

基于CEEMDAN和LSTM的股指价格组合预测方法：来自5个国家的数据分析

林昇铭¹, 王嘉宏^{1,2*}, 袁锦琛¹, 曾莹萍¹, 林持旺¹

¹福建理工大学计算机科学与数学学院, 福建 福州

²福建理工大学大数据挖掘与应用技术重点实验室, 福建 福州

收稿日期: 2024年1月25日; 录用日期: 2024年2月22日; 发布日期: 2024年2月29日

摘要

股票价格预测一直是研究者们挑战的领域。然而, 现有的基于深度学习的预测方法在预测不同国家的股票指数时无法达到统一优秀的效果。因此, 本文提出了一种名为CEEMDAN-PSO-LSTM的新模型来预测多个国家的股票指数收盘价。首先, 我们使用完全自适应噪声集合经验模态分解(CEEMDAN)方法将原始股票指数收盘价序列分解为多个本征模态函数(IMF)。然后, 得到的各个IMF通过利用粒子群优化算法(PSO)优化长短期记忆网络(LSTM)的超参数后的模型进行预测, 最终, 将各IMF的预测结果进行加和得到对原始收盘价序列的预测结果。为验证所提方法的可行性, 我们将其与LSTM、PSO-LSTM、EMD-LSTM和CEEMDAN-LSTM这四个模型进行对比, 并选取五支来自不同国家且具有代表性的股票指数作为数据集。通过各模型在各数据集上的实验表明本文所提方法表现优于其他四个模型, 这说明本文所提方法具有优良的可行性和普适性。

关键词

长短期记忆网络, 完全自适应噪声集合经验模态分解, 粒子群优化算法, 股票价格预测, 深度学习

A Combinatorial Prediction Method for Stock Index Price Based on CEEMDAN and LSTM: Data Analysis from Five Countries

Shengming Lin¹, Chia-Hung Wang^{1,2*}, Jinchen Yuan¹, Yingping Zeng¹, Chiwang Lin¹

¹College of Computer Science and Mathematics, Fujian University of Technology, Fuzhou Fujian

²Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou Fujian

Received: Jan. 25th, 2024; accepted: Feb. 22nd, 2024; published: Feb. 29th, 2024

*通讯作者。

文章引用: 林昇铭, 王嘉宏, 袁锦琛, 曾莹萍, 林持旺. 基于 CEEMDAN 和 LSTM 的股指价格组合预测方法: 来自 5 个国家的数据分析[J]. 计算机科学与应用, 2024, 14(2): 449-459. DOI: 10.12677/csa.2024.142045

Abstract

Stock price prediction has always been a challenging field for researchers. However, existing deep learning based prediction methods cannot achieve uniform and excellent results in predicting stock indices of different countries. Therefore, this article proposes a new model called CEEMDAN-PSO-LSTM to predict the closing prices of stock indices in multiple countries. Firstly, we use the complete ensemble empirical mode decomposition of adaptive noise (CEEMDAN) method to decompose the original stock index closing price sequence into multiple intrinsic mode functions (IMF). Then, the obtained IMF models are predicted by using the particle swarm optimization (PSO) algorithm to optimize the hyperparameters of the Long Short-Term Memory Network (LSTM). Finally, the prediction results of each IMF are summed to obtain the prediction results of the original closing price series. To verify the feasibility of the proposed method, we compared it with four models: LSTM, PSO-LSTM, EMD-LSTM, and CEEMDAN-LSTM, and selected five representative stock indices from different countries as the dataset. The experiments of each model on various datasets show that the proposed method performs better than the other four models, indicating that the proposed method has excellent feasibility and universality.

Keywords

Long Short-Term Memory Network (LSTM), Complete Ensemble Empirical Mode Decomposition of Adaptive Noise (CEEMDAN), Particle Swarm Optimization (PSO), Stock Price Prediction, Deep Learning

Copyright © 2024 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 引言

1.1. 研究背景

股票市场是一个信息错综复杂, 状态瞬息万变的市場。政治、经济、市场环境、政策环境、社会行为等都会影响股票价格的变动。因此, 如何把握股票市場行情信息和资产价格走势, 是股票投资者深入研究的课题。此外, 如何通过交易策略获得投资收益也是当下长期热门的研究[1]。

尽管有效市場假说认为股票市場是不可预测的, 但它仍然受到多种因素的影响, 例如经济政策、行业趋势、公司管理等, 这些因素都会影响股票价格。随着大数据时代的发展, 研究人员开始将人工智能应用于这一领域, 其中包括机器学习和深度学习技术[2]。然而机器学习预测方法存在一些局限性。相比之下, 深度学习网络具有强大的特征学习能力, 其网络层数更多, 结构更复杂, 能将数据中的浅层信息转化为更抽象的高特征信息, 因此, 深度学习对比机器学习, 有更好的预测效果并且已经在许多领域得到应用[3] [4]。

深度学习目前广泛应用于股票价格预测领域, 并产生了良好的结果。尽管使用这些方法在数据分析表现优异, 但时间序列存在不稳定和随机性, 以至于分析和预测时间序列依然具有挑战性。为了解决时间序列的存在的问题, 研究者们提出了时序分解方法, 以便于将时间序列分解成稳定且有序的子序列进行数据分析。经验模态分解(EMD)和 CEEMDAN 等时序分解方法可以作为降低时间序列复杂性的有用工

具。这些方法可以很容易地将高波动数据分解为相应的较小信号分量。

1.2. 国内外研究现状

由于投资者关注股票价格未来趋势，准确预测股票价格一直是金融领域备受关注的话题。然而，这仍然是一个难以解决的问题，许多研究者投身其中，并已取得不错的成果。传统的股市预测方法主要集中于时间序列分析。De Gooijer 等人[5]回顾了国际预报员协会和国际预报杂志期刊上发表的论文，发现其中超过 30% 的论文侧重于时间序列分析。这些方法包括自回归模型(AR)、移动平均模型(MA)、自回归移动平均模型(ARMA)。然而，这些方法都重点关注时间序列本身，而忽略了历史数据对序列的影响。因此，研究者探索了新的方向，将机器学习和深度学习应用于该领域。

Hall 等人[6]分析了人工神经网络(ANN)在约旦证券市场股票价格预测方面的可行性。Murat 等人[7]利用 MLP 模型对道琼斯指数的大盘数据进行预测。陈佳[8]则利用循环神经网络模型(RNN)研究中国股票市场。Wang 等人[9]使用 LSTM 神经网络并根据历史价格对股票价格进行预测，发现与其他人工智能算法相比，LSTM 的预测能力更好。宋刚等人[10]将 PSO 与 LSTM 相结合，提出一种 PSO-LSTM 模型来预测股票价格，PSO 算法对于神经网络有很好的优化效果[11]。

虽然上述方法在金融领域表现不错，但对于更为复杂的金融时间序列，普通的深度网络模型表现就不够出色。研究者发现时序分解方法与深度学习模型相结合，有很好的预测效果。Wang 等人[12]将时序分解方法与 PSO-LSTM 模型相结合，发现时序分解方法能有效提高股票预测的精度。颜轲越等人[13]则用了不同的信号分解方法用于股票价格预测，发现 CEEMDAN 有很好的提升效果。

1.3. 研究方法

本研究将时间序列分析与深度学习模型相结合，提出了一种名为 CEEMDAN-PSO-LSTM 的模型，用于预测不同国家股票指数的收盘价格。首先利用 CEEMDAN 算法对原始股票价格时间序列进行时序分解，得到 IMF。在 LSTM 模型的基础上，通过 PSO，对 LSTM 网络的超参数进行寻优。然后各个 IMF 分别利用优化后的模型进行预测，最终将得到的各个预测值通过加和处理得到最终的预测值，该预测值就是模型对原始股票价格时间序列的预测结果。本文选取五个国家的其中一支代表性的股票指数作为测试数据集，对本文提出的预测方法和 LSTM、PSO-LSTM、EMD-LSTM、CEEMDAN-LSTM 这四个方法的预测结果进行对比实验，以此验证本文所提出方法的有效性和实用性。

2. 模型介绍

在本节，我们会详细介绍本文所提出的方法。我们将 CEEMDAN，PSO，LSTM 神经网络应用在股票预测领域，构建了一个名为 CEEMDAN-PSO-LSTM 的模型来预测股票价格指数。CEEMDAN-PSO-LSTM 组合模型的股价预测流程图见图 1。该混合模型主要分为 PSO 算法部分、CEEMDAN 分解部分、LSTM 神经网络部分以及模型评估部分。该模型的各部分描述如下：

1) 输入为原始股票指数的收盘价记作 $I_{original}$ ，首先将其数据预处理后作为 PSO 算法的输入。本论文的数据预处理分为两步，其一是数据归一化，具体方法为 MinMax 归一化，其公式如下：

$$x_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (1)$$

此时，式中的 x_i 为 $I_{original}$ ，计算得到的 x_i 记作 I_{input} 。其二是划分数据集，将归一化后的数据集分割为训练集和测试集以及验证集，本实验训练集和测试集的选取规则为数据集的 70% 为训练集，30% 为测试集，随机 10% 的数据集作为验证集。

2) 利用 PSO 算法优化 LSTM 神经网络超参数, 具体步骤如下: 1) 初始化粒子群参数。2) 初始化粒子的速度和位置。随机生成一个粒子, 该粒子的位置为 $X_i = (h, b, l)$, 其中 h 为隐含层神经元个数, b 为 batch_size 大小, l 为学习率大小。该粒子的速度为 $V_i = (V_{i1}, V_{i2}, V_{i3})$ 。3) 确定 PSO 算法的适应度函数。适应度函数值公式如下:

$$Fit = \frac{1}{N} \sum_{k=1}^N (I_{input} - h_k)^2 \quad (2)$$

其中 N 为测试集的长度, I_{input} 为输入数据。4) 比较适应度函数值并更新个体最优值和群体最优值。通过与初始位置的适应度值比较更新个体最优值 P_{id} 和群体最优值 P_{gd} , 根据公式(3)和(4)不断更新粒子的速度和位置, 直到达到最大迭代次数, 确定最优位置, 其参数 h, b, l 即需要优化的 LSTM 神经网络的超参数。获取最优参数。公式(3)和(4)如下:

$$V_{id}^{t+1} = \omega \cdot V_{id}^t + c_1 r_1 \cdot (P_{id}^t - X_{id}^t) + c_2 r_2 \cdot (P_{gd}^t - X_{id}^t) \quad (3)$$

$$X_{id}^{t+1} = X_{id}^t + V_{id}^t \quad (4)$$

其中, w 为惯性权重; c_1, c_2 为学习因子, 取值范围为(0, 2], r_1, r_2 为[0, 1]之间均匀分布的随机数。 P_{id}^t 和 P_{gd}^t 分别表示第 i 个粒子在 t 时刻的个体最优值和全局最优值。 V_{id}^t 和 X_{id}^t 分别表示为第 i 个粒子在 t 时刻的速度和位置。

3) CEEMDAN 算法分解原始序列。在获取 LSTM 神经网络的最优参数之后, 再将 $I_{original}$ 经过 CEEMDAN 算法进行序列分解, 分解后得到 n 个 IMF ($IMF1, IMF2, \dots, IMF_n$)。然后, 将这 n 个 IMF 各自通过公式(1)进行数据预处理, 得到 $(IMF1_{input}, IMF2_{input}, \dots, IMF_n_{input})$, 将该结果作为 LSTM 网络的输入。

4) LSTM 网络预测。上述过程中已经得到经过 PSO 算法优化过后的超参数以及经过 CEEMDAN 算法分解后的收盘价的 IMF 分量。对 n 个 IMF 分别经过 n 个 LSTM 网络进行预测, 得到 n 个预测值 $(\hat{p}_1, \hat{p}_2, \dots, \hat{p}_n)$, 输出这 n 个预测结果。

5) 融合预测结果。融合函数是混合方法的核心。目前, 有许多融合函数, 如和、加权和、加权积等。这些融合函数的作用是将几个结果合并为最终结果。本实验的融合函数是将 n 个 IMF 各自的预测值 p_i 反归一化后进行求和。得到最终对于原始序列的预测值 P , 公式如式(5)和式(6)所示:

$$p_i = \hat{p}_i \cdot (\max(IMFi) - \min(IMFi)) + \min(IMFi) \quad (5)$$

$$P = \sum_{i=1}^n p_i \quad (6)$$

6) 评估模型。我们将最终得到的预测值与股指时间序列数据的实际值进行比较, 并计算 RMSE、MAE、MAPE 和 R^2 的值和记录运算时间 CPU_TIME。用这五个值来评估模型的性能。它们的公式如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{m=1}^n (I_{original} - P)^2} \quad (7)$$

$$MAE = \frac{1}{n} \sum_{m=1}^n |I_{original} - P| \quad (8)$$

$$MAPE = \frac{100\%}{n} \sum_{m=1}^n \left| \frac{I_{original} - P}{I_{original}} \right| \quad (9)$$

$$R^2 = 1 - \frac{\sum_{m=1}^n (I_{original} - P)^2}{\sum_{m=1}^n (I_{original} - \bar{I}_{original})^2} \quad (10)$$

在评估模型好坏程度上，RMSE、MAE、MAPE 数值越小，模型的预测结果与真实值之间的误差越小。 R^2 越接近 1 代表预测值与真实值拟合度越大，模型性能越好。

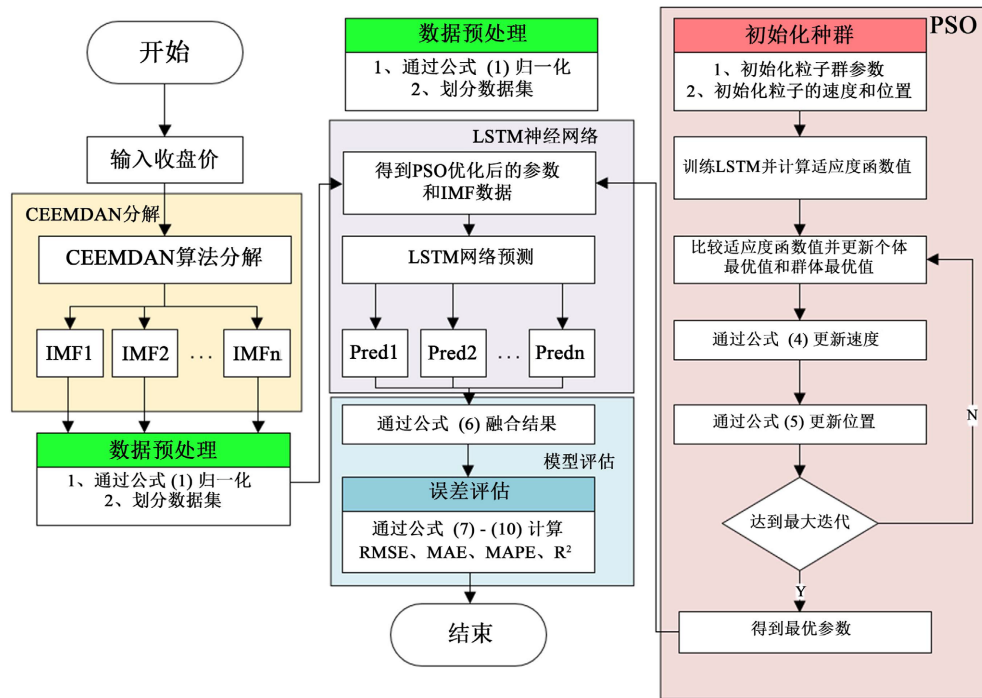


Figure 1. Price of the stock prediction process of CEEMDAN-PSO-LSTM combined model
图 1. CEEMDAN-PSO-LSTM 组合模型预测股票价格流程

3. 实验结果分析

本节详细介绍了本文实验部分的具体内容，为了更好的验证本文所提方法的预测结果，本次实验除了本文所提的方法之外，还与 LSTM、PSO-LSTM、EMD-LSTM、CEEMDAN-LSTM 这四个方法进行对比，评估的指标为 RMSE、MAE、MAPE、 R^2 和 CPU_TIME。通过对比各模型的评价指标可以看出本文所提方法更加可靠。本次实验的硬件环境为 CPU: 12th Gen Intel(R) Core(TM) i3-12100F@3.30 GHz，软件环境为 window10, Python 3.8.0, TensorFlow-cpu 2.1.0。

3.1. 数据集选取

本实验的数据集来源为：<https://www.tushare.pro/>，表 1 展示了所选取数据集的名称、所属地区以及选取的时期。

本文选取了五个不同国家的各一支具有代表性的股票指数。本实验通过选取不同国家的股票指数并进行分析，旨在发现国家差异性导致的预测结果的差别。选取的五个国家分别来自亚洲的中国和日本，欧洲的德国和英国以及北美洲的美国。选取的股票指数是上证指数(SSEC)、标准普尔 500 指数(S&P500)、英国富时 100 指数(FTSE 100)、德国法兰克福指数(DAX)、日经 225 (N225)。选取日期从

Table 1. The region and period of the dataset**表 1.** 数据集的地区和时期

指数	地区	国家	市场	数据量
SSEC	亚洲	中国	中国	2916
S&P500	北美洲	美国	美国	3017
FTSE 100	欧洲	英国	混合	3030
N225	亚洲	中国	日本	2937
DAX	欧洲	德国	德国	3042

2012/01/01~2024/01/01 约 3000 天左右的各个股票指数的收盘价。本实验训练集和测试集的选取规则为前 70% 为训练集，后 30% 为测试集。随机选取数据集的 10% 为验证集。本文后续实验利用 RMSE 作为模型的损失函数，选取的 epoch 均为 200，在整个训练周期过后保存验证集在损失函数上表现最好的那一次迭代结果作为输出。

3.2. 回看天数选择

本实验采用滚动预测建模方式，用最近 look_back 天的值预测下一天的值，look_back 称为回看天数。但是不同股票指数利用不同回看天数进行预测的结果也不同，因此我们选取回看天数为 5~10、15、20、25、30。利用本文所提模型在各数据集上进行实验，计算其 MAPE，通过对比各 MAPE 的值，选取表现最好的回看天数。图 2 展示了各数据集不同回看天数选取下的 MAPE 的值。从图 2 中我们可以看到，模型在 SSEC、N225、DAX、FTSE 100 这四个数据集上选取回看天数为 6 天时预测效果最好，模型在 S&P500 数据集上选取回看天数为 15 天时预测效果最好。因此，最终本实验选取的回看天数为 6 天。

3.3. 模型参数设置

本文所提方法利用 PSO 优化 LSTM 层的超参数，对需要优化的超参数的取值范围设置如下：学习率大小取值范围[0.001, 0.01]，批处理大小取值范围[0, 50]，隐藏层单元个数取值范围[30, 300]。PSO 算法涉及到的参数初始化设置如表 2。本文所提方法的 PSO 算法部分在该初始化设置下对本文的 LSTM 模型的超参数进行寻优。

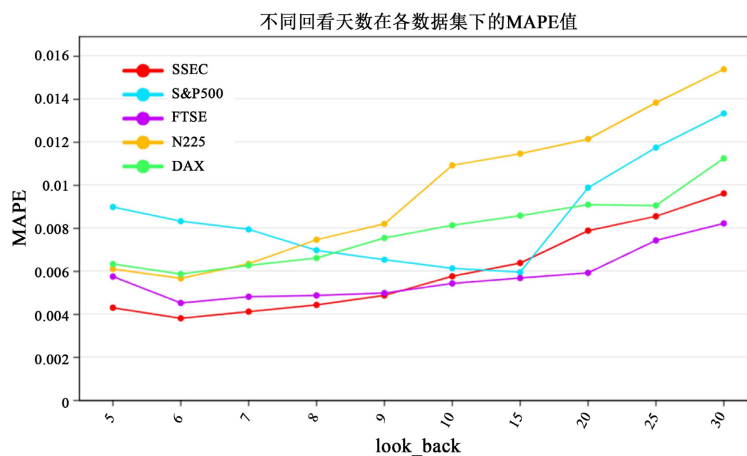
**Figure 2.** MAPE of different look_back for different datasets**图 2.** 不同回看天数在各数据集下的 MAPE 值

Table 2. The parameters settings of PSO algorithms
表 2. PSO 算法的参数设置

参数	数值
w	0.8
r_1	0.3
r_2	0.2
c_1	2
c_2	2
d	3
N	2
batch_size	[0, 50]
learning_rate	[0.001, 0.01]
neurons	[30, 300]

本实验各个网络模型的超参数设置如表 3 所示，其中 Proposed 为本文所提出的 CEEMDAN-PSO-LSTM 模型。本文后续实验都在表 3 参数设置下完成。

Table 3. The value of parameters of each method
表 3. 不同方法的参数值

方法	迭代次数	回看天数	学习率	Dropout	神经元	批大小	优化器	损失函数
LSTM	200	6	0.01	0.1	150	16	Adam	MSE
EMD-LSTM	200	15	0.001	0.1	200	15	Adam	MSE
CEEMDAN-LSTM	200	10	0.001	0.2	150	22	Adam	MSE
PSO-LSTM	200	6	0.00448	0.1	213	18	Adam	MSE
Proposed	200	6	0.00448	0.1	213	18	Adam	MSE

3.4. 实验结果及分析

本论文各方法的对比实验结果如表 4 所示。表中加粗部分即为表现效果最好的值，从表 4 中我们可以发现，本文所提方法在各个数据集上 RMSE、MAE、MAPE、 R^2 表现都优于其他模型。但是在 CPU_TIME 上的表现就比较差，原因是 PSO 算法需要时间寻找模型的最优解，并且 CEEMDAN 对原始序列进行分解需要时间，因此在 CPU_TIME 上表现最好的是最传统的 LSTM 模型。这在未来是可以优化的一个方向。图 3 和图 4 展示了本文所提方法在 SSEC 和 S&P500 这两个数据集上的预测结果与真实值之间的情况。从图中我们可以明显看到，所提方法在各数据集上的预测值与真实值都十分接近。这一观察结果表明我们的模型能够准确地预测目标变量，并且预测结果与真实值之间存在着密切的一致性。

通过观察表 4 我们可以看到对比 PSO-LSTM 和 LSTM，经过 PSO 算法寻优后的 LSTM 模型预测精度在各个数据集都有所提高，这证明用 PSO 算法对 LSTM 超参数进行寻优是有效的。而将 EMD 与 LSTM 相结合后的模型表现对比 LSTM 提升就比较大，并且表现同样优于 PSO-LSTM。这是因为神经网络对平缓的序列具有更好的预测效果，精度的提升很好的证明了时序分解方法与神经网络相结合的方法应用在金融预测领域效果显著。

我们还可以关注 EMD-LSTM 与 CEEMDAN-LSTM 的差别,我们可以发现 CEEMDAN-LSTM 表现效果都优于 EMD-LSTM。这是由于 CEEMDAN 算法可以解决 EMD 算法模式混叠的问题,其分解出的 IMF 更适合被预测,因此与 CEEMDAN-LSTM 模型性能都优于与 EMD-LSTM 模型。但是 CEEMDAN-LSTM 模型在 CPU-TIME 上花费了更长的时间,这是因为 CEEMDAN 算法对比 EMD 算法需要更长的时间来分解原始序列,因此 CEEMDAN-LSTM 模型需要花费更多的运算时间。

图 5 展示了本实验中各个股票指数的 R^2 雷达图。通过观察图 5,我们可以清晰地看到本文所提出的方法在各个国家的股票指数上具有出色的 R^2 值表现。这表明本文方法在不同股票指数上均展现出统一的优秀水平。我们可以特别关注到,在 SSEC 数据集上,本文所提出的方法相对于 CEEMDAN-LSTM 方法的改进并不十分显著。然而,在 N225 和 DAX 数据集上,相较于 CEEMDAN-LSTM 方法,本文所提出的方法展现出了明显的性能提升。这进一步证明了本文所提出的模型具有普适性,适用于预测各种类型的时间序列数据。因此,该模型非常适合用于预测各个国家的股票指数。

Table 4. Prediction results of all comparison methods for five stock indices

表 4. 所有比较方法对五支股票指数的预测结果

指数	方法	RMSE	MAE	MAPE	R^2	CPU_TIME
SSEC	LSTM	49.42934	35.31324	0.011873	0.96442	23.1411
	EMD-LSTM	23.95438	18.13328	0.005658	0.99013	223.3267
	CEEMDAN-LSTM	22.28871	17.04217	0.005321	0.99117	251.4321
	PSO-LSTM	37.32633	28.67844	0.008327	0.98231	87.2213
	Proposed	19.8797	15.97543	0.004864	0.99265	312.2543
S&P500	LSTM	81.26213	61.74108	0.016926	0.98123	24.8193
	EMD-LSTM	42.55853	36.1823	0.009753	0.98923	210.2357
	CEEMDAN-LSTM	38.19496	30.15249	0.008044	0.99133	224.3283
	PSO-LSTM	56.5458	43.5343	0.011476	0.98332	86.2342
	Proposed	29.68454	23.67231	0.005786	0.99469	340.2357
FTSE	LSTM	104.35234	69.87963	0.010234	0.96587	29.5431
	EMD-LSTM	57.52134	45.23177	0.006992	0.98942	221.8734
	CEEMDAN-LSTM	52.41531	38.58732	0.005913	0.99123	241.3278
	PSO-LSTM	85.87344	56.34467	0.008754	0.98153	92.2211
	Proposed	46.6532	32.53211	0.004776	0.99321	322.3211
N225	LSTM	490.01216	357.3127	0.014324	0.97535	30.4321
	EMD-LSTM	241.63233	186.28561	0.007233	0.99113	287.3124
	CEEMDAN-LSTM	188.62348	147.78641	0.005864	0.99225	333.7655
	PSO-LSTM	348.58122	269.38194	0.010947	0.98543	93.3532
	Proposed	167.06532	131.31241	0.004611	0.99523	418.7983
DAX	LSTM	232.01216	165.22418	0.012901	0.97533	30.6483
	EMD-LSTM	152.60336	126.33213	0.009394	0.98978	221.2357
	CEEMDAN-LSTM	124.58732	91.52134	0.006912	0.99134	236.6912
	PSO-LSTM	197.34513	135.26432	0.010557	0.98175	86.2312
	Proposed	101.54321	79.89764	0.005321	0.99404	347.1238

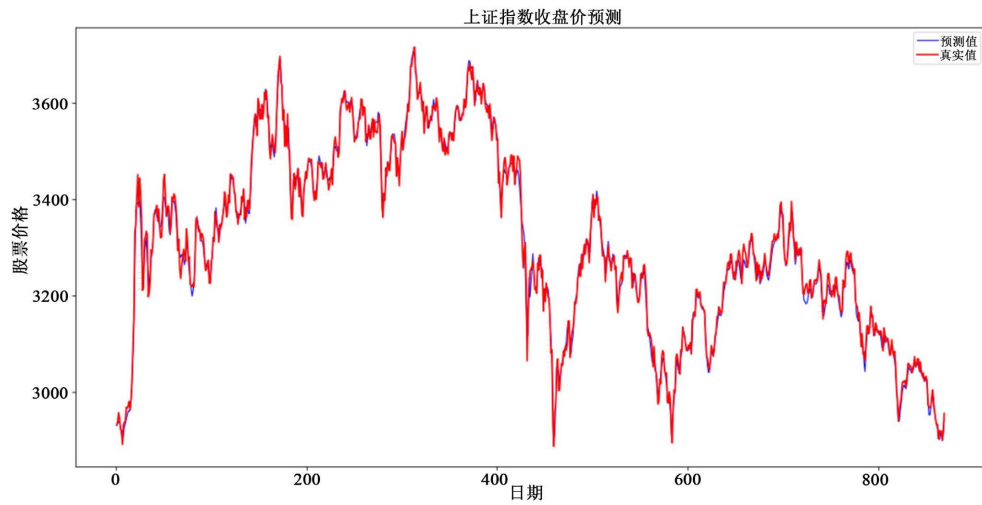


Figure 3. The close price and prediction of SSEC of proposed algorithm
图 3. 所提算法在 SSEC 上的收盘价的预测值和真实值

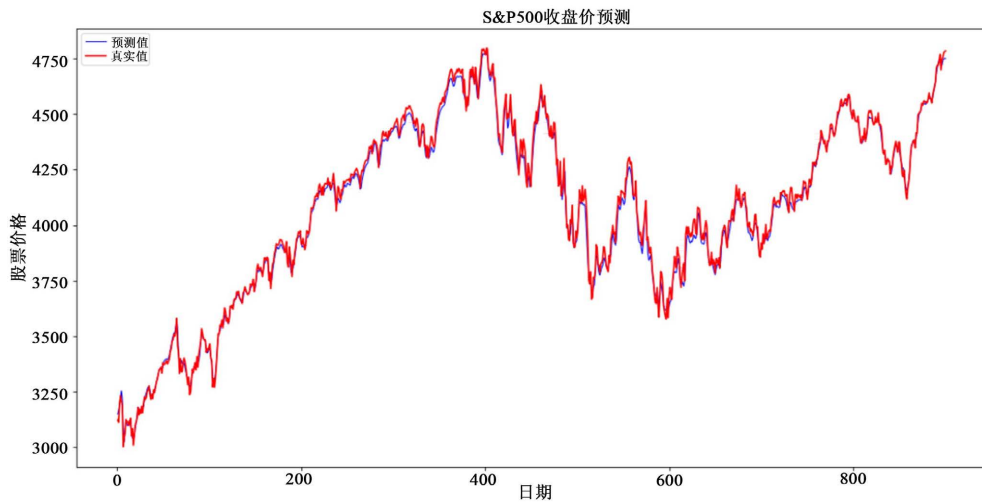


Figure 4. The close price and prediction of S&P500 of proposed algorithm
图 4. 所提算法在 S&P500 上的收盘价的预测值和真实值

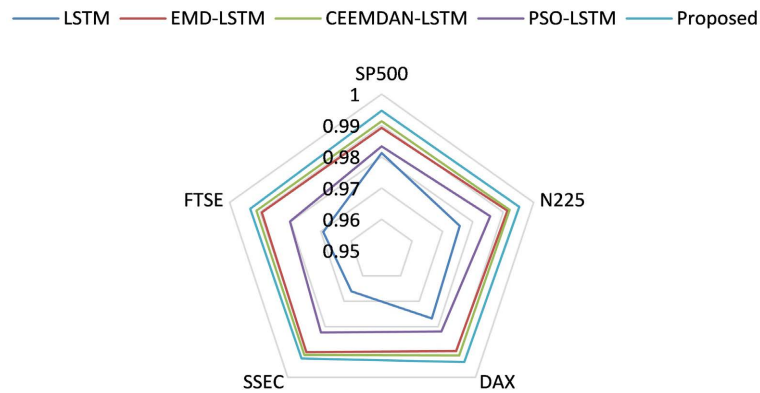


Figure 5. Radar chart of R^2 in this experiment
图 5. 本实验关于 R^2 的雷达图

4. 结论

在这项研究中,我们提出了一种新的神经网络模型,命名为 CEEMDAN-PSO-LSTM。我们的模型将时间序列分解算法与神经网络相结合,并利用 PSO 算法对神经网络的超参数进行优化。通过结合时间序列分解算法,我们的模型能够有效预测具有较大波动性的序列,并通过 PSO 算法进一步增强模型的预测能力。在实验分析部分我们还分析了 CEEMDAN 算法分解出的各个 IMF 分量对预测结果的影响。并且将我们的模型与其他四种深度学习方法进行了比较。我们选取了五个不同国家的数据集,使用的评估指标为 RMSE、MAE、MAPE、 R^2 以及 CPU_TIME。实验结果表明,我们的方法在预测各国股票指数方面表现更佳,验证了其有效性和普适性。实验结果分析显示,结合时间序列分解方法可以更好地预测非线性和非平稳的时间序列,而结合 PSO 算法对 LSTM 网络进行优化可以减少人为调参的影响。

尽管我们提出的模型在股票预测领域提供了可行的方法,但本文的实验仍有改进的空间。首先,当预测经过时序分解方法得到的 IMF1 分量时,由于该序列的复杂性和随机性,模型的预测效果较差,这是我们未来需要解决的问题。其次,我们可以对基础的 LSTM 神经网络模型进行改进。最后,可以尝试使用其他的优化算法来优化神经网络模型的参数。

基金项目

本课题得到福建省科学技术厅面上基金科研项目(2021J011070)、福建理工大学科研启动基金资助项目(GY-Z18148)的资助。

参考文献

- [1] Wang, C.-H. and Lee, C.-J. (2019) Data Analysis of Portfolio Optimization Using Artificial Neural Network in China's Stock Market. *Proceedings of 2019 IEEE 5th International Conference on Big Data Intelligence and Computing (DATACOM 2019)*, Kaohsiung, Taiwan, 18-21 November 2019, 33-38.
- [2] 侯亚妮. 基于 Bi-LSTM 深度学习的股票价格预测[J]. *统计学与应用*, 2021, 10(3): 538-546.
- [3] Wang, C.-H., Cai, J.B., Ye, Q., Suo, Y.F., Lin, S.M. and Yuan, J.C. (2023) A Two-Stage Convolution Network Algorithm for Predicting Traffic Speed Based on Multi-Feature Attention Mechanisms. *Journal of Intelligent & Fuzzy Systems*, **45**, 5181-5196. <https://doi.org/10.3233/JIFS-231133>
- [4] Wang, C.-H., Ye, Q., Cai, J.B., Suo, Y.F., Lin, S.M., Yuan, J.C. and Wu, X.J. (2024) A Novel Data-Driven Integrated Detection Method for Network Intrusion Classification Based on Multi-Feature Imbalanced Data. *Journal of Intelligent & Fuzzy Systems*. <https://content.iospress.com/articles/journal-of-intelligent-and-fuzzy-systems/ifs234402>
- [5] Gooijer, J.G. and Hyndman, R.J. (2005) 25 Years of IFF Time Series Forecasting: A Selective Review. *Tinbergen-Institute*, **22**, 443-447. <https://doi.org/10.2139/ssrn.748904>
- [6] Hammad, A.A.A., Ali, S.M.A. and Hall, E.L. (2007) Forecasting the Jordanian Stock Price Using Artificial Neural-network. *Intelligent Engineering Systems through Artificial Neural Networks*, **17**, 1-6.
- [7] Sezer, O.B., Ozbayoglu, A.M. and Dogdu, E. (2017) An Artificial Neural Network-Based Stock Trading System Using Technical Analysis and Big Data Framework. *ACM SE '17: Proceedings of the SouthEast Conference*, New York, April 2017, 223-226. <https://doi.org/10.1145/3077286.3077294>
- [8] 陈佳. RNN 神经网络在股指预测中的应用研究[D]: [硕士学位论文]. 天津: 天津科技大学, 2020.
- [9] Wang, C.-H., Wu, X.J. and Chen, Y.-T. (2021) An Empirical Analysis for Forecasting Stock Index Based on LSTM Neural Network. *Proceedings of the 5th International Conference on Electronic Information Technology and Computer Engineering (EITCE2021)*, Xiamen, 22-24 October 2021, 636-641. <https://doi.org/10.1145/3501409.3501524>
- [10] 宋刚, 张云峰, 包芳勋, 等. 基于粒子群优化 LSTM 的股票预测模型[J]. *北京航空航天大学学报*, 2019, 45(12): 2533-2542.
- [11] Wang, C.-H., Zhao, Q.G. and Tian, R. (2023) Short-Term Wind Power Prediction Based on a Hybrid Markov-Based PSO-BP Neural Network, *Energies*, **16**, Article 4282. <https://doi.org/10.3390/en16114282>
- [12] Wang, C.-H., Yuan, J.C., Zeng, Y.P. and Lin, S.M. (2024) A Deep Learning Integrated Framework for Predicting Stock Index Price and Fluctuation via Singular Spectrum Analysis and Particle Swarm Optimization. *Applied Intelli-*

gence, **54**, 1770-1797. <https://doi.org/10.1007/s10489-024-05271-x>

- [13] 颜轲越, 李莹, 王祎萌. 基于信号分解和机器学习模型的股票价格预测[J]. 计算机科学与应用, 2022, 12(4): 1080-1088.