

基于大数据的燃气管道工程造价预测模型

李 艳（临沂中石油昆仑能源有限公司，山东 临沂 276000）

摘要：随着城市化进程的加快和能源需求的不断增长，燃气管道工程的建设规模日益扩大，工程造价管理的复杂性和重要性也日益凸显。传统的工程造价方法往往依赖于人工经验和历史数据，难以准确预测和应对复杂多变的市场环境和项目需求。大数据技术的兴起为工程造价管理提供了新的解决方案，通过深度挖掘和分析历史工程造价数据，大数据技术能够揭示数据背后的隐藏规律和趋势，提高工程造价的预测精度和管理效率。本文旨在构建基于大数据的燃气管道工程造价预测模型，以提高造价预测的精度和可靠性。

关键词：大数据；燃气管道工程；工程造价；预测模型

0 引言

基于大数据的燃气管道工程造价预测模型研究，旨在利用大数据技术优化工程造价管理流程，提高预测准确性和效率，降低投资风险。这一研究不仅能够提升燃气管道工程的经济效益和社会效益，还能够推动工程造价管理的智能化、精细化发展。通过构建高效、准确的预测模型，可以为决策者提供有力的数据支持，促进工程项目的科学决策和顺利实施。

1 燃气管道工程造价的构成

1.1 直接成本

燃气管道工程的直接成本是构成工程造价的基础部分，涵盖了多个关键方面。材料成本占据重要比例，燃气管道所使用的管材、管件等材料，其质量、规格和市场价格波动都会直接影响造价。不同材质的管道，如钢管、PE 管，因性能、使用寿命不同，价格差异明显。优质的管材虽成本高，但能保障工程质量与安全，减少后期维护成本。人工成本也是直接成本的关键组成。从管道的铺设、焊接到设备安装，都需要专业技术工人。人工成本受地区经济水平、劳动力市场供需关系以及施工难度等因素影响。复杂地形或特殊施工环境下，人工投入大、施工效率低，人工成本相应增加。设备成本同样不可忽视。施工过程中需使用多种机械设备，如挖掘机、起重机、焊接设备等。设备的购置、租赁费用以及使用过程中的燃料消耗、维修保养费用，都会计入工程造价。先进的设备有助于提高施工效率和质量，但前期投入较大，在成本核算时需综合考量。

1.2 间接成本

间接成本在燃气管道工程造价中同样占据重要地位，管理费用是其中的重要一项，它包含了工程建设过程中各级管理人员的薪酬、办公费用、差旅费等。高效的项目管理团队能保障工程顺利推进，但相应的

管理成本也不容忽视。合理规划管理架构、提高管理效率，有助于在保证工程质量的前提下控制管理费用。规费是间接成本的另一关键部分，规费是按照国家法律法规规定必须缴纳的费用，如社会保险费、住房公积金等。这些费用是保障施工人员权益、维护社会稳定的主要支出，严格按照规定比例计取，是工程造价的法定组成部分。此外，间接成本还包括不可预见费用，燃气管道工程建设过程中，可能会遇到各种不可预见的情况，如恶劣天气、地质条件变化等。这些突发情况可能导致工程进度延误、施工方案调整，从而增加额外费用。预留一定比例的不可预见费用，能有效应对工程中的不确定性，确保项目顺利完成，这部分费用也需合理估算并纳入工程造价。

2 基于大数据的工程造价预测模型

2.1 数据驱动的全面性优势

基于大数据的工程造价预测模型借助海量数据，实现对各类影响因素的全面考量。传统预测方法往往因数据有限，只能关注部分关键因素，而大数据涵盖工程历史数据、市场动态信息、地理环境资料等多源数据。这些丰富的数据维度，使得模型能够捕捉到众多细微但可能对造价产生重大影响的因素。例如，通过分析不同年份、不同区域类似项目数据，可发现季节变化对材料运输成本的影响，以及当地政策法规调整对人工费用的作用。这种全面性极大提升了预测的准确性，为项目预算提供更可靠依据，减少因考虑不周导致的造价偏差。

2.2 强大的适应性与动态性

大数据环境下的预测模型具备强大的适应性与动态性，随着数据不断更新与积累，模型能够自动学习新数据中的规律和趋势。市场材料价格波动频繁、新技术新工艺不断涌现，传统模型难以及时适应这些变

化，而大数据模型可实时整合最新数据，调整预测参数。比如，当新型管材投入市场，其价格和性能数据被纳入模型后，模型能迅速反映出对工程造价的影响。同时，在项目实施过程中，模型可根据实际发生的数据进行动态调整，持续优化预测结果，更好地满足项目全生命周期造价管理需求。

2.3 挖掘潜在关系与精准预测

大数据技术能够挖掘数据间潜在的复杂关系，这是实现精准预测的关键。在工程造价领域，各因素间并非简单线性关系。通过大数据分析算法和模型可以发现看似不相关因素之间的内在联系。例如，当地经济发展水平与居民对燃气品质要求的提升，会间接影响管道材质和施工工艺选择，进而影响造价。传统方法难以察觉这类隐藏关联，基于大数据的模型通过深度挖掘这些潜在关系，构建更精准的预测模型。利用机器学习算法对海量历史数据进行训练和优化，不断提高预测的精准度，为项目决策提供科学、精确的造价预测支持。

3 数据收集与预处理

3.1 数据收集的多渠道策略

为构建有效的基于大数据的燃气管道工程造价预测模型，需从多个渠道广泛收集数据。历史项目资料是重要的数据来源，涵盖已完成燃气管道工程的设计图纸、施工记录、招投标文件及结算报告等。这些资料详细地记录了项目的各项参数与成本信息，能反映不同条件下的造价情况。市场信息平台也是关键渠道，实时提供材料价格、人工费用等动态数据，帮助了解市场价格波动趋势。此外，地理信息系统（GIS）数据不可或缺，包含地形地貌、地质条件等信息，这些因素会显著影响施工难度与成本。通过多渠道收集数据，确保数据的全面性与丰富性，为后续的分析和建模提供坚实基础。

3.2 数据收集方法的多样性

在数据收集过程中，采用多种方法以获取高质量数据。对于结构化数据，如企业内部数据库中的项目资料，可通过数据库查询与接口调用的方式直接获取，确保数据的准确性与完整性。网络爬虫技术则用于收集互联网上公开的市场价格信息、行业动态等非结构化数据，通过编写特定程序，按照规则自动抓取相关数据。实地调研与问卷调查针对一些难以通过网络获取的数据，如施工现场的实际情况、施工人员对成本影响因素的看法等。这种直接与实际参与者交流的方式，能获取一手信息，弥补其他方法的不足。多种收

集方法相互补充，保障数据来源的广泛与可靠。

3.3 数据预处理的关键步骤

收集到的数据往往存在质量问题，需要进行预处理。数据清洗是首要步骤，旨在去除重复数据，避免对模型训练产生干扰；处理缺失值，可根据数据特点采用均值、中位数填充或模型预测等方法；识别并修正异常值，防止其对整体数据分布产生误导。数据转换通过标准化和归一化操作，将不同量级和分布的数据统一到特定范围，提升模型训练效率和稳定性。特征工程则是从原始数据中提取和选择有代表性的特征，去除冗余信息，降低数据维度。例如，通过相关性分析找出与工程造价密切相关的因素作为关键特征。经过这些预处理步骤，数据质量得到提升，更适合后续的模型构建与分析。

4 基于大数据的燃气管道工程造价预测模型构建

4.1 模型类型选择

构建基于大数据的燃气管道工程造价预测模型，模型类型的选择至关重要，诸多模型各有特点与适用场景，需综合考量。神经网络模型，如多层次感知机（MLP），具有强大的非线性映射能力，能自动学习数据中的复杂关系，对于燃气管道工程造价这种受众多复杂因素影响的问题有较好的适应性。决策树模型则以直观的树状结构进行决策分析，可解释性强，能清晰展示各因素对造价的影响路径。支持向量机（SVM）在处理小样本、非线性及高维数据时表现出色，能有效避免过拟合问题。对比这些模型，结合燃气管道工程数据特点，若数据量丰富且注重模型精度与复杂关系捕捉，神经网络模型可能更合适；若追求模型的可解释性与快速决策分析，决策树模型有优势；而数据量有限时，SVM 或许是不错的选择，最终需通过实验评估确定最适宜的模型。

4.2 模型训练流程

模型训练是构建准确预测模型不可或缺的核心环节，数据集在经过预处理后，会被按照一定的比例细致地划分为训练集、验证集和测试集。训练集承担着模型参数学习与调整的重任，它是模型获取知识的源泉；验证集则在训练的各个阶段对模型进行严格的评估，其作用是监控模型状态，有效防止过拟合现象的出现；而测试集，则作为最终衡量模型性能的标尺。在确定了模型类型之后，接下来便是设置模型的初始参数，比如对于神经网络模型而言，需要明确网络的层数、每一层中神经元的数量以及所采用的激活函数等关键要素。随后，利用训练集中的数据对模型展开

迭代训练，这一过程中，会借助优化算法，如随机梯度下降法，不断地对模型参数进行调优，使得模型在训练集上的损失函数值能够逐步降低。在训练推进的过程中，会定期利用验证集来评估模型的性能表现，一旦发现验证集上的性能指标不再有所提升，甚至出现下降的趋势，这便是一个信号，表明模型可能出现了过拟合，此时就需要对模型的结构或者参数进行相应的调整。就这样，不断地进行训练与调整，直到模型在验证集上的表现趋于稳定。

4.3 模型评估指标

为准确判断基于大数据的燃气管道工程造价预测模型的有效性，需借助一系列评估指标。均方误差（MSE）衡量预测值与真实值之间误差的平方的平均值，能直观反映模型预测结果的整体偏差程度，MSE值越小，说明模型预测越准确。平均绝对误差（MAE）计算预测值与真实值误差的绝对值的平均值，相比MSE，MAE对误差的反应更直接，能体现预测结果的平均误差大小。决定系数（R²）反映模型对数据的拟合优度，取值范围在0到1之间，R²越接近1，表明模型对数据的解释能力越强，预测效果越好。此外，还可采用平均绝对百分比误差（MAPE），它以相对误差的形式评估模型，能更好地反映不同规模数据下的预测精度。通过综合运用这些评估指标，全面、准确地评估模型性能，为模型的优化与改进提供依据。

5 模型的优化与改进策略

5.1 数据层面优化

数据是模型的基础，从数据层面优化能显著提升模型性能。一方面，增加数据量是关键。积极拓展数据来源，与更多燃气管道工程企业、相关行业机构合作，获取大量不同地区、不同规模、不同施工条件下的项目数据。丰富的数据能让模型学习到更全面的特征和规律，增强泛化能力。另一方面，提升数据质量同样重要。加强数据收集过程中的质量把控，对收集到的数据进行多次核对与验证。进一步完善数据清洗流程，采用更先进的算法和技术处理缺失值、异常值，确保数据的准确性和完整性。此外，通过数据增强技术，如对部分数据进行变换、扩展等操作，增加数据的多样性，为模型训练提供更优质的数据支持，从而提升模型对各种情况的适应能力。

5.2 算法与模型结构调整

对算法和模型结构进行合理调整是优化模型的重要方向，深入研究不同大数据模型算法，结合燃气管道工程特点，探索更适合的算法或对现有算法进行改进。例

如，对于选定的深度学习模型，尝试不同的优化器，如Adam、Adagrad等，找到能使模型更快收敛且达到更好性能的优化算法。同时，对模型结构进行微调。增加或减少模型的层数、神经元数量，通过实验观察模型性能变化，找到最优的结构组合。还可以引入注意力机制、残差连接等技术，增强模型对关键信息的捕捉能力，解决梯度消失等问题，使模型能够更好地学习数据中的复杂关系，提高预测的准确性和稳定性。

5.3 融合多源信息与模型集成

融合多源信息并采用模型集成策略能有效提升模型的优化效果，在燃气管道工程造价预测中除了传统的工程数据，还可融合地理信息、气象数据、政策法规等多源信息。这些不同类型的数据从多个角度反映了影响造价的因素，通过合适的方法将其整合到模型中，能丰富模型的特征维度，提升预测能力。另外，采用模型集成技术，将多个不同的预测模型进行组合。例如，将神经网络模型、决策树模型和支持向量机模型的预测结果进行融合，通过加权平均、投票等方式得到最终预测结果。这种集成方式可以综合各个模型的优势，弥补单一模型的局限性，提高模型的鲁棒性和泛化能力，从而实现对燃气管道工程造价更精准的预测。

6 结束语

综上所述，基于大数据的燃气管道工程造价预测模型展现出显著优势。通过多渠道数据收集、科学的预处理及合理的模型构建与优化，能够有效提升预测的精准度与可靠性。这一模型为燃气管道工程的成本控制与决策提供了有力支持，助力行业在项目规划与执行中更科学地管理造价，推动燃气管道工程领域的持续健康发展。

参考文献：

- [1] 贺诗雨.燃气管道改造工程造价控制探讨[J].中国石油和化工标准与质量,2024,44(16):75-77.
- [2] 单克,杨光,孟伟,等.智慧燃气管道完整性管理实施内容与体系构建[C]//中国土木工程学会燃气分会,2021:10.
- [3] 张志猛.燃气工程造价中的成本控制策略分析[J].住宅与房地产,2019,(31):54.
- [4] 陈毓飞.燃气管道风险预测方法的研究与实现[D].北京邮电大学,2019.
- [5] 张玥.城市燃气管道项目工程造价控制的策略研究[J].绿色环保建材,2018,(11):175+178.
- [6] 张永明.大数据思维模式下的地下燃气管道风险等级评估[J].城市燃气,2016,(03):30-36.