

基于深度置信网络的轴承剩余寿命预测探析

魏德宝 汤青峰 杨芸菲 黄晗淇 石佳
(西南交通大学 四川 成都 610097)

摘要:滚动轴承作为大型机械设备的重要零部件,其健康状态对设备的正常运行有着决定性的影响,故对轴承进行寿命预测的意义重大。传统的轴承的剩余寿命预测以基于物理模型和统计模型的方法为主,但其预测误差较大,效率较低,特征反映较弱。本文针对以上问题,提出一种基于深度置信网络的轴承剩余寿命预测模型,此模型将深度置信网络和轴承健康因子相结合,有效避免了梯度问题,且无需人工设置大量参数,处理效率得到了很大的提高。

关键词:滚动轴承;寿命预测;深度置信网络;轴承健康因子

0 引言

本文建立的模型以西安交通大学发布的 XJTU-SY 轴承数据集为例,首先提取原始轴承震动信号的频域特征值,极大降低了轴承原始振动信号的复杂程度。再将其输入已经训练完成的深度置信网络内,输出一系列轴承健康因子值 HI,最终利用支持向量回归进行回归拟和,实现轴承寿命预测,且准确率稳定在 99.4% 左右,本文所用模型的特征提取能力与寿命预测效果良好,所预测结果具有较好的可参考性。

1 研究背景

1.1 社会背景

近期,我国提出“中国制造 2025”战略,国务院也发布了“新一代人工智能发展规划”,二者同时将保障大型设备仪器的安全性和可靠性视为重要研究方向。这就要求我国制造企业要在对大型设备的性能与健康状态的监测方面寻求创新,致力于提高大型设备的安全性与可靠性。而对设备零部件的剩余寿命预测是设备质量管理中的重要环节,其可以极大提高大型设备的安全性与可靠性。

1.2 现实需要

轴承作为大型旋转机械设备最重要的零部件之一,其健康状态直接影响到设备整体的正常运行,被称为“工业的关节”。而且很多精密设备的轴承使用寿命并不长,一旦运行时间超过使用寿命,轴承的运行精度就会急剧下降,从而导致机械设备发生故障。剩余寿命(Remaining Useful Life, RUL)预测作为复杂装备系统可靠性分析中的两类重要问题之一,可以有效避免生产事故,有效减少生产损失和成本。

2 基于深度置信网络的轴承剩余寿命预测模型的建立过程

2.1 轴承原始振动信号分析

Bearing1_1 轴承原始振动信号如图 1 所示,从图中大致可分出轴承正常运

行(0 ~ 80min),开始退化(80 ~ 90min),退化加剧(90 ~ 120min),完全失效(120min)几个状态。

- (1) 正常运行阶段:振幅较小,轴承的运行较为平稳;
- (2) 故障发生阶段:轴承退化现象出现,振幅在增大和平稳间不停切换;
- (3) 故障发展阶段:轴承退化现象加剧,振幅分布趋向杂乱,其中的“突变值”增多;
- (4) 完全失效阶段:轴承完全失效,振幅突然剧烈增加,同时轴承会产生巨大的噪声。

2.2 模型简介

本文以西安交通大学发布的 XJTU-SY 轴承数据集为例,首先获取轴承的水平和垂直振动信号后利用快速傅里叶变换提取出 5 个子频谱带的特征值,将其进行归一化处理输入已经训练完成的深度置信网络内,输出一系列轴承健康因子值,根据输出值利用支持向量回归进行回归拟和,从而最终得到轴承剩余寿命预测模型,进行寿命值计算。

2.3 特征提取与归一化处理

由于轴承的原始振动信号过于复杂,直接将其输入人工神经网络会使训练过程过于漫长,而对其采用特征提取

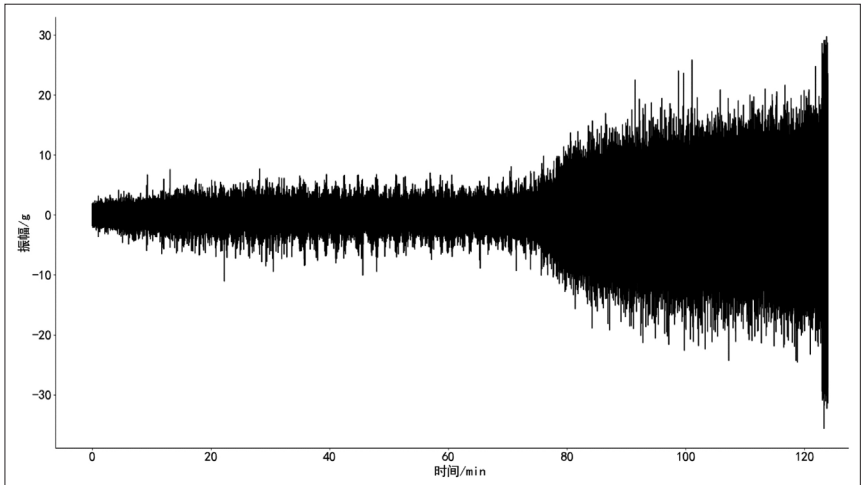


图 1 Bearing1_1 轴承原始振动信号图

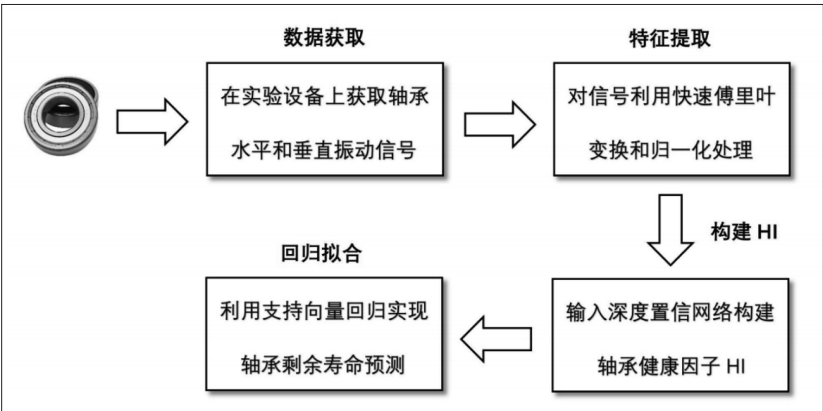


图 2 基于深度置信网络的轴承剩余寿命预测模型流程图

可以压缩原始数据，极大地减轻了数据驱动模型的计算压力。特征的好坏对机器学习算法的泛化能力至关重要。传统信号特征的具有诸多缺点，会导致深度置信网络提取轴承的特征维数较高，直接分析维数较高的特征不利于得到轴承退化趋势。

西安交通大学的雷亚国教授经研究发现轴承的振动信号频域中的带谱能量值具有很好的稳定性，提取方法通用性较高，其过程也无需调节任何参数。从轴承振动的角度分析，轴承的健康状态与退化、损坏阶段的信号特征频率有很大差异，选择带谱能量值作为特征能够准确地显示二者的区别。而且由于轴承损坏时实际情况复杂多变，往往会导致频率特征的偏移，而子频带能量值可以有效避免这一影响。

设 i 时刻的轴承振动频谱为 $Y_t(f)$ ，频谱的最高频率为 F_m ，则可将其分为若干个子频带： $[i \times F_0, (i+1) \times F_0], (i=0, 1, \dots, N-1)$ ，则每一个子频谱能量值可由公式 (4-1) 计算：

$$E_{t,i} = \frac{1}{F_0} \int_{i \times F_0}^{(i+1) \times F_0} Y_t(f) df \quad (1)$$

利用快速傅里叶变换 (Fast Fourier Transform, FFT) 处理初始信号，提取 5 个子频带的能量值，这样，每一小段信号中可提取 10 个能量值。

轴承从正常到失效的全寿命周期内的振动信号，都可以作为原始数据，通过 DBN 对其进行特征提取，把底层原始数据逐层抽象到高层数据特征，并对数据进行归一化与训练集和测试集的划分。

归一化处理可有效减轻人工神经网络的训练压力，对能量向量进行归一化公式为：

$$E_{t,i}' = \frac{E_{t,i} - E_{t,i_{min}}}{E_{t,i} - E_{t,i_{max}}} \quad (2)$$

将归一化的训练集输入到训练好的目标 DBN 网络，获取深度置信网络输出特征。

2.4 深度置信神经网络结构

深度置信神经网络是一个基于概率的生成模型，与其他传统的判别模型的神经网络不同，生成模型是同时建立

在数据和标签上的联合分布。它由多个 RBM 层组成，其结构图如图 2 所示。

其中，上一个 RBM 的隐层即为下一个 RBM 的显层，上一个 RBM 的输出即为下一个 RBM 的输入，依次类推，在最后一层使用 BP 网络进行输出。

该模型训练过程常采用无监督逐层训练 (unsupervised layer-wise training) 方法，如图 4 所示。

(1) 预训练阶段：分别单独无监督地训练每一层 RBM 网络，确保特征向量映射到不同特征空间时都尽可能多地保留特征信息，此时 h^1 可以看作 h^2 的可见层。

(2) 数据微调阶段：每一层 RBM 网络只能确保自身

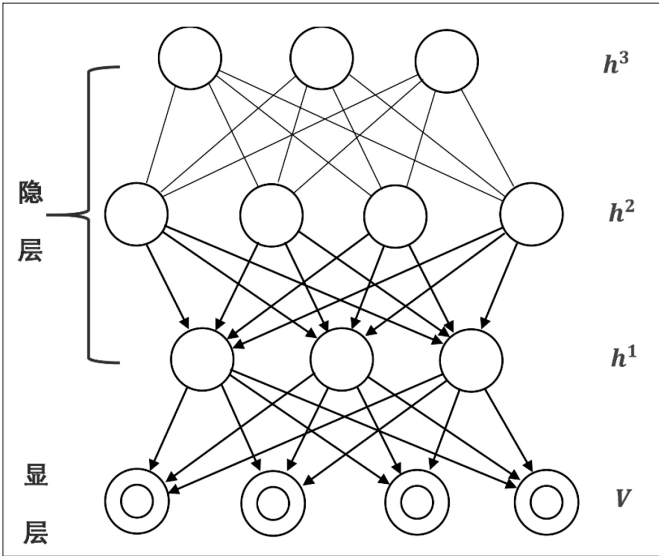


图 3 深度置信神经网络结构图

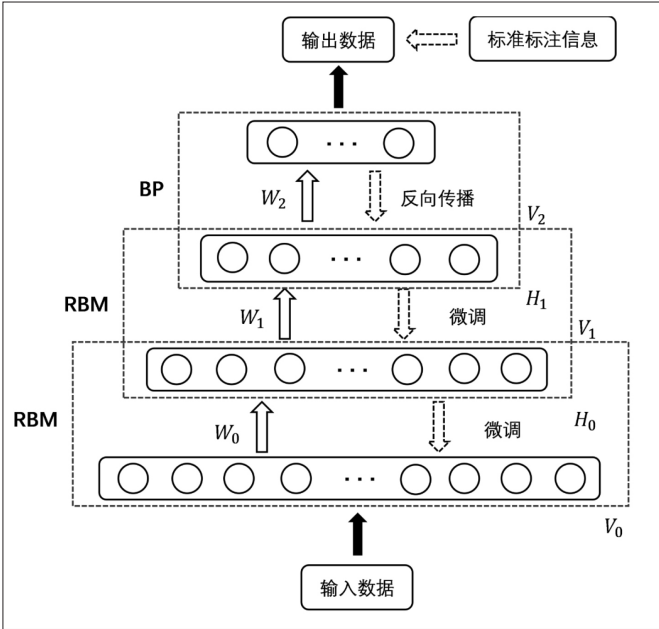


图 4 深度置信神经网络训练过程图

层内的权值对该层特征向量映射达到最优，而此阶段的要求是整体网络达到最优，所以反向传播网络还会将错误信息自顶向下传播至每一层 RBM，微调整个 DBN 网络。RBM 网络训练模型的过程可以看作对一个深层 BP 网络权值参数的初始化，使 DBN 克服了 BP 网络因随机初始化权值参数而容易陷入局部最优和训练时间长的缺点。

2.5 轴承健康因子的构建

轴承健康因子 (Health Indicator, HI) 的构建方便了后续的支持向量回归拟合工作。由于轴承的全寿命较长，故差值大的数据输入网络时很容易产生梯度问题，将轴承寿命放缩到 [0,1] 区间内，作为轴承健康因子代表轴承的寿命。

设训练集中轴承的初始寿命为 RUL_0 ， t 时刻时轴承的剩余寿命为 RUL_t ，该时刻 HI_t ：

$$HI_t = \frac{RUL_t}{RUL_0} \tag{3}$$

利用构建 HI 对轴承进行 RUL 预测，属于间接数据驱动预测方法，也是基于退化状态识别的数据驱动方法，是将退化状态识别中的不同性能变量映射出相应的 HI。假设 $x_1, x_2, \dots, x_n$ 为直接测量参数， f_h 和 f_D 表示 HI 映射函数和退化状态函数。则其构建原理可表示为：

$$HI_t = f_h(x_1, x_2, \dots, x_n) \tag{4}$$

$$RUL_t \propto f_D(HI_t) \tag{5}$$

2.6 支持向量回归

支持向量回归 (Support Vector Regression, SVR) 是一种具有较强泛化能力的机器学习回归算法，在数据回归问题上有着较好的表现。SVR 构建模型遵从结构风险最小化原则，在经验风险最小化的同时降低模型的复杂度，进而提高模型的泛化能力。

支持向量回归解决问题的基本思路是将样本数据从一个非线性映射 从输入空间映射至高维特征空间，再在高维特征空间中对样本数据进行线性回归，寻找最优回归超平面，即使得到超平面最远的样本点的“距离”最小，该拟合最优回归函数为：

$$f(x) = w\phi(x) + b \tag{6}$$

对于样本 (x, y) ，传统回归模型通常直接基于模型输出 $f(x)$ 与真实输出 y 之间的差别来计算损失，意味着当且仅当 $f(x)=y$ 时损失才为零。与此不同，支持向量回归能容忍 $f(x)$ 与 y 之间的差别绝对值大于 ϵ 时才计算损失。

于是，SVR 问题可以简化成：

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m l_{\epsilon}(f(x_i) - y_i) \tag{7}$$

通过引入拉格朗日乘子 $\mu_i \geq 0, \hat{\mu}_i \geq 0, \alpha_i \geq 0, \hat{\alpha}_i \geq 0$ 可得到 SVR 的解形如：

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) x_i^T x + b \tag{8}$$

根据支持向量回归可将神经网络输出的轴承健康因子 HI 拟合，最终得到轴承寿命预测曲线。

3 实验验证与效果分析

3.1 原始数据集划分

为了训练深度置信网络，提高特征提取的效果，本文使用留出法 (hold-out) 将原始数据集划分为两个互斥的集合，为训练集 S 和测试集 T，在 S 上训练出模型后用 T 来评估其测试误差，得出模型的预测效果。本文将 3 种工况下的 15 个轴承划分为包含 9 个轴承的训练集和包含 6 个轴承的测试集，如表 1 所示。

3.2 滚动轴承特征提取

对滚动轴承的原始振动信号进行分解，得到若干分量后从中提取所需特征值构建模型。本文提取的特征为频域特征，以轴承 Bearing3_4 为例，绘制其竖直方向和水平方向振动信号随时间的变化趋势图，如图 5 所示。

表 1 数据集划分

工况	数据集	样本总数	划分类别
1	Bearing1_1	123	训练集
	Bearing1_2	161	训练集
	Bearing1_3	158	训练集
	Bearing1_4	122	测试集
	Bearing1_5	52	测试集
2	Bearing2_1	491	训练集
	Bearing2_2	161	训练集
	Bearing2_3	533	训练集
	Bearing2_4	42	测试集
	Bearing2_5	339	测试集
3	Bearing3_1	2538	训练集
	Bearing3_2	2496	训练集
	Bearing3_3	371	训练集
	Bearing3_4	1515	测试集
	Bearing3_5	114	训练集

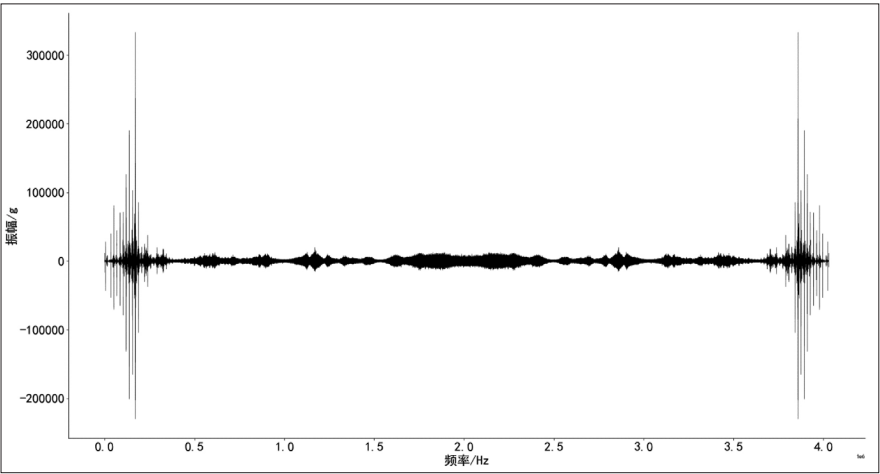


图 5 轴承 Bearing3_4 频域特征

利用快速傅里叶变换处理初始信号，提取 5 个子频带的能量值，这样，每一小段信号中可提取 10 个能量值。

为了减轻人工神经网络的训练压力，对能量向量进行归一化处理：

$$E_{t,i}' = \frac{E_{t,i} - E_{t,i\min}}{E_{t,i} - E_{t,i\max}} \quad (9)$$

3.3 网络初始化

训练网络之前要提前进行网络初始化，大致内容包括设定网络结构和其他初始化参数：

(1) 网络结构：DBN 源于最初的浅层神经网络，则关于隐层节点数的选取可参考 BP 隐含层节点数确定的经验公式：

$$n = \sqrt{m + p} + a \quad (10)$$

其中，m 为输入特征参数，p 为输出特征参数，n 代表隐层节点数，a 为区间 [1,10] 内的整数。

(2) 初始化参数：模型中需要初始化的参数主要有连接权重 w，显层偏置 a，隐层偏置 b，可以采用极小值随机初始化，本文使用经验公式初始化：

$$w = 0.1 \times \text{randn}(n,m) \quad (11)$$

$$a = 0.1 \times \text{randn}(1,n) \quad (12)$$

$$b = 0.1 \times \text{randn}(1,m) \quad (13)$$

表 2 DBN 模型参数确定（训练集）

参数	设置值	参数	设置值
数据总数	7032	向前堆叠学习率	0.1
隐层结构	[860 532 380]	向前堆叠动量项	0.9
迭代次数	40	后项微调学习率	2
批量大小	30	后项微调动量项	0.5

表 3 DBN 模型参数确定（测试集）

参数	设置值	参数	设置值
数据总数	2184	向前堆叠学习率	0.5
隐层结构	[900 600 900]	向前堆叠动量项	0.5
迭代次数	10	后项微调学习率	2
批量大小	50	后项微调动量项	0.5

其中，randn 表示可产生标准正态分布的随机数或矩阵的函数，randn(m,n) 代表产生 n 行 m 列的标准正态分布矩阵。

(3) 激活函数：

本文采用线性修正单元 (Rectified Linear Unit,ReLU), 作为 DBN 的神经元激活函数, 如式 (14) 所示。

$$ReLU = \max(0, x) = \begin{cases} 0, & x \leq 0 \\ x, & x \geq 0 \end{cases} \quad (14)$$

分别对训练集和测试集实验数据进行训练并预测，以 Bearing3_4 为例，得出训练集与测试集在实验中确定的具

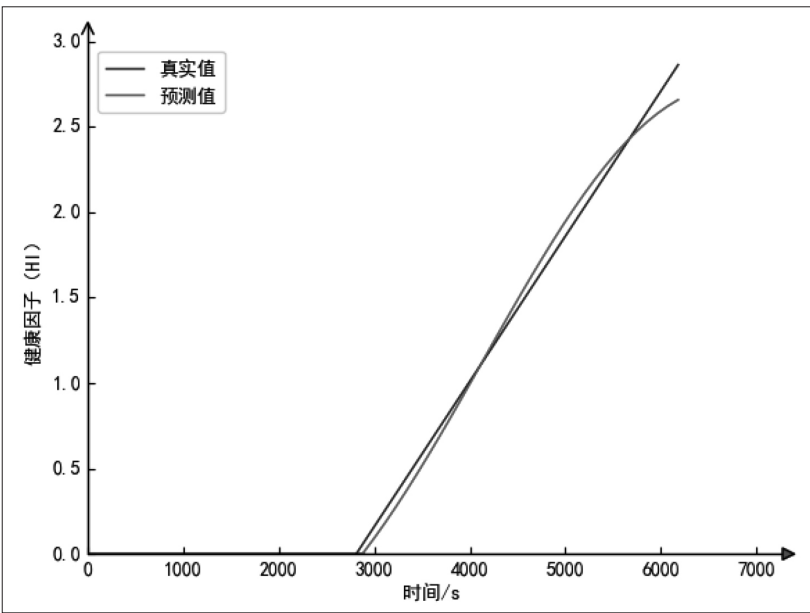


图 6 滚动轴承 Bearing3_4RUL 预测结果（测试集频域数据）

体参数如表 2 所示。

3.4 预测结果及效果分析

对深度置信网络进行初始化后，将提取出的训练集数据频域特征灌入网络中，对其进行逐层贪婪学习完成训练。训练过后再将提取出的测试集数据频域特征输入深度置信网络中，输出轴承健康因子 HI。经过以上步骤后利用支持向量回归对 HI 进行拟合回归，以滚动轴承 Bearing3_4 为例，得到如下基于频域数据的预测曲线，并在图中做出真实寿命值。

从图 6 可以看出，本文采用的基于深度置信网络的轴承剩余寿命预测方法所得出的预测值与真实值很接近，且在全周期内的预测误差浮动不大，相对比较稳定，说明该预测模型有良好的预测能力，能得到较为准确的预测结果。

为了直观地表示出轴承剩余寿命预测的效果，本文给出滚动轴承 Bearing3_4 的开始退化以及退化加剧阶段各 20 组数据的预测值，并与真实值进行对比，如下表所示。并计算出它们的准确率，如下式所示，其中 y_i 为真实值， $\hat{y}_i$ 为预测值。

$$A = 1 - \frac{y_i - \hat{y}_i}{y_i} \quad (15)$$

从表 5 中可以看出，在滚动轴承开始退化阶段进行的预测，真实值和预测值之间的相对误差呈减小趋势，准确率在不断提高。

从表 5 中可以看出，在滚动轴承退化加剧阶段进行的预测，真实值和预测值之间的相对误差仍然较小，准确率趋于稳定，预测精度较高，说明此阶段数据仍然存在参考价值。

3.5 对比实验分析

为进一步说明模型的预测效果，采用滚动轴承 Bearing3_4 原始数据中退化加剧阶段的 100 组数据，分别

表 4 开始退化阶段轴承 HI 真实值与预测值对比表（测试集）

组别	真实值	预测值	准确率	组别	真实值	预测值	准确率
1128	0.5001	0.503600	92.8%	1138	0.5431	0.546462	93.8%
1129	0.5043	0.507887	92.9%	1139	0.5474	0.550748	93.9%
1130	0.5086	0.512173	93.3%	1140	0.5517	0.559320	94.0%
1131	0.5129	0.516459	93.1%	1141	0.5560	0.563606	94.1%
1132	0.5172	0.520745	93.2%	1142	0.5603	0.567893	94.1%

表 5 退化加剧阶段轴承 HI 真实值与预测值对比表（测试集）

组别	真实值	预测值	准确率	组别	真实值	预测值	准确率
1401	1.6843	1.673717	93.8%	1411	1.7314	1.716579	93.9%
1402	1.6886	1.678004	93.8%	1412	1.7357	1.720865	93.9%
1403	1.6928	1.682290	93.8%	1413	1.7400	1.725151	93.9%
1404	1.6971	1.686576	93.8%	1414	1.7443	1.729437	94.4%
1405	1.7014	1.690862	93.8%	1415	1.7486	1.733723	94.0%

表 6 实验预测值误差指标对比表

预测模型	预测误差 MAPE
DBN-SVR	0.0367
PSO-BP	0.0687
PSO-LS-SVR	0.0976

用两种应用较广泛的 RUL 预测模型：基于 PSO-BP 的滚动轴承 RUL 预测模型以及基于 PSO-LS-SVR 滚动轴承 RUL 预测模型进行剩余寿命预测，并将二者与本文使用的模型进行效果对比分析。利用以上两种模型与本文所用模型对数据进行预测，得到预测结果误差对比如表 6 所示。

从上表中可以看出，对于滚动轴承剩余寿命数据，本文所用模型的预测值相对于 PSO-BP 模型预测值和 PSO-LS-SVR 模型预测值都更接近于真实值，预测误差最小且预测结果比较稳定，因此本文所采用的轴承剩余寿命预测模型性能更优，具有一定的工程实际意义。

4 结语

本文建立的滚动轴承剩余寿命预测模型将深度置信网络与 BP 网络相结合，建立轴承健康因子值 HI，最终利用支持向量回归进行回归拟和，实现寿命预测功能，且预测准确率稳定在 99.4% 左右，与其他两种使用广泛的预测网

络对比后可知，本文提出的轴承剩余寿命预测模型提取特征以及预测性能更优，具有一定的工程实际意义。

参考文献：

[1] 国务院，国务院关于印发新一代人工智能发展规划的通知[EB/OL].(2017-07-20).http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.

[2] 高耀智. 基于 1 (1/2) 维谱与小波包相结合的滚动轴承的故障诊断[D]. 湘潭大学,2009.

[3] 雷亚国,韩天宇,王彪,李乃鹏,闫涛,杨军.XJTU-SY 滚动轴承加速寿命试验数据集解读[J]. 机械工程学报,2019,55(16):1-6.

[4] 张娜. 基于深度信念网络的滚动轴承寿命预测方法研究[D]. 湖南大学,2018.

[5] 刘霞,秦华锋.基于深度置信网络的假手指静脉图像检测算法[J]. 重庆工商大学学报(自然科学版),2018,35(05):33-39.

[6] 张智儒. 面向情感分析的人脸表情识别系统的研究与实现[D]. 东北大学,2015.

[7] 刘立群,韩路路,张路娜,陈水明. 基于 BP 神经网络的风电功率预测方法[J]. 自动化应用,2021(06):3-6.

[8] 王林. 面向资源成本的云应用性能优化决策方法及应用[D]. 东北大学,2015.

作者简介：魏德宝（2001-），男，汉族，山东省聊城市，本科生在读，西南交通大学学生，研究方向：工业工程；
汤青峰（2001-），男，汉族，湖北省荆州市，本科在读，西南交通大学学生，研究方向：工业工程；
杨芸菲（2001-），女，汉族，山东省泰安市，本科在读，西南交通大学学生，研究方向：工业工程；
黄晗淇（2001-），女，汉族，福建省龙岩市，本科在读，西南交通大学学生，研究方向：工业工程；
石佳（2000-），男，汉族，陕西省宝鸡市，本科在读，西南交通大学学生，研究方向：工业工程。

