

# 基于Bi-LSTM深度学习的股票价格预测

侯亚妮

北方工业大学理学院统计系, 北京  
Email: 1076304044@qq.com

收稿日期: 2021年5月29日; 录用日期: 2021年6月13日; 发布日期: 2021年6月25日

## 摘要

股票价格的预测一直受到金融投资者及学者的广泛关注, 同时也是学者的研究重点。股票价格的非线性、波动性等特点使得使用传统统计学方法进行股票预测的准确率较不理想。长短时记忆循环网络(LSTM)模型在处理时间序列数据上有很大的优势, 为了进一步优化预测模型, 本文提出了双向传播长短时记忆循环网络(Bi-LSTM), 通过双向传播历史数据来对股票价格进行预测。另外, 还引入传统神经网络BP、LSTM、GRU与本文所提Bi-LSTM模型进行对比, 并引入MSE、RMSE、MAE、 $R^2$ 四种误差评估方法来对模型进行评估, 验证模型的预测精度。实证结果表明, 基于Adam优化算法的Bi-LSTM网络优于BP、LSTM、GRU等传统预测模型, 精确度有明显提升, 验证了本模型的有效性和可行性。

## 关键词

股票价格预测, 深度学习, Bi-LSTM模型

# Stock Price Prediction Based on Bi-LSTM Deep Learning

Yani Hou

Department of Statistics, College of Science, North China University of Technology, Beijing  
Email: 1076304044@qq.com

Received: May 29<sup>th</sup>, 2021; accepted: Jun. 13<sup>th</sup>, 2021; published: Jun. 25<sup>th</sup>, 2021

## Abstract

Stock price prediction has been widely concerned by financial investors and scholars, and it is also the research focus of scholars. The non-linearity and volatility of stock prices make the accuracy of stock prediction using traditional statistical methods less than ideal. LSTM model has great advantages in processing time series data. In order to further optimize the prediction model, this

paper proposes a bidirectional propagating long and short time memory cycling network (Bi-LSTM), which can forecast stock prices by bidirectional propagating historical data. In addition, the traditional neural network BP, LSTM and GRU are introduced to compare with the Bi-LSTM model proposed in this paper, and four error evaluation methods of MSE, RMSE, MAE and  $R^2$  are introduced to evaluate the model to verify the prediction accuracy of the model. The empirical results show that the Bi-LSTM network based on ADAM optimization algorithm is superior to BP, LSTM, GRU and other traditional prediction models, and the accuracy is significantly improved, which verifies the effectiveness and feasibility of this model.

## Keywords

Stock Price Forecast, Deep Learning, Bi-LSTM Model

Copyright © 2021 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

随着经济的发展和人民生活水平的提高,股票已经走入大众的视野。作为国家宏观经济的“晴雨表”,股票的市场价格直接关系到金融市场的稳定与发展。股票价格的复杂性、波动性以及高噪声、非线性性等特征使得准确预测股票价格具有一定的难度。

文献[1]基于传统的 ARIMA 模型来对股票的价格进行预测,预测结果在一定程度上可以代表股票的价格走势,但对于长期的趋势表现出了明显的局限性。随着人工智能及大数据的兴起,许多学者提出了使用机器学习算法来对股价进行预测[2] [3],预测结果明显优于传统统计时间序列预测模型。文献[4]将随机森林应用到股票预测中,发现集成学习算法在股票预测中具有一定的作用。文献[5]使用 SVM 模型来对上证指数进行预测,相比于传统统计分析方法,SVM 模型的预测值与实际值的误差率较低。为了解决机器学习模型的泛化能力不足的问题,采用深度学习来发现数据之间的潜在关系并进行建模,使模型在预测方面取得了较好的性能。文献[6]将 GARCH 模型与 BP 神经网络模型进行对比,发现无论是否考虑股票的波动性,GARCH 模型的预测精度均优于 BP 模型。随后有学者将深度学习模型应用到股票预测中[7],文献[8]提出了 LSTM 模型,有效的解决了循环神经网络在时间序列数据的预测中存在梯度消失及梯度爆炸的问题。文献[9] [10]基于 LSTM 神经网络来对沪深 300 指数的收盘价进行预测,试验结果表明 LSTM 模型对于股票价格的预测有显著的效果。但是 LSTM 模型只能通过历史数据来对未来某一时刻的值进行预测,学者们又提出了 Bi-LSTM 模型[11],通过对历史数据与未来数据同时学习来预测未来某一时刻的值,使得模型的预测误差明显降低。

为了进一步提高股票价格预测的准确度,本文引入基于 Adam 优化算法的 Bi-LSTM 模型来对中国建设银行的股票价格进行预测,研究表明基于 Adam 优化算法的 Bi-LSTM 网络的精确度明显优于 BP、LSTM、GRU 等传统预测模型。

## 2. 模型和方法

### 2.1. Bi-LSTM 的基本原理

LSTM (Long short-term memory)是循环网络 RNN 的一种,用来处理时间序列模型,LSTM 模型采取了

特殊的门结构，解决了 RNN 模型在训练过程中的梯度消失以及梯度爆炸等问题。LSTM 模型的基本单元结构如图 1 所示，通过 sigmoid 函数对数据进行激活，充当门控信号；通过门控机制来控制信息的累积或遗忘。

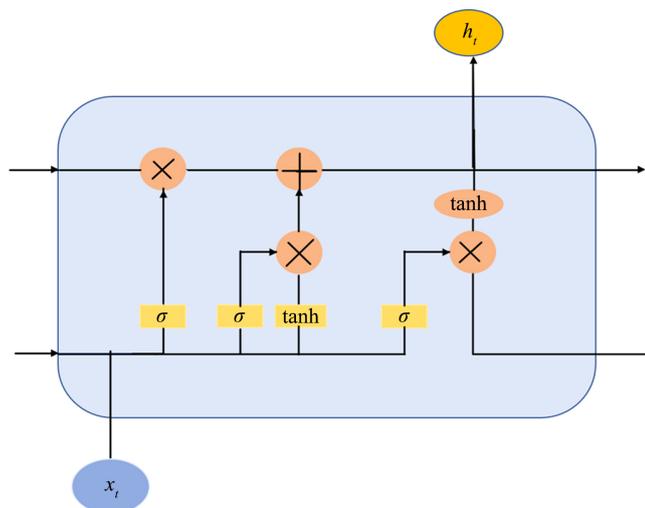


Figure 1. Basic unit structure of LSTM  
图 1. LSTM 基本单元结构图

LSTM 模型包含三个门结构：

- 1) 输入门：决定在  $t$  时刻是否有信息输入到记忆细胞。
- 2) 遗忘门：决定将记忆细胞中的值遗忘掉还是保留下来。
- 3) 输出门：决定信息是否要从记忆细胞中输出。

LSTM 的这种门控结构，能够更好适应于时间序列数据的处理，有效地解决长短时间序列变化上的问题。

LSTM 的计算公式如下：

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f)$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c)$$

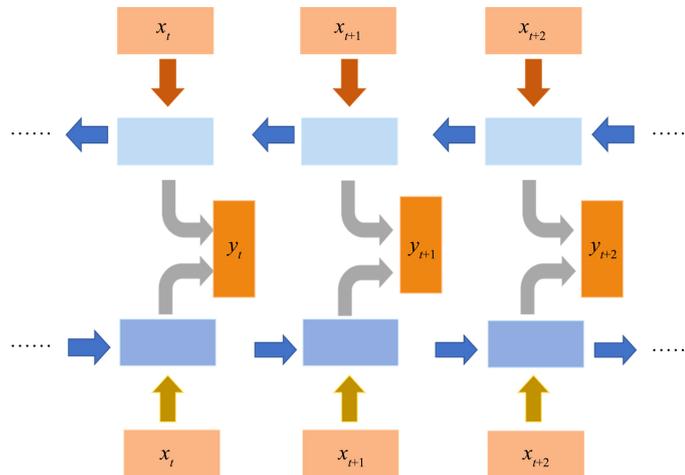
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

$$O_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = O_t * \tanh C_t$$

其中  $\sigma(\cdot)$  和  $\tanh(\cdot)$  为激活函数，计算公式分别为： $\sigma(x) = \frac{1}{1+e^{-x}}$ ， $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ ， $x_t$  为  $t$  时刻的输入， $C_t$  为  $t$  时刻的细胞状态， $\tilde{C}_t$  为  $t$  时刻的临时细胞状态， $h_t$  为隐藏层的状态， $f_t$ 、 $i_t$ 、 $O_t$  为三个门控。

由于 LSTM 结构只能通过  $t$  时刻之前的信息来对  $t$  时刻的信息进行推断和预测，因此无法获得更多的相关信息来进行分析。为了提高预测的准确度，本文引入了 Bi-LSTM 模型来对股票价格进行预测。对  $t$  时刻前后的股票价格信息进行双向读取，使用前向及后向的历史数据来共同预测  $t$  时刻的股票收盘价。Bi-LSTM 的基本结构如图 2 所示：



**Figure 2.** The basic structure of Bi-LSTM  
**图 2.** Bi-LSTM 的基本结构

## 2.2. Adam 优化算法

自适应矩估计(Adam)算法是将动量(momentum)和均方根传播(RMSprop)结合在一起的一种优化算法,通过对偏置进行矫正来使得参数更加平稳,使用 Adam 优化算法进行优化的步骤如表 1 所示:

**Table 1.** Adam optimization algorithm to optimize the steps  
**表 1.** Adam 优化算法进行优化的步骤

将参数初始化: $m_t = 0, n_t = 0$
计算 $t$ 时刻的梯度: $g_t = \nabla \theta f_t(\theta_{t-1})$
更新参数: $m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$ (momentum 项) $n_t = \beta_2 n_{t-1} + (1 - \beta_2) g_t^2$ (RMSprop 项)
计算修正偏差: $\hat{m}_t = m_t / (1 - \beta_1^t)$ $\hat{n}_t = n_t / (1 - \beta_2^t)$
更新权重: $\theta_{t+1} = \theta_t - \alpha * \hat{m}_t / (\sqrt{\hat{n}_t} + \epsilon)$

其中  $m_t$  为一阶动量项,  $n_t$  为二阶动量项, 均初始化为 0,  $\hat{m}_t, \hat{n}_t$  分别为修正后的值,  $\theta_{t+1}$  为第  $t+1$  次迭代模型的参数,  $\alpha$  为学习率。在研究过程中,  $\beta_1, \beta_2, \epsilon$  的取值分别为 0.9, 0.999,  $1 \times 10^{-8}$ 。

## 3. 性能评价指标

股票价格预测模型通常采用均方误差(MSE)、均方根误差(RMSE)、平均绝对误差(MAE)、 $R^2$  来对预测的精确度进行评价。MSE、RMSE、MAE 的值越小, 模型的预测精度越高。其中  $R^2$  表示决定系数, 范围在 0~1 之间,  $R^2$  的值越接近于 1 表示真实值与预测值之间的相关性越高。这些误差评估方法的统计指标定义如下:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (5)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (6)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (7)$$

## 4. 算例分析

### 4.1. 基于 Adam 优化算法的 Bi-LSTM 风功率预测模型构建

Bi-LSTM 是信息可以双向传播的 LSTM 结构, 本文采用前 9 天股票的开盘价、最高价、最低价、收盘价、成交金额作为模型的输入, 基于 Bi-LSTM 神经网络来对第 10 天的收盘价进行预测。

下面描述整个过程的详细步骤:

1) 将选取的数据分成训练集和测试集。在建立 Bi-LSTM 模型之前需要先对数据进行归一化处理, 归一化公式:  $data = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$ , 通过数据归一化可以消除不同量级对预测结果产生的影响。

2) 将归一化后的数据处理为可以批量运算的矩阵来训练 Bi-LSTM 神经网络, 确定模型参数, 对测试集进行预测, 将预测结果进行反归一化处理后与真实收盘价进行比较。

3) 采用多种误差评估方法来对模型的预测效果进行度量。并将 Bi-LSTM 模型的预测误差与 BP、LSTM、GRU 模型的预测误差进行比较。

### 4.2. 实验数据

本文研究所提出的模型及误差评定方法应用建行 2011 年 1 月 4 日~2021 年 5 月 25 日共 2522 条数据。其中包括五个不同的变量, 分别为开盘价、最高价、最低价、收盘价、总金额。首先对数据进行归一化处理, 将前 2018 条数据作为训练集, 将后 504 条数据作为测试集。

### 4.3. 数据描述

将建行数据的收盘价做折线图, 如图 3 所示。

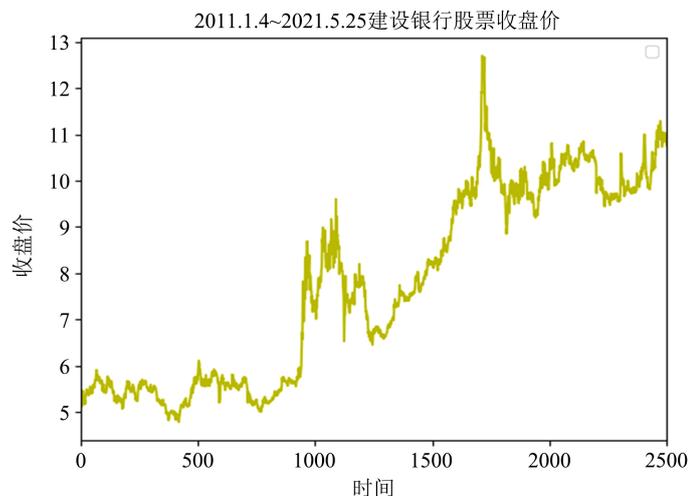


Figure 3. Line chart of Construction Bank stock  
图 3. 建设银行股票收盘价折线图

## 5. 参数设计

本节对本文提出的方法及 BP、LSTM、GRU、Bi-LSTM 四个不同模型进行了研究。所有的模型的实现都是基于 python 编程，其中 BP、LSTM、GRU 基于 pytorch 库实现。将 Bi-LSTM 模型的学习率设置为 0.001，batch\_size 设置为 64，本文通过前九天的股票价格来预测第十天的股票收盘价，将学习步长设置为 9，股票数据有 5 个不同的变量，将 Bi-LSTM 模型的 input\_size 设置为了 5，hidden\_size 设置为 512，output\_size 设置为 1，经过多次重复实验，将 Bi-LSTM 模型的 num\_epochs 设置为 20。

## 6. 实验结果分析

对本文数据分别进行 BP、LSTM、GRU、Bi-LSTM 建模，图 4 为预测值与真实值的对比图。可以看出基于 Bi-LSTM 预测模型的预测值与真实值近似重合。

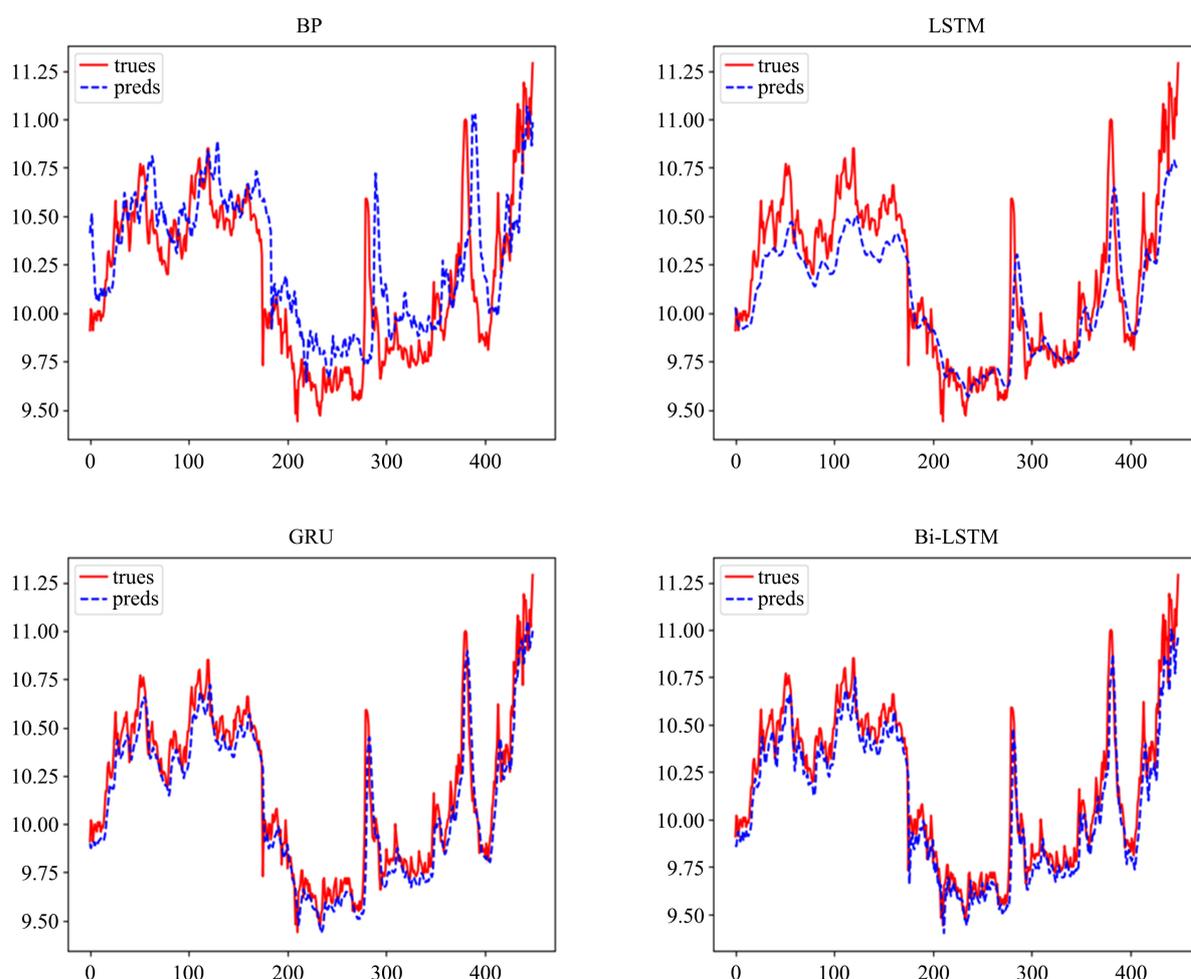
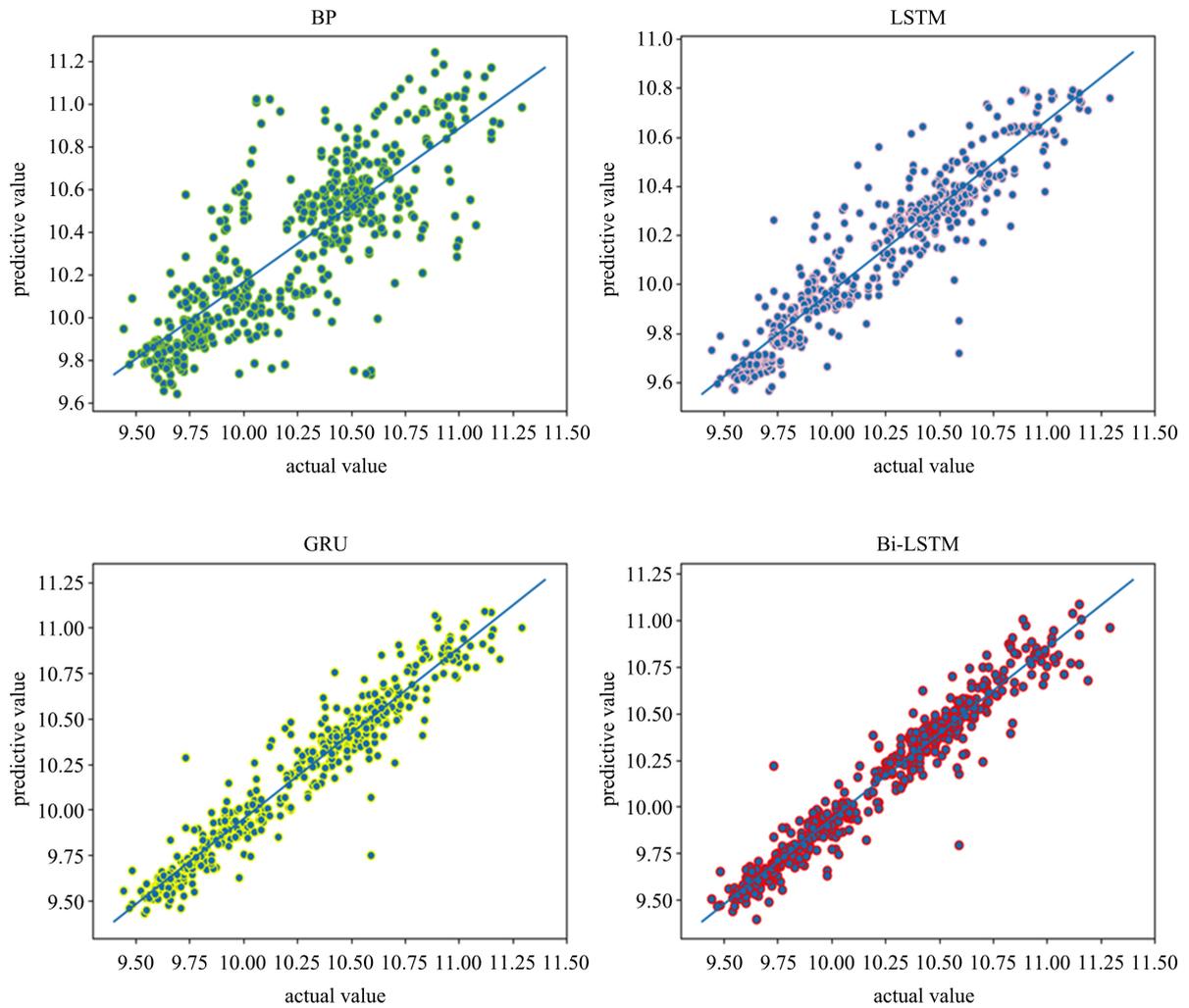


Figure 4. Comparison of predicted values and real values of different models

图 4. 不同模型预测值与真实值对比图

将四个模型的预测值与真实值进行线性拟合做出散点图，如图 5 所示。从拟合散点图中可以看出，本文所提方法预测值与真实值非常接近。回归方程的各参数如表 2 所示，其中 a 表示回归系数，b 表示截距， $R^2$  表示决定系数。

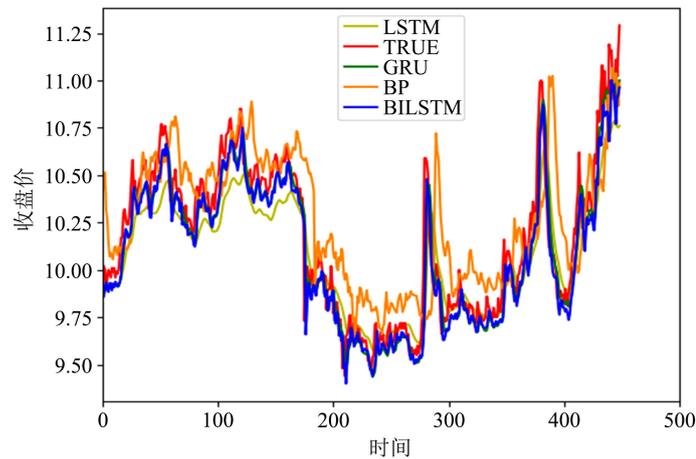


**Figure 5.** Fitting result graph of real value and predicted value  
**图 5.** 真实值与预测值拟合结果图

**Table 2.** Regression equation parameter table  
**表 2.** 回归方程参数表

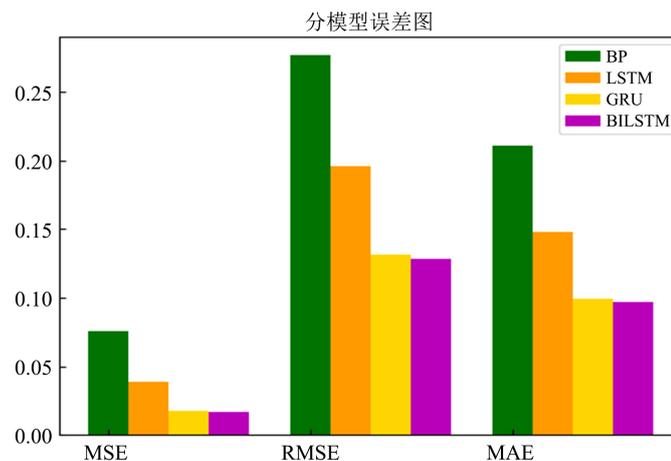
	BP	LSTM	GRU	Bi-LSTM
a	0.6958	0.7169	0.9145	0.9376
b	3.0125	2.9956	0.7887	0.5723
$R^2$	0.5286	0.7716	0.7813	0.8988

为验证 Bi-LSTM 模型的准确度，本部分，将本文所提出的预测方法与 BP、LSTM、GRU 这三个模型进行比较，将预测结果进行对比分析。图 6 为本文所提模型与三个传统模型的预测结果，从图中，可以看出基于 BP 模型预测出来的结果与真实值的偏差最大，其余模型的预测值与真实值较为接近，其中 Bi-LSTM 模型预测的结果与真实值最为接近，本文所提出的 Bi-LSTM 预测模型能更好地对股票的收盘价进行预测。



**Figure 6.** Comparison of predicted and true values of different models  
**图 6.** 不同模型的预测值与真实值对比图

为了进一步研究所提方法在不同时期的表现, 对不同模型预测性能进行评估, 分别使用不同的模型误差评估方法 MSE、RMSE、MAE 对四个模型的预测结果进行分析。表 3 和图 7 展示了误差结果, 误差评估方法的值越小, 模型的预测效果越好。由结果可以看出, 本文所提模型 Bi-LSTM 的 MSE、RMSE、MAE 的值分别为 0.0166、0.1282、0.0969, 均优于模型 BP、LSTM、GRU 的结果。说明本文所提模型的预测精度最为准确。



**Figure 7.** Error graph of different models  
**图 7.** 分模型误差图

**Table 3.** Model prediction error table  
**表 3.** 模型预测误差表

Models	MSE	RMSE	MAE
BP	0.0753	0.2767	0.2110
LSTM	0.0387	0.1959	0.1481
GRU	0.0173	0.1313	0.0993
Bi-LSTM	0.0166	0.1282	0.0969

## 7. 结语

股票价格的预测在金融市场占据重要的地位,本文基于 Adam 优化算法的 Bi-LSTM 模型来对股票的收盘价进行预测。引入几种经典模型(BP、LSTM、GRU)来对本文预测模型结果进行分析,通过 MSE、RMSE、MAE 误差评估方法来对所有模型进行评估,误差评估方法 MSE 的结果表明,本文所提模型与 BP、LSTM、GRU 相比,精确度分别提高了 77.95%、57.11%、2.36%。由于本文输入数据比较单一,可能会使得模型预测存在一定的局限性。下一步将通过论坛、贴吧等网站提取用户对股票的评论、新闻报道等相关信息,并对这些信息进行处理后加入到股票价格预测模型中,从多方面来对股票数据进行分析进而提高模型的预测精度。

## 参考文献

- [1] 许舒雅,梁晓莹. 基于 ARIMA-GARCH 模型的股票价格预测研究[J]. 河南教育学院学报(自然科学版), 2019, 28(4): 20-24.
- [2] 张倩倩,林天华,祁旭阳,赵霞. 基于机器学习的股票预测研究综述[J]. 河北省科学院学报, 2020, 37(4): 15-21.
- [3] 毛月月,张秋悦. 基于机器学习的股票预测方法研究[J]. 现代计算机, 2020(23): 44-47.
- [4] 张潇,韦增欣. 随机森林在股票趋势预测中的应用[J]. 中国管理信息化, 2018, 21(3): 120-123.
- [5] 张晶华,莫文柯,甘宇健. 基于 SVM 的上证指数预测研究[J]. 软件导刊, 2017, 16(8): 156-159+163.
- [6] 崔文喆,李宝毅,于德胜. 基于 GARCH 模型和 BP 神经网络模型的股票价格预测实证分析[J]. 天津师范大学学报(自然科学版), 2019, 39(5): 30-34.
- [7] Bahadur, S.T., et al. (2020) Stock Price Forecasting with Deep Learning: A Comparative Study. *Mathematics*, 8, 1441. <https://doi.org/10.3390/math8091441>
- [8] 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(6): 36-42.
- [9] 矫丰霞. 基于 LSTM 神经网络的股票价格预测研究[J]. 数字技术与应用, 2021, 39(3): 220-222.
- [10] 陈伟斌,林奕真,王宗跃. 股票信息挖掘与 LSTM 预测[J]. 集美大学学报(自然科学版), 2020, 25(5): 385-391.
- [11] 曾安,聂文俊. 基于深度双向 LSTM 的股票推荐系统[J]. 计算机科学, 2019, 46(10): 84-89.