

# 多种机器学习方法在股票预测中的优势对比

加楚懿

西安外国语大学经济金融学院, 陕西 西安

收稿日期: 2024年12月9日; 录用日期: 2024年12月23日; 发布日期: 2025年1月22日

## 摘要

股票价格的预测是金融领域的一个重要研究课题, 随着机器学习技术的发展, 各种模型在股票预测中的应用也越来越广泛。本文通过比较BP神经网络(BP)、极限学习机(ELM)和长短期记忆网络(LSTM)三种常见机器学习方法的表现, 分析其在股票预测中的优劣。通过对比实验数据的均方误差(MSE)、均方根误差(RMSE)、平均绝对误差(MAE)、和决定系数( $R^2$ )等评价指标, 本文验证了LSTM在处理时间序列数据方面的优势, 特别是在股票价格预测的任务中表现出色, 而BP神经网络和ELM各自具有特定场景下的应用价值。

## 关键词

股票预测, 机器学习, BP神经网络, 极限学习机(ELM), 长短期记忆网络(LSTM)

# Comparison of Advantages of Multiple Machine Learning Methods in Stock Prediction

Chuyi Jia

School of Economics and Finance, Xi'an International Studies University, Xi'an Shaanxi

Received: Dec. 9<sup>th</sup>, 2024; accepted: Dec. 23<sup>rd</sup>, 2024; published: Jan. 22<sup>nd</sup>, 2025

## Abstract

Predicting stock prices is an important research topic in the field of finance, and with the development of machine learning technology, various models are increasingly being applied in stock price prediction. This paper compares the performance of three common machine learning methods, BP neural network (BP), extreme learning machine (ELM), and long short-term memory network (LSTM), to analyze their advantages and disadvantages in stock price prediction. By comparing the evaluation indicators such as mean square error (MSE), root mean square error (RMSE), average absolute error (MAE), and coefficient of determination ( $R^2$ ) of the experimental data, this paper verifies the

advantage of LSTM in handling time series data, especially in the task of stock price prediction, and the excellent performance of LSTM in this task. Meanwhile, BP neural network and ELM each have their own application value in specific scenarios.

## Keywords

Stock Prediction, Machine Learning, BP Neural Network, Extreme Learning Machine (ELM), Long Short-Term Memory Network (LSTM)

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着全球金融市场的迅速发展,投资者对准确预测股票价格的需求变得愈发迫切。然而,股票市场本质上是一个复杂的动态系统,受多种因素影响,如经济数据、市场情绪、政治事件等,导致其价格波动具有高度的非线性和不确定性。传统的时间序列预测方法,如 ARIMA 模型,虽然在一定程度上能够捕捉数据中的趋势和季节性,但由于其线性假设和对复杂关系的捕捉能力不足,往往无法满足高精度预测的需求[1]。

近年来,机器学习技术的兴起为解决这一问题提供了新的思路。BP 神经网络(Backpropagation Neural Network)作为一种经典的多层感知器模型,具有强大的非线性映射能力,广泛应用于模式识别和预测任务。极限学习机(Extreme Learning Machine, ELM)则以其快速学习和良好的泛化能力在大数据处理上展现出优势。长短期记忆网络(Long Short-Term Memory, LSTM)作为一种改进的循环神经网络,专为捕捉序列数据中的长期依赖关系而设计,在处理金融时间序列数据方面表现突出[2]。

本研究在对比分析这三种模型的基础上,进一步探讨它们在股票价格预测中的实际应用效果。通过对不同模型的预测性能、收敛速度以及对市场波动的响应能力进行全面评估,本研究不仅为投资者提供更为科学的决策支持,还为机器学习模型在金融领域的实际应用提供了有价值的参考。研究结果将为未来开发更精确、稳定的股票预测模型奠定基础,促进金融科技的发展与创新。

## 2. 数据介绍与处理

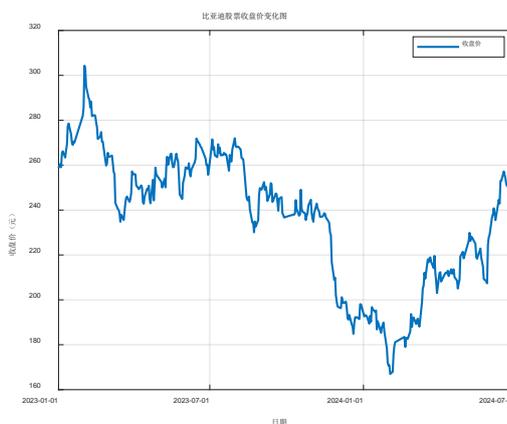


Figure 1. BYD stock closing price change chart (Yuan)

图 1. 比亚迪股票收盘价变化图(元)

(一) 数据介绍

本研究使用的股票数据集来自比亚迪(股票代码: SZ.002594), 涵盖了 2022 年 8 月 1 日至 2024 年 8 月 1 日的日线 K 线数据。

对比亚迪股票的收盘价数据并进行可视化, 如图 1 所示。首先, 代码从 CSV 文件中加载比亚迪股票的历史数据, 提取日期和收盘价, 并将日期转换为 datetime 格式。如图 2 所示, 通过正态分布曲线拟合收盘价的分布情况, 以便观察其分布特征。计算收盘价的基本统计信息, 包括均值、标准差、最小值、最大值和数据个数, 为后续分析提供参考。

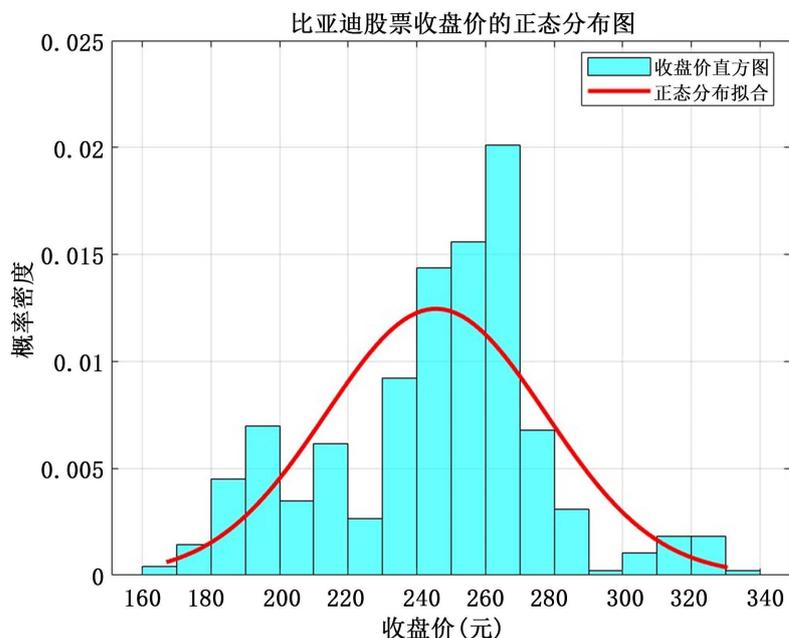


Figure 2. The closing price of BYD stock follows a normal distribution

图 2. 比亚迪股票收盘价正态分布

Table 1. BYD stock related statistical indicators

表 1. 比亚迪股票相关统计指标

| 统计指标 | 数值     |
|------|--------|
| 均值   | 245.55 |
| 标准差  | 32.02  |
| 最小值  | 167.00 |
| 最大值  | 330.54 |
| 数据个数 | 487    |

如表 1 所示, 根据统计指标的分析, 比亚迪股票的收盘价均值为 245.55 元, 表明整体价格水平相对稳定。标准差为 32.02 元, 显示出收盘价的波动性适中, 说明价格在均值附近有一定的波动。最小值为 167.00 元和最大值为 330.54 元, 反映出在观察期内收盘价的变化范围较大, 表明该股票在不同市场条件下经历了较大的价格波动。数据个数为 487, 提供了一个相对充足的样本量, 以支持上述统计分析的可靠性和有效性。

## (二) 数据处理

首先, 从数据集中提取收盘价作为特征和目标变量。特征变量是指输入到模型中的数据, 而目标变量则是模型需要预测的结果。因为本文的研究目标是基于历史收盘价预测未来的收盘价, 所以收盘价既作为特征又作为目标。

其次, 对收盘价数据进行 Min-Max 归一化处理。该方法通过将数据缩放到[0, 1]的范围内, 以减少不同特征之间的数值差异, 避免模型训练时出现数值不稳定的问题。具体公式为:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

再次, 使用滑动窗口法将数据转换为模型所需的输入格式, 采用前 7 天的收盘价数据作为特征, 以预测第 8 天的收盘价。这种方法可以有效地捕捉时间序列数据中的时序特征。

最后, 将数据集划分为训练集和测试集, 其中 80% 的数据用于训练, 20% 的数据用于测试。这一划分方式确保模型能够在未见数据上进行评估, 从而提高其泛化能力。训练集将用于模型参数学习, 而测试集则用于模型性能的验证[3]。

## 3. 模型介绍

神经网络是一种受生物神经系统启发的计算模型, 通过模拟人脑的工作方式来处理信息。它由多个层次的节点组成, 包括输入层、隐藏层和输出层[4]。输入层接收外部数据, 隐藏层通过加权求和和激活函数处理数据, 输出层则生成最终的预测结果。神经网络的核心在于通过前向传播计算输出, 再通过反向传播调整各层之间的权重, 从而逐步减小预测误差, 实现学习的目的。

神经网络在各种应用中表现出色, 包括图像识别、语音识别和自然语言处理等。特别是在深度学习的推动下, 神经网络能够处理极为复杂的任务, 如自动驾驶和金融预测等。通过不断的迭代训练和优化, 神经网络能够从数据中提取有用的特征, 并做出高精度的预测或分类。

BP(反向传播神经网络)是一种经典的前馈神经网络, 由输入层、隐含层和输出层组成[5]。其工作原理包括前向传播和反向传播两个阶段。在前向传播阶段, 输入层接收数据并逐层传递到输出层, 使用激活函数进行非线性变换计算输出值。在反向传播阶段, 网络计算输出层的误差, 并通过链式法则将误差反向传播至每一层, 以更新每个神经元的权重和偏置, 通常使用梯度下降法等优化算法来最小化损失函数。尽管 BP 网络可以处理复杂的非线性关系, 适用于分类和回归任务, 但其训练时间较长且容易过拟合。

ELM(极限学习机)是一种单隐含层前馈神经网络, 其最大特点是极高的训练速度。ELM 通过随机生成隐含层的权重, 而不需要进行传统的迭代训练。学习过程非常简洁: 将输入数据与随机生成的权重进行线性组合后, 使用线性回归方法计算输出层的权重, 从而实现一次性学习。这使得 ELM 在处理大规模数据集时表现出色, 并在多个基准测试中优于传统神经网络。尽管 ELM 训练速度快且泛化能力强, 但由于仅使用单隐含层, 其在处理复杂任务时可能受到限制。

LSTM(长短期记忆网络)是一种特殊的递归神经网络(RNN), 旨在解决传统 RNN 在处理长序列时出现的梯度消失和梯度爆炸问题[6]。LSTM 通过引入记忆单元和门控机制(包括输入门、遗忘门和输出门)来控制信息的流动, 使其能够选择性地记住或遗忘信息[7]。该机制使 LSTM 能够有效捕捉序列数据中的长短期依赖关系, 适用于时间序列预测和自然语言处理等任务[8]。虽然 LSTM 的模型结构相对复杂, 训练过程可能较慢, 但其在处理长序列数据时的建模能力极强, 能够有效应对复杂的时间序列任务[9]。

## 4. 模型创建、训练和对比分析

### (一) 模型的创建和训练

本文利用 MATLAB 的 `feedforwardnet` 函数定义了 BP 网络结构, 选择了 20 个隐藏层节点。该网络使用前向传播算法进行训练, 通过输入训练集数据, 模型调整其权重和偏置以最小化输出误差。训练完成后, 模型便可对测试集进行预测, 展示其在回归任务中的能力。

极限学习机(ELM)模型特点在于其高效的训练速度。首先, 随机初始化了输入层到隐藏层的权重, 并使用 Sigmoid 激活函数计算隐藏层的输出。接着, 利用伪逆方法(`pinv`)计算输出层的权重。通过将输入特征与隐藏层输出相结合, ELM 能够迅速生成预测结果, 避免了传统迭代训练的复杂性。

长短期记忆网络(LSTM)专门用于捕捉时间序列数据的特征。在 Matlab 中定义 LSTM 网络的结构, 包括输入层、LSTM 层、全连接层和回归层。训练过程中采用 Adam 优化算法, 并设置适当的超参数以提升训练效果。数据经过重塑后, 输入到 LSTM 中进行训练, 最终模型能够对测试集进行高效预测, 展现其在时间序列分析中的优势[10]。

将均方误差(MSE)作为训练函数的损失函数, 优化器则是 Adam。

Adam 优化器(Adaptive Moment Estimation)是一种基于一阶梯度的优化算法, 结合了 Momentum 和 RMSProp 的优点。Adam 通过维护每个参数的梯度的移动平均(即一阶矩)和梯度平方的移动平均(即二阶矩), 自适应地调整每个参数的学习率。

BP 神经网络的训练状态如图 3:

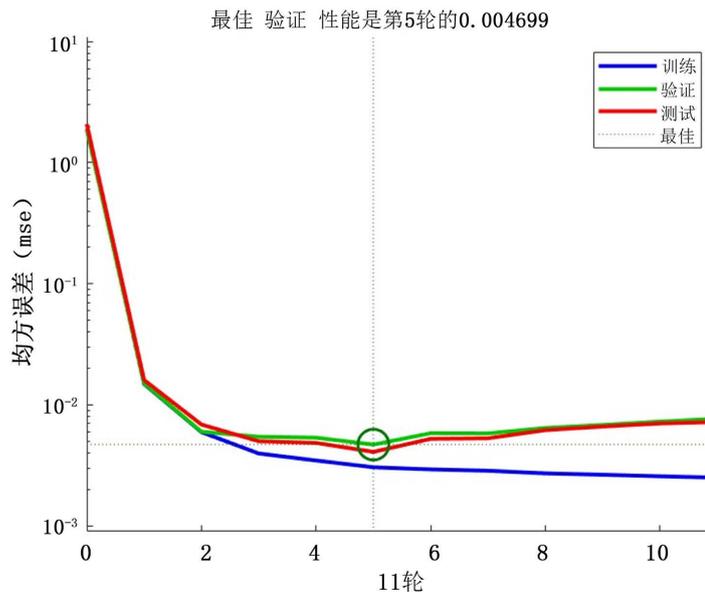
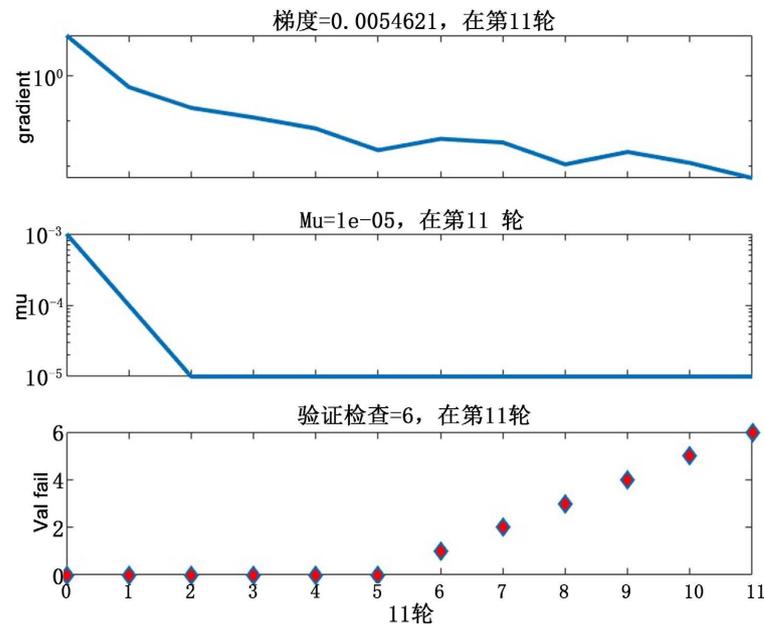
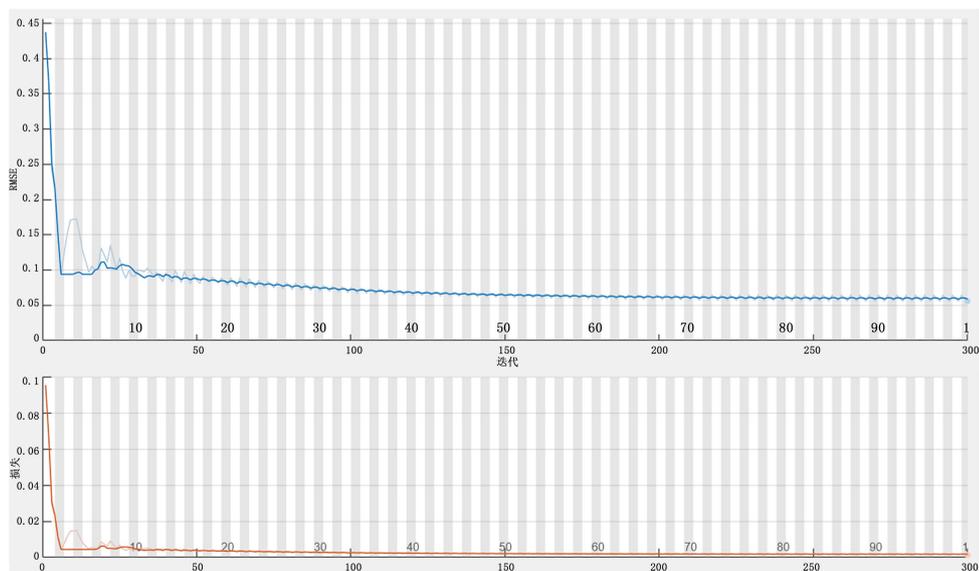


Figure 3. Mean square error graph of BP neural network training  
图 3. BP 神经网络训练均方误差图

从图 4 中可以看出, BP 神经网络在训练过程中表现良好。在第 5 轮时, 验证集的均方误差(MSE)达到了最低值 0.004699, 表明模型在这一轮的表现最佳。梯度逐渐趋于平稳, 学习率( $\mu$ )随着训练过程的深入逐步减小, 这显示出模型在不断优化并收敛。同时, 验证集的失败次数较少, 说明模型的泛化能力较好, 没有出现明显的过拟合现象。总体来说, 该模型训练效果良好, 收敛速度快, 且具备较好的泛化能力。LSTM 模型训练状态如图 5 所示:



**Figure 4.** Gradient and learning rate graph of BP neural network training  
**图 4.** BP 神经网络训练梯度与学习率图



**Figure 5.** LSTM model training diagram  
**图 5.** LSTM 模型训练图

从图 5 中可以看出，LSTM 模型的训练过程较为顺利。在上图中，随着训练轮次的增加，RMSE (均方根误差) 逐渐下降并趋于平稳，表明模型在不断优化并逐步收敛。在下图中，损失值也迅速下降，并在少数轮次后趋于平稳，显示出模型的误差较小，且训练过程没有出现明显的震荡或过拟合现象。总体来看，LSTM 模型在此次训练中表现出良好的收敛性和稳定性。

## (二) 模型对比分析

本文选用了均方误差(MSE)、均方根误差(RMSE)、平均绝对误差(MAE)和决定系数( $R^2$ )作为评估模型性能的指标[11]。MSE 和 RMSE 主要用于衡量预测值与真实值之间的误差，MSE 反映了误差的平方平

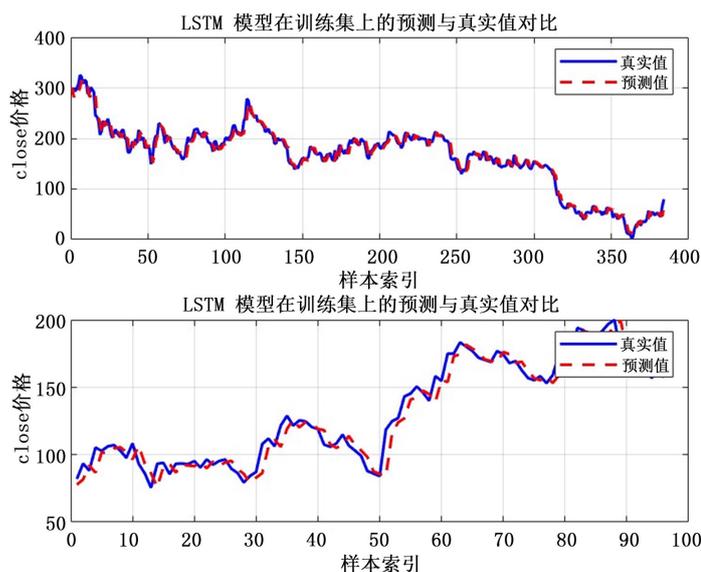
均，能够突出较大误差的影响；RMSE 则是 MSE 的平方根，提供了与原始数据单位一致的误差度量。MAE 则通过计算预测误差的绝对值平均，提供了对误差的直观理解。R<sup>2</sup> 则表示模型解释的变异性比例，其值越接近 1，说明模型的预测能力越强。

**Table 2.** Various indicators of different models

**表 2.** 不同模型的各项指标

| 模型   | MSE    | RMSE   | MAE    | R <sup>2</sup> |
|------|--------|--------|--------|----------------|
| BP   | 0.0036 | 0.0601 | 0.0453 | 0.9293         |
| ELM  | 0.0037 | 0.0609 | 0.0465 | 0.9275         |
| LSTM | 0.003  | 0.0546 | 0.0401 | 0.9417         |

如表 2 所示，根据模型的评估指标结果，LSTM 模型在各项性能指标中表现优越。均方误差(MSE)方面，LSTM 的值为 0.003，明显低于 BP 神经网络的 0.0036 和 ELM 的 0.0037，表明其在数据拟合能力上更强。均方根误差(RMSE)同样显示出 LSTM 的优势，RMSE 仅为 0.0546，相比之下，BP 的 RMSE 为 0.0601，ELM 为 0.0609，这进一步验证了 LSTM 在误差控制上的能力。此外，平均绝对误差(MAE)上，LSTM 的 MAE 为 0.0401，低于 BP 的 0.0453 和 ELM 的 0.0465，显示出其更小的预测误差和更好的稳定性。在决定系数(R<sup>2</sup>)方面，LSTM 达到 0.9417，远高于 BP 的 0.9293 和 ELM 的 0.9275，表明其能更有效地解释数据的变异性，具备更强的泛化能力。综上所述，LSTM 模型在时间序列预测任务中展现了明显优势，虽然 BP 和 ELM 模型也能提供合理预测，但在准确性和稳定性方面有所不足。



**Figure 6.** Comparison between LSTM model prediction and real value

**图 6.** LSTM 模型预测与真实值对比图

从图 6 中可以看出，LSTM 模型在训练集和测试集上的预测值与真实值之间的拟合程度较高。上图显示了模型在训练集上的表现，红色的预测曲线与蓝色的真实值曲线基本重合，说明模型在训练过程中捕捉到了数据的主要趋势。下图则展示了模型在测试集上的表现，同样可以看到预测值与真实值的匹配程度较好，尽管在某些波动较大的区域，预测值稍有偏差。总体而言，LSTM 模型在训练集和测试集上都展现出了较强的预测能力，能够有效地捕捉到时间序列数据的特征[12]。

## 5. 总结与展望

本文通过比较 BP 神经网络、极限学习机和长短期记忆网络在股票价格预测中的应用,对比不同模型的均方误差(MSE)、均方根误差(RMSE)、平均绝对误差(MAE)和决定系数( $R^2$ )等指标,分析了 BP 神经网络、极限学习机(ELM)和长短期记忆网络(LSTM)三种机器学习方法在股票预测领域的优劣势。BP 神经网络因其结构简单,适合处理小规模数据集,但在高维数据和长序列预测中表现欠佳,易陷入局部最优且训练速度较慢。极限学习机具有快速训练和强泛化能力的优势,适合实时预测和大规模数据处理,但其模型稳定性依赖于初始参数,可能导致预测结果波动。LSTM 作为递归神经网络的改进版本,专门针对时间序列数据,能够有效解决传统 RNN 在长序列中遇到的梯度消失问题,适应性强,通常提供更为准确的预测结果。

最终的研究结果表明 LSTM 在处理时间序列数据时表现突出,特别是在股票价格预测任务中,展现了其卓越的性能。总体来看,BP 神经网络和 ELM 在特定场景下仍具备应用价值,而 LSTM 则为高精度的股票预测提供了更为可靠的解决方案。本文的研究结果为金融领域中机器学习模型的实际应用奠定了基础,有助于提升投资决策的科学性与准确性[13]。

本文的研究为机器学习方法在股票价格预测领域的运用提供了一定的理论与实践依据,但仍然存在一些不足之处,因此,未来的研究可以在多个方面进行改进与扩展:

第一,选取更多的、先进的机器学习模型进行分析,例如卷积神经网络和图神经网络等模型,将多模型进行集成,充分利用各个模型的优势;

第二,弥补特征工程方面的不足,深入挖掘影响股票价格的其他潜在特征;

第三,探索深度学习方法如变换器和自注意力机制,应对时间序列预测任务;

第四,扩大数据集,使其能充分反映长期市场趋势,进而减少过拟合风险和时间跨度限制;

第五,关注模型的实时预测能力和在线学习方法,提升模型的适应性;

第六,提升模型的可解释性,帮助投资者理解决策过程,以及进行跨市场研究,探讨模型的普适性,为金融科技的发展提供更广泛的应用场景,推动股票价格预测领域的研究与实践发展。

## 参考文献

- [1] 费振华. 基于机器学习的不平衡数据下个人信用评级预测模型研究[J]. 长江信息通信, 2024, 37(4): 112-114.
- [2] 丁国辉, 刘宇琪, 王言开, 等. 基于翻转网络的低相关性序列数据预测研究[J]. 计算机工程, 2024, 50(2): 78-90.
- [3] 孙晔, 武文华, 樊哲良, 等. 基于 RBF 神经网络的 FPSO 系泊力预测方法和原型应用[C]//中国海洋工程学会. 第十六届中国海洋(岸)工程学术讨论会论文集(上册). 2013: 7.
- [4] 陈思文, 孔亚琪, 刘宇. 基于生成式人工智能的学业评价应用研究——以 ChatGPT 为例[J]. 软件工程, 2023, 26(10): 27-31.
- [5] 刘兆建. 基于 LSTM 的网络流量预测方法研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2023.
- [6] 李楠, 刘豪, 闵亮. 基于 BP 神经网络的个性化跌倒检测研究[J]. 微型电脑应用, 2024, 40(6): 35-37+41.
- [7] Fischer, T. and Krauss, C. (2018) Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions. *European Journal of Operational Research*, **270**, 654-669. <https://doi.org/10.1016/j.ejor.2017.11.054>
- [8] Zheng, Y., Liu, F. and Zhang, C. (2019) Stock Price Prediction Using LSTM Neural Network. *Journal of Computational and Theoretical Nanoscience*, **16**, 2392-2397.
- [9] Chen, Y., Zhang, H. and Zhao, L. (2020) Stock Market Prediction Using a Hybrid Model Based on LSTM and GRU. *Expert Systems with Applications*, **139**, Article ID: 112839.
- [10] Xu, J., Li, D. and Wu, H. (2022) Stock Price Prediction Using a Hybrid Model of CNN and LSTM. *Applied Soft Computing*, **114**, Article ID: 108051.
- [11] 贾秀燕, 孙秋霞, 李勃. 基于 K-means 聚类与 PLS 回归模型的交通速度短时预测[J]. 青岛大学学报(自然科学版), 2023, 36(1): 42-48+53.
- [12] 李伟东, 张学军. 基于深度学习的股票价格预测模型研究[J]. 计算机科学与应用, 2020, 10(3): 198-205.
- [13] 王建辉, 赵文博. 神经网络在股票市场预测中的应用综述[J]. 经济管理, 2021, 43(2): 102-109.