

基于 LSTM 分类器的航空发动机 预测性维护模型

蔺瑞管, 王华伟*, 车畅畅, 倪晓梅, 熊明兰
(南京航空航天大学民航学院, 江苏 南京 210016)

摘 要: 利用传感器数据进行预测性维护是航空发动机故障预测与健康管理(prognostic and health management, PHM)的关键问题。针对发动机剩余寿命预测准确性低的问题,提出基于长短期记忆网络(long short-term memory network, LSTM)分类器的预测性维护模型。LSTM 分类器通过门控单元对长时间序列信息进行充分筛选,并将有效信息用于时间序列预测。首先,采用滑动时间窗口制备训练样本。其次,将预处理后的样本输入 LSTM,预测设备在特定时间窗口内的失效概率。然后,通过调整窗口大小,得到最优性能的二分类模型,以更好地适应预测维护需求。最后,利用美国航空航天局 C-MAPSS 数据集验证了该模型的有效性,相比于已有分类方法,其在剩余使用寿命分类方面更加准确。

关键词: 故障预测与健康管理; 预测性维护; 长短期记忆网络; 时间窗; 二分类

中图分类号: TP 399

文献标志码: A

DOI:10.12305/j.issn.1001-506X.2022.03.39

Predictive maintenance model of aeroengine based on LSTM classifier

LIN Ruiguan, WANG Huawei*, CHE Changchang, NI Xiaomei, XIONG Minglan
(School of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: Predictive maintenance using sensor data is a key issue in aeroengine prognostic and health management (PHM). Aiming at the problem of low accuracy of remaining useful life prediction of aeroengine, a predictive maintenance model based on long short-term memory network (LSTM) classifier is proposed. The LSTM classifier fully screens the long time sequence information through the gating unit, and uses the effective information for time sequence prediction. Firstly, a sliding time window is used to prepare training samples. Secondly, the pre-processed samples are input into the LSTM to predict the failure probability of the equipment in a specific time window. Then, by adjusting the window size, a two-class model with the best performance is obtained to better adapt to predictive maintenance requirements. Finally, the National Aeronautics and Space Administration C-MAPSS data set is used to verify the effectiveness of the model. Compared with the existing classification methods, the proposed model is more accurate in remaining useful life classification.

Keywords: prognostic and health management (PHM); predictive maintenance; long short-term memory network (LSTM); time window; binary classification

0 引 言

作为飞机的关键部件之一,航空发动机的工作条件通常非常复杂,任何意外故障都可能导致灾难性后果^[1]。随

着传感器技术的最新发展,以及通信系统和机器学习技术的显著进步,预测性维护(predictive maintenance, PdM)已经成为航空发动机故障预测与健康管理(prognostic and health management, PHM)领域的研究热点^[2]。通过建立

收稿日期:2021-02-17; 修回日期:2021-06-16; 网络优先出版日期:2021-11-01。

网络优先出版地址: <http://kns.cnki.net/kcms/detail/11.2422.TN.20211101.1524.006.html>

基金项目:国家自然科学基金(U1833110)资助课题

* 通讯作者。

引用格式: 蔺瑞管, 王华伟, 车畅畅, 等. 基于 LSTM 分类器的航空发动机预测性维护模型[J]. 系统工程与电子技术, 2022, 44(3): 1052-1059.

Reference format: LIN R G, WANG H W, CHE C C, et al. Predictive maintenance model of aeroengine based on LSTM classifier[J]. Systems Engineering and Electronics, 2022, 44(3): 1052-1059.

航空发动机的 PdM 模型,管理者可以更有效地计划维护活动,以减少发动机停车时间并降低平均维护成本,保证发动机运行的可靠性和安全性^[3]。

随着工业中实际需求的日益增长,PdM 在近十年中受到了学者的极大关注^[4-6]。通常,航空发动机预测维护框架包括两个相互联系的关键部分:系统剩余使用寿命(remaining useful life, RUL)的预测和维修决策^[7]。在航空发动机的性能退化过程中,传感器数据之间存在紧密的时间相关性^[8]。RUL 预测就是根据该航空发动机的历史传感器时间序列数据,辨识隐含其中的运行规律,进而应用该规律对航空发动机剩余使用寿命进行预测^[9]。

提高 RUL 预测的准确性不仅可以提高安全性和可靠性,降低平均维护成本,并为航空发动机维修决策提供参考。基于预测方法的研究主要分为两大类:基于物理模型的 PdM 框架^[10]和基于数据驱动的 PdM 框架^[11]。但是由于设备结构日益复杂,再加上各种环境的影响,很难用物理模型去准确地预测 RUL^[12]。随着大数据时代的到来,以及计算机技术的不断发展,基于人工智能的数据驱动方法已经成为发动机 RUL 预测领域的研究热点。

在最近的研究中,已经开发了许多机器学习技术(尤其是深度学习),并成功应用于预测各种复杂系统的 RUL。深度学习算法可通过对历史性能退化状态序列的学习,逼近传感器数据的时间相关性规律,从而预测 RUL。在深度学习中,循环神经网络(recurrent neural network, RNN)包含递归隐藏层,非常适用于利用时间序列信息预测 RUL。但是,在 RNN 处理长时间序列数据时,会出现梯度消失现象,导致其在实际应用中性能受限^[13]。为了解决 RNN 的这一问题,文献^[14]提出了长短期记忆(long short-term memory network, LSTM)网络,其作为一种 RNN 的变体,将长期内存保存在单元状态,可有效处理发动机性能退化数据的长期依赖关系,适用于解决 RUL 预测问题,因此备受关注。

Che 等^[15]提出了一种结合多种深度学习算法的 PHM 模型,通过深度置信网络(deep belief network, DBN)和 LSTM 的集成来估计 RUL。主要思想是使用构造的健康因子(health indicator, HI)和目标标记来训练神经网络模型,根据预测的 HI,通过设置阈值以获得发动机 RUL 的估计值。Tamilselvan 等^[16]提出了一种使用 DBN 分类器的新型多传感器健康诊断方法。Guo 等^[17]提出了一种基于递归神经网络的健康指标,利用单调性和相关度量从原始特征集中选择最敏感的特征,用于进行 RUL 预测。Hinchi 等^[18]提出一种基于卷积和长短期记忆的深度 RUL 估计框架。首先,使用卷积层直接提取传感器数据的局部特征,然后利用 LSTM 层获得退化过程并估计 RUL。Aldulaimi 等^[19]提出了一种用于 RUL 估计的混合深度神经网络模型,该方法使用 LSTM 路径提取时间特征,而同时使用卷积神经网络(convolutional neural networks, CNN)提取空间特征,对复杂系统具有较好的预后效果。Yuan 等^[20]利用 LSTM 在复杂操作,混合故障和强噪声情况下预测发动机的 RUL。张妍等^[21]提出由多层感知器(multi-layer perceptron, MLP)和进化算法组

成的框架,并利用跨步时间窗口和分段线性模型来估计机械组件的 RUL。针对航空发动机性能退化和多状态参数时间序列预测的问题,车畅畅等^[22]构建了基于多尺度排列熵算法和 LSTM 的 RUL 预测模型。针对 RUL 预测精度低的问题。张永峰等^[23]提出基于一维 CNN 和双向长短期记忆(bi-directional long-short term memory, BD-LSTM)的集成神经网络模型,并于其他深度学习模型进行了比较。

在以上 RUL 预测的研究中,一般可以概括为以下 3 个步骤:①提取原始数据特征;②建立设备 HI 曲线;③预测设备 RUL^[17]。该方法的关键是对原有 HI 曲线进行向后的多步预测,即当 HI 值超过预先设定的阈值时视为失效,从而可计算得到 RUL。提出的模型性能严格取决于故障阈值定义,这在实践中并不简单,不仅需要大量专家经验参与,而且模型预测的鲁棒性和准确性也会受到阈值取值的影响。此外,这些研究为预测 RUL 值的回归问题,其准确性严格取决于预测范围(从当前时间到实际系统故障时间的时间段)。因此,若使用预测准确性较低的 RUL 值可能会导致错误的决策。

针对以上问题,本文提出一种新的航空发动机预测性维护模型。采用滑动时间窗口方法标记训练样本,充分表征了多元传感器数据的退化信息。运用 LSTM 分类器强大的特征提取能力处理时间序列数据,估计系统将来在特定时间窗口内发生故障的概率。与以往研究相比,所提出的预后方法不需要设定故障阈值,而是将预测 RUL 转化为二分类问题,即预测设备在特定时间窗口内是否会失效,有效提高了维修决策的准确性。通过分析窗口大小对模型性能的影响,得到最优性能的模型参数。由于这些时间窗口是根据运营计划者的要求定义的,因此所提出的方法对进一步的维修决策具有重要的研究意义。

1 模型算法描述

1.1 LSTM 分类器

LSTM 是在 RNN 基础上的改进,通过多个序列的组合和前后连接,RNN 能够根据当前信息和历史信息来进行预测。然而,随着神经网络的复杂度逐渐提高,RNN 往往出现信息过载和局部过优化的问题。作为 RNN 的变体,LSTM 能够利用门控制单元使网络的信息提取更有选择性,从而有效地提高信息的利用率和时间序列预测的准确率^[24]。LSTM 通过引入长时记忆单元、输入门、遗忘门、输出门、短时记忆单元等概念,让整个网络模型的运行时间更短、误差更小。

图 1 说明了在时间 t 时的 LSTM 层结构,LSTM 层通过使用单元状态 c_t 以及隐藏状态 h_t 和各种门来提供更丰富的内部状态。假设 N 个相同类型的发动机在运行期间受到 m 个传感器的监视。然后,以矩阵形式表示每个部件 $i(i=1,2,\dots,N)$ 在其生命周期中获取的传感器数据: $\mathbf{X}_i=[X^1, X^2, \dots, X^N]$, $\mathbf{X}_i \in \mathbf{R}^{m \times N}$, 其中 $\mathbf{X}^i=[x^i_1, x^i_2, \dots, x^i_m]$ 是在时间 t 的传感器测量值, N 表示序列长度。在训练阶段,构建的 LSTM 采用传感器测量序列 \mathbf{X}_i ,并判断 RUL 真实

值是否属于某个确定的时间窗口。接下来,在测试阶段,在时间 t 处,构造的 LSTM 分类器将传感器测量值 \mathbf{X}^t 的向量作为输入数据,并输出 RUL 属于确定时间窗口的概率。

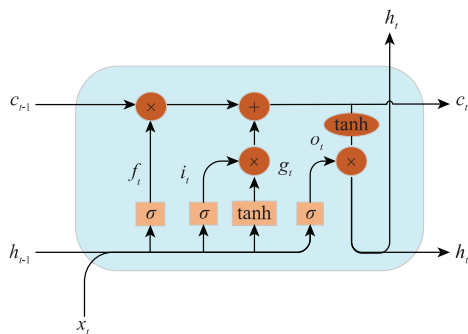


图 1 LSTM 单元结构

Fig. 1 LSTM unit structure

首先,遗忘门 f_t 控制 LSTM 层哪些长期记忆可以被遗忘:

$$f_t = \sigma(\mathbf{W}_f \mathbf{h}_{t-1} + \mathbf{U}_f \mathbf{X}^t + \mathbf{b}_f) \quad (1)$$

接下来,输入门计算可以从输入中获取的信息,并了解其中哪些部分应该存储到单元状态中:

$$\mathbf{g}_t = \tanh(\mathbf{W}_g \mathbf{h}_{t-1} + \mathbf{U}_g \mathbf{X}^t + \mathbf{b}_g) \quad (2)$$

$$i_t = \sigma(\mathbf{W}_i \mathbf{h}_{t-1} + \mathbf{U}_i \mathbf{X}^t + \mathbf{b}_i) \quad (3)$$

然后,更新单元状态中的长期记忆:

$$\mathbf{c}_t = \mathbf{c}_{t-1} \otimes \mathbf{f}_t + \mathbf{g}_t \otimes \mathbf{i}_t \quad (4)$$

最后,使用输出门根据输入,单元状态和先前的隐藏状态更新当前隐藏层的状态:

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \mathbf{h}_{t-1} + \mathbf{U}_o \mathbf{X}^t + \mathbf{b}_o) \quad (5)$$

$$\mathbf{h}_t = \mathbf{o}_t \otimes \tanh(\mathbf{c}_t) \quad (6)$$

在以上公式中, $\mathbf{W}_f, \mathbf{W}_g, \mathbf{W}_i, \mathbf{W}_o$ 是当前隐藏层和先前隐藏层之间的隐藏层权重值,而 $\mathbf{U}_f, \mathbf{U}_g, \mathbf{U}_i, \mathbf{U}_o$ 是当前输入层和当前隐藏层之间的权重值; $\mathbf{b}_f, \mathbf{b}_g, \mathbf{b}_i, \mathbf{b}_o$ 是偏差向量; \otimes 是逐元素乘法运算符; σ 是 Sigmoid 函数; \tanh 是激活函数。

1.2 基于 LSTM 分类器的预测性维护框架

在生产实践中,通常需要长期提供预测信息,以制定不同的维护计划。此外,由于技术和后勤方面的限制,无法在任何时间和任何地方执行航空发动机的维护操作。因此,运营计划者需要先了解设备在确定时间段内的故障概率,进而根据这些预测信息做出相应的预测性维修决策。

为了解决这一问题,本节提出一个基于 LSTM 分类器的预测性维护模型,该模型包含从数据预处理、模型的训练和测试到提供确定时间窗口内故障概率的整个过程,如图 2 所示。

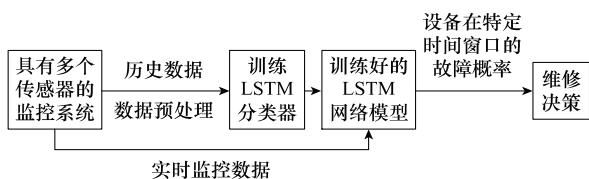


图 2 基于 LSTM 分类器的预测性维护流程

Fig. 2 Predictive maintenance process based on LSTM classifier

首先,针对航空发动机全寿命周期的预测性维护过程,利用具有多个传感器的监控器系统,采集各项发动机运行性能参数,建立数据集。其次,将预处理后的数据集分成训练集和测试集。将训练集和测试集输入到 LSTM 分类器中,进行模型训练,通过调整时间窗口得到最优性能的 LSTM 模型。最后,预测设备在特定时间窗口的故障概率,以指导发动机维修决策。

2 实验数据处理与模型构建

2.1 数据集描述

本文使用美国国家航空航天局的 C-MAPSS 涡扇发动机退化数据集进行模型仿真^[25]。C-MAPSS 数据集包括 4 个子集:FD001,FD002,FD003 和 FD004。其具有不同数量的运行条件和故障模式,每个子数据集进一步分为训练集和测试集。表 1 列出了 C-MAPSS 数据集的构成部分。第 1 行和第 2 行分别代表每个引擎的发动机单元编号和退化周期,第 3 行描述了发动机的运行设置,而最后的 21 列由来自 21 个传感器的多元时间数据组成。在以后的实验中,将考虑用 FD001 数据集对所提出模型进行验证和评估。在该数据集中,包括 100 个从运行到失效的时间序列,其中包含 20 631 个不同时间的测试数据,所有序列中测试数据的最大数量和最小数量分别为 303 和 31。

表 1 C-MAPSS 数据集的构成

Table 1 The composition of C-MAPSS data set

列号	内容
1	发动机单元编号
2	时间循环
3	时间循环 1
4	时间循环 2
5	时间循环 3
6	传感器测量值 1
⋮	⋮
26	传感器测量值 21

2.2 数据预处理

在训练 LSTM 网络之前,有必要对来自多个传感器源的异构数据进行预处理。

(1) 数据归一化

输入数据是从多个传感器源获得的,这些传感器源的范围不同。为了使用这些异构数据来训练 LSTM 分类器,有必要对它们进行归一化。本文采用 Min-Max 方法对训练集与测试集数据进行归一化,这种归一化将确保所有功能在所有工作条件下的平等贡献^[26];归一化的数据将在 $[0,1]$ 之间。

(2) 时间窗处理

为了对模型进行训练和测试,需要对训练集和测试集的数据进行标记。本文采用滑动时间窗口来定义数据标签。对于几个连续时间序列的数据,采用滑动时间窗以获得特征向量,可以获取更多有用的时间信息,这些信息可能会大大改善 RUL 分类性能^[27]。具体而言,就是利用先前的时间步长预测下一个时间步长,窗口沿时间方向滑动一个时间单位,便构造出单个训练样本,最后一个时间节点对应的 RUL 作为该训练样本的标签。

所提出的方法根据时间窗口来定义数据标签,在该时间窗口中,运营计划者需要故障预测信息来安排维护和生产活动。例如,运营计划者需要系统在规定时间内 w_1 中发生故障的概率,则将数据标记为两个类别。第一类记为 Deg0,表示设备 RUL 大于 w_1 的情况,即 $RUL > w_1$ 。第二类为 Deg1,表征其中设备 RUL 小于 w_1 ,即 $RUL \leq w_1$ 。如果 RUL 属于给定的类别,则其对应元素将设置为 1,而输出数

组的其余元素设置为 0。

2.3 网络模型参数设置

本文提出的深层 LSTM 分类模型由 Python 3.7.6 keras/tensorflow 深度学习库构建,处理器为 Intel(R) Core (TM) i7-9700 CPU @ 3.00 GHz 3.00 GHz,内存为 8 GB,操作系统为 Windows 10。图 3 显示了深度 LSTM 中 3 种类型的层:输入层,隐藏层和输出层。

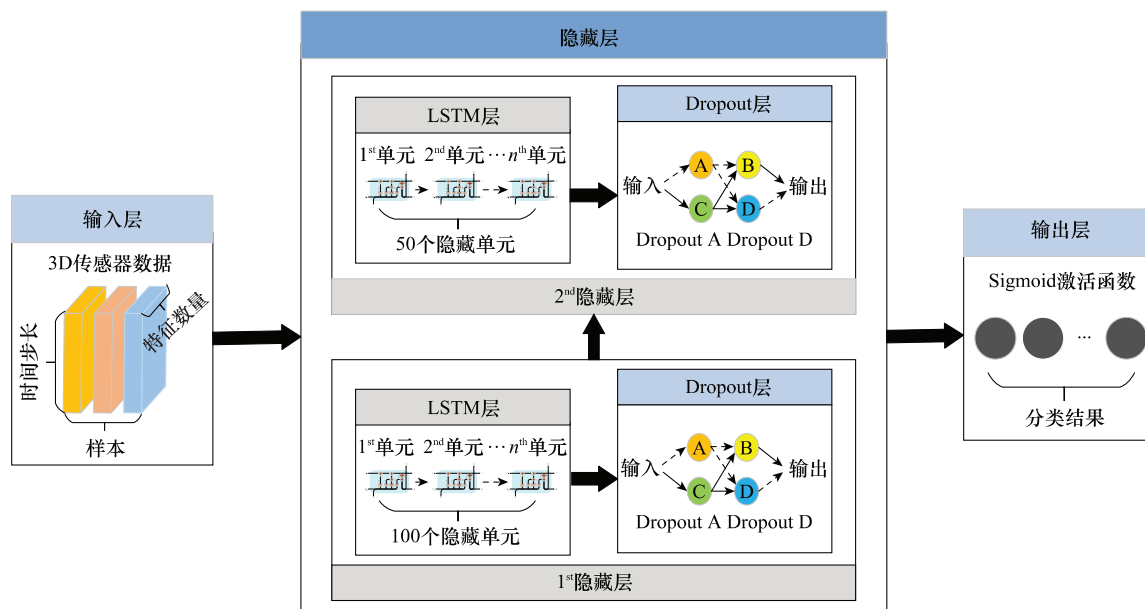


图 3 提出的深层 LSTM 分类模型

Fig. 3 The proposed deep LSTM classification model

输入层是将数据带入 LSTM 的网络层。输入数据为三维格式,即[样本,时间步长,特征数量]。这里的时间步长是指每个特征的信息能够传递给下一个特征的长度^[28]。为了对全部设备的 RUL 预测值进行二进制分类,时间步长需要满足测试集中记录数据的最小长度。由于数据集 FD001 中记录的最小长度为 31,则本文设定的时间步长为 30。

隐藏层介于输入层和输出层之间,是模型训练和测试的核心部分。在构建的隐藏层中,先后设置了 100 个和 50 个单元的层。另外,在每个 LSTM 层之后应用 Dropout,以减少神经网络训练数据的过拟合,从而提高网络的特征提取能力。

输出层是包含一个前馈神经网络的全连接层。该层用作网络和输出之间的原型。其允许将隐藏层输出处的三维张量转换为分类器输出处的一维数组。在本文中,将分类器输出定义为两个元素的向量,这些特征描述了观察结果属于两类的概率:Deg0($RUL > w_1$), Deg1($RUL \leq w_1$)。然后,在输出层中有两个单元,并使用“Sigmoid”激活函数。输出层提供了两个类别(Deg0 和 Deg1)上的概率分布。

为了训练 LSTM 分类器,将目标函数的损失(loss)定义为“binary_crossentropy”,该函数专门用于解决两类别分类问题。接下来,本文采用 Adam 优化算法^[29],其是随机

梯度下降算法的扩展式,具有计算效率高、内存需求小以及对大数据适用性高等优点,被广泛用于深度学习模型。为了评估模型的性能,将度量功能定义为“binary_accuracy”。类似于目标函数,其为所有分类问题提供了所有预测的平均准确率。

2.4 性能评价指标

如表 2 所示,对于二分类问题来说,根据预测的结果得到混淆矩阵,对角线元素显示每个类别的正确观测值。

表 2 二分类模型的混淆矩阵

Table 2 Confusion matrix of the binary classification model

实际	预测	
	1	0
1	真正类(true positive, TP)	假正类(false positive, FP)
0	假负类(false negative, FN)	真负类(true negative, TN)

本文选择准确率(Accuracy)、精确率(Precision)、召回率(Recall)、F 值(F1-score)来衡量二分类中不同模型算法的好坏,同时选择 ROC(receiver operating characteristic)曲线和 ROC 曲线下面积(area under ROC curve, AUC)来完善二分类的评价指标。Accuracy 表示分类正确的样本数占总样本数的比例。但是,对于不平衡数据的表现很差;Precision 表示预测为正类的样本中真正类所占的比例;Recall 是在所

有正类中被预测为正类的比例; F1-score 是 Precision 和 Recall 的调和平均, 一般用来衡量分类器的综合性能。ROC 曲线又称作“受试者工作特性曲线”, 横坐标为假正率 (false positive rate, FPR), 纵坐标为真正率 (true positive rate, TPR), 曲线越靠近左上角的点, 效果越好。AUC 定义为 ROC 曲线下的面积, 取值范围一般为 (0.5, 1.0)。AUC 就越大, 表示模型分类性能越好。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (7)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (8)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (9)$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (11)$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{FN}} \quad (12)$$

表 3 LSTM 分类器的相关参数

Table 3 Related parameters of LSTM classifier

Epochs	Batch_size	sequence_length	Validation_split	Dropout	LSTM_1 units	LSTM_2 units
100	200	30	0.05	0.2	100	50

3.2 RUL 分类结果分析

首先, 时间窗 w_0 和 w_1 的大小划分如表 4 所示, 分析 w_1 对发动机 RUL 分类效果的影响; 其次, 将准备好的 C-MAPSS FD001 数据集分成训练集、验证集和测试集; 然后, 利用滑动时间窗方法对数据集进行标记, 将标

2.5 概率混淆矩阵

在本文中, 为了预测观察值属于每个类别的概率, 将混淆矩阵重新定义为概率混淆矩阵, 以评估所提出算法的性能^[7]。该矩阵的元素 \hat{M}_{ij} 是观测值 x 的预测值为 j 而真实值为 i 的平均概率:

$$\hat{M}_{ij} = \frac{\sum_x P((\text{PL}_x = j) \cap (\text{TL}_x = i))}{\text{Count}(\text{TL}_x = i)} \quad (13)$$

式中: TL 表示真实标签; PL 表示预测标签; $\sum_x P((\text{PL}_x = j) \cap (\text{TL}_x = i))$ 是观测值 x 的预测标签为 j 而其真实标签为 i 的概率。

3 实验结果

3.1 实验设定及说明

定义 LSTM 分类器的相关参数如表 3 所示。在模型训练过程中, 本文采用 keras.callbacks 中的回调机制来加快训练过程并保存最佳分类模型。当损失值或准确率达到了最优值后, 就停止训练, 进而提高训练效率和模型拟合精度。

准化后的数据集输入建立的深层 LSTM 分类器中, 预测设备在特定时间窗口内的失效概率。通过设置时间窗口大小, 输出如表 4 所示的 9 组实验结果。其中, 一个 Epoch(时期) 表示使用训练集的全部数据对模型进行一次完整训练。

表 4 分类模型的实验结果

Table 4 Experimental results of classification model

组号	w_0	w_1	时期	准确率	精确率	召回率	F 值	AUC	时间/s
1	15	20	30	0.95	0.87	0.81	0.84	0.943	134
2	15	30	25	0.98	0.96	0.96	0.96	0.976	110
3	15	35	30	0.99	0.96	1	0.98	0.968	281
4	15	40	24	0.99	1	0.96	0.98	0.973	103
5	15	45	12	0.97	0.9	1	0.95	0.962	57
6	15	50	27	0.96	0.91	0.97	0.94	0.968	118
7	15	70	19	0.93	0.93	0.91	0.92	0.933	90
8	15	90	28	0.89	0.94	0.85	0.89	0.925	126
9	15	100	25	0.83	0.92	0.82	0.87	0.902	117

图 4 为 RUL 二分类性能评价指标的可视化展示。可以看出, 当 $w_1 = 35$ 和 $w_1 = 40$ 时, 模型的分类准确率最高为 0.99。然后, 随着时间窗口增大, Accuracy 逐渐减小到 0.83, 降幅为 16.2%; 当 $w_1 = 35$ 和 $w_1 = 40$ 时, F1-score 取值达到最大值 0.98, 可知两组实验具有相近的综合分类性能。由图 4 可知, 当 $w_1 > 40$ 时, F1-score 随 w_1 增大而逐渐减小; 另外, 第 3 组的 AUC 取值最大为 0.976, 这与第 4 组的结果无明显差别, $w_1 = 40$ 时的二分类 ROC 曲线如图 5 所示; 图 6 为模型运行时间随 w_1 的变化趋势 (红色虚线为平均运行时间), 可以看出第 3 组的运行时间最长为 281, 第 5 组的运行时间最小为 57, 两组的 AUC 取值相近, 因此在选择模型时可优先选择第 5 组。另外, 与第 2 组相比, 第 4 组具有更小运行时间, 即更高运行效率。

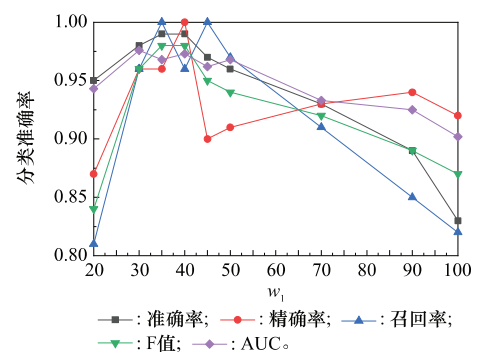
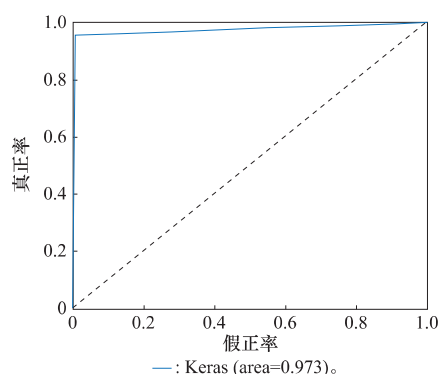
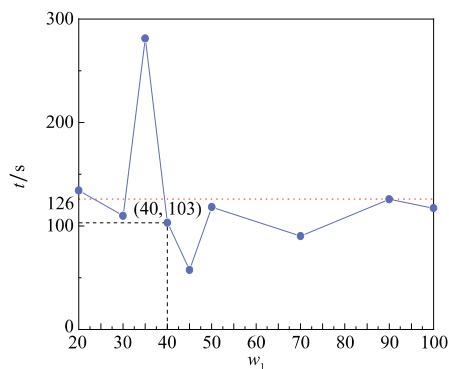
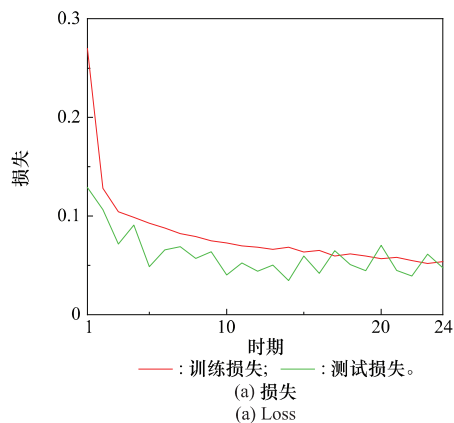
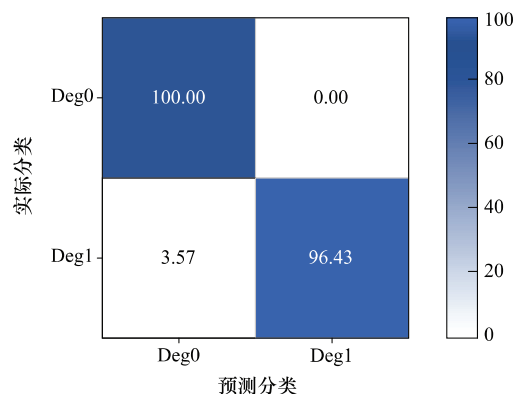


图 4 RUL 二分类性能评价结果

Fig. 4 RUL binary classification performance evaluation results

图5 $w_1 = 40$ 时的二分类 ROC 曲线Fig. 5 Two-class ROC curve when $w_1 = 40$ 图6 模型运行时间随 w_1 的变化趋势Fig. 6 Trend of model running time with w_1

综合以上分析可知,本文选择第4组为最优二分类模型,即时间窗口为 $w_0 = 15$ 和 $w_1 = 40$ 。图7描述了 $w_1 = 40$ 时的模型训练过程,随着时间的增加,训练集和验证集的 loss 逐渐减少。在运行 20 Epochs 以后,训练集与验证集的 loss 误差趋于平稳,并在 Epoch=24 时达到最优的模型训练性能。训练集和验证集的 Accuracy 呈现出与 loss 相反的变化趋势,同样在 Epoch=24 时达到最优值。图8为二分类模型测试集的概率混淆矩阵($w_0 = 15$, $w_1 = 40$)。当系统属于 Deg1 时,对于测试集,系统预测状态为 Deg0 的概率非常低,仅为 3.57%,而预测状态为 Deg1 的概率为 96.43%;当系统属于 Deg0 时,对于测试集,系统预测状态都为 Deg0,这表明此时模型具有优良的分类性能。

图7 $w_1 = 40$ 时的模型训练过程Fig. 7 Model training process when $w_1 = 40$ 图8 测试集的概率混淆矩阵($w_0 = 15$, $w_1 = 40$)Fig. 8 Probability confusion matrix of the test set ($w_0 = 15$, $w_1 = 40$)

3.3 与其他方法的比较

本文将提出的方法与 Python scikit-learn 库中现有方法进行了对比,包括逻辑回归、决策树、随机森林(random forest, RF)、支持向量分类(support vector classification, SVC)、K 近邻(K-nearest neighbors, KNN)、高斯朴素贝叶斯分类器(Gaussian naive Bayesian classifier, Gaussian NB),各种方法的窗口大小设置均为 $w_0 = 15$ 和 $w_1 = 40$ 。

在 Logistic Regression 中,分类方式参数为“ovr”,并采用“lbfgs”作为求解器,利用海森矩阵对损失函数进行迭代优化,最大迭代次数为 100;在 Decision Tree 中,特征选择标准为“entropy”,决策树最大深度为 4,最小叶子节点为 1,最小内部节点为 2,不考虑叶子节点的权重;在 Random Forest 中,特征选择标准为“entropy”,决策树最大深度为 6,最小叶子节点为 1,最小内部节点为 2,决策树个数为 50,并行工作数为 1;在 SVC 中,核函数为“rbf”,函数维度为 3,核函数参数为“auto”,不限制最大迭代次数,停止训练的误差值为 0.001;在 KNN 中,数的大小为 30,树的距离度量为欧几里德度量,并行工作数为 1,近邻数为 13,预测的权重函数为平均加权;在 Gaussian NB 中,先验概率 priors=None,即获取各类别的先验概率。

表 5 中比较了所提方法和在发动机测试数据集 FD001 中的性能,性能指标的可视化对比结果如图 9 所示。与其他方法相比,本文所提方法具有最大的 Accuracy、Precision、Recall 和 F1-Score。特别地,F1-Score 为反映模型分类性能的综合指标,由图 9 可以直观地看出所构建的 LSTM 分类器较现有方法具有显著的优越性,尤其适用于多元长序列传感器数据的处理过程。对比其他方法中最优方法的性能,所提方法的准确率提高了 5.31%,而 F1-Score 提高了 10.11%。这意味着所提出的方法具有最好的分类性能,表明了所提方法对发动机 RUL 分类问题的有效性。

表 5 与现有方法的性能对比

Table 5 Performance comparison with existing methods

编号	方法	准确率	精确率	召回率	F 值
1	逻辑回归	0.92	1	0.68	0.81
2	决策树	0.92	0.95	0.72	0.82
3	RF	0.91	0.94	0.68	0.79
4	SVC	0.92	0.95	0.72	0.82
5	KNN	0.92	0.95	0.72	0.82
6	Gaussian NB	0.94	0.83	0.96	0.89
7	本文方法	0.99	1	0.96	0.98

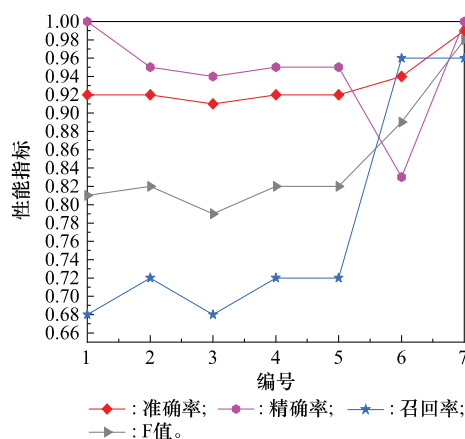


图 9 与其他方法的性能指标对比

Fig. 9 Comparison of performance indicators with other methods

4 结 论

(1) 本文提出了基于 LSTM 分类器的航空发动机预测性维护模型,与一般预测 RUL 值的方法不同,所提方法提供了设备 RUL 落入特定时间窗口的概率。

(2) 采用滑动时间窗口定义训练样本标签,然后将预处理后的数据集输入建立的深层 LSTM 分类器中,预测设备在特定时间窗口内的失效概率。通过分析 w_1 对故障概率的影响,得到最优性能的 LSTM 分类模型,以更好地适应实际维护需求。在特定的时间窗口内,维护工程师可以根据 RUL 分类信息来安排维护和生产活动。

(3) 提出的模型在美国国家航空航天局的 C-MAPSS 的数据集上进行了验证,评价指标均优于其他现有分类模型,验证了 LSTM 分类器的有效性。同时,更加准确的 RUL 分类模型可降低维护成本,提高维护效率。

参考文献

- [1] GANIN Y, USTINOVA E, AJAKAN H, et al. A remaining useful life prediction method with long-short term feature processing for aircraft engines[J]. Applied Soft Computing, 2020, 93: 106344.
- [2] CARVALHO T P, SOARES F, VITA R, et al. A systematic literature review of machine learning methods applied to predictive maintenance [J]. Computers & Industrial Engineering, 2019, 137: 106024.
- [3] 马奇友, 刘可薇, 杜坚, 等. 基于深度长短期记忆网络的发动机叶片剩余寿命预测[J]. 推进技术, 2021, 42(8): 1888-1897.
- [4] MA Q Y, LIU K W, DU J, et al. Prediction of residual life of engine blades based on deep short term memory network[J]. Journal of Propulsion Technology, 2021, 42(8): 1888-1897.
- [5] SNCHEZ S M, FRANGOPOL D M, PADGETT J, et al. Maintenance and operation of infrastructure systems: review [J]. Journal of Structural Engineering, 2016, 142(9): F4016004.
- [6] WANG K S, WANG Y. How AI affects the future predictive maintenance: a primer of deep learning[C]//Proc. of the International Workshop of Advanced Manufacturing and Automation, 2018.
- [7] GOURIVEAU R, MEDJAH K, ZERHOUNI N. From prognostics and health systems management to predictive maintenance[M]. London: Wiley, 2016.
- [8] NGUYEN K T P, MEDJAH K. A new dynamic predictive maintenance framework using deep learning for failure prognostics[J]. Reliability Engineering & System Safety, 2019, 188: 251-262.
- [9] YU W, KIM I Y, MECHEFSKE C. An improved similarity based prognostic algorithm for RUL estimation using an RNN auto encoder scheme[J]. Reliability Engineering & System Safety, 2020, 199: 106926.
- [10] XIA J, FENG Y W, LU C, et al. LSTM-based multi-layer self-attention method for remaining useful life estimation of mechanical systems[J]. Engineering Failure Analysis, 2021, 125: 105385.
- [11] LEI Y G, LI N P, GONTARZ S, et al. A model-based method for remaining useful life prediction of machinery [J]. IEEE Trans. on Reliability, 2016, 65(3): 1314-1326.
- [12] SI X S, WANG W B, HU C H, et al. Remaining useful life estimation-a review on the statistical data driven approaches[J]. European Journal Operational Research, 2011, 213(1): 1-14.
- [13] LIU Y C, HU X F, ZHANG W J. Remaining useful life prediction based on health index similarity[J]. Reliability Engineering & System Safety, 2019, 185: 502-510.
- [14] ZHANG J J, WANG P, YAN R Q, et al. Long short-term memory for machine remaining life prediction[J]. Journal of Manufacturing Systems, 2018, 48: 78-86.
- [15] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [16] CHE C C, WANG H W, FU Q, et al. Combining multiple deep learning algorithms for prognostic and health management

- of aircraft[J]. Aerospace Science and Technology, 2019, 94: 105423.
- [16] TAMILSELVAN P, WANG P. Failure diagnosis using deep belief learning based health state classification[J]. Reliability Engineering & System Safety, 2013, 115:124-135.
- [17] GUO L, LI N P, JIA F, et al. A recurrent neural network based health indicator for remaining useful life prediction of bearings[J]. Neurocomputing, 2017, 240:98-109.
- [18] HINCHI A Z, TKIOUAT M. Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network[J]. Procedia Computer Science, 2018, 127:123-132.
- [19] ALDULAIMI A, ZABIHI S, ASIF A, et al. A multimodal and hybrid deep neural network model for remaining useful life estimation[J]. Computers in Industry, 2019, 108: 186-196.
- [20] YUAN M, WU Y T, LIN L. Fault diagnosis and remaining useful life estimation of aero engine using LSTM neural network[C]//Proc. of the IEEE International Conference on Aircraft Utility Systems, 2016: 135-140.
- [21] 张妍, 王村松, 陆宁云, 等. 基于退化特征相似性的航空发动机寿命预测[J]. 系统工程与电子技术, 2019, 41(6):1414-1421.
- ZHANG Y, WANG C S, LU N Y, et al. Aeroengine life prediction based on similarity of degradation characteristics[J]. Systems Engineering and Electronics, 2019, 41(6):1414-1421.
- [22] 车畅畅, 王华伟, 倪晓梅, 等. 基于多尺度排列熵和长短时记忆神经网络的航空发动机剩余寿命预测[J]. 交通运输工程学报, 2019, 19(5):106-114.
- CHE C C, WANG H, NI X M, et al. Aero-engine remaining life prediction based on multi-scale permutation entropy and long-short-term memory neural network[J]. Journal of Traffic and Transportation Engineering, 2019, 19(5):106-114.
- [23] 张永峰, 陆志强. 基于集成神经网络的剩余寿命预测[J]. 工程科学学报, 2020, 42(10):1372-1380.
- ZHANG Y F, LU Z Q. Remaining useful life prediction based on an integrated neural network[J]. Chinese Journal of Engineering, 2020, 42(10):1372-1380.
- [24] ZHANG Y Z, XIONG R, HE H W, et al. Long short-term memory recurrent neural network for remaining useful life prediction of lithiumion batteries[J]. IEEE Trans. on Vehicular Technology, 2018, 67(7): 5695-5705.
- [25] SAXENA A, KAI G, SIMON D, et al. Damage propagation modeling for aircraft engine run-to-failure simulation[C]//Proc. of the International Conference on Prognostics and Health Management, 2008.
- [26] PEEL L. Data driven prognostics using a Kalman filter ensemble of neural network models[C]//Proc. of the International Conference on Prognostics and Health Management, 2008.
- [27] XIA T B, SONG Y, ZHENG Y, et al. An ensemble framework based on convolutional bi-directional LSTM with multiple time windows for remaining useful life estimation[J]. Computers in Industry, 2020, 115: 103182.
- [28] ZHANG B, ZHANG S H, LI W H. Bearing performance degradation assessment using long short-term memory recurrent network[J]. Computers in Industry, 2019, 106:14-29.
- [29] VERSTRAETE D, DROGUETT E, MODARRES M. A deep adversarial approach based on multi-sensor fusion for semi-supervised remaining useful life prognostics[J]. Sensors, 2020, 20(1): 176-186.

作者简介

蔺瑞管(1993—),男,博士研究生,主要研究方向为航空器可靠性。
王华伟(1974—),女,教授,博士,主要研究方向为航空器可靠性工程、航空器安全性分析技术。
车畅畅(1994—),男,博士研究生,主要研究方向为航空器可靠性。
倪晓梅(1992—),女,博士研究生,主要研究方向为适航管理。
熊明兰(1997—),女,博士研究生,主要研究方向为适航管理。