

Short Term Demand Forecasting of Bike Sharing Stations Based on LSTM Model

Yixuan Yang^{1*}

¹School of Computer Science and Engineering (School of Cyber Security), University of Electronic Science and Technology of China, 611731, Chengdu, China

Abstract. With the increasingly prominent role of shared bicycles in urban traffic, station demand forecasting has become the key to scheduling optimization. In view of the research gap in the targeted application of LSTM model in this field, this study collected the Toronto shared single vehicle operation data, built the LSTM prediction model, used to predict the demand of each station in the next three hours, and evaluated the prediction performance of the model from a multidimensional perspective. In this study, the overall prediction results of 140 time points in the future of 15 stations and the prediction of 40 time points in the future of specific stations are visualized. The experimental results demonstrate that the LSTM model attains high levels of fitting accuracy for the demand of shared bicycle stations, which is significantly better than ARIMA model. It can provide a scientific reference for the intelligent scheduling of shared vehicles, and expanding the input characteristic dimension can further improve the prediction accuracy.

1 Introduction

Nowadays, with the increase of urban population, the problem of traffic congestion has become increasingly serious. Therefore, bike sharing has become a key component of the urban transport networks because of its convenient and environmentally friendly characteristics. However, there are a series of operational challenges behind the flourish development of shared bicycles: such as during peak hours, the supply of shared bicycles in some areas exceeds demand, and users often cannot find available vehicles; Instead of peak hours, some areas have a substantial quantity of idle vehicles, resulting in resource waste.

Therefore, it has become a key issue to achieve accurate demand forecasting of shared bikes so as to improve the efficiency of shared bikes and the users' experience.

At present, researchers and experts have established a variety of methods to predict short-term traffic flow, among which the most commonly used ones include traditional statistical models such as autoregressive integral moving average model (ARIMA), Kalman filter (KF) [1], machine learning architectures such as LSTM (Long Short-Term Memory) networks, CNN (Convolutional Neural Network), and RNN (Recurrent Neural Network) [2], and hybrid prediction models [3]. The LSTM model is a recursive neural network developed and optimized on the basis of the recurrent neural network (RNN) architecture, which improves

* Corresponding author's email: 2023080905027@std.uestc.edu.cn

the ability of processing sequence data, so it is widely used in time series data prediction [4, 5].

The existing dynamic linear model has certain flexibility and adaptability for the related research of shared bicycles, but its processing ability for complex nonlinear relationships is limited [6]. After that, the study predicted the hourly starting and ending point (OD) demand of shared vehicles based on the hurdle model. This model can handle the special situation that there is little demand for a specific time and place, but if the characteristics of the data set do not meet the assumptions of the model, the prediction effect will be affected [7]. The method based on graph convolution neural network proposed by beiluz A and edrici a has advantages in dealing with the spatio-temporal characteristics of the demand for shared bicycles, and has a corresponding improvement in the prediction accuracy compared with the traditional model [8]. However, there is a gap in the contemporary studies on the forecasting of bike-sharing demand grounded in the LSTM model with the ability to process serial data, so this study will use this model to try to make a more accurate demand prediction.

This study takes the Toronto shared bicycle data in January 2018 as a reference. After preprocessing the data, an LSTM model is established to predict the short-term (3 hours) demand for shared bicycles at each station in the future. The objective of this research is to supply dependable predictive information for the shared bicycle dispatching system, so as to effectively improve the low utilization of shared bicycles caused by improper scheduling

2 Methods

2.1 Data source

The data used in this study are from the Toronto bike sharing system and downloaded from the kaggle data website. The original data covers the records of all shared single vehicle trips within this range from January 1 to March 31, 2018, with a total of 178559 trips recorded. Each data corresponds to a separate trip, including information as shown in Table 1. All trip data are sorted by start time.

Table 1. Original data information attribute information

Field Name	Description	Data Type
trip_id	Trip number	Int64
trip_duration_seconds	Trip duration (Unit: seconds)	Int64
from_station_id	Starting station number	Int64
trip_start_time	Departure time	Object
from_station_name	Starting station name	Object
trip_stop_time	Arrival time	Object
to_station_id	Destination number	Int64
to_station_name	Destination name	Object
user_type	User type	Object

2.2 Indicator Selection and Description

This study mainly uses the starting time of each trip and the starting station ID information to predict the short-term single vehicle demand at each station. Firstly, the data were preprocessed (outliers and non-conforming data were deleted), and then the data in January and February after processing were used as the training set, and the data in March after processing were used as the test set for research.

In order to predict the shared single vehicle demand of each station in the next 3 hours, the preprocessed data is divided into a window every 3 hours, and a time window in the future is predicted according to the past eight-time windows.

2.3 Method Introduction

Long Short-Term Memory (LSTM) networks represent a type of deep learning model that is widely applied to the processing of sequential data. In contrast to traditional Recurrent Neural Networks (RNNs), LSTM incorporates three gating mechanisms—an input gate, a forget gate, and an output gate—as well as a cell state. These structural designs allow LSTM to address the issue of long-term dependencies in sequential data more effectively. The operational principles of the three gates are elaborated below:

Forget gate: The sigmoid activation function is applied to compute the current input x_t and the hidden state h_{t-1} (the hidden state from the last time step), resulting in a vector whose values fall within the range of 0 to 1.

A value of 0 in the vector indicates that the corresponding component of the historical cell state is to be discarded, while a value of 1 denotes that this component is retained;

Input gate: New memory states are derived by merging the preserved information from the historical cell state with the new information that requires encoding into the current state;

Output gate: The final output is produced by integrating the processed cell state and hidden state information through a series of nonlinear transformations.

The architectural structure of the LSTM network is presented in Figure 1.

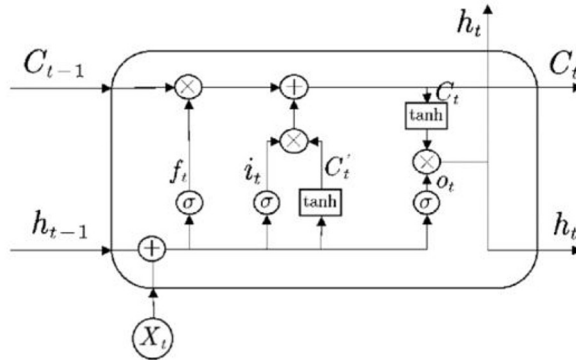


Fig. 1. LSTM structure [9]

In Figure 1, C_t is the cell state (memory state), x_t is the input information, and h_{t-1} is the hidden state (based on C_t).

2.4 Model evaluation indicators

In this study, the LSTM model was evaluated from the distribution of prediction error values and the values of mean square error, mean absolute error and root mean square error. That is to detect the deviation of the predicted value of each station object, and also consider the average error of the predicted value of all stations, so as to get a more comprehensive analysis of the model.

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \quad (1)$$

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (3)$$

At the same time, through a comparative analysis of the forecasting outcomes from the LSTM model and the ARIMA model [10], this experiment analyzes the fitting ability of the two models for the demand prediction of shared bicycles in the next 3 hours.

3 Results and Discussion

3.1 Analysis of Predictive Results

As the popular sites have richer data, which have more accurate prediction and greater impact on the operation system, this study conducts research and prediction for the top 15 sites with the largest demand. The data is normalized to unify the data range and accelerate convergence.

The LSTM model constructed in this experiment has two hidden layers (128 hidden units in each layer), and the discard rate is set to 0.2. After 300 rounds of training, the final loss is 0.008690.

Figure 2 shows the overall prediction results of LSTM model.

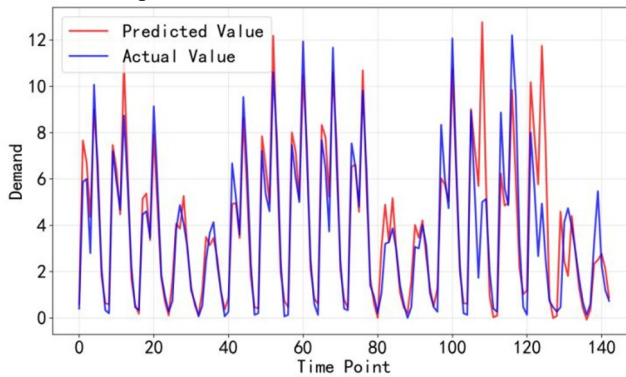


Fig. 2. Comprehensive forecasting outcomes derived from the LSTM model (Picture credit: Original)

The forecasting outcomes of the ARIMA model are presented in Figure 3, and the fitting effect of ARIMA model on the demand of shared single vehicle is significantly weaker than that of LSTM model.

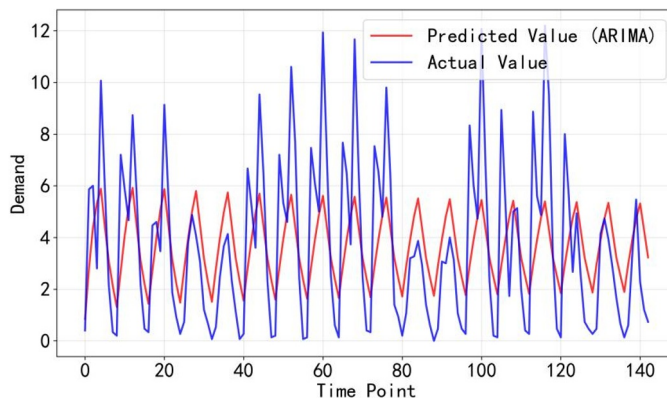


Fig. 3. Comprehensive forecasting outcomes derived from the ARIMA model (Picture credit: Original)

By visualizing the evaluation data of LSTM model prediction results, the corresponding results in Figure 4 are obtained. From the above distribution diagram, it is found that the experimental prediction error value approximately presents a normal distribution centered on 0, which is in line with the prediction expectation. The histogram below shows that the mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE) are 6.20, 1.67 and 2.49 respectively, which are small, indicating that the prediction effect is ideal.

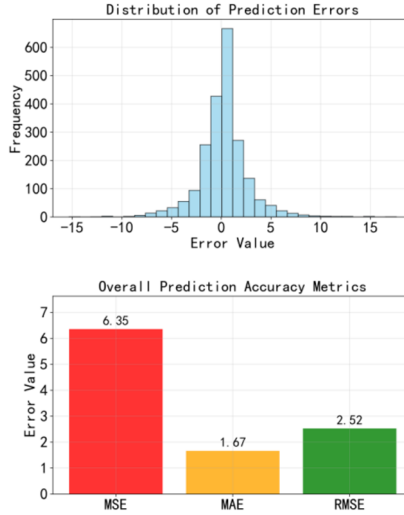


Fig. 4. LSTM model evaluation (Picture credit: Original)

To assess the effectiveness of the Long Short-Term Memory model more comprehensively and accurately, this study selected the prediction results of typical sites for separate analysis, focusing on the site 7042 with the largest demand and the site 7077 with the smallest demand. From the information observation in Figure 5, it can be seen that the demand forecast for the first 40 time points of station 7042 is relatively accurate, showing a good ability to fit the peak and cycle changes, but the prediction ability for the peak in the demand forecast for station 7077 is obviously weak, and even at the 36th time point, the predicted value is half of the actual value. Combined with the analysis of the two sites, it shows that the model has the defect of insufficient prediction ability when facing small data samples.

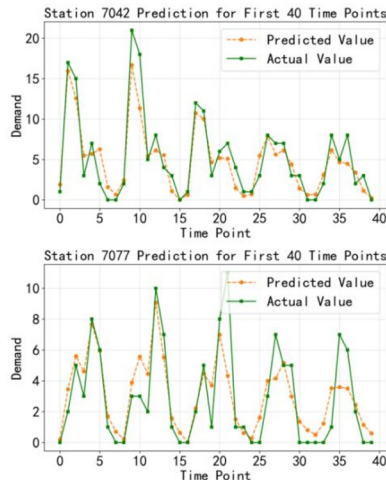


Fig. 5. Prediction comparison of LSTM model stations 7042 and 7077 (Picture credit: Original)

3.2 Discussion

In order to explore the accurate prediction method for the demand for bike sharing in the next three hours, this experiment uses LSTM model to forecast the requirement of 15 shared bicycles stations in the next 3 hours. The results show that the model is more suitable for the overall demand trend prediction: the average forecast demand is 3.72, and the deviation from the actual average demand of 3.68 is only 0.04, indicating that LSTM model has basic effectiveness in the task of three-hour-ahead ridership prediction of bike sharing.

Further analysis of the forecasting performance of different stations reveals that the precision of the forecasts of stations with sufficient data is better, including peak and cycle prediction. However, the model has the defect of inaccurate peak prediction for sites with small training data sets. This is mainly due to the fact that the model is difficult to fully capture the key characteristics of demand mutations in small sample scenarios, resulting in insufficient fitting ability for irregular fluctuations. In view of this defect, it can be optimized by increasing the amount of processed sample data or adding characteristic quantities such as weather, holidays and whether it is a business district as the input value of the LSTM model, so as to provide a more comprehensive decision basis for the model, which in turn lessens the divergence between the predictive outputs of the model and the actual number of riders, and improve the robustness of the model in complex scenarios.

4 Conclusion

Aiming at the core issue of short-term prediction of bike sharing demand in a 3-hour time window and operation optimization of shared bicycles, this study empirically analyzes the demand of 15 shared bicycles stations in the next three hours through the LSTM model, and finds that there are differences in the future demand of different shared bicycles stations (including the dimensions of demand scale, time periodicity and peak characteristics). Therefore, the accurate prediction of future ridership demand at each station obtained from the LSTM model is helpful to provide scientific suggestions for the dynamic scheduling of the shared single vehicle system and improve the vehicle utilization rate, so as to realize the efficient operation of the shared single vehicle operation system.

This paper's core contribution is to provide a new model idea for shared vehicle demand forecasting. By applying the LSTM model—skilled in sequence data processing—to this field, it offers new technical support and holds important practical reference value for enhancing the intelligent functions of shared vehicle operation systems.

The current research still has limitations: for the smaller stations in the training data set, the peak data predicted by the model is not accurate. Future research should try to increase the dimensionality of the model's input features, reduce the deviation between the predicted value and the actual demand by enriching the feature information, and continuously elevate the applicability and ruggedness of the model architecture in complex scenarios.

References

1. F. J. Li, Q. Zong. The Forecasting of the Elevator Traffic Flow Time Series Based on ARIMA and GP. *Adv. Mater. Res.*, 2074(588-589): 1466-1471 (2012). DOI:10.4028/www.scientific.net/AMR.588-589.1466
2. Q. Xi, W. Zhang, M. Du, et al. Passenger flow forecasting approaches for urban rail transit: a survey. *Int. J. Gen. Syst.*, 52(8): 919-947 (2023). DOI:10.1080/03081079.2023.2231133

3. J. Tang, Y. Zhu, S. Yang, et al. Predicting Urban Traffic Under Extreme Weather by Deep Learning Method with Disaster Knowledge. *Appl. Sci.*, 15(17): 9848 (2025). DOI:10.3390/APP15179848
4. Q. Zhao, X. Feng, L. Zhang, et al. Research on Short-Term Passenger Flow Prediction of LSTM Rail Transit Based on Wavelet Denoising. *Mathematics*, 11(19): 4204 (2023). DOI:10.3390/MATH11194204
5. Y. Liu, Z. Liu, R. Jia. DeepPF: A deep learning based architecture for metro passenger flow prediction. *Transp. Res. Part C*, 101: 18-34 (2019). DOI:10.1016/j.trc.2019.01.027
6. M. A. H., E. Mohammed, R. A. H. Dynamic linear models to predict bike availability in a bike sharing system. *Int. J. Sustain. Transp.*, 14(3): 232-242 (2020). DOI:10.1080/15568318.2019.1611976
7. L. Vera, A. A. R. Predicting hourly origin - destination demand in bike sharing systems using hurdle models: Lisbon case study. *Case Stud. Transp. Policy*, 9(4): 1836-1848 (2021). DOI:10.1016/J.CSTP.2021.10.003
8. A. Behroozi, A. Edrisi. Predicting travel demand of a bike sharing system using graph convolutional neural networks. *Public Transp.*, 17(1): 1-37 (2024). DOI:10.1007/S12469-024-00371-W
9. D. Chen, C. Xiong, M. Zhong. Improved LSTM Based on Attention Mechanism for Short-term Traffic Flow Prediction. *Proceedings of the 10th International Conference on Information Science and Technology (ICIST)*. Bath, London, and Plymouth, UK. 2020. 71-76. DOI:10.1109/ICIST49303.2020.9202045
10. X. Wei, K. S. The Comparison Between ARIMA and Long Short-term Memory for Highway Traffic Flow Prediction. *J. East. Asia Soc. Transp. Stud.*, 13: 1817-1834 (2020). DOI:10.11175/easts.13.1817