

基于机器学习与规则推理的SSD故障预测 方法研究及对比分析

汪子尧¹, 田 瑜², 黄俊杰³, 谭 捷⁴, 杨文婧^{3*}

(1. 北京大学计算机学院, 北京 100871; 2. 军事科学院战略评估咨询中心, 北京 100091;
3. 国防科技大学计算机学院, 湖南长沙 410073; 4. 军事科学院, 北京 100091)

摘要: 随着云计算、大数据及人工智能应用的快速演进, 数据中心规模持续扩张, 存储系统的可靠性已成为影响其稳定运行与服务可用性的关键因素。固态硬盘(Solid-State Drive, SSD)作为数据中心存储系统的关键组成部分, 因其高吞吐、低时延、低功耗等特性被广泛部署于数据中心核心存储层, 但在大规模、长周期运行条件下, SSD故障呈现出突发性强、演化模式复杂等特征, 对业务连续性与数据安全构成严峻挑战。为提高SSD故障预测的准确性与实用性, 本文提出基于分类模型与特征工程的机器学习预测方法, 以及基于显式规则引擎和动态特征补偿的规则推理预测方法。机器学习预测方法通过多阶段特征工程与集成学习, 在数据完备场景下实现了0.968的宏平均 F_1 分数, 但其“黑盒”特性在某种程度上限制了工业应用。规则推理预测方法通过构建多算法融合的显式规则引擎, 并引入基于SHAP (SHapley Additive exPlanations)值的动态特征补偿机制, 在数据完整情况下达到0.988的准确率; 在8个特征缺失的极端场景下仍保持0.941的准确率, 展现出强鲁棒性。实验结果对比分析表明, 机器学习预测方法在数据完备时预测精度高, 规则推理预测方法则在可解释性、实时性与缺失数据适应能力方面更具优势。本文进一步探讨了两类方法的融合路径, 为构建兼具感知能力与推理透明性的下一代智能运维系统提供了理论支撑与实践参考。

关键词: SSD故障预测; 规则推理; 机器学习; 特征工程; 实时预测

基金项目: 国家自然科学基金(No.62402499)

中图分类号: TP311

文献标识码: A

文章编号: 0372-2112(2026)01-0115-10

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250975

Research and Comparative Analysis of SSD Failure Prediction Methods Based on Machine Learning and Rule-Based Reasoning

WANG Ziyao¹, TIAN Yu², HUANG Junjie³, TAN Jie⁴, YANG Wenjing^{3*}

(1. School of Computer Science, Peking University, Beijing 100871, China;

2. Strategic Assessments and Consultation Institute, Academy of Military Sciences, Beijing 100091, China;

3. College of Computer, National University of Defense Technology, Changsha, Hunan 410073, China;

4. Academy of Military Sciences, Beijing 100091, China)

Abstract: With the rapid evolution of cloud computing, big data, and artificial intelligence applications, the scale of data centers continues to expand, and the reliability of storage systems has become a critical factor affecting their stable operation and service availability. As a key component of data center storage systems, solid-state drives (SSDs) are widely deployed in the core storage layers of data centers owing to their advantages of high throughput, low latency, and low power consumption. However, under large-scale and long-term operating conditions, SSD failures are characterized by strong suddenness and complex evolution patterns, posing severe challenges to service continuity and data security. To enhance the accuracy and practicality of failure prediction, this paper investigates a machine learning prediction methodology based on classification models and feature engineering, alongside a rule-based reasoning prediction approach utilizing an explicit rule engine and dynamic feature compensation. The machine learning methodology, through multi-stage feature engineering and ensemble learning, achieves a macro-average F_1 -score of 0.968 under complete data conditions; however, its “black-box” nature somewhat limits its industrial applicability. In contrast, the rule-based reasoning approach constructs an explicit rule engine integrating multiple algorithms and introduces a dynamic feature compensation mechanism based on SHAP (SHapley Additive exPlanations) values. This method attains an accuracy of 0.988 with complete data and maintains an accuracy of 0.941 under extreme conditions with eight missing features, demonstrating strong robustness. Comparative analysis of ex-

perimental results indicates that the machine learning methodology excels in prediction accuracy with complete data, while the rule-based reasoning approach offers superior interpretability, real-time performance, and adaptability to missing data. This paper further explores potential pathways for integrating these two methodologies, providing theoretical support and practical references for constructing next-generation intelligent operation and maintenance systems that possess both perceptual capability and transparent reasoning.

Keywords: SSD failure prediction; rule-based reasoning; machine learning; feature engineering; real-time prediction
Foundation Item(s): National Natural Science Foundation of China (No.62402499)

0 引言

随着云计算、大数据与人工智能技术的飞速发展,全球数据总量呈指数级增长,对数据中心的存储系统在容量、性能及可靠性方面提出了更高的要求^[1-2]。固态硬盘(Solid-State Drive, SSD)凭借其高吞吐、低延迟、强抗振性等优势,已逐步取代传统机械硬盘(Hard Disk Drive, HDD),成为现代数据中心的核心存储介质^[3]。然而,大规模部署下的SSD仍面临因电荷泄漏、擦写损耗、读干扰等物理机制引发的故障风险^[2]。行业报告显示,企业级SSD在5年运行周期内的年化故障率(Annualized Failure Rate, AFR)可高达数个百分比,其突发性故障对业务连续性与数据安全构成了严重威胁^[4]。

为应对这一挑战,工业界长期依赖基于S.M.A.R.T.(Self-Monitoring, Analysis and Reporting Technology)参数的阈值预警机制。该方法虽简单易行,但其静态阈值设定难以适应复杂多变的工作负载与环境,导致误报率居高不下,无法满足关键业务对故障预测精准性的严苛需求^[5-6]。近年来,基于机器学习(Machine Learning, ML)的预测方法为解决该问题提供了新思路。研究者们利用随机森林(Random Forest, RF)、梯度提升决策树(Gradient Boosting Decision Tree, GBDT)等集成学习算法,通过对海量S.M.A.R.T.日志进行训练,显著提升了预测准确率,在公开数据集上 F_1 -score可达90%以上^[1,7]。更进一步,深度学习模型被引入以捕捉故障的时序动态特性。例如,Koh等人^[8]提出了一种融合LSTM(Long Short-Term Memory, LSTM)与Transformer的时序上下文注意力网络(Temporal-Contextual Attention Network, TCAN),通过联合建模S.M.A.R.T.指标的时间演化特征及多属性间的依赖关系,实现了对数据中心SSD故障的高精度预测;Wang等人^[9]进一步提出了一种基于LSTM-xLSTM的SSD故障预测方法,通过引入梯度引导的分区机制,增强了模型对异构盘型差异及关键故障特征的建模能力。然而,这类预测方法存在固有局限:其一,决策过程不透明,导致运维人员难以理解和信任预测结果,阻碍了其在生产环境中的实际采纳^[10];其二,模型性能严重依赖于高质量、高完整度的训练数据,而对实际

运维中因采集中断、固件升级等因素导致的数据缺失异常敏感^[11];其三,复杂的模型结构常带来较高的推理延迟,难以满足对实时性要求极高的场景^[12]。

针对上述机器学习方法面临的挑战,基于规则推理(Rule-Based Reasoning, RBR)的预测方法以及可解释人工智能(eXplainable AI, XAI)技术受到广泛关注。规则推理系统通过显式的逻辑规则进行决策,不仅天然具备可解释性,且在对缺失数据的鲁棒性和推理效率方面展现出独特优势^[13-14]。早期的规则推理系统依赖于领域专家手工编写规则,耗时且难以覆盖所有故障模式。为此,规则学习算法,如C4.5、PART(Partial Decision Tree)、JRIP(Java Repeated Incremental Pruning),被广泛应用于从数据中自动提取精准且可理解的判别规则^[15-17]。Marcílio和Eler^[18]的研究表明,将SHAP(SHapley Additive exPlanations)值作为特征重要性度量基准,能够为规则优先级分配提供量化依据,增强规则推理系统的科学性与客观性。在已有研究中,学界和业界也提出了多种规则系统与机器学习相结合的混合方法。例如,Kwon和Lee^[19]提出了一种自适应交易策略的混合决策支持系统,将深度强化学习算法与规则系统相结合。

尽管已有研究取得了显著进展,但仍面临以下难点:(1)机器学习方法虽精度高但可解释性差,难以满足工业运维对决策透明度的需求;(2)规则推理方法可解释性强,但在复杂模式识别和时序特征利用方面存在局限;(3)现有研究多集中于数据完备的理想场景,对实际部署中常态存在的数据缺失问题关注不足,缺乏系统性的鲁棒性评估。

针对上述问题,本文基于作者在两个阶段的连续性研究工作,对基于机器学习与基于规则推理的SSD故障预测方法进行了深入的研究和对比分析。第1阶段研究聚焦于通过精细化的数据预处理、特征工程与集成学习策略,力求在准确率指标上实现突破。第2阶段研究则转向解决模型的可解释性、鲁棒性与实时部署问题,创新性地提出了一个基于多规则学习算法融合并引入动态特征补偿机制的轻量级预测框架。

本文的主要贡献如下:

(1)提出了一种基于特征选择和增强的集成分类方法,在公开数据集上将SSD故障预测的宏平均 F_1 -

score 提升至 0.968, 显著优于基线方法。

(2) 开发了一套基于 C4.5、PART 和 JRIP 等多规则学习技术融合的可解释预测系统, 通过显式规则引擎有效解决了“黑盒”模型在实际部署中的可信性问题。

(3) 创新性地设计了基于 SHAP 值的动态特征补偿机制, 使系统在面临多达 8 个特征缺失的极端情况下仍能保持 0.941 的预测准确率, 显著提升了系统的鲁棒性。

(4) 通过在同一基准数据集上进行大量对比实验, 系统分析了机器学习与规则推理方法在预测精度、可解释性、缺失数据鲁棒性等方面的性能表现及其权衡关系, 为不同应用场景下的工程实践提供了科学的选择依据。

尽管本文以 SSD 故障预测为研究焦点, 但提出的方法具有良好的通用性, 在架构上具备向内存、传统机械硬盘 (HDD) 等存储设备故障预测扩展的潜力。通过适配不同存储介质的监控指标与特征体系, 以及规则知识迁移, 该方法可为构建统一的异构存储智能管理系统提供有效技术支撑。

1 SSD 故障预测相关工作

SSD 故障预测研究呈现出从传统阈值方法, 到机器学习模型与可解释规则推理系统协同演进的格局。本节系统梳理 SSD 故障预测领域的关键技术发展脉络, 重点分析传统 S.M.A.R.T. 监测体系、基于机器学习的方法以及基于规则推理的方法三大类研究进展。

1.1 传统 S.M.A.R.T. 监测体系及其局限性

早期 SSD 故障预测主要基于 S.M.A.R.T. 属性阈值分析技术。该技术通过监控硬盘运行参数 (如重新分配扇区计数、读写错误率、擦写次数等), 在参数超过预设阈值时触发预警。Schroeder 等人^[4]对大规模部署环境中闪存设备的研究表明, SSD 的故障受磨损、设备年龄、读操作扰动和工艺特征等多种因素共同影响, 使得传统基于单一静态阈值的监测方法难以对故障进行准确判别。

传统方法的根本局限性在于其“事后检测”特性, 依赖静态阈值判断的传统监测机制通常只有在异常信号已经较为明显时才触发告警, 更接近于事后检测, 难以满足故障提前预测的需求。Pinheiro 等人^[5]基于大规模磁盘驱动器运行数据的研究表明, 尽管 S.M.A.R.T. 参数与故障具有较强相关性, 但仅依赖静态阈值判断仍难以实现对磁盘故障的有效预测。此外, SSD 故障通常表现为性能指标随时间逐步劣化并最终超过预设阈值的过程, 即退化具有明显的渐进性特征, 失效判定本质上依赖于阈值越界^[20]。由此可见, 传统基于阈值的监测方法更偏向于对已显现的异常进行

检测, 难以及时捕捉故障发生前的早期退化信号。

近年来, 行业组织尝试对 S.M.A.R.T. 标准进行扩展。开放数据中心委员会 (Open Data Center Committee, ODCC) 在 2025 年发布的《NVMe 子系统故障预测——健康度指标行业标准草案》中提出了覆盖控制器、固件、NAND 介质等全组件的健康度评估体系^[21]。然而, 这些扩展属性在业界支持程度不一, 且缺乏统一的故障映射模型, 限制了其实际应用效果。

1.2 基于机器学习的 SSD 故障预测方法

为克服传统方法的局限性, 研究者开始将机器学习技术应用于 SSD 故障预测。根据技术演进, 这些方法可分为浅层机器学习模型和深度学习模型两个阶段。

1.2.1 基于浅层模型的预测方法

早期研究主要采用 RF、GBDT 等集成学习算法。Xu 等人^[1]利用阿里巴巴提供的 SSD 数据集, 通过特征选择和时序窗口构建, 将故障预测的 F_1 -score 提升至 0.85 以上。研究表明, RF 模型因其对缺失数据的鲁棒性和高维特征处理能力, 在该任务中表现尤为突出^[22]。这些方法的性能严重依赖特征工程质量。同时, 数据不平衡处理成为关键挑战, 过采样和欠采样技术被广泛用于改善模型对少数类 (故障样本) 的识别能力。

1.2.2 基于深度学习的预测方法

随着时间序列数据的丰富, 深度学习模型开始应用于 SSD 故障预测。Koh 等人^[8]提出的时序上下文注意力网络, 将 LSTM 的时序建模能力与 Transformer 的注意力机制相结合, 通过交叉注意力同时刻画 S.M.A.R.T. 指标的时间演化特征与多属性之间的依赖关系。该方法采用特征分组策略, 针对单调累积型指标与随机波动型指标分别建模, 在不依赖强特征筛选的前提下充分利用全部监测属性, 有效提升了对数据中心 SSD 潜在故障的识别精度与模型泛化能力。Wang 等人^[9]针对 SSD 数据中普遍存在的盘型异构、样本极度不平衡以及属性和故障相关性随时间变化等问题, 提出了一种基于 LSTM-xLSTM 的 SSD 故障预测方法。该方法通过按盘型进行模型驱动的数据分区, 缓解不同 SSD 型号间的数据分布差异; 同时引入梯度引导的特征加权机制, 使模型在训练过程中动态强化与故障高度相关的关键属性, 并结合双动态损失函数提升对少数类故障样本的学习能力, 从而显著增强了模型在复杂实际数据中心场景下的预测稳定性与实用性。

然而, 深度学习方法面临模型复杂度高、推理延迟大、可解释性差等挑战。尽管其在预测准确率方面取得了明显提升, 但模型内部决策过程仍难以解释, 对其工程化部署和大规模应用形成一定制约。

1.3 基于规则推理的可解释预测方法

为解决机器学习模型的可解释性问题, 规则推理

方法近年来受到关注。这类方法通过生成显式逻辑规则,使预测过程透明可追溯。

规则学习算法可分为命题规则学习和关系规则学习两大类^[23]。对于 SSD 故障预测这类特征关系相对简单的场景,经典算法如 C4.5^[15]、PART^[16]和 JRIP^[17]被广泛应用于从数据中提取判别规则。研究表明,规则推理系统在噪声数据和缺失值场景下表现出较强鲁棒性^[24]。Marcílio 和 Eler^[18]将 SHAP 值作为特征重要性度量基准,为规则优先级分配提供了量化依据。然而,传统规则推理系统面临规则集膨胀、计算复杂度高等挑战,特别是在需要实时响应的工业场景中^[25]。

近年来,工业界开始探索规则推理系统与机器学习模型的融合。Kwon 和 Lee^[19]提出了一种自适应交易策略的混合决策支持系统,将深度强化学习算法与规则系统相结合提高了决策能力。杨鑫文等人^[26]则通过分析写入特征波动与热量传递的关联度,构建了基于物理规则的 SSD 故障预测系统。

值得注意的是,规则推理系统在实际部署中需要解决规则更新与概念漂移问题。Tabebordbar 等人^[27]提出一种自主的数据标注规则自适应方法,可根据随时间从动态数据标注环境中收集的反馈调整规则,增强了系统在复杂场景下的数据标注质量与可持续性。

2 基于分类模型与特征工程的预测方法

为实现 SSD 故障的高精度预测,本节系统性地应用了机器学习技术路线,设计了一种基于机器学习的 SSD 故障预测方法。其核心思想在于通过多层次特征工程挖掘 S.M.A.R.T. 数据中的潜在故障信息,并利用集成学习模型构建高性能分类器。

2.1 方法框架

该方法遵循典型的机器学习流程,如算法 1 所示,主要包括数据预处理、特征工程、模型训练与评估 4 个阶段。其目标是从海量 S.M.A.R.T. 监控数据中

算法 1 SSD 故障预测方法整体流程

1. 构建实验数据集
2. 数据预处理阶段
3. 处理缺失值(如丢弃或填充)
4. 处理异常值
5. 特征工程阶段
6. 特征选择
7. 特征构造
8. 模型选择阶段(如 RF、DT 或 GBDT)
9. REPEAT
10. 模型训练阶段
11. 模型评估阶段
12. 超参数调优
13. UNTIL 模型性能达到预期
14. 预测阶段

构建高精度的故障分类器。

2.2 实验设置与基准建立

实验采用阿里巴巴数据中心公开的 SSD S.M.A.R.T. 数据集,涵盖超过 50 万个 SSD 设备。原始数据经过严格的预处理流程,包括高缺失率特征过滤(缺失率>82%)、基于分位数的异常值处理,以及针对严重样本不平衡(正常:故障≈96:1)的随机欠采样。为建立性能基准,本研究首先以决策树(Decision Tree, DT)模型结合 Wikipedia 推荐的关键 S.M.A.R.T. 属性进行预测,其结果作为基线(如表 1 所示),6 款不同型号 SSD 的宏平均 F_1 -score 介于 0.474~0.711 之间,显示出传统先验知识在复杂预测任务中的局限性。

表 1 基线模型的宏平均 F_1 -score

Table 1 Macro-average F_1 -score of baseline models

SSD 模型	MA1	MA2	MB1	MB2	MC1	MC2
基线	0.474	0.499	0.482	0.517	0.711	0.646

2.3 特征工程与模型选择的系统性对比

为提升性能,本研究系统对比了多种特征选择策略与机器学习模型。特征选择方面,对比了保留所有特征、基于 Pearson 相关系数、基于 Spearman 相关系数以及依赖 Wikipedia 先验知识 4 种方法。模型方面,选取了 DT、RF 与 GBDT 这 3 种代表性集成学习算法。

实验结果表明,特征选择策略对预测性能具有决定性影响。如表 2 所示,采用 RF 模型并保留所有预处理后的特征时,所有 SSD 型号的宏平均 F_1 -score 均获得大幅提升,相较于基线提升幅度介于 37%~103% 之间。相比之下,严格依赖 Wikipedia 关键特征的策略性能最差,证实了在高质量数据基础上,数据驱动式特征利用远优于静态先验知识。

表 2 不同特征选择方法下最优模型的宏平均 F_1 -score 对比

Table 2 Comparison of macro-average F_1 -scores of optimal models with different feature selection methods

SSD 模型	保留所有特征 (RF)	Pearson 相关系数(RF)	Wikipedia 关键特征(DT)
MA1	0.964	0.952	0.474
MA2	0.980	0.968	0.499
MB1	0.965	0.954	0.482
MB2	0.979	0.966	0.517
MC1	0.971	0.960	0.711
MC2	0.931	0.920	0.646

在模型选择上,RF 模型展现出综合优势。在相同特征输入下,RF 模型的宏平均 F_1 -score 达到 0.968,显著优于 GBDT 和 DT 模型。深入分析发现,RF 虽然 F_1 -score 最高,但其受试者工作特征曲线(Receiver Operating Characteristic curve, ROC)下面积(Area Under

the ROC Curve, AUC)为0.65,表现一般,表明模型在极力捕捉故障样本的同时,也产生了较多的误报,这揭示了高召回率与低误报率之间的固有矛盾。

2.4 关键因素影响分析

本研究进一步探讨了数据采样方法与特征增强技术的影响。数据采样方面,采用随机采样划分训练集与测试集(7:3)的性能,显著优于基于时间顺序的划分(用前5个月数据预测第6个月),后者因时间概念漂移问题导致所有模型性能下降0.2以上,凸显了离线评估中数据划分策略的重要性。特征增强方面,针对SSD故障的时序特性,本研究构造了基于5天时间窗口的统计特征(均值、标准差等)。然而,在已包含54个高信息密度特征的场景下,特征增强并未带来预期中的性能提升,甚至因维度灾难导致部分模型性能轻微下降,这表明特征工程的价值存在上限,并非特征越多越好。

综上所述,基于机器学习的预测方法在数据完备、特征工程精细的条件下,能够实现极高的预测精度(F_1 -score达0.968)。然而,其性能严重依赖于高质量的数据预处理与特征选择,且模型本身作为“黑盒”,存在可解释性差、对数据完整性敏感的内在局限,为其工业部署带来了挑战。

3 基于规则推理的可解释预测方法

为解决机器学习模型的可解释性、实时性与数据缺失鲁棒性等工业部署难题,本节提出一种基于规则推理的可解释故障实时预测系统。本节工作不追求复现机器学习模型的极致精度,而是致力于通过显式的逻辑规则构建一个决策过程透明、在复杂现实场景下稳定可靠的轻量级实时预测框架。

3.1 系统架构

本研究提出的SSD故障实时预测方法由4个核心模块构成:多源数据预处理模块负责原始数据的规范化处理,混合规则集构建模块通过集成学习生成判别性规则,自适应特征重要性规则触发模块解决实际应用中的特征缺失问题,可视化决策支持系统则实现预测过程的可解释交互。各模块通过特征传递与反馈机制形成闭环优化体系,其中数据流经预处理后进入规则学习阶段,最终通过动态阈值调整实现实时预警。该系统采用规则引擎替代传统“黑盒”模型,实现从数据输入到故障预警的端到端透明推理。

3.2 核心创新与系统特色

本系统的特色集中于两点:一是构建了多算法融合的显式规则引擎,二是设计了基于SHAP值的动态特征补偿机制。

规则引擎通过集成C4.5^[15]、PART^[16]和JRIIP^[17]这

3种规则学习算法,从数据中自动提取判别规则。不同于单一算法,多算法融合策略有效克服了单一算法的模式捕获盲区,形成了覆盖范围更广、判别特性互补的混合规则集。通过施加覆盖率(0.006)、精度(0.9)与复杂度(<15)三重约束,最终从616条候选规则中遴选出312条核心规则,奠定了系统高精度与高可解释性的基础。

动态特征补偿机制是本系统应对现实数据缺失问题的核心创新。该机制利用SHAP值精确量化每个特征对规则判别力的贡献度。当发生特征缺失时,系统并非简单地将规则失效,而是计算缺失特征的归一化重要性指标 $\hat{I}_m = I_m^{\text{SHAP}}/\bar{I}_m$,并通过与动态阈值 γ 的比较,智能决策是否触发该规则。这一机制使系统具备了“优雅降级(graceful degradation)”的能力,而非“灾难性失效(catastrophic failure)”。

3.3 性能评估与特色分析

在数据完备场景下,系统在测试集上取得了0.988的准确率与0.974的精确度,证明了规则引擎本身具备强大的故障识别能力,且其决策过程完全透明,可追溯至具体规则条件。

在规则集规模影响方面,实验表明性能与规则数量呈非线性增长。如表3所示,当规则从80条扩展至300条时,召回率从54.20%大幅提升至87.64%,准确率升至98.77%,而推理时延仅从276 ms增至486 ms,仍满足工业实时性要求(<500 ms),体现了系统在效率与性能间的良好平衡。

表3 基于规则集扩展的故障预测性能优化分析

Table 3 Performance optimization analysis of failure prediction based on rule set expansion

规则集规模	准确率	精确度	召回率	时延/ms
80	0.939 5	0.970 6	0.542 0	276
100	0.945 8	0.971 4	0.590 6	295
120	0.954 2	0.972 6	0.629 4	314
140	0.961 7	0.972 5	0.692 1	333
160	0.969 3	0.972 9	0.732 0	352
180	0.974 6	0.973 1	0.759 3	371
200	0.980 1	0.973 6	0.794 6	390
220	0.983 5	0.973 6	0.824 8	409
240	0.985 9	0.973 5	0.844 7	428
260	0.987 2	0.973 5	0.861 9	447
280	0.987 5	0.973 7	0.870 9	466
300	0.987 7	0.973 9	0.876 4	486

在缺失数据鲁棒性方面,本系统表现出显著优势。实验结果表明,如图1及表4所示,即使在多达8个特征缺失的极端情况下,通过将动态补偿阈值 γ 设定为0.9,系统仍能保持0.941的准确率。系统性能随缺

失特征数增加呈可控的渐进式衰减,而非断崖式退化,这验证了所提出的动态补偿机制在缓解数据不完

整性影响方面的有效性,同时克服了机器学习模型对数据完整性高度敏感的局面。

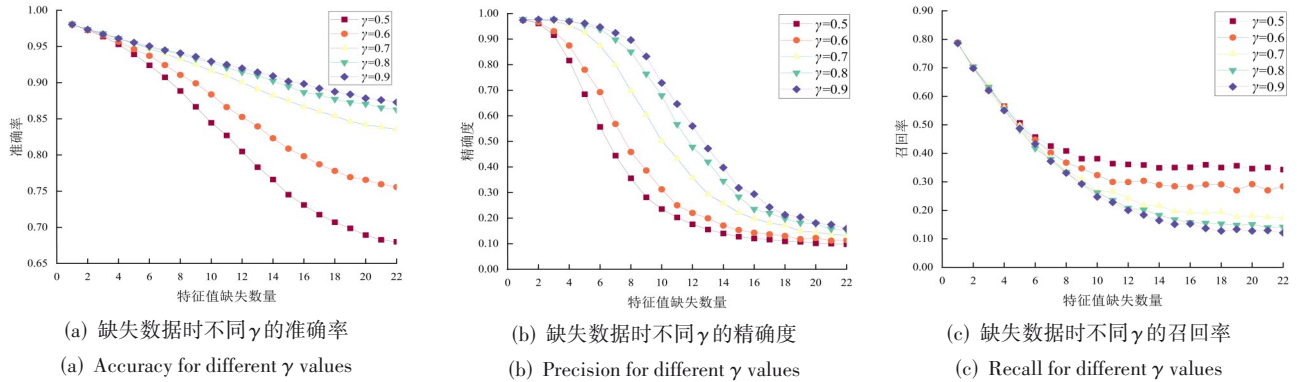


图1 缺失数据时不同 γ 的准确率、精确度、召回率

Figure 1 Accuracy, precision, and recall under different γ values with missing data

表4 规则推理系统在不同缺失特征数下的性能($\gamma=0.9$)

Table 4 Performance of the rule-based reasoning system with different numbers of missing features ($\gamma=0.9$)

缺失特征数	准确率	精确度	召回率
0	0.988	0.974	0.876
4	0.961	0.969	0.550
8	0.941	0.896	0.331

为进一步探究动态特征补偿机制中动态阈值 γ 的系统性影响,我们在缺失特征数从0~22的范围内,对 $\gamma \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$ 进行了系统性实验。如图1所示,系统在各 γ 值下均呈现平稳的性能衰减规律,进一步体现了机制的稳健性。具体而言,在缺失特征数较少(0~4)时, γ 对准确率与召回率影响较小;当 $\gamma \geq 0.7$ 时,系统整体性能趋于稳定。尤其在 $\gamma \geq 0.8$ 的情况下,不同缺失规模下的性能曲线表现出明显的收敛特性,说明该机制在较高阈值下具备良好的自适应与容错能力。表4进一步给出了 $\gamma=0.9$ 时,系统在不同缺失特征数下的准确率、精确度与召回率。结果显示,即使面临4个或8个特征缺失,系统准确率仍保持在0.94以上,精确度波动范围较小(0.896~0.974),表明规则推理系统在面对不完整输入时仍能保持稳定的判断能力。这一特性体现了动态补偿机制在较高 γ 设置下所具有的“可控性能降级”属性,能够有效支撑工业环境中对故障预测稳定性与一致性的严格要求。

在系统实际运行过程中,单条SSD监控记录可能同时触发多条规则,从而引发预测冲突。为确保输出决策的一致性与可靠性,本系统设计了一套基于优先级与证据加权的规则冲突消解机制,其核心包括以下2个层次:

(1)静态优先级初始化。每条规则根据其在独立验证集上的分类置信度(即精度)及其历史触发频

率,被赋予一个基础优先级。置信度越高、触发记录越稳定的规则,所获优先级也越高。

(2)动态加权投票判决。当多条规则同时被触发时,系统不仅依据静态优先级,还综合考虑各规则的实时置信度及其与当前实例的条件匹配程度(例如特征值与规则条件区间之间的相对距离),进行多规则加权投票。

最终,系统选取综合权重最高的规则作为故障判定输出。该策略不仅有效解决了多规则触发时的决策冲突问题,还通过引入动态证据融合机制,提升了对边界样本和复杂情形的判别能力,从而在保证系统决策一致性的同时,增强了结果的可解释性与可靠性。

综上所述,本节所构建的规则推理系统,凭借其内在的可解释性、确定的实时推理性能以及对缺失数据具备的强鲁棒性,形成了对机器学习方法的重要补充。该系统为在要求高安全性与决策透明度的工业场景中部署SSD故障预测系统,提供了一条可靠且实用的技术路径。

4 对比分析与讨论

4.1 综合性能对比分析

我们从预测准确性、可解释性、鲁棒性、实时性、资源消耗、场景适应性等6个维度对两种方法进行了系统对比(如表5所示)。本次对比基于同一基准数据集展开,确保了评估的一致性与公正性。

从预测准确性角度看,在数据完备场景下,基于机器学习的预测方法在 F_1 -score上略胜一筹(分别为0.968和0.922),这主要得益于其强大的非线性拟合能力与特征交互感知。然而,在准确率这一关键运维指标上,基于规则推理的预测方法表现更优(分别为0.988和0.968)。这反映出两类方法在优化目标上的

根本差异:机器学习方法追求召回率与精确度的平衡,而规则推理方法因其显式逻辑结构更利于实现低误报率,更贴近工业场景中对“宁可漏报、不可误报”的可靠性要求。

表5 两种方法的综合性能对比

Table 5 Comprehensive performance comparison of the two methods

评估维度	基于机器学习的预测方法	基于规则推理的预测方法	优势方法	说明
F_1 -score(数据完备)	0.968	0.922	机器学习	机器学习在综合分类指标上表现更优,适合追求高检测率的场景。
准确率(数据完备)	0.968	0.988	规则推理	规则推理在整体分类正确率上更高,利于控制误报。
可解释性	差(黑盒模型)	优(白盒规则)	规则推理	规则推理的每个决策可追溯至明确的逻辑条件,适合高可信场景。
缺失数据鲁棒性(8特征缺失下准确率)	<0.85(下降明显)	0.941(保持稳定)	规则推理	规则推理通过动态补偿机制实现优雅降级,对数据不完整更具容忍性。
实时推理延迟	较高(通常>1 s)	低(<500 ms)	规则推理	规则引擎执行效率高,满足工业实时性要求。
模型/系统复杂度	高(尤其深度学习)	中低(规则集+轻量引擎)	规则推理	规则系统结构简单,易于理解与维护。
资源消耗(内存与计算)	高	低	规则推理	适合边缘部署与资源受限环境。
扩展维护成本	低(自动学习)	中高(需规则维护与更新)	机器学习	机器学习可通过数据自动迭代,规则系统需人工介入维护。
典型适用场景	数据完备、离线分析、高精度预测、云平台	数据缺失常见、实时预警、高可解释性、边缘设备		二者形成互补,实际部署可结合使用。

注:表中加黑字体表示更优方法。

从可解释性角度看,基于规则推理的预测方法具有显著优势。其决策过程完全透明,每个预测结果都可以追溯到具体的规则条件(如“IF 属性A>阈值X AND 属性B<阈值Y THEN 故障”),极大地增强了运维人员对系统的信任。相比之下,机器学习方法特别是集成模型与深度学习,其决策过程难以直观解释,属于典型的“黑盒”模型,在金融、医疗等高可信要求场景中部署受限。

从鲁棒性角度看,在数据缺失场景下,基于规则推理的预测方法展现出明显更强的适应能力。本实验模拟了特征缺失比例从0~20%的多种情况,结果显示,当缺失特征数达8个(约占总特征15%)时,规则推理方法仍能保持94.1%的准确率,而机器学习方法在同等缺失情况下准确率下降超过15%。这得益于规则推理方法所提出的动态特征补偿机制,使系统能够根据特征重要性实现“优雅降级”,而非“灾难性失效”。

从实时性表现看,规则推理方法在推理效率上显著占优。在相同硬件环境下,其对单条样本的推理延迟稳定控制在500 ms以内,满足工业实时预警的典型要求(<1 s)。相比之下,机器学习方法,尤其是深度学习模型,受模型复杂度影响,往往具有更高的推理时延,在边缘或强实时场景中面临一定挑战。

从资源消耗角度看,规则推理系统模型轻量,内存与计算开销小,更适合部署在资源受限的边缘设备。机器学习模型(尤其是深度网络)则对计算资源与存储

空间要求较高,通常需部署在云端或高性能服务器。

从场景适应性看,机器学习方法在数据完备、模式复杂的场景中具有更强感知能力;而规则推理方法在数据缺失频繁、决策透明性要求高、实时响应迫切的工业现场中更具实用价值。

4.2 方法特性分析

在技术特性层面,基于机器学习与基于规则推理的预测方法展现出截然不同的优势特征。

基于机器学习的预测方法核心优势在于其强大的自适应学习能力,该方法能够从海量监控数据中自动发掘复杂的故障模式,无需依赖显式的规则编程,显著降低了领域知识依赖。在处理高维特征方面,基于机器学习的预测方法表现出卓越的包容性,能够有效捕捉多个S.M.A.R.T.属性之间微妙的非线性关系,这对于识别复合型故障模式尤为重要。此外,数据驱动的特性使机器学习预测方法具备持续优化能力,随着运维数据的不断积累,模型性能可得到逐步提升。在泛化能力方面,经过充分训练的机器学习模型对新型号SSD具有良好的适应能力,能够快速部署到异构存储环境中。

基于规则推理的预测方法的独特价值体现在其内在的可解释性机制。系统的每个预测决策都基于明确的逻辑规则,使得运维人员能够清晰地理解故障判断的依据,这种透明性极大地增强了系统在实际生产环境中的可信度。在异常诊断支持方面,基于规则推

理的预测方法能够提供完整的推理链条,为故障根因分析提供直接依据,显著缩短了问题定位时间。基于规则推理的预测方法另一重要优势在于其良好的领域知识融合能力,便于将专家经验以规则形式固化到系统中,形成数据驱动与知识引导的双重保障。在系统稳定性方面,规则行为具有确定性和可预测性,不会出现机器学习模型可能存在的随机波动,确保了运维决策的一致性。

4.3 应用场景匹配分析

基于系统性对比实验结果,本研究针对不同应用场景提出具体的方法选择建议。在数据质量高、完整性好的离线分析环境中,推荐采用基于机器学习的预测方法,其能够充分利用完备数据中的深层特征,实现故障预测精度的最大化。对于需要发现新型故障模式的探索性研究场景,基于机器学习的预测方法数据驱动特性使其具备更强的模式发现能力,有助于识别尚未被专家经验覆盖的故障前兆。在具备充足计算资源的云平台环境中,基于机器学习的预测方法能够发挥其模型复杂度的优势,通过深度特征提取实现精准预测。在对模型准确性要求极高且可接受“黑盒”特性的特定场景中,基于机器学习的预测方法提供的性能优势往往具有决定性意义。

基于规则推理的预测方法在对决策可解释性要求严格的生产环境中展现出独特价值。其透明的决策过程便于运维人员理解和验证,特别适合金融、医疗等对系统可靠性要求极高的领域。在数据质量不稳定、存在部分缺失的实际工业现场,基于规则推理的预测方法凭借其动态特征补偿机制,能够保持稳定的预测性能,避免因数据不完整导致的系统性能急剧下降。对于需要快速根因分析的运维场景,基于规则推理的预测方法提供的明确推理路径能够直接指导维护决策,显著提升故障处理效率。在计算资源受限的边缘部署环境中,基于规则推理的预测方法的轻量级特性使其成为理想选择,能够在满足实时性要求的同时保证预测可靠性。

针对大型数据中心的复杂需求,建议采用分层混合部署策略。在边缘节点部署轻量级规则推理系统,实现实时预警和快速响应;在中心平台部署高性能机器学习模型,进行深度数据分析和模型优化。这种架构既保证了系统在边缘侧的实时性和鲁棒性,又充分利用了中心平台的计算资源实现精准预测。通过建立规则推理与机器学习模型之间的协同机制,可以实现故障预警与根因分析的无缝衔接,构建完整的智能运维解决方案。此外,该混合架构还支持渐进式升级,便于根据业务发展需求灵活调整系统配置。

4.4 从AI范式融合的视角审视与展望

前述对比中2种方法所呈现出的性能与特性差异,其根源可追溯至人工智能领域两大学术流派——连接主义与符号主义的技术分野。本文的工作,在SSD故障预测这一具体工业场景下,为这两大范式的能力边界与互补性提供了清晰的实证。

基于机器学习的预测方法是连接主义的当代典范。其通过调整模型内部海量参数(如神经网络权重),以“哑符号”的数值计算方式从数据中分布式地学习复杂映射函数。这一范式赋予了其强大的感知与模式识别能力,使之能够发现并拟合出人脑难以预设的复杂故障前兆模式,从而在数据完备时达到0.968的 F_1 -score。然而,其“黑盒”特性、对数据完整性的高度依赖以及较高的计算复杂度,也正是连接主义范式在追求端到端学习过程中所付出的固有代价。

基于规则推理的预测方法则深深植根于符号主义传统。它将知识表示为“IF-THEN”形式的显式逻辑符号,并通过符号操作进行推理。这使其决策过程天然具备可解释性与可追溯性,运维人员可以像阅读说明书一样理解系统的决策逻辑。其对数据缺失的强鲁棒性(在8个特征缺失下准确率仍达0.941)源于其符号处理能力:当部分信息缺失时,系统仍可基于现有符号和逻辑进行部分推理,而非像连接主义模型那样因输入向量的不完整而性能急剧退化。

认识到这一本质,解决工业级SSD智能运维复杂需求的最有前景之路,并非在两种范式或方法中做“二选一”的取舍,而在于推动它们的深度有机融合,构建神经-符号集成系统(Neuro-Symbolic Integrated Systems)。基于本研究的实证发现,我们展望以下3条融合路径:

(1)符号引导的连接主义(Symbol-Guided Connectionism)。利用规则推理系统产生的可解释符号知识,来引导和约束机器学习模型。例如,可将高重要性规则筛选出的关键特征子集作为模型的优先输入,降低维度灾难与过拟合风险;或将核心规则作为领域知识注入模型训练过程,通过设计特定的损失函数或正则化项,使模型的输出更符合物理意义与专家经验,提升其泛化能力与可靠性。

(2)连接主义赋能的符号主义(Connectionism-Enabled Symbolism)。利用机器学习强大的表示学习与模式发现能力,来增强和扩展符号系统。例如,使用深度学习模型自动从海量原始数据中识别和定义新的、更精细的故障“符号”或状态概念,为规则推理系统的知识库提供持续、自动的知识原料,克服传统符号系统在知识获取上的瓶颈。

(3)分层协同的混合智能架构(Hierarchical Col-

laborative Hybrid Intelligence Architecture)。在系统工程层面,设计云边协同的混合架构。在边缘侧部署轻量级、低延迟的规则推理引擎(符号主义),处理大部分已知、确定的故障模式,确保实时响应与决策透明。在云端部署高性能的机器学习模型(连接主义),负责对边缘不确定案例进行深度分析,挖掘潜在的新型故障模式并持续优化与更新边缘侧的规则库。通过智能路由机制,实现两种范式在任务层面的高效协同。

综上所述,通过推动符号主义的可解释逻辑推理与连接主义的强大数值感知能力在SSD故障预测领域深度融合,我们有望最终构建出既强大又可信、既精准又稳健的下一代智能运维系统,为实现真正可靠、可用的工业人工智能奠定坚实的技术基石。

5 结束语

本文系统研究并对比了基于机器学习与规则推理的两类SSD故障预测方法。实验表明,机器学习方法在数据完备时具有更高的综合预测精度(宏平均 F_1 -score 达 0.968),而规则推理方法在可解释性、实时性以及鲁棒性方面表现显著更优,其提出的动态特征补偿机制在 8 个特征缺失下仍能保持 94.1% 的准确率。研究从神经-符号融合的视角,为两类方法的优势互补与系统集成提供了实证依据与可行路径。

未来工作将聚焦于:(1)神经-符号融合机制的深化与轻量化实现;(2)面向时序退化过程的动态规则演化;(3)跨设备、跨场景的规则迁移与联邦学习架构。通过推动感知与推理能力的深度融合,构建更健壮、可解释的下一代智能运维系统。

参考文献

- [1] Xu Fan, Han Shujie, Lee P P C, et al. General feature selection for failure prediction in Large-scale SSD deployment[C/OL]//2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), 2021: 263-270. <https://ieeexplore.ieee.org/document/9505157>.
- [2] Botezatu M M, Giurgiu I, Bogojeska J, et al. Predicting disk replacement towards reliable data centers[C/OL]//The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016: 39-48. <https://dl.acm.org/doi/10.1145/2939672.2939699>.
- [3] Meza J, Wu Q, Kumar S, et al. A large-scale study of flash memory failures in the field[J]. ACM Sigmetrics Performance Evaluation Review, 2015, 43(1): 177-190.
- [4] Schroeder B, Lagisetty R, Merchant A. Flash reliability in production: the expected and the unexpected[C/OL]//The 14th usenix conference on file and storage technologies, 2016: 67-80. <https://dl.acm.org/doi/10.5555/2930583.2930589>.
- [5] Pinheiro E, Weber W D, Barroso L A. Failure trends in a large disk drive population[C/OL]//Proceedings of the 5th USENIX conference on File and Storage Technologies, 2007: 2. <https://api.semanticscholar.org/CorpusID:2420428>.
- [6] You Wenyan, Dong Jiayuan, Feng Xingdi, et al. SSD failures in Large-scale data centers: what why and how [C/OL]//2024 International Conference on Networking, Architecture and Storage (NAS). 2024: 1-8. <https://ieeexplore.ieee.org/abstract/document/10781343>.
- [7] Li Jing, Ji Xinpu, Jia Yuhan, et al. Hard drive failure prediction using classification and regression trees[C]//2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 2014: 383-394. <https://ieeexplore.ieee.org/document/6903596>.
- [8] Koh C, Kang J S, Kim T, et al. Temporal-contextual attention network for solid-state drive failure prediction in data centers[J]. IEEE Access, 2024, 12: 154455-154466.
- [9] Wang Xiaofei, Zhang Yang, Chen Junyan, et al. Proactive SSD failure prediction with a gradient-guided LSTM-xLSTM hybrid model[C]//2025 IEEE International Conference on Cluster Computing. Piscataway: IEEE, 2025: 11186457.
- [10] Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead[J]. Nature Machine Intelligence, 2019, 1(5): 206-215.
- [11] Abu-Soud S M. A novel approach for dealing with missing values in machine learning datasets with discrete values[C]//2019 International Conference on Computer and Information Sciences. Piscataway: IEEE, 2019: 8716430.
- [12] Menghani G. Efficient deep learning: A survey on making deep learning models smaller, faster, and better[J]. ACM Computing Surveys, 2023, 55(12): 1-37.
- [13] van der Waa J, Schoonderwoerd T, van Diggelen J, et al. Interpretable confidence measures for decision support systems[J]. International Journal of Human-Computer Studies, 2020, 144: 102493.
- [14] Góra G, Skowron A, Wojna A. Explainability in RIONA algorithm combining rule induction and instance-based learning[C/OL]//2023 18th Conference on Computer Science and Intelligence Systems. 2023: 491-502. <https://ieeexplore.ieee.org/document/10305962>.
- [15] Quinlan J R. C4.5: Programs for machine learning[M]. CA, USA: Morgan Kaufmann Publishers Inc, 1993: 236-239
- [16] Frank E, Witten I H. Generating accurate rule sets without global optimization[C]//Proceedings of the Fifteenth International Conference on Machine Learning. New

- York: ACM, 1998: 144-151.
- [17] Govada A, Thomas V S, Samal I, et al. Distributed multi-class rule based classification using RIPPER[C]//2016 IEEE International Conference on Computer and Information Technology. Piscataway: IEEE, 2016: 303-309.
- [18] Marcílio W E, Eler D M. From explanations to feature selection: Assessing SHAP values as feature selection mechanism[C]//2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images. Piscataway: IEEE, 2020: 340-347.
- [19] Kwon Y, Lee Z. A hybrid decision support system for adaptive trading strategies: Combining a rule-based expert system with a deep reinforcement learning strategy[J]. Decision Support Systems, 2024, 177: 114100.
- [20] Li Peng, Liu Kai, Dang Wei, et al. Reliability assessment of NAND SSD based on acceleration degradation test[C]//2017 IEEE International Conference on Industrial Engineering and Engineering Management. Piscataway: IEEE, 2017: 1945-1949.
- [21] ODCC2505002 NVMe 子系统故障预测——健康度指标行业标准草案[S].
- [22] Breiman L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.
- [23] Fürnkranz J, Gamberger D, Lavrač N. Foundations of rule learning (1st ed)[M]. Berlin: Springer, 2012.
- [24] Cohen W W. Fast effective rule induction[M]//Machine Learning Proceedings 1995. Amsterdam: Elsevier, 1995: 115-123.
- [25] Kwasny S C, Faisal K A. Overcoming limitations of rule-based systems: An example of a hybrid deterministic parser[C]//Konnektionismus in Artificial Intelligence und Kognitionforschung. Berlin, Heidelberg: Springer, 1990: 48-57.
- [26] 杨鑫文, 麦钰岚, 郭巧玉. 一种固态硬盘的寿命预测方法及系统: CN202410517838.1[P]. 2024-05-28.
- [27] Tabebordbar A, Beheshti A, Benatallah B, et al. Feature-based and adaptive rule adaptation in dynamic environments[J]. Data Science and Engineering, 2020, 5(3): 207-223.

作者简介



汪子尧 男, 2000年3月出生于北京市。现为北京大学计算机学院博士研究生。主要研究方向为人工智能。
E-mail: ziyawang@stu.pku.edu.cn



谭捷 女, 1991年12月出生于河北省石家庄市。现为军事科学院副研究员。主要研究方向为人工智能和软件工程。
E-mail: j.tanjie@outlook.com



田瑜 女, 1992年5月出生于江苏省无锡市。现为军事科学院战略评估咨询中心助理研究员。主要研究方向为计算机视觉、大语言模型。
E-mail: tianyu10@alumni.nudt.edu.cn



杨文婧 女, 1988年10月出生于湖南省长沙市。现为国防科技大学计算机学院研究员、博士生导师。主要研究方向为以数据为中心的具身智能、空间计算。
E-mail: wenjing.yang@nudt.edu.cn



黄俊杰 男, 1990年9月出生于湖南省怀化市。国防科技大学计算机学院副研究员。主要研究方向为多模态学习、模型驱动深度学习、视觉增强与感知等领域。
E-mail: jjhuang@nudt.edu.cn