

基于 ABC-LSTM-GRU 的时间序列分解与预测模型

朱子敬, 何利文

(南京邮电大学物联网学院, 江苏 南京 210003)

✉ zhon0713@qq.com; helw@njupt.edu.cn



摘要:针对金融时间序列数据的高噪声、时间依赖性等问题,提出了一种人工蜂群算法-长短期记忆-门控单元(ABC-LSTM-GRU)混合模型。该模型综合利用长短期记忆网络(LSTM)和门控循环单元(GRU)循环神经网络,更全面地捕捉时间序列中的长期和短期关系。在特征处理阶段,通过相关性分析对特征进行筛选,同时采用奇异谱分析(SSA)对数据进行分解,得到高频、中频和低频三个部分。在模型的超参数优化中,采用了改进后的人工蜂群算法(ABC),以提高模型的性能。为验证 ABC-LSTM-GRU 混合模型的有效性,选择 NIFTY-50 股票指数进行实证分析。实验结果对比显示,ABC-LSTM-GRU 混合模型在时间序列预测方面的表现更佳,相较于 LSTM 与 GRU 模型,其在均方根误差(RMSE)指标上分别降低了 28.3%与 21.5%,显示出更为准确的预测性能。

关键词:GRU; LSTM; ABC; SSA; 股市预测

中图分类号:TP391 **文献标志码:**A

Time Series Decomposition and Prediction Model Based on ABC-LSTM-GRU

ZHU Zijiang, HE Liwen

(School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

✉ zhon0713@qq.com; helw@njupt.edu.cn

Abstract: This paper proposes a hybrid model of Artificial Bee Colony-Long Short Term Memory-Gated Recurrent Unit (ABC-LSTM-GRU) to address the issues of high noise and time dependence of financial time series data. This model integrates LSTM and GRU networks to comprehensively capture the long and short term relationships in time series. In the feature processing stage, features are screened by correlation analysis, and the data is decomposed by Singular Spectrum Analysis (SSA) into three parts, high frequency, medium frequency and low frequency. In the hyperparameter optimization of the model, an improved Artificial Bee Colony (ABC) algorithm is used to improve the performance of the model. In order to verify the effectiveness of ABC-LSTM-GRU model, NIFTY-50 stock index is selected for empirical analysis. The comparison of experimental results shows that ABC-LSTM-GRU hybrid model has better performance in time series prediction, and compared with LSTM and GRU model, it reduces the Root Mean Square Error (RMSE) index by 28.3% and 21.5%, respectively, demonstrating more excellent prediction performance.

Key words: GRU; LSTM; ABC; SSA; stock market forecast

0 引言 (Introduction)

股市作为经济体系的核心,对全球经济具有重要意义。在股

票市场中,准确预测股票价格一直是投资者和金融机构极为关注且极具挑战性的问题。然而,由于股票市场具有高度复杂性和不

确定性,所以股票价格的准确预测一直是一项充满挑战的任务。

近年来,深度学习在各领域的预测任务中展现出卓越性能,其在金融时间序列预测任务中同样表现优异。不同的深度学习学习方法,如卷积神经网络(CNN)^[1]、长短期记忆网络(LSTM)^[2-3]、双向长短期记忆网络(BI-LSTM)^[4]、门控循环单元(GRU)^[5-6]等,在股票价格预测中取得了显著成果。

此外,降维^[7]、超参数优化^[8]、时间序列分解^[9]等方法在一定程度上提高了股价预测的准确率。研究人员还提出各种混合深度学习模型^[10],通过融合不同的技术和方法提升预测性能,即降低均方根误差(RMSE)等误差指标。例如,林勇^[11]应用 MF-DCCA 方法分析了百度指数对人民币汇率的可预测性,结合鲸鱼优化算法-季节性趋势残差-长短期记忆(WOA-STL-LSTM)混合模型,证明其在人民币汇率预测中,相比传统模型具有更好的表现。这些混合模型在实验中取得了显著的进展,为研究如何提高股票价格预测性能提供了新的思路。

本文选取 NIFTY-50 股票指数进行预测,首先对输入数据进行特征处理,其次对预测数据进行分解,通过构建 LSTM-GRU 模型对分解后的数据进行预测,同时使用改进后的人工蜂群算法(ABC)获取最佳的超参数,以此获取更加精确的预测效果。

1 理论基础(Theoretical basis)

1.1 LSTM

LSTM 是一种专为解决传统循环神经网络(RNN)难以处理长期依赖关系的问题而设计的架构,其核心在于引入了输入门、遗忘门和输出门等关键组件,通过巧妙的门控制机制实现对信息流动的精准调控,其网络结构图如图 1 所示。

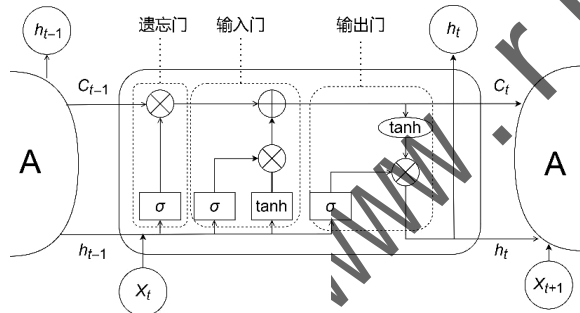


图 1 LSTM 网络结构图

Fig. 1 LSTM network structure

首先,输入门决定了当前时刻有多少新信息应该被纳入模型。通过考虑当前输入和前一个时间步的隐藏状态,输入门生成一个 0~1 的值,指示新信息的重要性。

其次,遗忘门决定了前一个时刻的记忆状态中有多少信息需要被遗忘。类似于输入门,遗忘门通过当前输入和前一个时间步的隐藏状态进行计算,生成一个 0~1 的值。

再次,细胞状态(Cell State)通过输入门和遗忘门的输出进行更新。这一细胞状态允许 LSTM 有效地保留并更新长期信息。

最后,输出门决定了当前时刻隐藏状态中有多少信息应该被输出。通过考虑当前输入和前一个时间步的隐藏状态,输出门生成一个 0~1 的值,控制信息的输出。

1.2 GRU

门控循环单元(Gated Recurrent Unit, GRU)相比于 LSTM 采用了更简化的结构,包含更新门(Update Gate)和重置门(Reset Gate)两个关键组件,其结构图如图 2 所示。

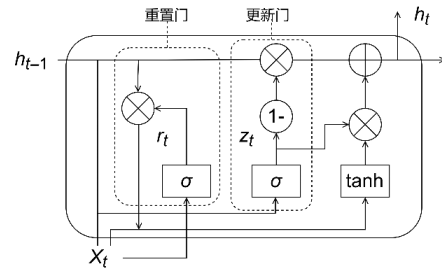


图 2 GRU 神经网络结构图

Fig. 2 Diagram of GRU neural network structure

GRU 的主要计算公式如下:

$$z_t = \sigma(w_z \cdot [h_{t-1}, X_t]) \quad (1)$$

$$r_t = \sigma(w_r \cdot [h_{t-1}, X_t]) \quad (2)$$

$$\tilde{h}_t = \tanh(w \odot [r_t \cdot h_{t-1}, X_t]) \quad (3)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4)$$

其中: X_t 表示 t 时刻的输入, w 、 w_z 、 w_r 表示模型学习的权重矩阵, h_{t-1} 表示上一时刻的隐藏状态, $\tilde{h}_t$ 表示候选隐藏状态, z_t 、 r_t 分别为更新门和重置门, σ 是 Sigmoid 函数, $\tanh$ 是双曲正切函数。

首先,重置门决定了模型在当前时间步应该“忘记”多少过去的信息。通过考虑当前输入和前一个时间步的隐藏状态,重置门生成一个 0~1 的值,控制前一个时间步的隐藏状态的遗忘程度。

其次,更新门决定了模型在当前时间步应该保留多少当前的信息。类似于重置门,更新门通过考虑当前输入和前一个时间步的隐藏状态生成一个 0~1 的值。

最后,通过这两个门的协同作用,GRU 实现了隐藏状态的灵活更新,既可以捕捉序列中的短期依赖关系,又能在一定程度上处理长期依赖关系。

1.3 ABC 算法

ABC 算法是一种启发式优化算法,该算法受到了蜜蜂觅食行为的启发,模拟了蜜蜂群体中蜜蜂的搜索行为,通过不断地在搜索空间中探索和利用潜在解决方案,寻找最优解。ABC 算法主要包括三种蜜蜂的角色,即工蜂、观察蜂和侦查蜂。

工蜂负责在搜索空间中生成候选解,并通过评估目标函数的值度量解的质量。观察蜂通过相邻解的信息交流,选择其中较优的解进行跟随和进一步探索。侦查蜂负责在搜索空间中进行随机的探索,引入多样性,防止陷入局部最优解。ABC 算法的关键在于蜜蜂群体之间的信息传递和协作,以及对搜索空间的全局搜索和局部搜索之间的平衡。

2 ABC-LSTM-GRU 混合模型构建(Construction of ABC-LSTM-GRU hybrid model)

本文设计的 ABC-LSTM-GRU 混合模型总体架构如图 3 所示。

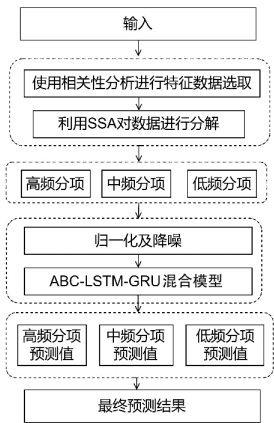


图 3 ABC-LSTM-GRU 混合模型总体架构

Fig. 3 Overall architecture of ABC-LSTM-GRU hybrid model

步骤 1: 对输入数据进行相关性分析。通过相关性分析的结果,筛选出对后续分解和预测具有重要影响的数据。利用奇异谱分解(SSA)将数据分解为三个部分,分别对应低频、中频和高频成分。

步骤 2: 对于高频、中频、低频三个部分的数据,首先进行MinMax 归一化,确保它们具有相似的数值范围;其次对每个频率部分分别应用小波降噪通过设定适当的阈值去除高频噪声,同时保留信号中的重要特征;最后将处理后的数据转化为PyTorch 张量,这样可以更方便地应用深度学习模型进行时间序列预测。

步骤 3: 使用 ABC-LSTM-GRU 混合模型对不同频率的数据分别进行预测。模型的层次结构采用 LSTM-GRU,其层次结构如图 4 所示。为了优化模型性能,针对每个频率的数据分别应用 ABC 算法获取最佳的超参数组合。通过对不同频率的预测结果进行重构,得到综合考虑各频率信息的最终预测值。

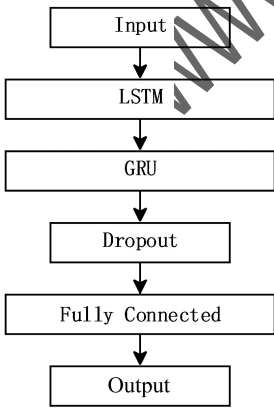


图 4 LSTM-GRU 模型层次结构

Fig. 4 LSTM-GRU model hierarchy

3 实证研究(Empirical research)

3.1 特征选取与预处理

金融序列数据选取 NIFTY-50 股票指数,数据来自开源

数据集网站,数据的时间跨度为 2016 年 5 月 2 日至 2021 年 4 月 30 日,总计 1 263 个交易日。获取的股票数据包含股票代码、交易日期、开盘价(Open)、最高价(High)、最低价(Low)、最后成交价(Last)、收盘价(Close)、成交量(Volume)、交易额(Turnover)、交易次数(Trades)、交割量(Deliverable Volume)、交割量占比(%Deliverble)、加权平均价格(VWAP)。

步骤 1: 获取数据后对所有字段使用斯皮尔曼秩相关系数进行相关性分析,结果如表 1 所示。斯皮尔曼秩相关系数是一种非参数统计量,用于度量两个变量之间的单调关系,其值范围为-1~1。值为 1,表示完全正相关关系;值为-1,表示完全负相关关系;值为 0,表示无相关关系。

Last、VWAP、Low、High、Open 的相关系数接近 1,表明它们与目标变量 Close 之间存在很强的正相关关系,即它们的值通常会一起升高或下降。

Deliverable Volume 的相关系数较低,说明其与 Close 之间的相关性不强。

Turnover、Trades、Volume 的相关系数为 0.5~0.7,表示与 Close 之间存在一定的正相关关系。

%Deliverble 的相关系数为-0.6,表明其与 Close 之间存在反相关性,即当特征值升高时,Close 的值可能会下降。

综上所述,根据斯皮尔曼秩相关系数进行筛选,选择保留与 Close 最接近的特征,即 Open、High、Low、Last、VWAP。

表 1 各字段的斯皮尔曼秩相关系数

Tab.1 Spearman rank correlation coefficients for each field

字段名	斯皮尔曼秩相关系数
Last	0.999 901
VWAP	0.999 577
Low	0.999 002
High	0.998 715
Open	0.997 239
Turnover	0.740 700
Trades	0.651 785
Volume	0.516 677
Deliverable Volume	0.190 801
%Deliverble	-0.612 009

步骤 2: 首先对目标数据进行 SSA,将其分解为三种频率的数据,即高频、中频和低频。其次按照股市数据的划分比例(8:2),对分割后的数据进行两个关键的处理步骤:一是归一化处理,确保不同特征之间的数值范围一致,避免模型偏向某些特征的情况;二是进行小波降噪,旨在去除数据中的噪声,提取潜在趋势信息,以及降低数据的复杂性。最后将处理后的数据转化为 PyTorch 张量,便于后续的模式预测。

3.2 超参数优化

LSTM-GRU 模型的超参数优化采用改进后的 ABC 算法,流程如图 5 所示,具体步骤如下。

步骤 1:随机生成初始参数集。在算法开始时,随机生成一个初始的参数集,包括学习率(learning_rate)、隐藏层维度(hidden_dim)、丢弃率(dropout)、训练轮数(num_epochs)。

步骤 2:计算初始性能。通过目标函数计算初始性能,即 RMSE 的值,这实际上是为了得到初始的性能基准,为整个算法的迭代过程做准备。

步骤 3:雇佣蜂选择参数集。雇佣蜂选择初始的参数集,该参数集将用于后续的计算。在雇佣蜂阶段,根据初始的参数集生成新的参数集,评估性能,若 RMSE 值降低,则保存新的参数。

步骤 4:概率计算和观察蜂的选择。概率计算用于决定哪些蜜蜂将被选择进入下一阶段,即观察蜂阶段。这种机制引入随机性,能帮助算法更全面地探索搜索空间,提高优化效果。

步骤 5:观察蜂阶段。观察蜂选择雇佣蜂,生成新的参数集,评估性能。若 RMSE 值降低,则保存新的参数。未被选中的雇佣蜂仍然可以保留原参数集,并在下一轮迭代中继续参与。

步骤 6:选择精英蜂。根据观察蜂的参数性能选择精英蜂。若达到了设定的迭代次数,则结束算法,输出精英蜂的结果。

步骤 7:继续迭代。若迭代次数未达到设定值,则继续迭代。精英蜂的参数集将成为新一轮的雇佣蜂进入下一轮的优化过程,继续重复“步骤 3”至“步骤 6”。

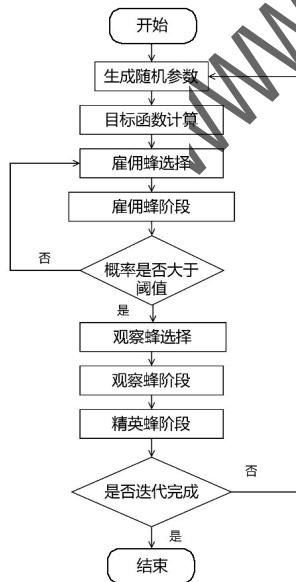


图 5 ABC 算法流程图

Fig. 5 ABC algorithm flow chart

对于当前模型的超参数优化,专注于调整 learning_rate、hidden_dim、dropout 和 num_epochs 四个参数。针对低频、中频和高频的数据分别进行了优化,优化的结果列于表 2 中。值得注意的是,其他模型预测中的超参数保持固定,learning_rate 设置为 0.02,hidden_dim 设置为 64,dropout 设置为 0.02,num_epochs 设置为 200。这样的设定有助于在超参数优化中突出目标参数的变化对模型性能的影响。

表 2 ABC 算法超参数优化

Tab.2 ABC algorithm hyperparameter optimization

参数名	初始范围	低频数据	中频数据	高频数据
learning_rate	[0.001,0.030]	0.012 97	0.013 70	0.014 79
hidden_dim	[16,128]	49	54	52
dropout	[0.01,0.03]	0.012 49	0.016 06	0.016 53
num_epochs	[50,400]	267	221	149

3.3 实验结果与分析

预测数据分为两个主要类别:一是基础股市数据(Raw Data,RD),包含 Open、High、Low、Volume 等原始信息;二是经过特征选取后的数据(Feature-Selected Data,FSD),包含 Open、High、Low、Last、VWAP 等精选特征。

为了评估模型的预测效果,采用了三个主要指标:RMSE、平均绝对误差(MAE)和平均绝对百分比误差(MAPE),这些指标的计算公式如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$
 (5)

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$
 (6)

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|$$
 (7)

其中: y_i 、 $\hat{y}_i$ 分别代表时间序列在时刻 i 的真实值和预测值, n 代表测试数据集的长度。

表 3 呈现了 LSTM-GRU 模型对不同频率数据的多指标评估结果,涵盖了在固定和优化超参数下的低频、中频、高频预测表现。高频分项的 RMSE、MAE 和 MAPE 指标均表现最佳,预测效果最出色。其次是中频分项,而低频分项的预测效果相对较差。此外,使用经过优化的参数进行的预测表现优于采用固定参数的预测,表明改进后的 ABC 算法在提升预测效果方面发挥了积极作用。

高频成分通常包含市场波动剧烈、瞬时变化显著的信息,这可能与市场的瞬时波动和快速变化有关。相比之下,低频成分可能包含趋势和长期变化,而中频成分则介于两者之间。由于低频成分受到噪声干扰较多,其预测效果相对较差,中频成分可能存在中期周期性波动,因此预测效果一般,而高频成分可能捕捉了市场中的短期波动,这些波动相对容易捕获。

表3 LSTM-GRU 模型对不同频率数据的
多指标评估结果

Tab.3 The multi indicator evaluation results of LSTM-GRU model on
data with different frequencies

分项	固定参数预测			优化后参数预测		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
低频 分项	28.833	20.057	0.013 17	23.979	17.729	0.011 84
中频 分项	17.306	12.296	0.007 91	14.716	10.611	0.007 35
高频 分项	11.968	8.780	0.006 32	10.135	6.514	0.004 49

表 4 展示了各模型的预测结果。从表 4 中可以明显观察到,在 LSTM 和 GRU 模型中,对于 FSD (Feature-Selected Data) 的预测效果均优于 RD (Raw Data),体现在 RMSE、MAE、MAPE 等指标均有明显提升,表明使用斯皮尔曼秩相关系数进行特征数据选取能够显著提升模型的预测性能。

此外,通过比较各个模型的预测值可以看出,构建的 ABC-LSTM-GRU 混合模型相较于独立使用 GRU 或 LSTM 模型,以及未经参数优化的 LSTM-GRU 模型,在股市预测中的表现更为优越。在未进行参数优化的情况下,LSTM-GRU 模型表现最优,GRU 次之,而 LSTM 表现相对较差。总体而言,通过 ABC 算法进行参数优化,不仅有助于提升模型的泛化性能,使其更好地适应未见过的数据,而且通过综合利用 LSTM 和 GRU 循环神经网络,并进行专门的超参数优化,显著提升了模型的整体性能。以上结果表明,通过巧妙地调整模型的超参数,可以使模型更有效地捕捉数据的特征,从而提高其泛化性能,模型能更准确地预测目标。

表 4 各模型的预测结果

Tab.4 The predicted results of each model

模型	数据类型	RMSE	MAE	MAPE
LSTM	RD	23.895	20.832	0.014 52
	FSD	17.390	15.492	0.010 25
GRU	RD	18.886	15.287	0.010 31
	FSD	14.389	12.249	0.007 92
LSTM-GRU	FSD	12.513	9.666	0.007 16
ABC-LSTM-GRU	FSD	11.091	7.224	0.004 78

4 结论 (Conclusion)

本文充分展示了在金融时间序列中,ABC-LSTM-GRU 混合模型表现出卓越的性能,它充分发挥了 LSTM 和 GRU 擅长捕捉长期依赖性和高效处理短期变化的特点。特征选取方法在模型预测中取得显著效果,通过相关分析筛选特征有效地降低了数据维度、去除了冗余信息及提高了模型的训练和预测效率。ABC 算法能自动搜索最佳的超参数配置,使得模型更好地适应数据,降低了过拟合风险,提高了模型预测的准确性,为其实际应用提供了更有效和可靠的方案。综合而言,ABC-LSTM-GRU 混合模型在金融时间序列预测中展现出出色的综合性能。

在未来的研究中,可以考虑引入金融文本信息或消费者情绪作为模型的额外输入特征,以进一步提高模型对市场变化的适应性和预测精度。此外,对时间序列进行更深入的分解也是一个可行的方向。通过将时间序列数据分解为趋势、季节性和残差三个部分,有望更全面地理解数据的内在结构,从而改善模型在不同时间尺度上的预测效果。这些方法的综合应用可以为股市预测模型性能的进一步提升提供有益的参考和改进方向。

参考文献 (References)

[1] NOURBAKHS Z, HABIBI N. Combining LSTM and CNN methods and fundamental analysis for stock price trend prediction[J]. Multimedia tools and applications, 2023,82:17769-17799.

[2] 李丽萍,曾丽芳,江绍萍,等. 基于 LSTM 神经网络的股票价格预测[J]. 云南民族大学学报(自然科学版),2023,32(4):528-532.

[3] ZHU T L, LIAO Y X, TAO Z. Predicting Google's stock price with LSTM model[J]. Proceedings of business and economic studies,2022,5(5):82-87.

[4] MANISH C, MANISH D. Financial models for Indian stock prices prediction using LSTM and bi-directional LSTM model[J]. IITM journal of management and IT, 2023,12(2):10-14.

[5] HU J C, CHANG Q Q, YAN S Y. A GRU-based hybrid global stock price index forecasting model with group decision-making[J]. International journal of computational science and engineering,2023,26(1):12-19.

[6] 毛增立. 基于 SSA-GRU 循环神经网络的地产股价格指数预测研究[D]. 哈尔滨:哈尔滨工业大学,2021.

[7] 刘子豪. 基于降维方法的信用风险评估模型研究及其应用[D]. 蚌埠:安徽财经大学,2021.

[8] 李晓寒,贾华丁,程雪,等. 基于改进遗传算法和图神经网络的股市波动预测方法[J]. 计算机应用,2022,42(5):1624-1633.

[9] 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(4):104-116.

[10] RANJAN P G, NARAYAN M M. An LSTM-GRU based hybrid framework for secured stock price prediction[J]. Journal of statistics and management systems, 2022, 25(6):1491-1499.

[11] 林勇. 基于深度学习和网络大数据的金融时间序列分析及预测[D]. 成都:四川师范大学,2022.

作者简介:

朱子敬(1999-),男,硕士生。研究领域:云计算与大数据,深度学习。

何利文(1968-),男,博士,教授。研究领域:网络与信息安全,云计算与大数据,图像处理。