

分类号: F830.91
密 级: 公开

单位代码: 10363
学 号: 2221010102

题 目 基于深度学习的股价预测与量化投资研究

英文并列

题 目 RESEARCH ON STOCK PRICE PREDICTION
AND QUANTITATIVE INVESTMENT BASED
ON DEEP LEARNING

学生姓名: 方言

指导教师: 张伟 副教授

专 业: 金融

研究方向: 金融大数据

论文答辩日期: 2025 年 5 月 14 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。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。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：方言

日期：2025 年 5 月 14 日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 _____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：方言

日期：2025 年 5 月 14 日

指导教师签名：张伟

日期：2025 年 5 月 14 日

基于深度学习的股价预测与量化投资研究

摘 要

股票市场作为现代金融体系的核心枢纽，其价格受经济、政治、市场情绪等多重因素影响，呈现出高噪声、非平稳性和高随机性等复杂特征。这为市场预测与分析带来了巨大挑战，也制约了投资决策的科学性和收益的稳定性。在此背景下，探索更精准的预测模型成为提升投资收益的关键。随着人工智能技术的发展，机器学习和深度学习等算法被引入股市研究，为股价预测开辟了新路径。本文在总结现有方法的基础上，创新性地从天气效应和网络舆情两个视角切入，构建融合深度学习技术的股价预测模型，并通过模拟交易验证其有效性。具体研究内容如下：

基于天气效应视角，本文构建了一种耦合 Savitzky-Golay 滤波与深度学习的混合预测模型 Savgol-BiLSTM-Att 进行股价预测，并通过模拟交易验证模型提升投资收益的能力。该模型创新性地将天气数据与股票数据相结合，利用 Savitzky-Golay 滤波算法对原始数据进行噪声抑制和特征提取，在保留关键特征信息的同时降低数据噪声和非平稳性对预测精度的干扰。在特征处理方面，模型将提取的有效特征映射至高维空间，通过双向长短期记忆网络（BiLSTM）捕捉时序依赖关系中的潜在特征。为进一步优化模型性能，引入时间注意力机制对

时序信息进行自适应加权，减少信息冗余，提升预测精度。在模型应用层面，基于预测结果设计了 TopkDropout 量化投资策略。实验结果表明，Savgol-BiLSTM-Att 的预测精度显著优于传统模型，TopkDropout 策略的夏普比率较沪深 300 指数高出 0.940078，各项评价指标均表现优异，充分验证了该模型在量化投资领域的应用潜力。

基于网络舆情视角，本文提出了一种融合 FMD 特征模态分解、QR 分位数回归与深度学习的区间预测模型 FMD-QRBiGRU 用于股价的预测以及真实交易模拟的收益与风险分析。该模型通过网络爬虫和自然语言处理技术，将网络舆情信息量化并构建成投资者情绪因子和关注度因子，挖掘其潜在信息价值。随后，采用 FMD 特征模态分解对原始收盘价数据进行自适应分解，提取多尺度特征并分离噪声成分，提升模型预测精度。同时，利用双向门控循环单元（BiGRU）学习原始行情数据、情绪因子、注意力因子以及 FMD 所获特征的长期依赖关系，并通过注意力机制实现样本间信息交互，挖掘多维特征。在模型应用层面，基于预测结果构建了区间约束交易策略。实证结果表明，FMD-QRBiGRU 模型对股票数据的建模能力显著优于主流基准模型，实现了 23.5235% 的年化收益率和 1.236564 的夏普比率，分别较沪深 300 指数高出 19.95% 和 1.16。此外，其年化波动率（0.16283）低于沪深 300 指数（0.192105），进一步验证了模型在风险可控下具备稳定的收益能力。

本研究以深度学习技术为核心，结合信号处理技术，融合多源数据提升股价预测精度，并设计适当的交易策略以提高投资收益。研究成果一方面揭示了债券市场波动的内在机制，另一方面为投资主体的决策优化与监管机构的政策制定提供了实证依据，兼具重要的理论价值与现实意义。

关键词：股价预测，量化投资，深度学习，信号处理，天气效应，网络舆情

RESEARCH ON STOCK PRICE PREDICTION AND QUANTITATIVE INVESTMENT BASED ON DEEP LEARNING

ABSTRACT

The stock market, as the core hub of the modern financial system, is influenced by multiple factors such as the economy, politics, and market sentiment, exhibiting complex characteristics like high noise, non-stationarity, and high randomness. This poses significant challenges for market prediction and analysis, and restricts the scientific nature of investment decision-making and the stability of returns. Against this backdrop, exploring more precise prediction models has become key to enhancing investment returns. With the development of artificial intelligence technologies, algorithms such as machine learning and deep learning have been introduced into stock market research, opening new pathways for stock price prediction. This paper, based on a summary of existing methods, innovatively approaches the issue from two perspectives: the weather effect and online public opinion, creating a stock price prediction model that integrates deep learning technologies and verifying its effectiveness through simulated trading. The specific research content is as follows:

Based on the perspective of weather effects, this paper constructs a hybrid prediction model Savgol-BiLSTM-Att that combines Savitzky-Golay filtering and deep learning to predict stock prices, and verifies the model's ability to improve investment returns through

simulated transactions. The model innovatively combines weather data with stock data, and uses the Savitzky-Golay filter algorithm to suppress noise and extract features from the original data, which reduces the interference of data noise and nonstationarity on the prediction accuracy while retaining the key feature information. In terms of feature processing, the model maps the extracted effective features to the high-dimensional space, and captures the latent features in the temporal dependency through the Bidirectional Long Short-Term Memory Network (BiLSTM). In order to further optimize the performance of the model, a temporal attention mechanism is introduced to adaptively weight the time series information, reduce information redundancy, and improve the prediction accuracy. At the level of model application, the TopkDropout quantitative investment strategy was designed based on the prediction results. The experimental results show that the prediction accuracy of Savgol-BiLSTM-Att is significantly better than that of the traditional model, and the Sharpe ratio of the TopkDropout strategy is 0.940078 higher than that of the CSI 300 index, and all evaluation indicators perform well, which fully verifies the application potential of the model in the field of quantitative investment.

Based on the perspective of network public opinion, this paper proposes an interval prediction model FMD-QRBiGRU that integrates FMD eigenmodal decomposition, QR quantile regression and deep learning for stock price prediction and return and risk analysis of real trading simulation. The model quantifies the network public opinion information through web crawler and natural language processing technology, and constructs it into investor sentiment factors and attention factors, and excavates its potential information value. Subsequently, the FMD eigenmodal decomposition was used to adaptively decompose the

original closing price data, extract the multi-scale features and separate the noise components, so as to improve the prediction accuracy of the model. At the same time, the Bidirectional Gated Recurrent Unit (BiGRU) is used to learn the long-term dependence of the original market data, emotional factors, attention factors and features obtained by FMD, and the information interaction between samples is realized through the attention mechanism to mine multi-dimensional features. At the level of model application, a range-bound trading strategy is constructed based on the prediction results. The empirical results show that the FMD-QRBiGRU model is significantly better than the mainstream benchmark model in modeling stock data, achieving an annualized return of 23.5235% and a Sharpe ratio of 1.236564, which are 19.95% and 1.16% higher than the CSI 300 Index, respectively. In addition, its annualized volatility (0.16283) is lower than that of the CSI 300 Index (0.192105), which further verifies that the model has stable returns under controllable risks.

This study employs deep learning and signal processing techniques to enhance stock price prediction accuracy by integrating multi-source data, while developing effective trading strategies to improve investment returns. The findings reveal bond market dynamics and provide empirical support for investor decision-making and regulatory policymaking, offering both theoretical and practical significance.

Fang Yan (Finance)

Supervised by Zhang Wei

KEY WORDS: stock price prediction, quantitative investment, deep learning, signal processing, weather effect, online public sentiment

目 录

摘 要.....	I
ABSTRACT.....	IV
第 1 章 绪论.....	1
1.1 研究背景及意义.....	1
1.1.1 研究背景.....	1
1.1.2 研究意义.....	3
1.2 国内外研究现状.....	4
1.2.1 股价预测方法相关研究.....	4
1.2.2 信号处理技术相关研究.....	8
1.2.3 量化投资相关研究.....	9
1.3 主要研究内容.....	10
1.3.1 研究内容.....	10
1.3.2 研究方法.....	11
1.3.3 技术路线.....	12
1.4 研究创新点.....	13
1.5 本章小结.....	13
第 2 章 理论基础与模型方法.....	14
2.1 量化投资.....	14
2.2 股市的可预测性相关理论.....	14
2.2.1 有效市场假说.....	14
2.2.2 行为金融理论.....	15
2.2.3 有限关注假说.....	15
2.3 信号处理基本理论.....	16
2.3.1 Savitzky-Golay 滤波算法.....	16
2.3.2 FMD 特征模态分解算法.....	17
2.4 深度学习模型基本理论.....	18
2.4.1 点预测模型.....	18

2.4.2	区间预测模型.....	25
2.5	预测模型训练.....	27
2.5.1	滑动窗口.....	27
2.5.2	损失函数.....	27
2.5.3	过拟合处理.....	28
2.5.4	参数优化.....	28
2.6	评估指标.....	29
2.6.1	点预测评估指标.....	29
2.6.2	区间预测评估指标.....	30
2.6.3	交易策略评估指标.....	31
2.6.4	统计检验.....	32
2.7	本章小结.....	32
第 3 章	基于天气效应与 Savgol-BiLSTM-Att 模型的股价预测与量化投资 ...	34
3.1	预测模型构建.....	35
3.1.1	模型构建.....	35
3.1.2	预测结果评估.....	36
3.2	实证分析.....	36
3.2.1	数据获取.....	36
3.2.2	特征选取及构建.....	38
3.2.3	数据处理.....	39
3.3	预测结果及分析.....	41
3.3.1	基准模型结果分析.....	41
3.3.2	本章模型结果分析.....	42
3.3.3	Kolmogorov-Smirnov 检验.....	44
3.4	量化投资应用分析.....	45
3.4.1	TopkDropout 策略的构建.....	45
3.4.2	回测结果分析.....	45
3.5	本章小结.....	46
第 4 章	基于网络舆情与 FMD-QRBiGRU 模型的股价区间预测与量化投资.	48

4.1	预测模型构建	50
4.1.1	FMD-QRBiGRU 预测模型构建	50
4.1.2	预测结果评估.....	53
4.2	实证分析.....	53
4.2.1	数据的获取.....	53
4.2.2	特征选取及构建.....	54
4.2.3	数据处理.....	59
4.3	预测结果及分析.....	60
4.3.1	点预测结果分析.....	60
4.3.2	区间预测结果分析.....	64
4.4	量化投资应用分析.....	66
4.4.1	策略构建.....	66
4.4.2	回测结果分析.....	66
4.5	稳健性检验.....	68
4.6	本章小结.....	71
第 5 章	总结与展望.....	73
5.1	本文工作总结	73
5.2	未来展望.....	74
	参考文献.....	76
	硕士期间发表的论文.....	85
	致谢	86

第1章 绪论

1.1 研究背景及意义

1.1.1 研究背景

作为国民经济体系的重要组成部分，我国股票市场近年来实现了跨越式发展，其资源配置功能日益凸显，对推动经济高质量发展起到了关键作用。伴随居民财富的持续积累，社会投资需求显著提升，股票市场的稳定运行不仅关乎宏观经济调控成效，更直接影响社会公众的财产性收入。因此，维护股票市场健康发展对于优化经济结构、提升民生福祉具有重要的战略意义。

股市波动作为资本市场运行的核心表征，不仅反映了经济的整体走势，更是衡量市场健康状况的关键依据。对股市波动的衡量一般用的是波动率，但波动率仅能反映价格在一定时期内的变化程度，难以全面刻画股价的实际变化情况及其内在价值。相比之下，股价数据作为股市运行最为直观的量化指标，承载着资产内在价值与市场动态变化的双重信息，是描述股市真实运行情况最直接且最有效的指标。基于此，本文对股市波动的探讨将主要立足于股价变动本身，重点研究股价预测问题。

在股票价格是否能够被预测的问题上，学术界存在不同的看法，主要围绕有效市场假说（EMH）与自适应市场假说（AMH）两种理论展开讨论。有效市场假说主张，在市场处于完全有效状态时，股票价格已经充分地将所有可获取的信息纳入其中，呈现出随机游走的特点，因此难以对其进行精准预测^[1]。然而，该理论在现实应用中面临着诸多挑战。一方面，行为金融学的研究成果表明，投资者在投资过程中普遍存在着认知偏差以及决策的非理性行为；另一方面，现实市场中存在着诸如交易成本、流动性约束等摩擦因素，这些因素会对市场效率产生影响。除此之外，套利限制以及投资者之间的异质性也在一定程度上削弱了市场的整体效率。鉴于这些实际情况，自适应市场假说应运而生。该理论认为，市场并非始终处于有效状态，而是存在着不断变化的套利机会，股票价格的变动也往往呈现出一些可以被识别的规律和模式。因此，自适应市场假说认为市场效率是动态变化的，在特定条件下，股票价格可能会呈现出可预测的特征。这一理论观

点为量化投资以及价格预测研究提供了重要的理论支撑,有力地推动了现代金融理论的进一步发展与完善^[2]。李斌和龙真(2023)等学者运用机器学习技术开展研究,证实了中国股票市场确实存在着显著的价格可预测性特征^[3]。

在股价预测领域,早期研究主要采用基本面分析和技术分析两种方法,值得注意的是,这两种经典方法均带有显著的经验导向特征,其有效运用高度依赖投资主体需同时具备系统的金融知识储备和长期的市场历练。当市场参与者缺乏这种复合型能力时,往往难以准确解构多维度的市场信号,进而触发典型的认知偏差现象,出现杀跌追涨等交易行为。计量经济学模型(如 ARIMA、GARCH)从统计学的角度出发,对股票价格进行定量预测,是另一种被广泛应用的预测方法。相较于基本面分析和技术分析方法对市场经验的依赖,计量经济学模型弱化了对市场结构的先验认知需求,转而聚焦于序列本身的统计特性,特别是要求建模数据满足平稳性及线性假设等约束条件。但在现实世界中,股价波动呈现出显著的非平稳性和非线性特征,这与传统计量经济学模型的假设存在根本性背离,使得其在实际应用场景中的解释力和预测精度均受到显著制约。近些年来,计算机技术的突破性发展为该领域带来新的研究视角,以模式识别为核心的人工智能技术,特别是具备自适应学习能力的深度神经网络结构(如 LSTM、CNN),展现出了在捕捉市场复杂结构的独特优势,为股票市场研究注入了新的活力。深度学习模型凭借其特有的多层神经网络架构,能够自动学习数据的高阶特征表示,有效识别数据的非线性关系和高维交互效应,并借助激活函数实现非线性映射,相较于计量经济学模型,其具有更强的拟合能力和泛化性能,尤其适用于处理复杂、非结构化的金融数据,这一特征使其在股价预测领域展现出显著优势,逐渐成为当前研究的热点方向^[4, 5]。

在信息技术的推动下,信号处理技术作为现代科学技术的核心组成部分,在多模态信息的解析与重构方面取得了突破性进展。相较于传统分析方法,信号处理技术具有显著优势,其方法灵活性使其可同时处理平稳与非平稳序列,且对数据分布没有严格限制,能够挖掘其他技术容易忽略的深层信息。时间和频率是刻画信号的重要参数,其中后者是用来描述波动的参数。值得注意的是,股票价格作为典型的时间序列数据,既具有显著的时间依赖性,又蕴含波动特征。这种特

性与信号处理的研究对象具有天然的契合度，因此，为了提高对股价预测的准确性，本研究将探索信号处理技术在股价时序预测中的创新应用。

另外，从股价预测的数据来源来看，传统方法主要依赖市场内部的历史价格数据，其基本假设是历史会重演，通过分析历史规律可推断股价未来走势。然而，从 1637 年荷兰郁金香事件到 2016 年中国股市两次熔断，再到 2020 年全球新冠疫情暴发等一系列极端事件，都反映出股票市场并不总是满足有效市场假说，投资者完全理性的假说也很难成立。研究表明，外部信息（如宏观经济指标、行业新闻和社交媒体舆情等）同样会对股价产生显著影响。由于股价波动受多因素共同作用，单一变量的预测模型往往存在局限性。因此，本研究将综合历史价格与多维外部数据（包括天气、文本和情绪指标等）构建多因子模型，以提升预测模型的准确性和鲁棒性。近年来，量化投资等金融科技与人工智能、深度学习等技术深度融合用于交易决策、风险控制等方面的研究成为新的趋势，所以本文还将研究的理论成果应用于实际投资进行探索。

综上所述，本研究聚焦于基于深度学习的股票市场预测模型优化及交易策略构建，旨在为投资者提供更精准的决策支持，实现收益最大化与风险可控的双重目标。从宏观层面看，精准的市场预测技术有助于监管部门完善市场监测体系，通过前瞻性风险预警机制维护金融稳定，促进资本市场健康发展，最终服务于实体经济的高质量发展。

1.1.2 研究意义

本研究构建的两种深度学习预测模型，通过引入创新性的研究方法和技术手段，为提升投资收益和降低风险提供了新的思路，在股票市场的分析和预测领域具有重要的理论意义和实践意义。

（1）理论意义

第一，在方法论方面，本研究融合了多学科的理论与技术。通过综合运用计算机科学、统计学、金融学等领域的先进技术，构建了系统化的市场分析框架，这种跨学科整合不仅为量化投资研究开辟了新路径，更创新性地将前沿技术应用于股票市场预测实践，拓展了相关领域研究的技术思路。第二，在理论认知方面，本研究对市场有效性理论提出了新的理解。通过实证分析，证实了证券市场存在

可预测性的特征，挑战了传统的有效市场假说，这为深入理解市场运行机制提供了重要的理论依据。

（2）实践意义

第一，在模型构建方面，本研究设计了基于深度学习技术的创新股价预测模型。这些模型能够有效识别和解析证券市场中的非线性特征和潜在规律，显著提升了价格预测的准确度。第二，在实践应用方面，本研究成果展现了突出的实用价值。通过大量实证检验，所提出的模型能够为投资决策提供可靠支持，在优化资源配置和风险控制方面表现出显著优势。第三，在市场建设方面，本研究具有积极的促进作用。通过提升市场信息的处理效率和决策支持能力，有助于优化资源配置效率，推动金融市场向更加规范和成熟的方向发展。

综上所述，本研究不仅在学术层面拓展了相关领域的研究深度，同时为市场参与者提供了切实可行的分析工具和决策参考，对于应对复杂市场环境下的投资挑战具有重要的指导意义。

1.2 国内外研究现状

准确理解股票市场的运行规律对投资决策具有重要价值。结合本研究内容，本节将系统梳理国内外在股价预测模型、信号处理方法以及量化投资策略等领域的研究进展。

1.2.1 股价预测方法相关研究

股票市场是经济社会发展的重要组成部分，也是连接企业和投资者的金融生态系统^[6]。从股市诞生之初，价格趋势预测始终是投资者关注的核心命题，这种持续探索推动着预测方法的演进。然而，股票市场是一个极其复杂的动态系统，受经济、政策等诸多因素的影响，股价具有高噪声、动态、非线性和非平稳等特征^[7,8]，因此，准确地预测股价一直是一项具有挑战性的工作^[9,10]。自 20 世纪以来，金融机构和研究人员致力于股票价格预测模型的开发，目前用于股市预测的方法可分为基本面分析、技术分析、定量预测三种。其中，定量预测主要运用数学模型和计算机技术来实现，具体可分为两方面：一是基于数理统计的计量经济模型，通过构建参数方程捕捉市场规律；二是依托人工智能技术（主要是机器学习、深度学习算法两种），通过数据驱动方式挖掘复杂市场特征。三类方法在理论框架与技术实现层面形成互补，共同推动着金融预测领域的发展。本文研究的

股价预测模型属于定量预测的一种,为此需要系统梳理现有定量预测方法的研究成果,以期在继承现有模型优势的基础上提出更优化的预测模型。

主流的计量经济模型有多元回归模型^[11]、差分整合移动平均自回归模型 (ARIMA)^[12, 13]和自回归条件异方差模型 (GARCH)^[14-17]等。这些计量经济模型是揭示金融市场内部运作的有效工具,在股票市场预测和分析中经常被使用。吴玉霞等(2016)研究了创业板市场股票价格预测问题,选取“华泰证券”250 期收盘价数据构建 ARIMA 模型进行实证分析。研究发现,该模型在短期动态和静态预测方面表现良好,能够有效捕捉股价变动规律,为市场参与者的投资决策提供了有价值的参考依据^[18]。Latha 等(2020)使用 ARIMA 模型对 S&P BSE Sensex 和 S&P BSE IT 指数的收益和波动进行预测。研究结果表明,尽管两个指数的平均收益都为正,但其数值都接近于零。预测值与实际值几乎相等,偏差较小,表明 ARIMA 模型能够较好地拟合历史数据,适用于中长期预测任务^[19]。李晶(2023)基于上证 50ETF 数据和期权合约的历史数据,建立 GARCH 模型进行波动率预测,并结合参数方程进行策略回测,在充分考虑现实交易成本的情况下,仍能实现较好的收益和风险对冲效果,为使用波动率实现期权对冲和风险管理方面提供了思路^[20]。方燕等人(2018)以沪深两市传媒板块指数为研究对象,实证检验了不同时间序列模型的预测性能。研究发现,传统 ARIMA 模型在应对异方差问题时存在明显局限,而 GARCH 模型虽能捕捉波动性但预测精度有限。为此,研究者提出了 ARIMA-GARCH 组合模型,通过整合两者的优势,更好地捕捉到市场中的动态变化,显著提升了预测效果。实证结果显示,相较于单一的模型,ARIMA-GARCH 混合模型不仅能有效处理股价序列的异方差和自相关问题,同时其预测精度也获得了显著提升^[21]。然而,传统的计量模型仅对平稳的金融数据具有较好的预测效果,而对于非线性、非平稳的数据则难以有效识别和捕捉其内在的动态变化规律,导致在股票价格预测领域的应用受到明显制约,预测精度和稳定性往往难以满足实际需求。

随着大数据技术的迅速发展,上述问题在基于人工智能算法的机器学习、深度学习模型中不复存在。目前,使用机器学习、深度学习模型进行多变量分析具有非线性、非平稳的特征数据,已成为金融和经济学中占主导地位和分析工具^[22, 23]。这类模型拥有强大的数据处理及特征提取能力,并可以提供更精确

的预测结果，在金融时间序列分析中表现出优越性能。具有代表性的技术包括：支持向量机（SVR）^[24, 25]、人工神经网络（ANN）^[26]、随机森林（RF）、反向传播神经网络（BPNN）^[27]、卷积神经网络（CNN）^[28]、循环神经网络（RNN）^[29]、长短期记忆神经网络（LSTM）^[30]、门控循环单元（GRU）^[31]等。其中，RNN 凭借其独特的循环结构在捕捉时序数据长期依赖关系方面展现出的显著优势，被广泛应用于时间序列的分析和预测任务。但在实际应用中，该模型极易出现梯度消失、梯度爆炸问题，使其难以完成长期记忆工作。随着深度学习的兴起，LSTM 通过引入门控机制，有效克服了这一难题^[32]，且实现了对信息的有效筛选和记忆，这种特性使得 LSTM 特别适合处理具有复杂时序特征的股票价格数据，相关实证研究也证实了其在股票预测任务中的卓越表现^[33-35]。李斌等（2019）系统地提出了 12 种不同的机器学习算法，基于 A 股市场的 96 项异象因子，构建股票收益预测模型及投资组合，这些算法涵盖了从经典的线性回归到复杂的深度学习等多种模型。研究发现，机器学习算法可以精准捕捉异象因子与超额收益之间的复杂关系，其构建的投资策略在实际应用中表现优异，相比于传统线性算法以及单因子策略，能实现更出色的投资绩效^[36]。王婷等（2020）构建 LSTM 模型预测股票短期趋势，并将 LSTM 与 CNN、RNN 和 MLP 等模型进行了比较，实验结果表明 LSTM 拥有最高的准确率和最低的拟合误差^[37]。双向长短期记忆网络（BiLSTM）能够前向和后向处理序列蕴含规律，在预测中能更充分利用信息。Barua 和 Sharma（2022）的研究发现，BiLSTM 在预测性能上明显优于 LSTM，且 CNN-BiLSTM 深度学习模型在样本外预测指数收盘价时，比单独使用 CNN 或 BiLSTM 效果更好^[38]。Zhang 等人（2023）耦合了 CNN、BiLSTM 和注意力机制，提出 CNN-BiLSTM-Attention 模型以提高股票价格和指数预测的准确性。通过预测中国股票指数（沪深 300 指数）的价格，发现 CNN-BiLSTM-Attention 比 LSTM、CNN-LSTM、CNN-LSTM-Attention 更准确。同时，为了研究所提模型在预测股票指数方面是否具有鲁棒性，选取了中国其他 3 个股票指数和 8 个国际股票指数进行测试，并证实了 CNN-BiLSTM-Attention 模型在预测股票价格方面的鲁棒性，且精度最高^[39]。以上研究成果从不同角度均验证了混合模型在处理时间序列预测任务时的有效性和可靠性。

此外,融合多源数据有望提升预测模型的性能。针对过度反应、代表性偏差等用传统的金融学理论不能解释的市场异象,以及对有效市场的挑战和补充,建立在投资者非理性假设基础上的行为金融理论逐渐形成,并得到了广泛认可和发展。行为金融学致力于探究投资者在决策过程中呈现的心理特质,并深入分析这些心理因素与市场非有效性之间的内在关联。投资者作为证券市场的主体,除了认知、文化、阅历等个人因素外,投资者的情绪成为影响金融市场的一个重要因素^[40]。DeLong 1990 首次揭示了市场套利行为的局限性。基于 DSSW 模型,该研究将部分投资者因外部因素干扰而产生认知偏差,进而导致股价偏离均衡价格的现象定义为投资者情绪效应。这一理论模型的提出,不仅首次建立了投资者情绪与股票收益之间的关联性分析框架,更为后续学者深入探究投资者情绪的作用机制提供了重要的基础^[41, 42]。Kenneth (2000) 通过应用 AAI 指数研究验证了应用投资者情绪预测 S&P500 的收益率的有效性^[43]。王美今等 (2004) 采用央视看盘指数构建投资者情绪指标,发现沪深股市投资者情绪的改变会对股市的波动性产生明显的反向修正作用^[44]。王道平等人 (2019) 分析中国投资者分类情绪及信心变化与中国股票市场波动影响,发现投资者对国内经济基本面和经济政策的信心变化是构成投资者情绪的关键要素,且投资者情绪与股市长期发展态势存在显著关联^[44]。Li 等人 (2020) 通过整合技术指标和新闻情绪,构建股票预测收益模型。研究结果表明,结合技术指标和新闻情感的预测模型表现最优,且金融领域特定的情感词典对预测性能提升效果显著^[45]。Zhang 等人 (2021) 研究了新兴市场 (以中国为例) 的投资者情绪与股市波动率的关系。结果表明,投资者情绪通过非线性模式显著影响股市波动率,且情绪指标能够有效提升波动率预测能力^[46]。李合龙等人 (2023) 构建 GARCH-MIDAS 模型研究投资者情绪对中国股市收益率波动的影响,发现投资者情绪对市场波动产生显著的正向影响,并能够解释 A 股中长期的波动趋势^[47]。

值得注意的是,这些时间序列预测方法都是针对点值时间序列展开的。相比之下,区间数据能够更全面地反映股票市场的内在结构特征。从信息含量的角度来看,点值数据只反映了某一时刻的具体值,而区间数据则同时包含了股价波动范围的上限和下限,能够更好地表征变量的不确定性特征及其变化规律,因而具有更高的信息价值^[10]。Erick 等人 (2021) 认为,区间预测通过量化了预测过程

中的不确定性,能够在特定的置信水平下将数据的不确定性约束在一个可控的范围内,这种预测方式相较于传统的点预测方法展现出更强的实用性和可靠性优势^[48]。而分位数回归(Quantile Regression, QR)凭借其灵活和高效的建模特性,近年来在预测领域受到广泛关注^[49]。例如,He 等人(2017)将 QR 与 SVR 相结合中提出了 SVQR 模型,通过来自新加坡的真实数据集对该模型的预测性能进行了验证。实验结果显示,通过优化配置 KSRQR 模型的核函数,该方法在电力负荷预测任务中展现出良好的应用前景^[50]。Hu 等(2022)将共形分位数回归(CQR)创新性地运用到 TCN 模型中,提出了一种 CTCQRN 区间预测模型。该模型在两个不同地理风电数据集上的实证研究表明,所提模型与基准模型相比具有明显的优势,显著提升了预测精度,有助于电力系统的经济安全运行^[51]。

1.2.2 信号处理技术相关研究

构建结合信号分解的混合模型能有效提升股价预测的准确性。随着数据的复杂化和多样化,预测波动大的复杂数据相对平稳数据更具挑战性,受到噪声干扰的数据会致其信噪比低,不进行任何处理便代入模型进行训练并预测很难取得较为准确的预测结果。为了减轻股价数据中的噪声影响进而提高预测效率,许多研究人员将深度学习算法与特征分解相结合,常见的有 EMD、CEEMDAN、VMD 等。谢游宇等(2023)将经验模态分解和奇异谱分析与 LSTM 模型和 SVR 模型进行结合,提出 EMD-SSA-LSTM-SVR 模型,提升了预测模型的效果^[52]。韩莹等(2023)为了提高金融时间序列的预测精度,引入 CEEMD 进行降噪处理,提出 CEEMD-LSTM-BLS(C-L-B)股票预测模型对农林牧渔行业股票价格进行实证分析。结果验证了 CEEMD-LSTM-BLS(C-L-B)股票预测模型效果最优^[53]。Lin 等人(2022)采用融合 LSTM 和 CEEMDAN 的混合模型对 CSI300、S&P500 和 STOXX50 指数的实际波动率 RV 进行预测。结果表明,CEEMDAN-LSTM 在新兴市场和发达市场预测 RV 方面均具有最佳性能^[54]。欧阳红兵等(2020)将小波分解算法与 LSTM 相结合对道琼斯工业指数的收盘价进行预测,实验验证了 LSTM 模型相较于其他基准模型拥有更高的预测精度,同时,小波分解算法能有效提高 LSTM 模型的泛化能力和预测精度^[55]。Qi 等人(2023)将 GRU 与 CEEMDAN 相结合,并使用小波阈值法对信号中的高频噪声进行去噪。构建的

GRU-CEEMDAN-wavelet 模型在 S&P500 和 CSI 300 股票指数的数据集上进行实证研究,结果表明,与其他基准模型相比,该模型具有最优预测精度^[56]。

尽管以上这些经典的信号分解算法具备简单高效的处理数据的能力,但也存在着一些缺点,如端点效应和模态混淆等。这些问题在一定程度上限制了信号分解的精度和可靠性。针对此局限性,Miao 等(2023)创新性地提出了特征模态分解(FMD)方法^[57]。该理论框架融合了解卷积理论的核心思想,通过动态优化滤波器参数的迭代机制,构建具有自适应特性的 FIR 滤波器系统。这种设计使得信号经过滤波处理后能够逐步逼近解卷积目标函数,从而实现对非线性复杂信号的高效分解。FMD 的核心在于利用自适应 FIR 滤波器组,将信号分解为多个独立的模态分量,每个分量都能清晰地反映信号的特定特征。与传统方法相比,FMD 不仅能够同时处理信号的周期性和冲击性,还对噪声等干扰成分具有很强的抗干扰能力,有效抑制这些干扰,从而提高信号分解的准确性和稳定性。这一创新理论的提出,为信号处理领域带来了新的思路和方法,也为解决传统算法的局限性提供了有力的工具。王晓真等人(2024)的研究结果表明,FMD 与 VMD 相比具有更好的特征提取性能,能够实现滚动轴承故障的精确诊断^[58]。根据信号领域的相关研究,Savitzky-Golay 滤波也非常适用于非线性、高噪声的数据信号降噪。Yang 等人(2023)构建 Savitzky-Golay-TCN 深度卷积神经网络,对中国金融市场的系统性风险指标进行预测,结果验证了 Savitzky-Golay 滤波做数据处理的有效性,与 TCN、CNN 和 LSTM 相比,Savitzky-Golay-TCN 深度卷积神经网络具有最优的预测精度^[59]。

滤波等信号处理技术的应用显著增强了模型的精度与稳定性,但在计算资源的优化配置方面仍需进一步探索。注意力机制通过动态分配信息权重,能够有效提升预测效率。Qiu 等(2020)^[60]、Ouyang 等(2021)^[61]、Abbasimehr 等(2022)^[62]的研究均能说明,加入注意力机制后,模型的预测性能得到明显提升。

1.2.3 量化投资相关研究

随着计算机技术和统计学的发展与结合,量化投资逐渐成为一种主流投资方式。量化投资是通过分析市场中存在的价格波动规律,由计算机发出指令,来进行投资的一种方法。在多数研究中,采用以下规则:当预测模型发出买入信号时,

算法便执行买入操作，并等待卖出信号以确定盈利或损失。近年来，量化投资领域正加速与智能算法深度融合。通过运用机器学习与深度学习技术，量化投资模型展现出对多维金融数据的高效解析能力，不仅实现了海量信息的精准化处理，更有效规避了传统人工决策中的认知偏差问题。这种技术融合显著抑制了非理性因素对投资行为的干扰，为构建科学化、系统化的投资方案提供了技术支撑。马一琳（2020）利用 LSTM 模型的个股预测结果改进 omega 模型，基于预测收益建立投资组合优化模型，结果表明该模型实证结果最优，且风险偏好对模型性能影响不大^[63]。潘唯一（2020）研究了机器学习在量化投资领域的应用，发现相较于传统方法，神经网络等先进算法可有效提升金融预测精度，同时，能为量化投资提供超额收益^[64]。何雨洵（2021）提出了一种结合金融市场情绪因素与历史交易数据的机器学习量化交易策略。该策略融合了自然语言处理（NLP）和 LSTM 两种深度学习模型，通过 Bert 模型分析舆情情感，LSTM 预测股价涨跌，并结合均值回归理论构建多因子框架。在沪深 300 指数成分股上的回测结果表明，该策略显著提升了投资绩效^[65]。林昱（2024）提出一种基于深度学习的股票预测框架，包含 PatchCNN、TE-EAN 和 Batch-MCRNN 三个创新模型。PatchCNN 通过“切片”技术将时间序列转化为类图像数据，结合 CNNMixer 架构预测收益率，其 TopkDrop 策略实现了 171.7% 的年化收益。TE-EAN 模型融合经验模态分解与时间注意力网络，有效捕捉股价涨跌趋势，AUC 达到 0.732，累计收益超 80%。Batch-MCRNN 框架利用多尺度卷积和批注意力机制预测重要交易点，回测年化收益达 68.3%，夏普比率为 2.913。这些模型通过优化数据处理和特征提取，显著提升了预测精度和投资收益^[66]。

1.3 主要研究内容

1.3.1 研究内容

基于以上研究背景，本研究构建了跨学科融合的创新研究框架：整合信号分解技术与深度学习算法，建立融合多源数据的股价预测模型，并将研究成果应用于量化投资实践，以期提升投资策略的收益，促进股市的健康发展。本文的技术路线图如图 1-1 所示，各章具体内容安排如下：

第 1 章，绪论。首先阐述了选题的研究背景及意义。其次，对股价预测、信号处理及量化投资等领域相关的国内外研究文献进行了系统性综述，通过对比分

析现有研究的优势与局限，确立了本研究的方向、思路和方法。然后，给出本研究的内容、方法、框架及创新点。

第 2 章，相关理论基础。介绍了量化投资相关理论，有效市场假说、行为金融学、有限关注假说等股市可预测性的相关理论，以及本文构建股票价格预测模型所涉及到的信号处理算法、深度学习相关理论，包括 Savitzky-Golay 滤波算法、FMD 特征模态分解算法、BiLSTM 相关理论、BiGRU 相关理论、Attention 机制等。同时介绍各种评价指标，为后续研究提供理论基础。

第 3 章，研究天气效应对股价波动及交易策略的影响。首先，收集股票收盘价以及上海地区的天气（温度、光照、云层覆盖率、AQI、雨、雪、风速）数据作为模型输入，通过滤波处理方法对收集到的天气数据和股价序列进行数据处理，构建基于股市天气效应的深度学习预测模型，并通过实证分析评估模型的预测能力及收益情况。本章的内容主要是从天气效应角度解释股价的波动，并进行量化投资研究。

第 4 章，研究网络舆情对股价波动及交易策略的影响。利用东方财富网股吧评论数据构建投资者情绪因子和投资者关注度因子，将其与 FMD 模态分解收盘价序列结果结合，作为模型输入，构建基于网络舆情的深度学习模型，实证分析评估模型的预测能力及收益情况。本章的内容主要是从网络舆情的情感和影响力两个角度解释股价的波动，并进行量化投资研究。

第 5 章，总结和展望。总结本文所有工作，并展望后续还待进一步完善的研究方向和内容。

1.3.2 研究方法

本研究旨在构建股价预测模型并应用于量化投资实践，为实现研究目标，采用了多种研究方法：

（1）理论分析法。系统梳理股价预测、量化投资等领域的相关文献，深入分析各类预测模型的优势与局限性，进而识别具有研究价值的优化方向。这一过程为本研究的模型创新提供理论支撑。

（2）实证研究法。本研究选取中国 A 股市场作为研究样本，运用深度学习技术，结合信号处理技术等方法，通过系统化的实验设计和数据测算，构建了股价预测模型，并据此构建量化交易策略。实证结果显示，跨学科技术融合显著提

升了模型预测准确度与投资收益,同时,所提出的深度学习模型为处理金融市场复杂性问题提供了创新性工具。

(3) 数据驱动分析法。本研究系统分析了影响股价波动的关键变量,并构建了融合多源数据的股价预测模型,有效提升了模型预测的精确度和结论的可靠性。

(4) 对比分析法。本研究通过与传统预测模型的横向比较,从经济解释力、预测精度及投资回报三个关键指标出发,系统验证了深度学习模型在证券市场预测中的优越性及其应用价值。

1.3.3 技术路线

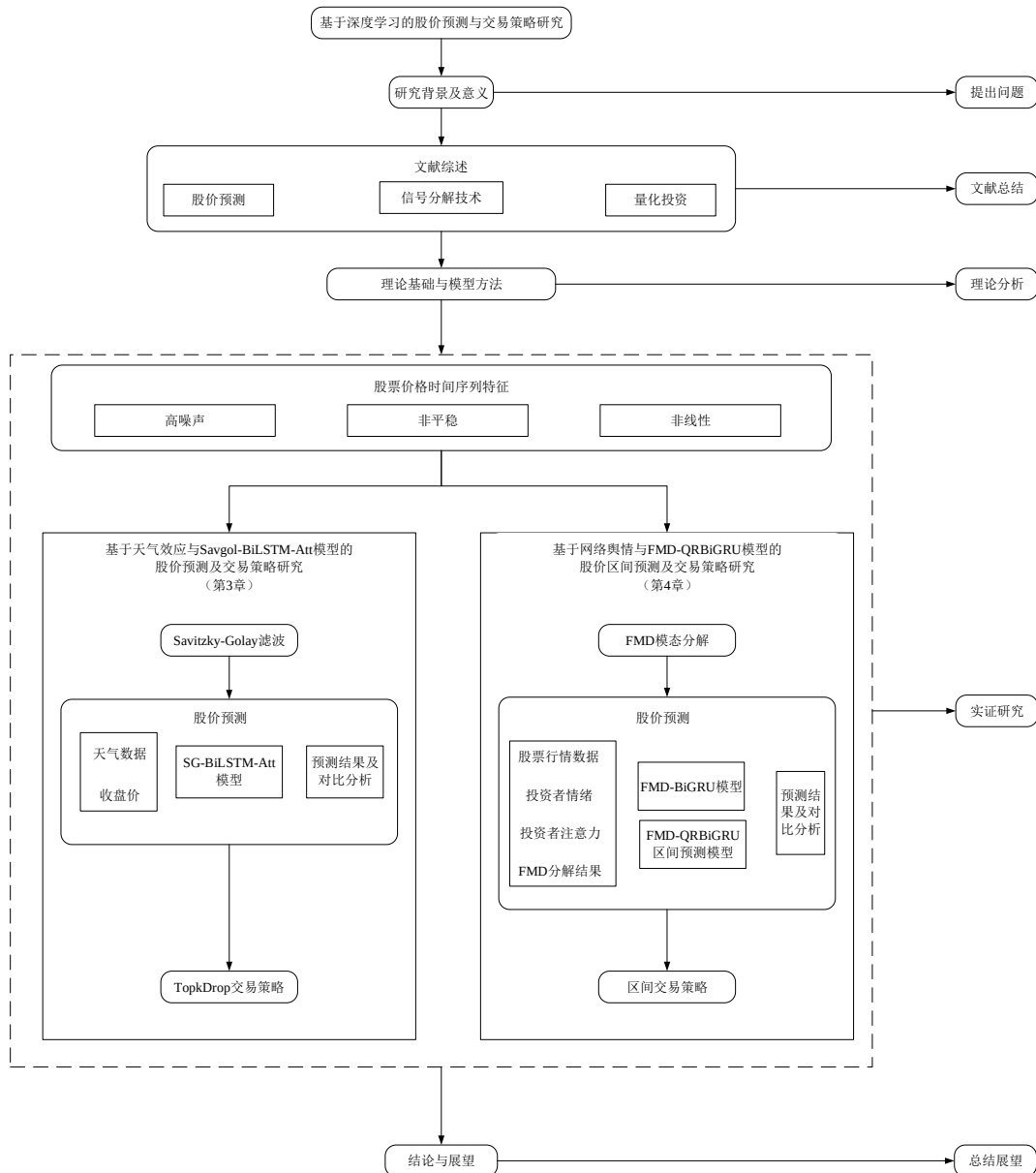


图 1-1 技术路线图

1.4 研究创新点

本文的主要创新之处主要体现在以下几点：

（1）基于金融市场天气效应提出了 Savgol-BiLSTM-Att 预测模型。该模型创新性地将天气因素引入股价预测研究，并结合 Savitzky-Golay 滤波技术优化特征提取，有效提升了预测精度。实证研究表明，该模型在中国 A 股市场前 30 大市值股票上的预测表现显著优于传统模型，且基于其构建的 TopkDrop 交易策略在回测中大幅跑赢沪深 300 指数，验证了模型的实际应用价值。

（2）基于网络舆情视角提出了 FMD-QRBiGRU 区间预测模型。该模型利用股吧评论数据构建投资者情绪和关注度因子，结合 FMD 模态分解技术对股价进行去噪和特征提取，实现了更稳健的股价预测。实证结果显示，引入网络舆情和 FMD 分解能显著降低预测误差，同时基于区间约束的交易策略在提升收益的同时有效降低了投资风险。

（3）提供了新的研究视角和方法。本文首次将天气效应、网络舆情与深度学习相结合，并创新性地运用信号处理技术（如 Savitzky-Golay 滤波、FMD 分解）优化金融数据特征提取。此外，从天气和舆情两个非传统维度探索股价波动规律，为量化投资研究开辟了新的思路，对提升预测精度和策略收益具有启示意义。

1.5 本章小结

本章首先介绍了研究的背景及意义，指出研究股价预测与量化投资问题的重要性。其次，通过整理与总结相关文献，梳理并确立了本研究的思路和方法。接着，给出了研究的核心思路及内容、技术路线图，以及本研究的创新点，为后续研究提供了丰富的支撑。

第2章 理论基础与模型方法

在本章中，主要介绍后续研究中所涉及的相关理论基础和模型方法，为下文的模型构建及实证分析奠定基础。重点阐述的理论包括量化投资、股市可预测性以及信号处理相关理论，同时详细介绍深度学习模型的相关方法。此外，本章还提供了模型训练的参数设置以及后文所需的各项评估指标。

2.1 量化投资

量化投资旨在通过定量的方法识别和利用市场中的投资机会，获得额外利润，基本流程包括数据收集、处理、模型构建、策略回测和实时交易。传统投资方式是投资者依靠自身的投资经验，进行长期价值投资，而这种投资方式极度依赖投资者的眼光和判断。与传统投资不同，量化投资依赖于数据驱动的决策过程，通常涉及大量历史数据分析和算法模型构建，因此更具客观性和精准性。量化投资策略分为多种类型，如套利策略、趋势跟随策略、价值投资策略和高频交易等，强调风险管理的重要性，使用止损、仓位管理和风险预算等技术降低潜在损失。

量化投资在金融市场中得到了广泛应用，尤其在对冲基金、资产管理公司和投资银行等机构中，随着计算能力的提升和数据获取的便利，量化投资的策略和模型不断演进，吸引了越来越多的投资者参与。量化投资在未来金融市场中将发挥越来越重要的作用。

2.2 股市的可预测性相关理论

2.2.1 有效市场假说

有效市场假说（Efficient Market Hypothesis, EMH）是建立在“理性人”和有效的市场竞争的假设基础上的，由 Fama 在 20 世纪 60 年代提出。该理论认为，金融市场是有效的，不存在信息不对称问题，市场价格反映了市场所有已知的信息。而市场的随机性使得投资者很难预测未来的价格变化，因此，投资者无法通过分析历史数据或其他信息来获得超额收益，长期投资和分散投资被认为是更为有效的投资策略。有效市场假说的三种形式（弱有效市场假说、半强有效市场假说、强有效市场假说）构成了完整的理论框架，为研究市场信息效率提供了重要的理论基础。

尽管有效市场假说在学术界得到了广泛的支持，但在实际市场中，仍然存在一些反例和争议，特别是在行为金融学的研究中，投资者的非理性行为和市场泡沫等许多主流金融学理论无法解释的异象挑战了这一理论的普遍适用性。

2.2.2 行为金融理论

行为金融学（Behavioral Finance）作为金融学的分支学科，主要探究心理认知因素对投资决策及市场运行的影响机制。区别于传统金融理论中“理性人”假设，该领域研究发现投资者在实际决策过程中往往受到情绪波动、认知偏差及群体效应等多重因素的干扰，从而导致非理性的投资行为。心理偏差（如过度自信、损失厌恶）、市场情绪（如贪婪和恐惧）、群体行为（如羊群效应）、框架效应等均会影响投资者的决策。行为金融学的研究从微观个体来对经济市场中发生的现象进行分析，揭示了市场并非总是有效的，投资者的非理性行为可能导致市场价格偏离其基本价值。这一理论为理解市场波动、资产定价和投资策略提供了新的视角，并对金融危机和市场异常现象提供更为合理的解释。

情绪因子能够有效捕捉市场非理性波动，弥补传统有效市场假说的解释局限，因此，将投资者情绪纳入股价预测模型和量化投资策略具有重要价值。

2.2.3 有限关注假说

有限关注假说（Limited Attention Hypothesis）是行为金融学中的一个重要理论，旨在解释投资者在信息处理和决策过程中所面临的认知限制。该假说认为，投资者的注意力是有限的，当面临市场繁杂的信息时，投资者对某一事件的关注程度必然会降低其对其他事件的关注程度，他们无法同时关注所有可用的信息，因此在做出投资决策时，往往只关注部分信息，而忽略其他重要的因素。投资者的注意力往往集中在一些显著或容易获取的信息上，例如热门股票、新闻头条或社交媒体上的讨论，而同时这种关注偏差可能会导致市场价格的非理性波动。除此之外，信息过载、时间延迟、市场反应等均会影响投资者对市场信息的获取，从而造成价格的短期波动。有限关注假说为理解投资者行为和市场动态提供了新的视角，强调了心理因素在金融决策中的重要性。这一理论帮助解释了许多市场现象，如资产价格的非理性波动、信息传播的非对称性以及投资者在面对复杂决策时的行为模式。

因此，在人工智能飞速发展的时代，网络新闻、股吧论坛等社交媒体引导并承载着人们越来越多的关注，如何量化投资者注意力，探索其对股价趋势及量化投资的影响意义重大。现有研究表明，互联网的搜索量^[67, 68]、发帖数量^[69]、新闻数量^[70]均能够在一定程度上反映投资者有限关注特征，为研究提供了新的数据来源和分析维度。

2.3 信号处理基本理论

2.3.1 Savitzky-Golay 滤波算法

Savitzky-Golay（简称 Savgol）滤波算法是一种用于平滑和微分数据的数字滤波技术，通过合理选择窗口大小和多项式阶数，可以有效地平滑数据并提取有用的特征信息。Savgol 滤波因其在保留信号特征方面的优越性被广泛应用，尤其在化学、物理和生物医学等领域的信号处理和数据分析中成为一种重要的工具。该算法基于时间域内的多项式最小二乘拟合，在去除信号噪声的同时保持其形状和宽度不变。此外，Savgol 滤波器能够在同一数据曲线上应用不同窗口宽度，以适应不同的平滑和滤波需求，特别是在处理时序数据时表现出色，同时对于来自非周期性和非线性源的噪声处理也十分有效。

对当前时刻 (y_t) 的前后一共 $2n+1$ 个观测值进行滤波，数据为 $y[i], i = -n, \dots, 0, \dots, n$ ，用 k 阶多项式对其进行拟合：

$$y = a_0 + a_1x + a_2x^2 + \dots + a_{k-1}x^{k-1} \quad (2.1)$$

同样，对于前后时刻（如 $t-1$ 、 $t+1$ 等时刻）的预测值，同样可以用式(2.1)计算得到 $2n+1$ 个方程，因此，拟合方程可以用矩阵形式来表示，其中，数据值向量为 Y ，系数矩阵为 X ，自变量矩阵为 A ，残差为 E ，式子如下：

$$Y_{(2n+1) \times 1} = X_{(2n+1) \times k} \times A_{k+1} + E_{(2n+1) \times 1} \quad (2.2)$$

要使得整个矩阵有解，必须满足 $2n+1 > k$ ，这样才能通过最小二乘法确定参数 $a_0, a_2 \dots a_{k-1}$ ，基于此，求得 A_{k+1} 的解 $\hat{A}$ 为：

$$\hat{A} = (X^T X)^{-1} X^T Y \quad (2.3)$$

那么，模型的滤波值 $\hat{Y}$ 为：

$$\hat{Y} = X\hat{A} = X(X^T X)^{-1} X^T Y \quad (2.4)$$

2.3.2 FMD 特征模态分解算法

FMD (Feature Modal Decomposition) 特征模态分解算法是一种用于信号处理和分析的技术，其核心在于将复杂的信号分解为多个模态，每个模态都能精准地反映信号的某一特定特征，从而实现对信号的有效分析和处理。在实际应用中，FMD 算法展现了强大的功能。它不仅能够显著提高信号处理的精度，还能为后续的特征提取和分析提供丰富的信息。在工程领域，它可以帮助工程师更精准地监测设备运行状态，提前发现潜在故障；在医学领域，它能够助力医生从复杂的生理信号中提取关键信息，为疾病的诊断提供有力依据；在金融领域，它更是能够从海量的市场数据中挖掘出有价值的信息，为投资决策提供参考。

在 FMD 的实现过程中，涉及多个关键步骤，包括 FIR 滤波器组设计、系数更新、故障周期估计以及模态个数选择等。具体而言，针对原始信号 x_n ，算法首先需要完成迭代参数的初始化，并采用汉宁窗口对 FIR 滤波器组进行初始化。在此基础上，以相关峭度 (CK) 作为目标函数，通过迭代的方式对滤波器组进行优化更新。这一过程不仅确保了滤波器能够逐步适应信号的特征，还为信号的分解提供了精准的工具。

$$\arg \max_{\{f_k(l)\}} \left\{ CK_M(u_k) = \sum_{n=1}^N \left(\prod_{m=0}^M u_k(n-mT_s) \right)^2 / \left(\sum_{n=1}^N u_k(n)^2 \right)^{M+1} \right\}$$

$$s.t. u_k(n) = \sum_{l=1}^L f_k(l)x(n-l+1) \quad (2.5)$$

式中， u_k 为第 k 个模态分量； $CK_M(u_k)$ 为相关峭度， $f_k(l)$ 为第 k 个长度为 L 的 FIR 滤波器； T_s 为输入周期； M 为移位顺序， N 为信号长度。根据 u_k 获得分解后的模态分量，其分解模式可以通过式(2.6)表示：

$$u_k = X * f_x \quad (2.6)$$

其中， $*$ 为卷积计算。此时，

$$CK_M(u_k) = \frac{u_k^H W_M u_k}{u_k^H u_k} = \frac{f_k^H X^H W_M X f_k}{f_k^H X^H X f_k} = \frac{f_k^H R_{XWX} f_k}{f_k^H R_{XX} f_k} \quad (2.7)$$

式中， R_{xwx} 为加权相关矩阵； R_{xx} 为相关矩阵。 CK 的最大化等价于下式：

$$R_{xx} f_k = R_{xx} f_k \lambda \quad (2.8)$$

在不断迭代过程中，第 k 个滤波器的系数会依据式(2.8)的解持续进行更新与调整，其优化目标在于使滤波输出信号的相关峭度值达到最大。当所有滤波器的系数完成更新后，通过计算相邻模态分量之间的相关系数，并构建相关矩阵 $CC_{(K \times K)}$ ，锁定相关系数最大的一对相邻模态分量，并计算其 CK ，同时淘汰其中 CK 较小的模态分量。最后，将迭代次数 K 增加 1，继续进行下一轮迭代，即 $K=K+1$ 。这一过程不断重复，直至满足预设的条件，从而完成整个信号分解的过程。

2.4 深度学习模型基本理论

2.4.1 点预测模型

(1) 支持向量回归

SVR 是目前机器学习领域用得较多的方法，由 Vapnik 在 1999 年提出，主要用于学习输入变量和连续输出变量之间的函数关系。即对于样本集合：

$$D = \{(x_1, y_1), \dots, (x_i, y_i), \dots, (x_n, y_n), x_i \in R^d, y_i \in R\} \quad (2.9)$$

要找到一个超平面使得所有样本到超平面的距离之和最小，如果训练集数据线性可分，即 SVR 模型的数学公式如下：

$$\min \frac{1}{2} \sum_{i,j=1}^n (\hat{\alpha}_i - \alpha_i)(\hat{\alpha}_j - \alpha_j) x_i^T x_j - \sum_{i=1}^n [y_i(\hat{\alpha}_i - \alpha_i) - \varepsilon(\hat{\alpha}_i + \alpha_i)], s.t. \sum_{i=1}^n (\hat{\alpha}_i - \alpha_i) = 0 \quad (2.10)$$

求解得到的预测函数为：

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \hat{\alpha}_i^*)(x_i^T x) + b^* \quad (2.11)$$

如果训练集的数据呈现非线性的特点，SVR 关键是通过引入非线性映射函数，将低维空间中具有非线性回归关系的数据集映射到高维空间，然后转化为线性回归关系，其回归函数可表示为：

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \hat{\alpha}_i^*) K(x_j, x_i) + b^* \quad (2.12)$$

其中 b 为偏差项。

$$b = \begin{cases} y_i - \varepsilon - \sum_{i=1}^n (\alpha_i^* - \hat{\alpha}_i^*) K(x_j, x_i) \\ y_i + \varepsilon - \sum_{i=1}^n (\alpha_i^* - \hat{\alpha}_i^*) K(x_j, x_i) \end{cases} \quad \alpha_i^* - \hat{\alpha}_i^* \in [0, C] \quad (2.13)$$

$K(x_j, x_i)$ 为核函数，包括多项式核、径向基核函数、神经网络核函数等。

阶次为 d 的多项式核：

$$K(x_j, x_i) = [(x_j, x_i) + 1]^d \quad (2.14)$$

径向基核函数：

$$K(x_j, x_i) = \exp\left(\frac{-\|x_j - x_i\|^2}{2\sigma^2}\right)^d \quad (2.15)$$

神经网络核函数：

$$K(x_j, x_i) = \tanh[c_1(x_j, x_i) + c_2] \quad (2.16)$$

(2) 循环神经网络模型

RNN 是一种对序列数据有较强的处理能力的神经网络模型。传统网络通常采用线性架构，数据从输入层经隐藏层传递至输出层，各层间通过全连接方式相连，但同层节点间不存在连接关系。与传统的神经网络不同，RNN 的独特之处在于其循环结构，如图 2-1 和图 2-2 所示，网络在每个时间步的处理过程中，不仅会接收当前时刻的输入信息，还会整合前一时刻的隐藏状态，以此生成当前时刻的隐藏状态输出。这种独特的循环连接方式使 RNN 具备记忆能力，能够有效捕捉序列数据中的时间依赖关系，因此在处理具有时间相关性的数据时展现出显著优势。从计算角度看，RNN 的基本组成单元是循环神经元，其运算过程包含两个关键步骤：首先对当前输入和前一时刻的隐藏状态进行线性组合，然后通过激活函数实现非线性转换。该过程的数学表达可描述为：

$$S_t = f(UX_t + WX_{t-1} + b) \quad (2.17)$$

$$O_t = VS_t + c \quad (2.18)$$

其中, S_t 是当前隐藏状态, $f(\bullet)$ 为非线性激活函数, U 和 W 为权重矩阵, b 为偏置项, O_t 是 t 时刻隐藏层的输出, V 是权重参数, c 是偏置。

该计算公式反映了网络状态随时间推移的演化规律, 其中前一时刻的隐藏状态承载着历史信息的记忆, 当前时刻的输入提供新的数据特征, 而偏置项则用于调整模型的输出基准。

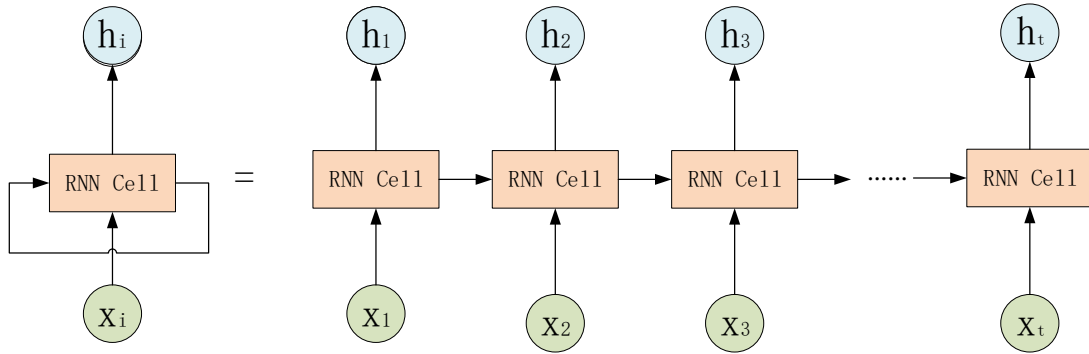


图 2-1 循环神经网络结构

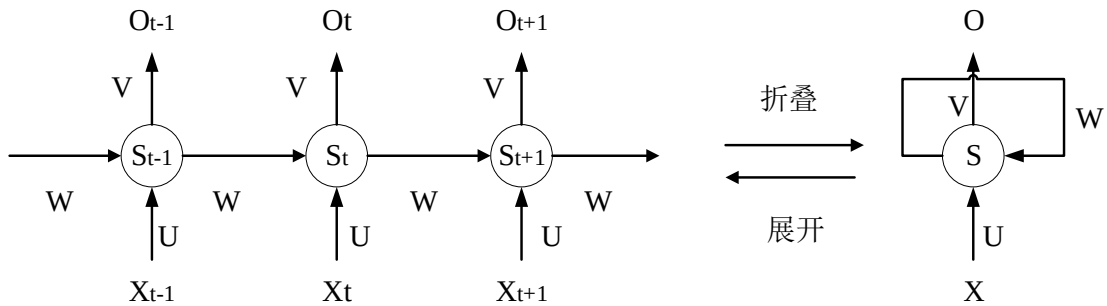


图 2-2 循环神经网络计算结构

(3) 双向长短期记忆神经网络

在实际应用中, RNN 模型的局限性逐渐显现, 尤其是在处理长序列数据时, 容易出现梯度消失或梯度爆炸的问题, 且随着时间步的增加, RNN 对早期信息的记忆能力会逐渐衰退, 导致其难以完成长期记忆工作。针对这些技术瓶颈, Hochreiter 和 Schmidhuber 在 1997 年创新性地提出了 LSTM。该模型通过引入门控单元系统 (包括输入门、遗忘门和输出门), 对 RNN 的隐层结构进行了优化, 从而有效缓解了梯度消失和梯度爆炸的问题。这种独特的门控机制使 LSTM 能够智能地筛选和保留重要信息, 在处理具有长期依赖特性的时序数据时表现出色。近年来, LSTM 被广泛应用于股价趋势预测等领域, 并取得了显著的研究成果。其结构如图 2-3 所示。

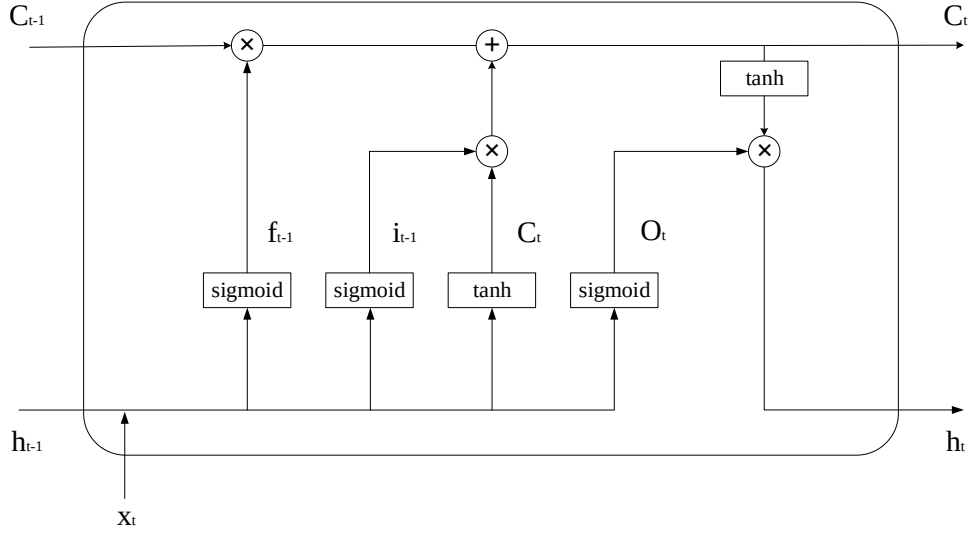


图 2-3 LSTM 结构图

LSTM 的基本结构由四个关键组件构成：遗忘门、输入门、输出门和内部记忆单元。其中，遗忘门 f_t 采用 sigmoid 激活函数进行调控，该机制通过分析前一时刻的隐状态和当前输入特征，生成一个 0 到 1 之间的权重系数，以此来决定哪些信息需要从细胞状态 C_{t-1} 中丢弃。公式如下：

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \quad (2.19)$$

其中，遗忘门输出值 f_t 介于 0 到 1 之间，其数值大小直接决定信息的保留程度，0 表示完全遗忘，1 表示完全保留。 h_{t-1} 代表 $t-1$ 时刻的隐藏状态， x_t 表示当前时刻的输入， W_f 和 b_f 分别为对应的权重矩阵和偏置向量， σ 代表 sigmoid 激活函数。

LSTM 的输入门 i_t 通过 sigmoid 函数和 $\tanh$ 函数的协同配合实现信息调控，根据上一个隐藏层的输出 h_{t-1} 和当前输入 x_t 以此来决定哪些新信息存入细胞状态。公式如下：

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \quad (2.20)$$

$$C_t = \tanh(W_c \times [h_{t-1}, x_t] + b_c) \quad (2.21)$$

$$C_t = f_t \times C_{t-1} + i_t \times C_t \quad (2.22)$$

其中, i_t 取值范围为 0~1, 控制新信息的重要性, σ 为 *sigmoid* 激活函数, $\tanh$ 是双曲正切激励函数, C_t 是候选记忆, C_t 表示当前时刻细胞状态, W_i 和 W_c 为对应的权重矩阵, b_i 和 b_c 是偏置项。

输出门 o_t 通过 *sigmoid* 函数生成调控系数, 将该系数与经过 $\tanh$ 函数归一化的细胞状态 C_t 相乘, 从而决定当前隐藏状态 h_t 的输出。公式如下:

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o) \quad (2.23)$$

$$h_t = o_t \times \tanh(C_t) \quad (2.24)$$

其中, h_t 表示当前时刻隐藏状态, 其数值由当前细胞状态 C_t 经门控机制处理后生成。

可以看出, 遗忘门的工作机制使得细胞状态 C_t 的梯度能够直接传递到前一时间步的 C_{t-1} , 这一特性正是 LSTM 有效缓解梯度消失和爆炸问题的关键所在。随着硬件设施的进步和预测精度要求的提高, BiLSTM 应运而生。该网络通过并行整合前向 LSTM 层和后向 LSTM 层, 实现了对时序数据的双向特征提取。这种结构不仅保留了传统 LSTM 处理序列数据的优势, 还能同时捕获历史和未来的信息, 显著提升了模型的表征能力和鲁棒性。BiLSTM 结构如图 2-4 所示。

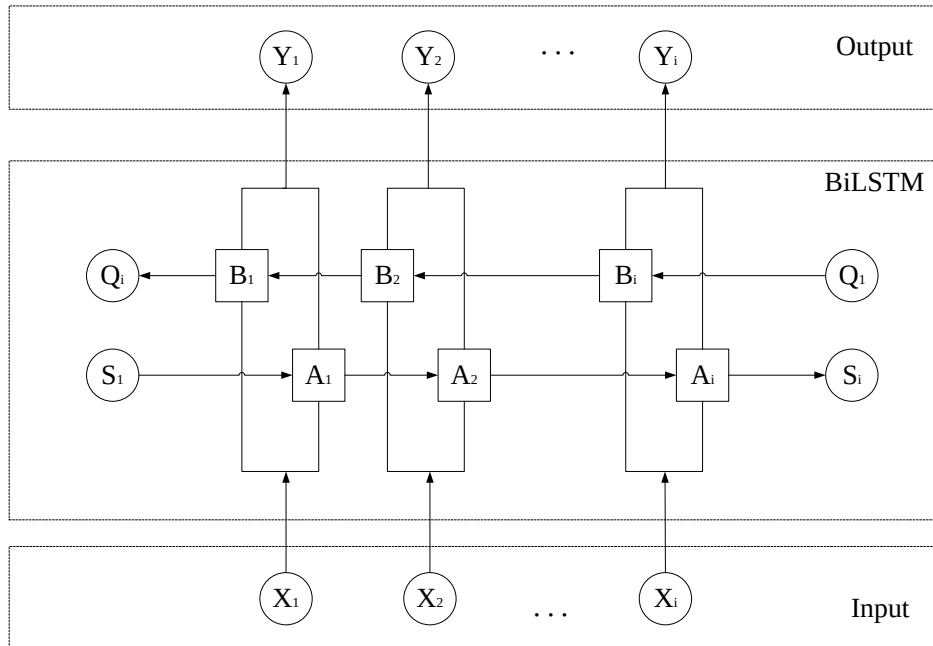


图 2-4 BiLSTM 结构图

(4) 双向循环控制单元

门控循环单元（GRU）是 Cho 等人^[71]在 2014 年提出的一种改进型循环神经网络，它在保留 LSTM 门控机制优势的基础上进行了结构优化。GRU 通过将传统 LSTM 中的遗忘门和输入门合并为单一的更新门，并引入重置门来控制历史信息的遗忘程度，不仅有效缓解了梯度消失和爆炸问题，还显著降低了模型参数数量。这种精简结构使 GRU 在保持良好时序建模能力的同时，计算效率得到明显提升。重置门决定是否将之前的状态忘记，由 $t-1$ 时刻的输出 h_{t-1} 与当前时刻的输入 x_t 决定重置门的输出 r_t 。 r_t 计算公式为：

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (2.25)$$

更新门 z_t 决定是否将隐藏状态 $\tilde{h}_t$ 更新为当前时刻的输入。 z_t 计算公式为：

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (2.26)$$

$$\tilde{h}_t = \varphi(r_t x_t \otimes U_h h_{t-1} + W_h x_t) \quad (2.27)$$

经过重置门和更新门后， h_t 由 h_{t-1} 与 $\tilde{h}_t$ 共同决定， h_t 计算公式为：

$$h_t = (1 - z_t) \otimes \tilde{h}_t + z_t \otimes h_{t-1} \quad (2.28)$$

其中， $\varphi(\bullet)$ 表示非线性激活函数； W 和 U 表示权重矩阵。

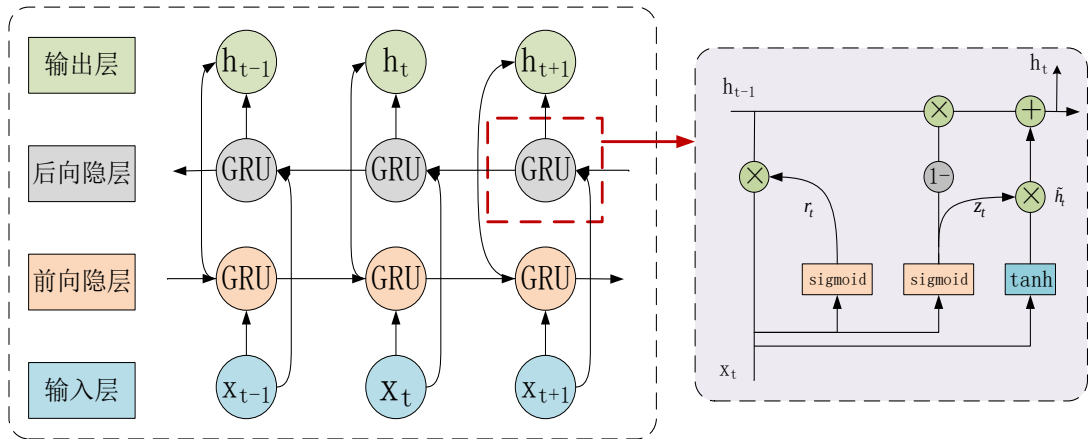


图 2-5 BiGRU 网络结构图

双向门控循环神经网络（BiGRU）是在 GRU 模型基础上发展而来，充分考虑数据之间前后变化的信息规律，通过引入门控机制以及双向结构，将两个 GRU 以相反时序连接，前向 GRU 获取输入序列的过往信息，后向 GRU 获取输入序

列的未来信息，增强了对序列数据的建模能力。将其应用到股价预测中，可以对输入的特征进行双向提取，预测当前股票价格的状态，提高模型性能。计算公式如，GRU 和 BiGRU 的网络结构如图 2-5 所示。

$$\begin{cases} \vec{h}_t = GRU(x_t, \vec{h}_{t-1}) \\ \tilde{h}_t = GRU(x_t, \tilde{h}_{t-1}) \\ \tilde{h}_t = \alpha_t \vec{h}_t + \beta_t \tilde{h}_t + b_t \end{cases} \quad (2.29)$$

式中， $GRU(\bullet)$ 表示门控循环单元； $\vec{h}$ 和 $\tilde{h}$ 为前向和后向的隐藏层的输出； α_t 和 β_t 为前向和后向的隐藏层各自的权重； b_t 为偏置。

(5) 注意力机制

注意力机制（Attention Mechanism, Att）在神经网络中是一种有限计算资源的优化分配机制，它使网络能够将资源集中在更重要的任务上。传统神经网络及其变体模型虽然具有记忆能力，但模型对于较早之前序列信息的提取和记忆能力有限，无法有效区分处理信息时各个信号的重要性。注意力机制可以把神经网络的隐藏层状态信息保存起来，通过为不同特征分配不同权重，能够聚焦于对当前任务最关键的信息，从而降低对其他信息的关注，提高处理效率和准确性。这解决了神经网络模型在处理长序列时可能出现的信息丢失问题。因此，引入注意力机制能够进一步提高股价预测的准确性。

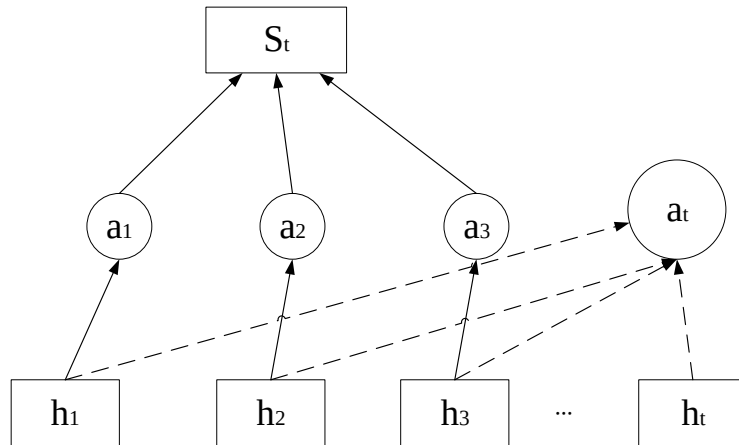


图 2-6 注意力机制结构图

对于一个*i*维的特征输入，首先通过密集层（全连接层）处理，生成与输入维度相对应的隐藏输出值 $\{h_1, h_2, h_3, \dots, h_t\}$ 。随后这些隐藏输出值经过 $softmax$ 函数的转换，形成一个权重总和为1的特征权重向量 $\{a_1, a_2, a_3, \dots, a_t\}$ 。值得注意

的是，这一转换过程并未改变特征的原始维度。最后将每个输入特征与对应的特征权重相乘，得到注意力层的输出结果，注意力机制结构图如图 2-6 所示。

计算公式为：

$$u_t = v_u^T \tanh(W_u h_t + U_u k_i) \quad (2.30)$$

其中 k_i 为初始化查询向量。注意力权重分布 α_t 计算公式为：

$$a_t = \frac{\exp(u_t)}{\sum_{i=1}^n \exp(u_i)} \quad (2.31)$$

通过权重 a_t 对各状态信息 h_t 进行加权求和，从而得到该时刻的注意力机制层的输出状态 S_t ，以影响模型对当前时刻的预测结果。

$$S_t = \sum_{t=1}^n a_t h_t \quad (2.32)$$

2.4.2 区间预测模型

分位数回归理论最早由 Bassett 等^[72]提出，目前在经济学等多个领域被广泛应用。该方法采用非对称加权残差最小化策略，不仅能估计条件均值，还能完整刻画因变量的条件分布特征。与传统最小二乘回归相比，分位数回归具有显著优势，它不要求误差项服从特定概率分布，这一数据驱动的特性使其在处理非标准分布数据时表现出更强的稳健性。设有 c 个解释变量 $U = \{U_1, U_2, \dots, U_c\}$ 作用于随机变量 S ，其分布函数可表示为：

$$F(s) = P(S \leq s) \quad (2.33)$$

对任意 $\tau, \tau \in [0, 1]$ ，有：

$$F^{-1}(\tau) = \inf \{s : F(s) \geq \tau\} \quad (2.34)$$

式中 $F^{-1}(\tau)$ 为 S 的第 τ 分位数； $\inf \{s\}$ 为集合 s 的下确界。

线性分位数回归模型中，给定解释变量 U 时，响应变量 S 的第 τ 个条件分位数 $Q_s(\tau|U)$ 可以表示为：

$$Q_s(\tau|U) = \beta_0(\tau) + \sum_{i=1}^c \beta_i(\tau) U_i = U' \beta(\tau) \quad (2.35)$$

其中 $\beta(\tau)$ 表示 τ 分位点下的回归系数向量。未知参数向量 $\hat{\beta}(\tau)$ 的估计量可通过求解得到:

$$\begin{aligned}\hat{\beta}(\tau) &= \min_{\beta} \sum_{i: S_i \geq U'_i \beta} \tau |S_i - U'_i \beta(\tau)| + \sum_{i: S_i < U'_i \beta} (1-\tau) |S_i - U'_i \beta(\tau)| \\ &= \min_{\beta} \sum_i \rho_{\tau}(S_i - U'_i \beta(\tau)), \tau \in (0,1)\end{aligned}\quad (2.36)$$

其中 $\rho_{\tau}(u)$ 为检验函数, 其通过使 $\rho_{\tau}(u)$ 最小, 来求解 $\beta(\tau)$ 。检验函数 $\rho_{\tau}(u)$ 为:

$$\rho_{\tau}(u) = u(\tau - I(u < 0)) = \begin{cases} u\tau, u \geq 0 \\ u(\tau - 1), u < 0 \end{cases}, \quad I(u < 0) = \begin{cases} 0, u \geq 0 \\ 1, u < 0 \end{cases} \quad (2.37)$$

从以上定义可知, 只要变化分位点就能够知道解释变量对被解释变量不同分位数的影响, 进而就能够更加全面地刻画被解释变量和解释变量的关系。在实际应用中, 解释变量与被解释变量之间的关系往往呈现出复杂多样的非线性特征。然而, 分位数回归模型的构建基于线性假设, 这种假设在一定程度上限制了对变量之间真实影响方式的刻画, 使其分析结果更倾向于线性关系的呈现。相比之下, BiGRU 能够有效模拟输入与输出之间的非线性结构, 从而为处理复杂的变量关系提供了一种更为灵活和精准的方法。因此, 本文整合 QR 和 BiGRU 形成的 QRBIGRU 模型实现了由解释变量至响应变量的非线性关联, 增强了预测的全面性。基本框架如图 2-7 所示。

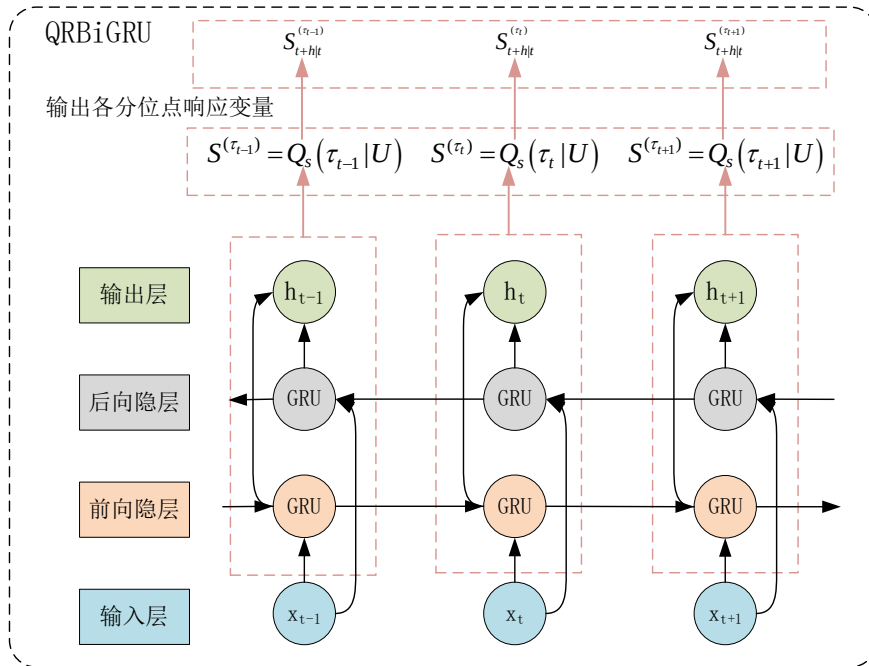


图 2-7 QRBIGRU 的基本框架

2.5 预测模型训练

2.5.1 滑动窗口

股票指标数据反映了股票交易信息随时间的动态变化,并呈现出一定的惯性趋势。为了确保模型训练的计算既有效又可行,本文采用了如图 2-8 所示的滑动窗口方法来划分数数据集。具体而言,滑动窗口的大小(即时间步长)设为 t ,滑动步长为 1。鉴于本文使用的是股票日度数据,即根据过去 t 个交易日的数据向量作为模型的输入,预测未来下一个股票交易日(即第 $t+1$ 日)的收盘价。一般来说,第 i 个滑动窗口的输入数据 $w_i(t)$ 为 $x_i(t)$ 至 $x_{i+30}(t)$,输出数据为 $x_{i+31}(t)$ 。对于下一个滑动窗口,输入数据为 $x_{i+1}(t)$ 至 $x_{i+31}(t)$,输出为 $x_{i+32}(t)$ 。重复步骤,并持续到时间序列的末尾。

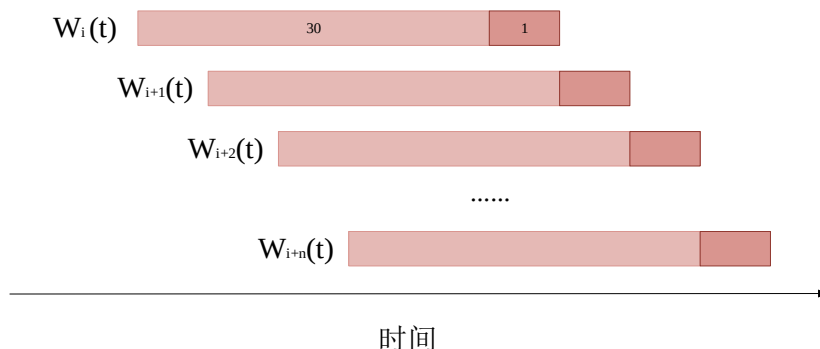


图 2-8 滑动窗口示意图

2.5.2 损失函数

在机器学习的庞大体系中,损失函数(Loss Function)扮演着至关重要的角色,它就像是衡量模型“成绩”的一把尺子。损失函数的作用在于衡量模型预测值与实际值之间的差距。简单来说,模型预测的结果越接近真实值,损失函数的值就越小;反之,差距越大,损失函数的值就越大。因此,模型训练的目标就是尽可能地减小这个差距,让损失函数的值降到最低。在监督学习中,损失函数的构建主要基于模型预测的输出与真实标签之间的差异。比如在图像识别中,模型需要判断一张图片是猫还是狗,而真实答案就是这张图片实际的类别。损失函数就会根据模型的判断与真实答案之间的差距来打分。而在无监督学习中,由于没有明确的“答案”,损失函数通常被定义为对数据的某种度量。例如,在聚类问题中,我们可以使用 K 均值损失函数或者高斯混合模型的负对数似然损失函数,来衡

量数据点之间的相似性或分布情况。在深度学习中，损失函数更是与优化器紧密相连。优化器的目标是通过调整模型的参数，让损失函数的值不断减小。在训练过程中，优化器会根据损失函数的梯度来更新模型参数，以便最小化损失函数的值。

在本文中，选择了均方误差（ MSE ）作为损失函数。均方误差是一种简单而有效的衡量方式，它通过计算预测值与真实值之间差值的平方平均值，直观地反映了模型的预测精度。通过不断优化 MSE ，模型能够更好地适应数据，从而在实际应用中发挥更大的价值。

2.5.3 过拟合处理

过拟合，顾名思义，就是模型对训练数据“学得太过认真”，以至于它不仅记住了数据中的规律，还记住了数据中的噪声和细节。这样的模型在面对未知数据时，往往会表现出极差的适应性，甚至可能出现错误的预测。这不仅削弱了模型的泛化能力，也让其在实际应用中的价值大打折扣。Dropout 是一种正则化技术，在神经网络的训练时，它会随机地将一部分神经元的输出值置为零。通过这种方式，神经元之间的依赖关系被削弱，模型不得不学会从其他途径寻找解决问题的方法，从而降低了模型的复杂度。Dropout 的另一个好处是，它让模型在训练过程中始终保持一种“不确定性”。这种不确定性让模型在面对各种情况时都能保持一定的灵活性。最终，经过 Dropout 处理的模型在面对未知数据时，能够更加从容地应对，而不是被训练数据中的细节所束缚。

在本文中，为避免神经网络模型出现过拟合，在训练过程中引入 Dropout，不仅有效降低了模型的复杂度，还显著提高了模型的泛化能力。

2.5.4 参数优化

神经网络采用反向传播机制来计算神经元间连接权重和偏置等参数 θ 的梯度，并利用梯度下降优化算法沿着梯度的反方向更新参数。常见的梯度下降优化算法有随机梯度下降（SGD）、动量法和 Adam 等。其中，Adam 算法融合了动量优化和自适应学习率的优势，通过计算每个参数的一阶矩（平均值）和二阶矩（方差）的指数移动平均，实现对各参数学习率的动态调整。这种自适应优化机制能够有效应对不同参数的梯度变化特性，显著提升模型参数 θ 的训练效率。

在本文中，选择了 Adam 算法作为梯度下降优化算法。通过 Adam 算法的智能优化，神经网络模型能够在复杂的环境中快速找到最优路径，从而更高效地完成训练任务。这不仅提高了模型的训练效率，还让模型在面对复杂多变的数据时，能够更加灵活地调整自身结构，更好地适应实际应用的需求。

2.6 评估指标

为了让研究成果不仅仅停留在理论层面，而是能够切实地反映其对实际投资决策产生的影响和价值，本研究将模型的预测方法和结果应用到真实的投资场景中。这一过程中，需要借助一些客观且可量化的指标，来全面和客观地评估基于模型的投资策略是否有效。鉴于此，本文选取了一系列常用的评价指标，这些指标不仅能评估模型的预测能力（包括点预测和区间预测），还可以用来对所构建的股票组合策略的投资效果进行评价。除了常见的误差评估指标外，本研究还从统计学的角度对模型的预测能力进行了评估。

2.6.1 点预测评估指标

本文采用 MAE 、 $RMSE$ 、 $MAPE$ 、 MSE 、 $RRMSE$ 、 ARV 、 $SMAPE$ 、 R^2 这 8 种评价指标评估点预测模型的表现。其计算公式分别为：

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.38)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.39)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (2.40)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.41)$$

$$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\max(y) - \min(y)} \quad (2.42)$$

$$ARV = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{\hat{y}_i} \right| \quad (2.43)$$

$$SMAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{(|\hat{y}_i| + |y_i|) / 2} \quad (2.44)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2} \quad (2.45)$$

其中，式中 $\hat{y}_i$ 和 y_i 分别是预测股价和实际股价， n 是测试集中的样本数量。

2.6.2 区间预测评估指标

下面将重点介绍用于评估预测区间的关键指标（*PICP* 和 *PINAW*），这些指标能够帮助我们更全面地衡量模型在不确定性量化方面的表现。

（1）区间覆盖率（*PICP*）：用于评估预测区间覆盖未来股票价格的真实值的概率。*PICP* 越接近置信水平，意味着模型的股价预测能力越出色，区间预测的可靠性越高。

$$\zeta_{PICP} = \frac{1}{N} \sum_{t=1}^N \delta_t \quad (2.46)$$

$$\delta_t = \begin{cases} 1, y_t \in [L_t^\alpha, U_t^\alpha] \\ 0, y_t \notin [L_t^\alpha, U_t^\alpha] \end{cases} \quad (2.47)$$

其中， L_t^α 和 U_t^α 为预测区间的上下限。

（2）区间宽度（*PINAW*），用于衡量预测区间的精确程度，该指标在保证覆盖率的条件下越小，说明区间预测的精度越好。

$$\zeta_{PINAW} = \frac{1}{N \times \Psi} \sum_{t=1}^N U_t^\alpha - L_t^\alpha \quad (2.48)$$

$$\Psi = y_{\max} - y_{\min} \quad (2.49)$$

其中， Ψ 为归一化因子； $y_{\max}$ 和 $y_{\min}$ 分别为当日股票收盘价实际的最高价和最低价。

这两个指标共同构成了对区间预测效果的综合评价体系，分别从可靠性和精确性两个维度反映模型的预测性能。

2.6.3 交易策略评估指标

本文依据预测结果计算 *Annual Return*、*Annual Volatility*、*Max Drawdown*、*Sharpe Ratio*、*Alpha* 和 *Beta* 共六类投资评价指标来检验模型用于量化投资的有效性。具体投资评价指标计算公式如下：

(1) 年化收益率是指投资组合在一年内的平均收益率：

$$Annual\ Return = (1 + \bar{r}_p)^{252} - 1 \quad (2.50)$$

$\bar{r}_p$ 为平均每日收益率

(2) 年化波动率是衡量投资组合收益波动性的指标，通常使用标准差来计算：

$$Annual\ Volatility = \sqrt{\frac{m}{n-1} \sum_{i=1}^n (r_p - \bar{r}_p)^2} \quad (2.51)$$

其中， m 是根据所选数据周期计算的全年数据量。 n 是交易日数目， r_p 是策略交易的日收益率。本文中 $m = 252$ 。

(3) 最大回撤是指在一定时间里，投资组合可能遭受的最大价值下降，它可以用来衡量投资组合潜在损失。

$$Max\ Drawdown = \frac{Max_t (\max_{\tau \leq t} P(\tau) - P_t)}{\max_{\tau \leq t} P(\tau)} \quad (2.52)$$

其中， $\max_{\tau \leq t} P(\tau)$ 是直到时间 t 的最大投资组合价值， P_t 是在时间 t 的投资组合价值， Max_t 表示在所有时间点 t 上最大化这个比值。在量化投资策略评估中，最大回撤是一个非常重要的风险指标，因为它可以帮助投资者了解在最糟糕的情况下可能会遭受多大的损失。

(4) 夏普比率是每承担一单位总风险所获得的超额回报。它衡量的是风险调整后的收益。

$$Sharpe\ Ratio = \frac{E(R_p) - R_f}{\sigma_p} \quad (2.53)$$

R_p 表示投资策略收益率； R_f 表示无风险收益率，本文设置为 2%； σ_p 表示投资策略收益率标准差。

(5) α 衡量的是投资相对于其 β 预测的市场回报的超额回报。

$$\alpha = R_p - [R_f + \beta_p (R_m - R_f)] \quad (2.54)$$

R_p 为年化收益率, R_f 为无风险利率, R_m 为基准年化收益率, β_p 为该策略的 β 值。

(6) β 值衡量的是投资相对于市场基准的波动性或系统性风险。

$$\beta = \beta_p = \frac{\text{cov}(r_p, r_m)}{\text{var}(r_m)} \quad (2.55)$$

r_p 指的是策略每日收益, r_m 指的是基准的每日收益。

2.6.4 统计检验

除了误差评估指标外, 本研究还采用两样本 Kolmogorov-Smirnov 检验, 从统计学角度评估模型的预测能力。K-S 检验是一种用于检测两个独立分布之间相似情况的非参数统计方法。以检验投资者情绪是否影响股价为例, 设加入投资者情绪指数后的样本 x 分布函数为 $F(x)$, 未加入投资者情绪指数指标的样本 y 分布函数为 $G(y)$, 如果经 K-S 检验认为 $F(x)$ 与 $G(y)$ 分布差异很大, 则认为投资者情绪是影响股价的主要因素之一, 反之则认为投资者情绪不是影响股价的主要因素。

$$H_0: F(x) \equiv G(y) \leftrightarrow H_1: F(x) \neq G(y) \quad (2.56)$$

检验统计量为:

$$D_N = \max \{ |F_p(X_{(i)}) - G_q(Y_{(j)})| \} \quad (2.57)$$

D_N 为两个样本分布的经验分布函数之间的最大距离。K-S 检验的原假设 H_0 为两分布函数为同一分布函数, 即两样本分布相同, 拒绝原假设表明两样本分布不同。

2.7 本章小结

本章深入探讨了研究的理论基础和模型方法。在理论基础方面, 先后介绍了量化投资、股市可预测性、信号处理相关理论。在模型方法方面, 着重介绍了深度学习模型的相关理论, 包括点预测模型和区间预测模型。最后还给出了预测模

型训练的参数设置以及后文需要用到的评估指标，包括点预测评估指标、区间预测评估指标、交易策略评估指标、以及统计检验。

第3章 基于天气效应与 Savgol-BiLSTM-Att 模型的股价预测 与量化投资

随着行为金融学的发展，学者们发现天气作为影响投资者情绪的因素之一，对于人们在金融市场的交易决策和交易行为具有显著影响^[73-75]。1993 年，Saunders 首次揭示了天气变化对股票市场的影响，他认为纽约证券交易所附近的多云天气对投资者情绪产生了显著的负面影响，云层覆盖率与股指之间存在长期显著的相关关系，而我们在阴天和晴天经历的情绪差异表明了天气会影响情绪。此后，天气变化对投资者交易行为以及股票市场表现的影响逐渐成为学者们关注的焦点^[76-78]。Güven 和 Hoxha（2015）研究了天气因素对幸福感及风险决策的影响机制，发现阳光作为外生变量能显著提升个体幸福感，而幸福感增强会导致风险规避倾向加剧。他们通过构建预测模型证实，幸福的人群在金融决策中更偏好稳健投资，且对长期经济指标（如通胀）的预期更为乐观^[79]。Dong 等人（2021）研究了每个地区和月份的天气（阳光、风、雨、雪和温度）与每日指数回报之间的关系，发现这五个天气变量都对股票回报有很强的影响，天气通过影响投资者情绪进而影响股票回报，并且天气的情绪影响比以前记录的更强烈、更普遍^[80]。进一步地，Dong 等人（2022）运用开盘前天气数据（日照、风力、降水、降雪及温度）及其与市场回报的统计关系进行每日指数收益预测，评估了基于全球 49 个国家每日天气条件的指数级交易策略收益性，结果发现基于天气条件构建的交易策略能产生显著超额收益（年均超额收益 9%），且证实了多种天气变量相较于单个天气变量会对全国股票收益产生了更重要的经济影响^[81]。近年来，也有研究表明空气污染对中国股市收益存在负向影响，并且较高的空气质量会使市场倾向于获得更高的收益率、更低的换手率和波动率^[82-84]。

因此，本章特从影响投资者情绪最稳定因素之一的天气着手，提出了一种基于天气效应的 Savgol-BiLSTM-Att 模型来预测股价，并通过投资应用来评判模型的应用价值，来探究天气效应对股票市场的影响。

3.1 预测模型构建

3.1.1 模型构建

基于以上天气对金融市场影响的研究现状,本章提出了一种基于天气效应的 Svagol-BiLSTM-Att 模型预测股价,并使用固定策略 Topkdropout 进行回测,以验证所使用模型的可行性。Svagol-BiLSTM-Att 模型的算法流程图如图 3-1 所示,具体步骤如下:

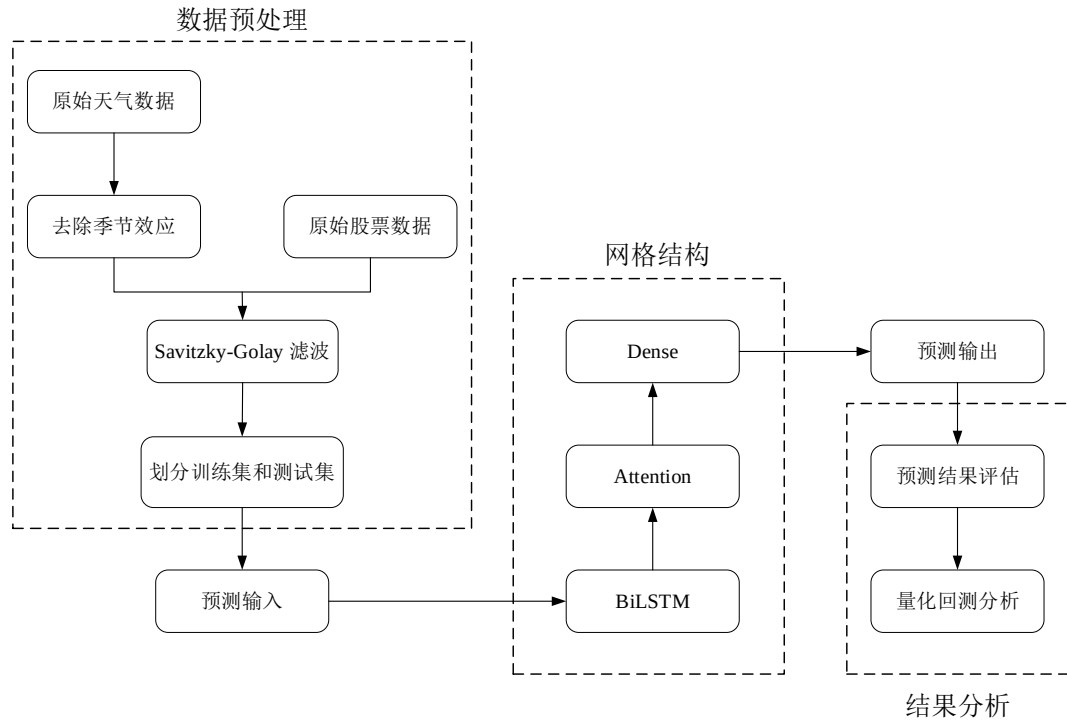


图 3-1 Svagol-BiLSTM-Att 预测结构图

步骤 1: 数据的采集与预处理。首先,利用 python 网络爬虫技术收集上海市的天气资料,并经过量化赋值等数据处理方式对原始数据进行处理,得到天气指标:温度、光照、云层覆盖率、AQI、雨、雪、风速。同时对温度等天气指标进行去季节化处理,剔除季节效应干扰气温等对股票价格的影响。

步骤 2: 数据清洗与特征工程。使用 Savgol 滤波对天气数据和股价数据进行平滑去噪,再对其进行归一化处理,并划分训练集和测试集。

步骤 3: 深度学习模型构建。将进行预处理之后的序列输入到 BiLSTM 层进行训练,以学习内部动态变化模式,提取时序特征。再引入了注意力机制,自动为 BiLSTM 层提取的特征分配不同的权重,以探索深层时间相关性。最后,通过 Dense 层得到模型输出值。

步骤 4: 模型验证与评估。对预测结果进行反归一化处理得到所需要的预测值, 采用 MSE 等指标评估模型性能, 并通过 K-S 检验验证预测模型的有效性。

步骤 5: 投资策略回测。根据预测结果制定固定策略 Topkdropout 并进行回测, 验证预测模型在量化投资方面的应用价值。

3.1.2 预测结果评估

本章中使用第 2 章中展示的其中四个评估指标来衡量每个模型的预测结果, 分别为 MSE 、 $RMSE$ 、 $RRMSE$ 和 R^2 。同时, 采用两样本 K-S 检验, 从统计学角度评估模型的预测能力。

3.2 实证分析

本章探讨机器学习与深度学习模型在分析天气变量与股票价格波动关系方面的潜力, 并通过 Topkdropout 策略的回溯测试来验证其性能。研究基于中国 A 股市场历史数据和天气信息, 分为模型拟合、预测分析和策略设计三部分。模型拟合环节构建并比较了 9 种模型的预测精度。预测分析环节重点评估各模型在预测天气与股价关联性方面的表现。最后, 在策略设计环节, 选取了 Svagol-BiLSTM-Att 模型, 探讨其在实际股市中的应用有效性。

3.2.1 数据获取

本章主要为了探究天气效应对我国金融市场的预测性能, 选取了中国 A 股证券市场市值排名前 30 的股票进行实证分析, 因为市值前 30 的股票具有机构持仓比例高 (例如贵州茅台、宁德时代等)、市场影响大、流动性高、数据噪声小 (便于排除其他干扰因素, 例如财务造假等) 等特征。样本区间为 2013 年 4 月 11 日至 2023 年 7 月 25 日, 选取近十年的收盘价数据及天气数据, 对于上市不足 10 年的股票, 数据收集自其上市日期起。天气数据选取的是上海地区的数据, 这是因为上海是中国的金融中心, 主要投资者都集中于此, 且股票的主要交易席位也位于上海。对于每个数据集, 研究将交易日 80% 的数据作为训练集, 剩余 20% 为测试集, 如表 3-1 所示。所有股票数据源自 wind 金融数据库和天气网站 (<https://lishi.tianqi.com/>、<https://www.aqistudy.cn/>)。

收盘价作为市场参与者普遍认可的价格, 蕴含着大量有价值的信息, 对于预测未来交易日股票价格的走势至关重要^[85, 86]。基于此, 本文将股票下一个交易日的收盘价作为预测对象。部分股票的收盘价数据及其划分情况如图 3-2 所示。

表 3-1 训练集和测试集的时间划分

序号	股票代码	总数	训练集	数量	测试集	数量
1	000001.SZ	2369	2013.11.4- 2020.8.9	1895	2020.8.10- 2023.7.25	474
...	...	...	...	...	...	...
24	300750.SZ	1245	2018.6.11- 2022.7.15	996	2022.7.18- 2023.7.25	249
25	300760.SZ	1161	2018.10.16- 2022.8.8	928	2022.8.9- 2023.7.25	233
26	300999.SZ	676	2020.10.15- 2022.12.29	540	2022.12.30- 2023.7.25	136
27	601658.SZ	879	2019.12.10- 2022.11.3	703	2022.11.4- 2023.7.25	176
29	603288.SH	2304	2014.2.11- 2021.8.26	1843	2021.8.27- 2023.7.25	461
30	688981.SH	735	2020.7.17- 2022.12.14	588	2022.12.15- 2023.7.25	147

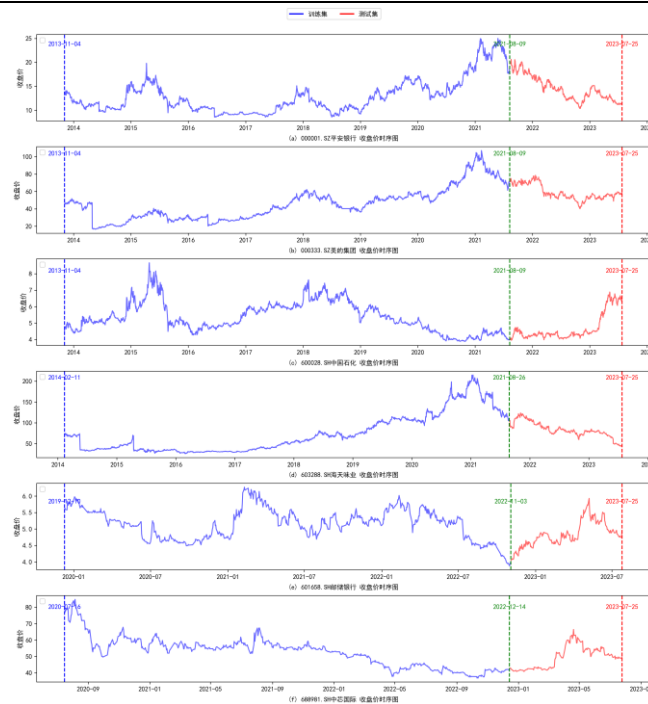


图 3-2 部分股票收盘价趋势图和数据集拆分

3.2.2 特征选取及构建

特征选择是深度学习和数据挖掘中的重要步骤，它对模型性能和预测准确性有很大的影响，特征选择的好坏直接关系到模型的泛化能力和效率。本章共设置两类输入特征，分别为股票收盘价数据、天气数据。具体如下：

（1）股票收盘价数据。从 wind 数据库获取。部分股票收盘价的统计特征如表 3-2 所示。

表 3-2 部分股票收盘价统计特征

	000001.SZ	000333.SZ	600028.SH	603288.SH	601658.SH	688981.SH
count	2369	2369	2369	2304	879	735
mean	13.224	47.941	5.228	74.973	5.086	52.180
std	3.481	17.910	0.928	42.019	0.455	8.916
median	12.55	48.97	5.03	69.38	5.14	36.83
min	8.57	16.57	3.88	25.94	3.88	53.71
max	25.01	107.11	8.67	214.98	6.27	84.82

（2）天气变量数据

本研究的天气数据为上海市的日度数据，选取的天气变量包括：温度(temp)、光照(sunshine)、云层覆盖率(skc)、AQI、雨(rain)、雪(snow)、风速(wind)。数据库中没有直接提供云层覆盖率，而是披露了与云层覆盖率直接相关的天气状况如晴、晴间多云等。因此本章借鉴陆静^[87]的量化方法，采用 1-5 分的离散标度对云层覆盖程度进行转换赋值。具体赋值标准如表 3-3 所示。

表 3-3 skc 赋值

天气状况	赋值	云层覆盖率
晴	1	0-10%
晴间多云	2	10%-30%
多云	3	30%-50%
阴或小雨	4	50%-80%
中雨或雾或雪	5	80%-100%

由于极端的天气状况也会对市场、通讯或其他商业活动造成损害，所以本章也对雨和雪等天气进行了检验。雨、雪、光照用虚拟变量表示，1 代表下雨、下雪、光照充足，0 代表无雨、无雪、光照不足。

3.2.3 数据处理

(1) 缺失值处理

本文的实验对象是中国 A 股证券市场市值排名前 30 的股票，选取 2013 年 4 月 11 日至 2023 年 7 月 25 日期间的全部交易日数据作为研究样本。考虑到不同股票的上市时间存在差异（部分股票在 2010 年前已上市，而部分如宁德时代等则在 2014-2023 年间陆续上市），对于在样本期开始后上市的股票，统一采用其实际上市日期至 2023 年 7 月 25 日的数据作为有效观测值。这种处理方法既保证了数据的时间连续性，又避免了因上市时间差异导致的数据缺失问题。

(2) 季节性指标处理

对于天气指标中的季节性变量，本研究采用差异化处理策略：首先，虚拟变量（如云层覆盖率）已通过赋值转换消除了季节性影响；其次，如下图 3-3(a) 的所示，温度等天气变量呈现显著季节性特征。针对日频天气数据特性，不可以采用 X_{11} 等处理季节数据的方法来处理，因此本文采用周度均值差分法进行去季节化处理，即用每日观测值减去当周均值，达到平均化的目的，去季节化后温度数据如图 3-3(b) 所示。该方法有效消除了温度数据的周期性波动，从而更准确地捕捉其与股价的真实关联。

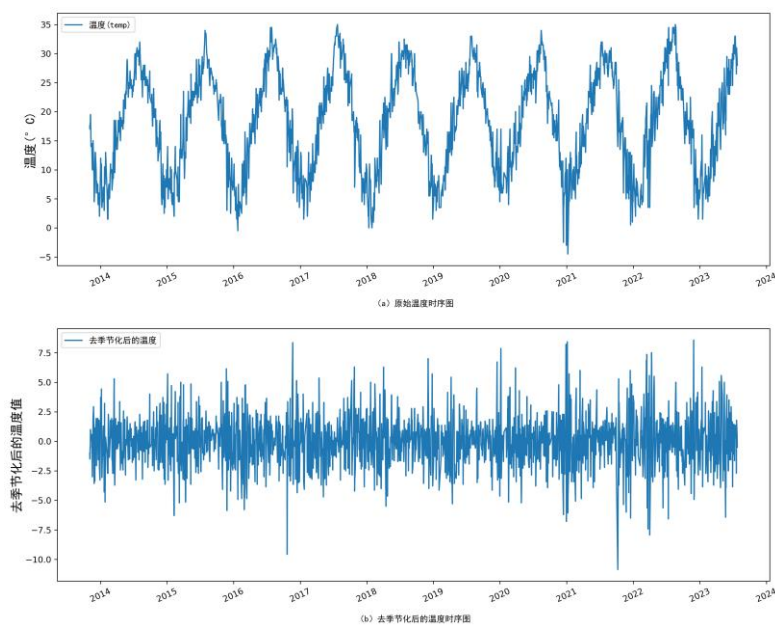


图 3-3 温度的季节性趋势图

（3）Savgol 滤波处理

金融数据通常包含很多噪声，这种噪声可能来自市场的短期波动或者是数据采集的不精确。与其他平滑技术相比，Savgol 滤波器的一个显著优势在于，它在平滑数据的同时，能够更好地保留数据的原始特征。在金融领域，尤其是股票价格分析中，这些特征——如趋势的转折点、价格的峰值和谷值——往往蕴含着关键信息，它们是投资者判断市场走向的重要依据，也是交易决策的核心参考。而 Savgol 滤波器在处理这些数据时，表现得尤为出色。与简单的移动平均等滤波器不同，Savgol 滤波器并不只是简单地“磨平”数据。它在去除噪声的同时，尽可能地保留了更多的原始信息。这种能力使得它能够减少或消除那些与数据本质无关的干扰，从而降低模型的复杂性。简而言之，它让数据变得更加清晰、简洁，同时也让基于这些数据构建的模型更加易于理解和解释。

图 3-4 是经过 Savgol 处理前后的四支股票部分收盘价序列对比，可以看到 Savgol 滤波器能在去噪基础之上保留了原始数据的趋势和特征。

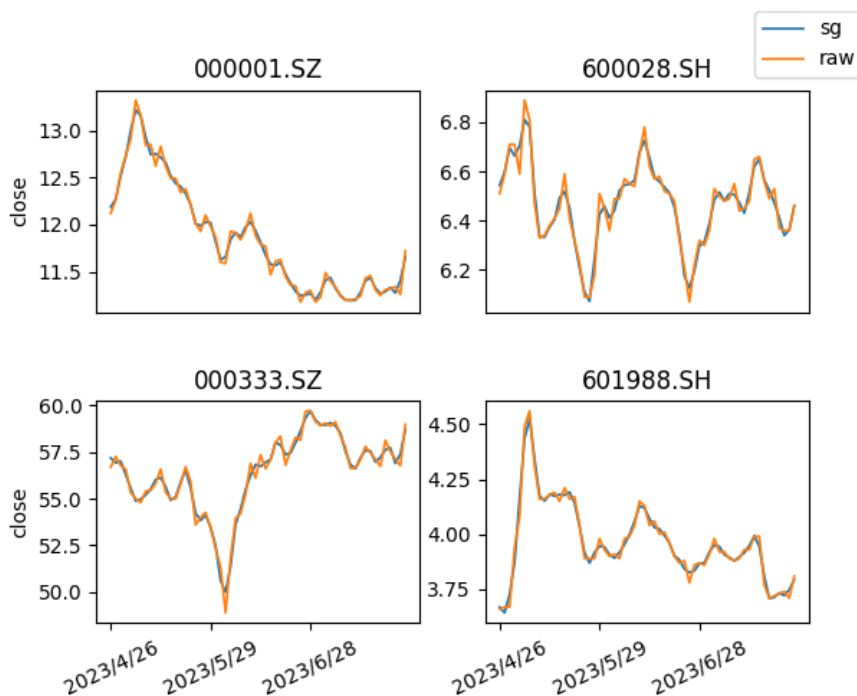


图 3-4 Savgol 前后的收盘价对比图

（4）归一化处理

为了避免数据的尺度问题对模型训练和预测的影响，使得模型训练更加稳定，提高模型的精度和收敛速度，在预测前，将采取公式中 *Min-Max* 归一化方式将数据映射到 $[0, 1]$ 区间，相关的公式为：

$$X_{norm} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (3.1)$$

其中， X 是原始数据， $X_{\min}$ 和 $X_{\max}$ 分别是数据的最小值和最大值。

数据收集完成后，以下小节将对实证结果进行讨论。

3.3 预测结果及分析

3.3.1 基准模型结果分析

在探究本文提出的混合模型之前，我们首先建立了三个基准模型（SVR、LSTM 和 BiLSTM-Attention）的未经滤波处理的预测性能基线。为系统验证天气因素对股市的影响机制，研究设计了对照实验，比较各模型在引入天气特征前后的预测精度变化。实验采用 MSE 、 $RMSE$ 、 $RRMSE$ 、 R^2 四个评价指标来衡量预测精度。限于篇幅，本章以中国石化（600028.SH）作为典型案例展示详细分析结果。

表 3-4 展示了不同模型的预测结果，可以看出，传统机器学习模型 SVR 在金融时间序列预测方面表现欠佳，其较高的误差（ MSE 、 $RMSE$ 、 $RRMSE$ ）以及较低的拟合优度（ R^2 ）表明该模型存在明显的改进空间。相比之下，基于深度学习算法的 LSTM 和 BiLSTM_Att 模型在所有评估指标上均显著优于 SVR 模型，例如，LSTM_close 和 BiLSTM_Att_close 的预测误差 MSE 显著降低，模型拟合优度 R^2 明显提升。值得注意的是，即使引入天气变量后，这种性能差距依然保持。可以得出，深度学习模型在误差方面相比机器学习模型有显著降低，拟合优度也有明显提升，表明基于深度学习算法的模型预测效果优于机器学习模型。而深度学习的优势源于其处理复杂非线性关系的能力：一方面，股价波动不仅受历史价格影响，还涉及经济因素、政策变化和市场情绪等多维度的非线性相互作用；另一方面，深度学习特有的特征重构机制能够更好地捕捉这些复杂关系，这正是传统机器学习方法的局限性所在。因此，采用深度学习框架进行残差修正和特征重构，对提升股价预测精度具有重要的实践价值。

另外，引入天气变量后，各模型的预测性能均获得提升，SVR_close_wea、LSTM_close_wea 和 BiLSTM_Att_close_wea 的 MSE 值较未加入天气变量的基础模型（SVR_close、LSTM_close 和 BiLSTM_Att_close 模型）均有不同程度降低。

这一现象表明天气因素确实包含影响股价波动的有效信息，其纳入能够增强模型对股价时序特征的捕捉能力，从而提升预测准确性。特别是在深度学习框架下，天气变量与神经网络的特征提取机制形成有效互补，进一步优化了模型对复杂市场动态的建模效果。

表 3-4 未平滑模型的预测结果

Model	<i>RMSE</i>	<i>MSE</i>	<i>RRMSE</i>	R^2
SVR_close	1.351882	1.827586	0.628774	-1.00714
SVR_close_wea	1.351734	1.827186	0.628685	-1.00712
LSTM_close	0.133435	0.017805	0.055930	0.967580
LSTM_close_wea	0.129773	0.016841	0.055823	0.968647
BiLSTM-Att_close	0.179998	0.032399	0.07406	0.932535
BiLSTM-Att_close_wea	0.168462	0.02838	0.071789	0.948028

3.3.2 本章模型结果分析

本节通过实证研究验证了 Savgol-BiLSTM-Att 混合模型的优越性能。基于真实市场数据与天气数据，该模型首先采用 Savitzky-Golay 滤波器对原始时序进行平滑处理，继而通过 BiLSTM-Attention 架构进行预测建模。如表 3-5 所示，通过系统比较平滑处理前后的预测效果，结果表明：经 Savgol 预处理的模型在预测精度上显著优于未经处理的基准模型，其预测结果与真实股价展现出更高的一致性。

表 3-5 平滑后模型的预测结果

Model	<i>RMSE</i>	<i>MSE</i>	<i>RRMSE</i>	R^2
Savgol-SVR_close_wea	1.335957	1.784781	0.621901	-1.08908
Savgol-LSTM_close_wea	0.127175	0.016173	0.053521	0.971329
Savgol-BiLSTM-Att_close_wea	0.116541	0.013582	0.04961	0.976389

图 3-5 展示了不同模型下不同输入情况下的各评价指标对比，由图可知，各模型的评价指标值在加入天气变量、加入天气变量及进行 Savgol 平滑处理后依次递减。Savgol-BiLSTM-Att_close_wea 模型在 *RMSE*、*MSE* 等指标上表现最佳，并拥有最高的 R^2 值。

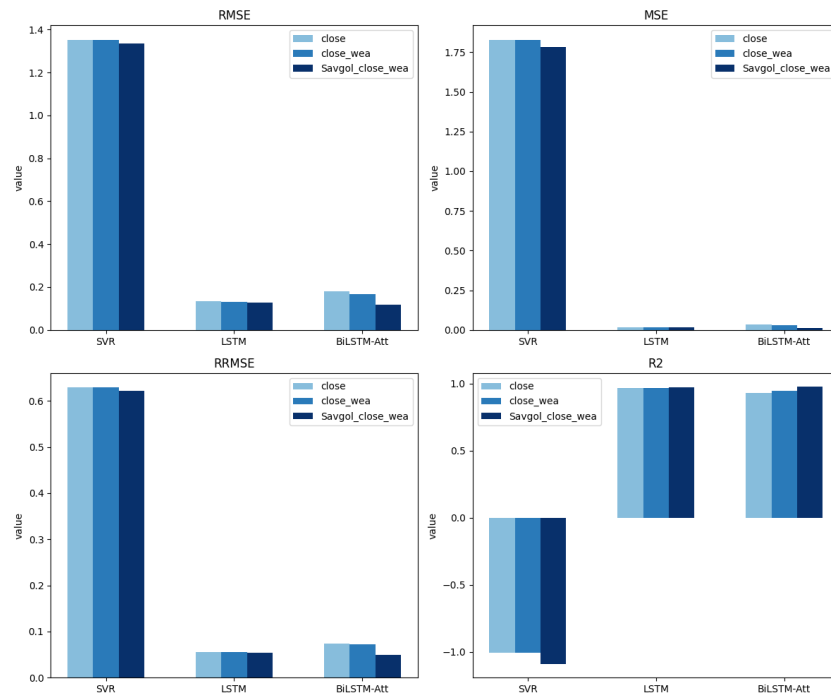


图 3-5 不同模型的评价指标对比

SVR 的模型仅能大致捕捉时间序列的总体趋势，残差较大，波动性高，方向不稳定，预测结果并不理想，所以图 3-6 仅展示除 SVR 外的 6 种模型拟合结果。相比之下，LSTM 等深度学习方法不仅在趋势上有所改善，而且在准确性上也更佳。此外，由于金融数据的噪声和非平稳性，单一的机器学习或深度学习模型未能达到最佳拟合效果。通过 Savitzky-Golay 对数据进行预处理并引入注意力机制，能显著提升模型性能。

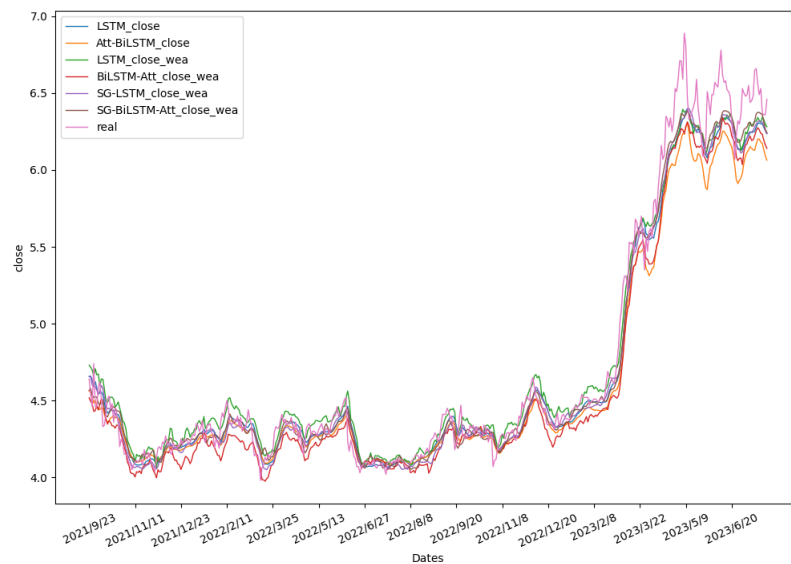


图 3-6 预测结果图

3.3.3 Kolmogorov-Smirnov 检验

基于前文实验结果，天气因素、注意力机制和 Savitzky-Golay 滤波对模型预测性能的提升效果已得到初步验证。例如，SG-BiLSTM-Att_close_wea 模型的 $RMSE$ 值较 LSTM_close 模型减少了 0.016894， MSE 值减少了 0.004223， R^2 值从 0.967586 增至 0.976389，表现出性能提升。为深入分析模型间的统计显著性差异，本研究采用两样本 Kolmogorov-Smirnov 检验对模型残差序列进行对比分析。K-S 检验零假设(H_0)为两模型残差序列分布相同， p 值小于 0.05 则拒绝 H_0 ，表明两序列分布不同，即模型间存在显著差异。

表 3-6 模型差异的 K-S 检验

	BiLSTM- Att_close	LSTM	BiLSTM-Att	Savgol- SVR	Savgol- LSTM	Savgol- BiLSTM-Att
BiLSTM-Att_close	1					
LSTM	0.000	1				
BiLSTM-Att	0.000	0.000	1			
Savgol-SVR	0.000	0.000	0.000	1		
Savgol-LSTM	0.001	0.000	0.000	0.000	1	
Savgol-BiLSTM-Att	0.000	0.000	0.000	0.000	0.013763	1

表 3-6 显示了五种模型通过 K-S 检验的 p 值差异。所有模型自身对比的 p 值为 1，而 Savgol_LSTM_close_wea 与 Savgol-BiLSTM-Att_close_wea 之间的 p 值为 0.013763，其他模型间的 p 值近似为 0。所有检验 p 值均小于 0.05，拒绝零假设，表明不同模型的残差序列分布存在显著差异。

Savgol-SVR_close_wea、Savgol-LSTM_close_wea 和 Savgol-BiLSTM-Att_close_wea 模型之间的差异性指出机器学习与深度学习模型在性能上的显著区别，深度学习模型较传统 SVR 具有显著优势。此外，BiLSTM-Att_close 与 BiLSTM-Att_close_wea 模型间的显著差异证实了天气因素作为有效特征具有统计显著性。

3.4 量化投资应用分析

3.4.1 TopkDropout 策略的构建

模型的预测性能是应用的保证,模型的应用则是预测的最终目的。本节基于 Savgol-BiLSTM-Att_close_wea 模型的预测结果,设计了 TopkDropout 动态组合策略,并通过评价指标评估其应用价值。该策略首先根据模型预测结果筛选出 topk 只最具潜力的股票构成候选组合,然后通过 Dropout 机制随机剔除部分股票以控制风险暴露,最终形成优化的投资组合。这种设计既保留了高潜力股票的优势,又通过随机剔除实现了风险分散,有效平衡了收益与风险的关系。具体的投资策略构造过程如下:

(1) 基于已训练好的 Savgol-BiLSTM-Att_close_wea 模型预测得到的 30 只股票的收盘价格序列,计算得到每日的预测对数收益率序列,第*i*天股票的预测对数收益率公式为:

$$\hat{R}_i = \ln\left(\frac{p_i}{p_{i-1}}\right) \quad (3.2)$$

其中, p_i 为在股票第*i*天的预测收盘价。

(2) 采用定量调仓机制,构建 TopkDropout 动态组合策略,其核心操作流程为:每日维持固定持仓规模(topk=10),根据测试集对数收益率排序,卖出组合内收益表现最差的 n_drop 只股票(n_drop=3),同时从股票池中买入未持仓股票中收益预期最高的 n_drop 只股票。该策略通过定期淘汰弱势标的、纳入强势标的的机制,实现了组合的持续优化。

(3) 在考虑交易手续费的情况下,模拟 TopkDropout 策略交易。并将在该种策略下的评价指标与沪深300指数进行比较,结合投资组合的 Alpha 和 Beta 值,以确定这种策略是否能产生超额回报。手续费设为 0.001。

3.4.2 回测结果分析

本节使用沪深 300 指数构建简单购买持有策略,并以此作为基准参照,验证 TopkDropout 策略的有效性,投资评价指标的具体计算结果如表 3-7 所示。

表 3-7 列示了 TopkDropout 策略和沪深 300 指数的投资指标结果,可以发现:从年化收益率来看,TopkDropout 策略在考察期内表现优于沪深 300 指数,实现

了正收益，而沪深 300 指数则出现了负收益；同时，TopkDropout 策略的夏普比率超出简单购买持有沪深 300 指数的 0.940078，表明在考虑风险后，TopkDropout 策略提供了正的风险调整后回报。尽管 TopkDropout 策略波动性较大，但其超额收益最大回撤在 8% 左右，而沪深 300 指数的最大回撤达到了 30.9827%，TopkDropout 策略的最大回撤远小于沪深 300 指数，表明其在最坏情况下的潜在损失更小。从 α 和 β 值来看，正的 α 值表明 TopkDropout 策略在调整市场风险（ β ）后，仍能获取超额收益。总而言之，基于预测模型构建的 TopkDropout 策略在收益能力、风险控制和超额收益三个方面均展现出显著优势，验证了模型的实际应用价值。

表 3-7 TopkDropout 策略评价结果

投资评价指标	TopkDropout 策略	沪深 300 指数
年化收益率	0.070903	-0.115015
年化波动率	0.375652	0.175333
最大回撤	-0.089694	-0.309827
夏普比率	0.129139	-0.810939
α	0.000731	
β	0.947531	

3.5 本章小结

本章创新性地提出了 Savgol-BiLSTM-Att 混合预测模型，该模型整合了 Savitzky-Golay 滤波器、深度学习网络和注意力机制三大技术模块，并首次系统性地将天气因素考虑在股价预测中进行研究。模型构建分为三个关键阶段：（1）数据预处理阶段：对天气序列进行去季节化处理，并采用 Savgol 滤波消除噪声干扰；（2）特征提取阶段：通过 BiLSTM 网络双向捕捉时序特征，并利用注意力机制动态分配特征权重；（3）预测输出阶段：通过全连接层生成最终预测结果，并评估模型的预测性能。基于该模型，设计了 TopkDropout 动态投资策略以 Savgol-BiLSTM-Att 模型在实际股市中的应用有效性。

在模型性能评估方面，考虑了天气指标的 Savgol-BiLSTM-Att 模型在四个评价指标上均为最优，值分别为 0.116541、0.013582、0.04961、0.976389。结果显

示，股票市场确实存在天气效应，天气指标的加入可以提高模型的预测精度；Savgol 滤波的加入去除了原始数据的噪声、毛刺等，能大大降低模型的预测误差。

在预测模型的投资应用价值方面，采用 TopkDropout 投资策略进行了实证测试。结果显示 Savgol-BiLSTM-Att 模型在降低误差和提高投资效能方面表现出色，相较于买入持有沪深 300 指数策略，实现了更高的收益和更低的风险，证明了其在量化投资领域的优越应用价值。因此，本研究的 Savgol-BiLSTM-Att 模型不仅在技术层面上创新，也在投资实践中展现了显著优势。

第4章 基于网络舆情与 FMD-QRBiGRU 模型的股价区间预测与量化投资

第3章围绕金融市场的天气效应展开研究,揭示了天气变化通过影响投资者情绪,进一步作用于股价波动及投资者决策的传导路径,即“天气—投资者情绪—投资者行为—股票市场”的影响机制。在这一机制中,投资者的心理状态作为投资者情绪的重要且合理代理变量,扮演着关键角色。而除了天气效应外,网络舆情作为市场情绪的重要载体,是影响投资者心理预期与决策行为的关键因素之一。它不仅能够改善金融市场中的信息不对称状况,其包含的情绪倾向和市场传播效应还会显著改变投资者的风险偏好,从而引导其交易行为并引发股价变动。然而,与可直接量化的财务指标不同,网络舆情是一个复杂且抽象的概念^[88],为系统考察其对金融市场的影响,本章对网络舆情的情感和影响力这两个关键属性进行了定量测度,以期更精准、全面地理解网络舆情对金融市场的作用机制。

(1) 网络舆情情感

在信息技术快速发展的当下,互联网文本数据因其及时性和丰富性,已成为捕捉市场情绪的重要载体。相较于传统的金融指标或问卷调查指标,基于文本的情绪分析能够更直接地刻画投资者的心理状态。因此,运用文本挖掘技术构建市场情绪指标,已成为当前金融研究的重要方向^[89]。大量实证研究证实,通过文本分析提取的市场情绪指标对股票收益具有显著的解释力。例如, Kim 等人(2023)从《纽约时报》中提取“摘要”文本数据,应用 FinBERT 模型对其进行情绪分析,计算情绪得分,并结合 LSTM 模型,对标普 500 指数进行股价预测。实证结果表明,融入情绪得分的 LSTM 模型在误差指标上表现最优,不仅证实了将新闻情绪纳入股价预测的重要性,也揭示了心理因素对金融市场的潜在影响。此外,该研究还论证了使用摘要数据而非全文本新闻进行分析的合理性^[90]。

在信息技术迅猛发展的浪潮推动下,投资者获取有效信息的渠道被不断拓宽。股吧作为网络上极具代表性的金融社交媒体平台,已然成为投资者交流观点、抒发情感与分享知识的重要阵地,同时也为研究投资者情绪提供了关键的数据源^[91]。与传统调查问卷等信息获取渠道相比,股吧评论数据凭借其庞大的用户基

数、高度的社交互动性以及实时性强等显著优势，使得基于此构建的投资者预期指标具备更强的有效性与更明确的指向性。

表 4-1 文本挖掘度量方式

数据来源	转换方式	代表性文献
东方财富网股吧评论文本	神经网络	赵传军等（2022） ^[94]
微博文本	BERT 分类模型	崔炎炎等（2022） ^[95]
《纽约时报》中提取“摘要” 文本数据	FinBERT 得分模型	Kim 等（2022） ^[90]

注：作者根据相关文献整理所得

因此，众多研究者通过文本挖掘与深度学习相结合的方法，对百度贴吧、微博以及东方财富股吧等网络平台中的投资者发帖内容进行情感分析（见表 4-1）。这种方法旨在建立科学有效的投资者情绪量化指标体系，从而深入挖掘网络舆情数据中蕴含的信息价值。例如，熊艳（2022）基于 2013-2017 年雪球和股吧论坛对的发帖数据，从信息传递与情绪宣泄双重视角研究了社交媒体对股价行为的影响，发现股吧论坛的情绪宣泄效应在我国股市占据主导地位，且在创业板表现最为显著。该研究证实了投资者情绪通过社交媒体渠道对资本市场产生的系统性影响，股票价格的波动与情绪指数之间存在显著的相关性^[92]。范小云等（2022）使用机器学习方法研究了中文金融文本情绪对宏观经济和金融市场的预测能力，发现新闻情绪能有效预测宏观经济指标，而股吧情绪则对股票市场的收益率、交易量和波动性具有显著预测作用。该研究表明投资者情绪通过不同文本渠道对资本市场产生差异化影响，其中社交媒体情绪对市场波动的预测效果尤为突出^[93]。上述研究均表明，股吧情绪不仅显著影响投资者的情绪与决策行为，还可以应用于市场趋势预测。

（2）网络舆情影响力

投资者的注意力是每只股票的稀缺资源，由于时间和精力限制，投资者实际上只能专注于有限的信息，而能够吸引投资者关注的事件更容易进入其决策系统。行为金融研究证明，投资者注意力作为认知、决策过程的重要组成部分，在投资决策中占据重要地位。特别是在市场效率较低的新兴市场，投资者的注意力在信息传播和解释中发挥着更重要的作用，并可能导致股价的非理性和异常波

动。例如, Gang 等人(2022)通过构建聚合投资者注意力指标 RF 和 GBDT, 在上交所 50 只成份股的样本中进行实证分析。结果表明, 注意力指数的构建在样本内和样本外都提供了具有经济意义的股票收益, 提高了模型的预测能力。此外, 注意力指数与收益波动性负相关。结果证明了投资者注意力代理的有效性, 并揭示了这些代理在资产定价中的出色预测能力^[96]。Barber 和 Odean(2008)研究了注意力驱动交易行为, 发现个体投资者因搜索成本限制更倾向于买入高关注度股票(如新闻热点股或高波动股)。其提出的“注意力-偏好”双阶段决策模型揭示了有限注意力对投资者行为的系统性影响, 认为注意力先限定选择范围, 偏好再决定最终选择^[97]。由此可以得出, 网络舆情能够通过吸引投资者注意力影响股价波动, 资产配置行为也会因为投资者关注度发生系统性调整, 这种变化最终将反映在金融市场整体价格形成过程中。

综上所述, 考虑到中国证券市场投资者结构呈现鲜明的散户化特征, 专业金融知识相对匮乏且投机倾向显著, 投资者群体表现出明显的从众心理和行为偏差, 其投资决策容易受到市场舆论的过度影响, 非理性交易行为频发。因此, 本章将以股票市场的特征数据和网络舆情资料为研究对象, 将网络舆情信息量化、构建成有效的投资者情绪因子和投资者关注度因子, 结合 FMD 特征模态分解技术、以及 QR 分位数回归方法, 构建深度学习预测模型来预测股价趋势; 并通过区间约束的交易策略进一步论证所构建预测模型的相对优越性。

4.1 预测模型构建

4.1.1 FMD-QRBiGRU 预测模型构建

本章采用 BiGRU 模型作为预测的基础模型, 同时引入 FMD 特征模态分解算法和 QR 分位数回归技术, 构建一种基于网络舆情的 FMD-QRBiGRU 的区间预测模型, 并应用于股票价格预测, 进而构建股票投资策略组合进行量化交易回测。研究使用了五个基准模型, 即 SVR、BPNN、LSTM、AtBiLSTM 和 BiGRU。研究框架由四个部分组成: 数据收集、数据处理、实证预测和交易模拟, 如图 4-1 所示。

第一阶段, 数据收集。本文收集三类核心数据: 股票日度收盘价数据(预测目标变量); 开盘价、成交量等市场数据(传统交易特征); 东方财富网股吧论坛的投资者评论文本(网络舆情数据)。

第二阶段，数据处理。（1）文本情感分析。对股吧文本进行情绪分析，以提取投资者情绪因子。（2）关注度量化。构建多维投资者关注度指标体系。（3）信号分解。使用 FMD 算法对股票收盘价进行多尺度特征分解。详细说明放在“数据处理”一节中。

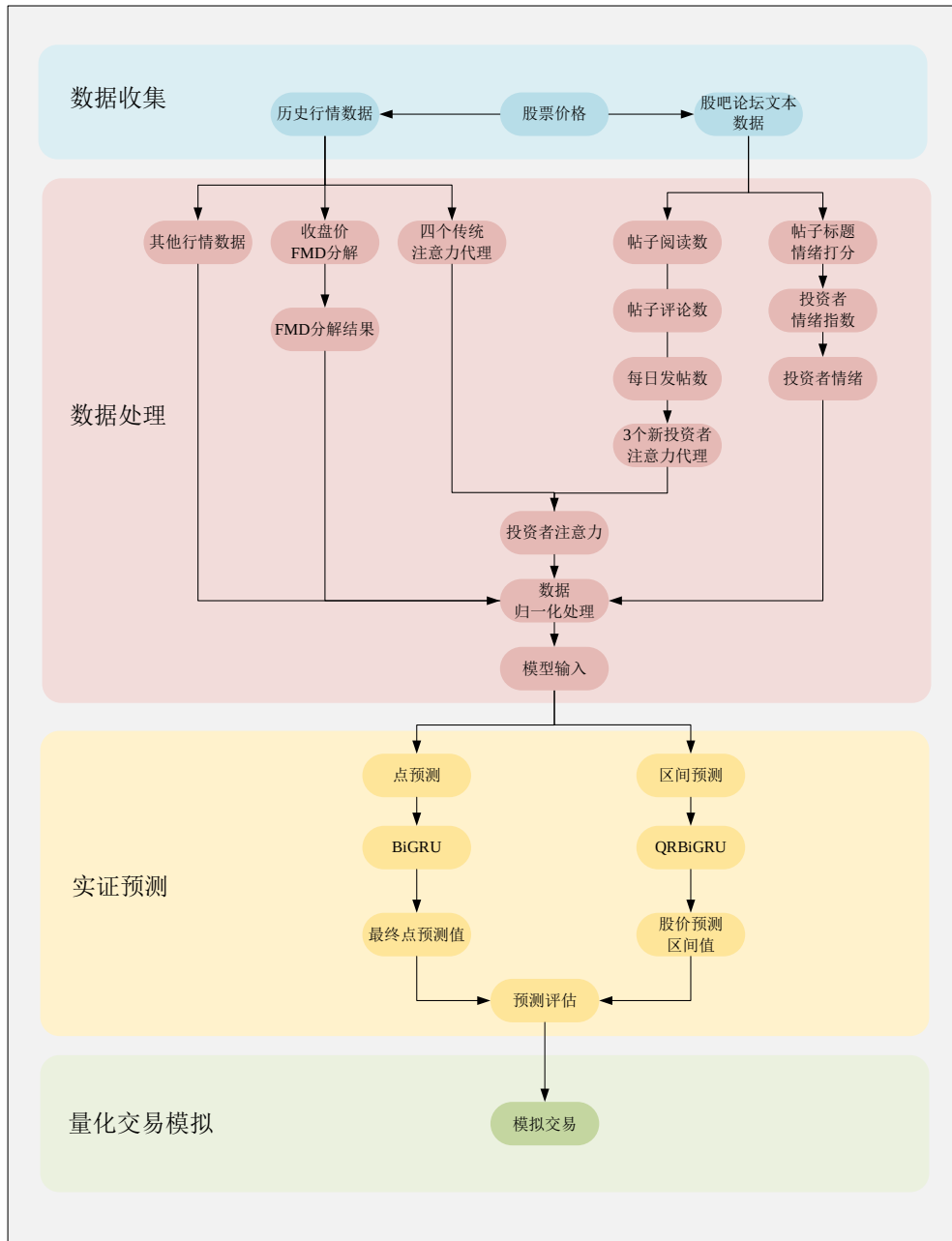


图 4-1 本节研究框架

第三阶段，实证预测。在深度学习模型的支持下，使用历史行情数据、情绪指数因子、投资者关注度因子、以及 FMD 分解结果来预测股价，包括点预测和区间预测，并通过误差评估指标和统计检验来评估模型的预测能力。

FMD-BiGRU 点预测模型、FMD-QRBIGRU 区间预测模型基本框架如图 4-2、图 4-3 所示。

第四阶段，量化交易模拟。结合点预测和区间预测的结果，本文构建基于区间约束的投资策略，通过保险建议来指导投资者的决策，以获得更大收益。

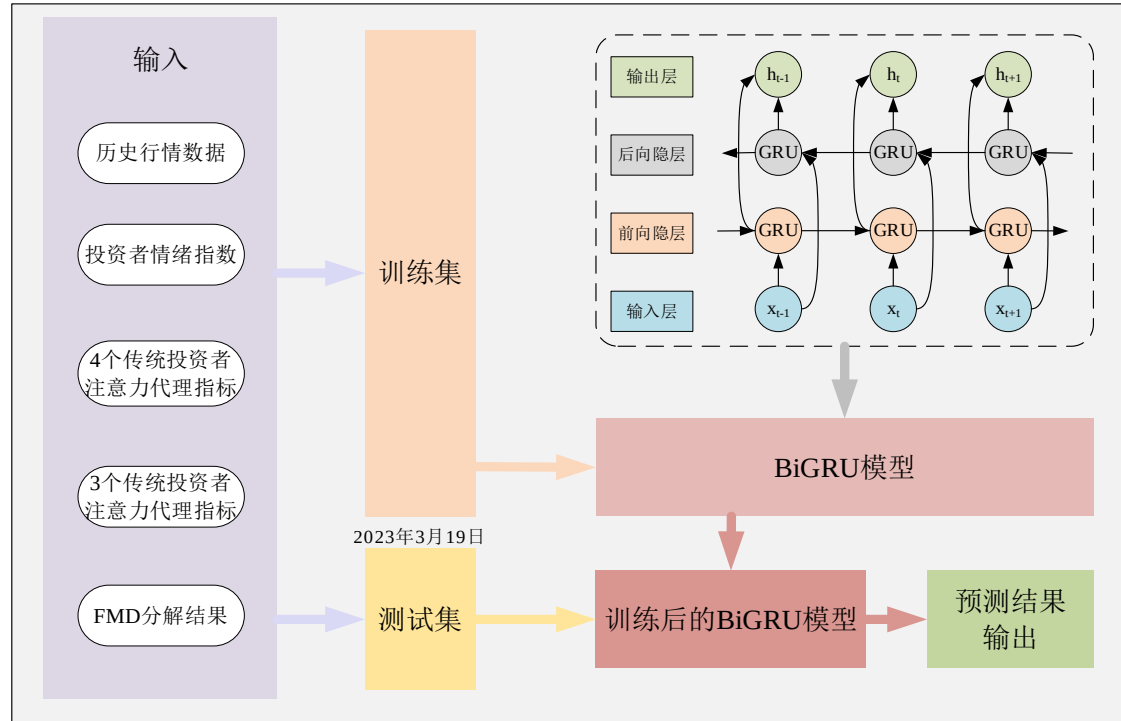


图 4-2 FMD-BiGRU 点预测模型的基本框架

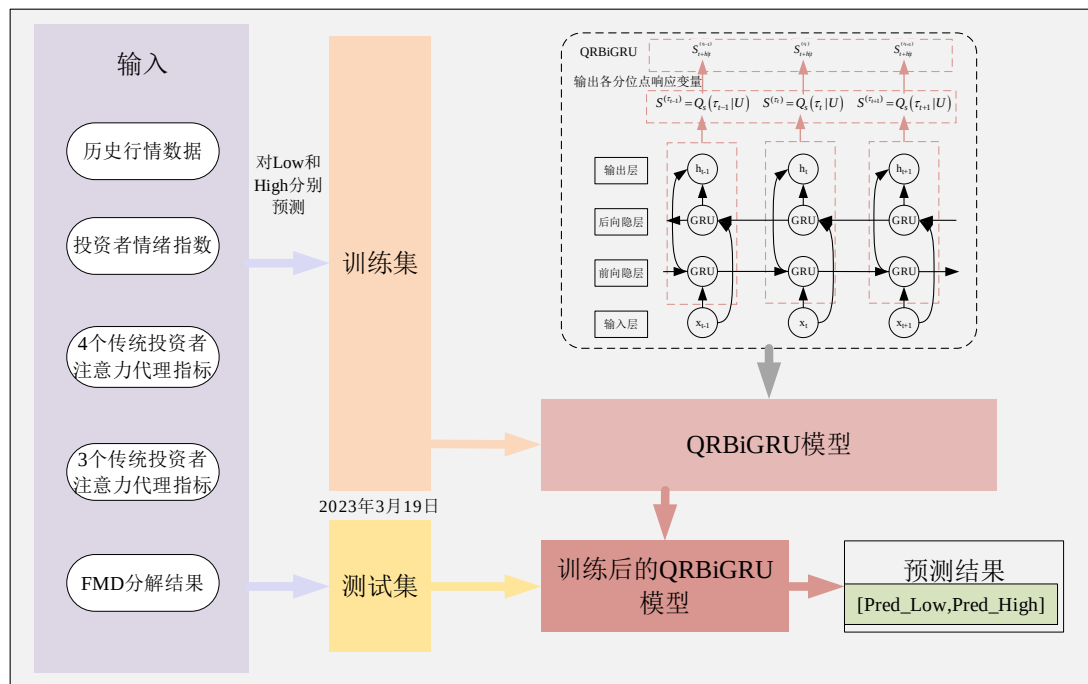


图 4-3 FMD-QRBIGRU 区间预测模型的基本框架

4.1.2 预测结果评估

本章采用 MAE 、 $RMSE$ 、 $MAPE$ 、 MSE 、 $RRMSE$ 、 ARV 、 $SMAPE$ 、 R^2 这 8 种评价指标对点预测结果进行综合评价。引入 $PICP$ 和 $PINAW$ 两个评价指标对预测区间的覆盖率和宽度作出评判。同时，采用两样本 K-S 检验，从统计学角度评估模型的预测能力。

4.2 实证分析

在本节中，研究对所建立模型的综合性能进行验证，实证研究直观地展示了所提出的模型的模拟过程并设置一些对比模型，通过预测结果的比较分析突出了所设计模型的优秀预测能力。最后将模型预测类比为真实的投资场景，从多个角度对 FMD-QRBiGRU 预测模型和其他循环神经网络及变体模型在 A 股市场上的应用效果进行实证分析。限于篇幅，本节主要展示 601398.SH 工商银行的实验结果。

4.2.1 数据的获取

首先，为了研究投资者情绪、投资者注意力以及 FMD 分解对我国不同市场回报的预测性，本文参照申万证券一级行业分级标准，综合市场流动性与产业分类，选取医药生物、非银金融、计算机、电子、银行和农林牧渔共 6 个行业市场作为研究对象。从产业分类来看，农林牧渔、电子板块分属一、二产业，医药生物板块在第一、二产业中均有占比，剩下的非银金融、计算机、银行板块属于第三产业。从流动性的角度来看，农林牧渔、电子、银行的流通性相对较弱，非银金融、计算机、医药生物的流动性较强。本章从这 6 个行业板块中分别选取各自板块中总市值最大的 5 只共 30 只股票进行实证分析。

样本区间设置为 2019 年 2 月 20 日至 2024 年 3 月 26 日。对于每个数据集，研究将交易日 80% 的数据作为训练集，剩余 20% 为测试集。也就是说，以 2023 年 3 月 17 日为界，使用 2019 年 1 月 2 日至 2023 年 3 月 17 日的数据为训练集进行模型训练，测试集使用 2023 年 3 月 20 日至 2024 年 3 月 26 日的数据来进行实证测试，数据来源为 wind 数据库。

本章同样选择股票下一个交易日的收盘价作为预测对象。601398.SH 工商银行从 2019 年至 2024 年期间的收盘价趋势如图 4-4 所示。

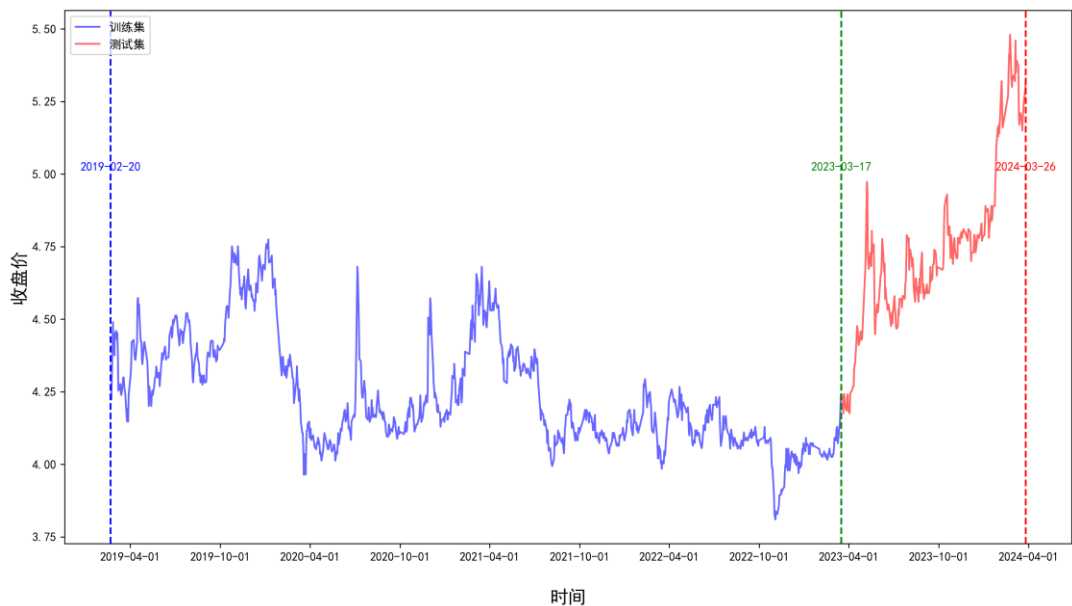


图 4-4 601398.SH 收盘价趋势图及数据集拆分

4.2.2 特征选取及构建

表 4-2 模型输入特征变量

数据分类	指标类别	指标名称	符号
股票行情数据	股票行情	开盘价	Open
		收盘价	Close
		最高价	High
		最低价	Low
		成交额	Turnover
		成交量	Volume
网络舆情数据	投资者情绪因子	情绪因子	Sentiment
	投资者注意力因子	极端收益	$E\ Ret$
		异常交易量	$Avol$
		过去收益	$PCRet$
		接近 30 天高点	$30dH$
		发帖数	$Count$
		阅读量	$Read$
		评论数	Com
FMD 分解数据	收盘价 FMD 分解数据	各模态分量	TMF1...IMFn

本章共设置三类输入特征，分别为股票行情数据、网络舆情数据、FMD 分解数据，其中股票行情数据包含 6 种类型，网络舆情数据包括投资者情绪因子、投资者注意力因子两种类型，FMD 分解数据是对收盘价进行 FMD 分解得到的若干模态分量。输入特征变量具体如表 4-2 所示：

(1) 股票行情数据包括每日的开盘价、收盘价、最低价、最高价、成交额、成交量这六个数据，数据从 wind 数据库获取。股票行情数据本身就能在一定程度上预测未来股价的走势，所以本章将其作为股价预测模型的部分输入特征。以 601398.SH 工商银行为例，表 4-3 展示了其部分数据及描述性统计结果。

表 4-3 描述性统计表（601398.SH）

Date	Open	High	Low	Close	Volume	Turnover
2019/2/20	4.2462	4.2614	4.2234	4.2386	106861146	595932400
2019/2/21	4.2386	4.2690	4.2158	4.2310	158310223	881970592
...	...	...	...	...	...	...
2024/3/25	5.25	5.29	5.24	5.29	237408100	1252612480
2024/3/26	5.29	5.33	5.26	5.33	243359624	1289061744
mean	4.3351	4.3659	4.3081	4.3386	220713481.5	1103632766
std	0.2895	0.2988	0.2829	0.2915	117821511.3	619973179.4
median	4.2386	4.2744	4.2108	4.2416	193958079	955282784
min	3.801	3.8099	3.7630	3.8099	55628021	276623120
max	5.48	5.52	5.41	5.48	1245663648	6850764288

(2) 投资者情绪因子的测度

本章选取的网络舆情数据来源于东方财富股吧论坛（<http://guba.eastmoney.com>）的投资者交流贴等文本数据，采用 Python 编程语言构建网络爬虫系统获取，具体实现过程如下：首先基于 Requests 库实现 HTTP 请求发送与响应获取功能；其次为规避网站反爬机制，采用动态 UserAgent 技术随机生成请求头信息，模拟真实用户的网页浏览行为；最终将采集到的交流贴信息，包括用户 ID、用户名、发帖时间、阅读数、评论数和标题文本等，存储至本地数据库。部分数据如表 4-4 所示，选择的发帖时间范围为 2019 年 1 月 2 日至 2024 年 3 月 26 日，剔除非交易日数据，约 9590550 条中文评论。

表 4-4 部分股吧文本数据及情绪得分（601398.SH）

用户 ID	用户名	标题文本	发帖时间	情绪得分
3168625222999740	抄底专业户必胜哥	真佩服自己！！！！太让我失望了	2024-03-26 14:13:17	0.0489
6556094395598280	希望全在股市	哪里跌的多，哪里就有我	2023-11-22 09:39:52	0.2273
1569054917932042	财源大展的段正淳	关注这只低价的肥羊	2023-11-03 10:01:47	0.7395
6808355787335260	贼胖儿	运气不错,[买入][买入][买入]	2023-07-03 09:44:27	0.9907
5406365716656368	要招财要进宝呀	比利息高就行	2023-06-30 14:20:36	0.6557
5601305857441036	慢积累	工行成五大行里最烂的了？	2022-04-13 10:34:16	0.0660
6331395838193322	股友 OkWUbq	垃圾分类，工行属于有毒垃圾	2022-04-01 11:28:42	0.0635

对于投资者情绪，本文使用 SnowNLP 对文本数据进行情绪分析。具体来说，SnowNLP 用于确定每天的各条用户评论文本的情绪得分，为了确保数据的可用性，将相同日期的情绪得分取平均，得到每日的投资者情绪指数数值，每日的情绪得分构成投资者情绪指数时间序列数据，计算公式如下。

$$Sent_{i,t} = \frac{1}{Count_{i,t}} \sum_{n=1}^{Count_{i,t}} S_t^n \quad (4.1)$$

其中， $Sent_{i,t}$ 为第 i 只股票第 t 天的投资者情绪指数， S_t^n 为第 t 天第 n 条文本的情绪得分， $Count_{i,t}$ 为第 i 只股票第 t 天的帖子数量。每日的情绪得分构成情绪因子时间序列数据。表 4-4 展示了 601398.SH 部分文本信息及情绪得分。表 4-5 给出部分公司的情绪因子数据，并绘制情绪因子曲线，如图 4-5。

表 4-5 部分公司情绪因子数据

日期	601398.SH 工商银行	002415.SZ 海康威视	...	600030. 中信证券	000876 新希望
2019/1/2	0.976993076	0.538467319	...	0.503023593	0.488585129
2019/1/3	0.829569391	0.570948711	...	0.597571514	0.576853624
2019/1/4	0.682145706	0.419603885	...	0.593874752	0.568044689
2019/1/7	0.369467051	0.484691839	...	0.582784466	0.591402259
2019/1/8	0.340207128	0.506387823	...	0.579087704	0.5064923
...	...	...	...	...	...
2024/3/20	0.437329443	0.41086098	...	0.533421783	0.486168348
2024/3/21	0.528000835	0.487291658	...	0.503166689	0.543897144
2024/3/22	0.39704047	0.429574293	...	0.516600884	0.45352823
2024/3/25	0.462836147	0.398921061	...	0.498172646	0.537963236
2024/3/26	0.370871856	0.409646708	...	0.553001439	0.469245148

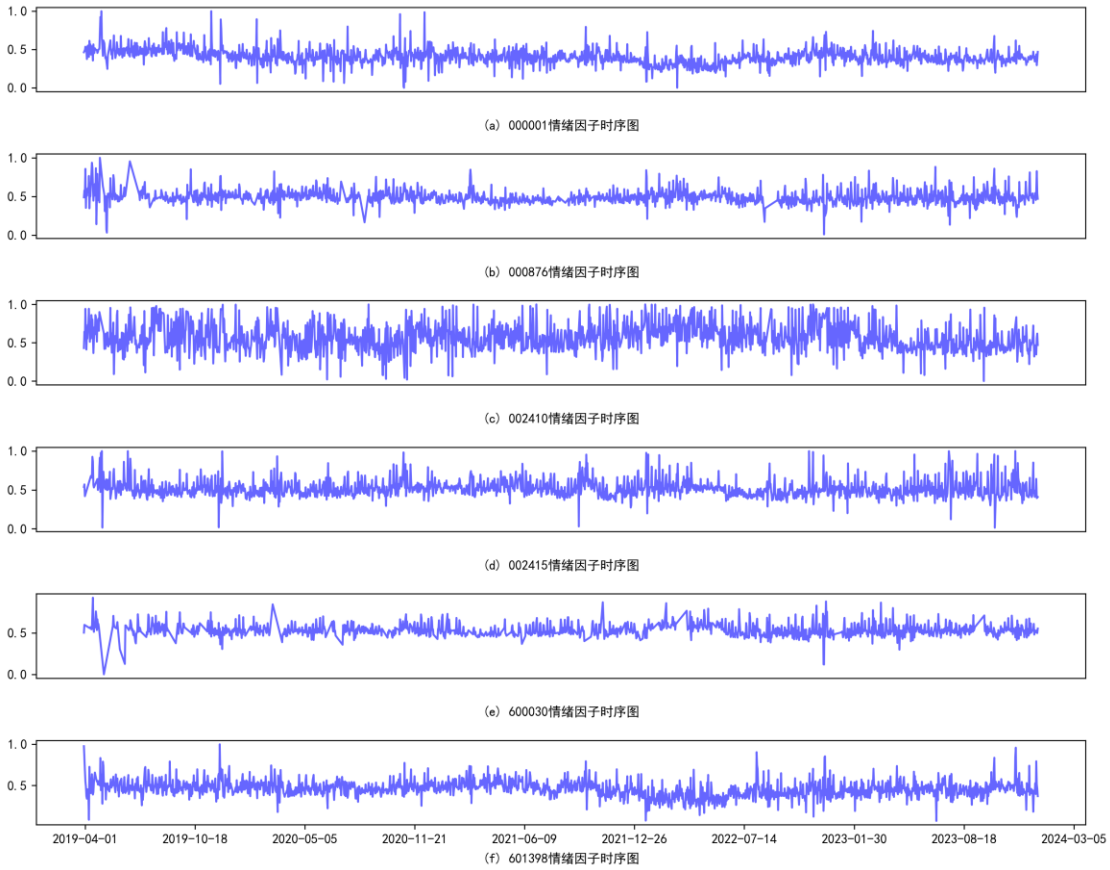


图 4-5 部分公司情绪因子时序图

(3) 投资者注意力因子的测度

投资者的关注度是通过不同的衡量指标来估计的，根据文献^[96]，本章首先计算了四个传统注意力代理。

第一个是极端收益（ $E Ret$ ），即股票的日回报率与前 30 天平均绝对回报率的比率。

$$ERet_{i,t} = \frac{Ret_{i,t}}{\frac{1}{30} \sum_{t=1}^{30} |Ret_{i,t-30}|} \quad (4.2)$$

其中， $Ret_{i,t}$ 为股票 i 在第 t 天的回报。

第二个代理变量是异常交易量（ $Avol$ ），即每只股票的每日交易量与前 30 天平均值的比率。

$$Avol_{i,t} = \frac{vol_{i,t}}{\frac{1}{30} \sum_{t=1}^{30} vol_{i,t-30}} \quad (4.3)$$

其中， $vol_{i,t}$ 是股票 i 在第 t 天的交易量。

第三个代理变量是过去收益（ $PCRet$ ），即每只股票在前 30 个交易日的每日累计收益。

$$PCRet_{i,t} = [\prod_{t=1}^{30} (1 + Ret_{i,t})] - 1 \quad (4.4)$$

第四个变量代理是接近 30 天高点（ $30dH$ ），即每只股票在第 t 天的当前价格与过去 30 天的最高价格之比。

$$30dH_{i,t} = \frac{P_{i,t}}{\max_{t=1}^{30} (P_{i,t-30})} \quad (4.5)$$

其中， $P_{i,t}$ 是股票 i 在第 t 天的收盘价。

除此之外，本文将股吧每日的发帖数，以及帖子的阅读量、评论数作为投资者关注度的补充代理变量，为新的投资者注意力代理。同样，将相同日期的阅读量、评论数取平均，得到每日的阅读量、评论数数值。具体计算公式如下。

$$Read_{i,t} = \frac{1}{Count_{i,t}} \sum_{n=1}^{Count_{i,t}} R_t^n \quad (4.6)$$

$$Com_{i,t} = \frac{1}{Count_{i,t}} \sum_{n=1}^{Count_{i,t}} C_t^n \quad (4.7)$$

其中， $Count_{i,t}$ 为第 i 只股票第 t 天的帖子数量， $Read_{i,t}$ 为第 i 只股票第 t 天的阅读量， $Com_{i,t}$ 为第 i 只股票第 t 天的评论数。 R_t^n 为第 t 天第 n 条文本的阅读量， C_t^n 为第 t 天第 n 条文本的评论数。根据以上公式计算出来的各指标前 30 个交易日没有值，数据从 2019 年 2 月 20 日开始，部分关注度因子数据如表 4-6 所示：

表 4-6 关注度因子部分数据（601398.SH）

日期	$E Ret$	$Avol$	$PCRet$	$30dH$	$Count$	$Read$	Com
2019/2/20	0	0.7486	0.0704	0.9755	25	5231.92	1.68
2019/2/21	-0.7849	1.0926	0.0686	0.9737	25	1964.52	1.4
2019/2/22	-1.0114	1.9789	0.0534	0.9720	78	1528.115	0.9487
2019/2/25	15.4935	3.9823	0.1221	1	20	3519.15	1.2
2019/2/26	-6.1372	2.5263	0.1077	0.9780	42	2977.4048	1.3810
...	...	...			...	...	...
2024/3/20	17.9159	0.7067	0.0163	0.9489	58	506.5862	3.0517
2024/3/21	9.3762	0.5959	0.0123	0.9526	30	564.4667	4.1
2024/3/22	7.2273	0.7747	0.0239	0.9580	57	508.3333	2.5088
2024/3/25	8.915200	0.7036	0.0256	0.9653	28	421.5357	9.3929
2024/3/26	22.7093	0.7488	0.0100	0.9726	81	408.9630	1.8642

4.2.3 数据处理

（1）FMD 特征模态分解

通过 FMD 对股票的收盘价进行分解，得到模态分量，并计算每个分量相关系数，最后将分解的模态分量作为新特征输入到预测模型中。以工商银行为例，分解后得到的模态分量如图 4-6 所示。FMD 分解可有效抑制分解过程中所产生的模态混叠问题，减少各模态分量间的信息泄露。

（2）数据归一化

例如特征当中的开盘价与成交额，不同特征之间的量纲并不相同，直接将原始数据作为模型输入特征进行预测会影响模型的预测效果。为了消除输入指标之

间量纲差异影响，同时提高模型的收敛效率与预测准确度，需对原始数据进行归一化处理。本章同样采用 *Min-Max* 归一化处理。

数据收集完成后，以下小节将对实验预测结果进行讨论。

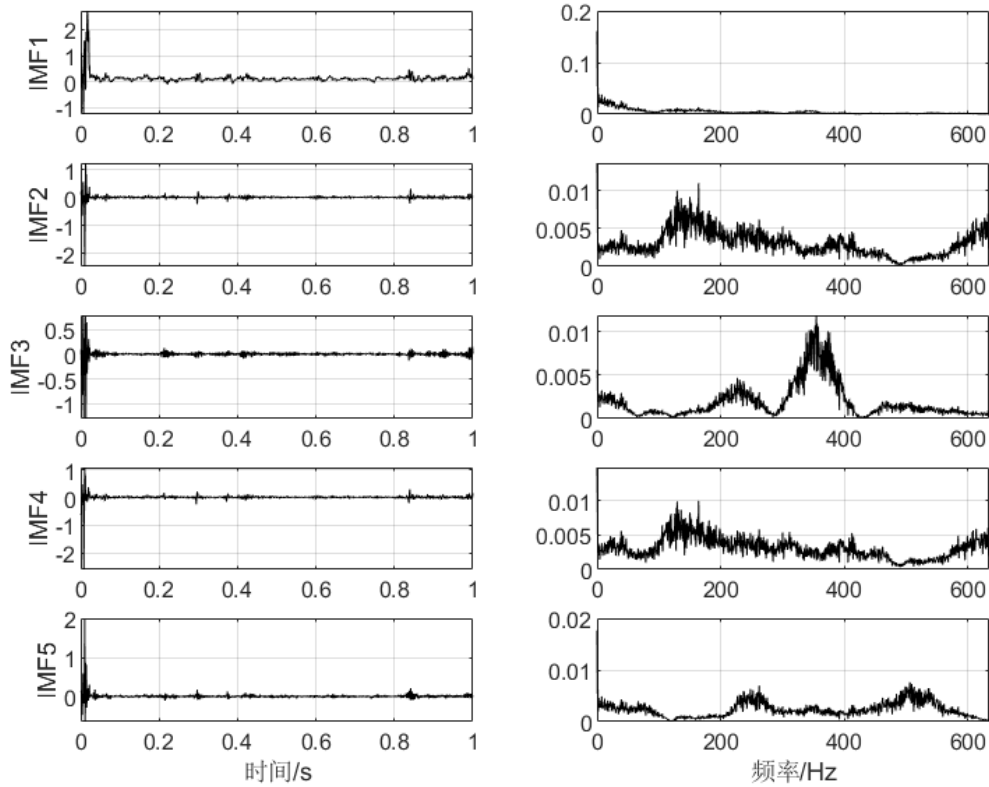


图 4-6 收盘价 FMD 分解所得各模态分量及功率谱

4.3 预测结果及分析

4.3.1 点预测结果分析

本节通过比较不同深度学习模型及其组合模型以验证所构建模型的有效性。设定模型的时间窗口为 30，利用单位交易日滑动窗口获得一系列输入特征与输出值。

通过上节特征选取及构建可知，深度学习模型输入特征共有三大类，因此本节设计 3 种预测模型特征输入方式，所有特征在输入模型进行预测之前均进行归一化处理。即将所有模型分为三部分，以每个股票交易日的股票基本行情数据为基准，构建基础预测模型（Model），在此基础上加入投资者情绪因子、投资者注意力因子构建预测模型（Model_X）以及融合 FMD 分解所得的各模态分量、投资者情绪因子、投资者注意力因子的预测模型（Model_X_Y），进行对比实验；

并从 $RMSE$ 、 MSE 、 MAE 、 $MAPE$ 、 R^2 、 $RRMSE$ 、 ARV 、 $SMAPE$ 八个误差评价指标上分析不同模型的预测性能。

表 4-7 各模型在测试集上的评价指标对比

指标	SVR	BPNN	LSTM	AtBiLSTM	BiGRU
$RMSE$	0.5464	0.5678	0.0774	0.0885	0.0540
MSE	0.2985	0.3224	0.0060	0.0078	0.0029
MAE	0.4616	0.4756	0.0563	0.0730	0.0316
$MAPE$	9.5129	9.8173	1.1760	1.5083	0.6260
R^2	-2.8405	-3.1479	0.8988	0.8678	0.9624
$RRMSE$	0.2656	0.2769	0.0351	0.0403	0.0239
ARV	0.1083	0.1121	0.0116	0.0153	0.0064
$SMAPE$	10.1142	10.4455	1.1697	1.5168	0.6313
$Rank$	4	5	2	3	1

粗体表示最小错误或最佳性能

首先，本研究全面考察了五个基础预测模型（即 SVR、BPNN、LSTM、AtBiLSTM 和 BiGRU）的性能，以说明不同模型的预测性能以及在本研究中使用 BiGRU 的理由。具体而言，使用工商银行股价数据集，以 $RMSE$ 、 MSE 、 MAE 、 $MAPE$ 、 R^2 、 $RRMSE$ 、 ARV 、 $SMAPE$ 为误差评价指标，比较了各预测模型的性能。不同模型间的比较结果如表 4-7 所示。图 4-7 展示了各预测模型在相同输入情况下的评价指标比较。

在比较五个基础预测模型的性能时，根据表 4-7、图 4-7 观察结果如下。首先，BiGRU 模型在预测股价方面优于其他模型，在 8 个误差评价指标上都有最优的结果，排名第一。其次，基于总体预测结果，LSTM 模型的性能与 BiGRU 相当；SVR、BPNN 的性能稍差。因此，本文优先选择 BiGRU 预测模型，并在此基础上进行区间预测。

基于基础模型 Model 预测，将五个模型中的每一个与提取的投资者情绪因子、投资者关注度因子进行匹配，并为它们提供相应的名称 SVR_X、BPNN_X、LSTM_X、AtBiLSTM_X 和 BiGRU_X。进一步地，再将其与 FMD 分解结果进

行匹配，分别以 SVR_X_Y、BPNN_X_Y、LSTM_X_Y、AtBiLSTM_X_Y 和 BiGRU_X_Y 命名。以此进一步比较和评估所提出方法的预测准确性和稳健性，以及验证投资者情绪因子、投资者注意力因子作为预测指标的有效性，和 FMD 分解能有效提升模型的预测性能。同时，表 4-8 显示了各模型的总体预测结果指标对比和排名；图 4-8 展示了各预测模型在不同输入情况下的评价指标比较；此外，由于 SVR 和 BPNN 模型预测值波动性较大，图 4-9 仅展示除此模型外其余各模型的预测拟合曲线。

表 4-8 各模型在测试集上的评价指标对比

模型	RMSE	MSE	MAE	MAPE	R^2	RRMSE	ARV	SMAPE	Rank
Model_X									
SVR_X	0.545	0.297	0.460	9.488	-	0.265	0.108	10.086	4
BPNN_X	0.567	0.321	0.475	9.773	-	0.276	0.112	10.404	5
LSTM_X	0.067	0.005	0.049	1.003	0.924	0.030	0.010	1.005	2
AtBiLSTM_X	0.075	0.006	0.056	1.144	0.905	0.034	0.012	1.151	3
BiGRU_X	0.049	0.002	0.028	0.544	0.970	0.022	0.006	0.549	1
Model_X_Y									
SVR_X_Y	0.540	0.292	0.455	9.371	-	0.262	0.107	9.954	4
BPNN_X_Y	0.562	0.316	0.470	9.705	-	0.275	0.111	10.355	5
LSTM_X_Y	0.066	0.004	0.047	0.980	0.927	0.030	0.010	0.979	2
AtBiLSTM_X_Y	0.068	0.005	0.049	1.015	0.923	0.031	0.010	1.017	3
BiGRU_X_Y	0.039	0.002	0.021	0.410	0.980	0.017	0.004	0.412	1

粗体表示最小错误或最佳性能

表 4-8、图 4-8 提供了以下几方面信息。首先，与基础模型（Model）相比，考虑了投资者情绪因子、投资者注意力因子的模型（Model_X）以及使用 FMD 分解结果的模型（Model_X_Y）均具有较低的误差值，模型的 MAE 值分别下降了 0.0012、0.0011、0.0076、0.0171、0.0041 和 0.0069、0.0052、0.0089、0.0238、0.0107。MAPE 值分别下降了 0.025、0.0442、0.1731、0.364、0.0818 和 0.1417、0.1128、0.196、0.4929、0.2164。其次，本文所使用的方法具有最低的误差值及

最高的 R^2 值，且在两个实验对照组中均排名第一。结果验证了股票市场对投资者情绪、注意力的敏感性，以及 FMD 分解在深度学习预测领域的重要性，证明了本章提出的模型在股票价格预测精度上有较大提升，保持输入最大优化情况下，能够更加准确的预测股价。

除了评价指标，本章还采用两样本 KS 检验对模型预测能力的差异进行统计检验，以证实上述结果的有效性。表 4-9 给出了 BiGRU 模型在不同输入方式下的 KS 检验 p 值的结果。K-S 检验零假设为两序列分布相同， p 值表示两个比较模型之间的显著差异， p 值小于 0.05 则表明两序列分布不同，即模型间存在显著差异。例如，第 2 行第 4 列的 p 值为 0.000，这意味着测试在 99% 的置信水平上拒绝零假设，即 BiGRU_X_Y 和 BiGRU 模型之间存在显著差异。K-S 检验的结果也证实了本文所考虑的影响因素以及 FMD 分解的有效性。

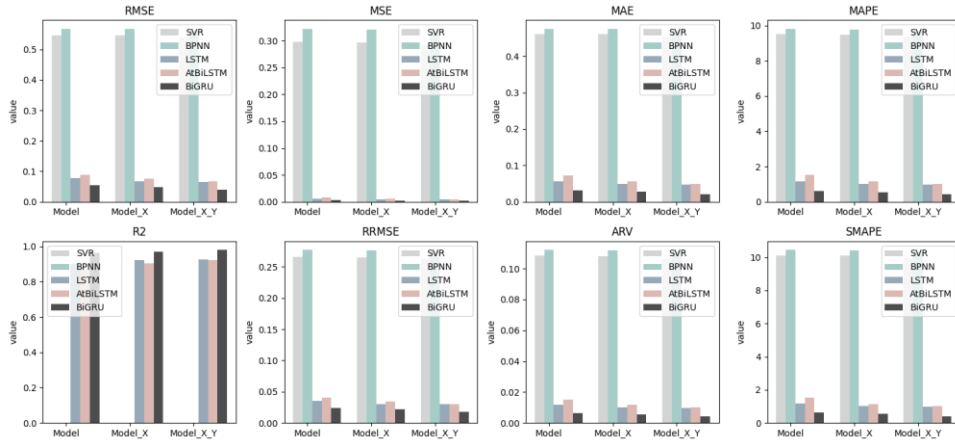


图 4-7 各预测模型相同输入情况下评价指标的比较

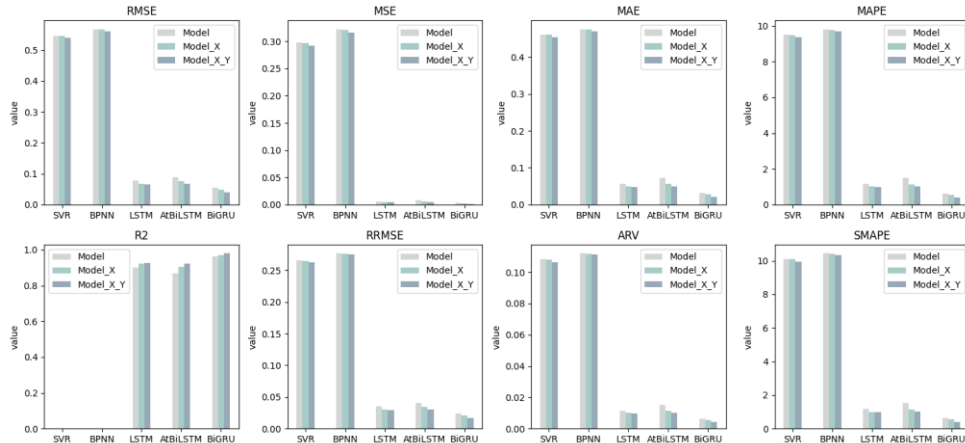


图 4-8 各预测模型不同输入情况下评价指标的比较

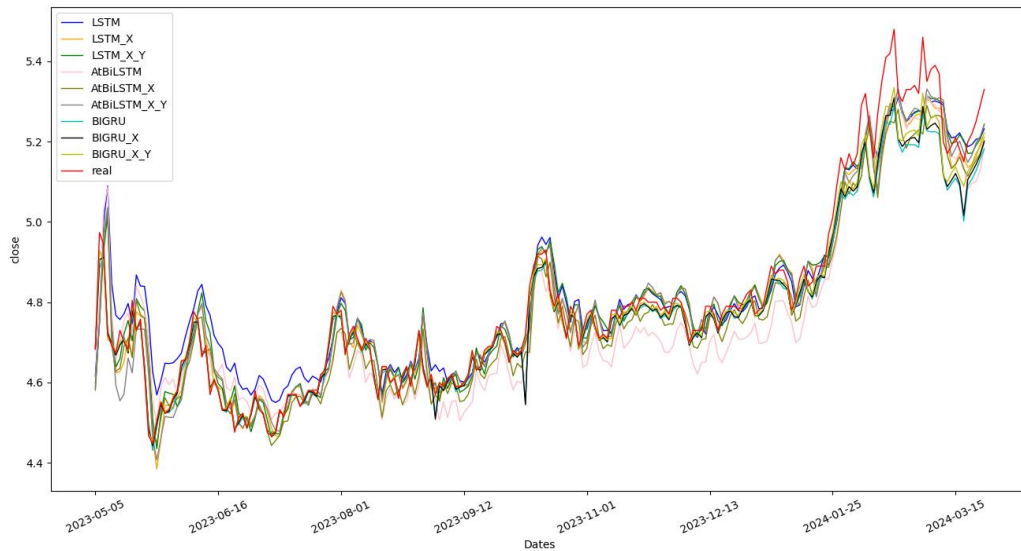


图 4-9 各模型预测收盘价趋势图

表 4-9 模型差异的 K-S 检验

	BiGRU	BiGRU_X	BiGRU_X_Y
BiGRU	1	0.001	0.000
BiGRU_X		1	0.000
BiGRU_X_Y			1

4.3.2 区间预测结果分析

鉴于点预测方法在刻画股价数据随机波动特征方面存在局限性,本研究将深入探索区间预测方法,提供更全面且精准的预测结果。为此,本文在沿用点预测性能最佳的 BiGRU 模型的基础上,将其与分位数回归相结合,开展区间预测工作。

在收集了点预测模型的预测收盘价后,本文再分别对股票价格的最低价和最高价进行区间预测,以获得股价波动的最终区间预测值。将股票对应的行情数据、投资者情绪因子、投资者注意力因子、FMD 分解结果作为输入,根据区间预测模型 QRBiGRU_X_Y,分别产生了股票日高和日低价格的预测值及其对应的上界和下界。本文以最低价区间预测结果的下界作为左端点,以最高价区间预测结果的上界作为的右端点,构成本文最终的区间预测值。QRBiGRU 区间预测模型使用与点预测模型一致的方式划分训练集和测试集,并使用滚动预测方式在预测

窗口进行预测。相关评价指标如表 4-10 所示，区间预测值与实际值如图 4-10 所示。

表 4-10 区间预测评价指标对比

指标	QRBIGRU_High_X_Y	QRBIGRU_Low_X_Y
<i>MAE</i>	0.014407	0.023498
<i>MSE</i>	0.000728	0.001677
<i>MAPE</i>	0.002815	0.004685
R^2	0.990837	0.977142
<i>PICP</i>	0.770161	0.669355
<i>PINAW</i>	0.012679	0.013535

从表 4-10、图 4-10 中可以观察到以下几点：第一，当实际波动性较强时，会显示更广泛的预期高点和低点，这反映了股票的总体趋势，表明了股票波动的变化。第二，QRBIGRU 模型预测的股票最高价和最低价与实际预测值之间的偏差很小，*MAPE* 小于 0.4%， R^2 分别达到 99.0837%和 97.7142%。第三，在 95% 置信水平下，该模型的*PICP* 指标均达到 70%左右，说明预测区间能较好地覆盖真实值曲线；*PINAW* 指标值均小于 0.014，说明区间宽度较窄，QRBIGRU 模型具有较优预测精度和稳定性。

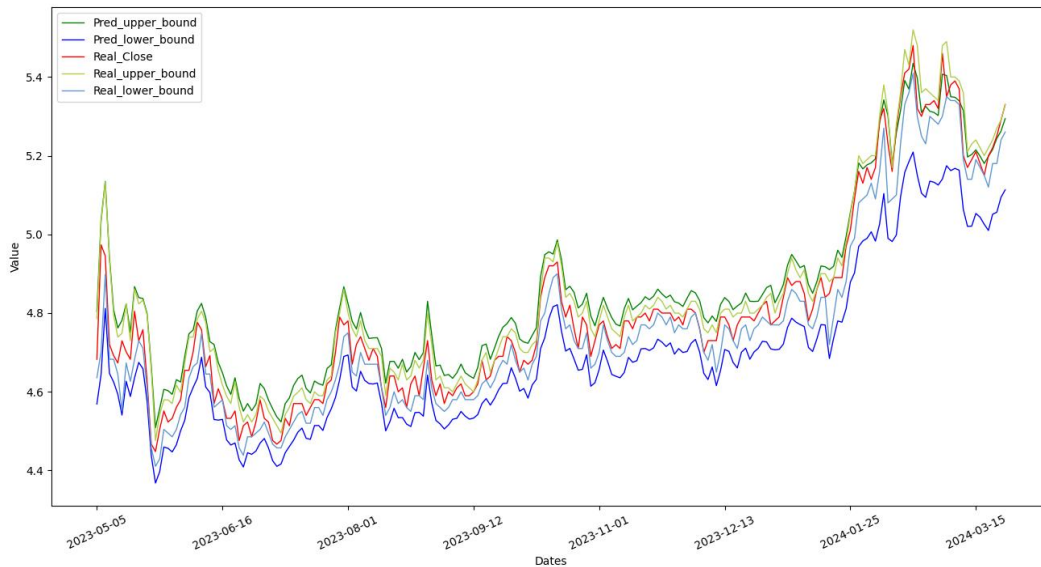


图 4-10 95%置信水平下分位数回归区间预测结果拟合图

4.4 量化投资应用分析

4.4.1 策略构建

在上文的实验中，使用误差评价指标、统计检验和区间估计指标来判断模型性能；然而，预测的准确性并不等于实际回报。在股票交易实践中，投资者的目标是制定有利可图的策略。因此，本文使用了一种长短期交易策略来展示所提出方法的盈利能力。实际策略是当预测收益率低于零时做空该头寸，当预测收益率高于零时做多该头寸。当预测回报为零时，维持头寸。

根据 2023 年 3 月 20 日至 2024 年 3 月 26 日的预测收盘价计算对应的预测收益率，第 t 天的预测收益率公式如下：

$$\hat{R}_t = \frac{\hat{C}_t - \hat{C}_{t-1}}{\hat{C}_{t-1}} \quad (4.8)$$

其中， $\hat{C}_t$ 表示第 t 天的收盘价， R_t 表示第 t 天的真实收益率。基于预测收益率 $\hat{R}_t$ ，可以制定以下投资策略。

为了减轻因错误投资造成的损失，使用区间预测的最高价和最低价作为保险。当预测收盘价在使用 QRBiGRU 模型预测的最高价和最低价之间时，即当满足 $\hat{L}_{t+1} \leq \hat{C}_{t+1} \leq \hat{H}_{t+1}$ 时，如果第二天的预测收益率是正的，那么在第二天开盘时做多该头寸，如果第二天的预测收益率是负的，那么就在第二日开盘时做空该头寸。只要 $\hat{R}_{t+1}$ 与 $\hat{R}_t$ 的正负方向相同，就可以实现正回报；然而，如果这两个值正负方向相反，投资就会产生负收益。否则，就认为第二天股票的波动性会更大，并且收益率预测出现错误，不执行任何交易。

4.4.2 回测结果分析

根据上述投资场景，本文模拟了 2023 年 5 月 5 日至 2024 年 3 月 26 日的真实股票交易，共计 218 天。为验证策略的有效性，本文使用同期的沪深 300 指数进行回测，并以此作为基准。

表 4-11 展示了不同模型在工商银行股票及沪深 300 指数上的模拟交易结果对比。其中，沪深 300 指数的投资评价指标采用消极投资策略（简单购买持有）计算得出。各模型的投资评价指标对比如图 4-11、图 4-12 所示。

表 4-11 各模型模拟交易结果

模型	Annual Return	Annual Volatility	Max Drawdown	Sharpe Ratio	Win Rate
Hs300					
	0.035757	0.192105	-0.452517	0.078773	0.503155
Model					
SVR	-0.002249	1.359428	-2.120090	-0.009012	0.532258
BPNN	0.130167	1.396117	-3.483488	0.080506	0.487903
LSTM	0.156610	0.145820	-2.078722	0.929464	0.564220
AtBiLSTM	0.149988	0.176722	-2.274569	0.734431	0.536697
BiGRU	0.224005	0.158838	-2.318804	1.210095	0.580645
Model_X					
SVR_X	-0.002277	1.353983	-2.117861	-0.009069	0.528226
BPNN_X	-0.052880	1.367244	-2.435963	-0.047047	0.471774
LSTM_X	0.159668	0.171348	-2.370016	0.806415	0.550459
AtBiLSTM_X	0.154357	0.175469	-2.562211	0.761297	0.564220
BiGRU_X	0.229082	0.166584	-2.166988	1.178695	0.588710
Model_X_Y					
SVR_X_Y	-0.003243	1.367927	-2.109188	-0.009685	0.524194
BPNN_X_Y	0.011642	1.439618	-2.378308	0.001094	0.500000
LSTM_X_Y	0.168631	0.16529	-2.488585	0.882577	0.541284
AtBiLSTM_X_Y	0.169533	0.176423	-2.381361	0.831259	0.564220
BiGRU_X_Y	0.235235	0.16283	-2.0396	1.236564	0.580645

粗体表示最小错误或最佳性能

结果表明，在考虑区间限制后，由于其双重保险结构，基于本章模型的投资策略享有最高的收益且承担了最低的风险，其年化收益率、夏普比率以及分别为 23.5235%、1.236564，优于所有模型，且比简单购买持有沪深 300 指数的年化收益率高出 19.9478%、夏普比率高出 1.157791。同时，风险指标度量值年化波动率为 0.16283，也小于沪深 300 指数的 0.192105。其次，在相同指标输入的情况

下, BiGRU 模型的各项指标最优; 在不同指标输入情况下, BiGRU_X_Y 模型的各项指标最优, 进一步验证了本文模型的有效性以及预测指标的有效性。总而言之, 采用 QRBiGRU_X_Y 区间预测模型的投资策略比持有沪深 300 指数收益更高、风险更低, 本章构建的模型在投资应用方面有良好的应用价值。

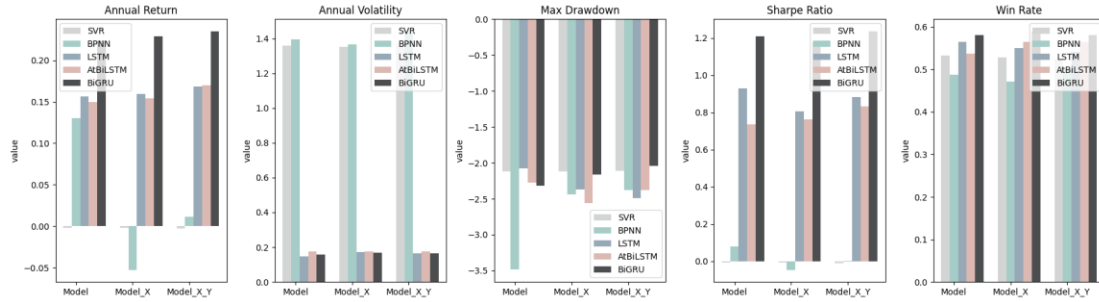


图 4-11 各模型在相同输入情况下指标的比较

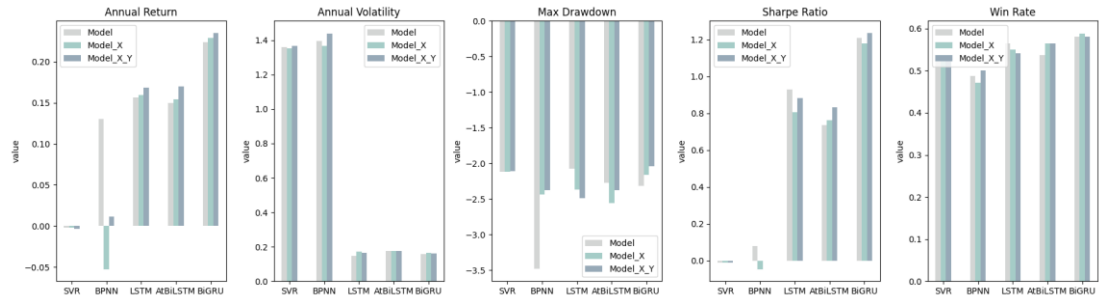


图 4-12 各模型在不同输入情况下指标的比较

以上结论可以从以下两个角度来解释。首先, 影响股票价格的波动的因素有很多。本章建议的策略全面考虑了多种指标, 并使用先进的深度学习预测模型 BiGRU 提取信息来进行股价预测。其次, 区间估计可以为交易提供保险。通过不完全依赖于点预测结果, 交易错误的频率降低, 获得更高的收益及更低的风险, 同时预测区间的置信水平可以防止不必要的损失。因此, 这种策略特别适合高波动性时期, 如股票大幅上涨和下跌, 可以引导投资者减少错误交易, 帮助投资者做出理性判断, 加强风险管理, 从而获得高回报。

4.5 稳健性检验

投资者情绪对金融市场有着显著影响, 股指市场亦不例外。白露莹 (2021) 研究了投资者情绪对沪深 300 股指期货定价偏差的影响机制, 通过 2017-2020 年的交易数据实证分析发现, 投资者情绪会通过市场流动性、波动性和投机活跃性等中介变量显著影响期货定价偏差, 表明非理性投资者行为在我国期货市场中仍

占据重要地位。该研究证实了投资者情绪不仅会影响股票市场,同样会对期货市场定价效率产生系统性影响,且这种影响比传统金融理论预期的更为持久和显著^[98]。为检验本章所提出方法的稳健性,在本小节中进行进一步的实验,即股指预测。本节选取了5只股指进行稳健性检验,分别为zssh000300沪深300指数、zssh000905中证500指数、zssz399001深证成指、zssz399006创业板指。中证500指数的点预测结果如表4-12所示,最高价、最低价的区间预测结果如表4-13所示,各预测模型的评价指标对比如图4-13、图4-14所示。

表4-12 各预测模型的指标对比(zssh000905)

模型	RMSE	MSE	MAE	MAPE	R^2	RRMSE	ARV	SMAPE	Rank
Model									
SVR	561.881	315709.999	444.480	8.242	-	7.195	0.073	7.726	4
BPNN	886.119	785207.443	731.320	13.085	-	11.160	0.119	12.359	5
LSTM	114.575	13127.406	97.748	1.763	0.901	1.532	0.017	1.745	2
AtBiLSTM	136.003	18496.709	119.417	2.174	0.860	1.824	0.021	2.144	3
BiGRU	27.968	782.183	18.003	0.324	0.995	0.382	0.003	0.324	1
Model_X									
SVR_X	561.302	315060.187	443.970	8.232	-	7.188	0.073	7.717	4
BPNN_X	855.245	731443.768	692.052	12.356	-	10.81	0.114	11.748	5
LSTM_X	94.793	8985.764	74.350	1.353	0.932	1.283	0.013	1.343	2
AtBiLSTM_X	103.773	10768.909	86.394	1.556	0.919	1.388	0.015	1.541	3
BiGRU_X	21.93	481.054	14.888	0.269	0.997	0.302	0.003	0.269	1
Model_X_Y									
SVR_X_Y	561.235	314984.921	443.922	561.235	-	7.187	0.073	7.717	4
BPNN_X_Y	788.491	621717.288	627.229	11.148	-	9.982	0.103	10.647	5
LSTM_X_Y	79.665	6346.536	57.368	1.056	0.952	1.097	0.011	1.052	2
AtBiLSTM_X_Y	87.006	7569.995	67.126	1.214	0.943	1.185	0.012	1.213	3
BiGRU_X_Y	13.384	179.130	9.915	0.178	0.999	0.182	0.002	0.178	1

粗体表示最小错误或最佳性能

表 4-13 最高价、最低价的区间预测结果

指标	QRBIGRU_High_X_Y	QRBIGRU_Low_X_Y
<i>MAE</i>	16.797548	13.602786
<i>MSE</i>	838.449585	367.783569
<i>MAPE</i>	0.002912	0.002496
R^2	0.994708	0.997952
<i>PICP</i>	0.826613	0.911290
<i>PINAW</i>	0.011383	0.010829

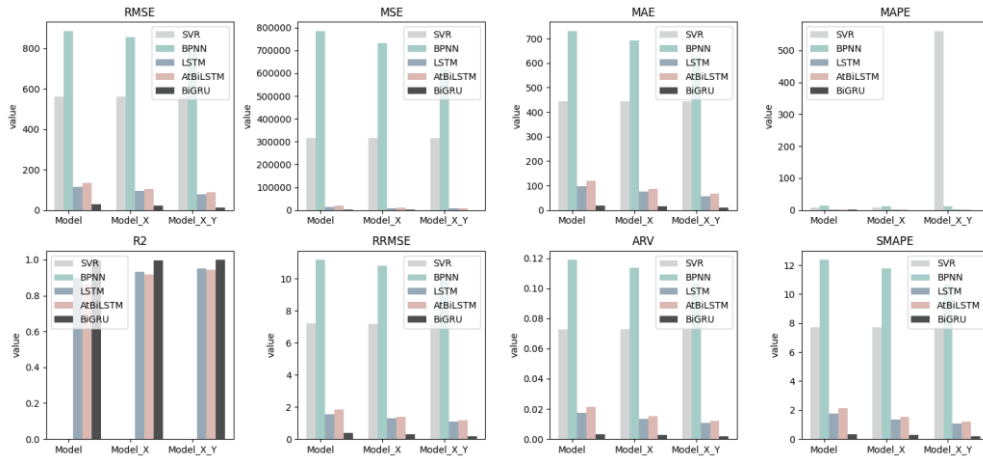


图 4-13 各模型在相同输入情况下指标的比较

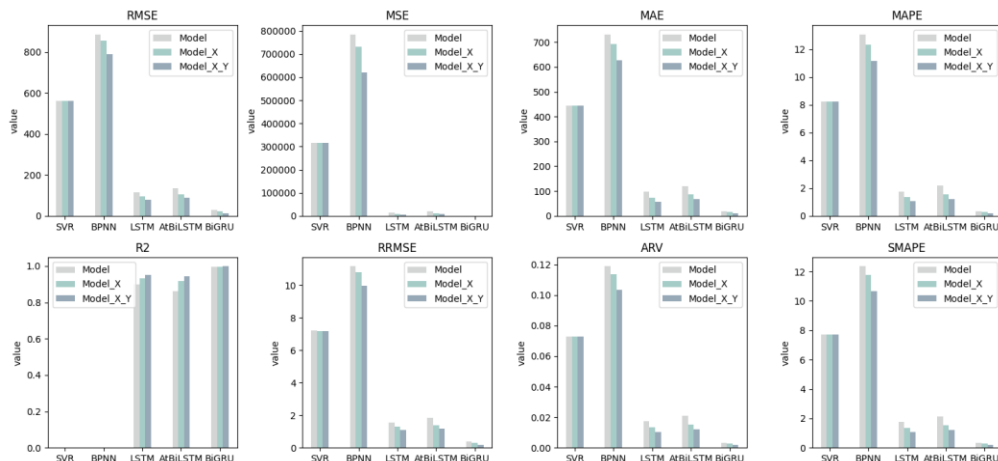


图 4-14 各模型在不同输入情况下指标的比较

基于实验结果，可以得出与以上实验相似的结论：首先在模型性能比较上，BiGRU 在单个预测模型的比较中具有绝对的性能优势，验证了其在时序数据建

模中的优越性。其次，模型具有多因子增强效应，引入投资者情绪、投资者注意力、FMD 分解特征后，在这三种情况（Model、Model_X、Model_X_Y）下的所有模型的整体预测性能都得到了提高，表明市场非结构化信息对股价预测具有重要补充作用。第三，预测精度优化上，本研究提出的组合预测模型（FMD-QRBiGRU）在所有对比模型中均取得最低预测误差，证实了其卓越的预测能力。第四，具有区间预测优势，该模型在区间预测中表现出高覆盖率和合理的区间宽度，在保证预测准确性的同时增强了结果的可信度，为量化投资提供了更科学的决策依据。

本节实验的结果支持了先前实验得出的结论，即使用的预测模型结合本研究中提出的多指标信息，可以有效地进行股价研究，从而做出相对高收益低风险的投资决策。

4.6 本章小结

随着互联网技术的快速发展，网络舆情已成为影响金融市场的重要因素。本章创新性地 will 网络舆情划分为情感和影响力两个属性进行量化，并构建了融合深度学习、投资者情绪、投资者注意力以及 FMD 分解技术的混合预测模型进行研究，包括预测和交易策略的两部分。具体而言，首先，在数据层面，本章通过文本情感分析技术从股吧评论中提取投资者情绪因子，结合七个维度的投资者注意力因子，与股票历史行情数据及 FMD 分解特征进行多源数据融合，为预测模型提供更全面的信息输入以提高预测准确性。其次，本研究突破了传统点预测的局限，提出了基于区间估计的创新交易策略。

在实证分析中，本章以工商银行为代表，对 6 个行业板块中共 30 只股票进行了股价预测和模拟交易实验。预测结果表明，所提出的方法优于其他基准模型。此外，在对 5 只指数进行价格预测的稳健性检验中，所提出的预测方法继续表现良好。实验表明，结合投资者情绪因子、投资者注意力因子以及 FMD 分解能显著减少预测误差，且具有稳健性。基于预测的模拟交易结果表明，所提出的具有区间约束的交易策略提高了工商银行等股票的投资收益率，并且降低投资风险。这一策略可以有效地引导投资者进行谨慎和准确的交易，以更低的交易成本获得更大的投资回报。

如国内外研究现状所述，关于预测股价以及量化交易投资模型的文献非常多。与这些研究相比，本章引入了一种创新的方法，利用区间预测结果作为约束，并结合投资者情绪、投资者注意力、FMD 分解，以降低仅基于点预测的交易策略中的投资风险。通过模拟交易，研究表明，投资者情绪、投资者注意力、FMD 分解可以提高模型预测精度；并且通过区间约束，可以提高投资收益率、降低投资风险。

第5章 总结与展望

5.1 本文工作总结

股票市场作为投资领域的重要组成部分，一直受到众多投资者的青睐。在复杂多变的市场环境下，投资者们借助各种技术工具，力求精准把握市场动态，预测其发展趋势。合理利用市场行情信息，制定科学的交易策略，不仅有助于投资者提升投资收益，还能有效规避潜在风险。相反，一旦出现判断失误或策略不当，投资者往往会遭受经济损失。因此，深入探索提升股票市场投资收益、降低风险的方法，对投资者具有深远的现实意义。本文综合运用信号处理技术和深度学习算法，对股价波动进行了深入研究，并构建量化投资策略以验证所提方案的可行性和有效性。本文的主要工作如下：

首先，本研究对股价波动、信号处理以及量化投资等领域的相关文献进行了梳理。详细阐述了有效市场假说、行为金融学理论以及有限关注假说等经典金融学理论，同时介绍了常见的信号处理算法和主流的股价预测模型。此外，对预测模型的训练流程以及预测评估指标进行了详细阐释，为后续研究工作的开展搭建了坚实的理论框架。

在此基础上，首先考虑了金融市场的天气效应，创新性地将金融市场的天气效应纳入股价预测的研究范畴，提出了 Savgol-BiLSTM-Att 新型深度学习模型。该模型的核心创新在于充分考量了股票市场与天气因素之间的潜在联系，并将 Savitzky-Golay 滤波技术融入深度学习模型中，从而更高效地提取金融数据的关键特征，显著增强了模型的预测性能。通过对中国 A 股证券市场市值排名前 30 的股票进行实证分析，研究发现股票市场确实存在天气效应，引入天气指标和 Savitzky-Golay 滤波后，模型的预测精度得到了有效提升。与众多传统预测模型相比，Savgol-BiLSTM-Att 模型在股价预测方面展现出明显优势；同时基于模型预测结果构建的 TopkDrop 交易策略在回测期间各项评价指标均显著优于同期沪深 300 指数，有力地证明了其在实际投资决策中的应用价值。

然后，基于网络舆情视角，进一步研究了网络舆情对股价波动及交易策略的影响。研究过程中，利用东方财富网股吧评论数据，构建了投资者情绪因子和投

投资者关注度因子,运用 FMD 模态分解方法对股票收盘价进行分解去噪,提取特征序列,并结合股票行情数据作为模型输入,构建了 FMD-QRBiGRU 区间预测模型,实现了对股价的有效预测。实证结果表明,该模型融合了投资者情绪因子、投资者注意力因子以及 FMD 分解技术,能显著减少预测误差,且具备良好的稳健性。同时,提出的具有区间约束的交易策略能有效提升股票投资的收益水平,降低投资风险。

文章的最后,本研究对基于天气效应和网络舆情构建的两种股价预测模型进行了全面总结。通过对比分析,明确得出从天气效应和网络舆情两个方面构建预测模型,对提高股价预测精度、实现更高投资收益具有显著效果的结论。同时,对研究过程中存在的不足之处进行了思考和分析。此外,结合当前人工智能时代的发展趋势,对股票投资收益提升领域未来的研究方向和热点问题进行了前瞻性展望,为后续相关研究提供了可能的思路 and 方向。

5.2 未来展望

本文以深度学习模型在股票投资领域的应用为切入点,选取中国股票市场作为研究对象,进行了一系列实证分析。通过对不同方法的对比分析,本文深入探讨了它们在实际应用中的有效性。研究的核心是将深度学习技术与信号处理相结合,旨在实现更精准的股价预测,并获得更高的投资收益。本研究所构建的模型在预测性能和量化回测方面均取得了较为理想的结果,为深度学习技术在股票投资领域的应用提供了新的视角,具备一定实践价值。然而,由于研究条件及个人时间精力有限,本研究仍有改进空间,具体如下:

(1) 多源异构数据的集成应用

本文分别探讨了天气效应、网络舆情对股价及量化投资的影响,而除此之外,股票价格还会受到政府政策、重大事件以及企业经营情况等多种因素影响。因此在后续研究中,可以将这些因素均考虑在内,并可以探索如何有效整合这些多源异构数据,以提高预测模型的准确度和鲁棒性。

(2) 预测模型的优化与调整

本文构建的股价预测模型包含大量超参数,而本文在确定股价预测模型的超参数时,主要依赖调参经验预设的方式进行参数设定。然而,这种人工调参方式可能无法获得最优参数组合。在后续研究中,建议采用更系统化的参数优化方法,

如网格搜索或随机搜索算法，通过自动化调参来寻找更优的超参数配置，从而进一步提升模型的预测性能。此外，随着人工智能与深度学习发展，数据获取和分析能力不断增强，Transformer 预测、Bert 文本情感分析等新兴预测模型的涌现也给股价预测带来了新的工具，可在后续研究中加以应用。

（3） 量化投资策略的优化

本文基于股价预测模型提出了两种量化投资策略，虽能部分体现了股价预测模型的潜力，但仍有优化空间。未来研究可从以下方向改进：首先，可尝试设计更多元的交易策略或优化现有策略参数，例如调整股票持仓权重分配方式；其次，可探索将预测模型与交易策略进行端到端联合训练，以投资收益率作为优化目标函数，从而更充分地挖掘模型的预测能力，实现收益最大化。这些改进方向有望进一步提升策略表现。

（4） 模型应用范围的扩展

本研究的实验集中在中国股票市场，而该市场存在诸如高昂的做空成本等特殊性，这些特征可能在一定程度上制约了模型预测性能的发挥以及投资策略的构建。因此，未来的研究可以将模型拓展至更广泛的金融市场，例如期货市场或国际股票市场等，通过在不同市场环境中的应用，进一步检验模型的稳定性和泛化能力，从而为模型的优化和策略的完善提供更全面的实证支持。

参考文献

- [1] Liu K Y, Zhou J N, Dong D Y. Improving stock price prediction using the long short-term memory model combined with online social networks [J]. Journal of Behavioral and Experimental Finance, 2021, 30: 100507.
- [2] Gu D L, Huang J. Multifractal detrended fluctuation analysis on high-frequency SZSE in Chinese stock market [J]. Physica A: Statistical Mechanics and its Applications, 2019, 521: 225-235.
- [3] 李斌, 龙真. 中国股票市场可预测性研究: 基于机器学习的视角 [J]. 管理科学学报, 2023, 26(10): 138-158.
- [4] Patel J, Shah S, Thakkar P, et al. Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques [J]. Expert Systems With Applications, 2015, 42(1): 259-268.
- [5] Roberto C, Angel M D R, Daniel P, et al. Web Traffic Time Series Forecasting Using LSTM Neural Networks with Distributed Asynchronous Training [J]. Mathematics, 2021, 9(4): 421-421.
- [6] Zhong X, Enke D. Predicting the daily return direction of the stock market using hybrid machine learning algorithms [J]. Financial Innovation, 2019, 5(1): 1-20.
- [7] Polanco-martínez J M. Dynamic relationship analysis between NAFTA stock markets using nonlinear, nonparametric, non-stationary methods [J]. Nonlinear Dynamics, 2019, 97(1): 369-389.
- [8] Wang X B, Wang X, Li B, et al. The nonlinear characteristics of Chinese stock index futures yield volatility:Based on the high frequency data of CSI300 stock index futures [J]. China Finance Review International, 2019, 10(2): 175-196.
- [9] Ren S S, Wang X, Zhou X, et al. A novel hybrid model for stock price forecasting integrating Encoder Forest and Informer [J]. Expert Systems With Applications, 2023, 234: 121080.
- [10] 陈炜, 徐慧琳, 汪寿阳等. 基于误差修正与分解的区间值股价时间序列预测研究 [J]. 系统工程理论与实践, 2023, 43(02): 383-397.

- [11]张强, 杨淑娥. 噪音交易、投资者情绪波动与股票收益 [J]. 系统工程理论与实践, 2009, 29(03): 40-47.
- [12]陈小玲. 基于 ARIMA 模型与神经网络模型的股价预测 [J]. 经济数学, 2017, 34(04): 30-34.
- [13]Nguyen V, Robert S. Applying Hybrid ARIMA-SGARCH in Algorithmic Investment Strategies on S&P500 Index [J]. Entropy, 2022, 24(2): 158-158.
- [14]王美今, 孙建军. 中国股市收益、收益波动与投资者情绪 [J]. 经济研究, 2004, (10): 75-83.
- [15]鲁训法, 黎建强. 中国股市指数与投资者情绪指数的相互关系 [J]. 系统工程理论与实践, 2012, 32(03): 621-629.
- [16]崔建福, 李兴绪. 股票价格预测: GARCH 模型与 BP 神经网络模型的比较 [J]. 统计与决策, 2004, (06): 21-22.
- [17]Yu X L, Huang Y. The impact of economic policy uncertainty on stock volatility: Evidence from GARCH–MIDAS approach [J]. Physica A: Statistical Mechanics and its Applications, 2021, 570: 125794.
- [18]吴玉霞, 温欣. 基于 ARIMA 模型的短期股票价格预测 [J]. 统计与决策, 2016, (23): 83-86.
- [19]Latha C M, Venkataramanaiah M, Rao K S N. S&P BSE Sensex and S&P BSE IT return forecasting using ARIMA [J]. Financial Innovation, 2020, 6(1): 47.
- [20]李晶. 基于 GARCH 模型的上证 50ETF 期权风险对冲策略研究 [J]. 经济问题, 2023, (03): 68-75.
- [21]方燕, 耿雪洋, 秦珊珊. 沪深两市传媒板块指数价格预测研究——基于 ARIMA—GARCH 模型的分析 [J]. 价格理论与实践, 2018, (01): 102-105.
- [22]Wang Y S. Market Index and Stock Price Direction Prediction using Machine Learning Techniques: An empirical study on the KOSPI and HSI. Business Intelligence and Data Mining [J], 2014, 9(2): 145-160.
- [23]Chen W, Zhang H Y, Kumar M M, et al. Mean–variance portfolio optimization using machine learning-based stock price prediction [J]. Applied Soft Computing Journal, 2021, 100: 106943.

- [24] 张晨希, 张燕平, 张迎春等. 基于支持向量机的股票预测 [J]. 计算机技术与发展, 2006, (06): 35-37.
- [25] Chen Y B, Tan H W. Short-term prediction of electric demand in building sector via hybrid support vector regression [J]. Applied Energy, 2017, 204: 1363-1374.
- [26] Miranda H B, Amorim S V, Herbert K. Practical machine learning: Forecasting daily financial markets directions [J]. Expert Systems With Applications, 2023, 233: 120840.
- [27] Zhang D H, Lou S. The application research of neural network and BP algorithm in stock price pattern classification and prediction [J]. Future Generation Computer Systems, 2021, 115: 872-879.
- [28] A. B G, Flávio B, C. S L H, et al. Order book mid-price movement inference by CatBoost classifier from convolutional feature maps [J]. Applied Soft Computing Journal, 2022, 116: 108274.
- [29] Persio L D, Honchar O. Recurrent Neural Networks Approach to the Financial Forecast of Google Assets [J]. International Journal of Mathematics and Computers in Simulation, 2017, 11: 7-13.
- [30] 彭燕, 刘宇红, 张荣芬. 基于 LSTM 的股票价格预测建模与分析 [J]. 计算机工程与应用, 2019, 55(11): 209-212.
- [31] 吴彬, 张勇, 唐颖军. 融合 SDE 算法和双路加权门控循环神经网络的股价走势预测模型 [J]. 小型微型计算机系统, 2021, 42(07): 1371-1376.
- [32] Yu Y, Si X S, Hu C H, et al. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures [J]. Neural computation, 2019, 31(7): 1235-1270.
- [33] Pan S Y, Long S W, Wang Y M, et al. Nonlinear asset pricing in Chinese stock market: A deep learning approach [J]. International Review of Financial Analysis, 2023, 87: 102627.
- [34] Andrés V, Werner K. Gold Volatility Prediction using a CNN-LSTM approach [J]. Expert Systems With Applications, 2020, 157: 113481.
- [35] Wang H Y, Wang J X, Cao L H, et al. A Stock Closing Price Prediction Model Based on CNN-BiSLSTM [J]. Complexity, 2021, 2021(1): 5360828.

- [36] 李斌, 邵新月, 李玥阳. 机器学习驱动的基本面量化投资研究 [J]. 中国工业经济, 2019, (08): 61-79.
- [37] 王婷, 夏阳雨新, 陈铁明. 基于多类别特征体系的股票短期趋势预测 [J]. 计算机科学, 2020, 47(S2): 491-495.
- [38] Barua R, Sharma A K. Dynamic Black Litterman portfolios with views derived via CNN-BiLSTM predictions [J]. Finance Research Letters, 2022, 49: 103111.
- [39] Zhang J L, Ye L S, Lai Y Z. Stock Price Prediction Using CNN-BiLSTM-Attention Model [J]. Mathematics, 2023, 11(9): 1985.
- [40] 赵澄, 叶耀威, 姚明海. 基于金融文本情感的股票波动预测 [J]. 计算机科学, 2020, 47(05): 79-83.
- [41] 熊熊, 许克维, 沈德华. 投资者情绪与期货市场功能——基于沪深 300 股指期货的研究 [J]. 系统工程理论与实践, 2020, 40(09): 2252-2268.
- [42] Khaled O, Kuntara P. A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news [J]. Journal of Financial Economics, 2022, 144(1): 273-297.
- [43] Kenneth L. F, Statman M. Investor Sentiment and Stock Returns [J]. Financial Analysts Journal, 2000, 56(2): 16-23.
- [44] 王道平, 贾昱宁. 投资者情绪与中国股票市场过度波动 [J]. 金融论坛, 2019, 24(07): 46-59.
- [45] Li X D, Wu P J, Wang W P. Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong [J]. Information Processing and Management, 2020, 57(5): 102212.
- [46] Zhang W G, Gong X, Wang C, et al. Predicting stock market volatility based on textual sentiment: A nonlinear analysis [J]. Journal of Forecasting, 2021, 40(8): 1479-1500.
- [47] 李合龙, 任昌松, 丘润文等. 基于 GARCH-MIDAS 的混频投资者情绪对股市波动的影响 [J]. 系统管理学报, 2023, 32(05): 1036-1045.
- [48] Erick M, Luiz C O F, M. D M L. Point and interval forecasting of electricity supply via pruned ensembles [J]. Energy, 2021, 232: 121009.

- [49] Guo H, Huang B, Wang J H. Probabilistic load forecasting for integrated energy systems using attentive quantile regression temporal convolutional network [J]. *Advances in Applied Energy*, 2024, 14: 100165.
- [50] He Y Y, Liu R, Li H Y, et al. Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory [J]. *Applied Energy*, 2017, 185: 254-266.
- [51] Hu J M, Luo Q X, Tang J W, et al. Conformalized temporal convolutional quantile regression networks for wind power interval forecasting [J]. *Energy*, 2022, 248: 123497.
- [52] 谢游宇, 王万雄. 基于 EMD 和 SSA 的股票预测模型 [J]. *计算机工程与应用*, 2023, 59(18): 285-292.
- [53] 韩莹, 张栋, 孙凯强等. 结合长短时记忆网络和宽度学习的股票预测新模型研究 [J]. *运筹与管理*, 2023, 32(08): 187-192.
- [54] Lin Y, Lin Z X, Liao Y, et al. Forecasting the realized volatility of stock price index: A hybrid model integrating CEEMDAN and LSTM [J]. *Expert Systems With Applications*, 2022, 206: 117736.
- [55] 欧阳红兵, 黄亢, 闫洪举. 基于 LSTM 神经网络的金融时间序列预测 [J]. *中国管理科学*, 2020, 28(04): 27-35.
- [56] Qi C Y, Ren J Y, Su J. GRU Neural Network Based on CEEMDAN-Wavelet for Stock Price Prediction [J]. *Applied Sciences*, 2023, 13(12): 7104.
- [57] Miao Y H, Zhang B Y, Li C H, et al. Feature Mode Decomposition: New Decomposition Theory for Rotating Machinery Fault Diagnosis [J]. *IEEE Transactions on Industrial Electronics*, 2023, 70(2): 1949-1960.
- [58] 王晓真, 彭勃, 王家忠等. 参数优化 FMD 的滚动轴承早期故障诊断 [J]. *组合机床与自动化加工技术*, 2024, (06): 131-134.
- [59] Yang X T, Zou A K, Cao J D, et al. Systemic risk prediction based on Savitzky-Golay smoothing and temporal convolutional networks [J]. *Electronic Research Archive*, 2023, 31(5): 2667-2688.

- [60] Qiu J Y, Wang B, Zhou C J. Forecasting stock prices with long-short term memory neural network based on attention mechanism [J]. PloS one, 2020, 15(1): 0227222.
- [61] Ouyang Z S, Yang X T, Lai Y Z. Systemic financial risk early warning of financial market in China using Attention-LSTM model [J]. North American Journal of Economics and Finance, 2021, 56: 101383.
- [62] Abbasimehr H, Paki R. Improving time series forecasting using LSTM and attention models [J]. Journal of Ambient Intelligence and Humanized Computing, 2021, 13: 673-691.
- [63] 马一琳. 基于深度神经网络的股价预测及投资组合研究 [D]. 南京: 东南大学, 2020.
- [64] 潘唯一. 基于机器学习技术的量化投资策略研究 [D]. 北京: 中国科学院大学, 2020.
- [65] 何雨洵. 深度学习 NLP 技术和 LSTM 算法模型在量化投资策略中的应用研究 [D]. 上海: 上海财经大学, 2021.
- [66] 林昱. 基于深度学习的股市预测方法研究及其在投资决策中的应用 [D]. 成都: 西南财经大学, 2024.
- [67] 张学勇, 吴雨玲. 基于网络大数据挖掘的实证资产定价研究进展 [J]. 经济学动态, 2018, (06): 129-140.
- [68] 张学勇, 唐国梅. 行业关注度与股票横截面收益率——基于百度行业搜索指数的研究 [J]. 经济学(季刊), 2022, 22(03): 773-794.
- [69] 石勇, 唐静, 郭琨. 社交媒体投资者关注、投资者情绪对中国股票市场的影响 [J]. 中央财经大学学报, 2017, (07): 45-53.
- [70] 饶育蕾, 彭叠峰, 成大超. 媒体注意力会引起股票的异常收益吗?——来自中国股票市场的经验证据 [J]. 系统工程理论与实践, 2010, 30(02): 287-297.
- [71] Cho K H, Merrienboer B V, Bahdanau D, et al. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches [J]. Proceedings of the 8th Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-8), 2014, 103-111.

- [72] Bassett G, Jr., Koenker R. Regression quantiles [J]. *Econometrica*, 1978, 46(1): 33-50.
- [73] Chang T, Nieh C-C, Yang M J, et al. Are stock market returns related to the weather effects? Empirical evidence from Taiwan [J]. *Physica A: Statistical Mechanics and its Applications*, 2005, 364: 343-354.
- [74] Li M G, Li W R, Wang F, et al. Applying BERT to analyze investor sentiment in stock market [J]. *Neural Computing and Applications*, 2020, 33(10): 1-14.
- [75] Sergey S. How weather impacts expressed sentiment in Russia: evidence from Odnoklassniki [J]. *PeerJ Computer science*, 2022, 8: 1164.
- [76] C K M, L F B, Oscar Y, et al. A warm heart and a clear head. The contingent effects of weather on mood and cognition [J]. *Psychological science*, 2005, 16(9): 724-731.
- [77] Hu T Y, Xie X Y, Li J. Negative or positive? The effect of emotion and mood on risky driving [J]. *Transportation Research Part F: Psychology and Behaviour*, 2013, 16: 29-40.
- [78] Waldemar T, Urszula M, Grzegorz M, et al. The Influence of Investors' Mood on the Stock Prices: Evidence from Energy Firms in Warsaw Stock Exchange, Poland [J]. *Energies*, 2021, 14(21): 7396-7396.
- [79] Guven C, Hoxha I. Rain or shine: Happiness and risk-taking [J]. *Quarterly Review of Economics and Finance*, 2015, 57: 1-10.
- [80] Dong M, Andréanne T. Does the Weather Influence Global Stock Returns? [J]. *Critical Finance Review*, 2021, 10: 207-249.
- [81] Dong M, Andréanne T. Global weather-based trading strategies [J]. *Journal of Banking and Finance*, 2022, 143: 106558.
- [82] 郭永济, 张谊浩. 空气质量会影响股票市场吗? [J]. *金融研究*, 2016, (02): 71-85.
- [83] 万孝园, 陈欣. 雾霾对中国股市收益的影响 [J]. *投资研究*, 2016, 35(01): 81-94.

- [84] 徐纯鸽. 空气质量对中国股票市场交易行为的影响 [D]. 上海: 上海财经大学, 2020.
- [85] Roberto C R, Francisco G, Karolina M. Stock market trading rule based on pattern recognition and technical analysis: Forecasting the DJIA index with intraday data [J]. *Expert Systems With Applications*, 2015, 42(14): 5963-5975.
- [86] 黄婷婷. 机器学习在金融领域的应用 [D]. 合肥: 中国科学技术大学, 2018.
- [87] 陆静. 中国股票市场天气效应的实证研究 [J]. *中国软科学*, 2011, (06): 65-78+192.
- [88] 田洪彬. 网络舆情对企业价值的影响研究 [D]. 济南: 山东财经大学, 2024.
- [89] 张飞鹏, 徐一雄, 陈曦等. 基于新闻文本情绪的区间值股票回报预测研究 [J]. *计量经济学报*, 2024, 4(01): 204-230.
- [90] Kim J, KIM H S, Choi S Y. Forecasting the S&P 500 Index Using Mathematical-Based Sentiment Analysis and Deep Learning Models: A FinBERT Transformer Model and LSTM [J]. *Axioms*, 2023, 12(9): 835.
- [91] 徐英皓. 网络新闻舆情对股指期货收益率的影响研究 [D]. 上海: 上海财经大学, 2023.
- [92] 熊艳. 论坛发帖与股价行为: 情绪宣泄还是信息传递? [J]. *中央财经大学学报*, 2022, (05): 29-45.
- [93] 范小云, 王业东, 王道平等. 不同来源金融文本信息含量的异质性分析——基于混合式文本情绪测度方法 [J]. *管理世界*, 2022, 38(10): 78-101.
- [94] 赵传君, 刘金峰, 武美龄等. 基于数值与文本特征的 LSTM 股票价格预测研究 [J]. *山西大学学报(自然科学版)*, 2023, 46(05): 1058-1063.
- [95] 崔炎炎, 刘立新. 网络舆情赋能金融科技股票收盘价预测研究 [J]. *统计研究*, 2022, 39(06): 148-160.
- [96] Gang C, W. G J, Dehua S, et al. Machine learning to establish proxies for investor attention: evidence of improved stock-return prediction [J]. *Annals of Operations Research*, 2022, 318(1): 103-128.
- [97] Barber B M, Odean T. All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors [J]. *The Review of Financial Studies*, 2008, 21(2): 785-818.

- [98] 白露莹. 投资者情绪对股指期货定价偏差的影响研究 [D]. 成都: 四川大学, 2021.

硕士期间发表的论文

- [1] 方 言 ， 张 伟 . Research on stock quantitative investment using Savgol-BiLSTM-Att prediction model based on weather effects. (在投)
- [2] 方言, 张伟. Research on Stock Price Interval Prediction and Quantitative Investment Based on Online Public Sentiment and FMD-QRBiGRU Model. (在投)

致谢

行文至此，意味着我的研究生生涯即将画上句点，而这也标志着我在安徽工程大学七年求学时光的终结。这一路走来离不开众多师长、亲友的扶持与陪伴。在此，我谨以最诚挚的谢意，向所有给予我帮助的人致以深深的感激。

感谢我们伟大的祖国。感恩祖国母亲提供的安全舒适的生活环境，让我有机会接受高等教育并顺利完成学业，逐渐成为一名有理想、有能力的新时代青年。

感谢我的母校安徽工程大学。从2018年9月7日入学的第一天起，“尚德敏学，唯实唯新”的校训始终指引我前行。安徽工程大学数理与金融学院浓厚的学术氛围和严谨的治学理念，一直激励我奋发向上。

感谢我的导师张伟副教授。张老师对本论文的写作付出了许多，从选题到后期的修改不断给我最细致的建议，批注细化到每一句话、每一个符号。当我在写作时感到困惑不解时，张老师总是不厌其烦给予我最耐心的解答。张老师不仅教会我如何做研究，更以身作则诠释了“学高为师，身正为范”的真谛。在此，我要向张老师致以最诚挚的谢意！

同时，我也要感谢数理与金融学院所有传授我知识的老师和辅导员魏老师，感谢他们给我学习、生活和求职上的关怀和帮助。我还要感谢金融研22班的每一位同学，尤其是我的三位室友，我们彼此鼓励、共同成长的点点滴滴，让我的硕士生涯充满温暖与色彩。

树高千尺不忘根深至沃土。我还要感谢我的父母和家人。他们的理解、鼓励和爱，让我在每一次迷茫时都能重拾信心，是我一路走来的最大动力。

最后，感谢坚强勇敢的自己。那些无数次情绪崩溃的瞬间，那些无数个含泪坚持的时刻，最终都成为了照亮前路的星火。

研究生阶段的结束恰是人生新篇章的开始，我将带着这份感恩继续前行。