

一种基于可解释人工智能技术的航空发动机健康衰退分析框架

张婉渝

成都天奥集团有限公司研发中心, 四川 成都

收稿日期: 2024年11月5日; 录用日期: 2024年12月30日; 发布日期: 2025年1月7日

摘要

航空发动机作为复杂且关键的机械设备, 其健康状态直接影响到飞行安全与运营效率。鉴于目前健康状态监测技术在可解释性和准确性上的局限性, 本文提出了一种基于Shapley值的可解释人工智能分析框架。该框架通过解析机器学习模型的预测结果, 揭示了发动机健康衰退的根本原因。首先, 运用Shapley值分析法评估各传感器数据特征的边际贡献, 识别了影响航空发动机健康的关键因素及其衰退路径。接着, 应用HDBSCAN聚类算法识别潜在故障风险的发动机。最后, 利用SkopeRules从聚类中提取易于理解且准确度高的规则集合, 作为复杂模型决策的替代方案, 增强了模型输出结果的透明度和可靠性。实验结果表明, 该框架能够揭示航空发动机健康衰退的演变路径, 为实现精准维护提供可靠的理论支撑。

关键词

Shapley值, 航空发动机, 聚类, 健康状态监测, SkopeRules

An Explainable Artificial Intelligence Framework for Analyzing Health Degradation of Aero-Engines

Wanyu Zhang

R&D Center, Chengdu Spaceon Group Co., Ltd., Chengdu Sichuan

Received: Nov. 5th, 2024; accepted: Dec. 30th, 2024; published: Jan. 7th, 2025

Abstract

As a complex and critical mechanical equipment, the health status of aircraft engines directly affects

文章引用: 张婉渝. 一种基于可解释人工智能技术的航空发动机健康衰退分析框架[J]. 人工智能与机器人研究, 2025, 14(1): 1-13. DOI: 10.12677/airr.2025.141001

flight safety and operational efficiency. Given the limitations of current health status monitoring technologies in terms of interpretability and accuracy, this paper proposes an interpretable artificial intelligence analysis framework based on Shapley values. This framework reveals the fundamental cause of engine health decline by analyzing the prediction results of machine learning models. Firstly, the Shapley value analysis method was used to evaluate the marginal contribution of each sensor data feature, identifying the key factors affecting the health of aircraft engines and their degradation paths. Next, the HDBSCAN clustering algorithm is applied to identify engines with potential fault risks. Finally, SkopeRules is used to extract a set of rules from clustering that are easy to understand and highly accurate, as an alternative to complex model decision-making, enhancing the transparency and reliability of the model output results. The experimental results indicate that the framework can reveal the evolutionary path of health decline in aircraft engines, providing reliable theoretical support for achieving precise maintenance.

Keywords

Shapley Values, Aero-Engine, Clustering, Health Monitoring, SkopeRules

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

航空发动机作为飞机的关键动力源，其健康状况直接关乎飞行安全。航空发动机长期在高温、高压、高转速环境下运行，频繁遭受剧烈振动和冲击，极易导致发动机健康衰退，进而产生发动机推力不足、燃油效率降低等问题，严重时甚至可能引发灾难性事件。因此，通过航空发动机健康衰退研究，及时发现并预防故障隐患，对于保障飞行安全具有重要意义。

早期航空发动机健康衰退研究主要依赖于复杂的物理模型，涵盖材料、结构、热力学和流体力学等多个方面。例如，丁小飞等[1]构建了双转子动力学模型，追溯了振动的根源并提出振动抑制策略。栾孝驰等[2]结合声发射和振动分析技术，以提高对航空发动机故障识别能力。尽管此类方法具有较高的准确性和可靠性，但建模过程复杂且耗时，并且对专业性要求极高。

随着数据处理技术的快速发展，人工智能已成为重要的分析工具。通过对庞杂的发动机运行数据进行剖析并提取关键特征，从而构建发动机健康状态预测模型。孙同敏[3]通过将深度置信网络与最小二乘支持向量机相结合，实现了对航空发动机健康状态的自动化评估。张世杰等[4]采用多层感知机和时空图卷积网络来诊断航空发动机气路故障，为航空发动机健康管理探索了新的方向。张博文等[5]改进了一维卷积神经网络，在标准卷积层前加入了概率密度信息层，进而提高了轴承健康监测的准确性。

尽管人工智能技术以其信息处理优势能够实现高精度预测，但在航空工业这一对安全要求极其苛刻的领域中，任何微小的概率误差都可能酿成灾难。因此，提升模型的透明度、增强其可靠性和稳健性对于推动人工智能技术在航空工业中的广泛应用显得尤为重要。所以，周志杰等[6]提出的基于规则的方法，能够清晰地展现系统的核心运作逻辑。在此基础上，程晓玉等[7]基于置信规则库建模技术，以增强模型的可解释性。但 Zhijie Z 等[8]为了解决置信规则库面临的解释性损失问题，引入了一种包含“搜索强度”、“信念分布的可解释性约束”以及“规则激活因子”的新型优化策略。与此同时，王诗彬等[9]则将先验知识嵌入稀疏表示模型，构建高透明度核心网络，从而有效改善了模型的可解释性。

尽管现有研究展示了人工智能在航空发动机健康管理中的显著成效，并初步探索了模型的可解释性。

但仍存在若干不足，目前提升决策透明度的策略主要依赖于专家经验制定的规则和设计可解释的模型结构。传统规则易于理解但难以适应高维动态特征空间；而基于可解释模型结构的方法虽能较好处理高维数据，但其解释机制易随模型结构变化而失效。为克服这些局限，本文构建了一种基于 Shapley 值的可解释人工智能框架。该框架通过预测航空发动机健康状态，利用 Shapley 值进行特征重要性分析，并结合聚类技术和规则挖掘算法，形成一种模型无关且高度可解释的技术方案。此框架不仅能够深入解析发动机健康衰退机制，还具有广泛应用于工业装备健康管理及故障预诊断的潜力。

2. 研究方法

为了深入剖析航空发动机健康状态退化的内在机制，本文构建了一种基于可解释人工智能技术的分析框架如图 1 所示。该框架首先对原始数据进行预处理并提取关键统计特征，然后建立健康状态预测模型。利用 Shapley 值方法量化各传感器特征对预测结果的边际贡献，识别出影响发动机健康衰退的关键因素及其路径。接着，通过降维和 HDBSCAN [10] 聚类算法识别潜在故障风险群体，并采用 SkopeRules [11] 从这些群体中抽取易于理解和高准确度的规则集，以增强模型决策的透明度与可靠性。随后部分将详细介绍所用核心技术：Shapley 值、非线性降维、HDBSCAN 聚类及 SkopeRules 规则挖掘。

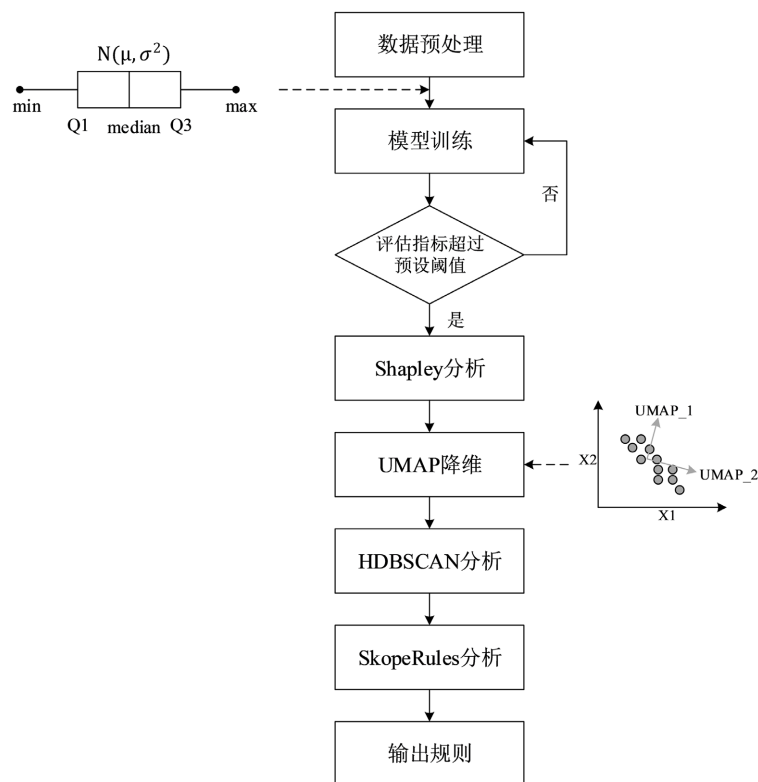


Figure 1. Explainable artificial intelligence analysis framework
图 1. 可解释人工智能分析框架

2.1. Shapley

Shapley 值的概念起源于博弈论，其理论基础在于对参与者个体贡献的量化分析，即在一个协作环境中，每位参与者的边际贡献取决于参与者进入合作序列的位置及其对于最终成果的影响程度。随着理论框架的不断演进，Shapley 值逐渐成为解析模型内部机制的关键手段之一，并由 Lundberg [12] 等学者进一

步优化。将这一方法应用于机器学习领域时，特别是在构建航空发动机健康状态监测系统的过程中，Shapley 值提供了一种有效的方法来评估各个输入变量(如温度、压力及振动水平)对于最终预测结果的重要性。例如，针对特定发动机样本 x 的健康状况预估值可以表达为基准预测值与所有相关特征对应的 Shapley 值之和：

$$f(x) = E[f(X)] + \sum_{i=1}^N \phi_i(f, x) \quad (1)$$

其中， $E[f(X)]$ 表示基准预测值， X 为测试样本集合， $\phi_i(f, x)$ 表示第 i 个特征对发动机样本 x 健康状态预测结果的具体影响度量，可通过下述公式精确计算得出：

$$\phi_i(f, x) = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{N!} [f_x(S \cup \{i\}) - f_x(S)] \quad (2)$$

其中， F 代表所有特征集合， N 代表特征总数， S 表示不含特征 i 的任意子集， f 则是预测模型， $f_x(S \cup \{i\}) - f_x(S)$ 反映了添加特征 i 后相对于原有子集 S 预测输出的边际增益变化。

进一步地，为了从全局视角考察某一特定特征在整个数据集内对模型表现的平均影响力，定义特征 i 的整体 Shapley 值如下所示：

$$\phi_i(f) = \frac{1}{|X|} \sum_{x \in X} \phi_i(f, x) \quad (3)$$

通过 Shapley 分析方法，不仅能够量化单个因素的重要程度，也能揭示出各因素间复杂的交互作用模式，为后续开展针对性维护策略制定提供了坚实的理论依据。

2.2. UMAP

UMAP [13] (Uniform Manifold Approximation and Projection) 是一种基于流形学习和拓扑数据分析的非线性降维技术。核心原理是将高维数据中的局部邻域结构映射到低维空间中，以便更好地进行分析和可视化。UMAP 的算法分为两个关键步骤：

1) 计算高维空间中的局部相似性

UMAP 首先在高维空间中为每个数据点构建一个局部邻域图，邻域图通过构建每个数据点的 k 最近邻点实现。对于高维空间中的点 x_i 和 x_j ，其局部相似性计算公式为：

$$p_{ji} = \frac{\exp\left(-\|x_i - x_j\|^2 / \sigma_i^2\right)}{\sum_{k \in N(i)} \exp\left(-\|x_i - x_k\|^2 / \sigma_i^2\right)} \quad (4)$$

其中， σ_i 是以 x_i 中心的高斯核的带宽超参数，通常根据邻域内点的密度来确定， $N(i)$ 表示 x_i 的 k 近邻集合。为了使矩阵对称，使用以下公式计算最终的相似性矩阵， n 是数据点的数量：

$$p_{ij} = \frac{p_{ji} + p_{ij}}{2n} \quad (5)$$

2) 优化低维嵌入

在低维空间中，UMAP 试图找到一个嵌入 y_j ，使得低维空间中的分布 q_{ij} 尽可能接近高维空间中的分布 p_{ij} ，然后通过最小化高维和低维空间分布之间的交叉熵来优化数据在低维空间中的布局，并通过梯度下降等优化算法进行求解，以找到最佳的低维嵌入。优化目标为：

$$C(P, Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \quad (6)$$

2.3. HDBSCAN

HDBSCAN 作为一种基于密度的空间聚类技术，在 DBSCAN 的基础上进行了改进，能够更有效地处理不同密度级别的聚类及噪声点，从而揭示数据内部的潜在子群。相较于传统的 k-means [14] 与模糊 C 均值(fuzzy c-means) [15] 等聚类方法，HDBSCAN 的优势在于无需预先设定聚类的数量。具体实现步骤如下：

1) 计算互达距离：给定半径 ε 和最小样本量超参数，对于每一对点 p 和 q ，计算它们之间的互达距离。

$$core_dist_\varepsilon(p) = \min(\varepsilon, dist_k(p)) \quad (7)$$

$$reachability_dist = \max(core_dist_\varepsilon(p), core_dist_\varepsilon(q), d(p, q)) \quad (8)$$

其中， $dist_k(p)$ 表示 p 点到第 k 个邻点的距离， $core_dist_\varepsilon(p)$ 是 p 点的核心距离， $d(p, q)$ 为点 p 和 q 之间的欧式距离。

2) 构建最小生成树：基于互达距离构建一个邻域图，其中每个点都与它的近邻点相连，连接边的权重为该点与邻点的互达距离，然后基于 Kruskal 等算法构建最小生成树。

3) 生成层次聚类树：基于构建好的最小生成树，从树中选择权重最小的边，将这条边所连接的两个顶点合并成一个新的簇。重复此过程直到所有点都被合并，以此形成一个自底向上的层次聚类树。

4) 切割树并选择最优簇：为了选择最优的簇集合，评估不同层次上的簇稳定性，选择表现出较高稳定性的层次作为最终的簇集合。在最终确定的簇中，不属于任何簇的点被视为噪声点。

2.4. SkopeRules

SkopeRules 是一种用于学习规则的机器学习方法，通过生成一组易于理解的逻辑规则进行分类。规则生成和选择过程可以描述为下步骤：

1) 初始化：选择一个特征子集 F 和一个初始规则集合 $R = \Phi$ 。

2) 规则生成：对于每个特征 $f_i \in F$ ，生成一系列候选规则 r_{ij} ， j 表示规则序号。

3) 规则评估：对于每条规则 r_{ij} ，计算性能指标，包括精确率、召回率、F1 分数。

4) 规则选择：选择 F1 分数最高的规则加入规则集合 R 。

5) 重复规则生成、评估、选择步骤，直到没有更好的规则可添加或者达到预设的停止条件，最后输出规则集合。

3. 实验

3.1. 数据集

本研究采用的数据来自 NASA Ames 研究中心、苏黎世联邦理工学院及 PARC 联合开发的 CMAPSS 动态模型生成的 N-CMAPSS [16] 数据集，动态模型结构如图 2 所示。该数据集模拟了实际飞行条件下涡轮扇发动机从运行到故障的过程，包含 8 个子集，6825 个飞行周期，涵盖 90 个发动机单元，每个单元平均约 75 个周期。数据集中包括 18 个传感器指标的持续监测读数以及每个操作周期内固定的辅助变量，变量描述详见表 1。数据集中航空发动机的健康状态被标记为“健康”或“不健康”。本研究旨在利用 N-CMAPSS 数据集验证所提方法的有效性和准确性，以解析航空发动机的健康衰退机制。

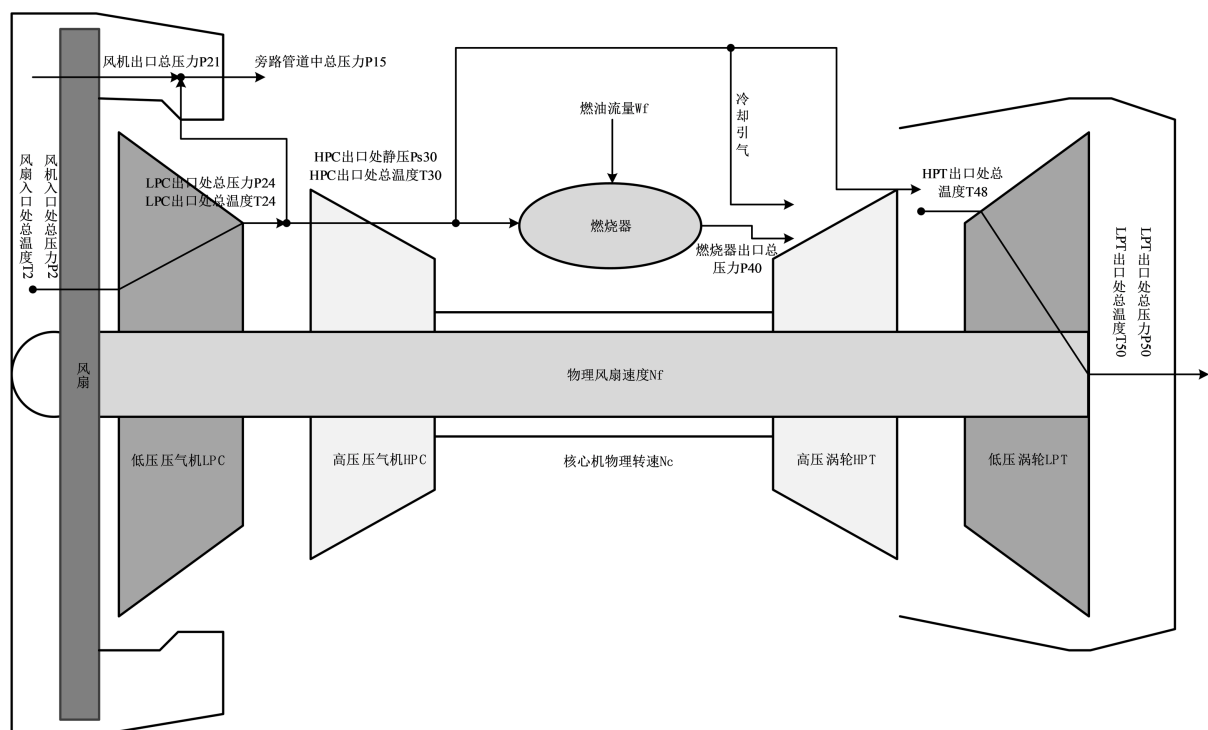


Figure 2. Schematic diagram of a turbofan engine
图 2. 涡轮风扇发动机示意图

Table 1. Variable descriptions
表 1. 变量描述

变量符号	描述	单位
hs	健康状态	-
RUL	剩余使用寿命	年
cycle	飞行周期	-
Fc	飞行等级	-
alt	海拔	ft
March	飞行马赫数	-
TRA	节气门开启角度	%
T2	风扇入口处总温度	°R
Wf	燃油流量	pps
Nf	物理风扇速度	rpm
Nc	核心机物理转速	rpm
T24	低压压气机(LPC)出口处总温度	°R
T30	高压压气机(HPC)出口处总温度	°R
T48	高压涡轮(HPT)出口处总温度	°R
T50	低压涡轮(LPT)出口处总温度	°R

续表

P15	旁路管道中总压力	psia
P2	风机入口处总压力	psia
P21	风机出口总压力	psia
P24	低压压气机(LPC)出口处总压力	psia
Ps30	高压压气机(HPC)出口处静压	psia
P40	燃烧器出口总压力	psia
P50	低压涡轮(LPT)出口处总压力	psia

3.2. 实验步骤

本研究专注于航空发动机健康状态评估,结合 Shapley 值分析以识别关键影响因素。基于分析结果,采用 UMAP 技术进行数据降维,并运用 HDBSCAN 聚类算法识别潜在故障发动机,进一步通过 SkopeRules 方法对聚类结果提取规则,进而实现复杂模型透明化,为不同衰退程度的发动机提供维护策略与优化建议。具体实验步骤如下:

1) 数据预处理:从 18 个时间序列变量中提取了 7 个关键统计特征:平均值、标准差、最小值、四分之一位数、中位数、四分之三位数和最大值,例如 LPT 出口处总温度的特征表示为 T50_mean、T50_std、T50_min、T50_Q1、T50_median、T50_Q3、T50_max。此外,还加入了飞行周期和飞行等级两个辅助变量,共 128 个特征。通过对不同发动机在多个运行周期内的这些特征进行提取,构建了一个包含 6825 个样本的数据集,并进行归一化处理。最终,按 8:2 的比例随机划分为训练集(5460 个样本)和测试集(1365 个样本)。

2) 模型训练:为有效提取航空发动机健康衰退的关键信息,本文构建了两个模型:神经网络模型 M1 和 XGBoost 分类模型 M2。M1 采用三层前馈神经网络,包含两个 48 神经元的隐藏层和一个 2 节点的输出层,使用 ReLU 作为激活函数。M2 基于 XGBoost 算法,设置 100 棵决策树、最大深度为 4,并引入 L1 正则化以增强泛化能力。经过训练优化,两个模型在 F1 分数上均超过 90%。

3) Shapley 值分析:基于训练完成的 M1 和 M2 模型对测试集进行预测。然后,计算测试集中每个特征对于预测结果的 Shapley 值。通过对 Shapley 值的分析,揭示各个特征变量对于预测结果的贡献度和排序,并对两个模型的解释结果进行对比分析,从而求证 Shapley 值分析法的模型无关性。

4) 聚类分析:首先,运用 UMAP 方法分别对原始特征以及 M1 模型的 Shapley 值进行降维处理。随后,利用 HDBSCAN 算法对降维后的结果进行聚类分析,探索发动机健康状态中的潜藏子群体。最后,探究各子群体中发动机的剩余使用寿命特性。

5) SkopeRules 分析:基于聚类分析的结果,运用 SkopeRules 算法方法来探索聚类子群中的潜藏模式,并从中推导出高精度透明化的规则集合。

3.3. 实验结果及讨论

3.3.1. Shapley 分析

在实施 Shapley 值分析的过程中,为了全面评估 M1 和 M2 模型的性能,采用了精确率、召回率、F1 分数作为评估指标。如表 2 所示,健康与不健康类别预测结果的 F1 分数均超过 97%,充分验证了两个模型在评估发动机健康状态方面的精确性。

Table 2. Evaluation metrics for M1 and M2 models
表 2. M1 与 M2 模型评估指标

	标签	精确率	召回率	F1 分数
M1	0 (不健康)	0.982	0.992	0.986
	1 (健康)	0.984	0.966	0.974
M2	0 (不健康)	0.990	0.990	0.990
	1 (健康)	0.973	0.982	0.977

在航空发动机健康状态预测方面，M1 和 M2 模型均表现出显著的有效性。然而，对于处于非健康状态的发动机，这两个模型未能揭示其健康衰退的具体原因。为此，本研究引入 Shapley 值分析法以量化并排序各特征对两模型预测结果的贡献度，进而探究发动机健康衰退的根本因素。图 3 与图 4 分别展示了基于 M1 和 M2 模型下，不同特征按照其重要性进行排序的结果。在此，每个数据点代表一个独立的预测案例；横轴显示了特定特征的 Shapley 值，用以衡量该特征对模型预测边际效应的影响程度；纵轴则标注了具体特征数值。此外，通过颜色编码增强了可视化效果：红色表示较高特征值的数据点，蓝色则对应较低特征值，从而直观地反映了各个特征对模型输出影响的不同程度。

通过对图 3 与图 4 的综合分析，可以得出结论：飞行周期(cycle)在所有因素中占据核心地位，是决定发动机健康状态的关键因素。随着飞行周期的数量增长，发动机健康状态呈现恶化趋势。此外，飞行等级(Fc)作为次要影响因素，细分为三个级别：1、2、3，分别对应短途、中途和长途飞行任务，表 3 详细列出了在实验条件下各飞行等级所对应的飞行条件。其中，长途飞行对发动机造成的磨损相对较大，而短途飞行则对健康状态的影响较小。除上述关键因素外，其他特性对于发动机健康状况的影响显得较为有限，未能形成显著的相关性。图 3 与图 4 之间关于飞行周期及飞行等级所揭示的趋势表现出高度的一致性，仅在一些影响力较弱的因素上存在细微差别。这一发现有效验证了 Shapley 值分析作为一种模型无关解释方法的有效性，无论是在基于神经网络还是集成学习构建的预测模型框架下，该方法都能够一致地识别出影响发动机健康状态的主要因素。

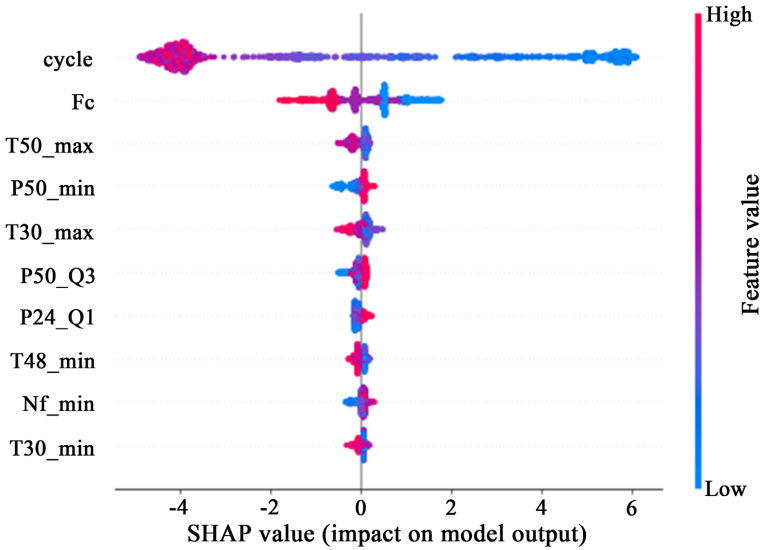


Figure 3. Shapley value distribution based on the M1 model
图 3. 基于 M1 模型的 Shapley 值分布图

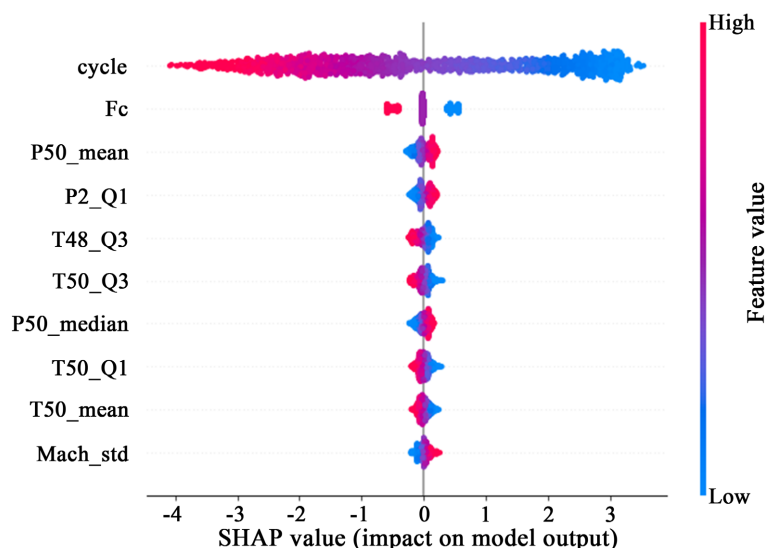


Figure 4. Shapley value distribution based on the M2 model
图 4. 基于 M2 模型的 Shapley 值分布图

Table 3. Flight level classification
表 3. 飞行等级划分

飞行等级	飞行时长(小时)	飞行次数
1	1~3	18
2	3~5	149
3	>5	185

3.3.2. HDBSCAN 分析

在上述 Shapley 分析中, 我们得知飞行周期是航空发动机健康衰退最重要的因素, 并且随着飞行周期的逐渐增加, 其健康状态呈现逐渐衰退的趋势, 表明了健康与不健康状态之间边界的模糊性, 两者间缺乏一个绝对的阈值区分。为深入探究介于这两种明确状态之间的其他潜藏状态, 本文采用 HDBSCAN 算法对样本数据进行聚类分析。旨在识别健康与不健康类别之外的样本群体, 并揭示这些样本群内在的分布规律。

为了更为直观地展示聚类结果, 本文采用 UMAP 方法对特征进行降维, UMAP 关键参数如表 4 所示:

Table 4. Key parameter settings for UMAP
表 4. UMAP 关键参数设置

参数名称	参数值
最近邻点数量	200
样本间最小距离	0.01
嵌入的维度	2
距离度量方法	correlation
迭代次数	500

根据上述 UMAP 方法降维得到低维特征向量 UMAP_1 和 UMAP_2, 并采用 HDBSCAN 算法对测试

样本进行聚类，HDBSCAN 算法的关键参数如表 5 所示：

Table 5. Key parameter settings for HDBSCAN
表 5. HDBSCAN 关键参数设置

参数名称	参数值
最小样本数	200
核心点最小邻居数	10
距离度量方法	euclidean
最近邻方法	kd_tree

HDBSCAN 聚类算法应用于低维特征向量的结果如图 5 所示，揭示了除健康和亚健康状态之外存在难以明确分类的航空发动机样本。这一发现支持了随着飞行周期增加，部分发动机表现出介于完全健康与完全不健康之间的过渡状态的观点。然而，图 5 中显示的簇间边界模糊及较低的内聚性，特别是对于那些分布更为分散的“其他”类别发动机，表明将这些样本归为同一类别并不十分合理，也反映出原始特征信息未能得到充分利用。为此，本研究进一步采用 HDBSCAN 算法对经 Shapley 值降维处理后的特征向量进行了聚类分析，其结果见图 6。

从图 5 和图 6 的对比结果可以看出，图 6 所呈现的聚类效果更好，类内距离更小，类间距更大，簇之间呈现明确的可分性，能够直观地反映存在潜藏风险的航空发动机群体。此对比结果有效证实了 Shapley 值对于特征关键信息的有效攫取，同时也表现出该方法在聚类算法中的有效运用。

3.3.3. SkopeRules 分析

在上述实验结果中，我们识别出一类具有潜在风险的航空发动机，其特征表现为剩余使用寿命介于健康与亚健康状态之间。基于此发现，本研究根据平均剩余使用寿命将航空发动机重新分类为健康、亚健康和不健康三类。为进一步提高模型决策过程的透明度，并直观描述各类别形成的条件，采用 SkopeRules 方法对这三种健康状态进行规则提取，其结果见图 7。所获得的规则集旨在替代原有的神经网络模型，以降低部署成本和推理时间，增强预测模型在实际应用中的可解释性和实用性。

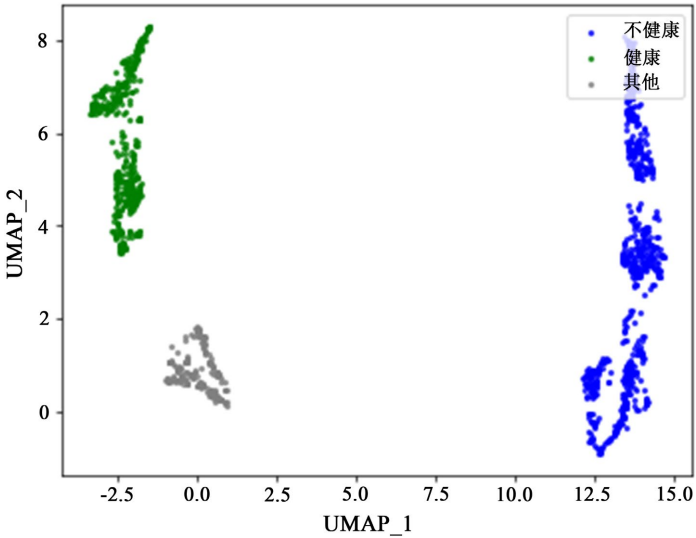


Figure 5. Clustering of engine health status using dimensionality reduction of original features
图 5. 基于原始特征降维的发动机健康状态聚类图

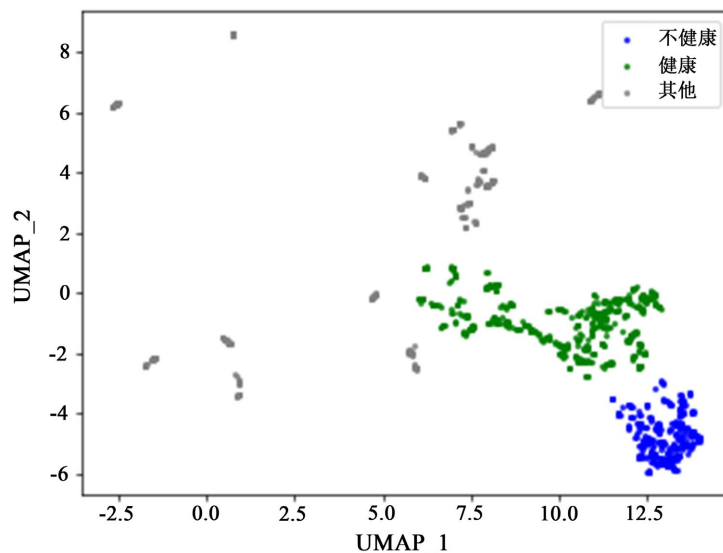


Figure 6. Clustering of engine health status based on dimensionality reduction using Shapley values
图 6. 基于 Shapley 值降维的发动机健康状态聚类图

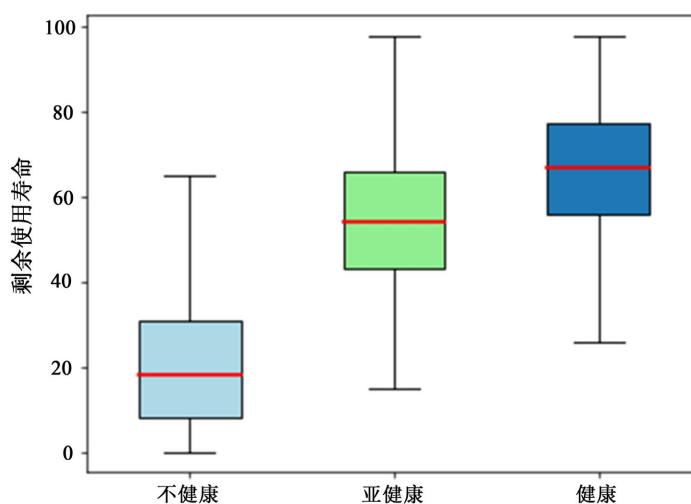


Figure 7. Box plot of remaining useful life for subgroups of engine health status
图 7. 发动机健康状态子群的剩余使用寿命箱线图

本文采用 Python 中的 SkopeRules 工具对聚类结果进行描述性规则提取，其关键参数设置如表 6 所示。

Table 6. Key parameter settings for SkopeRules
表 6. SkopeRules 关键参数设置

参数名称	参数值
树的最大去重深度	3
树的数量	100
最小精确度	0.5
最小召回率	0.5

通过 SkopeRules 对发动机健康状态三大类别提取的规则如表 7 所示：

Table 7. Rule descriptions for engine health status
表 7. 发动机健康状态的规则描述

健康状态	规则	精确率	召回率	F1 分数
健康	飞行周期数 ≤ 34.5 飞行等级 ≤ 1.5	0.701	0.933	0.801
亚健康	飞行周期数 ≤ 36.5 飞行等级 > 1.5	0.995	0.950	0.972
不健康	飞行周期数 > 36.5	0.998	0.975	0.986

从表 7 可以看出，飞行周期数是决定航空发动机健康状态的关键因素，这与 Shapley 分析结果一致，验证了其结论的可靠性。当飞行周期数超过 36.5 时，发动机更易进入不健康状态。由于单个飞行周期内执行的任务类型多样，仅凭飞行周期数难以准确判断发动机健康状况，还需结合飞行等级来辅助识别。较高飞行等级意味着更长的飞行时间和更多班次，加速了发动机的健康退化，从而处于亚健康状态。相对于 M1 和 M2 模型的预测精度，在规则集下的健康和亚健康类别的预测精度略有下降，而不健康类别的 F1 分数保持不变，三类别的 F1 分数均超过 80%。尽管整体分类精度有所牺牲，但三分类提供了更为细致的状态描述，有助于针对不同健康状态制定差异化维护计划，从而提升资源利用效率及维修效果。

4. 结论

本研究提出了一种基于 Shapley 值的可解释人工智能框架，旨在深入探讨航空发动机健康状态退化机制。通过该框架的应用，实现了对各个因素对航空发动机健康影响的量化评估，揭示了健康衰退机制。研究结果表明，在所有考虑因素中，飞行周期数对发动机健康状态的影响最为显著。此外，研究证实了无论是采用神经网络模型还是基于决策树算法的集成学习模型，运用 Shapley 值分析所获得的结果均表现出高度一致性。进一步地，将 Shapley 值与 UMAP 降维方法及 HDBSCAN 聚类算法相结合，有效地从健康发动机中甄别出潜在风险较高的亚健康发动机群体，并通过 SkopeRules 工具对处于不同健康水平的发动机群体进行规则归纳，明确了飞行周期是区分发动机是否开始进入不健康阶段的关键指标，而飞行等级高低是判断发动机是否维持健康状态或已步入亚健康状态的重要依据。这不仅验证了所提框架在不同预测模型之间具有良好的无关特性，还识别出了发动健康衰退的关键因素，明确了健康衰退过程中的关键转折条件，为实施航空发动机的预防性维护提供了坚实的理论依据。

5. 展望

本文利用 N-CMAPSS 数据集，评估了基于 Shapley 值的可解释人工智能框架在航空发动机健康状态预测中的有效性，揭示了发动机退化机理，并加深了对健康衰退过程的理解，从而为实现经济效益的预测性健康管理提供了支持。尽管如此，Shapley 值计算的高成本限制了其实用性，且在非线性模型中，其解释力相较于线性模型显得更加抽象。因此，未来的研究需聚焦于提高计算效率及开发更直观的非线性特征解释方法，以促进可解释人工智能技术的发展与广泛应用。

基金项目

中电天奥产业发展基金“风电场智慧运维大数据平台”(202201090404)。

参考文献

- [1] 丁小飞, 廖明夫, 韩方军, 等. 某型航空发动机振动值波动故障诊断和排除[J]. 振动测试与诊断, 2024, 44(2): 274-279.
- [2] 栾孝驰, 佟鑫宇, 沙云东, 等. 基于振动与声发射敏感参数识别的主轴承故障诊断方法[J]. 推进技术, 2024, 45(12): 269-281.
- [3] 孙同敏. 基于 DBN-SVM 的航空发动机健康状态评估方法[J]. 控制工程, 2021, 28(6): 1163-1170.
- [4] 张世杰, 胡家文, 苗国磊. 基于工况识别与自训练时空图卷积的航空发动机气路故障诊断[J]. 推进技术, 2024, 45(11): 245-254.
- [5] 张博文, 庞新宇, 程宝安, 等. 基于 PIRD-CNN 的航空发动机轴承故障诊断方法研究[J]. 振动与冲击, 2024, 43(18): 201-207.
- [6] 周志杰, 曹友, 胡昌华, 等. 基于规则的建模方法的可解释性及其发展[J]. 自动化学报, 2021, 47(6): 1201-1216.
- [7] 程晓玉, 韩鹏, 贺维, 等. 一种新的基于可解释性置信规则库的飞轮健康状态评估模型[J]. 航空学报, 2023, 44(S1): 177-189.
- [8] Zhou, Z., Cao, Y., Hu, G., Zhang, Y., Tang, S. and Chen, Y. (2021) New Health-State Assessment Model Based on Belief Rule Base with Interpretability. *Science China Information Sciences*, **64**, 209-223.
<https://doi.org/10.1007/s11432-020-3001-7>
- [9] 王诗彬, 王世傲, 陈雪峰, 等. 可解释性智能监测诊断网络构造及航空发动机整机试车与中介轴承诊断应用[J]. 机械工程学报, 2024, 60(12): 90-106.
- [10] McInnes, L., Healy, J. and Astels, S. (2017) HDBSCAN: Hierarchical Density Based Clustering. *The Journal of Open Source Software*, **2**, Article No. 205. <https://doi.org/10.21105/joss.00205>
- [11] Gardin, F., Gautier, R., Goix, N., et al. (2018) Skope-Rules: Machine Learning with Logical Rules in Python.
<https://github.com/scikit-learn-contrib/skope-rules>
- [12] Lundberg, S.M., Allen, P.G. and Lee, S.-I. (2017) A Unified Approach to Interpreting Model Predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 4768-4777
- [13] McInnes, L., Healy, J. and Melville, J. (2018) UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.
- [14] Sinaga, K.P. and Yang, M. (2020) Unsupervised K-Means Clustering Algorithm. *IEEE Access*, **8**, 80716-80727.
<https://doi.org/10.1109/access.2020.2988796>
- [15] Lin, Y. and Chen, S. (2021) A Centroid Auto-Fused Hierarchical Fuzzy C-Means Clustering. *IEEE Transactions on Fuzzy Systems*, **29**, 2006-2017. <https://doi.org/10.1109/tfuzz.2020.2991306>
- [16] Chao, M., Kulkarni, C., Goebel, K. and Fink, O. (2021) Aircraft Engine Run-to-Failure Dataset under Real Flight Conditions. NASA Ames Prognostics Data Repository.
<https://ti.arc.nasa.gov/tech/dash/groups/pcoc/prognostic-data-repository/#turbofan-2>