

基于 FCM 深度学习模型的证券股价研究

郭泉吕¹, 孙 荣²

(1. 重庆工商大学数学与统计学院, 重庆 400067; 2. 重庆工商大学社会经济应用统计重庆市重点实验室, 重庆 400067)

摘要: 股票市场预测是一个复杂且充满挑战的领域, 序列常表现出高噪声、非线性和非平稳性等特性。为了提高预测的准确性, 提出一种新的方法, 即结合模糊 C-均值(FCM)聚类算法来识别和利用股价预测序列中的局部趋势特征, 分析中综合考虑股票的关键市场数据, 包括开盘价、最高价、最低价、收盘价、成交量和成交额, 作为预测模型的输入特征。通过实验, 比较不同滑动窗口数对模型预测能力的影响的实证分析, 可以发现, 融合了 FCM 聚类和 LSTM-Transformer 组合模型的 FCM-LSTM-Transformer 方法的预测精度比单一深度学习模型和 LSTM-Transformer 组合模型均要高, 评价指标达到最优, 决定系数 R^2 分别提升了 2.75%、2.40%、2.19%。结果表明, 该模型处理股票市场数据的复杂性方面更具明显优势。

关键词: 金融时间序列; 模糊 C-均值聚类; 深度学习模型; FCM-LSTM-Transformer

中图分类号: F832.5 **文献标志码:** A **文章编号:** 1671-1807(2025)12-0044-09

2024 年 9 月, 中国股市迎来了显著的上涨行情, A 股总市值达 98.24 万亿元创历史新高。越来越多的人积极融入金融投资领域, 因其流动性高、投资收益高等特点, 将其视为一种重要的投资渠道。然而, 股票价格的走势受到多种因素的影响, 股票数据也因此表现出非平稳、非线性和高噪声等特点。如何更准确地预测股票价格仍是众多学者研究的热门课题^[1]。

大数据背景下, 数据挖掘技术被广泛应用于股票价格预测领域。目前, 股票价格预测的方法有很多, 其中基于统计学模型, 刘洁^[2]运用自回归移动平均法(ARMA)模型对格力电器共 240 期股票收盘价进行了线性拟合与预测, 结果显示 ARMA 模型短期预测效果较好。曹沥方和张若麟^[3]通过建立自回归求和移动平均法(ARIMA)探究牧原股份股价变动规律和趋势, 研究表明能为短期投资决策提供参考。Kim 和 Sayama^[4]在此基础上进一步改进融合网络测量结果, 实证该模型下可以更好地预测标准普尔 500 指数未来变化趋势。赵建喜等^[5]利用不同次数的指数平滑法(ETS)实现序列数据的去噪、曲线段特征匹配, 提高了预测精度。Sun^[6]在标准普尔 500 指数收盘价预测问题上, 比较了 ARIMA 模型和 ETS 指数平滑法等方法, 结果表明统计学模型下线性拟合度高。He 等^[7]构建了一种新型的基

于季节趋势分解的树突状神经元模型(STLDNM), 对 16 个真实股票数据进行评估预测, 发现该模型下的预测精度明显优于其他统计学模型。统计学模型本质上捕捉序列数据中的线性关系, 无法捕捉非线性关系, 机器学习模型能很好地弥补其处理非线性数据的不足。基于机器学习模型, 张磊^[8]和 Dash 等^[9]分别对支持向量机模型进行了改进和微调, 优化模型参数, 更好地捕捉非线性的股价数据的变化特征, 该方法可以实现更高精度的时序预测。王惠莹和郝泳涛^[10]将股价预测问题视为分类问题, 结合股票的相关经济特征以及股价历史走势, 基于 GS-RF 股价走势预测模型预测股价走势。黄颖和杨会杰^[11]用 XGBoost 分析动量因子特征重要性, 比较 XGBoost 和 LSTM 模型对黄金价格涨跌趋势预测效果, 结果 LSTM 模型预测效果优于 XGBoost。曾海潇^[12]以建立基于变权组合模型, 结合机器学习方法和深度学习方法实现多解释变量预测, 但机器学习算法往往难以捕捉序列数据中更深层次的关系, 比如序列数据中的长期依赖关系。为了解决这一问题, 深度学习模型逐渐应用于该领域, 深度神经网络(DNN)^[13]能更有效地表示复杂的高维函数。此外, 递归神经网络(RNN)、长短期记忆网络(LSTM)^[14]和门控循环单元(GRU)^[15], 可以解决计量经济学方法在处理非线性关系方面的局限性, Transformer 通过自注意力机制,

收稿日期: 2025-01-07

作者简介: 郭泉吕(2001—), 男, 重庆人, 硕士研究生, 研究方向为金融时间序列预测; 孙荣(1973—), 男, 四川达州人, 硕士, 教授, 研究方向为保险精算、应用统计。

能够同时处理序列中的所有位置,捕捉序列数据中长时间距离依赖关系^[16-17]。这对于股票价格预测尤为重要,因为股票价格往往受到多种因素的影响,包括历史价格、交易量及其他市场因素,这些因素之间存在着复杂的相互关系和长期依赖性。

在深度学习领域,股价预测模型面临一系列挑战,这些挑战限制了模型的预测能力和准确性。当前的研究方法往往侧重于利用单变量数据,即基于历史收盘价来预测未来的收盘价。这种方法忽视了股票市场数据的多维特性,如开盘价、最高价、最低价、成交量和成交金额等,这些因素与收盘价之间可能存在复杂的相互作用,它们在预测股价时能够提供额外的、有价值的信息。此外,当尝试将多变量股票信息作为特征输入,以预测单变量输出时,现有模型在捕捉输出序列中的局部波动特征方面存在困难。这种局限性可能导致模型无法充分理解股价的动态变化,从而影响预测的精确度。

本文以中信证券的历史交易数据为例,综合考虑股票开盘价、最高价、最低价、收盘价、成交量及成交金额,建立各种深度学习模型进行股价预测,旨在提高模型的预测精度,减少模型的收敛时间。研究的核心创新在于计算股票预测序列(收盘价)的延迟损益百分比,并运用模糊 C-均值(FCM)聚类算法将时间序列数据聚类,以捕捉并反映预测序列的日损益变化,增强模型对局部序列特征的捕捉能力。此外,通过组合不同预测模型,有效解决了单一模型可能遇到的梯度消失和梯度爆炸问题,从而优化了模型的学习效率和预测性能。

1 研究方法

本文的主要方法包括模糊 C-均值聚类算法和深度学习算法,要理解所提出的方法,首先必须掌握构成模型的组件及其学习机制。其中深度学习算法的架构包括 LSTM、GRU、Transformer、LSTM-Transformer、FCM-LSTM-Transformer。

1.1 模糊 C-均值聚类

模糊 C-均值聚类(fuzzy C-means, FCM)^[18]是一种将模糊数学中模糊集理论融合在聚类分析的算法。它基于相似性度量对数据点进行分组,以捕捉序列数据中潜在趋势。其满足目标函数值等式:

$$\begin{cases} \min J_{\text{FCM}}(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m \| \mathbf{x}_i - \mathbf{v}_j \|^2 \\ m > 1 \\ \sum_{j=1}^c u_{ij} = 1, u_{ij} > 0 \end{cases} \quad (1)$$

式中:令 $d_{ij} = \| \mathbf{x}_i - \mathbf{v}_j \|$, 表示数据点 i 到第 j 个聚类中心的欧式距离; \mathbf{x}_i 为第 i 个特征向量; \mathbf{v}_j 为第 j 个聚类中心; c 为聚类中心个数; \mathbf{U} 为隶属度矩阵; \mathbf{V} 为聚类中心矩阵; u_{ij} 为数据点 i 到第 j 个聚类中心的隶属度; m 为一个隶属度因子,也叫模糊指数; $J_{\text{FCM}}(\mathbf{U}, \mathbf{V})$ 为各类内据点到聚类中心的加权距离平方和,权重是隶属度的 m 次方。

求解约束优化问题,得到 \mathbf{U} 、 \mathbf{V} 的最优迭代解:

$$\begin{cases} u_{ij} = \frac{1}{\sum_{k=1}^c \frac{\| \mathbf{x}_i - \mathbf{v}_j \|^2}{\| \mathbf{x}_i - \mathbf{v}_k \|^2}^{2/m}} \\ \mathbf{v}_j = \frac{\sum_{i=1}^n u_{ij}^m \mathbf{x}_i}{\sum_{i=1}^n u_{ij}^m} \end{cases} \quad (2)$$

式中: k 为第 k 个聚类中心。

式(2)表明第 i 个样本到第 j 个聚类中心的最佳隶属度 u_{ij} 取决于该样本点到第 j 个聚类中心的距离与其到所有聚类中心距离的比值之和。

1.2 长短期记忆网络

递归神经网络(RNN)是传统前馈神经网络的高级演变,考虑时间序列数据的短期相关性更高效地处理序列输入数据,但是针对长时间维度的数据时会出现梯度消失、梯度爆炸问题。长短期记忆(LSTM)神经网络引入门控机制,可以很好地解决上述问题。LSTM 模型存储单元由三部分组成,包括遗忘门、输入门和输出门。模型内部结构如图 1 所示。

LSTM 神经网络将这三个不同的门集成到 RNN 架构中,进而增强了性能,提高了处理长时间序列的能力。

$$\mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \quad (3)$$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) \quad (4)$$

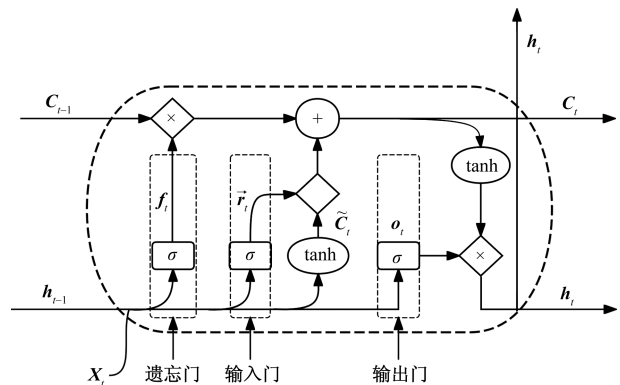


图 1 LSTM 神经元内部结构

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (5)$$

$$C_t = f_i C_{t-1} + i_t \tilde{C}_t \quad (6)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t \tanh C_t \quad (8)$$

式中： σ 为 sigmoid 函数； W_f 、 W_i 、 W_c 、 W_o 分别为遗忘门、输入门、记忆单元、输出门的权重矩阵； b_f 、 b_i 、 b_c 、 b_o 分别为对应的偏置向量； x_t 为当前时刻的输入值； h_{t-1} 为上一时刻的单元输出； \tanh 为激活函数； \tilde{C}_t 是新的候选值向量； C_{t-1} 是历史单元状态； C_t 是当前细胞状态。

1.3 门控循环单元

门控循环单元(CRU)神经网络是对 LSTM 神经网络的简化,舍弃原有的遗忘门 f_t 、输入门 i_t 及输出门 o_t ,引入重置门 r_t 与更新门 z_t 。这样既简化提高了计算效率,同时保留了捕捉时间序列数据中长期依赖关系的能力。模型结构如图 2 所示。

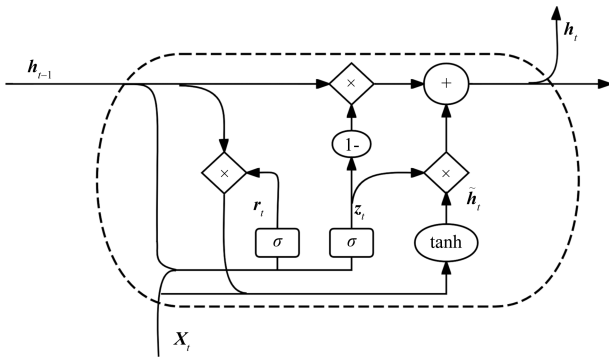


图 2 GRU 神经元内部结构

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \quad (9)$$

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \quad (10)$$

$$\tilde{h}_t = \tanh(W[h_r h_{t-1}, x_t]) \quad (11)$$

$$h_t = (1 - z_t)h_{t-1} + z_t \tilde{h}_t \quad (12)$$

1.4 Transformer 模型

在股价预测时序领域中,transformer 模型基于自注意力机制,捕捉不同时间点股价之间的长距离依赖关系,能够有效避免梯度消失或梯度爆炸的问题。

Transformer 模型的结构主要由编码器-解码器组成。输入的历史序列首先经过词编码和位置编码,然后输入到编码器中,目标序列为解码器的输入。编码器-解码器由多个编码层和解码层堆叠组成,每个编码层和解码层都包含多个注意力和归一化等神经网络层。自注意力机制层允许模型在处理序列时同时考虑所有位置的信息,这有助于捕捉长距离依赖关系和上下文信息。

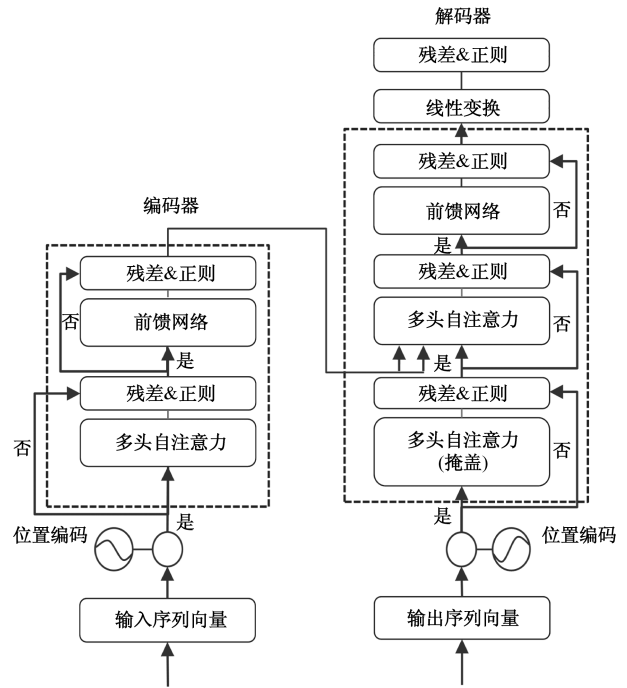


图 3 Transformer 模型框架

1.5 LSTM-Transformer 模型

针对时间序列预测问题中,有三种时间序列预测数据,分别是单变量输入单变量输出(univariate-univariate, S)、多变量输入多变量输出(multivariate-multivariate, M)以及多变量输入单变量输出(multivariate-univariate, MS)。通过融合 LSTM 和 Transformer 两种神经网络模型,可以使得复合模型更具灵活性,同时提高模型预测的鲁棒性,如图 4 所示。

由于本文模型采用股价多变量输入单变量输出的时间序列数据,以股价的相关信息作为特征变量来预测股价变化趋势,通过 LSTM 模型获得输入序列数据的长期依赖关系,结合 Transformer 模型的编码器中位置编码信息,结合作为 Transformer 模型的输入,更好地捕捉序列中的全局依赖关系^[19-20]。由于序列数据的特征,最后用全连接层代替解码器输出序列数据。

1.6 评估指标

选取 MAE、MAPE、MSE、RMSE 及 R^2 评价指标来评价不同模型之下预测拟合效果的优劣。

(1) 平均绝对误差(mean absolute error, MAE)。

$$MAE(y_{true}, y_{predict}) = \frac{1}{N} \sum_{i=1}^N |y_{true} - y_{predict}| \quad (13)$$

(2) 平均绝对百分比误差(mean absolute per-

centage error, MAPE)。

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_{\text{true}} - y_{\text{predict}}}{y_{\text{true}}} \right| \times 100\% \quad (14)$$

(3) 均方误差 (mean square error, MSE)。

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_{\text{true}} - y_{\text{predict}})^2 \quad (15)$$

(4) 均方根误差 (root mean square error,

RMSE)。

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{\text{true}} - y_{\text{predict}})^2} \quad (17)$$

(5) 可决系数 (R^2)。

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_{\text{true}} - y_{\text{predict}})^2}{\sum_{i=1}^N (y_{\text{true}} - \bar{y})^2} \quad (17)$$

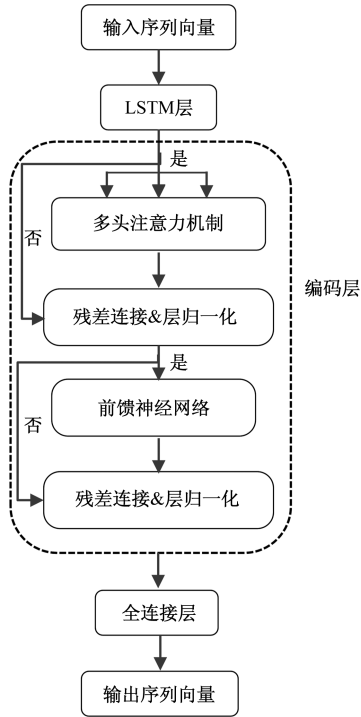


图 4 LSTM-Transformer 组合模型

2 实证分析

实证分析基于 Python 3.9 和 Pytorch 1.10.1 进行神经网络模型的搭建。以中国上市证券公司的中心证券作为研究对象,实现多变量特征输入,单变量特征输出。选择 SVR、LSTM、GRU、Transformer、LSTM-Transformer、FCM-LSTM-Transformer 等深度学习模型进行时间序列预测比较分析,并比较不同滑动窗口数、模型组合下的模型评价指标,探究最优的股价序列预测模型。

2.1 组织架构

FCM-LSTM-Transformer 流程如图 5 所示。

2.2 数据集

以中国上市证券公司的中心证券为研究对象,数据时间跨度从 2021 年 1 月 4 日至 2024 年 6 月 28 日,共计涵盖 844 个交易日的历史数据,选取开盘价、收盘价、最高价、最低价、成交量和成交额等 6 个关键指标作为多变量输入,并以证券的收盘价作为

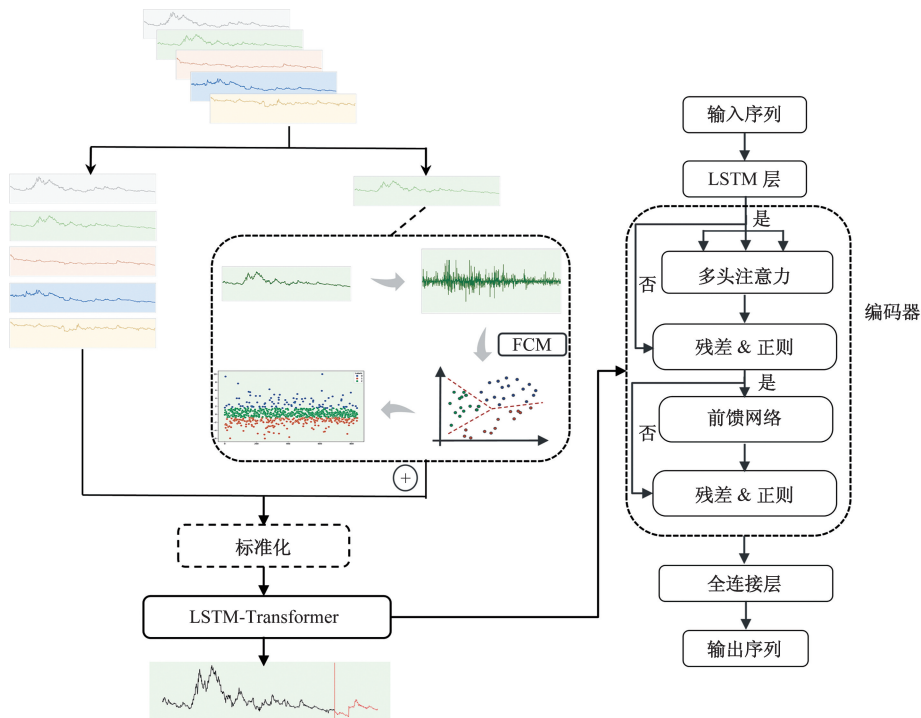


图 5 FCM-LSTM-Transformer 流程

单变量输出。在数据预处理阶段,首先对历史交易数据中的缺失值进行检查,并采用前后加权线性平均的方法对缺失交易日的数据进行填充。并对填充好的数据进行归一化处理,通过式(18)进行数据缩放,以确保最终输出在 0~1,同时保留时序数据的特征。

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (18)$$

通过以上的数据预处理后,为了实现多变量时间序列数据输入、单变量输出,采用滑动窗口方法对数据进行重构。具体而言,通过以固定步长依次滑动窗口,提取窗口内的数据作为模型的输入数据。这种滑动窗口的数据处理方式不仅能够有效捕捉时间序列的局部特征,还能保持数据之间的连续性。并将 2021 年 1 月 4 日至 2023 年末实验数据集以 8:2 划分为训练集和验证集,用于训练模型以及模型参数的调整,2024 年交易日数据作为测试集评估模型的性能。

2.3 模型参数

统一选择 Adam 优化器和 MSEloss 损失函数来训练所有的模型。Adam 优化器是一种自适应学习率的优化算法,其在训练过程中能够动态调整学习率,从而加快模型的收敛速度。MSEloss 函数用于计算模型预测值与真实值之间的均方误差,以评估模型的性能。

为了探究这些模型的性能和差异,对每个模型进行手动参数优化,经过多次调试下,确定模型的最佳超参数组合。本文选择的 LSTM、GRU、Transformer、LSTM-Transformer、FCM-LSTM-Transformer 5 种不同模型,这些模型训练时在结构和复杂度上存在显著差异。图 6 为绘制的模型损失曲线,可以发现 LSTM、LSTM-Transformer 两个模型在同样的参数设定下,训练损失和验证损失随迭代次数的变化情况。前者损失值迅速下降趋于稳定,后者损失值也逐渐减小且伴随着波动,这表明 LSTM-Transformer 模型经过 100 个训练轮数

(epoch)后仍未见平稳。后三个模型是基于自注意力机制的,模型更为复杂。所以依据模型的复杂度调节学习率 lr、正则化率 dropout 及模型的迭代次数 epoch。模型具体参数如表 1 所示。

2.4 分析结果

根据图 7 的分析,该证券收盘价的前一日损益百分比数据点分布揭示了股价变动的三种趋势,即损益百分比变动大于 1%、-1%~+1% 及小于 -1%。这些趋势能够有效地捕捉到收盘价中的局部序列特征。进一步的分析表明,Label_0 和 Label_1 的数量显著多于 Label_2 的数量,这反映在 2021 年 1 月 4 日至 2024 年 6 月 28 日期间,中信证券股价整体呈下跌趋势。

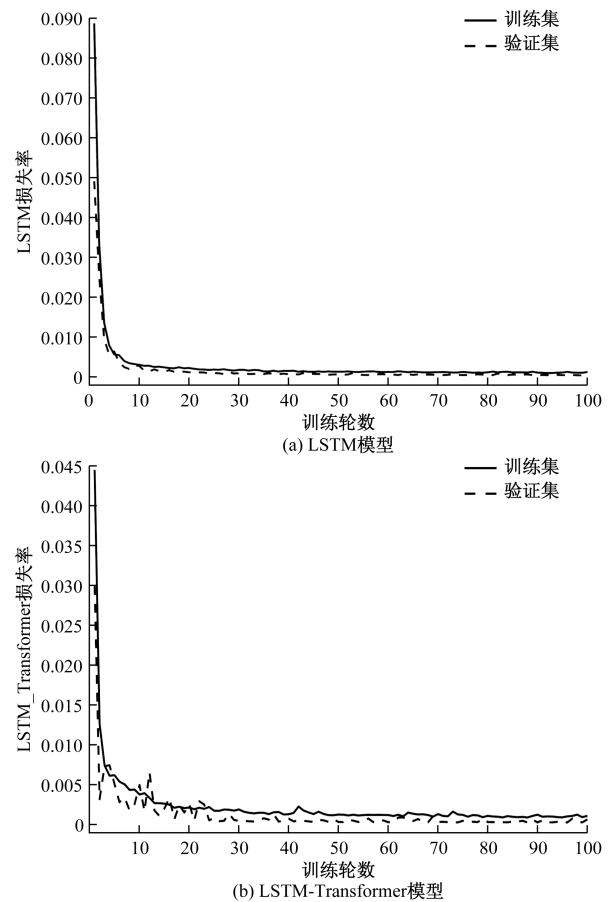


图 6 LSTM 和 LSTM-Transformer 模型损失曲线

表 1 模型参数设定

变量	LSTM	GRU	Transformer	LSTM-Transformer	FCM-LSTM-Transformer
批次大小	32	32	32	32	32
隐藏神经元大小	64	64	64	64	64
输入特征维度	6	6	6	6	7
输出特征维度	1	1	1	1	1
时间步长	16/32/64	16/32/64	16/32/64	16/32/64	16/32/64
dropout 率	0.5	0.5	0.1	0.1	0.5
学习率	0.001	0.001	0.001	0.000 5	0.001
训练轮数	100	100	200	300	300

为了验证不同深度学习模型、不同输入序列长度对股价预测的效果。选取 MAE、MAPE、MSE、RMSE 及 R^2 作为评价指标,这些指标被广泛应用于序列预测模型性能的评估,能够全面反映模型的准确度和可靠性。选择 LSTM、GRU、Transformer、LSTM-Transformer 及 FCM-LSTM-Transformer 模型在不同的输入序列长度下进行比较。

如表 2 所示,随着 time_step 的变化,也即是不同的时间序列输入长度,预测模型的性能指标呈现不同的变化趋势。在 time_step 为 16 时,FCM-LSTM-Transformer 的各个评价指标均优于其他对照组模型,其 MAE 仅为 0.181、RMSE 为 0.259、 R^2 达到 95.51%,在金融时间序列的数据中表现良好。且随着 time_step 的不断增大,大部分模型的预测性能普遍呈现下降趋势,这是因为 time_step 更长,输入模型的数据量越大,涉及更多的不确定因素,模型难以准确捕获。

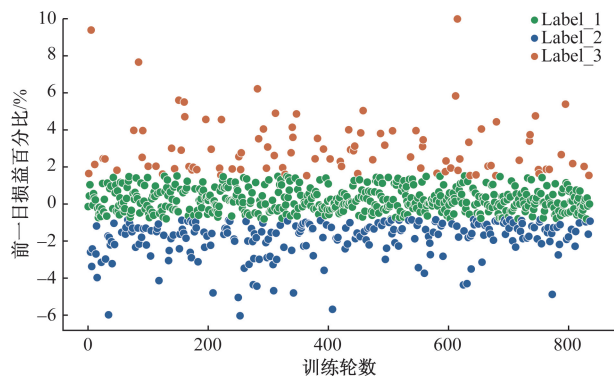


图 7 损益百分比聚类结果

同时,可以发现经过加入目标预测值的前一日损益百分比标签后的 FCM-LSTM-Transformer 模型,可以发现伴随着延迟损益百分比的输入,使得原本的组合模型 LSTM-Transformer 模型可以更加准确地捕捉到预测目标序列(收盘价)的局部和整体的变化趋势。随着 time_step 从 16 增加至 64,FCM-LSTM-Transformer 模型相较于 LSTM-Transformer 模型,MAE 分别减少了 11.27%、17.69%、17.41%,RMSE 减少了 8.8%、13.53%、11.07%, R^2 分别提升了 2.75%、2.4%、2.19%,这表明通过融合序列的局部特征,可以有效提高模型预测的准确度。

图 8 展示了在时间步长(time_step)设定为 16 时,不同深度学习模型与支持向量回归(SVR)机器

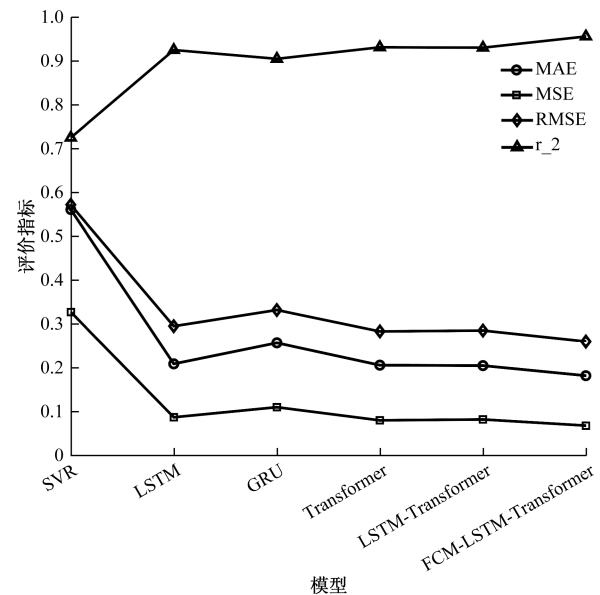


图 8 时间步长为 16 的模型评价指标

表 2 模型评价指标结果

Time_step	模型	MAE	MAPE/%	MSE	RMSE	R^2 /%
16	LSTM	0.208	1.082	0.086	0.294	92.42
	GRU	0.256	1.341	0.109	0.331	90.41
	Transformer	0.205	1.071	0.079	0.282	93.04
	LSTM-Transformer	0.204	1.057	0.081	0.284	92.95
	FCM-LSTM-Transformer	0.181	0.941	0.067	0.259	95.51
32	LSTM	0.246	1.292	0.105	0.324	90.16
	GRU	0.215	1.121	0.093	0.304	91.32
	Transformer	0.225	1.169	0.083	0.289	92.18
	LSTM-Transformer	0.226	1.179	0.092	0.303	91.36
	FCM-LSTM-Transformer	0.186	0.973	0.068	0.262	93.56
64	LSTM	0.197	1.032	0.080	0.283	91.45
	GRU	0.262	1.372	0.127	0.356	86.40
	Transformer	0.227	1.184	0.088	0.297	90.53
	LSTM-Transformer	0.224	1.175	0.089	0.298	90.49
	FCM-LSTM-Transformer	0.185	0.970	0.070	0.265	92.479

学习模型的预测性能对比。通过可视化的 MAE、MSE、RMSE、 R^2 评价指标,可以直观发现深度学习模型对于处理多变量时间序列数据效果要优于机器学习模型,能够捕捉序列间的长期依赖关系,也进一步表明 FCM-LSTM-Transformer 模型在预测精度方面具有显著优势。

从以上的分析结果得知,模型在 time_step 为 16 时,预测性能最佳。图 9 绘制测试集的真实数据以及各个模型预测数据的曲线,从整体趋势上看所有模型都能够准确的捕捉到序列的大致位置的实际趋势变化,原始数据呈现不断波动,预测值与原始数据的波动趋势大体吻合。这表明模型在把握整体波动方面表现良好,基本能够反映股价随时间的变化模式。当出现急剧的转折点,模型预测值的波动性与实际值相比,呈现较强的不匹配趋势。从局部放大图可以更加直观地看出,SVR 模型虽然可以刻画真实曲线整体趋势变化,但是预测精度低,难以实现准确地预测效果,深度学习模型整体比机器学习模型预测精度要高,FCM-LSTM-Transformer 模型对于整体趋势以及转折点的变化趋势上捕捉更为准确。

如图 10 所示,可以清晰地观察到,聚类算法和深度学习相结合的 FCM-LSTM-Transformer 模型,其预测值与实际股价数据之间的拟合度相较于 LSTM-Transformer 模型更好,其预测点紧密地围绕着回归直线分布。表明该模型在股价预测任务上性能优越。

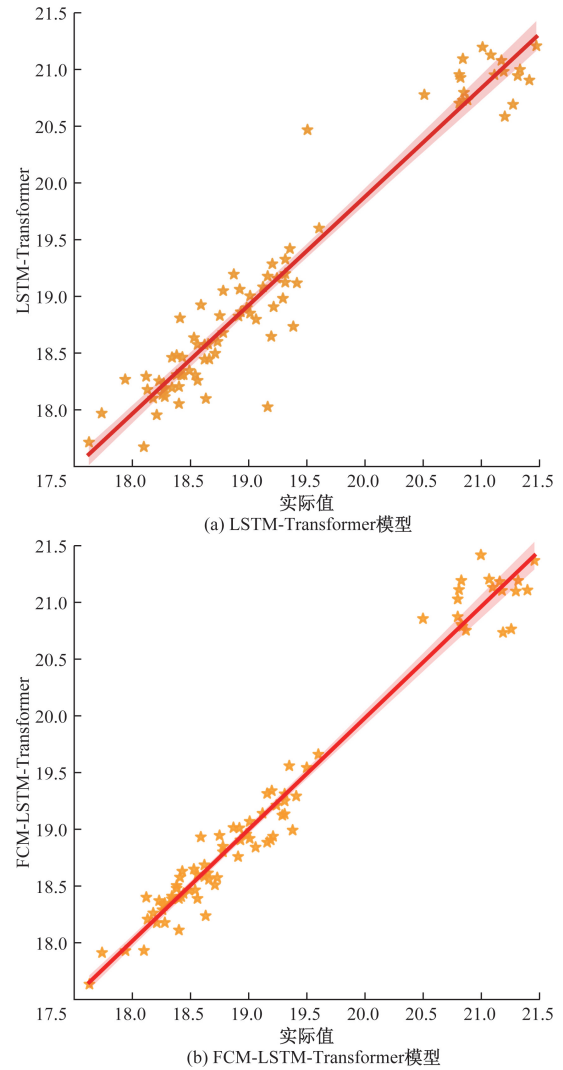


图 10 数据点回归拟合

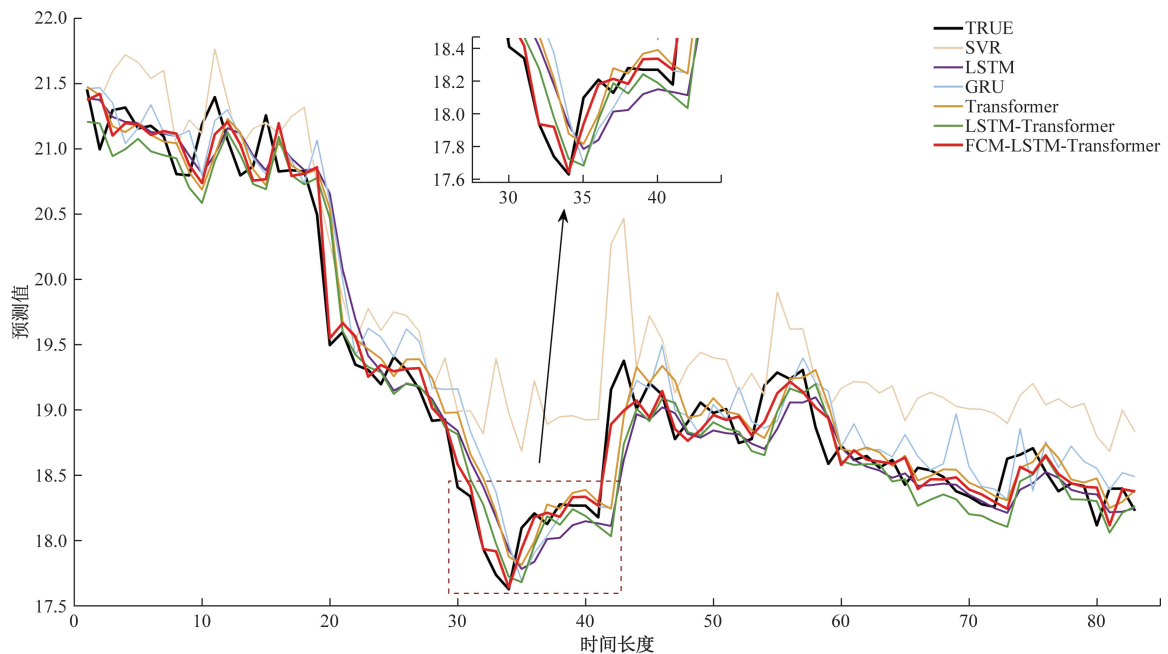


图 9 时间步长为 16 的模型预测值

3 结论

鉴于股票市场的高度流动性、波动性和非平稳性,为了提升股价变动趋势预测的效率与准确性,本研究创新性地提出了一种结合聚类算法与深度学习模型的预测方法——FCM-LSTM-Transformer。该模型以中信证券的股价序列数据为基础,运用模糊 C-均值聚类算法来识别并捕捉股价变动的局部特征,同时将开盘价、最高价、最低价、收盘价、成交量和成交金额等关键历史市场数据作为特征变量,整合输入至深度学习模型中。通过这种方式,模型能够从局部序列特征中洞察股价的波动趋势,有效弥补了单一模型在预测精度上的不足。

然而,股价的变动是一个复杂的现象,其预测需要综合考虑多种因素。除了本文模型所考虑的股价历史数据等内部因素,还应当将企业新闻事件、市场环境等外部因素纳入考量。通过这种多维度的分析,可以更全面地理解股价变动的驱动因素,从而提高预测的准确性,为投资者提供更为科学的决策支持,有效降低投资风险。

参考文献

- [1] GANDHMAL P D, KUMARK. Systematic analysis and review of stock market prediction techniques[J]. Computer Science Review, 2019, 34: 100190.
- [2] 刘洁. ARMA 模型在股票价格预测中的应用——以格力电器为例[J]. 中国管理信息化, 2021, 24(11): 153-155.
- [3] 曹沥方, 张若麟. 基于 ARIMA 模型的股价预测——以牧原股份为例[J]. 财富时代, 2024(2): 54-56.
- [4] KIM M, SAYAMAH. Predicting stock market movements using network science: an information theoretic approach[J]. Applied network science, 2017, 2: 1-14.
- [5] 赵建喜, 李雪飞, 易丹辉, 等. 基于曲线段特征匹配的股价预测研究[J]. 数学的实践与认识, 2018, 48(1): 75-82.
- [6] SUN Z. Comparison of trend forecast using ARIMA and ETS models for S&P500 close price[C]//Proceedings of the 2020 4th International Conference on E-Business and Internet. Singapore: Association for Computing Machinery, 2020: 57-60.
- [7] HE H, GAO S, JIN T, et al. A seasonal-trend decomposition-based dendritic neuron model for financial time series prediction [J]. Applied Soft Computing, 2021, 108: 107488.
- [8] 张磊. 基于支持向量机的股票市场趋势分析及预测研究[D]. 南京: 南京邮电大学, 2020.
- [9] DASH R K, NGUYEN T N, CENGIZ K, et al. Fine-tuned support vector regression model for stock predictions[J]. Neural Computing and Applications, 2023, 35(32): 23295-23309.
- [10] 王惠莹, 郝泳涛. 基于技术指标和随机森林的股价走势预测算法[J]. 现代计算机, 2021, 27(27): 43-47.
- [11] 黄颖, 杨会杰. 基于 XGBoost 和 LSTM 模型的金融时间序列预测[J]. 科技和产业, 2021, 21(8): 158-162.
- [12] 曾海潇. 基于 LightGBM-GRU 的新能源股票价格预测模型[D]. 成都: 西南大学, 2022.
- [13] 左川, 王宇, 李振. 深度学习在股票投资中的应用[J]. 郑州大学学报(哲学社会科学版), 2022, 55(2): 56-62.
- [14] SHI Z, CHEHADE A. A dual-LSTM framework combining change point detection and remaining useful life prediction[J]. Reliability Engineering and System Safety, 2021, 205: 107257.
- [15] XIAO J, DENG T, BI S. Comparative analysis of LSTM, GRU and transformer models for stock price prediction[C]//Proceedings of the International Conference on Digital Economy, Blockchain and Artificial Intelligence. Guangzhou: Association for Computing Machinery, 2024: 103-108.
- [16] 辛洲扬. 基于 LSTM 与 Transformer 模型的股价预测[D]. 济南: 山东大学, 2022.
- [17] WANG C, CHEN Y, ZHANG S, et al. Stock market index prediction using deep Transformer model[J]. Expert Systems with Applications, 2022, 208: 118128.
- [18] BEZDEK J C. Pattern recognition with fuzzy objective function algorithms [M]. New York: Plenum Press, 1981.
- [19] 王娟, 韩怡, 何巍, 等. 基于 LSTM-Transformer 的突发事件网民情感风险预测研究[J]. 科技传播, 2024, 16(19): 126-129.
- [20] CAO K, ZHANG T, HUANG J. Advanced hybrid LSTM-transformer architecture for real-time multi-task prediction in engineering systems[J]. Scientific Reports, 2024, 14(1): 4890.

Research on Stock Prices Based on FCM-Deep Learning Model

GUO Quanlü¹, SUN Rong²

(1. School of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China; 2. Chongqing Key Laboratory of Socio Economic Application Statistics, Chongqing Technology and Business University, Chongqing 400067, China)

Abstract: Forecasting the stock market is a difficult and intricate task, as price series often display traits like significant noise, nonlinearity and non-stationarity. In order to improve the accuracy of predictions, a new method that combined the fuzzy *C*-means (FCM) clustering algorithm to identify and utilize local trend features in stock price prediction sequences was proposed. In the analysis, key market data of stocks, including opening price, highest price, lowest price, closing price, trading volume, and trading amount, was comprehensively considered as input features for the prediction model. Through experiments, an empirical analysis was conducted to compare the impact of different sliding window sizes (16, 32, 64) on the model's predictive capability. It is found that the FCM-LSTM-Transformer method, which integrates FCM clustering with the LSTM-Transformer combination model, achieves higher prediction accuracy than both the standalone deep learning models and the LSTM-Transformer combination model. The evaluation metrics MAE, MAPE, MSE and RMSE reach their minimum errors, and the coefficient of determination R^2 improved by 2.75%, 2.4% and 2.19%, respectively. These results indicate that the proposed model has a significant advantage in handling the complexity of stock market data.

Keywords: financial time series; fuzzy *C*-means clustering; deep learning model; FCM-LSTM-transformer