

基于 LSTM 循环神经网络的核电设备状态预测

龚 安,马光明,郭文婷,陈 臣

(中国石油大学(华东)计算机与通信工程学院,山东 青岛 266580)

摘 要:核电站的规模随着经济的发展日益扩大,核电设备运行状态的研究已成为数据挖掘的重要研究领域。核电设备是高可靠性和高安全性的复杂系统,多年的设备运行产生了大量的时间序列数据。为了解决核电设备运行状态难以准确预测等问题,提出了一种基于长短期记忆(long short-term memory, LSTM)循环神经网络的核电设备状态预测方法。首先除去原数据中噪点明显的的数据,然后使用 z-score 标准化方法对数据进行预处理,然后实现 LSTM 的网络结构设计、网络训练和预测,最后对预测结果进行比较分析。考虑到核电设备各个部件运行产生的数据种类繁多,选择与核电设备运行状态相关的重要数据主泵电机绕组温度作为研究对象。通过与 GRU、RNN 等模型进行对比实验,表明了该算法对核电设备的运行状态有更高的预测精度。

关键词:核电设备;时间序列数据;循环神经网络;状态预测;深度学习

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2019)10-0041-05

doi:10.3969/j.issn.1673-629X.2019.10.009

Nuclear Power Equipment Status Prediction Based on LSTM Recurrent Neural Network

GONG An, MA Guang-ming, GUO Wen-ting, CHEN Chen

(School of Computer & Communication Engineering, China University of Petroleum, Qingdao 266580, China)

Abstract: Nuclear power plants grow in size as the economy grows. The research on the running state of nuclear power equipment has become an important field of data mining. Nuclear power equipment is a complex system with high reliability and security. Years of equipment operation have produced a large amount of time series data. In order to solve the problem that it is difficult to predict the operational status of nuclear power equipment accurately, a nuclear power equipment state prediction method based on long-term memory (LSTM) recurrent neural network is proposed. Firstly, the data with obvious noise points in the original data are removed, and then pre-processed by z-score standardization method. After that, the network structure design, network training and prediction of LSTM are implemented. Finally, the predicted results are compared and analyzed. Considering the variety of data reflecting the operating status of nuclear power equipment, choose the main pump motor winding temperature related to the operating status of nuclear power equipment as a experiment object. Through comparison experiments with GRU and RNN, the proposed algorithm has higher prediction accuracy for the operating status of nuclear power equipment.

Key words: nuclear power equipment; time series data; recurrent neural networks; state forecast; deep learning

0 引 言

随着中国经济的迅速发展,核电站的发展规模日益扩大^[1],为保证对核电系统安全状态的及时掌控,核电设备运行状态的预测变得越来越重要。然而,基于核电数据的保密性等原因,通过核电数据对核电设备运行状态进行预测,还未发现相关的研究,因此研究核电设备运行状态的预测方法很有意义。

核电系统属于复杂结构系统,结构越复杂,系统数据的复杂性越大。普通的预测方法如灰色预测模型^[2]、比例风险模型^[3]、非线性滤波器模型^[4]都局限于因果回归,并不能有效反映时序数据和设备运行状态之间的关系。

近年来,深度学习模型在时序数据的研究越来越深入^[5],其中 LSTM 神经网络^[6]通过设备产生的时序

收稿日期:2018-11-05

修回日期:2019-03-07

网络出版时间:2019-04-24

基金项目:国家油气重大专项(2017ZX05013-001)

作者简介:龚 安(1971-),男,副教授,硕导,CCF 会员(62929M),研究方向为大数据智能处理;马光明(1993-),男,硕士研究生,研究方向为数据挖掘。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20190424.1055.062.html>

数据预测其运行状态已经取得了不少成功的案例。例如, Noriaki Hirose 等提出基于 LSTM 的滚动摩擦预测模型^[7], 对机械系统的滚动摩擦进行研究, 实验结果精确预测了设备的滚动摩擦系数; Chen Zaifa 等将 LSTM 模型和经验模式分解算法相结合, 用 LSTM 模型处理轴承数据的本征模函数分量, 并将实验结果与支持向量机相对比, 更好地实现了机械状态的单步预测^[8]; AbdElRahman ElSaid 等使用 LSTM 模型实现了飞机发动机机振动值的预测^[9], 并由此建立了飞机发动机状态预警系统, 避免了过渡振动对飞机的不利影响; Zhao Rui 等使用 LSTM 获取磨损数据的依赖关系, 提出基于 LSTM 评估的生活工具健康检测系统^[10], 并通过对比实验展示了 LSTM 模型的优越性。

在上述研究的基础上, 文中采用 LSTM 神经网络对核电设备运行状态进行预测, 选择某核电站数据中的主泵绕组温度作为实验数据进行对比实验, 包括数据预处理、网络结构设计、神经网络的训练和预测等算法的实现。实验结果表明该方法对核电设备运行状态有更高的预测精度。

1 相关理论

1.1 数据预处理

数据预处理是数据挖掘的重要步骤, 同时也是必不可少的一环。为了更有效地挖掘出知识, 必须为数据挖掘模型提供简洁、干净、有效的数据。然而, 实际应用系统中获得的原始数据通常含有脏数据, 存在杂乱性、重复性、不完整性等方面的问题。

通过数据预处理对应的模块, 以领域知识作为指导, 对原始数据进行处理, 摒弃与数据挖掘目标无关的属性, 为数据挖掘的核心算法模型提供更准确和更有针对性的数据, 以提高数据挖掘的效率和准确度。

数据预处理相关理论与方法很多, 然而, 对于复杂设备时序数据的预处理, 目前还没有相关理论方法的深入研究。为了提高实验精度, 除了手动除去噪点之外, 实验采用 z-score 标准化方法对实验数据进行预处理, 其公式可表示为:

$$z = (x - \mu) / \sigma \quad (1)$$

x 为具体的实验数值; μ 为平均数; σ 为标准差。

1.2 LSTM 神经网络的相关理论

循环神经网络在网络结构设计中引入时序数据的概念, 使其在时序数据处理中具有更强的适应性。RNN (recurrent neural network) 虽然能够有效处理时间序列数据, 但仍存在以下两个问题: (1) RNN 存在梯度消失和梯度爆炸问题, 无法有效处理时间维度过长的时间序列数据; (2) RNN 模型的训练过程中需要预先设定延迟窗口的长度, 但是这一参数的最优值难以

获取。

LSTM 神经网络是众多 RNN 的变体之一, 在普通 RNN 的基础上, 在各隐藏层神经单元中加入记忆单元, 实现时间序列上的记忆信息可控。每次序列数据在隐藏层各单元之间传递时会通过遗忘门、输入门、输出门等可控门, 控制时序数据中之前数据和当前数据的记忆和遗忘程度, 从而使神经网络具有长期记忆功能。通过这种方式, LSTM 神经网络有效改善了循环神经网络梯度消失、长期记忆力不足等缺点^[11-12], 能够有效地利用长距离的时序数据。

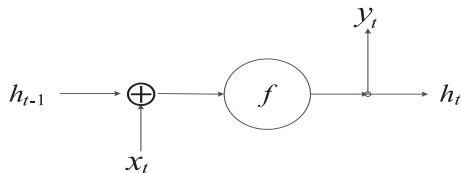


图1 RNN 隐藏层细胞结构

对于给定的时间序列 $\vec{x} = (x_1, x_2, \dots, x_n)$, 使用 RNN 模型, 如图 1 所示, 应用式 2 和式 3 可以计算出 $\vec{h} = (h_1, h_2, \dots, h_n)$ 和输出序列 $\vec{y} = (y_1, y_2, \dots, y_n)$ 。

$$h_t = f(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (2)$$

$$y_t = W_{hy}h_t + b_y \quad (3)$$

其中, W 表示权重系数矩阵; b 表示偏置向量; f 表示激活函数。

LSTM 神经网络的神经元细胞结构如图 2 所示^[13], 计算公式可表示为:

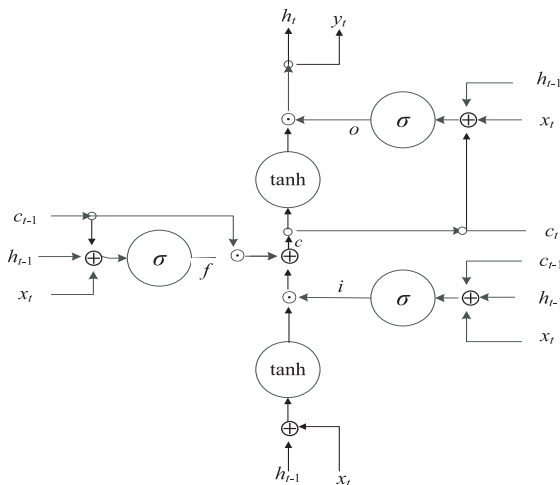


图2 LSTM 隐藏层细胞结构

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (4)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (5)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (6)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (7)$$

$$h_t = o_t \tanh(c_t) \quad (8)$$

其中, i, f, c, o 分别表示输入门、遗忘门、细胞状态、输出门; W 和 b 分别对应权重系数矩阵和偏执系数; σ 和 \tanh 分别对应 sigmoid 和双曲正切激活

函数。

训练过程可分为四个步骤:

- (1) 由式 4 ~ 式 8 计算 LSTM 细胞的输出值;
- (2) 计算每个 LSTM 细胞的误差;
- (3) 根据误差计算每个权重的梯度;
- (4) 应用基于梯度的优化算法更新权重。

LSTM 有许多变体,最成功的是门限循环单元 (gated recurrent unit, GRU)。GRU 模型保留了 LSTM 的长期记忆能力,将 LSTM 细胞中的输入门、遗忘门、输出门替换为更新门和重置门,并将细胞状态和输出两个向量合二为一。GRU 结构相对简单,参数数量少,训练效率较高,在实际应用中,两者的可比性很强。

1.3 梯度下降算法

梯度下降算法是机器学习领域使用最广泛的优化算法,在许多流行的深度学习库中都包含了不同版本的梯度下降算法的实现。基于梯度下降的优化算法有多种,比如 Adam^[14-15]、AdaGrad^[16]、RMSProp^[17] 等。文中选用随机梯度下降算法 (stochastic gradient descent, SGD)^[18]。SGD 算法的优化过程为:在给定的样本 M 中,随机取出副本 N 代替原始样本 M 来作为全集,对模型进行训练。这种训练方式优势明显,如果样本抽取在合适范围内,既会求出结果,并且速度更快。

SGD 算法能够有效解决训练过程随实验样本数量加大而变得异常缓慢的问题,有效提高收敛速度并减少占用的设备资源。

2 构建 LSTM 神经网络的预测模型

根据设备运行数据的特点,结合第一节提到的相关理论,给出基于 LSTM 模型的核电设备运行状态预测方法,包括模型的搭建、训练和预测。

2.1 预测模型框架的搭建

考虑到实验数据是单变量时间序列数据,构建的 LSTM 神经网络预测模型如图 3 所示。该模型由五个功能模块组成,包括输入层、隐藏层、输出层、网络训练和网络数据预测。输入层对实验数据集进行初步处理:划分数据集,数据标准化,数据分割等以满足 LSTM 神经网络的输入要求;隐藏层用 LSTM 细胞搭建循环神经网络;遗忘门决定保留多少上一时刻的神经元状态到当前时刻神经元状态,输入门决定保留多少当前时刻的输入到当前时刻的神经元状态,输出门决定当前时刻的神经元状态输出的多少;输出层提供预测结果。网络训练使用 SGD 算法更新权重,在保证精度的前提下提高了优化效率。网络数据预测使用迭代的方法进行数据预测。

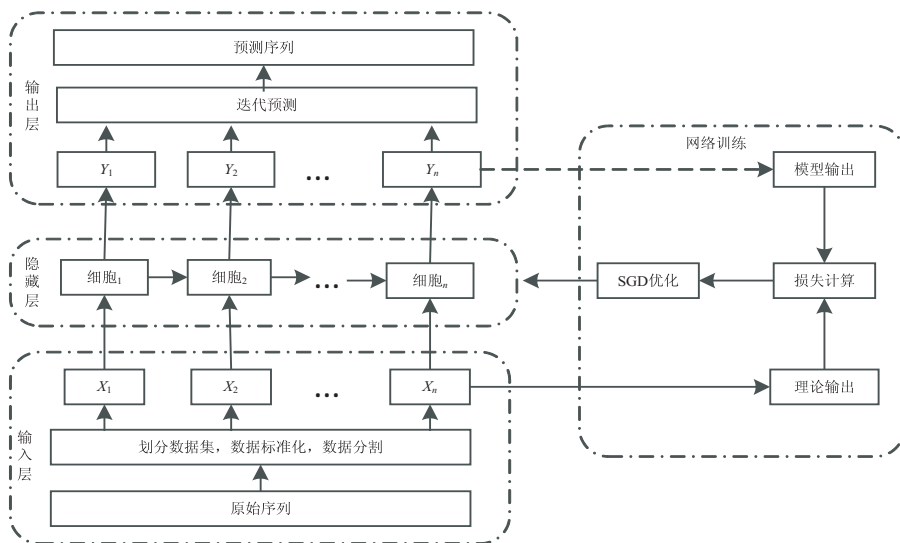


图 3 LSTM 模型框架

2.2 LSTM 模型的训练和预测

首先,定义原始的设备状态时间序列数据为:

$$f = \{f_1, f_2, \dots, f_n\} \quad (9)$$

为了适应隐藏层输入的特点,应用数据分割方法对数据集 f 进行处理。设定分割窗口长度 (数据步长) 为 L , 分割后的数据集为:

$$F = \{F_0, F_1, \dots, F_n\} \quad (10)$$

$$F_p = \{f_{p*L+1}, f_{p*L+2}, \dots, f_{p*L+L}\}, 0 \leq p \leq t \quad (11)$$

然后,将实验划分为训练集 train_x , train_y 和测

试集 test_x , test_y 两部分。

训练集可表示为:

$$F_{\text{train_x}} = \{f_{p*L+1}, f_{p*L+2}, \dots, f_{p*L+m}\} \quad (12)$$

$$F_{\text{train_y}} = \{f_{p*L+m+1}, f_{p*L+m+2}, \dots, f_{p*L+L}\}, 1 < m < t \quad (13)$$

测试集可表示为:

$$F_{\text{test_x}} = \{f'_{p*L+1}, f'_{p*L+2}, \dots, f'_{p*L+m}\} \quad (14)$$

$$F_{\text{test_y}} = \{f'_{p*L+m+1}, f'_{p*L+m+2}, \dots, f'_{p*L+L}\}, 1 < m < t \quad (15)$$

train_y, test_y 分别为 train_x, test_x 的理论输出。
接下来将训练集 train-x, train-y 输入到隐藏层神经网络。实验设定训练集与测试集的大小比值为 9 : 1, 则模型隐藏层输出 C 和 train-y 均为 (0.9 * t, L - m) 的二维数组。选择均方误差作为误差计算公式, 训练过程的损失函数可表示为:

$$\text{loss} = \sum_{i=1}^{(0.9 * t) * (L - m)} (c_i - y_i)^2 / (0.9 * t) * (L - m)$$

(16)

把损失函数最小化作为实验的优化目标, 选择不同的学习率和训练步长, 应用 SGD 优化算法不断更新权重, 得到最终的预测模型。
应用训练好的 LSTM 模型进行预测, 将测试集 test_x 输入到模型, 利用测试集 test_y 作为理论输出和模型的实际输出来计算模型的预测精度。

3 实验验证

本节使用核电数据作为实验数据集, 应用第一节的相关理论和第二节的模型构建方法展开实验验证, 具体包括数据准备、实验结果、对比实验分析三部分。

3.1 实验数据集

实验采用主泵绕组温度作为实验对象。主泵绕组温度是反映核电设备运行状态的重要参数, 主泵绕组温度的高低决定着核电设备实时的运行状态。数据时间范围: 从 2010 年 1 月 17 日 13 点 56 分到 2017 年 02 月 16 日 15 点 0 分, 数据量: 采样频率为 1 分钟 1 条数据, 总数据量为 3 725 343。

部分时间序列数据如图 4 所示。

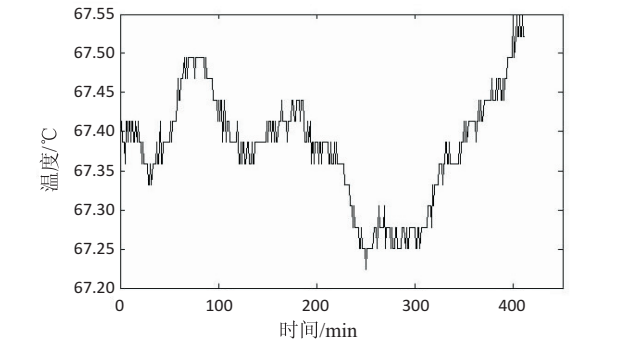


图 4 主泵电机绕组温度

3.2 度量指标

误差度量方式的选择对模型误差的计算有很大影响。常见的误差度量方式有: 均方误差 (mean squared error, MSE), 计算参数估计值与参数真值之差平方的期望值; 均方根误差 (root mean squared error, RMSE), 计算均方误差的算术平方根; 平均绝对误差 (mean absolute error, MAE), 计算绝对误差的平均值。对于预测模型精度的计算, 通常选择平均绝对误差作为度

量标准。MAE 的计算公式可表示为:

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |y_t - f_t|$$

(17)

其中, f_t 和 y_t 分别为设备运行状态时间序列数据在 t 时刻的实际值和模型的输出值; T 为数据点个数。通过对测试集计算 MAE 值来评估模型的预测精度。

3.3 实验平台和环境

实验所使用的计算机配置如下: 处理器为 (英特尔) Intel (R) Core (TM) i7-4720HQ CPU @ 2.60 GHz (2601 MHz); 内存为 8.00 GB; 操作系统为 Ubuntu 16.04; 程序设计语言为 Python3.5.2; 集成开发环境为 PyCharm Community Edition 2016.3; 实验程序中所用到的 RNN, LSTM 和 GRU 模型由 Python 的 TensorFlow 程序包实现。

3.4 实验结果

3.4.1 获取精度较高的模型参数

根据经验确定模型参数。分别选用 5、10、15 和 20 作为数据步长, 选用不同的学习率 ($\eta = 0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1$) 来训练 LSTM 模型。通过测试数据来检验模型精度, 实验结果如表 1 所示。

表 1 不同参数下 LSTM 模型的 MAE

学习率	L = 5, MAE	L = 10, MAE	L = 15, MAE	L = 20, MAE
0.001	0.043 4	0.042 7	0.043 0	0.045 0
0.002	0.037 9	0.037 1	0.040 3	0.040 8
0.005	0.035 7	0.035 4	0.034 5	0.036 3
0.010	0.034 6	0.034 6	0.034 2	0.033 8
0.020	0.031 8	0.033 9	0.030 5	0.031 1
0.050	0.036 1	0.036 9	0.038 6	0.036 9
0.100	0.055 8	0.057 2	0.059 0	0.061 0

由表 1 可看出, 当数据步长为 15、学习率 $\eta = 0.02$ 时, LSTM 模型的预测精度最高, MAE 为 0.030 5。

3.4.2 LSTM 模型与 RNN 模型和 GRU 模型的对比

为了验证 LSTM 模型在不同类型的循环神经网络中的优势, 实验中改变模型的隐藏层神经单元, 构建 RNN 和 GRU 网络模型, 并采用相同的多组参数对 RNN 模型和 GRU 模型进行训练, 得到预测精度最高的 RNN 模型和 GRU 模型。实验结果表明, RNN 模型预测精度最高时 MAE = 0.034 5, GRU 模型预测精度最高时 MAE = 0.032 6。

LSTM 与 RNN 对比结果如图 5 所示。

LSTM 与 GRU 对比结果如图 6 所示。

实验结果表明, LSTM 模型的预测精度明显高于 RNN 模型, 略高于 GRU 模型。

4 结束语

文中深入探究了设备状态预测的主要方法, 针对

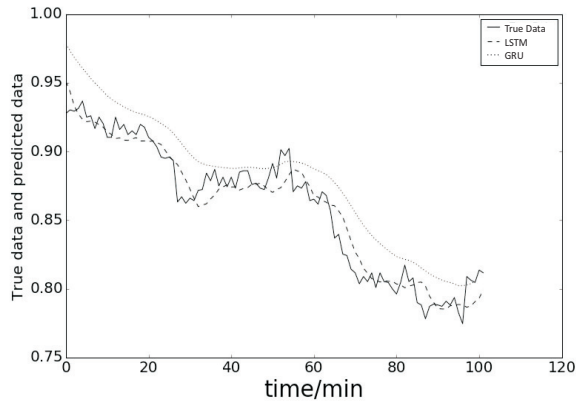


图 5 LSTM 与 RNN 对比结果

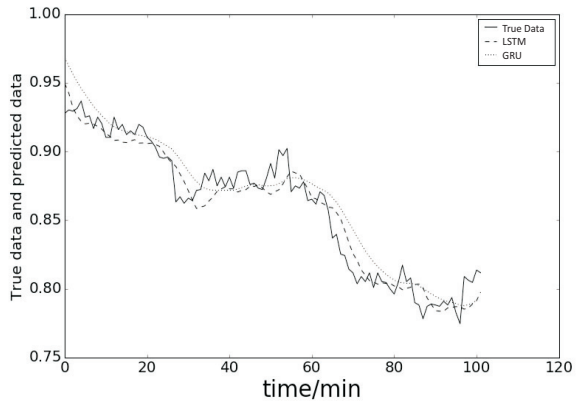


图 6 LSTM 与 GRU 对比结果

所获得的核电数据的特点和核电设备运行状态预测的需求,提出了基于 LSTM 神经网络的核电设备运行状态预测方法。对比实验结果表明:基于 LSTM 模型的预测方法与 RNN 模型和 GRU 模型相比,有更高的预测精度;LSTM 模型在训练过程中模型精度对学习率的取值较为敏感,过高和过低都会影响模型的预测性能;与其他循环神经网络(RNN 和 GRU)相比,LSTM 的预测精度更高,但训练过程耗时更多。

总的来说,验证了 LSTM 模型在核电数据分析领域的适用性,扩展了深度学习的应用范畴。基于当前的工作,后续将从以下方面继续研究:扩展隐藏层的数目,检验多隐藏层的实验结果;寻找有效的参数优化方法,提高参数优化效率和预测精度。

参考文献:

[1] 徐小杰,程覃思. 我国核电发展趋势和政策选择[J]. 中国能源,2015,37(1):5-9.

[2] 曹艳华,陈春良,宋敬华. 装甲装备自主式保障关键要素分析[J]. 装甲兵工程学院学报,2010,24(2):7-11.

[3] CHEN X, JIA Y, ZHANG Y. An optimal replacement model for condition based maintenance[C]//Proceedings of international conference on maintenance engineering. Beijing: Science Press, 2006:757-761.

[4] WANG W, JIA Y. A multiple condition information sources

based maintenance model and associated prototype software development[C]//Proceedings of 14th condition monitoring and diagnostic engineering management. [s. l.]: [s. n.], 2001:889-898.

[5] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553):436-444.

[6] GRAVES A. Long short-term memory[M]. Berlin: Springer, 2012:1735-1780.

[7] HIROSE N, TAJIMA R. Modeling of rolling friction by recurrent neural network using LSTM[C]//IEEE international conference on robotics and automation. Singapore: IEEE, 2017:6471-6478.

[8] CHEN Zaifa, LIU Yancheng, LIU Siyuan. Mechanical state prediction based on LSTM neural network[C]//Proceedings of the 36th Chinese control conference. Dalian, China: [s. n.], 2017:26-28.

[9] ELSAID A E R, WILD B, HIGGINS J, et al. Using LSTM recurrent neural networks to predict excess vibration events in aircraft engines[C]//IEEE 12th international conference on e-science. Baltimore, MD, USA: IEEE, 2016:260-269.

[10] ZHAO Rui, WANG Jinjiang, YAN Ruqiang, et al. Machine health monitoring with LSTM networks[C]//10th international conference on sensing technology. Nanjing, China: IEEE, 2016:1-6.

[11] MA Xiaolei, TAO Zhimin, WANG Y, et al. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data[J]. Transportation Research Part C Emerging Technologies, 2015, 54:187-197.

[12] SRIVASTAVA N, MANSIMOV E, SALAKHUTDINOV R. Unsupervised learning of video representations using LSTMs[C]//Proceedings of the 32nd international conference on machine learning. Lille, France: JMLR, 2015:843-852.

[13] GREFF K, SRIVASTAVA R K, KOUTNIK J, et al. LSTM: a search space odyssey[J]. IEEE Transactions on Neural Networks & Learning Systems, 2017, 29(10):2222-2232.

[14] 王鑫, 吴际, 刘超, 等. 基于 LSTM 循环神经网络的故障时间序列预测[J]. 北京航空航天大学学报, 2018, 44(4):772-784.

[15] KINGA D, ADAM J B. A method for stochastic optimization[C]//3rd international conference for learning representations. San Diego: [s. n.], 2015.

[16] DUCHI J, HAZAN E, SINGER Y. Adaptive subgradient methods for online learning and stochastic optimization[J]. Journal of Machine Learning Research, 2011, 12(7):2121-2159.

[17] YEUNG S, RUSSAKOVSKY O, NING J, et al. Every moment counts: dense detailed labeling of actions in complex videos[J]. International Journal of Computer Vision, 2018, 126(2-4):375-389.

[18] AMARI S I. Backpropagation and stochastic gradient descent method[J]. Neurocomputing, 1993, 5(4-5):185-196.