

# 基于大数据的铁路电力机车运行故障预测及改进策略

申 震

国家铁路局装备技术中心 北京 100070

**摘要:** 文章探讨了基于大数据的铁路电力机车运行故障预测及改进策略。通过收集和分析电力机车运行过程中的海量数据,利用先进的算法和模型对故障进行精准预测。在此基础上,提出预防性维护、故障诊断与排除以及运行优化等改进策略,旨在降低电力机车的故障率,提升运行效率和安全性。实践表明,这些策略能够显著降低维修成本,延长机车使用寿命,为铁路运输的可持续发展提供有力支持。

**关键词:** 大数据; 铁路电力机车; 运行故障; 改进策略

## 1 铁路电力机车运行故障概述

铁路电力机车作为现代铁路运输的主力军,其稳定运行对于保障交通动脉的畅通至关重要。在运行过程中,电力机车可能会遭遇多种故障。这些故障主要包括电气系统故障、机械部件损坏、控制系统失灵以及信号接收问题等。电气系统故障可能导致机车牵引力下降或完全丧失,影响运行速度和安全性。机械部件如转向架、轮对等的损坏,则可能引发脱轨等严重事故。控制系统失灵会使机车无法准确执行调度指令,造成运行混乱。而信号接收问题则可能导致机车误判前方路况,发生碰撞等危险。

## 2 基于大数据的铁路电力机车运行故障预测模型

### 2.1 数据采集与处理

在铁路电力机车运行故障预测模型中,数据采集与处理是基础且关键的一步。数据采集主要来源于电力机车上的各类传感器、监控设备和运行日志。这些数据源涵盖了机车的电气系统、机械部件、控制系统以及信号接收等多个方面。传感器数据包括电流、电压、温度、湿度等实时参数,而监控设备和运行日志则记录了机车的运行状态、维护记录和故障历史等关键信息。数据采集完成后,需要进行数据预处理,数据预处理包括数据清洗、数据整合和数据变换等环节。数据清洗主要是去除数据中的噪声、异常值和缺失值,确保数据的准确性和完整性。数据整合则是将来自不同数据源的数据进行合并,形成统一的数据集<sup>[1]</sup>。数据变换则是对数据进行标准化和归一化处理,消除不同量纲对预测结果的影响,提高模型的泛化能力。在数据采集与处理过程中,还需要考虑数据的安全性和隐私保护。对敏感数据进行加密和匿名化处理,设置严格的访问权限,确保只有授权用户才能访问数据,防止数据泄露和非法访问。

### 2.2 特征选择与提取

特征选择与提取是基于大数据的铁路电力机车运行故障预测模型中的核心步骤。特征选择是从大量的原始特征中筛选出最具有代表性、与故障最相关的特征。这可以通过相关性分析、主成分分析(PCA)和特征选择算法等方法实现。相关性分析可以找出与故障发生最相关的特征,主成分分析则可以对高维数据进行降维,减少冗余信息,提高模型运行效率。特征选择算法如递归特征消除(RFE)和基于树的特征选择等,可以根据特征的重要性进行排序,选择出最优的特征子集。特征提取则是从原始数据中提取出具有区分度的特征,这些特征能够更好地描述电力机车的运行状态和故障模式。常用的特征提取方法包括时域特征提取、频域特征提取和深度学习特征提取等。时域特征提取可以提取出如均值、方差、峰度等统计特征,频域特征提取则可以提取出如频谱、功率谱等频域特征。深度学习特征提取则利用神经网络等深度学习算法,从原始数据中自动学习出高层次的特征表示。在特征选择与提取过程中,需要注意避免过拟合和欠拟合的问题。过拟合是指模型在训练数据上表现良好,但在测试数据上表现较差,这通常是由于选择了过多的特征或过于复杂的模型导致的。欠拟合则是指模型在训练数据和测试数据上表现都不好,这通常是由于选择了过少的特征或过于简单的模型导致的。

### 2.3 模型建立与训练

在基于大数据的铁路电力机车运行故障预测模型中,模型建立与训练是关键步骤之一。决策树是一种基于树形结构的分类算法,通过递归地划分数据集,构建出能够预测故障的分类树。支持向量机则是一种基于最大间隔分类的二分类算法,可以扩展到多分类问题,通过找到能够将不同类别的数据点分开的最优超平面,实现故障的分类预测。随机森林则是一种基于多个决策树的集成学习方法,通过组合多个决策树的预测结果,提

高模型的稳定性和准确性。神经网络则是一种基于神经元和连接权重的非线性模型，通过训练数据调整连接权重，实现故障的分类和预测。除了机器学习算法外，深度学习算法也在铁路电力机车运行故障预测中得到了广泛应用，深度学习算法如卷积神经网络（CNN）、循环神经网络（RNN）和长短期记忆网络（LSTM）等，可以通过多层非线性变换，从原始数据中自动学习出高层次的特征表示，实现更准确的故障预测<sup>[2]</sup>。在模型建立与训练过程中，需要进行模型选择和参数调整，模型选择包括选择合适的算法和模型结构，参数调整则包括调整算法的参数和模型的超参数，以优化模型的性能。还需要进行模型训练和验证。模型训练是利用训练数据集对模型进行训练，使其能够学习到数据的特征和规律。模型验证则是利用验证数据集对模型进行评估，以检验模型的准确性和可靠性。

#### 2.4 预测结果分析与评估

预测结果分析与评估是基于大数据的铁路电力机车运行故障预测模型中的最后一步，也是确保模型准确性和可靠性的关键步骤。通过对预测结果进行分析和评估，可以了解模型的性能表现，找出模型的不足之处，为进一步优化提供依据。预测结果分析包括对比实际故障数据和预测数据，计算模型的准确率、召回率、F1值等指标。准确率是指模型正确预测故障的比例，召回率是指模型能够预测出实际发生故障的比例，F1值则是准确率和召回率的调和平均数，用于综合评估模型的性能。除了准确率、召回率和F1值等指标外，还可以利用混淆矩阵、ROC曲线和AUC值等工具对预测结果进行深入分析。混淆矩阵可以展示模型在不同类别上的预测结果，ROC曲线则可以展示模型在不同阈值下的性能表现，AUC值则是ROC曲线下的面积，用于评估模型的性能。在预测结果分析与评估过程中，需要注意避免过拟合和欠拟合的问题。过拟合是指模型在训练数据上表现良好，但在测试数据上表现较差，这通常是由于模型过于复杂或选择了过多的特征导致的。欠拟合则是指模型在训练数据和测试数据上表现都不好，这通常是由于模型过于简单或选择了过少的特征导致的。在预测结果分析与评估过程中，需要综合考虑模型的准确率、召回率、F1值等指标，以及模型的复杂度和特征选择等因素，进行综合考虑和权衡。通过对预测结果的分析与评估，可以找出模型的不足之处，如特征选择不当、模型复杂度过高或过低等问题。针对这些问题，可以进一步优化特征选择、调整模型参数和结构，以提高模型的准确性和可靠性。还可以将优化后的模型应用于实际电力

机车的运行故障预测中，验证模型的实用性和效果。基于大数据的铁路电力机车运行故障预测模型是一个复杂而系统的过程，需要综合考虑数据采集与处理、特征选择与提取、模型建立与训练以及预测结果分析与评估等多个环节。

### 3 铁路电力机车运行故障预测的应用实践

#### 3.1 应用案例介绍

在铁路电力机车运行故障预测领域，应用实践已经取得了显著的成果。某大型铁路公司引入了基于大数据和机器学习的电力机车运行故障预测系统。该系统首先对电力机车运行过程中的各种数据进行全面采集，包括电流、电压、温度、湿度等实时参数，以及机车的运行状态、维护记录和故障历史等信息。然后，通过特征选择与提取技术，从原始数据中提取出对故障预测有重要意义的特征。接下来，利用机器学习算法和深度学习算法，构建出能够准确预测电力机车运行故障的模型<sup>[3]</sup>。在实际应用中，该系统通过实时监测电力机车的运行状态，对潜在的故障进行预警。当系统检测到某些特征值出现异常时，会立即触发预警机制，通知相关人员对电力机车进行检查和维护。通过这种方式，该系统成功避免了多次潜在的故障发生，确保了铁路运输的安全和高效运行。该系统还具备故障类型识别和故障程度评估的功能，通过对故障数据的分析和挖掘，系统可以准确识别出故障的类型，如电气系统故障、机械部件损坏、控制系统失灵等，并评估故障的严重程度。这为维修人员提供重要的参考信息，帮助他们快速定位故障点，制定有效的维修方案，缩短维修时间和成本。

#### 3.2 预测效果分析

对于上述铁路电力机车运行故障预测系统的应用实践，可以从多个方面对其预测效果进行分析。首先，从准确率方面来看，该系统在故障预测方面表现出了较高的准确性。通过对历史数据的训练和验证，系统能够准确识别出潜在的故障点，并给出相应的预警。在实际应用中，系统成功避免多次潜在的故障发生，证明其预测结果的可靠性。其次，从召回率方面来看，该系统也表现出良好的性能。召回率是指系统能够预测出实际发生故障的比例。在实际应用中，该系统对大部分实际发生的故障都进行有效的预警，说明其召回率较高，能够全面覆盖潜在的故障风险。另外，还可以从F1值、混淆矩阵、ROC曲线和AUC值等指标对系统的预测效果进行更深入的评估。F1值是准确率和召回率的调和平均数，用于综合评估系统的性能。混淆矩阵可以展示系统在不同类别上的预测结果，ROC曲线则可以展示系统在不同阈

值下的性能表现，AUC值则是ROC曲线下的面积，用于评估系统的分类性能。通过对这些指标的分析，可以更全面地了解系统的预测效果，并找出其不足之处，为进一步优化提供依据。总的来说，该铁路电力机车运行故障预测系统在实际应用中表现出了良好的预测效果。通过实时监测电力机车的运行状态，对潜在的故障进行预警，该系统成功避免了多次潜在的故障发生，确保了铁路运输的安全和高效运行。

#### 4 基于预测结果的铁路电力机车运行改进策略

##### 4.1 预防性维护策略

基于铁路电力机车运行故障的预测结果，实施预防性维护策略是提升机车可靠性和延长使用寿命的关键。具体实践中，铁路部门应建立完善的预防性维护计划，根据预测结果确定维护的优先级，对于预测故障概率高且影响严重的部件，应优先安排维护。结合机车的运行计划和维护资源，制定详细的维护时间表，确保维护作业在不影响机车正常运行的前提下进行。还应加强对维护人员的培训，提升他们的专业技能和应对复杂故障的能力，确保维护作业的质量和效率。预防性维护策略的实施，不仅能够有效降低电力机车的故障率，减少因故障导致的运行中断和延误，还能延长机车的使用寿命，降低整体的维护成本。

##### 4.2 故障诊断与排除策略

在铁路电力机车运行故障预测的基础上，故障诊断与排除策略旨在快速、准确地定位并修复故障，确保机车能够迅速恢复运行。一方面，当预测模型发出故障预警时，应立即启动故障诊断程序。利用先进的故障诊断技术，如振动分析、热成像检测、频谱分析等，对机车的关键部件进行实时监测和分析，以快速确定故障的具体位置和原因<sup>[4]</sup>。另一方面，根据故障诊断结果，制定针对性的维修方案，并立即组织维修人员进行修复作业。应建立完善的故障记录和分析系统，对每次故障的发生、诊断、修复过程进行详细记录和分析，以积累经验教训，为今后的故障诊断和排除提供参考。故障诊断与排除策略的实施，能够显著提高电力机车故障处理的效率和准确性，减少因故障导致的运行中断时间，提升

铁路运输的整体服务水平。

##### 4.3 运行优化策略

基于铁路电力机车运行故障的预测结果，运行优化策略旨在通过调整机车的运行参数、优化运行计划和提升运行效率，来降低故障发生的风险和提升整体运行效能。具体而言，运行优化策略应包括以下几个方面：一是根据预测结果，对机车的运行速度、加速度、制动性能等关键参数进行动态调整，以减少对机车部件的磨损和疲劳，降低故障发生的概率。二是优化机车的运行计划，合理安排机车的发车间隔、运行路径和停靠站点，以减少机车的运行时间和能耗，提升整体运行效率。三是加强对机车司机的培训和管理，提高他们的驾驶技能和应急处理能力，确保机车在运行过程中能够保持稳定、安全的状态。运行优化策略的实施，不仅能够降低电力机车的故障率，提升运行效率，还能减少能源消耗和环境污染，推动铁路运输向更加绿色、高效的方向发展。

##### 结束语

基于大数据的铁路电力机车运行故障预测及改进策略是提升铁路运输效能和安全性的重要途径。通过深入挖掘数据价值，能够更准确地预测故障，制定更加科学合理的维护计划和运行策略。未来，随着技术的不断进步和数据的持续积累，有理由相信，铁路运输将更加高效、安全、环保，为人们的出行和货物的运输提供更加优质的服务。

##### 参考文献

- [1]袁金斗.高压电力客户用电安全在线检查技术研究及展望[J].电子技术应用,2024,50(09):1-8.DOI:10.16157/j.issn.0258-7998.245175.
- [2]刘兴明.基于大数据和人工智能的新能源运维优化研究[J].中国新技术新产品,2024,(14):37-39.DOI:10.13612/j.cnki.cntp.2024.14.046.
- [3]袁青乐.机车电气系统故障的快速判断与排除[J].建筑工程技术与设计,2017,(22):2383-2383.
- [4]梁杰,饶天贵.基于故障树的机车故障诊断系统设计[J].计算机应用与软件,2020,37(6):328-333.