

分类号

UDC

密级

编号 10741

兰州财经大学

LANZHOU UNIVERSITY OF FINANCE AND ECONOMICS

硕士学位论文

(专业学位)

论文题目：基于VMD-CBAM-BiGRU模型的股票预测研究

研究生姓名：王龙

指导教师姓名、职称：郭精军、教授

学科、专业名称：统计学、应用统计

研究方向：大数据分析及应用

提交日期：2024年6月3日

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名： 王龙 签字日期： 2024.6.3

导师签名： 郭峰 签字日期： 2024.6.3

导师(校外)签名： _____ 签字日期： _____

关于论文使用授权的说明

本人完全了解学校关于保留、使用学位论文的各项规定， 同意
(选择“同意”/“不同意”)以下事项：

1.学校有权保留本论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文；

2.学校有权将本人的学位论文提交至清华大学“中国学术期刊(光盘版)电子杂志社”用于出版和编入 CNKI《中国知识资源总库》或其他同类数据库，传播本学位论文的全部或部分内容。

学位论文作者签名： 王龙 签字日期： 2024.6.3

导师签名： 郭峰 签字日期： 2024.6.3

导师(校外)签名： _____ 签字日期： _____

Research on stock prediction based on VMD-CBAM-BiGRU model

Candidate: Long Wang

Supervisor: Jingjun Guo

摘要

股票价格预测一直是金融领域中备受关注的话题，其研究意义深远而广泛，不管是对于投资者还是市场监管者而言均有着重要的现实意义。当前，随着数据科学、机器学习和人工智能的快速发展，对于股票变动趋势的预测正越来越精确以及多样化，股票市场研究正朝着更加复杂和多层次的方向发展。

本文所构建的 VMD-CBAM-BiGRU 模型为股票预测领域提供了一种兼具高精确性和稳定性的预测方法。首先，本文在基本行情指标的基础上引入大量的技术指标丰富特征集合，而后通过相关性分析筛选出对股价预测有显著效果的特征作为特征集合。在一系列基础模型中选择 BiGRU 作为基础预测模型，由于模型涉及大量超参数，合适的超参数可以显著提高模型的预测效果，而人工寻参效率低下，因此使用改进后的动态粒子群算法进行超参数寻优，提高模型的预测能力。由于股价序列具有高度非平稳性和强噪声以及大量特征等特点，考虑到 VMD 算法具有的序列降噪效果以及 CBAM 算法处理多特征数据中的优势，将 VMD 算法和 CBAM 算法引入 BiGRU 就基础模型进行降噪和特征提取处理，从两方面的优化中提高 VMD-CBAM-BiGRU 模型的预测效果。

本文从三个方面进行了模型的实证分析，首先在上证综指、沪深 300、中国平安和贵州茅台四支股票上进行试验，验证了 VMD 算法和 CBAM 算法对于模型 BiGRU 显著的提升作用，证实了本文所构建 VMD-CBAM-BiGRU 模型在股票预测领域的稳定性和有效性，然后进行跨时期股价预测验证了模型在中短期预测中的可行性和有效性，最后使用所构建的模型作为交易策略进行模拟交易验证了模型在投资应用中的实用效果，从而为股市投资者和监管部门预测股价动向提供了一种优秀的预测工具。

关键词： 股价预测 粒子群算法 变分模态分解 注意力机制 双向门控神经网络

Abstract

Stock price prediction has always been a topic of concern in the financial field, and its research significance is far-reaching and extensive, which has important practical significance for both investors and market regulators. At present, with the rapid development of data science, machine learning and artificial intelligence, forecasts of stock movements are becoming more accurate and diversified, stock market research is moving towards a more complex and multi-layered direction.

VMD-CBAM-BiGRU model is constructed to provide a method with high accuracy and stability for stock forecasting. First of all, a large number of rich feature sets of technical indicators are introduced on the basis of basic market indicators, and then selects the features that have significant effect on stock price prediction as the feature set through correlation analysis. BiGRU is selected as the basic prediction model among a series of basic models. Since the model involves a large number of hyperparameters, suitable hyperparameters can significantly improve the prediction effect of the model, while manual parameter seeking is inefficient, particle swarm optimization will be used to optimize the hyperparameters and improve the prediction ability of the model. Since stock price sequence has high non-stationary, strong noise and a large number of features, VMD algorithm and CBAM algorithm are introduced into BiGRU to carry out noise reduction and feature extraction processing

on the basic model, and improve the prediction effect of VMD-CBAM-BiGRU model from the two aspects of optimization.

Empirical analysis of the model are conducted from three aspects. First, experiments were conducted on four stocks, namely Shanghai Composite Index, Husen-Shenzhen 300, Ping An of China and Kweichow Moutai, to verify that VMD algorithm and CBAM algorithm significantly improved BiGRU model. The stability and effectiveness of the VMD-CBAM-BiGRU model in the field of stock prediction was verified. Then, the feasibility and effectiveness of the model in the short and medium term prediction was verified by inter-period stock price prediction. Finally, the model was used as a trading strategy for simulated trading to verify the practical effect of the model in investment application. Thus, it provides an excellent forecasting tool for stock market investors and regulators to predict the trend of stock prices.

Keywords: Stock price prediction ; Particle swarm algorithm; Variational modal decomposition; Attention mechanism,; Bidirectional gated neural network

目 录

1 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 股价预测研究现状.....	4
1.2.2 群智能算法研究现状.....	8
1.3 主要内容.....	10
1.4 创新之处.....	11
2 相关理论与研究方法	13
2.1 股票市场相关概述.....	13
2.1.1 股票价格的影响因素.....	13
2.1.2 股票技术指标概述.....	15
2.2 群智能优化算法.....	17
2.2.1 群智能优化算法的介绍.....	17
2.2.2 动态粒子群算法.....	20
2.3 分解集成理论.....	21
2.3.1 经验模态分解.....	21
2.3.2 变分模态分解(VMD).....	23
2.4 深度学习理论.....	25
2.4.1 BP 神经网络.....	25
2.4.2 循环神经网络.....	28
2.4.3 双向门控循环神经网络(BiGRU).....	30
2.4.4 卷积注意力机制.....	32
3 VMD-CBAM-BiGRU 模型的构建	35
3.1 Pearson 相关性分析.....	35
3.2 模糊熵.....	35

3.3 超参数优化.....	37
3.4 模型评价指标.....	38
3.5 VMD-CBAM-BiGRU 模型的构建思路.....	39
3.6 VMD-CBAM-BiGRU 模型的构建流程.....	40
4 实证分析	42
4.1 数据预处理.....	42
4.1.1 数据来源及描述.....	42
4.1.2 数据标准化处理.....	42
4.1.3 数据集切分.....	43
4.2 实验分析.....	43
4.2.1 上证综指.....	43
4.2.2 沪深 300.....	46
4.2.3 中国平安.....	49
4.2.4 贵州茅台.....	52
4.3 不同跨度时期下股价预测效果分析.....	55
4.4 模拟交易.....	57
5 总结与展望	59
5.1 总结.....	59
5.2 研究展望.....	59
参考文献	61
致谢	65

1 绪论

1.1 研究背景及意义

1.1.1 研究背景

股票市场一直以来都是充满变数和挑战的领域，投资者们追逐着股票价格的波动，试图在市场中获取最大的利润。股票价格的预测一直是金融领域的研究热点之一，越来越多的投资者希望能够更好地把握市场走势，做出明智的投资决策。

股票市场的研究始于 17 世纪，当时阿姆斯特丹证券交易所的成立标志着股票市场正式成为金融体系的一部分。随着时间的推移，股票市场研究逐渐演变为一个复杂而多样的领域，吸引了学者、投资者和政策制定者的广泛关注。17 世纪的荷兰阿姆斯特丹是股票市场的发源地，最早的股票交易所于 1602 年成立。这一时期，荷兰东印度公司的股票开始在市场上交易，投资者们通过购买和出售股票来分享公司的利润。这一创新带动了其他国家的模仿，形成了股票市场的初步雏形。随着工业革命的兴起，股票市场开始进入蓬勃发展阶段。19 世纪中叶，铁路、煤炭和其他产业的兴起推动了大规模的资本需求，股票市场成为企业筹集资金的主要途径。伦敦证券交易所和纽约证券交易所等世界各地的交易所相继成立，标志着股票市场的国际化趋势。20 世纪初，随着投机活动和股市泡沫的不断出现，股票市场面临了多次金融危机，如 1929 年的“黑色星期四”和 1987 年的“黑色星期一”。这些危机引起了对市场监管的关注，各国政府纷纷加强对金融市场的监管和规范，推动了证券法律和监管机构的建立。20 世纪末，信息技术的迅速发展成为股票市场研究的重要推动力。电子交易系统的普及使得交易更加便捷高效，全球范围内的股票市场开始相互联系。同时，投资者和研究者能够更容易地获取和分析大量的市场数据，这为股票价格的预测研究提供了更多的可能性。股票市场研究方法的演变是研究背景中的一个关键方面。早期的研究主要集中在基本面分析，即通过研究公司的财务状况、盈利能力和行业地位等因素来评估股票的价值。随着时间的推移，技术分析逐渐崭露头角，投资者开始关注历史价格和交易量等图表模式，试图从中预测未来的市场走势。近年来，机器学习和人工智

能的发展为股票市场研究注入了新的活力。大数据的普及使得研究者能够利用庞大的数据集进行模型训练,通过深度学习和神经网络等技术挖掘隐藏在数据中的规律。自然语言处理技术的应用也使得研究者能够从新闻、社交媒体等非结构化数据中获取对市场情绪和预期的洞察。这些先进的技术为投资者提供了更全面、准确的信息,提高了股票价格预测的精度。当前,股票市场研究呈现出几个主要趋势。首先,数据驱动的方法变得日益重要,大数据分析和机器学习等技术的应用使得研究者能够更全面地理解市场动态。其次,人工智能技术的不断创新为股票价格预测提供了新的思路和工具,包括基于深度学习的模型和算法。此外,社交媒体分析和情感计算等新兴领域的研究也在不断拓展,为投资者提供更多关于市场情绪和预期的信息。

股票市场研究作为金融领域的核心问题,经历了漫长的历史演变。从最初的交易所创建到全球化和信息技术的崛起,股票市场的研究不断拓展其研究范畴和深度。不同时期的金融危机和市场波动推动了监管体系的建立,而现代技术的应用则为股票市场的进一步研究提供了新的机遇。当前,随着数据科学、机器学习和人工智能的快速发展,股票市场研究正朝着更加复杂和多层次的方向发展。数据科学的兴起为股票市场研究注入了新的活力。大数据分析不仅使研究者能够处理更庞大的数据集,而且还提供了更全面的信息视角。通过对市场行为、交易量、资金流向等多方面数据的分析,研究者可以更好地理解市场的内在规律,为投资者提供更为准确的决策支持。深度学习和神经网络技术的应用对股票价格预测的提高起到了关键作用。这些先进的模型能够学习复杂的非线性关系,从而更好地捕捉市场的动态变化。通过对历史价格、技术指标和其他影响因素的深度学习,模型能够更准确地预测未来的股票价格走势,为投资者提供更具实用性的市场分析。

股票价格预测作为金融领域的核心问题,吸引了学术界和产业界的广泛关注。通过不断创新和引入先进技术,研究者在预测精度和实用性上取得了显著的进展。然而,由于市场的不确定性和复杂性,股票价格预测仍然是一个充满挑战的领域,需要不断深入研究和创新。

1.1.2 研究意义

股票价格预测一直是金融领域中备受关注的话题，其研究意义深远而广泛。从投资者角度来看，准确的股票价格预测可以为其提供更明晰的市场方向，帮助其做出更明智的投资决策。从市场监管者的角度来看，对股票价格的准确预测有助于发现市场异常和操纵行为，维护市场的公正和透明。

股票价格预测研究的理论意义如下所示：

(一) 股票价格预测的研究推动了金融学和市场理论的不断进步。通过对市场数据的深入分析和模型的创新，学者们能够更好地理解市场的运作规律和投资者行为。股票价格预测研究为金融学提供了丰富的实证材料，促进了金融理论的发展和完善。

(二) 本文构建的模型同时针对股票价格序列的高噪声、多特征和长记忆性等特点使用相应的模型进行结合预测，从各方面提高股票价格预测的效果，为股票预测领域提供了较好的预测工具。

股票价格预测研究的现实意义如下所示：

(一) 投资者在市场中面临复杂的信息和多变的市场条件，通过准确的股票价格预测，投资者能够更好地了解市场走势和潜在的投资机会。这对于制定投资策略、调整仓位和选择入市和出市时机都具有重要意义。准确的预测能够帮助投资者规避潜在的风险，最大程度地实现投资收益。

(二) 股票价格预测对于资产配置优化至关重要。投资组合的构建需要考虑各种资产类别的预期收益和风险，而股票作为重要的资产类别之一，其价格的预测直接关系到整个投资组合的表现。通过对股票价格的准确预测，投资者可以更有效地分散投资风险，实现资产配置优化，提高整体投资组合的收益水平。

(三) 对股票价格的准确预测有助于市场监管机构及时发现市场异常和操纵行为。通过监测市场价格与实际价值的偏离，监管机构可以及时采取措施，维护市场的公正、公平和透明。此外，对股票价格的准确预测也提高了信息披露的价值，促使上市公司更加规范地披露信息，降低信息不对称对市场的不良影响。股票价格预测对企业融资和发展同样具有深远的影响。企业通过发行股票融资，其股票价格的表现直接关系到融资的成本和融资规模。准确的股票价格预测有助于企业更好地选择合适的融资时机和方式，提高融资的效率。

1.2 国内外研究现状

1.2.1 股价预测研究现状

长期以来,学术界对股价的研究始终是一个热门话题。基本面分析法和技术分析,这两种传统的投资预测方法都曾是股票价格预测研究中的中流砥柱,到现在为止,投资者们还在不断对其进行着深入地探讨和研究。其中,基本面分析法主要是对影响股票价格走势的一些基本要素进行分析。而另外一种传统方法是技术分析法,技术分析是另一类广泛应用于股价预测的方法,其主要依据是股价走势和交易量的图形模式。图形模式的识别和统计是技术分析的核心,包括头肩顶底、双顶底、突破形态等。技术分析注重历史数据的图表和指标,通过观察这些模式来预测未来的市场走势。然而,技术分析被一些学者批评为缺乏理论基础和科学性,其依赖于主观判断和历史数据的重复性,对于市场的非理性和突发事件缺乏应对能力。后来研究者开始采用基于统计的方法进行股价预测。典型的方法包括时间序列分析、回归分析和蒙特卡洛模拟等。时间序列分析通过对历史股价数据的趋势和周期性进行建模,来预测未来的股价变化。回归分析则通过考察与股价相关的各种因素,如财务指标、宏观经济数据等,建立回归模型进行预测。吴玉霞(2016)等通过 ARIMA 模型对创业板股票趋势进行了短期预测取得了较好的预测效果。张贵生(2016)等构建基于梯度因子的 G-ARMA-GARCH 模型,通过对上证综指、深证成指、香港恒生和标准普尔 500 指数收盘价收益率的实证研究表明,融合了梯度因子信息的 G-ARMA-GARCH 模型预测精度显著提高,而且具有良好的稳定性。罗志丹等利用一种基于跟踪微分器的泰勒展开预测(Taylor Expansion Forecasting,TEF)模型与 ARIMA 相结合的混合预测模型,预测未来股票价格的走势。这些方法在某些情况下能够提供一定的预测精度,但受限于对数据分布和关系的假设,对于复杂的非线性关系和非稳态的市场来说,其表现较为有限。

近年来,随着大数据技术的兴起和计算能力的提升,基于机器学习和深度学习的方法在股价预测领域得到了广泛应用。机器学习和深度学习等方法不仅能够处理大规模的数据,而且能够挖掘数据中的复杂模式和非线性关系。Zhang(2019)等采用改进的多目标优化算法和非线性集成策略的两阶段预测系统,显著地提高

了风速预测能力。董曼茹(2023)等将双层 Attention 机制引入到 LSTM 神经网络结构,构建 ATT-LSTM-ATT 模型应用于 CPI 预测,有效提升了模型对 CPI 预测的精准度。除此之外,在数据密集型的金融业中,机器学习的影响同样也尤为深远。最近几年,在金融数据分析中,有越来越多的学者使用机器学习技术对金融时间序列进行预测,并且都取得了比传统分析方法更好的结果。谢游宇(2023)等为了提高金融时间序列的预测精度,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,验证了经验模态分解和奇异谱分析对于预测模型的提升效果。杨静凌(2022)等在 Elman 神经网络的基础上引入集合经验模态分解算法和 Adaboost 算法,使用 Adaboost 算法优化 Elman 神经网络,然后预测各分量并加总预测结果,最终表明融合 EEMD 算法和 Adaboost 算法的 Elman 神经网络在股价预测中能发挥更好的预测效果。何泳(2023)等为了提高股票预测的精准度并减少运行时间,提出了一种改进的非支配排序遗传算法与极致梯度提升树模型相结合(INSIGA-III-XGBoost)的股票预测模型,与其他模型相比精确度更高、运行用时大幅减少,并且该模型具有可解释性。韩莹(2023)等针对股票序列不平稳特点,引入互补集成经验模态分解(CEEMD)进行降噪处理,提出 CEEMD-LSTM-BLS(C-L-B)股票预测模型。选取农林牧渔行业股票价格,提出的新模型与现有模型相比,可以很好的解决拟合差、时滞等问题。赵帅斌(2023)等为了提高预测精度,针对股票价格数据的时序性特征,提出用结合经验模态分解(EMD)、投资者情绪和注意力机制的双向长短期记忆神经网络来对股票价格进行涨跌预测,取得了较好的预测效果。邓德军(2023)等提出一种基于 E-V-ALSTM 混合深度模型的股票价格预测方法,充分发挥 EMD、VMD 和注意力机制的作用,提高了 LSTM 模型在股价预测中的效果。刘洋(2023)等基于多模态深度学习方法,通过分析旅游评论中消费者情绪,预测酒店股票的价格走势,结合 Yelp 的旅游评论数据进行实证研究,并与相关基线模型作比较。实验结果表明,本文提出的多模态模型具有优越性。方义秋(2022)等提出了一种联合 RMSE 损失 LSTM-CN 的方法,创新性地通过联合两个模型的 RMSE 损失函数,融合了 LSTM 学习长期时间序列依存关系的优点和 CNN 提取数据中深层特征的优点,在预测效果上具有良好的可行性和普适性。徐月梅(2021)等引入卷积神经网络和双向长短时记忆模型挖掘财经新闻中的新闻事件类型和新闻情感倾向,提出一种

深度融合股市财务数据、新闻事件特征及新闻情感特征的股票预测模型,模型的预测准确率进一步提升。张倩玉(2021)等提出一种结合深度学习和分解算法的股票价格预测模型引入自适应噪声的完整集成经验模态分解(CEEMDAN)算法和注意力机制提高模型拟合能力。綦方中(2020)等针对基于 BP 神经网络的股票价格预测模型在价格预测时存在较大误差的问题,在 BP 神经网络方法的基础上引入了主成分分析方法(PCA)和改进的果蝇算法(IFOA),提出一种基于 PCA-IFOA-BP 神经网络的股票价格预测模型,提高了 BP 神经网络的预测精度。包振山(2020)等提出 LSTM-GA 模型,将遗传算法(GA)解决调参问题来保证模型预测的平衡性得出改进模型的各方面指标均优于单独的 LSTM 模型。陈凯杰(2022)等提出 EE-MFCBA/L-L 股指收益率预测模型使用 MFCBA/L 模型实现对模态分量的预测,最后采用 LSTM 模型非线性集成各分量的预测结果.实证结果表明,所提出的模型可以更好地适应收益率变化特征。张婷婷(2021)等采用粒子群算法对核极限学习机的参数进行寻优,建立粒子群算法优化后的核极限学习机股票指数预测模型。结果表明,所构建的模型在股票指数预测的结果中,相比传统的极限学习机模型表现出预测性能的优越性和稳健性。邹婕(2023)等在融合自注意力机制(SA)和 GRU,构建 SA-GRU 模型的基础上,引入降维处理技术随机森林(RF)算法,将经过降维的股票因子数据作为 SA-GRU 模型的输入,实验得出 RF-SA-GRU 模型在股指预测方面依旧具有更好的预测性能。Thomas(2018)等则是将研究的重心放在分析时间序列以及提取信息上,引入了长短期记忆神经网络(LSTM)进行股价预测分析,发现了其在时间序列分析中的优越性,也进一步表明了 LSTM 由于其独特的结构和处理时间信息的能力在股价预测及分析研究中都能取得较好的结果。闫政旭(2021)等使用 Pearson 相关系数法在特征集合进行相关性分析得到最优特征子集,然后运用随机森林模型在股指以及多支个股上进行实证分析得到了更好的预测效果。王伦(2020)等使用深度森林算法对股票市场中的上涨股票进行预测,挑选预测趋势为上涨状态的股票进行量化投资并取得了超过其他算法的收益率。张雪芳(2023)等使用 XGBoost 算法对股票指数的涨跌趋势进行了预测取得了较高的精确度。李新尧(2023)使用 SSA 算法对沪深 300 股票指数进行降噪处理,然后将处理后的数据通过 LSTM 模型进行训练并预测,实验表明经降噪处理后的数据对于 LSTM 模型的预测效果有着很大的提高。杨青(2019)等则在基础的 LSTM 模

型上进行了优化,加深了神经网络的层数以提高模型的泛化能力,在实证研究中选择了多种股票数据和多种数据频率,运用优化后的 LSTM 进行股价预测,并与传统时序法、机器学习模型进行对比,发现优化后的 LSTM 可以在大多数股价预测中取得更好的效果。除了 LSTM 神经网络模型之外,还有人工神经网络模型、卷积神经网络、数据挖掘算法、多层感知机、递归神经网络、半监督深度学习方法等等机器学习方法将股票、期货等金融产品的价格预测精度提升到了新的高度。欧阳红兵(2020)等考虑对输入数据进行优化,引入了小波分解算法,进一步增强了模型对股价的预测能力。景楠(2021)等则考虑将不同的深度学习模型结合起来,不同的模型处理不同的部分,创造性地将 CNN 模型和 LSTM 模型结合起来,两种模型各司其职,分别处理各自更擅长的部分,取得了优秀的股价预测结果。而赵红蕊(2021)等在结合 LSTM 和 CNN 的基础上在 CNN 端引入注意力机制 CBAN 卷积模块,有效提升了特征提取能力。无独有偶,贺毅岳(2020)等以及 Lin(2021)等不约而同地考虑在进行股价预测之前先对数据进行降噪处理,同时使用了 CEEMDAN 算法,并运用 LSTM 模型对降噪后的数据进行预测,在股价预测问题上将预测精度提升到了一个更高的高度。牛红丽(2022)等将 Bagging 算法和 GRU 模型融合起来对股指数据进行预测,相比于单一模型极大地减小了预测误差且提高了稳定性。邹婕(2023)等将自注意力算法 SA 和基础模型 BiGRU 融合起来对股票价格趋势进行预测,极大程度提高了预测的精确度和稳定性。Lu W(2020)等提出了一种 CNN-BiLSTM-AM 方法来预测第二天的股票收盘价,使用 CNN 提取输入数据的特征,利用 AM 方法捕捉过去不同时间特征状态对股票收盘价的影响,最后通过 BiLSTM 模型对处理后的股价时间序列进行预测,与其他方法相比所提模型大大提高了股价时间序列的预测效果。Yu(2020)等选取了一种局部线性嵌入降维算法(LLE)来减小股价影响因素的维度,并与 BP 神经网络相结合,实现了高精度的股价预测。Bose(2021)等通过多元自适应回归样条(MARS)和深度神经网络(DNN)相结合的方法,在不影响精度的情况下减少了维度特征。除此之外,还有学者极具创造力地运用不同种类的数据、不同的特征选取方法为股价预测提供了不一样的思路。Zhou 等(2020)采用除历史交易数据外的多种异构数据对股价变化方向进行预测,实现了专注于活跃和非活跃股票的独特预测模式。Peng(2021)等应用了三种特征选择方法来缩小特征集,旨在消除指标

中的冗余信息。

综上所述,根据以往学者的研究,由于深度学习技术完全由数据驱动,根据大量数据进行训练学习,不易受到外界条件的限制,深度学习技术在股票预测领域取得了较好的预测效果,是当下预测股票价格趋势的有力工具。卷积神经网络在处理多特征数据时能起到较好的效果,长短期记忆网络则对于时间序列数据中的长记忆特性能发挥较好的作用,这些深度学习模型对于非平稳多特征的股价时间序列数据的预测表现出很好的效果,但以往学者的研究中较少有人同时将数据序列降噪、特征提取和长记忆性处理结合起来以进一步提高股票价格的预测效果,因此本文将序列降噪效果优秀的 VMD 算法、多特征处理效果优秀的 CBAM 算法与长记忆性处理效果优秀的 BiGRU 模型融合起来对股票变动趋势进行预测。

1.2.2 群智能算法研究现状

群智能优化算法是一类基于群体智能的计算方法,通过模拟群体中个体之间的合作和信息共享来实现问题的求解。这类算法模拟了自然界中群体协同行为的原理,包括鸟群、蚁群、鱼群等动物群体的行为。也正因为如此,群智能优化算法在优化领域得到了快速地发展。在各种优化算法中,学者们还观察了许多自然界中生物的社会行为,如鱼、鸟和昆虫的生物特性,以此为基础提出了许多新的算法。例如,人工蜂群算法(ABC)、萤火虫算法(FA)、果蝇优化算法(FOA)、灰狼优化算法(GWO)、鲸鱼优化算法(WOA)、天牛须搜索算法(BAS)以及最新提出的麻雀搜索算法(SSA)等。程杰(2023)等提出一种基于 GWO-CNN-BiLSTM 的超短期风电预测方法,通过灰狼优化算法(grey wolf optimizer, GWO)对组合模型进行优化,使该组合模型参数能实时适应风电历史数据。最后,仿真结果验证了所提出方法的有效性和优越性。张新生(2023)等使用 GWO 算法优化 KELM 正则化系数(C)和核函数参数(γ),对比分析得出 LASSO-GWO-KELM 模型预测值与实际值拟合优于其他模型,能够更为准确地预测工业碳排放量。史加荣(2024)等提出了一种基于变分模态分解、双向长短期记忆网络以及鲸鱼优化算法的混合深度学习模型,以用于短期风电功率预测,使用 WOA 对 BiLSTM 中的参数进行优化,提高了混合模型的效率和预测性能。王瑞(2023)等提出基于自适应噪声完备经验模态分解和鲸鱼算法优化支持向量回归(CEEMDAN-WOA-SVR)的铁路沿线风速

预测方法,采用鲸鱼优化算法(WOA)优化支持向量回归(SVR)模型的惩罚因子和核参数,提高了 SVR 模型的表现能力。王少钦(2023)等提出一种 EMD-SSA-BiLSTM 模型,通过麻雀搜索算法(SSA)对其超参数的设置进行优化,从而实现更好的短期风速预测效果。赵全明(2023)等为了进一步提高输出功率预测的精度和鲁棒性,采用深层神经网络技术,提出了基于参数寻优的麻雀搜索算法—门控循环单元(SSA-GRU)超短期混合风电功率预测模型,采用 SSA 对整个模型进行参数最优值搜索,将原有单一模型加入优化迭代组成混合算法模型,提高了模型的预测性能。张成龙(2020)等提出了一种多种群自适应果蝇优化算法,引入自适应惯性权重,并应用于广义回归神经网络参数优化,提高了模型在轴承寿命预测中的精度和效率。除了受到自然界中生物群体行为的启发,一些算法还源自于自然进化的原理以及物理规则,例如遗传算法(GA)、模拟退火(SA)等等。施爱平(2023)等将遗传算法(GA)加入长短期记忆神经网络(LSTM)预测模型的参数调节环节中,建立基于霜霉病—气象时序数据的手动调参 LSTM 模型和 BP 神经网络模型,提高了 LSTM 模型的预测精度。全一鸣(2023)等提出基于分形优化的变分模态分解(VMD)和遗传算法(GA)改进的反向传播神经网络(BP)模型的短期风速预测方法,采用遗传算法改进的 BP 神经网络显著提高预测效果,降低风速预测误差。陈石毓(2023)等提出新的负荷预测方法,先对数据进行模糊 C 均值聚类,然后将聚类结果分别传入由遗传算法、自注意力机制和卷积长短时记忆神经网络相结合的模型进行预测,经遗传算法优化后的模型提高了预测精度。许多群体智能优化算法不但可以获得更好的性能指标,还可以用于各个学科以及各种领域,并在这些领域中都取得了难以替代的成果。Keun(2021)等使用 GA-XGBoost 混合模型,通过遗传算法 GA 在几十个股票技术指标中进行最优特征数量的选取,将获得的最优特征集和原始特征集进行比较。验证了特征工程对预测结果有显著的提高。Chandar(2020)等使用灰狼优化算法 GWO 对 ENN 神经网络中的大量参数进行寻优处理,经过实证发现 GWO-ENN 模型相比于其他模型具有更高的预测精度。Vaishali(2021)等从新闻数据中提取特征结合指标数据对未来几天的股价使用集成深度学习框架进行预测从而提高了模型的预测表现。Kumar(2021)提出了九种新的综合模型用于预测盘中股价,其中包括 BP、ANN、TDNN、RBFNN 等模型,以及 GA、PSO、ABC 等群体智能优化算法,结果表明 PSO-BPNN 模型的预测准

确率最高。除此之外，也有学者在现有算法的基础之上进行改进，以获得更好的收敛效果和预测精度。

综上所述，群智能优化算法被广泛应用于模型参数的优化中，对于模型效果有着显著的提升，深度学习模型的构件中则含有大量参数，人工选择参数不仅耗时较长还不能达到较好的效果，引入群智能优化算法后则可以在很大程度上提高模型的效率。但大多群智能优化算法的全局寻优能力有限，很多学者并未考虑寻求方法提高优化算法的全局寻优能力，因此通过一定的方法提高优化算法的全局寻优能力是提高模型预测效果的有效途径，本文选择了群智能优化算法中全局寻优能力较强的粒子群算法，并且对算法增加了动态改进从而进一步提高了粒子群算法的全局寻优能力。

1.3 主要内容

本论文主要研究的内容是通过双向门控循环神经网络(BiGRU)来对股票价格进行预测，选取了 A 股市场中的上证综指、沪深 300、中国平安和茅台四支著名股指和股票进行实验验证，在双向门控循环神经网络(BiGRU)的基础上融合了变分模态分解(VMD)和注意力模块(CBAM)构建了 VMD-CBAM-BiGRU 模型，通过一系列实验验证了变分模态分解(VMD)和注意力模块(CBAM)对双向门控循环神经网络(BiGRU)模型在股票预测中的提升作用。具体内容安排如下。

第一部分为绪论及文献综述。这一部分首先介绍本文基于 VMD-CBAM-BiGRU 模型对股票市场中股价预测研究的相关背景以及意义。然后针对国内外股价预测以及群智能优化算法领域的相关文献进行概述分析，并从这些文献中总结出可借鉴应用的地方。

第二部分为对涉及到的模型的相对应理论的介绍。这部分对股价的影响因素、群智能优化算法以及深度学习理论的基础知识进行归纳概述，从而为模型的构建及实验奠定理论依据。

第三部分为对 VMD-CBAM-BiGRU 模型的构建，分别涉及相关系数筛选特征、模糊熵重构分量、参数的设定及优化、模型评价指标、VMD-CBAM-BiGRU 模型的构建思路以及 VMD-CBAM-BiGRU 模型的构建流程等六部分。通过使用 Pearson 相关性分析对基本行情指标以及技术指标进行筛选，剔除掉无关变量；

然后使用模糊熵对变分模态分解(VMD)后的各子序列进行重构以提高预测效率；再通过改进后的粒子群优化算法对 BiGRU 模型中的各项超参数进行寻优提高模型表达效果；最后构建 VMD-CBAM-BiGRU 模型。

第四部分为实证分析和对比研究。这一部分做了三个不同模块的实证分析，首先是将构建好的 VMD-CBAM-BiGRU 模型以及一系列对比模型代入四支股票数据集进行试验，得到各模型的评价指标结果以及拟合效果图，通过对比分析确定 VMD 算法和 CBAM 算法对于 BiGRU 模型的提升效果，从而印证 VMD-CBAM-BiGRU 模型在股价预测中的稳定性和优越性；然后是从次日收盘价、三日收盘价以及七日收盘价三个时间跨度的收盘价分别作为预测对象来分析所构建 VMD-CBAM-BiGRU 模型的长中短期的预测能力差异；最后则是将本文所构建 VMD-CBAM-BiGRU 模型作为投资策略与其他交易策略进行对比分析验证模型实际应用中的效果。

第五部分为论文内容的总结以及展望部分。总结了本文的主要研究内容以及研究意义，并对本文所构建模型的改进空间进行展望。

1.4 创新之处

本文在选择合适模型的基础上同时注重模型融合以及模型优化，通过将能从不同视角提升模型预测效果的模型进行融合，使用改进后的粒子群优化算法进行基础模型的优化，从两方面并行提升模型的预测能力，使投资者可以规避风险获取更多收益、管控部门提前做出股市预警。因此，本文主要创新点包括以下内容：

第一，本文选择 VMD 作为原始股价的分解算法，有效避免了 EMD 分解算法出现的模态混叠和端点效应等问题，并且在 VMD 算法分解序列后，使用模糊熵算法计算各模态分量的熵值，依据熵值将子序列进行重构，提高了模型预测效率，降低了模型运算的复杂度。选择 CBAM 注意力模块对模型的特征进行处理，增强了模型的特征提取能力。

第二，选择全局寻优能力较强的粒子群算法进行基础模型 BiGRU 的超参数寻优处理，并且本文在粒子群算法的基础上做了动态的改进，进一步提升了粒子群算法的全局寻优能力。

第三，选择具有双向时间层的 BiGRU 模型作为预测的基础模型，能够出色

地处理股票时间序列数据，通过将 VMD 算法、CBAM 算法和基础模型 BiGRU 进行结合，构建了 VMD-CBAM-BiGRU 模型，融合三个模型各自的特点后进一步提高了股价预测的精确度和稳定性，为股票预测领域提供了新的思路。

2 相关理论与研究方法

2.1 股票市场相关概述

2.1.1 股票价格的影响因素

股票价格的波动受到多方面因素的影响,这些因素涵盖了公司内部的基本面因素、宏观经济环境的因素以及市场行为的技术面因素。深入了解这些因素对股票价格的影响,有助于投资者更好地理解市场,做出更明智的投资决策。

(一) 基本面因素

(1) 公司业绩

公司业绩是影响股票价格最为直接的基本面因素之一。投资者通常通过分析公司的财务报表,关注其营收、利润、现金流等指标,以评估公司的盈利能力和财务健康状况。强劲的业绩通常会推动股价上涨,反之则可能导致股价下跌。

(2) 行业前景

不同行业的前景差异很大,而股票价格往往受到所处行业的整体影响。新兴产业的快速发展和前景看好可能推动相关股票价格上涨,而传统行业的不景气可能导致股价下跌。因此,投资者需要密切关注不同行业的发展趋势。

(3) 宏观经济因素

宏观经济因素对整个股市有着深远的影响。国家的经济增长率、通货膨胀水平、货币政策等都直接影响企业的盈利和市场的投资氛围。较为健康的宏观经济环境通常有利于股市的稳健发展,反之可能导致股市波动。

(4) 利率水平

利率水平是资本市场中一个至关重要的基本面因素。低利率环境通常促使投资者倾向于将资金配置到股票市场,因为相较于固定收益工具,股市的回报率更有吸引力。相反,高利率可能会使固定收益工具更具吸引力,导致股市流动性减弱,股价下跌。

(5) 政治因素

政治因素对股票市场同样具有显著影响。政治的不稳定性、政策的不确定性可能导致投资者对市场的信心下降,引发抛售潮,从而影响股票价格。政府政策

的变化和宏观政治背景都需要成为投资者关注的焦点。

(6) 公司治理

公司治理质量是公司内部基本面因素中的一个关键方面。良好的公司治理结构能够提高公司的透明度、降低经营风险，使投资者更有信心投资。投资者通常更愿意购买那些治理结构健全、信息披露完善的公司股票。

(二)技术面因素

(1) 技术指标

技术指标是投资者用来分析股票价格走势的重要工具。常见的技术指标包括移动平均线、相对强弱指标(RSI)、随机指标等。这些指标通过对历史价格和成交量的分析，为投资者提供了预测市场走势的参考依据。

(2) 成交量

成交量是技术分析中一个关键的因素，反映了市场中买卖双方的活跃程度。成交量的变化可以帮助投资者判断市场的力量和趋势的持续性。在股价上涨时伴随着较大的成交量，可能意味着市场的强势，反之可能意味着弱势。

(3) 趋势分析

趋势分析是技术分析的核心。通过观察股票价格的长期趋势，投资者可以更好地判断市场的发展方向。上升趋势、下降趋势和横盘市场都是趋势分析的重要概念，对投资者选择合适的交易时机具有指导意义。

(4) 交易信号

技术分析不仅提供对市场趋势的认识，还可以生成一些交易信号。例如，当短期移动平均线上穿长期移动平均线时，可能产生买入信号。这些信号对投资者制定具体的买卖策略有着指导作用。

(5) 市场情绪

市场情绪是技术面因素中的一个非常重要的组成部分。投资者的情绪波动会影响其交易决策，从而影响股票价格。恐慌性抛售、过度乐观等情绪波动可能导致市场剧烈波动，是投资者需要警惕的因素。

(三)其他因素

(1) 外部事件

外部事件对股票价格的影响无法忽视。自然灾害、地缘政治紧张、全球性事件等都可能引发市场的剧烈波动，甚至引发系统性风险。投资者需要密切关注外部环境的变化，及时调整投资策略。

(2) 全球市场走势

全球市场的走势也对股票价格产生重要影响。国际间的资本流动、全球经济格局的演变都会波及到各国股市。投资者需要综合考虑全球市场的动向，以更全面地把握股票价格的未来走势。

(3) 市场流动性

市场流动性是指市场上可以迅速买卖大量股票而不对股价造成过大波动的程度。较高的市场流动性通常有利于投资者进行交易，而低流动性可能导致交易成本上升，影响投资者的买卖决策。

(4) 投资者情绪和行为

投资者的情绪和行为对市场也有重要影响。投资者的羊群行为，即跟风效应，可能导致市场出现过度买入或过度卖出的情况。理性分析和对市场情绪的合理把握对投资者至关重要。

综上所述，股票价格的波动是一个复杂而多层次的过程，受到基本面因素、技术面因素以及其他外部因素的共同影响。投资者在制定投资策略时，应该全面考虑这些因素，并结合自身风险偏好和投资目标，制定合理的买卖策略。理性分析市场，深入了解股票价格背后的因素，有助于投资者更好地应对市场的波动，取得更为稳健的投资收益。

2.1.2 股票技术指标概述

股票技术指标是投资者在股票市场中用来分析和预测价格趋势的工具，通过使用特定计算公式对股票基本行情指标进行计算后得到，这些指标基于过去的价格和交易量数据，旨在揭示市场的技术特征，为投资者提供决策依据，股票市场中常用的几类技术指标包括能量类指标、运动趋向类指标、超买超卖类指标、成交量类指标和均线类指标等。

(1) 能量类指标

在股票市场中，关于“能量类”指标可能指的是与股票价格、成交量、趋

势等相关的技术指标,这些指标通过测量市场的能量或活力来提供投资者有关股票走势的信息,主要的能量类技术指标包括 CR、VR、MASS、NVI、PSY 和 BRAR 等,这些能量类指标提供了对市场活力和趋势的不同视角。

(2) 超买超卖类指标

超买和超卖是股票市场中的两种技术分析概念,用于指导投资者判断股票价格可能的转折点。超买指的是股票或其他资产的价格相对于其内在价值被过度推高,可能会发生价格回调或下跌;而超卖则表示股票或资产的价格相对于其内在价值被过度低估,可能迎来反弹或上涨,常见的超买超卖类指标包括 WR、ROC、KDJ、BIAS、SKDJ、MTM 等。

(3) 均线类指标

均线类技术指标是股票市场中常用的分析工具之一,通过计算股票价格的平均值,以图形的方式展示价格趋势,将一定时期内的股票价格加总,然后除以该时期的长度得到的平均值。根据计算方式和时间跨度的不同,均线主要分为简单移动均线(SMA)和指数加权移动均线(EMA)两种。

(4) 成交量能类指标

成交量是股票市场中的重要指标之一,指在一定时间内市场上交易的股票数量,通常以股或手为单位。成交量可以用来衡量市场的活跃度,大量交易通常伴随着市场的波动和趋势变化,它反映了交易市场上的活跃度和资金流向,在技术分析中具有重要的地位。成交量能类股票技术指标通过对成交量数据的分析,帮助投资者更好地理解市场趋势、价格变动和潜在的买卖信号。常见的成交量能类指标主要包括 OBV、VOL 和 AMO 等。

(5) 运动趋向类指标

运动趋向类股票技术指标主要用于帮助投资者判断股票价格的趋势方向和趋势的强度。其在股票价格涨跌幅度之内,分析供求关系之间的均衡点,即供需关系受价格的变动由均衡到失衡的循环过程,从而判断股票价格趋势判断这些指标通过分析一定时期内的价格变动情况,提供有关市场走势的信息。常用的运动趋向类指标包括 MACD、DPO 和 DMI 等。

2.2 群智能优化算法

2.2.1 群智能优化算法的介绍

群智能优化算法是一类基于群体智能的计算方法,通过模拟群体中个体之间的合作和信息共享来实现问题的求解。这类算法模拟了自然界中群体协同行为的原理,包括鸟群、蚁群、鱼群等动物群体的行为。群智能优化算法在解决各种优化问题方面表现出色,如粒子群算法、蚁群算法、鱼群算法等。群智能优化算法来源于生物学中对群体智能行为的研究。通过模拟自然界中群体的协同行为,这类算法试图解决复杂问题,如优化问题、搜索问题和模式识别问题。群智能优化算法具有分布式、并行、自适应的特点,能够应对复杂、非线性和高维度的问题。以下是一些常用群智能优化算法的概述。

(1) 人工蜂群算法

人工蜂群算法(Artificial Bee Colony, ABC)是一种基于蜜蜂觅食行为的群智能优化算法。该算法模拟了蜜蜂群体在寻找食物和传递信息过程中的行为,通过群体中蜜蜂的协同努力来解决问题基本原理是通过模拟蜜蜂的觅食和信息传递行为,从而实现问题的优化。在蜂群中,蜜蜂会根据花蜜的丰富程度和距离等因素,选择合适的路径进行觅食。觅食的蜜蜂在找到花蜜后,会返回蜂巢,并通过舞蹈等方式向其他蜜蜂传递信息,引导它们找到花蜜。人工蜂群算法的运行过程可以分为初始化、觅食和信息传递三个主要阶段,在初始化阶段随机生成一群蜜蜂,每只蜜蜂代表一个潜在解决方案,然后初始化蜜蜂的位置,即问题的解空间中的一个点。在觅食阶段蜜蜂根据当前位置计算适应度值,即目标函数的值,根据适应度值选择蜜蜂进行觅食,觅食的方式可以是在当前解的附近搜索新的解,对每只蜜蜂进行适应度值的评估,记录找到的更好解;在信息传递阶段,觅食结束后,根据蜜蜂找到的更好解,更新蜜蜂的位置,通过信息传递,蜜蜂之间共享找到的优秀解,以引导其他蜜蜂在下一次觅食中更有可能找到更好的解,重复进行觅食和信息传递的过程,直到满足停止条件(如达到最大迭代次数、找到满意解等)。

(2) 鲸鱼优化算法

鲸鱼优化算法的基本原理是通过模拟鲸鱼群体在寻找猎物的过程中的协同

行为,实现优化问题的求解。在鲸鱼群体中,鲸鱼通过合作和个体的调整,有效地寻找到食物。算法的灵感来源于鲸鱼的协同觅食行为以及它们在寻找猎物时的追踪策略。鲸鱼优化算法的运行过程可以分为初始化、觅食、追踪和调整三个主要阶段,在初始化阶段随机生成一群鲸鱼,每只鲸鱼代表一个潜在解决方案,初始化鲸鱼的位置,即问题的解空间中的一个点;在觅食阶段鲸鱼根据当前位置计算适应度值,即目标函数的值,根据适应度值选择鲸鱼进行觅食,觅食的方式可以是在当前解的附近搜索新的解,对每只鲸鱼进行适应度值的评估,记录找到的更好解;在追踪阶段在鲸鱼中选择适应度值最好的个体,作为引导鲸鱼,引导鲸鱼通过一定的规则追踪其他鲸鱼,以引导它们朝着更优的解进行移动,根据追踪的结果,调整每只鲸鱼的位置,通过合作和信息共享,群体中的鲸鱼共同调整位置,以更好地探索解空间,重复进行觅食、追踪和调整的过程,直至满足停止条件(如达到最大迭代次数、找到满意解等)。

(3) 麻雀搜索算法

麻雀搜索算法是一种群智能优化算法,灵感来源于麻雀的觅食行为。这一算法通过模拟麻雀个体之间的协作和信息传递,以及麻雀在寻找食物时的策略,实现对优化问题的求解。麻雀搜索算法基于麻雀在觅食时的协同行为和信息传递策略。麻雀在食物稀缺的环境下通过相互协作,有效地寻找到食物。算法尝试模拟这种集体智能行为,通过协作和信息传递来引导群体更好地探索解空间。麻雀搜索算法的运行过程可以分为初始化、觅食、信息传递和调整四个主要阶段;在初始化阶段随机生成一群麻雀,每只麻雀代表一个潜在解决方案,初始化麻雀的位置,即问题的解空间中的一个点;在觅食阶段麻雀根据当前位置计算适应度值,即目标函数的值,根据适应度值选择麻雀进行觅食,觅食的方式可以是在当前解的附近搜索新的解,对每只麻雀进行适应度值的评估,记录找到的更好解;在信息传递阶段,在麻雀中选择适应度值最好的个体,作为引导麻雀,引导麻雀通过一定的规则传递信息,以引导其他麻雀朝着更优的解进行移动;在调整阶段根据信息传递的结果,调整每只麻雀的位置,通过合作和信息共享,群体中的麻雀共同调整位置,以更好地探索解空间,重复进行觅食、信息传递和调整的过程,直至满足停止条件(如达到最大迭代次数、找到满意解等)。

(4) 遗传算法

遗传算法(Genetic Algorithm, GA)是一种模拟自然选择和遗传机制的群体搜索和优化算法。它通过模拟生物进化的过程,利用自然界的演化机制对问题的解进行搜索和优化。其基本原理是受到达尔文的进化论启发,通过模拟自然界中的选择、交叉和变异等遗传机制,以实现问题空间中最优解的搜索。遗传算法适用于解决优化问题,其中个体被编码为染色体,进化的过程通过选择、交叉和变异来模拟。遗传算法的运行过程可以分为初始化、选择、交叉、变异和替代五个主要阶段:在初始化阶段随机生成初始群体,每个个体表示问题的一个可能解,对每个个体进行染色体编码,将问题的解表示为基因序列;在选择阶段根据每个个体的适应度值(目标函数值)进行选择,适应度越高的个体被选中的概率越大,以模拟自然选择的过程,常用的选择方法包括轮盘赌选择、锦标赛选择等;在交叉阶段从选中的个体中随机选择一对进行交叉操作,产生新的个体,交叉的方式可以是单点交叉、多点交叉、均匀交叉等,在变异阶段对产生的新个体进行变异操作,以引入一些随机性,避免陷入局部最优解,变异操作可以是基因位的变换、插入或删除等;在替代阶段将新产生的个体替代原有群体中适应度较差的个体,形成下一代群体,重复进行选择、交叉、变异和替代的过程,直至满足停止条件(如达到最大迭代次数、找到满意解等)。

(5) 粒子群算法

粒子群算法(Particle Swarm Optimization, PSO)是一种群体智能优化算法,灵感来源于鸟群或鱼群等生物体的群体行为。该算法通过模拟群体中个体粒子的协同和信息共享来实现问题的优化。其基本原理是通过模拟群体中个体粒子的协同行为,以寻找问题的最优解。每个粒子表示问题的一个潜在解决方案,它在解空间中移动,并根据个体最优和群体最优的信息进行调整。粒子群算法在全局搜索和局部搜索之间寻找平衡,通过个体间的合作和信息传递来引导搜索过程。粒子群算法的运行过程可以分为初始化、移动、更新和迭代四个主要阶段;在初始化阶段随机生成一群粒子,每个粒子代表问题的一个可能解,为每个粒子随机初始化位置和速度;在移动阶段根据当前位置计算每个粒子的适应度值,即目标函数的值,更新粒子的速度和位置,根据当前速度、个体最优和群体最优进行调整;在更新阶段更新每个粒子的个体最优位置,记录使得适应度值最小(或最大)的位置,更新群体最优位置,记录使得适应度值最小(或最大)的位置;在迭代阶段重

复进行移动和更新的过程，直至满足停止条件(如达到最大迭代次数、找到满意解等)。

2.2.2 动态粒子群算法

1995年，美国学者 Kennedy 和 Eberhart 共同提出了粒子群算法，其基本思想源于对鸟类群体行为进行建模与仿真的研究结果的启发。它的核心思想是利用群体中的个体对信息的共享使整个群体的运动在问题求解空间中产生从无序到有序的演化过程，从而获得问题的可行解。粒子群算法首先初始化一群随机粒子，每个粒子都以一定的速度在 D 维空间中游走更新，每次迭代后都会产生每个粒子的局部最优值以及种群的全局最优值，粒子也根据每次迭代所产生的局部最优值和全局最优值来更新自身的速度和位置，直到一定的迭代次数后达到整个种群的最优值。粒子的更新公式如算式(1)和(2)所示：

$$v_{id}^k = wv_{id}^{k-1} + c_1r_1(pb_{est_{id}} - x_{id}^{k-1}) + c_2r_2(g_{best} - x_{id}^{k-1}) \quad (1)$$

$$x_{id}^k = x_{id}^{k-1} + v_{id}^k \quad (2)$$

粒子第 k 次迭代的速度由第 k-1 次粒子的速度、局部最优值和种群的全局最优值所决定。在速度更新公式中，w 代表粒子每次迭代时速度的权重系数，也叫惯性权重，其值对于最优值的收敛具有很大的影响，通常取[0, 1]范围内的值。c₁为局部学习因子，其值对于个体局部最优的收敛具有较大作用；c₂为全局学习因子，其值影响种群全局最优值的收敛，通常情况下取 c₁=c₂=2。粒子第 k 次迭代后的位置等于粒子第 k-1 次迭代后的位置加上粒子第 k 次迭代时的速度。

动态粒子群算法是粒子群算法的一种改进版本，主要用于解决动态环境下的优化问题。在动态环境中，问题的最优解可能随时间发生变化，传统的优化算法可能无法适应这种变化。动态粒子群算法通过引入机制来应对问题随时间变化的情况，动态粒子群算法基于粒子群算法，通过引入动态性的机制，使得算法能够在问题的解空间中适应变化。在动态粒子群算法中，个体粒子的位置和速度的更新不仅受到当前的适应度值和历史最优值的影响，还受到环境的动态变化影响，以更好地适应问题的变化。动态粒子群算法主要分为四个阶段，在初始化阶段随

机生成一群粒子，每个粒子代表问题的一个可能解，为每个粒子随机初始化位置和速度；在移动阶段，根据当前位置计算每个粒子的适应度值，即目标函数的值，更新粒子的速度和位置，根据当前速度、个体最优和群体最优进行调整，引入动态机制，根据问题的变化对粒子进行调整；在更新阶段更新每个粒子的个体最优位置，记录使得适应度值最小(或最大)的位置，更新群体最优位置，记录使得适应度值最小(或最大)的位置；在迭代阶段重复进行移动和更新的过程，直至满足停止条件(如达到最大迭代次数、找到满意解等)。

在粒子群算法的速度更新公式中， w 、 c_1 和 c_2 均为固定常数，容易让粒子群算法陷入局部最优解。 w 值应该由大到小变化，在算法寻优初期，需要较大的速度增强算法的全局寻优能力，在寻优后期，需要减小速度增强算法的局部寻优能力使其更快达到全局最优解。 c_1 值关系着算法的局部寻优能力，在寻优前期应该较小，后期变大， c_2 值关系着算法的全局寻优能力，在寻优前期应该尽可能大，后期变小。因此本文针对粒子群算法容易陷入局部最优值的缺陷进行了针对性的改进，使 w 、 c_1 和 c_2 值进行动态的变化，从而增强算法的寻优能力。因此本文针对粒子群算法中的固定值 w 、 c_1 和 c_2 值引入余弦函数进行了动态的改进，如公式(3)、(4)和(5)所示：

$$w = 0.8 \cos\left(\frac{\pi \cdot t}{T}\right) \quad (3)$$

$$c_1 = 2 \cos\left(\frac{\pi \cdot (1 - \frac{t}{T})}{2}\right) \quad (4)$$

$$c_2 = 2 \cos\left(\frac{\pi \cdot \frac{t}{T}}{2}\right) \quad (5)$$

2.3 分解集成理论

2.3.1 经验模态分解

经验模态分解(Empirical Mode Decomposition, EMD)是一种自适应的信号分

解方法，由黄锷在 1998 年提出。它在信号处理领域取得了显著的成就，特别适用于非线性和非平稳信号的分析。EMD 的核心思想是将信号分解成一组本地特征尺度的固有模态函数(Intrinsic Mode Functions, IMF)，每个 IMF 代表了信号中的一个局部振动或波动。EMD 的基本原理建立在两个基本假设之上：假设一：信号是由多个不同频率的振动叠加而成的；假设二：每个振动模态(IMF)的频率随时间变化。根据这两个假设，EMD 通过一系列迭代步骤将信号分解为 IMF。每个 IMF 的频率范围相对较窄，因此更容易反映信号中的局部特征。IMF 满足以下两个条件：条件一：在整个信号段上，极值点的个数和零交叉点的个数之差最多为 1；条件二：在任何给定点，局部均值为零。EMD 基本步骤如下：

步骤 1：极值点的提取。首先，从信号中提取局部最大值和局部最小值，形成信号的上、下包络线。

步骤 2：插值。对提取的极值点进行插值，得到信号的上、下包络线。插值可以采用线性插值或其他插值方法。

步骤 3 提取局部均值。将原信号减去其上包络线，得到一个局部均值为零的信号。

步骤 4：判断是否满足 IMF 条件。判断上一步得到的信号是否满足 IMF 的定义，如果是，则认为找到一个 IMF；如果不是，则重复步骤 2~4，直到满足条件。

步骤 5：重复步骤 1-4。重复以上步骤，直到原信号被分解为一系列 IMF。

每个 IMF 都是一个本地特征，具有明显的频率和振幅调制特性。如图 2.1 所示，通过将信号分解为 IMF，可以更好地捕捉信号中的局部振动和波动，使得信号的分析更具细节和准确性。EMD 分解相比于小波分解等传统分解方法具有很多优点，比如适用于非线性和非平稳信号的分析，无需先验知识，是一种数据驱动的方法，还能够提供信号的局部特征，更容易捕捉细节。但 EMD 分解并不是较为完美的分解算法，它同时也存在一些不可避免地缺点，比如存在模式混叠和端点效应等问题，需要采取一些改进方法来解决，并且算法的计算复杂度较高，可能导致计算量过大。

经验模态分解作为一种灵活而强大的信号分解方法，对于处理非线性和非平稳信号具有独特的优势。通过将信号分解为局部特征，EMD 为信号分析提供了

一种全新的视角。然而，仍然需要在实际应用中注意其存在的问题，并结合改进方法进行使用，以更好地适应各种复杂信号的分析需求。为了解决 EMD 存在的模式混叠和端点效应等问题，一些改进的方法被提出，如集合经验模态分解(Ensemble Empirical Mode Decomposition, EEMD)和变分模态分解(Variational Mode Decomposition, VMD)。

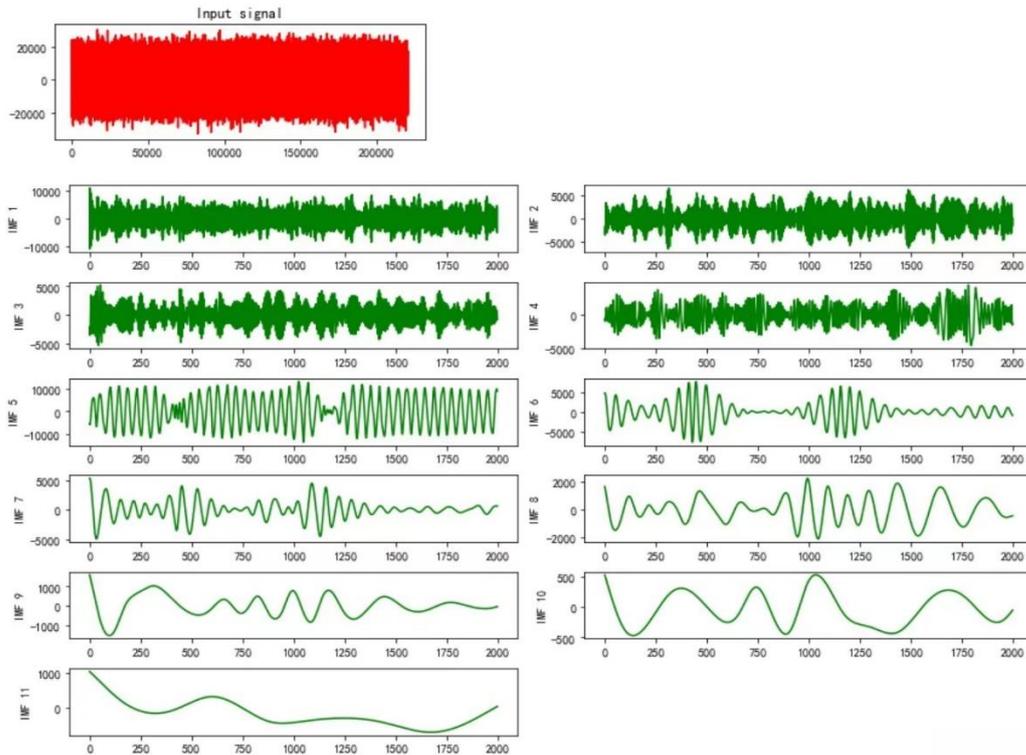


图 2.1 原序列分解为各频率 IMF 分量

2.3.2 变分模态分解(VMD)

变分模态分解(Variational Mode Decomposition, VMD)是一种信号分解方法，由 Dragomiretskiy 和 Zosso 于 2014 年提出。相较于经验模态分解(EMD)，VMD 引入了变分原理，通过优化问题的方式得到信号的不同频率分量。它是一种自适应的、完全非递归的信号处理方法，可以通过控制带宽来抑制模式混叠现象。与 EMD 原理不同的是，VMD 假设每个模态的中心频率带宽有限，并在各模态之和等于原始信号的约束下，使各模态估计带宽之和最小。继 Dragomiretskiy 和 Zosso(2013)之后，约束变分问题表示为：

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s. t. \sum_{k=1}^K u_k(t) = y(t) \end{cases} \quad (6)$$

其中 $y(t)$ 代表股票价格， u_k ($k = 1, 2, \dots, K$)代表所有模态分量， ω_k 是它们的中心频率， t 为时间脚本， $\delta(t)$ 为狄拉克函数； $*$ 为卷积运算且 $j^2 = -1$

为了将有约束问题转化为无约束问题，引入二次惩罚项和拉格朗日乘子 λ ，得到如下增广拉格朗日函数：

$$\begin{aligned} L(\{u_k(t)\}, \{\omega_k(t)\}) = & \alpha \sum_{k=1}^k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \|y(t)\|_2^2 - \\ & \sum_{k=1}^K u_k(t) \|_2^2 + \langle \lambda(t), y(t) - \sum_{k=1}^K u_k(t) \rangle \end{aligned} \quad (7)$$

其中， α 是平衡参数，模态 u_k 在频域中可以使用公式(8)进行更新，中心频率 ω_k 可以使用公式(9)进行更新， λ 可以使用公式(10)进行更新。

$$\hat{u}_k^{n+1} = \frac{\hat{y}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (8)$$

$$\hat{\omega}_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (9)$$

$$\hat{\lambda}^{n+1}(p) = \hat{\lambda}^n(\omega) + \tau(\hat{y}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)) \quad (10)$$

相比于 EMD 分解算法，VMD 将信号分量的获取过程转移到变分框架内，引入了正则化项，更好地保持了信号的频率分离性，采用一种非递归的处理策略，通过构造并求解约束变分问题实现原始信号的分解，并且分解得到的模态之间是正交的，有利于更精确地表示信号。能有效避免模态混叠、过包络、欠包络、边界效应等问题，具有较好的复杂数据分解精度及较好的抗噪声干扰等优点。

2.4 深度学习理论

2.4.1 BP 神经网络

人工神经网络(Artificial Neural Network, ANN)是一种模拟人脑神经元网络结构的计算模型,其中 Backpropagation(BP)神经网络是最为经典和广泛应用的一种类型。BP 神经网络在模式识别、预测、分类等领域取得了显著的成功。BP 神经网络是一种前馈式神经网络,它由输入层、隐藏层和输出层组成。每一层包含多个神经元,神经元之间存在权重,通过这些权重进行信息传递。BP 神经网络的核心思想是通过学习调整权重,使网络的输出更接近于期望输出,从而实现模式识别和函数逼近等任务。

在 BP 神经网络中,每个神经元都有一个激活函数,常用的激活函数包括 Sigmoid 函数、ReLU 函数等。神经元的输入是上一层神经元的输出,通过权重相乘后进行加权求和,再经过激活函数得到输出。数学表达式如下:

$$y = f(\sum_{i=1}^n w_i \cdot x_i + b) \quad (11)$$

其中, y 为神经元的输出, w_i 为权重, x_i 为输入, b 为偏置项, f 为激活函数。

BP 神经网络的网络结构包含输入层、隐藏层和输出层。如图 2.2 所示,输入层接收外部输入,输出层产生最终的输出,隐藏层在输入层和输出层之间进行信息处理。输入层接收外部输入,每个输入对应一个神经元;隐藏层位于输入层和输出层之间,进行信息的中间处理,可以有多个隐藏层;输出层产生网络的最终输出,每个输出对应一个神经元。

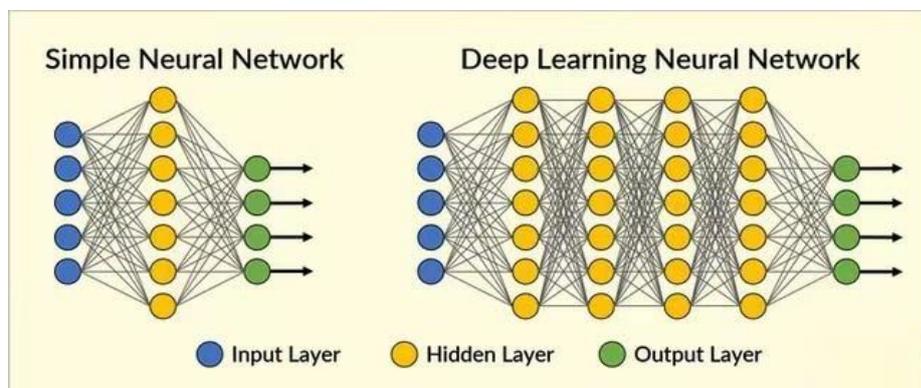


图 2.2 BP 神经网络结构

BP 神经网络的学习算法主要依赖于误差反向传播(Backpropagation,简称 BP)

算法。该算法通过计算网络输出与期望输出之间的误差，并将误差反向传播至网络的每一层，从而调整权重和偏置项。在进行误差反向传播之前，需要进行前向传播计算。如图 2.3 所示，前向传播是从输入层到输出层逐层计算神经元的输出，得到整个网络的输出。具体步骤如下：

- 步骤 1：将输入值传递至输入层。
- 步骤 2：计算隐藏层每个神经元的输出。
- 步骤 3：将隐藏层的输出作为输入传递至输出层。
- 步骤 4：计算输出层每个神经元的输出。

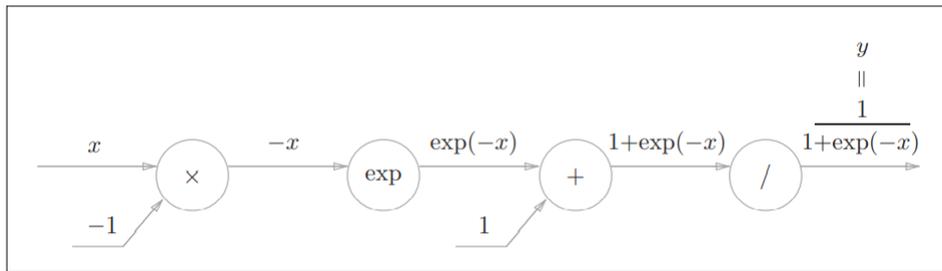


图 2.3 BP 神经网络前向传播过程

计算输出层的误差是指网络输出与期望输出之间的差异，常使用均方误差(Mean Squared Error, MSE)作为误差度量：

$$E = \frac{1}{2} \sum_{i=1}^n (d_i - y_i)^2 \tag{12}$$

其中， d_i 为期望输出， y_i 为实际输出， n 为输出层神经元数目。

误差反向传播的过程如图 2.4 所示，步骤如下：

步骤 1：计算输出层的误差梯度。计算输出层每个神经元的梯度，即误差关于输出的偏导数。

步骤 2：传播误差至隐藏层。根据链式法则，将输出层的误差梯度传播至隐藏层，计算隐藏层每个神经元的梯度。

步骤 3：更新权重和偏置项。根据梯度下降法则，利用计算得到的梯度更新网络的权重和偏置项。

步骤 4：重复迭代。重复进行前向传播和误差反向传播，直到网络收敛或达到设定的迭代次数。

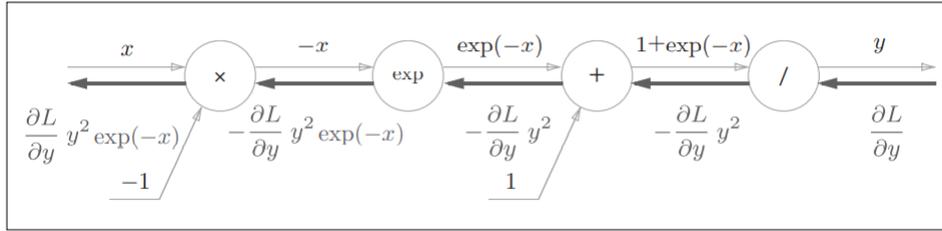


图 2.4 BP 神经网络的误差反向传播过程

BP 神经网络的训练过程主要包括初始化、学习率、正则化、批量训练和收敛判据几个主要模块；在训练开始之前，需要对网络的权重和偏置项进行初始化。常用的初始化方法包括随机初始化和 Xavier 初始化等，学习率是梯度下降法中的一个关键参数，它决定了权重和偏置项在每次迭代中的更新步长。合适的学习率可以加速网络的收敛，但过大的学习率可能导致震荡或发散。为防止过拟合，可以采用正则化方法，常见的有 L1 正则化和 L2 正则化；正则化项会在损失函数中引入权重的惩罚项，使得网络更加平滑和泛化性更好；BP 神经网络通常采用批量训练(Batch Training)方式，即每次迭代使用整个训练集进行权重更新。与随机梯度下降相比，批量训练通常更加稳定；训练过程中需要设置合适的停止条件，通常包括达到预定的迭代次数或误差降至某个阈值以下。

BP 神经网络广泛应用于各个领域，够学习复杂的模式，用于图像识别、语音识别等领域；在金融、经济学等领域，BP 神经网络可以用于预测股票价格、货币汇率等；也可以应用在自动控制系统中，如机器人运动控制、飞行器控制等；在医学影像处理、疾病诊断等方面，BP 神经网络也取得了显著成就。但 BP 神经网络仍然存在一些局限性，比如在深度神经网络中，由于误差反向传播过程中梯度的连续相乘，可能导致梯度消失或梯度爆炸问题。这使得网络在深度学习中难以训练；BP 神经网络通常需要大量的标记数据进行训练，否则容易过拟合。在某些应用场景下，数据获取成本较高，这就限制了 BP 神经网络的应用；BP 神经网络的训练时间较长，特别是在深度网络中。为了取得良好的性能，通常需要较长的训练时间和更大的计算资源。

BP 神经网络作为一种经典的人工神经网络模型，以其强大的模式学习和逼近能力在各个领域取得了广泛的应用。然而，也存在着训练时间长、容易陷入局部极小值等问题。随着深度学习的兴起，更加复杂的神经网络结构和优化算法不断涌现，BP 神经网络仍然是人工神经网络领域中不可或缺的一部分。

2.4.2 循环神经网络

循环神经网络(Recurrent Neural Network, 简称 RNN)源自于 1982 年由 Saratha Sathasivam 提出的霍普菲尔德网络。循环神经网络是一类在序列数据上表现出色的神经网络模型。相较于传统的前馈神经网络, RNN 具备处理序列信息的能力, 使其在自然语言处理、语音识别、时间序列分析等领域取得显著成果。

循环神经网络是一种对序列数据有较强的处理能力的网络。在网络模型中不同部分进行权值共享使得模型可以扩展到不同样式的样本, 比如 CNN 网络中一个确定好的卷积核模板, 几乎可以处理任何大小的图片。将图片中分成多个区域, 使用同样的卷积核对每一个区域进行处理, 最后可以获得非常好的处理结果。同样的, 循环网络使用类似的模块(形式上相似)对整个序列进行处理, 可以将很长的序列进行泛化, 得到需要的结果。

RNN 的目的就是用来处理序列数据的, 序列数据是在时间上排列的一系列值的集合, 如文本中的单词、时间序列中的股价、语音信号中的音频样本等。RNN 的设计目的就是为了更好地处理这种时序信息, 使网络能够对序列中的先前信息进行记忆和利用。在传统的神经网络模型中, 是从输入层到隐含层再到输出层, 层与层之间是全连接的, 每层之间的节点是无连接的。但是这种普通的神经网络对于很多问题都无能为力。比如你要预测句子的下一个单词是什么, 一般需要用到前面的单词, 因为一个句子中前后单词并不是独立的。

RNN 的核心在于其循环结构, 如图 2.5 和图 2.6 所示, 在每个时间步, RNN 接收当前输入和前一时间步的隐藏状态, 并输出当前时间步的隐藏状态。这种结构使得网络能够在不同时间步之间传递信息, 从而更好地处理序列数据。循环神经元是 RNN 的基本单元, 其计算过程包括对当前输入和前一时间步的隐藏状态进行线性变换和激活函数处理。数学表达式如下:

$$h_t = f(W_{hh}h_{t-1} + W_{xh}x_t + b_h) \quad (13)$$

其中, h_t 为当前时间步的隐藏状态, W_{hh} 和 W_{xh} 为权重矩阵, b_h 为偏置项, f 为激活函数。这个公式表示当前时间步的隐藏状态由前一时间步的隐藏状态、当前时间步的输入和偏置项共同决定。

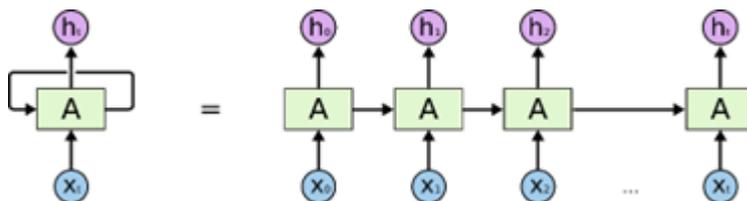


图 2.5 循环神经网络结构

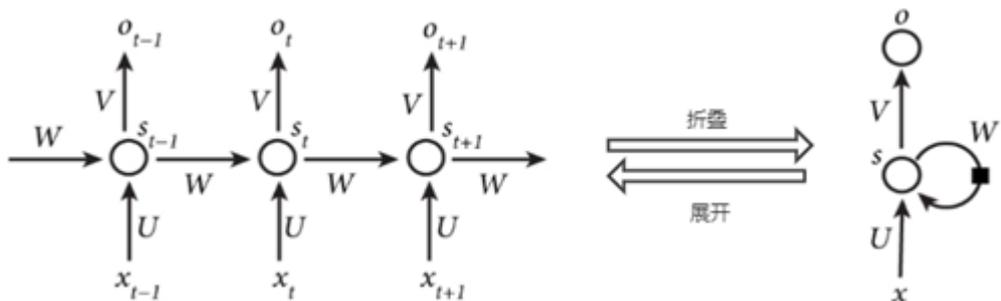


图 2.6 循环神经网络计算结构

循环神经网络的输入是序列数据，每个训练样本是一个时间序列，包含多个相同维度的向量。网络的参数如何通过训练确定？这里就要使用解决循环神经网络训练问题的 Back Propagation Through Time 算法，简称 BPTT，RNN 的学习算法通常采用误差反向传播(BPTT)算法。BPTT 将循环神经网络展开为一个前馈神经网络，然后利用误差反向传播的方法更新网络的参数。

循环神经网络的建模过程包括以下几个步骤：

步骤 1：数据预处理。首先需要对输入数据进行预处理，如去除噪声、标准化、向量化等，以提高模型的训练效果和泛化性能。

步骤 2：确定输入和输出。根据具体的应用场景，确定 RNN 的输入和输出，如文本分类任务中，输入可以是一个句子或一个文本段落，输出可以是一个分类标签。

步骤 3：序列长度处理。RNN 在处理序列数据时需要对序列长度进行处理，如截断、填充等，以保证输入数据的维度一致。

步骤 4：网络结构设计。根据具体的应用场景，设计合适的网络结构，如选择不同的循环单元，设置隐藏层的大小、层数等参数，以及选择合适的激活函数等。

步骤 5：模型训练。使用反向传播算法和梯度下降算法对模型进行训练，根据损失函数的变化来调整模型参数，以提高模型的性能和泛化能力。

步骤 6：模型评估。使用测试集对模型进行评估，选择合适的评估指标，如准确率、召回率、F1 值等来评估模型的性能和泛化能力。

步骤 7: 超参数调优。RNN 中存在大量的超参数, 如学习率、批次大小、正则化参数等, 需要通过实验和交叉验证等方法来调优, 以提高模型的性能和泛化能力。

步骤 8: 模型应用。将训练好的模型应用于实际场景中, 如对新数据进行分类、回归、生成等任务。

循环神经网络具有很多优点, 可以很好地处理序列数据, 如文本、语音、时间序列等, 具有循环连接和历史信息处理的特点, 能够保留历史信息并利用历史信息来影响当前时刻的输出和隐藏状态; 具有很高的灵活性, 可以根据不同的应用场景设计不同的网络结构和模型参数; 在处理序列数据时通常具有很好的预测性能, 能够学习到序列数据中的规律和特征, 并用于分类、回归、生成等任务。但同时循环神经网络也存在一些缺陷, 比如 RNN 在训练时需要处理序列数据中的时序信息, 因此计算复杂度较高, 训练时间也较长; 存在梯度消失和梯度爆炸等问题, 这会导致模型训练困难或无法收敛; 在处理长序列数据时, 由于历史信息的累积会导致梯度消失或梯度爆炸, 因此无法有效地记忆较长的历史信息。计算过程具有依赖性, 因此难以有效地并行化, 无法充分利用 GPU 等硬件加速器的计算能力。

循环神经网络作为一种能够处理序列数据的重要模型, 在多个领域取得了显著的成就。然而, 传统 RNN 仍然面临梯度消失和梯度爆炸的问题, 限制了其在长序列上的表现。随着 LSTM、GRU 等结构的提出, RNN 在一定程度上得到了改善。

2.4.3 双向门控循环神经网络(BiGRU)

门控循环神经网络(Gated Recurrent Neural Network, GRU)模型是为了更好地处理时序数据中的长期依赖性而进化出的一种更高级的循环神经网络(Recurrent Neural Network, RNN)。GRU 模型相比于长短期记忆网络(LSTM), 通过减少门控制组件有效减少参数量, 模型训练效率大幅提高。GRU 模型通过门保留了时间序列的重要特征, 缓解了循环神经网络(RNN)模型梯度消失问题, 在训练任务中保留了模型的记忆功能。

图 2.7 为 GRU 模型单个神经元结构展开图, t 时刻的输入值 x_t 和 $t - 1$ 时刻的

隐藏状态 h_{t-1} 共同构成 t 时刻神经元的输入值；输出值由 t 时刻隐藏状态的输出值 y_t 和传输到 $t+1$ 时刻的隐藏状态 h_t 所构成。

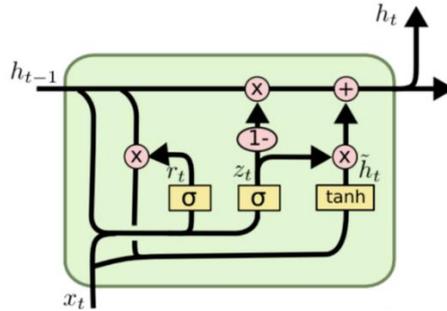


图 2.7 GRU 神经元结构图

GRU 模型的计算公式如下所示：

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{14}$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \tag{15}$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [r_t * h_{t-1}, x_t] + b_h) \tag{16}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{17}$$

$$y_t = \sigma(W_o \cdot h_t) \tag{18}$$

其中，公式(14)为更新门公式，公式(15)为重置门公式，公式(16)为候选隐藏状态更新公式，公式(17)为最终隐藏状态更新公式，公式(18)为输出值公式。其中 W_r 、 W_z 、 $W_{\tilde{h}}$ 、 W_o 分别为更新门、重置门、更新和输出公式的权重， b_r 、 b_z 、 b_h 分别为更新门、重置门和候选隐藏状态公式的偏置。

双向门控循环神经网络(BiGRU)是一种在循环神经网络(RNN)和门控循环神经网络(Gated Recurrent Neural Network, GRU)模型基础上发展而来的模型，通过引入门控机制以及双向结构，增强了对序列数据的建模能力。BiGRU 的核心在于其双向结构，即同时考虑正向和反向的信息。在序列中的每个时间步，BiGRU 会维护两组隐藏状态，分别表示正向和反向的信息。这样，模型能够充分利用整个序列的信息，更好地捕捉上下文关系，类似于 GRU，BiGRU 也采用门控机制。每个方向的 GRU 单元都包括更新门、重置门和新状态的计算，这些门控制着信息的流动，使得网络能够更好地处理长序列，避免梯度消失和梯度爆炸问题。BiGRU 的主要结构就是双向 GRU 层，其包括正向和反向两个方向的 GRU 单元。每个方向的 GRU 单元都独立地计算隐藏状态，但在预测时会将两个方向的隐藏状态合并起来使用；在双向结构中，正向和反向的参数通常是共享的，即使用相

同的权重和偏置。这样的设计能够减少模型的参数数量，提高训练效率；在输出层，BiGRU 通常会将正向和反向的隐藏状态进行连接或求平均，然后通过全连接层生成最终的输出。输出可以是分类标签、回归值等，具体任务而定。

BiGRU 模型包含两个独立的 GRU 组块，通过从时间序列数据正反两个方向汇总信息来获取时序信息。图 2.8 为 BiGRU 结构图，BiGRU 模型神经元当前的隐藏状态是由当前的输入 x_t 、 $t-1$ 时刻前向隐藏状态 $\overrightarrow{h_{t-1}}$ 和反向隐藏状态 $\overleftarrow{h_{t-1}}$ 三个部分共同决定，BiGRU 隐藏状态更新公式如下：

$$\overrightarrow{h}_t = GRU(x_t, \overrightarrow{h_{t-1}}) \tag{19}$$

$$\overleftarrow{h}_t = GRU(x_t, \overleftarrow{h_{t-1}}) \tag{20}$$

$$h_t = W_t \overrightarrow{h}_t + V_t \overleftarrow{h}_t + b_t \tag{21}$$

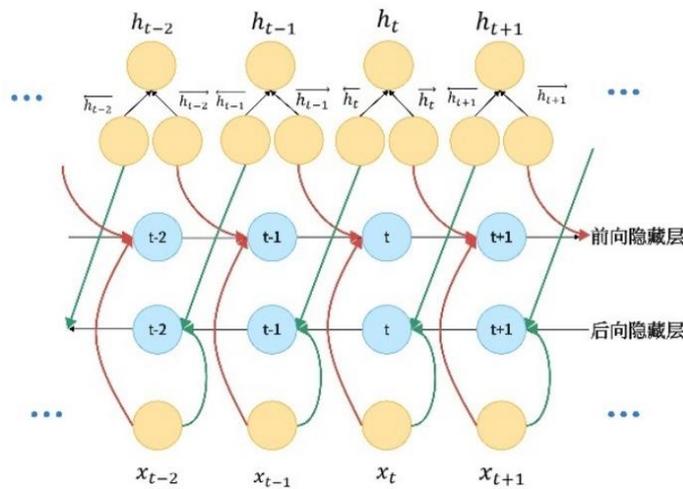


图 2.8 BiGRU 模型网络结构图

公式(19)为 t 时刻神经元前向隐藏层状态更新公式，公式(20)为 t 时刻神经元反向隐藏层状态更新，公式(21)为 BiGRU 模型 t 时刻隐藏状态加权更新公式。其中， W_t 表示 t 时刻神经元的前向隐藏状态 \overrightarrow{h}_t 的权值； V_t 分别表示 t 时刻神经元反向隐藏状态 \overleftarrow{h}_t 的权值； b_t 是 t 时刻隐藏状态的偏置。

2.4.4 卷积注意力机制

CBAM(Convolutional Block Attention Module)理论是一种用于提高卷积神经网络(CNN)性能的方法。它通过引入注意力机制来增强神经网络对输入特征的关

注度，从而提高网络在图像分类、目标检测等任务上的性能。CBAM 的核心思想是在卷积神经网络的基本块(block)中引入两种注意力机制，即通道注意力(Channel Attention)和空间注意力(Spatial Attention)，如图 2.9 所示，这两种注意力机制共同作用，使得网络能够更加有效地学习和利用输入图像的重要特征。

通道注意力机制旨在捕捉输入特征图中不同通道之间的关联性，如图 2.10 所示，通过学习每个通道的权重，网络可以更好地选择哪些通道对于当前任务更为重要。通道注意力的计算过程通常包括全局平均池化(Global Average Pooling)用于获取每个通道的平均值，然后通过全连接层学习权重。空间注意力机制则旨在捕捉输入特征图中不同空间位置之间的关系，如图 2.11 所示，通过学习每个空间位置的权重，网络可以更好地关注图像中的重要区域。空间注意力通常通过使用 3x3 的卷积核来获取每个位置的信息，并利用全连接层进行权重学习。

CBAM 模块的输出是将通道注意力和空间注意力结合在一起的特征图，这样网络在学习过程中能够更加精确地关注重要的特征，提高对输入图像的代表能力。CBAM 的引入可以在不增加网络参数和计算复杂度的情况下，显著提高卷积神经网络的性能。总体而言，CBAM 理论通过引入通道注意力和空间注意力机制，使得网络能够更加智能地学习和利用输入图像的特征信息。这种注意力机制的引入可以增强网络对关键特征的感知能力，提高模型的泛化性能，从而在图像处理任务中取得更好的性能。

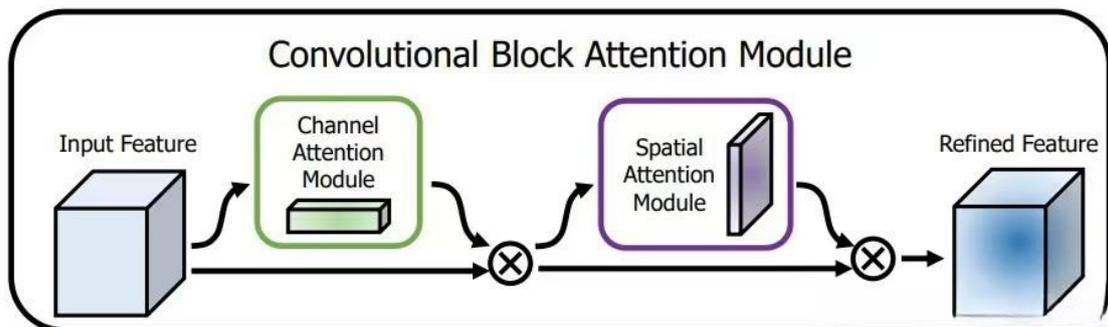


图 2.9 卷积注意力机制模块

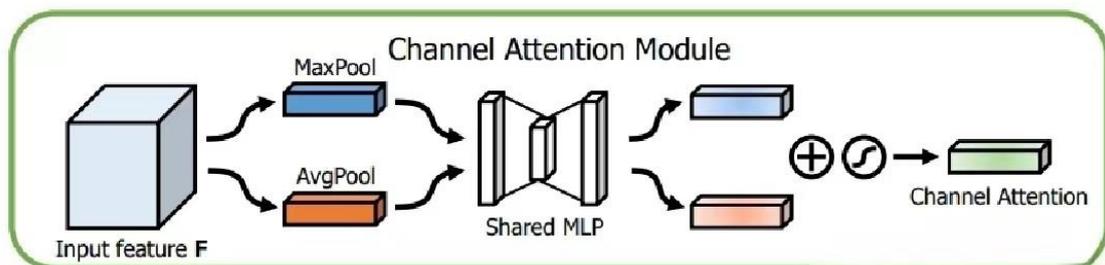


图 2.10 通道注意力模块

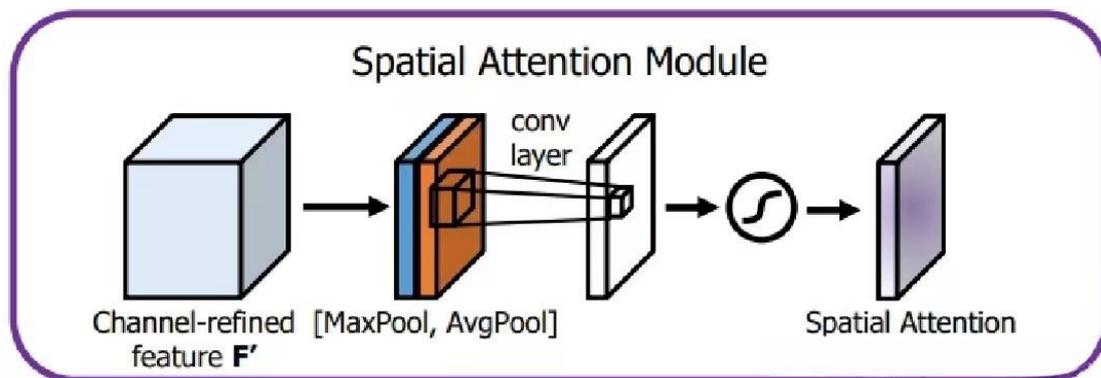


图 2.11 空间注意力模块

3 VMD-CBAM-BiGRU 模型的构建

3.1 Pearson 相关性分析

Pearson 相关性分析是一种用于衡量两个连续变量之间线性关系的统计方法，它的核心是计算 Pearson 相关系数。这一方法旨在揭示变量之间的线性依赖关系，其值介于-1 和 1 之间，Pearson 相关系数的计算基于协方差和两个变量标准差的比值。

本文在指标体系的选择上不仅包括开盘价、最高价、最低价、当日收盘价、涨跌值、涨跌幅、成交量和成交额等 8 种基本行情指标，还纳入了 MA、MACD、MTM、BIAS 等多达 14 种各类技术指标。本文将次日收盘价的价格作为预测因变量。由于本文所构建的指标体系指标众多，为了消除无关指标对预测效果的影响，本文选择皮尔逊相关系数法对特征指标集和预测因变量进行相关程度分析，计算出每个特征指标与预测因变量的相关系数，剔除掉相关系数较小的无关变量，筛选出对次日收盘价有较大影响的特征指标作为最优特征集。本文通过皮尔逊特征选择将相关系数小于 0.3 的无关变量剔除掉，剔除掉的无关变量如表 3.1 所示，最终得到的最优特征集包括开盘价、最高价、最低价、成交量、成交额、当日收盘价、MA、MACD、RSI、ROC、VR、BOLL 和 ADTM 等 13 项特征指标。

表 3.1 无关变量相关系数

无关变量	相关系数
涨跌	0.071
涨跌幅	0.071
MICD	0.004
MTM	0.235
BIAS	0.199
KDJ	0.211
WR	-0.142
OBV	-0.04
VHF	0.109

3.2 模糊熵

经过 VMD 算法分解后会得到多个子序列，有的子序列之间频率接近表达的

信息相似度较高, 不经处理对每个子序列进行建模训练会导致运算量极大, 因此为减轻模型的运算量, 提高运算效率, 本文采用模糊熵算法对每个 VMD 分解后的子序列进行熵值计算从而将熵值接近的子序列重构为新的模态分量序列, 根据重构后的子序列进行后续模型的训练分析。

FE (Chen 等, 2007) 来源于 Zadeh 模糊集的概念, 它使用模糊隶属函数代替 Heaviside 函数来确定模糊边界。熵值表示序列的复杂度; 熵值越大, 序列越复杂。由于有限元可以更好地评价序列的复杂度 (Chen 等, 2009), 因此本研究采用有限元来评价所有模态的复杂度。计算过程如下:

(1) 给定的 N 维时间序列 $[u(1), u(2), \dots, u(N)]$

(2) 定义相空间维数 $m (m \leq N - 2)$ 和相似容限度 r , 重构相空间 $X(i) = [u(i), u(i + 1), \dots, u(i + m - 1)] - u_0(i), i = 1, 2, \dots, N - m + 1$, 其中 $u_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} u(i + j)$

(3) 引入隶属度函数 $A(x) = \begin{cases} 1, & x = 0 \\ \exp[-\ln(2) \cdot x], & x > 0 \end{cases}$, 其中 r 为相似容限度,

对于 $i = 1, 2, \dots, N - m + 1$, 计算 $A_{ij}^m = \exp[-\ln(2) \cdot (d_{ij}^m/r)^2], j = 1, 2, \dots, N - m + 1$, and $j \neq i$

通常情况下, 较大的 m 能更细致地重构系统的动态演化过程。相似容限 r 的取值也是一个值得考虑的问题, 过大的相似容限会导致信息丢失, 相似容限值越大, 丢失的信息越多, 而太小的相似容限度则会增加结果对噪声的敏感性, 一般定义 r 为 $r \cdot SD$, 其中 SD (Standard Deviation) 为原一维时间序列的标准差 [1,2]。

以上证综指为例, 经过变分模态算法 (VMD) 分解后的模态分量如下图 3.1 所示, 此时原股价序列被分为了八个频率各不相同的模态分量, 但由于部分分量相似度较大, 并且分量太多也影响模型的运算效率, 因此本文通过模糊熵方法确定各模态分量的值, 然后将熵值接近的分量相加为新的重构分量, 经过重构后的模态分量如图 3.2 所示, 从一定程度上提高了模型的预测效率及复杂度。

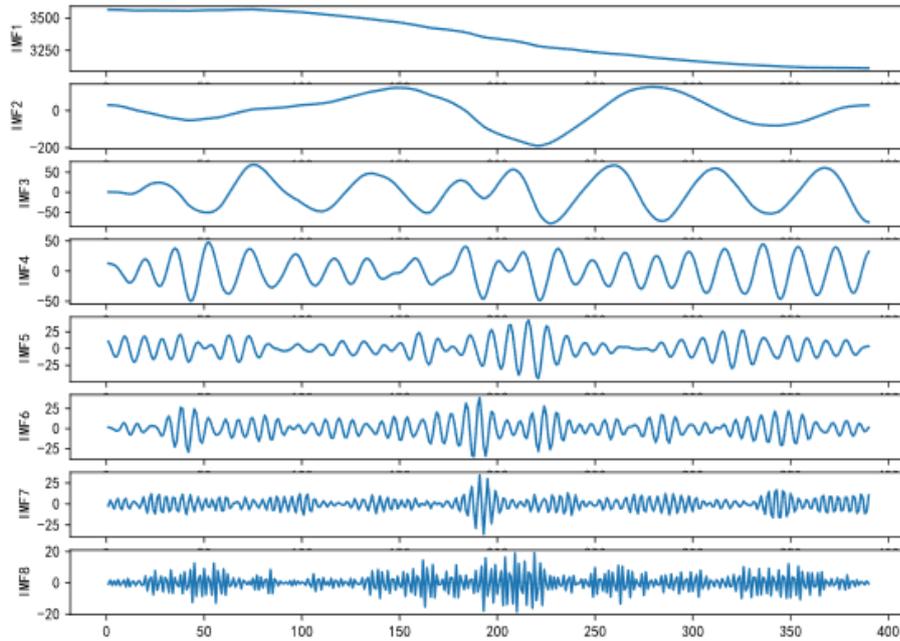


图 3.1 VMD 算法分解后的序列

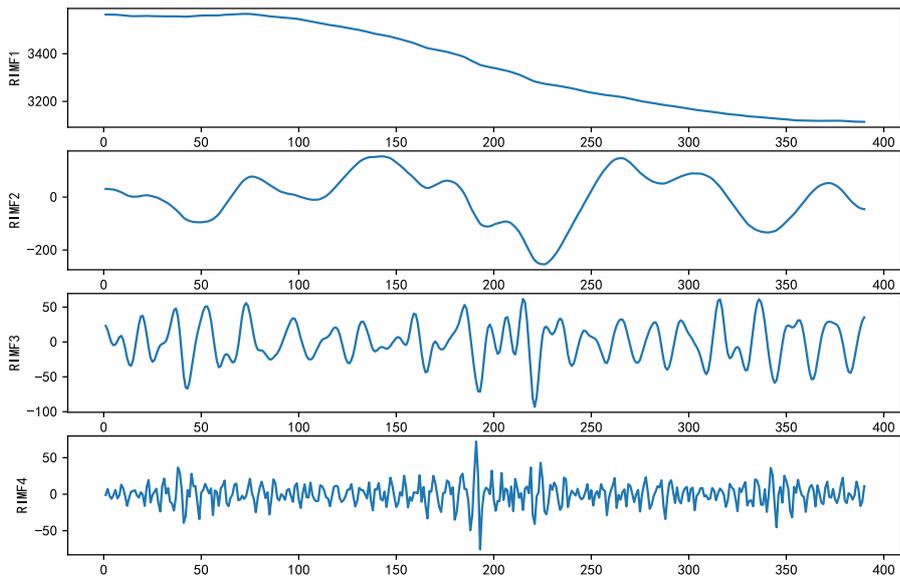


图 3.2 模糊熵重构后的序列

3.3 超参数优化

由于基础模型 BiGRU 含有隐藏单元数、学习率、丢失率以及批量大小等超参数，超参数选取的合适与否对于 BiGRU 模型的表达能力具有较大的影响，而人工寻参不仅耗费时间长而且效果不好，因此本文选择使用群智能算法中全局寻优能力较强的粒子群算法来寻求最佳超参数组合，为了进一步增强粒子群算法的全局寻优能力，本文在算法基础上创新性地加入了动态的改进从而使超参数组合

更加接近最优组合值。

在使用动态粒子群算法进行模型超参数寻优时，迭代批次 T 设置为 100，学习率寻优区间设置为[0.001, 0.01]，隐藏神经元数量区间设置为[100, 300]，GRU 层神经元数量区间设置为[100,300]，丢失率区间设置为[0.1,0.3]，适应度函数选择 MSE 函数，对以上各超参数进行寻优，参数优化的结果如下表所示：

表 3.2 模型最优参数

模型	神经网络层	最优参数
BiGRU	GRU 层神经元	(180)
	隐藏层神经元	(240)
	学习率	(0.0065)
	丢失率	(0.15)
GRU	GRU 层神经元	(200)
	隐藏层神经元	(220)
	学习率	(0.007)
	丢失率	(0.2)

3.4 模型评价指标

对于本文所提模型及其他对比模型，均采用拟合优度 R^2 、平均绝对误差(MAE)、平均绝对百分比误差(MAPE)和均方根误差(RMSE)等四项评价指标作为评判模型优劣的标准。四项评价指标的含义及公式如下所示：

拟合优度 R^2 (R-squared)是一种用于评估回归模型拟合程度的统计指标，提供了一种度量模型解释观测数据变异性的方式，取值范围在 0 到 1 之间。 R^2 值越接近于 1，说明模型的拟合效果越好，自变量对因变量的解释程度越高。如公式(22)所示：

$$R^2 = 1 - \frac{\sum(y_t - \hat{y}_t)^2}{\sum(y_t - \bar{y}_t)^2} \quad (22)$$

平均绝对误差(MAE, Mean Absolute Error)是一种用于评估预测模型性能的指标，度量了模型预测值与实际观测值之间的平均绝对差异，MAE 的值越小说明真实值和预测值的偏差越小，模型的预测效果越好。如公式(23)所示：

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_t - \hat{y}_t| \quad (23)$$

平均绝对百分比误差(MAPE, Mean Absolute Percentage Error)是一种用于衡

量预测模型精度的指标,表示模型的预测值相对于实际观测值的平均百分比误差,MAPE 的值越小,表示模型的预测越准确。MAPE 常用于评估时间序列数据的预测模型,尤其是在需要百分比误差度量的应用场景中。如公式(24)所示:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (24)$$

均方根误差(RMSE, Root Mean Squared Error)是一种用于度量预测模型误差的指标,计算了模型预测值与实际观测值之间的平方误差的平均值,并取其平方根。RMSE 的值越小,表示模型的预测越准确。如公式(25)所示:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_t - \hat{y}_t)^2} \quad (25)$$

3.5 VMD-CBAM-BiGRU 模型的构建思路

由于股票价格时间序列具有高度非平稳性、长记忆性、噪声强和数据波动大等特点,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,需要对股票价格时间序列进行降噪处理,在以往的学者研究中,通常采用小波分解和经验模态分解(EMD)等算法来对数据序列进行降噪处理,但这些经典的降噪算法也存在着一些缺点,小波分解采用的小波基函数是固定的,这意味着它可能不够适应信号的局部特性。在处理股价时间序列这种非平稳信号时,可能会导致信息的丢失,并且波分解会将信号分解为多个分辨率层次的细节信息,但这些层次的分辨率不一定对所有信号都是合适的。EMD 则可能产生模态混叠,即一个模态函数中包含了多个频率成分。这可能导致对信号成分的解释不够清晰,频率的分解比较依赖于数据的具体情况,而 VMD 基于数学上的变分原理,通过最小化一个成本函数来得到本征模态函数。这使得 VMD 的理论更加严密,并且对初始条件相对不敏感,由于它是基于数学优化原理的,因此对于不同的初始条件具有较好的稳定性,总的来说, VMD 在理论上更为严谨,且在一些实际应用中表现出更好的性能,尤其在处理股价时间序列这种非线性和非平稳信号时。因此本文将 VMD 算法运用到股价时间序列中以期望达到降低噪声影响的目的。

另一方面,在股价预测领域很多学者均应用 RNN、LSTM、GRU 等神经网络进行股价预测分析,这是因为股价时间序列具有长记忆性的特点,各时间

节点之间存在着或强或弱的关联，普通的机器学习模型以及 BP 神经网络等模型很难提取到股价时间序列数据中的时间特征，循环神经网络则可以很好的关注股价序列中所蕴含的时间信息，在常用的循环神经网络模型中，LSTM 和 GRU 模型在预测时间序列数据中的效果极为接近，但不同的是 GRU 模型相对于 LSTM 模型结构更为精简，提高了训练效率，又由于 BiGRU 模型为双向结构，能够同时从两个时间方向提取数据序列中的时间特征，相比于 GRU 模型在处理时间序列数据上更加精准高效，因此本文采用 BiGRU 模型作为预测的基准模型，BiGRU 模型中包含大量的超参数，这些超参数的取值对于模型的表现效果有着显著的影响，为了使 BiGRU 模型以最佳性能进行训练预测，本文采用群智能优化算法中全局寻优能力较强的粒子群算法进行模型参数的优化，为了增强粒子群算法的全局寻优能力，本文又在粒子群算法的基础上进行了动态的改进，通过改进后的动态粒子群算法对模型最优参数进行寻找。本文在股票基本行情指标的基础上添加了大量的技术指标，为了模型对大量指标特征下股价数据的预测效果，本文将 CBAM 融入模型的预测，CBAM 将通道注意力机制和空间注意力机制结合起来对数据特征进行权重分配，进一步提高了模型的特征提取能力，将经过 CBAM 处理后的特征向量输入 BiGRU 模型进行训练从而提高模型的预测表达能力。

综上所述，本文通过 VMD 算法对原始股价时间序列进行分解再重构为高频、中频、低频和趋势项等模态分量实现降噪效果，然后通过 CBAM 模型对包含大量指标特征的股价数据进行特征提取，最终通过提取时间特征较强的 BiGRU 模型进行最终的训练及预测得出结果。

3.6 VMD-CBAM-BiGRU 模型的构建流程

VMD-CBAM-BiGRU 模型构建的主要流程如下图 3.3 所示，模型的构建主要通过下下几个步骤来实现。

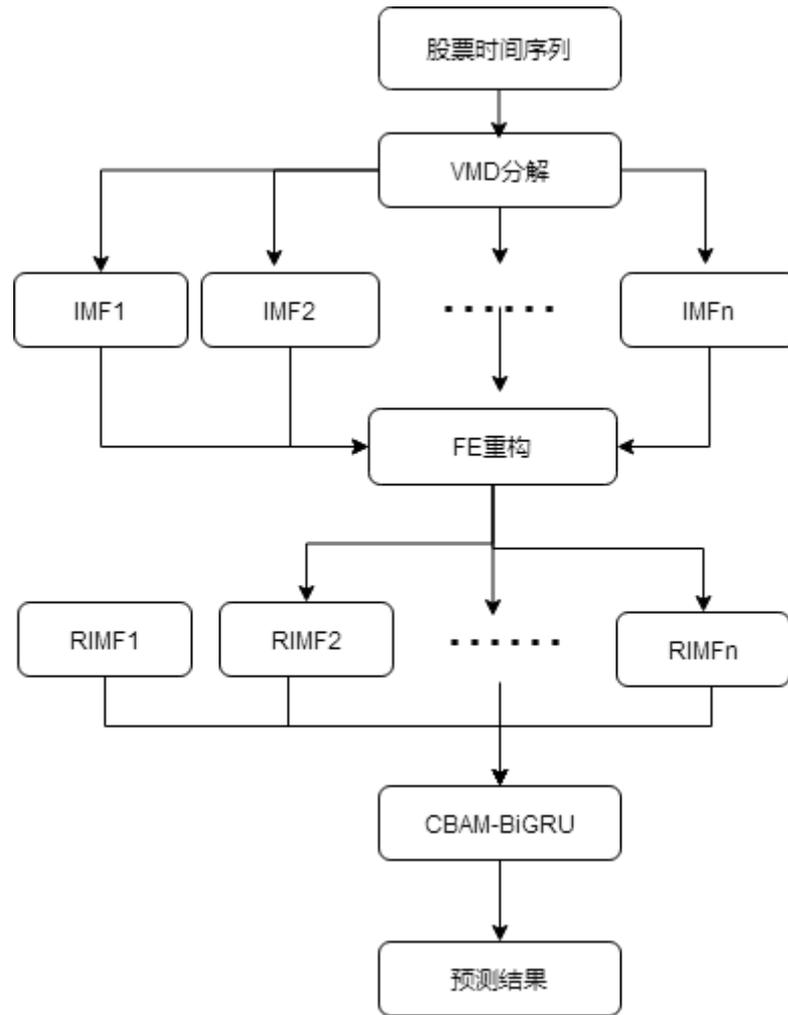


图 3.3 VMD-CBAM-BiGRU 模型构建流程

步骤 1: 首先对所有特征指标数据进行标准化处理以消除量级差异的影响。

步骤 2: 使用 VMD 算法将原始股票价格序列分解为趋势序列和各频率分量等多个子序列。

步骤 3: 计算各模态分量的模糊熵值, 依据模糊熵值的接近程度将 VMD 算法分解后的子序列重构为小于 VMD 分解序列数量的不同频率的模态分量。

步骤 4: 对重构后的所有数据集均进行相应时间步数的滑动处理。

步骤 5: 将重构后的四个模态分量分别作为新的子被解释变量, 分别使用 CBAM-BiGRU 模型对四个数据集进行训练并得出预测结果。

步骤 6: 将各重构模态分量的预测结果相加得出最终的模型预测结果, 并与其他对比模型的预测结果进行对比分析来验证本文所构建模型的优越性和可行性。

4 实证分析

4.1 数据预处理

4.1.1 数据来源及描述

在数据的选择上，为了更好的说明本文所构建模型的可靠性以及对不同使用对象的普适性，本文选取 A 股市场上两支具有代表性的股指作为实验对象来说明所构建模型对市场监管者监测股市整体趋势变化而具备的作用，另外还选取了 A 股市场上两支具有代表性的个股进行实验来说明所构建模型对投资者在股市中合理投资、规避风险等方面所具备的作用。

本文所有数据均于东方财富 choice 软件获取，为了验证所构建模型的稳定性和泛化性，本文从 A 股市场中选取了上证指数和沪深 300 两支著名股票指数以及茅台和中国平安两支个股分别进行实验，每支股票均选取时间范围为 2015 年 1 月 1 日到 2022 年 12 月 31 日七年间的 1900 余条数据，分别使用所构建模型及对比模型在数据集上进行预测分析。本文所选取的四只股票的次日收盘价的统计描述如下表 4.1 所示。

表 4.1 四支股票价格描述性分析

	个数	均值	标准化	最小值	25%值	50%值	75%值	最大值
上证综指	1939	3216	369	2464.36	2976.88	3195.45	3403.13	5166.35
沪深 300	1939	3975.22	634.15	2853.76	3461.65	3860.43	4308.41	5807.71
茅台	1939	927.19	644.39	137.24	304.58	687.91	1625.86	2479.11
中国平安	1939	47.53	16.74	19.33	29.89	46.13	62.31	81.21

4.1.2 数据标准化处理

在获取数据后，由于各特征指标之间量级相差过大，直接代入模型进行训练分析会影响预测效果，数据标准化是一种常见的数据预处理步骤，旨在使不同尺度或单位的数据具有相同的比例，以便更好地适用于某些机器学习算法或统计分析方法。而因此在进行模型训练前需要对数据集进行标准化处理来消除量级差距的影响，本文通过对数据集进行 Min-Max 处理实现标准化过程，公式如下所示：

$$X_{\text{normalized}} = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (26)$$

4.1.3 数据集切分

为了对四支股票数据进行模型的训练及预测，首先将所有数据的时间步数设置为 10 天，次日收盘价为预测对象，将前 10 天的特征指标数据与第 11 天的预测对象值作为一条数据，即通过前十天的数据预测第十一天的结果，依次滑动得到时间步为 10 的数据集，然后本文依据 8: 2 的比例将数据集划分为训练集和测试集，训练集包含 1500 余条数据，测试集包含 300 余条数据，模型先于训练集上进行训练后提高对该股票数据集的适应能力，而后将测试集代入训练好的模型进行预测分析得出模型的各项评价指标结果，通过评价指标结果和预测效果图来直观的显示模型的预测效果。

4.2 实验分析

4.2.1 上证综指

上证综指为上海证券交易所的大盘指数，反映了在上海证券交易所上市的全部股票的价格变化，分析上证综指的价格走向对于投资者做出投资决策具有较强的指导意义。在基础预测模型的选取上，本文选取了深度学习技术中的 BP 神经网络、GRU 和 BiGRU 模型进行比较分析，从对比实验得出的评价指标结果表可以看出，BP 模型表现效果最差，这是因为相比于其他模型而言，BP 神经网络无法关注股价时间序列中的时间信息，从图 4.1 可以直观地看出，BP 神经网络模型对于股价变化的跟随效果较弱，尤其在极值点出的吻合程度较差；GRU 模型相比于 BP 神经网络而言加入了时间层次上的分析，对于股票时间序列序列而言更为适合，从评价指标来看，GRU 模型的拟合程度达到了 90.9%，相比于 BP 神经网络提升了 3%，MAE、MAPE 和 RMSE 分别为 51.536、1.508 和 63.414，相比于 BP 神经网络误差分别降低了 29%、13.4%和 15.3%；BiGRU 模型则在 GRU 模型的基础上加入了双向的时间层，对于处理时间序列数据而言更为准确高效，在上证综指的实验中，BiGRU 模型的拟合程度达到了 95.5%，相比于 GRU 模型

有 4.6% 的提升, MAE 值、MAPE 值和 RMSE 分别为 36.573、1.084 和 44.307, 相比于 GRU 模型预测误差分别减少了 29%、42% 和 19.1%, 从预测效果图 4.1 也可以看出, BiGRU 模型相比于其他几个模型更适合进行股票价格时间序列的预测分析, 因此本文将 BiGRU 模型作为基础预测模型进行预测。

为了进一步提高模型的预测效果, 本文在基础模型 BiGRU 的基础上引入能有效降低噪声影响的 VMD 算法以及能更好提取关键特征信息的 CBAM 算法, 在 VMD 算法和 BiGRU 模型融合前, 根据模糊熵方法将子序列重构为如图 3.2 所示的重构序列, 然后分别将两种算法和基础模型 BiGRU 相结合进行实验, 从评价指标结果表 4.2 可以看出, VMD 和 CBAM 两种算法引入基础模型后均能提高模型的预测表现, 两种算法的提升效果十分接近, 加入 VMD 算法后的模型拟合程度相比于基础模型 BiGRU 提升了 1.8% 达到了 97.3%, MAE、MAPE 和 RMSE 分别为 26.813、0.806 和 34.465, 预测误差分别降低了 26.5%、25.6% 和 22.2%, CBAM 在引入模型后拟合程度相比于基础模型 BiGRU 提升了 2.8% 达到了 98.3%, MAE、MAPE 和 RMSE 分别为 20.607、0.623 和 27.305, 预测误差分别降低了 43.5%、42.5% 和 38.3%, 从拟合效果图 4.2 也可以看出, VMD-BiGRU 模型和 CBAM-BiGRU 对股价变化的跟随程度均好于基础模型 BiGRU, 在极值点处的预测结果拟合程度也高于 BiGRU 模型, 从而验证了 VMD 算法和 CBAM 算法分别和基础模型融合后对 BiGRU 模型预测效果的显著提升。

经过上述实验验证后, 本文同时将 VMD 算法和 CBAM 算法融入基础模型 BiGRU 构建 VMD-CBAM-BiGRU 模型, 期望能进一步提升模型的预测能力, 将上证综指数据集代入 VMD-CBAM-BiGRU 模型进行训练并预测得到评价指标结果和预测效果拟合图, 从评价指标可以看出, VMD-CBAM-BiGRU 模型的拟合程度达到了 99.2%, 相比于 VMD-BiGRU 模型和 CBAM-BiGRU 模型分别提高了 1.9% 和 0.9%, MAE、MAPE 和 RMSE 分别为 5.415、0.163 和 7.292, 相比于 VMD-BiGRU 模型分别降低了 79%、79.7% 和 78.8%, 相比于 CBAM-BiGRU 模型分别降低了 73.7%、73.8% 和 73.3%, 相比于基础模型 BiGRU 则降低了 84.9%、92.1% 和 83.5%, 从拟合效果图 4.3 也可以看出 VMD-CBAM-BiGRU 模型对于次日股票收盘价的预测结果紧跟股价波动的变化, 且误差达到很小的程度。

综上所述, 通过在上证综指上进行的一系列的实验, 验证了 VMD 算法和

CBAM 算法对于基础预测模型的提升作用，并验证了将两种模型同时融合所构建的 VMD-CBAM-BiGRU 模型在预测股价时间序列上的优越性和准确性。

表 4.2 上证指数评价指标结果

	R^2	MAE	MAPE	RMSE
VMD-CBAM-BiGRU	0.992	5.415	0.163	7.292
VMD-BiGRU	0.973	26.813	0.806	34.465
CBAM-BiGRU	0.983	20.607	0.623	27.305
BiGRU	0.955	36.573	1.084	44.307
GRU	0.909	51.536	1.508	63.414
BP 神经网络	0.873	59.925	1.743	74.89

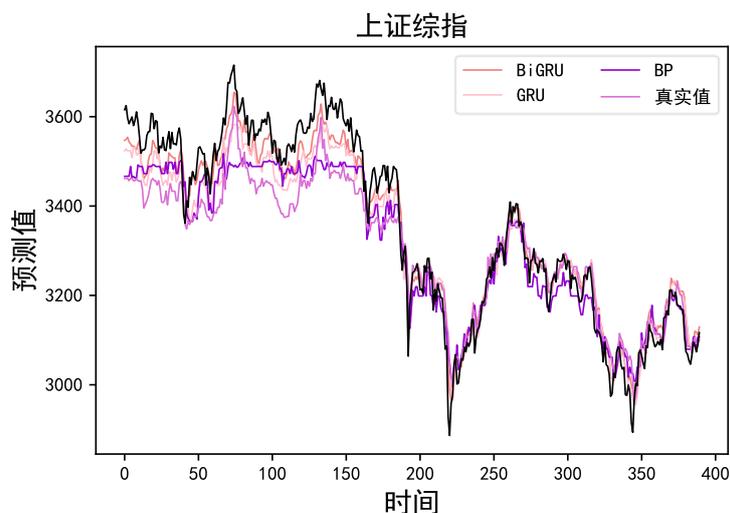


图 4.1 上证指数拟合效果图

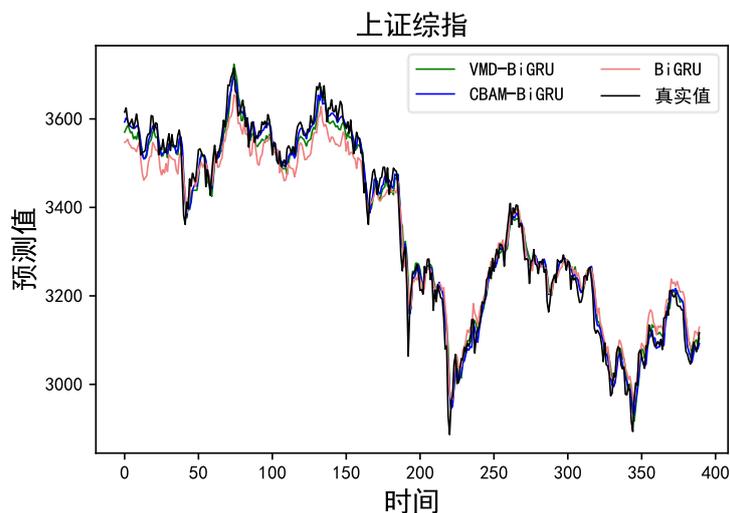


图 4.2 上证指数拟合效果图

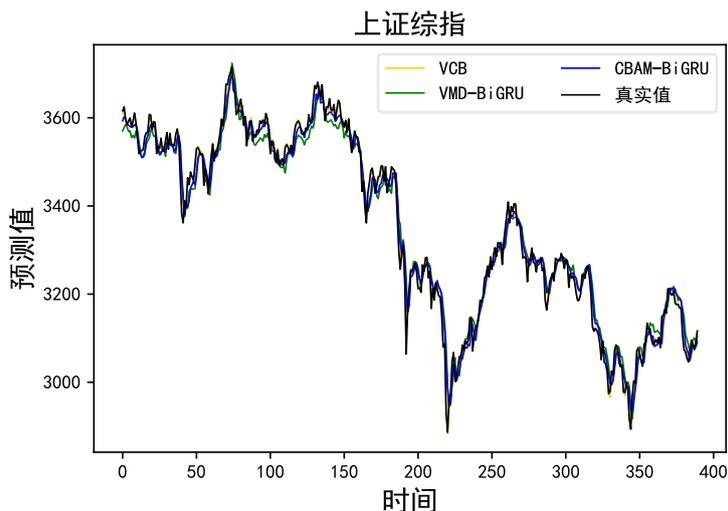


图 4.3 上证指数拟合效果图

4.2.2 沪深 300

为了进一步验证所构建模型的股价预测能力，本文又于 A 股市场中选取了著名的沪深 300 股指进行实验研究。对 BP 神经网络、GRU 和 BiGRU 模型进行比较分析，从对比实验得出的评价指标结果表可以看出，BP 神经网络的预测效果最差，从图 4.4 可以直观地看出，GRU 模型相比于 BP 神经网络更为贴近股价的变化趋势，对于股票时间序列而言更为适合，从评价指标来看，GRU 模型的拟合程度达到了 92.4%，相比于 BP 神经网络提升了 4.5%，BiGRU 模型由于在 GRU 模型的基础上加入了双向时间层，在对沪深 300 股价预测的实验中拟合程度达到了 94.8%，相比于 GRU 模型有 2.4% 的提升，MAE 值、MAPE 值和 RMSE 分别为 93.684、2.096 和 108.23，相比于 GRU 模型预测误差分别减少了 18.3%、16.6% 和 17.5%，从预测效果图也可以看出，BiGRU 模型相比于其他几个模型更适合进行股票价格时间序列的预测分析，验证了 BiGRU 模型作为基础模型的优越性。

为了进一步提高模型的预测效果，在基础模型 BiGRU 的基础上引入 VMD 算法和 CBAM 算法，根据模糊熵方法将 VMD 分解后的子序列重构为如图 4.7 所示的重构序列，然后分别将两种算法和基础模型 BiGRU 相结合进行实验，从评价指标结果表 4.3 可以看出，VMD 和 CBAM 两种算法对于 BiGRU 模型均有较大的提升且效果相近，VMD-BiGRU 模型和 CBAM-BiGRU 模型分别较 BiGRU 提升了 3.5% 和 3.7%，从拟合效果图 4.5 也可以看出，VMD-BiGRU 模型和 CBAM-

BiGRU 对股价变化趋势的拟合程度均优于基础模型 BiGRU，验证了 VMD 算法和 CBAM 算法分别和基础模型融合后对 BiGRU 模型预测提升效果。

经过上述实验验证后，同时将 VMD 算法和 CBAM 算法融入基础模型 BiGRU 构建 VMD-CBAM-BiGRU 模型，将沪深 300 数据集代入 VMD-CBAM-BiGRU 模型进行训练并预测得到评价指标结果和预测效果拟合图，从评价指标可以看出，VMD-CBAM-BiGRU 模型的拟合程度达到了 99.2%，相比于 VMD-BiGRU 模型和 CBAM-BiGRU 模型分别提高了 1.3% 和 1.1%，MAE、MAPE 和 RMSE 分别为 20.335、0.461 和 25.949，相比于 VMD-BiGRU 模型分别降低了 58.4%、59% 和 57.4%，相比于 CBAM-BiGRU 模型分别降低了 48.2%、54.3% 和 54.5%，从拟合效果图 4.6 可以看出 VMD-CBAM-BiGRU 模型对于次日股票收盘价的预测结果紧跟股价波动的变化，预测误差在所有模型中最小。

综上所述，所构建 VMD-CBAM-BiGRU 模型于 A 股市场中沪深 300 股指的预测中仍取得了优秀的预测效果，进一步验证了模型的稳定性。

表 4.3 沪深 300 评价指标结果

	R ²	MAE	MAPE	RMSE
VMD-CBAM-BiGRU	0.996	20.335	0.461	25.949
VMD-BiGRU	0.983	48.975	1.126	60.995
CBAM-BiGRU	0.985	43.98	1.01	57.02
BiGRU	0.948	93.684	2.096	108.23
GRU	0.924	114.795	2.515	131.4
BP 神经网络	0.881	144.921	3.17	164.345

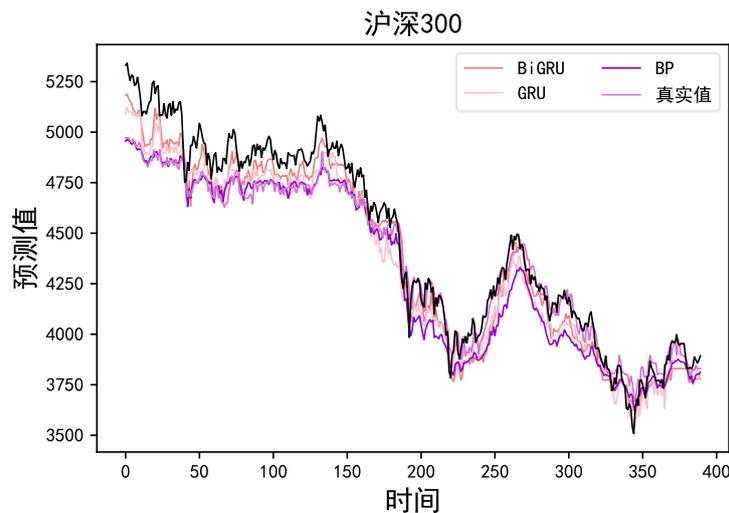


图 4.4 沪深 300 拟合效果图

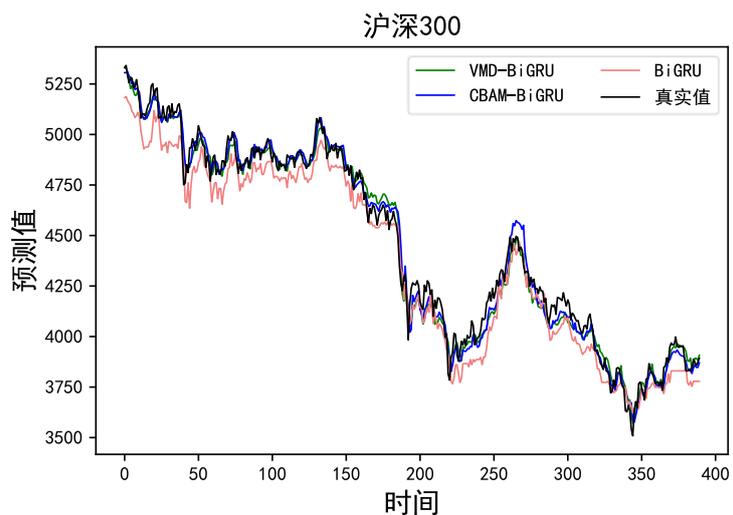


图 4.5 沪深 300 拟合效果图

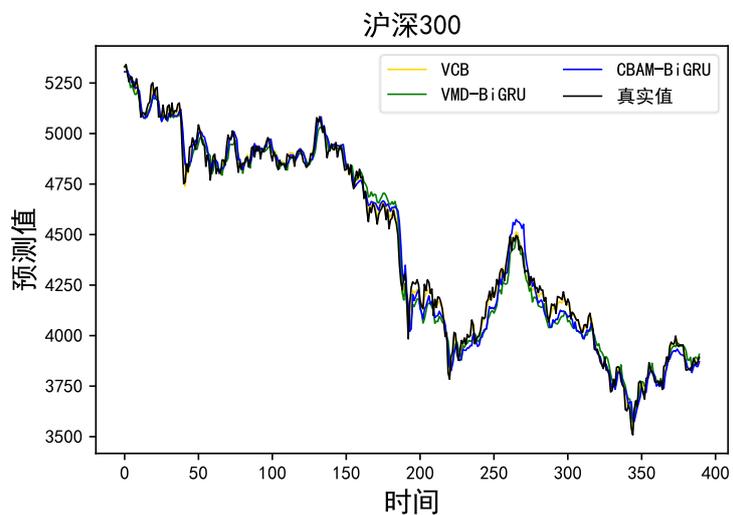


图 4.6 沪深 300 拟合效果图

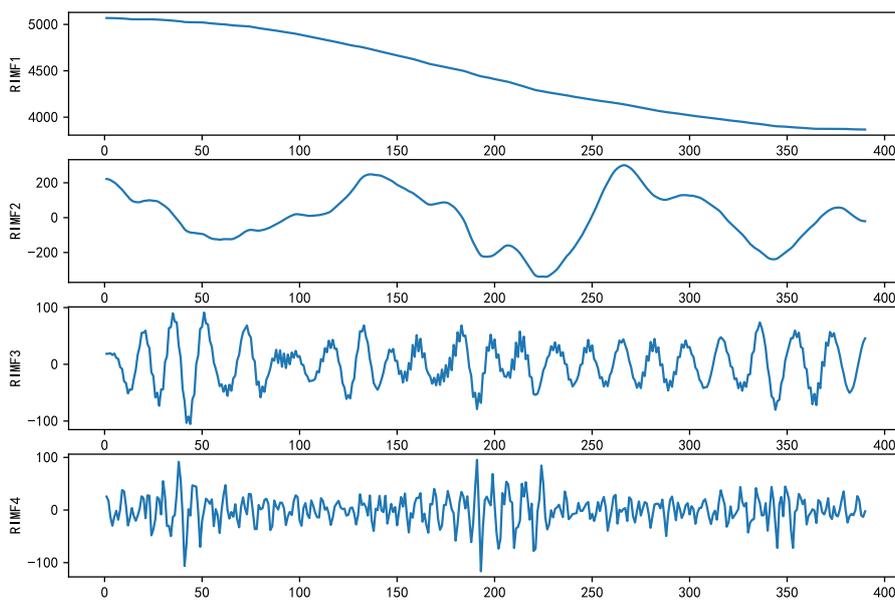


图 4.7 沪深 300 重构子序列

4.2.3 中国平安

中国平安保险是中国第一家股份制保险企业,已经发展成为金融保险、银行、投资、医疗健康等业务为一体的多元综合金融服务集团,于上海证券交易所上市,在 A 股市场中是极具代表性的一支个股,因此选取中国平安作为本文构建模型的实验对象来进一步验证 VMD-CBAM-BiGRU 模型在股价时间序列预测中的稳定性和优越性。首先验证基础模型 BiGRU 的优越性,通过对 BP 神经网络、GRU 和 BiGRU 模型进行比较分析,从评价指标结果表 4.4 可以看出, BP 神经网络模型在三种模型中表现较差,与 GRU 和 BiGRU 模型存在较大的差距。从图 4.8 可以直观地看出, BP 神经网络模型对于股价涨跌趋势变化的预测能力较差; GRU 模型相比于 BP 神经网络而言对股价波动更为敏感,从评价指标表 4.4 来看, GRU 模型的拟合程度达到了 91.7%, 相比于 BP 神经网络提升了 6%, MAE、MAPE 和 RMSE 分别降低了 20.9%、21.9%和 25%; BiGRU 模型则在对中国平安股票的实验中, BiGRU 模型的拟合程度达到了 94.2%, 相比于 GRU 模型有 2.5%的提升, MAE 值、MAPE 值和 RMSE 分别为 0.961、2.106 和 1.263, 相比于 GRU 模型预测误差分别减少了 18.4%、13.5%和 19.7%, 从预测效果图 4.9 可以看出, BiGRU 模型相比于其他模型对于中国平安股价的预测效果更加稳定精确, 从而验证了选择 BiGRU 模型作为基础模型的优越性。

为了进一步提高模型的预测效果,在模型 BiGRU 的基础上引入 VMD 算法和 CBAM 算法,分别将两种算法和基础模型 BiGRU 相结合进行实验,另外在 VMD 算法接入模型前需要使用模糊熵重构子序列,如图 4.11 所示;从评价指标结果表 4.4 可以看出, VMD 和 CBAM 两种算法引入基础模型后均能提高模型的预测表现,两种算法的提升效果十分接近, VMD 算法的拟合程度相比于基础模型 BiGRU 提升了 3.6%达到了 97.8%, MAE、MAPE 和 RMSE 分别为 0.634、1.461 和 0.797, 预测误差分别降低了 34%、30.6%和 36.8%, CBAM 在引入模型后拟合程度相比于基础模型 BiGRU 提升了 3.4%达到了 97.6%, MAE、MAPE 和 RMSE 分别为 0.647、1.492 和 0.812, 预测误差分别降低了 34%、29%和 35.7%, 从拟合效果图也可以看出, VMD-BiGRU 模型和 CBAM-BiGRU 对中国平安股价变化的跟随程度均好于基础模型 BiGRU, 在极值点处的预测结果也高于 BiGRU 模型, 从而验证了 VMD 算法和 CBAM 算法分别和基础模型融合后对中国平安

股价预测效果的显著提升。

经过上述实验验证后,同时将 VMD 算法和 CBAM 算法融入基础模型 BiGRU 构建最终的 VMD-CBAM-BiGRU 模型,从而进一步提升模型的预测能力,将中国平安数据集代入 VMD-CBAM-BiGRU 模型进行训练并预测得到评价指标结果和预测效果拟合图,从评价指标可以看出,VMD-CBAM-BiGRU 模型的拟合程度达到了 99.5%,相比于 VMD-BiGRU 模型和 CBAM-BiGRU 模型分别提高了 1.7% 和 1.9%,MAE、MAPE 和 RMSE 分别为 0.279、0.655 和 0.361,相比于 VMD-BiGRU 模型分别降低了 55.9%、55.1%和 54.7%,相比于 CBAM-BiGRU 模型分别降低了 56.8%、56.8%和 55.5%,相比于基础模型 BiGRU 则降低了 70.9%、68.8%和 71.4%,从拟合效果图也可以看出 VMD-CBAM-BiGRU 模型对于中国平安次日股票收盘价的预测结果紧跟股价波动的变化,误差在所有模型中最小,最能反映股价变化的趋势。

在中国平安个股股价的预测分析中,充分验证了 VMD 算法和 CBAM 算法与 BiGRU 模型融合后构建的 VMD-CBAM-BiGRU 模型在预测股价时间序列上的稳定性和精确性。

表 4.4 中国平安评价指标结果

	R²	MAE	MAPE	RMSE
VMD-CBAM-BiGRU	0.995	0.279	0.655	0.361
VMD-BiGRU	0.978	0.634	1.461	0.797
CBAM-BiGRU	0.976	0.647	1.492	0.812
BiGRU	0.942	0.961	2.106	1.263
GRU	0.917	1.145	2.464	1.573
BP 神经网络	0.853	1.449	3.155	2.100

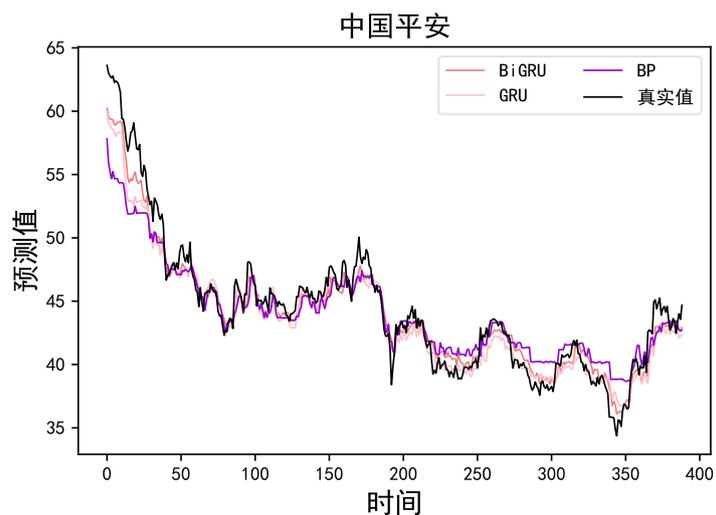


图 4.8 中国平安拟合效果图

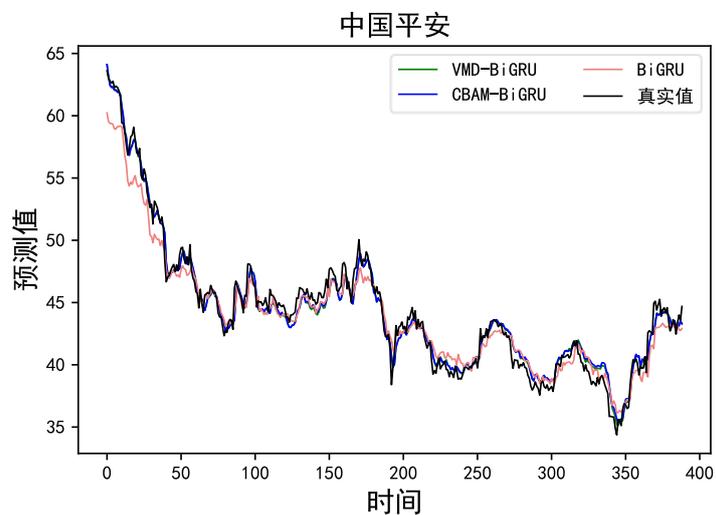


图 4.9 中国平安拟合效果图

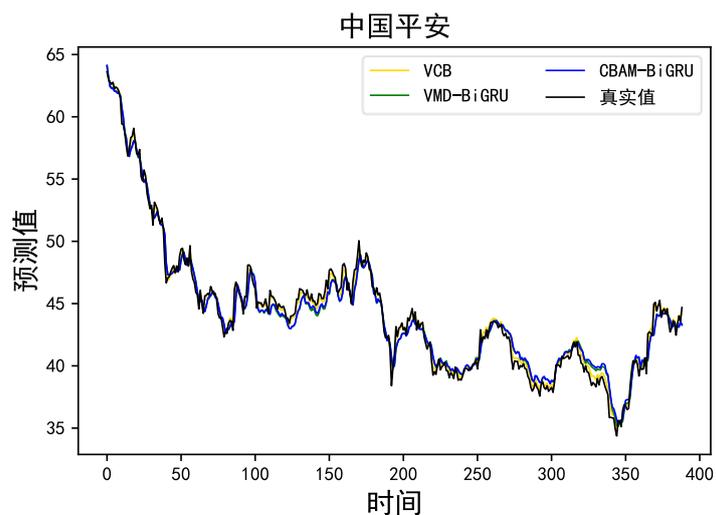


图 4.10 中国平安拟合效果图

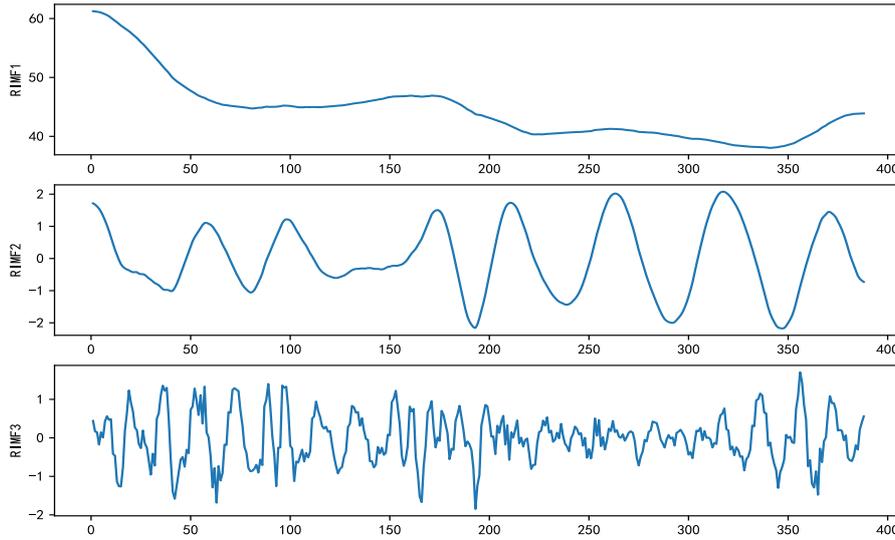


图 4.11 中国平安重构子序列

4.2.4 贵州茅台

为了进一步验证所构建模型对股票市场中个股的预测效果,本文又选取了贵州茅台个股进行实验验证,贵州茅台酒股份有限公司涉足产业包括白酒、保健酒等多个领域,主导产品贵州茅台酒历史悠久、源远流长,具有深厚的文化内涵,于上海证券交易所上市,是 A 股市场中极具代表性的一支股票。首先使用 BP 神经网络、GRU 和 BiGRU 模型对茅台股价进行预测,从评价指标结果表 4.5 可以看出,BP 神经网络模型在三种模型中表现相对较差,与 GRU 和 BiGRU 模型存在一定的差距。从图 4.12 可以地看出,GRU 模型和 BiGRU 模型对股价变化的跟随效果较好,能较准确地预测极值点地变化,反观 BP 神经网络则表现乏力,不能准确的跟随股价的涨跌趋势,从评价指标来看,GRU 模型的拟合程度达到了 89%,相比于 BP 神经网络提升了 5.4%,相比于 BP 神经网络各项误差分别降低了 19.6%、19.8%和 18.3%;BiGRU 模型在对贵州茅台股票的的实验中拟合程度达到了 92.2%,相比于 GRU 模型有 3.2%的提升,MAE 值、MAPE 值和 RMSE 相比于 GRU 模型分别减少了 16.4%、14.8%和 15.6%,从预测效果图也可以看出,BiGRU 模型相比于其他几个模型对于贵州茅台股价的预测效果更加稳定精确,验证了将 BiGRU 模型作为预测基础模型的优越性。

为了进一步提高模型的预测效果,在基础模型 BiGRU 的基础上引入 VMD 算法以及 CBAM 算法,在 VMD 算法将茅台价格序列分解为子序列后,使用模

糊熵方法重构子序列,然后分别将两种算法和基础模型 BiGRU 相结合进行实验,从评价指标结果表 4.5 可以看出, VMD 和 CBAM 两种算法引入基础模型后均能提高模型的预测表现,两种算法的提升效果十分接近, VMD-BiGRU 模型的拟合程度相比于基础模型 BiGRU 提升了 3.6%达到了 95.8%, MAE、MAPE 和 RMSE 的预测误差分别降低了 26.9%、25.3%和 26.9%, CBAM 在引入模型后拟合程度相比于基础模型 BiGRU 提升了 4.7%达到了 96.9%, MAE、MAPE 和 RMSE 的预测误差分别降低了 37.9%、36.3%和 36.87%,从拟合效果图 4.13 也可以看出, VMD-BiGRU 模型和 CBAM-BiGRU 对贵州茅台股价变化的跟随程度均好于基础模型 BiGRU,在极值点处的预测准确度也高于 BiGRU 模型,从而验证了 VMD 算法和 CBAM 算法分别和基础模型融合后对贵州茅台股价预测效果的显著提升。

在经过上述实验验证后,同时将 VMD 算法和 CBAM 算法融入基础模型 BiGRU 构建 VMD-CBAM-BiGRU 模型,从而进一步提升模型的预测能力,将贵州茅台数据集代入 VMD-CBAM-BiGRU 模型进行训练并预测得到评价指标结果和预测效果拟合图,从评价指标可以看出, VMD-CBAM-BiGRU 模型的拟合程度达到了 98.7%,相比于 VMD-BiGRU 模型和 CBAM-BiGRU 模型分别提高了 2.9%和 1.8%, MAE、MAPE 和 RMSE 相比于 VMD-BiGRU 模型分别降低了 55.33%、54.1%和 46%,相比于 CBAM-BiGRU 模型分别降低了 47.3%、46.1%和 37.5%,相比于基础模型 BiGRU 则降低了 67.3%、65.7%和 60.58%,从拟合效果图也可以看出 VMD-CBAM-BiGRU 模型对于贵州茅台次日股票收盘价的预测结果紧跟股价波动的变化,误差远远小于其他对比模型。

在贵州茅台个股股票的预测分析中,进一步验证了 VMD 算法和 CBAM 算法与 BiGRU 模型融合后构建的 VMD-CBAM-BiGRU 模型在预测个股股价时间序列上的稳定性和优越性和稳定性。

表 4.5 茅台评价指标结果

	R²	MAE	MAPE	RMSE
VMD-CBAM-BiGRU	0.987	11.288	0.671	17.27
VMD-BiGRU	0.958	25.275	1.461	32.017
CBAM-BiGRU	0.969	21.46	1.247	27.663
BiGRU	0.922	34.576	1.958	43.82
GRU	0.89	41.31	2.3	51.959
BP 神经网络	0.836	51.382	2.867	63.551

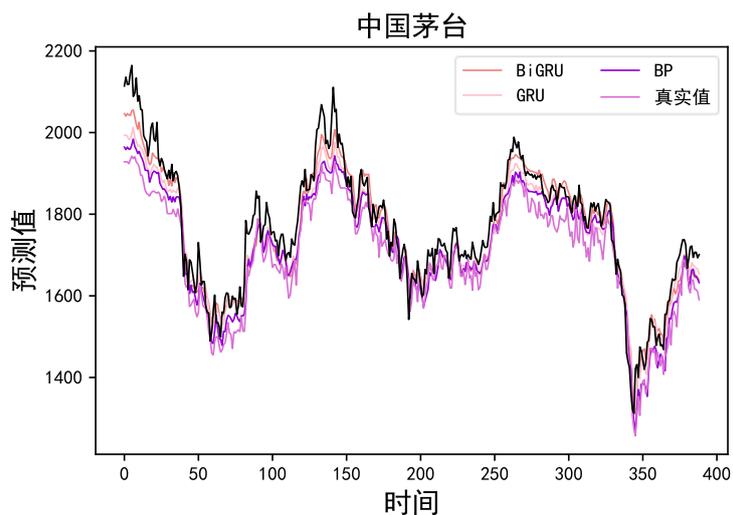


图 4.12 茅台拟合效果图

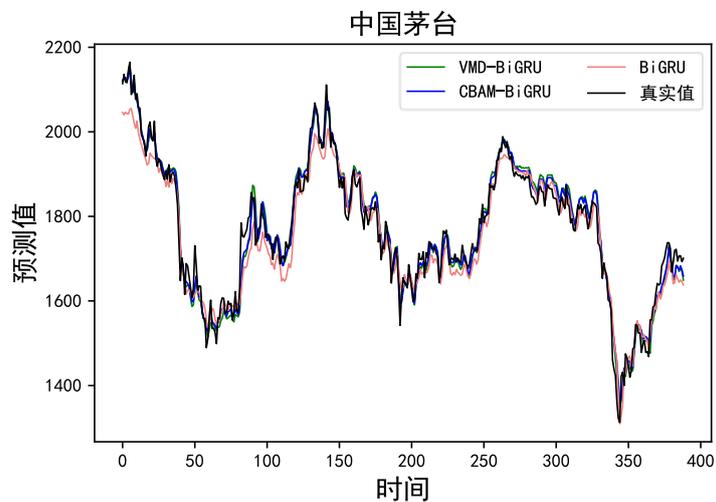


图 4.13 茅台拟合效果图

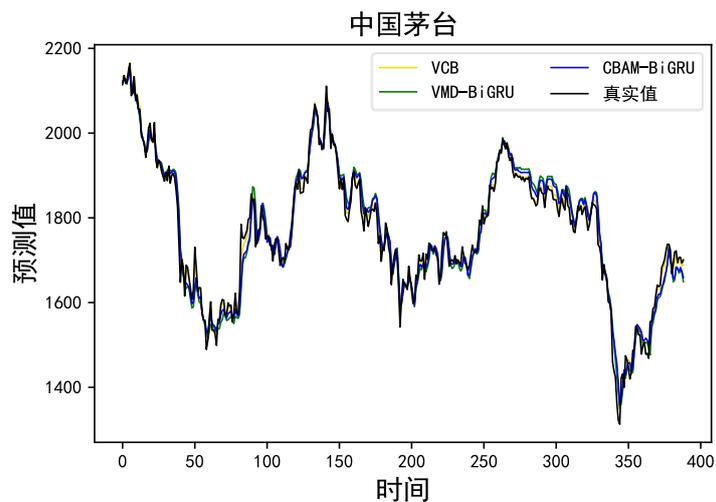


图 4.14 茅台拟合效果图

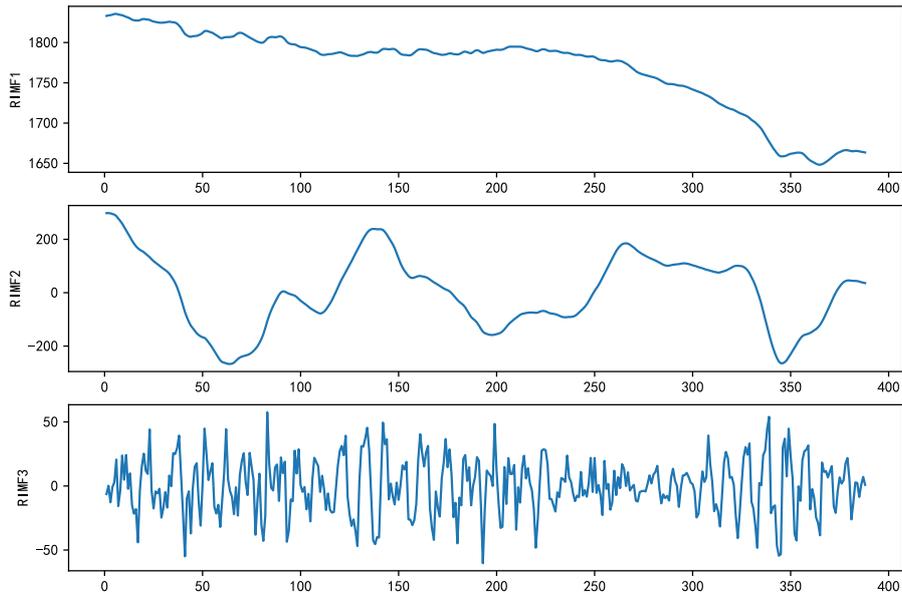


图 4.15 茅台重构子序列

4.3 不同跨度时期下股价预测效果分析

经过上一节在四支股票上的实验，验证了本文所构建的 VMD-CBAM-BiGRU 模型在股票预测中的稳定性以及优越的性能，在上述实验中预测对象均为次日收盘价，为了验证所构建模型在不同时期跨度下的预测效果，将次日收盘价、三日收盘价和七日收盘价分别作为预测对象进行模型的预测实验，最终得到如下表 4.6 所示的不同时期的股价评价指标结果以及如图 4.16、4.17 和 4.18 所示的不同跨度时期的预测拟合效果图。从评价指标结果可以看出，随着时间跨度变大，模型的预测效果逐渐变弱，拟合程度从次日收盘价到七日收盘价降低了 14% 左右，各项评价指标平均增大了 90%，但是相对于七日收盘价较差的预测效果，三日收盘价的预测效果虽然有所降低但仍有 94% 的拟合程度，在股价预测中仍能发挥较优秀的作用。

总的来说，本文所构建的 VMD-CBAM-BiGRU 模型在对次日收盘价以及三日收盘价的股价预测中均能取得较好的预测表现，在日期跨度较大的七日收盘价的股价预测中表现一般，因此本文所构建模型比较适合中短期股价预测并能取得优越的预测表现。

表 4.6 不同日期跨度下股价预测评价指标结果

	R^2	MAE	MAPE	RMSE
次日收盘价	0.998	5.415	0.163	7.292
三日收盘价	0.941	38.188	1.153	50.862
七日收盘价	0.851	66.115	2.006	83.728

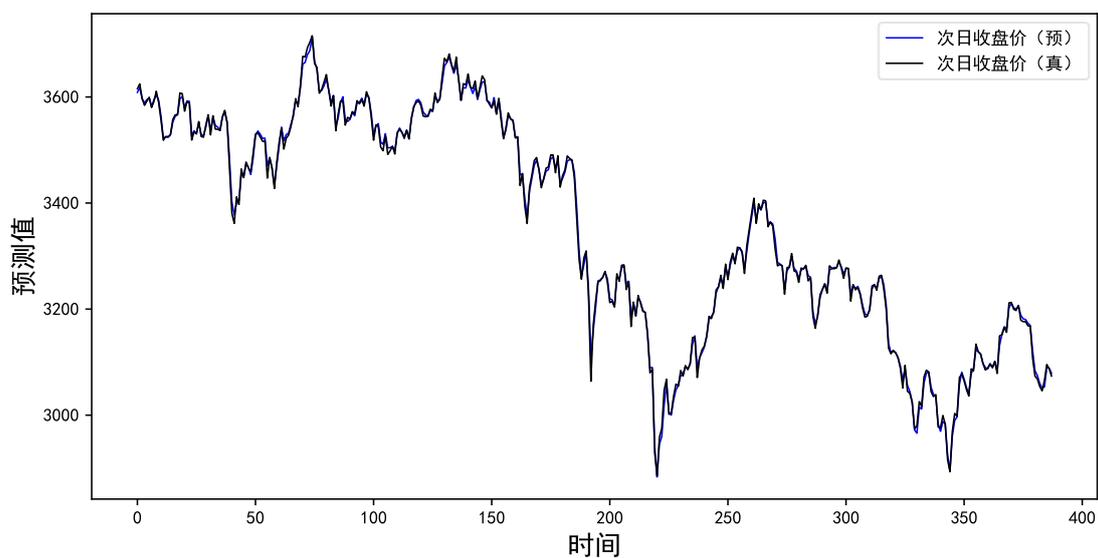


图 4.16 次日收盘价拟合效果图

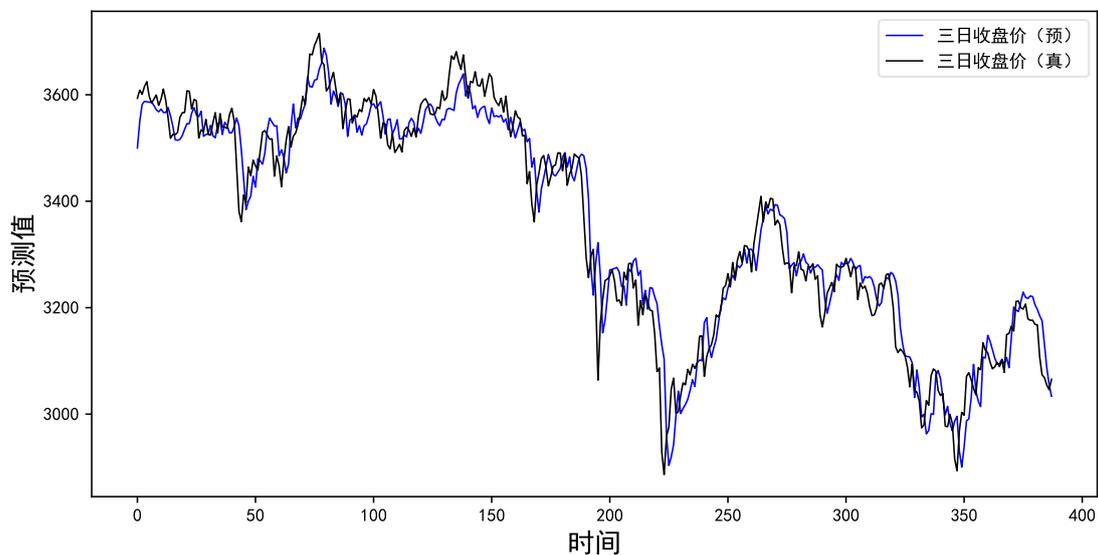


图 4.17 三日收盘价拟合效果图

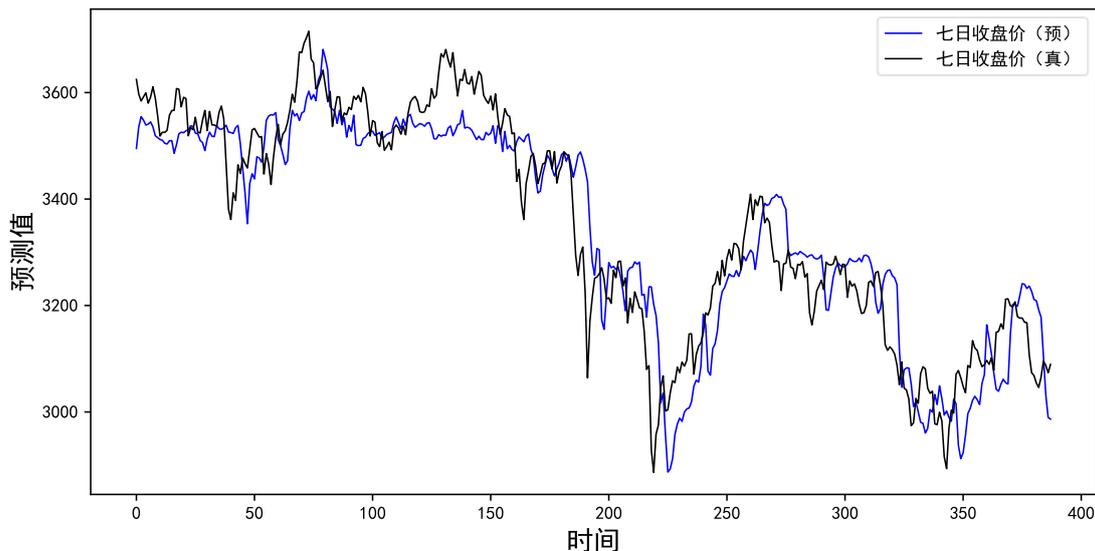


图 4.18 七日收盘价拟合效果图

4.4 模拟交易

在 4.2 和 4.3 两节内容中，分别验证了所构建 VMD-CBAM-BiGRU 模型在预测股价序列中优秀且稳定的预测效果以及在中短期股价预测中的稳定性。但模型在股价预测中达到很好的预测精度并不意味着也一定能在实际投资交易中取得良好的收益或者降低风险，因此在本节进行模拟交易来验证模型策略在实际应用中的效果。本节设计了持入卖出型策略以及基于各种模型的策略两种方式进行模拟交易，持入卖出型策略即从买入节点开始持有股票，持有到最后卖出；基于模型的策略为从买入节点开始持有股票直到预测股价开始下跌前卖出。本模拟交易实验选择夏普率作为交易策略的评价指标，夏普率越大则表明策略获得的收益越高。

在 4.2 节中的四支股票测试集上进行模拟交易的实验后得出如下表 4.7 所示的结果，从表中可以看出，相比于其他各模型策略以及持入卖出型策略，本文所构建的 VMD-CBAM-BiGRU 模型作为投资策略得到的夏普率最高并且获利更为稳定，验证了 VMD-CBAM-BiGRU 模型作为投资策略在股市投资中能获取较高且稳定的收益。

表 4.7 不同交易策略下的模拟交易效果

股票	投资策略	夏普率
上证综指	VMD-CBAM-BiGRU	0.674
	VMD-BiGRU	0.582
	CBAM-BiGRU	0.595
	BiGRU	0.137
	GRU	0.084
	BP 神经网络	-0.038
	持有卖出策略	-0.093
沪深 300	VMD-CBAM-BiGRU	0.536
	VMD-BiGRU	0.352
	CBAM-BiGRU	0.375
	BiGRU	0.097
	GRU	0.073
	BP 神经网络	-0.014
	持有卖出策略	0.039
中国平安	VMD-CBAM-BiGRU	0.721
	VMD-BiGRU	0.531
	CBAM-BiGRU	0.577
	BiGRU	0.215
	GRU	0.196
	BP 神经网络	-0.353
	持有卖出策略	0.082
贵州茅台	VMD-CBAM-BiGRU	0.437
	VMD-BiGRU	0.251
	CBAM-BiGRU	0.228
	BiGRU	0.096
	GRU	-0.017
	BP 神经网络	-0.098
	持有卖出策略	-0.079

5 总结与展望

5.1 总结

本文从选择基础预测模型入手,针对一系列基础模型在上证综指、沪深 300、中国平安以及茅台等四支著名股票上进行实验验证,最终选择处理时间序列数据优势较为明显的 BiGRU 模型作为本文的基础模型。BiGRU 模型的结构具有很强的延展性,可以接入其他深度学习模块,为了进一步增强 BiGRU 模型的预测性能,同时考虑到股价时间序列具有大量的指标,因此最终选择将处理特征数据优势明显的 CBAM 模块接入 BiGRU 模型;考虑到股价序列的强噪声和高度非平稳性,为了达到降噪目的,本文选择变分模态分解算法对股价序列进行分解处理,并通过模糊熵进行序列重构提高模型预测效率;另外由于基础模型 BiGRU 包含大量超参数,超参数的好坏影响着模型的表达效果,因此本文使用粒子群算法对模型进行超参数寻优,并在优化算法基础上就进行了动态的改进增强了参数的全局寻优能力。

最终经过一系列实验验证了构建出的 VMD-CBAM-BiGRU 模型在股票预测领域优越的预测性能,并且经过不同时间跨度的预测实验验证了所构建模型在中短期股价预测中的实用性和精确性,最后使用 VMD-CBAM-BiGRU 模型作为交易策略与其他各交易策略作对比验证了模型在股票投资中优越稳定的获利效果。

5.2 研究展望

本论文的研究工作以 VMD-CBAM-BiGRU 模型为基础,利用该模型对股价进行预测,证实了该方法的有效性和可行性。但是本文还有很多地方有待深入研究和完善。

第一,在群智能优化算法方面,本文仅使用了动态粒子群算法这一优化算法,并未引入其他智能优化算法进行对比,在参数寻优的过程中由于设备性能限制也未能进行更大范围区间的参数寻优。

第二,在分解算法的选择上,本为仅使用了 VMD 算法进行序列分解,未尝试与更多分解算法进行对比分析。

第三，在股价影响因素的选择上，本文虽在基础行情指标的基础上加入了大量的技术指标因素，但未纳入情感指标因素，未来会尝试纳入更多类型指标以期进一步提高股票预测效果。

参考文献

- [1] L.Zhang, Y.Dong, J.Wang. Wind speed forecasting using a two-stage forecasting system with an error correcting and nonlinear ensemble strategy[J]. IEEE Access, 2019, 7(99): 1-1.
- [2] Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions[J]. European Journal of Operational Research, 2018, 270(2): 654-669.
- [3] Y.Lin, Y.Yan , J.Xu, et al. Forecasting stock index price using the CEEMDAN-LSTM model[J]. The North American Journal of Economics and Finance, 2021,57: 101421.
- [4] W.Lu , J.Li, J.Wang, et al. A CNN-BiLSTM-AM method for stock price prediction[J]. Neural Computing and Applications, 2021, 33: 4741-4753.
- [5] Z.Yu, L.Qin, Y.Chen, et al. Stock price forecasting based on LLE-BP neural network model[J]. Physica A: Statistical Mechanics and Its Applications, 2020, 553: 124197.
- [6] Bose A, Hsu C H, Roy S S, et al. Forecasting stock price by hybrid model of cascading multivariate adaptive regression splines and deep neural network[J]. Computers and Electrical Engineering, 2021, 95: 107405.
- [7] Z.Zhou, M.Gao, Q.Liu, et al. Forecasting stock price movements with multiple data sources: Evidence from stock market in China[J]. Physica A: Statistical Mechanics and its Applications, 2020, 542: 123389.
- [8] Peng Y, Albuquerque P H M, Kimura H, et al. Feature selection and deep neural networks for stock price direction forecasting using technical analysis indicators[J]. Machine Learning with Applications, 2021, 5: 100060.
- [9] Yun K K, Yoon S W, Won D. Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process[J]. Expert Systems with Applications, 2021, 186: 115716.
- [10] Kumar Chandar S. Grey wolf optimization-Elman neural network model for stock price prediction[J]. Soft Computing, 2021, 25: 649-658.

- [11] Ingle V, Deshmukh S. Ensemble deep learning framework for stock market data prediction (EDLF-DP)[J]. *Global Transitions Proceedings*, 2021, 2(1): 47-66.
- [12] Chandar S K. Hybrid models for intraday stock price forecasting based on artificial neural networks and metaheuristic algorithms[J]. *Pattern Recognition Letters*, 2021, 147: 124-133.
- [13] 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.
- [14] 张贵生,张信东.基于梯度因子的 ARMA-GARCH 股票价格预测模型研究[J].*山西大学学报(哲学社会科学版)*, 2016, 39 (01): 115-122.
- [15] 罗志丹,刘英,郭伟.基于跟踪微分器的泰勒展开与 ARIMA 混合模型对股票价格的短期预测[J].*数学的实践与认识*, 2019,49 (23): 67-77.
- [16] 董曼茹,唐晓彬.基于双层 Attention 机制的 LSTM 模型对 CPI 的预测研究[J/OL].*中国管理科学*,2024, 1-14.
- [17] 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.
- [18] 杨静凌,唐国强,张建文.基于 EEMD-Elman-Adaboost 的中美股票价格预测研究[J].*运筹与管理*, 2022, 31 (11): 194-199.
- [19] 何泳,李环.改进的 NSGA-III-XGBoost 算法在股票预测中的应用[J].*计算机工程与应用*, 2023, 59 (18): 293-300.
- [20] 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.
- [21] 赵帅斌,林旭东,翁晓健.基于经验模态分解与投资者情绪的 Attention-BiLSTM 股价趋势预测模型[J].*计算机应用*, 2023, 43 (S1): 112-118.
- [22] 邓德军,徐洪珍,韦诗玥. E-V-ALSTM 模型的股价预测[J].*计算机工程与应用*, 2023, 59 (06): 101-112.
- [23] 刘洋,张雯,胡毅等.基于多模态深度学习的酒店股票预测[J].*数据分析与知识发现*, 2023, 7 (05): 21-32.
- [24] 方义秋,卢壮,葛君伟.联合 RMSE 损失 LSTM-CNN 模型的股价预测[J].*计算机工程与应用*, 2022, 58 (09): 294-302.

- [25] 徐月梅, 王子厚, 吴子歆. 一种基于 CNN-BiLSTM 多特征融合的股票走势预测模型[J]. 数据分析与知识发现, 2021, 5 (07): 126-137.
- [26] 张倩玉, 严冬梅, 韩佳彤. 结合深度学习和分解算法的股票价格预测研究[J]. 计算机工程与应用, 2021, 57 (05): 56-64.
- [27] 綦方中, 林少倩, 俞婷婷. 基于 PCA 和 IFOA-BP 神经网络的股价预测模型[J]. 计算机应用与软件, 2020, 37 (01): 116-121.
- [28] 包振山, 郭俊南, 谢源等. 基于 LSTM-GA 的股票价格涨跌预测模型[J]. 计算机科学, 2020, 47 (S1): 467-473.
- [29] 陈凯杰, 唐振鹏, 吴俊传等. 基于分解-集成和混频数据采样的中国股票市场预测研究[J]. 系统工程理论与实践, 2022, 42 (11): 3105-3120.
- [30] 张婷婷, 唐振鹏, 吴俊传. 基于优化 KELM 模型的股票指数预测方法[J]. 统计与决策, 2021, 37 (13): 148-150.
- [31] 邹婕, 李路. RF-SA-GRU 模型的股价预测研究[J]. 计算机工程与应用, 2023, 59 (15): 300-309.
- [32] 闫政旭, 秦超, 宋刚. 基于 Pearson 特征选择的随机森林模型股票价格预测[J]. 计算机工程与应用, 2021, 57(15): 286-296.
- [33] 王伦, 李路. 基于 gcForest 的多因子量化选股策略[J]. 计算机工程与应用, 2020, 56(15): 86-91.
- [34] 张雪芳, 温馨. 基于 XGBoost 的股指涨跌预测策略研究[J]. 计算机与数字工程, 2023, 51(03): 686-689.
- [35] 李新尧. 基于 SSA-LSTM 神经网络的股票价格预测研究[J]. 信息系统工程, 2023(03): 48-50.
- [36] 杨青, 王晨蔚. 基于深度学习 LSTM 神经网络的全球股票指数预测研究[J]. 统计研究, 2019, 36(03): 65-77.
- [37] 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.
- [38] 景楠, 史紫荆, 舒毓民. 基于注意力机制和 CNN-LSTM 模型的沪铜期货高频价格预测[J/OL]. 中国管理科学, 2020, 1-13.
- [39] 赵红蕊, 薛雷. 基于 LSTM-CNN-CBAM 模型的股票预测研究[J]. 计算机工程与

- 应用,2021,57(03):203-207.
- [40] 贺毅岳, 李萍, 韩进博. 基于 CEEMDAN-LSTM 的股票市场指数预测建模研究[J]. 统计与信息论坛, 2020, 35(06): 34-45.
- [41] 程杰, 陈鼎, 李春等. 基于 GWO-CNN-BiLSTM 的超短期风电预测[J]. 科学与技术, 2023, 23 (35): 15091-15099.
- [42] 张新生, 魏志臻, 陈章政等. 基于 LASSO-GWO-KELM 的工业碳排放预测方法研究[J]. 环境工程, 2023, 41 (10): 141-149.
- [43] 史加荣, 王双馨. 基于 VMD-BiLSTM-WOA 的短期风电功率预测[J]. 陕西科技大学学报, 2024, 42 (01): 177-185.
- [44] 王瑞, 马祯, 李磊. 基于 CEEMDAN-WOA-SVR 的高铁沿线超短期风速预测方法[J]. 中国铁道科学, 2023, 44 (06): 80-86.
- [45] 王少钦, 林婧姝, 郭明浩. 基于 EMD-SSA-BiLSTM 的桥址区短期风速预测[J/OL]. 计算机仿真, 2023, 1-8.
- [46] 赵全明, 李珂, 王笑欢等. 基于 SSA-GRU 神经网络的超短期风电功率预测[J]. 传感器与微系统, 2023, 42 (11): 151-155.
- [47] 张成龙, 郑凯, 刘杰. 基于小波包能量谱和改进 FOA-GRNN 的轴承寿命预测[J]. 组合机床与自动化加工技术, 2020, (07): 73-76.
- [48] 施爱平, 钱震威, 李英豪等. 基于 GA-LSTM 的酿酒葡萄霜霉病预测方法研究[J]. 中国农机化学报, 2023, 44 (10): 144-151.
- [49] 全一鸣, 喻敏, 王文波等. 基于分形优化的 VMD 和 GA-BP 的短期风速预测[J]. 太阳能学报, 2023, 44 (07): 436-446.
- [50] 陈石毓, 李壮举, 刘浩等. SA-GA-CNN-LSTM 新型冷暖系统负荷预测方法研究[J]. 中国测试, 2023, 49 (09): 115-122.
- [51] 牛红丽, 赵亚枝. 利用 Bagging 算法和 GRU 模型预测股票价格指数[J]. 计算机工程与应用, 2022, 58 (12): 132-138.
- [52] 邹婕, 李路. 基于随机森林的 SA-BiGRU 模型的股票价格预测研究[J]. 中国物价, 2023, (11): 52-56.

致谢

写在这儿，三年的研究生旅程已经到了终途，一路走来，这条求学之路艰辛但也充满了乐趣，在此特别感谢三年来研究生生涯的引路人郭精军老师，郭老师在学习和生活中给予我无与伦比的关怀，令我时刻充满着前进的动力，郭老师浑厚的学术造诣和高尚的品德是我一直以来的前进目标，每周两次的组会中总能让我有新的知识收获和思考，在论文撰写的过程中，从开题到最后的收笔，郭老师即使工作再忙也会抽出空闲在每个环节都给予我细致的指导，帮助我渐渐完善论文，另外在生活中遇到麻烦事郭老师也会很热心的开导我，给予我最大的帮助，使我走出困境，今后也将以老师为榜样前行，另外，还特别感谢父母在读研期间给予我在物质和精神双层面的支持，分享我的喜悦，抚慰我的忧虑，让我能够没有后顾之忧地投入到学术的海洋中，也很感激这三年中对我带来帮助的所有朋友和同学，感谢各位的一路陪伴，让我能在求学期间丰富多彩，文已致此，祝大家前程似锦，一帆风顺！