

基于误差修正和CEEMDAN-IGWO-ELM的股票价格预测建模

谢汇钦, 颜七笙*

东华理工大学理学院, 江西 南昌

收稿日期: 2024年4月29日; 录用日期: 2024年5月22日; 发布日期: 2024年5月31日

摘要

针对股票价格非平稳、非线性等特性引发预测精度低的问题, 引入Halton Sequence搜索算法、莱维飞行与等级制度策略对灰狼优化算法(GWO)进行改进, 提出一种基于误差修正和CEEMDAN-IGWO-ELM股票价格预测模型。首先将股票交易数据进行归一化处理, 作为极限学习机(ELM)的输入对股票价格进行预测得到初始预测结果, 进而得到误差序列。然后利用PE自适应地确定自适应噪声完备集合经验模态分解(CEEMDAN)的参数, 对误差序列进行分解, 利用IGWO算法优化ELM模型可调参数对每个子序列建模预测, 叠加各子序列预测结果对初始预测序列进行误差修正, 得到最终股票预测值。仿真实验与Diebold-Mariano检验结果表明, 与其他预测模型相比, 所建立模型具有更高的预测精度和优越性。

关键词

自适应噪声完备集合经验模态分解, 灰狼优化算法, 极限学习机, 误差修正, 股票价格预测

Stock Price Prediction Modeling Based on Error Correction and CEEMDAN-IGWO-ELM

Huiqin Xie, Qisheng Yan*

School of Science, East China University of Technology, Nanchang Jiangxi

Received: Apr. 29th, 2024; accepted: May 22nd, 2024; published: May 31st, 2024

Abstract

Aiming at the problem of low prediction accuracy caused by non-static and non-linear characteristics of stock price, the grey wolf optimization algorithm (GWO) is improved by introducing Halton

*通讯作者。

Sequence search algorithm, Levy flight and hierarchy strategy. A stock price prediction model based on error correction and CEEMDAN-IGWO-ELM was proposed. Firstly, the stock transaction data was normalized as input for extreme learning machine (ELM) to predict the stock price to obtain the initial prediction result and obtain an error sequence. Then, PE is used to adaptively determine the parameters of complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and decompose the error sequence. The IGWO algorithm is used to optimize the adjustable parameters of ELM model and predict each subsequence. After stacking the prediction results of each, the Error subsequences obtain the final stock prediction result. The experimental results and Diebold-Mariano test show that compared with other prediction models, this model has better prediction accuracy and superiority.

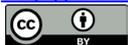
Keywords

Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), Grey Wolf Optimization (GWO), Extreme Learning Machine (ELM), Error Correction, Stock Price Prediction

Copyright © 2024 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

股票市场是宏观经济的“晴雨表”，反映当前的市场经济状况。投资者希望通过预测方法能够准确的预测股票价格的波动趋势，为投资者提供参考，做出最佳决策，避免较大的投资风险和损失的出现。但股票数据具有非线性、不平稳的特点，影响因素众多，且影响机制复杂，使得股票价格的预测更加困难。

目前，股票预测方法主要分两类：一类是传统的统计学方法，通过数学模型对股票价格进行预测；另一类是机器学习预测模型。传统时间序列预测方法通过线性模型的拟合对股票价格进行预测，如隐马尔可夫模型[1]、自回归滑动平均模型混合模型(ARMA) [2]、ARIMA [3]等方法。随着大数据和人工智能技术的发展，机器学习模型逐步运用在股票预测之中，如随机森林(RF) [4]、支持向量机(SVM) [5]、神经网络[6] [7]等模型，相较于传统的统计学模型，机器学习模型在股票预测方面有着更好的预测精度与泛化性能[8]。极限学习机(ELM)算法是以神经网络为基础的一种机器学习方法，其学习速度快，训练时间短，泛化能力强的优点，在交通预测[9]、卫星导航[10]、疾病预测[11]等许多方面预测效果较好。张文煜等[12]通过灰狼算法和 ELM 构建风速多步预测模型，对复杂的风速序列进行预测取得较好的预测结果。武博[13]构建基于 LSTM 神经网络的股票价格预测模型，与传统时间序列预测模型 ARIMA 相对比，LSTM 对非线性问题的预测性能更好，预测结果也是最优。同时学者对单一预测模型进行组合研究，韩莹等[14]针对 LSTM 易陷入局部最优问题，提出宽度学习与深度学习进行组合建立新预测模型，通过实验证明了该模型在多个精度指标有明显提升。朱菊香等[15]利用卷积神经网络(CNN)提取数据特征输入到 LSTM 网络，CEEMDAN 分解 PM2.5 序列，其模型能够准确预测 PM2.5 的浓度。通过实验表明，选择合适的模型进行组合，组合模型有着更好的预测精度和准确性[16] [17] [18]。

通过智能优化算法优化机器学习模型的参数可以提升模型性能。灰狼优化算法(GWO) [19]是基于自然界中灰狼群的一种仿生智能优化算法，通过狼群的等级制度、群体协作捕食的特性，模拟灰狼追踪、包围，攻击猎物等行为达到优化求解的目的。陈颖等[20]提出灰狼优化算法(GWO)优化支持向量机(SVR)校正预测模型，与其他模型对比，GWO-SVR 收敛速度更快，预测误差更小。zhang 等[21]通过 SSA 确定

LSTM 模型参数。试验结果表明 SSA-LSTM 具有更高的预测精度。罗聪等[22]利用反向初始化、非线性收敛因子、权重更新策略改进 GWO 对最小二乘支持向量机(LSSVM)参数惩罚因子和和函数宽度进行寻优, 实验结果验证该模型的优越性与稳定性。

一些学者针对灰狼算法的收敛速度慢、容易陷入局部最优解等问题, 结合其他优化算法的优势进行算法改进。Daniel Ocran 等[23]在 2022 年提出一种融合 HGWO 与 PSO 的鲁棒混合启发式算法, 结合 PSO 与 GWO 的优势, 在性能测试中表现出优良的勘探和开发能力, 其算法性能优于 PSO、GWO 和传统遗传算法(GA)。Bikash Kumar Saw 等[24]采用混合灰狼粒子群算法(GWO-PSO)对多个可重构性分布式电源进行并行规划, 在 IEEE33 和 69 节点 RDN 上进行算例分析对比自适应粒子群算法(PSO)取得更优的性能。Masume Khodsuz 等[25]研究针对棒状和对极状电极对避雷器(SA)动态模型的影响, 采用 PSO-GWO 混合算法调整 SA 动态模型参数, 实验表明 PSO-GWO 混合算法的优化性能优于其他优化算法, 提供更加准确的实验结果。

CEEMDAN 算法能够解决模态混叠和噪声残留问题, 结合改进 GWO 收敛速度快、搜索能力强以及 ELM 方法, 本文提出一种基于误差修正和 CEEMDAN-IGWO-ELM 股票价格预测模型, 引入 *Halton Sequence* 搜索算法、莱维飞行与等级制度策略对 GWO 进行改进。首先对中国石油股票交易数据进行归一化处理, 利用 ELM 对股票价格进行预测得到初始预测结果, 与真实股票价格作差得到误差序列, 然后利用 PE 自适应地确定 CEEMDAN 的参数, 再对误差序列进行分解, 运用 IGWO 优化 ELM 的参数预测误差子序列, 将各子序列的预测结果叠加后对初始预测结果进行修正, 最终得到股票预测结果。仿真实验结果和 DM 检验结果表明, 与其他模型相对比, 本文所建立的模型在股票预测方面有着较高的预测精度和一定的优越性。

2. 研究方法

2.1. 自适应噪声完备集合经验模态分解(CEEMDAN)

经验模态分解(Empirical Mode Decomposition, EMD) [26]是一种针对非线性、非平稳复杂时间序列的分解算法, 但存在模态混叠问题。针对此问题, 不少学者在 EMD 算法基础上提出几种不同的改进, WU [27]等提出集成经验模态分解(EEMD), 引入多个高斯白噪声对原始序列叠加进行 EMD 分解, 对并行 IMF 分量求平均最终得到 IMF 分量, EEMD 在一定程度上抑制了模态混叠问题的发生, 但加入的白噪声影响不能消除达到忽略不计的程度。互补集成经验模态分解(Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, CEEMD)引入互为相反数的正负白噪声[28], 与 EEMD 分解达到相同精度的迭代次数大大减少, 但在重构信号中仍然存在添加的白噪声的影响。Torres 等人提出自适应噪声完备集合经验模态分解(CEEMDAN) [29], 通过加入自适应高斯白噪声, 解决 EMD 和 EEMD 算法存在的模态混叠和噪声残留问题。CEEMDAN 具体分解步骤如下。

步骤 1: 在原始信号 $x(t)$ 添加 N 次满足正态分布的白噪声 $n_i(t)$, 得到 N 个预处理序列 $e_i(t)$ ($i = 1, 2, \dots, N$)。

$$e_i(t) = x(t) + \varepsilon_0 n_i(t) \quad (1)$$

式中 ε_0 是高斯白噪声的权值系数。

步骤 2: 对 N 个预处理序列 $e_i(t)$ ($i = 1, 2, \dots, N$) 进行 EMD 分解得到的第一模态分量求均值, 得到第一个 IMF 分量, IMF_1^i 为 $e_i(t)$ 序列进行 EMD 分解得到的第一模态分量。

$$IMF_1 = \frac{1}{N} \sum_{i=1}^N IMF_1^i \quad (2)$$

$$r_1 = x(t) - IMF_1 \quad (3)$$

步骤 3: 定义 $E_k(\cdot)$ 为对序列进行 EMD 分解得到的第 k 个模态分量, 对残余分量 r_1 添加白噪声, 得

到 N 个新序列 $r_i(t) = r_1(t) + \varepsilon_1 E_1(n_i(t))$, 再次进行 EMD 分解, 求均值得到第二个 IMF 分量。

$$IMF_2 = \frac{1}{N} \sum_{i=1}^N E_1(r_1(t) + \varepsilon_1 E_1(n_i(t))) \quad (4)$$

$$r_2 = r_1(t) - IMF_2 \quad (5)$$

步骤 4: 重复步骤 3, 得到 k 个 IMF 模态分量和第 k 个残余分量。

$$IMF_k = \frac{1}{N} \sum_{i=1}^N E_1(r_{k-1}(t) + \varepsilon_{k-1} E_{k-1}(n_i(t))) \quad (6)$$

$$r_k = r_{k-1}(t) - IMF_k \quad (7)$$

直到残余分量为单调函数, 则停止分解。即原信号 $x(t)$ 可以表示为:

$$x(t) = \sum_{i=1}^k IMF_i + R(t) \quad (8)$$

式中: IMF_i 表示第 i 个模态分量, $R(t)$ 是最终的残差分量。

2.2. 灰狼优化算法(GWO)

灰狼在自然界中是一种群居性动物, 在狼群中有着严格的等级制度, 不同等级的灰狼在群体中分工明确。自上到下: 第一层 α 狼, 属于狼群的头狼, 主要负责狼群的狩猎、休息等决策; 第二层是 β 狼, 与 α 狼是从属关系, 帮助 α 狼制定决策或其他活动, 当 α 狼空缺时, β 狼代替狼成为新的狼; 第三层是 δ 狼, 听从 α 狼和 β 狼以及管理着 δ 狼; 第四层狼是 ω 狼, 处于狼群社会的最底层, 主要负责执行以上三层狼的决策及保持狼群内部的平衡。

在 GWO 算法中, 我们将其 α 狼比作是最优解, β 狼是次优解, 第三优是狼 δ , ω 狼是候选解, 在迭代过程中, 通过 α 狼, β 狼和 δ 狼的位置引导 ω 狼的位置进行更新, 使得狼群追踪、靠近并围捕猎物完成狩猎行为。

灰狼在追踪行为过程中位置利用以下公式进行更新

$$\bar{D} = |\bar{C} \cdot \bar{X}_p(t) - \bar{X}(t)| \quad (9)$$

$$\bar{X}(t+1) = \bar{X}_p(t) - \bar{A} \cdot \bar{D}. \quad (10)$$

式中: \bar{D} 表示灰狼个体与猎物之间的距离; t 表示当前迭代次数; \bar{X}_p 表示为猎物的位置向量; \bar{X} 表示为灰狼当前的位置向量; \bar{A} 与 \bar{C} 是协同系数向量, 其计算公式如下:

$$A = 2a \cdot r_1 - a, C = 2 \cdot r_2 \quad (11)$$

$$a = 2 - 2 \cdot t/T \quad (12)$$

式中: a 为收敛因子, 伴随着迭代次数增加从 2 到 0 线性减小, t 和 T 分别是当前迭代次数和最大迭代次数; r_1 与 r_2 是两个每一维的取值都在 $[0, 1]$ 中的随机向量。

灰狼围捕目标猎物行为的公式如下:

$$\begin{cases} \bar{D}_\alpha = |\bar{C}_1 \cdot \bar{X}_\alpha - \bar{X}|, \\ \bar{D}_\beta = |\bar{C}_2 \cdot \bar{X}_\beta - \bar{X}|, \\ \bar{D}_\delta = |\bar{C}_3 \cdot \bar{X}_\delta - \bar{X}|. \end{cases} \quad (13)$$

式中: \overline{D}_α 、 \overline{D}_β 、 \overline{D}_δ 分别代表的是灰狼 α 、 β 、 δ 与其他灰狼个体的距离; \overline{X}_α 、 \overline{X}_β 、 \overline{X}_δ 分别是 α 、 β 、 δ 狼各自的位置; \overline{C}_1 、 \overline{C}_2 、 \overline{C}_3 都为随机向量; \overline{X} 表示灰狼的位置

$$\begin{cases} \overline{X}_1 = \overline{X}_\alpha - \overline{A}_1 \cdot \overline{D}_\alpha \\ \overline{X}_2 = \overline{X}_\beta - \overline{A}_2 \cdot \overline{D}_\beta \\ \overline{X}_3 = \overline{X}_\delta - \overline{A}_3 \cdot \overline{D}_\delta \end{cases} \quad (14)$$

$$\overline{X}(t+1) = \frac{\overline{X}_1 + \overline{X}_2 + \overline{X}_3}{3} \quad (15)$$

式中: \overline{X}_1 表示的是 ω 狼向 α 狼靠近的步长与方向; \overline{X}_2 表示的是 ω 狼向 β 狼靠近的步长与方向; \overline{X}_3 表示的是 ω 狼向 δ 狼靠近的步长与方向, 由此最终确定 ω 狼的位置。

2.3. 极限学习机(ELM)

极限学习机与传统神经网络算法不同, 采用随机生成输入层权值与隐藏层的阈值, 确定隐藏层神经元个数即可建立极限学习机模型, 具有训练参数少、学习速度快等优点。训练样本为:

$R = \{(x_j, y_j) | j = 1, 2, \dots, N; x_j \in R^N, y_j \in R^M\}$, 则一个隐藏层的极限学习机算法表示为:

$$y_j = \sum_{i=1}^L \beta_i g(\omega_i x_j + b_i), j = 1, 2, \dots, N \quad (16)$$

式中: β_i 为输出权值; ω_i 为输入权值; b_i 为隐藏层神经元阈值; $g(\cdot)$ 为隐藏层的激活函数; x_j 为输入值, N 为样本数, L 为隐藏层神经元个数。

可用矩阵形式表示为:

$$T = H \cdot \beta \quad (17)$$

H 为隐藏层的输出矩阵; β 为输出权值矩阵。

通过将样本数据训练集代入, 计算得到输出权值矩阵:

$$H = \begin{bmatrix} g(\omega_1 x_1 + b_1) & g(\omega_2 x_1 + b_2) & \cdots & g(\omega_L x_1 + b_L) \\ g(\omega_1 x_2 + b_1) & g(\omega_2 x_2 + b_2) & \cdots & g(\omega_L x_2 + b_L) \\ \vdots & \vdots & \vdots & \vdots \\ g(\omega_1 x_N + b_1) & g(\omega_2 x_N + b_2) & \cdots & g(\omega_L x_N + b_L) \end{bmatrix} \quad (18)$$

输出层权值 β 计算公式如下:

$$\beta = H^* \cdot T \quad (19)$$

求解到的 β 是最小且唯一。

H^* 是 H 的 Moore-Penrose 广义逆。

再将样本测试集数据代入 ELM 模型计算预测结果, 即为:

$$Y = H \cdot \beta \quad (20)$$

2.4. 排列熵

排列熵(Permutation Entropy, PE)用来衡量时间序列的复杂程度, PE 越大, 说明该时间序列越复杂; PE 越小, 说明该时间序列越规律。通过分解得到的子序列分量, 计算子序列的排列熵反映分量的复杂程度, 排列熵计算原理如下所述。

对于一维时间序列 $\{X(i), i=1, 2, \dots, N\}$ 进行相空间重构, 得到重构矩阵

$$\begin{pmatrix} x(1) & x(1+\tau) & \cdots & x(1+(m-1)\tau) \\ x(2) & x(2+\tau) & \cdots & x(2+(m-1)\tau) \\ \vdots & \vdots & & \vdots \\ x(K) & x(K+\tau) & \cdots & x(K+(m-1)\tau) \end{pmatrix} \quad (21)$$

式中: K 表示重构向量个数; m 为嵌入维数; τ 为延迟时间。重构矩阵中每一行 $\{x(j), x(j+\tau), \dots, x(j+(m-1)\tau)\}$ 代表一个重构分量, 然后按照升序方式对每个重构分量中每个元素进行重排可得: $x(i+(j_1-1)\tau) \leq \dots \leq x(i+(j_m-1)\tau)$, 式中 j_1, j_2, \dots, j_m 表示重构元素所在的列。对于每一个时间序列 $X(i)$ 进行重构得到重构矩阵的每一行都可以用一组符号序列。

$$S(l) = (j_1, j_2, \dots, j_m) \quad (22)$$

式中 $l=1, 2, \dots, k$, 且 $k \leq m!$, m 维相空间映射可得到 $m!$ 不同的符号序列 $\{j_1, j_2, \dots, j_m\}$, $S(l)$ 只是其中一种。若 k 种不同的符号序列出现的概率分别为 P_1, P_2, \dots, P_k , 则排列熵表示为:

$$H_{PE}(m) = -\sum_{j=1}^k P_j \ln P_j \quad (23)$$

将其进行归一化处理, 得: $0 \leq (H_{PE} = H_{PE}/(\ln m!)) \leq 1$, 式中 H_{PE} 为归一化后的排列熵。

3. 改进灰狼优化算法

为了增强优化算法的效果, 使得模型的预测效果更好, 提高预测精度, 提出改进灰狼优化算法 (Improve GreyWorf Optimization, IGWO), 增强算法的全局搜索能力, 主要对以下三方面进行改进。

3.1. Halton Sequence 搜索算法

灰狼优化算法是采用随机初始化的方式构建种群初始位置信息, 这样所得到的种群位置随机性较高, 有可能无法均匀分布在整个解空间, 导致算法陷入局部最优解和重复运算。针对上述问题, 本文引入 *Halton Sequence* 搜索算法产生随机数进行对狼群的位置初始化, 步骤如下:

首先设定一个质数 p 为基数, 则对于任意整数 $m, m \in (1, n)$, 其都可以表示以质数 p 为基的形式, 其表示形式为:

$$m = \sum_{i=0}^n b_i p^i = b_n p^n + b_{n-1} p^{n-1} + \dots + b_1 p^1 + b_0 \quad (24)$$

基于公式(24), 得到基本反序函数 $\theta_p(n)$, 其表示形式为:

$$\theta_p(n) = b_1 p^{-1} + b_1 p^{-2} + \dots + b_m p^{-m-1} \quad (25)$$

则对于任意整数 m 和维度 $d \in [1, D]$, 选择质数表中的质数 $[p_1, p_2, \dots, p_D]$, 可得到 *Halton Sequence* 初始化灰狼群, 其如下公式表示为:

$$H(i) = [\theta_{p_1}(i), \theta_{p_2}(i), \dots, \theta_{p_D}(i)] \quad (26)$$

分别对 *Halton Sequence* 初始化和随机初始化进行仿真实验, 设定灰狼种群数为 400, 解的维度为二维, 得到图 1 灰狼种群分布散点图, 通过观察两个种群分布图, 可以得知 *Halton Sequence* 初始化灰狼种群分布在空间中更加均匀, 且没有出现灰狼个体位置重叠现象; 随机灰狼群初始化种群分布会出现灰狼位置重叠现象, 相较 *Halton Sequence* 初始化会出现位置空缺现象, 导致算法重复计算或搜索不全。

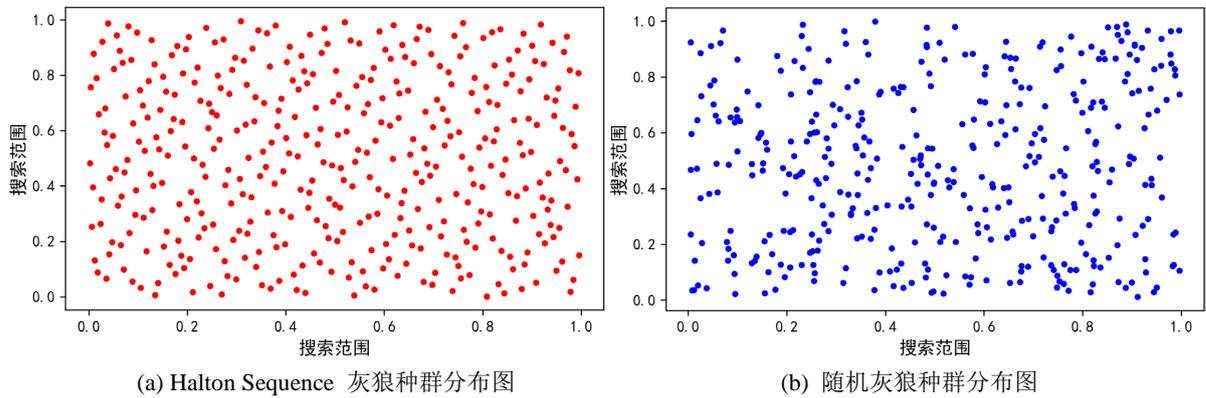


Figure 1. Comparison of Random Population Distribution and Halton Sequence Initialized Population Distribution
图 1. 随机种群分布与 Halton Sequence 初始化种群分布对比

3.2. 莱维飞行

为了提高优化精度, 利用 Levy 飞行对灰狼算法的寻优过程进行优化[30] [31], 其更新公式如下:

$$x'_i(t) = x_i(t) + l \oplus Levy(\lambda) \tag{27}$$

式中: $x'_i(t)$ 表示是迭代第 t 次的第 i 个解; \oplus 表示点乘; l 表示是控制步长的权重, $l = 0.01(x_i(t) - x_b(t))$, 其中 $x_b(t)$ 表示是当前迭代次数灰狼位置中的最优解; $Levy(\lambda)$ 表示是服从莱维分布的路径, 且满足: $Levy \sim u = t^{-\lambda}$, $1 \leq \lambda \leq 3$ 。由于莱维分布较为复杂, 通常使用 Mantegna 算法进行模拟计算, 步长计算公式为:

$$s = \frac{u}{|v|^{1/\gamma}} \tag{28}$$

式中, u 和 v 服从正态分布, $u \sim N(0, \delta_u^2)$, $v \sim N(0, \delta_v^2)$, $\delta_v = 1$, $\delta_u = \left\{ \frac{\Gamma(1+\gamma)\sin(\pi\gamma/2)}{\gamma \cdot \Gamma[(\gamma+1)/2] \cdot 2^{(\gamma-1)/2}} \right\}^{1/\gamma}$, 式中 γ 取值范围为 $[0, 2]$, 这里实验取值为 1.5。

对 Levy 飞行进行仿真实验, 仿真实验步长为 1000, 图 2 为 Levy 飞行仿真实验结果, Levy 飞行特殊的行走方式, 偶尔飞行步长较大, 具有更大的搜索范围, 有利于提高优化算法的全局搜索能力。

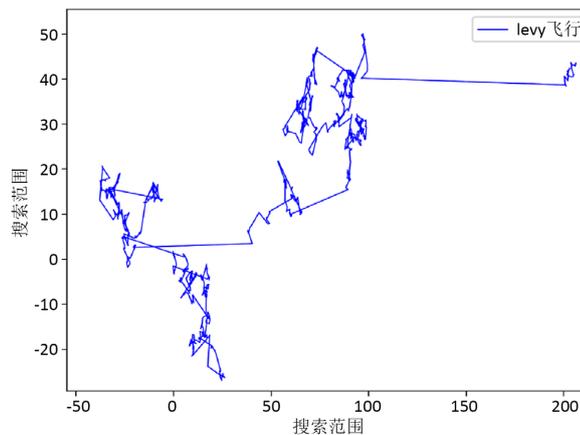


Figure 2. Simulation experiment results of Levy flight
图 2. 莱维飞行仿真实验结果

基于此仿真实验, 对灰狼优化算法的种群更新方式进行改进, 公式(3)为改进后 α 狼位置更新策略, β 狼与 δ 狼更新策略不变。

$$\overline{X}_1' = Levy(\overline{X}_\alpha + \overline{A}_1 \cdot \overline{D}_\alpha) - \overline{A}_1 \cdot \overline{D}_\alpha \tag{29}$$

$$\overline{X}_2 = \overline{X}_\beta - \overline{A}_2 \cdot \overline{D}_\beta \tag{30}$$

$$\overline{X}_3 = \overline{X}_\delta - \overline{A}_3 \cdot \overline{D}_\delta \tag{31}$$

3.3. 等级制度

灰狼群中有着等级制度, 但是在算法过程中并没有很好的体现这一方面, 为了更好的彰显出灰狼群中三只头狼的等级制度, 提出灰狼算法的等级制度改进, 其本质是在灰狼群更新位置时, 在三只头狼的位置前赋予一个等级权重, 分别为: η_1, η_2, η_3 , 其权重占比为: $\eta_1 : \eta_2 : \eta_3 = 4 : 3 : 2$ 。

3.4. 改进 GWO 算法流程

通过上述对灰狼优化算法进行改进, 改进后灰狼优化算法具体步骤如下:

- 1) 采用 *Halton Sequence* 搜索算法初始化灰狼种群位置。
- 2) 计算灰狼群中每只灰狼个体的适应度值。通过计算灰狼群的适应度值, 选择表现好的三只灰狼记做 $X_\alpha, X_\beta, X_\delta$, 指引其余灰狼个体更新搜索位置, 其中 α 狼位置通过公式(29)进行更新。
- 3) 计算更新后灰狼群个体适应度值, 比较新最优灰狼个体适应度值与上一代灰狼群最优灰狼个体的适应度值, 若新最优个体适应度值优于上一代个体最优适应度值则替换最优个体适应度值, 反之保持不变。
- 4) 若迭代次数达到最大迭代次数, 进行步骤 5, 反则返回步骤 2。
- 5) 将最优灰狼个体位置信息和适应度值输出。

4. IGWO 性能测试

4.1. 测试函数及参数设置

为验证 IGWO 算法的有效性, 本文选取了 4 个标准测试函数进行仿真实验, 通过与灰狼优化算法(GWO)、粒子群算法(PSO)进行对比。算法实验平台均为 Windows10、24 GB 内存、64 bit 系统中, 采用软件 python 3.10 (64-bit)进行。表 1 为标准测试函数的名称、表达式、维度、搜索范围和极值, 其中测试函数 1 为标准单峰测试函数, 可用于测试算法的收敛能力与求解精度; 测试函数 2~3 为标准多峰测试函数, 用于测试算法是否能够跳出局部极值搜索到全局极值; 测试函数 4 为固定维数的多峰测试函数。实验设置算法种群规模为 20, 最大迭代次数设为 300。

Table 1. Standard unimodal test function

表 1. 标准测试函数

函数名称	表达式	维度	搜索范围	极值
Sphere	$F_1(x) = \sum_{i=1}^n x_i^2$	30	[-100, 100]	0
Rastrigin	$F_2(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	[-5.12, 5.12]	0
Griewank	$F_3(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[-600, 600]	0
Shekel	$F_4(x) = -\sum_{i=1}^7 \left[(x - a_i)(x - a_i)^T + c_i \right]^{-1}$	4	[0, 10]	-10.4028

4.2. 实验结果对比分析

为了避免实验结果的偶然性, 提高实验的准确性, 将三种算法分别对每个基准测试函数独立运行 30 次, 最后得到实验结果数据: 平均值、标准差、最优值。其中最优结果用粗体标出, 实验结果数据如表 2 所示。其中标准差对比算法的稳定性; 平均值衡量算法的求解精度; 最优值体现算法的寻优效果。

由表 2 可知, 在单峰测试函数 F_1 , IGWO 算法对比 GWO 和 PSO, 寻优效果较好, PSO 寻优结果较

Table 2. Algorithm optimization results
表 2. 算法寻优结果

函数	GWO			PSO			IGWO		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
F_1	6.72E-38	3.23E-37	9.44E-44	1.71E+00	3.79E+01	1.08E+00	7.23E-246	0	1.78E-261
F_2	0	0	0	1.98E+02	1.06E+01	1.73E+02	0	0	0
F_3	0	0	0	1.78E+02	2.16E+01	1.26E+02	0	0	0
F_4	-4.1784	1.3002	-8.1453	-8.6313	3.0334	-10.3994	-8.2967	2.7426	-10.4016069

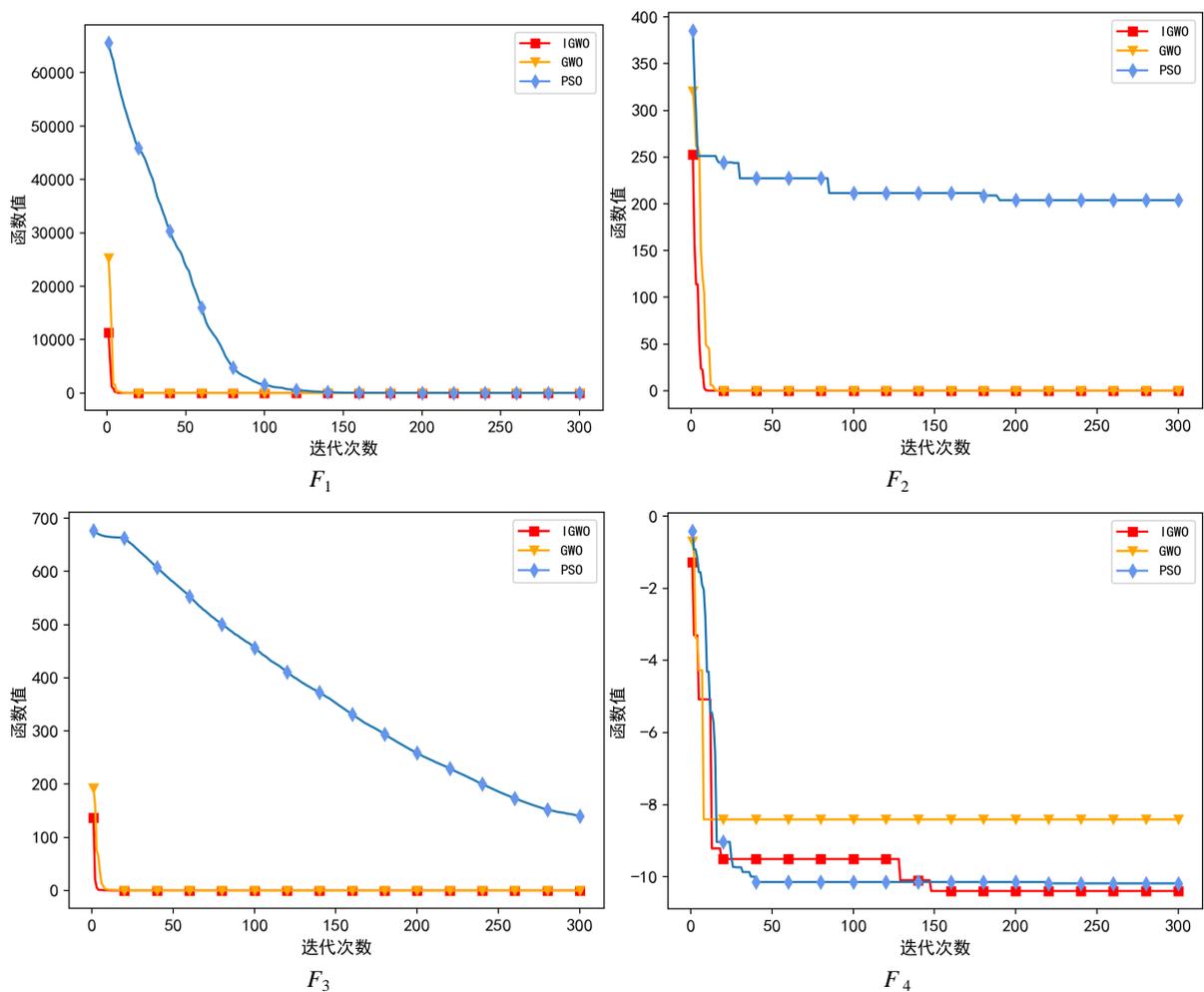


Figure 3. Benchmark function iteration curve
图 3. 准测试函数迭代曲线

差, 且 IGWO 对比 GWO 算法搜索精度更高; 在多峰测试函数 F_2 、 F_3 中, GWO、IGWO 均能寻优到最优值, 且 IGWO 和 GWO 对比 PSO 更加稳定; 在固定维数的多峰测试函数 F_4 中, GWO 算法搜索精度较差, PSO 与 IGWO 寻优效果优于 GWO。

为更好的体现算法的效果, 图 3 展示标准测试函数的迭代曲线, 从图 3 中可以看出, IGWO 算法具有较好的收敛精度和较好的收敛速度, 在单峰测试函数 F_1 中, IGWO、GWO 收敛速度要快于 PSO, 结合表 2, IGWO 算法的收敛精度要高于 GWO, 表明改进后 GWO 算法提高了收敛精度; 在多峰测试函数 F_2 、 F_3 中, PSO 算法的寻优效果明显要差于 GWO、IGWO, 且在函数 F_2 中 PSO 算法陷入局部最优值; 在固定维数的多峰测试函数 F_4 中, IGWO、PSO, GWO 均陷入局部最优值, 最终 IGWO 跳出局部最优值, 寻优到全局最优值, 表明 IGWO 具有更好的全局搜索能力。

综上所述, 通过表 2 和图 3 的寻优效果可知, 本文所建立了 IGWO 算法在收敛速度和收敛精度均要优于其他对比算法, 改进后的 IGWO 全局搜索能力更强, 寻优效果更好。

5. 模型建立

5.1. 实验数据

本文选用中国石油股票 2010 年 11 月 1 日至 2021 年 1 月 29 日作为股票数据对象, 其中包含收盘价、最高价、最低价、开盘价、成交金额, 次日收盘价股票数据, 共 2493 条股票样本数据, 具体数据来源于文献[32]。

5.2. 实验模型

建立 Error-CEEMDAN-IGWO-ELM 模型, 首先对中国石油股票交易数据进行归一化处理后通过极限学习机(ELM)对次日股票收盘价进行预测, 得到初始预测结果, 与真实值作差得到误差序列, PE 自适应确定 CEEMDAN 参数, 对误差序列进行分解, 通过 IGWO 对 ELM 参数寻优后对误差子序列进行预测, 再对预测后子序列相加重构即为误差预测值, 对初始预测结果进行修正, 得到最终股票预测结果。具体建模步骤如下:

1) 对中国石油股票交易数据进行归一化处理后作为变量输入 ELM 模型, 次日收盘价作为输出, 得到初始预测序列。其中 BP 神经网络、ELM 模型实验均将前 1993 条股票数据样本作为训练集, 后 500 条股票数据样本作为测试集。

2) 将 ELM 模型预测结果与真实次日收盘价数据作差得到误差序列。

3) 利用 CEEMDAN 对误差序列进行分解得到若干个模态分量和一个残余分量, CEEMDAN 参数由 PE 自适应确定, 然后进行交叉分组, 逐一对分组后的误差子序列作为 IGWO-ELM 模型的输入进行预测, 其中交叉分组是指一个模态分量中 500 个样本数据均分为 5 组, 每组 100 个样本数据, 依次将这 100 个样本数据作为样本数据测试集, 其余为样本数据训练集。经过交叉分组预测后的预测分量进行相加重构即为误差预测序列。

4) 将得到的误差预测序列用于对 ELM 初始预测结果进行修正, 得到最终股票预测结果。图 4 为误差修正和 CEEMDAN-IGWO-ELM 预测模型流程图。

5.3. 数据预处理

5.3.1. 归一化

中国石油股票数据中收盘价、最高价、最低价、开盘价、成交金额五种数据的单位与数量级有差别, 对股票数据进行归一化处理。

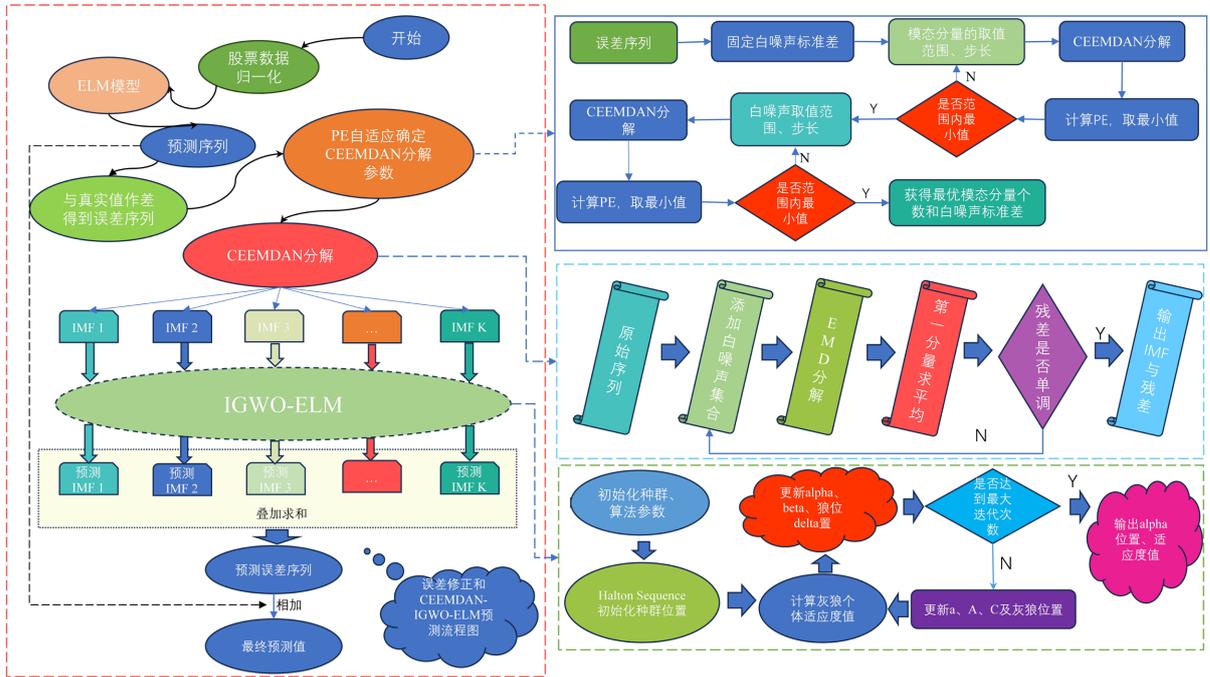


Figure 4. Flow chart of prediction model based on error correction and CEEMDAN-IGWO-ELM
 图 4. 基于误差修正和 CEEMDAN-IGWO-ELM 预测模型流程图

$$x = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (32)$$

式中: x 为输入数据, x_{\max} 、 x_{\min} 为输入数据的最大值与最小值。

5.3.2. PE 确定 CEEMDAN 参数

本文选用 CEEMDAN 分解算法对误差序列进行分解, 在 CEEMDAN 分解算法中, 白噪声标准差与总集合次数的范围分别为(0.01~0.5)和(100~300), 分解效果较好。实验中 CEEMADN 加入的总集合次数为 100, 利用排列熵对 CEEMDAN 分解的模态分量个数 K 与白噪声进行择优选择, 首先固定 $\delta = 0.1$, 设

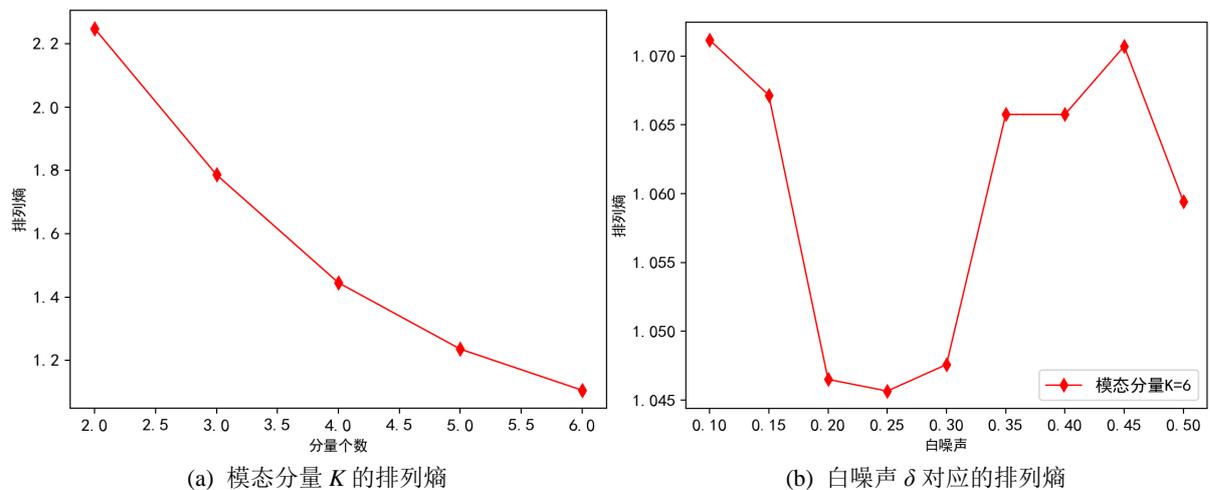


Figure 5. Permutation entropy of modal component K and white noise δ
 图 5. 模态分量 K 和白噪声 δ 的排列熵

置 K 的寻优范围为(2~6), 步长: 1, 由图 5(a)可知, 当 $K = 6$ 时排列熵最小, 则最优模态分量 $K = 6$, 其次, 固定模态分量 $K = 6$, 设置 δ 的寻优范围为(0.1~0.5), 步长为 0.05, 由图 5(b)可知, 当 $\delta = 0.25$ 时, 排列熵最小, 则最优白噪声标准差 $\delta = 0.25$ 。

通过 PE 自适应确定 CEEMDAN 分解模态分量个数与白噪声标准差后对误差序列进行 CEEMDAN 分解, 得到 6 个 IMF 分量和 1 个残差序列, CEEMDAN 分解结果如图 6 所示。经过 CEEMDAN 分解后的子序列更加简单, 更有利于模型的预测。

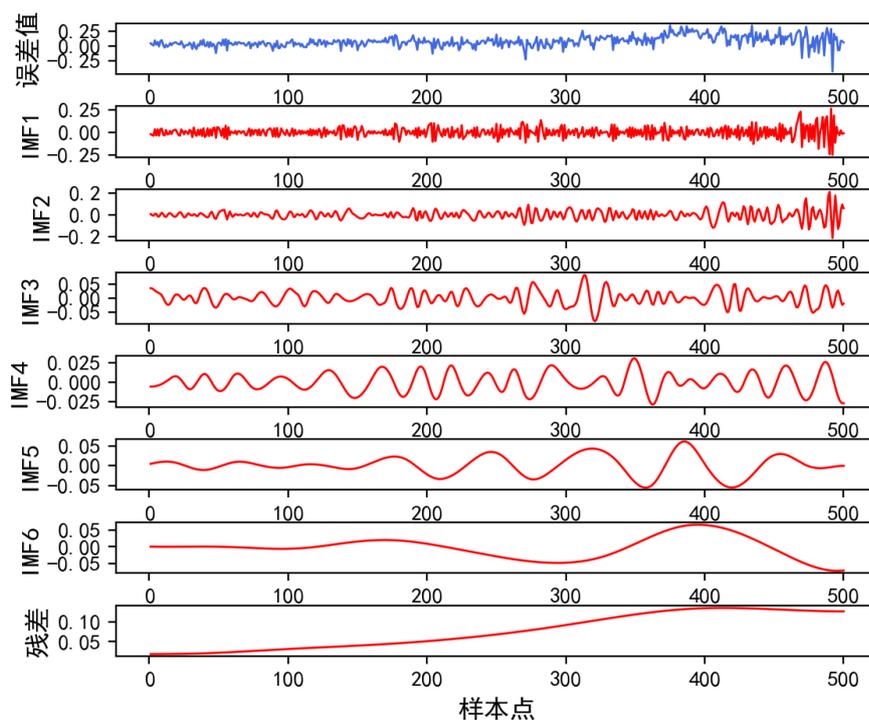


Figure 6. CEEMDAN decomposition results

图 6. CEEMDAN 分解结果

5.4. 评价指标标准

为了评价模型预测性能的效果, 本文采用平均绝对误差(MAE)、平均绝对百分比误差(MAPE)、均方误差(MSE)、均方根误差(RMSE)和拟合优度决定系数(R^2)这 4 个指标作为评价的标准, 其具体计算公式如下:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - y'_i| \quad (32)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - y'_i}{y_i} \right| \quad (34)$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2 \quad (35)$$

$$R^2 = \frac{\left[\sum_{i=1}^N (y_i - \bar{y})(y'_i - \bar{y}') \right]^2}{\left[\sum_{i=1}^N (y_i - \bar{y})^2 \right] \left[\sum_{i=1}^N (y'_i - \bar{y}')^2 \right]} \quad (36)$$

式中: N 为样本个数; y_i 为真实值; y'_i 为预测值; \bar{y} 为平均真实值; \bar{y}' 为平均预测值。

6. 实验结果与对比分析

6.1. 基于误差修正和 CEEMDAN-IGWO-ELM 模型的股票预测结果分析

在 Windows 10 Version 22H2 (x64 版) 环境中, 运用软件 python 3.10 (64-bit), 基于中国石油股票数据, 对股票次日收盘价进行预测, 并对比其他模型分析预测结果。其中 ELM 模型与 BP 神经网络的隐藏神经元个数均为 5, 隐藏层的权值与阈值取值范围均为 $[-1, 1]$, 隐藏层激活函数选用 sigmoid 函数, BP 神经网络更新次数为 100, 学习率为 0.3。GWO 与 PSO 部分参数设置如下: 种群大小为 20, 最大迭代次数为 300, PSO 学习因子 $c_1 = c_2 = 2$, PSO 惯性权重计算方式: $\omega = ws - (ws - we) * \frac{t}{T}$, 其中 $ws = 0.9$, $we = 0.4$, t, T 分别是当前迭代次数和最大迭代次数, 两种优化算法以均方根误差(RMSE)作为种群的适应度函数。

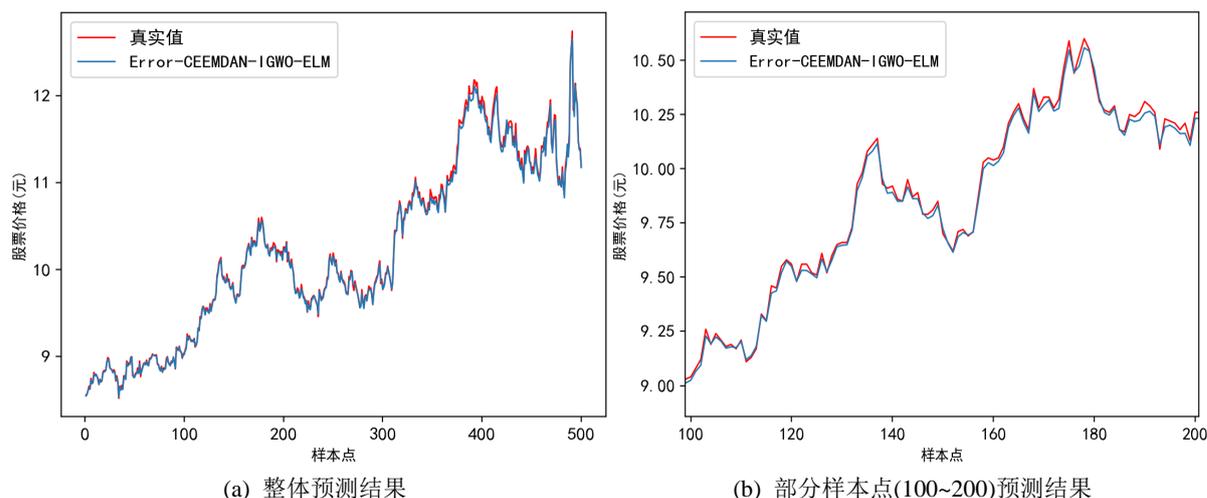


Figure 7. Error CEEMDAN-IGWO-ELM prediction results

图 7. Error-CEEMDAN-IGWO-ELM 预测结果

将 IGWO-ELM 预测得到的误差预测子序列进行相加重构, 得到的误差预测序列对 ELM 初始预测结果进行修正, 最终得到股票价格的预测结果(图 7)。图 7 中基于误差修正和 CEEMDAN-IGWO-ELM 模型的股票预测曲线与真实股票价格曲线拟合程度较高, 预测值与真实值较为接近, 表明 Error-CEEMDAN-IGWO-ELM 模型在股票预测较高的预测精度。

6.2. 模型对比分析

为了更好验证基于误差修正和 CEEMDAN-IGWO-ELM 模型的预测效果, 将该模型与 BP 神经网络、ELM、基于误差修正的 CEEMDAN-ELM (Error-CEEMDAN-ELM)、基于误差修正和 CEEMDAN-GWO-ELM (Error-CEEMDAN-GWO-ELM)、基于误差修正和 CEEMDAN-PSO-ELM (Error-CEEMDAN-PSO-ELM) 模型预测结果进行对比, 预测结果如下图所示。

由图 8 可知, 相较于 BP 神经网络, ELM 模型的预测结果更加接近股票真实值, 误差相较 BP 神经网络小, 预测趋势拟合程度较好。Error-CEEMDAN-ELM 模型与 ELM 模型预测结果对比如图 9 所示, 相较于 ELM 预测结果, Error-CEEMDAN-ELM 模型的预测结果有着不错的提升, 证明误差修正与 CEEMDAN 分解对 ELM 模型在股票预测方面可以提升模型的预测性能, 提高预测精度。

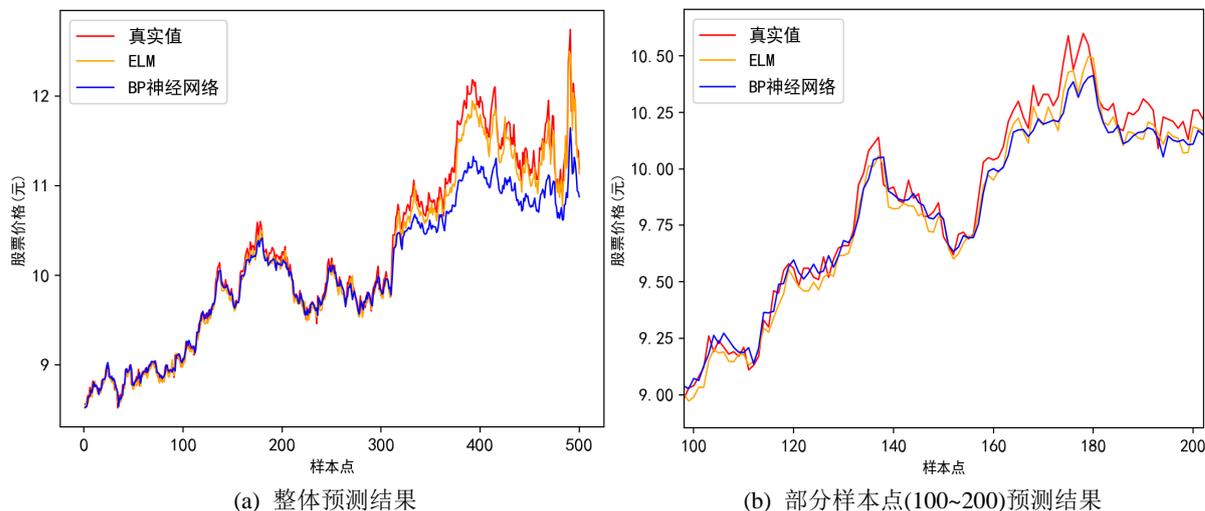


Figure 8. ELM and BP prediction results
图 8. ELM、BP 预测结果

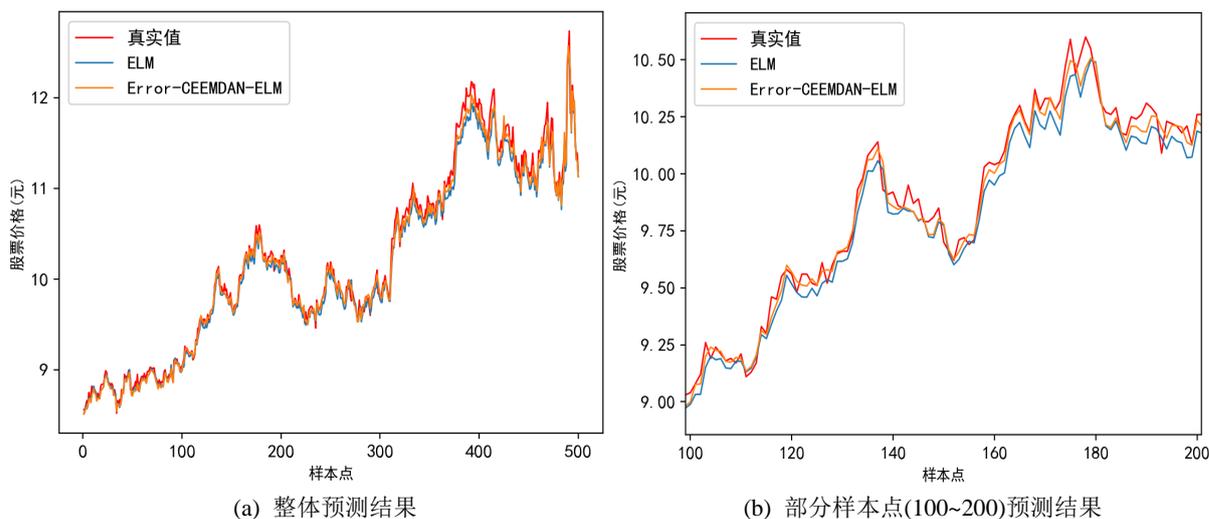


Figure 9. ELM and Error CEEMDAN-ELM prediction results
图 9. ELM 与 Error-CEEMDAN-ELM 预测结果

误差序列通过 CEEMDAN 分解后, 采用 PSO、GWO 优化算法对 ELM 模型参数寻优预测, 如图 10 所示。通过优化算法优化后的模型预测效果上相较于 Error-CEEMDAN-ELM 模型与真实值更加接近, 表明优化算法可以提升预测模型的性能, 更加准确的预测股票价格, 与真实价格曲线趋势更加吻合, 预测精度更高。

由图 11 可知, 通过 IGWO 优化预测模型参数和 GWO 优化预测模型在股票预测结果上有着相似的预测效果, 整体上预测曲线与真实价格曲线比较吻合, 在局部的预测结果上, IGWO 优化模型的预测结果要优于 GWO 优化模型的预测结果, 与真实股票价格更加接近, 误差更小, 说明 IGWO 相比 GWO 在模型优化上效果较好。

6.3. 各模型预测结果误差分析

各模型的预测结果进行误差分析, 模型评价指标如表 3 所示。

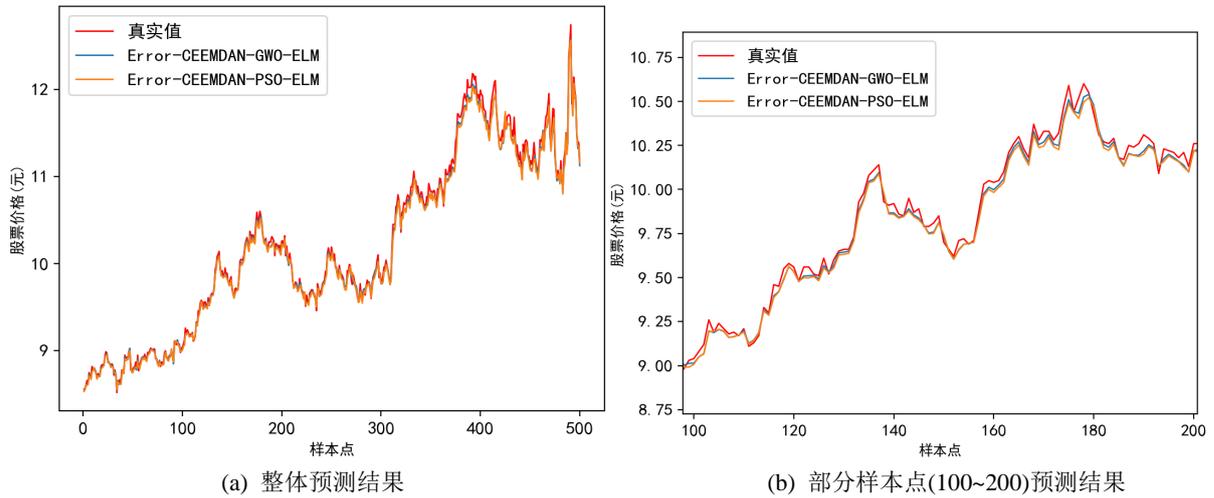


Figure 10. Prediction results of optimization algorithm
图 10. 优化算法预测结果

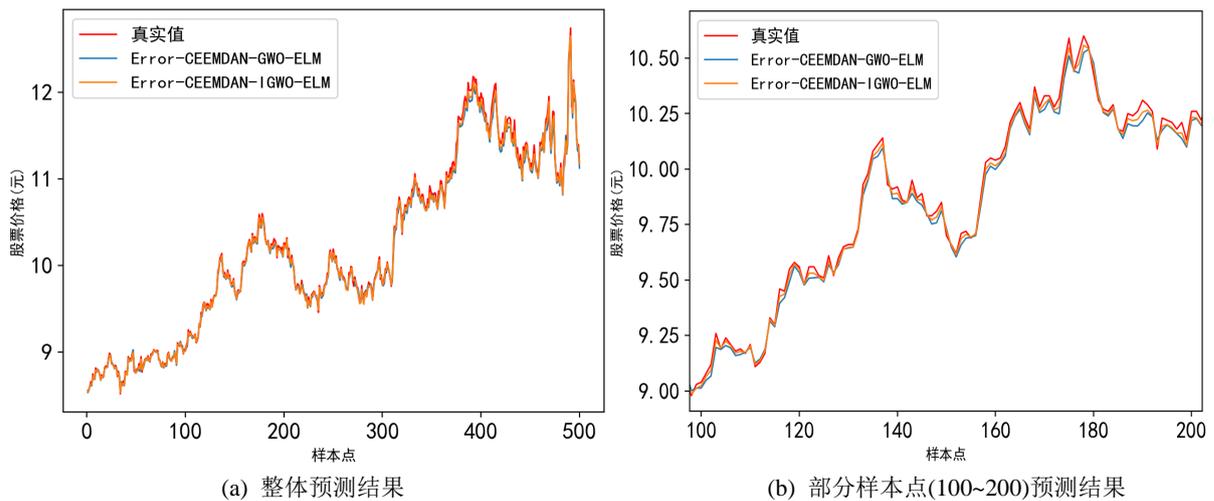


Figure 11. Optimization of prediction results using IGWO and GWO algorithms
图 11. IGWO 与 GWO 算法优化预测结果

Table 3. Evaluation Index Values for Each Model
表 3. 各模型评价指标值

模型	MAE	MAPE	MSE	R^2
BP 神经网络	0.1426	0.1168	0.0392	0.9504
ELM	0.1189	0.0149	0.0353	0.9702
Error-CEEMDAN-ELM	0.0701	0.0075	0.0132	0.9808
Error-CEEMDAN-PSO-ELM	0.0686	0.0073	0.0105	0.9901
Error-CEEMDAN-GWO-ELM	0.0582	0.0057	0.0085	0.9922
Error-CEEMDAN-IGWO-ELM	0.0437	0.0045	0.0033	0.9965

由表 3 可知, Error-CEEMDAN-IGWO-ELM 模型的 MAE、MAPE、MSE、 R^2 评价指标值分别为: 0.0437、0.0045、0.0033、0.9965, 对比其他模型的评价指标值, 均是最优, 由此说明基于误差修正和 CEEMDAN-

IGWO-ELM 模型在股票预测方面有着较高的预测精度。与基础的 ELM 模型相比较, 三种误差分别降低了 0.0752、0.0104 和 0.032, R^2 也从 0.9702 提升到 0.9965, 表明误差修正和 IGWO 优化 ELM 模型参数具有较好的效果; 经过 IGWO、GWO 与 PSO 优化算法对模型参数进行寻优后, 相较于未寻优模型的 MAE 分别降低了 0.0264、0.0119 和 0.0015, 说明寻优算法对模型参数进行寻优有一定效果, GWO 算法与 PSO 算法优化效果相比优化效果较好。Error-CEEMDAN-IGWO-ELM 模型的各项指标值与 Error-CEEMDAN-GWO-ELM 模型对比具有一定的优化, 表明 IGWO 算法具有一定的改进效果。文献[32]对中国石油股票价格预测, PSO-ELM 模型的 MAPE 为 0.0095, 对比本文提出的 Error-CEEMDAN-IGWO-ELM 模型对中国石油股票价格预测有着更小的预测误差。因此, 总体来说本文提出的 Error-CEEMDAN-IGWO-ELM 模型对于股票预测有着更高的预测精度。

6.4. DM 检验

为了进一步验证所建立的预测模型的优越性, 引入一种假设检验方法: DM 检验, 通过 DM 检验进行验证两个预测模型的性能是否在一定的显著性水平下存在显著性差异, 验证所建立的 Error-CEEMDAN-IGWO-ELM 预测模型与其他预测模型对比其模型预测性能, 表 4 给出基于 MAE 为损失函数的各模型 DM 检验结果, DM 检验结果显示, Error-CEEMDAN-IGWO-ELM 模型与其他模型具有显著性差异, 验证了所建立的 Error-CEEMDAN-IGWO-ELM 模型的优越性。

Table 4. Diebold-Marianotestresults
表 4. DM 检验结果

模型	DM
BP 神经网络	2.5314*
ELM	2.1832**
Error-CEEMDAN-ELM	1.8114*
Error-CEEMDAN-PSO-ELM	1.5978**
Error-CEEMDAN-GWO-ELM	1.4979***

注: *表示 10% 的显著性水平; **表示 5% 的显著性水平; ***表示 1% 的显著性水平。

7. 结论

股票价格的波动趋势和未来的走势预测一直是投资者关心的重点之一, 有效的预测股票未来的波动趋势对于投资者具备有一定的参考价值。本文提出一种基于误差修正和 CEEMDAN-IGWO-ELM 的股票预测模型, 为提高股票预测精度, 引入 Halton Sequence 搜索算法、莱维飞行、等级制度对灰狼优化算法 (GWO) 进行改进, 提高优化算法的效果, 通过 CEEMDAN 分解算法, IGWO 算法对 ELM 模型参数寻优以及误差修正, 通过仿真实验及对比分析, 得出以下结论:

- 1) 针对股票价格数据的非线性, 选取多个股票影响因子归一化后作为 ELM 模型的输入对次日股票收盘价进行预测, 其中股票数据包含收盘价、成交金额等数据, 为股票价格预测提供基础。
- 2) 利用 PE-CEEMDAN 对误差序列进行分解, 得到的多个模态分量和残差分量更能反映误差序列的变化特征, 解决模态混叠和噪声残留问题, 提高预测效果。
- 3) 通过寻优算法对 ELM 模型的参数进行寻优, 仿真实验表明寻优算法对 ELM 模型参数有着一定的效果, 且改进后的 GWO 算法的全局寻优能力要优于 GWO、PSO。

4) 通过仿真实验及对各模型的对比分析和 DM 检验, 验证本文提出的基于误差修正和 CEEMDAN-IGWO-ELM 的股票预测模型的可行性与优越性, 且预测效果较好, 预测曲线也与真实价格曲线更加吻合。

基金项目

国家自然科学基金项目(71961001)。

参考文献

- [1] 王晨. 基于隐马尔科夫模型的股票价格指数预测[D]: [硕士学位论文]. 济南: 山东大学, 2018.
- [2] 王莹. 基于 ARMA 模型的股票价格的分析及预测[J]. 生产力研究, 2021(9): 124-127.
- [3] 吴玉霞, 温欣. 基于 ARIMA 模型的短期股票价格预测[J]. 统计与决策, 2016(23): 83-86.
- [4] 邹婕, 李路. 基于随机森林的 SA-BiGRU 模型的股票价格预测研究[J]. 中国物价, 2023(11): 52-56.
- [5] 钟琳, 颜七笙. 基于误差修正和 VMD-ICPA-LSSVM 的短期风速预测建模[J]. 南京信息工程大学学报, 2024, 16(2): 247-260. <https://doi.org/10.13878/j.cnki.jnuist.20230421002>
- [6] Ishwarappa, A.J. (2021) Big Data Based Stock Trend Prediction Using Deep CNN with Reinforcement LSTM model. *International Journal of Systems Assurance Engineering and Management*. <https://doi.org/10.1007/s13198-021-01074-2>
- [7] 李若晨, 肖人彬. 基于改进狼群算法优化 LSTM 网络的舆情演化预测[J/OL]. 复杂系统与复杂性科学: 1-13. <https://kns.cnki.net/kcms/detail/37.1402.n.20221009.1057.002.html>, 2023-12-29.
- [8] Sun, Y., Sun, Q.S. and Zhu, S. (2022) Prediction of Shanghai Stock Index Based on Investor Sentiment and CNN-LSTM Model. *Journal of Systems Science and Information*, **10**, 620-632. <https://doi.org/10.21078/JSSI-2022-620-13>
- [9] 蔡浩, 李林峰, 李涵, 等. 基于极限学习机的短期交通流预测混合优化模型[J]. 交通运输系统工程与信息, 2023, 23(5): 75-82, 183.
- [10] 周彦, 王山亮, 杨威, 等. 基于 PSO-ELM 的卫星导航欺骗式干扰检测[J]. 导航定位与授时, 2022, 9(5): 153-161.
- [11] 顾爽, 陈真诚, 张振强, 等. 基于极限学习机的胰岛素评价模型[J]. 科学技术与工程, 2021, 21(10): 3935-3939.
- [12] 张文煜, 马可可, 郭振海, 等. 基于灰狼算法和极限学习机的风速多步预测[J]. 郑州大学学报(工学版), 2024, 45(2): 89-96. <https://doi.org/10.13705/j.issn.1671-6833.2023.05.008>
- [13] 武博. 基于 LSTM 模型的股票价格预测[D]: [硕士学位论文]. 大连: 大连理工大学, 2022.
- [14] 韩莹, 张栋, 孙凯强, 等. 结合长短时记忆网络和宽度学习的股票预测新模型研究[J]. 运筹与管理, 2023, 32(8): 187-192.
- [15] 朱菊香, 任明煜, 谷卫, 等. 基于 CEEMDAN-IGWO-CNN-LSTM 空气质量预测建模[J/OL]. 计算机仿真: 1-11. <https://link.cnki.net/urlid/11.3724.tp.20230922.1501.004>, 2023-12-29.
- [16] 孙存浩, 胡兵, 邹雨轩. 指数趋势预测的 BP-LSTM 模型[J]. 四川大学学报(自然科学版), 2020, 57(1): 27-31.
- [17] Wang, Q. and Zhang, Y.W. (2022) Research on PM2.5 Pollution Prediction Method in Hefei City Based on CNN-LSTM Hybrid Model. *Journal of Physics: Conference Series*, **2400**, Article ID: 012006. <https://doi.org/10.1088/1742-6596/2400/1/012006>
- [18] 尤睿凡. 基于时间序列模型与机器学习的组合模型的股票价格指数预测研究[D]: [硕士学位论文]. 济南: 山东大学, 2021.
- [19] Mirjalili, S., Mirjalilis, M. and Lewis, A. (2014) Grey Wolf Optimizer. *Advances in Engineering Software*, **69**, 46-61. <https://doi.org/10.1016/j.advengsoft.2013.12.007>
- [20] 陈颖, 张灿, 肖春艳, 等. 基于 GWO-SVR 的土壤镉元素含量含水率校正预测模型研究[J]. 光学学报, 2020, 40(10): 174-181.
- [21] Zhang, Y., He, D. and Wu, Q. (2023) Forecasting of PM2.5 Concentration Time Series Based on SSA-LSTM Model. *International Conference on Statistics, Data Science, and Computational Intelligence (CSDSCI 2022)*, Qingdao, 19-21 August 2022, 373-380. <https://doi.org/10.1117/12.2656939>
- [22] 罗聪, 范力栋, 赵学文, 等. 基于 VMD-IGWO-LSSVM 的覆冰预测模型研究[J]. 电网与清洁能源, 2021, 37(6): 9-17.
- [23] Daniel, O., Sunday, S.I. and Eric, B. (2022) A Compositional Function Hybridization of PSO and GWO for Solving

- Well Placement Optimisation Problem. *Petroleum Research*, **7**, 401-408. <https://doi.org/10.1016/j.ptlrs.2021.12.004>
- [24] Kumar, B.S., Kumar, A.B., Jobanputra, J.H. and Kolhe, M.L. (2022) Solar-DG and DSTATCOM Concurrent Planning in Reconfigured Distribution System Using APSO and GWO-PSO Based on Novel Objective Function. *Energies*, **16**, Article 263. <https://doi.org/10.3390/en16010263>
- [25] Khodsuz, M. and Mashayekhi, V. (2023) Grounding System Impedance Influence on the Surge Arrester Frequency-Dependent Model Parameters Using PSO-GWO Algorithm. *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, **42**, 1456-1476. <https://doi.org/10.1108/COMPEL-07-2022-0229>
- [26] Huang, N.E., Shen, Z., Long, S.R., *et al.* (1998) The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **454**, 903-995. <https://doi.org/10.1098/rspa.1998.0193>
- [27] Wu, Z. and Huang, N.E. (2009) Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Advances in Adaptive Data Analysis*, **1**, 1-41. <https://doi.org/10.1142/S1793536909000047>
- [28] Yeh, J.R., Shieh, J.S. and Huang, N.E. (2010) Complementary Ensemble Empirical Mode Decomposition: A Novel Noise Enhanced Data Analysis Method. *Advances in Adaptive Data Analysis*, **2**, 135-156. <https://doi.org/10.1142/S1793536910000422>
- [29] Torres, M.E., Colominas, M.A., Schlotthauer, G. and Flandrin, P. (2011) A Complete Ensemble Empirical Mode Decomposition with Adaptive Noise. 2011 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, 22-27 May 2011, 4144-4147. <https://doi.org/10.1109/ICASSP.2011.5947265>
- [30] 李阳, 李维刚, 赵云涛, 等. 基于莱维飞行和随机游动策略的灰狼算法[J]. 计算机科学, 2020, 47(8): 291-296.
- [31] Bhatt, B., Sharma, H., Arora, K., Joshi, G.P. and Shrestha, B. (2023) Levy Flight-Based Improved Grey Wolf Optimization: A Solution for Various Engineering Problems. *Mathematics*, **11**, Article 1745. <https://doi.org/10.3390/math11071745>
- [32] 钟琳, 颜七笙. 基于 PSO-ELM 的中国石油股票价格预测建模[J]. 江西科学, 2022, 40(1): 11-16.