

# 基于TBA融合模型的股票指数预测

韩迪<sup>1</sup>, 郭维<sup>1</sup>, 廖凯<sup>2</sup>, 孙传一<sup>1</sup>, 汪勃澄<sup>1</sup>, 林坤玲<sup>2</sup>

1) 广东金融学院信用管理学院, 广东广州 510521; 2) 广东金融学院金融信用大数据研究中心, 广东广州 510521

**摘要:** 由于股票市场的波动性和复杂性特点, 股指预测一直是金融预测研究中的难点. 长短期记忆(long short-term memory, LSTM)网络模型常用于金融指数的预测中, 但该模型在长时间序列上易导致数据信息利用不充分. 利用双向长短期记忆(bidirectional LSTM, BiLSTM)网络模型、时间卷积网络(temporal convolutional network, TCN)和注意力机制协同提高了模型识别以及提炼长时间序列数据特征的能力, 构建一种新型股指预测融合模型TCN-BiLSTM-attention(简称TBA模型). 以中国境内近30年的公开股指数据集为例, 将TBA模型与目前金融类主流的机器学习、神经网络预测算法以及kaggle竞赛排行前列的模型在上证指数、沪深300指数与创业板指数进行预测对比和消融实验. 结果显示, 相较于对照实验组的平均预测误差, TBA模型有明显降低且表现稳定, 兼具准确性与鲁棒性. 研究结果可广泛用于基于时间序列的多种金融预测场景.

**关键词:** 数字经济; 股指预测; 长短期记忆网络; 时间卷积网络; 注意力机制; 消融实验

**中图分类号:** TP183; F208; TP3391 **文献标志码:** A **DOI:** 10.3724/SP.J.1249.2023.06665

## Stock index forecasting based on TBA fusion model

HAN Di<sup>1</sup>, GUO Wei<sup>1</sup>, LIAO Kai<sup>2</sup>, SUN Chuanyi<sup>1</sup>,  
WANG Bocheng<sup>1</sup>, and LIN Kunling<sup>2</sup>

1) School of Gredit Management, Guangdong University of Finance, Guangzhou 510521, Guangdong Province, P. R. China

2) Finance Credit Big Data Research Center, Guangdong University of Finance, Guangzhou 510521, Guangdong Province, P. R. China

**Abstract:** Due to the volatility and complexity of stock market, stock index prediction has always been a challenge in the field of financial forecasting. Long short-term memory (LSTM) network model is commonly used in financial index forecasting, however, this model has some limitations in long time series which may lead to insufficient use of data information. By using bidirectional long short-term memory network model (BiLSTM), temporal convolutional network (TCN) and attention mechanism, a novel fusion model named TCN-BiLSTM-attention (hereinafter referred to as TBA model) for stock index forecasting is constructed to further improve the ability of recognition and extraction of long time series data features of the model. Taking the public stock index datasets within China for nearly 30 years as an example, the TBA model is compared and ablated with the current mainstream machine learning and neural network prediction algorithms in finance as well as the top ranked models in Kaggle. The experimental results show the TBA model has significantly lower average prediction error and more stable performance compared with the average error baseline of the control experimental group in the forecasting of SSE, CSI 300 and GEI multi-day

**Received:** 2022-12-17; **Accepted:** 2023-02-15; **Online (CNKI):** 2023-07-12

**Foundation:** Educational Commission Key Program of Guangdong Province of China under Grants (2020ZDZX3066); Philosophy and Social Science Plan of Guangdong Province (GD21YYJ02, GD19CJLJ01)

**Corresponding author:** Associate professor GUO Wei (18-090@gduf.edu.cn)

**Citation:** HAN Di, GUO Wei, LIAO Kai, et al. Stock index forecasting based on TBA fusion model [J]. Journal of Shenzhen University Science and Engineering, 2023, 40(6): 665-673. (in Chinese)



indices, thus this model can be used in a variety of financial forecasting scenarios based on time series.

**Key words:** digital economy; stock index prediction; long short-term memory network; temporal convolutional network; attention mechanism; ablation experiment

股票市场作为资本市场的重要组成部分,在开拓投融资渠道、优化社会资金配置上起到关键作用.股票价格与股票价格指数预测(以下简称股指预测)是金融领域的焦点问题之一.准确预测股指走势也是金融预测问题中的一个重难点.由于股票市场具有复杂性和波动性,且金融时间序列作为一种特殊的时间序列,兼具时间序列的共性与金融数据的特殊性,其波动规律难以被把握<sup>[1]</sup>.为提高预测金融时间序列变化的准确度,已有学者提出了各类预测方法.现有的预测方法从算法类型上可以分为数理统计方法和机器学习方法.

以数理统计方法为主的线性预测模型的开发与优化使金融时间序列的预测研究取得了很大进步.研究发现,差分整合移动平均自回归模型具有较好的股价预测能力,且能有效预测风险资产的价格波动区间<sup>[2-3]</sup>.广义自回归条件异方差模型能有效预测股指期货收益波动率<sup>[4]</sup>.上述工作对于金融时间序列走势的判断简单易行且准确性较高,但大部分线性预测模型是通过拟合股价及其影响因素之间的时序关系来预测股价未来走势,缺乏捕捉金融时间序列中非线性关系的能力,应用场景较局限.

随着机器学习的快速发展,基于循环神经网络(recurrent neural network, RNN)改进的长短期记忆(long short-term memory, LSTM)网络模型已被大量用于金融时间序列预测. CHEN 等<sup>[5]</sup>使用 LSTM 模型令股票收益预测的准确率比随机预测提升了 12.9%; SELVIN 等<sup>[6]</sup>证明 LSTM 模型预测股票价格的表现优于线性模型;杨青等<sup>[7]</sup>发现 LSTM 模型对股指的预测兼具精确性与稳定性.可见, LSTM 模型能够挖掘使用变量之间的深层次关系并调用大量潜在有价值的非线性序列数据,从而提高了金融时间序列预测的准确度,逐步取代了线性预测模型.

但是, LSTM 模型仅能抓取时间序列的单向特征,且处理长时间序列的能力较弱.研究发现,基于 LSTM 模型构建的组合模型能够进一步提高金融时间序列的预测精度<sup>[8-11]</sup>.双向长短期记忆(bidirectional LSTM, BiLSTM)网络模型的正向和反向的训练层能有效利用时间序列的双向关系,提高了预测准确率<sup>[12]</sup>.因为 BiLSTM 模型能够捕获深层次关联信息<sup>[13]</sup>并且利用金融时间序列的双向关系<sup>[12]</sup>,受此

启发本研究使用 BiLSTM 模型作为基础模型,针对金融时间序列的特征,从特征提炼与特征识别两个角度优化 BiLSTM 模型.为提高 BiLSTM 模型提炼特征的能力,采用卷积神经网络(convolutional neural network, CNN)下的一维时间卷积网络(temporal convolutional network, TCN)处理金融时间序列.模型融合 TCN 后能有效提高提炼金融时间序列特征的能力并减少预测误差<sup>[14-16]</sup>.为提高 BiLSTM 模型识别特征的能力,将近年来广泛应用于图像识别<sup>[17-18]</sup>、自然语言处理<sup>[19]</sup>以及推荐系统<sup>[20]</sup>等领域的注意力机制融入 BiLSTM 模型.模型融合注意力机制后,能向重要特征赋予较高的权重,实现信息处理资源的高效分配<sup>[21-22]</sup>.

本研究构建了一种性能准确且稳定的股指预测融合模型,记为 TCN-BiLSTM-attention (TBA)模型.该模型以 BiLSTM 层捕获金融时间序列的深层次关系;再以 TCN 层提炼金融时间序列的数据特征,降低整体运算量;最后以注意力层识别数据特征,为重要特征赋权.为验证 TBA 模型的准确性与鲁棒性,选取覆盖范围广且具有一定区分度的数据集作为实验数据,并通过消融模型与主流模型进行对照实验.在 FlyAI 竞赛平台关于“A 股股票趋势走向预测”的比赛中, TBA 模型在发稿前仍保持第 2 名的成绩(和第 1 名预测误差小于 1%).

本研究贡献主要有以下两点:①根据金融时间序列的特殊性以及共性,设计了 TBA 融合模型,有效提高了金融股指预测的准确性;②为 TBA 模型设计了自适应机制,可根据数据特点动态调整参数.在处理不同数据集时调整复合模型的权重比例,保证模型性能的稳定.

## 1 模型构建

TBA 模型整体架构如图 1, 整个流程通过 TCN 层提炼股指特征信息, BiLSTM 层提取时序特征并挖掘股指深层次关系,注意力层识别并赋权重要特征.其中,  $x_1, x_2, \dots, x_t$  为输入的带有时间序列的股指数据;  $\alpha_1, \alpha_2, \dots, \alpha_t$  为 LSTM 模型单元输出;  $\nu$  为隐藏层 BiLSTM 得到的输出; dense 为全连接层;  $y$  为模型的预测结果.

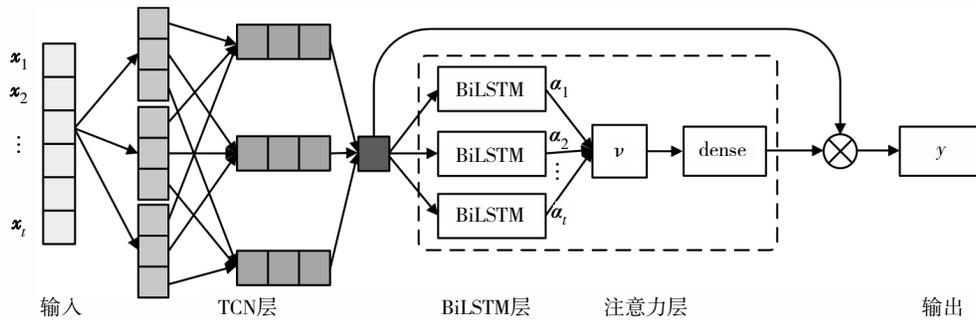


图1 基于时间卷积网络、双向长短期记忆网络与注意力机制融合模型的TCN-BiLSTM-Attention股指预测模型

Fig. 1 TCN-BiLSTM-Attention stock index prediction model based on the fusion model of temporal convolutional neural network, bidirectional long short-term memory network and attention mechanism, which can adjust parameters adaptively according to different stock index data characteristics.

### 1.1 双向长短期记忆网络 BiLSTM

HERMANS 等<sup>[23]</sup>研究表明, RNN 模型的每一层都是一个循环网络, 它接受前一层的隐藏状态作为输入再次输入到神经元, 时间任务执行分层处理. 这种串联结构可保持数据中的依赖关系, 大幅减少了需要训练的神经网络参数. 但 RNN 模型的反向传播算法相对简单, 在长时间序列下, 可循环利用的神经元权重参数会带来很大弊端, 如产生梯度爆炸或梯度消失, 造成长期依赖. LSTM 模型能有效解决 RNN 模型存在的问题<sup>[24]</sup>. LSTM 模型可通过遗忘门、输入门和输出门来实现对信息的存储和更新 (门控由 sigmoid 函数和点乘运算实现), 通过激活函数将输入信号映射到一个取值范围在 0~1 的概率值  $P$  上. 当  $P = 0$  时, 代表上一时刻特征变化得比较缓慢, 无需传递梯度更新参数; 当  $P = 1$  时, 代表上一时刻特征变化较大, 因此, LSTM 模型能够很好地传递梯度, 有效降低了发生梯度消失的概率. 虽然 LSTM 模型通过门控机制维持长期依赖, 解决了 RNN 梯度消失的问题, 但它只对当前时刻之前的历史状态存在依赖关系, 不能利用当前时刻之后的信息. 由于当前的状态也可能和未来的信息存在依赖, 因此, 本研究以正向和反向 LSTM 层构成 BiLSTM 模型以同时利用时间序列的历史信息和未来信息, 从而捕获单个 LSTM 模型无法捕获的附加特征. BiLSTM 模型架构如图 2.

根据金融时间序列的特点, 正向 LSTM 层顺序输入特征信息, 得到正向 LSTM 层隐藏状态; 反向 LSTM 层逆序输入特征信息, 得到反向 LSTM 层隐藏状态; 将正向和反向的隐藏状态进行拼接得到 BiLSTM 模型输出的隐藏状态. BiLSTM 模型输出的隐藏状态在  $t$  时刻的输出为

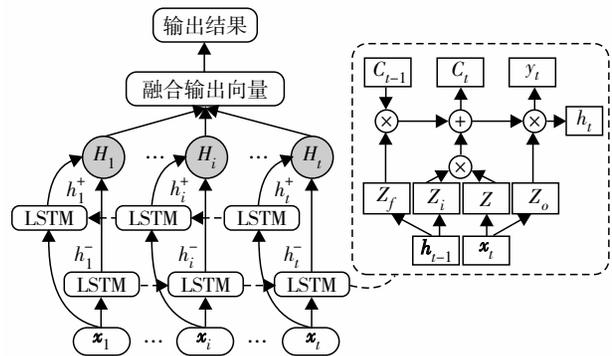


图2 由2个方向相反的LSTM层(正向和反向LSTM层)构成的BiLSTM模型架构

Fig. 2 Diagram of the BiLSTM model architecture consisting of two LSTM layers with opposite directions (positive and negative LSTM layers).

$$H_t = h_t^+ \oplus h_t^- \tag{1}$$

其中,  $h_t^+$  和  $h_t^-$  分别为  $t$  时刻正向和反向 LSTM 层的隐藏状态,  $h_t^+ = f(U^1 h_t + W^1 x_t + b^1)$ ,  $h_t^- = f(U^2 h_t + W^2 x_t + b^2)$ ;  $f$  为隐藏层激活函数;  $W^1$  和  $W^2$  分别为正向 LSTM 层与反向 LSTM 层在  $t$  时刻的权重矩阵;  $U^1$  和  $U^2$  分别为  $t - 1$  时刻和  $t + 1$  时刻的权重矩阵;  $b^1$  和  $b^2$  为偏置向量. 由式(1)可见, 相比单向 LSTM 模型的输出  $h_t^+$ ,  $H_t$  包含更多有效的附加特征, 同时也充分利用了时间序列的历史信息与未来信息.

### 1.2 TCN

CNN 的局部连接、权值共享和池化操作使其拥有较少的连接和参数, 更易于训练<sup>[25-26]</sup>, 这令 CNN 逐渐成为处理时间序列问题时通用的特征提取工具. 本研究旨在探究金融时间序列类数据与图形、语音和文本等二维数据类型的差异, 而 TCN 更加擅长捕捉时序上的依赖关系, 因此更适合采用一维

TCN 进行处理, 同时也可以捕捉到局部信息.

本研究构建的 TCN 模型由输入层、4 个卷积层、2 个池化层与输出层组成, 基本结构如图 3.

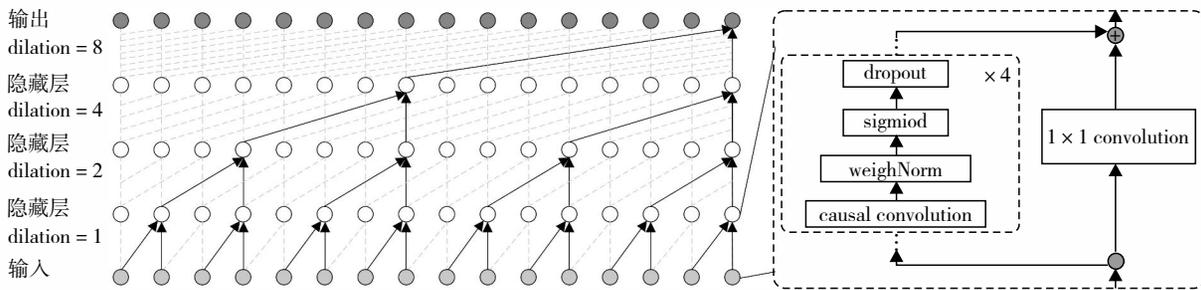


图3 时间卷积网络的基本结构

Fig. 3 The structure of temporal convolutional neural network.

TCN 模型可表示为 1 个一维全卷积网络 (one dimensional fully convolutional networks, 1D FCN) 和因果卷积 (causal convolution) 的结合, 即

$$TCN = 1D\ FCN + \text{causal convolution} \quad (2)$$

与传统 CNN 相比, FCN 将最后的全连接层替换为卷积层, 如图 3 中 causal convolution 所示.

### 1.3 注意力机制

尽管 BiLSTM 模型通过隐藏层在一定程度上缓解了梯度消失问题, 但对于更长的时间序列的泛化能力较为有限. 为保持模型预测的稳定性和准确性, 并更好地关注金融时间序列中相对较新的数据, 本研究将基于时间序列特征的注意力机制引入 BiLSTM 模型的隐藏层和输出层之间, 使得不同隐藏层的输出权重采用乘法门来控制 (类似于 LSTM 中的遗忘机制).

注意力机制可对重要的信息赋予较高权重, 提高信息的利用程度以解决特征识别能力较差的问题, 进而提高模型的预测效果. 对 BiLSTM 模型提取的特征矩阵进行初始化, 获得向量查询  $Q$ 、键  $K$  和值  $V$ , 其中,  $Q$  和  $V$  之间的权重由  $Q$  和  $K$  之间的相似度决定, 则注意力机制定义为

$$\text{attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (3)$$

其中, softmax 是一个将相似度映射到  $0 \sim 1$  的函数;  $\sqrt{d}$  为缩放因子, 用于在高维空间中防止内积过大导致梯度消失.

为更符合金融时间序列的特点, 本研究优化后的注意力机制的设计思路如图 1 中虚线标注的注意力层所示. 注意力层对全连接层 (dense) 整合的 TCN 层与 BiLSTM 层输出的特征序列向量进行加权

其中, 卷积层提取局部特征; 池化层对提取的特征进行压缩、简化网络复杂度, 减少过拟合; 输出层进行分类输出.

后输出股指预测结果.

### 1.4 TBA 模型构建

TBA 模型由 TCN 层、BiLSTM 层与注意力层构成. 其中, TCN 层包括 1 个输入层、4 个一维卷积层、2 个最大池化层和 1 个输出层; BiLSTM 层包括正向与反向 LSTM 层. 模型算法伪代码见图 4.

由图 4 可见, TCN 层对金融时间序列数据进行

<p>输入: 数据集 <math>D</math>, 即股票指数实例集合; 数据特征 <math>d</math></p> <p>输出: 股票指数收盘价预测值 <math>y</math></p>
<p>采用 Z-score 标准将数据集 <math>D</math> 作归一化处理</p> <pre> 1 repeat 2   for each <math>d</math> in <math>D</math> do 3     <math>c_j^k \leftarrow \text{Conv}(d)</math> // <math>c_j^k</math> 为 TCN 提取的特征值 4     if <math>c_j^k</math> 满足 BiLSTM 模型的输入要求 5       <math>X \leftarrow c_j^k</math> // <math>X</math> 为整合后的数据集 6       for <math>k = 1</math> to <math>m</math> in <math>X</math> do 7         <math>h_{t+}</math> <math>\leftarrow \{b^1, h_t, U^1, W^1\}</math> 8         <math>h_{t-}</math> <math>\leftarrow \{b^2, h_t, U^2, W^2\}</math> 9         <math>H_t \leftarrow h_{t+} \oplus h_{t-}</math> 10        if 注意力层获得来自 BiLSTM 模型的 <math>H_t</math> then 11          <math>\alpha_i \leftarrow \text{attention}(Q, K, V)</math> // <math>\alpha_i</math> 为注意力权重 12          <math>y \leftarrow \text{dense}(c_j^k, H_t, \alpha_i)</math> 13        初始化 Adam, earlystop, <math>L_2</math> 等参数 14        设定参数权重列表 <math>Z_i</math>, 约束条件 loss 15        for <math>j</math> in <math>Z_i</math> do 16          更新参数 <math>\leftarrow</math> 迭代操作 17        until epoch 达到最优损失值 18      end for 19    until epoch 达到最优状态                 </pre>

图4 TBA 模型的股指预测算法伪代码

Fig. 4 Pseudo code of stock index forecasting algorithm for TBA model.

局部感知与卷积计算,得到局部关键股指预测特征矩阵后,经全连接层整合后交由输出层分类输出(第2~3行);BiLSTM层获得经TCN层处理后相对稳定且高质量的信息后,通过正向与反向LSTM层捕获时间序列的特征,整合输出特征序列向量(第4~9行);同时,将注意力层引入BiLSTM模型的隐藏层与输出层之间,并对TCN层与BiLSTM层的输出的特征序列向量进行加权后输出股指预测的结果(第10~12行);最后,TBA模型设计了自适应机制(第14行),使TBA模型在处理不同数据集时,能根据数据特征动态调整模型间的比重,使模型可自行选择更为关注全局或局部特征。

## 2 特征工程

### 2.1 数据集选取

金融预测的数据集主要存在两个特点<sup>[27]</sup>:①数据维度存在多样性,即针对某个问题的研究用到的数据集包括指数和个股等多个维度;②内容存在多样性,即使用市场指数数据集进行预测时,选取不同市场的指数或抽取成分股组成数据集调用。考虑到数据集的覆盖范围与区分度,本研究选取上海证券交易所股票价格综合指数(Shanghai stock exchange composite index,简称SSE)、沪深300指数(CSI 300)和创业板指数(growth enterprise index,GEI),分别从维度与内容两个角度验证预测股票指数时TBA模型的准确性和鲁棒性(股指数据来源自wind数据库<http://www.wind.com.cn>):

1) SSE由所有在上海证券交易所挂牌上市的股票(包括A股与B股)为样本,以加权平均法计算,反映了上海证券交易所内所有股票整体价格的涨跌变化情况。

2) CSI 300由沪深市场中规模大且流动性好的最具代表性的300只证券组成,反映沪深市场最具代表性的股票整体价格的涨跌变化。

3) GEI由创业板的所有股票为样本,以加权平均法计算,反映具有高成长性的中小企业和新兴公司股票价格的整体趋势。

### 2.2 数据预处理

数据集特征构建是模型预测能力的天花板,本研究经去伪存真选取处理后的开盘价、前收盘价、最高价、最低价和收盘价5个关键特征作为股指预测的输入变量,建立的数据集特征信息表如表1。

表1 数据集特征信息

Table 1 Dataset characteristics information

字段名称	字段类型	备注	样例
Date	string	时间	2020-01-01
Open	float	开盘价	3 310.644
Closed	float	前收盘	3 325.110
High	float	最高价	3 319.126
Low	float	最低价	3 184.964
Closing	float	收盘价	3 196.768

由于本研究选取的样本数据的值域变化范围较大,为确保模型的准确性(同时避免较小的特征失效),以及降低整体运算量以提高模型的速度,需先对数据集做归一化处理。通过计算每个特征的均值和标准差,将所有特征值转化为标准正态分布的形式。整体来看,归一化前后特征值趋势相同,但归一化处理后特征值的分布较归一化之前更加集中,有效避免了特征失效的问题。

## 3 实验思路与模型选择

### 3.1 TBA模型预测思路

本研究使用股指开盘价、前收盘价、最高价、最低价和收盘价5个指标的日度数据作为模型输入。TBA模型通过学习股票指数历史交易数据信息,对未来1、5和10d内指数收盘价进行预测。

### 3.2 对照实验模型选择及预测思路

在金融领域主流的预测模型中,支持向量回归(support vector regression,SVR)<sup>[28]</sup>是经典的机器学习算法,以非线性近似能力而闻名;RNN<sup>[29]</sup>和多层感知机(multilayer perceptron,MLP)<sup>[30]</sup>作为神经网络算法,也被广泛用于金融预测任务;Informer模型在长期序列预测上表现非常出色<sup>[31]</sup>;极致梯度提升(extreme gradient boosting,XGBoost)算法<sup>[32]</sup>也常在各类人工智能竞赛中表现优异。因此,本研究选取了这5种模型作为对照组,统一设置神经网络参数,并以股指开盘价、前收盘价、最高价、最低价和收盘价等指标作为输入,以未来1d的股指收盘价作为输出,进行预测实验。

### 3.3 消融实验模型选择及预测思路

由于TBA模型是综合了TCN、LSTM、BiLSTM、TCN-BiLSTM和BiLSTM-attention这5种模型的特点和优势,所以为展示TBA融合模型相较于单一模型

的性能优势, 本研究还设计了一系列消融实验, 将 TBA 模型与这 5 个前置模型进行对比, 从而验证组合模型的有效性.

## 4 实验验证与结果分析

本研究将 TBA 模型应用于 SSE、CSI 300 与 GEI 预测, 以验证模型的准确性、鲁棒性, 以及模型自适应机制的有效性.

### 4.1 参数设定

TBA 模型的自适应机制会根据数据特点调整参数, 从而提高金融预测的稳定性与准确度. 模型训练使用 adam 优化器, 学习率为 0.001, 并以 0.02 的衰减来更新学习率. 预测 SSE 时, TCN 层的激活函数为 tanh; 预测 GEI 与 CSI 300 时, TCN 层的激活函数为 Relu. TBA 模型的超参数设置如表 2.

表 2 TBA 模型超参数设置  
Table 2 TBA model hyperparameters

参数名称	含义	取值
batch_size	每轮训练的样本数量	64
time_step	时间步长	1
input_dim	输入特征的维数	5
dropout	每次训练随机丢弃神经元的比率	0.1
epochs	训练次数	500
kernel_size	卷积层中卷积核的长度	1
max_pooling_size_1d	最大池化层窗口大小	32
global_max_pooling_1d	全局最大池化层窗口大小	64

### 4.2 评价指标

为客观评价 TBA 模型性能, 选用均方根误差 (root mean square error, RMSE) 与百分比误差绝对值的标准差<sup>[7]</sup> (standard deviation of absolute percentage error, SDAPE) 两个指标对模型的预测精度与稳定性进行评价. RMSE 值越小代表模型预测精度越高, SDAPE 值越小表示模型预测稳定性越高. RMSE 和 SDAPE 的计算公式分别为

$$E_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2} \quad (4)$$

$$\sigma_{\text{SDAPE}} = \sqrt{\frac{1}{n} \sum_{t=1}^n \left( \left| \frac{\hat{y}_t - y_t}{y_t} \right| - E_{\text{MAP}} \right)^2} \quad (5)$$

其中,  $\hat{y}_t$  为预测值;  $y_t$  为实际值;  $n$  为预测样本个

数;  $t$  为数据序列编号;  $E_{\text{MAP}}$  为平均绝对百分比误差,

$$E_{\text{MAP}} = \frac{1}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t - y_t}{y_t} \right| \quad (6)$$

### 4.3 对照试验

对照实验旨在回答两个问题: ① TBA 模型股指预测的误差是否较主流算法有提升? ② TBA 模型在其他数据集上是否具有普适性, 自适应机制是否有效?

图 5 给出了 TBA、RNN、MLP、XGBoost、SVR 和 Informer 六个模型对 SSE、CSI 300 和 GEI 预测时的 RMSE 与 SDAPE 来验证 TBA 模型的预测精度与稳定性 (详细数值结果可扫描论文末页右下角二维码查看补充材料表 S1 至表 S3). 由图 5 可见, TBA

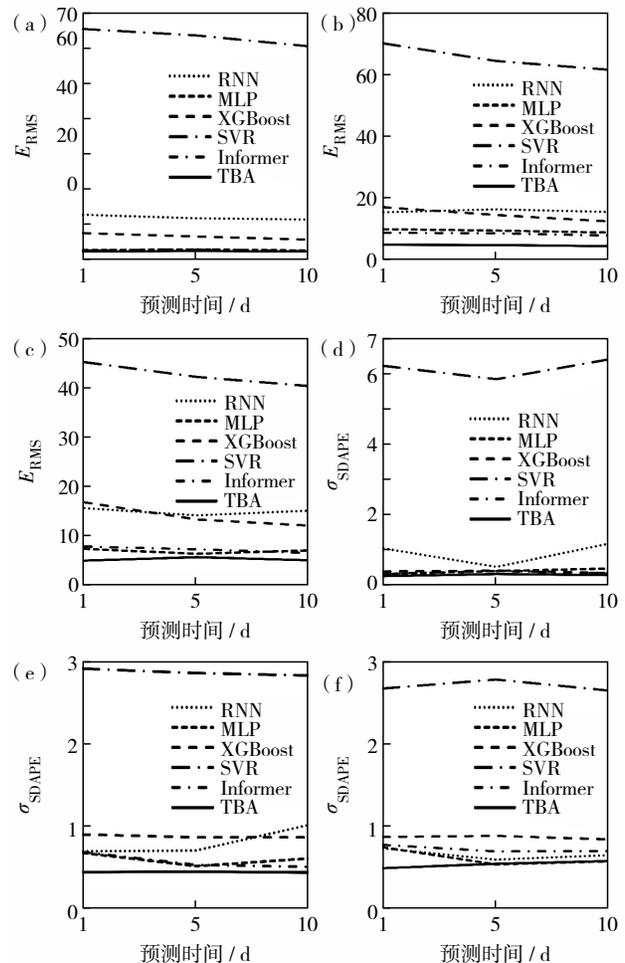


图 5 不同模型 (a) 对 SSE 预测的  $E_{\text{RMS}}$ , (b) 对 CSI 300 预测的  $E_{\text{RMS}}$ , (c) 对 GEI 预测的  $E_{\text{RMS}}$ , (d) 对 SSE 预测的  $\sigma_{\text{SDAPE}}$ , (e) 对 CSI 300 预测的  $\sigma_{\text{SDAPE}}$ , (f) 对 GEI 预测的  $\sigma_{\text{SDAPE}}$

Fig. 5 (a)  $E_{\text{RMS}}$  for SSE, (b)  $E_{\text{RMS}}$  for CSI 300, (c)  $E_{\text{RMS}}$  for GEI, (d)  $\sigma_{\text{SDAPE}}$  for SSE, (e)  $\sigma_{\text{SDAPE}}$  for CSI 300, and (f)  $\sigma_{\text{SDAPE}}$  for GEI by using different models.

模型对SSE、CSI300与GEI指数预测的总体效果最好, 且兼具准确性与稳定性.

在预测精度方面, 当预测时长为1、5和10 d时, TBA模型在SSE预测中的RMSE分别为2.061、2.157和2.038; CSI300预测中的RMSE分别为4.558、4.421和4.106; GEI预测中的RMSE分别为4.897、5.645和5.016. TBA模型的RMSE平均值为3.822, 相较于RNN、MLP、XGBoost、SVR和Informer的RMSE平均值分别降低了10.167、2.318、7.717、53.021与2.028.

在预测稳定性方面, 当预测时长分别为1、5和10 d时, TBA模型在SSE指数预测中的SDAPE分别为0.229%、0.290%和0.261%; CSI300预测中的SDAPE分别为0.429%、0.434%和0.425%; GEI指数预测中的SDAPE分别为0.496%和0.546%和0.581%. TBA模型的SDAPE平均值为0.410%, 比RNN、MLP、XGBoost、SVR和Informer的SDAPE平均值分别降低了0.372%、0.116%、0.286%、3.486%与0.128%.

综上, TBA模型股指预测的误差值较主流算法对比有明显下降, 其整体预测稳定性优于主流算法. TBA模型在预测不同数据集时仍表现优秀, 层与层之间的自适应机制有效, 具有普适性.

#### 4.4 消融实验

消融实验旨在回答TBA融合模型的性能与前置模型的性能相比提升有多大的问题.

本研究通过消融试验比较TBA与TCN、LSTM、BiLSTM、TCN-BiLSTM和BiLSTM-attention模型在对SSE、CSI 300和GEI上不同指标的预测性能, 从而验证融合算法的必要性, 结果如图6(详细数值结果可扫描论文末页右下角二维码查看补充材料表S4至表S6). 由图6可见, TBA模型在消融实验中表现最优, 其预测精度与稳定性均优于对比模型.

由图6可见, 在对SSE、CSI 300与GEI的预测结果中, TBA模型大部分结果为最优值, 总体预测效果最佳. TBA模型的平均 $E_{RMS}$ 值与SDAPE均小于其他模型. TBA模型的RMSE平均值为3.822, 比TCN、LSTM、BiLSTM、TCN\_BiLSTM与BiLSTM\_attention的 $E_{RMS}$ 值平均值分别降低了6.169、8.328、2.898、0.941与2.607. TBA模型的SDAPE平均值为0.410%, 比TCN、LSTM、BiLSTM、TCN\_BiLSTM与BiLSTM\_attention的SDAPE平均值分别降低了0.384%、0.799%、

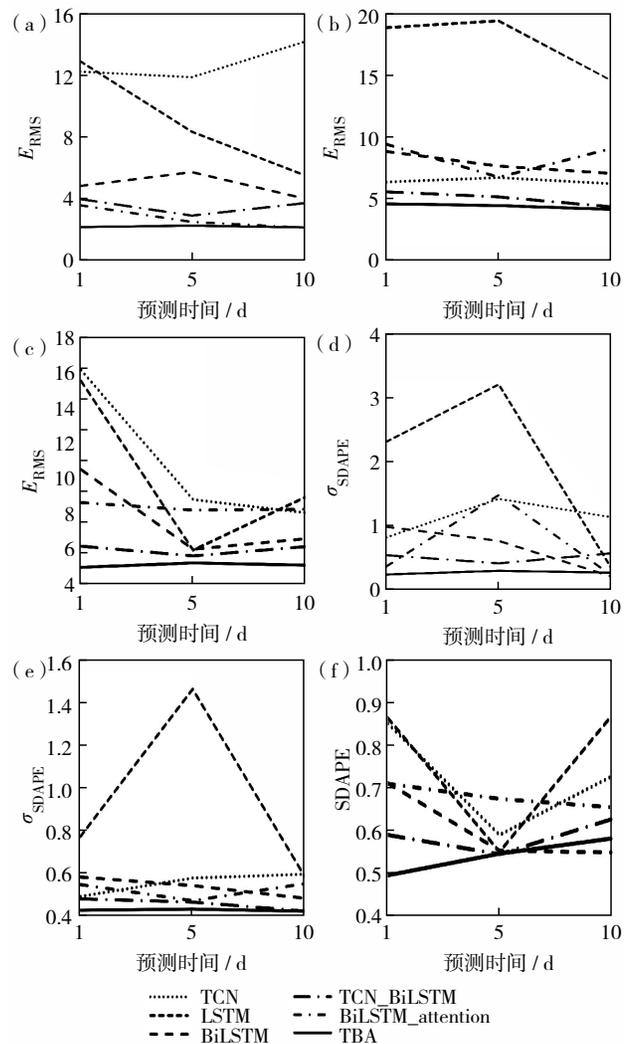


图6 不同消融模型对股指预测的结果 (a)对SSE预测的 $E_{RMS}$ ; (b)对CSI 300预测的 $E_{RMS}$ ; (c)对GEI预测的 $E_{RMS}$ ; (d)对SSE预测的 $\sigma_{SDAPE}$ ; (e)对CSI 300预测的 $\sigma_{SDAPE}$ ; (f)对GEI预测的 $\sigma_{SDAPE}$

Fig. 6 (a)  $E_{RMS}$  for SSE, (b)  $E_{RMS}$  for CSI 300, (c)  $E_{RMS}$  for GEI, (d)  $\sigma_{SDAPE}$  for SSE, (e)  $\sigma_{SDAPE}$  for CSI 300, and (f)  $\sigma_{SDAPE}$  for GEI by using different ablation models.

0.183%、0.104%与0.217%.

综上, 实验中6个模型的平均 $E_{RMS}$ 值从低到高依次为TBA、TCN\_BiLSTM、BiLSTM\_attention、BiLSTM、TCN和LSTM模型, 证明BiLSTM解决了单层LSTM无法捕获时间序列的某些附加特征的问题, 有效提高了模型的预测精度与稳定; 加入TCN后有效提高了BiLSTM的特征提炼能力; 加入注意力机制后有效提高了BiLSTM的特征识别能力. 最终, TBA融合模型的性能与前置模型的性能相比有较明显提升.

## 结 语

基于股指数据的特点构造一种基于TCN、BiLSTM与注意力机制的多特征融合的强性能预测TBA模型,并采用该模型分析开盘价、前收盘价、最高价、最低价与前收盘价5个指标对未来股指走势进行预测.融合TCN模型与注意力机制进一步提高了BiLSTM模型在金融时序数据上性能,同时为TBA模型设计了自适应机制,使该模型能在处理不同数据集时自行调整超参数,动态平衡TCN模型与注意力机制的比例.将TBA模型应用于中国3个主要的股指SSE、CSI300与GEI预测中,并将预测结果与其他5种主流的对照模型相比分析得出以下结论:TBA模型股指预测的误差值与主流算法对比有较明显的降低,具有较好的准确性;在预测不同数据集时表现稳定,鲁棒性较强.融合模型都比单独使用模型的性能得到有效提升,其自适应机制有效.因此使TBA模型可以运用在多种基于时间序列的金融预测场景,兼具准确性与鲁棒性.

基金项目:广东省自然科学重点领域专项基金资助项目(2020ZDZX3066);广东省哲学社会科学规划资助项目(GD21YYJ02, GD19CJLJ01)

作者简介:韩迪(dihan@gdudf.edu.cn),广东金融学院副教授、博士.研究方向:人工智能、推荐系统.

引 文:韩迪,郭维,廖凯,等.基于TBA融合模型的股票指数预测[J].深圳大学学报理工版,2023,40(6):665-673.

## 参考文献 / References:

- [ 1 ] 次必聪,张品一.基于ARIMA-LSTM模型的金融时间序列预测[J].统计与决策,2022,38(11):145-149.  
CI Bicong, ZHANG Pinyi. Financial time series forecasting based on ARIMA-LSTM models [J]. Statistics & Decision, 2022, 38(11): 145-149. (in Chinese)
- [ 2 ] WADI S A L, ALMASARWEH M, ALSARAIREH A A, et al. Predicting closed price time series data using ARIMA model [J]. Modern Applied Science, 2018, 12(11): 181-185.
- [ 3 ] 姚金海.基于ARIMA与信息粒化SVR组合的股指预测研究[J].运筹与管理,2022,31(5):214-220.  
YAO Jinhai. Study on stock index prediction based on ARIMA and information granular SVR combination [J]. Operations Research and Management Science, 2022, 31(5): 214-220. (in Chinese)
- [ 4 ] 沈慈慈,王伟杰,侯为波.基于GARCH类模型沪深300股指期货波动率预测研究[J].淮北师范大学学报自然科学版,2022,43(4):23-29.  
SHEN Cici, WANG Weijie, HOU Weibo. A study on volatility prediction of Shanghai and Shenzhen 300 stock index futures based on GARCH models [J]. Journal of Huabei Normal University Natural Science, 2022, 43(4): 23-29. (in Chinese)
- [ 5 ] CHEN Kai, ZHOU Yi, DAI Fangyan. A LSTM-based method for stock returns prediction: a case study of China stock market [C]//2015 IEEE International Conference on Big Data. Piscataway, USA: IEEE, 2015: 2823-2824.
- [ 6 ] SELVIN S, VINAYAKUMAR R, GOPALAKRISHNAN E A, et al. Stock price prediction using LSTM, RNN and CNN-sliding window model [C]// International Conference on Advances in Computing, Communications and Informatics. Piscataway, USA: IEEE, 2017: 1643-1647.
- [ 7 ] 杨青,王晨蔚.基于深度学习LSTM神经网络的全球股票指数预测研究[J].统计研究,2019,36(3):65-77.  
YANG Qing, WANG Chenwei. A study on forecast of global stock indices based on deep LSTM neural network [J]. Statistical Research, 2019, 36(3): 65-77. (in Chinese)
- [ 8 ] 孙存浩,胡兵,邹雨轩.指数趋势预测的BP-LSTM模型[J].四川大学学报自然科学版,2020,57(1):27-31.  
SUN Cunhao, HU Bing, ZOU Yuxuan. A BP-LSTM trend forecast model for stock index [J]. Journal of Sichuan University Natural Science Edition, 2020, 57(1): 27-31. (in Chinese)
- [ 9 ] SONG D, BUSOGI M, CHUNG BAEK A M, et al. Forecasting stock market index based on pattern-driven long short-term memory [J]. Economic Computation and Economic Cybernetics Studies and Research, 2020, 54(3): 25-41.
- [ 10 ] LI Zhenwei, HAN Jing, SONG Yuping. On the forecasting of high-frequency financial time series based on ARIMA model improved by deep learning [J]. Journal of Forecasting, 2020, 39(7): 1081-1097.
- [ 11 ] 李辉,化金金,邹波蓉.基于RF-LSTM组合模型的股票价格预测[J].河南理工大学学报自然科学版,2022,41(1):136-142.  
LI Hui, HUA Jinjin, ZOU Borong. Stock price prediction based on RF-LSTM combination model [J]. Journal of Henan Polytechnic University Natural Science, 2022, 41(1): 136-142. (in Chinese)
- [ 12 ] SIAMI-NAMINI S, TAVAKOLI N, NAMIN A S. A comparative analysis of forecasting financial time series using ARIMA, LSTM, and BiLSTM [EB/OL]. (2019-11-21) [2022-11-01]. <https://arxiv.org/abs/1911.09512>.
- [ 13 ] 夏林中,叶剑锋,罗德安,等.基于BERT-BiLSTM模型的短文本自动评分系统[J].深圳大学学报理工版,2022,39(3):349-354.  
XIA Linzhong, YE Jianfeng, LUO De'an, et al. Short text automatic scoring system based on BERT-BiLSTM model [J]. Journal of Shenzhen University Science and

- Engineering, 2022, 39(3): 349-354. (in Chinese)
- [14] 李洋, 董红斌. 基于CNN和BiLSTM网络特征融合的文本情感分析[J]. 计算机应用, 2018, 38(11): 3075-3080.  
LI Yang, DONG Hongbin. Text sentiment analysis based on feature fusion of convolution neural network and bidirectional long short-term memory network [J]. Journal of Computer Applications, 2018, 38(11): 3075-3080. (in Chinese)
- [15] 国显达, 那日萨, 崔少泽. 基于CNN-BiLSTM的消费者网络评论情感分析[J]. 系统工程理论与实践, 2020, 40(3): 653-663.  
GUO Xianda, NA Risa, CUI Shaoze. Consumer reviews sentiment analysis based on CNN-BiLSTM [J]. Systems Engineering - Theory & Practice, 2020, 40(3): 653-663. (in Chinese)
- [16] BAI Shaojie, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [EB/OL]. (2018-04-19) [2022-11-01]. <https://arxiv.org/abs/1803.01271>.
- [17] KARMAKAR P, TENG S W, LU Guojun. Thank you for attention: a survey on attention-based artificial neural networks for automatic speech recognition [EB/OL]. (2021-02-14) [2022-11-01]. <https://arxiv.org/abs/2102.07259>.
- [18] CIRINO P T, BARNES M A, ROBERTS G, et al. Visual attention and reading: a test of their relation across paradigms [J]. Journal of Experimental Child Psychology, 2022, 214: 105289.
- [19] 夏林中, 罗德安, 刘俊, 等. 基于注意力机制的双层LSTM自动作文评分系统[J]. 深圳大学学报理工版, 2020, 37(6): 559-566.  
XIA Linzhong, LUO De'an, LIU Jun, et al. Attention-based two-layer long short-term memory model for automatic essay scoring [J]. Journal of Shenzhen University Science and Engineering, 2020, 37(6): 559-566. (in Chinese)
- [20] HAN Di, CHEN Yijun, ZHANG Shuya. Implicit social recommendation algorithm based on multilayer fuzzy perception similarity [J]. International Journal of Machine Learning and Cybernetics, 2022, 13(2): 357-369.
- [21] NIU Zhaoyang, ZHONG Guoqiang, YU Hui. A review on the attention mechanism of deep learning [J]. Neurocomputing, 2021, 452: 48-62.
- [22] 任欢, 王旭光. 注意力机制综述[J]. 计算机应用, 2021, 41(增刊1): 1-6.  
REN Huan, WANG Xuguang. Review of attention mechanism [J]. Journal of Computer Applications, 2021, 41 (Suppl. 1): 1-6. (in Chinese)
- [23] HERMANS M, SCHRAUWEN B. Training and analysing deep recurrent neural networks [C]// Advances in Neural Information Processing Systems. San Francisco, USA: Curran Associates, Inc., 2013: 1: 9.
- [24] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.
- [25] 李炳臻, 刘克, 顾佼佼, 等. 卷积神经网络研究综述[J]. 计算机时代, 2021, 2021(4): 8-12, 17.  
LI Bingzhen, LIU Ke, GU Jiaojiao, et al. Review of the researches on convolutional neural networks [J]. Computer Era, 2021(4): 8-12, 17. (in Chinese)
- [26] 严春满, 王铖. 卷积神经网络模型发展及应用[J]. 计算机科学与探索, 2021, 15(1): 27-46.  
YAN Chunman, WANG Cheng. Development and application of convolutional neural network model [J]. Journal of Frontiers of Computer Science and Technology, 2021, 15 (1): 27-46. (in Chinese)
- [27] HAN Di, HUANG Yifan, JING Xiaotian, et al. AND: effective coupling of accuracy, novelty and diversity in the recommender system [C]// The 17th International Conference on Mobility, Sensing and Networking (MSN). Piscataway, USA: IEEE, 2021: 772-777.
- [28] GUPTA D, PRATAMA M, MA Zhenyuan, et al. Financial time series forecasting using twin support vector regression [J]. PloS One, 2019, 14(3): e0211402.
- [29] ROUT A K, DASH P K, DASH R, et al. Forecasting financial time series using a low complexity recurrent neural network and evolutionary learning approach [J]. Journal of King Saud University, 2017, 29(4): 536-552.
- [30] ECER F, ARDABILI S, BAND S S, et al. Training multi-layer perceptron with genetic algorithms and particle swarm optimization for modeling stock price index prediction [J]. Entropy, 2020, 22(11): 1239.
- [31] ZHOU Haoyi, ZHANG Shanghang, PENG Jieqi, et al. Informer: beyond efficient transformer for long sequence time-series forecasting [C]// Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, USA: AAAI Press, 2021: 11106-11115.
- [32] 黄卿, 谢合亮. 机器学习方法在股指期货预测中的应用研究——基于BP神经网络、SVM和XGBoost的比较分析[J]. 数学的实践与认识, 2018, 48(8): 297-307.  
HUANG Qing, XIE Heliang. Research on the application of machine learning in stock index futures forecast: comparison and analysis based on BP neural network, SVM and XGBoost [J]. Mathematics in Practice and Theory, 2018, 48(8): 297-307. (in Chinese).

【中文责编: 英子; 英文责编: 木柯】



补充材料