Intelligent transportation model based on least squares method and linear regression

Yukai Wang, Huiling Zhao, Kaiwei Ren, Bingrui Ding*

Tianjin College, University of Science and Technology Beijing, Tianjin, China

ABSTRACT

This article covers four questions about the development of high-speed rail in China.

For question 1, we use a linear regression model to predict the total mileage of high-speed rail in the future. Based on the time and total high-speed rail mileage data for the past 10 years in the appendix, and using the LinearRegression model in the Sklearn library for fitting and forecasting. By fitting the model, we obtain slope and intercept parameters that describe the relationship with time. To verify the accuracy of the model. Further, we used the trained model to predict the total mileage of high-speed rail in the next 50 years and printed out the results. In addition, we save the prediction results to a CSV file for subsequent analysis and application.

For question 2, we analyzed the relationship between high-speed rail passenger traffic and years from 2008 to 2021 through linear regression models. In the fitting process, we obtain the regression coefficient and intercept, which describe the linear relationship between high-speed rail passenger traffic and the year. By plotting scatterplots and regression straight lines, the fit of the data and the regression model was visualized. The results show that with the growth of time, the passenger volume of high-speed rail shows an increasing trend. By outputting the linear regression equation, we get a specific mathematical expression to describe the relationship between high-speed rail passenger traffic and year.

For question 3, the optimization coefficient is mainly used to construct an efficiency model that comprehensively considers factors such as high-speed rail operating mileage, passenger traffic and passenger turnover. By defining the function of computational efficiency and the optimization objective function, the scipy. optimize. minimize() function is used to optimize and obtain the coefficient of the optimal solution. Finally, the final optimization equation is output, indicating the degree to which various factors affect efficiency. The goal of the entire procedure is to find the best coefficients to maximize the overall efficiency.

Finally, we wrote a report on the future development of China's high-speed railway, covering construction planning, technological innovation, policy support and market demand, and provided prospects and suggestions for future development.

This study is of great significance for understanding the development trend of high-speed rail passenger traffic and predicting future passenger traffic. In addition, the method can also be applied to data analysis and prediction in other fields to provide reference and guidance for decision-making.

Keywords: Linear Regression Models; Linearregression Model; Regression Coefficient; Optimization Factor

1 INTRODUCTION

China high-speed railway refers to the railway system with high speed, large capacity and high service level planned, built and operated by the State Railway Administration of China. Since 2007, China's high-speed railway construction has experienced rapid development,

becoming one of the longest, fastest and most advanced high-speed railway networks in the world.

The construction of China's high-speed railway has not only greatly improved people's travel conditions, but also driven the development of the surrounding economy, becoming an important support for China's economic transformation and modernization [1]. At the same time, China's high-speed railway also occupies an important position in the world, affecting the development direction and trend of global high-speed railway construction [2].

Today, as one of the most developed high-speed railway networks in the world, China's high-speed rail has become an important part of China's modern transportation infrastructure. Indicators such as the mileage of high-speed rail are the basis for evaluating the progress and quality of high-speed rail construction [3]. The speed and passenger volume of high-speed rail can reflect whether a country is a "strong country" in terms of transportation.

Solve the following problems through mathematical modeling:

According to the relevant data of the total mileage of high-speed rail in the past 10 years, analyze the change of total mileage in the next 50 years, and analyze the reasons for this change.

Analyze the relationship between the operating speed of China's high-speed rail and passenger traffic [4].

With the data of China's high-speed rail mileage and passenger traffic known, build a mathematical model to optimize the efficiency of high-speed rail transportation.

Write a report on the future development of China's high-speed railway (including but not limited to the current situation of China's high-speed railway and suggestions for future development, no less than 800 words)

2 PROBLEM ANALYSIS

Analysis of question 1

Analysis of the research significance of question 1:

Question 1 belongs to the prediction problem of the total mileage of high-speed rail, and its research significance is mainly reflected in the following aspects:

Planning and decision support: By predicting the future development trend of the total mileage of high-speed rail, it can provide important reference information for relevant departments and decision-makers, so as to help plan future transportation, infrastructure construction and economic development. This is of great significance for the formulation of the overall national development strategy [5].

Economic and industrial development: As a fast, safe and efficient mode of transportation, the development of high-speed rail has a positive impact on promoting regional economic development, improving transportation efficiency and improving people's travel conditions. Forecasting the total mileage of high-speed rail can help companies and investors make sound decisions and promote the development and investment of related industries [6].

Social life and people's well-being: The development of high-speed rail can shorten the distance between urban and rural areas, improve the convenience and comfort of people's travel, and improve people's quality of life. Forecasting the total mileage of high-speed rail can give people an early understanding of the future development of the high-speed rail network, providing more convenient choices for daily life and travel [7].

Analysis of the characteristics of the data given in the annex:

The total mileage data of high-speed rail from 2008 to 2021 is given in the annex. Through the analysis of these data, the following characteristics can be obtained:

Growth trend: The data shows that the total mileage of high-speed rail has shown a clear growth trend in the past 2008-2021. This may mean that the development of high-speed rail has received attention from the government and has achieved remarkable results in promoting transportation upgrades.

Annual fluctuations: Although the overall growth trend is shown, the data also shows certain annual fluctuations. It may be affected by factors such as economic development, policy adjustments and construction progress, which makes the annual growth rate different.

Analysis of the results required by question 1:

Question 1 requires forecasting the total mileage of high-speed rail over the next 50 years. The forecast results can provide a reference for the future development of high-speed rail and are of great significance to policymakers, enterprises and investors.

By establishing a time series analysis method based on the total mileage data of high-speed rail operation in the past decade, linear regression modeling using the least squares method is used to predict future development based on past trends [8].

Analysis of question 2

The analysis of the research significance of question 2 includes the following aspects:

Practical problem solving: The research of question 2 can help us understand and master the relationship between high-speed rail operation speed and passenger traffic, so as to provide scientific basis for high-speed rail operation management and planning. By establishing mathematical models, future passenger traffic can be predicted to provide guidance for high-speed rail transportation demand and resource allocation.

Improvement of data analysis ability: The analysis of the characteristics of the data presented in the annex is an important step in the research of Question 2. By analyzing the distribution, trends, outliers and other characteristics of the data, the data itself can be better understood and provided as a reference for subsequent modeling and prediction.

Decision support: The research results in Question 2 can provide a basis for high-speed rail related decisions. Through the analysis of the forecast results, we can better understand the impact of high-speed rail operation speed on passenger traffic, and provide scientific support for formulating operational strategies, optimizing resource allocation and forecasting future demand.

Based on the analysis above:

Since no specific data and the specific results required for question 2 are given, further analysis of model selection and interpretation of predicted results is not possible. However, using the fitted mathematical model, future passenger traffic can be estimated based on the predicted operating speed, thus providing a reference for the planning and decision-making of the high-speed rail system [9].

Analysis of question 3

Problem 3 belongs to the problem of high-speed rail transportation efficiency optimization, which needs to be solved by establishing a mathematical model. This problem is solved using the following mathematical analysis:

Linear programming: A linear programming model can be built to optimize the transportation efficiency of high-speed rail. By setting the objective function and constraints, various indicators (such as high-speed rail operating mileage, passenger traffic, passenger turnover, etc.) are taken into account, and the optimal solution is found to improve the efficiency of high-speed rail transportation.

Time series analysis: You can use the time series analysis method to analyze the data over the years and dig out its inherent trends and periodic changes. Based on the trend of historical data, it is possible to predict the future high-speed rail transportation demand, and formulate corresponding transportation plans accordingly, so as to optimize the high-speed rail transportation efficiency.

Regression analysis: A regression model can be established to analyze the relationship between high-speed rail transportation efficiency and various indicators. Through regression analysis, the influence of different indicators on the efficiency of high-speed rail transportation can be determined, and the basis for formulating appropriate optimization strategies can be provided.

For the analysis of the characteristics of the data given in the annex, the following aspects are considered:

Data trend: analyze the change trend of various indicators over the years, understand the overall trend of high-speed rail transportation development, determine its growth or decline trend, and determine possible influencing factors.

Correlation of indicators: study the correlation between various indicators, and explore their internal connections and mutual influences. For example, whether there is a positive correlation between the mileage of high-speed rail operations and passenger traffic and passenger turnover can help find the key elements to optimize efficiency.

Indicator change range: analyze the change range of various indicators to understand their contribution to the efficiency of high-speed rail transportation. By evaluating the magnitude of the movement, it is possible to determine the focus and direction of the optimization strategy.

For the analysis of the results required in question 3, consider the following aspects:

Improvement of high-speed rail transportation efficiency: The established mathematical model should be able to give a set of optimization schemes to improve the efficiency of high-speed rail transportation. This includes increasing the operating mileage of high-speed rail, increasing passenger traffic, improving passenger turnover, and other indicators.

Prediction and planning: Based on the analysis of historical data and mathematical models, the demand for future high-speed rail transportation can be predicted and reasonable planning schemes can be formulated. This helps to allocate resources and manage operations in advance, and supports the sustainable development of the high-speed rail industry.

3 MODEL ASSUMPTIONS

Assume that the data given by the question is true and reliable;

It is assumed that the system is stable for a short period of time, that is, the properties and characteristics of the system remain unchanged for the period of time under investigation.

It is assumed that the behavior of the system is influenced only by the variable factors in the appendix, and other factors can be ignored or left unchanged.

Assume that elements or components of the same type in a system have similar behaviors and characteristics that can be grouped into the same category for analysis and modeling.

Assume that some part or factor in the system becomes a bottleneck in overall performance, limiting the maximum efficiency or capability of the system.

Suppose that the system transitions between different equilibrium states over the period of time under consideration, rather than just tending to a fixed equilibrium point.

4 DEFINITIONS AND SYMBOL DESCRIPTIONS

4.1 Solution of question 1

4.1.1 Source of data

X	Time data for the last 10 years
Y	Total high-speed rail mileage data at the corresponding time
Model	Linear regression model object
future_mileage.csv	Save the CSV file name of the prediction result
slope	Linear regression coefficient, the slope of a regression line
intercept	The intercept of the linear regression model
w1, w2, w3	Weight coefficient
w1_opt, w2_opt, w3_opt	The optimal weight coefficient obtained by optimization

5 MODEL EVALUATION

5.1 Establishment and solution of the model of problem 1

For solving this problem, we can use the linear regression model of least squares method to model and predict the total mileage of high-speed rail.

First, we take time as the independent variable X and the total high-speed rail mileage as the dependent variable Y. We have data for the last 10 years, i.e. $X1$, $X2$, ..., $X10$ and the corresponding Y1, Y2, ..., Y10.

The steps are as follows:

1. Data preparation:

Take the time of the last 10 years as the independent variable X:X = [1, 2, ..., 10].

The total high-speed rail mileage in the last 10 years is used as the dependent variable Y:Y = [Y1, Y2, ..., Y10], here because the data for 2022 is missing [10].

Linear regression model fitting:

Use the least squares method to fit the data. Find the line Y = a*X + b that best fits the data, where a is the slope and b is the intercept.

Model parameter estimation:

Calculate the values for slope a and intercept b to minimize the overall error. It can be calculated using formulas:

$$a = (n\sum(Xi^{*}Yi) - \sum Xi^{*}\sum Yi)/(n^{*}\sum(Xi^{2}) - (\sum(Xi)^{2}) \quad (1)$$

$$b = \frac{\sum Yi - a^{*}\sum Xi}{n} \quad (2)$$

Where n is the sample size and Σ represents the summation sign. Finally, a=3590.28; b=5071.67.

Figure 1 below shows the least squares linear regression model, where the abscissa represents the year (2012-2021) and the ordinate represents the total mileage of high-speed rail.

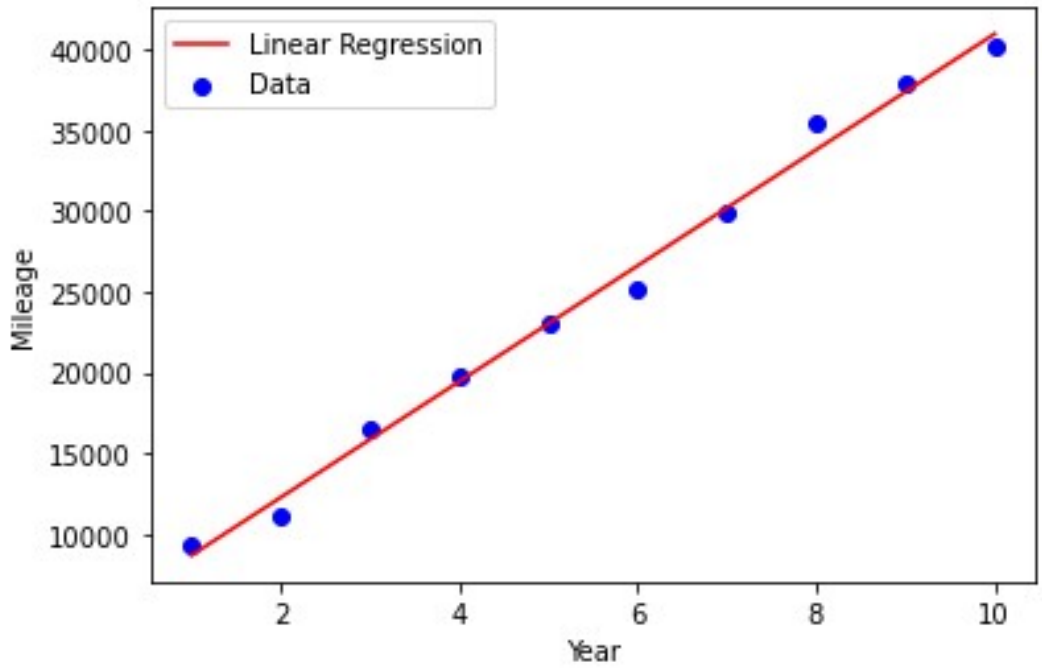

Figure 1: Least squares linear regression model

Forecast total mileage for the next 50 years:

Using the obtained model parameters a and b, the time in the next 50 years is used as input X, and the total mileage of the corresponding high-speed rail predicted in the model is Y.

The final high-speed rail operating mileage for the next 50 years is saved in future_mileage.csv file [11-13].

Table 1 below shows the forecast results of high-speed rail operating mileage data for the next 50 years:

Table 1: High-speed rail operating mileage forecast results from 2023 to 2072

Num	Year	Mileage
1	2023	44564.73
2	2024	48155.01
3	2025	51745.29
4	2026	55335.57
5	2027	58925.85
6	2028	62516.13
7	2029	66106.41
8	2030	69696.68
9	2031	73286.96
10	2032	76877.24
11	2033	80467.52
12	2034	84057.80
13	2035	87648.08
14	2036	91238.36
15	2037	94828.64
16	2038	98418.92
17	2039	102009.19
18	2040	105599.47
19	2041	109189.75
20	2042	112780.03
21	2043	116370.31
22	2044	119960.59
23	2045	123550.87
24	2046	127141.15
25	2047	130731.42
26	2048	134321.70
27	2049	137911.98
28	2050	141502.26
29	2051	145092.54
30	2052	148682.82
31	2053	152273.10
32	2054	155863.38
33	2055	159453.65
34	2056	163043.93
35	2057	166634.21
36	2058	170224.49
37	2059	173814.77
38	2060	177405.05
39	2061	180995.33
40	2062	184585.61
41	2063	188175.88
42	2064	191766.16
43	2065	195356.44
44	2066	198946.72
45	2067	202537.00
46	2068	206127.28
47	2069	209717.56
48	2070	213307.84
49	2071	216898.12
50	2072	220488.39

Figure 2 below shows the prediction diagram of the least squares linear regression model, where the abscissa represents the year (2012-2072) and the abscissa represents the total mileage of high-speed rail.

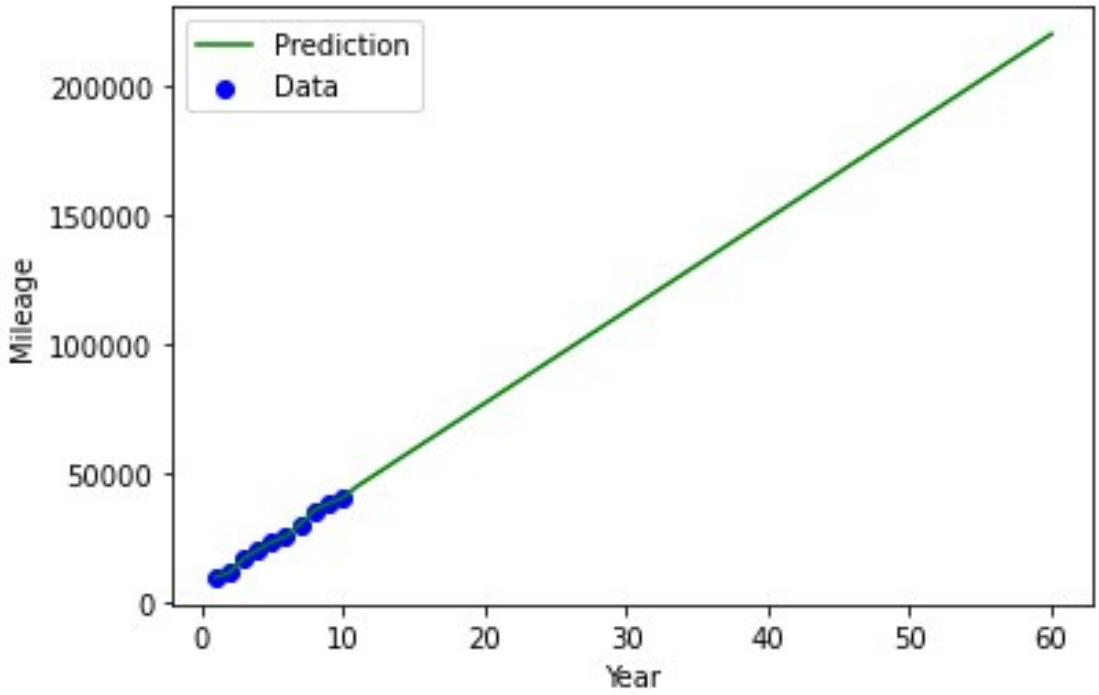

Figure 2: Least squares linear regression model prediction plot

5.2Establishment and solution of the model of problem 2

In question two, because the speed of high-speed rail is unknown, I will use a data-driven approach to model the relationship between them by analyzing historical high-speed rail speed and passenger traffic data. Here's how modeling is used:

Data analysis: Conduct preliminary exploratory analysis of passenger traffic data to understand the trend of passenger traffic [14].

Build a mathematical model: suppose that there is a functional relationship between the passenger volume of high-speed rail and the operating speed. Use regression analysis to model this relationship. where operating speed is the independent variable and passenger traffic is the dependent variable. Consider the linear regression model.

Model fitting: Apply the model to the data and obtain the parameter estimate of the model through the fitting technique. Use the fitted model to make predictions on the data.

Running speed prediction: Using the established model, the corresponding operating speed is deduced according to the preset passenger traffic data. Predictions can be made from model parameter estimates and new passenger traffic data [15].

It should be noted that since the operating speed of high-speed rail is unknown, the established model can only provide a relational model based on historical data, and there is still uncertainty about the actual situation in the future.

The process of establishing the program is as follows:

Import the necessary libraries:

Import the 'numpy' library for working with arrays and mathematical calculations.

Import the 'matplotlib.pyplot' library for charting.

Import the 'LinearRegression' model from the 'sklearn.linear_model' library for linear regression analysis.

Enter known high-speed rail passenger traffic data:

Define the year ('year') as an argument and use the 'np.arange' function to generate an array for the years 2008-2021.

Passenger traffic ('passenger_count') is defined as the dependent variable, giving passenger traffic data for the corresponding year.

Build a linear regression model:

Assign the year (X) and passenger traffic (y) to the variables X and y, respectively.

Using the 'LinearRegression' model, create a linear regression object named 'model'.

Fit the data using the 'fit' method to obtain the regression coefficients and intercepts.

where the regression coefficient slope=14419; intercept intercept=-28922676

4. Plot scatterplots and regression lines:

Use the 'plt.scatter' function to plot a scatter plot with year in abscum and passenger traffic (passenger_count) and blue in color.

Use the 'plt.plot' function to plot a regression line with the abscissa year and the ordinate calculated from the linear regression equation in red.

Add label information such as abscissa, ordinate, title, and legend.

Use the 'plt.show' function to display the chart.

Figure 3 below shows the relationship between high-speed rail operation speed and passenger traffic from 2008 to 2021:

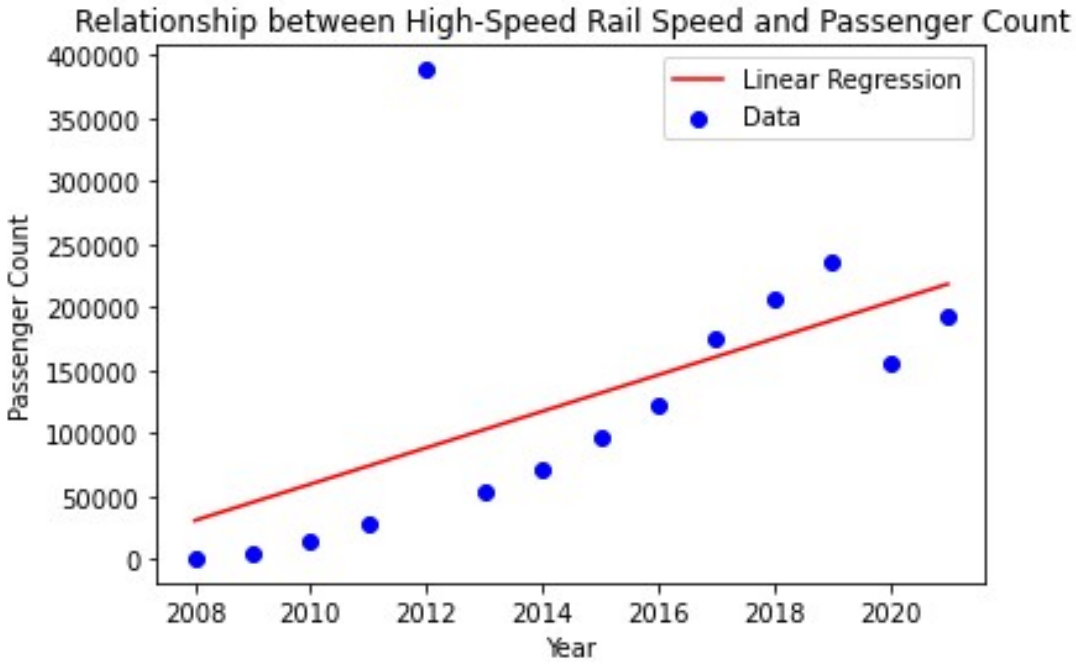

Figure 3: Relationship between high-speed railway operating speed and passenger traffic

Output linear regression equation:

Use 'model.coef_' to get the regression coefficient (slope).

Use 'model.intercept_' to get the intercept.

Output the regression equation to the console with the form "Linear regression equation: $y = slope * x + intercept$".

The resulting linear regression equation is:

$$y = 14419 * x - 28922676 \quad (3)$$

5.3Establishment and solution of the model of problem three

The modeling idea of this program is based on a linear weighted model to optimize the efficiency of high-speed rail transportation. By preprocessing the high-speed rail operation data over the years, the proportion of each factor in the total is calculated, and then the efficiency function and objective function are defined, and finally the optimal solution is solved by the optimization algorithm.

First of all, according to the high-speed rail operation data of the past years in the appendix, including the operating mileage, passenger traffic and passenger turnover. These data are key indicators for evaluating the efficiency of high-speed rail transportation [16].

Next, the indicators are normalized to calculate their relative weights in the population. This allows data from different years to be compared and eliminates orders of magnitude differences.

Figure 4 below shows the trend of high-speed rail operating mileage:

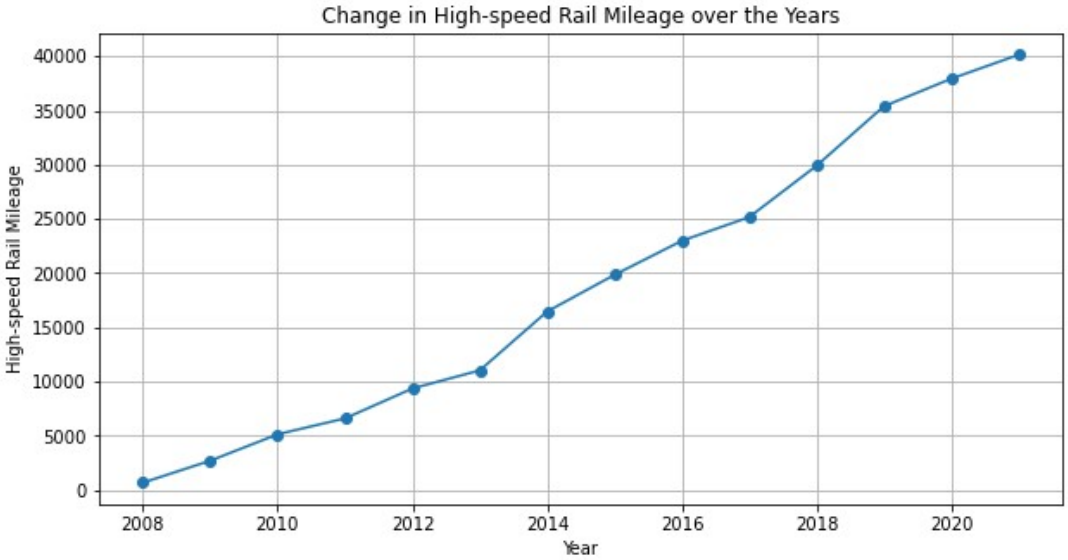

Figure 4: Trend chart of high-speed rail operating mileage

Figure 5 below shows the trend of passenger traffic:

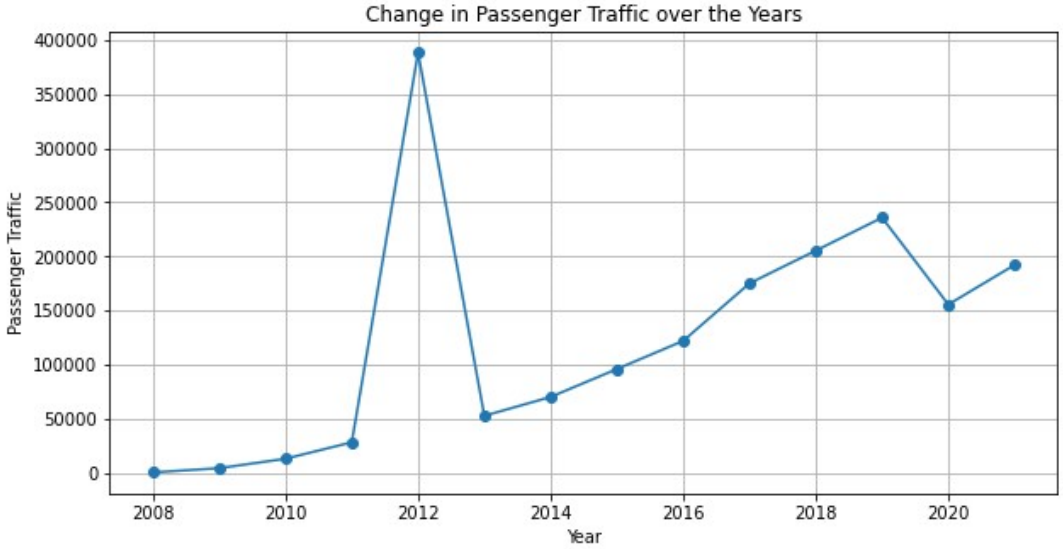

Figure 5: Trend chart of passenger traffic

Figure 6 below shows the trend of passenger turnover:

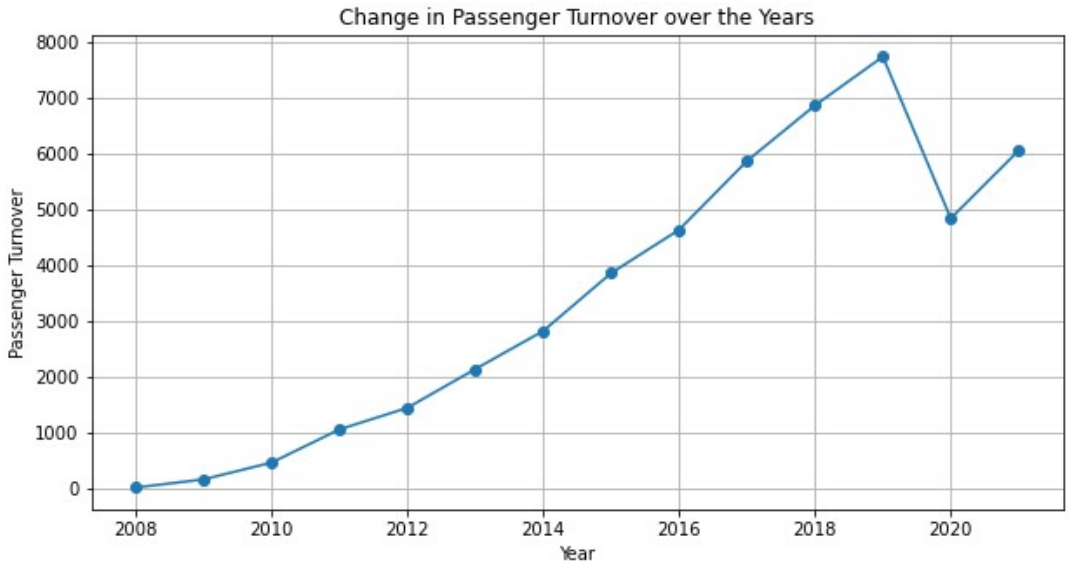

Figure 6: Trend chart of passenger turnover

After preprocessing the data, we need to define a function that calculates the efficiency of high-speed rail transportation. In this function, we use a linear weighted model that multiplies each factor with the corresponding weight coefficient and summes the result.

By adjusting the value of the weight coefficient, we can control how important each factor is in the efficiency calculation. This efficiency function will be used as the goal of optimization.

To achieve the goal of maximizing efficiency, we need to define an objective function. Because the optimization problem is usually a minimization problem, we take the negative number of the efficiency function as the objective function.

Thus, during optimization, the optimization algorithm will look for solutions that minimize the value of the objective function, even if the solution maximizes efficiency.

Next, we need to select an optimization algorithm to solve the optimal solution. The minimize function from the SciPy library is used here, which provides a variety of optimization algorithms to choose from.

We need to provide an initial guess and then find the optimal solution by minimizing the objective function. Ultimately, the algorithm returns an optimal set of weight coefficients that can be used to represent how important each factor is in the efficiency calculation.

Figure 7 below shows the histogram of the optimal solution coefficient:

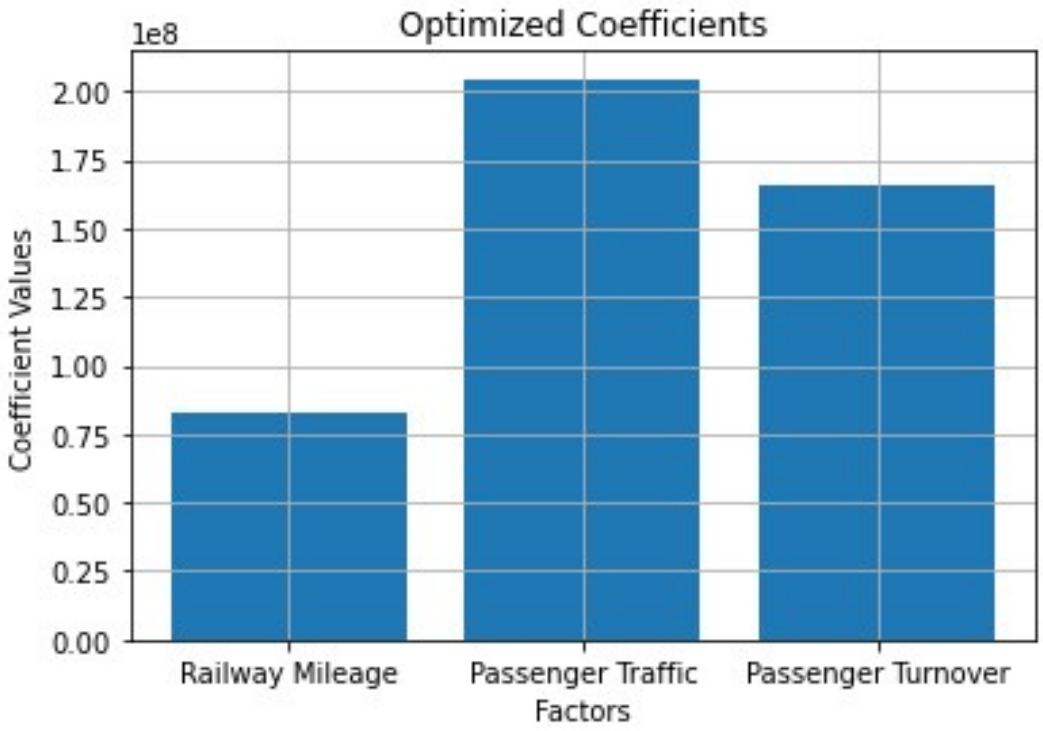

Figure 7: Histogram of optimal solution coefficients

Through the optimal solution obtained by the optimization algorithm, we can build the optimization equation. By substituting the weight coefficient in the optimal solution into the efficiency function, the final equation that can calculate the transportation efficiency of high-speed rail can be obtained.

This equation can be used to analyze data from different years, evaluate the efficiency of high-speed rail transportation, and make optimization and decision-making based on actual conditions.

Efficiency = 82768154.01 * The proportion of high- speed rail operating mileage to railway operating mileage204501238.91 * The proportion of passenger traffic in high-speed rail passenger traffic + 165536300.91 * Passenger turnover accounts for the proportion of railway passenger turnover.

In summary, the modeling idea of this question is to optimize the efficiency of high-speed rail transportation based on the linear weighting model. By preprocessing data, defining efficiency and objective functions, and using optimization algorithms to solve optimal solutions, we can quantify the efficiency of high-speed rail transportation and find the optimal solution to maximize it.

This modeling idea can be applied to the actual high-speed rail transportation management and decision-making, help optimize resource allocation, improve operational efficiency, and promote the development of the transportation field.

6 REPORT ON THE FUTURE DEVELOPMENT OF CHINA'S HIGH-SPEED RAILWAY

This report aims to analyze the future development prospects of high-speed railway in China and make corresponding recommendations. First, the development history and achievements of China's high-speed railway are reviewed. Then, the challenges and opportunities of high-speed rail development are discussed, and their impact on economic, environmental and social development is analyzed. Finally, suggestions were put forward to strengthen high-speed railway technology innovation, improve network planning, improve

service quality and strengthen international cooperation to promote the sustainable development of China's high-speed railway.

China's high-speed railway has developed rapidly since 2008 and has made remarkable achievements. At present, China's high-speed rail network has become one of the largest and fastest in the world. However, with the development of the economy and the improvement of people's living standards, China's high-speed railway still faces many challenges and opportunities.

The history and achievements of the development of high-speed railway

Looking back at the development of China's high-speed railway, starting from the Beijing-Shanghai high-speed railway to today's extensive high-speed railway network covering all parts of the country, China has made great breakthroughs in high-speed rail technology, construction speed and safety management. The opening of high-speed railway not only improves traffic efficiency, but also promotes regional economic development and people's living standards.

Challenges and opportunities for the development of high-speed rail

Despite the great achievements of high-speed rail, there are still some challenges. First, the connection between urban and rural areas is still not close enough, and some remote areas still do not benefit from the development of high-speed rail. Second, high-speed rail has higher operating costs and requires a better profit model and management approach. In addition, high-speed rail safety management, environmental protection and passenger service quality are also issues that need to be paid attention to. However, with the continuous innovation of technology and the support of policies, high-speed rail still has great development potential.

The economic, environmental and social impact of high-speed rail

The development of high-speed rail has had a profound impact on China's economy, environment and society. First of all, the convenient and rapid transportation brought by high-speed rail has promoted economic development and coordinated development between regions. Secondly, the construction of high-speed railways has led to the development of related industries and created jobs. In addition, the green travel and low carbon emission characteristics of high-speed rail help improve the environmental situation. Finally, the development of high-speed rail has also improved people's happiness and quality of life.

Suggestions for promoting the sustainable development of China's high-speed railways

In order to promote the sustainable development of China's high-speed railway, a series of measures need to be taken. First of all, strengthen technological innovation and research and development to improve the operational efficiency and safety performance of high-speed railways. Second, improve network planning, further improve the connection between urban and rural areas, and enable more areas to benefit from the development of high-speed rail.

In addition, more investment should be made to improve the quality of passenger services and enhance user experience. Finally, strengthen international cooperation, actively participate in international high-speed railway standards and technical exchanges, and enhance the influence of China's high-speed railway in the world.

6 CONCLUSION

The future development of China's high-speed railway has broad prospects. By overcoming the challenges and seizing the opportunities, China can further improve the development of high-speed railways. At the same time, the sustainable development of high-speed rail also requires the joint efforts and support of the government, enterprises and all sectors of society. It is believed that with the cooperation of all parties, China's high-speed railway will provide people with faster, more convenient and comfortable modes of transportation, and make greater contributions to economic and social development.

Model evaluation and promotion

For the model in question one:

Model advantages:

Simple and intuitive: A linear regression model is a simple and intuitive model that is easy to understand and interpret.

Explainability: Linear regression models can provide the degree to which each feature affects the target variable, helping us understand the relationship between variables.

High computational efficiency: Linear regression models have relatively low computational complexity and are suitable for large-scale datasets.

Strong interpretability of prediction results: The prediction results output by the linear regression model are continuous numerical values, which are easy to interpret and apply.

Model disadvantages:

Hypothesis limitations: Linear regression models are based on the assumption of linear relationships, and may not fit data well for nonlinear relationships.

Sensitive to outliers: Linear regression models are sensitive to outliers and may cause the model's predictions to deviate from the real situation.

Model improvements:

Introduce regularization: Using regularization terms (such as L1 regularization or L2 regularization) in the model can reduce the sensitivity of the model to outliers and improve the generalization ability of the model.

Ensemble method: Try to use ensemble learning methods, such as random forests or gradient boosting trees, to improve the overall prediction performance by combining the prediction results of multiple models.

For the model in question two:

In this code example, we fit known high-speed rail passenger traffic data using a linear regression model and evaluate the model. Here are the advantages, disadvantages, and possible improvements to the model:

Model advantages:

Simple and intuitive: The linear regression model is a simple and intuitive model that is easy to understand and implement.

Strong interpretability: linear regression models can provide the degree of influence of each feature on the target variable and help us understand the relationship between variables.

High computational efficiency: The computational complexity of linear regression models is relatively low and suitable for large-scale datasets.

Strong interpretability of prediction results: The prediction results output by the linear regression model are continuous numerical values, which are easy to interpret and apply.

Model disadvantages:

Assumption limitations: Linear regression models assume a linear relationship between the independent and dependent variables, and may not fit data well for nonlinear relationships.

Model improvements:

Feature engineering: By introducing more features, transformation features or interactive features, the fitting ability of linear regression models to nonlinear relationships can be improved.

Introduce regularization: Using regularization terms (such as L1 regularization or L2 regularization) in the model can reduce the sensitivity of the model to outliers and improve the generalization ability of the model.

For the model in question three:

Model advantages:

Be able to analyze the trend of high-speed rail operating mileage, passenger traffic and passenger turnover according to the past years.

By optimizing the objective function, the optimal coefficient value can be obtained to evaluate the efficiency of high-speed rail operation.

Use the matplotlib library to visualize the data and results.

Model disadvantages:

A linear relationship is assumed and other possible nonlinear relationships are not considered.

There may be problems with local optimal solution in optimizing the objective function, and the global optimal solution may not be obtained.

Model improvements:

Consider introducing more factors, such as economic development level, population density, etc., to more comprehensively evaluate the efficiency of high-speed rail operations.

You can try other optimization algorithms, such as genetic algorithm, particle swarm algorithm, etc., to get more accurate coefficient values.

Perform more detailed preprocessing and analysis of data, such as handling outliers, missing values, etc., to improve the robustness and reliability of the model.

REFERENCES

[1] Yan, H., Qi, Y., Ye, Q., & Yu, D. J. (2021). Robust least squares twin support vector regression with adaptive FOA and PSO for short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, *23*(9), 14542-14556.

[2] Saerens, B., Rakha, H., Ahn, K., & Van Den Bulck, E. (2013). Assessment of alternative polynomial fuel consumption models for use in intelligent transportation systems applications. *Journal of Intelligent Transportation Systems*, *17*(4), 294-303.

[3] Cong, Y., Wang, J., & Li, X. (2016). Traffic flow forecasting by a least squares support vector machine with a fruit fly optimization algorithm. *Procedia Engineering*, *137*, 59-68.

[4] Manogaran, G., Balasubramanian, V., Rawal, B. S., Saravanan, V., Montenegro-Marin, C. E., Ramachandran, V., & Kumar, P. M. (2020). Multi-variate data fusion technique for reducing sensor errors in intelligent transportation systems. *IEEE Sensors Journal*, *21*(14), 15564-15573.

[5] Mei, Z., Zhang, W., Zhang, L., & Wang, D. (2020). Real-time multistep prediction of public parking spaces based on Fourier transform–least squares support vector regression. *Journal of Intelligent Transportation Systems*, *24*(1), 68-80.

[6] Zhang, Y., & Huang, G. (2018). Traffic flow prediction model based on deep belief network and genetic algorithm. *IET Intelligent Transport Systems*, *12*(6), 533-541.

[7] Zheng, Y., & Peng, S. (2013). A practical roadside camera calibration method based on least squares optimization. *IEEE Transactions on Intelligent Transportation Systems*, *15*(2), 831-843.

[8] Li, X., Hu, Z., Shen, Y., Hao, L., & Shang, W. (2022). Distributed Intelligent Traffic Data Processing and Analysis Based on Improved Longhorn Whisker Algorithm. *IEEE Transactions on Intelligent Transportation Systems*.

[9] Boukerche, A., Tao, Y., & Sun, P. (2020). Artificial intelligence-based vehicular traffic flow prediction methods for supporting intelligent transportation systems. *Computer networks*, *182*, 107484.

[10] Boukerche, A., & Wang, J. (2020). Machine learning-based traffic prediction models for intelligent transportation systems. *Computer Networks*, *181*, 107530.

[11] Li, L., Qu, X., Zhang, J., Wang, Y., & Ran, B. (2019). Traffic speed prediction for intelligent transportation system based on a deep feature fusion model. *Journal of Intelligent Transportation Systems*, *23*(6), 605-616.

[12] Gan, S., Liang, S., Li, K., Deng, J., & Cheng, T. (2016). Long-term ship speed prediction for intelligent traffic signaling. *IEEE Transactions on Intelligent Transportation Systems*, *18*(1), 82-91.

[13] Deng, Y., Zhou, L., Wang, L., Su, M., Zhang, J., Lian, J., & Wei, J. (2020). Radio environment map construction using super-resolution imaging for intelligent transportation systems. *IEEE Access*, *8*, 47272-47281.

[14] Wang, D., Ma, X., Ma, D., & Jin, S. (2016). A novel speed–density relationship model based on the energy conservation concept. *IEEE Transactions on Intelligent Transportation Systems*, *18*(5), 1179-1189.

[15] Wang, Z., Chu, R., Zhang, M., Wang, X., & Luan, S. (2020). An improved hybrid highway traffic flow prediction model based on machine learning. *Sustainability*, *12*(20), 8298.

[16] Afandizadeh Zargari, S., Memarnejad, A., & Mirzahossein, H. (2021). Hourly origin–destination matrix estimation using intelligent transportation systems data and deep learning. *Sensors*, *21*(21), 7080.