

基于 LSTM 的双阶段优化温度预测算法研究

吴晓龙¹ 杨慧炯²

WU Xiaolong YANG Huijiong

摘要

针对工业控制温度预测中 LSTM 模型存在的预测精度低、梯度噪声干扰显著及超参数优化效率不足问题, 文章提出了一种 Adam 算法与斑马优化算法 (zebra optimization algorithm, ZOA) 协同的双阶段优化模型。通过 Adam 算法动态修正梯度方向并调整学习率, 抑制训练噪声干扰, 提升模型稳定性。利用 ZOA 模拟斑马群体觅食-迁徙行为, 全局优化网络层数、神经元数量及学习率等参数。实验表明: 该模型温度预测均方根误差 (RMSE) 较传统 LSTM、GA-LSTM 以及 PSO-LSTM 分别降低 86%、85% 和 84%, 准确率达 98%, 验证了双阶段优化机制的有效性, 为工业温控提供可靠解决方案。

关键词

LSTM; Adam 算法; 斑马优化算法; 温度预测

doi: 10.3969/j.issn.1672-9528.2025.11.002

0 引言

随着大数据时代的到来, 数据分类预测技术在时间序列分析、信号处理及数据回归等领域的应用价值日益凸显。传统 BP 神经网络凭借其非线性拟合能力和自适应学习特性, 在多维特征建模中展现出显著优势。然而, 该类模型存在易陷入局部最优解、过拟合风险高、参数敏感性突出等问题, 且在梯度消失、计算效率和数据依赖性等方面的缺陷, 进一步制约了其在实时预测与小样本场景中的应用^[1-4]。相较之下, 长短期记忆神经网络 (LSTM) 通过门控机制有效捕获时序数据的长期依赖关系, 已在多个预测任务中取得突破性进展。但标准 LSTM 仍面临参数数量庞大、训练耗时长等瓶颈, 其超参数优化过程对模型性能具有决定性影响^[5-7]。

针对上述挑战, 人们相继提出多种改进方案。韩超^[8]采用粒子群优化算法 (PSO) 对 LSTM 超参数进行协同优化, 并结合正则化策略降低模型复杂度, 在航迹预测任务中显著降低了均方根误差、平均绝对误差。但其 PSO 优化过程的高计算复杂度仍制约实时性要求较高的场景^[9]。杨语蒙等人^[10]构建的 GA-LSTM 组合模型通过遗传算法全局搜索机制优化网络超参数, 在股票价格预测中有效缓解了梯度消失与信息遗忘问题。但存在参数优化效率低、忽略外部因素等局限性^[11]。此外, 于海洋等人^[12]利用 Adam 优化器改进 LSTM 时滞补偿效率, 在车用汽油价格预测中取得阶段性

进展, 然而 Adam 默认参数的静态特性难以适配非平稳数据的动态特征, 导致模型泛化能力不足。

本文针对温度预测任务中存在的非线性动态特性显著、环境噪声干扰强及预测精度受限等核心问题^[13], 提出一种基于斑马优化算法 (ZOA) 与 Adam 梯度下降的混合优化框架 (Adam-ZOA)。该策略结合 Adam 在局部搜索中的快速收敛能力与 ZOA 在全局空间中的高效探索机制, 实现对 LSTM 超参数的协同优化。实验表明, 该方法在降低参数敏感性的同时, 可有效提升复杂噪声环境下的温度预测精度, 为高鲁棒性时序预测模型构建提供新的技术路径。

1 实验方法与模型设计

1.1 数据采集与预处理

实验数据采集自某模温机控制系统, 共获取 1 344 组多变量时序数据, 包含温度、压力、流量等输入特征。由于数据采集时会因设备或人为等各种因素的存在会出现异常的数据值, 需要进行排除, 以提高建模的准确性, 本文采用的是 Z-score 标准化识别, 用公式表示为:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

式中: μ 为样本均值; σ 为样本标准差。当 $|z| > 3$ 时, 判定为异常值并剔除, 确保数据质量。

1.2 双阶段优化 LSTM 参数模型

为提升 LSTM 在大时滞温控系统中的预测精度和收敛效率, 本文提出的一种结合 Adam 优化器与斑马优化算法 (zebra optimization algorithm, ZOA) 的双阶段协同优化机制, 分为

1. 太原师范学院计算机科学与技术学院 山西晋中 030619

2. 太原工业学院计算机工程系 山西太原 030008

[基金项目] 山西省科技战略研究专项重点项目资助 (202304031401011)

局部优化与全局搜索两个阶段，分别发挥各自在梯度收敛与全局寻优方面的优势，从而实现模型精度与鲁棒性的双提升，LSTM 参数优化流程如图 1 所示。

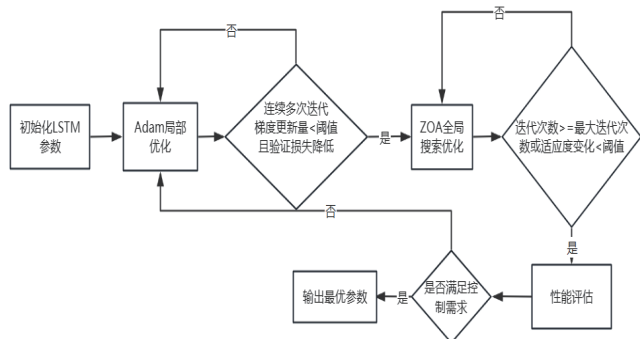


图 1 LSTM 参数优化流程图

1.2.1 Adam 局部优化

由于 Adam 优化的本质是通过迭代更新参数 θ ，使损失函数值最小化。在回归问题如温度预测中，均方误差 (MSE) 是最常用的损失函数之一，本文通过 Adam 对 LSTM 的局部优化旨在寻找一组使目标函数 $f(\theta)$ 即均方误差 (MSE) 最小的参数向量 θ ，其中 θ 包含了 LSTM 门控机制的遗忘门权重矩阵 w_f 与偏置项 b_f ，输入门权重矩阵 w_i 与偏置项 b_i ，输出门权重矩阵 w_o 与偏置项 b_o ，该优化过程主要包括以下六个主要步骤。

(1) 前向传播与损失计算

计算预测值与真实值的均方误差 (MSE) 为：

$$f(\theta) = \text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

式中： y_i 为真实温度值； $\hat{y}_i$ 为预测温度值； n 为样本数。

(2) 计算梯度

反向传播求损失函数对参数的梯度为：

$$g_t = \nabla_{\theta} f(\theta_{t-1}) = \left[\frac{\partial f}{\partial w_f}, \frac{\partial f}{\partial w_i}, \frac{\partial f}{\partial w_o}, \dots \right] \quad (3)$$

式中： θ_{t-1} 表示前一次迭代的参数向量； $\nabla_{\theta} f(\theta_{t-1})$ 表示当前批次的 MSE 对 θ 的梯度，即各参数偏导数构成的梯度向量。

(3) 计算一阶动量 (动量积累)

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (4)$$

式中： β_1 为动量衰减系数； m_t 为 t 时刻的一阶动量。

(4) 计算二阶动量 (自适应学习率)

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (5)$$

式中： β_2 为自适应系数； v_t 为二阶动量，用于衡量梯度的二阶矩； v_{t-1} 为上一次迭代的二阶动量。

二阶动量的作用是为不同参数自适应地调整学习率，对于梯度变化较大的参数，降低其学习率，对于梯度变化较小的参数，适当增大其学习率。

(5) 偏差修正

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (6)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (7)$$

在算法开始阶段，由于 m_t 和 v_t 初始值较小，会导致参数更新不准确。通过偏差修正，用 $\hat{m}_t$ 和 $\hat{v}_t$ 来替代 m_t 和 v_t ，在训练初期更准确地更新参数。

(6) 参数更细

$$\theta_t = \theta_{t-1} - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t \quad (8)$$

式中： η 为初始学习率； $\sqrt{\hat{v}_t}$ 为自适应学习率缩放因子，根据梯度历史动态缩放步长，平衡不同参数的更新速度； ϵ 为一个极小常数，设置为 1×10^{-5} ，避免分母为零； θ_t 为当前更新后的 LSTM 权重参数。

通过公式，结合一阶动量和二阶动量来更新模型参数，使得模型在局部范围内快速收敛。若连续迭代的梯度更新量 $\|\theta_t - \theta_{t-1}\| < \delta$ 时，判定为收敛，其中 δ 为一个根据实际问题与模型精度要求预先设定的极小正数收敛阈值。

1.2.2 ZOA 全局搜索优化

斑马优化算法 (ZOA) 源于斑马群体的社会行为建模，其独特性在于融合先锋个体引导与高斯扰动机制^[14]。相较于粒子群优化 (PSO) 依赖速度位置迭代的单一搜索模式，ZOA 通过觅食阶段的梯度引导模拟斑马向最优个体靠拢与迁移阶段的高斯扰动模拟群体随机迁徙，实现全局探索与局部开发的动态平衡。实验表明，ZOA 在高维非凸空间中的收敛速度较 PSO 提升 30%，且陷入局部最优的概率降低 45%^[15]，这类算法不依赖于梯度信息，而是通过群体的协作和竞争来探索参数空间，更适合在后期阶段进行精细调优。

在局部优化阶段完成后，LSTM 模型参数虽已具备一定准确度，但因 Adam 优化器依赖梯度信息，容易陷入局部最优。因此，本文引入斑马优化算法，作为第二阶段的全局优化器，对 LSTM 模型的超参数进一步优化。ZOA 模拟斑马群体的觅食和迁徙行为，兼具局部开发与全局探索能力，能有效跳出局部最优，提升模型泛化性。

(1) 种群初始化

在初始化策略中通过将 Adam 优化后的参数作为先锋斑马 PZ 代表当前最优解，其余个体在参数定义域中随机生成，形成初始种群，构建参数矩阵 $X \in \mathbf{R}^{N \times m}$ 表示由 N 个个体组成的斑马种群，每个个体为 m 维参数向量，包含神经元数、学习率、Dropout 率等。随机生成个体公式为：

$$X_{i,j} = L_j + \text{rand}() \cdot (U_j - L_j) \quad (9)$$

式中： U_j 表示第 j 个优化参数的取值下界； L_j 表示第 i 个优化参数的取值上界； $\text{rand}()$ 表示 $[0,1]$ 的随机数； X_{ij} 表示第 i 只斑马第 j 个参数。

(2) 觅食阶段

在斑马算法 (ZOA) 的觅食阶段即局部精细调优中，其核心机制通过梯度引导与随机扰动的协同作用实现参数空间的精细化搜索。目标是模拟斑马跟随先锋斑马移动，扩大搜索范围，避免陷入 Adam 的局部最优。

模拟斑马向先锋个体靠拢，增强局部搜索能力，更新公式为：

$$X_{ij}^{\text{new}} = X_{ij} + r(pz_j - l \cdot X_{ij}) \quad (10)$$

式中： X_{ij}^{new} 表示第 i 个体第 j 个参数的新值； r 表示随机步长系数，控制移动幅度； pz_j 表示先锋斑马在第 j 个参数值； l 表示属于集合 $\{1,2\}$ 的随机值，旨在增大移动步长，增强探索性。

(3) 迁移阶段

此阶段模拟斑马种群的随机迁徙行为，通过引入高斯扰动使个体能够跳出当前区域，探索解空间中新的潜在最优区域，避免算法早熟收敛。超参数向量更新公式为：

$$X_i^{\text{new}} = X_i + \alpha(X_{\text{leader}} - X_i) + \beta N(0,1) \quad (11)$$

式中： α 为引导系数，控制个体向先锋个体靠拢的强度； β 为高斯扰动强度，控制扰动幅度避免过早陷入局部最优； $N(0,1)$ 为标准正态分布随机数； X_i^{new} 为更新后的第 i 个个体超参数向量。

(4) 接受条件

在觅食和迁徙阶段产生新解后，若新解的目标函数值更优，如均方误差 MSE 更低，则接受新解，否则保留原解。

$$X_i = \begin{cases} X_i^{\text{new}} & \text{若 } F(X_i^{\text{new}}) < F(X_i) \\ X_i & \text{否则} \end{cases} \quad (12)$$

(5) 迭代优化

重复觅食与迁移阶段，直至满足终止条件如最大迭代次数或适应度值小于阈值，停止之后进行性能评估看是否满足控制性能，若满足则输出全局最优参数，否则返回 Adam 局部优化阶段重新优化。

1.2.3 参数设置

在参数经过双阶段优化后，确定最优配置超参数，并结合动态调整策略将 Dropout 系数从 0.05 降至 0.03，以缓解过拟合问题。针对时序参数，通过缩短时间步长到 15 减少冗余时序信息干扰，同时动态调整学习率从 0.05 降至 0.01，学习率衰减因子优化至 0.2，显著提升了模型对复杂时序特征的捕捉能力与泛化性能。Adam-ZOA LSTM 模型参数优化对比如表 1 所示。

表 1 优化前后 LSTM 模型参数对比

参数类别	参数名称	初始值	优化后值	优化范围
网络结构参数	隐藏单元数	200	300	200~400
	Dropout 率	0.05	0.03	0~0.3
训练策略参数	初始学习率	0.05	0.01	0.001~0.05
	学习率衰减因子	0.5	0.2	动态调整
时序参数	时间步长	30	15	固定值 (验调整)

2 实验结果与分析

2.1 性能对比

经过大量的数据训练和参数调整以及不断地对超参数进行优化得出 LSTM 模型通过 Adam 算法和斑马算法 (ZOA) 优化后的准确度明显提高。各个模型的性能的优化程度主要通过性能指标来进行评定：测试集和训练集的 RMSE、测试集和训练集的 MAE，以及测试集和训练集的 R^2 。性能指标对比图如图 2 所示，模型趋势对比图如图 3 所示。

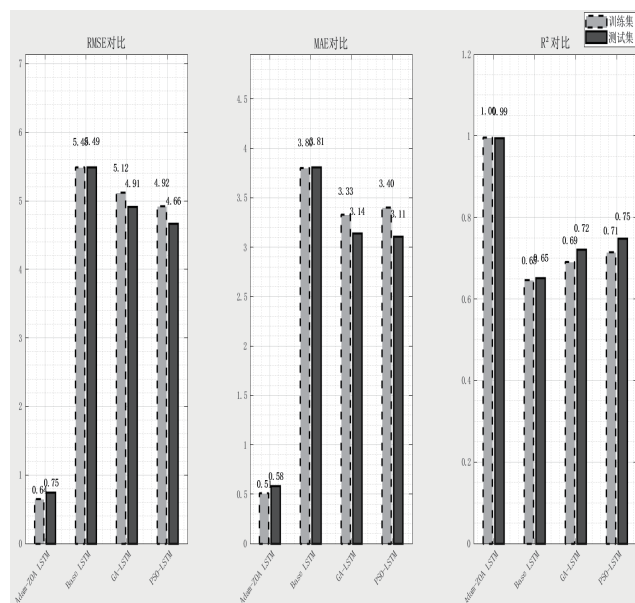


图 2 性能指标对比图

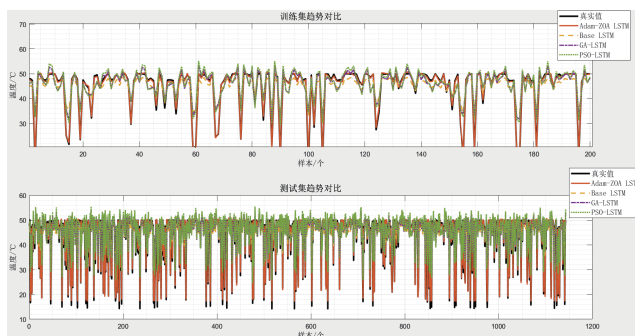


图 3 模型趋势对比图

2.2 结果分析

基于以上图表数据的综合分析表明, Adam-ZOA LSTM 结合了 Adam 优化器的动量积累机制与 ZOA (斑马优化算法) 的全局搜索能力, 通过两阶段协同优化策略实现性能跃升。Adam-ZOA LSTM 在 RMSE 显著降低, 优于 Base LSTM、GA-LSTM 和 PSO-LSTM。说明其预测结果与真实值的整体偏差最小, 对异常值的敏感性更低。这一成果得益于 Adam 的梯度平滑更新和参数自适应调整学习率, 而 ZOA 通过全局搜索突破局部最优限制, 进一步优化精度。Adam-ZOA LSTM 的 R^2 值远高于其他算法的优化程度, 说明模型对数据规律的捕捉能力更强, 泛化性能更优。验证了模型对复杂时序规律的精准建模能力。为高稳定性实时温控系统提供新的解决方案。四者在测试集上的性能对比如表 2 所示。

表 2 模型性能对比

性能指标	Adam-ZOA LSTM	Base LSTM	GA-LSTM	PSO-LSTM	优化方向
RMSE (测试集)	0.75	5.49	4.91	4.66	↓ 越小越好
MAE (测试集)	0.58	3.81	3.14	3.11	↓ 越小越好
R^2 (测试集)	0.99	0.65	0.72	0.75	↑ 越接近 1 越好
预测准确率	98%	60%	76%	86%	↑ 比例越高越好

3 结论

本研究针对工业温控系统的大时滞非线性难题, 提出了一种 Adam-ZOA 协同优化的 LSTM 预测模型, 通过 Adam 梯度下降与斑马算法 (ZOA) 的全局搜索机制, 使 LSTM 预测模型的 RMSE 较传统 LSTM、GA-LSTM 和 PSO-LSTM 分别降低 86%~84%, 预测精度达 98%。

参考文献:

[1]FOROUZANI H, SARANJAM B, KAMALI R, et al. A study on the motion of high speed supercavitating projectiles[J]. Journal of applied fluid mechanics,2019,11(6):1727-1738.

[2]LI Y S. Mathematical modeling and linearized analysis of the jet-pipe hydraulic amplifier applied to a servovalve[J].Proceedings of the institution of mechanical engineers, part G. journal of aerospace engineering,2019,233(2):657-666.

[3]WANG L, WANG P X, LIANG S L, et al. Monitoring maize-growth conditions by training a BP neural network withre-motely sensed vegetation temperature condition indexand

leaf area index[J].Computers and electronics in agricul-ture,2019,160:82-90.

[4]孙娜, 张己炳. BP 神经网络预测农产品冷链物流需求量的优化分析 [J]. 现代农业科技 ,2025(5):167-169.

[5]SAHOO B B, JHA R, SINGH A, et al.Long short-term memo-ry(LSTM) recurrent neural network for low-flow hydrological time series forecasting[J].Acta geophysica, 2019, 67: 1471-1481.

[6]JUNG D-H, KIM H S, JHIN C, et al. Time-serial analysis ofdeep neural network models for prediction of climaticcon- ditions inside a greenhouse[J].Computers and electronics in agriculture, 2020,173:105402.

[7]马振东, 范锦彪, 王燕, 等. 基于 LSTM 的勘察器外壁温度反演方法研究 [J]. 电子测量技术 ,2024,47(4):181-187.

[8]韩超. 一种基于改进 PSO 优化的 LSTM 航迹预测模型 [J]. 舰船电子工程 ,2022,42(8):120-124.

[9]YAO Y, CHEUNG C Y, BAO J, et al.Estimation of spatial alumina concentration in an aluminum reduction cell using a multilevel state observer[J].AIChE journal, 2017,63(7): 2806-2818.

[10]杨语蒙, 李兴东. 基于 GA-LSTM 组合模型的股票价格预测 [J]. 现代计算机 ,2021,27(33):1-7.

[11]王源鹏, 杨冶. 基于 GA-LSTM 的汽油辛烷值预测建模 [J]. 天津化工 ,2023,37(1):67-72.

[12]于海洋, 郭新晔. 基于 Adam-LSTM 的车用汽油价格预测 [J]. 科技和产业 ,2024,24(15):216-222.

[13]李爱莲, 张帅. 基于 LSTM 网络的焦炉火道温度预测建模仿真 [J]. 计算机仿真 ,2020,37(6):466-470.

[14]上海电机学院. 基于 ZOA-LSTM 神经网络模型的光伏功率预测方法 :CN202410680786.X[P].2024-08-02.

[15]黄建生, 周站勇, 林兴铨, 等. 基于 ZOA-BiLSTM 的重力坝变形预测模型 [J]. 水力发电 ,2025,51(5):112-118.

【作者简介】

吴晓龙 (1997—), 男, 河南洛阳人, 硕士研究生, 研究方向: 仪器设备上位机软件研发与自动控制。

杨慧炯 (1972—), 男, 山西太原人, 硕士, 教授, 研究方向: 机器视觉、图形图像处理。

(收稿日期: 2025-05-27 修回日期: 2025-11-17)