秦传东, 孟立群. 基于 VMD-SAPW-LSTM 模型的股价预测研究[J]. 智能计算机与应用,2026,16(5):20-27. DOI:10.20169/j.issn.2095-2163.24083001

基于 VMD-SAPW-LSTM 模型的股价预测研究

秦传东[1,2], 孟立群[1]

(1 北方民族大学 数学与信息科学学院, 银川 750021; 2 宁夏智能信息与大数据处理重点实验室, 银川 750021)

摘 要: 针对股票市场股价预测的高精度需求,本文创新性地提出了一种融合变分模态分解(VMD)、长短期记忆网络(LSTM)以及自注意力机制(SA)与位置加权注意力机制(PW)的复合预测模型—VMD-SAPW-LSTM。该模型首先利用 VMD 信号分解技术将复杂的股票收盘价时间序列分解为多个具有不同频率特性的模态分量,有效降低了数据的复杂性和非平稳性。随后,在 LSTM 网络框架中融入 SA 与 PW 机制,不仅增强了模型对时间序列数据中关键信息的捕捉能力,还通过位置信息的加权处理,进一步提升了模型对时间序列动态变化的敏感度。实验验证阶段,采用 Adam 优化器对模型参数进行精细调整,显著优化了模型的预测性能。实验结果显示,VMD-SAPW-LSTM 模型在股价预测任务中展现出卓越的表现,其预测准确性相较于传统方法及其他深度学习模型有了显著提升。本研究不仅为股票市场的精准预测开辟了新的路径,也深刻展示了深度学习技术在金融时间序列分析领域的巨大应用潜力和广阔前景。

关键词: 变分模态分解; 长短期记忆网络; 自注意力机制; 位置加权注意力机制; 股价预测

中图分类号: TP391 **文献标志码**: A **文章编号**: 2095-2163(2026)05-0020-08

Research on stock price prediction based on VMD-SAPW-LSTM model

QIN Chuandong[1,2], MENG Liqun[1]

(1 School of Mathematics & Information Science, North Minzu University, Yinchuan 750021, China;
2 Key Laboratory of Intelligent Information and Big Data Processing of Ningxia, Yinchuan 750021, China)

Abstract: To achieve a highly accurate solution for stock price prediction, an innovative composite forecasting model—VMD-SAPW-LSTM—has been proposed. This model integrates Variational Mode Decomposition (VMD), Long Short-Term Memory Network (LSTM), Self-Attention Mechanism (SA), and Position-Weighted Attention Mechanism (PW). The model first uses Variational Mode Decomposition (VMD) to decompose the complex stock closing price time series into multiple modal components with different frequency characteristics. This effectively reduces the complexity and non-stationarity of the data. Subsequently, integrating the Self-Attention (SA) and Position-Weighted (PW) mechanisms within the LSTM network framework not only enhances the model's ability to capture key information in the time series data but also improves its sensitivity to dynamic changes through weighted handling of positional information. During the experimental validation phase, the Adam optimizer was used to fine-tune the model parameters, significantly enhancing the model's predictive performance. Results show that the VMD-SAPW-LSTM model demonstrates outstanding performance in stock price prediction, with accuracy significantly improved compared to traditional methods and other deep learning models. This research not only opens new avenues for precise stock market forecasting but also highlights the vast potential and broad prospects of deep learning technologies in financial time series analysis.

Key words: Variational Mode Decomposition; LSTM; Self-Attention Mechanism; Position Weighted Attention; stock price prediction

0 引 言

股票市场作为全球经济的重要组成部分,其价格预测一直是金融研究的核心问题之一。传统时间序列分析方法在处理复杂金融数据时常常面临挑战,尤其是在捕捉数据中的非线性和动态变化特征方面。

近年来,深度学习技术的快速发展为股票价格预测带来了新的机遇。王雨涵等[1]提出了基于 BP

基金项目: 国家自然科学基金(61562001);宁夏回族自治区教育厅高等学校科学研究项目(NYG2024093); 北方民族大学研究生创新项目(YCX23096)。

作者简介: 秦传东(1976—),男,博士,副教授。主要研究方向:机器学习与大数据技术。

通信作者: 孟立群(1999—),女,硕士。主要研究方向:智能计算与大数据分析。Email: mlqqq_921@163.com。

收稿日期: 2024-08-31

神经网络(BPNN)和高阶模糊认知图(HFCMs)的股票价格预测模型。在深度学习应用于股票预测的研究中,长短期记忆网络(LSTM)因其优越的时间序列建模能力而受到广泛关注。LSTM 通过引入门控机制来解决传统递归神经网络(RNN)在处理长序列时面临的梯度消失和梯度爆炸问题,能够有效捕捉时间序列数据中的长期依赖关系。张瑞雪等[2]提出基于深度学习的股价预测研究,通过时间注意力机制和 LSTM 预测股价序列。吴峰等[3]提出了新型的双流 LSTM 神经网络,在股价趋势预测中表现优越。赵浩博[4]提出基于卷积和循环神经网络模型融合的股票开盘价预测研究方法。李婧琦[5]提出一种基于鲸鱼算法(WOA)优化长短期记忆网络(LSTM)的股票价格预测模型。邹婕[6]融合自注意力机制(SA)和 GRU 模型进行股价预测。许丽[7]提出了一种基于 LSTM 和新闻股票情感分析的组合优化模型 SVM_LSTM。

融合信号处理与深度学习的方法能更有效地挖掘股价序列蕴含的深层变化规律,实现金融时序数值和趋势的高精度预测。孙晨宇等[8]提出基于 CEEMDAN 分解的 LSTM-Attention 模型进行股市预测。孔繁苗等[9]提出 EMD-LSTM-LB 分频时序预测算法,将 EMD 分解后的序列分类,极端子序列由 LSTM 进行训练预测,非极端子序列由 LB 进行训练预测。史建楠等[10]通过 DMD 算法对股票数据进行分解计算,然后采用 LSTM 网络对基本面数据和模态特征进行价格建模预测。黄后菊等[11]建立 VMD-CSSA-LSTM 组合模型对股票价格进行预测。郭金录[12]基于 VMD-EEMD-LSTM 模型对沪深 300 进行指数预测。

本研究采用了变分模态分解(Variational Mode Decomposition, VMD)与深度学习算法相结合的方式对股票收盘价进行预测。首先,本文利用变分模态分解(VMD)技术对股票收盘价时间序列进行模态分解。随后,本文在长短期记忆神经网络(LSTM)模型中引入了自注意力机制(SA)和位置加权注意力机制(PW),形成了 SAPW-LSTM 模型。自注意力机制能够使模型更好地捕捉序列中的重要依赖关系,而位置加权注意力机制则进一步优化了序列中不同位置的权重,从而增强了模型对复杂时间序列的预测能力。

1　原理介绍

1.1　变分模态分解

变分模态分解 (Variational Mode Decomposition, VMD)是一种用于信号分析的工具,旨在将复杂信号分解成一组具有特定频率特征的固有模态函数(Intrinsic Mode Functions, IMFs)。VMD 分解的具体步骤如下:

1) 初始化模态函数和中心频率

初始时,通常选择模态函数为零或其他简单形式,随机选择中心频率[13]。

2) 构建 VMD 的目标函数,引入拉格朗日乘子

每个模态分量具有特定的中心频率和有限带宽,VMD 旨在最小化所有模态的带宽估计值之和。VMD 的信号分解过程实际上等同于求解一个变分约束问题[14],原始的约束变分问题表达式如下:

$$\begin{cases}\min\sum_k \left\| \partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)\cdot u_k(t)\right] e^{-j\omega_k t}\right\|_2^2 \\ \text{s.t.}\ \sum_k u_k = f\end{cases} \tag{1}$$

其中, $u_k(t)$ 为分解后得到的第 k 个固有模态函数(IMFs)分量; ω_k 为第 k 个 IMFs 对应的中心频率; $\delta(t)$ 为冲击函数; f 是原始输入信号。

通过 Hilbert 变换得到与 $u_k(t)$ 相关的解析信号进而获取单边谱,乘以 $e^{-j\omega_k t}$ 调节每个模态的频谱,使之调制到相应的基频带[15]。为求上述模型的最优解,引入二次惩罚项和拉格朗日乘数,以使约束变分问题不受约束。构造增强拉格朗日矩阵 $\mathscr{L}$ 后,式(1)的变分问题转化为下式:

$$\begin{aligned}L(\{u_k\},\{\omega_k\},\lambda) = \alpha\sum_k \left\| \partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right) * u_k(t)\right] e^{-j\omega_k t}\right\|_2^2 + \\ \left\| f(t) - \sum_k u_k(t)\right\|_2^2 + \\ \left\langle \lambda(t), f(t) - \sum_k u_k(t)\right\rangle\end{aligned} \tag{2}$$

其中, $\{u_k\}$ 为模态分量集合; $\{\omega_k\}$ 为中心频率集合; λ 为拉格朗日乘子; $*$ 为卷积运算符。

3)采用变分优化求解模态函数和中心频率

更新模态函数 $u_k(t)$:

$$\frac{\partial L}{\partial u_k} = -2\left[x(t) - \sum_{j=1}^{k} u_j(t)\right] + 2\alpha\left(\frac{\partial^2 u_k(t)}{\partial t^2} + u_k(t)\right) + \lambda(t) = 0 \tag{3}$$

更新中心频率 ω_k:

$$\frac{\partial L}{\partial \omega_k} = -2\alpha\int_{-\infty}^{\infty} u_k(t)\left(\frac{\partial u_k(t)}{\partial t} + 2\pi\omega_k u_k(t)\right)\mathrm{d}t = 0 \tag{4}$$

4) 更新拉格朗日乘子 λ

在每次迭代中,拉格朗日乘子 $\lambda(t)$ 需要根据当前的模态函数 $u_k(t)$ 的值来进行更新,更新公

式为：

$$\lambda^{(n+1)}(t) = \lambda^{(n)}(t) + \mu[x(t) - \sum_{k=1}^{k} u_k^{(n+1)}(t)] \tag{5}$$

其中，$\lambda^{(n)}$ 是第 n 次迭代的拉格朗日乘子，μ 是步长，$\sum_{k=1}^{k} u_k^{(n+1)}(t)$ 是第 $n+1$ 次迭代中的模态函数的和。

5）确定算法的收敛性，并输出分解结果

给定精度 $\varepsilon > 0$，若满足

$$\frac{\sum_{k} \| \hat{u}_k^{n+1} - \hat{u}_k^{n} \|_2^2}{\| \hat{u}_k^{n} \|_2^2} < \varepsilon \tag{6}$$

则停止迭代，否则返回 3）。

1.2 LSTM 神经网络

长短期记忆（LSTM）网络是循环神经网络（RNN）的拓展模型，其在一定程度上解决了传统 RNN 模型在处理长序列时出现的梯度消失或梯度爆炸的问题[16]，在处理具有长时依赖关系的序列数据时表现优异。图 1 展示了 LSTM 的一个记忆门以及 3 个门控单元，3 个门控单元分别为遗忘门、输入门与输出门，这些门控单元可以决定是否保留信息来调节信息流，帮助网络更好地处理长期依赖关系。遗忘门的输出值由 Sigmoid 函数得到，其根据 h_{t-1} 和 x_t 的输出组合来决定是否保留对应的信息，表达式如下：

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \tag{7}$$

其中，h_{t-1} 为上一时刻的输出；x_t 为当前时刻的输入；w_f 为遗忘门权重；b_f 为遗忘门偏执。

Sigmoid 函数计算公式如下：

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{8}$$

输入门会根据 Sigmoid 函数决定添加哪些新信息，随后 tanh 函数会创建 t 时刻记忆单元的候选状态 $\hat{C}_t$[17]，表现形式如下：

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{9}$$

$$\hat{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{10}$$

其中，tanh 函数的计算公式如下：

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{11}$$

旧状态 C_{t-1} 会与 f_t 相乘，遗忘掉由 f_t 所确定的需要遗忘的信息，之后加上 $i_t \otimes \hat{C}_t$，由此得到新的记忆细胞状态 C_t：

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \hat{C}_t \tag{12}$$

输出门最终决定产生什么输出信息，O_t 和隐含层 h_t 的更新公式如下：

$$O_t = \sigma(W_o \cdot [h_{t-1}, p_t] + b_o) \tag{13}$$

$$h_t = O_t \otimes \tanh(C_t) \tag{14}$$

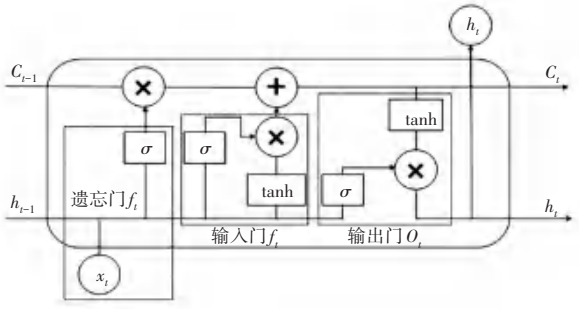


图 1 LSTM 模型结构

Fig. 1 LSTM model architecture

1.3 自注意力机制

自注意力机制（Self-Attention Mechanism）[18]是一种能够学习序列中各个元素之间关系的机制，对于序列中的每个元素，其会根据该元素与其他元素的相关性对其赋予权重，从而得到该元素在序列中的表示。

在每个时间步，自注意力模块可以对所有位置的特征进行加权求和[19]，选择性地聚集每个位置的输入特征，使模型可在垂直交叉堆叠的多层 LSTM 中捕获全局空间依赖性[20]。该机制可以克服长距离依赖带来的信息丢失，更善于捕获内部数据特征。

1.4 位置权重注意力机制

位置权重注意力机制（Positional Weight Attention Mechanism）[21]是注意力机制的一种变体，旨在通过加权处理来强调时间序列中关键时间点的信息，从而提高模型对股票收盘价预测的准确性[22]。该机制认为，在股票价格变动的过程中，不同时间点的信息对于预测未来收盘价的重要性是不同的。

2 原理介绍

2.1 数据来源

本文选取招商银行（600036. SH）2014 年 5 月 1 日至 2024 年 5 月 1 日的日度收盘价数据为主要研究对象，数据来自 Tushare 财经共享数据集，共获取 2 028 个交易日历史数据，如图 2 所示。

招商银行自 2014 年至 2024 年期间，股票整体呈现上升趋势，招商银行在资本市场中拥有庞大的市值规模，且其市盈率保持相对较低水平，显示出市

场对其未来成长潜力的乐观预期。在股价走势上，招商银行长期展现出上涨态势，自 2016 年以来股价实现数倍增长，这一表现不仅跨越了包括 2018 年中美贸易紧张局势在内的市场调整期（如熊市），也涵盖了如 2020 年全球经济复苏带动的牛市行情，因此其股价数据具有高度的市场代表性和研究价值。



图 2　招商银行近 10 年股价走势图

Fig. 2　Stock price trend chart of China Merchants Bank in recent 10 years

2.2　VMD 分解股价信息

进行 VMD 分解时，选择合适的 K 值对于算法的性能至关重要。如果 K 值设置得太小，可能无法捕捉到所有重要的模态，导致分解结果不完整；如果设置得太大，则可能会产生无意义的模态，增加计算复杂度和噪声干扰。

ADF（Augmented Dickey-Fuller）检验是一种广泛应用的统计方法，旨在检测时间序列数据中是否存在单位根，从而判断该序列是否为平稳过程。在统计推断中，ADF 检验的统计量越小，通常意味着拒绝原假设（即序列中存在单位根的假设）的置信度越高，意味着序列的平稳性越强，非平稳的可能性显著降低。

基于上述原理，本文采用 ADF 检测对各分量序列进行评估，选取 ADF 统计量整体最小的 K 值，以确保所得到的分量序列集合在统计上最为接近平稳状态，将非平稳性风险降至最低。表 1 展示了不同 K 值配置下，各分量序列通过 ADF 检测获得的统计量值，为确定最优 K 值提供了数据支持。

由表 1 可见，与 $K=5$ 时相比，当 $K=6$ 时，各个分量的 ADF 值均有所上升，因此，可以合理推断，当模态数量增加至 6 个时，VMD 呈现出过度分解的迹象，这表明进一步的模态增加并未带来信息分解上的实质性增益，反而可能引入了不必要的复杂性和噪声。

综上所述，本文选择带宽约束（也称惩罚因子）$\alpha=7\,000$、收敛停止条件 $\varepsilon=10^{-7}$、分解模态个数 $K=5$ 来分解招商银行（600036.SH）2 028 个交易日历史数据，分解效果如图 3 所示，将收盘价时间序列数据分解成 5 个模态和一个残差项。

表 1　不同 K 值下的 ADF 检验值

Table 1　ADF test values under different K values

	$K=2$	$K=3$	$K=4$	$K=5$	$K=6$	$K=7$
IMF1	-2.12	-4.74	-4.83	-5.55	-5.42	-0.42
IMF2	-11.30	-9.54	-9.32	-5.93	-5.93	-5.66
IMF3		-15.49	-14.28	-14.20	-13.35	-12.30
IMF4			-13.62	-14.62	-12.06	-16.77
IMF5				-13.08	-12.08	-13.80
IMF6					-11.08	-14.23
IMF7						-14.92

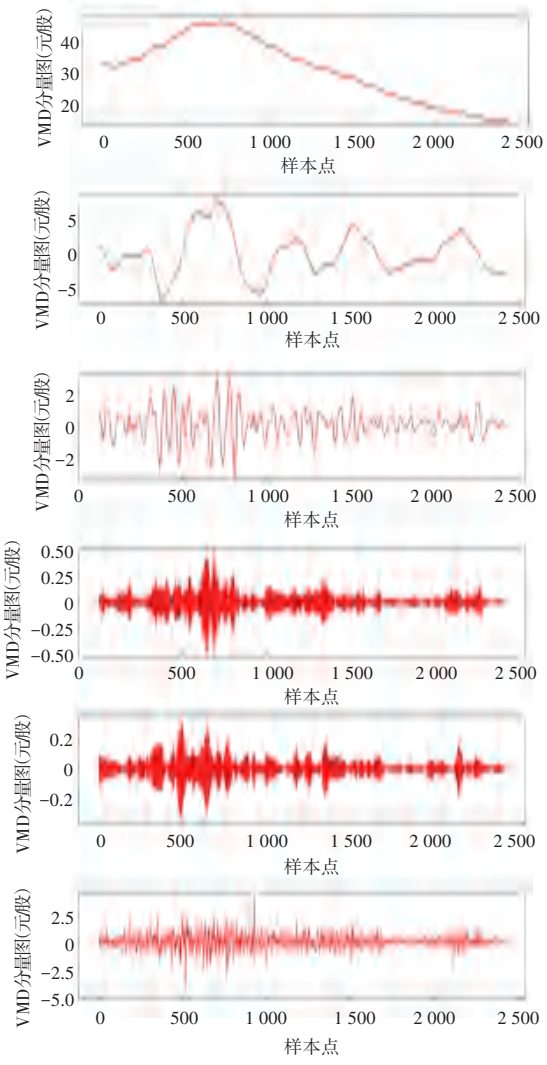


图 3　VMD 分解结果

Fig. 3　Results of Variational Mode Decomposition

2.3　数据预处理

对使用 VMD 分解后的每个时间序列应用 ADF（Augmented Dickey-Fuller）检验，判断序列是否平稳。若序列平稳，则不做处理，若序列不平稳，则进行一阶差分处理。

一阶差分处理(First Difference)是一种常见的时间序列预处理方法,用于使非平稳序列转化为平稳序列。非平稳序列可能表现出趋势性、季节性或者其他随时间变化的性质,这些性质会影响本文的分析和预测结果的准确性。对于时间序列 $\{y_t\}$,其一阶差分序列 $(y_t\}$ 可以定义为:

$$\Delta y_t = y_t - y_{t-1} \tag{15}$$

其中,Δy_t 表示在时间点 t 的一阶差分值;y_{t-1} 表示时间点 $t-1$ 的值;一阶差分是当前值与前一期值之间的差值,反应了序列在时间单位内的变化量。

在时间序列预测中,通常需要将过去的数据用作特征进行预测。股票价格的时间序列数据绝对值范围可能很大,归一化可以将数据压缩到统一的范围内,确保所有过去的特征在相同的尺度上,从而使模型在训练时更加稳定。

本文使用 MinMaxScaler 对差分后的序列进行归一化,将数据范围缩放到 [0, 1] 之间,以便模型更好地学习,公式如下:

$$x_{\text{Scaled}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{16}$$

其中,x 是原始数值;$x_{\min}$ 是特征最小值;$x_{\max}$ 是特征最大值;x_{Scaled} 是归一化后的数值。

2.4 模型评价指标

为定量评估模型的预测性能,本文采取均方根误差(RMSE)、平均绝对误差(MAE)和决定系数(R^2)3 种指标来检测模型的预测效果。

均方根误差(RMSE)是预测值与真实值之间差异的平方的均值,值越小,模型预测精度越高,计算公式如下:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i^*)^2} \tag{17}$$

平均绝对误差(MAE)是预测值与真实值之间差异的绝对值的平均值,值越小,模型预测精度越高,计算公式如下:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - y_i^*| \tag{18}$$

决定系数(R^2)是因变量的变异性能够由模型解释的比例,即模型对数据的拟合程度,越接近 1,模型拟合效果越好,计算公式如下:

$$R^2 = 1 - \sum_{i=1}^{n}(y_i - y_i^*)^2 \Big/ \sum_{i=1}^{n}(y_i - y^2)^2 \tag{19}$$

2.5 SAPW-LSTM 模型构建

自注意力机制(SA)允许模型在处理股票时间序列数据时,关注当前、过去和未来时间点的相关信息。但其本身并不直接考虑元素之间的位置关系。位置权重注意力机制(PW)通过给序列中的每个元素分配一个位置编码,使得模型能够感知到元素之间的顺序和位置差异。这对于股票预测尤为重要,因为股票价格的变化往往与时间序列中的位置信息密切相关。

结合自注意力机制和位置权重注意力机制,模型能够更全面地捕捉股票时间序列数据中的关键信息和模式[18]。这种组合方式有助于提升模型在复杂市场环境下的预测精度和稳定性。

因此,本文在 LSTM 模型的基础上,先后引入自注意力机制和位置权重注意力机制,构建 SAPW-LSTM 模型,用来分别预测由 VMD 分解后的各个分量,模型如图 4 所示。

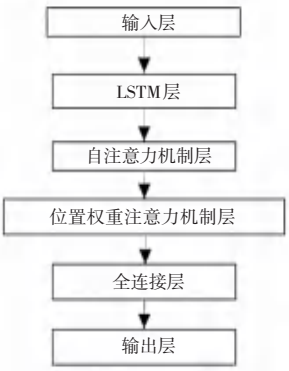


图 4 SAPW-LSTM 模型

Fig. 4 SAPW-LSTM model

接下来用 SAPW-LSTM 模型对分解后的招商银行收盘价数据依次进行预测,用前 80%的数据作为训练集,剩下 20% 的数据作为测试集,训练模型时,采用 Adam 优化器调整参数,损失函数为 MSE,迭代次数为 10-0,效果如图 5 所示。

如图 5 所示,随着 Adam 优化算法迭代次数的稳步增长,模型在各个模态上的损失值呈现出显著的下降趋势,这一趋势明确表明模型的学习能力在不断增强,进而有效地提升了其预测精度与泛化能力,显著改善了整体模型的预测效果。模型预测结果如图 6—7 所示。

如图 6—7 所示,实线代表真实值,虚线代表预测值,通过详尽的数据可视化分析,可以明确地观察到所构建模型在拟合效果上展现出了高度的准确性。这两幅图不仅直观地反映了模型对历史数据的良好捕捉能力,更重要的是,其在很大程度上成功地预测并模拟了未来收盘价的走势,为市场趋势分析提供了有力的支持。接下来计算模型的预测误差,结果见表 2。

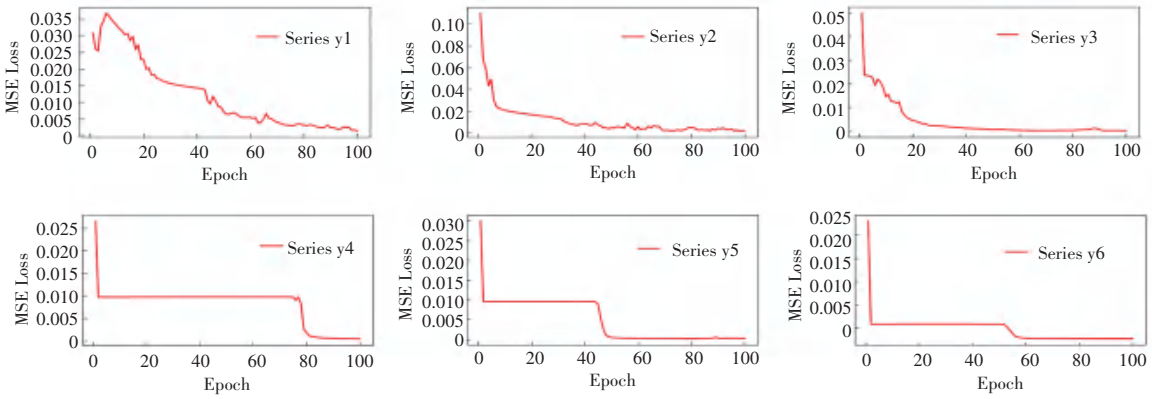


图 5　各个分量的 Adam 迭代效果

Fig. 5　Adam iteration effect of each component

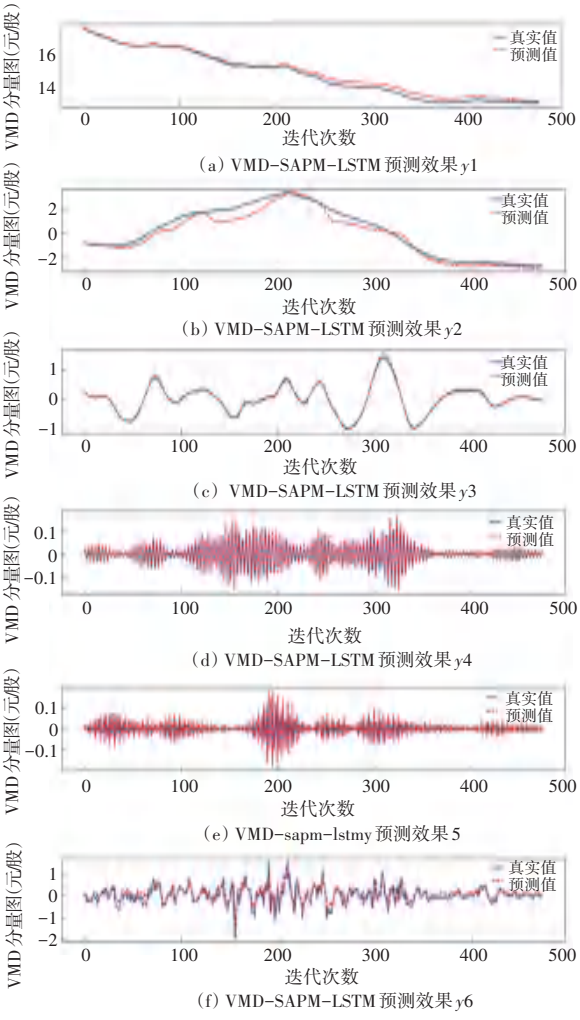


图 6　各个分量的预测效果

Fig. 6　Prediction effect of each component

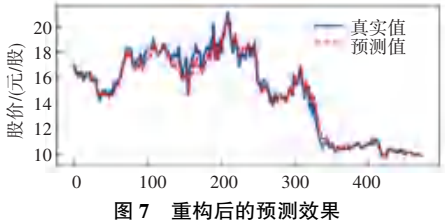


图 7　重构后的预测效果

Fig. 7　Prediction effect after reconstruction

表 2　各个序列的预测误差

Table 2　Prediction error of each sequence

序列	模态 1	模态 2	模态 3	模态 4	模态 5	模态 6	总数据
RMSE	0.003 1	0.495 7	0.036 9	0.009 0	0.006 0	0.298 0	0.474 0

表 2 中详细列出了每个模态的均方根误差(RMSE)计算结果，这一关键性能指标是衡量模型预测精度的重要标准。从表中数据可以看出，所有模态的 RMSE 值均稳定地保持在 0~0.5 的较小范围内，这一结果不仅验证了模型预测结果的高度一致性，也彰显了模型在复杂市场环境下依然能够保持低误差水平的高超性能。

2.6　不同模型效果对比

针对招商银行的历史收盘价数据，对不同时间序列建模技术进行了深入对比分析，本研究揭示了模型性能在多个关键评估指标上的显著差异，采用了均方根误差(RMSE)、平均绝对误差(MAE)和决定系数(R^2)作为核心评价指标，结果如表 3、图 8 所示。

表 3　不同预测模型对比分析

Table 3　Comparative analysis of different prediction models

模型	RMSE	MAE	R^2
RNN	2.327	1.431	0.686
LSTM	2.391	1.420	0.664
GRU	2.414	1.461	0.656
VMD-RNN	1.982	1.223	0.784
VMD-GRU	1.187	1.001	0.928
VMD-LSTM	0.712	0.788	0.975
VMD-SA-LSTM	0.581	0.688	0.983
VMD-PW-LSTM	0.482	0.601	0.988
VMD-SAPW-LSTM	0.473	0.587	0.988

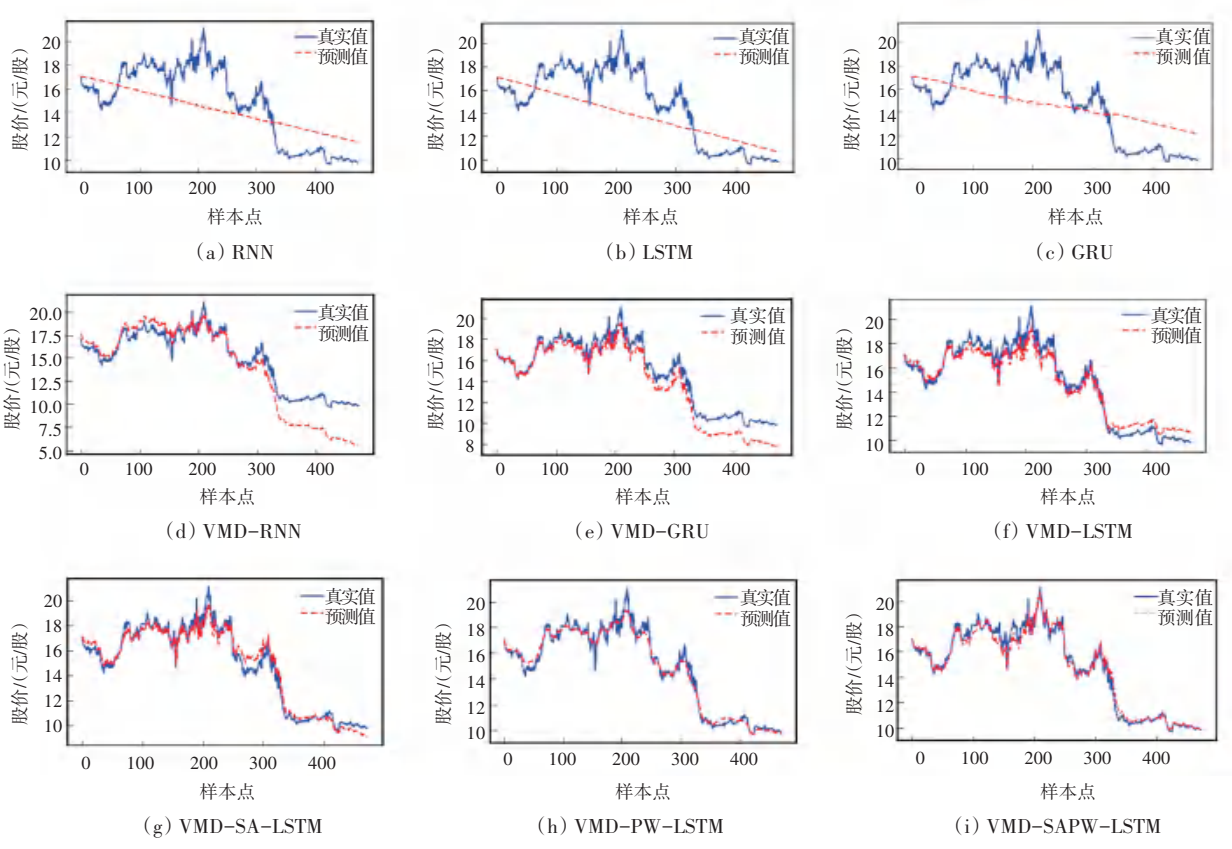


图 8 不同模型预测效果

Fig. 8 Prediction effect of different models

从实验结果来看,基础模型 RNN、LSTM 和 GRU 虽然在预测上展现了一定能力,但相比于融合 VMD 技术的模型,其 RMSE 和 MAE 值均较高,而 R^2 值较低,分别反映了预测误差的增大和模型拟合度的不足。值得注意的是,VMD 技术的引入显著提升了模型的预测性能,其中 VMD-LSTM 模型相较于其基础版本 LSTM,RMSE 从 2.391 5 降低至 0.712 3,MAE 从 1.420 7 降至 0.788 5, R^2 从 0.664 1 提升至 0.974 8,这一系列数据的变化充分展示了 VMD 在特征提取和降噪方面的有效性。

进一步地,VMD-SA-LSTM 和 VMD-PW-LSTM 模型通过引入注意力机制使得模型关注元素之间的相关性和位置,在 VMD-LSTM 的基础上进一步优化了预测精度,RMSE 分别降至 0.580 5 和 0.482 2, R^2 则分别提升至 0.983 3 和 0.988 5,特别是 VMD-SAPW-LSTM 模型,其 RMSE 低至 0.473 8,MAE 为 0.587 9, R^2 高达 0.988 9,这些数据不仅验证了复杂模型设计的优越性,也强调了多种策略融合在提升时间序列预测精度方面的重要性。

综上所述,本研究通过系统性的对比分析,证明了 VMD 技术在时间序列预测中的关键作用,以及注意力机制对于进一步优化模型性能的积极影响,为相关领域的研究和应用提供了有价值的参考。

2.7 与现有文献中的模型对比分析

为了全面评估 VMD-SAPW-LSTM 模型在金融市场预测中的稳定性和性能优势,本研究以上证指数为例,采用该模型进行了预测,并将预测结果与文献[6]中的 RF-SA-GRU 模型及文献[8]中一系列基于不同信号分解和深度学习方法的模型进行了对比分析。结果见表 4。

表 4 不同模型的上证指数预测效果

Table 4 Forecasting effect of Shanghai stock index with different models

模型	RMSE	MAE	R^2
RF-SA-GRU[6]	37.68	27.52	0.98
CEEMDAN-LSTM[8]	16.41	12.39	0.83
CEEMDAN-Seq2Seq[8]	17.41	13.47	0.81
EMD-LSTM-Attention[8]	16.34	13.18	0.83
EEMD-LSTM-Attention[8]	16.69	12.35	0.83
CEEMDAN-LSTM-Attention[8]	14.72	10.75	0.87
VMD-SAPW-LSTM	13.25	3.32	0.99

由表 4 可得,VMD-SAPW-LSTM 模型的 RMSE 值为 13.25,远低于其他对比模型,其中,表现次优的 CEEMDAN-LSTM-Attention 模型的 RMSE 为

14.72，表明该模型在预测误差控制上表现更为出色。同时，其 MAE 值仅为 3.32，进一步验证了模型预测值与实际值之间的高度一致性。R^2 值达到了 0.99，显著高于其他所有对比模型，体现了模型对上证指数走势的精准捕捉能力和高度拟合优度。

综上所述，本文提出的 VMD-SAPW-LSTM 模型在预测上证指数时，不仅误差小、精度高，而且稳定性强，显著优于现有文献中的其他模型。这一结果不仅验证了变分模态分解（VMD）、自注意力机制（SA）、位置加权注意力机制（PW）与 LSTM 相结合策略的有效性，也为金融市场预测领域提供了新的思路和方法，具有重要的学术价值和实际应用前景。

3　结束语

针对股票收盘价的时间序列，本文采用变分模态分解（VMD）技术对股票收盘价进行模态分解，随后在长短期记忆神经网络（LSTM）模型中引入自注意力机制（SA）和位置加权注意力机制（PW），构建 SAPW-LSTM 模型，对分解后的各子序列分别开展预测，最后将多个序列的预测值进行重构，得到最终的股票价格预测。该方法融合了信号处理技术和深度学习方法，旨在从原始数据中提取更具价值的有效特征，引入的两个注意力机制使得模型不仅关注每个元素之间的相关性，也关注元素之间的位置关系，位置对于金融时间序列至关重要，因此该方法能够更全面地捕捉股票时间序列数据中的关键信息与模式。

VMD-SAPW-LSTM 模型的引入不仅在理论上丰富了股票预测方法的工具箱，更在实践中展示了其在提升预测能力和解释市场行为方面的潜力，为该领域进一步研究和应用提供了新的思路和方法。

参考文献

[1] 王雨涵，张亚萌，魏国亮. 基于 BP 神经网络和高阶模糊认知图的股票价格预测[J]. 智能计算机与应用，2023，13（8）：100-106.

[2] 张瑞雪，郝泳涛. 基于深度学习的股价预测研究[J]. 电脑知识与技术，2023，19（33）：8-10.

[3] 吴峰，谢聪，姬少培. 基于双流 LSTM 神经网络的股价趋势预测[J]. 应用科学学报，2023，41（2）：344-358.

[4] 赵浩博，李锡祚. 基于卷积和循环神经网络模型融合的股票开盘价预测研究[J]. 智能计算机与应用，2019，9（6）：55-58.

[5] 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.

[6] 邹婕，李路. RF-SA-GRU 模型的股价预测研究[J]. 计算机工程与应用，2023，59（15）：300-309.

[7] 许丽，张利，李桂城，等. 基于 LSTM 和新闻情感的股票价格预测方法[J]. 智能计算机与应用，2022，12（5）：107-113.

[8] 孙晨宇，张树东. 基于 CEEMDAN 与 LSTM-Attention 的股市预测模型[J]. 计算机应用与软件，2023，40（12）：119-125.

[9] 孔繁苗，高鹭，李鹏程，等. EMD-LSTM-LB 分频时序预测算法[J]. 计算机工程与设计，2023，44（10）：3021-3030.

[10] 史建楠，邹俊忠，张见，等. 基于 DMD-LSTM 模型的股票价格时间序列预测研究[J]. 计算机应用研究，2020，37（3）：662-666.

[11] 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（自然科学版），2020，12（5）：1-13.

[12] 郭金录. 基于 VMD-EEMD-LSTM 模型的沪深 300 指数预测研究[J]. 现代财经（天津财经大学学报），2020，40（8）：31-44.

[13] GUO W，LIU Q，LUO Z，et al. Forecasts for international financial series with VMD algorithms [J]. Journal of Asian Economics，2022，80：101458.

[14] GAO Z，ZHANG J. The fluctuation correlation between investor sentiment and stock index using VMD - LSTM：Evidence from China stock market [J]. The North American Journal of Economics and Finance，2023，66：101915.

[15] WU J，HU Y，WU D，et al. An aquatic produ-ct price forecast model using VMD - IBES - LSTM hybrid approach [J]. Agriculture，2022，12（8）：1185.

[16] HOCHREITER S，SCHMIDHUBER J. Long short-term memory [J]. Neural Computation，1997，9（8）：1735-1780.

[17] ZHANG Y，WANG S，WANG H. A hybrid deep learning model for stock price prediction using LSTM and attention mechanism [J]. Knowledge-based Systems，2021，212：106636.

[18] LI Y，ZHENG Y，LIU Y，et al. Forecasting stock prices with LSTM and attention mechanism [J]. Expert Systems with Applications，2020，155：113491.

[19] 李晓明，张磊，张蕾. 基于自注意力机制的序列建模方法研究[J]. 计算机学报，2020，43（6）：1181-1193.

[20] 张磊，陈静，刘洋. 自注意力机制及其在深度学习中的应用综述[J]. 计算机科学，2019，46（11）：29-36.

[21] 刘伟，刘鹏，王志勇. 基于位置编码的注意力机制改进方法[J]. 计算机研究与发展，2019，56（10）：2104-2115.

[22] 张磊，李伟，周宁. 结合自注意力机制与位置编码的机器翻译模型[J]. 中文信息学报，2019，33（5）：60-68.