

# 关于LSTM-TCN模型结合递归特征消除法的股票预测

王文姣\*, 张 娜

兰州交通大学数理学院, 甘肃 兰州

收稿日期: 2022年9月17日; 录用日期: 2022年10月8日; 发布日期: 2022年10月17日

## 摘 要

近年来, 神经网络深度学习算法在金融领域研究中一直得到广泛应用。为了帮助股民深刻了解股票行情进而实现合理化投资, 在融合长短时记忆网络(LSTM)和时序卷积神经网络(TCN)的优势构建的LSTM-TCN模型基础上, 引入递归特征消除法, 通过反复训练得到可以明显降低预测误差的特征向量组合作为LSTM-TCN模型的输入; 其次利用LSTM网络提取数据局部时间特征, 再将输出的特征矩阵输入TCN网络进一步挖掘长时间特征, 从而构成LSTM-TCN联合模型来对股票收盘价进行预测, 可实现时序特征的充分挖掘, 获取更多的信息。中国平安股票数据的预测结果表明: LSTM-TCN联合模型的RMSE为0.773, MAPE为1.05%, 均小于其他传统机器学习模型以及LSTM、TCN模型, 说明了提出的LSTM-TCN联合模型在股价预测方面具有更高的预测精度。

## 关键词

神经网络, 股票预测, 收盘价, LSTM-TCN联合模型

# Stock Prediction Based on LSTM-TCN Model Combined with Recursive Feature Elimination Method

Wenjiao Wang\*, Na Zhang

School of Mathematics and Physics, Lanzhou Jiaotong University, Lanzhou Gansu

Received: Sep. 17<sup>th</sup>, 2022; accepted: Oct. 8<sup>th</sup>, 2022; published: Oct. 17<sup>th</sup>, 2022

## Abstract

In recent years, neural network deep learning algorithm has been widely used in the research of  
\*通讯作者。

financial field. In order to help investors deeply understand the stock market and realize rational investment, the Recursive Feature Elimination method is introduced on the basis of LSTM-TCN model established by combining the advantages of Long-Term And Short-Term Memory Network (LSTM) and Time-Series Convolution Neural Network (TCN), and the combination of feature vector that can significantly reduce prediction error is obtained through repeated training as the input of LSTM-TCN model; secondly, the LSTM network is used to extract the local time features of the data, and then its output feature matrix is input into the TCN network to further mine the long-time features, so as to form the LSTM-TCN joint model to predict the stock closing price, which can fully mine the time series characteristics and obtain more information. The prediction results of Ping An's stock data show that the RMSE and MAPE of the LSTM-TCN joint model are 0.773 and 1.05% respectively, which are lower than those of other traditional machine learning models, LSTM and TCN models, indicating that the LSTM-TCN joint model has higher prediction accuracy in stock price prediction.

## Keywords

Neural Network, Stock Prediction, Closing Price, The LSTM-TCN Joint Model

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

股票市场作为金融行业的支柱之一,有着不可替代的影响地位。一方面,股票市场是我国最重要的直接融资渠道。大部分上市公司都需要借助股票市场实现资本融资;另一方面,我国股票市场的发展也与我们生活水平息息相关。随着生活条件的进一步提高,越来越多的人将目光聚焦于股票投资中来。因此,股票市场的运转很好地诠释了国民经济的发展变化。股票行情总是受诸多因素影响而不断波动,如何从各方面因素入手来把握股票行情的动向值得不断的探讨研究。

随着大数据时代的来临,像支持向量机、随机森林等传统机器学习模型以及深度学习算法等技术逐渐被引入金融数据领域的研究中并得到应用。相较于传统机器学习模型而言,深度学习算法具有更强的数据识别能力。递归神经网络(RNN)是目前关于时间数据最常用的深度学习模型,而另一个重要的深度学习分支是广泛用于图像识别等领域的卷积神经网络(CNN)。其中,为提高时间数据的预测效果,又进一步发展出长短记忆神经网络(LSTM)和时序卷积神经网络(TCN)。由于具备良好的选择性、扩展记忆性以及时序内部影响的特性,LSTM和TCN模型在股票价格序列数据的预测研究中拥有广阔的应用前景。

本文基于深度学习算法对股票价格数据进行了预测研究,以股票收盘价为研究目标,提出了一种结合LSTM和TCN网络结构的联合预测模型。该模型主要采用端到端的网络结构[1],首先使用LSTM来提取数据的时间特征和局部特征,然后利用TCN进一步挖掘时间序列中的长记忆特征,充分捕获股票价格时间序列的数据特征从而更准确地预测股票价格数据。

本实验以Tensorflow作为神经网络框架,使用Python语言进行模型的代码实现。以沪深指数中最具代表性个股数据——中国平安为研究对象,对真实值和预测值通过图形拟合[2]和误差评估方法进行结果分析,并与LSTM、TCN、随机森林回归(RFR)和支持向量机回归(SVR)模型的预测结果进行对比实验,验证了LSTM-TCN模型的有效性。

## 2. 模型原理

### 2.1. 长短记忆神经网络

长短记忆神经网络(Long Short-Term Memory, LSTM)是为改进循环神经网络(RNN)的梯度消失和梯度爆炸的缺陷发展而来的一种循环神经网络, 其优势在于它通过引入门结构, 具体包括输入门、输出门和遗忘门, 可以有效实现信息的长期记忆和有效筛选, 从而克服 RNN 易遗忘前期信息的弊端。所以针对股票价格时间序列数据, LSTM 神经网络有着独特的优势[3]。LSTM 细胞的信息传递过程如下所示:

$$\begin{aligned} \text{遗忘门: } f_t &= \sigma(W_f [h_{t-1}, x_t] + b_f) \\ \text{输入门: } &\left\{ \begin{aligned} i_t &= \sigma(W_i [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_c [h_{t-1}, x_t] + b_c) \\ C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \end{aligned} \right\} \\ \text{输出门: } &\left\{ \begin{aligned} O_t &= \sigma(W_o [h_{t-1}, x_t] + b_o) \\ h_t &= O_t * \tanh(C_t) \end{aligned} \right\} \end{aligned}$$

其中:  $W_f, W_i, W_c, W_o$  分别表示各门的权重矩阵;  $b_f, b_i, b_c, b_o$  分别是各门的偏置;  $\sigma, \tanh$  分别表示 *sigmoid* 和双曲正切激活函数;  $\tilde{C}_t$  表示  $t$  时刻由输入门决定的部分更新信息, 以此计算  $t$  时刻的输出状态  $h_t$  和更新细胞状态  $C_t$ 。

### 2.2. 时序卷积神经网络

时序卷积神经网络(Temporal Convolutional Network, TCN)是卷积神经网络为适合时间序列预测需要而发展的一种新型预测算法。与 RNN 模型相比, TCN 可以通过增加网络层数、改变扩张系数和滤波器的大小来改变感受野, 可以更好地控制模型的内存大小, 进而控制模型的记忆长短。并且 TCN 的反向传播路径和序列的时间方向不同, 这可避免 RNN 中的梯度消失和梯度爆炸的问题[3]。

#### 2.2.1. 时序卷积

TCN 主要是通过对一维卷积进行变形从而发展为适用于时间问题的模型, 可通过增加网络层数来学习长时间跨度信息, 每一层网络均使用一维卷积核读取当前时间点前的所有信息, 将序列通过网络层层传递, 最后得到预测结果。网络越深, 可学习的信息时间跨度越长[4]。

时序卷积与普通的一维卷积存在以下两点不同, 具体结构见图 1(c)。

- 1) 使用因果卷积。特点在于网络输出长度等于输入长度, 保证未来信息不会用于预测。
- 2) 采用扩张卷积。特点在于可通过增加层数和扩张卷积核大小来增大感受野。越到上层, 卷积核越大, 输出的信息范围越广, 从而确保滤波器能够获得更大范围输入特征, 高效获取长时间跨度信息。

#### 2.2.2. 残差链接

TCN 网络是由多个残差链接模块组成(见图 1(a))。每一个残差模块(Residual block)又包括扩张因果卷积层(Dilated Causal Conv), WeightNorm 层, ReLU 激活函数和 Dropout 层四个部分, 其中: WeightNorm 层的作用是加快训练速度, 提高准确性[5]; ReLU 激活函数被添加到两个卷积层之后, 可有效避免训练过程中梯度消失的问题; 为防止模型过拟合现象还在每个卷积层之后通过 Dropout 引入了正则化; 加入  $1 \times 1$  卷积可以调整残差张量的宽度, 解决输入和输出可能出现的维度差异问题。残差模块的具体结构见图 1(b)。

### 2.3. 基于 LSTM-TCN 股价预测模型

本文结合 LSTM 与 TCN 的各自优势, 并引入全连接层模块[6], 针对股价预测问题构建了 LSTM-TCN

联合模型, 其整体结构如图 2 所示。

### 3. 递归特征消除法(RFE)

递归特征消除法(Recursive Feature Elimination, RFE)是一种包装式的特征选择方法[7], 其目的是搜寻最优特征组合。它的核心思想是通过反复构建基本模型, 进行多轮训练并在每一轮训练中, 根据特征系数消除最好(或最差)的特征, 然后对剩余的特征新子集进行下一轮训练, 直至遍历所有特征。在这个过程中, 特征被消除的次序即为特征的重要性排序。

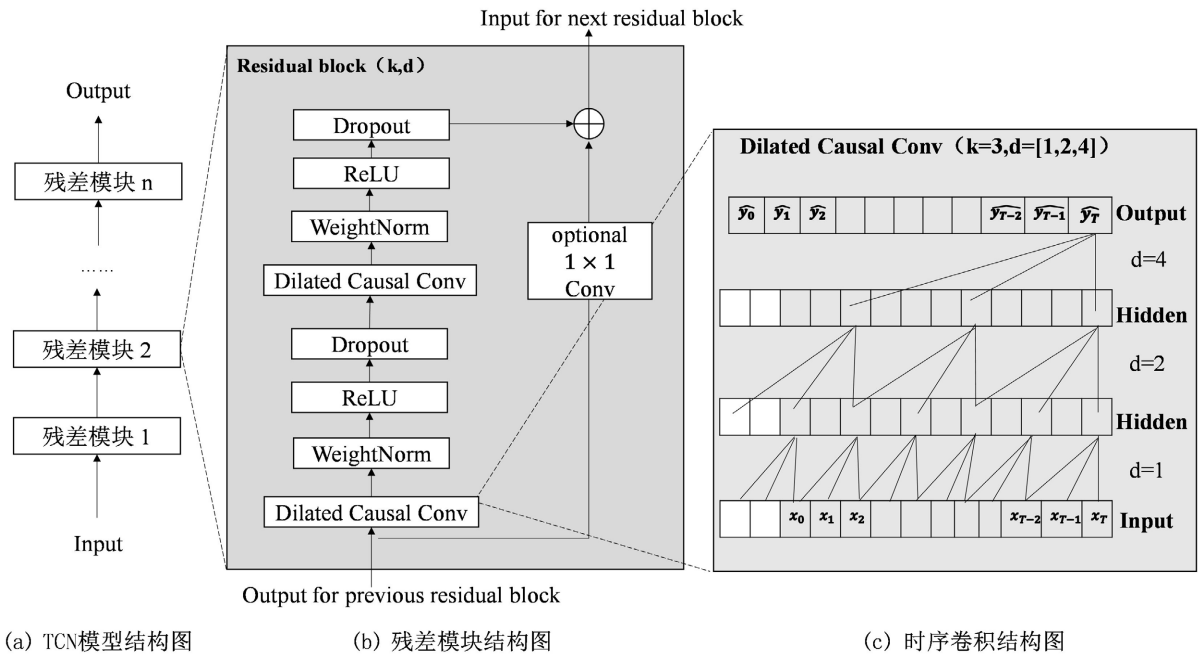


Figure 1. The structure diagram of TCN model

图 1. TCN 模型结构图

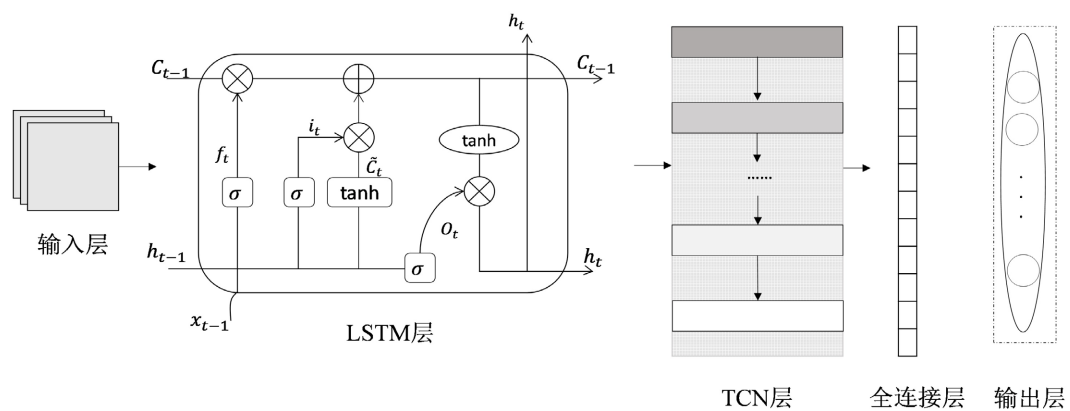


Figure 2. The structure diagram of LSTM-TCN model

图 2. LSTM-TCN 模型结构图

基于已有研究成果, 本文采用基于随机森林的特征选择方法, 利用递归特征消除方法的特点[8], 获得影响股价波动的各因素的重要性排序。其具体执行过程如下:

- 1) 确定包含  $k$  个特征的初始特征集合, 构建随机森林分类模型, 计算每个特征的重要性, 并利用交叉验证法衡量模型分类精度。
- 2) 从当前特征集合中剔除重要性最小的特征, 得到一个新的特征子集, 再次构建随机森林分类模型, 计算新的特征子集中每个特征的重要性, 并利用交叉验证法衡量模型分类精度。
- 3) 递归重复(2), 直至特征子集变为空集。最终特征剔除的顺序即为各特征重要性排序。
- 4) 在这个过程中可得到  $k$  个不同特征数量的特征子集, 选择分类精度最高的特征子集作为最优特征组合。

## 4. 实证研究

### 4.1. 数据获取

沪深 300 指数反映的是最具代表性 A 股上市公司股票行情的综合表现, 这类股票具有较强的稳定性、代表性和可操作性, 将其作为投资业绩的评价标准有利于投资者随时跟踪股票市场行情, 及时规避风险实现最优投资, 因此具有预测意义。通过证券之星网站可查询“沪深 300 指数”各成分股及其对应权重, 其中本文选择了沪深指数中权重最大的成分股—中国平安作为实验对象。利用 Tushare 财经接口包下载的中国平安 2021 年 1 月 4 日至 2022 年 2 月 28 日, 分别获取了股票的收盘价、开盘价、最高价、最低价、涨跌额、涨跌幅、成交量、成交额(等时序数据, 共包含 1 个预测目标(*close*)和 52 个待研究特征。

### 4.2. 数据处理

1) 数据预处理: 由于股票交易市场存在节假日休市情况, 可以观察到所获取原始数据集存在缺失、不连续等情况, 需对数据集进行插值操作, 保证时序数据的连续性。本文采用的是线性插值法进行填补。

2) 数据归一化处理: 由于各变量之间具有不同的量纲和单位, 为消除变量间的量纲影响, 保证实验模型的稳健性, 需要对数据进行归一化处理。常采用最大-最小归一法, 具体转换方式:

$$X = \frac{x - \max(x)}{\max(x) - \min(x)}$$

其中:  $x$  为待处理原始序列,  $X$  为归一化后的序列。

### 4.3. 特征选择

若使用全部数据进行训练, 会由于维度过大使得神经网络产生大量参数, 造成训练缓慢以及维数灾难等问题, 故为进一步排除不重要特征和特征维数爆炸等问题对模型的影响, 在数据处理的基础上依据第 2 节阐述的关于随机森林递归特征消除法的执行过程, 对各特征变量关于 *close* 进行最优特征组合的筛选, 最终保留了 421 个样本, 筛选出 6 个重要特征变量, 见表 1。

**Table 1.** The input variables

**表 1.** 输入变量

特征名称	符号说明
开盘价	<i>open</i>
最高价	<i>high</i>
最低价	<i>low</i>
连续两天的开盘价均值	<i>open_2_sma</i>
布林带中轨线	<i>boll</i>
布林带上轨线	<i>boll_ub</i>

#### 4.4. 实验评价指标

为了评估真实值与预测值之间的误差, 本文中选取了几个常用的误差评估指标作为衡量标准。一般来说 RMSE 越小, MAPE 越小, 表示测量精度越高。具体计算公式如下:

$$\text{均方根误差: } RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$\text{平均绝对百分误差: } MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

其中:  $n$  为样本容量,  $y$  为样本实际值,  $\bar{y}$  为样本均值,  $\hat{y}$  为样本预测值。

#### 4.5. 实验结果及分析

经过 RFE 算法处理构成了  $421 \times 7$  的原始数据集。本文按照为 7:3 比例将原始数据集划分为训练集和测试集, 即中国平安股票数据中 2021 年 1 月 4 日至 2021 年 10 月 24 日共 294 条样本为训练集, 用于模型训练; 2021 年 10 月 25 日至 2022 年 2 月 28 日共 127 条样本为测试集, 用于预测评估。

##### 4.5.1. LSTM-TCN 模型预测结果

不同的数据结构与参数大小会产生不同的预测效果, 经多次测试训练, 本文对模型框架进行以下设置应用:

###### 1) 输入数据(Input)

输入数据的张量形状设置为  $[N, 10, 6]$ , 其中  $N$  代表样本数量, 10 指利用过去 10 个时间点的样本数据来预测下一时间点, 6 代表每个时间点的特征数量。

###### 2) LSTM-TCN 结构设计

TCN 中卷积核与 LSTM 中单元的功能相似[9]。设置 LSTM 神经元数量为 64; 设置 TCN 滤波器数量为 64, 卷积核为 5, 同时为避免过拟合现象将 Drop\_rate 设置为 20%。

###### 3) 输出数据(Output)

输出数据形状为  $[N, 1]$ ,  $N$  代表样本数量。具体而言就是将 TCN 块的输出拉直作为全连接层的输入, 然后运用 Relu 激活函数生成各时间点的预测值。

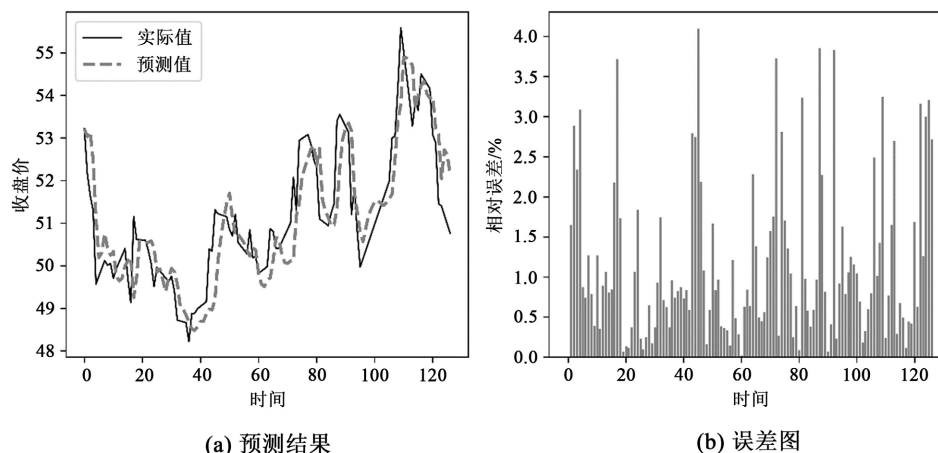
本文基于训练集对 LSTM-TCN 模型进行训练, 并对测试集样本进行预测, 最终利用图形拟合的形式绘制出预测曲线与实际曲线, 同时计算了各测试样本的预测相对误差, 得到图 3 的预测结果。

$$\text{相对误差} = \left| \frac{\text{实际值} - \text{预测值}}{\text{实际值}} \right| \times 100\%$$

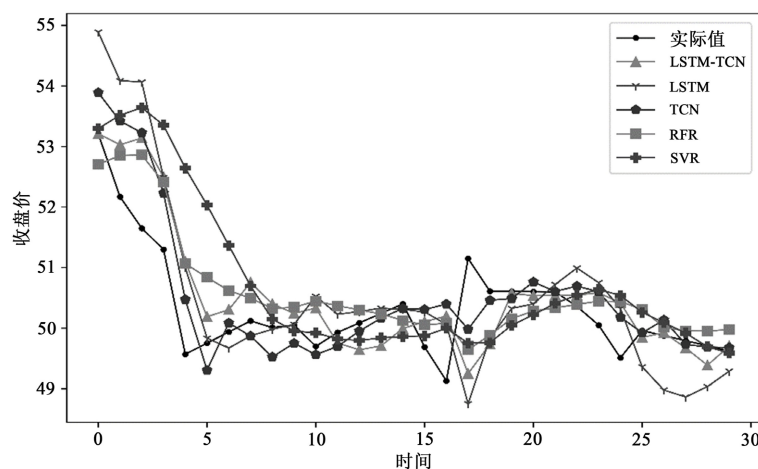
观察并分析上述拟合曲线, 不难看出: 预测曲线相较于实际曲线存在一定时间滞后性, 但所提模型对于该股票数据的预测输出与真实输出的波动趋势总体是一致的。针对有限的样本数据和实验环境, 存在预测误差是不可避免的。从误差图来看测试样本的相对误差基本控制在 5% 以内, 说明 LSTM-TCN 模型对于中国平安股价的预测还是比较准确的。

##### 4.5.2. 误差评估

为验证深度学习方法的优越性以及所提模型的有效性, 分别采用随机森林回归模型(RFR)、支持向量回归机(SVR)等传统机器学习模型以及单一结构的 LSTM、TCN 模型与 LSTM-TCN 模型进行对比实验, 用图形拟合的方式展示模型预测结果。由于测试集数据较多且各模型预测差异较小, 为更清晰地观察各曲线的预测差异, 仅绘制了前 30 个样本的预测曲线, 结果见图 4。



**Figure 3.** The prediction results of TCN model  
**图 3.** LSTM-TCN 模型预测结果



**Figure 4.** The comparison of local prediction curves of models  
**图 4.** 模型局部预测曲线对比

对比各模型预测曲线和实际曲线走势, 可知各模型的预测曲线都基本捕获了股票价格的波动趋势, 都具有一定的有效性预测。但 LSTM-TCN 联合模型的预测曲线整体更接近于实际曲线, 说明了所提出联合模型具有提升预测精度的效果。

为了更加直观、准确地验证各个模型的预测效果, 分别计算各模型在测试集上的预测误差, 得到表 2 结果。

**Table 2.** The comparison of model prediction errors  
**表 2.** 模型预测误差对比

模型	RMSE	MAPE/%
<b>LSTM-TCN</b>	<b>0.773</b>	<b>1.05</b>
LSTM	0.776	1.08
TCN	0.778	1.11
RFR	0.886	1.33
SVR	1.058	1.56

对比表 2 中评估的各模型预测误差, 探究深度学习模型和普通学习模型的优劣, 可得出以下结论:

1) 与传统机器学习模型(RFR 和 SVR)相比, LSTM、TCN 以及 LSTM-TCN 这些深度学习模型的预测精度均高于传统机器学习模型, 说明了深度学习算法具有更强的数据识别能力, 更能捕获股票时序数据特征从而进行更高精度的预测。

2) 与单一结构的 LSTM、TCN 模型相比, LSTM-TCN 模型的 RMSE 分别减少了 0.003 和 0.005, MAPE 分别减少了 0.03% 和 0.06%, 说明了 LSTM-TCN 联合模型可提升股价的预测精度, 值得推广。

## 5. 结论与展望

本文提出“LSTM-TCN”模型对股票市场进行建模预测, 选取代表性的中国平安个股数据为研究对象, 用实例验证了 LSTM-TCN 模型可通过学习和训练很好地完成股票价格走势的预测。结果表明:

1) LSTM-TCN 显示在样本外的预测能力显著高于其他已广泛应用的支持向量机等传统机器学习模型。单一的 LSTM 和 TCN 模型的股票预测结果较为理想但仍有改进之处。由于股票市场的复杂性以及随机性, 实证中的股价预测曲线与股票市场的股价实际值是存在一定误差的, 结合 LSTM 和 TCN 模型的优势所构建的 LSTM-TCN 模型能够进一步缩小该误差, 具有一定参考价值。

2) 合理选取模型输入变量可保证预测精度。由于股票价格的波动趋势受诸多因素的影响, 本文在实证研究部分运用递归消除法对涉及到的特征指标进行重要性评估并合理筛选了重要变量作为模型输入变量。但由于神经网络模型的灵活性, 对于不同的问题变量的选取也需要灵活多变。需要根据实际情况进行合理筛选, 否则易陷入局部极值[10]。

3) 文章选取的数据量有限, 研究结果不能完全代表整个股票市场的预测应用效果。可进一步扩充数据类型, 研究 LSTM-TCN 模型的适用性。

## 致 谢

本文对给予转载和引用权的资料、图片、文献、研究思想和设想的所有者, 表示感谢; 对导师在论文中提供的学术指导和帮助表示感谢; 对各大财经网站提供的股票交易数据致谢。

## 参考文献

- [1] 赵红蕊, 薛雷. 基于 LSTM-CNN-CBAM 模型的股票预测研究[J]. 计算机工程与应用, 2021, 57(3): 203-207.
- [2] 孙瑞奇. 基于 LSTM 神经网络的美股股指价格趋势预测模型的研究[D]: [硕士学位论文]. 北京: 首都经济贸易大学, 2015.
- [3] 严珂, 申恒乐. 基于 TCN-LSTM 的太阳能光伏发电预测方法与流程[EB/OL]. 国知局. [http://www.xjshu.com/zhuanli/55/201911125632\\_3.html](http://www.xjshu.com/zhuanli/55/201911125632_3.html), 2020-03-24.
- [4] 张妮. 基于 LSTM 神经网络的股价预测应用研究[J]. 现代商业, 2021(16): 116-118.
- [5] 周楚杰. 基于 LSTM 和 TCN 混合深度学习的风速短期预测模型[D]: [硕士学位论文]. 兰州: 兰州大学, 2019: 24-25.
- [6] 田红丽, 金硕, 闫会强. 一种基于 TCN 和新闻情感的股票预测方法[J]. 信息技术与信息化, 2021(6): 17-21.
- [7] 陈维杰, 江伟辉, 贾西贝. 基于 CNN-GRU 联合模型的股票指数价格预测[J]. 信息技术与信息化, 2021(9): 87-91.
- [8] Wang, G.D., Sun, L.X., Wang, W., Chen, T., Guo, M.T. and Zhang, P. (2020) A Feature Selection Method Combined with Ridge Regression and Recursive Feature Elimination in Quantitative Analysis of Laser Induced Breakdown Spectroscopy. *Plasma Science and Technology*, **22**, 16-25. <https://doi.org/10.1088/2058-6272/ab76b4>
- [9] 黄卫卫. 基于随机森林——递归特征消除的道路交通事故成因分析[J]. 电脑知识与技术, 2018, 14(14): 240-243.
- [10] 李骏琪, 杨益立. “CNN + TCN”模型在高频国债期货市场走势预测中的应用[J]. 统计与决策, 2022, 38(1): 145-148.