Operations Research and Fuzziology 运筹与模糊学, 2021, 11(2), 137-146
Published Online May 2021 in Hans. http://www.hanspub.org/journal/orf
https://doi.org/10.12677/orf.2021.112017

Hans汉斯

基于PyTorch框架LSTM深度学习股票预测系统

李泽艳[1]，陈银钧[2]

[1]贵州大学数学与统计学院，贵州 贵阳
[2]重庆大学数学与统计学院，重庆
Email: 1413460302@qq.com

收稿日期：2021年4月2日；录用日期：2021年5月3日；发布日期：2021年5月11日

摘 要

随着机器学习与深度学习的发展，传统的时间序列模型已经不能满足人们对于股票预测准确性的要求。因此，本文引入深度学习中基于PyTorch框架的LSTM循环神经网络模型对创业300指数的收盘价进行预测，通过设置迭代次数、遗忘门偏置值以及LSTM单元数，对比模型的预测误差。研究结果表明，迭代次数为200、LSTM单元数为2、遗忘门偏置值为0.4的LSTM模型对创业300指数收盘价走势的拟合误差最小，平均绝对百分比误差达到0.0109，为进一步使用PyTorch框架构建循环神经网络准确预测股价提供了依据。

关键词

股票预测，PyTorch，LSTM，创业300指数

LSTM Deep Learning Stock Prediction System Based on PyTorch Framework

Zeyan Li[1], Yinjun Chen[2]

[1]The College of Mathematics and Statistics, Guizhou University, Guiyang Guizhou
[2]The College of Mathematics and Statistics, Chongqing University, Chongqing
Email: 1413460302@qq.com

Received: Apr. 2nd, 2021; accepted: May 3rd, 2021; published: May 11th, 2021

Abstract

With the development of machine learning and deep learning, the traditional time series model

文章引用: 李泽艳, 陈银钧. 基于 PyTorch 框架 LSTM 深度学习股票预测系统[J]. 运筹与模糊学, 2021, 11(2): 137-146.
DOI: 10.12677/orf.2021.112017

has been unable to meet people's requirements for the accuracy of stock forecast. Therefore, this paper introduced the LSTM recurrent neural network model based on PyTorch framework in deep learning to predict the closing price of Entrepreneurship 300 Index, and compared the forecast error of the model by setting the number of iterations, the value of the forgetting gate bias and LSTM unit numbers. The research shows that the LSTM model with 200 iterations, 2 LSTM units and 0.4 forgetting gate bias value has the smallest fitting error for the closing price trend of Entrepreneurship 300 Index, and average absolute percentage error reaches 0.0109, providing a basis for further using PyTorch framework to construct recurrent neural network to accurately predict stock price.

Keywords

Stock Forecast, PyTorch, LSTM, Entrepreneurship 300 Index

1. 引言

股票市场是国民经济发展变化的“晴雨表”，是政治、经济、社会等诸多因素的综合反映。但股票的价格受到多种因素的共同影响，具有非线性、非平稳性、低信噪比以及长记忆性等特点，导致准确地预测股票价格的走势变得十分困难[1]。

传统的统计预测方法包括 ARIMA (移动平均自回归) [2]模型、GARCH (广义自回归条件异方差) [3]模型、GM (灰色系统理论) [4]模型等模型，该类模型更多地适用于平稳且线性的时间序列数据，因此不适用于股票价格预测。机器学习的方法包括决策树[5]、支持向量机[6]、逻辑回归[7]等，由于机器学习算法的自学能力非常强、对噪声数据的鲁棒性和容错性较强，适合处理非线性数据，因此该方法目前已广泛应用于图像识别、语音识别、情绪分类等领域，但是其在金融行业特别是股票预测方面的应用还不够深入，而深度学习作为机器学习的一个新的研究方向，更接近于人们对于人工智能的要求，被广泛应用于金融和经济领域，尤其在股票数据的预测中表现优异。陈可心和黄刚[8]提出由 BiLSTM 和 CLSTM 混合构建的 CStock 模型，该模型结合新闻和股价走势进行预测，不但利用了股票市场中的交易数据，同时考虑到财经以及政治新闻对于股票市场的影响。实验结果表明，CStock 模型在一定程度上能够准确有效地对股票走势进行预测。

韩山杰和谈世哲[9]研究发现基于 TensorFlow 框架 LSTM 循环神经网络模型比 BP 神经网络模型在对苹果公司股价的预测上准确性更高，耗时更短，更高效。本文提出了基于 PyTorch 框架 LSTM 循环神经网络模型，不单单针对某支股票价格进行预测，而是选取创业 300 指数从开盘以来的交易数据，即 2012 年 7 月 2 日到 2020 年 11 月 6 日共计 2032 个日交易数据进行拟合预测。由于数据时间跨度比较大，运用传统的 RNN 模型可能出现梯度爆炸和梯度消失的现象，从而对网络的稳定性造成极大的影响，使得预测误差较大。而 LSTM 循环神经网络模型可以很好地解决上述问题，并且利用时间序列向前和向后两个时间方向上的上下文关系，使预测准确度得到较大地提升。同时，该模型引入了 Dropout 策略，在一定程度上解决了深层网络模型带来的训练难、收敛速度慢和过拟合等问题，对实际问题的预测具有更高的准确性，因此基于 LSTM 深度学习方法对股票价格预测有较强的现实意义。

2. 模型介绍

2.1. LSTM 长短期循环神经网络

长短时记忆模型(Long Short-Term Memory, LSTM)是循环神经网络的变体。尽管在理论上，RNN 可以处理任何长距离依赖问题，但实际上，其对于梯度消失、爆炸等问题很难实现，对此，LSTM 通过引入门机制和记忆单元提供了解决方案，这里增加了遗忘门、更新门和输出门三个门机制。LSTM 单元结构如图 1 所示。

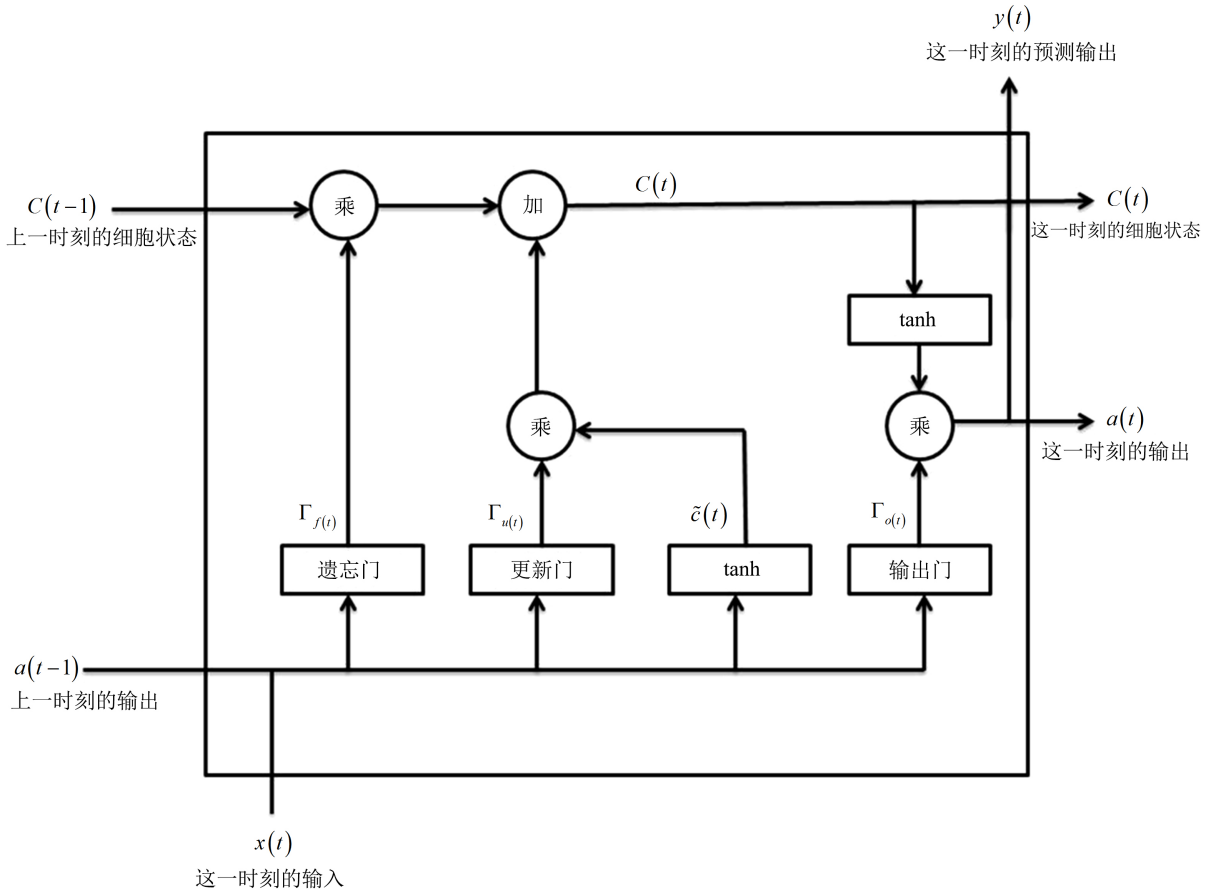

Figure 1. LSTM neural network unit structure diagram
图 1. LSTM 神经网络单元结构图

从图 1 可以看到，LSTM 模型的 1 个神经元包含了 1 个细胞状态(cell)和 3 个门(gate)机制。细胞状态(cell)是 LSTM 模型的关键所在，类似于存储器，是模型的记忆空间。细胞状态随着时间而变化，记录的信息由门机制决定和更新。门机制是让信息选择式通过的方法，通过 sigmoid 函数和点乘操作实现。sigmoid 取值介于 0~1 之间，乘即点乘则决定了传送的信息量(每个部分有多少量可以通过)，当 sigmoid 取 0 时表示舍弃信息，取 1 时表示完全传输(即完全记住)。

LSTM 拥有三个门，来保护和控制细胞状态：遗忘门(forget gate)、更新门(update gate)和输出门(output gate)。

LSTM 中保存的历史信息由遗忘门、更新门和输出门控制，在 t 时刻，各门状态的数学表达式为：

$$\tilde{c}(t)=\tanh\left(W_c\left[a(t-1),x(t)\right]+b_c\right) \tag{1}$$

$$\Gamma_{f(t)}=\delta\left(W_f\left[a(t-1),x(t)\right]+b_f\right) \tag{2}$$

$$\Gamma_{u(t)}=\delta\left(W_u\left[a(t-1),x(t)\right]+b_u\right) \tag{3}$$

$$\Gamma_{o(t)}=\delta\left(W_o\left[a(t-1),x(t)\right]+b_o\right) \tag{4}$$

$$c(t)=\Gamma_f*c(t-1)+\Gamma_u*\tilde{c}(t) \tag{5}$$

$$a(t)=\Gamma_o*\tanh\left(c(t)\right) \tag{6}$$

其中，$x(t)$ 为 t 时刻的输入数据，$a(t)$ 是 t 时刻 LSTM 单元的输出状态值，$\tilde{c}(t)$ 是 t 时刻记忆单元的候选值，$\Gamma_{f(t)}$ 是遗忘门 t 时刻的状态值，$\Gamma_{u(t)}$ 是更新门 t 时刻的状态值，$\Gamma_{o(t)}$ 是输出门 t 时刻的状态值。W 为相应的权值，b 是对应的偏置参数。记忆单元的状态值由更新门和遗忘门共同调节。

2.2. Dropout 介绍

Dropout 的数学形式表达式为：

$$y=f\left(W*d(x)\right) \tag{7}$$

$$d(x)=\begin{cases}mask*x, & \text{训练阶段}\\(1-p)x, & \text{其他}\end{cases} \tag{8}$$

其中，p 为 Dropout 率，mask 为以 $1-p$ 为概率的贝努力分布的二值向量。

从式(8)可以看出，Dropout 与 L1 和 L2 范式正则化不同，Dropout 并不会修改代价函数，而是修改深度网络本身。Dropout 随机“删除”网络中的一些隐藏神经元，保持输入输出神经元不变。这样，对于一个网络而言，Dropout 便是用相同的数据训练了多个不同的神经网络，产生了多个不同程度的拟合状态。但这些网络共用一个损失函数，相当于对神经网络本身进行了优化，求取了所有状态的平均值；同时，减少了神经单元之间的相互协同关系，增加了网络的鲁棒性。

2.3. 框架介绍

在开始深度学习项目之前，选择一个合适的框架非常重要，因为选择一个合适的框架能起到事半功倍的作用。众所周知，股票市场是一个受外界影响大，具有极强的不稳定性与随机性的系统，为投资者带来收益的同时，也可能会为投资者带来极大的亏损。因此，为了尽可能地扩大收益规避风险，对股票走势的准确预测变得尤为重要。

目前流行的深度学习主流框架有 PyTorch 和 TensorFlow 两大框架，针对本文创业 300 指数数据集，通过十折交叉验证结果发现，基于 PyTorch 框架的预测误差均小于 TensorFlow 框架，即基于 PyTorch 框架的预测准确性更高。与此同时，TensorFlow 框架更多运用于工业研究，语言系统设计复杂，在迭代次数较大时，计算时间较长；而 PyTorch 框架更多运用于学术界，应用十分灵活，接口沿用 Torch，契合用户思维，设计上更直观，追求尽量少的封装，建模过程更透明，便于多次调参和调整迭代次数进行最优模型的选择。综上所述，本文选择 PyTorch 框架构建 LSTM 循环神经网络模型对创业 300 指数收盘价进行预测。

2.4. 性能指标

为了衡量网络的稳定性与预测的准确性，本文根据股票的预测价格与实际价格，选取平均绝对百分

比误差 MAPE 作为评估指标，计算公式如式(9)所示：

$$\mathrm{MAPE}=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i-y_i}{y_i}\right| \tag{9}$$

其中，$\hat{y}_i$ 为股票的预测价格，y_i 为股票的实际价格，N 为训练的天数。

3. 实证分析

3.1. 数据的介绍

选取创业 300 指数从 2012 年 7 月 2 号至 2020 年 11 月 6 号的日交易数据，特征变量为开盘价 X_1、最高价 X_2、最低价 X_3、成交量 X_4、总金额 X_5 和涨跌幅 X_6，输出变量或预测变量为收盘价 Y。

3.2. 数据预处理

因为数据的不同输入变量数值上差异明显，并且会对模型训练造成严重的影响，因此首先对数据进行归一化处理，消除量纲对模型造成的影响。

3.3. 算法实现

3.3.1. 模型建立

本文提出的 LSTM 循环神经网络模型的结构如图 2 所示，该模型由输入层、LSTM 隐藏层、激活层和输出层四部分构成。

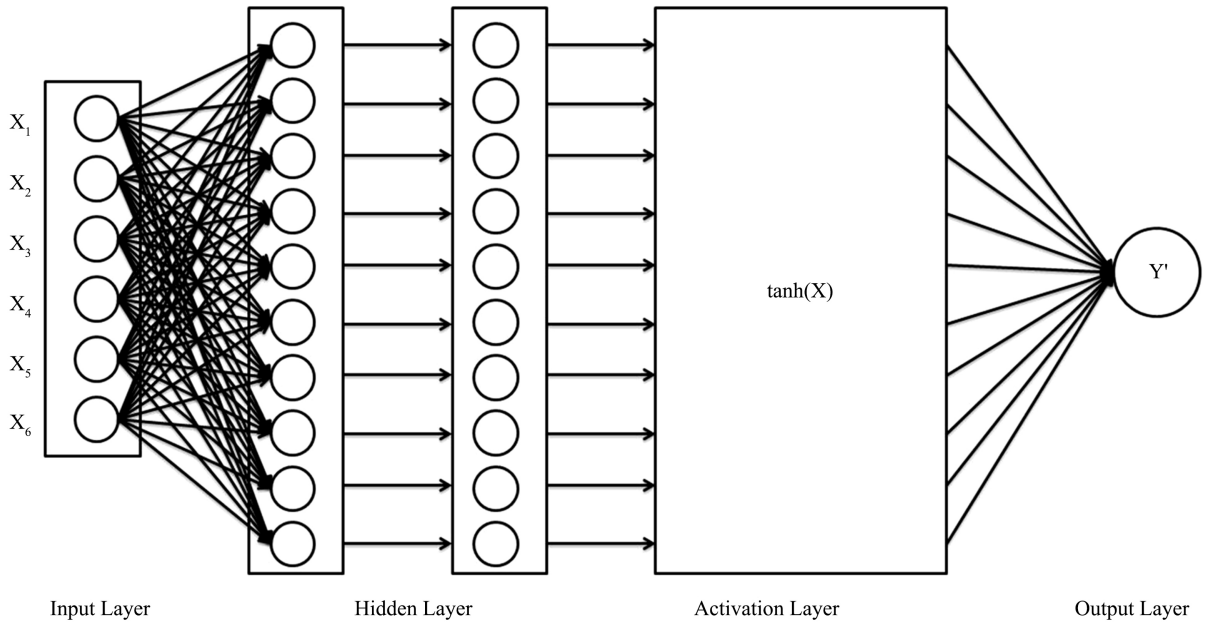

Figure 2. LSTM prediction model
图 2. LSTM 预测模型

3.3.2. 算法步骤

1) 数据预处理，将数据 $X=\{X_{11},X_{12},\cdots,X_{1n},X_{21},\cdots,X_{2n},\cdots,X_{6n}\}$ 归一化处理；

2) 初始化神经单元的细胞状态 $c(0)$ 和 $a(0)$，将预处理后的数据输入第一层 LSTM 神经元；

3) 按照式(1)~(4)，计算出当前神经元记忆候选值 $\tilde{c}(t)$，更新门状态 $\Gamma_{u(t)}$，遗忘门状态 $\Gamma_{f(t)}$ 和输出门状态 $\Gamma_{o(t)}$；

4) 根据式(5)计算当前神经元的记忆状态值 $c(t)$；

5) 根据式(6)计算当前神经元的输出值 $a(t)$；

6) 保留 $c(t)$、$a(t)$ 并将其运用到下一个时刻的 LSTM 神经元的计算中；

7) 重复 3~6 步骤，直到向前层、向后层的 LSTM 神经元均学习完全部的时间序列；

8) 重复以上步骤，直到最后一层 LSTM 输出非线性数据特征 $\hat{y}_i$。

3.4. 结果分析

3.4.1. 不同迭代次数的预测误差分析

Data1 我们选择开盘价、最高价、最低价、成交量、成交额这五个变量作为输入变量，Data2 加入涨跌幅变量即选择开盘价、最高价、最低价、成交量、成交额和涨跌幅这六个变量作为输入变量，以收盘价为输出变量建立 LSTM 模型进行预测。改变迭代次数观察不同输入变量个数下的平均绝对百分比误差。设置遗忘门偏置为 1，LSTM 单元数为 2，表 1 为输出预测误差结果对比。

Table 1. Comparison table of prediction errors of different iterations
表 1. 不同迭代次数的预测误差对比表

迭代次数	Data1 (输入维度 = 5，遗忘门偏置= 1.0，LSTM 单元数 = 2)	Data2 (输入维度 = 6，遗忘门偏置 = 1.0，LSTM 单元数 = 2)
10 次	0.0472	0.0439
50 次	0.0360	0.0322
100 次	0.0303	0.0214
200 次	0.0210	0.0189

由表 1 的结果可知，输入的维数越多，预测平均绝对百分比误差越小；证明涨跌幅这个变量的输入可以有效地降低预测误差，即引入涨跌幅变量具有实际意义。随着迭代次数的增加，预测误差在逐渐降低，而且从误差数据中可以看出在迭代 100 次时，网络已经比较稳定，但在迭代 200 次时，预测准确率达到最优。

3.4.2. 不同遗忘门偏置的预测误差分析

基于 Data2 的数据，设置迭代次数为 200，LSTM 单元数的值为 2，设置遗忘门偏置的值依次为 1、0.7、0.4 进行实验，预测误差结果如表 2 所示。

Table 2. Prediction error table with different forget bias
表 2. 不同遗忘门偏置的预测误差表

迭代次数	遗忘门偏置 = 1，LSTM 单元数 = 2	遗忘门偏置 = 0.7，LSTM 单元数 = 2	遗忘门偏置 = 0.4，LSTM 单元数 = 2
200 次	0.0189	0.0223	0.0174

由表 2 的前两列可以看出，Data2 的遗忘门偏置适当减小，即忘记部分信息，网络训练效果却有所下滑；从后两列可以看出，在 Data2 的遗忘门偏置变得更小，甚至忘记超过一半的信息时，网络训练效果却有所提高，并且高于完全保留信息时的网络，因此需进一步研究确定最优参数。

3.4.3. 不同 LSTM 单元数的预测误差分析

基于 Data2 的数据，设置迭代次数为 200，在遗忘门偏置为 1 和 0.4 的条件下，依次设置 LSTM 的单元数为 2、7、14 建模进行预测，所得到的预测误差结果如表 3 所示。

Table 3. Prediction error table for different LSTM cell numbers
表 3. 不同 LSTM 单元数的预测误差表

迭代次数	遗忘门偏置 = 1			遗忘门偏置 = 0.4		
	LSTM 单元数 = 2	LSTM 单元数 = 7	LSTM 单元数 = 14	LSTM 单元数 = 2	LSTM 单元数 = 7	LSTM 单元数 = 14
200 次	0.0186	0.0239	0.0482	0.0109	0.0159	0.0385

由表 3 的结果可知，在 LSTM 单元数增加的情况下，网络训练模型反而降低了，可以看出，其股票行情在 7 天内的关联程度比 14 天内的关联程度高，在 2 天内的关联程度比 7 天内的关联程度要高。在相同的单元数的情况下，遗忘门偏置应选择比较小的数，预测效果会比较好；在 LSTM 单元数较大的情况下，应该选择较小的遗忘门偏置，以免记忆太多的无效信息。

综上所述，得到创业 300 指数收盘价预测准确率最高的模型为输入变量为 6 维，迭代次数为 200，遗忘门偏置为 0.4，LSTM 单元数为 2 的基于 PyTorch 框架 LSTM 循环神经网络模型。

3.4.4. 可视化结果

选取 Data2，即选择开盘价、最高价、最低价、成交量、成交额和涨跌幅这六个变量作为输入变量，迭代次数为 200 次，分别输出 LSTM 模型一(遗忘门偏置 = 1、LSTM 单元数 = 2)、LSTM 模型二(遗忘门偏置 = 0.7、LSTM 单元数 = 2)、LSTM 模型三(遗忘门偏置 = 0.4、LSTM 单元数 = 2)这三个模型的 Loss 值随迭代次数变化的折线图见图 3~5 与模型预测结果的可视化图见图 6~8。

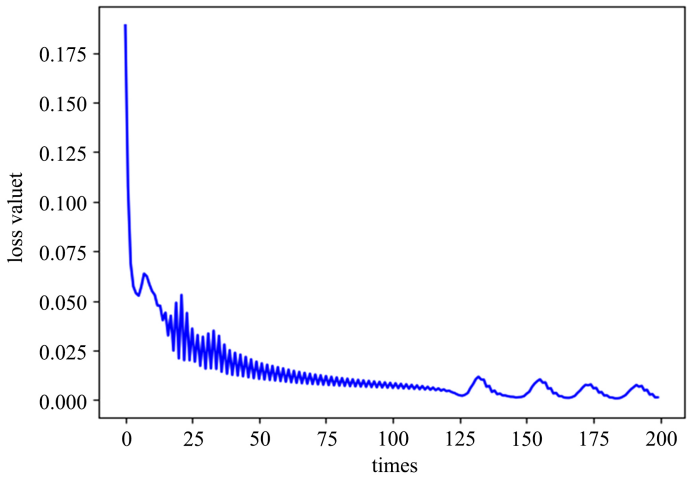

Figure 3. Loss function diagram (LSTM model 1)
图 3. Loss 函数图(LSTM 模型一)

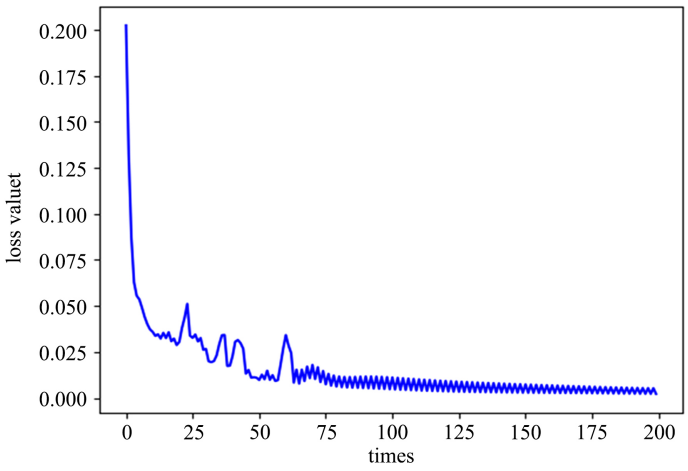

Figure 4. Loss function diagram (LSTM model 2)
图 4. Loss 函数图(LSTM 模型二)

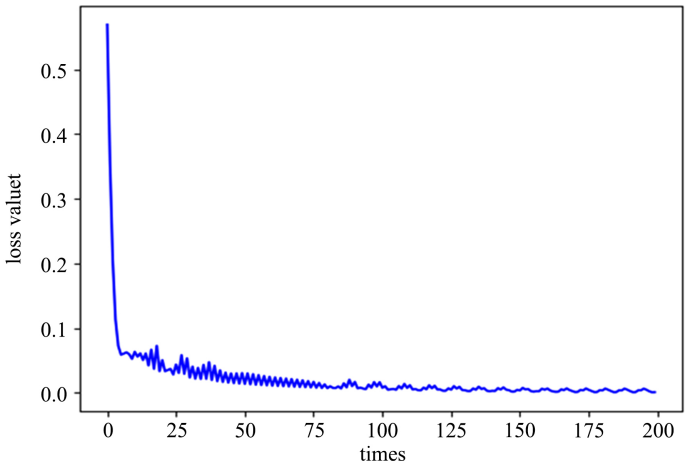

Figure 5. Loss function diagram (LSTM model 3)
图 5. Loss 函数图(LSTM 模型三)

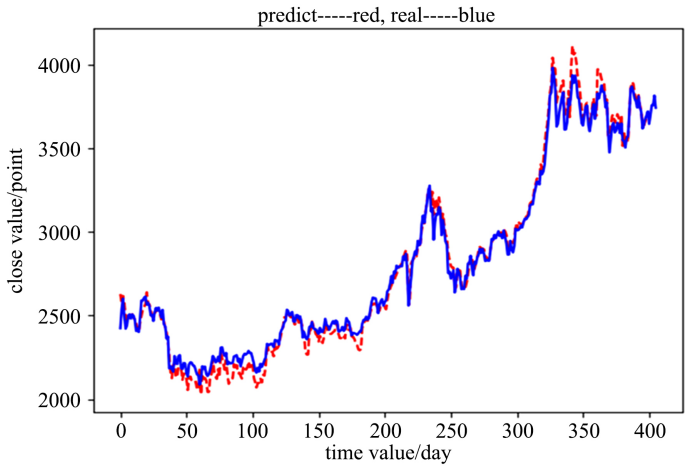

Figure 6. LSTM prediction of model 1
图 6. 预测拟合图(LSTM 模型一)

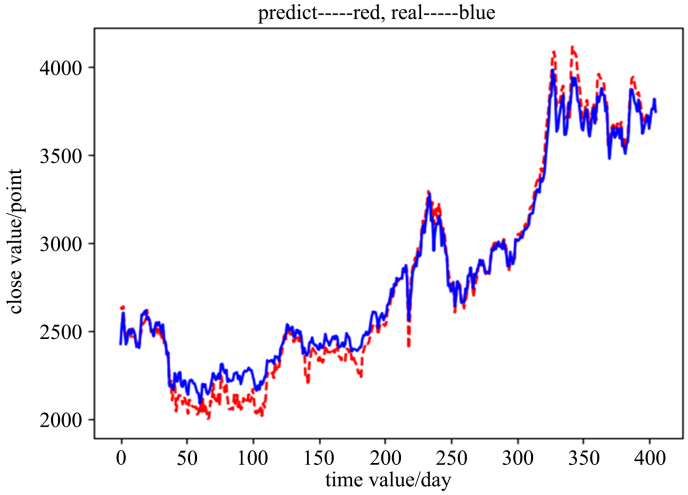

Figure 7. LSTM prediction of model 2
图 7. 预测拟合图(LSTM 模型二)

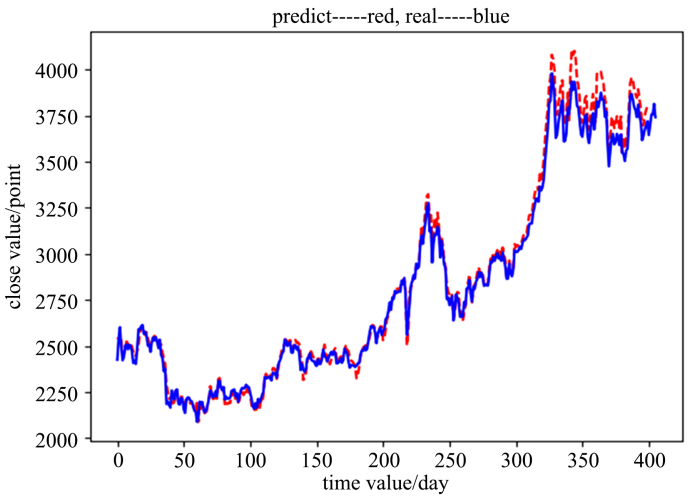

Figure 8. LSTM prediction of model 3
图 8. 预测拟合图(LSTM 模型三)

从以上可视化的结果可以看出，遗忘门偏置 = 1、LSTM 单元数 = 2 时，随着迭代次数的增加，损失值直线下降，但是在迭代次数大于 100 之后，损失值上下震荡明显，说明网络不稳定，从预测拟合图中可以看出，模型前期拟合效果较差，中期与后期相对较好，说明网络训练不稳定；遗忘门偏置 = 0.7、LSTM 单元数 = 2 时，Loss 可视化的结果可以看出，25 到 100 次迭代区间，损失值忽高忽低，100 次之后，损失值上下震荡明显，说明网络极不稳定，从预测拟合图中可以看出，模型前期和后期的拟合效果较差，误差较大；遗忘门偏置 = 0.4、LSTM 单元数 = 2 时，随着迭代次数的增加，网络损失值逐渐收敛，在迭代 100 次之后，损失值趋于最小值，说明网络基本稳定，在迭代 200 次时，损失值达到最小，从预测拟合图中可以看出，模型前中期走势的拟合效果都比较好，只有后期拟合稍有误差，整体拟合效果优于前两个模型。因此，可视化结果与上文预测平均绝对百分比误差输出结果吻合，更进一步证明输入变量为 6 维，迭代次数为 200，遗忘门偏置为 0.4，LSTM 单元数为 2 的基于 PyTorch 框架 LSTM 循环神经网络模型的预测准确率最高，拟合效果最好。

4. 结论

PyTorch 框架具有设计直观，简易封装，集成度高等特点，使其能够灵活地调节参数，提高建模分析的效率。使用 PyTorch 框架设计并实现了用于股票预测的深度学习流程以及算法框架，得到了有较高准确率的预测模型。

通过实验验证了 PyTorch 框架构建 LSTM 循环神经网络模型的优势，分析了迭代次数、遗忘门偏置值以及 LSTM 神经元个数对网络模型的影响。首先，通过设置不同的迭代次数可知，迭代次数在 100 次后，网络输出误差有明显的降低，说明该网络是一个较轻量级的网络；为了预测的准确率，选择迭代 200 次进行后面的建模分析。其次，通过设置不同的遗忘门偏置可知，偏置值为 1 和 0.4 时预测效果优于偏置值为 0.7 时的模型。因此，在遗忘门偏置值为 1 和 0.4 时，分别设置不同的 LSTM 神经元个数建模可知，遗忘门偏置为 0.4，LSTM 单元数为 2 的基于 PyTorch 框架 LSTM 循环神经网络模型的预测误差最小，平均绝对百分比误差达到 0.0109；最后，该结果通过 Loss 函数图与预测拟合图得到了进一步证明。为进一步使用 PyTorch 框架构建循环神经网络准确预测股价提供了依据，具有广阔的应用前景。

基金项目

国家自然科学基金(11761020)；贵大培育[2019] 62。

参考文献

[1] 于海姝, 蔡吉花, 夏红. ARIMA 模型在股票价格预测中的应用[J]. 经济师, 2015(11): 156-157.

[2] 石佳, 刘威, 冯智超, 等. 基于 ARIMA 模型的股市价格规律分析与预测[J]. 统计学与应用, 2020, 9(1): 101-114.

[3] 崔旭盛, 李鑫, 宗建新, 刘灵娣, 靳鹏博, 董学会. 基于 GARCH 族模型的中药材市场连翘价格波动分析[J]. 北方园艺, 2021(4): 144-150.

[4] 侯瑞环, 徐翔燕. 基于改进 GM(1, 1)模型的中长期人口预测[J]. 统计与决策, 2021, 37(1): 186-188.

[5] 王禹. 基于 Cart 树和 Boosting 算法的股票预测模型[D]: [硕士学位论文]. 哈尔滨: 哈尔滨理工大学, 2018.

[6] 马雪姣. 基于支持向量回归机模型的价格预测[D]: [硕士学位论文]. 郑州: 郑州大学, 2018.

[7] 阎馨, 吴书文, 屠乃威, 朱永浩, 付华. 基于逻辑回归和增强学习的瓦斯突出预测[J/OL]. 控制工程: 1-7[2021-03-07]. https://doi.org/10.14107/j.cnki.kzgc.20200044

[8] 陈可心, 黄刚. CStock: 一种结合新闻与股价的股票走势预测模型[J]. 计算机技术与发展, 2020, 30(9): 18-22.

[9] 韩山杰, 谈世哲. 基于 TensorFlow 进行股票预测的深度学习模型的设计与实现[J]. 计算机应用与软件, 2018, 35(6): 267-271+291.