

基于IndRNN-LSTM模型的股票价格预测

扈文

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2021年12月17日; 录用日期: 2022年1月6日; 发布日期: 2022年1月19日

摘要

结合独立循环神经网络和长短期记忆网络建立IndRNN-LSTM模型, 选取道琼斯指数的18个指标, 对道琼斯指数的开盘价格进行预测。另外, 分别采用CNN-LSTM、IndRNN、LSTM、SVM、BP和CNN神经网络模型对开盘价格进行预测, 并将七种模型的预测结果进行比较, 结果表明IndRNN-LSTM模型的预测精度较高, 能更好地预测股票的走势。

关键词

独立循环神经网络, 长短期记忆网络, 道琼斯指数

Stock Price Prediction Based on Independently Recurrent Neural Network and Long Short-Term Memory Network

Wen Hu

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: Dec. 17th, 2021; accepted: Jan. 6th, 2022; published: Jan. 19th, 2022

Abstract

IndRNN-LSTM model is established by combining independently recurrent neural network and long short-term memory network. This paper selects 18 indexes of Dow Jones index to predict the opening price of Dow Jones index. CNN-LSTM, IndRNN, SVM and LSTM neural network models are also used to predict the opening price, and the prediction results of the five models are compared. The results show that the IndRNN-LSTM model has high prediction accuracy and can better predict the stock trend.

Keywords

Independently Recurrent Neural Network, Long Short-Term Memory Network, Dow Jones Industrial Average

Copyright © 2022 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

随着业务创新的全面铺开，股票市场深入推进，股票行业进入历史上最好的发展机遇期。因此人们越来越关注股票的走势，希望可以提高收益，降低风险。随着神经网络的发展，学者们开始尝试使用神经网络对股票价格进行预测。长期以来，许多学者对股票价格的变化进行了大量的研究。

周阳[1]首先建立了三因素预测模型，即交易信息，投资者关注与投资者情绪，货币政策这三个因素。为确保数据的真实性和全面性，选取了一系列相关数据作为这三个因素的代理指标，并利用主成分分析方法构建了相应的三个综合指数，然后建立 3 个长短期记忆网络模型进行预测，结果表明基于所有原始指标建立长短期记忆网络模型预测效果相对更优，也就是说 LSTM 在多变量股票价格预测方面有很好的效果。赵浩博和李锡祚[2]提出将股票数据和相关新闻相结合，使用卷积和循环神经网络的组合模型对股票开盘价进行预测。首先将道琼斯指数的相关新闻数据转换成向量，使用卷积神经网络对相关新闻进行特征提取，利用循环神经网络对股票数据进行训练，将两者的结果进行合并，并输入到深度神经网络中，建立深度神经网络模型对开盘价进行预测，结果表明此方法对股票的预测有强大的作用。钱赟[3]选取上证综合指数和深证综合指数，首先用主成分分析法对数据做降维处理，然后建立序列到序列(Seq2Seq)模型，对股票的收盘价格进行短期趋势预测，并且另外选取六只股票做普适性分析，结果表明，Seq2Seq 模型在股票短期趋势预测方面有较好的泛化能力。王东和王霄鹏[4]使用主成分分析法将股票的基础数据进行降维，再将股票相关技术指标 KDJ、MACD 与主成分分析降维后的数据结合，一同作为输入数据，建立 PCA-S-LSTM 模型，并进行预测。实验结果表明：PCA-S-LSTM 模型不仅减少了运行时间，且更好地预测了平安银行的收盘价。黄超斌和程希明[5]利用长短期记忆网络模型对七千多条上证综合指数进行预测分析，并将其预测结果与 GRU、RNN、CNN 和 BP 神经网络模型的预测结果进行对比，结果表明，长短期记忆网络模型有更好的性能和更高的预测精度。

2. LSTM 网络和 IndRNN 网络

2.1. LSTM 网络

递归神经网络(RNN)已被广泛应用于解决时间序列问题中，比如投资证券的管理、语音的识别、股票市场的预测等。与卷积神经网络(CNN)等前馈网络相比，RNN 具有递归连接，其中最后一个隐藏状态是对下一个状态的输入，更新方式如下：

$$h_t = \sigma(Wx_t + Uh_{t-1} + b) \quad (1)$$

其中 x_t 和 h_t 分别是时间步 t 处的输入和隐藏状态， W 是当前输入的权重， U 是循环输入的权重， b 是偏置向量， σ 是激活函数。

Hochreiter 和 Schmidhuber 提出了长短期记忆网络(LSTM)，LSTM 是在 RNN 神经网络的基础上改进而形成的。主要在 RNN 的循环结构中加入了门控结构，即输入门，遗忘门和输出门，它们精细运算神经元内部，解决长时间的信息依赖问题，以达到长时记忆的效果[6]。

LSTM 的内部门控结构如下：

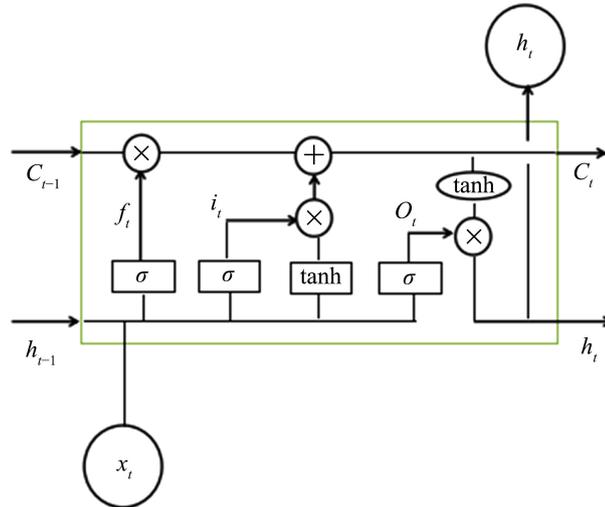


Figure 1. Gating structure of LSTM
图 1. LSTM 的门控结构

LSTM 最重要的结构是单元状态，它的运行类似一条传送带，如图 1 中横向穿过的一条直线。LSTM 中的遗忘门，输入门和输出门分别对应图 1 中 f_t ， i_t 和 O_t ，在运行的过程中 LSTM 会受到三个门控制的影响。

上一时刻的输出 h_{t-1} 与当前时刻的输入 x_t 会进入到遗忘门 f_t 中，由遗忘门来决定哪些信息需要丢弃，主要计算方法为输出的值与上一时刻的记忆进行相乘运算实现遗忘功能，计算公式如下：

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

遗忘门由 sigmoid 层决定，如图 1 中 σ 层，它根据上一时刻的输出和当前输入来为单元状态中的数字输出一个 [0,1] 的数字，1 表示完全保留，0 表示完全抛弃。 C_{t-1} 是上一时刻的状态值，它会随着时间的变化不断更新。

接着输入门 i_t 决定单元状态需要储存哪些信息，它由两个步骤完成，首先由 sigmoid 层决定哪些信息需要更新，其次在 tanh 层输入新的候选值，其计算公式如下：

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

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

接下来通过遗忘门和输入门选择的信息进行更新，得到下一时刻的新信息。计算公式如下：

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

最后输出门 O_t 先经过 sigmoid 层决定哪些信息可以输出，其计算公式如下：

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

接下来将单元状态中的信息输入到 \tanh 函数中与 sigmoid 的门限值相乘，最终得到输出的信息，计算公式如(7)所示。

$$h_t = o_t * \tanh(C_t) \tag{7}$$

2.2. IndRNN 网络

通过对 RNN 进行改造后，递归神经网络可以实现良好的记忆能力，但是构建一个深度的 LSTM 或 GRU 的递归神经网络实际上是困难的，所以仍然存在梯度消失和爆炸问题。

2018 年澳大利亚伍伦贡大学联合中国科技大学提出一种新型的循环神经网络——独立循环神经网络 (Independently Recurrent Neural Network, IndRNN)。IndRNN 可以调节时间梯度反向传播，有效地解决梯度消失和爆炸问题，并且可以用来处理长期序列。而且可以有效地堆叠多层 IndRNN，构建比现有 RNN 更深的网络。IndRNN 与传统的 RNN 网络相比最大的差别在于，IndRNN 中使用阿达玛积处理更新状态，其更新方式如下：

$$h_t = \sigma(Wx_t + U \odot h_{t-1} + b_n) \tag{8}$$

其中 \odot 表示阿达玛积。IndRNN 中一层中每个神经元都独立的，神经元之间的连接可以通过叠加两个或更多层的 IndRNN 来实现。IndRNN 的基本结构[7]如下图 2：

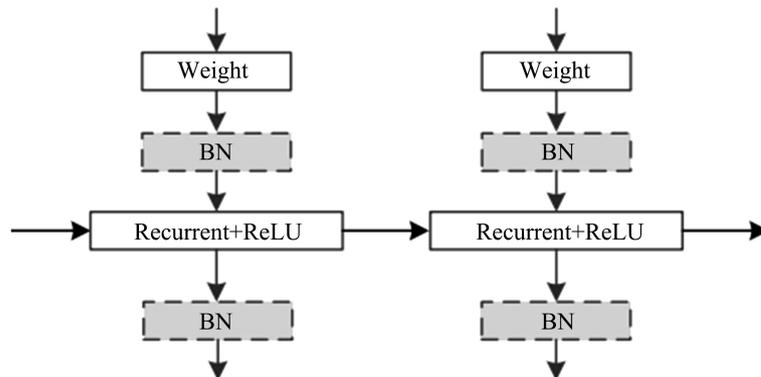


Figure 2. Structure of IndRNN
图 2. IndRNN 的基本结构

3. 实证分析

3.1. 数据选取

文章标题中尽量避免使用生僻的英文缩写。本文使用 Python3.7 编程环境，主要使用 Keras 包实现 LSTM 和 IndRNN 的构建。数据选取道琼斯指数 2010 年 7 月 9 日至 2021 年 6 月 22 日共 2752 个交易日的数据。分别用开盘价、收盘价、最高价、最低价、成交量、MA.MA1、MA.MA2、MA.MA3、MA.MA4、MA.MA5、MA.MA6、KDJ.K、KDJ.D、KDJ.J、MACD.DIFF、MACD.DEA、MACD.MACD、CCI.CCI 等 18 个指标预测股票开盘价。将数据按照 7:3 的比例进行划分，其中前 70% 的数据作为训练集，后 30% 的数据作为测试集，预测窗口大小为 10，即使用前 10 个历史数据预测未来一天的股票开盘价。道琼斯指数的开盘价格如图 3 所示。

考虑到 IndRNN 模型和 LSTM 模型与其它模型相比具有较长的时间记忆能力，而且能较好地拟合股票开盘价格与各影响因素之间的非线性特征，因此采用 IndRNN-LSTM 模型进行道琼斯指数开盘价格预测。

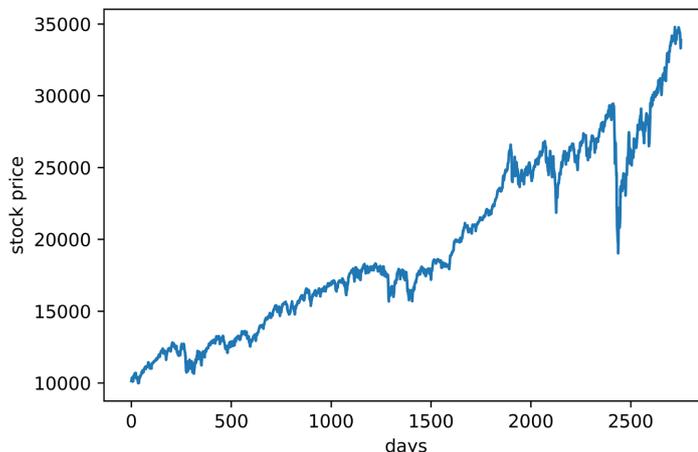


Figure 3. Opening price of Dow Jones index
图 3. 道琼斯指数开盘价

3.2. 数据处理

由于数据之间的量纲不同，所以首先将数据进行标准化处理，消除量纲的影响，本文采用最大最小值(minmax)标准化方法，其计算方法为 $x' = \frac{x - x_{\max}}{x_{\max} - x_{\min}}$ ，其中 x' 为标准化的数据， x 为原始数据， x_{\max} 为原始数据中的最大值， x_{\min} 为原始数据中的最小值。

3.3. 建立 IndRNN-LSTM 模型

经过多次实验本文选择构建 4 层神经网络，前 3 层为 IndRNN，每层有 512 个神经元。第 4 层为 LSTM，有 256 个神经元，每层神经网络使用 ReLU 作为激活函数。第 5 层为 Dropout 层，防止过拟合。最后一层为 Dense 层。将神经网络的训练损失函数设置为平方绝对误差(mean absolute error, MAE)。数据按照 7:3 划分训练集和测试集后，前 1915 个数据作为训练样本，后 815 个数据作为测试样本。考虑到 Adam 优化器计算高效，非常适合大规模的数据，所以 IndRNN-LSTM 模型将选择 Adam 作为优化器，并且设置迭代次数为 1000 次，预测结果如图 4 所示。

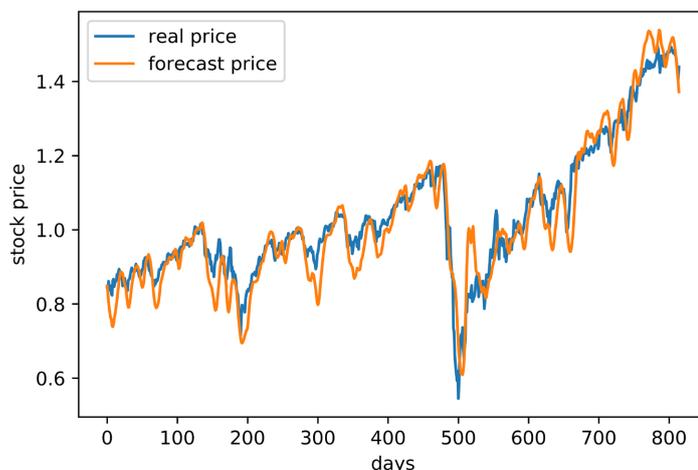


Figure 4. Prediction results of IndRNN-LSTM
图 4. IndRNN-LSTM 预测结果

从以上预测结果可以看出，从始至终，IndRNN-LSTM 模型都保持较好的预测效果，能精准的学习出股票的涨跌趋势，对价格拐点反应较灵敏，预测误差较小，证明了 IndRNN-LSTM 模型在股票开盘价格预测上的适用性。

3.4. 建立对比模型

为了验证 IndRNN-LSTM 模型的效果，选取了 IndRNN 模型、LSTM 模型、CNN-LSTM 组合模型和 SVM 模型进行预测并与 IndRNN-LSTM 模型对比。IndRNN 模型建立 3 层，每层有 512 个神经元，仍然选用 Adam 为优化器，ReLU 为激活函数，迭代次数为 1000 次，预测结果如图 5 所示。

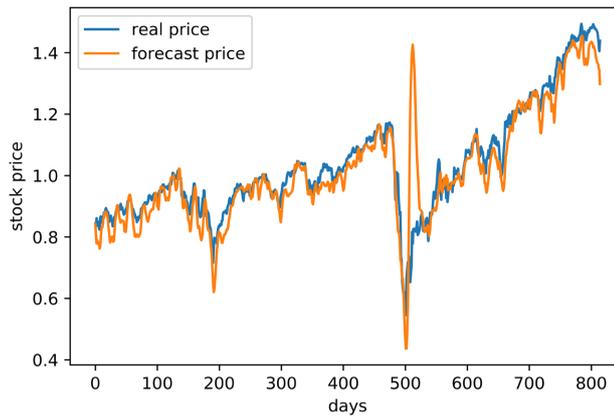


Figure 5. Prediction results of IndRNN
图 5. IndRNN 预测结果

从图 5 中可以发现，IndRNN 模型的预测误差较小，在大多数情况下能较好地对股票开盘价格进行预测，但是在第 500 个数据处，可以发现 IndRNN 模型的预测趋势与实际价格走势相差很大，在实际值下降时，IndRNN 模型还能较好的学习趋势，但当实际值上升后，预测值上升的更高，远远大于实际值，因此 IndRNN 模型对股票开盘价的预测走势不如 IndRNN-LSTM 模型好。

LSTM 模型建立 3 层，每层有 512 个神经元，选用 Adam 为优化器，ReLU 为激活函数，迭代次数为 1000 次，预测结果如图 6 所示。

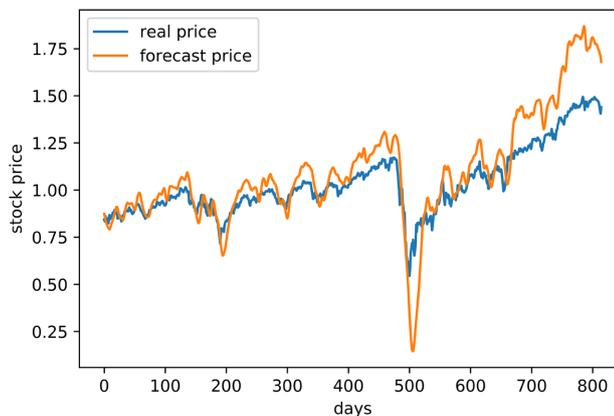


Figure 6. Prediction results of LSTM
图 6. LSTM 预测结果

从以上预测结果可以看出,随着时间推移,LSTM模型的预测误差在慢慢变大,在第500个数据处,由于LSTM模型不能有效提取信息,所以在股票开盘价已经开始上涨的时候,模型的预测价格仍然在下降,且从700以后预测值明显高于实际值,导致预测误差增大,所以LSTM模型的预测准确率明显不如IndRNN-LSTM模型。

CNN模型中第一层为卷积层,滤波器大小为32,卷积核大小为3,步长为1。第二层为最大池化层,池化窗口为1。另外再分别添加一层卷积层和一层池化层,且参数与前两层相同。在池化层后再添加一层压平层,将多维输入一维化,迭代次数同样为1000。预测结果如图7所示。

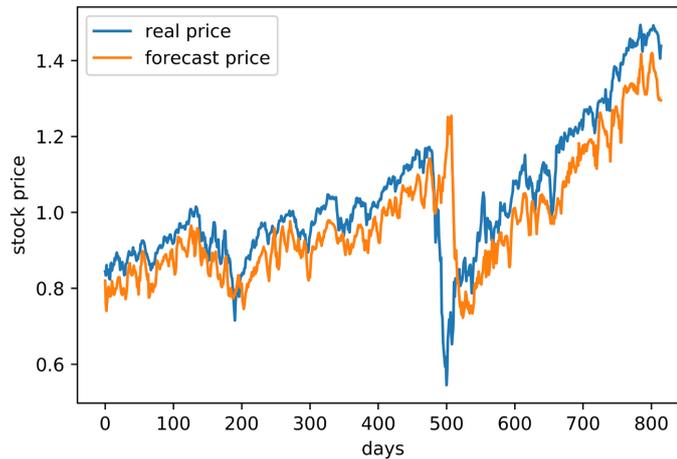


Figure 7. Prediction results of CNN
图 7. CNN 预测结果

从以上结果可以看出预测值大多数情况下都比实际值偏小,且在第500个数据处真实值在下降,而预测值却在上升,所以CNN模型的预测误差较大。

为了更好地对比试验效果,CNN-LSTM模型中的卷积神经网络与CNN模型参数相同,滤波器大小为32,步长为1。后三层为LSTM,每层有512个神经元,选用Adam为优化器,ReLU为激活函数,迭代次数为1000次,预测结果如图8所示。

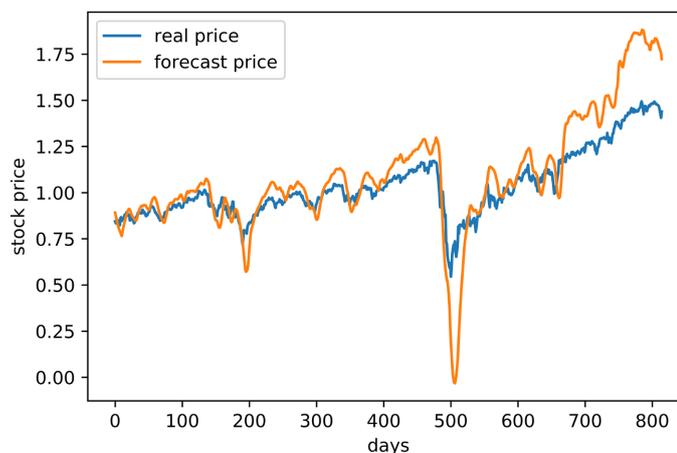


Figure 8. Prediction results of CNN-LSTM
图 8. CNN-LSTM 预测结果

通过比较图 6 与图 8 的预测结果可以发现, CNN-LSTM 模型与 LSTM 模型的预测趋势相似, 但是 CNN-LSTM 模型的预测误差稍大一些, 从图中可以看出, CNN-LSTM 模型在第 500 个数据处的预测误差远大于 IndRNN 模型在第 500 个数据处的预测误差。

SVM 模型中参数 C 设置为 1, 核函数选择 RBF 函数, 预测结果如图 9 所示。

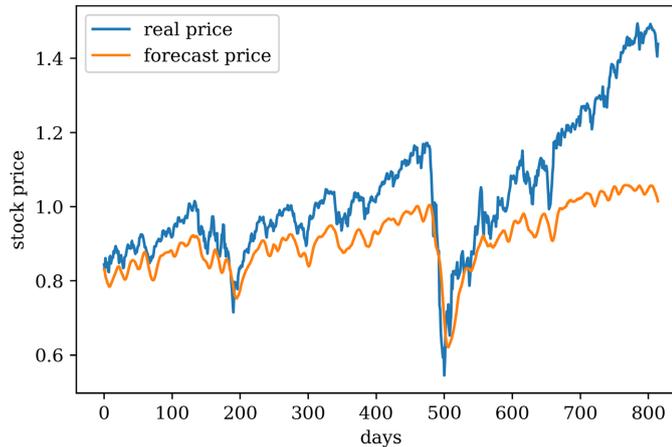


Figure 9. Prediction results of SVM
图 9. SVM 预测结果

从以上预测结果中可以看到使用 SVM 模型预测的效果远不如其他模型的预测效果好, 预测值明显低于实际值, 且随着时间推移, 预测效果越来越差。与神经网络相比, SVM 没有特征提取的过程, 无法发现数据间的深层规律, 这使得 SVM 在道琼斯指数开盘价预测的表现上远不如 IndRNN-LSTM 模型。

BP 模型的预测结果如下, 从整体上看 BP 模型的预测误差较大, 在每一时刻, 预测值都低于实际值, 且在从第 700 个数据开始预测值远远低于实际值。预测结果如图 10 所示。

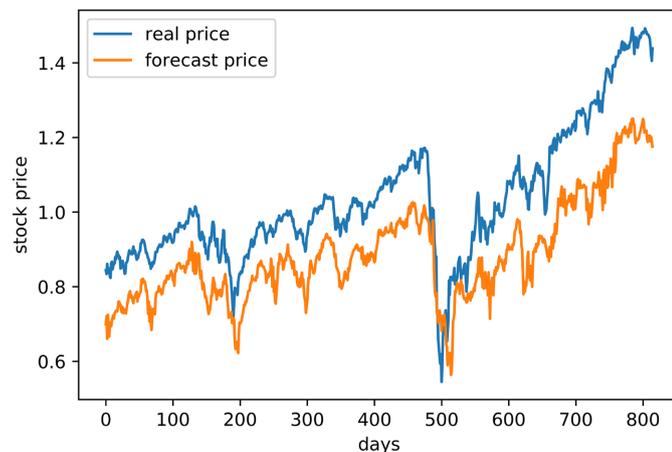


Figure 10. Prediction results of BP
图 10. BP 预测结果

3.5. 评价指标

为了更精准的分析预测的准确度, 本文采用均方根误差(RMSE)、均方误差(MSE)、平均绝对误差(MAE)

对预测结果进行评价。均方误差反映的是预测值与真实值之间差异程度的期望值，当它们完全吻合时均方误差为 0，此时为完美模型，如果误差越大，该值越大，计算公式如下：

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2 \tag{9}$$

均方根误差是用来衡量观测值与真实值之间的偏差，误差越大，该值越大，计算公式如下：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2} \tag{10}$$

平均绝对误差能更好地反映预测值误差的实际情况，误差越大，该值越大，计算公式如下：

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{X}_i - X_i| \tag{11}$$

公式(9)~(11)中 \hat{X}_i 表示预测值， X_i 表示真实值。

Table 1. Prediction error
表 1. 预测误差

	IndRNN-LSTM	IndRNN	CNN-LSTM	LSTM	SVM	BP	CNN
MSE	0.003182	0.006520	0.026705	0.020482	0.025673	0.024997	0.011964
RMSE	0.056408	0.080746	0.163417	0.143115	0.160228	0.158104	0.109379
MAE	0.202241	0.210390	0.331746	0.318163	0.352225	0.384277	0.289082

由表 1 可知，各模型的误差走势与图 3~7 中的预测价格曲线与实际价格曲线之间的误差走势是一致的。在 6 个对比模型中 IndRNN 模型的预测效果相对较好一些，LSTM 模型与 CNN 模型的预测价格能力次之，SVM、BP 与 CNN-LSTM 模型的预测价格能力较差。而 IndRNN-LSTM 模型与其他模型相比预测误差显著下降，且由图 4 可以直观的看出预测价格与真实价格的曲线较接近，其预测精度最高。

SVM 的预测效果较差是由于 SVM 需要借助特征提取算法才能有效地对数据进行降维或信息提取，如果无法找到合适的特征提取算法，那么就很难有好的预测性能。而 BP 神经网络和卷积神经网络的预测结果都小于实际值，由于 BP 和 CNN 没有长期记忆能力，无法记忆历史趋势，所以导致预测结果较差。但是 BP 神经网络与其他神经网络相比运算速度较快。相比之下，IndRNN、LSTM 和 CNN-LSTM 可以对股票数据进行特征提取且神经网络由较强的学习能力，因此神经网络相比于 SVM 模型有更高的预测精度。IndRNN-LSTM 的预测精度显著高于其它模型可能有以下原因，首先 IndRNN-LSTM 模型有较好的特征提取能力；其次 IndRNN-LSTM 模型有较好的长期记忆能力；并且使用 IndRNN-LSTM 组合模型对股票开盘价格进行预测可以将预测任务进行分解，利用 IndRNN 模型的长期记忆性对股票数据的历史走势进行分析，并提取有效信息。然后通过 LSTM 模型对未来股票的开盘价格进行预测，取得了较好的效果。通过上述实证分析，基于 IndRNN-LSTM 的神经网络适用于道琼斯指数建模分析。

4. 总结

本文构建了 IndRNN-LSTM 模型，该模型利用 IndRNN 与 LSTM 神经网络的长时间记忆性对股票价格进行预测，从预测结果可以发现，IndRNN-LSTM 模型能够有效提取股指数据中的有效信息，能挖掘其潜在的特征，较好地学习数据间的非线性关系，因此 IndRNN-LSTM 模型能较好地预测出股票开盘价的

走势。另外,本文还建立了 CNN-LSTM、LSTM、SVM、BP、CNN 模型,进行对比实验。在对比试验中发现 IndRNN-LSTM 模型对于道琼斯指数开盘价预测精度更高,且模型在开盘价格波动较大时仍有较好的拟合能力,可以更准确地预测股票的开盘价格,从总体上来说,本文构建的 IndRNN-LSTM 模型对投资者进行股票开盘价预测有一定的参考作用。

另外此模型还有一些不足,股票的涨跌除了与交易数据和技术指标有关,还受到政府政策、自然灾害等偶发因素的影响,而这些影响因素复杂多变,未能体现在模型中。同时由于模型中缺少文本类数据,导致无法对道琼斯指数开盘价格中的重要转折点进行解释。因此,在未来的研究中不仅要关注股票的交易数据和技术指标,还可将历史重大事件作为模型的输入,对其进行特征提取,从而提高 IndRNN-LSTM 模型的预测精度。

参考文献

- [1] 周阳. 基于 LSTM 模型的上证综指价格预测研究[D]: [硕士学位论文]. 南京: 南京邮电大学, 2019.
- [2] 赵浩博, 李锡祚. 基于卷积和循环神经网络模型融合的股票开盘价预测研究[J]. 智能计算机与应用, 2019, 9(6): 55-58+64.
- [3] 钱赟. 基于主成分分析与序列到序列学习模型的股票收盘价格预测[D]: [硕士学位论文]. 上海: 上海师范大学, 2020.
- [4] 王东, 王霄鹏, 杨川东. 一种基于主成分 LSTM 模型在股票预测中的研究[J]. 重庆理工大学学报(自然科学), 2021, 35(2): 282-288.
- [5] 黄超斌, 程希明. 基于 LSTM 神经网络的股票价格预测研究[J]. 北京信息科技大学学报(自然科学版), 2021, 36(1): 79-83.
- [6] 王玉堃. LSTM 和 GRU 神经网络在股指高频数据预测中的研究[D]: [硕士学位论文]. 济南: 山东大学, 2020.
- [7] Li, S., Li, W., Cook, C., *et al.* (2018) Independently Recurrent Neural Network (IndRNN): Building A Longer and Deeper RNN. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 5457-5466. <https://doi.org/10.1109/CVPR.2018.00572>