

基于大数据的建筑工程质量检测风险预警模型构建

高建春 吴少波

云南科仑工程质量检测有限公司 云南 昆明 650299

摘 要：建筑工程质量检测对于保障工程安全至关重要，然而传统检测方式面临诸多局限。在此背景下，本文先阐述大数据在建筑工程质量检测中的多方面应用价值；接着介绍模型构建流程，包括确定预警指标、数据预处理、模型构建、验证与优化；随后探讨关键技术，涵盖数据存储与管理、机器学习算法、模型评估指标及实时预警技术。通过整合大数据与多种技术，旨在提升建筑工程质量检测水平，实现质量风险的及时预警与有效管控。

关键词：大数据；建筑工程；质量检测；风险预警；模型构建

引言

在建筑工程领域，质量检测与风险预警是保障工程安全与质量的关键环节。随着大数据技术的蓬勃发展，其强大的数据整合、分析与挖掘能力为建筑工程质量检测带来了新的契机。通过构建基于大数据的建筑工程质量检测风险预警模型，可全面整合工程全生命周期数据，实现实时监测、精准评估与及时预警。本文将深入探讨该模型的应用价值、构建流程、关键技术等内容，为提升建筑工程质量检测水平、保障工程安全提供有力支持^[1]。

1 大数据在建筑工程质量检测中的应用价值

大数据技术可整合建筑工程全生命周期各类数据，涵盖规划、设计、施工到竣工验收阶段，包括设计图纸、施工记录、材料检测报告、环境监测数据等，打破数据孤岛形成全面系统的数据集合，综合分析能全面了解质量状况，发现潜在质量风险因素。借助传感器、物联网等技术，大数据可实现建筑工程质量实时监测，采集施工过程各项数据如混凝土强度、结构变形、温度湿度等并传输至数据处理中心分析，数据异常时及时发出预警信号，提醒相关人员采取措施，动态管控质量风险，避免质量事故扩大。大数据分析技术能深度挖掘海量建筑工程质量数据，发现数据关联关系与潜在规律，通过建立数据分析模型准确评估质量风险等级，为质量检测人员和管理者提供决策依据，据此确定重点检测部位与项目，优化检测方案，提高检测效率与针对性。大数据记录建筑工程全生命周期质量相关信息，为质量追溯提供有力支持，出现质量问题时可快速查询记录，定位问题产生的时间、环节与责任主体，便于责任认定与整改，同时大数据分析能总结质量问题产生原因与规律，为后续建筑工程质量管理提供经验，避免类似问题重复出现，从而有效提升建筑工程质量检测水平^[2]。



图 1 大数据在建筑工程质量检测

2 建筑工程质量检测风险预警模型构建流程

2.1 确定预警指标

依据建筑工程自身特性以及质量要求来筛选指标，建筑工程涵盖多种结构类型与施工内容，不同部分对质量要求各异，所以要从整体质量把控角度出发，选取能全面反映质量状况的指标。结构变形指标是重要一项，建筑结构在施工及使用过程中，变形情况直接影响其安全性与稳定性，通过监测结构变形数据，可及时察觉潜在风险。例如，对某桥梁工程进行监测，发现其桥墩在施工过程中的沉降量超出设计允许值的 5% 时，可能存在结构安全隐患。材料强度指标也不可或缺，建筑材料强度关乎建筑承载力，若材料强度不达标，易引发结构破坏等质量问题^[2]。如某建筑项目使用的钢筋，经检测其屈服强度低于标准值 8%，导致部分梁柱出现裂缝。施工工艺指标同样关键，施工工艺是否规范合理，对工程质量影响巨大，不恰当的施工工艺可能导致质量隐患。确定这些指标时，要确保其具备可测量性，能通过相应检测设备或技术手段获取准确数据；同时要有可操作性，在实际检测工作中能够方便实施，便于检测人员操作与记录^[3]。

2.2 数据预处理

数据预处理是数据分析与建模流程中的关键环节，需对收集到的原始数据执行系统性处理操作，主要涵盖数据清洗与数据转换两个核心步骤。数据清洗阶段需识别并处理数据中存在的噪声与异常值，通过缺失值检测机制定位数据集中缺失的字段，根据业务需求选择填充策略或直接删除缺失记录；同时运用统计方法或领域知识识别偏离正常范围的异常数据点，采用修正、平滑或剔除等方式确保数据质量符合分析标准。数据转换环节则聚焦于数据格式的标准化处理，针对文本类非结构化数据，需通过编码技术将其转换为数值型特征向量，常见方法包括独热编码、词嵌入等；对于日期时间等特殊类型数据，需分解为年、月、日等结构化字段；数值型数据则需进行标准化或归一化处理，消除量纲差异对模型训练的影响。

2.3 模型构建

模型构建是风险预警系统开发的核心环节，需根据问题特性选择适配的建模方法，常用方法包括神经网络、决策树、支持向量机及集成学习算法等，其中神经网络通过多层非线性变换捕捉复杂特征关系，决策树则基于特征阈值划分构建可解释性规则。确定方法后，需将预处理后的数据划分为训练集与测试集，通常按 7:3 或 8:2 比例随机分割，确保两类数据分布一致且无重叠。训练阶段使用训练集数据迭代优化模型参数，神经网络通

过反向传播算法调整权重，决策树通过信息增益或基尼系数确定分裂节点，过程中需监控过拟合现象并采用正则化、早停等策略控制模型复杂度。参数调优环节可运用网格搜索、随机搜索或贝叶斯优化等方法，在预设参数空间内寻找最优组合。完成训练后，使用独立测试集评估模型泛化能力，采用准确率、召回率、F1 值、ROC 曲线等指标量化性能表现，同时分析混淆矩阵定位误分类样本。若测试结果未达预期，需返回调整模型结构或参数重新训练，最终输出满足精度要求的工程风险预警模型，确保其在实际应用中具备稳定可靠的预测能力。

2.4 模型验证与优化

完成模型构建后，必须通过实际工程数据对模型进行全面而严格的验证，以量化评估其预测准确性。这一过程不仅要求使用独立于训练集的测试数据，以避免信息泄露导致的评估偏差，更需重点考察模型在未见数据上的泛化能力，确保其能在实际工程环境中稳定、准确地预测风险。若验证发现模型预测偏差超出可接受范围，则需启动优化流程，通过深入分析误差来源，定位模型缺陷，并针对性地实施优化措施。这一过程往往涉及多次迭代，包括补充关键特征、调整模型超参数、增加数据样本量或切换更适配的建模方法等，直至模型在各项评估指标上达到预设阈值，且在不同场景下保持稳定预测能力。表 1 将详细展示模型验证与优化的具体流程及其关键步骤，为实际工程质量风险监控提供有力支持。

表 1 模型验证与优化流程及关键步骤

阶段	流程步骤	具体内容
验证阶段	1. 数据准备	使用独立于训练集的测试数据，如某建筑工程后续施工阶段的 200 组数据
	2. 模型预测	运用构建好的模型对测试数据进行预测，获取预测结果
	3. 对比评估	对比模型预测结果与真实风险发生情况，量化评估预测准确性，如预测准确率为 85%
优化阶段（若验证偏差超出范围）	1. 误差分析	通过特征重要性分析、残差分布诊断等方法定位模型缺陷
	2. 确定原因	判断误差来源，如特征工程不足、参数设置不当或算法选择偏差
	3. 优化措施	针对性实施优化，如补充关键特征、调整模型超参数（如神经网络学习率调整为 0.005）、增加数据样本量或切换建模方法
再验证阶段	1. 重新训练	根据优化措施重新训练模型
	2. 再次验证	使用新的测试数据再次验证模型性能，如准确率提升到 88%
	3. 闭环迭代	形成“验证 - 优化 - 再验证”的闭环迭代过程，直至模型满足预设阈值且稳定

3 建筑工程质量检测风险预警模型关键技术

3.1 数据存储与管理技术

建筑工程在全生命周期内产生的数据量极为庞大，涵盖设计、施工、监理等多环节，涉及结构参数、材料性能、施工工艺等各类信息，这对数据存储与管理技术提出极高要求。需采用分布式文件系统与数据库管理

系统相结合的方式。分布式文件系统具备强大的扩展性和容错性，能将数据分散存储在多个节点上，通过并行处理机制，实现对海量建筑工程数据的高效存储与快速访问，满足大规模数据实时读写需求，确保数据完整性和可用性^[3]。数据库管理系统则侧重于数据的组织与管理，通过建立科学合理的数据模型，对建筑工程数据进

行结构化存储,将不同类型、不同来源的数据进行有序整合。同时,利用数据库管理系统提供的查询语言和索引机制,能够快速定位所需数据,为后续的数据分析提供便利。二者协同作用,分布式文件系统解决海量数据存储与访问难题,数据库管理系统实现数据有效组织与便捷查询,共同为建筑工程质量检测风险预警模型提供坚实的数据基础,保障模型能够基于全面、准确的数据进行风险分析与预警。

3.2 机器学习算法

建筑工程数据具有多维性、复杂性和动态性等特点,包含结构设计参数、施工过程记录、材料性能指标等多方面信息,且数据随工程进展不断更新变化。不同机器学习算法在原理、性能和适用场景上存在差异,需依据建筑工程数据特性与预警需求进行选择。支持向量机通过寻找最优分类超平面实现数据分类与回归,在小样本数据场景下能较好地捕捉数据特征,对建筑工程中部分关键但样本量少的检测指标分析具有优势。随机森林算法基于多棵决策树集成,可处理高维数据,在处理建筑工程中大量不同类型检测数据时,能有效降低过拟合风险,且对数据中的噪声干扰有较强抵抗能力。

3.3 模型评估指标

在建筑工程质量检测风险预警模型构建完成后,为精准衡量其性能优劣,需选取恰当的模型评估指标,这些指标从不同维度反映模型对风险的识别与判断能力,为模型优化和应用提供关键依据。准确率是基础评估指标之一,它通过计算模型预测正确的样本数在总样本数中所占比例,直观呈现模型整体预测的准确程度。在建筑工程质量检测场景中,该指标可体现模型对各类风险状况判断的准确性,数值越高表明模型预测结果与实际情况相符程度越高。例如,某建筑工程质量风险预警模型在测试集上的准确率为90%,说明该模型在整体上能够较为准确地预测工程质量风险。召回率侧重于衡量模型正确预测出正样本的能力,在建筑工程质量检测里,正样本通常代表存在质量风险的情况,召回率高意味着模型能尽可能多地识别出实际存在的风险,减少漏检情况。例如,某模型在测试集上的召回率为85%,说明该模型能够识别出85%的实际存在的质量风险。F1值作为综合评估指标,是准确率和召回率的调和平均数,它兼

顾了模型的精确性和召回能力,避免单一指标可能存在的偏差。

3.4 实时预警技术

实时预警技术在建筑工程质量检测风险预警体系中至关重要,是达成风险预警模型核心目标的关键环节。构建完备的实时预警机制,需先将精心构建的风险预警模型与高效可靠的实时数据采集系统深度融合。实时数据采集系统持续不间断地收集建筑工程在施工过程中的各类关键数据,涵盖结构应力、变形位移、环境温度、材料强度等多维度信息,确保数据获取的及时性与全面性^[4]。这些实时采集的数据会被迅速传输至风险预警模型进行实时分析处理。模型依据预设的预警指标范围对数据进行比对判断,一旦监测数据超出正常范围,即判定为潜在风险状况。此时,系统会立即触发预警机制,以多种高效便捷的方式向相关人员发送预警信息。短信方式能确保信息快速直达,即便相关人员处于移动状态也能及时接收;邮件方式可提供详细的预警内容及相关数据报告,方便后续查阅与分析;系统提示则能在工程现场的管理系统中直接弹出,引起现场人员的即时关注。

结语

综上所述,基于大数据的建筑工程质量检测风险预警模型的构建,是提升建筑工程质量与安全性的的重要举措。通过整合全生命周期数据、应用机器学习算法、构建严谨的模型评估与实时预警机制,该模型能够精准识别与及时预警潜在质量风险。未来,随着技术的不断革新与实际应用经验的累积,该模型将进一步优化与完善,为建筑工程质量提供更为可靠的保障,推动建筑行业向更高质量、更安全的方向发展。

参考文献

- [1] 刘景德.数字化转型背景下建筑工程质量风险的监控与预警机制[J].产品可靠性报告,2025(5):85-86.
- [2] 王国英,刘永灿.基于机器学习的建筑工程进度预测与风险预警模型研究[J].计算机应用文摘,2025,41(15):118-120.
- [3] 廖金凤.建筑工程检测信息化管理平台的构建与应用研究[J].开放系统世界(香港),2025,1(4):75-77.
- [4] 陈妙妮.建筑工程造价指标智能化测算技术应用及优化策略[J].住宅与房地产,2025(26):71-73.