

# PCC-BiLSTM-GRU-Attention 组合模型 预测方法<sup>①</sup>



高凯悦, 牟莉, 张英博

(西安工程大学 计算机科学学院, 西安 710048)  
通信作者: 高凯悦, E-mail: 903119649@qq.com

**摘要:** 传统预测模型在处理多元时间序列时, 常常难以捕捉其非线性动力系统的复杂变化规律导致预测精度较低. 针对此问题, 本文将 PCC-BiLSTM-GRU-Attention 组合模型的预测方法进行了探讨和验证. 该方法首先使用 Pearson 相关系数 (PCC) 进行相关性检验并删除无关特征, 实现了对多元数据的降维选优. 其次使用双向长短期记忆神经网络 (BiLSTM) 双向提取时序特征. 最后使用 GRU 神经网络融合注意力机制 (Attention), 进一步学习双向时序特征的变化规律, 精准捕捉关键时刻的信息. 为了验证该方法在多元时间序列中的可行性, 本文以股票价格预测作为实验场景, 并与 BP 模型、LSTM 模型、GRU 模型、BiLSTM-GRU 模型、BiLSTM-GRU-Attention 模型进行对比. 验证结果表明: 本文探讨的 PCC-BiLSTM-GRU-Attention 组合模型的预测方法相比其他模型具有较高的预测精度, 其平均绝对百分比误差 (MAPE) 达到了 2.484%, 决定系数达到了 0.966.

**关键词:** 相关系数; 双向长短期记忆神经网络; 门控循环单元网络; 注意力机制; 多元时间序列; 深度学习

引用格式: 高凯悦, 牟莉, 张英博. PCC-BiLSTM-GRU-Attention 组合模型预测方法. 计算机系统应用, 2022, 31(7): 365-371. <http://www.c-s-a.org.cn/1003-3254/8580.html>

## Prediction Method of PCC-BiLSTM-GRU-Attention Combined Model

GAO Kai-Yue, MOU Li, ZHANG Ying-Bo

(School of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China)

**Abstract:** When dealing with multivariate time series, traditional prediction models are often difficult to capture the complex variation of nonlinear dynamic systems, which results in low prediction accuracy. To solve this problem, this study discusses and verifies the prediction method of the PCC-BiLSTM-GRU-Attention combined model. In the method, Pearson correlation coefficient (PCC) is first used for correlation tests and irrelevant features are deleted to achieve dimensionality reduction and optimization of multivariate data. Then, bidirectional long short-term memory (BiLSTM) neural network is used to extract time series features. Finally, GRU neural network is integrated with the attention mechanism to further learn the change rule of bidirectional time series features and accurately capture the critical moment information. To verify the feasibility of this method in multivariate time series, this study takes stock price prediction as the experimental scene and compares it with the BP model, LSTM model, GRU model, BiLSTM-GRU model and BiLSTM-GRU-Attention model. The verification results show that the prediction method of the PCC-BiLSTM-GRU-Attention combined model has higher prediction accuracy than other models, with the mean absolute percentage error (MAPE) reaching 2.484% and the determination coefficient 0.966.

**Key words:** correlation coefficient; bidirectional long short-term memory (BiLSTM) neural network; gated loop unit network; attention mechanism; multivariate time series; deep learning

<sup>①</sup> 基金项目: 陕西省 2020 年技术创新引导专项 (2020CGXNG-026)

收稿时间: 2021-10-11; 修改时间: 2021-11-08; 采用时间: 2021-11-19; csa 在线出版时间: 2022-05-31

多元时间序列是指以相同间隔按指定采样率对某过程进行采集的一组变量,且按时间先后顺序排列。在实际处理多元时间序列时,需要分析各变量之间潜在的复杂关系和变化规律,才能对其做出有效预测。但现有方法常因为多元时间序列复杂的变化规律难以捕捉,所以导致预测精度较低。如何对其做出有效预测,成为国内外诸多学者关注的问题。

早期传统多元时间序列预测方法主要采用灰色模型<sup>[1]</sup>、自回归综合移动平均模型 (ARMA)<sup>[2]</sup>、自回归移动平均模型 (ARIMA)<sup>[3]</sup> 等。这些方法虽然结构简单,但都属于线性方法,不能较好反映时间序列中非线性因素的影响。为了克服这一缺陷,人们将神经网络引入对多元时间的序列的预测。例如, Meng 等人<sup>[4]</sup> 使用 BP 神经网络实现了对股价走势的预测。这种模型可以较好反映出多元时间序列中非线性因素对结果的影响,但无法联系多元时序数据间的潜在的内部关联。邓玉婧等人<sup>[5]</sup> 采用了循环神经网络 (RNN) 实现了对航班客座率的预测,证明了 RNN 神经网络在时序预测方面的有效性,这种模型可以解决无法联系多元时序数据关系的缺点,但在预测较长序列时经常会出现梯度爆炸的缺陷,使模型无法从训练数据中得到较高的学习效率。廖志伟等人<sup>[6]</sup> 使用 LSTM 神经网络实现了对短期电煤价格预测,这种模型在 RNN 神经网络的基础上增加了不同的门控单元来实现长期信息记忆功能,成功减缓了 RNN 神经网络梯度爆炸问题。周莽等人<sup>[7]</sup> 使用 GRU 神经网络实现了对短期电力负荷的预测,这种模型是 LSTM 模型的一种流行变体,不仅同样可以克服 RNN 模型梯度爆炸问题,其结构较 LSTM 模型还更为简易,更易训练。BiLSTM 神经网络作为近年来神经网络中的新型模型<sup>[8,9]</sup>,它创新性地将一个正向 LSTM 模型和一个反向 LSTM 模型相结合,实现了特征之间潜在关系的双向挖掘。与 LSTM 相比, BiLSTM 模型在预测多元时序序列时,其预测精度明显提高。江知航等人<sup>[10]</sup> 使用 BiLSTM 神经网络对棉花价格进行了预测,但没有考虑多维特征与预测目标的相关性以及无关特征对预测结果的影响,导致预测精度并未大幅提升。陈宇韶等人<sup>[11]</sup> 在预测股票价格前使用了 PCC 相关系数,在庞杂多维特征中分析出与预测目标高度相关的特征,消除了无关特征对预测结果的影响,降低了模型复杂度的同时有效地改善了最终的预测效果。

基于以上,为了更有效地对多元时序序列进行预

测,本文将 PCC-BiLSTM-GRU-Attention 组合模型的预测方法进行了探讨和验证。具体来说,该方法首先利用 PCC 相关系数对输入特征进行筛选,删除无关变量,避免无关变量对结果精度的影响。其次使用 BiLSTM 神经网络双向获取输入特征之间潜在的关联关系。最后使用 GRU 神经网络融合 Attention 机制,进一步深入学习双向时序特征的变化规律,更精准的捕获关键时刻的信息以及非线性动力系统的复杂变化规律。为了验证此方法在多元时间序列中的预测效果,本文以股票价格预测<sup>[12-15]</sup> 为实验场景,以各模型的平均绝对百分比误差值和决定系数值作为衡量标准进行对比实验,验证结果表明,相比于其他 5 种模型, PCC-BiLSTM-GRU-Attention 组合模型的预测方法拟合程度更高,预测效果更好。

## 1 PCC-BiLSTM-GRU-Attention 组合预测模型

### 1.1 PCC 相关系数

根据文献 [16] 可知,在预测多元时间序列前,建立初始化指标体系后进行相关性筛选可以有效减少的数据复杂度和无关变量对预测结果的影响。Cai 等人<sup>[17]</sup> 根据 PCC 相关系数计算出变量间的相关程度来删除冗余变量。因此本文选择 PCC 相关系数对实验进行特征选择, PCC 相关系数的公式为:

$$r_{X,T} = \frac{\sum_{i=1}^k X_i Y_i - \frac{\sum_{i=1}^k X_i \sum_{i=1}^k Y_i}{k}}{\sqrt{\sum_{i=1}^k X_i^2 - \frac{\left(\sum_{i=1}^k X_i\right)^2}{k}}} \sqrt{\sum_{i=1}^k Y_i^2 - \frac{\left(\sum_{i=1}^k Y_i\right)^2}{k}} \quad (1)$$

其中,  $r_{X,Y}$  为变量  $X, Y$  的相关系数,  $X_i$  为自变量,  $Y_i$  为因变量。变量的相关程度可由相关系数的绝对值表示,具体如表 1 所示。

相关系数绝对值	0-0.2	0.2-0.5	0.5-1
相关程度	无关	弱相关	强相关

### 1.2 BiLSTM 神经网络

BiLSTM 神经网络由两个 LSTM 神经网络叠加组成<sup>[18]</sup>。LSTM 神经网络是另一种 RNN 神经网络,在处

理时间序列方面具备显著优势<sup>[19]</sup>,它具有遗忘门、输入门、输出门和细胞状态等部分,其结构如图1所示.细胞状态用来存储和传递时序特征信息,3种门结构在训练过程中会学习如何去保存、遗忘和输出时序特征信息<sup>[20]</sup>.这种结构使得在LSTM神经网络中早期单元的特征信息一样可以传递至后期单元的细胞状态中,实现了LSTM神经网络可长时间记忆特征信息的优势.

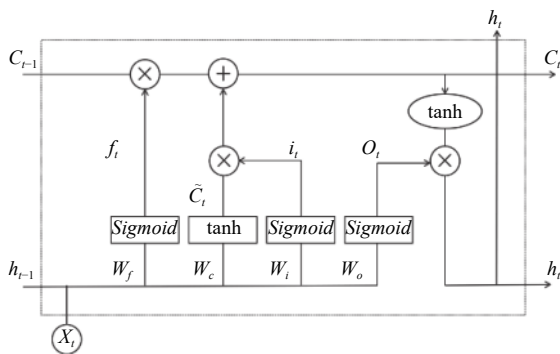


图1 LSTM神经网络单元结构图

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

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

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

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

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

$$h_t = o_t \cdot \tanh(C_t) \quad (7)$$

其中,  $W_f$ 、 $W_i$ 、 $W_c$ 、 $W_o$ 和 $b_f$ 、 $b_i$ 、 $b_c$ 、 $b_o$ 分别对应各门结构的权重矩阵与偏置,其运行原理首先是遗忘门通过式(2)计算当前细胞状态需要遗忘的信息 $f_t$ ,然后输入门通过式(3)计算当前细胞状态需要存放的信息 $i_t$ ,并通过式(4)、式(5)在tanh层中创建候选值加入到细胞状态中并且将单元状态从 $C_{t-1}$ 更新为 $C_t$ ,最后输出门通过式(6)、式(7)计算出当前LSTM单元的输出 $h_t$ .

图2是BiLSTM神经网络结构图,它由一个前向LSTM计算层一个为后向LSTM计算层.这种结构可对时序序列关系进行较长时间的双向挖掘,提取双向时序特征.

### 1.3 GRU神经网络

GRU神经网络对LSTM内部进行了有效简化,引

入了重置门、更新门的概念,这使其在保留原LSTM神经网络预测效果的同时变得更易训练<sup>[21]</sup>,其结构如图3所示.

$$r_t = \text{Sigmoid}(W_r \cdot [h_{t-1}, X_t] + b_r) \quad (8)$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t * h_{t-1}, X_t] + b_h) \quad (9)$$

$$z_t = \text{Sigmoid}(W_z \cdot [h_{t-1}, X_t] + b_z) \quad (10)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (11)$$

其中,  $X_t$ 代表当前时刻的输入,  $h_{t-1}$ 代表上一时刻GRU单元的隐藏状态,  $h_t$ 代表当前GRU单元计算出的隐藏状态,  $W_r$ 、 $W_h$ 、 $W_z$ 和 $b_r$ 、 $b_h$ 、 $b_z$ 为各门结构对应的权重矩阵和偏置,GRU神经网络单元的运行原理首先是通过式(8)计算出重置门的取值,再通过式(9)计算候选状态.若重置门取值接近0,则候选值会选择忽略前一个隐藏状态的信息,这有效地使隐藏状态具有丢弃不相关的特征信息的能力.计算出候选值后,通过式(10)使更新门控制特征信息需要传递多少信息至当前隐藏状态,实现了单元的长期记忆功能,最后通过式(11)计算出当前隐藏状态的最值.

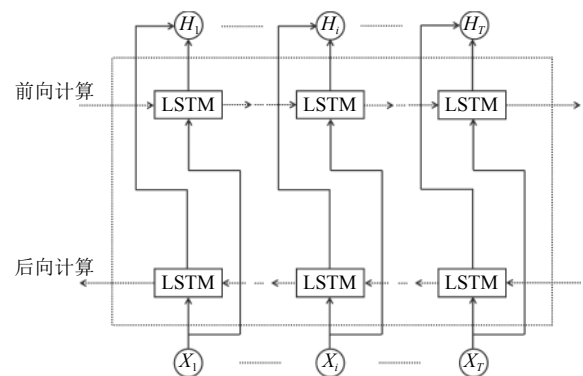


图2 BiLSTM神经网络结构图

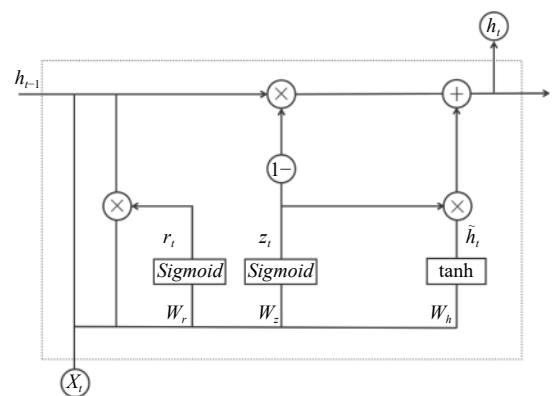


图3 GRU神经网络单元结构图

### 1.4 Attention 注意力机制

Attention 机制通过不同特征对于预测结果影响的大小来为神经网络中的隐藏单元赋予不同权重,使对预测结果有突出影响的因素获得更高关注,为后续再次处理数据提供更好的参考<sup>[22-24]</sup>,大大减少了对资源的消耗,提高了预测精度,其结构具体如图4所示。

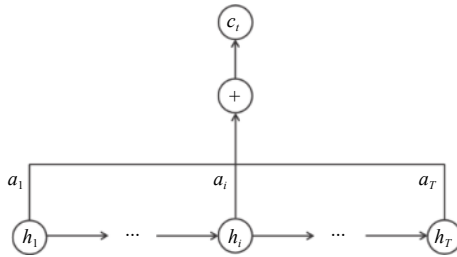


图4 Attention 机制网络结构图

$$e_i = \tanh(W h_i + b) \quad (12)$$

$$a_i = \frac{\exp(e_i)}{\sum_{j=1}^T \exp(e_j)} \quad (13)$$

$$c_t = \sum_{i=1}^T a_i h_i \quad (14)$$

其中,  $W$ 、 $b$  为权重系数和偏置系数,  $h_i$  为 GRU 神经网络中  $t$  时刻中输出的第  $i$  个隐藏单元状态值. Attention 机制首先根据 GRU 神经网络输出每个隐层单元的值  $h_i$  计算式 (12) 后得到每个隐层单元的得分  $e_i$ , 再通过式 (13) 对注意力得分值进行归一化, 最终计算式 (14) 将归一化后的得分值  $a_i$  和隐层单元  $h_i$  进行加权求和得到  $t$  时刻这一时间步长内的隐藏层状态值  $c_t$ .

### 1.5 PCC-BiLSTM-GRU-Attention 组合预测模型

PCC-BiLSTM-GRU-Attention 组合模型的预测方法主要由 6 个阶段构成, 具体层级结构如图5所示。

(1) PCC 相关系数: 将原始数据中的自变量逐一与预测的因变量进行 PCC 相关性分析, 删除与因变量无关的特征, 具体表示为:

$$X_i = [x_1, \dots, x_i, \dots, x_N, p] \quad (15)$$

$$X_i = [x_1, \dots, x_i, \dots, x_n, p] \quad (16)$$

其中,  $X_i$  表示某时刻中的第  $i$  天的输入序列,  $x_i$  为第  $i$  个特征值,  $N$  为未降维选优时的总特征数,  $n$  为降维后的总特征数.  $p$  为某时刻第  $i$  天的预测值, 未经过筛选时

$X_i$  可表示为式 (15), 经过 PCC 相关性分析筛选出不相关的特征后,  $X_i$  可表示为式 (16).

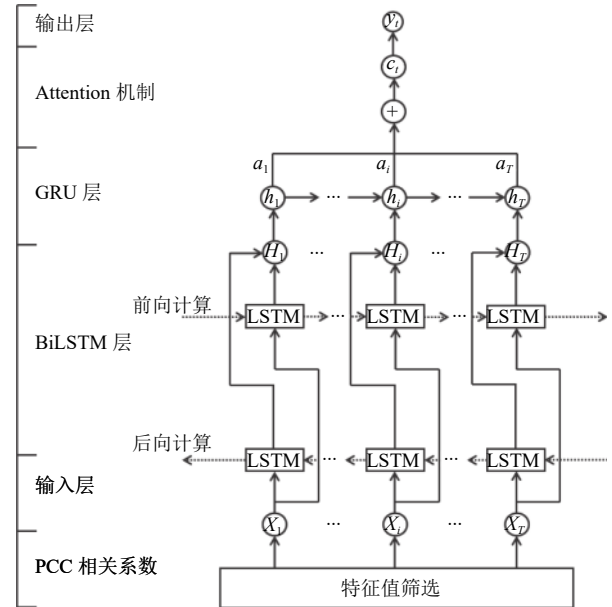


图5 PCC-BiLSTM-GRU-Attention 组合模型预测方法的层级结构

(2) 输入层: 将经过特征值筛选后的特征以及因变量作为本文模型的输入, 在  $t$  时刻中的输入序列具体为:

$$X_t = [X_1, \dots, X_i, \dots, X_T]^T \quad (17)$$

其中,  $X_t$  为  $t$  时刻的输入序列,  $T$  为时间步数.

(3) BiLSTM 层: 在前向层中输入正向  $X_t$  序列  $[X_1, \dots, X_i, \dots, X_T]^T$  计算出该层第  $i$  个输出的隐藏状态  $\vec{H}_i$ , 在后向层中输入反向  $X_t$  序列  $[X_T, \dots, X_i, \dots, X_1]^T$  计算出该层第  $i$  个输出的隐藏状态  $\overleftarrow{H}_i$ :

$$\vec{H}_i = \overrightarrow{LSTM}(\vec{H}_{i-1}, X_i, \vec{C}_{i-1}) \quad (18)$$

$$\overleftarrow{H}_i = \overleftarrow{LSTM}(\overleftarrow{H}_{i+1}, X_i, \overleftarrow{C}_{i+1}) \quad (19)$$

$$H_i = [\vec{H}_i \oplus \overleftarrow{H}_i] \quad (20)$$

$$H_t = [H_1, \dots, H_i, \dots, H_T]^T \quad (21)$$

其中, 通过式 (18) 使用对应  $t$  时刻的输入序列  $X_t$ 、第  $i-1$  个输入时间步的前向 LSTM 层中的隐藏单元状态  $\vec{H}_{i-1}$  以及第  $i-1$  个输入时间步的前向 LSTM 层中的细胞状态  $\vec{C}_{i-1}$ , 计算出在  $t$  时刻的前向 LSTM 层中第  $i$  个输出的隐藏状态  $\vec{H}_i$ . 通过式 (19) 使用对应  $t$  时刻的输



入序列 $X_t$ 、第 $i+1$ 个输入时间步的后向 LSTM 层中的隐藏单元状态 $\overleftarrow{H}_{i+1}$ 以及第 $i+1$ 个输入时间步的后向 LSTM 层中的细胞状态 $\overleftarrow{C}_{i+1}$ , 计算出在 $t$ 时刻的后向 LSTM 层中第 $i$ 个输出状态 $\overleftarrow{H}_i$ , 通过式(20)计算出 BiLSTM 层在 $t$ 时刻的第 $i$ 个隐藏单元状态值, 将式(21)中 $H_t$ 最终作为 $t$ 时刻上提取的双向时序序列。

(4) GRU: 将 BiLSTM 层中输出的双向时序序列 $H_t$ 作为 GRU 层的输入序列进行进一步学习:

$$h_i = GRU(H_{i-1}, H_i) \quad (22)$$

$$h_t = [h_1, \dots, h_i, \dots, h_T] \quad (23)$$

其中, 记 $h_i$ 为 $t$ 时刻中第 $i$ 个 GRU 层的隐层状态, 式(23)中 $h_t$ 为最终在 $t$ 时刻上 GRU 层的隐层状态。

(5) Attention 机制: 将 GRU 层输出的隐层状态 $h_t$ 作为 Attention 机制的输入, 每个隐层单元的得分 $e_i$ 、归一化后得分 $a_i$ 和最终 Attention 层输出 $c_t$ 由式(12)~式(14)计算而来。

(6) 输出层: 选择 Sigmoid 函数作为激活全连接层的函数, 输出最终 $t$ 时刻最终预测值, 记作 $y_t$ :

$$y_t = \text{Sigmoid}(\omega c_t + b) \quad (24)$$

## 2 实验分析

### 2.1 数据来源

本文选取某银行从 2006 年 1 月 5 日至 2019 年 6 月 5 日这一时段内所有交易日的股指数据作为实验数据集。采集数据主要包括: 开盘价、最高价、最低价、涨幅率、振幅率、总手数、交易金额、换手率、成交次数和收盘价, 最终将股票收盘价作为目标变量。此数据集大且全面, 具有多元特性, 在连续的时间序列上呈现一定的变化规律。本文将数据集选取前 80% 的数据为训练集, 后 20% 的数据为测试集。

### 2.2 数据预处理

本文对数据的预处理分为两个步骤, 归一化和特征数据的降维选优:

(1) 为了消除特征参数之间数值与单位的差异性, 以对每个特征进行同等对待, 本文对原始特征数据进行归一化处理来提升预测精度, 归一化公式如下:

$$x_i = \frac{x_i - \min}{\max - \min} \quad (25)$$

其中,  $\min$  代表特征 $x$ 在时序序列中的最小值,  $\max$  代表在时序序列中的最大值。

(2) 将归一化后的特征数据进行 PCC 相关性分析。如表 2 可以看出各特征与股票收盘价的 PCC 相关程度, 为了消除无关变量对模型预测带来的噪声影响, 决定选取开盘价、最高价、最低价、总手数、交易金额和成交次数这些具有相关性的因素作为神经网络输入参数, 删除无关因素来实现对特征数据的降维选优。

表 2 不同特征与股票价格的相关系数

特征	PCC相关系数
开盘价	0.998
最高价	0.999
最低价	0.999
涨幅率	0.004
振幅率	-0.155
总手数	0.534
交易金额	0.542
换手率	-0.189
成交次数	0.564

### 2.3 参数设置

基于本文提出方法训练中的优化目标为平均绝对百分比误差 MAPE 最小化, 即决定采用 Adam 优化算法对模型中的权重进行迭代更新。模型各参数选用经验数据且通过多次实验进行调整, 最终确定 BiLSTM 神经网络层的隐层节点个数为 64 个, GRU 神经网络层的隐层节点个数为 16 个, 迭代次数为 100 次, batch-size 大小为 64, 滑动窗口长度为 5, 学习率为 0.01。

### 2.4 评价指标

本文选用的评价指标为平均绝对百分比误差 (MAPE) 和决定系数 ( $R^2_{\text{score}}$ ) 两种指标来评价模型的优劣性, 两种指标公式如下:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_{\text{true}}(i) - y_{\text{pre}}(i)}{y_{\text{true}}(i)} \right| \times 100\% \quad (26)$$

$$R^2_{\text{score}} = 1 - \frac{\sum_{i=1}^n (y_{\text{true}}(i) - y_{\text{pre}}(i))^2}{\sum_{i=1}^n \left( y_{\text{true}}(i) - \frac{1}{n} \sum_{i=1}^n y_{\text{true}}(i) \right)^2} \quad (27)$$

其中,  $n$  为样本个数,  $y_{\text{true}}(i)$  为 $i$ 时刻的实际值,  $y_{\text{pre}}(i)$  为 $i$ 时刻的预测值。MAPE 的值代表预测值与真实值产生的误差值与真实值之间的比例, 当比例越接近 0% 时, 表明预测误差越小。  $R^2_{\text{score}}$  的值表示预测值曲线与真实值曲线的拟合优度, 当  $R^2_{\text{score}}$  的值越接近 1 时拟

合优度越好.

### 2.5 多模型对比

为了验证 PCC-BiLSTM-GRU-Attention 组合模型的预测方法的预测性能, 将该方法与 BP、LSTM、GRU、BiLSTM-GRU、BiLSTM-GRU-Attention 等神经网络模型放在股票价格预测场景中进行对比实验, 观察不同模型对某银行测试集的预测效果. 为了减少结果中的偶然因素对各神经网络精度的影响, 本文对每种神经网络进行了多次实验, 记录下每种模型在测试集预测的最优结果, 实验结果如表 3 所示.

表 3 不同神经网络预测结果对比

特征	MAPE (%)	决定系数
BP	9.769	0.623
LSTM	7.946	0.698
GRU	7.853	0.758
BiLSTM-GRU	4.700	0.938
Attention-BiLSTM-GRU	2.856	0.946
PCC-Attention-BiLSTM-GRU	2.484	0.966

从表 3 可观察到, PCC-BiLSTM-GRU-Attention 组合模型的预测方法在 MAPE 与决定系数两种评估指标上计算的结果均优于其他神经网络. 该方法的 MAPE 指标为 2.484%, 与 BP、LSTM、GRU、BiLSTM-GRU、BiLSTM-GRU-Attention 等神经网络相比, 其 MAPE 指标分别下降了 7.285%、5.462%、5.369%、2.216%、0.372%. 该方法的决定系数指标为 0.966, 与 BP、LSTM、GRU、BiLSTM-GRU、BiLSTM-GRU-Attention 等神经网络进行对比后, 决定系数分别提高了 0.343、0.268、0.208、0.028、0.02, 由此可见, 实验的结果验证了 PCC-BiLSTM-GRU-Attention 组合模型的预测方法可以有效提高对多元时间序列的预测. 此外, 将 PCC-BiLSTM-GRU-Attention 组合模型的预测方法与 BiLSTM-GRU-Attention 神经网络模型相比较, MAPE 指标降低了 0.372%, 决定系数提高了 0.02, 证明了使用 PCC 相关性系数剔除无关变量可有效减少其对预测精度的影响. 将 BiLSTM-GRU-Attention 神经网络模型与 BiLSTM-GRU 神经网络模型相比较, MAPE 指标降低了 1.844%, 决定系数提高了 0.008, 证明了 Attention 机制神经网络的加入能够有效捕捉关键时刻的信息. 图 6 为 PCC-BiLSTM-GRU-Attention 组合模型的预测方法与其他模型的预测结果图, 可以看出该方法相较于其他模型预测的结果同真实值的波动更为拟合.

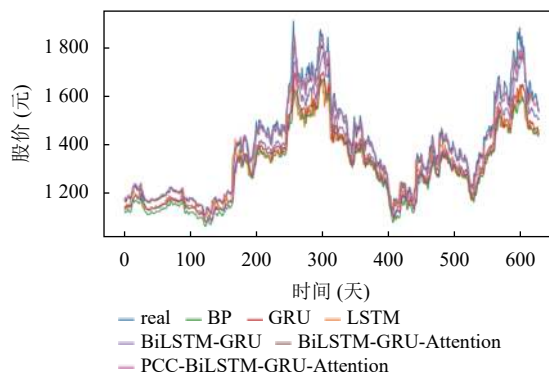


图 6 不同模型预测结果图

### 3 结论

本文将 PCC-BiLSTM-GRU-Attention 组合模型的预测方法进行了探讨和验证. 该方法首先利用 PCC 相关系数将特征数据进行筛选, 删除冗余无关特征对实验结果精度的影响, 实现多元数据的降维选优. 其次选取 BiLSTM 神经网络双向学习时序序列的变化规律, 再使用 GRU 融合 Attention 机制进一步深度学习双向时序特征的变化规律, 同时更精准的捕获关键时刻的信息, 最后通过全连接层输出最后预测值. 本文结合 PCC 相关系数、BiLSTM 神经网络、GRU 神经网络、Attention 机制进一步提高了对多元时序序列的预测精度. 将该方法通过实验与 BP、LSTM、GRU、BiLSTM-GRU、BiLSTM-GRU-Attention 五种模型进行对比, 验证结果表明, PCC-BiLSTM-GRU-Attention 组合模型的预测方法在 MAPE 和决定系数两种指标上有显著优势, 与真实值具有更高的拟合性, 实际预测精度最佳.

#### 参考文献

- 成枢, 郭祥琳, 冯东恒. 改进的 GM-AR 组合模型在地铁沉降预测中的应用. 测绘工程, 2018, 27(5): 59-63.
- Medel CA. A comparison between direct and indirect seasonal adjustment of the Chilean GDP 1986-2009 with X-12-ARIMA. Journal of Business Cycle Research, 2018, 14(1): 47-87. [doi: 10.1007/s41549-018-0023-3]
- Menculini L, Marini A, Proietti M, et al. Comparing prophet and deep learning to ARIMA in forecasting wholesale food prices. Forecasting, 2021, 3(3): 644-662. [doi: 10.3390/forecast3030040]
- Meng S, Shang L, Wu YD, et al. Prediction and analysis of stock price trend based on BP neural network. International Journal of Computational and Engineering, 2017, 2(4):

- 206–209.
- 5 邓玉婧, 武志昊, 林友芳. 基于多粒度时间注意力 RNN 的航班客座率预测. 计算机工程, 2020, 46(1): 294–301.
  - 6 廖志伟, 陈琳韬, 黄杰栋, 等. 基于特征空间变换与 LSTM 的中短期电煤价格预测. 东北大学学报 (自然科学版), 2021, 42(4): 483–493. [doi: [10.12068/j.issn.1005-3026.2021.04.005](https://doi.org/10.12068/j.issn.1005-3026.2021.04.005)]
  - 7 周莽, 高僮, 李晨光, 等. GRU 神经网络短期电力负荷预测研究. 科技创新与应用, 2018, (33): 52–53, 57. [doi: [10.3969/j.issn.2095-2945.2018.33.021](https://doi.org/10.3969/j.issn.2095-2945.2018.33.021)]
  - 8 李泽文, 胡让, 刘湘, 等. 基于 PCA-DBILSTM 的多因素短期负荷预测模型. 电力系统及其自动化学报, 2020, 32(12): 32–39.
  - 9 Chen QP, Xie QS, Yuan QN, *et al.* Research on a real-time monitoring method for the wear state of a tool based on a convolutional bidirectional LSTM model. Symmetry, 2019, 11(10): 1233. [doi: [10.3390/sym11101233](https://doi.org/10.3390/sym11101233)]
  - 10 江知航, 王艳霞, 颜家均, 等. 基于 BILSTM 的棉花价格预测建模与分析. 中国农机化学报, 2021, 42(8): 151–160.
  - 11 陈宇韶, 唐振军, 罗扬, 等. 皮尔森优化结合 Xgboost 算法的股价预测研究. 信息技术, 2018, (9): 84–89.
  - 12 邓飞燕, 岑少琪, 钟凤琪, 等. 基于 LSTM 神经网络的短期价格趋势预测. 计算机系统应用, 2021, 30(4): 187–192. [doi: [10.15888/j.cnki.csa.007855](https://doi.org/10.15888/j.cnki.csa.007855)]
  - 13 Budiharto W. Data science approach to stock prices forecasting in Indonesia during Covid-19 using long short-term memory (LSTM). Journal of Big Data, 2021, 8(1): 47. [doi: [10.1186/s40537-021-00430-0](https://doi.org/10.1186/s40537-021-00430-0)]
  - 14 包振山, 郭俊南, 谢源, 等. 基于 LSTM-GA 的股票价格涨跌预测模型. 计算机科学, 2020, 47(S1): 467–473.
  - 15 Yu ZX, Qin L, Chen YJ, *et al.* Stock price forecasting based on LLE-BP neural network model. Physica A: Statistical Mechanics and its Applications, 2020, 553: 124197. [doi: [10.1016/j.physa.2020.124197](https://doi.org/10.1016/j.physa.2020.124197)]
  - 16 林娜娜, 秦江涛. 基于随机森林的 A 股股票涨跌预测研究. 上海理工大学学报, 2018, 40(3): 267–273, 301.
  - 17 Cai JC, Xu K, Zhu YH, *et al.* Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest. Applied Energy, 2020, 262(6): 114566–190.
  - 18 张维, 袁绍欣, 陶建军, 等. 基于多元因素的 Bi-LSTM 高速公路交通流预测. 计算机系统应用, 2021, 30(6): 184–190. [doi: [10.15888/j.cnki.csa.007969](https://doi.org/10.15888/j.cnki.csa.007969)]
  - 19 宋予佳, 张健, 邢珺. 基于长短期记忆网络构建短时车流预测模型. 公路, 2019, 64(7): 224–229.
  - 20 Zhao Z, Chen WH, Wu XM, *et al.* LSTM network: A deep learning approach for short-term traffic forecast. IET Intelligent Transport Systems, 2017, 11(2): 68–75. [doi: [10.1049/iet-its.2016.0208](https://doi.org/10.1049/iet-its.2016.0208)]
  - 21 Wang YS, Liao WL, Chang YQ. Gated recurrent unit network-based short-term photovoltaic forecasting. Energies, 2018, 11(8): 2163. [doi: [10.3390/en11082163](https://doi.org/10.3390/en11082163)]
  - 22 朱星嘉, 李红莲, 吕学强, 等. 一种改进的 Attention-based LSTM 特征选择模型. 北京信息科技大学学报, 2018, 33(2): 54–59.
  - 23 Yang F, Zhang HY, Tao SM. Travel order quantity prediction via attention-based bidirectional LSTM networks. The Journal of Supercomputing, 2021. [doi: [10.1007/s11227-021-04032-8](https://doi.org/10.1007/s11227-021-04032-8)]
  - 24 Qin JQ, Zhang Y, Fan SX, *et al.* Multi-task short-term reactive and active load forecasting method based on attention-LSTM model. International Journal of Electrical Power & Energy Systems, 2022, 135: 107517.

(校对责编: 牛欣悦)