Journal of Engineering, Project, and Production Management
2026, 16(5), 2025-355

Predicting Cost-Benefit Trade-Off Path in Urban Rail Transit Construction Schemes Using a LightGBM Model

Dewen Hou
Associate Professor, School of Rail Transportation, Hunan Railway Professional Technology College, ZhuZhou, 412001, China, E-mail: hnrpcdan@sina.com (corresponding author).

Project Management
Received December 23, 2025; revised February 25, 2026; accepted March 9, 2026
Available online June 17, 2026

Abstract: Urban rail transit projects are characterized by large investment scales, long industrial chains, and diversified benefits, and their cost-benefit trade-offs directly affect project sustainability. Based on cost-benefit theory and machine learning methods, this paper first clarifies the research questions of quantifying cost-benefit trade-off relationships and predicting dynamic trade-off paths for urban rail transit projects, and defines the decision-making connotation of the trade-off path as the dynamic change trend of cost-benefit balance under the joint action of multi-dimensional influencing factors. Then, it constructs a cost-benefit trade-off prediction framework integrated with the LightGBM model. Using panel data of 32 newly-built rail transit projects in 15 first-and second-tier cities (Chengdu, Wuhan, Nanjing, Beijing, Shanghai, Guangzhou, Shenzhen, etc.) in China from 2018 to 2023, including detailed construction cost accounting, operational benefit statistics and urban macroeconomic data, 12 characteristic indicators were identified from three dimensions: construction costs, financial benefits, and national economic benefits. Model parameters were optimized through grid search to achieve an accurate prediction of cost-benefit trade-off paths. The results show that the LightGBM model's coefficient of determination (R^2) is 0.892, and its Mean Squared Error (MSE) is 0.037, both of which are significantly better than those of traditional regression models. The added value benefits passenger flow density, unit cost, and land are the core factors affecting the trade-off path, with their weight accounting for more than 45%; based on the model prediction, a trade-off optimization strategy of “gradient cost control and diversified benefit improvement” is proposed, which is applicable to rail transit projects in different tier cities in China under the current urban development stage and can provide a scientific basis for decision-making in rail transit projects.

Keywords: Urban rail transit, lightgbm model, cost-benefit trade-off, path prediction, parameter optimization.

DOI 10.32738/JEPPM-2025-355

1. Introduction

1.1. Background

Urban rail transit is a high-capacity urban passenger transportation mode characterized by high speed and efficiency, safety and comfort, energy conservation, and environmental protection. Vigorously developing rail transit is widely recognized as an effective approach to addressing issues such as traffic congestion, environmental pollution, and land resource scarcity (Tan et al., 2025). At present, China’s urban rail transit is in a phase of large-scale and rapid development. Globally, China has built the largest rail transit network. Zhou (2014) and Guo (2022) found that in first-tier cities such as Beijing, Shanghai, Guangzhou, and Shenzhen, rail transit has already formed a certain scale, which has actively alleviated traffic congestion in these megacities. However, at the same time, urban rail transit construction has high requirements for technical equipment levels and involves complex structures, while its engineering construction environment is highly challenging and carries a certain degree of risk. Therefore, rail transit construction requires advanced development planning and substantial capital investment, which leads to a certain lag in achieving profitability (He et al., 2024; Wang, 2017).

People expect accelerated urban rail transit construction, as it promotes urban development and improves people’s quality of life. Nevertheless, everything has two sides: the huge construction costs, complex system structure, and the particularity of benefits have placed operators of urban rail transit in an awkward position (Qi, 2014; Wang, 2018). Globally, except for the profitable metro systems in Hong Kong (China) and Singapore, the operations of almost all other urban rail transit systems are in deficit, and they rely on government subsidies and other sources to maintain operations (Luo and Feng, 2021). Therefore, it is necessary to examine and study the comprehensive benefits and evaluation methods of urban rail transit.

Light Gradient Boosting Machine (LightGBM) is an efficient tree-based learning algorithm, which is particularly effective in processing large-scale data, with fast operation speed and low memory consumption (Ke, 2017). Tang (2021) and Quinto (2020) found that compared with traditional Gradient Boosting Decision Trees (GBDT), LightGBM is more efficient in handling high-dimensional data and can process datasets containing a large number of features, making it highly suitable for high-frequency data analysis. In addition, several key features of LightGBM significantly improve prediction accuracy. It uses a histogram-based splitting algorithm to optimize the decision tree construction process, which significantly enhances computational efficiency (Tang, 2021). Meanwhile, LightGBM supports parallel and distributed learning, further improving performance when processing large datasets (Luo, 2022).

Targeting the characteristics of high-dimensional sparse data, LightGBM introduces the Exclusive Feature Bundling (EFB) technique. This technique can identify and bundle highly correlated, mutually exclusive sparse features, treating them as a single unit for model training. Ke et al. (2014) found that this effectively compresses the feature space and reduces storage requirements and computational complexity. Furthermore, in terms of the decision tree growth strategy, LightGBM adopts a more flexible and efficient Leaf-wise growth method, which tends to prioritize expanding the currently most profitable leaf nodes. This not only accelerates the model training speed but also helps prevent overfitting to a certain extent (Liu et al., 2020). Finally, to meet the demands of the big data era, Yuan et al. (2014) and Li (2020) found that LightGBM has further strengthened its parallel learning capabilities. By optimizing strategies for data parallelism, feature parallelism, and voting parallelism, LightGBM can fully utilize the multi-core performance of modern computing hardware, significantly improving the model training speed in a distributed computing environment.

1.2. Research Status

The “cost-benefit analysis method” has long been a common approach for scholars to study economic benefits. First proposed by the Americans in 1936 when formulating flood management measures, this method was gradually applied to project planning and management in the field of urban rail transit as its application became increasingly widespread. However, the United States Government Accountability Office (1999) indicated that during the application process, there has never been a unified quantitative model for costs and benefits. It was not until 1950 that a unified methodological system was established (Lu, 2012). Owing to the advantages of cost-benefit analysis in evaluating the economic benefits of projects, especially large-scale public infrastructure, it has been recognized by scholars researching rail transit benefits. After continuous revision and improvement, it was eventually successfully applied in the field of rail transit. Massiani (2014) extracted the benefit of travel time savings solely from the economic benefits of urban rail transit and conducted in-depth research from a microeconomic perspective. In the research process, he analyzed the application of the “cost-benefit” model in the economic benefits of rail transit and the impact of travel time savings on residents and society. Graham et al. (2003) collected relevant data on urban rail transit systems in 17 cities worldwide and used cost-benefit analysis to explore in depth how the scale and density of rail transit affect its economic benefits. They concluded that the growth of rail transit scale can drive steady growth in economic benefits, while rail transit density only affects the magnitude of economic benefits and does not exhibit a stable proportional growth relationship like that of scale. Gwee, Currie, and other researchers studied the application of the cost-benefit analysis method in rail transit projects across 11 countries, compared the differences among these applications, and analyzed the causes and influencing factors of these differences.

Compared with foreign research, China’s rail transit research began relatively late. In 1997, China issued the Measures for Economic Evaluation of Railway Construction Projects (Planning Department of the Ministry of Railways, 1997). Currently, China has not yet established a complete set of its own evaluation methodology system for assessing the benefits of urban rail transit. The benefit evaluation of most related projects refers to or draws on existing research results in this field. Zhou and Lu (2012) conducted in-depth research on the national economic benefit indicators of urban rail transit, provided quantitative models for some indicators, and incorporated the social cost accounting methods commonly used in international transportation project evaluation into the calculation of rail transit economic benefits, providing a reference for other scholars studying the economic benefit indicators of rail transit. Sang (2009) proposed that the social and economic benefits of urban rail transit, which are difficult to measure quantitatively, can be analyzed through a combination of quantitative and qualitative methods, which enables a reasonable estimation of the benefits of rail transit to the national economy. Although some scholars have quantified certain benefits in rail transit systems and established quantitative method models in current research, the evaluation characteristics of these research results all show certain limitations, and their applicability in cost-benefit trade-off evaluation still needs to be improved and verified.

Given the advantages of machine learning in data processing capabilities and prediction accuracy, it has also been widely applied in the field of rail transit. The LightGBM model has already been studied in the rail transit field. Wang (2018) used the LightGBM model to predict vegetable prices, and the results showed that its accuracy was higher than that of other models. Wang (2017) built a model using Long Short-Term Memory (LSTM) to predict short-term passenger flow, and the results indicated that the prediction accuracy could reach 93%. Ma et al. (2015) constructed a model using the LSTM algorithm, conducted simulation experiments by inputting historical passenger flow data of a certain section in Xi’an, and compared it with other models. The results showed that LSTM predictions were more accurate.

This paper sets two core research questions: (1) How to construct a multi-dimensional indicator system to quantitatively characterize the cost-benefit trade-off of urban rail transit construction projects? (2) How to build a high-precision prediction model to capture the non-linear relationship between multi-dimensional factors and cost-benefit trade-offs, and predict the dynamic trade-off path? The trade-off path in this study is defined as the dynamic change trend of the cost-benefit trade-off coefficient with the adjustment of core influencing factors (e.g., passenger flow density, unit cost), which reflects the direction and extent of the change in cost-benefit balance, and its decision-making meaning lies in providing a clear adjustment direction for project decision-makers to optimize the cost-benefit structure. This paper aims to construct a cost-benefit trade-off system for urban rail transit and a calculation model for indicators, so as to provide some decision-making references for decision-makers of rail transit enterprises. Over time, it can also provide ideas and methods for maximizing benefits for operators in the later stage, which has guiding value and provides a theoretical reference and practical implementation significance for urban rail transit construction. The research procedure is shown in Fig. 1.

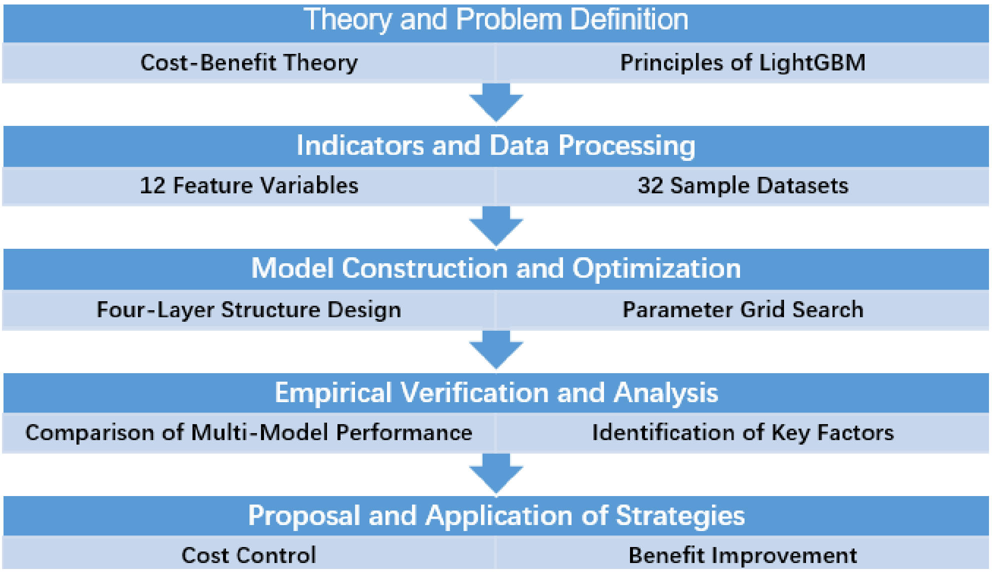

Fig. 1. Research procedure

2. Theoretical Basis and Construction of Indicator System

2.1. Principles of the LightGBM Algorithm

Light Gradient Boosting Machine (LightGBM) is a decision tree-based training algorithm, featuring fast training speed, low memory requirements, and high prediction accuracy. The LightGBM algorithm model overcomes the challenge of balancing high accuracy and high efficiency when dealing with large-sample, high-dimensional data through the Gradient-based One-Side Sampling (GOSS) algorithm, Exclusive Feature Bundling (EFB) algorithm, and histogram algorithm. It predicts using an ensemble of Gradient Boosting Decision Trees (GBDT), with core optimizations including discretizing continuous features using a histogram algorithm to reduce computational complexity. Constructing decision trees using a leaf-wise growth strategy to improve model accuracy. Introducing gradient-based one-sided sampling and exclusive feature bundling techniques to handle high-dimensional sparse data.

The objective function of the algorithm is shown in Eq. (1).

$$Obj(\theta) = \sum_{i=1}^{n} L\left(y_i, \hat{y}_i^{(t)}\right) + \sum_{k=1}^{K} \Omega\left(f_k\right) \tag{1}$$

Eq. (1) is the core objective function of the LightGBM algorithm, which consists of two parts: the loss function $L(y_i, \hat{y}_i^{(t)})$ and the regularization term $\Omega(f_k)$. The loss function measures the prediction error between the actual value y_i and the t-th iterative prediction value $\hat{y}_i^{(t)}$, and the regularization term is used to control the complexity of the k-th decision tree f_k to avoid overfitting. The algorithm achieves accurate prediction of the target variable by minimizing the objective function through continuous iterative optimization of decision tree parameters, which is the mathematical foundation for the LightGBM model to fit the cost-benefit trade-off relationship of urban rail transit projects.

Among them, $L(y_i, \hat{y}_i^{(t)})$ represents the loss function, $\Omega(f_k)$ denotes the regularization term, and K is the number of decision trees. Accurate prediction is achieved by minimizing the objective function through iterative optimization.

In machine learning, many algorithms require converting categorical features into multi-dimensional one-hot-encoded features. However, the LightGBM model does not need additional one-hot encoding, and all categorical features can be directly input into the model. Meanwhile, this algorithm model supports feature parallelism and data parallelism, which improves the operational efficiency of the algorithm, an advantage not possessed by many other machine learning algorithm models. In terms of feature processing, LightGBM abandons the traditional strategy of sorting features sample by sample, and instead adopts a histogram approximation algorithm. By aggregating continuous feature values into a fixed number of discrete bins, this algorithm reduces memory usage and shortens computation time. Specifically, the histogram algorithm divides the feature space into multiple discrete intervals and accumulates the gradient information for each interval, thereby avoiding the need to re-sort all data in each iteration. This improvement not only enhances computational efficiency but also significantly reduces the complexity of model training by decreasing the number of candidate points during data splitting.

The flow chart of the LightGBM algorithm is shown in Fig. 2.

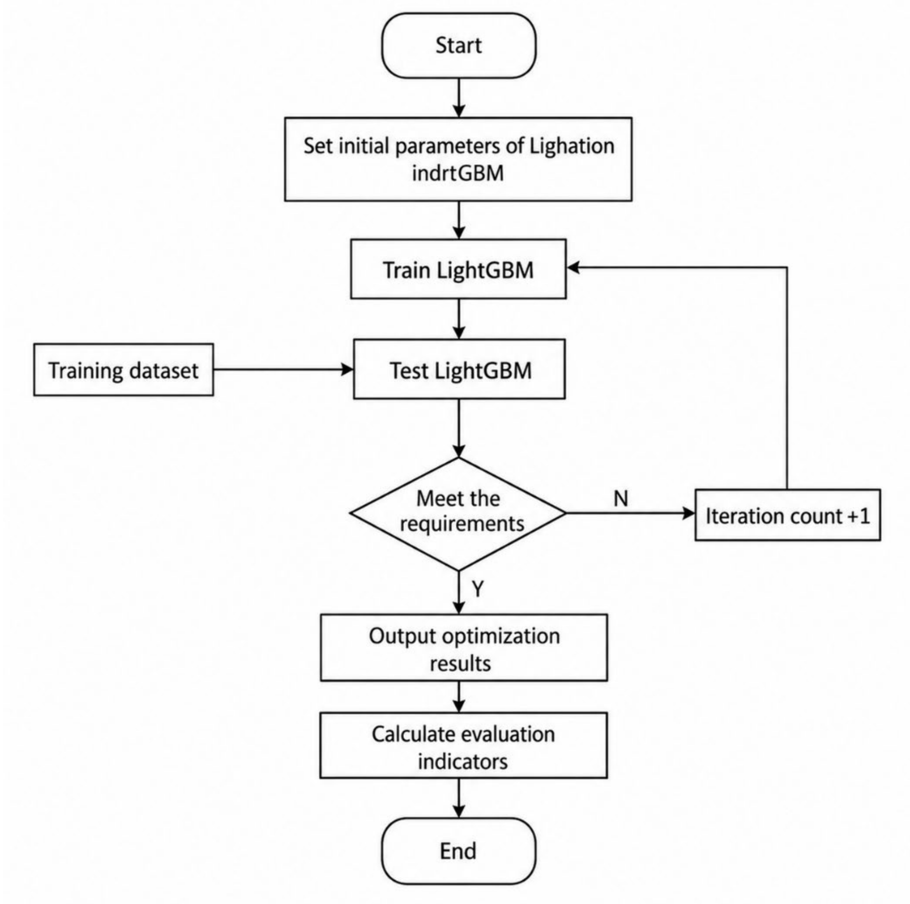

Fig. 2. LightGBM algorithm

2.2. Development of Cost-Benefit Indicator System

The principles of indicator design are shown as follows. Adhere to the systematic principle: cover all aspects of costs and benefits. Adhere to the quantifiable principle: prioritize indicators with monetizable value or standardized metrics. Adhere to the relevance principle: screen and evaluate indicators with strong correlation using the Pearson correlation coefficient. Feature Variables are shown in Table 1.

The cost-benefit trade-off coefficient is selected as the target variable, and its calculation formula is shown in Eq. (2).

$$T = \frac{\sum_{j=1}^{m} W_j B_j}{\sum_{i=1}^{n} C_i} \tag{2}$$

Eq. (2) is the calculation formula of the cost-benefit trade-off coefficient T, which is the core target variable of this study to measure the cost-benefit balance of urban rail transit projects. The numerator is the weighted sum of m benefit indicators B_j (weight W_j), which reflects the comprehensive benefits of the project considering the different importance of each benefit; the denominator is the sum of n cost indicators C_i, which reflects the total construction and operation costs of the project. When T>1, it means that the comprehensive benefits of the project are greater than the total costs, and a larger T value indicates a better cost-benefit trade-off effect. This formula quantifies the multi-dimensional cost-benefit relationship into a single coefficient, which provides a clear quantitative standard for the prediction of the trade-off path.

Wherein, T is the trade-off coefficient (T > 1 indicates that benefits are greater than costs, and a larger T means a better trade-off effect). B_j represents the j-th benefit indicator. W_j is the weight of the benefit indicator, and C_i denotes the i-th cost indicator. For reference, the benefit indicators include Net Present Value (NPV), time-saving cost, land value-added benefit, and environmental benefit. The entropy weight method is used to determine their weights as 0.35, 0.25, 0.30, and 0.10, respectively.

3. Model Development and Data Processing

The LightGBM Prediction Model Structure is shown in Table 2. Grid search combined with 5-fold cross-validation is used to optimize hyperparameters. The optimization range and optimal results are shown in Table 3.

Table 1. Feature variables

Dimension	Feature variables	Data source and calculation method
Construction Cost	Unit cost (RMB Yuan/km), Proportion of civil engineering (%), Land acquisition and resettlement cost (100 million RMB Yuan), Mechanical and electrical installation cost (100 million RMB Yuan)	Project bid documents, construction cost accounting reports, and local transportation department statistical yearbooks; unit cost = total construction cost / route length
Project Characteristics	Route length (km), Number of stations (unit), Degree of automation (Level 1-5), Route type (Metro = 1, Light rail = 2, Suburban railway = 3)	Project feasibility study reports, operation manuals; automation degree is classified according to Chinese urban rail transit operation standards
External Environment	Urban GDP (100 billion RMB Yuan), Population size (1 million people), Passenger flow density (10,000 passenger trips/(km·day)), Average land price (10,000 RMB Yuan/sq.m)	National Bureau of Statistics, local real estate transaction centers, rail transit operation companies; passenger flow density = daily passenger flow / route length

Table 2. Structure of the lightgbm prediction model

Model layer	Description of core functions
Data Input Layer	Receives 12 preprocessed feature variables (including variables related to construction cost, project characteristics, and external environment dimensions)
Feature Processing Layer	Converts input features through histogram discretization
Core Prediction Layer	Constructs 100 decision trees and adopts the Least Squares Regression loss function
Result Output Layer	Outputs the predicted value of the cost-benefit trade-off coefficient

3.1. Data Sources

Fifteen first- and second-tier cities, including Chengdu, Wuhan, Nanjing, Beijing, Shanghai, Guangzhou, Shenzhen, Chongqing, Tianjin, Hangzhou, Suzhou, Xi’an, Changsha, Wuxi, and Qingdao, were selected as the research objects. The research sample is 32 newly-built urban rail transit projects (including metro, light rail, and suburban railway) that were completed and put into operation in these cities from 2018 to 2023. The data sources include three aspects: (1) Official statistical data: National Bureau of Statistics, local transportation bureaus, and real estate transaction centers. (2) Project-specific data. Feasibility study reports, bid documents, construction cost accounting, and operation statistics of rail transit projects. (3) Enterprise operation data: Daily passenger flow, ticket revenue, and non-ticket revenue released by urban rail transit operation companies. All data have been verified for authenticity and consistency through cross-checking multiple data sources.

3.2. Data preprocessing

The data preprocessing procedure is explained below.

- Missing value handling: The K-nearest neighbor (KNN) algorithm was used to impute missing data, such as land acquisition and resettlement costs (missing rate < 5%).
- Outlier handling: The 3σ criterion was adopted to identify and eliminate 2 samples with abnormal unit costs.

- Feature standardization: For continuous variables, Z-score standardization was applied, with the formula: $x' = (x - \mu)/\sigma$; for categorical variables (route type), one-hot encoding was used for conversion.

Table 3. Optimal parameters of the LightGBM model

Parameter Name	Optimization Range	Optimal Value	Function Description
Learning rate (learning_rate)	0.01 – 0.3	0.08	Controls the contribution of each tree
Tree depth (max_depth)	3 – 10	6	Limits the complexity of decision trees
Minimum child samples (min_child_samples)	10 – 50	25	Prevents overfitting
Number of decision trees (n_estimators)	50 – 200	120	Total number of trees in ensemble learning
Subsample ratio (subsample)	0.6 – 1.0	0.8	Improves generalization through random sampling

3.3. Data partitioning

Samples were divided into a training set (22 samples) and a test set (10 samples) at a ratio of 7:3. The training set was used for model fitting, and the test set was used for performance verification. The basic statistical characteristics of the dataset are shown in Table 4.

4. Empirical Analysis and Result Verification

Three indicators, i.e., coefficient of determination (R^2), Mean Squared Error (MSE), and Mean Absolute Error (MAE), were used to compare the prediction performance of the LightGBM model with that of the Multiple Linear Regression (MLR) and Random Forest (RF) models. The results are shown in Table 5. The table shows that the LightGBM model performs optimally across all three evaluation indicators. This indicates that LightGBM can more accurately capture the non-linear relationship between costs and benefits, meeting the needs of trade-off path prediction.

Through the feature importance evaluation function of the LightGBM model, the influence weights of each input indicator on the trade-off coefficient are calculated, as shown in Table 6.

Three typical projects in the test set (metro in first-tier cities, light rail in second-tier cities, and suburban railways) were selected for trade-off path prediction, and the dynamic changes between actual values and predicted values were compared (Table 7). As shown in Table 7, the predicted results are highly consistent with the actual operational status of the projects, verifying the model's applicability across different scenarios in Chinese first- and second-tier cities.

5. Cost-Benefit Trade-Off Optimization Strategies

Cost control strategies are proposed from two perspectives: construction-phase costs and operational-phase costs (see Table 8). Strategies for improving benefits are formulated from three aspects: revenue structure, indirect benefits, and differentiated trade-offs. Specific measures are shown in Table 9 below.

6. Conclusion

This study employs the LightGBM model to predict the cost-benefit trade-off paths in urban rail transit construction schemes. A cost-benefit evaluation system covering 12 characteristic indicators has been constructed. The trade-off coefficient is defined as the target variable, enabling the quantitative expression of multi-dimensional relationships. The prediction framework built based on the LightGBM model exhibits excellent performance, with an R^2 of 0.892, which can accurately predict trade-off paths under different scenarios. Four core influencing factors, including passenger flow density and unit cost, have been identified, with a cumulative weight of 0.54, providing target points for path optimization. Strategies for gradient cost control and multi-dimensional benefit enhancement are proposed, which can effectively improve the trade-off effect of projects.

Some quantitative assumptions were made. The comprehensive benefits of urban rail transit projects can be quantified as the weighted sum of multiple benefit indicators, and the weights of each benefit indicator are stable within the research time window (determined by the entropy weight method). The non-linear correlation between the 12 input indicators and the cost-benefit trade-off coefficient is stable, and there is no structural break in the correlation during the research period.

All collected data are true and effective, and the missing values and outliers processed by the KNN algorithm and 3σ criterion do not affect the overall law of the dataset.

Table 4. Basic statistical characteristics of the dataset

Indicator Type	Indicator Name	Mean	Standard Deviation	Minimum	Maximum
Construction Cost	Unit Cost (10,000 Yuan/km)	52300	8700	38500	71200
	Land Acquisition and Resettlement Cost (100 Million Yuan)	18.6	7.2	6.3	35.8
Project Characteristics	Route Length (km)	28.5	10.3	12.1	56.7
	Passenger Flow Density (10,000 Passenger Trips/(km·day))	6.2	2.8	2.1	12.5
External Environment	Urban GDP (100 Billion Yuan)	1.5	0.8	0.6	3.2
Target Variable	Trade-off Coefficient T	1.28	0.35	0.76	2.15

Table 5. Prediction performance comparison

Model	R^2	MSE	MAE	Running Time (seconds)
Multiple Linear Regression (MLR)	0.628	0.125	0.287	1.2
Random Forest (RF)	0.795	0.068	0.192	8.7
LightGBM	0.892	0.037	0.135	3.5

6. Conclusion

This study employs the LightGBM model to predict the cost-benefit trade-off paths in urban rail transit construction schemes. A cost-benefit evaluation system covering 12 characteristic indicators has been constructed. The trade-off coefficient is defined as the target variable, enabling the quantitative expression of multi-dimensional relationships. The prediction framework built based on the LightGBM model exhibits excellent performance, with an R^2 of 0.892, which can accurately predict trade-off paths under different scenarios. Four core influencing factors, including passenger flow density and unit cost, have been identified, with a cumulative weight of 0.54, providing target points for path optimization. Strategies for gradient cost control and multi-dimensional benefit enhancement are proposed, which can effectively improve the trade-off effect of projects.

Some quantitative assumptions were made. The comprehensive benefits of urban rail transit projects can be quantified as the weighted sum of multiple benefit indicators, and the weights of each benefit indicator are stable within the research time window (determined by the entropy weight method). The non-linear correlation between the 12 input indicators and the cost-benefit trade-off coefficient is stable, and there is no structural break in the correlation during the research period.

All collected data are true and effective, and the missing values and outliers processed by the KNN algorithm and 3 σ criterion do not affect the overall law of the dataset.

This research is subject to the following limitations due to the above assumptions. The research sample is limited to 32 rail transit projects in 15 Chinese first- and second-tier cities from 2018 to 2023, and the model's applicability to third- and fourth-tier cities and overseas rail transit projects needs to be further verified. The model does not consider the impact of macro policy changes (e.g., urban planning adjustment, transportation subsidy policy), unexpected events (e.g., public health emergencies, natural disasters) and social factors (e.g., resident's travel habits) on the cost-benefit trade-off. The research time window is 2018-2023, and the stability of the model's prediction effect in the long term needs to be further tested with follow-up data.

Future research may expand the sample size and research scope, include third- and fourth-tier cities in China and overseas rail transit projects to improve the generalizability of the model. Add policy, social, and unexpected event factors to the indicator system, and build a dynamic prediction model that considers time-varying factors. Combine the LightGBM model with other machine learning algorithms (e.g., LSTM, XGBoost) to construct an ensemble model and further improve the trade-off path's predictive accuracy.

Table 6. Influence weights of each indicator on the trade-off coefficient

Influencing factor	Weight	Core impact description
Passenger Flow Density	0.21	In first-tier cities, the average daily passenger flow exceeds 10,000 passenger trips per kilometer, which significantly increases ticket revenue and advertising income, serving as the core driver for benefit improvement.
Unit Cost	0.16	Costs of civil engineering and mechanical & electrical equipment account for over 70% of the total cost; standardized construction can reduce the unit cost by 15%-20%.
Land Value-added Benefit	0.09	The average price of land around stations rises by 30%-50%; the "rail + property" model can effectively convert indirect benefits into tangible gains.
Urban GDP	0.08	Economically developed cities have strong consumption capacity, enabling non-ticket revenue to account for more than 30% of the total revenue.

Note: The cumulative weight of the four core factors reaches 0.54, making them the key control points for the cost-benefit-benefit trade-off path.

Table 7. Prediction and analysis of trade-off paths

Project Type	Predicted Trade-off Coefficient	Actual Value	Error	Path Characteristics
Metro Project in First-Tier Cities	1.98	2.15	8.0%	High Passenger Flow - High Land Value Increment - High Trade-off Efficiency
Light Rail Project in Second-Tier Cities	1.02	0.97	5.2%	Medium Cost - Medium Benefit - Marginal Balance
Suburban Railway Project	0.85	0.76	11.8%	High Cost - Low Passenger Flow - Trade-off Imbalance

Table 8. Gradient cost control strategy

Strategy level	Specific measures	Core objectives
Cost Optimization in the Construction Phase	Promote "standardization + modularization" construction and adopt unified station design standards Implement targeted measures for civil engineering based on geological classification, and replace traditional construction methods with TBM (Tunnel Boring Machine) construction in soft soil foundation areas	Reduce survey and design costs by more than 10%, and control the metro unit cost within 500 million Yuan per kilometer
Cost Control in the Operation Phase	Establish a dynamic cost monitoring platform, with a focus on controlling electricity costs and maintenance costs Adopt energy-saving traction systems \Implement predictive maintenance	Reduce unit energy consumption by 30%, lower equipment failure rate by 25%, and increase the operation cost control rate to more than 85%

Table 9. Strategies for enhancing multiple benefits

Strategy Type	Specific Measures	Core Objectives
Revenue Structure Optimization	Implement floating fares during peak hours Develop commercial complexes and advertising resources at stations	Increase the proportion of ticket revenue by 15% and raise non-ticket revenue to 35% of total revenue
Indirect Benefit Conversion	Allocate 40% of land transfer revenue within 500 meters of stations to project subsidies Reinvest land value-added benefits into construction costs	Shorten the investment payback period by 3-5 years
Differentiated Trade-off Strategies	First-tier cities: Increase route density to boost passenger flow density (maximize benefits) Second-tier cities: Prioritize route construction in densely populated areas (cost priority + benefit potential exploration) Suburban railways: Adopt the "PPP + special bonds" financing model	Improve the suburban railway trade-off coefficient to above 1.0 for Chinese first- and second-tier cities, meeting the cost-benefit optimization needs of rail transit projects in different types of cities under the current urban development stage

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of AI Tools

The author used DeepSeek solely for language editing and readability improvement. The author reviewed and verified all content and takes full responsibility for the accuracy and integrity of the manuscript.

References

Chen, W. (2017). Effective Improvement of Urban Traffic Problems by the Development of Metro Rail Transit. *Heilongjiang Science and Technology of Communications*, (10), 188-189. DOI: 10.3969/j.issn.1008-3383.2017.10.112

Graham, D. J., Couto, A., Adeney, W. E., and Glaister, S. (2003). Economies of Scale and Density in Urban Rail Transport: Effects on Productivity. *Transportation Research Part E: Logistics and Transportation Review*, 39(6), 443-458. DOI: 10.1016/S1366-5545(03)00017-6

Guo, H. B. (2022). Effective Improvement of Urban Traffic Problems by the Development of Metro Rail Transit. *Engineering Technology*, (10), 4.

He, J., Zhang, X., Li, Y., and Li, T. (2024). Research on Service Quality Evaluation of Urban Rail Transit Based on Passenger Demand. *Urban Mass Transit*, 27(4), 67-73. DOI:10.16037/j.1007-869x.2024.04.013

Ke, G., Meng, Q., Finley, T., Wang, T., Chen,W., Ma, W.,Ye, Q., and Liu T. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Neural Information Processing System*

Li, Z. X. (2020). Research on Supporting Technologies and Prototype Implementation of Parallel Learning Models for Financial Credit Evaluation. University of Electronic Science and Technology of China. Master's Degree Thesis.

Liu, H., Xiao, Q., Jiao, Z., Meng, J., Mu, Y., Hou, K., Yu, X., Guo, S., and Jia, H. (2020). LightGBM-Based Prediction of Remaining Useful Life for Electric Vehicle Battery under Driving Conditions. *2020 IEEE Sustainable Power and Energy Conference (iSPEC)*, 2577-2582, DOI: 10.1109/iSPEC50848.2020.9351029.

Lu, M. (2012). Research on the Comprehensive Benefits of Urban Rail Transit Systems. *Beijing Jiaotong University*. DOI: 10.7666/d.Y2429428

Luo, J. H. (2022). Application of High-Frequency Data in Cyclical Stock Price Trend Prediction Based on LSTM and LightGBM Ensemble Model. Jiangxi University of Finance and Economics. Master's Degree Thesis.

Luo, T. and Feng, J. (2021). Quantitative Evaluation of Safety and Economic Benefits of Rail Transit System Based on Grey Prediction Model. *Industrial Safety and Environmental Protection*, 47(9), 5. DOI: 10.3969/j.issn.1001-425X.2021.09.013

Ma X., Tao Z., Wang Y., Yu H., and Wang Y. (2015), Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, Transportation Research Part C: Emerging Technologies 54, 187-197. DOI: 10.1016/j.trc.2015.03.014

Massiani, J. (2014). A Micro-Founded Approach to the Valuation of Benefits of Freight Travel Time Savings. *Research in Transportation Economics*, (47), 61-69. DOI: 10.1016/j.retrec.2014.09.019

Planning Department of the Ministry of Railways. (1997). Interim Measures for Material Management of Joint-Venture Railways. Railway Purchase and Logistics, (6).

Qi, H. R. (2014). Research on Investment and Financing Modes of Urban Rail Transit. *Dalian Maritime University.* DOI: 10.7666/d.Y2601253

Quinto, B. (2020). *Next-Generation Machine Learning with Spark: Covers XGBoost*, Apress Berkeley, CA. DOI: 10.1007/978-1-4842-5669-5

Sang, Y. (2009). Analysis and Evaluation of the Social Benefits of Chengdu Metro. *Manager' Journal*, (16), 1.

Tan, D., Chen, K., Wu, D., Li, Y., Hu, S., and Zhang, C. (2025). Research on the Coordinated Development of Supply and Demand between Rail Transit and Land Use from the Perspective of Low Carbon. *Journal of Transportation Systems Engineering and Information Technology*, 25(4), 13-23. DOI: 10.16097/j.cnki.1009-6744.2025.04.002

Tang, C. F. (2021). Research and Implementation of Network Intrusion Detection Algorithm Based on LightGBM and Autoencoder. Xinjiang University. Master's Degree Thesis.

United States. Government Accountability Office, (1999). Mass Transit: Status of New Starts Transit Projects With Full Funding Grant Agreements. Retrieved from https://rosap.ntl.bts.gov/view/dot/33827 on August 19, 1999.

Wang, B. G. (2017). Analysis of the Economic Benefits of Urban Rail Transit. *China International Finance*, (23), 10-12. doi: 10.19516/j.cnki.10-1438/f.2017.23.007

Wang, M. (2017). Research on Short-term Traffic Flow Prediction Method Based on Machine Learning, Chang'an University. Master's Degree Thesis.

Wang, Y. B. (2018). Discussion on Methods for Increasing Economic Benefits in the Urban Rail Transit Industry. *Chinese Science and Technology Journal Database - Industry C*, (4), 00306-00307.

Wang, Z. (2018). Research and Implementation of Price Prediction Model Based on Machine Learning, Chang'an University. Master's Degree Thesis.

Ye, F., Wang, J., Li, Z., Jihan, Z., and Yang, C. (2021). Jane Street Stock Prediction Model Based on LightGBM. *2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 385-388. DOI: 10.1109/ICSP51882.2021.9408851

Yuan, L. M., Du, X. Y., Chen, Q. Z., Li, T. T., Jiang, H., Tang, B. L., Zhang, B. Q., Zhao, M., Qian, H. M., and Lin, C. F. (2022). Parallel Forecasting Method for Underwater 3D Temperature and Salinity Based on LightGBM Model. CN202111261491.1

Zhou, G. and Lu, J. (2012). Identification and Calculation of National Economic Benefits of Urban Rail Transit Projects. *Urban Rail Transit Research*, 15(12), 4. DOI: 10.3969/j.issn.1007-869X.2012.12.005

Dewen Hou is an Associate Professor at the School of Rail Transportation. He received his Postgraduate degree in business administration from Hunan University of Technology in Zhuzhou, Hunan. His research interests include the operation and management of rail transportation. In recent years, he has published several academic papers in related fields and participated in a number of provincial and national scientific research projects.