



基于自编码器与随机森林的轨道交通 电机轴承剩余寿命预测

李 伟^{1,2}

(1. 中车永济电机有限公司, 山西 永济 044502;
2. 轨道交通牵引电机山西省重点实验室, 山西 永济 044502)

摘 要: 针对轨道交通牵引电机滚动轴承寿命预测中故障样本稀缺、全寿命周期数据难以获取的问题, 文章提出一种结合自编码器(AE)与随机森林(RF)的轴承剩余使用寿命(RUL)预测方法。首先, 利用自编码器对非故障轴承进行无监督特征学习, 通过构造误差进行样本扩充, 以缓解故障样本不足造成的影响; 其次, 构建随机森林回归模型实现剩余寿命预测, 并使用均方误差(MSE)损失函数评价预测性能; 最后, 构建基于RF的多步回归预测模型, 通过迭代实现对轴承的RUL预测。采用某牵引电机的滚动轴承全寿命试验数据对该方法的准确性进行验证, 结果表明, 该方法可较准确预测滚动轴承的剩余寿命, 为轨道交通牵引电机轴承寿命预测提供了一种思路。

关键词: 滚动轴承; 剩余使用寿命; 自编码器; 随机森林; 多步预测

中图分类号: U264.1; TH133.33 **文献标志码:** A **文章编号:** 1000-128X(2026)01-0073-06

0 引言

牵引电机长时间工作于高载荷、强振动等严苛环境中, 滚动轴承极易发生疲劳损伤, 若未及时发现并更换故障轴承, 极有可能导致电机故障乃至车辆运行中断。因此, 开展对牵引电机滚动轴承剩余使用寿命(remaining useful life, RUL)的准确预测, 对保障列车运行安全、降低维修成本和提高运维效率具有重要意义。

近年来, 机器学习方法被广泛应用于滚动轴承的性能退化趋势预测^[1-2]。滚动轴承的退化包括无故障运行阶段和故障运行阶段, 前一阶段的各项监测指标会在小范围内波动, 当产生故障后, 各项监测指标急剧变化, 直至轴承完全失效^[3]。文献[4]利用希尔伯特熵有效识别了轴承的退化状态, 并建立了基于支持向量机的RUL预测模型; 文献[5]提出了一种基于随机过程的多阶段RUL预测方法, 使用卡尔曼滤波等方法计算不

同退化阶段下轴承的RUL预测结果; 文献[6]提出了基于均方谐波噪声比指标和改进正则化粒子滤波相结合的方法预测轴承的RUL; 为了提高轴承RUL预测精度, 文献[7]提出一种基于多尺度局部核回归异常值检测的退化阶段识别方法; 文献[8]的研究表明各寿命阶段样本内部不具备关联性, 并提出了一种改进均衡分布适配的退化阶段识别及预测方法; 文献[9]提出了一种基于状态划分和集成学习模型的RUL预测方法。值得指出的是, 现有方法虽然在不同角度取得了进展, 但大多数方法都依赖于完整的全寿命周期振动数据样本^[10-12]。实际应用中, 尤其是在轨道交通牵引电机领域^[13-14], 故障样本数量有限、获取代价高昂, 直接限制了现有方法的工程应用。因此, 如何在少样本或非完整退化数据条件下实现轴承寿命预测, 仍是当前研究中的关键挑战。

为此, 本文提出一种基于自编码器(autoencoder, AE)与随机森林(random forest, RF)相结合的牵引电机滚动轴承RUL预测方法。利用AE模型计算轴承样本

基金项目: 中国国家铁路集团有限公司科技研究开发计划(K2023J051)

通信作者: 李 伟, 男, 高级工程师, 主要从事复杂工况下电机轴承应用技术研究; E-mail: 156789934@163.com

引用格式: 李伟. 基于自编码器与随机森林的轨道交通电机轴承剩余寿命预测[J]. 机车电传动, 2026(1): 73-78. DOI:10.13890/j.issn.1000-128X.2026.01.010.

Citation: LI Wei. Remaining life prediction of rail transit motor bearings based on auto-encoder and random forest[J]. Electric drive for locomotives, 2026(1): 73-78. DOI:10.13890/j.issn.1000-128X.2026.01.010.

特征的均方误差(mean squared error, MSE), 以此判断轴承的故障状态; 随后构建随机森林回归模型并采用多步迭代策略, 实现轴承 RUL 的精准预测。

1 轴承寿命预测总体思路

滚动轴承在无故障运行阶段的各项监测指标呈现随机的小幅波动特点^[15], 如图 1 所示。若在该阶段对滚动轴承 RUL 进行预测, 会产生较大的误差。因此, 在开展寿命预测之前, 首先需准确识别轴承是否已进入故障退化阶段; 当轴承处于故障阶段后再启动 RUL 预测进程, 可有效提升预测的准确性与实用性。

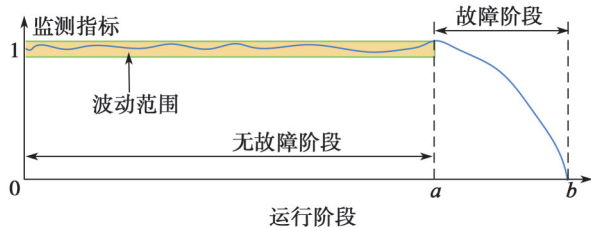


图1 轴承全寿命周期的特征变化

Fig. 1 Characteristic changes throughout lifecycle of bearings

2 基于AE的轴承故障识别

对轨道交通牵引电机滚动轴承进行故障识别的难点在于, 受其高安全性要求的限制, 无法直接大量获取其在线运行的故障振动数据。这一问题导致无法直接构建完备的故障样本库以支持模型的训练与测试。为应对上述挑战, 本文采用基于无监督学习的AE进行建模, 从而在缺乏故障样本的条件下实现对滚动轴承运行状态的有效判别。

2.1 AE模型及其故障识别原理

AE模型由编码器和解码器两部分组成, 典型结构如图2所示。工作过程分为2个步骤:

①编码过程: 将输入样本 x 映射到潜在空间 z , 可表示为

$$z = \sigma(Wx + b) \quad (1)$$

式中: W 为编码器权重矩阵; b 为编码器偏置项; σ 为编码器激活函数。

②解码过程: 对潜在空间变量 z 进行反映射, 得到输出样本 $\hat{x}$ 为

$$\hat{x} = \sigma'(W'z + b') \quad (2)$$

式中: W' 为解码器权重矩阵; b' 为解码器偏置项; σ' 为解码器激活函数。

AE训练包含前向传播、误差计算、反向传播及优化收敛4个步骤。模型训练完成后, 对于异常样本的重构误差会明显增大。评价AE训练结果性能的常用损失函数为均方误差(MSE), 其表达式为

$$L(x, \hat{x}) = \|x - \hat{x}\|^2 \quad (3)$$

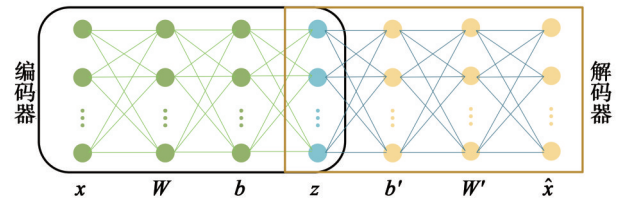


图2 自编码器典型结构

Fig. 2 Typical structure of AE

2.2 基于AE的样本扩充

轨道牵引电机滚动轴承在装车运行后, 短期内积累的振动数据样本量较少, 为获取更好的AE训练效果, 需要不断扩充其正常运行的振动数据样本。如图3所示为本文所提方法的非故障样本集扩充流程, 详细步骤为:

- ① 运行前期的 m 个振动样本作为基本的非故障样本集;
- ② 基于非故障样本集训练AE模型;
- ③ 将第 $m+1$ 个样本作为AE模型输入, 计算MSE;
- ④ 若MSE处于限值内, 则将第 $m+1$ 个样本纳入非故障样本集, 完成样本扩充;
- ⑤ 采用新的样本集对AE模型进行训练并更新;
- ⑥ 按同样的步骤, 依次对后续样本重复步骤③~步骤⑤;
- ⑦ 直至第 n 个样本, 计算MSE超限, 判断轴承故障, 非故障样本集扩充完毕。

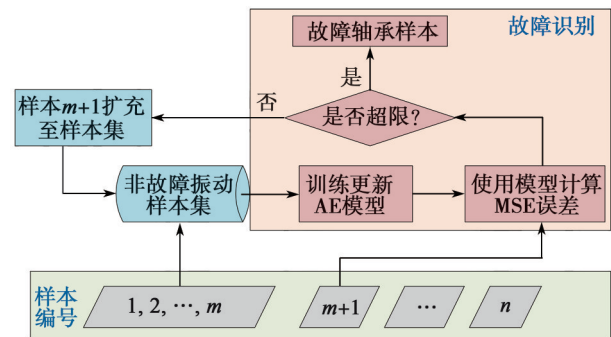


图3 非故障样本集扩充流程

Fig. 3 Fault-free sample set expansion process

非故障振动样本扩充完毕后, AE模型所依赖的训练数据更加全面且具有代表性, 能够更准确地刻画滚动轴承在正常工况下的特征分布, 提升模型对微小波动的鲁棒性, 从而有效降低无故障运行阶段监测指标随机小幅波动对故障识别准确性的干扰, 提高模型判别故障状态的可靠性与精度。

2.3 滚动轴承故障识别

基于AE在异常数据识别方面的优越性能, 本文设计并提出了一种基于AE的滚动轴承故障识别方法, 整

体流程如图4所示。

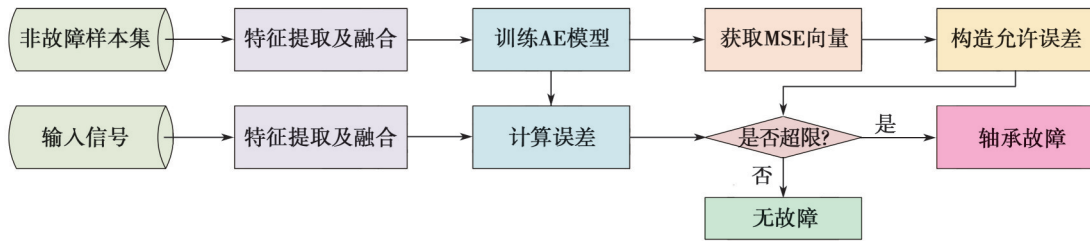


图4 基于AE的滚动轴承故障识别流程

Fig. 4 AE-based fault identification process for rolling bearings

具体步骤如下:

- ① 对非故障轴承的 k 个振动样本作为 AE 模型的训练样本;
- ② 对 AE 模型训练并获取其 MSE 向量 $L_i, i=1, 2, 3, \dots, k$, 构造允许误差为:

$$\left. \begin{aligned} E_{th} &= \left\{ E \mid u - 3\sigma_L \leq E \leq u + 3\sigma_L \right\} \\ u &= \frac{1}{k} \sum_{i=1}^k L_i \\ \sigma_L &= \sqrt{\frac{1}{k-1} \sum_{i=1}^k \left(L_i - \frac{1}{k} \sum_{i=1}^k L_i \right)^2} \end{aligned} \right\} \quad (4)$$

式中: μ, σ_L 分别为向量 L_i 的均值及标准差。

- ③ 将第 $k+1$ 个样本作为被识别样本输入 AE 模型, 获得计算误差 E_{k+1} ;

- ④ 通过 E_{k+1} 与允许误差 E_{th} 定义的范围判断轴承的故障状态。

3 基于RF的轴承寿命预测

牵引电机轴承产生故障后, 其 RUL 呈强非线性衰减趋势。RF 预测模型在揭示非线性系统演化规律方面具有独特的优势^[16-17], 本文将构建改进的随机森林多步迭代预测模型, 下面将具体分析具体原理与实现过程。

3.1 随机森林回归原理

对于具有 M 个特征的输入数据集 x 及输出数据集 y , 随机森林回归流程为:

- ① 采用 Bootstrap 方法对数据集 x 进行有放回抽样, 共抽取 s 次, 构成 s 个样本的训练集, 对应 s 个决策树;
- ② 随机选取 P 个特征, 根据误差最小化原则进行特征选择优化, 生成决策树, 其中误差计算选择 MSE;
- ③ 对所有决策树的预测结果取平均值, 作为最终的预测值。

3.2 改进随机森林寿命预测流程

基于 AE 故障识别结果, 构建迭代更新的多步 RUL 预测模型, 流程如图5所示。

具体步骤为:

- ① 基于 AE 的滚动轴承故障识别结果, 以样本 $x_{n-k} \sim x_n$ 作为 RF 模型的初始训练集;
- ② 对 RF 模型进行训练并得到 RF 预测模型;
- ③ 使用该模型预测得到下一时刻的 RUL 预测值 y_1 ;
- ④ 将 y_1 与相应的特征向量组合, 构建新样本 x_{n+1} 加入 RF 训练样本集, 重复步骤②, 获得更新后的 RF 预测模型;
- ⑤ 使用更新后的 RF 预测模型, 重复步骤③迭代得到轴承 RUL 预测值 y_2 ;
- ⑥ 重复步骤④、⑤, 总计预测 l 次以后可获得一系列迭代预测值 $y_1, y_2, \dots, y_l$ 。

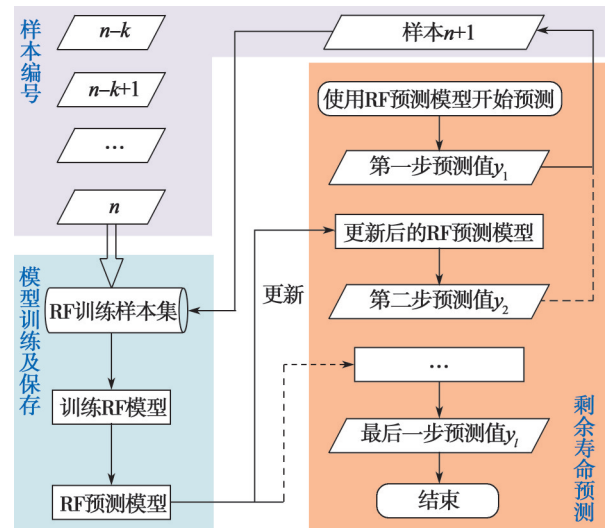


图5 改进RF的滚动轴承寿命预测

Fig. 5 Rolling bearing life prediction with improved RF

在模型改进之后, 可以通过持续扩充训练数据集, 逐步修正模型在初始阶段可能存在的偏差, 从而提高寿命预测模型精度。结合迭代更新策略, 模型能够在不同运行阶段动态适应数据变化, 实现对滚动轴承较长时间范围内 RUL 的稳定预测, 有效提升预测结果的连续性与可靠性。

4 试验验证

为验证本文提出轨道交通电机轴承寿命预测方法

的准确性,采用了某型牵引电机用滚动轴承的全寿命试验数据作为验证对象,对故障识别及寿命预测的结果进行分析。

4.1 数据样本

本次试验在某型轴承寿命试验机上开展,试验对象为某型牵引电机用 6311 轴承。试验台结构如图 6 所示,其中位置 1 为试验轴承,位置 4 为陪试轴承。

试验条件:轴向载荷 $F_a=2.5$ kN,径向载荷 $F_r=24$ kN,转速 $n=2\ 500$ r/min。采用 NI-9068 采集仪与 PCB 振动传感器采集数据,采样频率为 25.6 kHz。

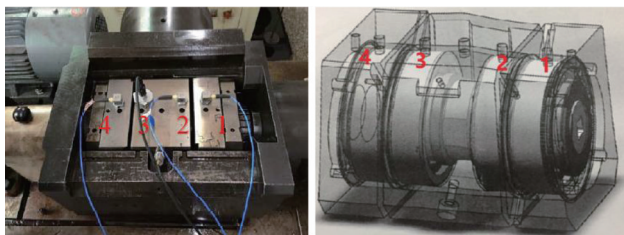


图6 轴承寿命试验台

Fig. 6 Bearing life test rig

4.2 模型参数及故障识别

本文设计了具有 4 层编码器及 4 层解码器的 AE 结构,其中编码器的核数分别为 4、8、16、32,解码器的核数分别为 32、16、8、4,激活函数为 ReLU。AE 模型在训练时,最大训练轮数为 100,学习率为 0.01。

从第 100 个样本起按图 3、图 4 所示的流程对 AE 模型进行训练并识别轴承故障,获取其训练及计算的 MSE,直至误差超出式(4)定义的范围。得到的 MSE 误差及误差允许范围如图 7 所示。

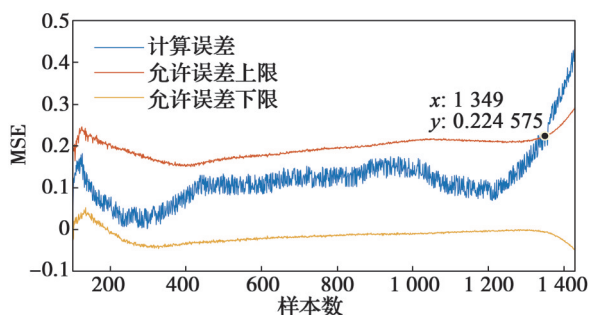


图7 样本的计算MSE及允许范围

Fig. 7 Calculated MSE and allowable range of samples

从图 7 可以看出,在第 1 349 个样本处,AE 模型评估 MSE 超出了误差允许范围,据此判定轴承进入故障状态。

4.3 寿命预测结果

选取第 1 299~1 349 号样本的计算误差 MSE 作为 RF 模型的训练数据集输入特征,记作 $x_j, j=1\ 299, 1\ 300, \dots, 1\ 349$ 。对应的 RF 模型的训练数据集输出以

$y_j, j=1\ 299, 1\ 300, \dots, 1\ 349$ 表示,见式(5)。

$$y_j = 1 - 0.05 \frac{x_j - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

式中: $x_{\min}$ 、 $x_{\max}$ 分别为 x_j 的最小值与最大值。

按照图 5 所示的流程启动轴承寿命预测,预测步长设为 50,轴承 RUL 预测结果如图 8 所示。从图 8 可见,预测曲线与真实 RUL 曲线整体趋势一致。计算其预测值与实际值的 MSE 误差为 0.062,预测精度满足工程要求。

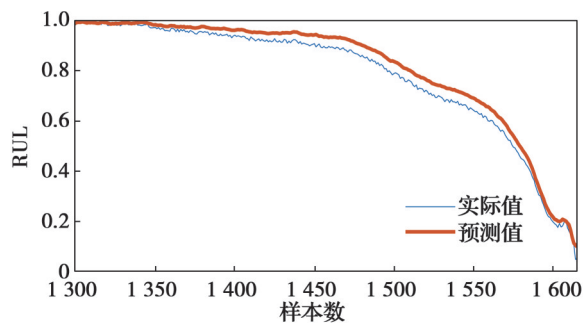


图8 轴承剩余使用寿命预测

Fig. 8 Prediction of RUL of bearings

5 结束语

针对轨道交通牵引电机滚动轴承故障样本稀缺、全寿命周期数据难以获取的工程难题,研究了无需完整故障数据即可实现寿命预测的技术路径,提出了一种基于自编码器(AE)与改进随机森林(RF)回归的剩余使用寿命(RUL)预测方法。该方法从被预测轴承的运行数据出发,使用自编码器模型对运行数据进行无监督特征学习及构造误差的方法进行轴承状态监测,这不仅解决了轨道牵引电机轴承故障样本获取困难的问题,而且为轴承寿命预测提供了准确的起始时间点。在此基础上,使用随机森林回归方法构建了轴承的多步寿命预测模型,通过对预测值的不断迭代得到预测序列,实现了未来较长一段时间内的轴承寿命预测,使预测结果更具有现实意义。采用一组轴承全寿命台架试验数据对该方法进行验证,预测值与实际值的均方误差为 0.062,低于工程经验阈值,表明该方法能够满足实际应用要求。

参考文献:

- [1] 裴洪,胡昌华,司小胜,等.基于机器学习的设备剩余寿命预测方法综述[J].机械工程学报,2019,55(8):1-13.
PEI Hong, HU Changhua, SI Xiaosheng, et al. Review of machine learning based remaining useful life prediction methods for equipment[J]. Journal of mechanical engineering, 2019, 55(8): 1-13.
- [2] 何正嘉,曹宏瑞,晁艳阳,等.机械设备运行可靠性评估的发展与思考[J].机械工程学报,2014,50(2):171-186.

- HE Zhengjia, CAO Hongrui, ZI Yanyang, et al. Developments and thoughts on operational reliability assessment of mechanical equipment[J]. Journal of mechanical engineering, 2014, 50(2): 171-186.
- [3] 许艳雷, 邱明, 李军星, 等. 基于SKF-KF-Bayes的滚动轴承剩余使用寿命预测方法[J]. 振动与冲击, 2021, 40(19): 26-31.
- XU Yanlei, QIU Ming, LI Junxing, et al. Remaining useful life prediction method of rolling bearing based on SKF-KF-Bayes[J]. Journal of vibration and shock, 2021, 40(19): 26-31.
- [4] 张龙龙. 基于多健康状态评估的滚动轴承剩余寿命预测方法研究[D]. 成都: 电子科技大学, 2014.
- ZHANG Longlong. Remaining useful life prediction of rolling elements bearings based on multiple health state assessment[D]. Chengdu: University of Electronic Science and Technology of China, 2014.
- [5] 范乐贤, 刘淑杰, 张洪潮. 综合多阶段退化和多不确定轴承剩余寿命预测[J]. 现代机械, 2022(6): 33-39.
- FAN Lexian, LIU Shujie, ZHANG Hongchao. Prediction of remaining useful life of bearing based on multistage degradation and multi-variability[J]. Modern machinery, 2022(6): 33-39.
- [6] 徐仁义, 王航, 彭敏俊, 等. 剩余寿命预测新方法及其在滚动轴承中的应用[J]. 振动、测试与诊断, 2022, 42(4): 636-643.
- XU Renyi, WANG Hang, PENG Minjun, et al. A new method for predicting remaining life and its application in rolling bearings[J]. Journal of vibration, measurement & diagnosis, 2022, 42(4): 636-643.
- [7] 谢雨洁, 肖友刚, 王田天, 等. 基于异常检测的轴承退化阶段识别方法[J]. 中南大学学报(自然科学版), 2022, 53(5): 1740-1749.
- XIE Yujie, XIAO Yougang, WANG Tiantian, et al. Degradation stage recognition of bearings based on outlier detection [J]. Journal of central south university(science and technology), 2022, 53(5): 1740-1749.
- [8] 吴昊年, 陈仁祥, 胡小林, 等. 改进均衡分布适配的滚动轴承寿命阶段识别[J]. 振动工程学报, 2021, 34(1): 194-201.
- WU Haonian, CHEN Renxiang, HU Xiaolin, et al. Life stage identification of rolling bearing based on improved BDA method[J]. Journal of vibration engineering, 2021, 34(1): 194-201.
- [9] 胡志辉, 王绪光, 王贡献, 等. 基于状态划分和集成学习的轴承剩余使用寿命预测模型[J]. 机电工程, 2024, 41(8): 1423-1430.
- HU Zhihui, WANG Xuguang, WANG Gongxian, et al. Remaining useful life prediction model of bearing based on state division and ensemble learning[J]. Mechanical & electrical engineering magazine, 2024, 41(8): 1423-1430.
- [10] 王仲, 姜娇, 宋洋, 等. 基于改进CNN-BGRU的滚动轴承剩余使用寿命预测模型[J]. 机电工程, 2025, 42(11): 2158-2167.
- WANG Zhong, JIANG Jiao, SONG Yang, et al. Prediction approach of remaining useful life for rolling bearing based on improved CNN-BGRU[J]. Mechanical & electrical engineering magazine, 2025, 42(11): 2158-2167.
- [11] 卢瑾, 张永平. 基于注意力机制的滚动轴承剩余使用寿命预测方法[J]. 机电工程, 2023, 40(4): 516-521.
- LU Jin, ZHANG Yongping. Remaining useful life prediction of rolling bearings based on attention mechanism[J]. Mechanical & electrical engineering magazine, 2023, 40(4): 516-521.
- [12] 李敬一, 苏翔. GWO优化CNN-BiLSTM-Attention的轴承剩余寿命预测方法[J]. 振动与冲击, 2025, 44(2): 321-332.
- LI Jingyi, SU Xiang. Residual life prediction method of bearings based on CNN-BiLSTM-Attention optimized by the grey wolf optimizer[J]. Journal of vibration and shock, 2025, 44(2): 321-332.
- [13] 张宗峰, 陈喜红, 张关震, 等. 关于铁路机车车辆滚动轴承剩余寿命的研究[J]. 机车电传动, 2023(6): 31-38.
- ZHANG Zongfeng, CHEN Xihong, ZHANG Guanzhen, et al. Research on remaining life of rolling bearings for rolling stock[J]. Electric drive for locomotives, 2023(6): 31-38.
- [14] 谭亚红. 自适应小波分析和多层卷积极限学习码器的轴承故障识别研究[J]. 机车电传动, 2021(6): 106-113.
- TAN Yahong. Fault identification of rolling bearing based on adaptive wavelet analysis and multiple layers convolution extreme learning auto-encoder[J]. Electric drive for locomotives, 2021(6): 106-113.
- [15] 徐峥, 梁伟阔, 谈芳吟, 等. 基于DL-Gamma的滚动轴承剩余寿命预测方法[J]. 兵器装备工程学报, 2023, 44(9): 155-162.
- XU Zheng, LIANG Weige, TAN Fangyin, et al. Residual life prediction method for rolling bearing based on DL-Gamma[J]. Journal of ordnance equipment engineering, 2023, 44(9): 155-162.
- [16] 耿明, 张海沧, 康丽齐, 等. 基于主成分分析和随机森林回归模型的工艺装备轴承剩余寿命预测[J]. 城市轨道交通研究, 2023, 26(4): 12-16.
- GENG Ming, ZHANG Haicang, KANG Liqi, et al. Prediction of residual life of jig rolling bearing based on PCA and random forest model[J]. Urban mass transit, 2023, 26(4): 12-16.
- [17] 张霞, 梁海波, 高原, 等. 基于PCA-RF协同的GRU网络滚动轴承退化趋势预测[J]. 现代制造工程, 2025(3): 132-140.
- ZHANG Xia, LIANG Haibo, GAO Yuan, et al. Rolling bearing degradation trend prediction based on PCA-RF-coordinated GRU network[J]. Modern manufacturing engineering, 2025(3): 132-140.

Remaining life prediction of rail transit motor bearings based on auto-encoder and random forest

LI Wei^{1,2}

(1. CRRC Yongji Electric Co., Ltd., Yongji, Shanxi 044502, China;

2. Shanxi Key Laboratory of Traction Motor for Rail Transit, Yongji, Shanxi 044502, China)

Abstract: Life prediction of rolling bearings in rail transit traction motors faces challenges such as scarcity of fault samples and difficulty in lifecycle data acquisition. To address these issues, this paper proposes a method for predicting remaining useful life (RUL) of bearings by integrating an auto-encoder (AE) with random forest (RF). Firstly, an AE was employed to perform unsupervised feature learning on fault-free bearings, and the sample expansion was realized through reconstruction error modeling to mitigate the impact of insufficient fault samples. Secondly, an RF regression model was built to predict the RUL of bearings, and the mean squared error (MSE) loss function was used to evaluate the prediction performance. Finally, a multi-step regression prediction model based on RF was constructed to predict the RUL of bearings through iteration. The accuracy of the proposed method was verified using lifecycle test data of rolling bearings for a specific traction motor. The results show that the proposed method can accurately predict the RUL of rolling bearings, offering a novel approach for life prediction of bearings in rail transit traction motors.

Keywords: rolling bearing; remaining useful life; auto-encoder; random forest; multi-step prediction