

基于大数据分析的建筑工程 造价风险预警机制研究

赵 君

(安徽德瑞工程咨询有限公司, 安徽 芜湖 241000)

摘 要 基于大数据分析研究建筑工程造价风险预警机制, 从市场环境、项目管理和工程技术三大维度识别风险因素, 分析其分类与特征。构建造价风险预警指标体系, 选取成本偏差率、价格波动指数、进度绩效指数和质量缺陷率等关键指标, 运用大数据分析方法对指标数据进行挖掘与处理。选择支持向量机(SVM)模型构建预警模型, 通过历史数据训练与优化模型, 确定预警阈值。研究表明: 基于大数据分析的建筑工程造价风险预警机制能有效预警造价风险, 为项目管理提供决策支持。

关键词 建筑造价风险; 大数据分析; 支持向量机; 风险指标体系; 动态预警

中图分类号: TU723

文献标志码: A

DOI: 10.3969/j.issn.2097-3365.2025.24.024

0 引言

建筑工程造价风险贯穿项目全过程, 受多种因素影响, 如原材料价格波动、设计变更、施工进度延误等, 可能导致造价失控。随着大数据技术的发展, 其海量性和动态性特点为造价风险预警提供了数据支持^[1]。然而, 当前造价风险预警研究多为定性分析或单一因素定量分析, 缺乏基于大数据的系统性、动态性预警机制。因此, 本研究基于大数据分析, 构建建筑工程造价风险预警机制, 旨在为项目管理提供科学决策支持, 促进建筑行业健康可持续发展。

1 建筑工程造价风险因素分析

1.1 风险因素识别

建筑工程造价风险因素复杂, 涵盖市场环境、项目管理和工程技术三大维度。在市场环境方面, 原材料价格波动(如钢材、水泥、木材)受经济形势、供求关系和政策调整影响, 波动频繁^[2]。钢材价格每吨上涨100元, 工程材料成本可能增加5%~10%。表1显示了常见建筑材料价格波动对造价的影响。

表1 常见建筑材料价格波动对造价的影响

材料类型	单位	平均用量	价格波动范围 (%)	造价影响 (%)
钢材	吨	1 000	±10%	±5%
水泥	吨	5 000	±8%	±3%
木材	立方米	200	±15%	±4%

同时, 劳动力市场变化和物价上涨导致建筑工人工资逐年上升。在项目管理维度, 设计变更因前期设计不充分、业主需求变更或地质条件不符, 会增加工程量、改变施工工艺或材料设备, 直接影响造价。施工进度延误使工期延长, 增加人工、材料、设备成本, 还可能产生违约金。其影响公式为:

$$\Delta C = \sum(N_i \times \Delta U_i) + \sum(Q_j \times \Delta P_j) + \text{违约金} \quad (1)$$

式(1)中, N_i 为用工人数, ΔU_i 为人工单价变化, Q_j 为材料用量, ΔP_j 为材料价格变化。在工程技术维度, 施工技术选择直接影响造价, 新型技术虽成本高但安全性更高。施工质量问题的返工、维修费用增加, 甚至影响使用寿命, 质量问题导致的返工和维修费用可能占总造价的3%~5%。

1.2 风险因素分类与特征

为了构建有效的预警机制, 建筑工程造价风险因素需按可控性与内外源性分类并分析其特征。可控风险(如施工进度延误、设计变更)可通过管理措施和技术手段控制, 其发生概率低但影响大, 且与其他风险(如设计变更与进度延误)相关。不可控风险(如原材料价格波动、自然灾害)受外部环境影响, 难以预测和缓解, 影响程度大, 需通过大数据分析评估其与其他风险(如原材料价格波动与人工成本变化)的关联性^[3]。内部风险(如施工质量问题)与项目管理水平和技术能力相关, 可通过优化管理流程和技术方案控制; 外部风险(如政策法规变化)受外部环境影响, 难以干预, 需通过综合评估其与其他风险的关联性。

2 基于大数据分析的建筑工程造价风险预警指标体系构建

2.1 预警指标体系的设计原则

构建建筑工程造价风险预警指标体系需遵循科学性、系统性、可操作性和灵敏性原则。科学性原则要求指标真实反映风险本质，如成本偏差率衡量成本控制效果；系统性原则强调指标体系全面覆盖风险方面，如结合成本偏差率、进度绩效指数和价格波动指数综合评估；可操作性原则指出指标需明确量化，便于数据采集，如价格波动指数提供具体数据支持；灵敏性原则要求指标能及时反映风险变化，如成本偏差率变化率助力预警^[4]。

2.2 预警指标的选取与分类

在建筑工程造价风险预警中，选取了成本偏差率（CV）、价格波动指数（PFI）、进度绩效指数（SPI）和质量缺陷率（QDR）作为关键指标，分别衡量成本控制、价格波动、进度执行及质量问题对造价的影响。这些指标被分为一级指标（成本风险、进度风险、价格风险、质量风险）和二级指标（CV、PFI、SPI、QDR）。权重分配基于各指标的影响程度，可通过专家打分法（如层次分析法，AHP）或数据分析法（如主成分分析法，PCA）确定。综合风险评分计算公式为：

$$\text{Risk Score} = \sum_{i=1}^n (w_i \times x_i) \quad (2)$$

式（2）中， w_i 为权重， x_i 为标准化指标值。预警指标的层级结构如图 1 所示。

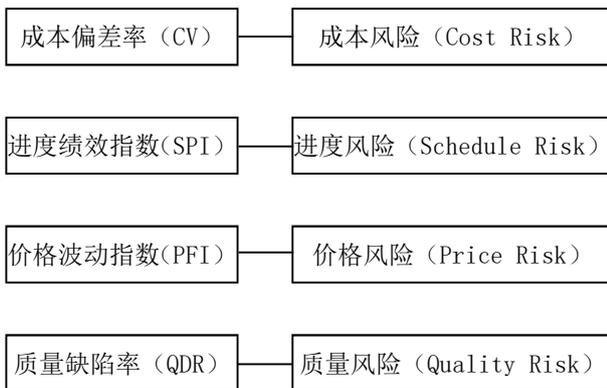


图 1 预警指标的层级结构

2.3 基于大数据的指标数据分析与处理

2.3.1 趋势分析：时间序列模型的应用

趋势分析借助时间序列模型识别造价数据中的长期趋势与季节性变化，常用 ARIMA 模型，其公式为：

$$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} \quad (3)$$

式（3）中， X_t 表示时间序列数据， μ 为常数项，

$\phi_1, \phi_2, \dots, \phi_p$ 是自回归系数，反映前期数据对当前数据的影响； $\theta_1, \theta_2, \dots, \theta_q$ 为移动平均系数，与白噪声误差项 ϵ_t 相关。应用时需先检验数据平稳性，若非平稳需差分处理，再用 AIC 或 BIC 准则确定 p 和 q 值。如今，LSTM 等深度学习方法也用于处理复杂造价数据趋势预测，能结合多变量输入，提高预警准确性。

2.3.2 关联分析：风险因素关系的量化

关联分析用于量化风险因素间的关系，常用皮尔逊相关系数，其公式为：

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4)$$

式（4）中， x_i 和 y_i 是两个风险变量的观测值， \bar{x} 和 \bar{y} 分别为它们的均值。该系数衡量两变量的线性相关程度，取值范围在 -1 到 1 之间。当 r_{xy} 接近 1 时，表示相关性较强。Spearman 秩相关系数则适用于非线性或非正态分布数据，通过计算变量的秩次进行相关性分析。此外，关联规则挖掘可发现多个风险因素之间的复杂关联，如“设计变更频繁”与“材料价格上涨”同时出现时，造价超支风险会显著增加。

2.3.3 聚类分析：风险群的识别

聚类分析采用 K-means 或层次聚类算法归类相似项目以识别风险群。K-means 的目标是最小化簇内平方和：

$$\min_{\{C_1, C_2, \dots, C_k\}} \sum_{j=1}^k \sum_{i \in C_j} \|x_i - \mu_j\|^2 \quad (5)$$

式（5）中， C_j 表示第 j 个簇， μ_j 为该簇的中心点。算法通过初始化簇心位置，不断迭代优化更新簇心和分配样本，直至收敛。层次聚类包括凝聚型和分裂型，通过不同链接方法（如单链接、完全链接、平均链接）度量距离。聚类结果可通过散点图、热力图等可视化技术展示，还可结合决策树或随机森林等分类算法挖掘各风险群的特征变量。

2.3.4 数据处理与综合风险评分

数据标准化通过公式：

$$x_i^{\text{norm}} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (6)$$

消除量纲差异，将数据缩放至 [0, 1] 区间，式（6）中， x_i 为原始数据， $\min(x)$ 和 $\max(x)$ 分别为数据的最小值和最大值，标准化后的值 x_i^{norm} 便于不同变量间的比较。权重确定可采用层次分析法或熵权法，反映各风险指标的重要性。基于此，综合风险评分计算公式为：

$$\text{Risk Score} = \sum_{i=1}^n w_i \cdot x_i^{\text{norm}} \quad (7)$$

式（7）中， w_i 为第 i 个风险指标的权重， x_i^{norm} 为标准化后的风险指标值。根据评分设定预警阈值，通常将风险分为三级：低风险 [0, 0.3]、中风险 [0.3, 0.6]、高风险 [0.6, 1]，实现对建筑造价风险的量化评估与

预警，为决策者提供及时准确的风险信息，助力采取有效措施降低风险损失。

3 基于大数据分析的建筑工程造价风险预警模型构建

3.1 预警模型的选择与建立

建筑工程造价风险具有复杂性、动态性和不确定性，需要选择能够处理多变量、非线性关系且适应性强的预警模型。选择了支持向量机（SVM）模型，因其在处理小样本、非线性数据时表现出色，且泛化能力强^[5]。SVM模型原理：假设输入数据集为 $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ ，其中 $x_i \in R_d$ 是输入特征向量， $y_i \in \{-1, 1\}$ 是对应的类别标签。SVM的目标是找到一个超平面 $w \cdot x + b = 0$ ，使得两类数据的间隔最大化。间隔的大小由 $\frac{2}{\|w\|}$ 决定，因此优化问题可以表示为：

$$\min_{w,b} \frac{1}{2} \|w\|^2 \text{ subject to } y_i(w \cdot x_i + b) \geq 1 \quad (8)$$

通过拉格朗日乘子法，转化为对偶问题：

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (9)$$

式(9)中， $\alpha_i \geq 0$ 是拉格朗日乘子。最终分类函数为：

$$f(x) = \text{sign}(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b) \quad (10)$$

式(10)中， $K(x_i, x)$ 是核函数，常用的有线性核、多项式核和径向基核（RBF）。预警指标与模型结合：本文的预警指标包括成本偏差率（CV）、价格波动指数（PFI）、进度绩效指数（SPI）和质量缺陷率（QDR）。这些指标作为输入特征向量：

$$x = [CV, PFI, SPI, QDR] \quad (11)$$

输入 SVM 模型中，通过历史数据训练模型，学习输入特征与造价风险之间的映射关系，从而实现对造价风险的预警。

3.2 模型的训练与优化

本研究利用历史造价数据和已知风险事件作为训练数据，数据来源包括已完成项目的成本记录、进度报告和质量检查报告等。采用支持向量机（SVM）模型进行训练，选择径向基核函数（RBF）：

$$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2) \quad (12)$$

通过调整参数 C （惩罚系数）和 γ ，优化模型性能。使用交叉验证方法对模型进行验证和评估，计算模型的平均准确率、召回率和 F1 分数，以确保模型的可靠性和稳定性。表 2 是模型优化过程中的部分结果。

最终选择 $C=100$ 和 $\gamma=0.1$ 的参数组合，模型的平均准确率达到 92%，召回率为 88%，F1 分数为 90%。

表 2 模型优化过程中的部分结果

参数组合	C	γ	准确率	召回率	F1 分数
1	1	0.1	0.85	0.80	0.82
2	10	0.1	0.90	0.85	0.87
3	100	0.1	0.92	0.88	0.90
4	10	0.01	0.88	0.82	0.85

3.3 预警阈值的确定

预警阈值的确定是风险预警系统的关键，本文结合统计分析、专家经验和风险偏好进行设置，并根据项目阶段和市场环境动态调整。具体而言，通过历史数据分布将风险等级划分为低、中、高三个级别，分别对应模型输出值的前 30%、中间 40% 和后 30%，阈值计算公式为：

$$T_k = \text{Percentile}(y, p_k) \quad (13)$$

式(13)中， T_k 为第 k 级阈值， y 为模型输出值， p_k 为对应百分位数。同时，邀请行业专家对阈值进行调整优化，并根据项目管理者风险偏好调整阈值以满足不同管理目标。在项目关键阶段或市场环境变化时，通过公式：

$$T_{\text{new}} = T_{\text{old}} \times (1 - \alpha) \quad (14)$$

动态调整阈值，式(14)中， α 为调整系数，以提高预警灵敏度。

4 结束语

基于大数据分析的建筑工程造价风险预警机制，通过识别市场环境、项目管理和工程技术三大维度的风险因素并分类，结合成本偏差率、价格波动指数、进度绩效指数和质量缺陷率等关键指标，利用大数据分析和 SVM 模型实现了造价风险的有效预警。研究表明，该机制能够精准识别风险，模型准确率和稳定性较高，并可通过动态调整阈值适应不同项目阶段和市场环境。未来可进一步优化模型并拓展应用范围，为建筑行业造价管理提供更有力的支持。

参考文献：

- [1] 董娜,卢泗化,熊峰.大数据背景下基于ABC-SVM的建筑工程造价预测[J].技术经济,2021,40(08):25-32.
- [2] 金海江.建筑工程造价咨询中的风险管理策略探究[J].中国科技期刊数据库工业A,2025(01):137-140.
- [3] 李青,许冬冬.建筑工程造价动态管理分析[J].新材料·新装饰,2025,07(05):191-194.
- [4] 邱静静.基于大数据分析的建筑工程项目风险管理优化[J].计算机产品与流通,2024(11):69-71.
- [5] 王红艳.基于大数据分析的建筑工程造价优化策略[J].中文科技期刊数据库(引文版)工程技术,2024(12):165-168.