

基于双通道输入的CNN-BiLSTM模型的 股价预测

卢钟行, 张琴*

重庆交通大学数学与统计学院, 重庆

收稿日期: 2026年3月8日; 录用日期: 2026年4月2日; 发布日期: 2026年4月9日

摘要

针对现有混合神经网络模型在处理多源时序数据时输入结构单一、难以有效融合不同频率特征的问题, 提出一种基于双通道输入的CNN-BiLSTM股价预测模型。该模型设计双通道输入结构, 分别处理5分钟级高频数据(开盘价、最高价、最低价、收盘价、成交量)和日级低频数据(日均价、日成交量、调整后收盘价): 高频通道通过两层卷积神经网络(CNN)提取日内局部波动特征, 低频通道直接输入双向长短期记忆网络(BiLSTM)捕捉长期趋势依赖, 最后将双通道特征进行门控融合以完成股价预测。以贵州茅台(600519.SH) 2020年1月至2025年12月的交易数据为实证对象, 采用50日滑动窗口预测未来5日收盘价。实验结果表明, 所提模型在测试集上取得预测准确率为98.07%, 显著优于仅使用日级数据的RNN、GRU、LSTM, 仅使用分钟级数据的CNN-LSTM、单通道CNN-BiLSTM以及LSTM-Transformer等基准模型, 验证了双通道输入结构在捕捉多尺度时序特征方面的有效性。研究为高频金融时序预测提供了新的模型输入端架构思路, 对投资者的量化决策具有积极的参考价值。

关键词

双通道输入, CNN-BiLSTM, 多尺度数据, 股价预测, 贵州茅台

Stock Price Prediction Using a Dual-Channel Input CNN-BiLSTM Model

Zhongxing Lu, Qin Zhang*

School of Mathematics and Statistics, Chongqing Jiaotong University, Chongqing

Received: March 8, 2026; accepted: April 2, 2026; published: April 9, 2026

*通讯作者。

文章引用: 卢钟行, 张琴. 基于双通道输入的 CNN-BiLSTM 模型的股价预测[J]. 应用数学进展, 2026, 15(4): 246-257.
DOI: 10.12677/aam.2026.154154

Abstract

To address the problem that existing hybrid neural network models have a single input structure and cannot effectively fuse features of different frequencies when processing multi-source time series data, this paper proposes a CNN-BiLSTM stock price prediction model based on dual-channel input. The model designs a dual-channel input structure to separately process 5-minute high-frequency data (open price, high price, low price, close price, volume) and daily low-frequency data (daily average price, daily volume, adjusted close price). The high-frequency channel uses a two-layer convolutional neural network (CNN) to extract intraday local fluctuation features, while the low-frequency channel directly inputs data into a bidirectional long short-term memory network (BiLSTM) to capture long-term trend dependencies. Finally, the dual-channel features are fused via a gating mechanism to complete stock price prediction. Taking the trading data of Kweichow Moutai (600519.SH) from January 2020 to December 2025 as the empirical object, a 50-day sliding window is used to predict the closing price of the next five days. Experimental results show that the proposed model achieves a prediction accuracy of 98.07% on the test set, significantly outperforming baseline models such as RNN, GRU, LSTM using only daily data, CNN-LSTM using only minute-level data, single-channel CNN-BiLSTM, and LSTM-Transformer, verifying the effectiveness of the dual-channel input structure in capturing multi-scale temporal features. This study provides a novel model input architecture for high-frequency financial time series prediction and has positive reference value for investors' quantitative decision-making.

Keywords

Dual-Channel Input, CNN-BiLSTM, Multi-Scale Data, Stock Price Prediction, Kweichow Moutai

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

股票市场作为金融体系的核心,其价格波动直接影响投资者收益与资本市场稳定。我国自1989年试点建立上海证券交易所以来,资本市场迅猛发展,投资者参与度持续提升[1]。然而,传统的股价分析方法(如基本面分析和技术指标分析)在非线性特征挖掘方面始终存在不足,且对快速变化的市场环境反应滞后[2]-[6],难以满足高频交易与实时决策的需求。

近年来,深度学习技术通过构建多层神经网络结构,能够从海量数据中自动提取复杂非线性关系,为股价预测提供了新的方法论。该技术在图像识别和自然语言处理领域的突破性进展,推动了其在金融时序预测中的应用探索[7]。股价数据的高噪声、动态性和非平稳性,与深度学习处理复杂非线性问题的能力高度契合,使得基于机器学习的股价预测方法逐渐成为该领域的主流研究方向[8]-[10]。

早期研究尝试将经典神经网络应用于股价预测。Kamijo和Tanigawa[11]利用RNN对东京证券交易所股票数据进行建模,验证了循环神经网络在时序预测中的可行性。Moghar和Hamiche[12]基于LSTM模型对Google和NKE资产进行预测,进一步证实了LSTM在捕捉长期依赖关系上的优势。Di Persio和Honchar[13]通过对比RNN、LSTM和GRU模型的性能,发现LSTM在股价预测任务中具有更高的准确率。这些研究为深度学习模型在金融时序预测中的应用奠定了基础。

随着研究的深入, 学者们发现单一模态数据的特征表达能力有限。Kraus 和 Feuerriegel [14]在分析公司临时公告文本数据时, 发现融合文本情感特征的 LSTM 模型能够显著提升股价波动预测精度。Peng 和 Guo [15]利用 LSTM 和 Transformer 对中国 A 股银行股进行预测, 创新性地将输入数据划分为多个时间窗口(如 2 年前、1 年前、一周前等), 通过多尺度特征提取提升了模型性能。这表明, 数据预处理与特征工程的优化是提高预测精度的关键环节, 且随着时间跨度的增加, 数据量增长的同时带来的长期时序模型会使得单一的神经网络模型的鲁棒性下降, 这为模型设计提出了新的挑战。

在混合模型研究方面, CNN-LSTM 因其独特的结构设计而备受关注。CNN 通过卷积核提取局部波动特征, LSTM 则建模时间维度的长期依赖关系, 二者协同工作显著提升了预测精度[16]-[18]。然而, 现有研究仍存在以下不足: (1) 特征融合机制单一, 传统方法简单拼接 CNN 与 LSTM 的输出, 忽略不同尺度特征的重要性差异; (2) 输入结构局限, 多数模型采用单通道输入, 将不同频率数据简单混合, 导致高频噪声与低频趋势相互干扰; (3) 长期依赖建模能力有限, LSTM 对超长期时序模式(如年度经济周期)的捕捉能力仍有提升空间。针对这些问题, 研究者通过不同的模型组合及多源数据融合(如新闻、公告等文本情感数据) [19] [20], 不断提升股价预测的准确率, 然而, 投资者通常关注中长期的持有收益, 如何有效地利用长期历史数据为投资者提供参考, 仍是做股价预测工作中关键的问题所在。

基于此, 本文将预测目标设定为未来第 5 个交易日的收盘价。这一设定主要基于以下两点考虑: 首先, 从实际应用角度看, 多数机构投资者和个人投资者的调仓周期并非以日为单位, 而是以周或更长时间为单位。预测未来一周(5 个交易日)后的价格, 对于周度资产配置、风险对冲以及量化策略的制定具有更直接的参考价值。其次, 从学术研究角度看, 相比于预测下一交易日的价格(即一步预测), 预测未来第 5 日的价格(即多步预测)对模型捕捉长程时序依赖的能力提出了更高要求。一个能够在此任务上表现优异的模型, 意味着其提取的特征更具鲁棒性, 能够更好地剥离短期市场噪音, 把握中长期趋势, 从而更具研究价值和挑战性。

针对以上挑战, 本文提出一种双通道输入的 CNN-BiLSTM 混合预测模型。该模型的核心创新在于输入端的结构性设计: 高频通道处理每 5 分钟的 K 线数据(开盘价、最高价、最低价、收盘价、成交量), 使单个样本的数据量达到传统日线模型的 48 倍, 为捕捉日内微观波动提供了丰富信息; 低频通道同时输入寻常模型未曾利用到的日级数据(日均价、日成交量、调整后收盘价), 确保模型能够把握长期趋势。通过分通道特征提取与门控融合, 实现多尺度信息的有效互补。以贵州茅台(600519.SH) 2020 年 1 月至 2025 年 12 月的交易数据为实证对象, 采用 50 日滑动窗口预测未来 5 日收盘价。结果表明, 本文模型在 RMSE、MAE、MAPE 等多项评价指标上均优于对比模型, 验证了其有效性和优越性。

2. 研究方法

针对传统混合神经网络模型在处理多源时序数据时输入架构单一、难以有效融合不同频率特征的问题, 提出一种双通道输入的 CNN-BiLSTM 股价预测模型。

该模型的核心思想是设计两个独立的数据输入通道并对齐好数据日期, 在传统的仅利用每日的开盘价、最高价、最低价、收盘价、成交量来进行股价预测的基础上, 升级为利用股价分析中使用范围最广的 5 分钟的 k 线数据, 并增加单日级的日均价、日成交量以及调整后的收盘价数据, 两种数据具备的不同特征在各自的数据通道中得到充分的提取, 再通过门控加权融合实现多尺度信息的互补, 最后进行最终的股价预测工作。模型总体架构如图 1 所示。

值得注意的是, 尽管模型的预测目标为第 5 个交易日的收盘价, 但在滑动窗口构建样本的过程中, 模型实际上学习了从过去 50 天数据到未来 5 天内每一天的映射关系。将研究焦点集中于第 5 天, 旨在最大化评估模型对中长期依赖关系的建模能力, 其性能表现可作为衡量模型有效性的关键指标。

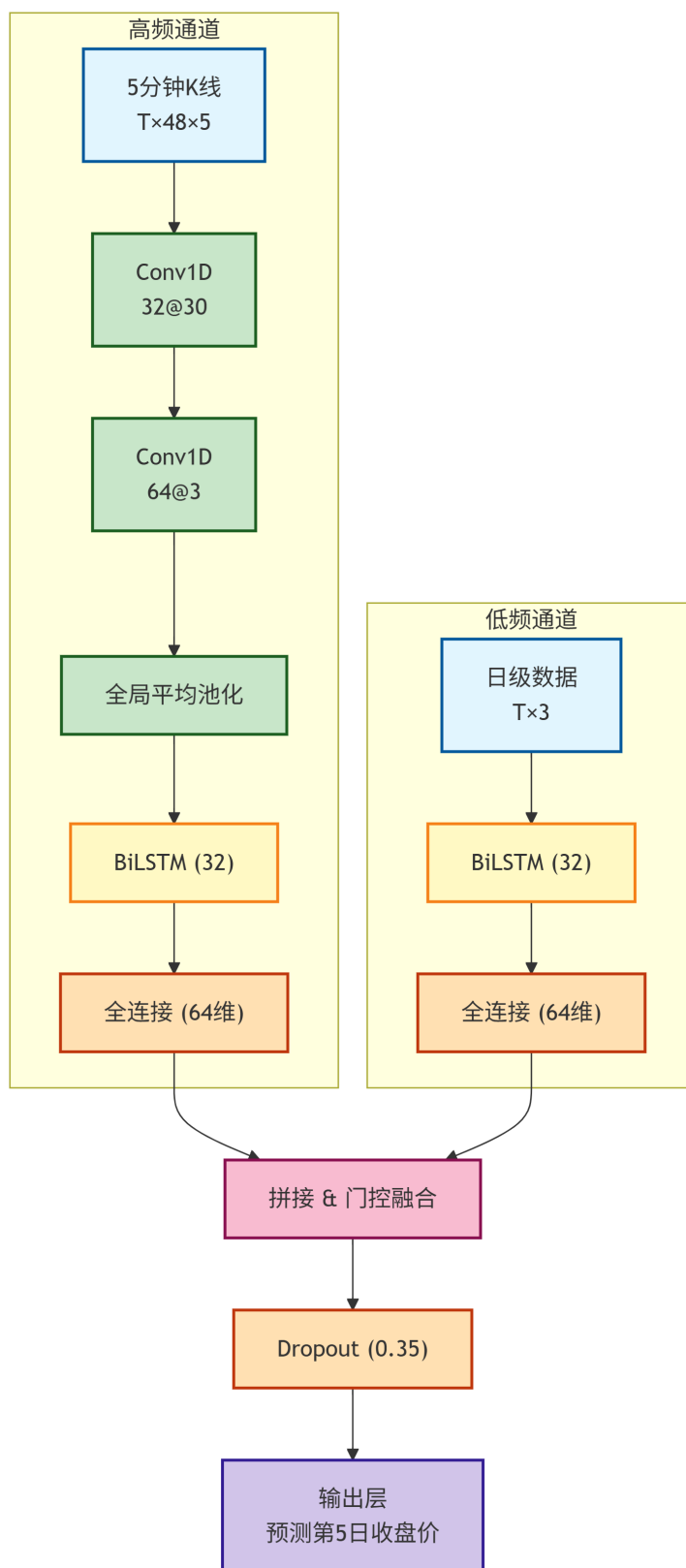


Figure 1. Dual-channel CNN-BiLSTM model structure diagram
图 1. 双通道 CNN-BiLSTM 模型结构图

模型主要包括四个模块:

5 分钟级高频通道: 输入 5 分钟 K 线数据(开盘价、最高价、最低价、收盘价、成交量), 通过卷积神经网络(CNN)提取日内局部波动特征, 再经双向长短期记忆网络(BiLSTM)捕捉跨日时序依赖。

每日级低频通道: 输入日级数据(日均价、日成交量、调整后收盘价), 直接通过 BiLSTM 捕捉长期趋势特征。

门控融合层: 将两个通道输出的特征向量映射到相同维度后, 通过可学习的门控机制动态融合, 实现特征互补。

输出层: 利用过去 50 个交易日的数据预测未来第 5 个交易日的收盘价。

2.1. 五分钟 K 线数据高频通道设计

五分钟的 K 线高频数据记录了每个交易日内部的微观价格变动, 蕴含丰富的市场情绪和短期交易模型。假设样本数量为 N , 每个样本包含连续 T 个交易日的数据(本文取 $T = 50$), 在此前提下, 每个交易日有 L 个 5 分钟 K 线记录(A 股市场交易时间为 4 小时, $L = 48$), 每个 K 线记录有 F_m 个特征(本文取开盘价、最高价、最低价、收盘价、成交量, 即 $F_m = 5$)。则分钟级输入数据可表示为四位数组即张量: $X_m \in \mathbb{R}^{N \times T \times L \times F_m}$ 。

为有效提取日内局部模式, 本文对每个交易日内部的分钟序列独立应用一维卷积网络(Conv1D), 并通过 TD (*TimeDistributed*) 层实现时间步的参数共享。采用两层卷积结构以逐层抽象特征, 第一层卷积设置了交大的卷积核(kernel_size = 30), 其物理意义在于覆盖连续的 150 分钟交易数据。在 A 股交易时间仅为 4 小时的背景下, 该设计旨在捕捉早盘趋势和午盘反转等持续时间长达半个交易日的中观市场形态, 而非聚焦于分钟级的微观噪声。第二层卷积核则采用较小的卷积核(kernel_size = 3), 在经过第一层池化后的缩短时间序列上, 用于精细组合这些中观特征, 提取更具判断力的高层抽象特征, 其设计思想也暗合双通道模型的设计思路。

$$H_m^{con1} = TD(Con(filters = 32, size = 5, act = 'relu', padding = 'same'))(X_m) \quad (1)$$

$$H_m^{con2} = TD(Con(filters = 64, size = 3, act = 'relu', padding = 'same'))(H_m^{con1}) \quad (2)$$

第一层卷积层输出形状为 $(N, T, L, 32)$, 第二层卷积层输出形状为 $(N, T, L, 64)$ 。随后, 采用全局平均池化(Global Average Pooling)沿时间步维度对每个交易日的卷积特征进行聚合, 得到每个交易日的日内特征向量:

$$H_m^{pool} = TD(GlobalAveragePooling1D())(H_m^{con2}) \quad (3)$$

此时输出形状为 $(N, T, 64)$, 即每个交易日用一个 64 维向量表示其日内波动模式。

为捕捉交易日之间的时序依赖, 将上述日内特征序列输入双向 LSTM (BiLSTM)。BiLSTM 由前向 LSTM 和后向 LSTM 组成, 能够同时利用过去和未来的上下文信息, 更全面地建模时序关系:

$$\vec{h}_t = \overline{\text{LSTM}}(h_t^{pool}, \vec{h}_{t-1}), \vec{h}_t = \overline{\text{LSTM}}(h_t^{pool}, \vec{h}_{t+1}) \quad (4)$$

$$h_t^{bi} = [\vec{h}_t; \vec{h}_t] \quad (5)$$

最终取最后一个时间步的隐藏状态作为分钟级通道的输出特征向量 $o_m \in \mathbb{R}^{64}$ (本文取 LSTM 单元数为 32, 双向后维度为 64)。随后通过一个全连接层将特征映射到与日级通道相同的投影维度, 准备融合:

$$o'_m = Dense(64, act = 'relu')(o_m) \quad (6)$$

o'_m 表示从原变量派生出来的新变量。

2.2. 日级数据低频通道设计

每日的日均价等低频数据反映股票的长期趋势和宏观周期变化。在数据对齐的前提下, 设每个样本包含 T 个交易日的日级数据, 每个交易日有 F_d 个特征(本文取日均价、日成交量、调整后收盘价, $F_d = 3$)。则日级输入数据可表示为矩阵 $X_d \in \mathbb{R}^{N \times T \times F_d}$ 。

由于日级数据本身已具有较低的时间频率且趋势性较强, 本文直接采用 BiLSTM 对其进行建模, 以捕捉长期依赖关系, 其公式表示如下:

$$\bar{s}_i = \overline{\text{LSTM}}(x_i^d, \bar{s}_{i-1}), \bar{s}_i = \overline{\text{LSTM}}(x_i^d, \bar{s}_{i+1}) \quad (7)$$

$$s_i^{bi} = [\bar{s}_i; \bar{s}_i] \quad (8)$$

同样取最后一个时间步的隐藏状态作为日级通道的输出特征向量 $o_d \in \mathbb{R}^{64}$, 并通过全连接层映射到与分钟级通道相同的投影维度:

$$o'_d = \text{Dense}(64, \text{act} = 'relu')(o_d) \quad (9)$$

2.3. 特征融合与输出层

两个通道输出的特征向量 $o_m \in \mathbb{R}^{64}$ 和 $o_d \in \mathbb{R}^{64}$ 分别编码了高频局部波动和低频长期趋势信息。为使模型能够有效结合两种尺度的特征, 本文采用门控融合方式, 让网络自适应学习两个通道的权重, 具体操作如下, 先将两个向量拼接, 再通过一个 sigmoid 层生成门控权重:

$$g = \sigma(\text{Dense}(64)([o'_m; o'_d])) \quad (10)$$

$$f = o'_m \odot g + o'_d \odot (1 - g) \quad (11)$$

其中 σ 为 sigmoid 激活函数, \odot 表示逐元素相乘。融合后的特征向量 $f \in \mathbb{R}^{64}$ 再通过一个全连接层进行非线性变换, 并施加 Dropout 防止过拟合:

$$g = \text{Dropout}(0.35)(\text{Dense}(64, \text{act} = 'relu')(f)) \quad (12)$$

输出层采用线性激活函数, 将融合后的特征映射为未来第五个交易日的收盘价预测值:

$$\hat{y} = \text{Dense}(1, \text{act} = 'linear')(z) \quad (13)$$

模型训练采用均方误差(Mean Squared Error, MSE)作为损失函数:

$$\Gamma = \frac{1}{N} = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (14)$$

其中 y_i 为真实收盘价格, \hat{y}_i 为预测值。

2.4. 模型训练与参数设置

模型使用 Adam 优化器进行训练, 初始学习率设为 0.001。训练轮数(epochs)设为 100, 批次大小(batch size)为 64。为防止过拟合, 引入早停机制(early stopping): 当验证集损失连续 10 轮不再下降时停止训练, 并恢复最优权重。所有实验均在 Python 3.9 环境中实现, 深度学习框架采用 TensorFlow 2.10。

模型的主要参数经过参考相关文献与实验的网格搜索而确定。

本文模型的主要超参数设置如表 1 所示:

Table 1. Main hyperparameters of the model
表 1. 模型主要超参数

参数名称	参数值	说明
窗口长度 T	50 天	每个样本包含的历史交易日数量
第一层卷积核数量	32	第一层 Conv1D 提取的日内特征数量
第二层卷积核大小	64	第二层 Conv1D 提取的日内特征数量
LSTM 单元数	32	每个单向 LSTM 的隐藏层单元数
投影维度	64	融合前两个通道特征的映射维度
Dropout 率	0.35	全连接层的随机失活比例
初始学习率	0.001	Adam 优化器的初始学习率
批次大小	64	每次迭代训练的样本数量
最大训练轮数	100	允许的最大迭代轮数
早停耐心值	10	验证损失连续 10 轮不下降即停止

窗口长度 T 50 天固定值。基于金融领域的经验法则, 认为约两个月的交易数据足以涵盖近期趋势。第一层卷积核数量 32 网格搜索{32, 64, 128}, 以验证集 MAPE 最低为准。32 维已足够捕捉主要的日内模式。

第一层卷积核大小 30 基于领域知识设计, 覆盖 2.5 小时交易数据, 旨在捕捉“半日”级别的中观市场形态。第二层卷积核数量 64 网格搜索{32, 64, 128}, 以验证集 MAPE 最低为准。增加维度以组合第一层提取的特征。第二层卷积核大小 3 固定为小尺寸, 用于在时间维度上精细组合第一层输出的高阶特征。LSTM 单元数 32 网格搜索{16, 32, 64}, 性能和效率的平衡点。投影维度 64 网格搜索{32, 64, 128}, 确保两个通道的特征在融合前有足够的表达能力。Dropout 率 0.35 网格搜索{0.2, 0.35, 0.5}, 综合照顾全部模型后确定。

2.5. 对比模型设置

为验证双通道输入结构和残差融合的有效性, 本文设置以下基准模型进行对比:

(1) LSTM 模型: 仅使用日级数据作为输入, 采用单层 LSTM (32 单元)直接预测下一交易日收盘价。输入形状为 $(N, T, 3)$, 其中 $T = 50$ 。

(2) CNN-LSTM 模型: 仅使用分钟级数据作为输入, 首先对每个交易日的 5 分钟 K 线序列应用 Conv1D (与本文分钟级通道相同)提取日内特征, 再输入 LSTM (32 单元)进行预测。输入形状为 (N, T, L, F_m) 。

(3) 单通道 CNN-BiLSTM 模型: 仅使用分钟级数据作为输入, 分钟级通道处理方式与本文分钟级通道完全相同(两层 CNN + BiLSTM), 但无日级通道和门控融合, 直接由 BiLSTM 输出预测。

(4) RNN 模型: 仅使用日级数据作为输入, 采用单层简单循环神经网络(RNN, 32 单元)直接预测下一交易日收盘价。输入形状为 $(N, T, 3)$ 。

(5) GRU 模型: 仅使用日级数据作为输入, 采用单层门控循环单元(GRU, 50 单元)直接预测下一交易日收盘价。输入形状为 $(N, T, 3)$ 。

(6) LSTM-Transformer 模型: 结合 LSTM 与 Transformer 的混合模型。首先使用 LSTM 层(32 单元, 返回序列)对日级数据 $(N, T, 3)$ 进行编码, 提取时序特征; 然后将 LSTM 的输出序列输入 Transformer 编码器层(包含 4 头注意力机制, 前馈网络维度 64), 最后通过全局平均池化和全连接层输出预测值。该模

型用于对比 Transformer 结构在时序建模中的表现。

所有对比模型均使用与本文模型相同的训练数据(贵州茅台 2020 年 1 月至 2025 年 12 月数据)、窗口长度(50 天)、预测步长(5 天)。训练集/测试集划分(前 80% 训练, 后 20% 测试)以及训练策略(早停、批次大小等), 以确保对比的公平性。最后利用消融实验来证明残差融合的有效性。

3. 实证分析

3.1. 数据来源与预处理

本文以贵州茅台(股票代码: 600519.SH)为研究对象, 数据来源于 Baostock 金融数据库, 选取 2020 年 1 月 1 日至 2025 年 12 月 31 日共计 6 年的交易数据。数据包含两个部分:

5 分钟 k 线数据: 每个交易日包含 48 根 5 分钟 K 线, 每根 K 线记录开盘价、最高价、最低价、收盘价和成交量, 共获取约 72,000 条数据记录(具体以实际实验获取为准)。

单日 k 线数据: 每个交易日包含开盘价、最高价、最低价、收盘价、成交量, 共获取约 1500 条的日线数据记录。

数据预处理步骤如下:

(1) **数据对齐:** 将 5 分钟数据按交易日分组, 确保每个交易日有完整的 48 根 K 线; 对于因停牌或数据缺失导致不足 48 根的交易日, 采用该交易日最后一条有效数据向前填充至 48 根。

(2) **特征构造:** 直接采用直接采用 5 分钟 k 线的开盘价、最高价、最低价、收盘价、成交量, 共 5 维特征来构造分钟级特征。并计算日均价的同时选取日成交量和复权收盘价作为日级特征, 共三维特征。

(3) **样本构建:** 采用滑动窗口方式生成样本, 窗口长度设为 $T = 50$ 天。对于第 t 个交易日, 将前 $t - 50$ 至 $t - 1$ 共 50 天的分钟级数据和日级数据作为输入, 第 5 天的收盘价作为预测目标。共生成有效样本 $N - 50 - 1$ 个, 其中 N 为有效交易日总数。

(4) **数据集划分:** 采用传统的 8 比 2 模式划分训练集和测试集。

(5) **数据标准化:** 对分钟级特征和日级特征分别进行 Z-score 标准化, 标准化参数仅基于训练集计算, 并应用于测试集。最后进行反标准化输出计算评估模型性能。

3.2. 评价指标

为全面评估模型性能, 采用四个回归评价指标, 均方根误差(RMSE)、平均绝对误差(MAE)、平均绝对百分比误差(MAPE)、决定系数(R^2)并计算预测准确率(PA), 定义为 $1 - \text{MAPE}$, 以百分比形式直观反映预测精度, 数值越高表示预测越准确。具体公式如下:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (15)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (16)$$

$$\text{MPAE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (17)$$

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

$$\text{PA} = (1 - \text{MAPE}) \times 100\% \quad (19)$$

3.3. 实验设置

所有实验均在一下环境中进行:

硬件: Intel Core i7-12700K CPU, 32GB RAM, NVIDIA RTX 3080 GPU

软件: Python 3.9, TensorFlow 2.10, Scikit-learn 1.2, pycharm2025.1

模型训练采用 Adam 优化器, 初始学习率 0.001, 批次大小 64, 最大训练轮数 100。使用早停机制, 当验证集损失连续 10 轮不下降时停止训练, 并恢复最佳权重。所有对比模型均采用相同的训练策略。对于对比模型中的自有超参数(如 LSTM/GRU/RNN 单元数、Transformer 头数等), 我们也遵循了与本文模型类似的基于验证集的网格搜索原则, 以尽可能确保对比的公平性。例如, LSTM 单元数同样在 {16, 32, 64} 中择优选取。最终报告中, 对比模型均采用其在验证集上性能最佳的超参数组合。

3.4. 实验结果与分析

3.4.1. 模型整体性能对比

为验证本文提出双通道 CNN-BiLSTM 模型的有效性, 将其与 6 个基准模型在测试集上进行对比, 实际下载到的日线数据为 1455 条, 5 分钟数据 69,840 条, 茅台股价(截止日期为 2025 年 12 月 31 日) 1389.72 元。最终结果如表 2 所示。

Table 2. Performance comparison of different models on the Kweichow Moutai stock data test set
表 2. 不同模型在贵州茅台股票数据测试集上的性能对比

模型	RMSE (元)	MAE (元)	MAPE (%)	R ²	PA (%)
RNN	49.27	39.12	2.63	0.307	97.37
GRU	40.32	31.30	2.10	0.526	97.90
LSTM	47.96	36.08	2.41	0.344	97.59
CNN-LSTM	51.56	42.39	2.87	0.241	97.13
单通道 CNN-BiLSTM	41.73	32.10	2.14	0.503	97.86
LSTM-Transformer	61.74	49.53	3.32	0.088	96.68
本文模型	38.51	28.97	1.93	0.577	98.07

从表 2 可以看出, 本文模型在所有指标上均表现最优, MAPE 仅为 1.93%, 预测准确率达到 98.07%, R² 为 0.577, 显著优于其他对比模型(因为茅台股价已经是破千元级别, 所以 R² 值在数据量接近 7 万条的情况下难以接近 1)。与排名第二的 GRU 相比, 本文模型的 MAPE 降低了 0.17 个百分点; 与仅使用分钟级数据的单通道 CNN-BiLSTM 相比, MAPE 降低了 0.21 个百分点。指的注意的是, 前面的单一模型都是使用的日线数据仅 1455 条数据来进行预测工作, 而股价分析经常使用的是 5 分钟级的 K 线数据, 而组合模型在数据量是前者是 48 倍左右的情况下, 预测准确率跟单一模型在数据量更少的情况下(也就是数据波动越小)仍能很接近, 也表明组合模型在股价预测工作方面的可靠性。在这之上, 本文模型以微弱的优势领先与其他模型, 这充分验证了双通道输入结构和门控融合机制在融合多尺度时序信息方面的有效性。

3.4.2. 预测效果可视化

为直观展示本文模型的预测效果, 绘制测试集上真实股价与预测股价的对比曲线(图 2), 以及预测误差分布直方图(图 3)。

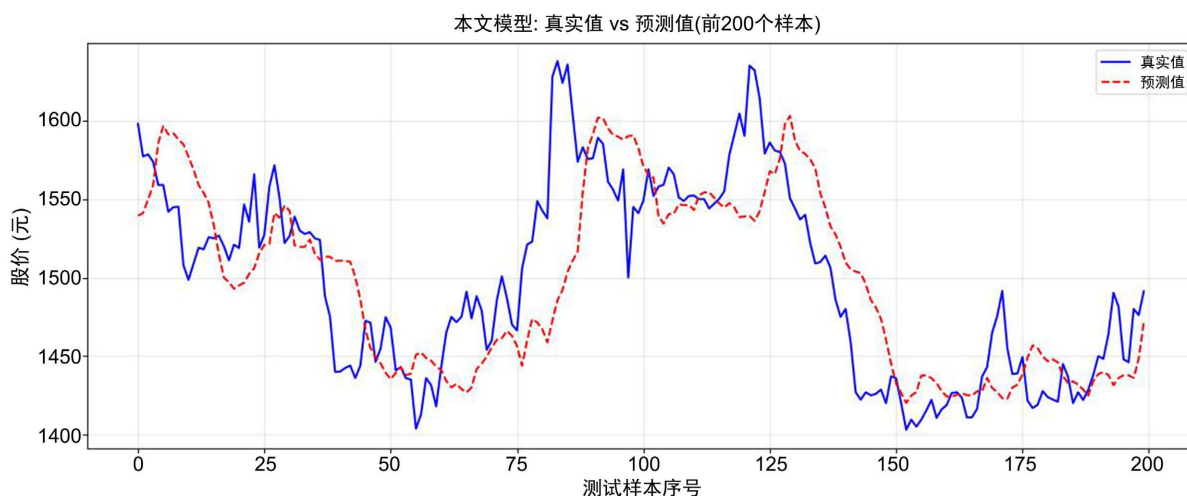


Figure 2. Comparison of true and predicted values on the test set (first 200 samples)

图 2. 测试集真实值与预测值对比(前 200 个样本)

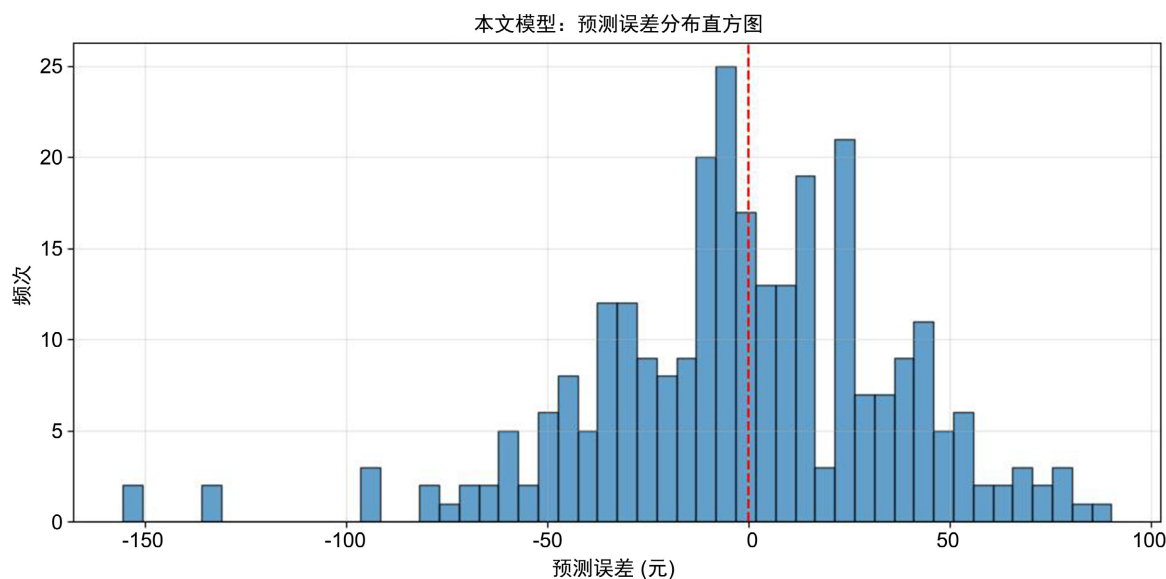


Figure 3. Histogram of prediction error distribution

图 3. 预测误差分布直方图

从图 2 可以看出, 本文模型的预测曲线与真实曲线高度重合, 即使在股价波动较大的区域也能较好地跟随趋势, 表明模型具有较强的拟合能力。图 3 显示预测误差基本服从均值为零的正态分布, 绝大多数误差集中在 ± 50 元以内, 说明模型无系统性偏差, 预测稳定性良好。

4. 结束语

股市因其高风险高收益特性, 股价预测一直是投资者与研究者共同关注的热点。本文提出一种基于双通道输入的 CNN-BiLSTM 股价预测模型, 通过 5 分钟级高频数据与单日级低频数据的双通道结构, 结合门控融合机制整合多尺度信息, 以贵州茅台(600519.SH) 2020 年 1 月至 2025 年 12 月交易数据为实证对象, 获取了超 7 万条数据, 最后采用 50 日窗口预测未来 5 日收盘价。实验结果表明, 本文模型在测试集上 MAPE 为 1.93%、预测准确率达 98.07%, 显著优于 RNN、GRU、LSTM、CNN-LSTM、单通

道 CNN-BiLSTM 及 LSTM-Transformer 等六种基准模型, 验证了其在捕捉股价在非平稳性与非线性特征方面的有效性。

本文仍存在一定局限性: 模型的关键超参数(如卷积核大小、数量等)虽结合了领域知识与网格搜索进行优化, 但仍在一定程度上依赖人工经验, 且最优参数组合可能存在数据集特异的过拟合风险, 模型的泛化能力有待在多只股票、多种市场环境下进一步验证。后续研究将在模型输入和工作端引入更多的特征输入, 增强预测结果的准确性, 引入对抗机制注意力模块等, 提高模型的鲁棒性, 扩大实证范围, 并融合更多外部数据以进一步提升模型性能。

致 谢

十分感谢李前程同学在代码编写时提供的指导与帮助。

参考文献

- [1] 赵迪, 蔡晓铭. 股市风云三十年[M]. 北京: 首都经济贸易大学出版社, 2021.
- [2] Baruník, J. and Křehlík, T. (2016) Combining High Frequency Data with Non-Linear Models for Forecasting Energy Market Volatility. *Expert Systems with Applications*, **55**, 222-242. <https://doi.org/10.1016/j.eswa.2016.02.008>
- [3] Zeng, Z., Kaur, R., Siddagangappa, S., et al. (2023) Financial Time Series Forecasting Using CNN and Transformer. arXiv: 2304.04912.
- [4] Althelaya, K.A., El-Alfy, E.M. and Mohammed, S. (2018) Evaluation of Bidirectional LSTM for Short- and Long-Term Stock Market Prediction. 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, 3-5 April 2018, 151-156. <https://doi.org/10.1109/iacs.2018.8355458>
- [5] Ghosh, A., Bose, S., Maji, G., et al. (2019) Stock Price Prediction Using LSTM on Indian Share Market. *EPiC Series in Computing*, **63**, 101-110.
- [6] 黄星源. 基于改进 LSTM 算法的股票价格趋势预测研究[D]: [硕士学位论文]. 西安: 西安工业大学, 2024.
- [7] Wang, C., Chen, Y., Zhang, S. and Zhang, Q. (2022) Stock Market Index Prediction Using Deep Transformer Model. *Expert Systems with Applications*, **208**, Article ID: 118128. <https://doi.org/10.1016/j.eswa.2022.118128>
- [8] Muhammad, T., Aftab, A.B., Ibrahim, M., Ahsan, M.M., Muhu, M.M., Khan, S.I., et al. (2023) Transformer-Based Deep Learning Model for Stock Price Prediction: A Case Study on Bangladesh Stock Market. *International Journal of Computational Intelligence and Applications*, **22**, Article ID: 2350013. <https://doi.org/10.1142/s146902682350013x>
- [9] Lu, W., Li, J., Li, Y., Sun, A. and Wang, J. (2020) A CNN-LSTM-Based Model to Forecast Stock Prices. *Complexity*, **2020**, Article ID: 6622927. <https://doi.org/10.1155/2020/6622927>
- [10] Nayyeri, N. and Mosavi, J. (2020) Deep Learning for Stock Market Prediction. *Entropy*, **22**, Article 840.
- [11] Kamijo, K. and Tanigawa, T. (1990). Stock Price Pattern Recognition—A Recurrent Neural Network Approach. 1990 IJCNN International Joint Conference on Neural Networks, San Diego, 17-21 June 1990, 215-221. <https://doi.org/10.1109/ijcnn.1990.137572>
- [12] Moghar, A. and Hamiche, M. (2020) Stock Market Prediction Using LSTM Recurrent Neural Network. *Procedia Computer Science*, **170**, 1168-1173. <https://doi.org/10.1016/j.procs.2020.03.049>
- [13] Di Persio, L. and Honchar, O. (2017) Recurrent Neural Networks Approach to the Financial Forecast of Google Assets. *International Journal of Mathematics and Computers in Simulation*, **11**, 7-13.
- [14] Kraus, M. and Feuerriegel, S. (2017) Decision Support from Financial Disclosures with Deep Neural Networks and Transfer Learning. *Decision Support Systems*, **104**, 38-48. <https://doi.org/10.1016/j.dss.2017.10.001>
- [15] Peng, Z. and Guo, P. (2022) A Data Organization Method for LSTM and Transformer When Predicting Chinese Banking Stock Prices. *Discrete Dynamics in Nature and Society*, **2022**, Article ID: 7119678. <https://doi.org/10.1155/2022/7119678>
- [16] 高欣. 基于 Attention 机制的 CNN-LSTM 概率预测模型的股指预测[J]. 现代信息科技, 2024, 8(12): 155-159, 163.
- [17] Lu, W., Li, J., Wang, J. and Qin, L. (2021) A CNN-BiLSTM-AM Method for Stock Price Prediction. *Neural Computing and Applications*, **33**, 4741-4753. <https://doi.org/10.1007/s00521-020-05532-z>
- [18] Livieris, I.E., Pintelas, E. and Pintelas, P. (2020) A CNN-LSTM Model for Gold Price Time-Series Forecasting. *Neural Computing and Applications*, **32**, 17351-17360. <https://doi.org/10.1007/s00521-020-04867-x>

- [19] Liu, S., Zhang, C. and Ma, J. (2017) CNN-LSTM Neural Network Model for Quantitative Strategy Analysis in Stock Markets. In: Liu, D., Xie, S., Li, Y., Zhao, D. and El-Alfy, E.S., Eds., *Neural Information Processing*, Springer, 198-206. https://doi.org/10.1007/978-3-319-70096-0_21
- [20] Li, H., Shen, Y. and Zhu, Y. (2018) Stock Price Prediction Using Attention-Based Multi-Input LSTM. *Proceedings of Machine Learning Research*, **95**, 454-469.