Long term predictions for traffic forecasting
How does the accuracy degrade with time?

Bas Verlooy

Supervisor: Elena Congeduti

[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 13, 2023

Name of the student: Bas Verlooy
Final project course: CSE3000 Research Project
Thesis committee: Elena Congeduti, George Losifidis

Abstract

This research investigates the application of Long Short-Term Memory (LSTM) models for predicting traffic conditions over different time horizons. The study focuses on capturing short-term traffic patterns at a granular 15-minute interval and examines the degradation of model performance as the prediction horizon increases. The results show that the LSTM models achieve an average Root Mean Square Error (RMSE) of 7.6 and Symmetric Mean Absolute Percentage Error (SMAPE) of 28.8% across all sensors, indicating their effectiveness in capturing traffic patterns. However, the models struggle to accurately predict sudden spikes in traffic and exhibit a performance decline when predicting beyond 60 minutes. Despite these challenges, the LSTM models demonstrate resilience in long-term forecasting. Future work could involve expanding the dataset, incorporating external factors such as weather data, and exploring alternative machine-learning approaches to improve accuracy and address the challenges of capturing abrupt traffic fluctuations. Overall, this study highlights the potential of LSTM models in traffic prediction and identifies areas for further research and refinement to enhance their performance in long-term forecasts.

1 Introduction

Traffic forecasting is a critical component in the management and operation of modern transportation systems. Reliable predictions can reduce congestion but also improve traffic safety and efficiency, contributing to the overall performance of the cities traffic flow. However accurately predicting traffic flows over a long period poses significant challenges. The dynamic, non-linear nature of traffic systems influenced by a number of external factors such as the weather, time of day and accidents make long term prediction a difficult task [1].

Methods such as time-series analysis have had some success in predicting traffic patterns, while failing to succeed in long term predictions [2]. Time-series analysis is a method in which past data trends are used to predict the next data points. Traditional models rely on several preconditions, the model structure is predefined, residuals obey the normal distribution and the time series is stationary [3]. These preconditions fail with irregularity of traffic flows which is why model-driven methods fail to deliver reliable performance in the long term [4].

To overcome these challenges, during the research Long Short-Term Memory (LSTM) networks were used. LSTM networks are a type of recurrent neural network that can learn and remember over long sequences, making them well-suited to handle the details and temporal dependencies inherent in traffic data[5]. This paper focuses on the use of LSTM networks in long-term traffic forecasting, providing a comprehensive evaluation of their performance against different time horizons.

The data utilized in this study was gathered by the municipality of The Hague in November 2019. The data collection was facilitated by detection loops embedded in the road surface. Although data for cars, bicycles, and trams was available, the decision was made to focus solely on car data. This choice was motivated by its relevance for comparison with existing literature and the significance of accurate predictions applied to car traffic.

This research paper investigates the efficiency of LSTM (Long Short-Term Memory) networks in dealing with long-term prediction tasks and analyzes the decline in forecasting accuracy over extended time horizons. While conventional models typically forecast traffic conditions for a 15-minute ahead timeframe, this study aims to extend the prediction horizon to 1 or more hours using LSTM networks [6]. The research question of this paper is as follows: 'How does the LSTM model handle long horizon predictions and how does accuracy degrade with time?'

In previous research conducted in the field of traffic prediction, the focus has primarily been on predicting traffic patterns within a relatively short-term period, typically up to 1 hour in advance [7]. This time frame allows for timely decision-making and enables quick responses to traffic congestion or incidents, which cannot be predicted over a long-time horizon.

In some cases, the average speed of vehicles has been the primary target of prediction models [8]. By estimating the average speed, insights into the overall flow and congestion levels on highways or high volume roads have been obtained. These predictions contribute to a better understanding of traffic conditions and assist in optimizing route planning strategies.

The majority of existing research in this domain has centered around predicting traffic conditions specifically on highways or roads with high traffic volumes. These areas have been of particular interest due to the frequency of congestion and the necessity for effective traffic management and planning[9]. The prediction models and algorithms developed in these studies have often focused on addressing the unique characteristics and challenges associated with high-volume road networks.

The paper is structured as follows, the methodology and results section details the approach using LSTM networks, including the data collection and preprocessing steps. It presents an evaluation of LSTM performance across different time horizons and provides an analysis of the accuracy of traffic forecasting. In the responsible research section, considerations regarding data privacy and ethical implications are addressed to ensure the study adheres to responsible research practices. The discussion section delves into the strengths and limitations of LSTM networks in traffic forecasting, drawing comparisons with traditional models and identifying potential areas for improvement. Lastly, the conclusion and future works section summarizes the key findings, explores the implications of LSTM networks in traffic management, and proposes avenues for future research to enhance the accuracy and reliability of long-term traffic predictions.

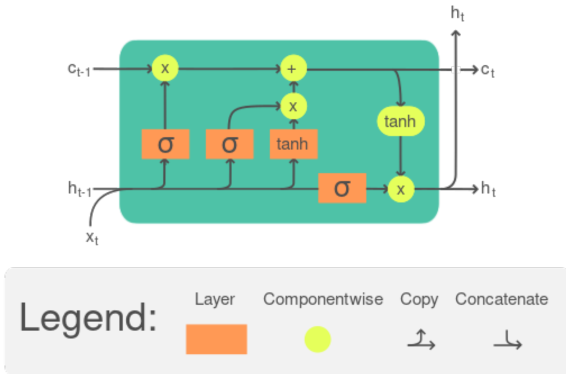

Figure 1: A single LSTM unit. [12]

.

2 Long-Term Traffic Prediction: An LSTM Approach

In this section, we aim to discuss the application of LSTM networks for long-term traffic predictions. Previous research in this field has mainly focused on short-term traffic forecasting. However, the objective of our study is to explore the long-term prediction capabilities of LSTM networks.

2.1 Understanding Long Short-Term Memory Networks

LSTM networks are a type of Recurrent Neural Network (RNN) architecture that have been designed to address some limitations of traditional RNNs[10]. RNNs are known to suffer from the vanishing gradient problem, which hampers their ability to learn from the information in distant parts of the input[11]. LSTM networks overcome this by introducing a memory cell that can maintain information in memory for long periods of time, making them particularly suitable for tasks that involve sequential data with long-term dependencies.

An LSTM unit has three main components: the input gate, the forget gate, and the output gate. These gates collectively determine how much information should be stored or discarded from the cell state at each timestep.

Input Gate: Decides how much of the newly computed information for the current timestep should be stored in the cell.

Forget Gate: Determines how much of the existing information in the cell state should be kept.

Output Gate: Decides what information from the current cell state should be output.

These mechanisms allow LSTM networks to selectively remember or forget things, leading to their ability to handle long-term dependencies in the data.

Hyperparameter tuning is a critical aspect of optimizing the performance of LSTM networks, especially when dealing with sequential or temporal data. While LSTM networks have shown their effectiveness in various tasks, it is important to clarify that the objective of this study is not to create the best LSTM network globally. Rather, the focus is on evaluating the performance of LSTM networks specifically in the context of long-term traffic prediction.

The potential of LSTM networks for long-term traffic prediction lies in their capability to capture intricate temporal dynamics and learn from historical trends [13]. However, to harness the full potential of these networks, careful consideration and tuning of hyperparameters such as the number of layers, units per layer, and learning rate are essential.

While achieving the highest possible accuracy with LSTM networks is desirable, this study places a higher emphasis on understanding their performance in longer-term prediction rather than solely pursuing optimal accuracy. By investigating the performance of LSTM networks in the specific context of long-term traffic prediction, valuable insights can be gained regarding their suitability and potential challenges in real-world traffic forecasting scenarios.

The mathematical validity of LSTM networks has been well-established in previous studies, which serve as a solid foundation for our research [14; 15]. These studies also provide a comprehensive understanding of the LSTM architecture, reinforcing our confidence in the effectiveness of LSTM networks for the task of long-term traffic prediction.

2.2 Methodology

Our research method involves using LSTM networks for our forecasting model. We decided to use LSTM due to its inherent ability to recognize long-term patterns, which is crucial in predicting long-term traffic trends.

The majority of existing studies have relied on datasets with high amounts of traffic. The dataset contains data from urban roads, in contrast to most papers that used data from highways or other busy roads with higher amounts of traffic. Our approach, conversely, uses a dataset that contains the quantity of traffic. This approach is based on the understanding that traffic volume can significantly impact long-term traffic management and planning and due to the available data.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(x_i - \hat{x_i})^2}{N}} \tag{1}$$

To train our LSTM network, we used the Root Mean Square Error (RMSE) as shown in Equation 1 as our learning metric, here x_i is the predicted value and $\hat{x_i}$ is the actual value. RMSE is a commonly used metric for regression models and offers an interpretation of how well our model can predict future data by understanding the average squared difference between the predicted and actual traffic volumes.

SMAPE (Symmetric Mean Absolute Percentage Error) and MAPE (Mean Absolute Percentage Error) are commonly used metrics for evaluating forecast accuracy, with the main difference lying in their calculation of the percentage error. MAPE is computed by taking the absolute difference between the forecasted and actual values, dividing it by the actual value, and multiplying by 100 to obtain the percentage error. However, MAPE's limitation lies in its asymmetry, as it can be more sensitive to under-forecasts than over-forecasts due to its denominator being the actual value.

To address this issue, SMAPE was introduced as an improved version of MAPE. SMAPE takes the absolute difference between the forecasted and actual values, divides it by the average of the forecasted and actual values, and multiplies by 100. By using the average of forecasted and actual values in the denominator, SMAPE ensures equal treatment of under- and over-forecasts. This approach provides a more balanced measurement of forecast accuracy.

The formulas for MAPE and SMAPE are as follows, where F_t is the predicted value and A_t is the actual value:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{|F_t - A_t|}{A_t}\right) \times 100 \tag{2}$$

$$SMAPE = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{|F_t - A_t|}{(A_t + F_t)/2}\right) \times 100 \tag{3}$$

SMAPE's use of the average of forecasted and actual values in the denominator reduces the bias towards under- or over-forecasts, providing a more balanced measurement of forecast accuracy. Moreover, SMAPE is less affected by zero or close to zero actual values, which can cause division by zero issues in MAPE.

It is worth noting that in cases where the actual value is 0, the calculation of SMAPE is excluded in this research. This exclusion is implemented to avoid skewed results and maintain the meaningfulness of the metric. When the actual value is 0 and the prediction doesn't match, the SMAPE formula would result in a value of 200%, leading to misrepresented results. Excluding these cases ensures a more accurate and realistic assessment of the forecast performance, focusing on situations where comparisons between the forecasted and actual values lead to more insightful values.

Overall, SMAPE is considered a more reliable and robust metric for evaluating forecast accuracy, particularly in datasets containing low values or exhibiting asymmetry in forecast errors. It offers a comprehensive and symmetric evaluation of forecast performance, enabling a more accurate assessment of forecasting models.

2.3 Problem Description

The problem we are addressing can be summarized as follows: Given historical traffic data, our objective is to predict future traffic volumes over an extended time period. We operate under the assumption that past traffic patterns contain valuable information for forecasting future trends. LSTMs, being capable of learning from historical data and capturing patterns, are well-suited for this task.

This research aims to contribute to the existing literature by focusing on the application of LSTM networks for long-term traffic volume predictions, which is an area that has not been extensively explored. The choice of using both RMSE (Root Mean Square Error) and SMAPE (Symmetric Mean Absolute Percentage Error) as evaluation metrics reflects our desire to balance the need for accuracy in traffic prediction (measured by RMSE) and the importance of handling proportional differences (addressed by SMAPE).

2.4 Contribution and Significance

The use of both RMSE and SMAPE metrics in this study allows for a comprehensive evaluation of the LSTM model's performance. RMSE, a commonly used metric for regression models, provides insights into how well the model predicts future traffic volumes by measuring the average squared difference between predicted and actual values. On the other hand, SMAPE offers a symmetric approach to handling both over- and under-forecasts, making it an improvement over the traditional MAPE metric. This is particularly advantageous when dealing with datasets that exhibit relatively low traffic volumes.

This study makes two primary contributions. Firstly, it expands the application of LSTM networks beyond the commonly studied area of short-term predictions and explores their utility for long-term traffic forecasting. Secondly, it uses the SMAPE on a dataset with lower values, these lower values can skew the calculations and thus have to be taken into account.

It is important to note that the aim of this research is not to revolutionize the field or introduce a groundbreaking approach. Instead, it explores an uncommon area in the literature, extending the known advantages of LSTM networks to a domain of application and refining evaluation techniques to better accommodate real-world traffic data. This paper seeks to fill an existing gap in the literature and serves as a foundation for future explorations and advancements in this field.

3 Experimental Setup and (future) Results

This section elucidates the specific configurations of our experimental setup and presents the results obtained from the application of our LSTM-based traffic forecasting model. The methodology is designed in a manner that allows for replication by a proficient reader for result validation.

3.1 Experimental Setup

The experiments were conducted on a MacBook Pro equipped with an M1 chip, 16GB of RAM, and running macOS 13.5. For the development and execution of the LSTM model, Python 3.9.7 was used along with the TensorFlow 2.12.0 library [1]. This setup is not recommended as Tensorflow is not supported out of the box and requires a separate install in order to work [2]. The performance of Tensorflow is also suboptimal and will not work as well as with non-ARM M1 Macbooks or Windows or Linux setups.

The dataset used in this study was collected in November 2019 from detection loops embedded in roads across 11 intersections in The Hague. The dataset comprises readings from a total of 130 sensors, providing a basis dataset for our analysis. The data is aggregated per 15 minutes and shows the number of vehicles passing a sensor during that time period.

For the preprocessing of this data, null values were addressed by forward filling (ffill), which is a method of filling null values in a time series dataset with the previous data point. This helps to maintain the continuity and integrity of

[1] https://www.tensorflow.org/

[2] https://developer.apple.com/metal/tensorflow-plugin/

the data, particularly important when dealing with time-series data. Additionally, the day and hour features were extracted from the timestamp in the dataset's index for further analysis.

Lastly, the traffic data was normalized before being fed into the model. Normalization is a common preprocessing step for neural networks, as it scales the data to a small range of values between 0 and 1. This helps to speed up the training process and can also help to avoid numerical instability issues. In our case, Z-score normalization was used, a method that standardizes the features by removing the mean and scaling to unit variance. This operation involves transforming each feature value into a score that reflects how many standard deviations it is from the mean of the feature. This ensures that our LSTM network would receive data that has been scaled to a comparable range, aiding in the training process by mitigating issues of scale across features and accelerating convergence.

The LSTM network was constructed with 2 LSTM layers and 1 dense layer, and a batch size of 32 was used for training the model. Although our model was set to train for 500 epochs, we utilized an Early Stopping callback with a patience of 10. This mechanism allows our model to stop training when it no longer improves, saving computational resources and preventing overfitting. The learning rate was set to 0.001, and the Adam optimizer was used during the learning process. The dataset was split into 22 days for training, 3 days for validation, and 5 days for testing.

3.2 Performance Evaluation Method

The primary focus of our method was on evaluating the model's capability to predict traffic volume at future time steps, where each step corresponds to 15-minute intervals. To achieve this, a time-lagged labeling technique was used, where the label for a given instance was the traffic volume at a future time step, i. For example, if $i = 1$, the model was tasked with predicting the traffic volume 15 minutes into the future; if $i = 5$, the model predicted the volume 75 minutes ahead, and so on.

This evaluation method was executed in a stepwise manner, advancing by one hour (i.e., four 15-minute steps) at a time. In this manner, we were able to cover a wide range of future time steps, starting from $i = 1$ (15 minutes into the future) up to $i = 41$ (10 hours and 15 minutes into the future).

To ensure a robust evaluation and mitigate the effects of randomness or noise created by the initial conditions of the LSTM, for each time step i, 20 individual runs of the LSTM prediction were performed. The final predicted value for each time step was then computed as the average of these 20 runs. This approach provided a more reliable estimate of the LSTM model's predictive performance at each future time step.

Through this method, we were able to provide a detailed performance analysis of our LSTM-based model, assessing its ability to make accurate predictions over various future time frames, and understanding the model's strengths and weaknesses in different prediction scenarios.

The subsequent sections present the results of this rigorous evaluation process, which provides valuable insights into the LSTM model's long-term traffic prediction capabilities. The insights gained through this method are key to understanding the potential and limitations of LSTM networks in the field of traffic forecasting and planning.

3.3 Results

The results section presents the evaluation of the LSTM model in predicting traffic conditions, with a specific focus on forecasting 15 minutes into the future. It is important to note that in this study, the data was aggregated per 15 minutes, which aligns with the chosen prediction timeframe. As such, the 15-minute prediction interval corresponds to one timestep in the LSTM model. The achieved results provide insights into the accuracy and performance of the LSTM model in capturing short-term traffic patterns, specifically at this granular 15-minute interval. The obtained RMSE value of 7.6 indicates the average deviation between the predicted and actual values within this time step, while the SMAPE value of 28.8% reflects the percentage difference between the predicted and actual values. These results contribute to understanding the model's effectiveness in capturing and predicting traffic patterns at this temporal resolution.

Table 1: Performance Metrics of LSTM Model for 15 minutes in the future

Metric	All sensors	Single sensor
RMSE	7.6	7.3
SMAPE	28.8	26.2

Figure 2 presents the mean average of all predictions for 130 sensors, illustrating the overall performance of the LSTM model. The graph showcases that the predicted values closely align with the actual values, indicating a strong correspondence. This suggests that the model captures the general trends and patterns in the traffic data effectively across multiple sensors. The RMSE value, which reflects the overall accuracy of the model's predictions, is consistent with the results obtained for all sensors combined. However, it is important to note that Figure 3 highlights the predictions for a single sensor. In this graph, despite the overall RMSE remaining the same, it becomes apparent that individual sensor predictions exhibit sporadic spikes that are challenging to detect accurately. These spikes can introduce additional errors, leading to a less smooth and more volatile prediction pattern for a single sensor. This finding underscores the complexity of predicting traffic conditions at a granular level and highlights the need for further investigation and refinement to capture and account for such localized variations.

The performance of the LSTM model is observed to degrade as the prediction horizon extends from 1 timestep to 5 timesteps in the future. This degradation is evident in the increasing RMSE and SMAPE values, as shown in Table 2. However, beyond 5 timesteps, the performance degradation becomes minimal, indicating that the model is able to maintain relatively stable accuracy in predicting traffic patterns over longer time horizons. To illustrate, Figure 4 showcases the predictions for 41 timesteps (equivalent to 10 hours) in the future. While the model captures the general flow of traffic, it struggles to accurately predict spikes or sudden variations in traffic, as evidenced by the deviations between the

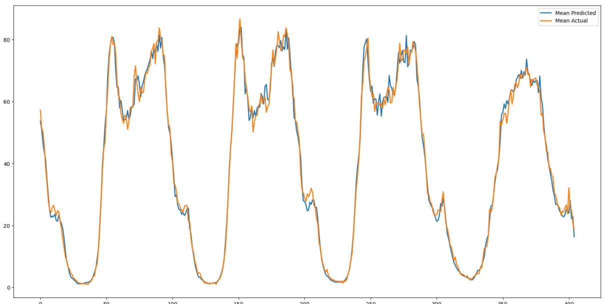

Figure 2: Predicting 15 minutes in the future, with 130 sensors. TODO: make x-axis correct

.

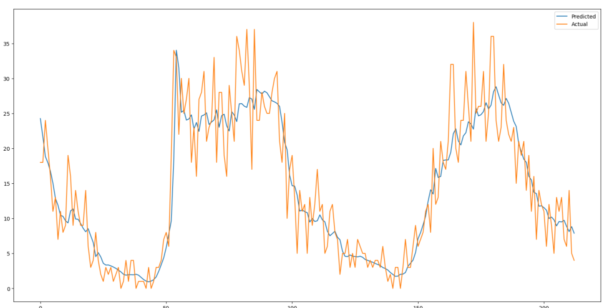

Figure 3: Predicting 15 minutes in the future, with 1 sensor. TODO: make x-axis correct

.

predicted and actual values. Notably, when comparing the predictions for a single timestep and 41 timesteps, the model performing the single timestep prediction demonstrates the ability to pick up sudden spikes in traffic, as observed around the 70th timestep, whereas the 41-timestep prediction fails to capture these abrupt changes and follows the smoother trend. This highlights the challenge of forecasting fine-grained traffic fluctuations over extended time horizons using the LSTM model.

Table 2: Performance Metrics of LSTM Model for different timesteps

Minutes\Metric	RMSE	SMAPE %
15	7.6	28.8
30	7.8	29.5
45	7.9	29.9
60	8.1	30.3
75	8.1	30.1
135	8.4	31.1
315	8.5	30.9
615	8.7	31.3

4 Responsible Research

Conducting research into LSTM-based traffic forecasting requires careful attention to both ethical considerations and reproducibility. These fundamental principles intertwine within

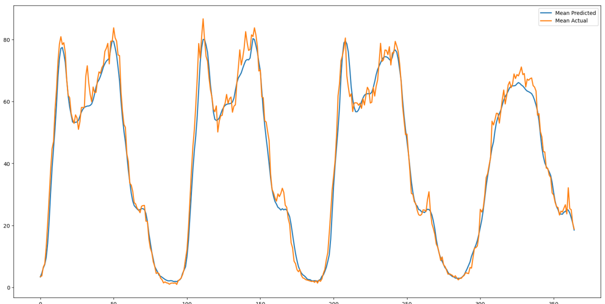

Figure 4: Predictions 10 hours in the future, with 130 sensors. TODO: make x-axis correct

.

the broader view of responsible computational research, including in the field of traffic forecasting.

Foremost among the ethical considerations is data privacy and security. Even though the traffic data used is aggregated into 15-minute intervals and therefore anonymized, which dilutes individual movement patterns into broader traffic flows, the potential risk of re-identification cannot be entirely dismissed. However, the process of high-level aggregation makes it exceptionally challenging to trace back to a specific individual, enhancing data privacy.

On the matter of reproducibility, a strong emphasis is placed on the transparency of methodology. While the specific dataset used is confidential, the aggregated dataset is available in the GitHub repository, furthermore it is underscored that the focus of the study lies more on the long-term LSTM-based forecasting methodology, which is universally applicable. The LSTM architecture, hyperparameters, data preprocessing techniques, and evaluation metrics are detailed exhaustively, thereby ensuring reproducibility given similar traffic datasets.

Despite the acknowledged challenges due to the stochastic nature of neural network training, multiple runs are incorporated and average performance metrics are reported, to provide a more reliable and reproducible assessment of the model.

Additionally, the adaptability of the model across varying traffic patterns, regions, and data quality is highlighted as a challenge for reproducibility and contextual validity. Continued model refinement is thus suggested as a necessary step.

By focusing on these ethical and reproducibility considerations, the study contributes to a more transparent and ethically conscious practice of AI research in traffic forecasting. The focus on methodological transparency over specific data underscores the study's potential for effective replication while prioritizing privacy and ethical considerations.

5 Discussion

The primary goal of this study was not necessarily to design the best-performing LSTM model for traffic prediction but rather to investigate the capability of LSTM models in capturing and predicting long-term traffic flow patterns. In this context, our findings have provided some insights.

The results demonstrate that the LSTM model can effectively capture and predict traffic patterns at a granary 15-minute interval. With an RMSE of 7.6 and SMAPE of 28.8% for the all-sensor model for a single timestep prediction, these results align with our expectations and serve as a validation of LSTM's application in short-term traffic forecasting.

It is important to note that the effectiveness of LSTM models can vary depending on the context and dataset size. For instance, an RMSE of 6.75 was achieved in a paper, which is lower than ours [16]. However, they were working with a dataset three times larger than ours, offering a wider base for the model to learn from, which likely contributed to their lower error rate.

Similarly, they reported a MAPE of 17.14% is significantly lower than our SMAPE of 28.8%. However, their data was derived from highway traffic, typically registering higher traffic volumes. When working with smaller values the (S)MAPE can be prone to high values. As MAPE is a relative error metric, larger traffic volumes can skew the error percentages. Therefore, the comparison of error metrics needs to take into account these contextual differences.

In the context of long-term horizon predictions, our study revealed that the LSTM model experiences a decrease in performance as the prediction horizon extends from 1 to 5 timesteps. Beyond 5 timesteps, however, the model's performance stabilized, suggesting that despite the initial decline, the LSTM can maintain a consistent level of predictive accuracy over longer periods.

Yet, it's evident from our results that while LSTM models demonstrate a reliable capability for traffic flow prediction, certain challenges persist. These notably include predicting granular-level traffic fluctuations and maintaining a high degree of accuracy over extended forecast horizons.

6 Conclusions and Future Work

This research aimed to assess the effectiveness of Long Short-Term Memory (LSTM) models in predicting traffic conditions, with a focus on longer-term horizons. The key research question revolved around the LSTM model's capacity to maintain accuracy and performance in forecasting traffic patterns at granular 15-minute intervals over extended periods.

The conclusions drawn from this research are that LSTM models do present a valid approach for short-term traffic forecasting. With an average Root Mean Square Error (RMSE) of 7.6 and Symmetric Mean Absolute Percentage Error (SMAPE) of 28.8% across all sensors, these models exhibited their capability to capture traffic patterns effectively.

However, the study also exposed some challenges in applying LSTM models to granular-level traffic prediction. Specifically, sudden spikes in traffic flow proved difficult to predict accurately. Moreover, as the prediction horizon extended from 15 minutes (1 timestep) to 60 minutes (4 timesteps), a noticeable degradation in the LSTM model's performance was observed. Yet, interestingly, the performance appeared to stabilize when predicting beyond 60 minutes, indicating a degree of resilience in the LSTM model's long-term forecasting ability.

Despite this, the model struggled to accurately capture sudden traffic fluctuations over extended periods. This highlights one of the key challenges in long-term traffic forecasting - managing the fine balance between capturing broader trends and detecting immediate, abrupt changes in traffic flow.

Looking ahead, there are several potential paths for future work in this area. Firstly, having a larger dataset could potentially improve the performance of the LSTM model, as observed in the lower error rates reported by Yu et al. [16], who used a dataset three times larger than ours.

Secondly, incorporating external factors, such as real-time weather data or event data, could enhance the model's ability to anticipate sudden changes in traffic flow. This could address the noted challenge of accurately predicting sudden traffic spikes.

Finally, exploring other machine learning approaches or refining the current LSTM model may help improve the accuracy of long-term forecasts, particularly in capturing abrupt changes in traffic flow.

In conclusion, this study underlines LSTM models' potential in traffic prediction tasks and identifies key challenges and areas for further investigation. As cities around the globe strive towards smarter, more efficient transportation systems, continued research in this area holds substantial promise for contributing to this paper.

References

[1] Eleni I. Vlahogianni, John C. Golias, and Matthew G. Karlaftis. Short-term traffic forecasting: Overview of objectives and methods. *Transport Reviews*, 24(5):533–557, 2004.

[2] Chun-Hsin Wu, Jan-Ming Ho, and D.T. Lee. Travel-time prediction with support vector regression. *IEEE Transactions on Intelligent Transportation Systems*, 5(4):276–281, 2004.

[3] J. Zheng and M. Huang. Traffic flow forecast through time series analysis based on deep learning. *IEEE Access*, 8:82562–82570, 2020.

[4] W. Yanan F. Xiaoshi and M. Zhang. Network traffic forecasting model based on long-term intuitionistic fuzzy time series. *Information Sciences*, 506:131–147, 2020.

[5] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9:1735–80, 12 1997.

[6] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang. A hybrid deep learning based traffic flow prediction method and its understanding. *Transportation Research Part C: Emerging Technologies*, 90:166–180, 2018.

[7] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias. Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies*, 43:3–19, 2014. Special Issue on Short-term Traffic Flow Forecasting.

[8] T. Bogaerts, A. D. Masegosa, J. S. Angarita-Zapata, E. Onieva, and P. Hellinckx. A graph cnn-lstm neural network for short and long-term traffic forecasting

based on trajectory data. *Transportation Research Part C: Emerging Technologies*, 112:62–77, 2020.

[9] H. Farah, J. Liu, X. Wang, Y. Li, X. Kang L., and Gao. Method of evaluating and predicting traffic state of highway network based on deep learning. *Journal of Advanced Transportation*, 2021:8878494, 2021.

[10] J. Yue W. Bao and Y. Rao. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS ONE*, 12, 2017.

[11] S. Hochreiter. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 06(02):107–116, 1998.

[12] G. Chevalier. Larnn: Linear attention recurrent neural network, 2018.

[13] N. K. Manaswi. *RNN and LSTM*, pages 115–126. Apress, Berkeley, CA, 2018.

[14] F. Kratzert, D. Klotz, C. Brenner, K. Schulz, and M. Herrnegger. Rainfall–runoff modelling using long short-term memory (lstm) networks. *Hydrology and Earth System Sciences*, 22(11):6005–6022, 2018.

[15] V. Venkatasubramanian. The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE Journal*, 65(2):466–478, 2019.

[16] H. Yu, Z. Wu, S. Wang, Y. Wang, and X. Ma. Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks. *Sensors*, 17(7), 2017.