Harnessing LSTM Networks for Traffic Flow Forecasting: A Deep Learning Approach

Ahmad Mustafa[1], Khurram Shehzad Khattak[1], Zawar Hussain Khan[2]
[1]Dept. of Computer Systems Engineering UET Peshawar Peshawar, Pakistan
[2]College of Computer Science and Engineering University of Ha'il Ha'il, KSA
*Correspondence: cadet.mustafa2011@gmail.com, khurram.s.khattak@gmail.com, z.hussain@uoh.edu.sa

Citation | Mustafa. A, Khattak. K. S, Khan. Z. H, "Harnessing LSTM Networks for Traffic Flow Forecasting: A Deep Learning Approach", IJIST, Vol. 07 Special Issue. pp 329-337, May 2025.

Received | April 17, 2025 **Revised |** May 18, 2025 **Accepted |** May 21, 2025 **Published |** May 23, 2025.

Accurate traffic flow forecasting in areas with different types of vehicles and varied driving behaviors is crucial for improving urban transportation systems and reducing congestion. In this paper, we introduce a Long Short-Term Memory (LSTM) approach to predict short-term traffic flow in such diverse conditions. Our model uses time-series data from real-world traffic sensors, capturing the patterns and dependencies that occur over time in mixed traffic environments. We tested the model using a dataset from seven days, with six days for training and one day for testing. The LSTM model achieved an R2 value of 0.96, a Mean Squared Error (MSE) of 2.82, and a Mean Absolute Error (MAE) of 1.13. These results demonstrate the effectiveness of LSTM networks in predicting traffic flow in complex traffic conditions, surpassing traditional machine learning models. This study provides valuable insights into using deep learning techniques for intelligent transportation systems (ITS).

Keywords: Traffic Flow Prediction, LSTM, Deep Learning, Congestion, Intelligent Transportation System

Introduction:

Accurate and timely traffic flow prediction is key to developing Intelligent Transportation Systems (ITS), which aim to improve urban mobility, reduce congestion, and enhance road safety [1][2]. Traffic flow forecasting is essential for traffic management, dynamic route planning, and decision-making, especially in busy urban areas. However, predicting traffic flow becomes particularly challenging in heterogeneous traffic, where various vehicle types—such as cars, buses, motorbikes, and bicycles—follow different driving behaviors and occupy different amounts of road space.

While traditional statistical methods work well in simpler traffic environments, they struggle to handle the complexities of heterogeneous traffic. Machine Learning (ML) techniques like Random Forests (RF) and Support Vector Machines have been applied to model the non-linear nature of traffic patterns. However, these models often fail to capture temporal dependencies and long-term interactions, especially in dynamic and diverse traffic conditions [3].

Recently, deep learning models, particularly Long Short-Term Memory (LSTM) networks, have shown promise for time series forecasting. LSTM networks are well-suited for traffic flow prediction because they can capture long-term dependencies and handle non-linear relationships in sequential data [4][5][6]. LSTMs are particularly effective at modeling the complex temporal patterns in traffic data, which is inherently sequential. By remembering past inputs, LSTMs can better predict future traffic conditions, even in environments where vehicle interactions are complex and traffic flows are multi-modal.

This paper proposes an LSTM-based deep learning model for short-term traffic flow forecasting, specifically designed to address heterogeneous traffic conditions. We use real-world traffic data collected over seven days, with six days for training and one day for testing. The dataset includes various traffic parameters, such as vehicle speeds, classifications, traffic density, and flow rates, allowing the model to learn the interactions between different vehicle types.

Our results show that LSTM networks can effectively model traffic flow in mixed traffic conditions, achieving an R2 value of 0.96, a Mean Squared Error (MSE) of 2.82, and a Mean Absolute Error (MAE) of 1.13. These results highlight the power of LSTM networks in capturing temporal dependencies in heterogeneous traffic and their superiority over traditional ML models for traffic flow prediction.

This study focuses on developing a predictive model for heterogeneous traffic flow using Long Short-Term Memory (LSTM) networks, aiming to forecast average vehicle speed across various vehicle types and conditions. The model's performance is assessed using key metrics such as R^2, Mean Squared Error (MSE), and Mean Absolute Error (MAE) to evaluate its accuracy and generalization. The impact of temporal and lagged features on prediction accuracy is also explored, highlighting the importance of effective data preprocessing. Additionally, the LSTM model is benchmarked against existing traffic forecasting approaches to demonstrate its effectiveness. Ultimately, the study aims to support Intelligent Transportation Systems (ITS) by providing insights that enhance traffic management and reduce congestion.

The rest of the paper is organized as follows: Section II reviews related work on traffic flow prediction using ML and deep learning methods. Section III describes the methodology, including data preprocessing, model design, and training. Section IV presents the experimental results and discusses model performance. Finally, Section V concludes the paper with insights and future research directions.

Related Work:

Traffic flow forecasting has become increasingly important in traffic management and urban planning. Various methods, including traditional statistical models and advanced machine learning (ML) techniques, have been proposed to predict traffic patterns. This section summarizes key contributions in this field, with a focus on the use of Long Short-Term Memory (LSTM) networks for processing heterogeneous traffic data.

Gao et al. [7] proposed a traffic flow forecasting method based on multitemporal traffic flow volume, utilizing the relationship between flow and speed. They evaluated five models—LSTM, backpropagation (BP), CART, KNN, and SVR—using real-world data from the Beijing Third Ring Expressway (June 1–15, 2009). The dataset included variables such as time, traffic flow, speed, lane number, and heavy vehicle composition. The results showed that LSTM outperformed other models, offering the highest prediction accuracy with an average Mean Absolute Percentage Error (MAPE) of 3.29 %, followed by BP (5.31 %) and CART (6.85%). KNN (11.13%) and SVR (11.62%) had the highest errors.

In [8], the authors combined Recurrent Neural Networks (RNNs) with Multi-Agent Systems (MAS) to detect highway traffic anomalies by comparing traffic flows with normal patterns. The model, using LSTM networks to capture periodic data and long-term dependencies, processes univariate time series generated from daily vehicle flow data. MAS is used to create datasets based on city events, with data transmitted through the Kafka Framework. The model detected traffic anomalies by comparing predicted and actual data, confirming its accuracy in identifying outliers.

Chu et al. [9] proposed a traffic flow prediction model using LSTM networks to improve prediction accuracy. The dataset spanned 8 months, and recorded daily and hourly traffic flow across 243 days. The performance of LSTM, Gated Recurrent Unit (GRU), and Stacked Autoencoders (SAEs) was compared using six evaluation metrics. LSTM outperformed other models, achieving the lowest Mean Absolute Error (MAE) of 6.83 and the highest R^2 value of 0.93, demonstrating superior accuracy.

Bouchemoukha et al. [10] introduced LSTM-TF, an LSTM-based model, designed to address challenges in traffic forecasting, such as spatial correlations and non-linear temporal dynamics. LSTM-TF incorporates multiple LSTM layers to handle both temporal and spatial features. Compared to models like HA, ARIMA, T-GCN, SVR, RNN, and GRU, LSTM-TF consistently outperformed across various prediction horizons. Using the SZ-taxi and Los-loop datasets, LSTM-TF reduced RMSE by 9.29%, 46.19%, and 6.01% in a 15-minute forecasting task, proving its effectiveness in handling non-stationary time series data.

In [11], the authors addressed the challenges of forecasting stochastic traffic flow time series by using deep learning techniques. They proposed an LSTM-based prediction model and compared it with traditional models—ARIMA and BPNN—using traffic data from OpenITS. The dataset, collected at 5-minute intervals over 10 days from an inductive loop on No. 1 Yuanda Road, Changsha, showed that LSTM outperformed ARIMA and BPNN with lower RMSE values and a MAPE of 4.82%, highlighting the superiority of deep learning in traffic flow forecasting.

The authors in [4] aimed to improve short-term traffic volume forecasting with deep learning models based on LSTM networks. They evaluated four models: simple LSTM, LSTM encoder-decoder, Conv-LSTM, and CNN-LSTM, using traffic data from Austin, Texas. Among these models, Conv-LSTM demonstrated the highest performance for a 15-minute time horizon, with an RMSE of 49.23 and a MAPE of 9.03%, confirming its accuracy for short-term forecasting.

Abduljabbar et al. [12] used LSTM networks to predict vehicle speed using real-time spatial and temporal data from traffic sensors. The LSTM model achieved high accuracy, ranging from 88% to 99% for outbound traffic and 96% to 98% for inbound traffic, over prediction horizons of 5 to 60 minutes. The model outperformed others, maintaining accuracy over distances up to 15 km, and was tested on data collected from Melbourne's Eastern Freeway.

In [5], the authors compared the performance of statistical models (VAR and ARIMAX) with LSTM for multivariate short-term traffic volume forecasting. The study included traffic volume, speed, and waiting time data, along with weather information from Austin, Texas. The VAR model outperformed the others with an accuracy of 91.46%. This comprehensive analysis of traffic patterns and environmental factors further confirmed the potential of LSTM for accurate forecasting.

In [13], the authors proposed M-B-LSTM, a hybrid deep learning model designed to tackle the challenges of stochasticity and distribution imbalance in traffic flow patterns. M-B-LSTM integrates an online self-learning network to balance data distributions, and a deep bidirectional LSTM (DBLSTM) to process forward and reverse contexts, reducing uncertainty. The model outperformed existing models on the Changchun and I90 datasets, demonstrating its ability to handle traffic flow's stochastic nature.

The study in [3] proposed an LSTM-based traffic flow forecasting model and compared it with traditional methods. The model was tested using data from the California Transportation Performance Measurement System (PeMS). The LSTM model achieved RMSE values of 8.4 for training and 9.86 for test data, and MAE values of 5.52 and 6.83, respectively. These results showed that LSTM outperformed classical statistical methods, demonstrating its effectiveness for traffic flow prediction.

The study in [6] introduced GLSTM-A, a hybrid LSTM model for efficient traffic flow prediction on edge devices without spatial data access. The GLSTM-A combines a Grid LSTM for long-term dependencies with a standard LSTM for short-term analysis, significantly improving prediction accuracy and memory efficiency. Tested on datasets from PeMS and Hyderabad, GLSTM-A outperformed other models, achieving the lowest RMSE, MAE, and MAPE values.

This paper builds on previous research by leveraging LSTM networks to forecast heterogeneous traffic flow. The proposed model addresses the limitations of earlier methods by employing deep learning techniques to capture complex temporal patterns in traffic data. The following sections present our methodology and results, demonstrating the effectiveness of LSTM networks in enhancing traffic flow predictions.

Methodology:

This section outlines the methodology used to forecast heterogeneous traffic flow using LSTM networks, covering data collection, preprocessing, model architecture, training procedures, and implementation Figure 1.

Data Collection:

Traffic data was systematically collected [14] over a span of seven consecutive days, from Monday to Sunday, between 9:00 AM and 5:00 PM, with minute-by-minute recordings of vehicle counts. The dataset comprises various traffic parameters, including time, vehicle count, traffic flow speed, and detailed counts of vehicle types such as cars, buses, bicycles, and motorbikes. Additional metrics such as flow rate, peak hour factor, traffic density, and time-distance headway were also recorded to ensure a comprehensive understanding of traffic dynamics. The data was obtained from a sensor node deployed along University Road in Peshawar, Pakistan, yielding a total of 2,880 data points, with Thursday's data reserved for testing purposes. In total, the dataset encompasses 17 distinct features, providing a rich foundation for in-depth traffic flow analysis.

Upon analyzing the collected data, we observed distinct traffic patterns. For example, traffic volume tended to be higher on Mondays, likely due to the start of the workweek, while Thursday saw the lowest traffic flow, possibly due to reduced commuter activity in the mid-week. These variations in traffic flow across different days of the week provided important insights into the temporal dependencies that the LSTM model could leverage for forecasting.

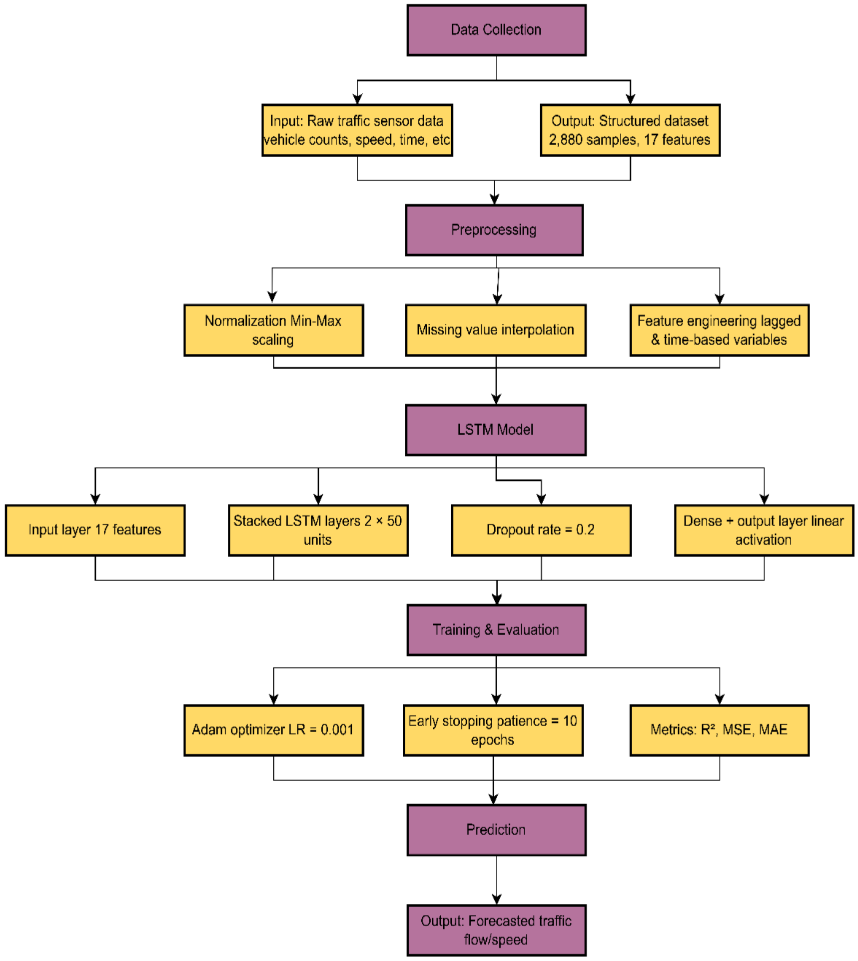

Figure 1 Flow Diagram

Data Preprocessing:

To ensure the dataset was suitable for model training, several preprocessing steps were undertaken. First, Min-Max normalization was applied to continuous variables such as traffic flow and speed to accelerate convergence during the training process. To address missing values and maintain data integrity, interpolation techniques were employed to estimate and fill in the gaps. Additionally, feature engineering was performed to enhance the dataset's predictive capacity. Temporal features such as the time of day and the day of the week were introduced to capture diurnal and weekly traffic variations. Furthermore, lagged variables from previous minutes were incorporated to enable the model to learn underlying temporal patterns in traffic flow behavior.

Model Architecture:

The proposed model architecture begins with an input layer that accepts 17 features, including traffic flow, speed, and time-based categorical variables. This is followed by two stacked Long Short-Term Memory (LSTM) layers, each comprising 50 units, designed to capture both short-term and long-term temporal dependencies inherent in traffic data. To mitigate the risk of overfitting, dropout layers with a rate of 0.2 were incorporated after each LSTM layer. A fully connected dense layer was then employed to map the learned temporal representations to the target variable. Finally, the output layer uses a linear activation function to predict the average speed of vehicles.

Training Procedures:

To evaluate the model's performance, two train-test split scenarios were explored. In the first scenario, the model was trained on six days of data, while Thursday's data—

completely unseen during training—was used for testing. In the second scenario, the dataset from all seven days was split into 70% for training and 30% for testing. The model was trained using a batch size of 32 for 100 epochs, with a learning rate of 0.001 and the Adam optimizer employed for optimization. To prevent overfitting, early stopping was implemented, halting the training process if there was no improvement in validation loss over 10 consecutive epochs. The model's effectiveness was ultimately assessed by evaluating its predictions of average vehicle speed on the test data.

Implementation:

The implementation of the LSTM model was carried out using Python, with TensorFlow and Keras serving as the primary frameworks for model construction and training. For data preprocessing tasks—including normalization and feature engineering—pandas and NumPy were utilized to efficiently manage and transform the dataset.

This methodology effectively combines LSTM networks with careful data preprocessing and engineering to forecast traffic flow patterns, demonstrating the model's capability to learn temporal dependencies in traffic data.

Results:

The performance of the LSTM model for forecasting heterogeneous traffic flow was evaluated using several metrics to assess its accuracy and generalization. Here's an overview of the evaluation results:

The model was initially trained using a dataset spanning six days—Monday, Tuesday, Wednesday, Friday, Saturday, and Sunday—where 70% of the data was used for training and the remaining 30% for testing. To further validate the model's generalization capability, an additional evaluation was conducted by training on the same six days and testing exclusively on data from Thursday, which was completely withheld during the training phase to ensure it remained unseen.

Figure 2 illustrates the training and validation loss curves over the course of 100 epochs. The loss values exhibited rapid convergence within the initial few epochs, followed by a sharp decline, suggesting that the model quickly learned the underlying traffic patterns. After approximately 10 epochs, both training and validation losses stabilized at low values, indicating that the model did not overfit the data and demonstrated strong generalization capability on unseen inputs.

The model's performance was quantitatively assessed using multiple evaluation metrics. The coefficient of determination (R^2) was calculated to be 0.9629, indicating that approximately 96.29% of the variance in average vehicle speed is explained by the model, reflecting strong predictive capability. The Mean Squared Error (MSE) was found to be 2.8194, suggesting that the model's predictions closely align with the actual values and exhibit low overall prediction error. Additionally, the Mean Absolute Error (MAE) was recorded at 1.1325, further confirming the model's accuracy by demonstrating minimal average deviation between predicted and observed values.

Conclusion:

The results highlight that the LSTM model effectively captures complex traffic flow dynamics, showing strong generalization to both training and unseen data. The high R^2 value, along with low MSE and MAE, confirm the model's accuracy and robustness in forecasting heterogeneous traffic flow.

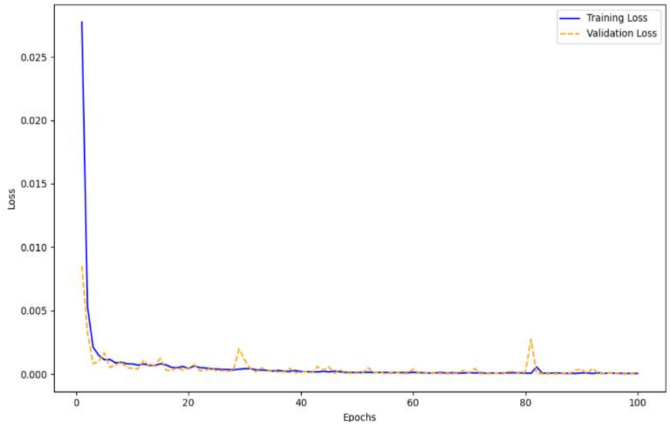

Figure 2. Training and Validation Loss Curves

Table 1. Performance Comparison of Lstm Models Using Different Data Splits

Metric	Model Trained on 6 days, Tested on Thursday	Model Trained on Seven days data (70% Train, 30% Test)
R^2	0.9629	0.9616
MSE	2.8194	3.1814
MAE	1.1325	1.0822

The comparison of both approaches is shown in Table 1. Although the R^2 values of both models are very similar, with only a small decrease in the 70%-30% split model (0.9616 vs. 0.9629), the model trained on six days and tested on Thursday performed better in terms of MSE, with a lower value of 2.8194 compared to 3.1814. However, the current model slightly improved in terms of MAE (1.0822 vs. 1.1325), showing a smaller average prediction error. These results suggest that both models are highly effective in predicting heterogeneous traffic flow, but the model trained on six days and tested on Thursday has a slight edge in explaining variance and minimizing prediction error.

In both configurations, the models used a simple train-test split approach rather than k-fold cross-validation. Given the similar performance of both models, this approach provided reliable insights into the model's robustness. Future work could consider using k-fold cross-validation to better capture the variability across different subsets of the data.

To further demonstrate the performance of the LSTM model, Figure 3 compares the predicted and actual average vehicle speed. The close agreement between the predicted and actual values highlights the high accuracy of the LSTM model in predicting traffic dynamics.

The daily vehicle counts varied, with Monday typically showing higher traffic volumes due to commuter rush, while Thursday exhibited lower counts, providing an opportunity to evaluate the model's ability to generalize across different traffic conditions. For testing, Thursday's data was excluded, allowing for validation on an unseen day's traffic pattern. These variations in traffic statistics are integral to the model's ability to learn and predict heterogeneous traffic patterns, as the LSTM network is capable of handling such daily fluctuations in traffic flow.

The results from our traffic flow prediction model, with an R^2 of 0.9629, MSE of 2.8194, and MAE of 1.1325, show strong predictive accuracy and align well with findings from other studies. For example, Gao et al. [7] reported that LSTM outperformed other models with a mean absolute percentage error (MAPE) of 3.29%, and Chu et al. [9] achieved an R^2 of 0.9307 and an MAE of 6.828492, both demonstrating LSTM's effectiveness in traffic forecasting. Similarly, Abduljabbar et al. [12] achieved high LSTM accuracy ranging from 88% to 99% using Melbourne traffic data, while the GLSTM-A model in [6] outperformed alternatives, with an RMSE of 1.155 and MAE of 0.882. Compared to these studies, our model performs similarly or better in terms of R^2 and MAE, further validating the strength and efficiency of LSTM-based architectures for traffic flow prediction. Our lower MSE and higher R^2 indicate strong model performance, especially when compared to traditional statistical models or simpler neural networks.

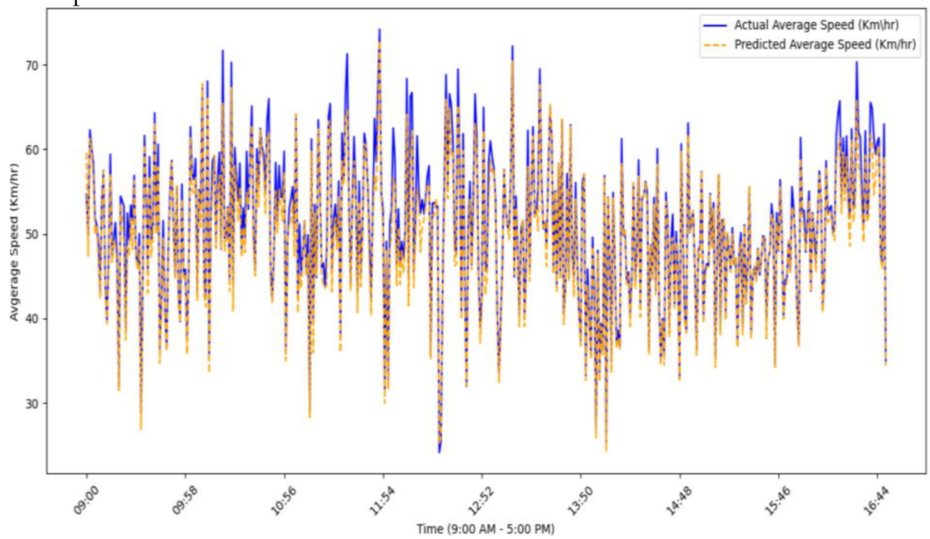

Figure 3. Predicted vs. Actual Average Speed per minute on Thursday 13-January-2022

Conclusion:

This study presents a thorough approach to forecasting heterogeneous traffic flow using LSTM networks, a powerful deep learning technique. The LSTM model was trained on a diverse dataset that includes various vehicle types, showcasing its ability to capture complex traffic patterns.

The model's evaluation results reveal strong accuracy, with an R-squared value of 0.9629, MSE of 2.8194, and MAE of 1.1325. These metrics highlight the model's effectiveness in predicting average vehicle speed, making it a valuable tool for Intelligent Transportation Systems (ITS). Accurate forecasting of heterogeneous traffic flow can lead to better traffic management, reduced congestion, and improved road safety.

Future research could integrate additional data sources, such as real-time traffic conditions and weather data, to further improve prediction accuracy. Expanding the model to explore different deep learning architectures could also provide valuable insights into the robustness of traffic flow forecasting.

In conclusion, this study demonstrates the potential of LSTM networks for traffic flow prediction, opening the door for advanced applications in smart transportation solutions that can create more efficient and safer road networks.

References:

[1] T. A. G. Ali Zeb, Khurram S. Khattak, Muhammad Rehmat Ullah, Zawar H. Khan, "HetroTraffSim: A Macroscopic Heterogeneous Traffic Flow Simulator for Road Bottlenecks," *Futur. Transp*, vol. 3, no. 1, pp. 368–383, 2023, doi: https://doi.org/10.3390/futuretransp3010022.

[2] A. N. K. ALI ZEB, KHURRAM S. KHATTAK, AREEB AGHA, , ZAWAR H. KHAN, M. ATHAR JAVED SETH, "ON-BOARD DIAGNOSTIC (OBD-II) BASED CYBER PHYSICAL SYSTEM FOR ROAD BOTTLENECKS DETECTION," *J. Eng. Sci. Technol.*, vol. 17, no. 2, pp. 0906–0922, 2022, [Online]. Available: https://jestec.taylors.edu.my/Vol 17 Issue 2 April 2022/17_2_07.pdf

[3] P. K. Baskar and H. Kaluvan, "Long short-term memory (LSTM) recurrent neural network (RNN) based traffic forecasting for intelligent transportation," *AIP Conf. Proc.*, vol. 2435, no. 1, Mar. 2022, doi: 10.1063/5.0083590/2823318.

[4] D. Haputhanthri and A. Wijayasiri, "Short-term traffic forecasting using LSTM-based

deep learning models," *MERCon 2021 - 7th Int. Multidiscip. Moratuwa Eng. Res. Conf. Proc.*, pp. 602–607, Jul. 2021, doi: 10.1109/MERCON52712.2021.9525670.
[5] A. W. B. Dissanayake, O. Hemachandra, N. Lakshitha, D. Haputhanthri, "A Comparison of ARIMAX, VAR and LSTM on Multivariate Short-Term Traffic Volume Forecasting," *28th Conf. Open Innov. Assoc. Fruct*, 2021, [Online]. Available: https://www.researchgate.net/publication/351389331_A_Comparison_of_ARIMAX_VAR_and_LSTM_on_Multivariate_Short-Term_Traffic_Volume_Forecasting
[6] R. Biju, S. U. Goparaju, D. Gangadharan, and B. Mandal, "Grid LSTM based Attention Modelling for Traffic Flow Prediction," *2024 IEEE 99th Veh. Technol. Conf.*, pp. 1–7, Jun. 2024, doi: 10.1109/VTC2024-SPRING62846.2024.10683344.
[7] Y. W. and S. L. Y. Gao, C. Zhou, J. Rong, "Short-Term Traffic Speed Forecasting Using a Deep Learning Method Based on Multitemporal Traffic Flow Volume," *IEEE Access*, vol. 10, pp. 82384–82395, 2022, doi: 10.1109/ACCESS.2022.3195353.
[8] S. M. Snineh, N. E. A. Amrani, M. Youssfi, O. Bouattane, and A. Daaif, "Detection of traffic anomaly in highways by using recurrent neural network," *5th Int. Conf. Intell. Comput. Data Sci. ICDS 2021*, 2021, doi: 10.1109/ICDS53782.2021.9626741.
[9] Q. Chu, G. Li, R. Zhou, and Z. Ping, "Traffic Flow Prediction Model Based on LSTM with Finnish Dataset," *2021 IEEE 6th Int. Conf. Intell. Comput. Signal Process. ICSP 2021*, pp. 389–392, Apr. 2021, doi: 10.1109/ICSP51882.2021.9408888.
[10] H. Bouchemoukha, M. N. Zennir, and A. Lahoulou, "Is Classical LSTM more Efficient than Modern GCN Approaches in the Context of Traffic Forecasting?," *Proc. - 2021 IEEE Int. Conf. Recent Adv. Math. Informatics, ICRAMI 2021*, 2021, doi: 10.1109/ICRAMI52622.2021.9585940.
[11] J. Zheng and M. Huang, "Traffic Flow Forecast Through Time Series Analysis Based on Deep Learning," *IEEE Access*, vol. 8, pp. 82562–82570, 2020, doi: 10.1109/ACCESS.2020.2990738.
[12] S. L. Rusul L. Abduljabbar, Hussein Dia, Pei-Wei Tsai, "Short-Term Traffic Forecasting: An LSTM Network for Spatial-Temporal Speed Prediction," *Futur. Transp*, vol. 1, no. 1, pp. 21–37, 2021, doi: https://doi.org/10.3390/futuretransp1010003.
[13] Q. Zhaowei, L. Haitao, L. Zhihui, and Z. Tao, "Short-Term Traffic Flow Forecasting Method with M-B-LSTM Hybrid Network," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 1, pp. 225–235, Jan. 2022, doi: 10.1109/TITS.2020.3009725.
[14] A. Asif Khan, Khurram S. Khattak, Zawar H. Khan, Thomas Aaron Gulliver, "Edge Computing for Effective and Efficient Traffic Characterization," *Sensors*, vol. 23, no. 23, p. 9385, 2023, doi: https://doi.org/10.3390/s23239385.