

基于注意力机制的 CEEMDAN-CNN-LSTM 股指价格预测

詹志滨, 张振华*

广东外语外贸大学 数学与统计学院, 广东 广州 510006

DOI:10.61369/ASDS.2025100010

摘 要 : 针对股价的高频不确定性、长期记忆性, 本研究提出了基于注意力机制特征赋权和自适应分解技术的 CEEMDAN-CNN-LSTM 神经网络模型。首先, 使用自适应分解技术 (CEEMDAN) 进行模态分解, 将原始时间序列分解为高低频段和趋势项。其次, 利用卷积神经网络 (CNN) 和长短期记忆神经网络 (LSTM), 独立对高低频信息进行信号提取和特征学习, 并通过注意力模块 (AM) 进行信号特征的权重自适应分配。最后, 集成各频段预测值作为最终结果。本研究选取沪深 300 和中证 500 股票指数进行实证分析, 结果显示, 基于注意力机制的 CEEMDAN-CNN-LSTM 模型, 既能避免突变信号的遗漏, 专注于序列依赖关系, 同时又能有效提取共性特征, 避免过拟合与预测结果偏右等问题。与其他主流预测模型相比, 其预测精度更高, 预测误差显著降低, 泛化能力更强。

关 键 词 : 注意力机制; CEEMDAN; CNN; LSTM; 股票指数

Attention-based CEEMDAN-CNN-LSTM Model for Stock-Index Price Forecasting

Zhan Zhibin, Zhang Zhenhua*

School of Mathematics and Statistics, Guangdong University of Foreign Studies, Guangzhou, Guangdong 510006

Abstract : To capture the high-frequency uncertainty and long-memory characteristics of stock-index movements, we propose an attention-based CEEMDAN-CNN-LSTM neural network that synergistically combines adaptive decomposition with feature-weight learning. First, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) is employed to disentangle the raw price series into high-frequency, low-frequency and trend components. Second, a convolutional neural network (CNN) and a long short-term memory network (LSTM) are trained separately on each frequency band to extract local patterns and long-range dependencies, while an attention module (AM) autonomously re-weights the extracted features according to their instantaneous relevance. Finally, the band-specific forecasts are aggregated to produce the ultimate prediction. Empirical experiments on the CSI300 and CSI500 indices demonstrate that the proposed model retains abrupt signals, preserves sequential dependence, and concurrently learns common factors without suffering from right-skewed or over-fitted forecasts. Compared with five mainstream benchmarks, the proposed approach significantly reduces prediction error and exhibits superior generalisation capability.

Keywords : Attention Mechanism; CEEMDAN; CNN; LSTM; stock index

引言

全球经济风云变幻, 我国经济正处于新旧动能转换的关键节点。习近平总书记强调金融安全关乎国家安全, 股票市场作为金融体系核心枢纽, 价格波动影响深远。它不仅关乎系统性金融风险防范, 对投资者而言, 精准预测股价更是实现资产增值、规避风险的核心要务。而机器学习技术凭借海量数据处理与复杂模式识别能力, 能深度挖掘市场多维度信息, 精准捕捉股价波动规律, 与股价预测需求高度契合, 为相关研究者开辟了科学决策新路径。

基金项目: 本文获广东省哲学社科规划项目 (No.GD24CTS03); 语言工程与计算广东省重点实验室招标项目 (No.LEC2020ZBKT002) 支持。

作者简介: 詹志滨, 广东外语外贸大学数学与统计学院, 硕士研究生, 研究方向: 时间序列分析、数据挖掘。

通讯作者: 张振华, 广东外语外贸大学数学与统计学院, 博士, 教授, 硕士生导师, 研究方向: 统计机器学习。

在特征学习上,学者们已经深刻意识到长短期记忆神经网络(LSTM)能够克服股价预测长期依赖性,但对如何充分挖掘有效特征这一核心问题,学者们展开了不同的研究。肖田田为挖掘股票间价格的潜在影响作用,通过K-means聚类对同类股票的价格趋势进行共性特征挖掘,再输入到目标股票价格的LSTM预测神经网络中,揭示了同类股票股价间互相预测的可行性^[1]。马致远为解决传统量化模型的多维度因子筛选问题,分别提出了基于主成分分析和等距映射算法改进的LSTM模型,提高了LSTM的预测性能^[2]。沈山山和李秋敏就时间特征提取角度,利用卷积神经网络(CNN)提取时间维度特征,并用注意力机制(AM)自动赋权,提高LSTM的预测效果^[3],而刘芸从LSTM结构上进行调整,搭建基于AM的双向长短期记忆网络(BiLSTM)模型来预测股价崩盘风险^[4]。

不过,随着对金融市场宏观结构和机器学习的深入研究,学者们开始将信号处理和深度学习等跨学科技术结合起来,构建混合或集成模型。比如刘铭和单玉莹基于经验模态分解(EMD)和LSTM结合的模型,对上证指数、深证指数和沪深300进行了预测,给出了不同波动性下的不同股指预测模型^[5];再如黄后菊和李波将信号去噪领域的变分模态分解(VMD)与麻雀搜索优化算法(CSSA)相结合,改良优化了LSTM股票价格预测模型^[6];又如贺毅岳等人将互补集成经验模态分解(CEEMD)与LSTM结合,重新提出一种优化预测模型^[7]。

为了显著提升预测精度并增强模型的稳健性,本研究针对股指时间序列数据的高复杂性,在CEEMD、CNN、LSTM的基础上^[8]改进分解技术并加入注意力机制模块,构造一种融合自适应分解和卷积神经网络-长短期记忆网络-注意力机制模块的混合模型(CEEMDAN-CNN-LSTM-AM)。模型构建的流程如下:首先,利用CEEMDAN技术对原始股指收盘价进行深入的模式分解,解构出k个反映局部特性的固有模式函数(IMF)分量,有效降低数据的内在复杂性。其次,针对每个IMF分量,独立构建并运用CNN提取时空节点特征,将特征传入LSTM进行预测,并通过注意力模块进行信号特征的权重自适应分配。最后,经由全连接层的处理,输出各IMF的预测结果,集成加总出最终股指收盘价的预测值。通过沪深300和中证500指数的实证分析,结果表明该混合模型优于其它预测模型,在预测精度、适用性以及模型泛化能力方面均表现出明显的优势,展现出强大的特征分解和提取能力。

一、研究方法

(一) 自适应分解 CEEMDAN

CEEMDAN是一种自适应信号分解算法,是对传统经验模态分解(EMD)和集合经验模态分解(EEMD)的改进,特别适用于非线性、非平稳信号的分析^[9]。该算法在每次分解阶段加入自适应噪声,并通过逐层平均IMF分量,消除重构造误差,提高分解一致性和稳定性。其算法步骤如下:

1. 对原始信号添加高斯白噪声,构造含噪信号:

$$x_i(t) = x(t) + c_i s_i(t) \quad (1)$$

其中, c_i 为高斯白噪声权值系数; $s_i(t)$ 为第 i 次处理时产生的高斯白噪声。

2. 对每个含噪信号进行EMD分解,提取第一层IMF并计算其均值:

$$IMF_1(t) = \frac{1}{K} \sum_{i=1}^K IMF_1^i(t) \quad (2)$$

其中, $IMF_1(t)$ 表示CEEMDAN分解得到的第1个模式分量; $r_1(t)$ 表示第1次分解后的余量信号:

$$r_1(t) = x(t) - IMF_1(t) \quad (3)$$

3. 对余量信号的重复上述分解过程,逐层提取IMF,直到残差为单调函数或满足停止条件。公式表示为:

$$r_j(t) = \frac{1}{K-1} \sum_{i=1}^{K-1} (r_{j-1}(t) - IMF_j^i(t) + c_{j-1} \delta_{j-1}(t)) \quad (4)$$

其中, $IMF_j(t)$ 表示CEEMDAN分解得到的第 j 模式分量;

$c_{j-1}(t)$ 表示对序列进行EMD分解后的第 $j-1$ 个IMF分量; c_{j-1} 表示CEEMDAN对第 $j-1$ 阶段余量信号加入噪声的权值系数; $r_j(t)$ 表示第 j 阶段余量信号。

4. 输出模式函数,得到若干IMF分量和一个最终残余项,完成信号分解。总原始信号的分解过程可表示为:

$$x(t) = \sum_{i=1}^n IMF_i(t) + r_n(t) \quad (5)$$

(二) 卷积神经网络 CNN

卷积神经网络(Convolutional Neural Network, CNN)通过离散卷积与层级非线性变换,能够高效捕获数据的局部相关性与平移不变特征,已广泛应用于计算机视觉、自然语言处理及金融时间序列建模^[10]等领域。其典型拓扑由交替堆叠的卷积层与池化层构成,并在卷积后引入激活函数以赋予网络非线性表达能力。

1. 卷积层

将卷积核视为在时序或图像平面上滑动的局部感知器,对输入片段执行加权求和与偏置后,经激活函数输出特征图。为缓解梯度消失并降低计算开销,通常采用ReLU(Rectified Linear Unit)作为非线性映射:

$$y_j^l = \max \left\{ 0, \sum_{i \in M_j} x_i^{l-1} \otimes W_{ij}^l + b_j^l \right\} \quad (6)$$

其中, x_i^{l-1} 为第 $(l-1)$ 层的第 i 个特征图, W_{ij}^l 与 b_j^l 分别表示第 l 层的第 j 个卷积核的权重与偏置, \otimes 代表离散卷积运算, y_j^l 为经过ReLU后的输出特征图。

2. 池化层

在保留显著特征的同时降低特征图分辨率,进而缩减参数量

并提升模型泛化性能。本文采用最大池化：

$$p^l = \text{Maxpool}(y^l) \quad (7)$$

其中， p^l 为池化后的输出，Maxpool 以不重叠窗口取最大值，实现局部特征的二次抽象与降维。

(三) 长短期记忆网络 LSTM

长短期记忆网络 (Long Short-Term Memory, LSTM) 是一类专门为捕获长距离时序依赖而设计的循环神经网络 (RNN) 变体。针对传统 RNN 在反向传播过程中因梯度指数级收缩或膨胀而导致的梯度消失 / 爆炸难题，LSTM 通过引入记忆单元与门控机制 (输入门、遗忘门、输出门) 对信息流进行显式且可学习的控制，从而稳定地建模长期动态。

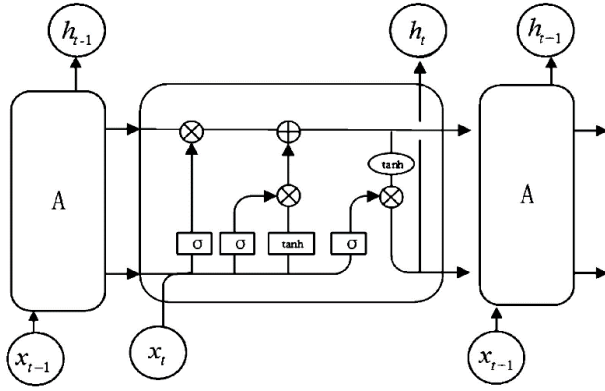


图1: LSTM 网络结构展开图

图1为 LSTM 的网络结构展开图，其中 x_t 表示当前 t 时刻的输入， h_t 表示当前 t 时刻细胞的状态值，其时间步 t 的前向计算可概括为：

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \quad (8)$$

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f) \quad (9)$$

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \quad (10)$$

$$C_t = \tanh(W_c * [h_{t-1}, x_t] + b_c) \quad (11)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot C_t' \quad (12)$$

$$h_t = o_t \odot \tanh(C_t) \quad (13)$$

其中， i_t 、 f_t 、 o_t 分别表示输入门、遗忘门与输出门的激活值； C_t' 为候选记忆， C_t 为更新后的细胞状态； $W*$ 与 $b*$ 为对应门的可学习参数； \odot 代表逐元素乘积， $\sigma(\cdot)$ 与 $\tanh(\cdot)$ 分别为 Sigmoid 与双曲正切激活函数。该结构通过门控向量对细胞状态进行选择保留、更新与输出，在保持端到端可微的同时，显著增强了网络对长程依赖的建模能力。

(四) 注意力机制 AM

注意力机制 (Attention Mechanism, AM) 通过可学习权重模拟人脑的选择性聚焦过程，对输入序列中的关键信息予以强化。给定输入序列 $H=[h_1, \dots, h_T]$ 生成查询 Q 与共享键值 h (键值共享简化)，随后计算相关性得分 S ，并经 Softmax 归一化得到

注意力权重 A ，最终输出上下文向量 $C = A \cdot h$ ，实现动态加权聚合。其计算流程如下：

$$S = v \cdot \tanh(W_q Q + W_k h), v \in R^{d_h} \quad (14)$$

$$A = \alpha_{t,d} = \text{softmax}(S) = \frac{\exp(e_{t,d})}{\sum_i \exp(e_i)} \quad (15)$$

$$C = Ah = \sum_t \sum_d \alpha_{t,d} \cdot h_{t,d} \quad (16)$$

其中 S 表示时刻 t 查询 Q 与共享键值 h 的相关性得分； W_q 、 W_k 为可训练权重矩阵； v 为可学习向量 (维度与隐藏层维度 d_h 一致)，通过向量 - 内积将矩阵压缩为标量得分； $\alpha_{t,d}$ 为归一化后的注意力权重； C 为融合全局信息的上下文向量。

在 CEEMDAN-CNN-LSTM 框架中，注意力模块置于 LSTM 输出端，对通道 - 时间二维特征图 $H \in R^{T \times D}$ 的每一位置 (t, d) 计算 softmax 权重 $\alpha_{t,d}$ ，并执行逐点加权，所得上下文向量 C 经展平后输入回归层。该机制沿通道 - 时间二维平面自适应分配权重，动态强化关键通道及关键时间窗的贡献，从而显著提升模型预测精度。

(五) CEEMDAN-CNN-LSTM-AM 组合框架

图2给出了面向股价序列的建模流程，具体步骤如下：

1. 自适应分解：利用 CEEMDAN 将原始序列分解为若干本征模态函数 (IMFs)，完整保留从高频噪声到低频趋势的多尺度信息。
2. 标准化：对各 IMF 序列进行归一化处理，以消除量纲差异。
3. 分量建模：对每个 IMF 独立构建 CNN-LSTM-AM 子模型，依次通过卷积层提取局部特征、LSTM 捕获长期依赖、注意力机制强化关键时序贡献，并经由全连接层输出单步预测结果。
4. 加性重构：将各 IMF 的预测值按时间步累加，得到最终股价预测序列，并通过逆标准化恢复原始尺度。
5. 性能评估：采用多误差指标与基线模型对比，验证框架有效性。

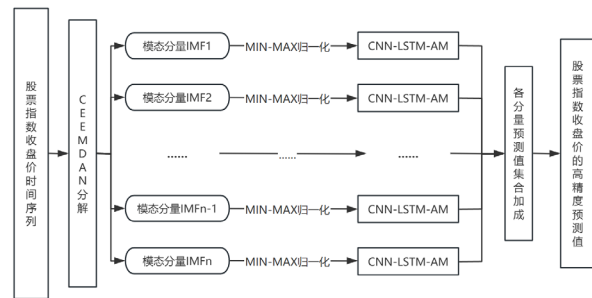


图2: 股指分解预测建模流程图

二、实证分析

(一) 数据来源

本研究使用的实验数据来源于锐思数据库 (<https://www.resnet.com/db>)，具体包括沪深300指数和中证500指数自2005

年1月4日至2024年7月1日的收盘价数据，共计4734个交易日。取前80%作为训练集，后20%作为测试集。

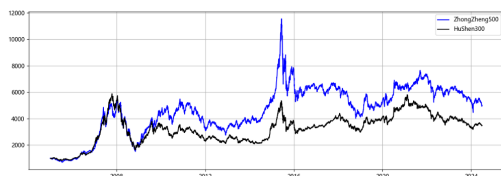


图3: 沪深300股指与中证500的收盘价走势图

(二) 数据预处理

本研究采用滑动窗口的方式，将选取过去30个交易日的沪深300指数收盘价来预测下一个交易日的收盘价，分别创建出3787个训练样本和947个测试样本。训练样本用于训练各模型的结构与参数，而测试样本则用于评估模型的预测性能。在每次训练模型前，为了提高模型的拟合精确度，需要先对输入的序列数据进行Min-Max归一化处理，消除量纲的影响。

(三) 参数设置

本研究所采用的CEEMDAN-CNN-LSTM-AM模型参数及其对比模型参数均采用网格搜索进行超参数寻优，最终自研模型的主要参数设置如下：批处理样本规模为32，迭代次数设置为50，生成器选择Adam优化器，学习率为0.001，损失函数为均方误差MSE，其余各模型的网络层超参数如表1所示。

表1: 各模型超参数设置

模型	超参数	取值
CNN-LSTM-AM	卷积层	1
	卷积核数	32
	卷积核大小	3
	其它隐藏层	2
	隐藏层神经元数	64
CNN-LSTM	卷积层	1
	卷积核数	32
	卷积核大小	3
	其它隐藏层	2
	隐藏层神经元数	128
LSTM	隐藏层	2
	隐藏层神经元数	128

其中，为充分验证CEEMDAN-CNN-LSTM-AM模型的预测效果，选取MLP、SVR、LSTM、CNN-LSTM、CNN-LSTM-AM等五种主流模型与CEEMDAN-CNN-LSTM-AM模型进行消融实验和对比实验。

(四) 评价指标

为系统评估模型的预测精度，本文采用平均绝对误差(MAE)与均方根误差(RMSE)、绝对百分比误差(MAPE)三项指标。设 y_i 为第 i 个样本的实际观测值， \hat{y}_i 为对应预测值，则各指标定义如下公式17-19所示，上述指标取值越小，表明模型预测值与真实值之间的偏离程度越低，预测性能越优。

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (17)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (18)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (19)$$

除上述指标外，进一步引入确定系数(R^2)来衡量模型的整体拟合优度。 R^2 值越接近1表示模型拟合得越好。

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (20)$$

(五) 沪深300股指序列的CEEMDAN分解

从图4可以看出，沪深300指数被分解为9个固有模态函数(IMF)和1个残余项，从IMF1到IMF9的频率逐渐下降。IMF1至IMF4包含了大量高频信号，可能代表突发事件或者市场情绪对股价序列的影响，而IMF5至IMF9的噪声较少，更多地反映了序列本身的惯性趋势特征。由此来看，采用CEEMDAN分解可以有效提高复杂序列预测的精确性和准确性。

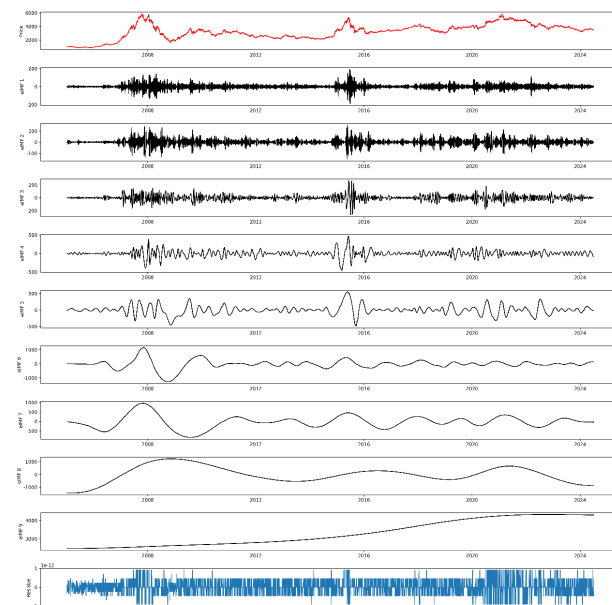


图4: 沪深300股指序列CEEMDAN自适应分解结果

(六) 沪深300的不同模型预测效果对比

图5给出了六组模型对沪深300在测试集上的预测效果图，相较于其它对比模型，CEEMDAN-CNN-LSTM-AM预测值与真实值的曲线更加贴合，预测滞后性和曲线右偏的问题得到显著改善，偏移程度最小，预测精度最高。不同模型的具体评价指标结果如表2所示。表2结果显示，CEEMDAN-CNN-LSTM-AM在所有指标上均显著优于其余五种模型。其MAE较CNN-LSTM-AM、CNN-LSTM、LSTM、MLP与SVR分别降低5.7%、36.2%、50.7%、43.8%与59.6%；RMSE分别减少17.0%、37.5%、46.6%、50.1%与64.6%；MAPE分别下降3.4%、34.1%、50.6%、40.8%与56.4%；决定系数 R^2 则分别提升1.0%、3.3%、5.3%、6.4%与14.8%。综合来看，CEEMDAN分解与注意力机制

的联合引入, 使该模型在股价预测任务中展现出更强的稳健性与更高的精度。

表2: 沪深300的不同模型评价结果

模型	MAE	RMSE	MAPE	R ²
CEEMDAN-CNN-LSTM-AM	78.320	91.081	1.831	0.979
CNN-LSTM-AM	83.066	109.721	1.896	0.969
CNN-LSTM	122.786	145.681	2.779	0.946
LSTM	158.973	170.472	3.711	0.926
MLP	139.427	182.409	3.092	0.915
SVR	193.886	257.069	4.203	0.831

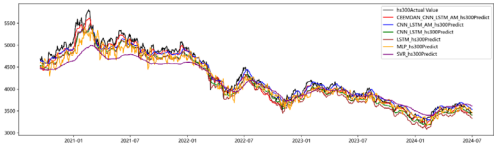


图5: 沪深300股指不同模型预测曲线图

(七) 中证500的不同模型预测效果对比

为了检验自研模型的稳健性和泛化能力, 本研究选取同一时间段的中证500指数, 进行相同的数据集处理和自适应分解, 进一步有效评估 CEEMDAN-CNN-LSTM-AM 模型的学习迁移能力。表3给出了六组模型在中证500指数测试集上的评价结果, 图6是不同模型的预测曲线对比图。

可以看到, 相较于沪深300, 中证500覆盖中盘成长股, 平均市值较小, 成分股中信息技术、工业、材料行业权重更高, 主要以成长型、高弹性为主, 跟踪的是高景气行业, 故中证500的整体波动幅度较大, 突变因素较多, 因此六组模型的学习拟合能力都有所下降。尽管如此, CEEMDAN-CNN-LSTM-AM 模型仍然能充分分解信号并提取关键特征, 显示出良好的预测性能, 可迁移能力和泛化能力都更为出色。

表3: 中证500股指的不同模型评价结果

模型	MAE	RMSE	MAPE	R ²
CEEMDAN-CNN-LSTM-AM	90.423	118.203	1.519	0.959
CNN-LSTM-AM	153.035	186.857	2.434	0.898
CNN-LSTM	188.371	216.705	2.992	0.863
LSTM	200.106	220.947	3.238	0.858
MLP	202.503	234.873	3.247	0.839
SVR	195.739	248.323	3.281	0.821



图6: 中证500股指不同模型预测曲线图

三、结束语

本研究提出多尺度深度预测框架 CEEMDAN-CNN-LSTM-AM。以沪深300为标的, 滚动窗口实验显示, 该模型全面碾压五种基线: 与次优的 CNN-LSTM-AM 相比, RMSE 降低 17.0%, MAE 下降 5.7%, MAPE 缩减 3.4%, R² 提高 1.0 个百分点至 0.979; 相较于传统 LSTM, RMSE、MAE、MAPE 分别锐减 46.6%、50.7%、50.6%, R² 提升 5.3 个百分点。结果表明, “分解-独立建模-注意力加权-重构”链路有效解耦复杂频率混叠, 显著抑制噪声并强化特征。该框架为高频风控与量化配置提供了精度更高、解释性更强的工具。

参考文献

[1] 肖田田. 基于 K-means-LSTM 模型的证券股价预测 [J]. 科技和产业, 2024, 24(03): 210-215. DOI: CNKI: SUN: CYK.0.2024-03-030.

[2] 马致远. 基于股票预测模型 LSTM 的降维比较 [J]. 科技和产业, 2025, 25(11): 8-16. DOI: CNKI: SUN: CYK.0.2025-11-002.

[3] 沈山山, 李秋敏. 基于注意力机制的 CNN-LSTM 短期股票价格预测 [J]. 软件, 2022, 43(02): 73-75. DOI: CNKI: SUN: RJZZ.0.2022-02-043.

[4] 刘荟. 基于深度学习 BiLSTM-Attention 模型的股价崩盘风险预测研究 [J]. 信息系统工程, 2025, (08): 146-149. DOI: CNKI: SUN: XXXT.0.2025-08-038.

[5] 刘铭, 单玉莹. 基于 EMD-LSTM 模型的股指收盘价预测 [J]. 重庆理工大学学报 (自然科学), 2021, 35(12): 269-276.

[6] 黄后菊, 李波. 基于 VMD-CSSA-LSTM 组合模型的股票价格预测 [J]. 南京信息工程大学学报, 2024, 16(03): 332-340.

[7] 贺毅岳, 李萍, 韩进博. 基于 CEEMDAN-LSTM 的股票市场指数预测建模研究 [J]. 统计与信息论坛, 2020, 35(06): 34-45.

[8] 丛敬奇, 成鹏飞, 赵振军. 基于 CEEMD-CNN-LSTM 的股票指数集成预测模型 [J]. 系统工程, 2023, 41(04): 104-116.

[9] 皇甫陈荫, 阮贺彬, 徐俊俊. 融合 CEEMDAN-CNN-LSTM 的风电机组多气象场景功率回归预测 [J]. 综合智慧能源, 2025, 47(09): 38-50.

[10] 朱婷, 颜七笙. 融合改进变分模态分解和 CNN-Bi LSTM-AM 的股价预测研究 [J]. 辽宁工业大学学报 (自然科学版), 2025, 45(02): 134-140. DOI: 10.15916/j.issn1674-3261.2025.02.012.