学校代码：10270　　学号：212502550

上海师范大学

硕士专业学位论文

基于多维度因子的 ATT-LSTM 模型的股票量化策略设计—以沪深 300 成分股为例

学　　院：商 学 院

专业学位类别：金融专业硕士

专 业 领 域：数据分析方向

论 文 类 型：交易策略设计

研 究 生 姓 名：曾 新 月

指 导 教 师：文 燕 平

完 成 日 期：2023 年 5 月 10 日

论文题目：基于多维度因子的 ATT-LSTM 模型的股票量化交易策略设计—以沪深 300 成分股为例

论文类型: 交易策略设计

专业方向：金融数据分析

摘要

中国股票市场近年来迅速发展，股票市场作为股权的流通市场，能促使中国的整个金融体系更加稳定，同时股价也反映了对应上市公司的经营情况，但实际中股市容易受到诸多因素影响。与此同时近年来量化交易也在逐渐盛行，因此探索如何利用股票市场的数据来对股票价格进行预测有重大意义，若能通过建立适当的模型取得较好的股价预测结果，则能为量化交易策略提供良好基础。

本文的研究样本为 2017-2021 年的沪深 300 所有成分股的日数据，为了更加完整地涵盖各种影响股价的因素，本文选择了四个维度(行情因素指标、关联市场因素、国内市场因素与技术指标)的 38 个影响因子，先后通过 IC 值与相关系数检验筛选出 18 个有效输入指标。在数据处理上，对收盘价数据选用 Db8 函数进行小波降噪处理，以及对所有数据进行异常值、缺失值、归一化预处理，将这些数据用来对沪深 300 成分股的收盘价进行预测，横向比较显示相比较于 LSTM 模型，ATT-LSTM 模型泛化能力更强，对股价的拟合优度 R^2 更高，MSE、RMSE、MAE 误差更小。

由于考虑到实际股票交易过程中的股市整体走势会对短期交易产生较大的影响，受到动量效应的启发，本文创建了均线距离指标 ΔD_{MA}，并通过分析其在上期与本期的正负值与变化趋势来确定开仓买入的标的数量。在量化策略构建过程中，首先利用 ATT-LSTM 模型对沪深 300 所有成分股次日收盘价进行预测，并计算收盘价对数收益率，将收益率排名靠前的标的作为股票池的待选标的，结合均线距离指标 ΔD_{MA} 确定每期买入股票数量，每三天强制换仓。在 2020 年 6 月 28 日至 2021 年 10 月 30 日期间策略回测的结果为年化收益率 24.44%，高于沪深 300 基准年化收益率 19.76%。同时交易的 65%个交易日都处于盈利状态，夏普比率达 1.24，波动率为 0.176%，因此可以得知策略能够有效抵御市场风险，收益情况较好。

因此，本文结合深度学习模型构建的股票量化交易策略能获得较好的收益，能给予各类投资者一定借鉴价值。

关键词： ATT-LSTM；均线距离指标；量化交易设计；沪深 300 成分股

Abstract

The Chinese stock market has developed rapidly in recent years. As a market for the circulation of equity, the stock market is beneficial for the overall stability of China's financial system. At the same time, stock prices also reflect the operating conditions of corresponding listed companies. However, the stock market is susceptible to many factors, and at the same time, quantitative trading is gradually becoming popular. Therefore, how to use stock market data to predict stock prices is the main content of research. By establishing appropriate models to obtain good prediction results, it can provide a good foundation for the design of quantitative trading strategies.

The research sample of this paper is the daily data of all constituent stocks of CSI 300 Index from 2017 to 2021. In order to more completely cover various factors affecting stock prices, this paper selects 38 influencing factors from four dimensions (market factor indicators, related market factors, domestic market factors and technical indicators), and successively selects 18 effective input indicators through IC value and correlation test, and uses Db8 function wavelet denoising for closing price, Subsequently, Outlier, missing value and normalization preprocessing were performed on all data, and then these data were used to predict the closing price of CSI 300 Index constituent stocks. Horizontal comparison shows that ATT-LSTM model has a higher goodness of fit R^2 for stock prices than LSTM model, and MSE, RMSE and MAE errors are smaller.

Due to the fact that the overall trend of the stock market in the actual stock trading process will have a significant impact on short-term trading, inspired by the momentum effect, this article creates the moving average distance indicator ΔD_{MA} and determines the number of open positions to buy by analyzing its positive and negative values and changing trends in the previous and current periods. In the process of building a quantitative trading strategy, first use the ATT-LSTM model to predict the next day's closing prices of all constituent stocks of CSI 300 Index, calculate the logarithmic return rate of closing prices, take the subject with the highest return rate as the subject to be selected in the stock pool, determine the number of stocks to be purchased in each period in combination with the average distance index ΔD_{MA}, and force the position change every three days, so as to increase the position when the market is optimistic and reduce the number of stocks to be purchased when the market is pessimistic, Finally, during the period from June 28, 2020 to October 30, 2021, the annual return rate of the

strategy backtesting results reached 24.44%, which was 19.76% higher than the benchmark annual return rate of CSI 300 Index. At the same time, from the proportion of profit days in a single day, 65% of the trading days are profitable, the Sharpe ratio is 1.24, and the volatility is 0.176%. Therefore, we can know that the risk of this strategy is low, the strategy built can effectively resist market risk, and the yield is good.

Therefore, the stock quantitative trading strategy constructed by combining deep learning models in this article can achieve good returns and provide certain reference value to various investors.

Key words: ATT-LSTM; Average Line Distance Index; Quantitative Transaction Design; CSI 300

目　录

第 1 章 绪论

1.1 研究的背景

股票市场不论对于哪个国家而言都对其经济与社会发展起着重要的作用，由于股票市场的变动而产生的一些金融活动会渗透在国家的各个方面，同时也同步会影响世界上其他国家的经济与社会发展。随着中国经济的迅速崛起，中国的股票市场也逐渐成为全球股票市场的重要组成部分。但相较于国外股票市场，我国股票市场发展时间还不够长，同时也不足够稳定，所以不论是个人还是机构投资者都可能在股票市场中碰壁。

近年来，随着数据采集和处理技术的不断提高，人工智能和及其学习等新技术的应用也逐渐渗透到股票分析领域。在这样的背景下，股票市场分析不仅利用了各种数学及数学应用，还运用了一些新技术与新算法，各种新的分析方法层出不穷。尤其是在算法层面，在金融市场的不同细分领域当中都有各类学者工程师们提出了各种算法作为技术支持。在中国的股票市场研究分析当中常被采用的方法是基本面分析与技术分析两种，前者主要是利用对宏观情况的把控以及对公司经营情况的分析来得出结论，而在技术分析当中常常是利用对过往历史的日数据或者分钟数据来建立模型并分析，通过对时间序列的把控来预测未来情况发展。

除了利用时间序列来进行分析，研究人员还采用了各种人工智能，机器学习相关的方式来对股票市场进行分析。机器学习模型此前已经在许多领域得到广泛应用，不论是医疗保健、金融服务、智能交通、电子商务还是社交媒体等领域。在医疗保健方面，机器学习模型可用于辅助医生和研究人员进行疾病诊断、治疗方案制定以及疾病趋势预测。在金融服务领域，机器学习模型可用于风险评估、欺诈检测以及客户分类等任务。在智能交通领域，机器学习模型可用于车辆导航、交通管制以及自动驾驶等方面。在电子商务和社交媒体方面，机器学习模型可用于基于用户行为和兴趣的推荐和个性化服务。总之，机器学习模型已成为现代社会中不可或缺的技术，对我们的生活和工作产生了重要的影响。

在中国股票市场的分析中，机器学习模型也逐渐成为重要的研究领域。这些模型可以分析海量数据，从中学习模式和趋势，并根据学习到的知识做出预测。尤其是在近年来，随着人工智能技术和大数据技术的不断进步，机器学习模型的应用范围和预测准确率也得到了极大的提升。中国的股票市场是一个不断变化和发展的市场，政策和经济环境的变化会影响股票价格的波动，因此需要不断更新的分析方法来应对市场的挑战。在这样的背景下，利用机器学习模型对中国股票

市场进行分析和预测，可以更好地理解市场的走势，为投资者和交易员提供决策支持，并促进市场的稳定和发展。本文希望在利用深度学习模型的基础上，综合考虑多个维度的影响因素来对股票收盘价进行预测，从而构建出一个对中国股票市场有借鉴意义的量化交易策略。

1.2 研究目的和意义

1.2.1 研究目的

我国股票市场成立于上世纪九十年代，股票市场对于各方都带来一定利益。一方面，公司在股票市场中进行股权融资能够有效的减少其纯债务的风险，另一方面，在股票市场当中也存在一定高风险高收益的投资机会，但是由于股票市场的波动较大，因此投资者可能会在分析的过程当中出现一定偏差从而带来大量损失，因此研究股票市场并基于此去适度预测分析是十分有必要的。

通过文献查阅发现在金融市场预测领域有许多应用 LSTM(Long Short-Term Memory，长短期记忆网络)模型的例子，LSTM 模型具有良好的预测性能，因此在学术领域有许多改进 LSTM 模型使其预测率更准确的例子，近年来兴起的注意力机制使得 LSTM 模型在股价预测方面更有说服力。另外在因子选择使用和处理方面，若要选择多维度因子数据，则需要考虑到因子使用效率的问题。

本文主要使用的模型是 ATT-LSTM 模型(Attention LSTM，基于注意力机制的 LSTM 模型)，首先查阅文献确定多维度因子，并且使用相关方法对多维度的指标进行筛选并处理，并利用处理好的输入特征与 ATT-LSTM 模型与 LSTM 模型结合从而预测股价。将两者的预测结果进行比较分析后确定更优的预测模型；在基于上述的对比研究确定出预测模型以后，将 ATT-LSTM 模型的预测结果用于本文构建的量化交易策略当中，在此之前需要先对数据进行降噪处理，本文的研究对象为沪深 300 所有成分股，将模型预测的结果进行计算，将收益率排名靠前的股票纳入股票池，并利用均线指标判断大盘乐观与悲观情况进行控仓，从而获得能有一定超额收益且能控制风险的量化策略。

1.2.2 研究意义

1.理论意义

在大量文献阅读的基础上，本文采用的模型为 ATT-LSTM 模型与 LSTM 模型，结合上述两种模型并选择利用多维度的因子进行股价预测，通过选用更好的预测模型来作为基础。同时本文不仅在因子使用筛选上有创新，能够更全面地考

虑对不同维度因子对股价的影响，在构建量化交易策略的过程中加入了均线距离指标 ΔD_{MA} 来控制整个交易的仓位与风险，增加了本策略的可实施性。

因此，本文将多维度因子通过 IC 值(Information Coefficient，因子信息指数)检验与相关系数检验并将其分别应用于 ATT-LSTM 与 LSTM 模型当中来对股票价格进行预测，选用效果更好的模型来进行最后的预测，选定好模型后将预测后的股价收益率排名前 35 的股票作为待选股票标的，并利用均线距离指标 ΔD_{MA} 来控仓，最后通过加入止损线来进行优化，因此在提高预测准确性的同时，对量化交易策略设计也具有重要意义。

2.现实意义

机器学习模型能够使得在对股价预测的过程中能有更理性、更迅速且更精准的结论，因此在构建量化交易策略的过程当中能够有更好的表现，有利于二级市场的投资者在风险可控的基础上获取适当利益。由于中国股市的发展，股票市场越来越具有可分析性，并且在股票市场投资能获得一定的超额收益，对于普通民众来说，增加了多样性资产配置的可能性，同时享受了经济增长的好处，能够适当地推动民生。股票市场的发展及流动性的增加，能够推动股市相关制度更规范。另外，本文使用沪深 300 成分股作为研究对象，在构建股票池时能考虑综合更多的行业，因此使用 IC 检验与相关系数检验对各维度数据进行处理并使用添加进 ATT-LSTM 模型能够给予想要从事量化投资领域的研究者一些思路。

1.3 研究内容与技术路线

1.3.1 研究内容

股票市场是中国金融市场的重要组成部分，它可以提供直接融资场所，并具有调节资源配置和促进宏观经济发展的能力。此外，股票市场还能为宏观经济、社会和投资者带来积极影响，但是股票市场投资仍存在较大风险，进行股市预测研究能够有效控制风险。只有通过深入研究，才能更好地理解市场动态、制定科学的投资策略并实现更好的风险控制，本文的研究内容如下：

1. 多维度指标的确定及处理

通过文献检索阅读，不仅从 LSTM 模型在股票市场的应用，也考虑其在期货市场中对股指期货的预测应用，如学者邱冬阳根据股指期货理论定价和他本身对于股指期货市场的实际认知，在因子筛选的过程中综合考虑了五个维度的因子，共计 89 个指标：他不仅选取了自身行情因素，还选用了可能会影响该股指的内在因素，关联市场因素，宏观经济因素，以及偶发事件因素，并通过采用维度删减的方法，通过组合不同维度的 LSTM 深度学习模型对沪深 300 股指期货进行

预测，发现 LSTM 模型能很好的描绘沪深 300 股指期货多维高频数据的特征。本文参考其的影响维度，从四个维度(行情因素指标、关联市场因素、国内市场因素、技术指标)共纳入在经济学意义上能够解释股价变动的变量 38 个。由于影响因子众多，虽然构建了多维度因子库，本文还需对因子进行 IC 检验与相关系数检验法进行筛选,使算法的运行速度加快，使特征个数减少并保留有效输入特征。

2. 小波变换降噪

本文使用的是沪深 300 成分股在 2017 年-2021 年的数据，由于股票收盘价时间序列数据中普遍存在噪声，当存在噪声时，会对模型预测精度有较大的影响，所以本文针对收盘价时间序列的数据处理层面选用了各类小波函数来对该数据序列进行降噪，并通过选用效果最好的小波函数将数据分解重构后的时间序列保留待用。

3. ATT-LSTM 模型股价预测

利用 ATT-LSTM 模型来预测股价，通过四个评价指标 MSE(Mean Square Error，均方误差)、RMSE(Root mean squared error, 均方根误差)、MAE(Mean Absolute Error, 平均绝对误差)、MAPE(Mean Absolute Percentage Error，平均绝对百分比误差)、R^2(Coefficient of determination，拟合优度)来确定 ATT-LSTM 模型的预测情况，同时还与加入 Attention 机制之前的 LSTM 模型的预测精度进行对比，来判断 ATT-LSTM 模型是否具有能够在二级市场量化交易预测股价使用。

4. 量化交易策略股票池设定

本文根据预测的股票价格计算收益率并排序，将排名靠前的股票作为投资标的均线距离指标优化量化择时交易，根据均线距离指标来反应换仓期大盘的趋势，从而确定股票池当中可供买卖的股数的设置，从而获得较优的交易策略，并进行模拟交易和策略回测，并对策略进行评价。

通过以上研究思路，能够完成较股价预测模型的实验，并且能够获得一定的超额收益，能够支撑其应用于二级市场，给投身于二级市场投资的个人及机构投资者一定建议。在预测准确的基础上还能在一定程度上规避系统性风险，能够避免投资者重大损失，同时也能够反向促进中国的股票市场往更成熟的未来发展。

1.3.2 技术路线

本文将以长短期记忆网络 LSTM 模型为基础，并添加了注意力机制来研究股价时间序列问题，通过选用不同的四维度指标为基础，并用 IC 值检验与相关系数检验筛选处理有效多维度数据。本文以沪深 300 成分股设置为股票池，将多维度中的所有处理过后的因子作为待选因子以选取输入因子，构建以 ATT-LSTM 为基础的多维度指标选股模型，并对股票价格进行预测，在不断调试参数以后选

用最优的结果，而后将预测模型获得收益率排名靠前的股票加入可供买卖的股票池中，通过均线距离指标 ΔD_{MA} 来对买入股票数量与买入股票仓位进行确定，并且通过量化回测系统来进行策略的整体实现。并且在评价本文的量化交易策略时，计划通过收益与风险情况来对策略进行评价，其中包括通最大回撤、收益率、年化收益率和夏普比率，并且为了使得模型具有更好的可行性，本文在加入了各种程度的止损线用以匹配不同风险偏好的投资者，具体的技术路线图如下所示：

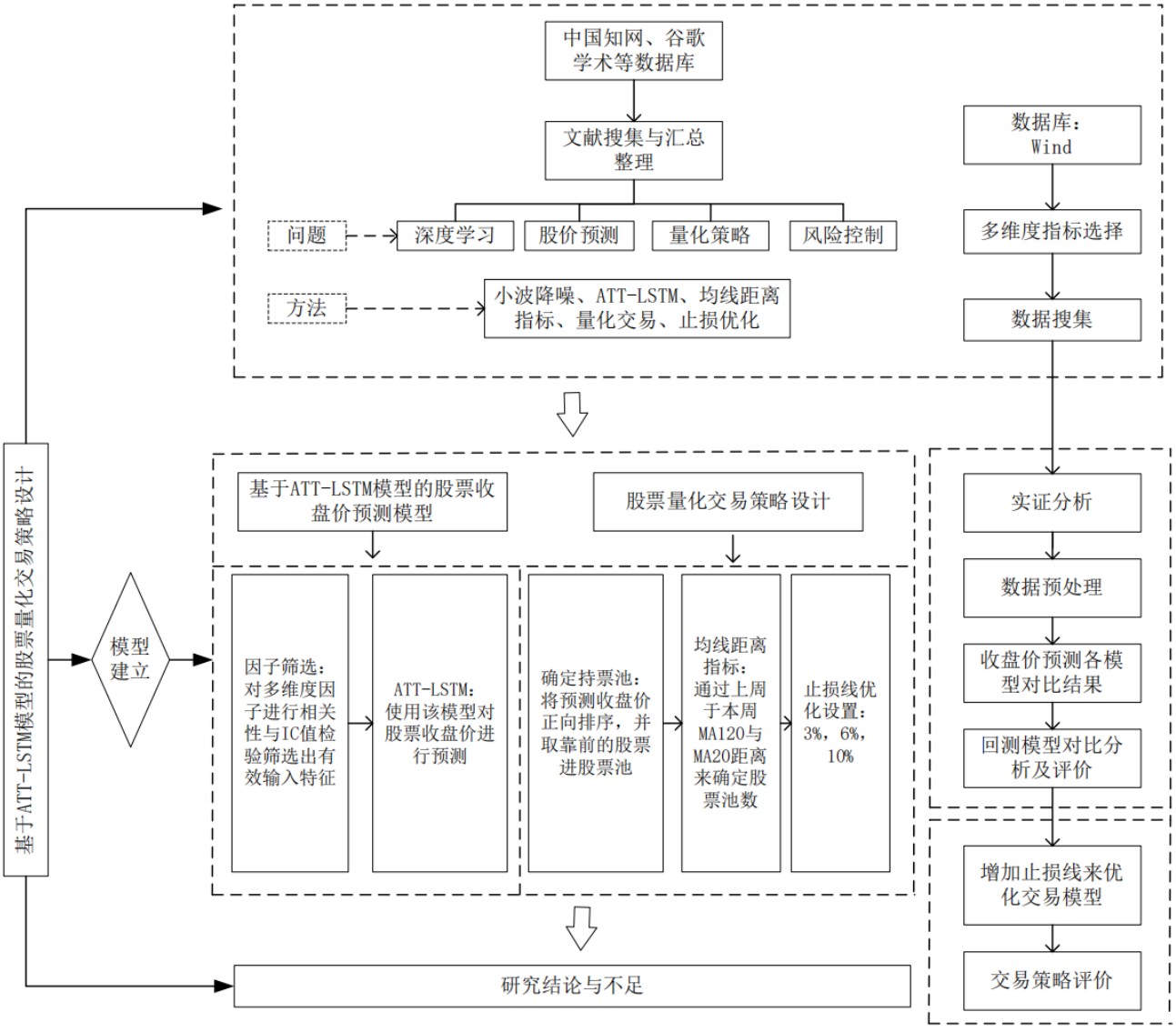

图 1 技术路线图

1.4 本文的主要贡献

大量文献表明，目前越来越多的深度学习模型应用在股价预测与量化交易领域，在分析预测股票市场的过程中，只要足够准确且及时便能使金融风险发生的可能性降低，对于普通民众来说进行证券市场投资就是一个更好的投资选择，这

部分的超额收益能够使百姓收益，同时也能够反向作用于中国的股票市场，使得中国的股票市场变得更成熟规范。

本文的主要贡献与创新点体现在下述三个方面：

1. 本文主要通过影响股价的多维度因子着手，从行情指标、关联市场因素、国内市场因素指标、技术指标等四个维度出发，初步筛选了具有一定相关性与可行性的因子 38 个，先后通过 IC 值检验与相关系数检验来确定最后的输入特征共 18 个，同时也覆盖了上述四个维度，能够从更广泛的角度来预测股价；

2. 在进行股价预测的过程中，利用筛选好的多维度因子放入模型当中预测，通过对比加入注意力机制前后的模型预测结果发现加入注意力机制后的 ATT-LSTM 模型能够在沪深 300 成分股当中具有更优的预测效果，不仅体现在有更高的拟合优度，同时也体现在有更小的误差，并且在多只股票预测当中能够有更强的泛化能力。

3. 同时本文在进行量化交易策略设计的过程当中，结合 ATT-LSTM 模型的预测结果来设定股票池，因为本策略是纯多头的策略，没有办法应对空头的情况，本文为了控制空头的风险，在构建量化交易的过程当中增加了均线距离指标 ΔD_{MA} 来控制换仓的股数，从而达到控制风险的效果。通过上述步骤的量化交易策略设计后进行回测，从而得到该模型量化策略的回测结果相比基准收益有更好的表现。另外，本文选取的股票标的为沪深 300 所有成分股，共涉及 29 个行业，在构建股票池时能对实盘投资有更好的借鉴意义。

第 2 章 相关理论回顾与文献综述

2.1 文献综述

2.1.1 机器学习在股价预测领域研究现状

机器学习的研究从上世纪八十年代开始兴起，极大地推动了 AI 的发展。神经网络作为机器学习研究的源头能够实现对复杂数据的高效处理和分析，其有非常多的应用场景，如图像和语音识别、医疗诊断、自动驾驶等等，同时包括了量化投资领域。

历年来有非常多的学者研究机器学习模型在股价预测当中的应用，Gen Cay 通过研究对比分析传统的移动平均统计学算法与神经网络算法预测股价市场走向结果的，发现利用神经网络模型有更高的准确性[1],也有学者不拘泥与对比传统方法与神经网络之间的准确性比较，从而转向研究不同神经网络之间的分析对比，如衍生出了卷积神经网络和循环神经网络，两者具有不同的特性，其中前者具有空间性分布数据特征，后者具有时间性分布数据的特征。

在神经网络的发展中，学者 Hochreiters 提出长短期记忆网络模型能够有效解决循环神经网络产生的梯度爆炸和梯度消失问题[2]。除此以外，Graves A 在研究中发现，LSTM 在预测股票价格的过程中能显著提高股票价格时间序列预测的精度，从而为 LSTM 模型应用在股票市场打下坚实的基础[3]。随后也有学者从改进神经网络的角度切入，如 Liu, Xiangwei[4]等学者采用的 BP 神经网络也能够很好的在股票市场的预测过程中提供较好的预测效果，从而推动了改进型的神经网络模型在股票市场量化投资领域的应用。对量化投资预测和机器学习的研究过程当中，DumoulinV 发现金融股票市场的时间序列数据是具有非线性特征的[5]。

除此以外，许多学者开始使用 XGBoost(eXtreme Gradient Boosting，极度梯度提升树)、GRU(Gate Recurrent Unit，循环神经网络)等算法来对股票价格进行预测。在对比研究过程中，shen 等学者发现 GRU 算法相比传统的神经网络算法应用于股票价格预测时能够有更好的预测效果[6]，同时黎镭等人利用 GRU 算法来预测多个标的股票收盘价时也有较好的表现，这说明 GRU 算法在股价预测领域有非常强大的泛化能力和强大的学习能力[7]。另外，相较于循环神经网络和长短期记忆模型，刘宁宁，张量采用主成分分析与 GRU 相结合的方式来对股票指数进行预测有更好的泛化效果与更高的预测精度[8]。为了解决利用 GRU 算法对股价预测产生较差效果问题，在融合自自注意力机制和 GRU 后，又引入了将为处理技术随机森林算法从而对股票因子进行降维筛选，从而提高模型预测精度。

尽管机器学习在股价预测领域取得了一些进展，但股市本身的不确定性和复

杂性，以及数据的噪声和偏差等问题仍然存在，因此预测的准确性也存在一定的局限性。

2.1.2 LSTM 模型股价预测领域研究现状

近年来众多学者都发现由于 LSTM 模型具有优良特性，并且在一定程度上解决了 RNN 的梯度爆炸和消失问题，LSTM 模型在股票市场的预测应用近年来渐渐增加，胡谦、Zhang Ru、李钦、季楚然、谢合亮、许丽、田雨等学者通过应用不同类型的因子于 LSTM 模型当中发现对股价预测的精度有显著提高[9-15]；黄建华、宋刚、方义秋、李晨阳、史建楠、姚远等学者通过对 LSTM 模型进行改进来对股价进行预测，其中近年兴起并有较好的预测结果的模型有 PSO-LSTM(基于自适应粒子群优化的长短期记忆股票价格预测模型)、CNN-LSTM(CNN 卷积层与 LSTM 集成模型)与 DMD-LSTM 模型(基于动态模态分解—长短期记忆神经网络)[16-21]；林杰、沈山山、邓瑶瑶、唐成、杨向前等利用 ATT-LSTM 模型对股价进行预测，其中增加 Attention 机制能够凸显重要模态对原始序列的影响力，能使预测精度和性能显著提升[22-26]。

另外，从国内各学者利用 LSTM 模型进行股票预测的同时选取的因子可以看出，单一因子或者少数因子的预测模型较多，诸如李钦仅从技术指标出发，选取 55 个具有预测里的技术指标因子来对 LSTM 模型进行构建与优化，最终构建沪深 300 指数预测模型能够获得一定超额收益[11]；季楚然利用 PCA(Principal Component Analysis,主成分分析)回归对特征进行筛选，并在加入技术指标和相关资产价格的特征选取基础上构建 LSTM 神经网络模型来预测上证 50 指数次日的收盘价，结果证明 LSTM 模型能够应用于股票市场的价格预测研究，并有良好的效果，这能够给予各类投资者关于如何将 AI(Artificial Intelligence)应用在量化投资领域一定的借鉴意义[12]；谢合亮构造了一种集合 Elastic-net 与 LSTM 的多因子量化投资模型，并提出目前国内外的研究现状主要集中在高频数据领域,而如何利用低频数据进行深度学习研究的文献较少，而且在因子选择方面从多个角度考虑，其中包括估值、杠杆、市值、财务、动量反转等角度，共选择 39 个相关股票因子[13]。许丽等考虑了新闻股票感情分析因素，构建了基于 LSTM 和新闻情感的预测模型，发现其 RSEM、MAE、MSE 比 XGBoost 模型均更低，模型预测精度更高[14]。田雨在融合投资者情绪的基础上利用 LSTM 模型对股价走势进行预测，发现融入投资者情绪的预测模型比为融入情绪特征的预测模型效果好[15]。

另外也有一些学者从改进 LSTM 模型的角度出发来构建预测精度更高的股票预测模型，诸如黄建华等在传统的 LSTM 模型上增加了改进粒子群算法来对

股票进行预测，并成功提高 PSO 算法的寻优性能，避免出现局部最优解，所提出的模型准确率大大提高，且具有普遍适用性[16];宋刚等提出一种基于自适应粒子群优化的长短期记忆股票价格预测模型，实验在沪市、深市、港股股票数据中预测结果有较高的预测出准确度且具有普遍适用性[17]；方义秋等在 LSTM 模型与 CNN(Convolutional Neural Networks，卷积神经网络)模型能够提取数据深层特征特点的基础上，还联合了两者的 RMSE 损失函数，新的联合模型在预测效果上具有良好的可行性和普适性的结论[18]；李晨阳在股票价格趋势预测中引入深度学习神经网络算法,将两种原理不同的神经网络架构 CNN 和 LSTM 相结合来对股价涨跌进行预测，基于 CNN-LSTM 模型构建的量化选股策略能够在不同情况下均取得超额收益[19]；史建楠等提出一种基于动态模态分解—长短期记忆神经网络（DMD-LSTM）的股票价格时间序列预测方法，其能够在特定的市场环境下实现更高的预测精度[20]；姚远等提出了一种融合了 HP 滤波（Hodrick-Prescott Filter）和 LSTM 神经网络模型的股指价格预测模型，实验结果表明,提出的 HP-LSTM 混合模型提高了股指价格预测精度与可解释性[21]。

邓瑶瑶采用了改进后的 WT-ILSTM-ATT 模型对股票收盘价进行预测，在结合趋势指标来构建量化投资策略[24]。唐成 Attention-LSTM 方法构建模型对上证 50 指数最高价涨跌趋势进行预测，结果显示所构建量化投资方法有较好的预测能力[25]。杨向前等提出了基于 VMD（变分模态分解）的 Attention-LSTM 模型,Attention 机制能够突显重要模态对原始序列的影响，因此建立的预测模型会更有效，具有更佳的预测精度和性能[26]。

2.1.3 量化交易近期研究现状

在量化交易策略的设计过程中主要分为两大类，一类是与选股模型结合的量化交易策略设计，一类是在股票池给定的情况下进行策略设计。前者在选股的过程当中可以与各种机器学习或传统模型进行结合进行预测，后者在策略设计角度可以通过增加各种不同类型的指标来进行。

不同学者在设计量化交易的策略上各有不同。Ticknor 等学者发现不论在神经网络模型还是传统的模型基础下，构建量化交易策略过程中可以设计出比单纯买入卖出效果更好的策略[27]。Enke 等学者在训练模型时选用的输入特征为相关的基本面指标，结果表明能较好的预测[28]。

叶伟睿[29]采用的是多因子的选股策略，设计过程中以 LSTM 模型为基础，并且利用特征工程来改进优化模型，评价结果表明该模型有较高的预测精度，量化策略有较好的投资收益，因此能够有效的应用于公募基金投资管理股票的进程中。刘天贶首先结合多因子选股模型和板块轮动思想构建选股策略，并且通过记

录交易明细等根据数据做出投资决策，按照固定规则进行买入卖出操作，最后通过设计智能牛市区间量化回测算法，在牛市当中有较好的收益率表现，提高交易策略的收益并降低风险[30]；钟正豪使用多机器学习模型融合的预测来进行选股，并且通过固定的投资策略 Topkdropout 模拟交易，结果证明在沪深 300 成分股上模型融合方法以及融合模型 Weighted_Ensemble 更稳定且有更好的超额收益表现[31]。

因此若期望使用选股与量化策略结合的方式来设计最佳策略，则需要在利用机器学习模型选股的基础上，构建合适的因子来辅助策略构建。

2.1.4 文献述评

本文从量化交易策略发展与基于不同因子的 LSTM 模型和改进型 LSTM 模型股价预测等方面，对国内外相关文献和研究成果进行整理和分析。随着时间的变化，各种新技术逐渐问世，为本文写作在思路和技术上提供了极大的帮助，但是现有的研究仍存在着进一步改进的空间：

第一，在因子维度选择上，传统的计量模型难以挖掘复杂的输入特征，大多学者仅对等单一维度的多因子指标进行研究，诸如单纯使用技术指标因子，或者单纯使用股价自身行情相关数据，却忽略了很多其他因素，综合考虑诸多因素的文章更是少之又少，本文在因子选择上综合考虑了股价行情因素、关联市场因素、国内市场因素以及丰富的技术指标类型来构建多维度因子筛选库，其中技术指标包含了摆动指标、波动指标、超买超卖指标、反趋向指标、量价指标、能量指标、强弱指标、趋向指标及压力支撑指标。

第二，在多维度数据处理上，PCA 主成分分析与 LASSO(Least absolute shrinkage and selection operator)降维都是比较主流的方式，但是鉴于本文股票标的数量较多，通过对单一标的得出的降维有效特征表达式会使得预测模型的泛化能力降低，所以需要最大程度保留输入特征的信息，仅通过输入特征与收盘价的 IC 值与相关系数来对特征进行筛选。

第三，在优化 LSTM 模型上，由于 LSTM 的良好性能使得众多学者从不同角度切入优化 LSTM 模型。在利用 Attention-LSTM 模型的众多研究当中，对数据的处理相对欠缺。但是不可否认的时加入注意力机制后的 LSTM 模型在股票市场预测的精度有显著提高。

第四，在量化交易策略的设计上，纯多头的趋势性择时交易策略面临一定的风险，即使有一定的机器学习模型作为选股的支撑，但还需要在构建策略过程中设定指标，本文通过增加均线距离指标 ΔD_{MA} 来控制纯多头交易的风险，同时也能保证一定的超额收益。

因此，为补充股票量化交易方面综合考虑多维度因子相关研究，本文利用相关性检验与 IC 值检验对多维度因子进行筛选处理使用，并将处理好的时间序列数据添加进 ATT-LSTM 模型当中获得较全面且较准确的预测模型，且基于均线距离指标 ΔD_{MA} 来达到更高的夏普比例。

2.2 相关理论

2.2.1 量化投资理论

在过去的几十年中，随着计算机技术的不断进步和数据的大规模应用，量化投资已经成为了全球投资行业的一个重要分支，其发展也越来越受到关注。美国是全球量化投资市场的领头羊，拥有众多的量化基金和公司。在美国，量化投资已经占据了投资市场的很大一部分，其中以对冲基金和私募股权为主。

近几年中国的量化投资也在发展迅速，尤其是在私募基金市场上。根据 Wind 数据，截止 2022 年底，中国私募基金中量化对冲类基金已经超过 2000 只。但是在中国能被称为真正的量化基金的标的还比较稀缺，但这并不妨碍我们仍然在这条路上持续探索。量化策略在很多方面都有十分明晰的优势，如可以更系统地对现有数据的不同维度进行全方位的评价，而非感性地进行处理，能够在基于对过往数据的分析中将结果应用于未来数据，从而有更高概率获取更多的收益，同时由于量化交易的自动纪律性，也能够在更短的时间内抢占投资机会。

量化投资目前已有多种形式，其中有应用最广泛的多因子选股策略，市场中性的 Alpha 策略、套利策略、市场波动较强时的 CAT 策略(Commodity Trading Advisor Strategy)以及多策略融合的组合策略。本文在预测股票价格时，选择利用机器学习的选股策略并且在交易策略设计的过程当中使用均线距离指标 ΔD_{MA} 来降低量化交易过程中的风险，并且带来更多的量化交易收益。

2.2.2 资产配置理论

在构建量化交易策略过程中，除了需要考虑配置的股票种类以及进场出场时机，还需要考虑一个非常重要的因素，即买入数量。常用的资产配置模型有两种，分别是等权重模型与最小方差模型：

(1) 等权重模型

等权重资产配置模型的特点是在构建投资组合的过程当中，不论买入哪一只股票都会给予其相同的权重，该模型相较于其他资产配置模型更简单,此时所有

股票标的的权重相加之和为 1。

(2) 最小方差模型

最小方差模型的特征即为可以综合考虑在获取一定收益的基础上还能有效的控制风险，是股票投资当中常被选择的资产配置模型。计算某投资组合 A 的收益情况：

$$r_A = \varpi_1 r_1 + \cdots + \varpi_n r_n = \sum_{i=1}^{n} \varpi_i r_i \tag{2-1}$$

其中 n 为该投资组合当中的股票数量，r_A 为投资组合的综合收益率，ϖ_i 为投资组合当中不同股票的买入权重，所有股票的权重相加为 1，另外 r_i 为单只股票的收益率。该投资组合的期望收益与方差计算如下：

$$E(r_A) = \varpi_1 E(r_1) + \cdots \varpi_n E(r_n) \tag{2-2}$$

$$\sigma_A^2 = \sum\sum \varpi_i \varpi_j \operatorname{cov}(r_i, r_j) \tag{2-3}$$

其中 $E(r_A)$ 为投资组合期望的收益，σ_A^2 反映了投资组合 A 的波动情况，出于控制风险的角度，最小化方差模型的公式如下：

$$\min \sigma_A^2 = \sum\sum \varpi_i \varpi_j \operatorname{cov}(r_i, r_j) \tag{2-4}$$

从实际运用角度出发，若采用最小化方差法进行调仓，则会大大增加策略进行过程中的计算量，影响策略的效率。本文计划构建的量化交易策略每 3 天进行一次换仓，属于短频换仓，因此选用等权重模型来对资金进行分配能够提高整体策略的运行效率。

2.2.3 行为金融学理论

(1) 动量效应

动量效应又称惯性效应，指股票的收益率具有延续原来运动方向的趋势，即过去一段时间收益率较高的股票在未来获得的收益率仍会高于过去收益率较低的股票。

(2) 反转效应

反转效应是指在较长的时间内，表现差的股票在其后的一段时间内有强烈的可能性进行相当大的逆转从而恢复到正常水平。相同地，若是某股票长期出于良好的涨幅，则倾向于在其后的时间内出现差的表现。

本文在量化交易设计的过程中创建了均线距离指标 ΔD_{MA}，考虑到在量化交易策略的交易频次相对较高且换仓时间较短的情况下，可将当时的情况认为是有强动量效应的状态，因此不论好坏的交易环境都会延续原来运动方向的趋势。

第 3 章 基于 ATT-LSTM 模型的股票问题分析及交易策略构思

3.1 股票量化交易问题的提出

自我国股票市场建立以来，量化投资交易策略也在逐渐盛行。许多私募股权公司都在积极使用各种量化策略来构建能够较好规避市场风险的方案，同时能最大程度的获取收益。但同时为了避免各种投机主义，以及防止我国股票市场在不成熟的情况下出现各种系统性风险或者黑天鹅事件导致的非理性下跌，传统的股票投资方式需要逐渐被优化。量化交易策略能够很好的摆脱人工手动买卖股票时候的非理性操作，能够及时的判断及完成换仓操作。

本文通过机器学习模型对沪深 300 成分股的收盘价进行预测，再利用均线距离指标 ΔD_{MA} 通过对大盘是否乐观做出初步的判断来筛选排名靠前的需要进行换仓的股票池，并且通过能够判断当下投资情况的均线距离指标来控制买入股票数量。通过使用预测精度较好的模型训练、预测及回测，希望本文策略的情况能够超过基准收益。所以在该过程中，本文需要着重研究的三方面为股票收盘价预测的精度、判断市场行情的准确性、量化策略设计的合理性。

3.1.1 股票多维度特征指标的选择

本文期望通过选择合适的多维度特征来对股票价格进行预测，因为除了确定相对应的深度学习模型，选择对股价预测模型有影响的输入特征亦十分重要。通过查阅大量文献可知，李钦[11]选用 55 个技术指标来对股价进行预测，季楚然[13]不仅使用了技术指标还增加了相关资产价格的特征来使用 LSTM 模型进行预测，还有学者许丽[14]在预测过程中增加了新闻股票感情分析因素，邱冬阳[32]通过选择 CSIF300(沪深 300 期货指数)的自身行情因素、内在因素、宏观因素、关联市场因素、偶发因素等共 5 个维度的指标进行预测，有良好的效果。

近年来机器学习模型已经逐渐取代传统模型应用于股价预测，多因子的预测效果显著提高。本文参考上述文献的影响维度，从股价行情因素指标、关联市场因素、国内市场因素、技术指标四个维度来构建股价预测模型当中的因子库。通过选取有效的多维度特征，后续预测模型的精度才会有更好的表现，同时更具备真实世界借鉴意义。

3.1.2 股票收盘价预测方法的选择

近三十年来，中国股票市场发展迅速发展。股票作为公司股权在二级市场流通的一种形式，本质上属于可买卖的资产商品。其次公司股票从发行开始，到可

流通阶段以及定价过程，全程股价都会由于各种因素而发生改变或者产生剧烈波动。股票价格一方面会因为公司本身运营的情况好坏而发生改变，另一方面可能会因为公司所在行业在宏观市场情况下的发展程度而改变。因此，股价本质上是一种波动性很强且非平稳的序列，如果采用单纯的时间序列模型来对股票价格进行预测效果会很受限制。

最初学者会使用相关时间序列的计量模型来对股票价格进行预测，但是传统的时间序列诸如 ARMA(Autoregressive moving average model，自回归滑动平均模型)、ARIMA(Autoregressive Integrated Moving Average model，整合移动平均自回归模型)等模型都会存在一定的缺陷，如存在滞后性且预测精度低等情况。后来学者们发现在除了使用传统的各种计量模型以外采用人工智能模型来预测效果上会有更好的表现。

在股价收盘价预测方面，林杰等提出了一种基于注意力机制的 LSTM 股价趋势预测模型，选用选取 42 只中国上证 50 从 2009 年到 2017 年的股票数据为实验对象，结果表明与传统的机器学习模型相比，添加了注意力机制的 LSTM 模型具有更好的预测能力。沈山山等提出了基于注意力机制的 CNN-LSTM 短期股票价格预测模型，先使用 CNN 来对数据序列进行卷积操作,以提取其特征分量。然后,利用长短期记忆网络对所抽取出的特征分量做序列预测。目前看来利用 LSTM 模型加入注意力机制的模型能够有效的提升股票收盘价预测的精度，本文在输入特征方面加入多维度数据从而达到改进的目的。

3.1.3 股票交易策略投资组合设计

目前来看，交易策略主流的有三种类型，第一种价值投资，这是一种偏向于长线的股票投资，通过对该股票对应的公司业务及财务进行分析，判断出该公司在未来能有较好的业务增长，不论是由于该行业在未来有较好的发展还是该公司个体有独特的优势，综合来看判断出随着公司的发展，未来一段时间后公司的市值会有较大的增幅。第二种是短线操作根据股价波动赚取差价，这种做法对个人投资者有较大的风险，但是对于机构投资者来说由于其可以采用融资融券的方式，所以使用较多。第三种是市场羊群效应，跟随市场短期板块热点进行买卖股票，完全不考虑任何其他方面的因素。

本文通过预测精度及效果较好的模型来预测股价，并通过预测的结果来确定需要进行换仓建仓的股票，但是纯多头的交易策略面对的风险很多元化，所以在构建策略过程中添加均线距离指标 ΔD_{MA} 来平滑风险，从而达到收益最大化与风险控制的效果，得到一定的超额收益。

3.2 基于 ATT-LSTM 模型的交易策略理论框架

在通过总结各大学者的文献内容后，本文计划采用 ATT-LSTM 模型来对股票价格进行预测，并且本文创新性地提出使用均线距离指标 ΔD_{MA} 来进行股票交易

策略设计。由于股票价格序列的非平稳性，所以需要先对股票价格数据进行一定的降噪处理，利用相关性检验对相关指标进行筛选处理，再基于 ATT-LSTM 与 LSTM 模型对股价进行预测，选取最适合的参数与模型进行股价预测，作为构建量化策略的基础。

3.2.1 股价序列降噪

任何宏观因素或者微观因素，都可能在任意时间造成股票价格的变化，根据前任的研究表明，股票价格并非是一个平稳的时间序列，如果在预测的过程当中采用传统的时间序列预测方式来进行，最终呈现的结果一定差强人意。因此在预测过程中首先要考虑的就是如何生成高质量的数据来对模型进行训练，即需要采用一定的模型来对时间序列进行特征提取。在选用不同方式来对股价时间序列进行处理时需要考虑的首要问题即为如何消除趋势性因素，常用的方法为差分法，但是其在使用的过程当中会存在很多的噪声，正因噪声的存在从而导致直接使用差分法处理后的股价时间序列并不能训练出很好的预测精度，因为需要在前期先对股价序列进行噪声处理。

本文计划采用小波降噪的方式来对股票价格时间序列进行降噪。小波降噪在信号学领域属于信号滤波问题，小波降噪显著优于传统低通滤波器的特点是它能成功保留信号特征。由此可见，小波降噪是特征提取与低通滤波器的综合产物,其流程图如下所示：

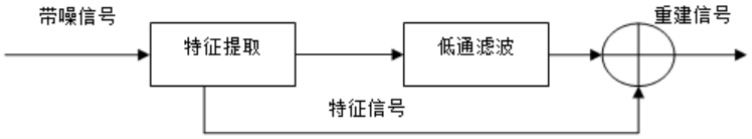

图 2 小波降噪流程图

彭燕首先对股票历史特征时间序列数据中的缺失值进行插值处理，然后将股票收盘价数据中的噪声数据使用小波降噪的方式降噪，结果表明利用小波降噪后的数据的预测结果更优。股票收盘价历史事件序列数据中的趋势性特征是重要的信息，噪声属于无效信息，所以在处理股价时间序列时使用小波降噪来进行处理，去除数据中的无效信息从而提高预测精度[33]。

3.2.2 ATT-LSTM 模型

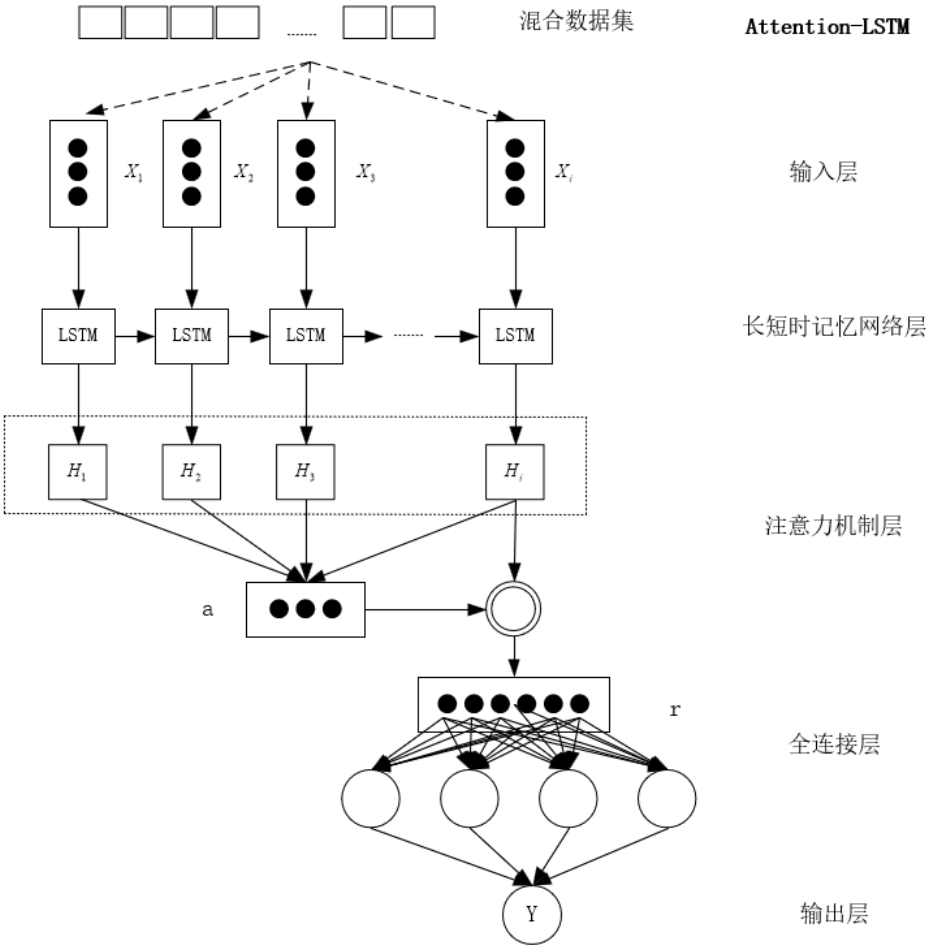

图 3 ATT-LSTM 神经网络架构

1. ATT-LSTM 模型

上图展示了 LSTM 模型在添加了注意力机制以后的预测动线，其中包括最初数据收集且整理的部分，以及数据输入层，将数据输入进模型后进入长短时记忆网络 LSTM 层，在经过长短时记忆网络 LSTM 层后还需要经过注意力机制层，最后再通过全连接层到达输出层。

加入注意力机制的 LSTM 模型应用，本文的注意力机制采用概率分配的方法，能够以分配概率的方法代替原先的随即权值分配。注意力机制能够使得模型预测的结果有权重影响，使用组合模型预测股价序列，并与之前未加入注意力机制的 LSTM 模型预测结果进行比较，对模型进行修正。

LSTM 模型改进了循环神经网络 RNN，解决了其再训练过程中梯度爆炸或者梯度消失等问题。由于所有的循环神经网络都存在重复模块的形式，这种重复的模块结构非常简单，例如 tanh 层或 sigmoid 层。LSTM 拥有三个门，分别是遗忘门(forget gate)，输入门(input gate)和输出门(output gate)，遗忘门负责选择遗忘过去的无效信息，输入门负责确定有效信息，输出门决定输出信息。

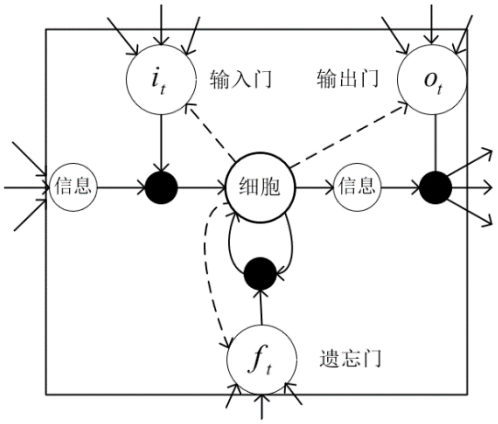

图 4 LSTM 单元结构示意图

1) 遗忘门：读取当前时刻输入的 x_t 和上一时刻的记忆单元状态信息 h_{t-1}，随后通过 sigmoid 函数来输出值。

$$f_t = \sigma(W_f \bullet [h_{t-1}, x_t] + b_f) \tag{3-1}$$

2) 获得整合后的新信息：输入门可以控制当前时刻的数据影响记忆单元状态值，将新的有效信息进行存放，首先计算输入门的值 i_t，随后计算当前时刻 t 的候选准备记忆单元信息 $\tilde{C}_t$，最后将原细胞状态与新的信息进行合并：

$$i_t = \sigma(W_i \bullet [h_{t-1}, x_t] + b_i) \tag{3-2}$$

$$\tilde{C}_t = \tanh(W_c \bullet [h_{t-1}, x_t] + b_c) \tag{3-3}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{3-4}$$

3) 输出信息：确定内容经过输出门后，于记忆单元状态信息经过 tanh 变换，得到当前时刻的记忆单元输出信息：

$$o_t = \sigma(W_o \bullet [h_{t-1}, x_t] + b_o) \tag{3-5}$$

$$h_t = o_t \times \tanh(C_t) \tag{3-6}$$

2. 注意力机制

近年来注意力机制逐渐被广泛应用，它能让模型在训练的过程当中把注意力集中在部分重要信息中，并且在此过程中不会存在其他诸如影响信息等副作用，同时又因为它具有独立性，因此能够应用于任何模型的训练中。

在注意力机制的运行过程中，能够自动的地对相对重要的信息匹配更大的权重，从而提升模型的训练精度，不仅可以应用在翻译、图像等领域，还可以运用

在预测时间序列数据上，因此为机器学习模型应用于股票价格预测提供了更好的选择。在预测的过程当中，由于会输入大量的因子，不同的因子会有不同程度的影响，因此注意力机制可以在预测过程当中起到非常大的作用，通过训练来决策各个因子的影响程度，从而调节它们的权重，使得最终结果有更高的准确性。

其中 $H \in R^{d \times N}$ 表示由 LSTM 模型输出向量 $[h_1, h_2, \cdots h_N]$ 所组成的矩阵，N 表示长度，本文中用预测周期 T 的数据预测第 T+3 天的涨跌趋势，最终产生注意力权重向量 α 与 r：

$$M = \tanh(W_h\ H) \tag{3-7}$$

$$\alpha = soft\max(\omega^T\ M) \tag{3-8}$$

$$r = H\alpha^T \tag{3-9}$$

在上式当中，$W \in R^{d \times N}$，$\alpha \in R^N$，$r \in R^d$。$W_h \in R^{d \times d}$ 和 $\omega^T \in R^d$ 市后续模型需要训练的参数矩阵。最终注意力机制输出的隐藏层数据 H 和 r 的拼接。最终输出的向量表示如下：

$$h^* = \tanh(W_p r + W_x h_N) \tag{3-10}$$

其中 $h^* \in R^d$，W_p 和 W_x 是后续模型需要训练的参数矩阵；输出变量 h^* 最后经过一个全连接层和 Softmax 分类器实现股票涨跌的预测。

3.3 均线距离指标 ΔD_{MA}

通过参考目前二级市场较多的研究报告，在每周构建买卖策略的过程当中通常会先对本周大盘的走高或走低趋势做一个判断。由于本文的交易策略是短期换仓，因此也有必要对当下的交易环境做一个初步判断来优化当前买入卖出情况，于是本文创新地构建了均线距离指标 ΔD_{MA} 来控制量化交易策略中的买入仓位。具体操作即为计算本周与上周的长短期移动平均线的距离来判断本周市场整体环境的大致走势，将此前已经预测出排名靠前的股票进行筛选从而达到优化的目的。

1. 移动平均线 MA(Moving Average)

移动平均线是通过计算连续 n 天的收盘价的算术平均值得出。具体计算方法是：

$$MA(n) = \frac{C_1 + C_2 + C_3 + \cdots C_n}{n} \tag{3-11}$$

其中 C 为前 n 日的收盘价，n 为移动平均周期数。

2. 计算本周与上周的长短期均线的距离

通过计算本周与上周上证综指长期均线 MA120 与短期均线 MA20 之间的距离，来反映当下市场的乐观与悲观情况。将上周记为 LD_{MA}，本周记为 D_{MA}。现实

意义，短期均线 MA20 在长期均线 MA120 下方，距离为负值，反之若在上方距离即为正值。

具体计算方法如下所示：

$$LD_{MA} = last\_MA(20) - last\_MA(120) \quad (3\text{-}12)$$

$$D_{MA} = MA(20) - MA(120) \quad (3\text{-}13)$$

$$\Delta D_{MA} = |LD_{MA}| - |D_{MA}| \quad (3\text{-}14)$$

以下通过区分 8 种情况来分析市场的投资环境以及短期的趋势变化，从而对量化交易过程当中的仓位设定与买入股票数量有一定的优化和借鉴作用。当短期均线在长期均线之下时，可能是在底部预备崛起也可能是目前正处在比较低落的位置，但是通过比较上期与本期两均线之间的距离即可知短期本次调仓期间市场的情况，具体情况如下表所示：

表 1 均线距离指标 ΔD_{MA} 大势参考表

	ΔD_{MA} <0	ΔD_{MA} >0
LD_{MA} & D_{MA} <0	短期看空	短期虽空但有回暖之势
LD_{MA} & D_{MA} >0	短期看多，较强	短期看多，较弱
LD_{MA} >0& D_{MA} <0	短期大势看空	短期大势看空
LD_{MA} <0& D_{MA} >0	短期大势看多	短期大势看多

1) LD_{MA} & D_{MA} <0 时，说明不论是上期还是本期 MA20 均线都在 MA120 均线下方，短期来看空头较明显，但若本周比上周的距离小，说明目前行情正在回暖；

2) LD_{MA} & D_{MA} >0 时，说明不论是上期还是本期 MA20 均线都在 MA120 均线上方，短期来看多头较明显，整体投资行情较好，但若本期的 D_{MA} 比上期 LD_{MA} 的距离更大，根据动量效应说明多头态势还会持续，反之则再整体投资环境较好的情况下有些许空头之势；

3) LD_{MA} >0 & D_{MA} <0 时，在上期多头行情较好的情况下本期短期均线 MA20 直接低于长期均线 MA120，说明本期的空头态势较强；

4) LD_{MA} <0 & D_{MA} >0 时，在上期空头行情较明显的情况下本期短期均线 MA20 高于长期均线 MA120，说明本期多头态势较强。

3.4 股票交易策略评价方法

本文使用的股价预测模型是 ATT-LSTM 模型，在前文的梳理过程中提及股价是非平稳的时间序列，因此需要对股票价格进行降噪，降噪的结果需要使用系列指标来进行评价。与此同时，更重要的是 ATT-LSTM 模型对收盘价的预测结果与与本文的量化交易策略的结果评价，对模型预测精度的评价从 MSE、RMSE、MAE 与 R^2 四个维度来评价。并且一般而言对策略的评价会从收益于风险的角度

出发，本文也不例外，综合使用年化收益率、波动率、最大回撤和夏普比例等四个指标来对本文的量化策略进行评价。

（1） 降噪评价

时间序列当中的噪声会影响预测的结果，本文使用小波降噪来对时间序列数据的噪声进行处理，通过 SNR 信噪比来对降噪的结果进行评价。一般来说，信噪比越大，说明信号中的噪声越小，数据质量越高。SNR 的计算公式如下:

$$SNR = 10\log[\frac{\sum_{i=1}^{N} X_i^2}{\sum_{i=1}^{N}(X_i - \hat{X}_i)^2}] \tag{3-15}$$

（2） 预测精度评价

本文通过四个评价指标均方误差 MSE、平均绝对误差 MAE、均方根误差 RMSE、拟合优度 R^2 来评价预测模型的精度，其中 R^2 越接近 1，MSE、RMSE、MAE 越小说明误差越小，模型的拟合效果越好，计算公式如下：

$$MSE = \frac{1}{N}(\sum_{i=1}^{N}(X_i - \hat{X}_i)^2) \tag{3-16}$$

$$RMSE = \sqrt{\frac{1}{N}(\sum_{i=1}^{N}(X_i - \hat{X}_i)^2)} \tag{3-17}$$

$$MAE = \frac{1}{N} \times \sum_{i=1}^{N} | X_i - \hat{X} | \tag{3-18}$$

$$R^2 = \sum_{i=1}^{N}(\hat{X}_i - \bar{X}_i)^2 / \sum_{i=1}^{N}(X_i - \bar{X}_i)^2 \tag{3-19}$$

其中 X_i 是实际股票收盘价时间序列数据， $\bar{X}_i$ 为实际收盘价数据的均值， $\hat{X}_i$ 为通过预测模型得出的股票收盘价序列数据。

（3） 量化交易策略评价指标

1) 年化收益率

年化收益率能够直观地反映出交易策略在一年交易周期中的收益率情况，能够清晰地和其他金融产品的收益率进行比较，年化收益率越高越好，计算公式如下：

$$\text{年化收益率} = \frac{\text{投资期间收益率}}{\text{投资天数}} \times 365 \times 100\% \tag{3-20}$$

2) 波动率

波动率指的是该交易策略在交易的过程当中，账户收益率的风险情况，理论上是波动率越小对于投资者而言投资风险越小，计算公式如下：

$$\sigma = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}(r_i - \bar{r})^2} \tag{3-21}$$

其中 r_i 为收益率时间序列， $\bar{r}$ 为收益率均值。

3) 最大回撤率

在评价交易策略的过程当中，除了利用账户的最终收益率以及波动率情况来衡量，还应考虑的是该策略在行进过程当中存在的最糟糕的风险位置。最大回撤率是在策略运行周期中，策略整体走到最低点时的收益率回撤幅度的最大值。最大回撤是一个重要的风险指标，在很多交易策略及金融产品当中是个非常重要的评价指标。

4) 夏普比例

夏普比率能够同时考虑交易策略的风险与收益，是一个具有综合性的实用指标。在调整风险后的收益率情况能够更真实的反映策略设计的好坏，其计算公式如下：

$$SharpRatio = \frac{E(r_A) - r_f}{\sigma_A} \tag{3-22}$$

其中 $E(r_A)$ 为投资组合的预期收益率，投资组合的标准差为 σ_A，股票市场中的无风险利率为 r_f。

第 4 章 基于多维度因子库的 ATT-LSTM 股价预测模型

在前三章内容中已经介绍了本文所采用的股价预测模型 ATT-LSTM 与利用均线技术指标的股票交易策略设计构想，包括多维度因子的选取及降维处理、利用小波降噪对非平稳的股票收盘价数据进行处理、使用模型 ATT-LSTM 模型预测股价、量化交易中均线距离指标 ΔD_{MA} 控仓的原因。接下来本文将在本章节介绍如何选取输入特征并对数据做预处理，同时使用 LSTM 与 ATT-LSTM 模型对股票价格进行预测，通过横向比较的效果来确定 ATT-LSTM 模型在多股票预测过程中有较好的有效性。

4.1 数据选取与处理

4.1.1 股票样本数据说明

近年来，运用深度学习的各种模型来对时间序列数据进行预测已有较多尝试，本文在选取数据时考虑了数据的质量，为了能够提取有效的数据特征同时避免预测误差较大，通过 Wind 数据库，选取 2017 年 1 月 1 日到 2021 年 10 月 31 日股票日品数据为研究对象，具体的标的选择的是具有代表性的沪深 300 成分股。本文通过选用沪深 300 成分股，可以构建一个相对完整的股票池，有下图可知本文的研究标的共涉及 29 个行业：

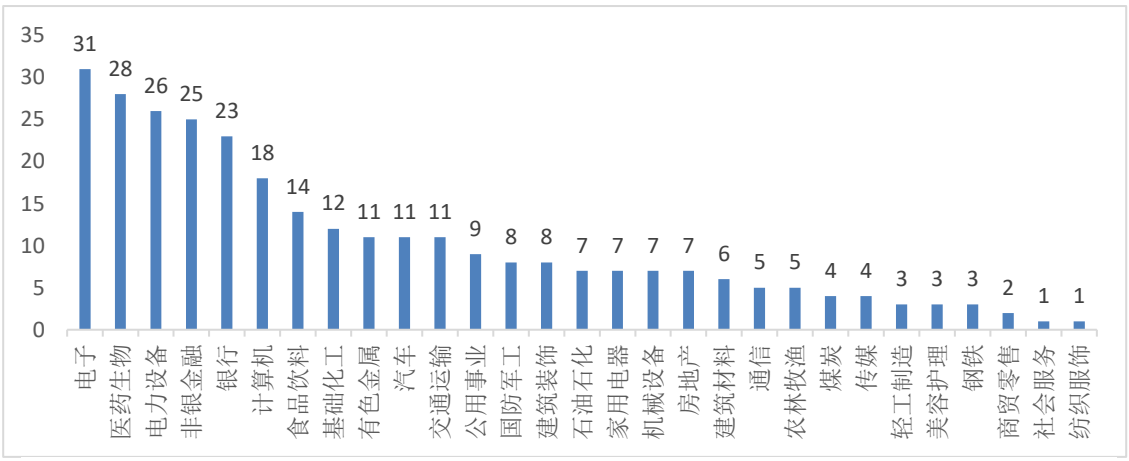

图 5 本文研究对象：沪深 300 成分股不同行业数量

本文选用的沪深 300 成分股的时间为 2022 年 6 月，沪深 300 成份股选股标准是根据股票所属行业，业绩，成长性等方面来决定的，沪深 300 成分股的投资优势在于业绩优于整体。根据申万一级行业分类来区分沪深 300 成分股所在的行业，具体各行业股票占比如下所示：

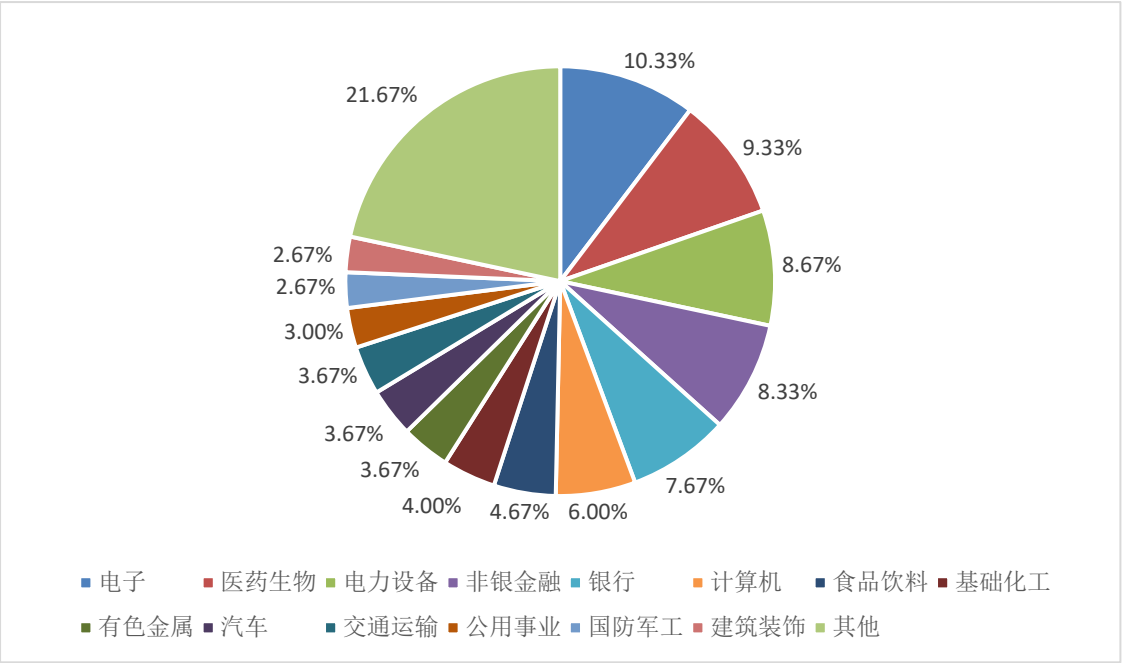

图 6 沪深 300 成分股所在行业占比

通过对股票的行业占比可以发现，在本文选取的 300 只股票当中，电子类，医药生物，电力设备，非银金融，银行类，计算机类的股票占比较多，占比分别为 10.33%，9.33%，8.67%，8.33%，7.67%，6.00%。食品饮料、基础化工、有色金属、汽车、交通运输次之，分别为 4.67%，4.00%，3.67%，3.67%，3.67%，而公用事业、国防军工、建筑装饰更少，分别为 3.00%，2.67%，2.67%。

4.1.2 指标数据说明

在股票市场使用深度学习模型进行价格预测从而量化投资交易，选用相关的影响因子非常重要，本文选取的因子从四个维度出发，具有实际且合理的可解释性意义。选用有效的因子在深度学习训练中起到至关重要的效果，从而才能应用在量化交易当中获得超额收益。

本文参考了邱冬阳学者研究 LSTM 模型在沪深 300 股指期货价格预测过程中采用的输入特征[32]，在构建初始多维度因子过程中综合考虑了多个维度，分别是行情因素指标、关联市场因素、国内市场因素、技术指标。

其中，行情因素指标是众多学者在利用机器学习模型的过程中率先使用的常规指标，包括开盘价、收盘价、最低价、成交量、成交额、换手率、振幅、前一天收盘价；由于不同国家之间的股票市场以及不同金融市场之间会存在一定的影响关系，本文还综合考虑了关联市场因素对中国股市的影响，其中包括了东京日经 225 指数、富时中国 50 指数、纳斯达克 100 指数、标普 500 指数、道琼斯工业平均指数、美元兑人民币汇率；本文的研究对象为沪深 300 成分股，很大程度上会受到二级市场大盘的影响，因此本文还添加了国内市场因素，即上证综指与

深证成指作为输入特征；在研究股价预测相关问题时，技术指标也是各大学者优先选用的输入特征，如李钦选用 55 个技术指标来对股价进行预测，本文选用了 23 个技术指标作为待选因子，其中涵盖了摆动指标、波动指标、超买超卖指标、反趋向指标、量价指标、能量指标、强弱指标、趋向指标、压力支撑指标。四个维度共计 38 个因子，汇总如下表所示：

表 2 多维度因子

指标维度		指标名称	指标代号
行情因素指标		开盘价	Open
		最高价	High
		最低价	Low
		成交量	Volume
		成交额	AMT
		换手率	Turn
		振幅	Swing
		前一天收盘价	Close
关联市场因素		东京日经 225 指数	N225
		富时中国 50 指数	CHA50CFD
		纳斯达克 100 指数	NDX100
		标普 500	SPX
		道琼斯工业平均指数	DJIA
		美元兑人民币汇率	HL
国内市场因素		上证综指	SZZZ
		深证成指	SZCZ
技术指标	摆动指标	SI 摆动指标	SI
	波动指标	STD 标准差	STD
	超买超卖指标	ADTM 动态买卖气指标	ADTM
	反趋向指标	KDJ 随机指标	KDJ
		CCI 顺势指标	CCI
		RSI-6	RSI6
		RSI-12	RSI12
		RSI-24	RSI24
		BIAS 乖离率	BIAS
		RPS 相对强度指标	RPS
	量价指标	SOBV 能量潮	SOBV
		PVT 量价趋势指标	PVT
	能量指标	PSY 心理指标	PSY
	强弱指标	阶段强势指标	Tech-20

表 2 多维度因子(续)

指标维度		指标名称	指标代号
技术指标	趋向指标	MACD	MACD
		MA5	MA5
		MA10	MA10
		MA20	MA20
		MA120	MA120
		BBI 多空指数	BBI
		DMA 平均线差	DMA
	压力支撑指标	BOLL	Boll
		CDP 逆势操作	CDP

4.1.3 输入特征选择与评价

在股票市场使用深度学习模型进行价格预测从而量化投资交易，选用相关的影响因子非常重要，本文选取的因子从四个维度出发，具有合理的经济学解释。选用有效的因子在深度学习训练中起到至关重要的效果，只有选择的输入特征足够好，才能应用在量化交易当中获得超额收益。

上文已经完成了本文多维度因子库的构建，以及完成了股票历史数据与指标的选取、对数据进行了简单的预处理以及选定预测模型的特征输入。

但是由于因子数量众多，还需要将已经本文早前确定的多维度因子进行筛选，本文计划通过不同因子与次日收益率从的 IC 值来做第一轮的因子筛选，其次将剩下的因子与收盘价进行相关系数检验，从而留下足够优质的输入特征用以模型训练。

1) IC 值检验

$$IC^T = corr(Y^T, R^{T+1}) \tag{4-1}$$

其中 Y^T 表示第 T 期的不同特征因子的数据序列，R^{T+1} 表示第 T+1 期的股价收益率序列，二者的相关系数即为第 T 期的因子 IC 值，用 IC^T 来表示。本文各个因子的 IC 值计算如下图所示:

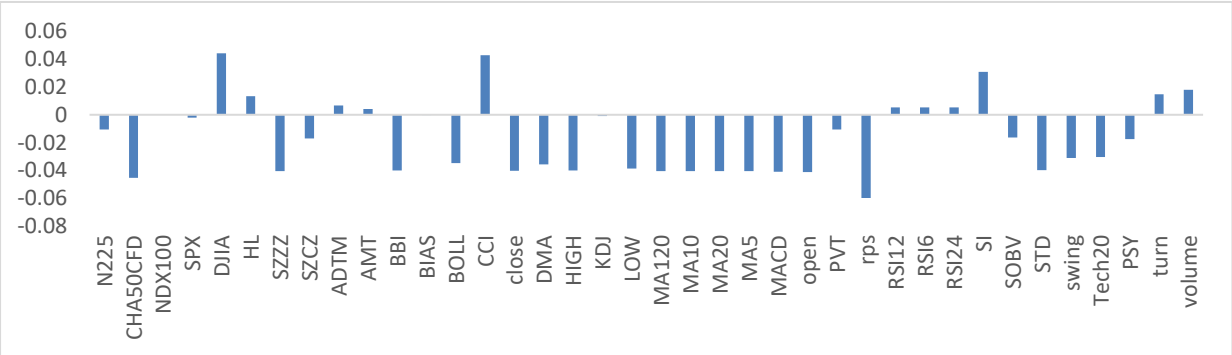

图 7 因子 IC 值计算结果

一般 IC 值均值的绝对值大于 0.02 则代表其为有效因子[35]，本文选用 0.03 作为筛选标准，因此根据对 IC 值的计算并筛选得到如下表结果，左侧为删除的指标，右侧为留下的输入特征：

表 3 多维度因子的 IC 值展示

|IC|	指标	IC 值	|IC|	指标	IC 值
<3% (删除)	N225	-0.01068	>3% (保留)	CHA50CFD	-0.04541
	NDX100	0.000397		DJIA	0.044139
	SPX	-0.00201		SZZZ	-0.04042
	HL	0.013281		BBI	-0.0401
	SZCZ	-0.01709		BOLL	-0.03475
	ADTM	0.006662		CCI	0.042776
	AMT	0.004241		close	-0.04029
	BIAS	-0.00013		DMA	-0.03567
	KDJ	-0.00101		HIGH	-0.04006
	PVT	-0.01059		LOW	-0.03861
	RSI12	0.005292		MA120	-0.04049
	RSI6	0.005292		MA10	-0.04049
	RSI24	0.005292		MA20	-0.04049
	SOBV	-0.01631		MA5	-0.04049
	PSY	-0.01958		MACD	-0.04105
	turn	0.014669		open	-0.04128
	volume	0.020988		rps	-0.05976
				SI	0.030943
				STD	-0.03984
				swing	-0.03108
				Tech20	-0.03049

2) 相关系数

在利用 IC 值对多维度因子进行筛选剔除之后，本文计划对剩下的因子利用相关系数进行筛选，通过计算各个指标与收盘价的相关系数来剔除相关系数过低的指标，从而得到更适配 ATT-LSTM 模型的输入特征，具体结果如下表：

表 4 相关系数计算汇总表

指标	相关系数
CHA50CFD	0.405704
DJIA	-0.36085
SZZZ	0.602291
BBI	0.993494
BOLL	0.977372

表 4 相关系数计算汇总表(续)

指标	相关系数
CCI	0.064278
close	1
DMA	0.229405
HIGH	0.999088
LOW	0.998872
MA120	0.991243
MA10	0.991243
MA20	0.991243
MA5	0.991243
MACD	0.233399
open	0.997441
rps	0.139261
SI	0.004296
STD	0.462387
swing	0.372811
Tech20	0.450757

由上表可知，CCI、RPS、SI 三个指标的相关系数的绝对值小于 0.2，故删除，最终剩下四个维度的 18 个输入特征。为了使得最终选择的输入特征的相关系数更具可视化，其相关系数热力图如下：

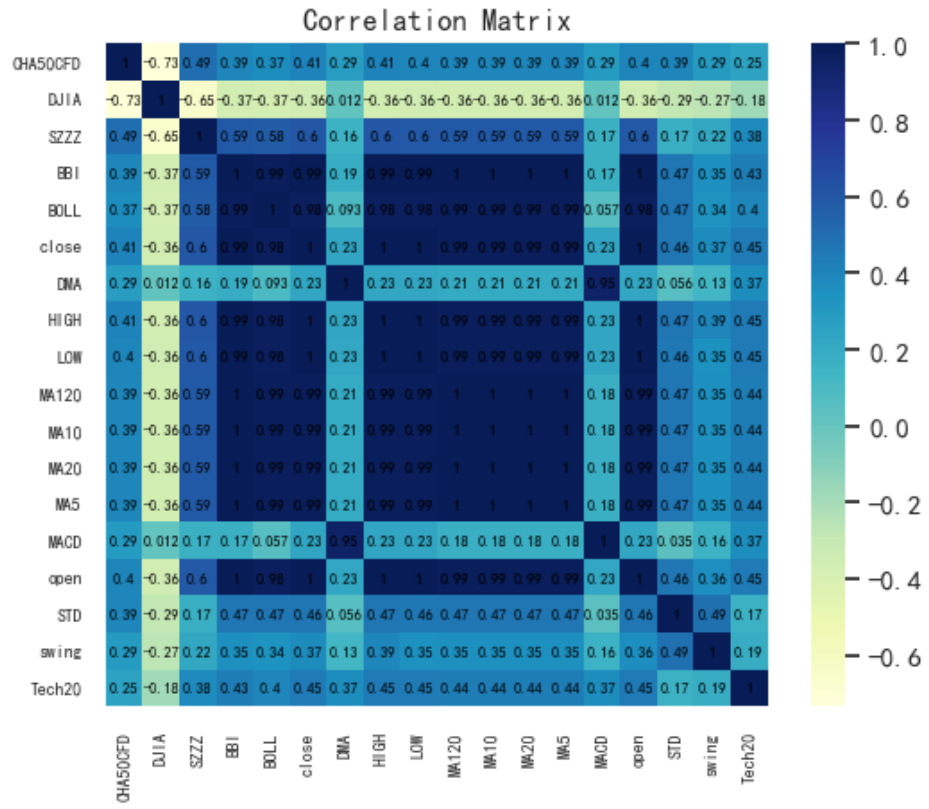

图 8 指标相关系数热力图

上图展示的是最终选定的 18 个输入特征与收盘价的相关系数情况，通过颜色可以判定两者的相关性。本文经过 IC 值与相关系数筛选后筛选出最后的输入特征，展示如下：

表 5　多维度输入特征最终表

维度	指标释义	指标代码
行情因素指标	前一天收盘价	close
	最高价	HIGH
	最低价	LOW
	开盘价	open
	振幅	swing
国内市场因素	上证综指	SZZZ
技术指标	STD 标准差	STD
	阶段强势指标	Tech20
	BBI 多空指数	BBI
	DMA 平均线差	DMA
	MA120	MA120
	MA10	MA10
	MA20	MA20
	MA5	MA5
	MACD	MACD
	BOLL	BOLL
关联市场因素	富时中国 50 指数	CHA50CFD
	道琼斯工业平均指数	DJIA

4.1.4 数据预处理

在利用机器学习模型进行训练前，需要对所有数据进行预处理，不论是股票收盘价还是各个输入特征的时间序列数据，由于时间跨度较长，会存在一定的缺失，并且由于数据来源问题还可能存在一定的异常值，另外不同的输入特征之间的计算方式不同因此会有较大的量纲差距，这些都会对模型训练产生一定的影响，因此首先需要对上述数据进行缺失值、异常值以及归一化的处理。

（1）　异常值处理

由于数据来源的问题以及真实市场当中各数据可能存在较大的波动性，使得收集来的股价数据与输入特征数据存在异常值。本文首先使用 3σ 原则查找异常值，当股票的历史数据在$(\mu-3\sigma,\ \mu+3\sigma)$区间以外时即被认为是异常值，本文对异常值采用的是删除处理。

（2）　缺失值处理

由于本身数据可能存在一定缺失，且上一步将异常值删除的处理导致数据序列当中存在缺失值。首先通过对缺失值进行查找，由于本文采用的数据均为万的导出，因此原先数据存在缺失的情况不多，整体上本文查找到的缺失值采用使用插值的方式来处理。

（3） 归一化处理

由于选取的股票数量众多，不同公司的股价及指标数据变量的取值差别极大，同时不同指标之间由于计算方式的差别也会有特别大的差距。在预测过程中，不论是较小还是较大的量纲数据都会影响预测的精度，所以本文在利用模型训练预测股价之前，先对股价序列数据和各个输入特征序列数据都进行了归一化处理，将所有数据值控制在[0,1]区间内，公式如下所示:

$$X_i = \frac{x_i - x_{\max}}{x_{\max} - x_{\min}} \tag{4-2}$$

其中 x_i 、 X_i 分别是在归一化处理前后的时间序列数据， $x_{\max}$ 和 $x_{\min}$ 分别表示原数据在时间序列当中的最大值和最小值。

4.1.5 股票收盘价数据小波降噪

在真实股票市场中，有非常多的影响因素会使得股票价格产生剧烈波动，不论是宏观经济环境还是相关行业政策法规，亦或者是投资者情绪等因素都可能会影响股票价格异常波动，当波动较大时股票收盘价便会存在大量无效数据，为了使得预测模型中使用更高质量的数据，在本小节中将所选取的沪深 300 成分股的 300 只股票中的收盘价数据进行小波降噪处理。

利用小波降噪来处理数据时，需要确定的是效果最好的小波降噪函数以及确定分解的层数。只有选择合适的小波函数，才能有效地减少在小波分解的过程当中对收盘价时间序列数据造成的信息损失。本文数据量较大，所以在确定小波分解过程中小波分解的层数时，统一采用 3 层分解层数。

表 6 小波函数选择表

评价指标	Db8	Coif3	Sym3
SNR	42.1512	41.9365	42.0951
RMSE	0.1145	0.1173	0.1152

其中，SNR 值越大说明降噪后数据质量越优质，RMSE 越小说明降噪过程中对数据的破坏越少。综合上表的评价结果本文选用 Db8 小波函数作为本文降噪函数。

由于共有 300 只股票的收盘价数据需要进行降噪处理，所以本文在此选择沪深 300 成分股中的 000001.SH 收盘价数据来进行小波降噪分析，数据时间为 2017

年-2021 年，对于该股数据，本文选择 db8 小波函数，并对数据进行 3 层分解。

最终降噪结果如下图所示：

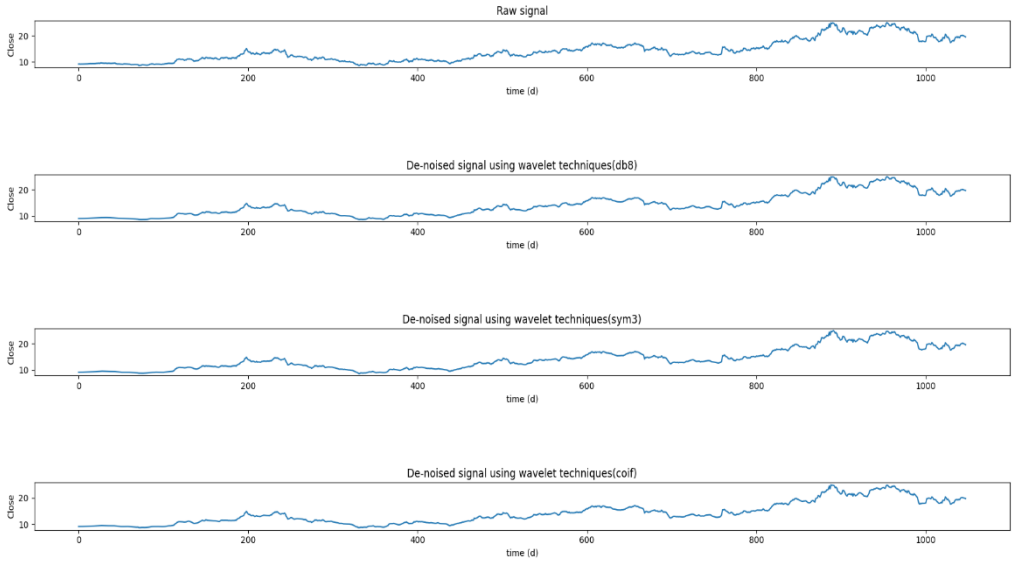

图 9 三种函数降噪结果

第一张图展示的是未降噪前的股票历史收盘价数据，后续三张图分别展示的是通过三种不同的降噪函数得到的更光滑后的股票收盘价数据。

4.2 基于 ATT-LSTM 模型的股票收盘价预测

在前文的步骤中，首先完成了多维度因子库的建立，而后完成了数据的预处理，其中包括处理数据的异常值、缺失值、归一化，收盘价小波降噪，同时对于大量的多维度因子进行 IC 值评价与相关系数筛选来确定最终输入特征，接下来将进入利用 ATT-LSTM 模型对沪深 300 成分股的收盘价进行预测的部分。

由于本文筛选沪深 300 成分股时使用的是 2022 年 7 月的数据，因此从 2017 年至今可能会存在部分股票数据不全的情况，剔除数据量不够的股票后，本文股票收盘价预测标的选用得是沪深 300 成分股共 273 只股票，其中涉及 29 个行业，并通过阅读大量文献，将行情因素指标、关联市场因素指标、国内市场因素指标及技术指标共 38 个指标加入本文研究得多维度因子库中，而后通过相关性检验、IC 值评估来确定 18 个有效的输入特征。本文在 LSTM 模型的基础上加入了 Attention 注意力机制层以此来提高整个模型的预测精度同时增加可解释性。

表 7 股价预测模型训练与测试区间

训练集	测试集
2017 年 1 月 1 日-2020 年 6 月 28 日	2020 年 6 月 29 日-2021 年 10 月 30 日

接下来本文选取 2017 年 1 月 1 日-2020 年 6 月 28 日的股票日数据，并将 2017 年 1 月 1 日-2020 年 6 月 28 日的数据作为训练集进行训练，2020 年 6 月 29 日-2021 年 10 月 30 日的数据作为测试集来对预测模型进行测试。本文将使用 ATT-LSTM 模型来对沪深 300 所有成分股的收盘价数据进行预测及评价基于多维度因子库筛选出并且使用其中的 18 个输入特征，分别涵盖了四个维度：行情因素、国内市场因素、关联市场因素与技术指标。

鉴于后文交易策略是以 3 天为一周期进行换仓，所以本文在利用 ATT-LSTM 模型预测时进行未来 3 天的预测，在训练过程中使用了前 5 天的数据。

在训练机器学习模型的过程中，通常会采用的方式是先调整不同的 epoch 来观察模型的损失函数是否符合预期，本文由于是将添加注意力机制前后的 LSTM 模型来进行横向对比，因此先从二者的损失函数表现来参考。

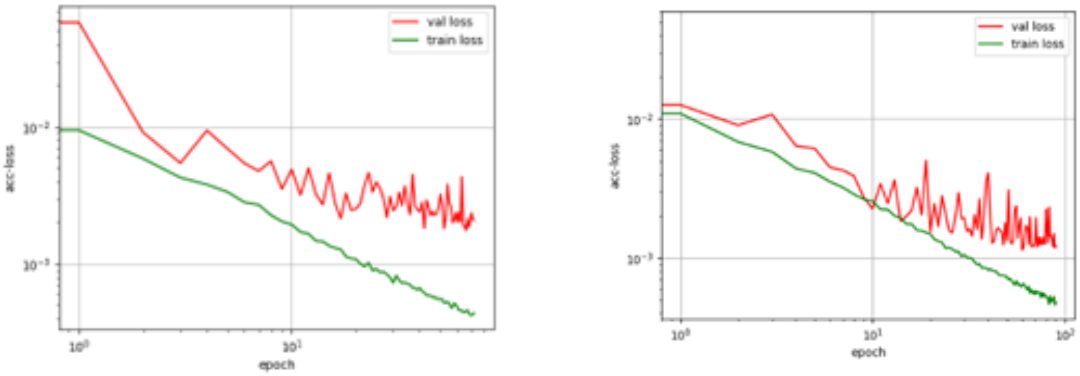

图 10 损失函数结果示意图(左 LSTM,右 ATT-LSTM)

由上图两种模型的训练集与验证集的损失函数可知，两者大趋势都向下降，可见两者都是有效的模型，但是相比较而言 ATT-LSTM 模型的损失函数曲线下降的更快且更稳定，因此 ATT-LSTM 模型的效果更优。但单从损失函数是无法评估预测模型效果的，因此还需要从拟合程度及误差情况来对模型下效果进行评价。

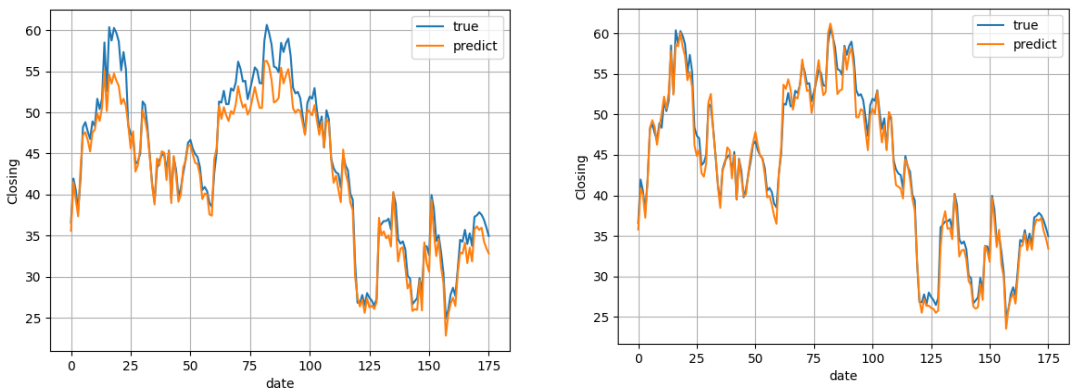

图 11 中国平安股价拟合结果示意图(左 LSTM,右 ATT-LSTM)

由上图可知加入了注意力机制的 LSTM 模型的拟合精度更高一些，但是考虑到模型的泛化能力，所以需要对沪深 300 的所有股票进行模型预测，通过初步预测发现有拟合效果与原来收盘价差异非常大的股票，通过剔除部分 R^2 测试异常的股票后得到最终的评价结果，最终将所有预测评价结果取均值后如下表所示：

表 8 测试集预测模型评价结果均值

模型	RMSE	MAE	MSE	R^2
LSTM	0.12382	0.08951	0.02093	0.8989
ATT-LSTM	0.10549	0.07776	0.01358	0.9318

上表的均值结果是将所有单只股票的预测评价结果取了平均值，因此会出现 MSE 均值不是 RMSE 均值平方的情况。很明显地可以看出，增加了注意力机制之后，LSTM 模型有更好的模型预测表现，不仅体现在有更高的拟合优度，R^2 达 0.9318，而单纯的 LSTM 模型的拟合优度为 0.8989，另外添加了注意力机制的 LSTM 模型在 MSE、RMSE、MAE 上的表现都更优于 LSTM 模型预测的结果，三者都更低，说明预测结果与原数据的误差更小。综合而言，ATT-LSTM 模型在本文预测的股票标的中的泛化能力更强，也有更好的预测效果，本文构建量化交易策略将会和 ATT-LSTM 模型的预测结果进行结合。

第 5 章 基于 ATT-LSTM 多维度模型的股票交易策略的构造与有效性评价

通过对上述股票收盘价预测模型进行比较可以发现，ATT-LSTM 模型在股价预测上有更优的结果，所以本文基于此模型的预测结果来进行交易策略的构造，同时在进行回测时选择在测试集中进行。接下来本文的交易策略构建基于 ATT-LSTM 模型对沪深 300 成分股进行预测，选出排序靠前股票后利用均线距离指标 ΔD_{MA} 来确定买入的股票数量，回测周期为 ATT-LSTM 模型测试集时间内，并通过各评价指标来评估交易策略的有效性。

5.1 股票策略的构建

5.1.1 确定建仓股票

本文选取的股票标的是沪深 300 成分股的 300 只股票，通过对该 300 只股票收盘价进行预测分析，选择数据的时间跨度为 2017 年 1 月 1 日到 2021 年 10 月 31 日，选择数据即为通过多维度因子库进行相关检验及 IC 检验筛选得到的 18 个有效特征，通过 ATT-LSTM 模型来对股票收盘价进行预测，训练集为 2017 年 1 月 1 日-2020 年 6 月 28 日，测试集为 2020 年 6 月 29 日-2021 年 10 月 30 日。

由于在过往研究中，只有基于历史的回测具有一定的超额收益，才会对真实股票市场的实盘有一定的借鉴意义，所以本文将利用测试集中的结果来进行回测。本文通过使用前 5 天的输入特征数据来预测未来 3 天的股票收盘价，本文计划通过对每次对收盘价预测后，将预测结果的对数收益率进行排序，将排名在前 35 名的股票加入买卖股票池。

5.1.2 均线距离指标 ΔD_{MA}

通过上文 ATT-LSTM 模型的预测结果展示，每次股价预测都会得到未来 3 天的股票价格，本文根据预测的收盘价数据计算对数收益率，将排名前 35 的股票作为投资标的放入待选股票池中，每只股票买入的数量等权分配。通过本文构建的均线距离指标 ΔD_{MA} 的上期与本期的正负变化以及绝对值变化来确定每期买入的股票数。以 3 天为一个周期进行换仓，依次来完成整个回测以及对比的过程。

均线距离指标的构建首先是基于短期均线与长期均线的距离与位置，其次是基于上期与本期的距离变化，从而得出市场大势的赚钱效应如何。如当短期均线

在长期均线下方时可知，对于短期投资而言目前的市场价格较低，但是本期较上期而言距离若缩短则说明当下的投资环境有转好的迹象，所以需要确定不同情况下的股票池，以匹配相对应的市场投资情况。

5.1.3 止损设置

在实际股票交易的过程中可能会存在一系列不可预测的风险，即使拥有相对精准的预测结果，但是实际操作过程中会存在各种意外，因为不论是何种类型的交易策略设计，加入适当的止损线都是有必要的，能够规避策略在极端情况下存在的风险，本文将止损点设置为持仓价值亏损的 3%、6%，10%，如具有一定风险偏好的投资者适合使用 10%的止损方式，设置不同的止损线可以适配不同风险偏好的投资者，能让本文的策略在真实股票交易市场更有实用价值。

最终本文的交易策略框架如下图所示，利用 ATT-LSTM 模型预测的结果进行收益率正向排序，并且利用本文创建的均线距离指标 ΔD_{MA} 来进行控仓，最后利用不同得到止损线来进行调整：

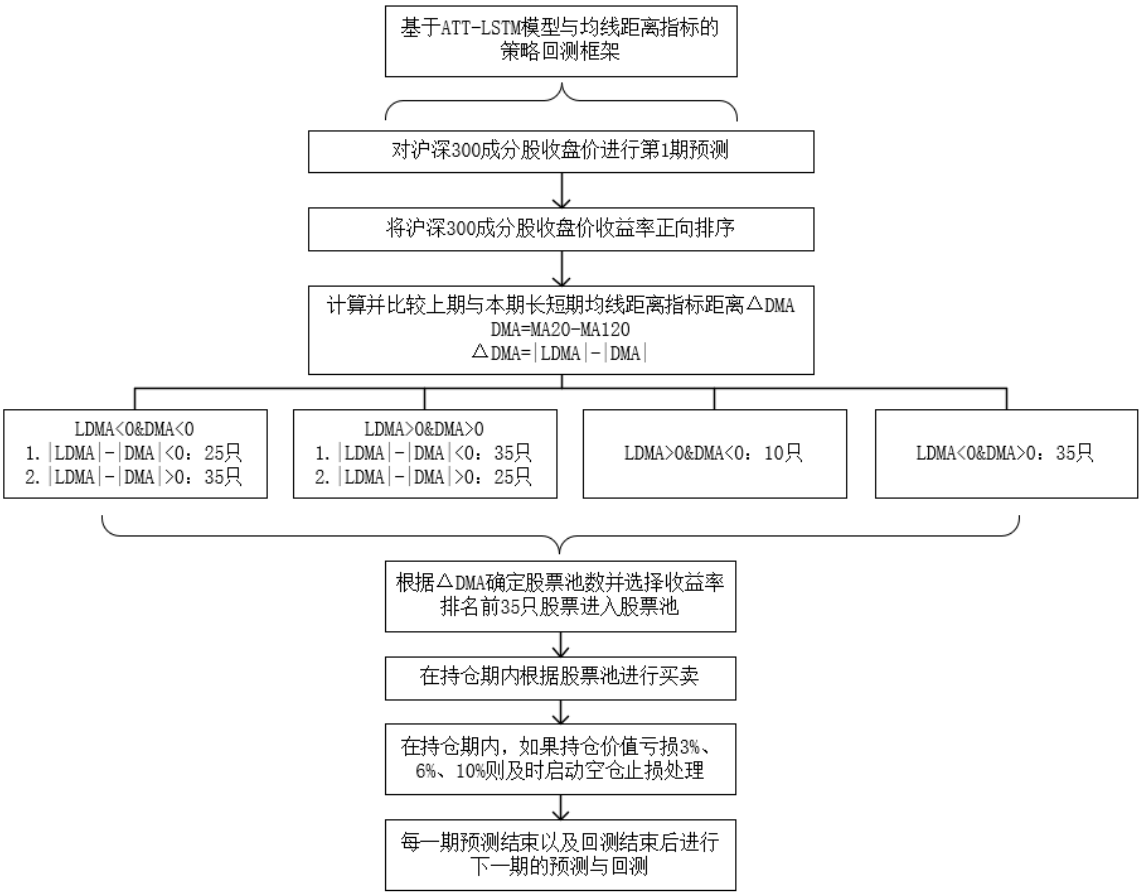

图 12 量化交易策略流程图

5.2 股票策略回测结果与有效性评价

本文基于多维度因子库的 38 个特征，对其进行一系列筛选最后确定 18 个输入特征，选取的数据集为 2017 年 1 月 1 日-2021 年 10 月 31 日的股票日数据，其中 2017 年 1 月 1 日-2020 年 6 月 28 日为训练集，2020 年 6 月 29 日-2021 年 10 月 31 日为测试集，本文的预测模型利用前 5 天的数据对未来 3 天进行预测，根据每期预测的结果来计算收益率，综合选择排名在前 35 名的股票进入股票池。并且通过本文的均线距离指标 ΔD_{MA} 进行判断，本期市场的投资赚钱效应如何，若本期市场大势乐观则增加买入股票的配资，否则反之，尽可能使收益最大化且风险最小化。

表 9 策略参数

回测参数	参数设置值
交易标的	沪深 300 成分股
收益基准	沪深 300 大盘收益
回测区间	2020.06.28-2021．10．30
数据来源	Wind
初始资金	￥1，000，000
止损方式	持仓止损的 3%、6%，10%
手续费	0.1%
换仓时间	每 3 天一次强制换仓

本文的股票交易策略是基于 ATT-LSTM 模型来预测股票价格，根据模型预测出各成分股的次日收盘价，并且筛选出预测收益率排名前 35 的股票进入股票池，同时引入均线距离指标 ΔD_{MA} 来对准备买入的股票数量进行限定，同时根据不同的情况进行不同的仓位设置，在回测过程中，如何通过均线距离指标来设定仓位十分重要，下表指的是在不同情况下筛选多少数量进股票池进行买入操作，本文的设定如下表所示：

表 10 均线距离指标不同控仓情形

	ΔD_{MA}<0(只)	ΔD_{MA}>0(只)
LDMA&DMA<0	25	35
LDMA&DMA>0	35	25
LDMA>0&DMA<0	10	10
LDMA<0&DMA>0	35	35

下图展示了基准收益率与策略单日收益率的对比情况，很明显的可以看出：基准单日收益率的波动较大，风险较大，而且亏损的天数也较多，相比较而言策略单日收益率的情况大多都在 0 以上，虽然单日盈利的上限没有特别高，但是属于处在稳定盈利的过程中。本文构建的策略在回测期间单日盈利天数为 211 天，

占整体回测交易日的 64.72%，可见策略的收益情况较好。

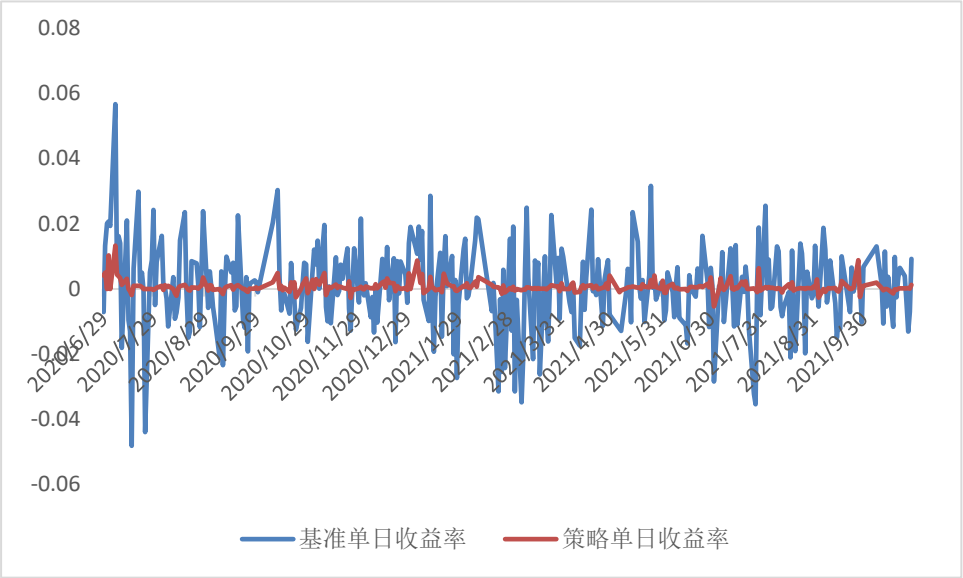

图 13 基准与策略单日收益率对比

对比基准和策略的单日收益率能够反映策略在每天的交易状态下的收益情况，但不论哪种策略都会持续一定的长周期，因此还需要评价策略在一段较长的事件内的累计收益率情况，下图展示了基准与策略累积收益率的情况，回测区间为 2020 年 6 月 29 日-2021 年 10 月 30 日：

图 14 基准与策略的累积收益率对比

上图清晰的反映出在回测期间，策略累积收益率一直处于缓慢上升的过程中，而基准累计收益率上下波动较明显，并且在回测的后半期一直在 18%左右上下波动，同时从曲线的趋势上来说，本文的策略能够逆势上涨，并且能够持续性的盈利，说明本文策略有较好的表现。本文构建的基于 ATT-LSTM 模型与均线距离指标 ΔD_{MA} 的策略在回测期间的累积收益率达 23.04%，同期基准累积收益率为

19.76%，将策略的累积收益率转换成年化收益率为 24.44%。

下表表示的是基于 ATT-LSTM 模型预测结果与上述策略构建的情况下的回测结果展示：

表 11 策略回测结果评价

评价指标	评价结果
年化收益率	24.44%
夏普比率	1.24
最大回撤	6.36%
波动率	0.176%

在回测过程中，本文使用的策略有较好的年化收益率 24.44%，同时也可以知道策略的波动率很小，能够稳健盈利，夏普比率为 1.24。因此可知本文在 ATT-LSTM 模型在对股价进行高精度预测的基础上，并且利用均线距离指标 ΔD_{MA} 进行控仓的策略有良好的收益与极低的风险，具备一定真实股票市场的参考意义。

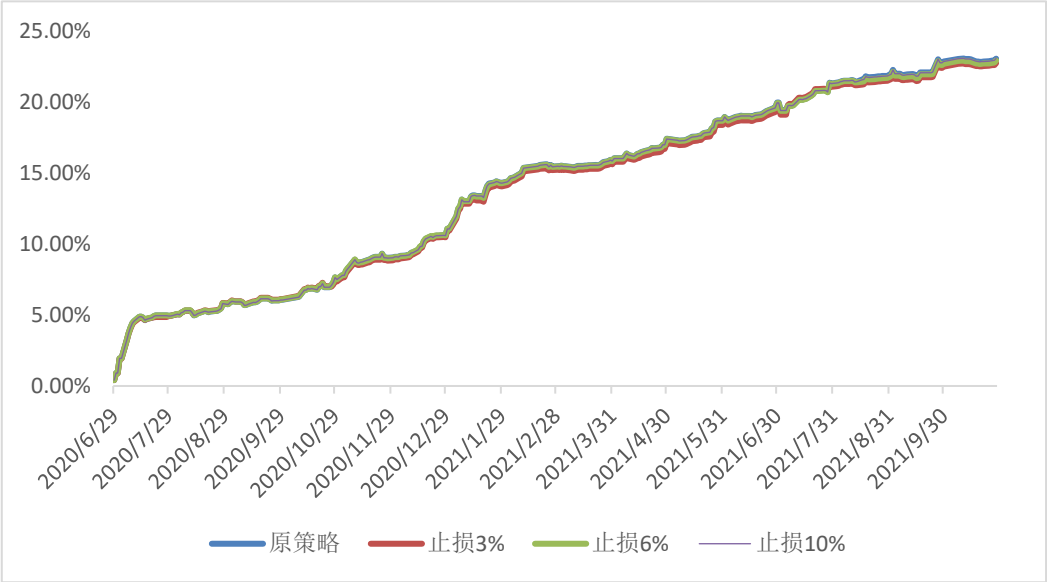

图 15 添加不同止损线的交易结果

就本文的回测区间而言，本文的原策略收益效果较好，原因在于盈利的天数较多，若止损设置比较小，则有可能在股价回调之前就已经卖出，因此会损失一部分盈利，但本文由于单个换仓周期的损失较小，不同的止损线设置对于本文的策略效果没有太大影响，但考虑到本文策略在实盘过程中会遇到各种各样其他极端且突发的情况，故分别设置了 3%、6%、10%的止损线可供不同风险偏好投资者的选择，增加了本文策略的普适性。

5.3 交易策略方案的风险提示

由于早期在股票市场的挣钱效应导致中国股票市场中个人投资者还是远高于国外股市比例，相对的机构专业投资者相对较少，因此市场整体的波动性较大。本文通过利用深度学习模型对股票价格的预测以及通过设置交易策略来完成历史回测，取得了一定的效果，但是同一量化投资策略并不一定适用于所有情况，投资者在实盘投资过程中需要考虑量化策略会带来的各种风险。

本文基于 ATT-LSTM 模型的股票交易策略的风险主要有以下几点：

第一，本文通过使用 ATT-LSTM 模型来预测股票收盘价，由于股票数量众多，则模型泛化能力有限，在实盘操作过程中的应用还需要具体情况具体分析；

第二，股票市场由多方因素影响，不论是国内、国外的宏观事件，还是特定行业的变化，亦或者是由于投资者心理变化产生的系列影响，股票价格都有可能产生巨大波动。本文通过均线距离指标 ΔD_{MA} 来控仓不一定是非常精准且及时的信号，参考的价值会产生一定的限制。

第三，股票价格及各指标数据虽然是时间序列数据，但是股票价格变动是十分剧烈的，会存在许多突发情况，即使在当下投资者在掌握了精准的模型预测结果时，也要有对该结果可能存疑的准备。

第 6 章 结论与展望

前文已经完成了本文的主要阐述，并且通过对 ATT-LSTM 模型预测结果的分析以及对引入均线距离指标 ΔD_{MA} 的量化策略的评价等来说明本文的综合策略设计有较好的效果，具有一定的合理性与借鉴意义。在第六章中，本文将对全文做适当总结与建议评价。

6.1 本文主要总结

本文主要采用 ATT-LSTM 模型和均线距离指标 ΔD_{MA} 构建量化交易策略，在构建股票量化交易策略中，全文一开始通过对多维度因子中的输入特征进行筛选并处理，随后将收盘价进行小波降噪处理，并在此基础上将添加注意力机制前后的 LSTM 预测模型进行对比，并在量化交易策略构建中采用了均线距离指标来控仓并回测的四个部分，能够给予投资者一定的现实借鉴意义。

第一部分基于多维度因子库的输入特征选择与确定，通过对前人大量文献的阅读，本文多维度因子库当中的因子包括行情因素指标、关联市场因素、国内市场因素、技术指标当中的摆动指标、波动指标、超买超卖指标、反趋向指标、量价指标、能量指标、强弱指标、趋向指标与压力支撑指标共 38 个，并通过相关性检验与 IC 检验筛选确 18 个有效输入特征。

第二部分是股票收盘价时间序列数据小波降噪，本文对股票收盘价的数据进行 6 层分解，最终选择了 Db8 小波函数，基于 Db8 小波函数的 3 层降噪效果最好，SNR 值为 42.15，RMSE 的值为 0.4135，优于另外两个函数。

第三部分是预测效果的评价，首先使用了中国平安股票数据作为参考来进行预测效果的评价，可以发现在加入了 Attention 机制之后的 LSTM 模型的损失函数收敛速度更快且数值更低，在拟合结果偏离预测结果并且 ATT-LSTM 模型的 R^2 高于单纯的 LSTM 模型，同时在 MSE、RMSE、MAE 上的表现都更优于 LSTM 模型预测的结果。综合而言，ATT-LSTM 模型在本文预测的股票标的中的泛化能力更强，也有更好的预测效果。

第四部分是基于 ATT-LSTM 模型和均线距离指标 ΔD_{MA} 的股票交易策略的历史回测，通过回测的结果可以发现该策略相较于大盘的基准收益率而言有较大优势，同时年化收益率达 24.44%，另外从单日收益率与夏普比率来看，本文构建的策略风险较低，所有交易日当中有 65%的天数是盈利的，所以本文构建的策略在有抵御风险的基础上有较好的收益表现，能够持续盈利。

6.2 本文的展望与不足

基于本文现已实现的部分可以有以下两点展望，第一是关于拓展本文的研究对象，本文的沪深 300 成分股的选择虽然横跨了 29 个行业，但是在预测过程中不具备动态调整的结构，在筛选沪深 300 成分股时只采用了 2022 年 6 月当期的成分股而不会随着股票市场的变化而动态地调整，因此若能持续地针对每期沪深 300 成分股的变化进行更新，策略则会有更好的适用性；第二是本文在构建多维度因子时考虑了四个维度，但实际引发股市变动的因素还有很多，如许多高频的因子，未来可以基于此考虑使用更多维度的数据，通过筛选后将其通过机器学习模型的训练来预测股价。

本文在设计和实操过程中搭建的整个预测与回测框架仍存在一些不足，首先在选择股票池进行预测时，选择的是沪深 300 的所有成分股，在模型训练当中虽然能够提升股票收盘价的预测精度，但是该模型相对来说泛化能力还有待提高，并不一定适用于所有的时间序列数据预测；关于交易策略的回测周期，本文的预测周期为 16 个月，更适合短期预测，每三天一次频繁换仓的短线交易相对于长线的价值投资会有更大的风险；关于本文的量化交易策略在设置的过程中本金为 100 万元人民币，在实际交易过程中，除了机构投资者，对于普通的个人投资者而言投资股票的金额有限，因此在购买股票相应头寸的过程当中可能会有一定的限制；由于本文的回测系统在买卖过程当中默认能以指定价格买入卖出，但在实际情况过程当中必然存在不能以较低原定价格买入与以较高原定价格卖出的情况，该部分在实际交易过程中会较大程度地影响最终策略的收益率，从而降低策略的效果。

参考文献

[1] Gen Cay, R. Non-linear Prediction of Security Returns with Moving Average Rules[J]. Journal of Forecasting, 1996(3): 43-4.

[2] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 11735-1780.

[3] Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures[J]. Neural Networks, 2005, 18(5-6): 602-610.

[4] Liu X, Ma X. Based on BP Neural Network Stock Prediction[J]. Journal of Curriculum & Teaching, 2012, 1(1).

[5] Dumoulin V, Belghazi I, Poole B, et al. Adversarially Learned Inference[J]. arXiv: Machine Learning, 2016.

[6] Shen Guizhu，Tan Qingping． Deep Learning with Gated Recurrent Unit Networks for Financial Sequence Predictions[J]. Procedia Computer Science, 2018, 131: 895-903．

[7] 黎镭，陈蔼祥，李伟书等． GRU 递归神经网络对股票收盘价的预测研究［J］． 计算机与现代化，2018(11) : 103-108．

[8] 刘宁宁，张量. 基于 PCA-GRU 的股票指数预测模型的研究[J]. 计算机应用研究, 2020, 37(S1): 113-115.

[9] 胡谦. 基于机器学习的量化选股研究[D]. 山东大学, 2016.

[10] Zhang Ru，Huang Chenyu，Zhang Weijian, Chen shaozhen. Multi Factor Stock Selection Model Based on LSTM. International Journal of Economics and Finance[J]: Volume 10, Issue 8. 2018. PP 36-36.

[11] 李钦. 基于 LSTM 的技术指标量化投资策略设计[D]. 广东外语外贸大学, 2021.

[12] 季楚然. 基于 LSTM 神经网络的股指价格预测研究[D]. 安徽财经大学, 2021.

[13] 谢合亮. LSTM 在多因子量化投资模型中的改进及应用研究[D]. 中央财经大学, 2019.

[14] 许丽,张利,李桂城,肖一凡,陈丽绵,唐艳. 基于 LSTM 和新闻情感的股票价格预测方法[J]. 智能计算机与应用, 2022, 12(05): 107-113.

[15] 田雨. 基于投资者情绪和 LSTM 的股价走势预测研究[D]. 上海师范大学, 2021.

[16] 黄建华, 钟敏, 胡庆春. 基于改进粒子群算法的 LSTM 股票预测模型[J]. 华东理工大学学报:自然科学版, 2022, 48(5):12.

[17] 宋刚，张云峰，包芳勋等. 基于粒子群优化 LSTM 的股票预测模型［J］. 北京航空航天大学学报, 2019(12): 2533－2542．

[18] 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.

[19] 李晨阳. 基于 CNN-LSTM 的股票价格预测及量化选股研究[D]. 西北大学, 2021.

[20] 史建楠, 邹俊忠, 张见等. 基于 DMD-LSTM 模型的股票价格时间序列预测研究［J］． 计算机应用研究, 2020(3) : 662－666.

[21] 姚远, 张朝阳. 基于 HP-LSTM 模型的股指价格预测方法[J]. 计算机工程与应用, 2021, 57(24): 296-304.

[22] 林杰, 康慧琳. 基于注意力机制的 LSTM 股价趋势预测研究[J]. 上海管理科学, 2020, 42(01): 109-115.

[23] 沈山山, 李秋敏. 基于注意力机制的 CNN-LSTM 短期股票价格预测[J]. 软件, 2022, 43(02): 73-75.

[24] 邓瑶瑶. 基于 WT-ILSTM-ATT 模型的股票交易策略设计[D]. 上海师范大学, 2021.

[25] 唐成. 基于 Attention-LSTM 深度学习方法的量化投资研究[J]. 商讯,2021(36): 155-157.

[26] 杨向前, 欧阳鹏. 基于 VMD 和 Attention-LSTM 的金融时间序列预测[J]. 软件, 2020, 41(12): 142-149.

[27] Ticknor J L. A Bayesian regularized artificial neural network for stock market forecasting[J]. Expert Systems with Applications, 2013, 40（14）：5501-5506.

[28] Enke D, Thawornwong S: The use of data mining and neural networks for forecasting stock market returns[J]. Expert Systems with Applications, 2005, 29(4): 927-940.

[29] 叶伟睿. 基于深度学习的公募基金量化系统研究与实现[D]. 辽宁大学, 2022.

[30] 刘天贶. 私募基金智能股票量化交易系统的研究与实现[D]. 辽宁大学, 2022.

[31] 钟正豪. 基于机器学习的 A 股量化交易策略研究[D]. 西南财经大学, 2022.

[32] 邱冬阳, 丁玲. 基于多维高频数据和 LSTM 模型的沪深 300 股指期货价格预测[J]. 重庆理工大学学报(社会科学), 2022, 36(03): 55-69.

[33] 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.

[34] 刘子豪. 基于降维方法的信用风险评估模型研究及其应用[D]. 安徽财经大学, 2021

[35] 余娜. 量化选股及量化择时策略研究[D].山东大学,2021.

[36] 张苗苗. 基于 PCA 降维、LSTM 和混频模型的上证综指价格预测研究[D].东华大学, 2021.

[37] 张怡. 基于 ARIMA 和 AT-LSTM 组合模型的股票价格预测[J]. 电脑知识与技术, 2022, 18(11): 118-121.

[38] 程孟菲, 高淑萍. 基于深度迁移学习的多尺度股票预测[J]. 计算机工程与应用, 2022, 58(12): 249-259.

[39] 许雪晨, 田侃. 一种基于金融文本情感分析的股票指数预测新方法[J]. 数量经济技术经济研究, 2021, 38(12): 124-145.

[40] 宋丽娜. 基于情感分析和 PCA-LSTM 模型的股票价格预测[J]. 中国管理信息化, 2021, 24(21): 159-161.

[41] 戴小雪, 张蜀林. 基于强化学习方法的股票交易应用[J]. 经营与管理, 2021(03): 23-27.

[42] 方红, 韩星煜, 徐涛. 改进型基于 LSTM 的股票预测方法[J]. 安徽大学学报(自然科学版), 2019, 43(06): 36-42.

[43] 毛景慧. 基于 LSTM 深度神经网络的股市时间序列预测精度的影响因素研究[D]. 暨南大学，2017.

[44] 杜睿. 基于 GRU 改进的 LSTM 门控制长短期记忆网络的股票交易策略设计[D]. 上海师范大学, 2020.

[45] 潘飞. 基于多尺度特征的 Attention-MCNN-LSTM 模型的构建和实证研究[D]. 浙江工商大学, 2021.

[46] 胡聿文. 基于优化 LSTM 模型的股票预测[J]. 计算机科学, 2021, 48(S1): 151-157.

[47] 汤如意. 基于 LSTM 神经网络的海龟交易模型优化研究[D]. 华东师范大学, 2021.

[48] 邸浩, 赵学军, 张自力. 基于 LSTM-Adaboost 模型的商品期货投资策略研究[J]. 南方金融, 2018(08): 62-76.

[49] 王宣承．基于 PCA 和神经网络的量化交易智能系统构建—以沪深 300 股指期货为例［J］. 投资研究，2014(9) : 23－39.

[50] 毕秀纯， 张曙光. 基于 LSTM 神经网络的黑色金属期货套利策略模型[J]. 中国科学技术大学学报，2018(2)：25-132.

[51] 胡聿文. 基于多技术指标和深度学习模型的股票趋势预测方法研究[D]. 江西财经大学, 2021.

[52] 王国长，梁焙婷，王金枝． 改进的自适应 PCA 方法在股票市场中的应用［J］． 数理统计与管理，2019(4) : 750－760.

附录

均线距离指标控制股票池数量的部分代码：

```
for d in self.datas:
    wait_list=stock_list[str(d.datetime.date(0))]
    if self.dma[0]<=0 and self.last_dma<=0:
        if abs(self.last_dma)-abs(self.dma[0])<0:
            wait_list=wait_list[0:25]
        else:
            wait_list=wait_list
    elif self.dma[0]>=0 and self.last_dma>=0:
        if abs(self.last_dma)-abs(self.dma[0])>0:
            wait_list=wait_list[0:25]
        else:
            wait_list=wait_list
    elif self.dma[0]<=0 and self.last_dma>=0:
        wait_list=wait_list[0:10]
    elif self.dma[0]>=0 and self.last_dma<=0:
        wait_list=wait_list
```