

# 大数据驱动的工程造价风险评估模型开发与应用

刘永恒

安徽天都建筑工程咨询有限公司, 中国·安徽 安庆 246000

**摘要:** 随着工程项目的复杂性和不确定性增加, 传统的工程造价风险评估方法逐渐暴露出局限性。大数据技术的应用为工程造价风险评估提供了新的思路和方法。论文基于大数据驱动的工程造价风险评估模型, 探讨了其理论基础、数据获取与处理、模型构建与优化等方面的内容。通过综合运用大数据、机器学习、数据挖掘等技术, 建立起能够实时、精准识别项目风险的模型。实证分析表明, 该模型能够有效提高工程项目的风险预测精度, 优化造价管理, 并为决策者提供科学的风险应对措施。

**关键词:** 大数据; 工程造价; 风险评估; 机器学习; 数据处理; 模型优化

## Development and Application of Big Data Driven Engineering Cost Risk Assessment Model

Yongheng Liu

Anhui Tiandu Construction Engineering Consulting Co., Ltd., Anqing, Anhui, 246000, China

**Abstract:** With the increasing complexity and uncertainty of engineering projects, traditional cost risk assessment methods are gradually exposing their limitations. The application of big data technology provides new ideas and methods for engineering cost risk assessment. The paper is based on a big data-driven engineering cost risk assessment model, exploring its theoretical basis, data acquisition and processing, model construction and optimization, and other aspects. By comprehensively utilizing technologies such as big data, machine learning, and data mining, a model capable of real-time and accurate identification of project risks is established. Empirical analysis shows that the model can effectively improve the accuracy of risk prediction in engineering projects, optimize cost management, and provide decision-makers with scientific risk response measures.

**Keywords:** big data; engineering cost; risk assessment; machine learning; data processing; model optimization

## 0 前言

在当下, 工程项目愈发复杂多变, 传统工程造价风险评估手段已难应对。鉴于此, 大数据技术应运而生, 为其开辟新径。论文聚焦大数据驱动的评估模型, 深挖理论、数据、构建等要点, 助力精准控风险、优造价。

## 1 大数据驱动的工程造价风险评估理论基础

### 1.1 工程造价风险评估相关理论

#### 1.1.1 风险管理理论

风险管理是识别、评估、控制和监控各种风险因素的过程, 其目的是尽量减少不确定性对项目目标的负面影响。在工程项目中, 风险管理理论通常包括风险识别、风险评估、风险控制和风险监控四个步骤。尤其在造价管理中, 风险评估环节至关重要, 它帮助项目团队识别可能影响工程预算、成本、进度的风险因素, 预测这些风险对项目造价的潜在影响, 并制定合理的应对措施。

#### 1.1.2 工程造价风险评估模型的发展

工程造价风险评估模型经历了从定性到定量, 再到混合模型的逐步发展。

**定性模型:** 早期的风险评估主要依赖专家经验和定性分析方法, 如德尔菲法、SWOT 分析等, 评估过程中主要

依赖主观判断, 适用于风险因素较为简单或不确定性较大的场景。

**定量模型:** 随着数据分析技术的发展, 定量模型逐渐兴起, 利用概率论、统计分析、回归分析等方法, 基于历史数据进行风险量化评估, 能够提供更准确的风险预测结果。

**混合模型:** 近年来, 混合模型结合了定性与定量的优点, 通过模糊综合评判、神经网络、贝叶斯网络等方法, 对复杂的风险因素进行综合分析。混合模型不仅能够处理结构化数据, 还能有效处理不确定性和模糊性, 提高了评估的准确性和实用性<sup>[1]</sup>。

#### 1.1.3 风险评估的指标体系

为了科学评估工程造价风险, 需建立合理的风险评估指标体系。一般来说, 主要包括以下几个方面的指标:

**成本风险:** 包括原材料价格波动、人工成本变动、供应链中断等因素, 这些都可能造成工程造价的超支。

**时间风险:** 涉及工期延误、施工进度滞后等因素, 时间的延误往往会带来额外的费用, 如人工和机械的租赁费用等。

**质量风险:** 质量问题可能导致返工、延误和资源浪费, 从而增加项目的直接和间接成本。

**外部风险:** 如政策法规变化、市场需求波动等外部因

素对工程造价的影响。通过对这些指标的量化评估,能够为项目管理者提供有效的风险预警。

## 1.2 大数据与风险评估的融合

### 1.2.1 大数据驱动的风险评估模型框架

大数据技术的引入,使得传统的工程造价风险评估方法发生了根本性的变化。基于大数据的风险评估模型通常由数据采集、数据处理、模型构建和风险预测四个主要部分组成。

**数据采集:**通过传感器、项目管理系统、历史数据库等渠道,收集工程造价、进度、质量、市场和政策等相关数据。

**数据处理:**对原始数据进行清洗、标准化、特征提取等处理,确保数据质量和可用性。

**模型构建:**基于机器学习、数据挖掘等技术,构建多元化的风险评估模型,综合考虑各种影响因素,预测项目的潜在风险。

**风险预测与控制:**通过分析评估结果,提出风险防范措施,并实时监控风险变化,进行动态调整。

### 1.2.2 大数据对传统风险评估方法的影响

大数据的应用使得传统的工程造价风险评估方法得到了显著提升。传统方法多依赖专家判断和定性分析,评估结果容易受到主观因素的影响,且在面对大量复杂数据时缺乏有效的处理手段。而大数据通过对海量数据的深度挖掘与分析,能够揭示出潜在的规律和趋势,使得风险评估更加精准和科学。

例如,传统的风险评估方法可能只能识别明显的成本波动因素,而大数据技术能够通过对历史项目数据的综合分析,识别出那些不易察觉但同样会影响成本的潜在风险,如季节性价格波动、突发自然灾害等。此外,大数据还能够实现实时监控,使得风险评估不再是静态的,而是动态的,可以随时调整应对措施。

## 2 数据获取与处理

### 2.1 数据源与数据类型

#### 2.1.1 工程项目历史数据

工程项目历史数据是进行造价风险评估的核心数据来源,通常包括项目的设计、施工、竣工以及维护阶段的详细信息。这些数据涵盖了成本、进度、质量控制、资源使用等方面的内容,能够为预测未来项目中的潜在风险提供宝贵的参考。例如,历史造价数据能够揭示某些风险因素(如原材料价格波动、施工延误等)对项目预算的影响,为模型提供准确的输入。

#### 2.1.2 环境数据、政策法规数据、市场数据

除了项目本身的数据,外部环境数据也是不可忽视的重要信息来源。环境数据包括气候条件、地理环境等因素,这些因素可能会影响施工进度与安全。政策法规数据则涵盖政府对建筑行业的相关规定、环保要求、税务政策等,可

能直接影响项目的成本和风险。市场数据则包括建筑材料、劳动力成本的变化趋势,以及宏观经济环境对工程造价的影响。这些外部数据为造价风险评估提供了动态背景。

#### 2.1.3 企业内部与外部数据的集成

在大数据驱动的风险评估中,数据的整合是关键环节。企业内部数据通常包括历史财务数据、供应链管理数据、项目执行过程中产生的实时数据等,而外部数据则包括行业标准、市场趋势、政策法规等。通过将内外部数据进行融合,能够获得更全面的视角,使得风险评估更加精准。

## 2.2 数据预处理方法

### 2.2.1 数据清洗与缺失值处理

在实际应用中,收集到的原始数据往往存在错误、重复或缺失等问题,这会影响模型的准确性。因此,数据清洗是必不可少的步骤。数据清洗的过程包括去除冗余数据、修正错误信息以及填补缺失值等。缺失值处理方法有多种,如均值填补、插值法、KNN 填补等,这些方法可以有效减少数据缺失带来的偏差,确保后续分析的可靠性。

### 2.2.2 数据标准化与归一化

为了避免不同量纲的数据影响模型的表现,数据标准化与归一化是预处理中的重要步骤。标准化将数据转换为均值为 0、方差为 1 的分布,归一化则将数据压缩到特定的范围(如 0 到 1)。这有助于消除特征之间的量纲差异,确保各项数据在模型中具有相同的影响力,从而提高模型的稳定性与预测能力。

### 2.2.3 特征选择与降维方法

特征选择和降维是大数据处理中常用的两种技术。特征选择旨在通过挑选最相关的特征,剔除冗余或不重要的变量,从而减少计算复杂性并提升模型的性能。常见的特征选择方法包括基于相关性的选择、信息增益、卡方检验等。而降维技术(如主成分分析 PCA)则通过将高维数据映射到低维空间,减少特征数目,同时保留数据的主要信息,避免维度灾难并提高模型的效率。

## 2.3 数据分析与建模方法

### 2.3.1 数据挖掘与模式识别

数据挖掘是从大量数据中提取有用信息的过程,尤其适用于揭示潜在的模式和规律。在工程造价风险评估中,数据挖掘可以通过分类、聚类、关联规则等技术,识别影响项目造价的关键因素和潜在风险。例如,通过对历史项目数据的分析,可以识别出材料价格波动、劳动力短缺等因素与项目风险之间的关系,帮助管理者在项目初期就采取预防措施。

### 2.3.2 机器学习与深度学习在风险评估中的应用

机器学习和深度学习是现代数据分析中的重要方法,尤其适用于大规模数据的自动化分析。通过训练机器学习模型(如决策树、支持向量机、随机森林等),可以实现对项目风险的高效预测。在工程造价风险评估中,机器学习能够

学习历史项目中的风险模式，并根据新的数据进行实时预测。此外，深度学习技术，如神经网络，可以通过多层次的模型结构识别复杂的非线性关系，进一步提高评估的精确度。

### 2.3.3 统计分析 with 风险预测方法

统计分析是工程造价风险评估中传统且重要的分析方法，通过回归分析、方差分析等统计方法，可以对项目中各类风险因素进行量化评估。风险预测方法如蒙特卡罗模拟、贝叶斯网络等，可以基于大量历史数据和风险变量，预测不同风险场景下的可能结果。这些方法能够提供关于项目造价和进度的概率分布，帮助决策者制定更合理的应对策略<sup>[2]</sup>。

## 3 工程造价风险评估模型的构建与优化

### 3.1 模型构建

#### 3.1.1 基于大数据的风险评估模型框架

基于大数据的工程造价风险评估模型通常包含数据采集、数据预处理、风险分析和结果预测等几个核心模块。在数据采集阶段，模型通过集成项目管理系统、市场动态、历史数据以及政策法规等多个数据源，构建全面的风险评估基础。数据预处理是将原始数据进行清洗、标准化、特征选择等处理，以确保数据的准确性和可用性。风险分析模块则通过应用数据挖掘、机器学习等方法，识别潜在的风险因素，并量化其可能对项目造价的影响。最终，模型通过预测与评估，将风险结果反馈给决策者，帮助其采取适当的应对措施。

#### 3.1.2 模型输入与输出变量设计

为了构建高效的风险评估模型，需要科学设计输入和输出变量。输入变量通常包括项目的基本信息（如项目规模、工期、工程类型等），历史项目数据（如以往项目的造价数据、进度数据等），外部数据（如市场价格、政策法规变化等）以及动态数据（如项目执行过程中出现的风险事件等）。这些输入变量为模型提供了全面的信息。输出变量则主要是项目风险的预测结果，通常包括成本风险、时间风险、质量风险等的概率分布，以及各类风险对项目造价的潜在影响。

#### 3.1.3 评估指标与权重的确定

在构建风险评估模型时，需要根据不同风险类型设定评估指标，并对每个指标的权重进行合理分配。常见的评估指标包括材料价格波动、劳动力成本、施工进度延误、质量问题等。权重的确定通常基于专家经验、历史数据分析或者通过多种方法（如层次分析法、模糊综合评判等）进行计算。合理的指标和权重设计能够确保模型评估的准确性和有效性。

### 3.2 模型优化

#### 3.2.1 模型准确性与可靠性的提升

优化工程造价风险评估模型的首要任务是提高其准确性和可靠性。首先，采用更多的数据来源，增加数据的多样性和完整性，有助于减少偏差，提升模型的泛化能力。其次，

通过引入更为先进的算法（如深度学习、集成学习等），可以提高模型的预测精度。最后，使用交叉验证、拟合控制等方法，确保模型在不同数据集上的稳定性和可靠性，从而提升模型的适用性。

#### 3.2.2 迭代优化与模型验证

为了不断提升风险评估模型的表现，需要通过迭代优化的方式进行改进。这一过程通常包括模型参数的调整、算法的更新以及特征工程的不断完善。在模型优化过程中，验证是关键步骤。通过对历史项目数据的验证，检测模型的预测能力是否符合实际情况。常用的验证方法包括留出法、交叉验证法等，通过不同的训练集和验证集组合，验证模型的泛化能力和稳定性。

#### 3.2.3 风险评估模型的可解释性与应用性

虽然复杂的机器学习和深度学习模型通常能提供较高的预测精度，但其“黑箱”特性往往使得模型缺乏可解释性。为了提高模型在工程管理中的应用性和透明度，需要采取措施增强模型的可解释性。常见的方法包括使用决策树、特征重要性分析、LIME（局部可解释性模型-agnostic 解释方法）等，帮助决策者理解模型如何根据输入数据进行决策。此外，优化后的模型应具备较强的适应性，能够根据不同类型的工程项目进行快速调整，以满足不同项目的需求。

### 3.3 模型应用与实证分析

#### 3.3.1 工程案例分析与风险评估结果展示

在模型应用阶段，通过选择典型工程案例进行实证分析，可以验证模型的实际效果。通过对历史项目数据的回顾分析，模型能够准确预测出可能的风险因素，并量化其对项目造价的影响。例如，在一个大型建筑项目中，模型可能会根据历史数据和实时市场信息预测出原材料价格波动、施工延误等风险因素，从而帮助管理者提前做好风险应对措施。模型预测的结果可以通过图表、报告等形式直观展示，便于决策者理解<sup>[3]</sup>。

#### 3.3.2 模型在不同类型工程项目中的适用性验证

为了检验风险评估模型的普适性，需要在不同类型的工程项目中进行应用验证。例如，在基础设施建设、房地产开发、工业工程等项目中，评估模型需要考虑各自不同的风险因素。通过对多个项目的风险评估与模型预测结果进行对比，可以验证模型是否具备较强的跨行业适应性，从而进一步提升其在实际项目中的应用价值。

#### 3.3.3 风险评估结果对造价控制的实际影响

模型的核心价值在于能够为工程造价管理提供决策支持。通过对潜在风险的精准识别和量化，决策者能够更有效地控制项目造价。例如，在模型预测出成本超支风险较高时，项目管理者可以提前采取采购策略调整、资源优化、进度控制等措施，从而降低造价风险并确保项目的顺利完成。实证分析表明，基于大数据的风险评估模型能够在多个案例中帮助实现成本的有效控制，显著提升项目的经济效益。

## 4 结语

大数据驱动的工程造价风险评估模型能够通过多源数据的集成和先进的分析技术,显著提高风险评估的准确性和动态性。与传统方法相比,基于大数据的模型更具全面性和实时性,能够及时揭示潜在风险并为项目管理提供决策支持。在实际应用中,模型优化的不断推进提升了其准确性和可解释性,从而为工程项目的造价控制和风险管理提供了强

有力的工具。

### 参考文献:

- [1] 陈怡冰,王莉香.合同管理对工程结算及工程造价的影响[J].散装水泥,2021(3):18-20.
- [2] 杨晓.工程合同管理中风险评估及对策[J].工程技术研究,2018(13):133-134.
- [3] 李东辉.试析工程合同管理中风险评估及对策[J].江西建材,2018(4):254+256.