

基于ARIMA模型的股票价格预测

张宇^{1*}, 肖黄梅¹, 张益豪², 赵静³

¹重庆对外经贸学院大数据与智能工程学院, 重庆

²重庆对外经贸学院跨境商务学院, 重庆

³重庆对外经贸学院党政办公室, 重庆

收稿日期: 2025年2月27日; 录用日期: 2025年3月18日; 发布日期: 2025年3月31日

摘要

本文基于ARIMA模型对平安银行股票的收盘价进行了预测研究。通过选取2024年1月2日至2024年11月30日的收盘价作为训练集, 2024年12月1日至2025年1月2日的收盘价作为测试集, 构建了ARIMA (2,1,1)模型。模型经过平稳性检验、参数定阶和诊断检验后, 预测了未来一个月的股票收盘价, 并与实际数据进行了对比。结果表明, ARIMA模型在短期股票价格预测中具有较高的准确性, 预测误差较小, 能够为投资者提供有价值的参考。

关键词

ARIMA模型, 股票, 平稳性检验

Stock Price Prediction Based on the ARIMA Model

Yu Zhang^{1*}, Huangmei Xiao¹, Yihao Zhang², Jing Zhao³

¹Big Data & Intelligence Engineering School, Chongqing College of International Business and Economics, Chongqing

²Global Business School, Chongqing College of International Business and Economics, Chongqing

³Party and Government Office, Chongqing College of International Business and Economics, Chongqing

Received: Feb. 27th, 2025; accepted: Mar. 18th, 2025; published: Mar. 31st, 2025

Abstract

In this paper, the closing price of Ping'an Bank stock is studied based on ARIMA model to predict the closing price. The ARIMA (2,1,1) model was constructed by selecting the closing prices from January 2, 2024 to November 30, 2024 as the training set and the closing prices from December 1, 2024 to

*通讯作者。

January 2, 2025 as the test set. The model predicted the closing price of stocks in the coming month after smoothness test, parameter fixed order and diagnostic test, and compared with the actual data. The results show that the *ARIMA* model has high accuracy in short-term stock price prediction with small prediction error, and can provide valuable reference for investors.

Keywords

ARIMA Model, Stock, Stationarity Test

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

股票市场作为现代金融体系的重要组成部分，其价格波动不仅反映了企业的经营状况和市场供需关系，还受到宏观经济环境、政策变化、市场情绪等多种因素的影响。股票价格的预测一直是金融领域的重要研究课题，吸引了学术界和业界的广泛关注。准确预测股票价格不仅有助于投资者制定合理的投资策略，降低投资风险，还能为市场监管者提供决策支持，维护金融市场的稳定。

随着金融市场的全球化和信息化，股票价格的波动性日益增强，传统的预测方法逐渐显示出局限性，早期的股票价格预测主要依赖于技术分析和基本面分析。技术分析通过研究历史价格和交易量数据，寻找价格波动的规律；基本面分析则通过分析企业的财务报表、行业前景和宏观经济指标，评估股票的内在价值。然而，这些方法往往依赖于分析者的主观判断，难以应对市场的复杂性和不确定性。为了克服传统方法的局限性，研究者们开始引入数学模型和统计方法进行股票价格预测。其中，时间序列分析是最常用的方法之一。*ARIMA* 模型作为时间序列分析的代表性模型，能够捕捉时间序列中的趋势、季节性和随机波动，广泛应用于股票价格预测。*ARIMA* 模型通过对历史数据的拟合，预测未来的价格走势，尤其适用于短期预测。

本文以平安银行股票为例，探讨 *ARIMA* 模型在股票价格预测中的应用。通过构建 *ARIMA* 模型，预测未来一个月的股票收盘价，并评估模型的预测效果，旨在为投资者提供一种有效的预测工具。本文的研究不仅有助于验证 *ARIMA* 模型在股票价格预测中的有效性，还为未来的研究提供了新的思路和方法。

2. 股票研究现状

股票市场是现代金融体系的核心枢纽，其价格波动蕴含海量信息。这一波动不仅直观反映企业的经营实况与市场供需格局，还与宏观经济环境的起伏、政策导向的调整以及市场情绪的变化紧密相连，受这些多元因素交织影响。股票市场的价格波动被视为经济发展的晴雨表，对股票价格的精准预测一直是众多研究学者努力的方向，如何对股价进行精准预测变得越来越重要。

姜淑瑜根据股票市场的特点和 LSTM 递归神经网络的特性，对浦发银行(600000)股价进行预测。实验结果表明，LSTM 模型预测股价，结果误差小，精准度高，具有良好的预测效果[1]。Mao 和 Wang 通过研究精准预测股价方向，实证分析真实股票收益率数据，探究 CNN-LSTM 模型在中国股票价格变动中的学习和预测性能，以及其预测结果在量化交易策略中的应用，研究过程涵盖 CNN-LSTM 模型搭建、股票数据处理、模型训练与测试，以及基于预测结果的量化策略改进，最终发现该模型在股票价格变动学习与预测方面性能良好，依据其预测结果构造的指标对量化交易策略有显著改进效果[2]。牛晓健和侯启

明通过对于真实股票收益率数据的实证分析,探究 CNN-LSTM 模型在中国股票价格变动中的学习和预测性能如何以及 CNN-LSTM 模型的预测结果如何应用于量化交易策略中。研究发现, CNN-LSTM 模型在股票价格变动的学习与预测上具有良好的性能,依据模型预测结果构造的指标对量化交易策略有明显的改进效果[3]。Peng 等人提出一种融合注意力机制的 GAN-LSTM-Attention 模型,通过结合 LSTM 和 CNN 模型,基于四支美股数据(标普 500 指数、苹果、AMD、谷歌)进行多指标实验与消融验证,证明该模型在股价预测中兼具高精度和鲁棒性[4]。Ayman 等人研究通过对比 ARIMA 与 XGBoost 模型对沙特电信公司股票周收盘价的预测,证实 XGBoost 在各项指标上均优于 ARIMA,揭示了机器学习在捕捉数据潜在趋势中的有效性及其相较于传统统计模型的优势,并强调需深化机器学习在金融预测领域的应用探索[5]。陈健和刘伟基研究提出结合 Hyperband 超参数优化算法与 LSTM 的 Hyperband-LSTM 模型,通过沪深 300、上证 50 和深证 100 进行实验验证,证明其较 Bayes-LSTM 和传统 LSTM 模型在股票价格预测中兼具更高精度、更低时间成本及跨市场稳健性[6]。

总体来看,股票研究正朝着数据驱动、技术融合和全球化视角的方向发展。无论是国内还是国外,大数据、人工智能和机器学习技术的应用都在不断深化,为股票研究提供了新的工具和方法。同时,随着全球对可持续发展的重视,ESG 投资成为国内外研究的共同重点。未来,股票研究将更加注重跨学科融合,结合传统方法与现代技术,以应对日益复杂的市场环境,为投资者提供更精准的决策支持。

3. ARIMA 模型介绍

ARIMA 模型是一种经典的时间序列预测模型,广泛应用于经济学、金融学、气象学、销售预测等领域。ARIMA 模型通过对时间序列数据的自回归(AR)、差分(I)和滑动平均(MA)进行建模,能够有效捕捉时间序列中的趋势、季节性和随机波动,从而对未来值进行预测。它用 $ARIMA(p, d, q)$ 的形式来表示,其中 p 代表自回归项的阶数, d 代表差分次数,而 q 代表移动平均项的阶数。

(1) 自回归(AR): 自回归部分表示当前值与过去若干期值之间的线性关系,模型假设当前值可以表示为过去 p 个值的加权和。自回归模型表示如下:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \quad (1)$$

其中, y_t 是当前值, $\phi_1, \phi_2, \dots, \phi_p$ 是自回归系数, ε_t 是误差项。

(2) 差分(I): 差分的主要目的是确保时间序列数据的稳定性。当时间序列中的数据表现出趋势性和季节性时,通过对时间序列数据进行一阶或多阶差分处理,可以消除它的趋势性和季节性的影响,让时间序列数据趋于平稳。公式表示如下:

$$\Delta y_t = y_t - y_{t-1} \quad (2)$$

(3) 滑动平均(MA): 这部分是用来捕捉时间序列数据的随机误差,在 ARIMA 模型中,通过将误差项的当前观测值与过去的观测值来实现的,可以帮助模型调整因自回归部分未能捕捉到的随机波动。模型假定误差项是独立同分布,并且误差项没有自相关性,公式表示如下:

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (3)$$

其中, $\theta_1, \theta_2, \dots, \theta_q$ 是滑动平均系数, ε_t 是误差项。

ARIMA 模型通过对上述三种项的组合得到,即:

$$ARIMA(p, d, q) = AR(p) + I(d) + MA(q) \quad (4)$$

4. ARIMA 模型在股票价格预测中的应用

本文使用 ARIMA 模型对平安银行股票的收盘价进行预测。选取 2024 年 1 月 2 日至 2024 年 11 月 30

日的收盘价作为模型的训练集，2024 年 12 月 1 日至 2025 年 1 月 2 日的收盘价作为测试集。用训练出来的 *ARIMA* 模型预测 2024 年 12 月 1 日至 2025 年 1 月 2 日的收盘价，并与真实数据做比较。最后，用评价指标 MAE、RMSE 和 MAPE 衡量模型的预测准确性。图 1 所示为平安银行股票在 2024 年 1 月 2 日~2024 年 11 月 30 日的每个交易日的收盘价数据的时间序列图，从图中可以看出股票价格变化幅度较大，整体呈上升趋势。

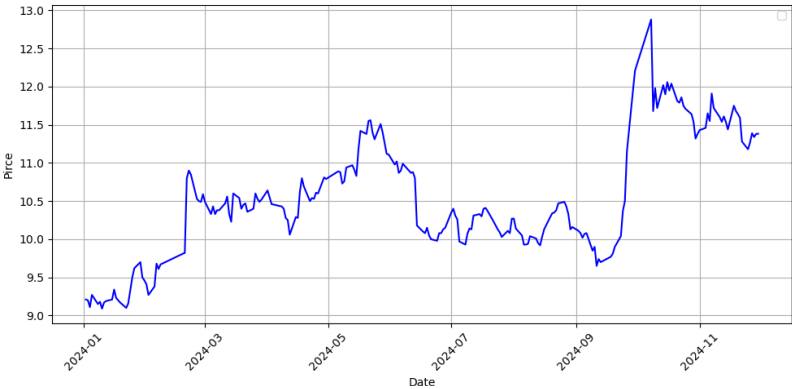


Figure 1. A time series chart of the closing price of Ping'an Bank's stock
图 1. 平安银行股票收盘价时间序列图

4.1. 平稳性检验

对训练集数据做单位根(ADF)检验，ADF 检验如表 1 所示：

Table 1. ADF test
表 1. ADF 检验

Augmented Dickey-Fuller Test	
Test Statistic	-2.2619
P-value	0.1846

由表 1 可知，ADF 检验的 $P\text{-value} = 0.1846 > 0.05$ ，存在单位根。因此，可以认为训练集数据的非平稳序列，需要做平稳化处理。对训练集数据进行一阶差分，得到序列 `diff_1`，并检验 `diff_1` 的平稳性。`Diff_1` 的时间序列图以及 ADF 检验的结果如图 2 和表 2 所示：

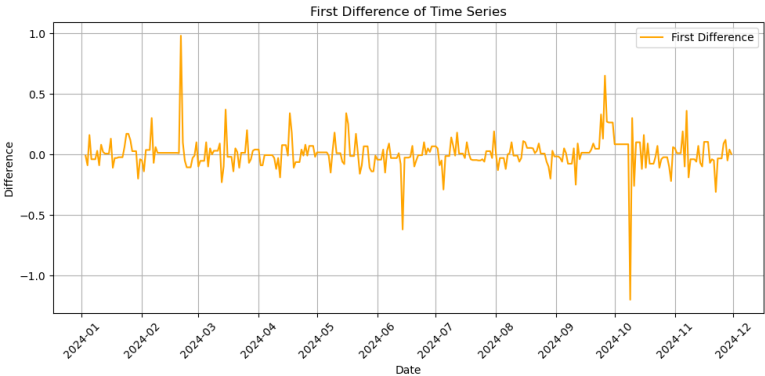


Figure 2. First-order difference time series plots
图 2. 一阶差分时间序列图

Table 2. First-order differential ADF test
表 2. 一阶差分的 ADF 检验

Augmented Dickey-Fuller Test	
Test Statistic	-10.7518e-10
P-value	2.6533e-19

由图 2 可知，一阶差分数据序列基本平稳，并且从表 2 可知 ADF 检验的 P 值 < 0.05 。因此，一阶差分后的序列 diff_1 已经平稳，取 $d = 1$ 。

4.2. 模型定阶

对一阶差分后的数据画出其自相关图(ACF)和偏自相关图(PACF)，如图 3 和图 4 所示。

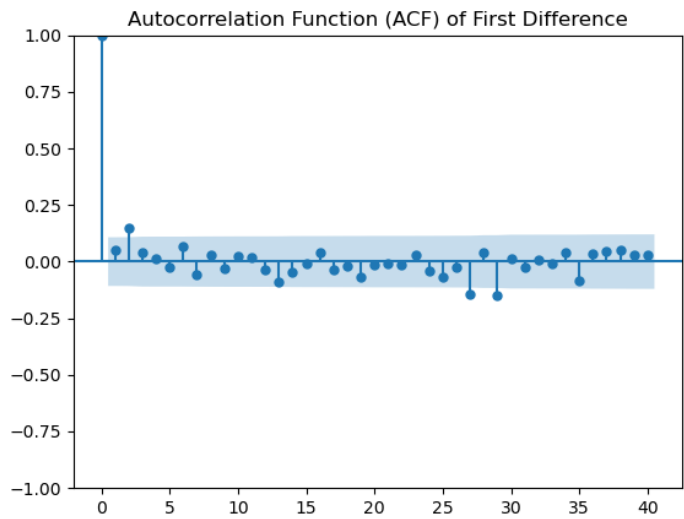


Figure 3. Autocorrelation plot of a first-order difference sequence
图 3. 一阶差分序列的自相关图

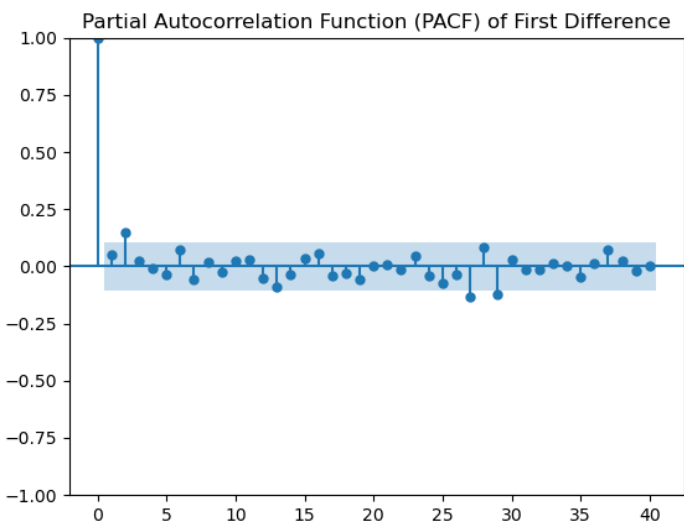


Figure 4. Partial autocorrelation plot of a first-order difference sequence
图 4. 一阶差分序列的偏自相关图

对一阶差分后的数据运用参数网格搜索, 尝试不同的 $ARIMA$ 模型参数 (p, d, q) 的组合, 以找到最小化赤池信息准则(AIC)的组合。AIC 是衡量模型拟合优度的一种准则, AIC 值越小, 表明模型效果越好。

如图 5 所示, 热力图的行表示自回归(AR)阶数, 列表示移动平均(MA)阶数, 每个单元格的颜色表示对应模型的 AIC 值。通过观察热力图, 可以找到 AIC 值最小的模型参数组合, 从而确定最优的 $ARIMA$ 模型。由图可见, 最优参数为(2,1,1)。

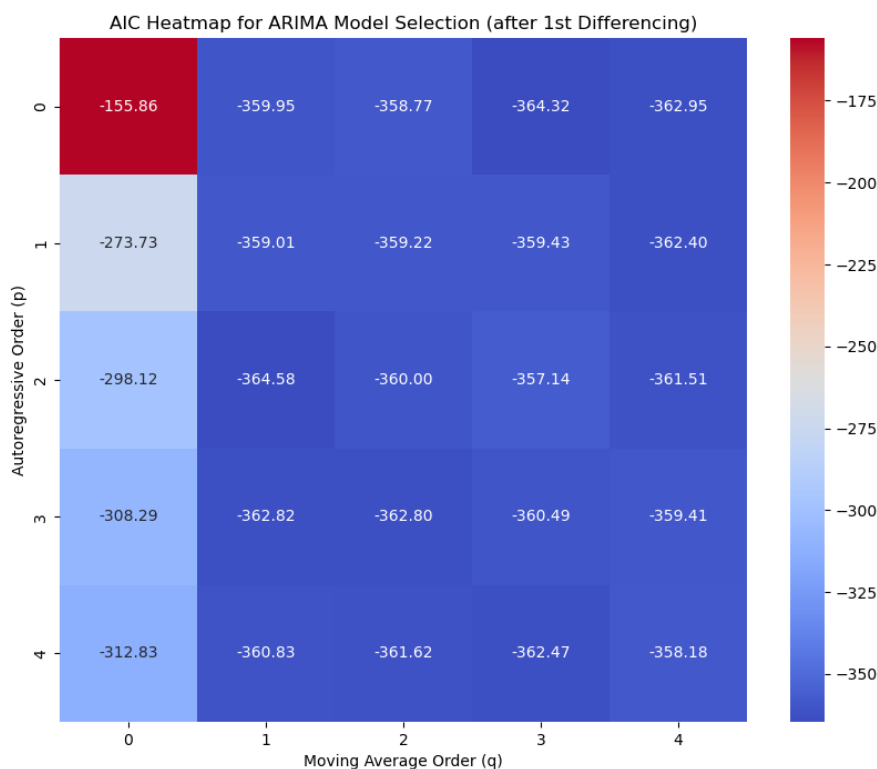


Figure 5. AIC heat map

图 5. AIC 热力图

接着, 利用极大似然估计对模型参数进行估计, $ARIMA$ (2,1,1)模型系数的估计值、置信区间以及检验的 P 值如表 3 所示:

Table 3. $ARIMA$ (2,1,1) parameter estimates and significance test

表 3. $ARIMA$ (2,1,1)各参数估计值和显著性检验

	coef	P-value	95%置信下限	95%置信上限
AR. L1	0.0490	0.0036	-0.0266	0.1246
AR. L2	0.1501	0.0024	0.0533	0.2469
ma. L1	0.0187	0.0000	-1.4116	-0.5874

由表 3 可知, $ARIMA$ (2,1,1)模型各参数对应的 P 值都小于 0.05, 说明各参数均显著, 确定模型为 $ARIMA$ (2,1,1)。将估计的参数代入到模型表达式中得到:

$$(1-B) = \frac{(1-0.0187B)\varepsilon_t}{1-0.049B-0.1501B^2} \quad (5)$$

4.3. 模型的诊断检验

对于所建立的 $ARIMA(2,1,1)$ 模型，还需要对该模型的有效性进行检验。通过分析残差序列图和 QQ 图来检验，只有通过检验，才能利用该模型进行有效的预测。画出残差序列图和 QQ 图，如图 6 和图 7 所示：

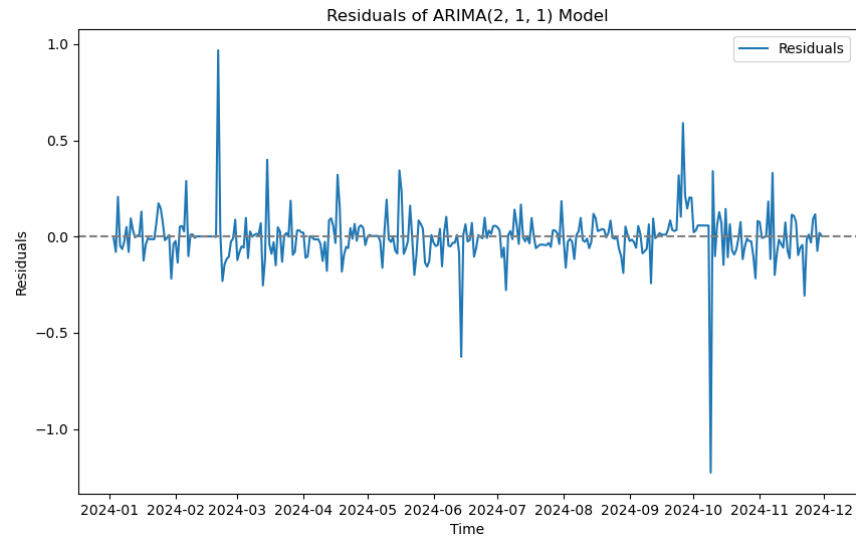


Figure 6. Residual sequence diagram

图 6. 残差序列图

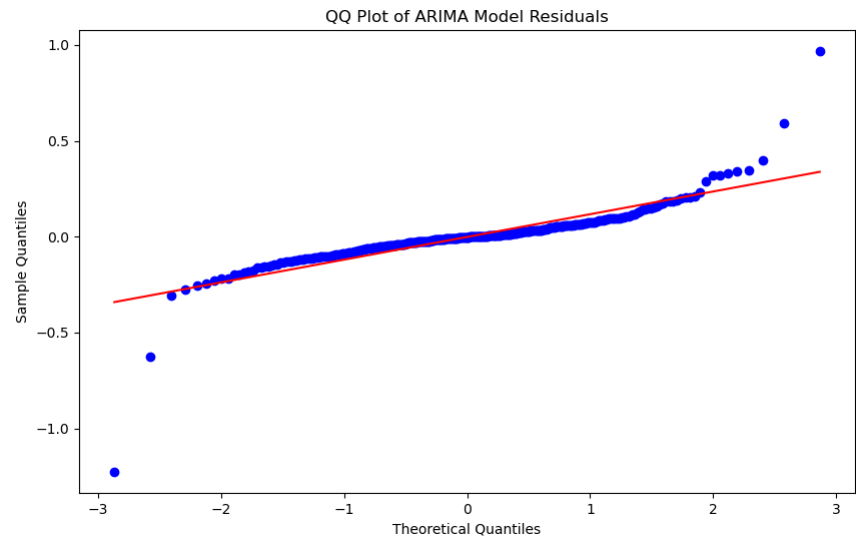


Figure 7. QQ plot of residuals

图 7. 残差 QQ 图

由图 6 和图 7 可知，残差序列是白噪声序列。因此，建立的 $ARIMA(2,1,1)$ 模型是可行的，可以用来预测平安银行股票的收盘价。

4.4. 模型的预测结果

通过建立的 $ARIMA(2,1,1)$ 模型对平安银行 2024 年 12 月 1 日~2025 年 1 月 2 日的数据进行预测，并

与实际值进行对比，预测结果的评价指标如表 4 所示。表 5 是预测的真实值和预测值。

Table 4. Evaluation indexes: MAE, RMSE and MAPE
表 4. 评价指标 MAE、RMSE 和 MAPE

评价指标	值
MAE	0.1759
RMSE	0.2113
MAPE	1.4967

Table 5. The predicted value and relative error of the prediction outcome
表 5. 预测结果的预测值和相对误差

时间	真实值	预测值	相对误差(%)	时间	真实值	预测值	相对误差(%)
2024/12/1	11.3867	11.3947	0.07	2024/12/18	11.65	11.5053	1.24
2024/12/2	11.39	11.4009	0.10	2024/12/19	11.59	11.5118	0.67
2024/12/3	11.49	11.4081	0.71	2024/12/20	11.62	11.5183	0.88
2024/12/4	11.46	11.4146	0.40	2024/12/21	11.65667	11.5247	1.13
2024/12/5	11.44	11.4212	0.16	2024/12/22	11.69333	11.5312	1.39
2024/12/6	11.66	11.4276	1.99	2024/12/23	11.73	11.5377	1.64
2024/12/7	11.6633	11.4341	1.97	2024/12/24	11.86	11.5442	2.66
2024/12/8	11.6667	11.4406	1.94	2024/12/25	11.92	11.5506	3.10
2024/12/9	11.67	11.4471	1.91	2024/12/26	11.86	11.5571	2.55
2024/12/10	11.79	11.4600	2.78	2024/12/27	11.83	11.5636	2.25
2024/12/11	11.73	11.4695	2.22	2024/12/28	11.87	11.5701	2.53
2024/12/12	11.85	11.4765	3.24	2024/12/29	11.91	11.8965	0.11
2024/12/13	11.56	11.473	0.75	2024/12/30	11.95	11.783	1.39
2024/12/14	11.5633	11.4794	0.73	2024/12/31	11.7	11.5895	0.94
2024/12/15	11.5667	11.4859	0.70	2025/1/1	11.565	11.5959	0.27
2024/12/16	11.57	11.4924	0.67	2025/1/2	11.43	11.6024	1.51
2024/12/17	11.53	11.4989	0.27				

如表 5 所示，大多数预测值与真实值之间的相对误差在 2%以内，也有很多预测数据在 1%左右。由表 4 可知 MAE、RMSE 和 MAPE 的值分别是 0.1759、0.2113、1.4967，从这些指标可以看出 $ARIMA(2,1,1)$ 模型对股票的短期预测具有较高的准确性，其中预测结果的 MAE RMSE 可以看出预测误差比较低，而 MAPE 的值表明了模型对股票数据的相对误差控制在较小的范围内。通过 $ARIMA(2,1,1)$ 模型对平安银行股票的收盘价预测，结果显著，可以为股票投资者提供一定的参考价值，帮助他们做出合理的决策。

5. 结语

本文利用 $ARIMA$ 模型对平安银行股票的收盘价进行了预测研究。通过对训练集数据的平稳性检验、模型定阶和诊断检验，确定了 $ARIMA(2,1,1)$ 模型为最优模型。模型预测结果显示，预测值与实际值之间的相对误差较小，MAE、RMSE 和 MAPE 等评价指标均表明模型具有较高的预测准确性。 $ARIMA$ 模型在

短期股票价格预测中表现出色，能够为投资者提供有价值的参考信息。然而，股票市场受多种因素影响，未来研究可以结合其他模型或引入更多影响因素，以进一步提高预测精度。

基金项目

2024-2025 年重庆对外经贸学院科研项目：股票市场波动的统计特征分析与预测模型研究项目 (KYZK2024042)。

参考文献

- [1] 姜淑瑜. 基于 LSTM 模型的股票价格预测[J]. 江苏商论, 2025(1): 83-86.
- [2] Mao, J. and Wang, Z. (2024) Deep Learning-Based Stock Price Prediction Using LSTM Model. *Proceedings of Business and Economic Studies*, 7, 176-185. <https://doi.org/10.26689/pbes.v7i5.8611>
- [3] 牛晓健, 侯启明. 基于 CNN-LSTM 模型的中国股票价格预测与量化策略研究[J/OL]. 贵州省党校学报, 1-18. <https://doi.org/10.16436/j.cnki.52-5023/d.20241128.005>, 2025-02-25.
- [4] Li, P., Wei, Y. and Yin, L. (2025) Research on Stock Price Prediction Method Based on the Gan-Lstm-Attention Model. *Computers, Materials & Continua*, 82, 609-625. <https://doi.org/10.32604/cmc.2024.056651>
- [5] Almaafi, A., Bajaba, S. and Alnori, F. (2023) Stock Price Prediction Using ARIMA versus Xgboost Models: The Case of the Largest Telecommunication Company in the Middle East. *International Journal of Information Technology*, 15, 1813-1818. <https://doi.org/10.1007/s41870-023-01260-4>
- [6] 陈健, 刘伟基. 基于 Hyperband-LSTM 模型的股票价格预测研究[J]. 金融管理研究, 2023(1): 65-85.