陈孝文, 苏攀, 王欣宇,等. 基于自注意力机制的时间卷积网络纸浆价格预测[J]. 智能计算机与应用,2026,16(5):113-118.
DOI:10.20169/j.issn.2095-2163.24072603

基于自注意力机制的时间卷积网络纸浆价格预测

陈孝文[1], 苏 攀[1], 王欣宇[2], 张新香[2]
(1 湖北中烟工业有限责任公司, 武汉 430040; 2 中南财经政法大学 信息工程学院, 武汉 430073)

摘 要: 为帮助烟草企业解决烟用材料采购过程中的成本和风险控制难题,本文提出一种基于自注意力机制的时间卷积网络(Temporal Convolutional Network,TCN)模型,用于科学预测烟用关键辅助材料原纸的上游原材料——纸浆期货的价格。该模块利用卷积模块捕捉长期依赖关系,同时利用自注意力机制学习重点时序段,实现纸浆价格的精准预测。实验结果表明:本文提出的 TCN-Attention 模型相比 LSTM、LSTM-Attention 和 TCN 模型,在纸浆价格预测中的精准度更高。该方法为期货价格预测提供了新思路,也为烟草企业合理制定内衬纸、接装纸的采购价格提供科学参考。

关键词: 自注意力机制; 时间卷积网络; 价格预测; 纸浆期货
中图分类号: TP183 **文献标志码**: A **文章编号**: 2095-2163(2026)05-0113-06

Pulp price prediction based on self-attention mechanism and temporal convolutional network

CHEN Xiaowen[1], SU Pan[1], WANG Xinyu[2], ZHANG Xinxiang[2]
(1 China Tobacco Hubei Industrial Co., Ltd., Wuhan 430040, China;
2 School of Information Engineering, Zhongnan University of Economics and Law, Wuhan 430073, China)

Abstract: To solve the problems of procurement, cost control, and risk management in tobacco enterprises, a Temporal Convolutional Network(TCN) model based on Self-Attention Mechanism was proposed to predict the price of pulp, the upstream raw material for cigarette paper, a key auxiliary material in tobacco manufacturing. The model employs a convolutional module to capture long-term dependencies and uses the self-Attention mechanism to learn the key temporal segments, thus completing the prediction of pulp prices. The experimental results show that the TCN-Attention model proposed in this paper has higher accuracy in pulp price prediction than the LSTM, LSTM-attention, and TCN models. This method provides a new idea for futures price prediction and a scientific reference for tobacco enterprises to reasonably formulate procurement prices for lining paper and tipping paper.

Key words: Self-Attention Mechanism; Temporal Convolutional Network; price prediction; pulp futures

0 引 言

卷烟配套原纸具有高透性、高抗张性等特点,是生产卷烟过滤嘴内外层的关键材料。作为卷烟内衬纸、接装纸的重要基础材料,其采购、生产、加工成本直接影响卷烟产品的生产成本与利润空间。纸浆作为卷烟配套原纸的上游主要原材料,其价格波动在很大程度上决定了卷烟配套原纸的最终生产成本。对纸浆价格进行监控预测,可辅助卷烟制造企业实现周期性采购、成本及风险控制,从而有效保障企业健康运营。

商品价格属于典型的时间序列数据,针对时间序列数据的预测问题,国内外学者已展开了大量的研究,内容涵盖农产品价格[1-3]、电力负荷[4,5]、碳排放权[6-7]、期货价格[8-11]等领域。在研究方法上,多采用神经网络模型进行时间序列预测,神经网络模型已成为该研究的主流趋势。根据所采用的神经网络模型架构,一般可分为递归架构神经网络和卷积架构神经网络。

在利用递归架构神经网络方面,陈立平等[3]利用长短期记忆网络(LSTM)对红枣期货价格进行预测,结果显示窗口期设为 20 的 LSTM 模型预测效果

作者简介: 陈孝文(1976—),男,硕士,经济师。主要研究方向:智能采购与供应链管理; 苏 攀(1977—),男,学士,工程师。主要研究方向:智能采购与供应链管理; 王欣宇(2001—),男,硕士研究生。主要研究方向:预测、算法设计与分析。
通信作者: 张新香(1975—),女,博士,教授,博士生导师。主要研究方向:大数据分析,商务智能,算法设计与分析。Email:zxxyx88@sina.com。

收稿日期: 2024-07-26

最佳。范俊明等[12]构建了多层 LSTM 网络对大豆期货价格进行预测,取得的预测效果优于自回归移动平均模型(ARIMA)、多层感知器模型(MLP)、支持向量机模型(SVR),同时发现,调整 LSTM 网络参数能显著影响预测效果。石榕[13]提出了一种引入注意力机制的门循环单元(GRU)模型对豆粕期货价格进行预测,结果显示引入注意力机制对模型预测效果有显著增强效应。虞文美等[14]利用 LSTM 对中国原油期货指数进行预测,发现时间窗口为 30 天的模型有着较高预测精度,预测结果能为金融机构进行风险识别和测度提供一定参考。李婧琦[15]利用鲸鱼算法优化 LSTM 对上证指数进行预测,发现该模型相较于 LSTM 、PSO-LSTM 及 SSA-BP 能实现更快、更准确的预测。

在卷积架构神经网络方面,路凯龙[16]利用卷积神经网络(CNN)预测沪铜期货价格,结果表明,CNN 在该任务上的表现逊于 GRU 模型。Bai 等[17]设计了通用的时序卷积网络架构,该网络在多个标准数据集上的表现优于 LSTM 等递归网络,为解决时序问题提供了新的出发点。袁文辉等[18]提出了基于特征选择策略的 TCN 模型并应用于电力负荷预测,该模型在仿真分析中取得了优于传统预测方法的精度。郑晓亮等[19]基于提出的 STE-TCN 模型对中短期电力负荷开展预测,结果表明该模型的预测精度较传统时序网络有所提升,且能显著提升预测时长。李望月等[20]提出基于改进 CEEMDAN-TCN 的模型对弗兰德伦地区陆上风电场数据集进行预测,发现该模型缓解了过拟合问题,能够更好地捕捉数据特征并能取得更优的预测精度。

梳理现有文献可知,鲜有研究在同一任务场景中同时应用递归架构神经网络和卷积架构神经网络,并系统对比两个网络架构优劣势,特别是在通用时序卷积架构神经网络方面的相关研究更少。本文将以纸浆期货价格预测为例,构建递归架构神经网络和通用时序卷积架构神经网络,同时引入自注意力机制,比较两个架构网络的预测性能,并验证引入自注意力机制在时序数据预测上的有效性和预测精度优势。

1 模型介绍

1.1 自注意力机制

注意力机制是深度学习中的常用的技术手段,引入注意力机制可以使模型自适应地调整对不同序列段的关注度,从而赋予不同权重,并将学习焦点放在重点特征信息上。文章引入自注意力机制以实现对长短期记忆网络(LSTM)和时间卷积网络(TCN)输出特征的权重分配。

自注意力机制的计算过程主要分为 3 个步骤:第一步自注意力机制将嵌入层的输入通过线性变换得到 $\boldsymbol{Q}$(Query)、$\boldsymbol{K}$(Key)、$\boldsymbol{V}$(Value)3 个矩阵;第二步计算矩阵 $\boldsymbol{Q}$ 与矩阵 $\boldsymbol{K}$ 之间的点积并同时缩放;最后,利用 Softmax 函数对权重进行归一化处理并与矩阵 $\boldsymbol{V}$ 相乘得到加权求和。整个过程的公式表示如下:

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \text{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d}}\right) \tag{1}$$

式中 d 为对应向量的维数。

1.2 时间卷积网络

时间卷积网络(TCN)是为了适应时间序列任务专门构建的卷积架构网络。TCN 的任务目标是:给定输入序列 $x_1,x_2,\cdots,x_t$,预测对应输出 $y_1,y_2,\cdots,y_t$,且约束条件为预测一段时间 t 内的输出 y_t,模型只能使用之前可观测到的输入 $x_1,x_2,\cdots,x_t$,目标函数为:

$$\hat{y}_1,\hat{y}_2,\cdots,\hat{y}_t = f(x_1,x_2,\cdots,x_t) \tag{2}$$

TCN 的主要结构由一维因果卷积模块、一维膨胀卷积模块和残差模块组成。

1.2.1 因果卷积模块

因果卷积模块在处理时间序列数据时规定了严格的时间约束,即某时刻 t 的信息只由 t 之前的信息确定,确保不会造成信息泄露,保证了因果关系。

1.2.2 膨胀卷积模块

经因果卷积模块处理数据后,还需要引入膨胀卷积模块以构造深度网络,从而实现长时记忆。经过膨胀卷积模块,神经元的感受也随网络层深度逐步扩大,因果膨胀卷积的结构示意图如图 1 所示,膨胀卷积运算公式为:

$$F(s) = (x *_d f)(s) = \sum_{i=0}^{k-1} f(i)\cdot x_{s-d\cdot i} \tag{3}$$

式中:$F(s)$ 为 s 的卷积,x 为一维时间序列元素,f 为滤波器,d 为膨胀因子,k 为滤波器大小,$s-d\cdot i$ 为时间维度方向。



图 1 因果膨胀卷积结构示意图

Fig. 1 Schematic diagram of dilated causal convolution

1.2.3　残差模块

残差模块由多个子残差模块组成，其主要解决跨层传递信息问题及梯度消失问题，结构如图 2 所示，残差模块的表达式为：

$$o = f(\Gamma(x) + x) \tag{4}$$

式中：$f(\cdot)$ 为激活函数，Γ 为残差变换。

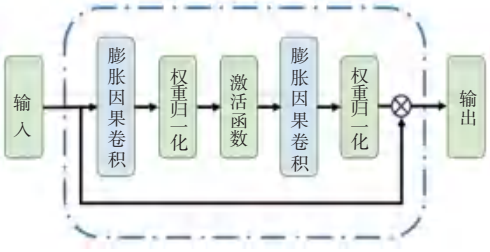


图 2　残差模块结构示意图

Fig. 2　Schematic diagram of residual module

2　价格预测设计

2.1　数据来源

纸浆期货的上市时间为 2018 年 11 月 27 日，所以本文实验使用自 2018 年 11 月 30 日至 2024 年 5 月 31 日的纸浆期货数据，数据源自上海期货交易所官网。

期货数据一般有多个合约，主力合约代表某一期货品种中交易量最大、流动性最好的合约，其价格走势往往决定整个品种的趋势。因此，预测期货数据时，使用主力合约数据进行预测，预测结果有着较为广泛的代表性，能最为精准地反应纸浆期货价格的未来走势。鉴于此，本文筛选整理原始数据得到纸浆主力合约期货数据，构建本文的数据源。

2.2　多步迭代策略

多步迭代策略是时间序列预测领域中较为常用的预测策略，迭代示意图如图 3 所示。



图 3　多步迭代示意图

Fig. 3　Schematic diagram of multi-step iteration

多步迭代策略的关键在于利用预测窗口预测未来的数据。以预测未来一天的数据为例，假设选定的时间窗口为 i，模型将利用最后 i 天的数据预测第 $i+1$ 天的数据，即

$$T_x = f(T_{x-1}, T_{x-2}, \cdots, T_{x-i}) \tag{5}$$

完成这样的一次预测操作被视为完成一次迭代。

第二次迭代时，模型将经预测得到第 $i+1$ 天的数据视为已知数据加入到序列的尾部得到新序列，同时时间窗口大小不变，时间窗口沿着序列的尾部方向滑动 t 个步长，在这个例子中 t 为 1，模型则将利用新序列最后 i 天的数据预测第 $i+1$ 天的数据，即

$$T_{x+1} = f(T_x, T_{x-1}, \cdots, T_{x-i+1}) \tag{6}$$

之后的迭代操作与第二次迭代同理。

2.3　实验方案

选取 2018 年 11 月 30 日至 2024 年 5 月 31 日的纸浆主力期货数据开展对比分析实验，目的在于精准预测纸浆期货价格的未来变动情况。利用本文构建的基于自注意力机制的 TCN 模型，与长短期记忆网络 LSTM、基于自注意力机制的 LSTM 以及 TCN 模型进行比较，以验证本文所提出的基于自注意力机制的 TCN 模型在纸浆期货价格预测任务上的有效性和准确性。纸浆价格预测实验流程如图 4 所示。

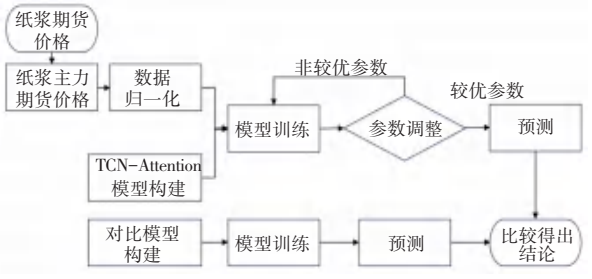


图 4　纸浆价格预测实验流程图

Fig. 4　Experiment flowchart for pulp price prediction

2.4　模型构建

2.4.1　数据归一化

为消除数据量纲与数值范围差异对模型训练的影响，采用线性归一化（Min-Max Scaling）方法对期货数据进行归一化处理，公式如下：

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{7}$$

其中，x 为原始数据；$x_{\max}$ 与 $x_{\min}$ 为数据集中的最大值与最小值；x_{norm} 为经过线性归一化后的数据。

2.4.2　模型训练和预测

在模型训练前将归一化后的数据按照 8∶2 划分为训练集与测试集，其中训练集包含 1 066 条数据，测试集包含 267 条数据。在模型训练时，将训练集输入至模型中并采取网格搜索法选取较优参数。在模型预测时，将测试集输入至训练好的模型中，经模型计算得到预测结果。

2.4.3　基于自注意力机制的时间卷积网络

由于期货价格易受到国际局势、行业政策、投资

者情绪等非产品因素的影响而产生非规律性的变化,同时在时间序列上的不同阶段,特征重要程度也有所不同。非周期性变化和特征重要程度的不同会影响模型的学习过程,进而影响模型的预测性能,因此本文引入注意力机制以降低这些因素的干扰。

本文构建的基于自注意力机制的时间卷积网络模型分为三部分。第一部分为 TCN 模块,由因果膨胀卷积模块、残差连接模块和跳跃连接模块组成;第二部分为自注意力机制模块;第三部分为全连接网络层。模型的整体架构如图 5 所示。

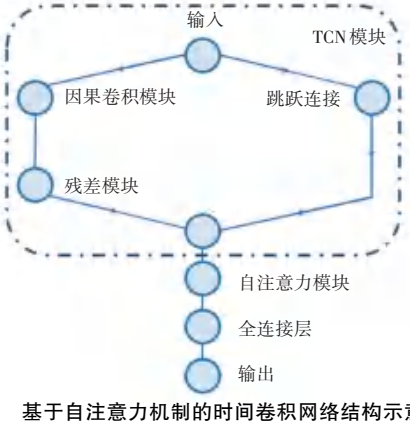


图 5 基于自注意力机制的时间卷积网络结构示意图

Fig. 5 Schematic diagram of Temporal Convolutional Network structure based on Self-Attention mechanism

2.4.4 评价指标

为了科学有效地衡量各模型的预测性能,本文选取均方根误差(RMSE)、绝对误差(MAE)以及平均绝对百分比误差(MAPE)作为评价指标,测量模型的预测精度。各指标的计算公式如下:

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{x}_i)^2} \tag{8}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|x_i-\hat{x}_i| \tag{9}$$

$$\mathrm{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{x}_i-x_i}{x_i}\right|\times 100\% \tag{10}$$

式中:x_i 代表纸浆期货的真实价格,$\hat{x}_i$ 代表纸浆期货的预测价格。

3 结果与讨论

3.1 实验结果

文本所构建的 TCN-attention、LSTM、LSTM-attention 和 TCN 模型在纸浆主力期货数据测试集上的预测效果如图 6—9 所示。使用 RMSE、MAE、MAPE 3 个指标能更清晰发现各模型的优劣,各模型的预测结果见表 1。

表 1 模型间对比实验结果

Table 1 Comparative experimental results between models

模型名称	RMSE	MAE	MAPE/%
LSTM	0.069 2	0.054 9	1.421 0
LSTM-attention	0.049 5	0.037 4	0.972 6
TCN	0.023 5	0.017 6	0.956 9
TCN-attention	0.022 8	0.016 7	0.911 9

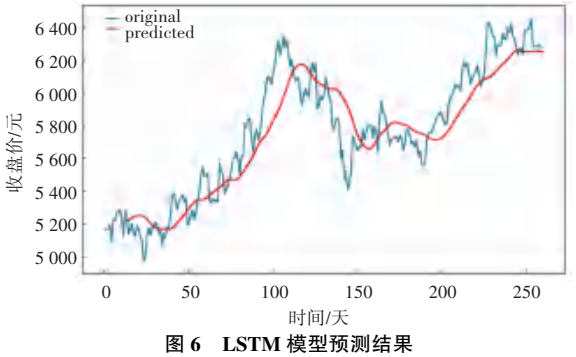


图 6 LSTM 模型预测结果

Fig. 6 Prediction results of LSTM model

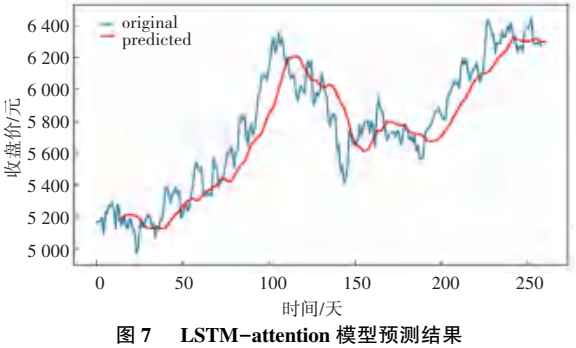


图 7 LSTM-attention 模型预测结果

Fig. 7 Prediction results of LSTM-attention model

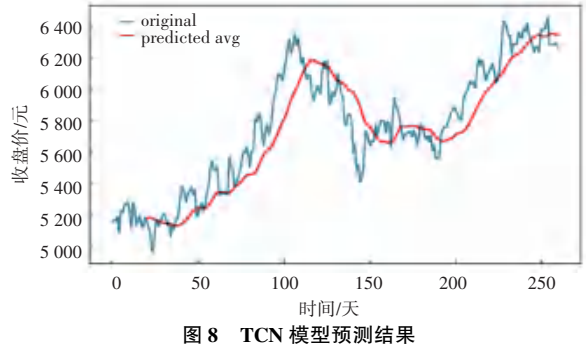


图 8 TCN 模型预测结果

Fig. 8 Prediction results of TCN model

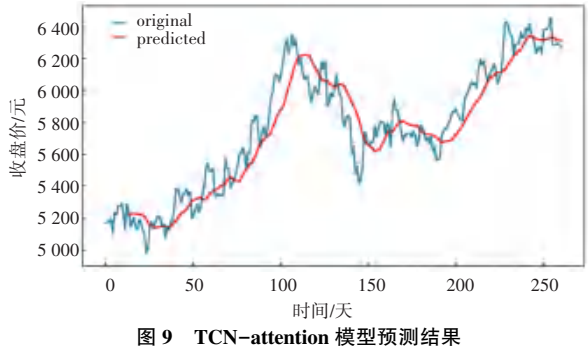


图 9 TCN-attention 模型预测结果

Fig. 9 Prediction results of TCN-attention model

观察表 1 中的数据可知，本文提出的 TCN-attention 模型在纸浆价格预测任务上表现出明显的优势。对比 LSTM、LSTM-attention 和 TCN 模型，TCN-attention 模型的预测结果在 RMSE 上的提升分别为 0.046 4、0.026 7 和 0.000 7，在 MAE 上的提升分别为 0.038 2、0.020 7 和 0.000 9，在 MAPE 上的提升分别为 0.509 1%、0.060 7% 和 0.045 0%。观察图 6—9（图中横坐标代表时间，单位为天；纵坐标代表收盘价格，单位为元），可以发现 TCN-attention 模型的预测结果更接近真实值。上述实验结果验证了 TCN-attention 模型的有效性。

为模型设定不同长度的时间窗口会对模型预测准确性产生一定影响[21-22]，调整时间窗口的长度，可以对期货价格数据中包含的信息进行过滤，使模型能够捕捉更为重要的价格变动因素。不同长度的时间窗口对模型的性能影响不同，时间窗口长度较小则能模型更好地反映市场倾向，捕获价格突变因素；时间窗口长度较长则可以更好地把控价格变化总体趋势。为了保证设定的时间窗口能使模型在一定程度上兼顾长短期的信息捕捉能力，在模型构建过程中，需要不断寻找合适长度的时间窗口来确保模型得到较优的预测结果。因此，进行关于时间窗口长度的研究具有重大意义。

表 2 为 TCN-attention 模型在不同窗口期下的纸浆主力期货数据测试集上的预测结果：

表 2　不同窗口期对比实验结果

Table 2　Comparison of experiment results for different window periods

窗口期(天)	RMSE	MAE	MAPE/%
3	0.023 1	0.017 1	0.928 4
5	0.022 7	0.016 8	0.916 0
7	0.022 8	0.016 7	0.911 9
15	0.024 3	0.018 1	0.976 7
30	0.025 9	0.019 6	1.068 4

结果显示模型窗口期为 7 时在 MAE、MAPE 上取得了最优的结果，模型窗口期为 5 时在 RMSE 上取得了最优的结果，而模型窗口期为 3、15、30 时，预测效果较差。因此，窗口期为 5 和 7 时能使得 TCN-attention 模型能够更好地兼顾长短期的信息捕捉能力并得到更好的预测效果。

3.2　讨论

本文构建了一种基于自注意力机制的时间卷积网络预测模型，并以 2018 年 11 月 30 日至 2024 年 5 月 31 日的纸浆期货数据为样本，测试了模型的预测精度及有效性，主要研究结论如下：

（1）比较 LSTM 与 LSTM-attention 模型、TCN 与 TCN-attention 模型，引入自注意力机制的模型预测效果更为准确，证明了在模型中引入自注意力机制的有效性。

（2）相较于 LSTM、LSTM-attention 与 TCN 模型，本文构建的 TCN-attention 模型在对纸浆期货价格的预测中取得了最好的效果。RMSE、MAE 及 MAPE 指标分别为 0.022 8、0.016 7、0.911 9%，同时取得最优值。

（3）相较于递归架构的网络，针对时序问题改进的通用时序卷积架构网络（TCN），能够更为有效地挖掘时序信息，同时有效建立长期依赖和解决梯度消失问题，在纸浆期货价格预测任务上表现出了更优的准确率。

（4）在 TCN-attention 模型上进行时间窗口长度的调整，发现调整窗口期长度对模型性能有较大影响，且窗口期为 5 和 7 时模型能够更好地兼顾长短期的信息捕捉能力并得到更好的预测效果。

4　结束语

为了帮助烟草企业更合理制订采购价格、科学安排生产、精准规避风险，本文提出了一种基于自注意力机制的时间卷积网络预测模型，对卷烟配套原纸的上游主要原材料纸浆价格进行监控预测。实验结果验证了本文提出的 TCN-attention 模型的预测精准性以及引入自注意力机制的有效性，同时比较了递归架构网络和通用时序卷积架构网络在该任务上的性能差异。

卷烟生产工艺繁杂，成本构成复杂、为了进一步助力烟草企业健康发展，需要对更多烟用原材料进行预测。同时影响烟用材料价格的因素众多，能源政策、环保政策、经济政策等都会极大程度影响烟用材料价格未来走势，因此考虑基于政策文本进行更高精度的预测，将是未来研究的重点。

参考文献

[1] 刘彦虹，刘合兵，尚俊平. 基于深度学习的农产品期货价格预测研究[J]. 河南科学，2024，42(3)：430-439.

[2] 卢超凡，史世凯，王鲁. 基于改进 LSTM 的苹果价格预测模型研究[J]. 山东农业大学学报(自然科学版)，2022，53(3)：491-496.

[3] 陈立平，邢小丹，张玉亭，等. 基于 LSTM 的红枣期货价格预测方法[J]. 农业与技术，2024，44(1)：162-165.

[4] 庞伟林，关兆雄，李沐栩. 一种混合 LSTM-SAM 的居民电力负荷预测模型[J]. 人工智能科学与工程，2024，46(2)：40-49.

[5] 梅锦超, 张鹏宇, 程斌, 等. 基于改进二次模态分解和 BiLSTM-Attention 的短期电力负荷预测[J]. 电工材料, 2024, 47(1): 100-104.
[6] 毛涵申. 基于 CS-BP 神经网络模型的碳排放权价格预测[J]. 信息技术与信息化, 2023, 48(9): 52-55.
[7] 沈蕾, 罗梦丝. 基于 LSTM 算法的碳排放权交易价格多因素预测研究[J]. 价格理论与实践, 2022, 45(7): 64-68.
[8] 邱冬阳, 丁玲. 基于多维高频数据和 LSTM 模型的沪深 300 股指期货价格预测[J]. 重庆理工大学学报(社会科学), 2022, 36(3): 55-69.
[9] 斯燕, 陈艺. 基于 LSTM 神经网络模型的铁矿石期货市场实证研究[J]. 中国集体经济, 2024, 562(2): 100-103.
[10]王毓谦. 深度学习框架下 LSTM 网络的期货价格预测模型研究[J]. 中小企业管理与科技(下旬刊), 2021, 35(12): 137-139.
[11] ZHANG T, TANG Z. Agricultural commodity futures prices prediction based on a new hybrid forecasting model combining quadratic decomposition technology and LSTM model [J]. Frontiers in Sustainable Food Systems, 2024, 8(1):1-17.
[12]范俊明, 刘洪久, 胡彦蓉. 基于 LSTM 深度学习的大豆期货价格预测[J]. 价格月刊, 2021, 452(2): 7-15.
[13]石榕. 基于注意力机制的 GRU 模型的豆粕期货价格预测[J]. 洛阳理工学院学报(自然科学版), 2023, 33(4): 87-91.
[14]虞文美, 方扶星. 基于 LSTM 神经网络的中国原油期货指数预测研究[J]. 哈尔滨师范大学自然科学学报, 2020, 36(2): 32-37.
[15]李婧琦. 基于鲸鱼算法优化 LSTM 的股票价格预测模型[J]. 智能计算机与应用, 2023, 13(2): 35-40.
[16]路凯龙. 基于深度学习方法的沪铜期货价格预测研究[J]. 全国流通经济, 2019, 28(33): 152-153.
[17]BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [J]. arXiv preprint arXiv,1803. 01271, 2018.
[18]袁文辉, 张仰飞. 基于特征选择策略和 TCN 的电力负荷预测方法[J]. 信息技术, 2024, 57(4): 9-14.
[19]郑晓亮, 束庆宇. 基于 STE-TCN 的中短期电力负荷预测[J]. 重庆工商大学学报(自然科学版),2024, 41(6):59-64.
[20]李望月, 樊重俊. 基于改进 CEEMDAN-TCN 模型的风电功率预测研究[J]. 智能计算机与应用,2024, 14(1): 112-118.
[21]王兴芬, 张彦博. 不同时间窗下铁矿石期货价格预测精度研究——基于 CNN-LSTM 注意力模型的分析[J]. 价格理论与实践, 2022, 457(11): 142-145.
[22]张杰, 甄柳琳, 翟东升. 基于多级注意力机制的大豆期货预测[J]. 系统工程, 2023, 41(2): 148-158.