Scooter-Share Travel Demand Forecast: A Context-Aware LSTM Recurrent Neural Network Approach

Dr. Junfeng Jiao (PI)

Dr. Yiming Xu

A publication of the USDOT Tier 1 Center:

Cooperative Mobility for Competitive Megaregions

At The University of Texas at Austin

DISCLAIMER: The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Technical Report Documentation Page

<table>
<tr><td>1. Report No. CM2-80</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Scooter-Share Travel Demand Forecast: A Context-Aware LSTM Recurrent Neural Network Approach</td><td colspan="2">5. Report Date
October 22, 2024</td></tr>
<tr><td colspan="2">6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Author(s)
Junfeng Jiao, Ph.D., The University of Texas at Austin
Yiming Xu, Ph.D., The University of Texas at Austin</td><td colspan="2">8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
The University of Texas at Austin, School of Architecture, 310 Inner Campus Drive, Austin, TX 78712</td><td colspan="2">10. Work Unit No.</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
USDOT 69A3551747135</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
Center for Cooperative Mobility for Competitive Megaregions
The University of Texas at Austin
310 Inner Campus Drive, Goldsmith Hall, 2.308
Austin, TX 78712</td><td colspan="2">13. Type of Report and Period Covered</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code</td></tr>
<tr><td colspan="4">15. Supplementary Notes</td></tr>
<tr><td colspan="4">16. Abstract
Shared micromobility has been popular in many cities in the U.S. The rise of shared micromobility brings significant operational challenges such as fleet management and demand forecasting. This project develops a Context-Aware Long Short-Term Memory (CALSTM) recurrent neural network to enhance the prediction of daily travel demand for scooter-sharing in Austin, Texas. The CALSTM model boosts prediction accuracy by integrating the impact of nearby points-of-interest (POIs) and daily weather conditions on scooter usage. It processes historical scooter-sharing demand and weather information through separate LSTM modules to extract temporal information. The outputs from these modules are combined through element-wise multiplication to establish temporal dependencies. Additionally, POI information is analyzed using a Multi-Layer Perceptron (MLP) to capture spatial dependencies. These spatial and temporal dependencies are then integrated by another MLP module to produce the forecast outputs. Case study experiments in Austin, TX, demonstrated that the CALSTM model significantly outperformed benchmark models, achieving improvements of 28% in Mean Absolute Error (MAE) and 19% in Root Mean Squared Error (RMSE) over traditional LSTM models. These results offer valuable insights for transportation planning and the enhancement of shared micromobility in urban settings.</td></tr>
<tr><td colspan="2">17. Key Words
Shared micromobility; Demand; Deep Leaning, Urban Mobility</td><td colspan="2">18. Distribution Statement
No restrictions to all.</td></tr>
<tr><td>19. Security Classif. (of this report)
N/A</td><td>20. Security Classif. (of this page)
N/A</td><td>21. No. of Pages</td><td>22. Price
N/A</td></tr>
</table>

ACKNOWLEDGEMENT OF SPONSORSHIP AND STAKEHOLDERS

We would like to thank Cooperative Mobility for Competitive Megaregions (CM2), a United States Department of Transportation (USDOT) Tier-1 University Transportation Center (UTC), for funding of this research.

LIST OF AUTHORS

Lead PI:

Junfeng Jiao, Ph.D.
The University of Texas at Austin
jjiao@austin.utexas.edu

Additional Researchers:

Yiming Xu, Ph.D.
The University of Texas at Austin
yiming.xu@utexas.edu

CONTENTS

ABSTRACT

Shared micromobility has been popular in many cities in the U.S. The rise of shared micromobility brings significant operational challenges such as fleet management and demand forecasting. This project develops a Context-Aware Long Short-Term Memory (CALSTM) recurrent neural network to enhance the prediction of daily travel demand for scooter-sharing in Austin, Texas. The CALSTM model boosts prediction accuracy by integrating the impact of nearby points-of-interest (POIs) and daily weather conditions on scooter usage. It processes historical scooter-sharing demand and weather information through separate LSTM modules to extract temporal information. The outputs from these modules are combined through element-wise multiplication to establish temporal dependencies. Additionally, POI information is analyzed using a Multi-Layer Perceptron (MLP) to capture spatial dependencies. These spatial and temporal dependencies are then integrated by another MLP module to produce the forecast outputs. Case study experiments in Austin, TX, demonstrated that the CALSTM model significantly outperformed benchmark models, achieving improvements of 28% in Mean Absolute Error (MAE) and 19% in Root Mean Squared Error (RMSE) over traditional LSTM models. These results offer valuable insights for transportation planning and the enhancement of shared micromobility in urban settings.

Keywords:
Shared micromobility; Demand; Deep Leaning, Urban Mobility

1. INTRODUCTION

Shared micromobility, including dockless bike-sharing and scooter-sharing, has been popular in many cities in the U.S. These modes are appealing due to their compact size and the convenience of on-demand access, positioning them as effective low-carbon options for short-distance travel. Recent trends show a remarkable expansion in shared micromobility's usage. According to a report by NACTO (2024), Americans undertook 133 million trips using these services in 2023, split between 65 million e-scooter trips and 68 million bike-sharing trips. This represents a 16% increase in usage compared to the previous year.

The rise of shared micromobility brings with it significant operational challenges, particularly in the distribution and rebalancing of e-scooter fleets. It is essential to ensure that scooters are strategically deployed across urban areas to meet high demand and avoid both scarcity and congestion. Achieving an optimal distribution requires precise demand forecasting, which also facilitates efficient fleet management. Moreover, operators must carefully balance affordability for users with profitability. Effective demand forecasting can further support the development of dynamic pricing strategies, aligning costs for users with revenue goals for service providers.

The travel demand forecasting problem has been extensively studied by researchers. Deep learning methods, such as Transformer (Vaswani et al., 2017), LSTM (Hochreiter and Schmidhuber, 1997), and GRU (Cho et al., 2014), have been proven effective in processing time series data due to their robust architectures and ability to capture sequential dependencies and complex non-linear relationships. These methods typically use historical demand data as input and conduct time series modeling to forecast future demand. However, forecasting for e-scooter sharing involves complexities beyond simple time series analysis due to significant spatial influences on demand. Studies indicate that shared micromobility usage is concentrated in urban downtowns near entertainment and retail centers, and on university campuses. Additionally, there is evidence of a strong linkage between micromobility demand and economic activity. Incorporating point-of-interest data can enhance prediction accuracy. Furthermore, research highlights the impact of weather conditions—like temperature, precipitation, and wind—on e-scooter usage (Noland, 2021; Xu et al., 2024; Xu et al., 2023). For instance, adverse weather conditions notably decrease the usage of e-scooter services (Noland, 2021). Therefore, including weather data is crucial for accurate demand forecasting in shared micromobility.

This project develops a context-aware long short-term memory (CALSTM) network to predict daily travel demand for scooter-sharing in Austin, Texas. The CALSTM model enhances prediction accuracy by incorporating the influence of nearby points-of-interest (POIs) and daily weather conditions on scooter usage. The findings from this study provide critical insights that can inform transportation planning efforts aimed at enhancing shared micromobility in urban areas. By identifying areas of high demand influenced by local economic activities, this research supports strategic decisions to promote scooter-sharing as a viable urban transport option.

This report consists of five sections. Section 2 reviews related literature on travel demand forecasting. Section 3 formally defines the research problem and introduces the details of the proposed method. Section 4 presents the application of the proposed method in forecasting travel demand for shared e-scooters in Austin, TX. Section 5 concludes this project and discusses limitations and future directions.

2. LITERATURE REVIEW

2.1 Shared E-scooter Usage

Recent studies have extensively analyzed the usage patterns of shared e-scooters, focusing on aspects such as trip distance, duration, and purpose, as well as their spatiotemporal distribution. For instance, the average trip duration across the U.S. and Canada in 2022 was 11 minutes, but this varied by city (Caspi et al., 2020; Jiao and Bai, 2020; McKenzie, 2019). In Austin, the typical e-scooter trip covered 0.77 miles and lasted 7.55 minutes, while in Washington DC, trips were shorter, averaging 0.6 km in distance and 5 minutes in duration (Jiao and Bai, 2020; McKenzie, 2019). Additionally, research has shown that most e-scooter trips are for non-commuting purposes like errands and leisure, underscoring the versatility of e-scooters in urban mobility (Bai et al., 2021; Caspi et al., 2020; Nikiforiadis et al., 2021; Shah et al., 2023).

Temporal usage patterns also differ from traditional traffic peaks (Liu et al., 2019; NACTO, 2024; Xu et al., 2022). For example, a peak in e-scooter use across the U.S. was noted between 4 pm and 6 pm, while in Indianapolis, peak times extended from 4 pm to 9 pm on weekdays and 2 pm to 7 pm on weekends (NACTO, 2024; Liu et al., 2019). Spatially, e-scooter trips are heavily concentrated in downtown and university areas, with significantly lower densities in suburban zones (Bai and Jiao, 2020; Hosseinzadeh et al., 2021; McKenzie, 2019, 2020; Xu et al., 2022).

Various factors influencing e-scooter usage have been identified. Demographic elements such as age, gender, income, education level, ethnicity, marital status, and residency significantly affect user preferences, with a trend toward higher usage among younger individuals, males, those with higher incomes and education levels, and singles (Cao et al., 2021; Christoforou et al., 2021; Laa and Leth, 2020; Lee et al., 2021; Sanders et al., 2020). Zonal demographic characteristics like population density, employment rates, and the presence of young adults and highly educated individuals also correlate positively with e-scooter usage (Bai and Jiao, 2020; Caspi et al., 2020; Merlin et al., 2021).

Moreover, the built environment plays a crucial role. Streets with better conditions and cycling amenities experience more e-scooter trips (Caspi et al., 2020; Hosseinzadeh et al., 2021). Proximity to city centers and transit stations enhances usage, and areas with diverse land uses and more commercial zones tend to see higher intensities of e-scooter trips (Bai and Jiao, 2020; Merlin et al., 2021; Hosseinzadeh et al., 2021).

Weather conditions also significantly impact e-scooter usage. Studies have shown that adverse weather, such as rain, snow, or extreme temperatures, leads to a noticeable decrease in e-scooter trips (Noland, 2021). Conversely, mild and sunny weather tends to increase usage rates, as more people choose e-scooters for transportation or leisure activities (Bai et al., 2021; Caspi et al., 2020). Weather-related data is thus crucial for predicting daily and seasonal fluctuations in e-scooter demand (Noland, 2021; Xu et al., 2024; Xu et al., 2023).

2.2 Travel Demand Forecasting Using Deep Learning

Deep learning models have increasingly been applied to predict travel demand, with significant attention on ridesourcing (Ke et al., 2021; Geng et al., 2019), taxi services (Liu et al., 2019; Xu et al., 2017; Yao et al., 2018), and bike-sharing systems (Lin et al., 2018; Li et al., 2019; Kim et al., 2019). These studies have shown that deep learning outperforms traditional statistical methods like ARIMA and machine learning techniques such as gradient-boosted trees and random forests in forecasting accuracy. Specific neural network architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs) like LSTM and GRU, and Transformers, are employed to handle time-varying data effectively (Chen et al., 2020; Pan et al., 2019; Li et al., 2021). Moreover, recent studies have combined different neural networks into an integrated spatiotemporal model to simultaneously capture both spatial and temporal patterns. Research indicates that spatiotemporal models generally surpass individual CNN or RNN models in performance due to their ability to simultaneously analyze spatial and temporal data, enhancing prediction accuracy (Tang et al., 2021; Geng et al., 2019; Yao et al., 2018; Li et al., 2021; Zhang et al., 2024). However, the complexity of these models leads to challenges such as extended training times and limited scalability when processing long data sequences (Xu et al., 2023; Cai et al., 2020; Xu et al., 2020). These limitations stem from the intricate architectures required to integrate and process multiple data dimensions effectively.

In the realm of shared micromobility, various deep learning frameworks have been employed for demand forecasting. These include RNN (Chen et al., 2020; Pan et al., 2019) and spatiotemporal models (Xu et al., 2023; Lin et al., 2018; Liu et al., 2019; Ai et al., 2019; Yang et al., 2020; Phithakkitnukoon et al., 2021; Ham et al., 2021; Song et al., 2023; Li et al., 2023; Ma and Liu, 2024; Liang et al., 2024; He and Shin, 2020; Liang et al., 2023). For instance, Chen et al. (2020) evaluated various RNN architectures for forecasting real-time demand in New York's station-based bike-sharing network. Lin et al. (2018) developed a spatiotemporal architecture that combines GCN and LSTM to analyze complex station-level correlations within large bike-sharing networks. Liang et al. (2023) introduced a domain-adversarial multi-relational graph neural network that utilizes multimodal historical data for predicting bike-sharing demand. These models, particularly spatiotemporal ones, utilize advanced graph neural networks to analyze spatial data alongside temporal data, enhancing accuracy by incorporating variables like zonal functionality, demographics, and built environment characteristics. Nevertheless, as the graph size increases, these models face significant computational challenges (Xu et al., 2023). Furthermore, while recent research has begun to address dockless micromobility systems, the majority of studies still focus on station-based systems, highlighting a notable research gap and potential for accuracy improvements in dockless micromobility forecasting.

3. METHODS

This project develops a context-aware long short-term memory (CALSTM) recurrent neural network to predict daily travel demand for scooter-sharing. In this section, we first describe the formal definition of the scooter-sharing demand prediction problem. Then we introduce the framework for the proposed CALSTM model and the details of each model component.

3.1 Problem Definition

In this project, we forecast the daily scooter-sharing travel demand at census tract level. The number of scooter-sharing uses in each area during a specific time interval is defined as the travel demand for scooter-sharing. The scooter-sharing demand of an area i at time t is denoted by x_t^i. The scooter-sharing demand of all areas at time t is $X_t = [x_t^1, x_t^2, \dots, x_t^N]$, where N is the number of areas. The historical demand is denoted by a matrix $\boldsymbol{X}^{N \times T} = [X_{t-T+1}, \dots, X_t]$, where T is the length of the input time sequence. The weather condition is denoted by C and the point of interest information is denoted by POI.

The scooter-sharing demand prediction problem can be defined as follows. Given historical demand $\boldsymbol{X}_t$, weather condition C, and point of interest information POI, find a mapping $f: \mathbb{R}^{N \times T} \rightarrow \mathbb{R}^{N \times M}$ that maps historical demand and other information to the demand in the next M time $\boldsymbol{Y}_t$:

$$\boldsymbol{Y}_t = [X_{t+1}, \dots, X_{t+M}] = f(\boldsymbol{X}_t, C, POI)$$

3.2 Model Framework

The framework of proposed context-aware long short-term memory (CALSTM) recurrent neural network is presented in Figure 1. The inputs of the model include historical scooter-sharing demand, weather information, and POI information. To extract temporal information, we use a LSTM module to process the historical demand and use another LSTM module to process weather information. The outputs of the two LSTM modules are integrated by element-wise multiplication to generate the temporal dependencies. The POI information is processed by a Multi-Layer Perceptron (MLP) module to extract spatial dependencies. The spatial dependencies and temporal dependencies are then process a MLP module to generate the forecasting output. Note that the two LSTM modules share the same architecture but use different parameters. This also applies to the two MLP modules.

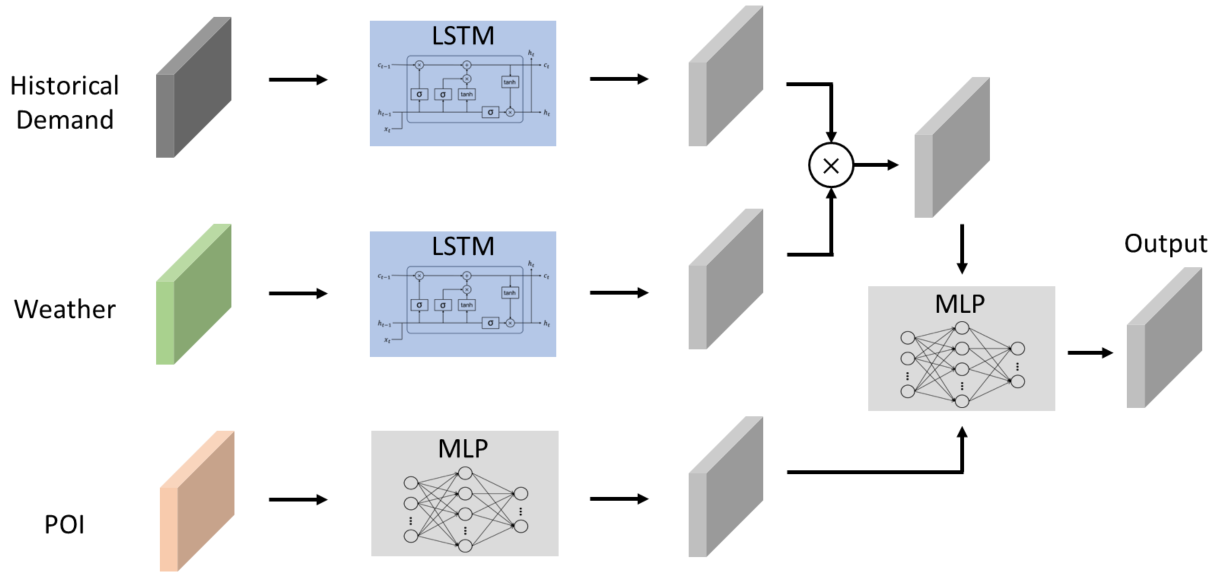

Figure 1. Framework of Context-aware Long Short-term Memory Network

3.3 Long Short-Term Memory

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) specifically designed to address the issue of long-term dependencies in sequence prediction problems. Traditional RNNs struggle with learning long-term dependencies due to the vanishing gradient problem, where gradients of the network's weights either grow or decay exponentially during backpropagation, making it difficult to train the network effectively over longer sequences. LSTM is designed aimed at mitigating the vanishing gradient problem.

An LSTM unit is intricately structured, comprising a cell and three regulatory gates: an input gate, an output gate, and a forget gate. The cell is crucial as it retains values over arbitrary time intervals, ensuring the continuity of information through the sequence of data. Each gate within the unit plays a pivotal role in the flow of information. The forget gate determines what information to discard from the previous state. It does this by analyzing both the previous state and the current input, then mapping this information to a value between 0 and 1. A value close to 1 indicates that the information is important and should be retained, while a value closer to 0 means it should be forgotten. Simultaneously, the input gate decides which new information is significant enough to update the current cell state. It uses a similar mechanism to the forget gate, evaluating the new data to decide which parts are stored and which are disregarded, ensuring that only relevant information is kept. Finally, the output gate controls the information that is emitted from the cell state. It assesses the current cell state, along with previous outputs, to assign a value between 0 and 1 to each piece of information, where 1 means fully passing the information through and 0 means blocking it. This selective output enables the LSTM network to utilize and maintain useful, long-term dependencies. These dependencies are critical as they allow the network to make informed predictions, affecting both current and future time-steps in the data sequence. The framework of a LSTM unit is presented in Figure 2.

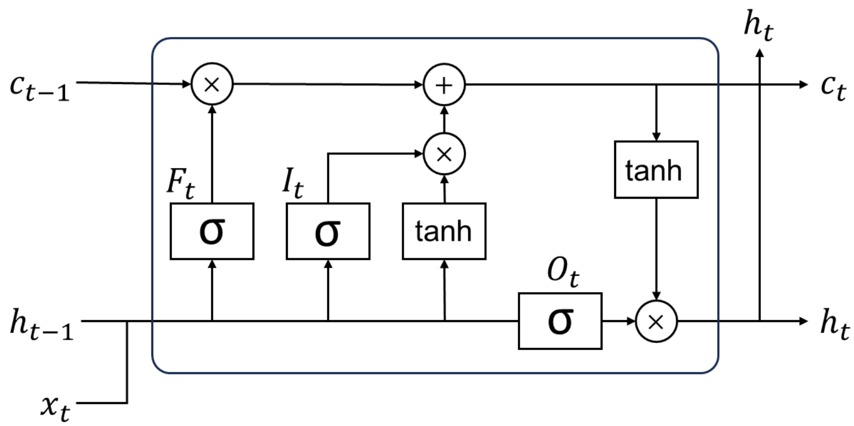

Figure 2. Framework of Long Short-term Memory Unit

In Figure 2, the input gate is I_t, the forget gate is F_t, and the output gate is O_t. The three gates are calculated by:

$$I_t = \sigma(x_t W_{xi} + h_{t-1} W_{hi} + b_i)$$

$$F_t = \sigma\left(x_t W_{xf} + h_{t-1} W_{hf} + b_f\right)$$

$$O_t = \sigma(x_t W_{xo} + h_{t-1} W_{ho} + b_o)$$

where W_{xi}, W_{xf}, W_{xo} and W_{hi}, W_{hf}, W_{ho} are weight parameters; b_i, b_f, and b_o are bias parameters; $\sigma(\cdot)$ is the sigmoid function.

The input node $\tilde{C}_t$ is calculated by:

$$\tilde{C}_t = \tanh(x_t W_{xc} + h_{t-1} W_{hc} + b_c)$$

where W_{xc} and W_{hc} are weight parameters; b_c are bias parameters; $\tanh(\cdot)$ is the hyperbolic tangent function.

The memory cell internal state C_t and the hidden state h_t is updated by:

$$C_t = F_t \otimes C_{t-1} + I_t \otimes \tilde{C}_t$$

$$h_t = O_t \otimes \tanh(C_t)$$

where $\otimes$ is the element-wise product operator.

3.4 Multi-Layer Perceptron

Multi-Layer Perceptron (MLP) is a class of feedforward artificial neural network that consists of at least three layers of nodes: an input layer, a hidden layer, and an output layer. Unlike other neural networks, MLPs use a supervised learning technique called backpropagation for training, which involves adjusting the weights of the network to minimize the difference between the actual output and the predicted output.

Each node in the network, except for the input nodes, is a neuron that uses a nonlinear activation function. MLPs are differentiated from other neural networks by their deep structure with one or more hidden layers. Each neuron in a hidden layer transforms the values from the previous layer with a weighted linear summation, followed by a non-linear activation function.

The capability of MLPs to learn non-linear models and non-linear decision boundaries is one of their primary advantages. They are widely used in applications where complex pattern recognition, such as speech recognition, image recognition, and machine translation, is essential. The depth and size of the MLP can be adjusted to handle different levels of complexity within the data, making MLPs a versatile tool in the realm of artificial neural networks. However, MLPs can be prone to overfitting, particularly with datasets of limited size, requiring techniques such as regularization, dropout, or sufficient training data to mitigate this risk.

4. EXPERIMENTS

We conducted a case study based on the dockless scooter-sharing demand in Austin, TX. During the case study, we compared the performance of our model against leading benchmark models to establish its relative effectiveness and advancements in the field. This section introduces the data, the model setting, and the model comparison results in detail.

4.1 Data

The data used in this study includes daily scooter-sharing data, the weather condition data, and the POI data. This section introduces the data collection process and data descriptions.

We collected the scooter-sharing trip information for Austin from the local government's data-sharing web portal[1]. This data includes scooter-sharing trips made between January 1, 2021, and March 31, 2022. The temporal and spatial distributions of the scooter-sharing trips are presented in Figure 3 and Figure 4.

[1] https://data.austintexas.gov/

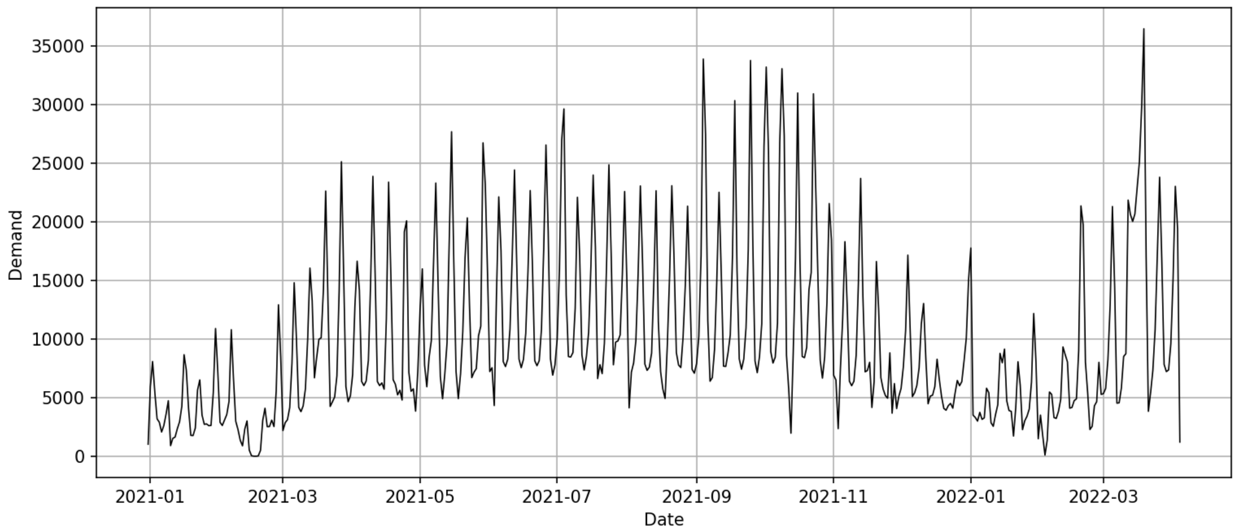

Figure 3. Temporal Distribution of Scooter-Sharing Trips

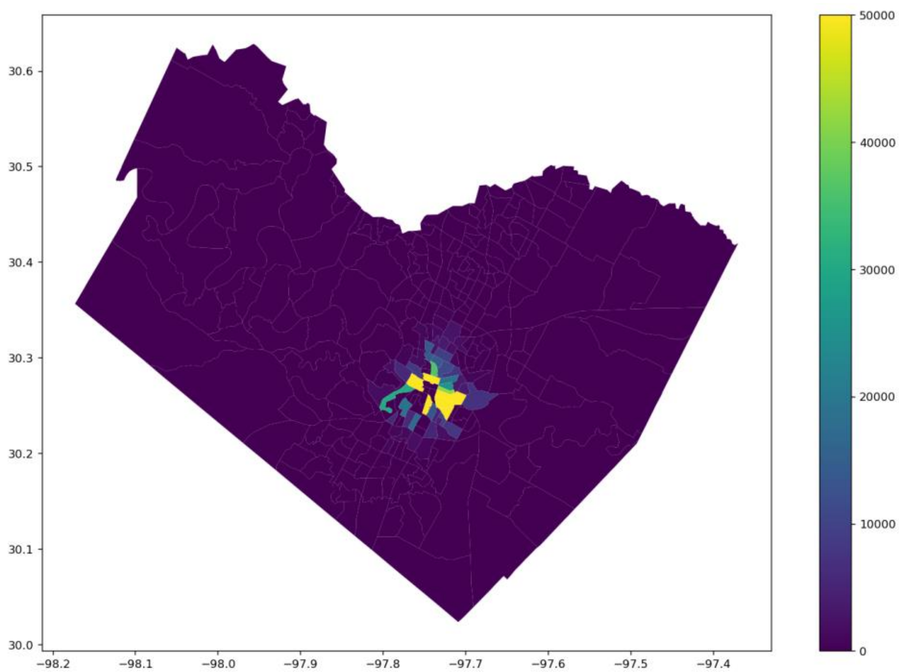

Figure 4. Spatial Distribution of Scooter-Sharing Trips

Weather data—comprising precipitation levels, average temperatures, and wind speeds—were sourced from the Global Historical Climatology Network (GHCN) database. Meanwhile, POI information was gathered from the City of Austin Open Data Portal and the Google Maps API. The collected POIs encompassed recreational facilities, tourist attractions, restaurants, and hotels.

4.2 Model Setting

The experiments were conducted on an NVIDIA RTX A6000 GPU. The batch size was set to 64. The input time sequence length was 14 and the output time sequence length was 1, which means we used the previous 14 days scooter-sharing demand to forecasting the demand in the next day. For the LSTM modules, the number of layers was set to 3 and the number of units per layer was set to 64. For the MLP modules, the number of layers was set to 3, the number of units was set to [32, 64, 32], and the dropout rate was set to 0.5. The model was trained using the RMSprop optimizer and L1 loss function, with an initial learning rate of 0.001. We allocated 60% of the data for training, 20% for validation, and the remaining 20% for testing. Training was conducted over 300 epochs, incorporating an early stopping mechanism to prevent overfitting.

4.3 Model Performance Comparison

We used two metrics, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), to evaluate the performance of the proposed model. The calculation of the two metrics is:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|$$

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

where y_i is the actual observations and $\hat{y}_i$ is the prediction. Note that smaller MAE and RMSE indicate better forecasting accuracy.

We compared the proposed CALSTM model with multiple machine learning models, including Historical Average (HA), Decision Tree (DT), Random Forest (RF), Multi-Layer Perceptron (MLP), and Long Short-Term Memory (LSTM). HA is a foundational statistical approach in time series forecasting that generates prediction by averaging past observations. DT is a machine learning model that uses a tree-like structure of decisions and their possible consequences to perform prediction tasks. RF is an ensemble learning method that constructs multiple decision trees during training and outputs the mean prediction of the individual trees. MLP and LSTM are deep learning methods introduced in Section 3.

The performance of the proposed model and the benchmark models is presented in Table 1. Note that all the models had been fine-tuned. According to the results, we found that the proposed CALSTM model overperformed all benchmark models. Compared with the LSTM model, the improvements of CALSTM model were 28% for MAE and 19% for RMSE. The RNN-based models, LSTM and CALSTM, significantly outperformed the classical machine learning models and the statistical models.

Table 1. Model Performance

Model	MAE	RMSE

HA	16.53	21.18
DT	10.93	16.03
RF	7.79	11.43
MLP	6.59	9.66
LSTM	1.90	2.69
CALSTM	**1.36**	**2.18**

5. CONCLUSIONS

This project introduces a Context-Aware Long Short-Term Memory (CALSTM) recurrent neural network, specifically designed to enhance the prediction of daily travel demand for scooter-sharing in Austin, Texas. By effectively leveraging the impacts of nearby POIs and daily weather variations on scooter usage, the CALSTM model substantially improves prediction accuracy over traditional approaches. It utilizes two separate LSTM modules to process historical scooter-sharing demand and weather data, extracting crucial temporal patterns from these inputs. The outputs of these LSTM modules are then synthesized through element-wise multiplication, establishing robust temporal dependencies essential for accurate demand forecasting. Additionally, the model employs a MLP to analyze POI data, thereby extracting spatial dependencies that further refine the accuracy of its predictions. Another MLP module integrates these spatial and temporal insights to generate the final forecast outputs. This layered approach ensures that both spatial and temporal factors are meticulously accounted for, enhancing the model's predictive capabilities.

The effectiveness of the CALSTM model was rigorously tested through a series of experiments as part of a case study in Austin, TX. The results were impressive, showing that the CALSTM model not only outperformed selected benchmark models but did so with substantial margins—achieving improvements of 28% in MAE and 19% in RMSE compared to conventional LSTM models. These significant enhancements highlight the model's potential to transform transportation planning and improve the operational efficiency of shared micromobility systems in urban areas.

Incorporating the influence of nearby points-of-interest (POIs) and daily weather conditions into forecasting models like the CALSTM can significantly improve accuracy because these factors directly affect user behavior in shared transportation systems, such as scooter-sharing.

POIs, such as commercial centers, tourist attractions, and recreational facilities, inherently draw varying volumes of traffic, which fluctuates daily and seasonally. Understanding the geographical distribution and the type of these POIs allows models to predict potential demand surges and declines more accurately. For example, a scooter-sharing station near a popular café or museum may experience spikes in demand during opening hours, while stations near office buildings might see increased demand at the beginning and end of workdays. By analyzing the spatial relationships between scooter stations and these POIs, forecasting models can anticipate changes in demand based on likely user destinations and origins.

Weather conditions also play a crucial role in the usage patterns of transportation services, especially in modes sensitive to environmental factors, like scooters. Adverse weather conditions such as rain or extreme temperatures can significantly reduce the likelihood of using scooters, while favorable weather can increase usage. By integrating real-time or forecasted weather data into predictive models, the accuracy of predicting daily travel demand can be enhanced. This integration allows the models to adjust demand forecasts based on anticipated weather changes, providing operators with the information needed to optimize fleet management and deployment strategies.

The findings from this study provide critical insights that could inform future urban planning and policy-making efforts, aiming to optimize the integration and management of shared micromobility solutions. By adapting to and predicting changes in urban mobility demand more effectively, city planners and transportation authorities can better accommodate the dynamic needs of urban populations, ensuring that shared transportation resources are utilized optimally to benefit both users and the environment.

6. FUTURE WORK

Building on the successes of the CALSTM model, several avenues for future research and development present exciting opportunities to further enhance the predictive accuracy and utility of our forecasting models for scooter-sharing demand. First, to refine the granularity of our demand forecasts, transitioning from daily to hourly predictions could significantly enhance operational strategies for scooter-sharing systems. Hourly forecasting would allow us to capture the intraday variability in user behavior, which is influenced by factors such as work schedules, meal times, and social events. This level of detail can aid in more dynamically allocating resources throughout the day and better managing peak and off-peak periods. Implementing hourly forecasts requires a detailed analysis of time-series data and may involve adjusting existing models to accommodate the faster cycles of demand fluctuation.

Secondly, the Transformer model, known for its effectiveness in handling sequences, offers a promising alternative to LSTM due to its ability to process entire sequences simultaneously and its enhanced capability for capturing long-range dependencies. By adopting Transformer models, we can potentially overcome some of the limitations inherent in LSTMs. This change could lead to improvements in learning complex temporal patterns more effectively, thus providing more accurate predictions.

Third, incorporating more comprehensive spatial information, such as the built environment and demographic data, could significantly refine the model's contextual awareness. Factors such as population density, urban design, public transit availability, and socioeconomic profiles of neighborhoods can influence scooter usage patterns. By integrating these spatial variables, our models could offer deeper insights into how different segments of the population engage with scooter-sharing systems under varying urban conditions. This approach would not only improve the accuracy of demand predictions but also assist in strategic planning and targeted marketing efforts, ensuring that services are tailored to meet the diverse needs of the community.

REFERENCES

1. NACTO. (2024). Shared Micromobility in 2023. New York, NY.
2. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. Advances in neural information processing systems 30
3. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780.
4. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078
5. Noland, R. B. (2021). Scootin'in the rain: Does weather affect micromobility?. Transportation research part A: policy and practice, 149, 114-123.
6. Xu, Y., Ke, Q., Zhang, X., & Zhao, X. (2024). ICN: Interactive convolutional network for forecasting travel demand of shared micromobility. GeoInformatica, 1-26.
7. Xu, Y., Zhao, X., Zhang, X., & Paliwal, M. (2023). Real-time forecasting of dockless scooter-sharing demand: A spatio-temporal multi-graph transformer approach. IEEE Transactions on Intelligent Transportation Systems, 24(8), 8507-8518.
8. Caspi, O., Smart, M. J., & Noland, R. B. (2020). Spatial associations of dockless shared e-scooter usage. Transportation Research Part D: Transport and Environment, 86, 102396.
9. Jiao, J., & Bai, S. (2020). Understanding the shared e-scooter travels in Austin, TX. ISPRS International Journal of Geo-Information, 9(2), 135.
10. McKenzie, G. (2019). Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in Washington, DC. Journal of transport geography, 78, 19-28.
11. Bai, S., Jiao, J., Chen, Y., & Guo, J. (2021). The relationship between E-scooter travels and daily leisure activities in Austin, Texas. Transportation research part D: transport and environment, 95, 102844.
12. Nikiforiadis, A., Paschalidis, E., Stamatiadis, N., Raptopoulou, A., Kostareli, A., & Basbas, S. (2021). Analysis of attitudes and engagement of shared e-scooter users. Transportation research part D: transport and environment, 94, 102790.
13. Shah, N. R., Guo, J., Han, L. D., & Cherry, C. R. (2023). Why do people take e-scooter trips? Insights on temporal and spatial usage patterns of detailed trip data. Transportation research part A: policy and practice, 173, 103705.
14. Liu, M., Seeder, S., & Li, H. (2019). Analysis of e-scooter trips and their temporal usage patterns. Institute of Transportation Engineers. ITE Journal, 89(6), 44-49.
15. Xu, Y., Yan, X., Sisiopiku, V. P., Merlin, L. A., Xing, F., & Zhao, X. (2022). Micromobility trip origin and destination inference using general bikeshare feed specification data.

Transportation Research Record, 2676(11), 223-238.

16. Hosseinzadeh, A., Algomaiah, M., Kluger, R., & Li, Z. (2021). E-scooters and sustainability: Investigating the relationship between the density of E-scooter trips and characteristics of sustainable urban development. Sustainable cities and society, 66, 102624.
17. Cao, Z., Zhang, X., Chua, K., Yu, H., & Zhao, J. (2021). E-scooter sharing to serve short-distance transit trips: A Singapore case. Transportation research part A: policy and practice, 147, 177-196.
18. Christoforou, Z., de Bortoli, A., Gioldasis, C., & Seidowsky, R. (2021). Who is using e-scooters and how? Evidence from Paris. Transportation research part D: transport and environment, 92, 102708.
19. Laa, B., & Leth, U. (2020). Survey of E-scooter users in Vienna: Who they are and how they ride. Journal of transport geography, 89, 102874.
20. Lee, H., Baek, K., Chung, J. H., & Kim, J. (2021). Factors affecting heterogeneity in willingness to use e-scooter sharing services. Transportation Research Part D: Transport and Environment, 92, 102751.
21. Sanders, R. L., Branion-Calles, M., & Nelson, T. A. (2020). To scoot or not to scoot: Findings from a recent survey about the benefits and barriers of using E-scooters for riders and non-riders. Transportation Research Part A: Policy and Practice, 139, 217-227.
22. Merlin, L. A., Yan, X., Xu, Y., & Zhao, X. (2021). A segment-level model of shared, electric scooter origins and destinations. Transportation Research Part D: Transport and Environment, 92, 102709.
23. Ke, J., Qin, X., Yang, H., Zheng, Z., Zhu, Z., & Ye, J. (2021). Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph convolutional network. Transportation Research Part C: Emerging Technologies, 122, 102858.
24. Geng, X., Li, Y., Wang, L., Zhang, L., Yang, Q., Ye, J., & Liu, Y. (2019, July). Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting. In Proceedings of the AAAI conference on artificial intelligence (Vol. 33, No. 01, pp. 3656-3663).
25. Liu, L., Qiu, Z., Li, G., Wang, Q., Ouyang, W., & Lin, L. (2019). Contextualized spatial–temporal network for taxi origin-destination demand prediction. IEEE Transactions on Intelligent Transportation Systems, 20(10), 3875-3887.
26. Xu, J., Rahmatizadeh, R., Bölöni, L., & Turgut, D. (2017). Real-time prediction of taxi demand using recurrent neural networks. IEEE Transactions on Intelligent Transportation Systems, 19(8), 2572-2581.
27. Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., ... & Li, Z. (2018, April). Deep multi-view spatial-temporal network for taxi demand prediction. In Proceedings of the AAAI conference on artificial intelligence (Vol. 32, No. 1).

28. Lin, L., He, Z., & Peeta, S. (2018). Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach. Transportation Research Part C: Emerging Technologies, 97, 258-276.

29. Li, Y., Zhu, Z., Kong, D., Xu, M., & Zhao, Y. (2019, July). Learning heterogeneous spatial-temporal representation for bike-sharing demand prediction. In Proceedings of the AAAI conference on artificial intelligence (Vol. 33, No. 01, pp. 1004-1011).

30. Kim, T. S., Lee, W. K., & Sohn, S. Y. (2019). Graph convolutional network approach applied to predict hourly bike-sharing demands considering spatial, temporal, and global effects. PloS one, 14(9), e0220782.

31. Chen, P. C., Hsieh, H. Y., Su, K. W., Sigalingging, X. K., Chen, Y. R., & Leu, J. S. (2020). Predicting station level demand in a bike-sharing system using recurrent neural networks. IET Intelligent Transport Systems, 14(6), 554-561.

32. Pan, Y., Zheng, R. C., Zhang, J., & Yao, X. (2019). Predicting bike sharing demand using recurrent neural networks. Procedia computer science, 147, 562-566.

33. Li, H., Wang, J., Ren, Y., & Mao, F. (2021). Intercity online car-hailing travel demand prediction via a Spatiotemporal Transformer Method. Applied Sciences, 11(24), 11750.

34. Tang, J., Liang, J., Liu, F., Hao, J., & Wang, Y. (2021). Multi-community passenger demand prediction at region level based on spatio-temporal graph convolutional network. Transportation Research Part C: Emerging Technologies, 124, 102951.

35. Zhang, X., Zhao, X., Xu, Y., Nilsson, D., & Lovreglio, R. (2024). Situational-aware multi-graph convolutional recurrent network (sa-mgcrn) for travel demand forecasting during wildfires. Transportation Research Part A: Policy and Practice, 190, 104242.

36. Cai, L., Janowicz, K., Mai, G., Yan, B., & Zhu, R. (2020). Traffic transformer: Capturing the continuity and periodicity of time series for traffic forecasting. Transactions in GIS, 24(3), 736-755.

37. Xu, M., Dai, W., Liu, C., Gao, X., Lin, W., Qi, G. J., & Xiong, H. (2020). Spatial-temporal transformer networks for traffic flow forecasting. arXiv preprint arXiv:2001.02908.

38. Ai, Y., Li, Z., Gan, M., Zhang, Y., Yu, D., Chen, W., & Ju, Y. (2019). A deep learning approach on short-term spatiotemporal distribution forecasting of dockless bike-sharing system. Neural Computing and Applications, 31, 1665-1677.

39. Yang, Y., Heppenstall, A., Turner, A., & Comber, A. (2020). Using graph structural information about flows to enhance short-term demand prediction in bike-sharing systems. Computers, Environment and Urban Systems, 83, 101521.

40. Phithakkitnukooon, S., Patanukhom, K., & Demissie, M. G. (2021). Predicting spatiotemporal demand of dockless e-scooter sharing services with a masked fully convolutional network. ISPRS International Journal of Geo-Information, 10(11), 773.

41. Ham, S. W., Cho, J. H., Park, S., & Kim, D. K. (2021). Spatiotemporal demand prediction model for e-scooter sharing services with latent feature and deep learning. Transportation research record, 2675(11), 34-43.

42. Song, J. C., Hsieh, I. Y. L., & Chen, C. S. (2023). Sparse trip demand prediction for shared E-scooter using spatio-temporal graph neural networks. Transportation research part D: transport and environment, 125, 103962.

43. Li, X., Xu, Y., Zhang, X., Shi, W., Yue, Y., & Li, Q. (2023). Improving short-term bike sharing demand forecast through an irregular convolutional neural network. Transportation research part C: emerging technologies, 147, 103984.

44. Ma, C., & Liu, T. (2024). Demand forecasting of shared bicycles based on combined deep learning models. Physica A: Statistical Mechanics and its Applications, 635, 129492.

45. Liang, Y., Zhao, Z., Ding, F., Tang, Y., & He, Z. (2024). Time-dependent trip generation for bike sharing planning: A multi-task memory-augmented graph neural network. Information Fusion, 106, 102294.

46. He, S., & Shin, K. G. (2020, April). Dynamic flow distribution prediction for urban dockless e-scooter sharing reconfiguration. In Proceedings of the web conference 2020 (pp. 133-143).

47. Liang, Y., Ding, F., Huang, G., & Zhao, Z. (2023). Deep trip generation with graph neural networks for bike sharing system expansion. Transportation Research Part C: Emerging Technologies, 154, 104241.

48.