

doi:10.11835/j.issn.1000.582X.2024.08.010

基于 APSO 的 LSTM 神经网络模型优化方法研究

袁琳娜, 杨良斌

(国际关系学院 网络空间安全学院, 北京 100091)

摘要: 多隐含层长短期记忆神经网络(long short-term memory, LSTM)循环神经网络权值与阈值更新依赖梯度下降算法, 模型收敛速度慢, 网络节点的权值计算易出现局部极值, 导致 LSTM 神经网络模型不能得到全局最优, 网络模型泛化能力下降, 限制 LSTM 循环神经网络的应用。因此, 利用加速粒子群优化算法(accelerated particle swarm optimization, APSO)的优化能力, 提出一种改进 LSTM 神经网络模型。该模型将均方根误差设计为适宜值函数, 并利用 APSO 算法构建寻优策略, 对各神经元节点间的权值进行全局优化, 提升模型的泛化和预测性能。通过经典 DataMarket 及 UCI 数据集的实验结果表明, APSO-LSTM 模型的预测精度较传统 LSTM 模型有显著提升, 验证了 APSO-LSTM 模型的有效性和实用性。

关键词: 神经网络; 权值优化; 适宜值; APSO-LSTM 模型

中图分类号: TP183

文献标志码: A

文章编号: 1000-582X(2024)08-103-09

LSTM neural network model optimization algorithm based on APSO

YUAN Linna, YANG Liangbin

(School of Cyber Science and Engineering, University of International Relations, Beijing 100091, P. R. China)

Abstract: Due to the slow convergence speed of the model with many hidden layers in the LSTM (long short-term memory) recurrent neural network, the updating of its weights and thresholds depends on the gradient descent algorithm, which may lead to the local extremum phenomenon in the weight correction of the network nodes, resulting in the reduction of the generalization ability of the LSTM neural network model. Based on this, this paper proposes an optimized LSTM neural network model based on APSO (accelerated particle swarm optimization) algorithm (APSO-LSTM). In this model, root mean square error is designed as an appropriate value function, and APSO algorithm is used to build an optimization system to optimize the weights of each neuron node globally, so as to improve the prediction performance of the model. The experimental results on the classic DataMarket and UCI datasets show that the prediction accuracy of APSO-LSTM model is significantly improved compared with the traditional LSTM model, which verifies the effectiveness of APSO-LSTM model.

Keywords: neural network; weight optimization; fitness value; APSO-LSTM model

收稿日期: 2020-07-11

基金项目: 国家安全高精尖学科建设科研专项(2019GA37)。

Supported by National Security High Precision and Advanced Discipline Construction Research Project(2019GA37).

作者简介: 袁琳娜(1998—), 女, 主要从事数据工程与科学方向研究, (E-mail)459131607@qq.com。

随着深度学习的迅猛发展,人工智能已成为众多学者研究的热点之一^[1]。目前,神经网络模型已被广泛应用于语音识别^[2]、机器翻译^[3-4]、视频动作识别^[5-6]及时序预测等任务。而循环神经网络RNN(recurrent neural networks)作为神经网络模型的重要分支,也在各领域取得了巨大成就。RNN适合处理时序数据及其潜在的时序联系,但仍存在梯度消失及梯度爆炸等缺陷^[7-8],为此,Hochreiter等^[9]于1997年提出了长短期记忆神经网络模型LSTM,用于改进传统的RNN模型,即通过3个门控单元有针对性丢弃冗余并抑制梯度消失。但LSTM神经网络模型仍存在学习效率不高、隐层过多时出现梯度消失等缺陷,故众多学者不断对LSTM(long short-term memory)模型予以改进。如门控循环神经网络(gated recurrent unit,GRU)是在LSTM基础上优化了网络结构,由3个门控单元变为2个门控单元,GRU模型现已在工业生产中被普遍运用。2018年,Li等^[10]结合CNN(convolutional neural networks)与LSTM模型,用以新闻文本分类,有显著成效。Song等^[11]提出一种改进的深度LSTM神经网络,该模型在一定程度上能预测趋势。

上述大多是以改进LSTM的网络结构为目标,当隐含层数过多时,LSTM神经网络的收敛速度较慢,模型参数的修正往往只能得到局部最优。目前,已有越来越多的研究着眼于群体智能算法,并用其优化神经网络的权值及阈值等参数。Lin等^[12]提出了一种基于遗传算法优化的LSTM神经网络模型(GA-LSTM)用于股票分析。ElSaid等^[13]提出基于蚁群算法优化的LSTM神经网络模型(ACO-LSTM)等。APSO算法是由Yang等^[14]在2010年提出的改进粒子群优化算法,较标准的PSO(particle swarm optimization)算法^[15-16],APSO更易获取全局最优解,且收敛速度更快。王永贵等^[17]针对算法易陷入早熟收敛的问题,修改了惯性权重的变化规律,避免早熟收敛,使种群在后期也有丰富特性。夏学文等^[18]提出具备反向学习和局部学习能力的粒子群优化算法,该算法将每次迭代中的最差粒子位置和每个个体的历史最差位置记录下来,当算法陷入停滞时,对部分粒子采用反向学习策略,向这些较差的信息学习,使粒子脱离局部最优。

本文提出基于APSO算法优化的LSTM神经网络模型APSO-LSTM(accelerated particle swarm optimized-long short-term memeroy),该模型通过将LSTM神经网络的预测结果的均方根误差设计为适宜值函数,并利用APSO算法构建寻优空间,对LSTM中网络节点间的权重值进行全局优化,提升模型泛化能力。同时,利用所提的APSO-LSTM神经网络模型在经典数据集上建立预测模型,验证该模型的有效性。

1 LSTM神经网络模型

LSTM神经网络模型是由遗忘门、输入门及输出门等3个门控单元和cell单元组成,通过门控单元来学习并记忆时序数据,在较大程度上抑制梯度消失,且维系长距离的时序信息依赖,擅于处理时序问题。标准LSTM的神经元结构如图1所示。

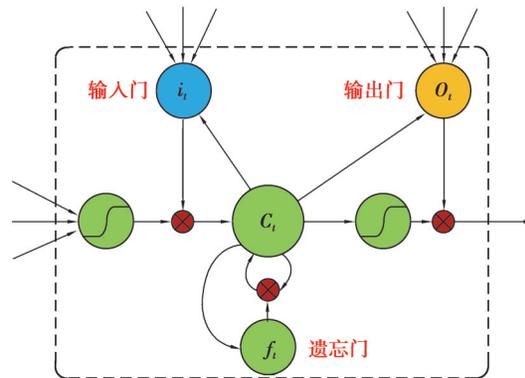


图1 LSTM神经元结构

Fig. 1 LSTM neuronal structure

设输入序列为 (x_1, x_2, \dots, x_t) ,则在 t 时刻:

1) INPUT GATE为输入门,主要处理输入数据

$$i_t = S(W_i * [h_{t-1}, x_t] + b_i), \quad (1)$$

2)FOGET GATE为遗忘门,主要负责决定当前神经元记忆的历史时序信息的保留程度

$$f_t = S(W_f * [h_{t-1}, x_t] + b_f) \quad (2)$$

3)OUTPUT GATE为输出门,表示该神经元的输出结果

$$o_t = S(W_o * [c_t, h_{t-1}, x_t] + b_o) \quad (3)$$

式中: x_t 表示 t 时刻神经元的输入; h_{t-1} 表示 $t-1$ 时刻该隐含层的输出状态; W_f 、 W_o 以及 W_c 分别表示 t 时刻该神经元的输入和输入门、遗忘门以及输出门的权重矩阵; b_f 、 b_o 、 b_c 分别表示输入门、遗忘门以及输出门的偏置向量; W_c 表示输入和cell单元之间的权重矩阵; c_t 表示 t 时刻cell单元输出为

$$c_t = f_t * c_{t-1} + i_t * \tanh(W_c * [h_{t-1}, x_t]), \quad (4)$$

h_t 表示 t 时刻该隐含层的输出状态信息

$$h_t = o_t * \tanh(c_t) \quad (5)$$

当LSTM网络有多个隐含层时, x_t 仅表示第1层隐含层神经元的输入, h_t 则表示 t 时刻其余隐含层对应的输入; \tanh 为激励函数,S表示Sigmoid函数。深度LSTM神经网络是以标准LSTM神经单元为基础,构建包含1个或多个隐含层的网络模型,通过遗忘门不断遗弃冗余,以存储有用的时序信息,保持长距离依赖关系,具有较强时序数据挖掘能力。

2 APSO算法

粒子群优化算法(particle swarm optimization, PSO)是由Eberhart^[15]提出的启发式优化算法,通过模拟鸟群觅食的生物过程,使粒子在合理的空间内按一定规则移动,在不断迭代中找到合适位置,找到1个可行解。PSO中,每个粒子均代表某个优化问题的候选解,其初始位置及速度是被随机设定的。

标准PSO中速度及位置的更新方程

$$V_{ij}^{t+1} = wV_{ij}^t + \alpha r_{1,ij}^t (y_{ij}^t - x_{ij}^t) + \beta r_{2,ij}^t (\hat{y}_j^t - x_{ij}^t), \quad (6)$$

$$x_{ij}^{t+1} = x_{ij}^t + V_{ij}^{t+1}, \quad (7)$$

式中: V_{ij}^t 为粒子 i 的速度; x_{ij}^t 为其当前所在的位置; y_{ij}^t 为该单个粒子在搜索过程中的历史最佳位置; \hat{y}_j^t 为整个种群的全局最佳位置,下标 j 是需要优化问题的空间维度数目^[13]; w 为惯性权重; $r_{1,ij}^t, r_{2,ij}^t$ 是2个随机数; α, β 为预设定的学习参数。式(6)中: wV_{ij}^t 为粒子的当前状态; $\beta r_{2,ij}^t (\hat{y}_j^t - x_{ij}^t)$ 表示粒子由全种群状态出发,进行全局寻优; $\alpha r_{1,ij}^t (y_{ij}^t - x_{ij}^t)$ 表示粒子由自身状态出发,进行局部寻优。

PSO算法在寻优过程中,易出现局部极值,且收敛速度较慢。在大多数优化情况下,不必运用粒子自身的历史最佳位置 y_{ij}^t 来更新粒子状态,这会增加训练开销且收敛较慢,可通过随机数技术实现粒子自身的多样性^[16]。因此,采用APSO算法对LSTM神经网络权值进行优化,该算法无须设定种群中粒子的初始速度,并使用随机数技术替代 $\alpha r_{1,ij}^t (y_{ij}^t - x_{ij}^t)$ 。相对标准的PSO算法,APSO算法的流程得以简化,避免随机设定的初始速度对粒子寻优不利,加快收敛速度,更易得到全局最优,下面介绍该算法的相关概念。

式(8)和式(9)为APSO算法中粒子的速度及位置更新方程

$$V_{ij}^{t+1} = wV_{ij}^t + \alpha r + \beta r_{2,ij}^t (\hat{y}_j^t - x_{ij}^t) \quad (8)$$

式(9)中, r 为取值范围介于[0,1]之间的随机数,这是由于初始化粒子速度及运用粒子历史最佳位置影响速度等处理步骤并不是必须的,一定程度上会带来臃肿。故运用随机数技术替代式(6)中的 $\alpha r_{1,ij}^t (y_{ij}^t - x_{ij}^t)$,进一步提高收敛速度,其余参数的含义与标准PSO算法一致。

$$x_{ij}^{t+1} = (1 - \beta) x_{ij}^t + \beta \hat{y}_j^t + \alpha r \quad (9)$$

同时,APSO算法也重写了粒子的位置更新公式。如式(8)所示,第3项中的 r 使系统更具移动性,若 α 取值合理,则APSO算法可彻底避免出现局部极值现象。一般情况下, α 取值[0.1,0.5]且 β 取值[0.2,0.7],可适用于大多数优化问题。式(8)中不包含有关粒子速度的参数,故APSO算法不必初始化粒子速度,简化算法流程。在迭代过程中由于粒子的随机性在逐步减小,可采用式(10)来优化参数 α 。

$$\alpha = \delta^t, \quad (10)$$

式中, $0 < \delta < 1$, 可将其设置为 0.1~0.99 间的任意参数, 这里 $t \in [0, t_{\max}]$, t_{\max} 为迭代的最大次数。

3 APZO-LSTM 神经网络模型

考虑到 LSTM 存在收敛速度慢, 且权值与阈值等网络参数的修正往往会陷入局部最优。而 APZO 算法是 PSO 算法的一种衍生优化模型, 寻优性能更强。故本文运用 APZO 优化算法对 LSTM 神经网络的权值进行调优, 可使网络中神经元节点间权重的取值更合理, 提高模型的泛化能力和表征能力。

本文根据 LSTM 神经网络中全部的神经单元数目, 决定 APZO 中粒子群的粒子数目, 以各神经元输出的均方根误差作为适宜度函数, 构建 APZO 寻优空间, 对训练完成后的网络原始权值进行全局寻优, 最终完成对整个 LSTM 网络的参数优化。

3.1 APZO-LSTM 模型的编码方式

本文所提的 APZO-LSTM 模型, 采用实数编码形式, 图 2 为具有 3 层隐含层的 LSTM 神经网络结构, 并据此论述 APZO 算法的编码方式。

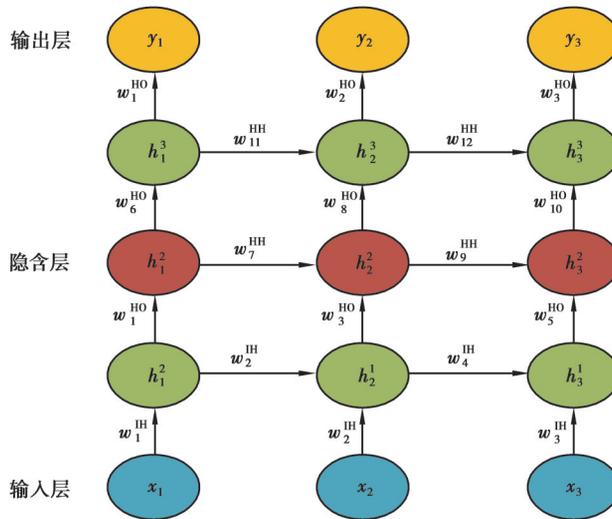


图 2 含 3 层隐含的 LSTM 神经网络结构图

Fig. 2 3-layer implicit LSTM neural network structure diagram

依据图 2 的 LSTM 网络结构图得到图 3, 表示 APZO-LSTM 模型中的某粒子编码方式, 如 w_1^{IH} 表示输入层第 1 个神经元与第 1 隐层的第 1 个神经元之间的权值, 以此类推, 粒子的编码方式囊括了所有网络节点的权值, 均为实数编码。

粒子	w_1^{IH}	...	w_3^{IH}
	w_1^{HH}	...	w_{12}^{HH}
	w_1^{HO}	...	w_3^{HO}

图 3 APZO-LSTM 模型编码图

Fig. 3 Model coding diagram of APZO-LSTM

3.2 APZO-LSTM 模型的适宜值函数

在 APZO-LSTM 模型中, 设计粒子的适宜值函数, 对种群中的粒子进行评价。不同于传统的 PSO 算法, 本文将 LSTM 模型输出值与实际值的均方根误差设计为粒子的适宜值函数, 衡量该粒子的重要性。均方根误差越小, 则该粒子对应的 LSTM 模型权值设置就越合理, 模型的泛化能力越强。

均方根误差,通常用 RMSE 表示。RMSE_i说明了粒子 *i* 在迭代过程中获取的全局最优解在训练集上的均方根误差值,也是 APSO 算法的适宜值函数。其计算过程如式(11)所示

$$RMSE_i = \sqrt{\frac{1}{Train_{num}} \sum_{j=1}^{Train_{num}} (d_{ij} - y_{ij})^2}, \quad (11)$$

式中:*n* 为粒子群的种群个数;Train_{num} 为训练集的数目;*d_{ij}* 是粒子 *i* 对应的 LSTM 模型的预测结果;*y_{ij}* 为实际值。

3.3 APSO-LSTM 模型流程实现

APSO-LSTM 网络是通过 APSO 对 LSTM 网络的初始权重值进行映射,将每个节点的各权重值都映射成粒子的某维属性,使每个粒子成为整个神经网络的候选权值解集合。运用 APSO 寻优网络,对网络的权值再全局调优,使神经元节点间权重的取值更合理,提高 LSTM 网络模型的预测性能和精度。图 4 为基于 APSO 优化的 LSTM 神经网络模型的流程图,*n* 为迭代次数,而 APSO-LSTM 模型的主要步骤说明如下:

- 1) 初始化 LSTM 神经网络及 APSO 算法。如 LSTM 网络的结构、节点数目、APSO 的种群数量、迭代次数等参数。
- 2) 训练 LSTM 神经网络,得出默认的最优权值。
- 3) 利用式(11)计算 APSO 种群中的各粒子 RMSE_i,构造寻优空间。
- 4) 搜寻全局最优粒子。将各粒子适宜值与最优粒子的适宜值比较,取适宜值较小者,设为新的全局最优粒子。
- 5) 依据式(8)~(10)对所有粒子的速度和位置予以更新。
- 6) 迭代次数 *n*+1,判断当前 *n* 是否大于迭代次数的最大值,若大于跳转步骤 7),否则跳转步骤 3)。
- 7) 若每次迭代中最优粒子的适宜值变化小于阈值,则中止迭代,否则继续迭代,直到满足算法停止条件。最后,输出全局最优粒子,该粒子对应 LSTM 网络最优的权值分布。

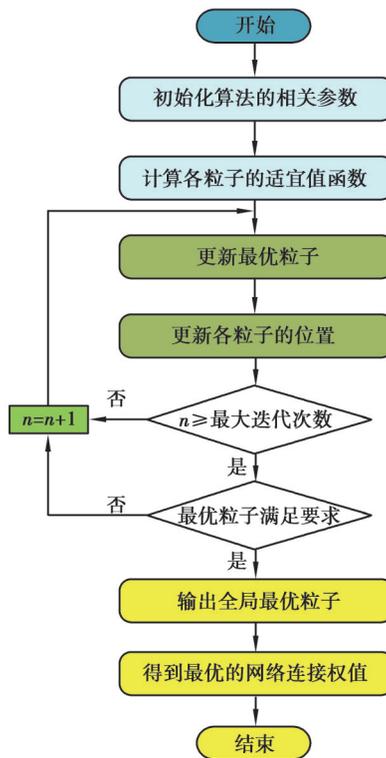


图 4 APSO-LSTM 模型流程图

Fig. 4 Model flowchart of APSO-LSTM

4 数值实验

建立了基于 LSTM、GA-LSTM、ACO-LSTM、PSO-LSTM 及 APSO-LSTM 的神经网络预测模型。GA-LSTM: 基于遗传算法优化权值 LSTM 模型; ACO-LSTM: 基于蚁群算法优化权值 LSTM 模型; PSO-LSTM: 基于 PSO 算法优化权值 LSTM 模型。通过对比各模型在不同数据集上的预测精度, 验证所提 APSO-LSTM 模型的性能。

实验的软件环境配置为: 深度学习框架 tensorflow1.10.0; 编程语言为 Python3; 计算机操作系统为 Ubuntu16.4。其硬件配置为: CPU 为 IntelCore i5-6300 3.20 GHz, 内存为 64 G, 显卡为 Nvidia GeForce GTX 1080Ti。

4.1 实验数据集

选用 DataMarket 中的洗发水销售数据集 (shampoo sales dataset, SSD)、日较低温度数据集 (minimum daily temperatures dataset, MDTD) 以及 UCI 数据库中的臭氧水平检测数据集 (ozone level detection dataset, OLD)。其中, SSD 刻画了 3 年期间的某洗发水销售数据, 实验主要预测其月销售量; MDTD 描述了墨尔本市连续 10 年的天气温度数据, 本实验主要对其日最低温度进行预测; OLD 数据集则记录了地面臭氧浓度数据, 本实验主要预测未来 1 段时间的臭氧浓度, 属于时序预测问题。

4.2 建立和训练预测模型

通过数据预处理、训练数据集及模型验证等步骤, 分别建立基于 GA-LSTM、ACO-LSTM、PSO-LSTM 及 APSO-LSTM 的神经网络预测模型。对预处理后的数据集, 取前 70% 作为训练集, 20% 作为验证集, 10% 为测试集。所建立 4 种模型的网络结构是一致的, 输入层和输出层均设为 1 层, 隐含层均设为 4 层。同时, 采用误差平方和 SSE 作为测试模型预测性能的指标, 如式(12)所示

$$SSE = \sum_{j=1}^n (y'_j - y_j)^2, \quad (12)$$

式中: n 为测试集的数目; y'_j 表示模型的预测结果; y_j 表示真实值。SSE 值越小, 则表明该模型的预测精度越高, 不同数据集下 3 种预测模型收敛速度对比分析如表 1 所示。

表 1 不同数据集下 3 种预测模型收敛速度

数据集	神经网络预测模型收敛步数(2%精度)		
	LSTM	PSO-LSTM	APSO-LSTM
SSD	341	310	280
MDTD	370	361	330
OLD	402	389	357
平均步数	371	353.3	332.3

由表 1 可知, 针对不同数据集, 以 5% 的精度为目标训练, APSO-LSTM 神经网络预测模型所需的收敛步数最少, 收敛速度最快。

4.3 实验结果及分析

神经网络模型训练完成后, 将上述 3 种数据集的测试集依次输入到各模型, 即可得到实验的仿真结果。表 2 为 GA-LSTM、ACO-LSTM、PSO-LSTM 及 APSO-LSTM 等 4 种模型在 SSD、MDTD、OLD 等数据集下的 SSE 值, 图 5 是依据表 2 绘制的折线图。由表 2 和图 5 可知, GA-LSTM 模型在 4 种不同数据集下的各 SSE 值均最大, ACO-LSTM 模型和 PSO-LSTM 模型次之, 而本文所提的 APSO-LSTM 模型在各项数据集上的 SSE 值均保持最低, 表明该模型的预测精度最高, 泛化能力与表征能力也更优。

图 6 表示在各数据集下 3 种模型的平均 SSE 值变化趋势。APSO-LSTM 模型的平均 SSE 值较 PSO-LSTM 模型下降 9.7%, 较 ACO-LSTM 模型下降 15.9%, 较 GA-LSTM 模型下降 22.6%。显然, 相对于同类的权值优化算法, APSO-LSTM 模型的预测性能更佳, 误差更小。ASPO 算法对 LSTM 神经网络的权值优化有促进作

用,易获得全局最优解。

表 2 不同数据集下 3 种预测模型的 SSE 值

Table 2 SSE values of 3 prediction models under different datasets

数据集	神经网络预测模型		
	LSTM	PSO-LSTM	APSO-LSTM
SSD	312.0	277.8	245.6
MDTD	292.7	267.1	243.2
OLD	369.4	330.3	301.8
平均 SSE 值	405.0	307.5	256.2

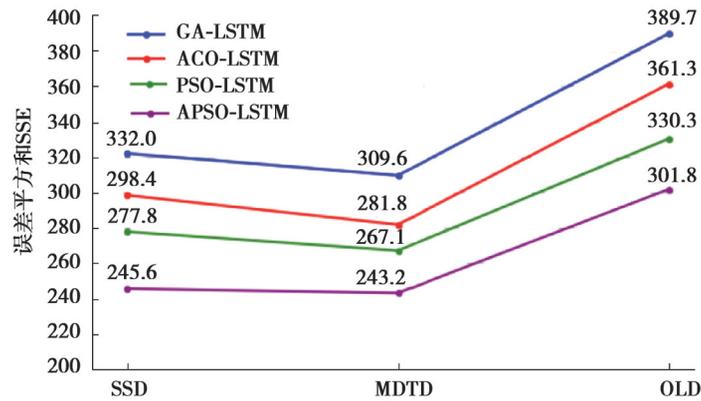


图 5 3 种模型 SSE 值对比图

Fig. 5 Comparison chart of SSE values for 3 models

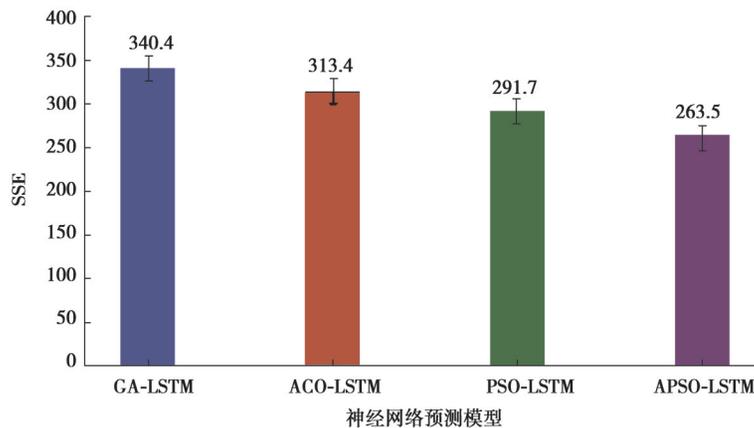


图 6 4 种模型的平均 SSE 值

Fig. 6 Average SSE values of 4 models

MDTD 数据集记录了墨尔本市连续 10 年的天气温度数据,本实验在测试过程中记录了墨尔本某月份连续 10 天日最低温度的实际值与 4 种模型的预测值,结果如表 3 所示。对比 APSO-LSTM 模型与其他 3 种模型的预测性能,增强实验说服力,图 7 与图 8 均是依据表 2 所绘的折线图。图 7 为 APSO-LSTM 模型关于墨尔本天气连续 10 天的预测结果。从图 7 可以看出,APSO-LSTM 模型预测曲线与实际温度曲线的最大误差不超过 1 °C,预测成效良好。

图 8 为 GA-LSTM、ACO-LSTM、PSO-LSTM 及 APSO-LSTM 等 4 种模型关于墨尔本天气连续 10 天的预测结果示意图。从图 8 可看出,GA-LSTM 模型的误差值最大,最大误差超过 3 °C,只有 APSO-LSTM 模型曲线最为接近实际值曲线,这进一步表明了所提 APSO-LSTM 模型的有效性和创新性。

表3 4种模型关于墨尔本天气的预测结果

天数	实际温度	GA-LSTM 预测值	ACO-LSTM 预测值	PSO-LSTM 预测值	APSO-LSTM 预测值
1	20.7	18.4	18.5	19.6	20.5
2	17.9	18.9	17.2	18.6	18.3
3	14.6	12.1	12.6	12.7	13.7
4	18.3	17.2	17.5	19	18.4
5	17.1	16	16.3	16.5	16.7
6	20.3	21.5	18.9	19.6	20.1
7	19.6	21.2	19.5	20.5	19.9
8	15.5	14.3	14.9	16.1	15.1
9	16.8	16.2	16.4	17.3	16.9

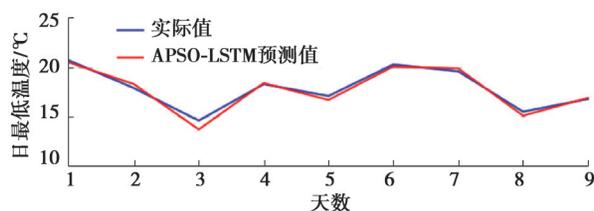


图7 APSO模型关于墨尔本天气的预测结果

Fig. 7 The prediction results of APSO model for Melbourne weather

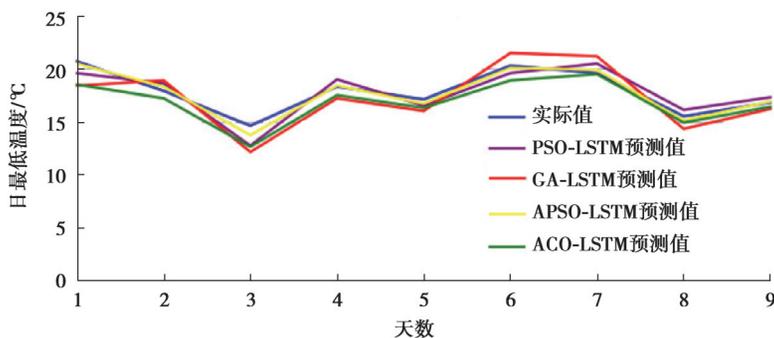


图8 4种模型关于墨尔本天气的预测结果

Fig. 8 Prediction results of four models for Melbourne weather

5 结论

对LSTM神经网络的优化已成为研究热点。本文针对标准LSTM神经网络中存在的收敛速度慢、权值的修正易陷入局部极值等问题,提出基于APSO优化的LSTM神经网络模型。实验结果表明,与GA-LSTM、ACO-LSTM、PSO-LSTM模型相比,APSO-LSTM模型在SSD、OLD及MDTD等3种数据集下的平均SSE值降低9.7%~22.6%,其预测误差最小,预测性能较同类算法有显著提升。因此,APSO-LSTM神经网络模型能充分学习数据中的信息关联,具有较强预测性能,也提供了新的神经网络优化思路。

参考文献

- [1] Wu J S, Long J, Liu M Z. Evolving RBF neural networks for rainfall prediction using hybrid particle swarm optimization and genetic algorithm[J]. *Neurocomputing*, 2015, 148: 136-142.
- [2] Graves A, Mohamed A R, Hinton G. Speech recognition with deep recurrent neural networks[C]//2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, Canada: IEEE, 2013: 6645- 6649.
- [3] Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks[C]//27th International Conference on Neural Information Processing Systems. New York: ACM, 2014: 3104-3112.
- [4] Meng F D, Lu Z D, Tu Z P, et al. Neural transformation machine: a new architecture for sequence-to-sequence learning[EB/OL]. (2015-06-22)[2020-06-12]. <http://arxiv.org/abs/1506.06442>.
- [5] Donahue J, Hendricks L-A, Rohrbach M, et al. Long-term recurrent convolutional networks for visual recognition and description[J]. *IEEE Transaction Pattern Analysis Machine Intelligence*, 2017, 39(4): 677-691.
- [6] Srivastava N, Mansimov E, Salakhutdinov R. Unsupervised learning of video representations using LSTMs[EB/OL]. (2015-02-16)[2020-06-12].<https://doi.org/10.48550/arXiv.1502.04681>.
- [7] Chatterjee S, Bandopadhyay S. Reliability estimation using a genetic algorithm-based artificial neural network:an application to a load-haul-dump machine[J]. *Expert Systems with Applications*, 2012, 39(12): 10943-10951.
- [8] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult[J]. *IEEE Transactions on Neural Networks*, 1994, 5(2): 157-166.
- [9] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [10] Li C B, Zhan G H, Li Z H. News text classification based on improved Bi-LSTM-CNN[C]//2018 9th International Conference on Information Technology in Medicine and Education (ITME). Hangzhou, China: IEEE, 2018: 8 90-893.
- [11] Song B N, Fan C X, Wu Y X, et al. Data prediction for public events in professional domains based on improved RNN-LSTM [J]. *Journal of Physics: Conference Series*, 2018, 976: 012007.
- [12] Lin M L, Chen C X. Short-term prediction of stock market price based on GA optimization LSTM neurons[C]//2018 2nd International Conference o Deep Learning Technologies. New York: ACM, 2018: 66-70.
- [13] ElSaid A, Higgins J. Optimizing long short-term memory recurrent neural networks using ant colony optimization to predict turbine engine vibration[J]. *Applied Soft Computing*, 2018, 73: 969-991.
- [14] Yang X S. Nature-inspired metaheuristic algorithms [M]. [s.l]: Luniver press, 2010.
- [15] Kennedy J, Eberhart R. Particle swarm optimization[C]//Proceedings of ICNN'95-International Conference on Networks Neural. Perth, Australia: IEEE, 1995: 194 2-1948.
- [16] Gandomi A H, Alavi A H. Krill herd: a new bio-in spired optimization algorithm[J]. *Communications in Nonlinear Science and Numerical Simulation*, 2012, 17(12): 4831-4845.
- [17] 夏学文, 刘经南, 高柯夫, 等. 具备反向学习和局部学习能力的粒子群算法[J]. *计算机学报*, 2015, 38(7):1397-1407.
Xia X W, Liu J N, Gao K F, et al. Particle swarm optimization algorithm with reverse-learning and local-learning behavior[J]. *Chinese Journal of Computers*, 2015, 38(7): 1397-1407. (in Chinese)
- [18] 霍林, 陆寅丽. 改进粒子群算法应用于 Android 恶意 应用检测[J]. *计算机工程与应用*, 2020, 56(7): 96-
Huo L, Lu Y L. Improved particle swarm optimization for android malware detection[J]. *Computer Engineering and Applications*, 2020,56(7): 96-101. (in Chinese)

(编辑 侯 湘)