场有效性不高的结论;Karadagl 和 Donmez(2012)检验了希腊、匈牙利、波兰、俄罗斯和土耳其期货市场的弱形式市场有效性,利用 2005 年 9 月至 2011 年 6 月的月度数据,他们进行了几种不同的测试(ADE, PP, KSS 单位根测试, IPS 和 UO 面板单位根测试)来检验市场有效性,发现线性单位根检验(ADE、PP)和线性面板检验(IPS 面板检验)显示市场满足弱有效状态,而非线性检验(KSS 单位根检验和 UO 面板单位根检验)则得出相反的结果。Birau(2013)采用 2007 年 1 月至 2011 年 12 月的数据研究了罗马尼亚和匈牙利资本市场的弱形式有效市场假说,得出两国资本市场均未达到若有效状态的结论。

美国股票市场作为全球发展最为成熟的金融市场之一,其市场有效性一直是学术界和业界关注的焦点;然而,对于美国股市有效性的研究,不同学者的结论存在分歧。Shiller(1984)通过对美国股市长期收益率的研究指出,长期收益率序列相比短期序列显示出更显著的负自相关性,而短期收益率的相关系数较小。这表明美国股市短期内表现出一定程度的弱式有效性,但从长期来看,并未达到弱式有效的标准;Toth 和 Kertesz(2006)观察到,在纽约证券交易所,即使是高频数据,时间依赖的互相关函数的不对称性呈现减弱趋势,其峰值位置向原点偏移,并且变得更为尖锐和高耸,这为市场有效性的提升提供了证据;Ito 和 Sugiyama(2009)利用状态空间模型来估计时变的自回归系数,他们发现在 1955 年至 2006 年的样本期间,美国这一成熟的股票市场展现了不同程度的有效性,并且指出市场在过去几年变得更加有效,表现为利差减小,内部流动性增强等。

而 Murray 等人(2024)基于 1927 年 1 月到 2022 年 12 月期间的月累计收益率数据,通过机器学习模型预测月度收益,基于模型预测构造投资组合,发现多空组合的收益在不进行任何处理、在全周期和子周期、在风险因子模型调整后以及对技术信号中性化后都是显著为正的,从而得出美国股票市场尚未达到弱有效状态的结论。

图 1. Murray 等人 (2024) 的文章架构



## 2.2 中国股票市场有效性的研究

对于国内 A 股市场而言，尽管国内金融股票市场起步较晚，发展不够完善且多变，也仍有众多学者对其有效性进行了研究，主要研究对象为有效市场假说的弱有效形式，且随着研究方法、时间段或发生时间的变化，研究结论也有所不同。

一部分学者持观点认为中国股市已经符合弱式有效市场的标准。2001 年，邓子来与胡健通过事件研究法分析了中国股市，并得出了当时市场已实现弱式有效性的结论，但还未达到半强式有效性；李晓明和张兵（2003）针对新兴股市的特性，运用了渐进有效性检验，并结合时变系数自回归模型（AR(2)）以及对波动集群和异方差性的考量，通过逐年度的检验，确认了 1997 年中国股市呈现弱式有效性；常嘉佳和马子红（2013）利用 GARCH 模型对 2000 至 2012 年的上证指数进行了分析，指出尽管上海股市与成熟市场存在差距，但认为其已达到弱式有效性；朱瑞在 2018 年的研究中，通过随机游走模型分析，得出了沪深股市达到弱式有效性的结论；2022 年，罗超予和杨启帆对沪深 300 指数的对数收益率进行了单位根和自相关平稳性检验，得出中国股市已实现弱式有效市场的结论。

尽管一些学者认为中国股市已经达到弱式有效性，但也有研究者持相反看法，认为中国股市尚未满足弱式有效市场的条件。吴市农（1996）选取了上海和深圳证券交易所的部分股票及两市的股指日收益率作为样本，通过检验这些序列是否存在自相关性来探讨市场有效性。研究结果指出，这些股票收益率序列显著地表现出自相关特性，具有统计规律性，表明市场并未达到弱式有效性；闫冀楠和张维（1997）采用单位根检验和独立性检验方法，对 1991 至 1996 年的上证周收盘价数据进行了分析。他们的研究结果显示，序列中存在单位根和一阶自相关，进一步支持了中国股市未实现弱式有效性的观点；2000 年，张亦春使用白噪声、随机游走和鞅过程等检验方法对 1993 至 2000

年的上证 A 股指数进行了分析，发现价格波动并不是随机的，表明市场未达到弱式有效性；王若凡（2009）对比了 2005 至 2008 年间沪、港、澳三地股市的交易指数。研究发现，尽管澳大利亚和香港股市呈现弱式有效性特征，但上海股市显示出信息传递的滞后、对外部冲击的持续性反应，以及更明显的投机性，认为沪市未符合弱式有效性的标准。这些研究表明，基于不同时间段、不同数据集、不同研究对象或不同研究方法，对于中国股市是否达到弱式有效性，可能会得出不同的结论。

本篇论文也将主要立足于中国 A 股市场，对中国股票市场的弱有效形式进行检验，本文选择在国内市场进行弱效市场假说的原因主要有以下几个方面：首先，从以上的文献综述内容来说，对于国内目前的市场有效性来说，先前的研究者在不同的历史时间段，采用不同的数据集、研究对象或者研究方法都可能导致得到不一致的结论，因此，本文也期望采用不同的验证方法（神经网络模型）对中国股票市场的有效性进行研究。

其次，从理论意义上来说，有效市场假说在一定程度上研究出了股票价格和与市场信息是否存在某种关系，由最初的提出到现在学者对其的不断改进，这一过程促进着全球金融领域的蓬勃发展，直至今日这一理论已演变为金融领域的一大重要理论之一，对这一关系的探讨能够为投资者和相关领域的研究人员提供有效的指引和依据。

最后，从现实意义来说，在弱有效市场假说下，技术分析应该是一种无效的的投资技术，技术分析方面的相关学术研究也广泛支持了这一说法。然而在业界的实际应用中，技术分析仍被投资经理广泛使用这说明。技术分析在业界长期以来的广泛使用表明，它的优点在学术研究中可能没有被充分发现和挖掘，因此有必要进行进一步的分析和验证。

## 2.3 神经网络模型的发展与应用

神经网络技术，模仿大脑神经元间的连接与互动，旨在开发出能够执行人工智能任务的机器学习算法。在这一领域的早期探索中，神经生理学家沃伦·麦卡洛克和数学家沃尔特·皮茨建立了一种简化的数学模型，该模型以生物神经元为基础，被称为 MP 模型。尽管 MP 模型的功能有限，仅能处理基础的逻辑判断任务，但它为后续神经网络的深入研究打下了坚实的基础。

神经网络通过模仿大脑神经元间的连接与互动，旨在开发出能够类似人工智能的机器学习算法。神经生理学家沃伦·麦卡洛克和数学家沃尔特·皮茨提出了一种以生物神经元为基础的简化数学模型，这个模型后来被称为 MP 神经元模型，尽管 MP 模型只能完成简单的逻辑判定，但也为后续的神经网络研究奠定了基础。1958 年，心理学家弗兰克·罗森布拉特创造了感知机，这是首个能够自动通过学习过程调整其权重的单层

神经网络模型，能够执行一些基础的图像识别任务。然而，1969 年，马文·明斯基等研究者提出感知器存在局限性，它们不能解决所谓的线性不可分问题，即无法利用单个模型来区分所有可能的数据类别，这也使得神经网络的研究陷入了第一次低谷。

直至 1986 年，Rumelhart 等人提出了误差反向传播算法（Back Propagation），简称 BP 算法。BP 算法使得多层神经网络的训练成为可能，有效克服了感知器在处理线性不可分问题上的不足，显著提升了神经网络的处理和学习能力，并推动了其在多个领域的应用。1989 年，Yann Lecun 等人开发了一个包含 7 层的卷积神经网络 LeNet-5，在手写数字识别任务中实现了 98% 的准确率。这一时期的神经网络研究获得了迅猛的进展，众多学者对神经网络的多种方法进行了深入探讨和研究，包括但不限于卷积神经网络（CNN）、循环神经网络（RNN）、玻尔兹曼机、Hopfield 网络以及自组织特征映射（SOM）网络等，每一种模型都在特定领域展现出了独特的优势。然而，尽管 BP 算法突破了神经网络发展的瓶颈，但这些复杂的高维非线性模型在实际计算中仍面临诸多挑战。例如训练深层神经网络时遇到的梯度消失或梯度爆炸问题，以及计算资源有限等因素。

2006 年，Hinton 带领团队通过运用限制玻尔兹曼机对神经网络的层次结构进行预训练，逐层地提炼出数据中的深层次特征。这种方法很快被扩展到其他的神经网络架构之中，有效地增强了模型对训练数据集之外数据的泛化性能。随着硬件的计算能力增强和算法的优化，研究者也能够构建并训练更加深层的网络。Hinton 等人随后重新界定了神经网络的范畴，提出了“深度学习”这一术语，并对这一领域进行了大力推广。2012 年，AlexNet 在 ImageNet 竞赛中实现了重大突破，这一成就凸显了通过深层网络结构进行非线性变换以提取复杂特征的能力，大幅提高了图像分类的准确度。尽管 AlexNet 的代码行数与之前的 LeNet-5 等浅层模型相比没有大幅增加，但其在深度学习领域的贡献在于对深层卷积神经网络的创新应用和实证验证。此后，硬件设施的持续进步和互联网带来的海量数据，都为神经网络技术的进一步发展提供了坚实基础和大力支持。

## 2.4 国内量化私募基金的发展与表现

中国量化私募行业发展历程大致可以分为三个主要阶段：2010 年到 2014 年的行业萌芽期、2015 年到 2021 年的快速成长期以及 2022 年以来的机遇期，而推动行业发展的核心是国内金融衍生工具的丰富和量化管理人策略多样化两个重要维度。

从最初的传统中低频多因子选股、ETF 套利、分级套利、期限套利等为主，再到当前的人工智能、另类数据、基本面量化等新技术，私募量化基金的策略逐步进入精细化和高频时代。根据 wind 对私募量化基金的分类，与股票市场相关的策略类型主要包括

量化选股、量化指数增强（例如 500 增强、1000 增强等）、量化中性策略等。伴随着国内量化投资领域的快速发展和不断涌出的新策略，私募量化产品的规模的持续扩张，且在产品的收益表现上超越市场平均水平，提升了市场对量化私募产品的认可和关注。

不同类型的量化私募基金策略的表现及其相对基准的超额表现如下图 2、图 3 所示，其中，中证 500 和 1000 指增产品的超额收益为同策略市场所有产品收益率的平均值减去对应指数的收益率，量化主动策略的超额收益为同策略市场所有产品收益率的平均值减去万得全 A 指数的收益率；而值得注意的是，量化中性策略由于已经做空对应的股指期货，因此直接计算其收益率，不再额外减去基准收益率。

图 2. 量化私募不同策略的净值曲线

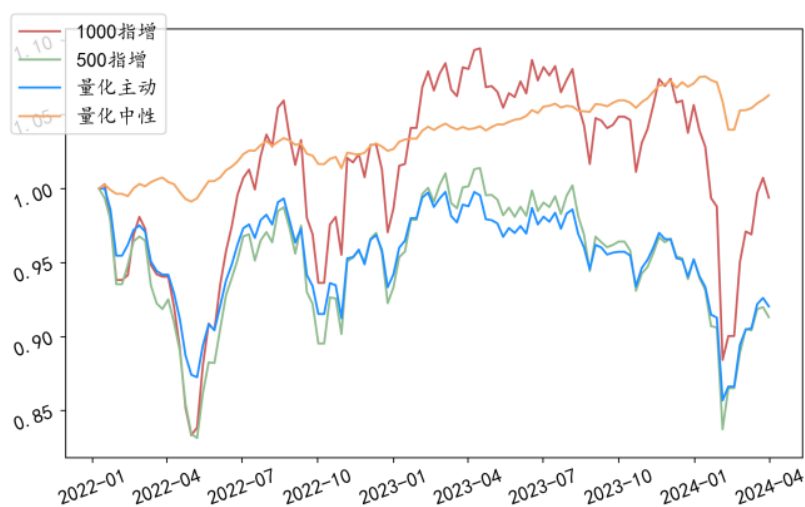
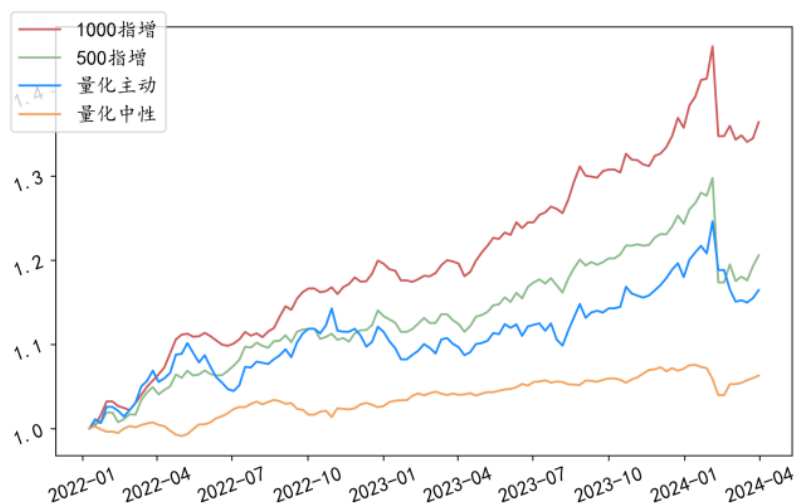


图 3. 量化私募不同策略的超额净值曲线



量化投资行业的投资方式和策略不断演变，这背后反映的是获取市场超额收益的难度在逐步增加。为了应对这一挑战，投资管理人必须利用更先进的技术和更完备的数据来构建有效的投资策略。从上图 3 中 2022 年到 2023 年的市场表现来看，不同量

化私募基金的策略整体表现良好，相对于基准指数的超额收益呈现出上升趋势；然而，进入 2024 年，许多量化私募策略的超额净值曲线遭遇了较大幅度的回撤，这可能提示了市场环境的变化或者策略适应性的挑战。

近年来，随着信息技术的突飞猛进，机器学习、深度学习等人工智能技术快速发展，越来越多的量化私募基金利用大数据时代丰富的股票特征信息，通过机器学习模型进行训练和预测，以期实现超额收益。然而，使用机器学习模型构建量化策略也存在一些挑战。一方面，机器学习模型通常依赖于历史数据进行训练，这可能导致模型在选择股票时趋向于偏好历史上表现较好的股票，从而可能加强这些股票未来表现的预测。这种基于历史数据的选股策略可能会引起市场上不同私募基金产品选股的趋同现象，即多个基金可能会选择相似的股票组合。同时，基于历史数据的机器学习模型也需要关注过拟合问题，确保模型不仅在历史数据上表现良好，也需要在未来的数据上保持有效。另一方面，市场条件的变化可能会影响模型的预测准确性，而机器学习模型的预测结果在不断变化的市场环境下是否能保持稳定也是值得关注的问题。

## 第三章 数据与样本构造

### 3.1 样本的构造

本篇论文研究中使用的股票数据来自 Wind 数据库，数据样本包括 2000 年 1 月-2024 年 3 月的股票月频数据。对于每条样本而言，它包含交易日期、股票代码、十二个特征值 ( $CR_1, CR_2, \dots, CR_{12}$ )、目标值，其中十二个特征值  $CR_k$  ( $1 \leq k \leq 12$ ) 表示过去  $k$  个月期间（从  $t-12$  月到  $t-12+k-1$  月）的股票累计收益率（如下图 4 所示）。这种表示法的动机是模拟投资者在第  $t-1$  月末观察一年期价格图的场景，投资者在做出投资决策时，通常会回顾过去一段时间内股票的表现，以寻找潜在的趋势和模式，因此我们通过分析过去 12 个月的累计收益率构建出反映股票历史表现的特征，

图 4. 特征  $CR_k$  的构造

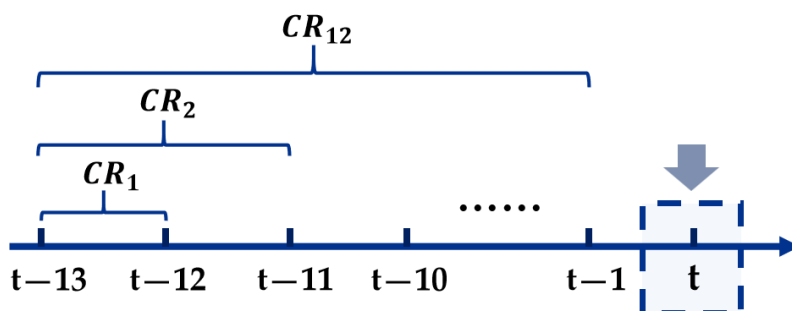


图 5. 全数据样本内容

性质	内容
时间区间	2000年1月-2024年3月
特征	$CR_1, CR_2, \dots, CR_{12}$
目标值	股票月收益率

对于每个月  $t$ ，该月股票池包含在  $t-1$  月最后一个交易日上市流通的所有股票。在构建样本特征时，我们特别关注了特征数据的完整性，由于构建样本特征时需要使用过去十二个月的累计收益率，为了确保所有样本所包含的特征数目相同，我们要求纳入样本池的每只股票在  $t-12$  到  $t-1$  的每个月都有非缺失的收益率。最后，由于在构建投资组合时是基于市值加权，我们需要样本中每只股票截至  $t-1$  月月底的流通股市值，我们将流通市值定义为月底流通股数量乘以每股的价格。

为了确保我们的神经网络模型不仅在样本内表现良好，而且具备样本外预测的能力，我们将全样本时间区间划分为两个不重合的部分：“模型选择期”和“主模型预测期”。



### 1. 模型选择期（2000 年 1 月–2015 年 12 月）

在这一阶段，我们的目标是使用了 2000 年 1 月至 2015 年 12 月的数据样本，训练多个备选的神经网络模型，并通过秩相关系数等指标来评估它们的性能。在模型选择期，我们尝试了不同的网络架构、损失函数和目标值等，以确定最终将用于预测的模型。这一过程允许我们评估不同模型的表现，并选择出在历史数据上表现最佳的模型作为我们的主模型。

### 2. 主模型预测期（2016 年 1 月–2024 年 3 月）

一旦确定了主模型，我们进入了主模型预测期。在这一阶段，我们采用了滚动时间窗口的方法来进行模型的训练和预测。与 Murray 等人（2024）使用的扩展时间窗口方法不同，我们的滚动时间窗口会随着时间的推移而动态更新。这意味着，从 2016 年 1 月开始，我们的模型训练集将向后滚动添加新的数据，从而确保模型始终是基于最新的信息进行训练的。其中，数据滚动的频率为每年进行一次滚动，例如首次使用 2000 年 1 月至 2015 年 12 月的数据进行训练，第二次使用 2001 年 1 月至 2016 年 12 月的数据进行训练等。

滚动时间窗口方法的优势在于通过不断纳入最新的数据，模型可以捕捉到最新的市场趋势和模式，适应市场结构和投资者行为随时间的变化，从而提高预测的时效性；且由于我国股票市场发展起步较晚，对于较为早期的股票数据而言，数据噪音较多，与当前的股票市场也有着较大的差别，因次选择滚动窗口在加入最新股票数据的同时，也剔除掉较为前期的数据，而不是 Murray 等人（2024）在美国市场使用的扩展窗口仅仅加入最新的数据。

## 3.2 样本的统计性描述

如下表 1 所示，为本篇论文用于神经网络模型训练及预测的特征值和原始目标值（即股票月收益率）的统计性描述。我们共有 689736 条样本数据，每条样本数据的包含 12 个特征（ $CR_1, CR_2, \dots, CR_{12}$ ）、目标值（month return）、股票代码和交易日期，对其中的十二个特征和目标值进行如下的统计性描述。

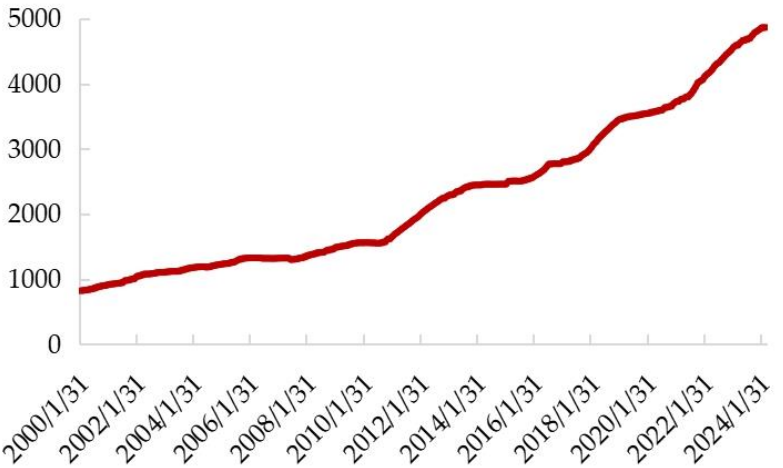
从表 1 我们观察到，对于十二个特征而言，随着  $k$  值的增大，即当我们考虑更长时间段的累计收益率时，累计收益率的平均值也呈现出逐步增大的趋势。这可能反映了市场在长期内的整体增长趋势，或者可能是由于某些股票在较长的时间窗口内表现出了持续的增长；而对于特征值及目标值整体而言，我们注意到数据的最大值明显高于最小值的绝对值。这种非对称性表明数据分布可能具有偏斜性，其中一些股票可能在某些月份表现出异常高的收益率，而下降的幅度则相对较小。

	$CR_1$	$CR_2$	$CR_3$	$CR_4$	$CR_5$	$CR_6$	$CR_7$	$CR_8$	$CR_9$	$CR_{10}$	$CR_{11}$	$CR_{12}$	ret
count	689736	689736	689736	689736	689736	689736	689736	689736	689736	689736	689736	689736	689736
mean	1.12%	2.28%	3.41%	4.54%	5.73%	6.86%	8.06%	9.34%	10.72%	12.07%	13.29%	14.65%	0.98%
std	14.05%	20.47%	25.48%	29.91%	34.16%	37.86%	41.77%	45.88%	50.22%	54.29%	58.37%	62.98%	13.43%
min	-0.60	-0.74	-0.83	-0.86	-0.89	-0.89	-0.92	-0.92	-0.91	-0.93	-0.92	-0.96	-0.59
25%	-7.01%	-9.95%	-11.97%	-13.57%	-14.92%	-16.29%	-17.40%	-18.37%	-19.50%	-20.58%	-21.76%	-22.72%	-6.75%
50%	0.00%	-0.14%	-0.52%	-0.57%	-0.50%	-0.62%	-0.55%	-0.46%	-0.48%	-0.61%	-0.77%	-0.92%	0.00%
75%	7.66%	11.22%	13.66%	16.07%	18.15%	20.25%	22.27%	24.38%	26.58%	28.42%	30.02%	31.71%	7.33%
max	6.79	6.92	6.20	5.71	4.21	4.50	5.20	7.86	7.61	8.07	8.98	11.17	2.79

表 1. 全样本统计性描述

由于本篇论文利用股票月度频率的数据进行模型的训练、预测以及组合的构造，我们统计了每个月样本数目的变化（如下图 6 所示）；随着时间的推移，每个月的样本数目呈现逐渐增加的趋势，这一现象不仅反映了市场扩容和数据积累的自然过程，而且为模型训练提供了日益丰富的数据资源。每个月样本数目的增加意味着模型可以从更多的历史信息中学习，从而可能提高预测的准确性和投资组合的构建效率。此外，新数据的增加也使我们能够更细致地捕捉市场的最新动态和个股表现，为投资决策提供更加坚实、准确的数据支持。

图 6. 月度样本数目变化



## 第四章 神经网络模型训练与预测

在本篇论文中，我们选择采用神经网络模型来检验弱市场有效性，基于以下几个关键原因：首先，神经网络模型因其强大的数据处理能力和模式识别特性，已经吸引了学术界和业界的广泛关注。这些模型能够发现简单特征之间的非线性或相互作用的复杂关系，并且从大量数据中学习到低层次的特征，这使得它们在金融市场分析中具有巨大的潜力，且与本篇论文希望通过简单的累计收益率特征构造有利可图的投资组合的目标一致；其次，在国内市场有效性的检验上，目前尚未有使用神经网络模型进行研究和检验的文章，因此我们参考 Murray 等人（2024）在美国股票市场上使用机器学习模型研究市场有效性的方法，通过这篇论文为国内股票市场有效性的研究提供新的视角和方法。

在本章节接下来的内容中，我们将借鉴 Murray 等人（2024）的方法，分为模型选择期和主模型预测期两部分进行介绍。在模型选择期，我们从模型架构、损失函数、损失函数中的样本权重和目标值四个维度构造了 80 个不同的神经网络模型，并且详细介绍了引入的五种模型架构的具体组成和作用，以及如何进行模型的训练和选择获得最终的主模型；主模型预测阶段，我们将使用在模型选择期确定的主模型进行进一步的滚动训练和预测，得到本篇论文最终的预测指标——神经网络预测收益率，记作 NNER（Neural Network Estimate Return），这个指标将作为我们评估市场有效性的关键工具，它反映了模型对于未来股票收益的预测能力。

### 4.1 模型选择期

在模型选择期，我们将从模型构建、模型训练和模型选择三个方面来展开研究。模型构建部分引入和介绍本篇论文将使用的多种神经网络架构；模型训练部分将对模型训练过程中损失函数、损失函数对每条样本的权重以及目标值提供多种选择，同时介绍模型训练所使用的数据集划分方法；最后，在模型选择方面，我们将根据一定规则选择出本篇论文后续分析所使用的主模型。

#### 4.1.1 模型构建

在该部分，基于 Murray 等人（2024）的采用的机器学习模型，我们也同样引入卷积神经网络（CNN）、长短期记忆网络（LSTM）和 CNN-LSTM。

Chung 等人（2014）通过实验发现，在音乐和语音信号建模的任务上，GRU（Gated Recurrent Unit）与 LSTM（Long Short-Term Memory）的性能不相上下。然而，GRU 在中央处理单元时间、参数更新速度以及模型泛化能力方面，相比 LSTM 更有优势。Yamak

(2019) 使用比特币的价格数据集探究和比较 GRU 和 LSTM 在时间序列预测问题上的表现效果, 结果发现门控循环单元 (GRU) 表现优于长短期记忆 (LSTM); Dey 等人 (2021) 对比了简单 RNN、LSTM 和 GRU 在不同股票走势和不同价格区间以及不同时间频率下的进行股价时序预测的效果, 结果表明, 与 LSTM 相比, GRU 产生的误差更小, 可靠性更高; Lin 等人 (2022) 根据价格历史预测股票价格的未来趋势, 均得出 GRU 预测效果更好的结论; Sako 等人 (2022) 发现 GRU 在货币汇率和股票市场指数的预测方面具有比 LSTM 和普通 RNN 更好的效果。因此, 除了 CNN、LSTM、CNN-LSTM, 我们还将引入门控循环单元 (GRU) 神经网络模型和 CNN-GRU 共同进行测试。

主要包括以下五种网络结构: 卷积神经网络 (Convolutional Neural Network, CNN)、长短期记忆网络 (Long-Short Term Memory, LSTM)、门控循环单元 (Gated Recurrent Unit, GRU)、CNN-LSTM、CNN-GRU。接下来, 我们将对每种网络结构进行详细介绍。

### ① 卷积神经网络 (Convolutional Neural Network, CNN)

卷积神经网络常用来处理具有网格状拓扑的数据, 例如我们的累积收益数据可以解释为一维网格。Kalchbrenner 等人 (2016) 利用一维卷积网络和解、编码器实现了英语到德语的翻译任务, 且发现其表现超越了基于循环神经网络的翻译模型; Bai 等人 (2018) 发现在引入扩展卷积和残差连接等架构元素后, 简单的卷积神经网络在不同的序列建模任务中比 LSTM 等循环神经网络更有效; Hoseinzade 和 Haratizadeh (2019) 也表明 CNN 在时间序列预测方面非常成功。

CNN 最关键的两个特征为局部感知 (Local Connectivity) 和参数共享 (Parameter Sharing)。局部感知是指网络中的隐藏单元仅与输入单元的一小部分相连, 而不是使用全连接。例如, 在图像处理中, 每个隐藏单元可能只与图像的一个局部区域相连, 进行局部的特征抽取; 参数共享指的是卷积层中每个神经元连接到数据窗口的权重是统一的, 每个神经元只专注于学习一种特征。这种权重共享机制不仅提高了特征抽取的效率, 还减少了模型训练的参数数量, 提升了模型的泛化能力。此外, CNN 在计算过程中采用的卷积操作, 相比传统的矩阵运算, 也显著提高了运算速度。

本篇论文的 CNN 网络结构主要包含如下图 7 所示的六层网络, 其中卷积层通过其卷积核的移动提取输入信息中的特征; 池化层对卷积层中提取的特征进行挑选和降维, 也能在一定程度上防止过拟合; 拉平层将多维的输入变成一维化, 作为全连接层前的过渡。

图 7. 卷积神经网络结构



## ② 长短期记忆网络 (Long-Short Term Memory, LSTM)

LSTM 是循环神经网络 (Recurrent Neural Network, RNN) 的一种, 循环神经网络是用于专门处理序列数据设计的神经网络类型。与普通的前馈神经网络不同, 循环神经网络的特点是其网络单元不仅接收输入和产生输出, 还存在一个内部的循环连接, 允许网络单元捕获并记忆来自之前时间点的信息; 这种结构使得循环神经网络十分适合处理涉及时序数据的时间序列任务, 例如机器翻译等。然而, 普通的循环神经网络存在对时序上的长期记忆存在不足的问题, 具体原因是其激活函数的输出值通常位于  $[0, 1]$  范围内, 信息在网络中的传递会随着时间增长而逐渐衰减。这种衰减效应导致普通神经网络模型在处理短期记忆时表现良好, 但不擅长进行长期记忆, 难以保持对较早时间点信息的记忆。Goodfellow 等人 (2016) 研究也指出, 传统的 RNN 在处理超过 5 个时间步长的关系时效果不佳。

为了应对循环神经网络在处理和训练长序列数据时遇到的梯度消失和梯度爆炸问题, Schmidhuber 及其学生 (1997) 引入了“记忆单元状态” (memory cell state) 的概念, 并据此发展出了长短期记忆网络, 即 LSTM。LSTM 中的记忆单元状态设计使其能够稳定地存储历史信息, 从而赋予了网络更强的记忆力, 使其能够更有效地处理长序列数据。LSTM 网络通过其内部的三个关键门控来管理信息的保留和更新, 以此来实现长期记忆和选择性遗忘。具体来说: “遗忘门”基于当前的输入数据  $x_t$  和前一时间步的隐藏状态  $h_{t-1}$  来决定哪些旧信息不再重要, 从而可以被网络遗忘; “输入门”在遗忘门处理之后起作用, 它决定在给定当前输入  $x_t$  和前一隐藏状态  $h_{t-1}$  下, 哪些新信息应该被存储进记忆单元状态  $c_t$  中; 最后, “输出门”根据当前的记忆单元状态  $c_t$  和输入  $x_t$ , 以及前一时间步的隐藏状态  $h_{t-1}$ , 确定在当前时间步应该输出的隐藏状态  $h_t$ 。LSTM 通过这些门控机制, 能够有效地处理长期依赖问题, 并避免梯度消失问题。

相比 RNN 只有一个传递状态  $h_t$ , LSTM 神经元在时间维度上向后传递了两个状态: “记忆单元状态”  $c_t$  和 “隐藏状态”  $h_t$ , 其中, 隐藏状态是记忆单元状态经过一个神经元和一道“输出门”后得到的。隐藏状态里包含的记忆, 实际上是记忆单元状态衰减之后的内容。除此之外, 记忆单元状态在一个衰减较少的通道里沿时间轴传递, 对时间跨

度较大的信息的保持能力比隐藏状态要强。因此，实际上隐藏状态 $h_t$ 里存储的主要是近期短时的记忆，而记忆单元状态 $c_t$ 里存储的主要是长时期的远记忆；记忆单元状态的存在，使得 LSTM 得以对长依赖进行很好地刻画（LSTM 网络具体结构见图 8）。

然而，由于 LSTM 相比 RNN 引入了很多新的内容，导致模型参数变多，也使得训练难度加大了很多。

图 8. LSTM 网络结构

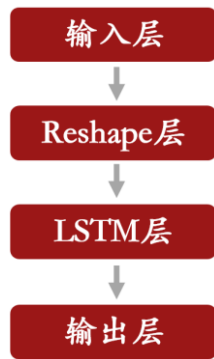
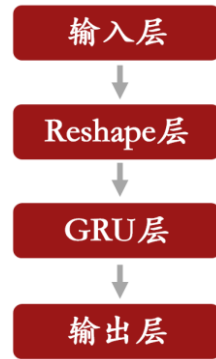


图 9. GRU 网络结构



### ③ 门控循环单元（Gated Recurrent Unit, GRU）

门控循环单元（GRU）由 Cho 等人于 2014 年引入，与 LSTM 相同，是为了解决普通 RNN 模型中长期记忆和反向传播中的梯度消失、爆炸等问题而提出来的；但相比 LSTM，GRU 内部少了一个“门控”，可以被视为 LSTM 的一个变种。

不同于 LSM 的三个门，GRU 仅使用了两个门来调控长期信息在神经元之间的流动和遗忘：首先，LSTM 中的“更新门”决定了从先前时刻的隐藏状态 $h_{t-1}$ 到当前时刻隐藏状态 $h_t$ 的信息保留量。更新门的值越大，允许从先前的记忆中流入的信息就越多，这与 LSTM 的记忆单元的作用紧密相关，可以看作是将输入门和遗忘门的功能结合，帮助网络捕捉时间序列中的长期依赖性；其次，LSTM 的“重置门”管理着前一时刻隐藏状态 $h_{t-1}$ 对当前候选隐藏状态 $\tilde{h}_t$ 的影响程度，即决定新输入信息与旧记忆如何融合。重置门的值越小，意味着更多的旧信息被舍弃，因此重置门有助于网络捕捉时间序列中的短期依赖性。

从以上的门控机制可以看出，与 LSTM 不同，GRU 不使用独立的线性自更新记忆单元来存储信息，而是在隐藏单元内部通过门控机制直接进行更新，仅保留了当前时刻的候选隐藏状态 $\tilde{h}_t$ 和隐藏状态 $h_t$ 。这种设计简化了网络结构，同时仍然保持了对时间序列数据中长期和短期依赖性的捕捉能力。

综上所述，相比于 LSTM，GRU 网络在门控机制上做了简化，这可能会使得它的记忆力相比 LSTM 稍弱。然而，简化的结构也带来了训练效率上的优势：GRU 模型由于参数较少，训练速度通常更快，且更易于收敛，同时仍能实现与 LSTM 相近的处理性能。鉴于 GRU 在训练效率和计算成本方面的优势，本篇论文决定采用 GRU 网络，以期在保证



模型性能的同时，也能考虑到实际的硬件资源和时间成本（具体结构见图 9）。

图 10. CNN-LSTM 网络结构



图 11. CNN-GRU 网络结构



#### ④ CNN-LSTM

第四种模型我们选择 CNN-LSTM 结构来结合 CNN 和 LSTM 共有的特点，CNN 擅长提取图像数据的局部特征，而 LSTM 擅长捕捉序列数据中的长期依赖关系，通过结合两者的优势，我们可以让模型同时考虑到数据的时序信息和空间信息，减少参数降低过拟合风险，从而提供更精确的预测、更出色的性能以及更高的训练效率。

在网络结构中的具体操作为，将 CNN 结构中的拉平层替换为 LSTM 层（具体结构见图 10）：首先，将特征输入到卷积神经网络的卷积层，卷积核将自适应地提取图像中的特征；其次，经过卷积层处理的特征会传递到最大池化层，这一步骤通过池化操作减少数据的维度，同时保留关键的特征信息；第三，这些降维后的特征被用作长短期记忆网络（LSTM）层的输入，LSTM 层负责训练神经网络，自动学习并识别序列数据中的特征；最终，通过 Adam 优化算法，将训练过程中产生的误差以反向传播的方式传递，逐层更新模型的参数，得到最终的模型。

#### ⑤ CNN-GRU

第五种神经网络模型与第四种类似，我们选择 CNN-GRU 结构来结合 CNN 和 GRU 共有的特点，并将 CNN 结构中的拉平层替换为 GRU 层（具体结构见图 11）。

在模型的参数选择上，我们主要依据 Murray 等人（2024）的设置，基于 Adam 优化器训练所有模型。具体来说，我们将学习率设置为 0.001，最大 epoch 数为 100；为了提高训练效率和防止过拟合，我们采用了早停机制，即如果连续 5 个训练周期的验证集损失值没有得到改善，则停止训练过程；同时，保留 30% 的验证集样本观测值来评估

模型的样本外性能。对于神经网络模型中每层结构的具体参数设置，Reshape 层将输入数据转换为三维数据，包含 batch\_size、时间戳和数据样本的维度，其中，我们将时间戳的数量设置为 12，维度设置为 1；卷积层有 64 个隐藏神经元，卷积核的大小为 5，步长设置为 1，并使用 ReLU 激活函数；Max-Pooling 层中的池大小设置为 2；LSTM 层有 64 个神经元，dropout 为 0.2，并使用 ReLU 函数进行激活；最终的输出层使用单神经元和线性激活函数。

#### 4.1.2 模型训练

机器学习根据其学习目标主要分为两大类：监督学习和无监督学习。监督学习的核心在于，它使用带有标签的数据集进行训练。在这种学习模式下，每个训练样本不仅包含一组特征（即输入变量），还包含一个与之对应的标签（即输出变量）。通过分析特征和标签之间的关系，监督学习算法能够学习到一个映射规则或函数  $f$ ，这个函数能够将输入特征映射到输出标签；当给定新的数据时，这个函数可以用来预测其对应的输出标签。相对地，无监督学习则处理的是没有标签的数据集。在无监督学习中，算法的目的是探索数据的内在结构和模式。这可能包括聚类分析（将数据分成不同的组）、异常检测（识别数据集中的异常值）、或降维（减少数据的复杂性同时保留关键信息）等任务。在本篇论文中，我们的目标是构建一个能够预测股票未来一个月收益的模型。为此，我们使用了过去十二个月的累计收益率作为输入特征，并利用下月的实际收益率作为标签，以此来训练我们的神经网络模型。由于我们的训练过程中每条样本都配有一个目标值，即下月的收益率，因此，我们的模型训练过程属于监督学习。

对于神经网络模型而言，其训练过程旨在通过学习输入特征和目标值之间的关系来构建一个能够准确预测的函数  $f$ ，主要包含以下几个步骤：首先是数据的输入与前向传播。前向传播指的是，将训练集数据输入到多层包含参数（权重）的神经网络模型后，数据在多层网络模型中逐层进行传递和计算，例如加权求和或非线性激活函数转换等，从而提取特征和目标值之间的关系。

第二步，基于模型输出的预测值和目标值计算损失函数值。前向传播结束后，模型将输出一个预测值，我们使用预先定义的损失函数来衡量这个预测值与实际目标值之间的差异，常见的损失函数包括均方误差、交叉熵等。损失函数的值越小，表示模型的预测越接近目标值；

第三步，基于损失函数值进行梯度计算与反向传播。反向传播指的是将损失函数对每个权重的梯度从输出层开始，反向传递回输入层，逐层计算每个权重对最终损失的贡献，与数据输入后的前向传播计算预测值形成对比。

第四步，基于反向传播的梯度值进行权重更新。在反向传播获得了每个权重参数的梯度后，模型将根据梯度下降等方法来调整权重值。权重更新的目标是减小损失函数



的值，从而提高模型的预测准确性；

接下来，神经网络模型将重复进行以上四个步骤进行迭代优化，每次迭代都包括前向传播、损失计算、反向传播和权重更新。通过多次迭代，模型逐渐学习到数据中的复杂关系，权重值也在不断更新中逼近最优解。为了确保模型在样本外的有效性，我们通常也会设置验证集来定期评估模型的性能，根据模型在验证集上的表现，选择最佳的模型参数或提前停止训练以防止过拟合。

因此，由以上的过程我们可以看出，在神经网络模型的训练过程中，模型架构的组成、损失函数的选择、目标值的设定和模型评估的方法等都至关重要。模型架构从根本上影响输入数据逐层输出的结果，损失函数决定了模型的优化目标和权重更新的方向，目标值的设定直接影响到模型训练结果的准确性，而模型的评估方法也会关系到模型在样本外的与曾能力等。因此，在接下来的内容中，我们将根据 Murray 等人（2024）的研究来选择和设置这些关键因素，以获得预测效果最好的神经网络模型。

### ① 损失函数的选择

在本文中，我们主要采用以下两种损失函数。首先，我们选择了均方误差（Mean Square Error, MSE）作为神经网络模型的损失函数。MSE 计算模型预测值  $f(x)$  与实际目标值  $y$  之间差值的平方和的平均值，其优点在于其函数曲线是光滑、连续的，并且在整个定义域内可导，这使得使用梯度下降算法进行优化变得简单有效。此外，MSE 的梯度随着误差的减小而减小，有助于模型更快地收敛到最优解。由于这些特性，MSE 成为了回归问题中最常用的损失函数之一。

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2$$

其次，由于本文的目标是基于模型预测值进行选股构造投资组合，我们也考虑使用信息系数（Information Coefficient, IC）作为损失函数，它描述了预测和实际股票收益之间的相关性，取值范围在-1 到 1。信息系数在实际投资领域中是一个衡量机构选股表现的重要指标，它通常被用来反映所挑选股票的某个因子值与其未来收益率之间的关联度，进而通过这个系数的高低评估该因子对于预测股票未来收益的有效性。

### ② 损失函数中样本权重的选择

由于本篇论文的目标是预测未来股票收益的横截面，如果在损失函数中给每个观察值赋予相同的权重，那么拥有更多股票的月份就会拥有更多的总权重。为了解决这个问题，除了等权重（EW）方法外，我们还考虑了两种不同的加权方法。第一个选择是每个月等权重，并且在每个月内，对每只股票赋予相等的权重（Equal Weight Per Month, EWPM）；第二种选择是每个月的权重相等，但根据上个月末的市值对给定月份内的每个

股票进行权重 (Equal Weight Per Month Value Weighted, EWPMVW)，这也与我们之后在构建投资组合时使用流通市值加权相一致。

值得注意的是，只有当损失函数为 MSE 时，我们选择对损失函数中的每条样本进行以上三种不同的加权选择 (EW、EWPM、EWPMVW)；而当损失函数为 IC 时，由于 IC 的计算方式为所有样本的预测值和目标值之间的相关系数，对每条进行加权没有实际的含义，因此只有所有样本等权这一种权重选择。

### ③ 目标值的选择

由于个股收益率可能存在很强的协变，为了避免机器学习模型的训练过程被这种时序的因素所影响，我们可以考虑原始收益率  $r$  的三种变形作为模型训练的目标值。首先考虑的是标准化的收益率  $r_{std}$ ，即用收益率减去平均收益率，然后用这个差除以股票截面收益率的标准差，这样可以使得每个月目标值的均值和标准差都为 0 和 1。

其次，由于  $r$  和  $r_{std}$  会使得机器学习模型在训练的过程中受极端值的影响较大，加上我们的目标是找出比其他股票表现更好（或更差）的股票从而构造投资组合，而不是预测收益差异的大小。因此，我们考虑使用  $r_{pctl} = \frac{rank_{i,t}}{N_t+1}$ ，其中， $N_t$  是第  $t$  个月的股票数量， $rank_{i,t}$  是股票  $i$  的收益率在第  $t$  月所有股票中的排名，其中收益最低的股票  $rank_{i,t} = 1$ ，收益最高的股票  $rank_{i,t} = N_t$ 。因此， $r_{pctl}$  每个月都是 0-1 之间的均匀分布。

最后，本篇论文还考虑使用  $r_{norm} = \theta^{-1}(r_{pctl}) = \theta^{-1}\left(\frac{rank_{i,t}}{N_t+1}\right)$  作为目标值的替换，其中  $\theta^{-1}[\cdot]$  表示标准正态累积分布函数的反函数。这种表示方法可以使得目标值不受极端值影响的同时，仍在每个月满足标准正态分布。

因此，综上所述，我们考虑了 80 个不同的机器学习模型，其中，模型分为三种模型架构 (CNN、LSTM、GRU、CNN-LSTM、CNN-GRU)，两种损失函数 (MSE、IC)，三种加权方案 (EW、EWPM、EWPMVW，只对 MSE 进行) 和四种因变量 ( $r$ 、 $r_{std}$ 、 $r_{pctl}$ 、 $r_{norm}$ ) 的所有组合。

### 4.1.3 模型选择

为了选择出对我们的样本数据预测效果最好的深度学习模型，且排除样本信息的泄露，我们仅使用 2000 年 1 月-2015 年 12 月期间的数据子集（而不是全部的数据集）进行模型选择期的数据集。具体来说，我们将模型选择其的数据子集划分为拟合月份和非拟合月份，拟合月份包括偶数年的偶数月 and 奇数年的奇数月，非拟合月份包括偶数年的奇数月 and 奇数年的偶数月，这样的划分是为了避免训练和测试集有所重合。接下

来，我们将每个模型分别在拟合月份上进行训练，在非拟合月份上进行预测，进一步基于模型的预测值来进行最终主模型的选择。

性质	内容	
时间区间	2000年1月-2015年12月	拟合月份：奇数年的奇数月、偶数年的偶数月 非拟合月份：奇数年的偶数月、偶数年的奇数月
特征	$CR_1, CR_2, \dots, CR_{12}$	
目标值	股票月收益率的不同形式	

表 2. 模型选择期的样本数据内容

由于对于神经网络模型而言，其训练过程是随机的，即对相同的数据进行两次重复的训练，会得到两个不同的预测函数。因此，参考 Murray 等人（2024）的方法，我们对每个模型在拟合月份上训练 30 次，来消除部分模型训练的不确定性，也方便进行模型的复现和检验。进一步，我们将在拟合月份上训练所得到的 30 个预测函数，应用于非拟合月份进行预测，这样可以确保模型在样本外的预测能力。最后，我们取这 30 个预测的平均值作为基于给定神经网络模型在模型选择期对应月份上的预测值。

基于神经网络模型对每个月的预测值，我们进一步计算模型预测值和实际收益率之间每月横截面的 spearman 秩相关性系数，再计算相关系数的时序平均值来评估模型的预测能力。其中，我们使用 spearman 秩相关性而不是 pearson 相关性，是因为我们后续在进行投资组合的构造和分析时，是依赖于股票在截面上的排序进行选择，因此我们更看重股票之间的相对收益，而不是获取绝对收益的能力；此外，由于模型进行训练时的目标值可能是原始股票收益率的变形（例如  $r_{std}$ ,  $r_{pctl}$ ,  $r_{norm}$ ），因此 pearson 相关性中的线性假设可能不成立。

如下表 3 所示，展示了 80 个模型经过上述过程后所得到的 spearman 秩相关系数。其中，秩相关系数最高的是网络结构为 CNN-GRU、损失函数为 IC（没有样本加权）以及因变量为  $r_{norm}$  的神经网络模型，我们也将该模型作为主模型，用于进行后续的模型训练和预测。

因变量	权重	CNN		LSTM		CNN-LSTM		GRU		CNN-GRU	
		MSE	IC	MSE	IC	MSE	IC	MSE	IC	MSE	IC
$r$	EW	1.31%	1.70%	-4.49%	-4.91%	-0.93%	-0.14%	-5.13%	-4.45%	-0.67%	-0.10%
	EWPM	-1.30%		-4.36%		-1.54%		-3.30%		-2.36%	
	EWPMVW	-2.59%		1.09%		-1.62%		-1.01%		-2.50%	
$r_{std}$	EW	7.30%	7.75%	6.51%	8.51%	7.64%	8.14%	6.74%	7.58%	7.84%	7.94%
	EWPM	6.62%		7.69%		5.33%		8.96%		6.90%	
	EWPMVW	4.76%		-2.80%		2.03%		-1.55%		1.23%	
$r_{PctI}$	EW	6.53%	8.31%	-0.19%	8.65%	2.58%	8.71%	7.59%	8.21%	3.99%	9.29%
	EWPM	4.77%		7.14%		0.74%		4.73%		3.03%	
	EWPMVW	-2.47%		1.63%		-4.79%		3.25%		-4.06%	
$r_{Norm}$	EW	8.55%	8.63%	7.55%	8.03%	8.57%	9.08%	7.09%	7.65%	8.65%	9.36%
	EWPM	7.51%		8.04%		7.20%		8.37%		7.79%	
	EWPMVW	3.08%		-2.42%		1.85%		-1.76%		1.04%	

表 3. 80 个模型预测值与收益率秩相关系数

4.2 主模型预测期

由 4.1 可知，本文最终进行预测和选股所使用的模型结构、特征、损失函数和目标值等如下表 3 所示。我们通过将其应用于模型预测期样本时间区间上的扩展窗口进行模型拟合训练，并使用在这些扩展窗口上拟合生成的函数来获取窗口后续时期的预测，我们将模型所得到的预测值命名为神经网络预测收益（Neural Network Estimate Return, NNER）。

性质	内容
时间区间	2016年1月-2024年3月
特征	$CR_1, CR_2, ..., CR_{12}$
目标值	$r_{norm}$
模型结构	CNN-GRU
损失函数	信息系数（IC）

表 4. 主模型预测期的样本数据内容与模型结构

进一步，本文定义 $NNER_{i,t}^{t_1,t_2}$ ，它表示使用 $t_1 \sim t_2$ 区间的样本数据训练神经网络主模型，基于模型对股票  $i$  在第  $t$  个月的目标值进行预测所得到的结果。举例来说， $NNER_{00001.SZ, 201601}^{200001, 201512}$  表示，使用 200001-201512 的数据训练模型，将模型拟合生成的函

数应用于股票“000001.SZ”在2016年1月的样本数据上，生成对该股票在该月的预测值。

本篇论文在模型训练过程中共采用了9个滚动窗口，包含200001-201512、200101-201612、200201-201712、200301-201812、200401-201912、200501-202012、200601-202112、200701-202212、200801-202312，对应的模型预测窗口为201601-201612、201701-201712、201801-201812、201901-201912、202001-202012、202101-202112、202201-202112、202301-202312、202401-202403。因此，我们可以获得如下的预测值：

图 12. NNER 定义

$$NNER_{i,t}^{Y_1,Y_2} = \begin{cases} NNER_{i,t}^{2000,2015}, & \text{if } 201601 \leq t \leq 201612 \\ NNER_{i,t}^{2001,2016}, & \text{if } 201701 \leq t \leq 201712 \\ NNER_{i,t}^{2002,2017}, & \text{if } 201801 \leq t \leq 201812 \\ NNER_{i,t}^{2003,2018}, & \text{if } 201901 \leq t \leq 201912 \\ NNER_{i,t}^{2004,2019}, & \text{if } 202001 \leq t \leq 202012 \\ NNER_{i,t}^{2005,2020}, & \text{if } 202101 \leq t \leq 202112 \\ NNER_{i,t}^{2006,2021}, & \text{if } 202201 \leq t \leq 202212 \\ NNER_{i,t}^{2007,2022}, & \text{if } 202301 \leq t \leq 202312 \\ NNER_{i,t}^{2008,2023}, & \text{if } 202401 \leq t \leq 202403 \end{cases}$$

如下表5给出了NNER在2016年1月-2024年3月模型预测期间全样本和分年度的统计性描述，计算了各横截面统计指标的时间序列平均值。在全样本区间上，平均每个月有3639个预测值，NNER的截面统计值在时序上的均值和中位数分别为0.000和0.001，都十分接近于0；截面平均标准差为0.6%；在分年度样本区间上，不难看出各统计指标和全样本区间上的较为一致，这也说明了模型预测值NNER的截面分布随时间推移也较为稳定。

	均值	标准差	最小值	1%	5%	25%	中位数	75%	95%	99%	最大值	数量
<b>201601-202403</b>	0.000	0.006	-0.060	-0.020	-0.010	-0.002	0.001	0.004	0.009	0.014	0.033	3639
<b>2016</b>	-0.001	0.006	-0.042	-0.018	-0.011	-0.005	-0.001	0.003	0.009	0.014	0.028	2708
<b>2017</b>	0.004	0.007	-0.054	-0.015	-0.006	0.001	0.005	0.008	0.013	0.018	0.035	2865
<b>2018</b>	0.002	0.006	-0.049	-0.015	-0.007	-0.002	0.001	0.005	0.011	0.018	0.037	3240
<b>2019</b>	0.001	0.006	-0.049	-0.017	-0.008	-0.002	0.001	0.004	0.009	0.014	0.032	3506
<b>2020</b>	-0.001	0.006	-0.066	-0.024	-0.012	-0.003	0.000	0.002	0.006	0.010	0.032	3612
<b>2021</b>	-0.003	0.007	-0.077	-0.031	-0.016	-0.005	-0.002	0.001	0.005	0.011	0.030	3868
<b>2022</b>	-0.004	0.007	-0.071	-0.028	-0.016	-0.007	-0.003	0.000	0.005	0.010	0.033	4314
<b>2023</b>	0.005	0.006	-0.068	-0.017	-0.005	0.003	0.006	0.009	0.012	0.015	0.039	4690
<b>202401-202403</b>	0.003	0.005	-0.060	-0.013	-0.005	0.000	0.003	0.006	0.011	0.014	0.025	4867

表 5. NNER 的描述性统计



第五章 基于模型预测进行截面选股的表现分析

在这一章节，我们将研究基于神经网络模型所得到的收益预测值 NNER 与未来股票真实收益横截面之间的关系，基于 NNER 构造投资组合，对投资组合在不同阶段和条件下的表现进行分析，以此来对有效市场假说进行检验。

5.1 基于 NNER 构造五分位投资组合

基对于每个月  $t$ ，我们根据  $t-1$  月月底预测所得到的 NNER 进行升序排列，将所有股票分类为五分位组合，每个投资组合第  $t$  个月收益计算为该投资组合中所有股票的加权平均收益率，权重为  $t-1$  月末的流通市值占投资组合总市值的比例，即流通市值加权。

表 5 展示了基于 NNER 排序所获得的各投资组合的表现，其中 5-1 投资组合（后文称之为 NNER 5-1 组合）表示做多第五组组头组合，做空第一组投资组合，收益计算方式为第五组的收益率减去第一组的收益率。虽然在 A 股市场不存在做空机制，但是我们在进行因子表现分析时，仍将多空组合看作重要的分析指标，一方面是体现了因子区分股票未来收益好坏的能力，另一方面也能与 Murray 等人（2024）在美国市场的相关研究相一致。从表 6 和图 9 我们可以看到，根据模型预测值构造的五分位组合在 2016 年 1 月 1 日-2024 年 3 月 31 日期间的年化收益率和夏普比率都是单调递增的，意味着模型预测值 NNER 具有不错的选股分组效果，能够有效的区分股票在未来的收益能力强弱；最终，NNER 5-1 组合的年化收益率、年化夏普比率分别达到 14.02%和 74.17%。

	年化收益率	年化波动率	最大回撤	夏普比率	卡玛比率
Group_1	-12.78%	24.28%	-72.36%	-52.63%	-17.66%
Group_2	-7.64%	23.01%	-53.14%	-33.19%	-14.37%
Group_3	-4.00%	21.71%	-53.05%	-18.42%	-7.54%
Group_4	0.69%	21.04%	-33.37%	3.27%	2.06%
Group_5	2.21%	20.06%	-42.25%	11.04%	5.24%
Group_5-1	14.02%	18.91%	-37.43%	74.17%	37.47%

表 6. NNER 投资组合表现

下图 13 展示了基于模型预测构造的 NNER 5-1 投资组合在 2016 年 1 月 1 日-2024 年 3 月 31 日期间的月累计收益率，我们可以看到组合累计收益率整体呈现稳定向上的趋势，也体现了模型预测值 NNER 的截面选股能力。

图 13. NNER 五分位组合和 NNER 5-1 组合年化收益率

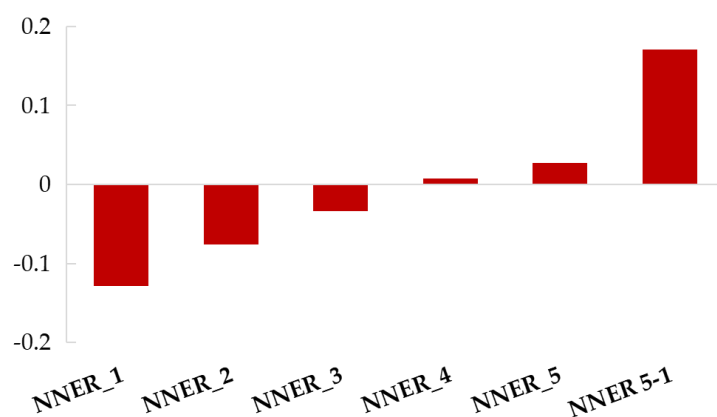
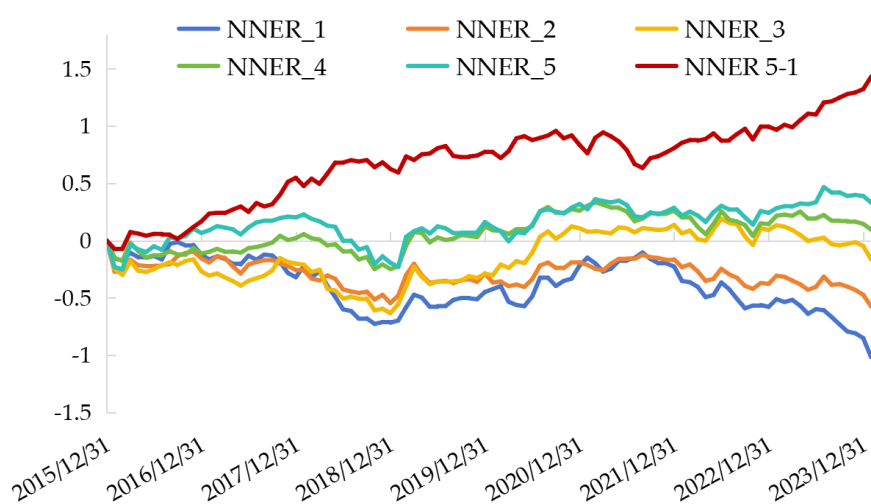


图 14. NNER 五分位组合和 NNER 5-1 组合月累计收益率



## 5.2 子周期内的预测能力

接下来,我们将研究基于深度学习模型的预测的预测能力是否只在某些子周期有效。从下表可以看到,除了 2021 年,NNER 5-1 组合在不同的子区间都能获得稳定的正收益,表明模型预测能力随着时间的推移是稳定的。



	年化收益率	年化波动率	最大回撤	夏普比率	卡玛比率
2016年	15.55%	20.06%	-15.32%	0.78	1.01
2017年	44.77%	15.82%	-7.25%	2.83	6.18
2018年	6.39%	19.27%	-12.70%	0.33	0.50
2019年	11.08%	17.87%	-16.21%	0.62	0.68
2020年	4.66%	17.07%	-14.44%	0.27	0.32
2021年	-6.76%	24.34%	-29.65%	-0.28	-0.23
2022年	17.49%	16.25%	-10.10%	1.08	1.73
2023年	35.48%	16.00%	-9.10%	2.22	3.90
20240101- 20240331	77.42%	15.07%	-4.63%	5.14	16.71

表 7. NNER 5-1 投资组合子区间表现

### 5.3 风险调整后 NNER 投资组合的表现

改进后的有效市场假说表示，如果收益是对风险的补偿，那么技术分析也是有利可图的。因此，我们需要对探究系统性风险因子的暴露是否可以解释 NNER 排序组合的收益。

1964 年，夏普等学者提出了资本资产定价模型（Capital Asset Pricing Model, CAPM）。这一模型确立了资产收益与其承担风险之间的联系，对现代金融市场的定价理论产生了深远的影响。尽管 CAPM 模型在理论上具有重要意义，但它在实际应用中遇到了一些挑战，例如它难以解释实际市场上股票收益率与流通市值之间存在的负相关关系等。鉴于这些局限性，资产定价的研究开始从单因子模型逐渐向多因子模型转变，以期增强模型的解释力和对未来收益的预测准确性，提供更为全面的资产定价视角。

Fama 和 French 在 1993 年首次提出三因子模型，在 CAPM 的基础上加入了规模（Small-Minus-Big, SMB）和价值（High-Minus-Low, HML）两个因子作为对市场风险因素的补充，认为小市值（Small）和高账面市值比（High）的股票收益率高于大市值（Big）和低账面市值比（Low）的股票，使得三因子模型能够更完整的描述和解释股票所包含的投资风险和截面收益。Fama-French 三因子模型如下，其中， $E[R_i]$ 表示股票  $i$  的预期收益率， $R_f$ 表示无风险收益率， $E[R_M]$ 为市场组合预期收益率， $E[R_{SMB}]$ 和 $E[R_{HML}]$ 分别为规模因子（SMB）以及价值因子（HML）的预期收益率， $\beta_{i,MKT}$ 、 $\beta_{i,SMB}$ 和 $\beta_{i,HML}$ 为个股  $i$  在相应因子上的暴露。

$$E[R_i] - R_f = \beta_{i,MKT}(E[R_M] - R_f) + \beta_{i,SMB}E[R_{SMB}] + \beta_{i,HML}E[R_{HML}]$$

随着 Fama-French 三因子模型在市场中的应用，研究者们发现该模型在解释市场

现象方面仍存在局限，例如未能涵盖市场中的动量效应。Jegadeesh 和 Titman (1993) 中提出了股票动量效应的概念，他们发现，历史上表现良好的股票在未来一段时间内往往能继续保持较高的收益率，而那些历史收益率较低的股票则可能继续表现不佳。因此，Carhart (1997) 在 Fama-French 三因子（市场因子、规模因子及价值因子）的基础上加入了动量因子（Up-Minus-Down, UMD）。Carhart 四因子模型如下：

$$E[R_i] - R_f = \beta_{i,MKT}(E[R_M] - R_f) + \beta_{i,SMB}E[R_{SMB}] + \beta_{i,HML}E[R_{HML}] + \beta_{i,UMD}E[R_{UMD}]$$

为了帮助投资者更准确地预测资产的未来收益，提供更精细化的风险管理工具，Fama 和 French (2015) 对原有的三因子模型进行了扩展，提出了包含五个因子的资产定价模型，新增了两个因子：首先是盈利能力因子（Robust-Minus-Weak, RMW），这个因子通过比较高盈利能力股票组合与低盈利能力股票组合的收益率差异来定义，盈利能力的高低是基于公司毛利润与账面价值之比来衡量；第二个是投资因子（Conservative-Minus-Aggressive, CMA），该因子衡量的是高投资水平股票组合与低投资水平股票组合之间的收益率差异，投资水平则是通过企业的年资产增长率与前一年总资产的比值确定。五因子模型不仅提高了对资产定价的解释力度，还提供了深入理解市场运作机制的新视角。例如，它揭示了高盈利能力且低投资的公司通常会有更高的预期回报率，反映了市场对这类公司的特别偏好或者是对其潜在风险的补偿等。Fama-French 五因子模型如下：

$$E[R_i] - R_f = \beta_{i,MKT}(E[R_M] - R_f) + \beta_{i,SMB}E[R_{SMB}] + \beta_{i,HML}E[R_{HML}] + \beta_{i,RMV}E[R_{RMV}] + \beta_{i,CMA}E[R_{CMA}]$$

Hou、Xue 和 Zhang (2015) 基于实体投资的经济学理论，提出了包含四个因子的资产定价模型，通常被称为 HXZ 四因子模型。这个模型也被称作 q 因子模型，其理论基础来源于经济学家托宾提出的 q 理论。在 q 理论中，q 比率被定义为公司实物资产价值与其市场重置成本的比值。根据 q 理论，当 q 值超过 1 时，企业倾向于增加对新资产的投资以追求经济利益；而当 q 值低于 1 时，意味着企业的实物资产被市场低估，企业可能会减少投资。q 理论从公司金融的视角解释了投资与盈利率对股票预期收益率的影响：投资水平较高的公司往往有较低的折现率和预期收益率，而盈利率较高的公司则有较高的折现率和预期收益率。在托宾的 q 理论基础上，HXZ 四因子模型进一步纳入了投资因子和盈利因子，形成了其独特的资产定价框架。与主要从个人投资者角度分析最优投资组合的传统模型不同，HXZ 四因子模型认为投资和盈利率是股票预期收益率的关键决定因素。该模型不仅涵盖了投资因子（INV）和盈利因子（ROE），还综合了市场因子和市值因子（ME）；其中，投资因子的计算方法为公司总资产的年度变化与上一年总资产的比值，而盈利因子则采用公司的净资产收益率（ROE）作为衡量标准。

与 Fama-French 五因子模型和 Carhart 四因子模型不同，HXZ 四因子模型并未包

含价值因子和动量因子。在价值因子方面，根据 Tobin 的  $q$  理论， $q$  比率较高的公司倾向于进行更大规模的投资，这使得这些公司更可能被视为成长型公司，而投资规模较小的公司则被视为价值型公司。因此，HXZ 模型中的投资因子已经涵盖了价值因子的经济含义。对于动量因子，当公司公布的财务报告超出市场预期时，其股价往往会上涨，这通常反映了公司盈利能力的增强；相反，当财务报告未达预期时，股价下跌，反映了盈利能力的减弱。因此，会计收益率与股票收益率之间存在类似的影响关系，HXZ 模型认为盈利因子可以替代动量因子的作用。Hou-Xue-Zhang 四因子模型如下：

$$E[R_i] - R_f = \beta_{i,MKT}(E[R_M] - R_f) + \beta_{i,ME}E[R_{ME}] + \beta_{i,INV}E[R_{INV}] + \beta_{i,ROE}E[R_{ROE}]$$

接下来，为了方便观察，我们将 NNER 5-1 组合的日频收益率进行年化并且乘以一百去掉百分号后，对经典的 Carhart 四因子模型、Fama-French 五因子模型和 Hou-Xue-Zhang 四因子模型进行时序回归，非零的回归截距系数表示因子模型捕获的系统风险因子不能完全解释 NNER 5-1 组合的收益。

### 5.3.1 Carhart 四因子模型回归

从下表 9 可以看出，将 NNER 5-1 组合收益率对 Carhart 四因子模型回归后的截距系数（即日频的  $\alpha$ ）为 0.06%，在 95% 的置信度上是显著的，调整后的  $R^2$  仅为 11.23%，说明 Carhart 四因子模型不能完全解释 NNER 投资组合带来的收益。值得注意的是，动量因子（UMD）的回归系数也较为显著，主要原因可能是模型的输入为过去十二个月的累计收益率，一定程度上反映的就是个股的动量或反转效应，因此对动量因子的回归系数较为显著是合理的。

	alpha	MKT	SMB	HML	UMD	R-adj
Group_1(low)	-12.5728**	1.0352***	0.1827***	-0.2318***	0.1205***	0.7108
Group_2	-6.992	1.0576***	0.1709***	-0.0841***	-0.1281***	0.7367
Group_3	-3.2792	1.0269***	0.1451***	-0.0323	-0.1273***	0.7503
Group_4	0.9729	1.0191***	0.0421*	-0.0086	-0.0689***	0.7501
Group_5(high)	1.7682	0.9947***	0.0379*	0.0797***	-0.1271***	0.7469
Group_5-1	14.3409*	-0.0404	-0.1448***	0.3115***	-0.2475***	0.1145

表 8. NNER 组合-Carhart 四因子模型回归结果

### 5.3.2 Fama-French 五因子模型回归

首先，我们将 NNER 5-1 投资组合的收益在 2016 年 1 月-2024 年 3 月整个测试区间内对 FF 五因子进行回归。可以从下图看到，回归后的截距系数为 0.05%，且在 95% 的置信度上显著，调整后的  $R^2$  仅为 10.33%，说明 Fama-French 五因子模型不能完全解释

NNER 投资组合带来的收益。值得注意的是，在该因子模型回归中，账面市值比因子的回归系数十分显著。

	alpha	MKT	SMB	HML	RMW	CMA	R-adj
Group_1(low)	-12.1435*	1.0224***	0.0663*	-0.2254***	-0.2687***	-0.1354	0.7095
Group_2	-6.7966	1.0506***	0.1673***	-0.0868*	-0.0391	0.0935	0.7338
Group_3	-3.2673	1.0194***	0.1553***	-0.0023	-0.0429	0.0157	0.7465
Group_4	0.5849	1.0282***	0.1395***	0.0045	0.1701***	0.0719	0.7514
Group_5(high)	1.2869	1.0052***	0.1512***	0.0970**	0.2139***	0.1000	0.7451
Group_5-1	13.4304*	-0.0172	0.0849	0.3224***	0.4825***	0.2354*	0.1053

表 9. NNER 组合-FF 五因子模型回归结果

### 5.3.3 Hou-Xue-Zhang 四因子模型回归

Hou-Xue-Zhang 四因子模型回归结果如表 8 所示：我们可以看出，将 NNER 5-1 组合收益率对 HXZ 四因子模型回归后的截距系数（即日频的 alpha）为 0.07%，在 99% 的置信度上是显著的，调整后的 R<sup>2</sup> 仅为 6.13%，说明 HXZ 四因子模型不能完全解释 NNER 投资组合带来的收益。值得注意的是，在 HXZ 四因子模型的回归中，各因子的回归系数都较为显著，而其中规模因子（ME）的回归系数为-0.3305，表明组合收益与市值大小成反比，也为论文后续基于神经网络模型选股分析量化私募基金的业绩表现提供了思路。

	alpha	MKT	ME	INV	ROE	R-adj
Group_1(low)	-13.9449**	1.0712***	0.2706***	-0.2870***	0.0272	0.6980
Group_2	-5.7454	1.0461***	0.1261***	-0.1349**	-0.2493***	0.7352
Group_3	-1.2373	1.0064***	0.0554*	-0.0888*	-0.3217***	0.7528
Group_4	2.0597	1.0013***	0.0190	-0.0796	-0.1448***	0.7514
Group_5(high)	3.8705	0.9617***	-0.0611**	0.0220	-0.2423***	0.7434
Group_5-1	17.8154**	-0.1094***	-0.3316***	0.309***	-0.2695***	0.0635

表 10. NNER 组合-HXZ 四因子模型回归结果

## 5.4 技术指标中性化后 NNER 投资组合的表现

由于在弱有效市场中，所有的历史信息，包括交易量、股价、K 线图等，都被认为会迅速地反映在当前的股票价格中，基于这些历史信息的技术分析，理论上不应该为投资者提供额外的信息优势或超额收益。因此，为了验证本文所研究的神经网络模型的预测能力没有被技术信号所包含，在本部分的内容中，我们将分析 NNER 在对技术信

号进行控制后的选股能力和收益表现。

在市场实践中，技术指标被交易者和分析师广泛采用（Schwager, 1989; Lo and Hasanhodzic, 2010），这些指标主要基于历史的价格和交易量数据来预测未来可能的价格走向。多项研究已经检验了基于不同技术指标构建的交易策略的盈利能力，包括过滤规则（Fama 和 Blume, 1966）、移动平均线（Brock 等人, 1992; Zhu 和 Zhou, 2009）、动量（Conrad 和 Kaul, 1998; Ahn 等人, 2003）和自动模式识别（Lo 等人, 2000）等。Neely 等人（2014）研究了技术指标直接预测股票风险溢价的能力，并将其与宏观经济变量的表现进行了比较。本文也将使用 Neely 等人（2014）提出的 14 个技术信号对 NNER 的选股能力进行研究，每一个技术信号都是一个在每期设定为 1 或 0 的指标，表示在当期看多或看空该只股票。14 个技术信号分为以下三大类：

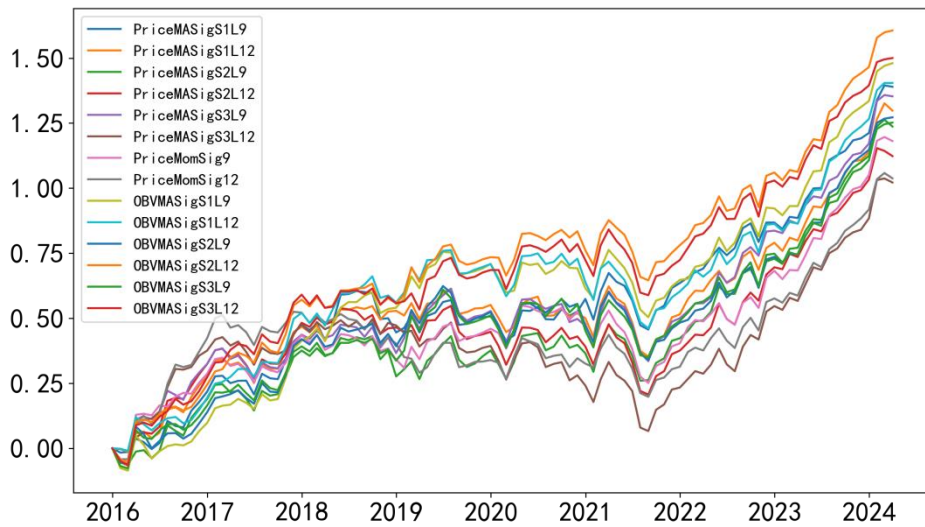
- (1) **价格移动平均信号 (Price Moving Average Signal)**：如果过去  $S$  个月的移动平均价格大于等于过去  $L$  个月的移动平均价格，则将价格移动平均信号设置为 1，否则将信号设置为零，其中  $S$  在  $\{1, 2, 3\}$  中取值， $L$  在  $\{9, 12\}$  中取值。命名方式为：当  $L=1, S=9$  时，记作 PriceMASigS1L9。
- (2) **价格动量信号 (Price Momentum Signal)**：如果当前价格大于或等于  $M$  个月前的价格，则将价格动量信号设置为 1，否则将信号设置为 0，其中  $M$  在  $\{9, 12\}$  中取值。命名方式为：当  $M=9$  时，记作 PriceMASig9。
- (3) **能量潮移动平均信号 (On-Balance Volume Moving Average Signal)**：能量潮 (On-Balance Volume, OBV) 是一种通过累加每日成交量变化来衡量市场情绪的技术指标。它的计算方法如下：今天的 OBV 等于昨天的 OBV 加上或减去今天的成交量。如果今天的收盘价高于昨天的收盘价，则今天的成交量记为正数；如果今天的收盘价低于昨天的，则记为负数；如果今天的价格没有变化，则成交量计为零。这种计算方法的基本思想是，将股价上涨时伴随的成交量看作是市场人气的积累，因此进行加法操作；而股价下跌时的成交量则被视为市场人气的流失，因此进行减法操作，以此来反映股票的市场人气。对于能量潮移动平均信号而言，若过去  $S$  个月的移动平均 OBV 大于或等于过去  $L$  个月的 OBV 移动平均值，则将能量潮移动平均信号设置为 1，否则将信号设置为零，其中  $S$  在  $\{1, 2, 3\}$  中取值， $L$  在  $\{9, 12\}$  中取值。命名方式为：当  $L=1, S=9$  时，记作 OBVMASigS1L9。

图 15. 14 个技术信号

- Moving Average Price:  $I_{MA(s) \geq MA(l)}, s \in \{1,2,3\}, l \in \{9,12\}$
- Momentum:  $I_{P_t \geq P_{t-m}}, m \in \{9,12\}$
- On-balance Volume (OBV):  $I_{MAOBV(s) \geq MAOBV(l)}, s \in \{1,2,3\}, l \in \{9,12\}$

接下来，我们将对每个技术信号进行中性化后，基于 NNER 构建投资组合，分析其选股能力和收益表现。技术信号中性化的具体做法为：将每一个技术信号值为 1 和 0 的股票分为两大组，分别在每组内对 NNER 进行排序后构建五分位投资组合。最后，计算两组的第五组收益率的和减去两组的第一组收益率的和，记为中性化后的 NNER 5-1 组合的收益率。下图为对技术信号中性化后 NNER 5-1 组合的累计收益率。

图 16. 技术信号中性化后 NNER 5-1 组合的月累计收益率



下表为对每一个技术信号中心化后 NNER 5-1 组合的表现，可以看到，每个组合都有较为显著的正收益和夏普比率；这说明，神经网络模型的预测能力没有被 Neely 等人（2014）所提出的 14 个技术信号所包含。

	年化收益率	年化波动率	最大回撤	夏普比率	卡玛比率
PriceMASigS1L9	16.48%	16.44%	-20.13%	100.26%	81.86%
PriceMASigS1L12	15.27%	16.14%	-25.78%	94.63%	59.23%
PriceMASigS2L9	14.32%	16.75%	-20.36%	85.53%	70.35%
PriceMASigS2L12	12.89%	16.24%	-31.58%	79.35%	40.81%
PriceMASigS3L9	15.76%	17.63%	-27.37%	89.44%	57.59%
PriceMASigS3L12	11.38%	17.11%	-38.04%	66.50%	29.92%
PriceMomSig9	13.60%	16.71%	-27.33%	81.40%	49.78%
PriceMomSig12	11.77%	16.17%	-30.83%	72.81%	38.18%
OBVMASigS1L9	17.67%	16.84%	-24.96%	104.92%	70.78%
OBVMASigS1L12	16.59%	16.96%	-28.91%	97.80%	57.38%
OBVMASigS2L9	14.70%	17.37%	-27.78%	84.67%	52.94%
OBVMASigS2L12	19.47%	16.71%	-21.30%	116.51%	91.38%
OBVMASigS3L9	14.40%	17.53%	-26.16%	82.15%	55.03%
OBVMASigS3L12	17.95%	16.81%	-23.13%	106.81%	77.63%

表 11. 技术指标中性化后 NNER 5-1 组合的表现

## 第六章 基于模型预测结果分析私募量化基金表现

2024 年初，国内不同策略的量化私募基金都经历了大幅的回撤，由下表各策略在过去几年的表现对比可以看出，2024 年第一季度的年化超额收益均为负，与过去三年的连续正收益形成对比。指数增强、量化选股和量化中性产品都受到了不小冲击，也引发了众多市场参与者的思考与分析。

		2022年	2023年	20240101-20240331
中证500指增	年化收益	3.07%	2.11%	-4.01%
	最大回撤	-1.35%	-1.82%	-9.55%
	夏普比率	140.36%	103.44%	-39.36%
中证100指增	年化收益	4.62%	2.91%	-1.32%
	最大回撤	-1.37%	-1.80%	-7.80%
	夏普比率	216.16%	123.59%	-15.98%
量化选股	年化收益	2.83%	1.29%	-2.78%
	最大回撤	-5.16%	-2.93%	-7.73%
	夏普比率	83.45%	44.90%	-44.82%
量化中性	年化收益	0.44%	0.96%	-1.06%
	最大回撤	-1.97%	-0.52%	-3.35%
	夏普比率	34.85%	120.37%	-41.37%

表 12. 不同私募量化基金策略超额表现

由于普通投资者无法看到私募量化基金的完整持仓，因此在本章节，我们期望基于本文神经网络模型的选股和投资组合构造能够一定程度上反映出市场上私募量化基金的持仓情况。

下表为基于本文神经网络模型预测构造的五分位投资组合的第五组（称为 NNER-5 组合）的净值序列与量化私募基金各策略净值序列的相关性（算法同 2.4 图 2），可以看出，二者在收益率上有较高的相关性，其中量化中性策略的相关性高达 80.8%；净值序列较高的相关性也意味着二者持仓的整体表现有较高的相关性。因此，我们认为基于本文神经网络模型预测的选股，可以一定程度上代表市场上量化私募基金整体的持仓方向。

	500指增	1000指增	量化选股	量化中性
NNER-5组合	0.401	0.605	0.257	0.808

表 13. 量化私募各策略与 NNER 第五组合净值曲线的相关性



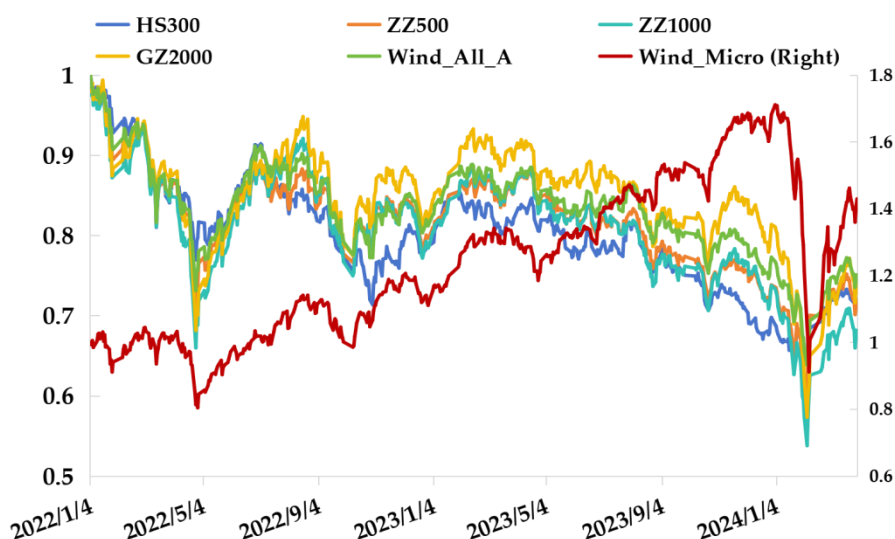
## 6.1 市场指数的表现

在本部分，我们主要基于国内股票市场的代表性指数，通过分析各指数的表现，挖掘私募量化基金 2024 年年初大幅回撤可能的原因。

图 17. 各市场指数净值曲线



图 18. 各市场指数净值曲线



上图为中国 A 股市场几只具有代表性的指数（沪深 300、中证 500、中证 1000、国证 2000、万得全 A 指数，图 17 展示各指数之间的关系）2021 年至今的表现，可以看到市场各股票指数也在 2024 年年初经历了大幅的回撤。股票市场在图中的过去几年时间里走过了牛熊市，各指数收盘价也经历了几次较为明显的涨跌。其中，三次较为大幅的下跌分别为 2022 年 1 月 1 日至 2022 年 5 月 1 日、2022 年 8 月 15 日至 2022 年 10 月 10 日、2024 年 1 月 1 日至 2024 年 2 月 7 日。如下表所示，对比三次大幅下跌中各指数的表现，我们发现在 2024 年年初的这次大幅下跌中，中证 1000 和国证 2000 指数的下跌幅度远大于沪深 300、中证 500 指数和万得全 A 指数；而在另外两次的大幅下跌中，三只宽基指数的下跌幅度较为接近。进一步引入万得微盘股指数（主要包含全部 A 股中市值居于后 400 的个股），可以看出，在前两次大幅下跌中，万得微盘股指数的表现与万得全 A 指数和其他宽基指数的表现较为接近，但是在 2024 年年初的这次下跌中，跌幅达到了 21.8%，远高于其他各指数。

		沪深300	中证500	中证1000	国证2000	万得微盘股指数	万得全A
20220101-20220501	年化收益	-3.11%	-4.08%	-5.03%	-4.76%	-2.60%	-3.79%
	最大回撤	-23.05%	-28.83%	-34.03%	-31.86%	-23.05%	-26.88%
	夏普比率	-57.25%	-66.44%	-70.81%	-66.68%	-33.00%	-66.16%
20220815-20221010	年化收益	-3.31%	-4.15%	-5.93%	-5.73%	-4.06%	-4.21%
	最大回撤	-9.77%	-12.00%	-16.68%	-16.73%	-12.26%	-12.18%
	夏普比率	-113.28%	-102.49%	-121.25%	-109.34%	-77.34%	-122.27%
20240101-20240207	年化收益	-1.02%	-5.31%	-10.60%	-12.78%	-21.80%	-6.32%
	最大回撤	-6.10%	-17.55%	-26.67%	-29.29%	-41.31%	-16.85%
	夏普比率	-23.99%	-66.18%	-113.23%	-137.34%	-172.83%	-103.53%

表 14. 各市场指数三次下跌行情的表现

基于对市场各指数表现的分析，我们进一步计算了不同的量化私募策略表现与各指数表现的相关性。从下表可知，对于 500 指增、1000 指增和量化选股策略而言，随着指数所代表包含的股票市值减小（不包括万得微盘股指数），各策略与指数指数之间的相关性上升；而对于量化中性策略而言，其与万得微盘股指数的相关性极高。由于中证 1000 指数、国证 2000 和万得微盘股指数代表的是 A 股市场的中小盘股票，因此，我们将基于本文的神经网络模型选股进一步探究中小盘股票在 2024 年年初量化私募基金大幅下跌行情中发挥的作用。

	沪深300	中证500	中证1000	国证2000	万得微盘股	万得全A
500指增	0.319	0.554	0.654	0.796	0.357	0.620
1000指增	-0.112	0.148	0.299	0.505	0.680	0.234
量化选股	0.474	0.677	0.781	0.897	0.223	0.752
量化中性	-0.748	-0.619	-0.496	-0.284	0.939	-0.525

表 15. 私募量化各策略与各市场指数净值曲线的相关性

6.2 基于模型预测结果的分析

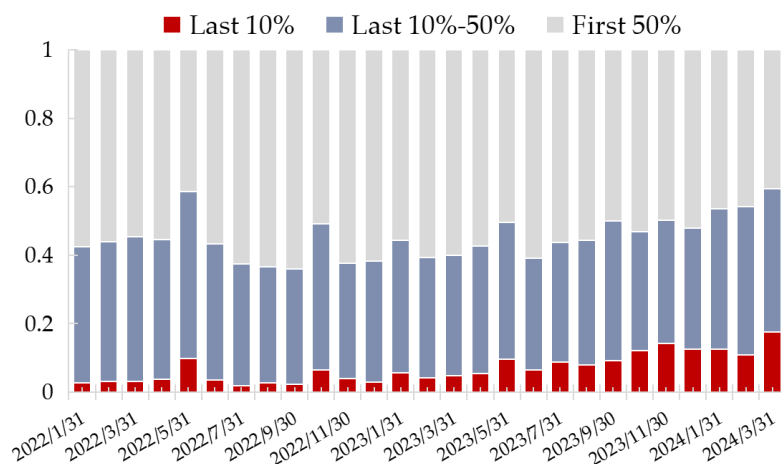
在这一部分，我们将基于第四章中依据神经网络模型预测构造的投资组合中的第五组（NNER-5 组合）从中小盘股票占比和收益率回归分析两个方面来分析私募量化基金净值下跌的原因。

6.2.1 模型选股中的中小盘股票数量占比

下图为 NNER-5 组合对于不同市值股票的持有比例，其中，我们选择观察每个月末市值大小位于最后 10%、后 10%-50%、前 50%的股票。可以看到，组合对中小盘股票的

持有比例（后 10%和后 50%）呈现上升趋势；且在 2023 年下半年持续上升，对于市值大小位于后 50%的股票持有比例逐渐超过 50%。

图 19. 第五组对于不同市值股票的持有比例



接下来，我们分析中小盘股占比上升的原因。首先，量化投资主要依托统计和数学模型来分析数据，寻找市场中存在的历史统计规律和模式，进一步基于这些历史规律和模式，构建投资策略和模型来减少主观判断和情绪对投资决策的影响，提高了投资的科学性和系统性。由下表 11 可以看到，近两年来，随着指数所包含的成分股市值减小，指数表现越好；尤其对于 Wind 微盘股指数而言，在过去两年的每个子区间内都有显著的正收益，2023 年上半年和下半年的年化收益都超过 40%。因此，对于量化模型选股而言，由于过去几年中小盘和微盘股上涨趋势较好，小市值风格长期占优，基于历史统计进行决策的量化因子和模型便会不约而同地向中小盘和微盘股集中，导致其在组合中的数量占比提升。

	2022上半年	2022下半年	2023上半年	2023下半年
沪深300	-19.80%	-24.93%	-3.47%	-22.70%
中证500	-25.67%	-17.08%	-0.53%	-19.60%
中证1000	-26.07%	-19.39%	3.12%	-21.10%
国证2000	-24.02%	-13.48%	6.49%	-15.32%
Wind微盘股	3.44%	21.41%	41.20%	47.21%
Wind全A	-20.69%	-18.75%	1.70%	-17.08%

表 16. 2022 和 2023 年子区间各市场指数表现

其次，由于以机器学习为代表的人工智能技术在量化领域得到越来越多的应用，导致不同量化私募基金产品之间的选股逐渐趋同到中小盘股票中，大量资金的涌入又

会进一步促进其整体的表现，使得过去表现较好的中小盘股票表现持续向好，从而吸引到更多的量化资金涌入中小盘和微盘股，导致量化管理人逐步进行小市值的侧重，中小盘股票在量化私募基金产品持有股票中的数量占比提升。

### 6.2.2 NNER-5 组合收益率序列的回归分析

进一步，下表为将 NNER-5 组合的收益率对 Fama-French 五因子模型进行回归后的结果。从回归结果可以看出，NNER-5 投资组合再规模因子（SMB）上的暴露显著为正，回归系数达到 6.078，且 p 值小于 0.001；这说明 NNER-5 组合在小市值股票上有较大的暴露。

	coef	std_err	t	P> t	0.025	0.975
<b>const</b>	0.0011	0.003	0.353	0.724	-0.005	0.007
<b>MKT</b>	1.0052	0.014	71.271	0.000	0.978	1.003
<b>SMB</b>	0.1512	0.025	6.078	0.000	0.102	0.200
<b>HML</b>	0.0970	0.032	3.047	0.002	0.035	0.159
<b>RMW</b>	0.2139	0.047	4.593	0.000	0.123	0.305
<b>CMA</b>	0.1000	0.020	1.617	0.106	-0.021	0.221
<b>Adj. R-squared</b>	0.7450					

表 17. NNER-5 组合收益率与 FF 五因子模型的回归

## 6.3 量化私募基金下跌的原因

基于以上对于市场指数表现和 NNER-5 组合的分析，我们得知量化私募基金对于中小盘和微盘股有较大的风格暴露，而中小盘股和微盘股在 2024 年年初由于雪球产品的敲入和基差走扩等因素迎来了大幅回撤，从而导致了量化私募基金的下跌。

### 6.3.1 雪球产品的敲入

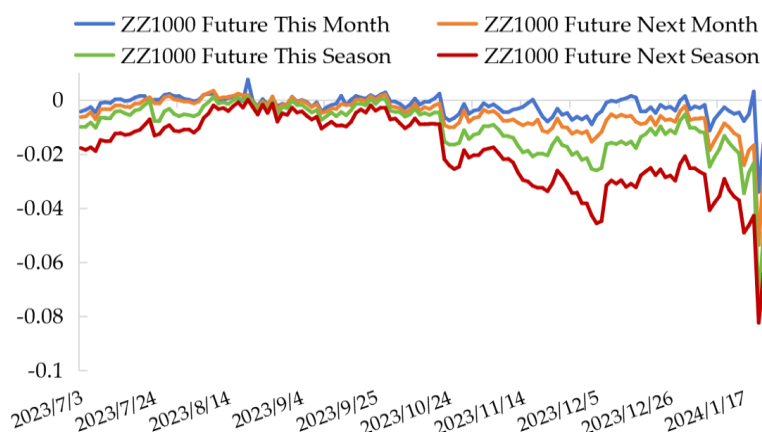
雪球产品是对带有雪球结构收益特征产品的统称，本质是一种奇异期权，设置有敲入敲出条件，最后的收益取决于挂钩标的资产的表现和敲入敲出事件是否发生。若标的资产价格上涨到一定程度（敲出价），雪球产品提前终止，投资者获得存续期间的固定收益；若标的资产价格下跌到一定程度（敲入价），则要根据到期日标的资产价格决定损益情况，投资者可能会承担标的资产下跌的风险；若标的资产从未触碰到敲入敲出价格，投资者获得整个产品期间的固定收益（如下表 18 所示）。

	敲出	未敲出
未敲入	产品提前结束 & 获得票息	产品持有至到期结束 & 获得票息
未敲入	产品提前结束 & 获得票息	产品持有至到期结束 & 获得标的跌幅

表 18. 雪球产品收益情况

对于投资者而言，买入雪球产品意味着向证券公司卖出了一个带触发条件（敲入和敲出）的看跌期权，投资者获得的雪球产品收益主要来自于证券公司支付的票息。而对于证券公司等卖出雪球产品的主体而言，为了支付投资者的票息收益，需要进行避险对冲交易，最常用的即为 Delta 对冲。由于作为看跌期权的买方，证券公司面临负的 Delta。因此，当雪球敲入后，标的价格下跌时，证券公司需要卖出股指期货来进行动态对冲，从而导致股指期货的价格下跌快于现货价格，进一步带来基差（期货价格 - 现货价格）走阔，贴水程度加深。如下表所示，我们使用基差率（= 期货价格 / 现货价格 - 1，基差率负值越大，表示期货贴水程度越深）来衡量中证 1000 股指期货的贴水程度，可以看到 2023 年下半年开始，中证 1000 股指期货的贴水程度逐渐加深。

图 20. 中证 1000 股指期货贴水程度（基差率）



### 6.3.2 量化中性策略的下跌

对于量化私募中性策略而言，其运作核心在于构建多空头寸，目的是获取与市场波动无关的相对收益，捕获所选股票组合相对于市场指数的超额表现，也就是 Alpha 收益。这种策略的操作方式是，一方面购买一组股票形成多头组合，另一方面持有空头工具，如股指期货，以此来抵消市场整体波动带来的风险，即 Beta 风险。股指期货对冲是当下最主要的手段；在量化中性策略的基础上，量化私募基金还衍生出了 DMA 产品，

该类产品本质上也是做多股票，同时做空股指期货，以量化中性策略的方式赚取 alpha 收益，但是其与普通量化中性策略产品的区别在于它能够利用杠杆使得策略的收益和回撤都进行翻倍的效果。中性策略通过使用股指期货等衍生品工具对冲了系统性风险，这种策略的表现通常与股市的整体走势关联不大，不会因为股市的大幅上涨而获得异常高的收益，同样不会在股市大跌时遭受重大损失，旨在实现与市场波动相对独立的稳健回报；而 2024 年年初，中性策略也未能幸免于难，主要原因也是其对中小盘和微盘股的风格暴露，这也与前文量化中性策略与万得微盘股指数和 NNER-5 组合的相关性较高的结论一致。接下来，我们将分析量化中性策略下跌的原因以及中小盘股票在其中起到的作用。

首先，量化中性策略由于对冲市场 Beta 的需求，拥有对于股指期货的空头头寸。由 6.3.1 的分析可知，由于雪球敲入导致股指期货价格下跌、贴水程度加深，导致产品的对冲成本上升。因此，为了减小对冲成本上升带来的损失，量化私募基金选择降低现货持有头寸来减少对冲的需求和高成本。

其次，由上图 20 可以看出，2024 年 1 月中旬，股指期货进入极端深贴水区域，由于当股指期货基差走阔到历史极值时，未来一定会收敛，这会给中性产品带来负贡献，加了杠杆的 DMA 产品便是更多倍的负贡献。因此，量化私募管理人为了避免巨大的损失，需要进行现货的减仓操作。

有上述两点我们可以知道，2024 年初，量化私募基金对持有的股票现货等进行减仓，减仓的体量对于大盘股来说可能不算很多，但是对于市值较小的中小盘和微盘股来说，由于其市值较小，会带来较大的抛压，稍有资金流动就可能导致其价格产生较大波动，从而容易出现大涨大跌的情况。而这也导致了小盘股在今年年初的严重下跌，从而引起了对中小盘股有风格暴露的私募量化基金各策略的净值大幅回撤。

## 第七章 结论及展望

基于从历史价格图中容易观察到的过去累计收益数据,我们通过基于神经网络模型生成的预测(NNER)是否可以预测未来股票的截面收益来测试弱有效市场假说的弱形式。我们首先使用2000年1月-2015年12月区间来确定用于生成基于神经网络的预测的最佳模型。我们发现,使用具有门控循环单元的卷积神经网络作为网络架构,信息系数作为损失函数,并将未来股票收益的归一化度量作为因变量,可以使得模型在2000年1月-2015年12月期间的表现最佳。我们使用样本早期的数据来确定最佳的神经网络模型,克服了对数据挖掘和结果的样本外有效性的担忧。

我们使用所选择的神经模型来生成2016年1月-2024年3月区间未来股票回报的样本外预测。投资组合分析表明,基于神经网络模型的预测能有力地预测未来股票地截面收益。因子分析(Hou-Xue-Zhang四因子、Carhart四因子和Fama-French五因子模型)和其他风险度量表明,与基于神经网络的预测相关的平均回报的变化不能完全被风险补偿所解释。因此,基于神经网络模型,我们的研究结果证明了中国A股市场还未达到弱有效状态。

进一步,基于神经网络模型预测构造的投资组合,本篇论文对2024年年初量化私募基金净值下跌的原因进行了分析:由于量化私募基金近年来对于中小盘股票的选择偏好,导致2024年年初中小盘和微盘股下跌时受到较大的冲击,从而带来了产品净值的大幅回撤。

最后,本篇论文做出了如下几点贡献:首先,我们证明了基于简单的技术指标和神经网络模型可以构造出有利可图的投资组合,从而为中国A股市场未达到弱有效状态提供了证据,丰富了对中国市场有效性研究的相关文献;其次,本文也从特征构造和模型两方面丰富了使用神经网络模型来解释和预测股票截面收益的相关文献;最后,本文基于模型预测对2024年初量化私募基金大幅回撤给出了定量和定性的分析。

## 参考文献

- Aktan, C., Iren, P., & Omay, T. (2019). Market development and market efficiency: evidence based on nonlinear panel unit root tests. *The European Journal of Finance*, 25(11), 979 – 993. <https://doi.org/10.1080/1351847X.2018.1560346>
- Carhart, M.M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, 52, 57-82.
- Cho, K., Merrienboer, B.V., Gülçehre, Ç., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder – Decoder for Statistical Machine Translation. *Conference on Empirical Methods in Natural Language Processing*.
- Chung, J., Gülçehre, Ç., Cho, K., & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *ArXiv*, abs/1412.3555.
- Coronel-Brizio, H.F., Hernández-Montoya, A.R., Huerta-Quintanilla, Mexico, M., Aplicada, D.D., Cinvestav-Merida, & México (2007). Assessing symmetry of financial returns series. *Physica A-statistical Mechanics and Its Applications*, 383, 5-9.
- Dey, P., Hossain, E., Hossain, M.I., Chowdhury, M.A., Alam, M.S., Hossain, M.S., & Andersson, K. (2021). Comparative Analysis of Recurrent Neural Networks in Stock Price Prediction for Different Frequency Domains. *Algorithms*, 14, 251.
- Ehsani.S, and innaimaa. (2022). Factor momentum and the momentum factor. *Journal of Finance*, 77(3).1877-1919.
- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1), 1-22.
- Fama, E.F. (1970). Efficient Capital Market: A Review of Theory and Empirical Work [J]. *Journal of Finance*, 27:383-417.
- Feng, G., He, J., Polson, N. G., and Xu, J. (2023). Deep learning in characteristics-sorted factor models. *Forthcoming in the Journal of Financial and Quantitative Analysis*.
- Freyberger, J., Neuhierl, A., and Weber, M. 2020. Dissecting characteristics nonparametrically. *Review of Financial Studies*, 33(5), 2326-2377.
- Goodfellow, I., Bengio, Y., and Courville, (2016). *A. Deep learning*. MIT press.
- Green, J., Hand, J. R., and Zhang, X.F. (2017). The characteristics that provide independent information about average us monthly stock returns. *Review of Financial Studies*, 30(12), 4389-4436.
- Harvey, C.R., Liu, Y., and Zhu, H. (2016). ... and the cross-section of expected returns. *Review of Financial Studies*, 29(1), 5-68.
- Hinton, G., E., Salakhutdinov, R., & R. (2006). Reducing the dimensionality of data with neural networks. *Science*.
- Hinton, Geoffrey E, Osindero, Simon, Teh, Yee-Whye. (2006). A fast learning algorithm for deep belief nets.[J]. *Neural computation*, 18(7).
- Hoseinzade, E. and Haratizadeh, S. (2019). Cnnpred: ONN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, 129(9).273-285.



- Hou, K., C. Xue and L. Zhang (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, Vol. 28(3), 650 – 705.
- Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*. 48(1), 65-91.
- Jiang, J., Kelly, B. T., and Xiu, D. (2022). (re-)imag(in)ing price trend.
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25(2).
- Lecun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., & Hubbard, W., et al. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4), 541-551.
- Lee, C.-C., Lee, J.-D. and Lee, C.-C. (2010). Stock prices and the efficient market hypothesis: Evidence from a panel stationary test with structural breaks, *Japan and the World Economy* 22(1), 49 – 58.
- Lewellen, J. and Nagel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics*. 82(2), 289-314.
- Lin, Z., Tian, F., & Zhang, W. (2022). Evaluation and Analysis of an LSTM and GRU Based Stock Investment Strategy.
- Lo, A.W., Mamaysky, H., and Wang, J. (2000). Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *Journal of Finance*, 55(4), 1705-1765.
- McLean, R. D. and Pontif, J. (2016). Does academic research destroy stock return predictability? *Journal of Finance*, 71(1), 5-32.
- Minsky, M., & Papert, S. (1969). *Perceptrons*. M.I.T. Press.
- Neely, C. J., Rapach, D.E., Tu, J., and Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management Science*, 60(7), 1772-1791.
- P.A. Samuelson. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6 (1965), pp 41-49.
- Peter T. Yamak, Li Yujian, and Pius K. Gadosey. (2020). A Comparison between ARIMA, LSTM, and GRU for Time Series Forecasting. In *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence (ACAI '19)*. Association for Computing Machinery, New York, NY, USA, 49 – 55. <https://doi.org/10.1145/3377713.3377722>
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65, 386-408.
- Rumelhart, D., Hinton, G. & Williams, R. (1986). Learning representations by back-propagating errors. *Nature* 323, 533 – 536. <https://doi.org/10.1038/323533a0>
- S. Hochreiter and J. Schmidhuber. (1997). "Long Short-Term Memory," in *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 15 Nov. doi: 10.1162/neco.1997.9.8.1735.
- Sako, K., Mpinda, B. N., & Rodrigues, P. C. (2022). Neural Networks for Financial Time Series Forecasting. *Entropy (Basel, Switzerland)*, 24(5), 657. <https://doi.org/10.3390/e24050657>
- Scott Murray, Yusen Xia, and Houping Xiao. (2024). Charting by Machines. *Journal of Financial Economics*, 153, ISSN 0304-405X.
- Sharpe, William F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk,

- Journal of Finance, 19 (3), 425-442
- Shiller, R.J. (1984). Stock Prices and Social Dynamics.
- Toth, B. and Kertesz, J. (2006). Increasing market efficiency: Evolution of cross-correlations of stock returns, Physica A-360(2), 505-515.
- Warren S.McCulloch, Walter Pitts.(1943). A logical calculus of the ideas immanent in nervous activity[J]. The bulletin of mathematical biophysics, 5(4):115-133.
- 常嘉佳, 马子红. (2013). 中国股票市场的弱有效性: 基于上证指数的分析[J]. 思想战线, 39(S2): 66-68.
- 邓子来, 胡健. (2001). 市场有效理论及我国股票市场有效性的实证检验[J]. 金融论坛, (10): 44-50.
- 罗超予, 杨启帆. (2022). 中国股票市场有效性研究[J]. 金融, 12(2): 195-201.  
<https://doi.org/10.12677/FIN.2022.122020>.
- 王若凡. (2009). 沪、港、澳股市有效性比较研究[D]: [硕士学位论文]. 南京: 南京理工大学.
- 吴世农.(1996). 我国证券市场效率的分析[J].经济研究, 04:13-19+48.
- 闫冀楠, 张维. (1998). 关于上海股市收益分布的实证研究[J].系统工程, 16(1): 21—25.
- 张兵, 李晓明. (2003). 中国股票市场的渐进有效性研究[J]. 经济研究, (1): 54-61.
- 张亦春, 刘猛. (2008). 开放经济下中央银行外汇市场干预的有效性分析[J]. 西部金融, (7): 20-23.
- 朱瑞.(2018). 我国沪深股市弱式有效性实证研究[J].时代金融,2018(20):198+200.
- Kalchbrenner, N., Espeholt, L., Simonyan, K., Oord, A.V., Graves, A., & Kavukcuoglu, K. (2016). Neural Machine Translation in Linear Time. ArXiv, abs/1610.10099.
- Bai, S., Kolter, J.Z., & Koltun, V. (2018). An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. ArXiv, abs/1803.01271.

## 北京大学学位论文原创性声明和使用授权说明

### 原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文作者签名：                    日期：        年    月    日

### 学位论文使用授权说明

（必须装订在提交学校图书馆的印刷本）

本人完全了解北京大学关于收集、保存、使用学位论文的规定，即：

- 按照学校要求提交学位论文的印刷本和电子版本；
- 学校有权保存学位论文的印刷本和电子版，并提供目录检索与阅览服务，在校园网上提供服务；
- 学校可以采用影印、缩印、数字化或其它复制手段保存论文；
- 因某种特殊原因需要延迟发布学位论文电子版，授权学校☐一年/☐两年/☐三年以后，在校园网上全文发布。

（保密论文在解密后遵守此规定）

论文作者签名：                    导师签名：

日期：        年    月    日