Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 817-834 **e-ISSN: 2656-8632**

RESEARCH ARTICLE OPEN ACCESS

Advanced Traffic Flow Optimization Using Hybrid Machine Learning and Deep Learning Techniques

Mohammed El Kaim Billah, Abdelfettah Mabrouk
Department of Computer Science ESTSB, ELITES Lab, Chouaib Doukkali University, El Jadida, Morocco

Corresponding author: Mohammed El Kaim Billah (e-mail: mohammed.kaimbillah@gmail.com), **Author(s) Email**: Abdelfettah Mabrouk (e-mail: mabroukdes@gmail.com)

ABSTRACT. Road traffic congestion remains a persistent and critical challenge in modern urban environments, adversely affecting travel times, fuel consumption, air quality, and overall urban livability. To address this issue, this study proposes a hybrid ensemble learning framework for accurate short-term traffic flow prediction across signalized urban intersections. The model integrates Random Forest, Gradient Boosting, and Multi-Layer Perceptron within a weighted voting ensemble mechanism, wherein model contributions are dynamically scaled based on individual validation performance. Benchmarking is performed against traditional and advanced baselines, including Linear Regression, Support Vector Regression, and Long Short-Term Memory (LSTM) networks. A real-world traffic dataset, comprising 56 consecutive days of readings from six intersections, is utilized to validate the approach. A robust preprocessing pipeline is implemented, encompassing anomaly detection, temporal feature engineering especially time-of-day and day-of-week normalization, and sliding window encoding to preserve temporal dependencies. Experimental evaluations on 4-intersection and 6-intersection scenarios reveal that the ensemble consistently outperforms all baselines, achieving a peak R^2 of 0.954 and an RMSE of 0.045. Statistical significance testing using Welch's t-test confirms the reliability of these improvements. Furthermore, SHAP-based interpretability analysis reveals the dominant influence of temporal features during high-variance periods. While computational overhead and data sparsity during rare events remain limitations, the framework demonstrates strong applicability for deployment in smart traffic systems. Its predictive accuracy and model transparency make it a viable candidate for adaptive signal control, congestion mitigation, and urban mobility planning. Future work may explore real-time streaming adaptation, external event integration, and generalization across heterogeneous urban networks.

Keywords: **Urban Traffic Control; Traffic Light Controller; Traffic Flow Prediction; Deep Learning; Machine Learning; Ensemble Model**.

1. INTRODUCTION

In the era of smart cities, optimizing traffic flow has become a critical aspect of urban planning and management. Traffic congestion is a pervasive problem in urban areas worldwide, causing frustration, wasted time, increased fuel consumption, and environmental pollution [1]. As populations grow and urbanization continues, traffic congestion becomes even more severe, posing significant challenges to transportation systems and urban planners [2]. Several factors contribute to traffic congestion, including the exponential growth of urban populations [3], [4], [5], resulting in a corresponding rise in the number of vehicles on the road, especially during peak hours [6], [7], [8]. Additionally, many cities grapple with limited road infrastructure that fails to keep pace with the escalating number of vehicles, exacerbating congestion during rush hours and major events [9], [10], [11]. Inefficient traffic management strategies, characterized by poorly synchronized traffic signals, a lack of real-time monitoring, and ineffective traffic flow control [12], [13], [14], further contribute to congestion hotspots [15], [16]. Moreover, human behavior, encompassing car dependency [17], [18], commuting patterns, and individual driving habits, significantly influences traffic congestion patterns. The adverse effects of congestion on productivity, air quality, and overall quality of life underscore the critical need for effective solutions. One such solution gaining prominence is traffic flow prediction [19], [20], [21], which plays a vital role in proactively managing congestion by informing decision-makers about expected traffic patterns. This allows for adjustments in traffic signal timings, resource deployment, and the optimization of public transportation services. At a 4-way intersection, traffic relies on well-coordinated signals or stop signs dictating the right-of-way for each road. Traffic lights alternate phases, guiding vehicles through the intersection seamlessly. Consideration for turning movements ensures left-turning vehicles find protected paths while right-turning ones yield appropriately. Pedestrian crossings are marked and synchronized with signals, providing safe passage. Traffic volume governs signal durations, and congestion management strategies come into play in high-density scenarios.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

On the other hand, six-way intersections, with their added complexity, demand sophisticated signalization, lane configuration, and pedestrian safety measures [22]. These intersections serve as intricate hubs where engineers employ advanced techniques to optimize traffic flow, ensuring a harmonious coexistence of diverse movements and modes of transportation. Predicting traffic flow is greatly enhanced by integrating machine learning (ML) and deep learning (DL) techniques, revolutionizing transportation authorities' congestion management. ML algorithms can analyze vast amounts of historical traffic data to identify patterns and trends, enabling more accurate predictions of future traffic conditions [23]. DL models, with their ability to process complex data and learn hierarchical representations, further refine traffic flow predictions by capturing intricate relationships between various factors influencing traffic patterns, such as time of day, weather conditions, and special events [24]. By harnessing ML and DL technologies, transportation authorities can develop advanced traffic prediction systems that provide real-time insights into traffic dynamics. These systems can dynamically adapt to changing traffic conditions, allowing authorities to implement responsive measures to alleviate congestion and optimize traffic flow.

Aim of the Study: This study aims to develop an ensemble machine learning architecture that improves traffic flow prediction and optimization in smart cities. The model integrates multiple ML and DL techniques to provide more accurate forecasts, enabling adaptive traffic control measures. This study makes the following contributions:

1. Novel Ensemble Framework: Introduces a hybrid ensemble learning approach that combines ML and DL models to optimize traffic flow predictions.
2. Comprehensive Hyperparameter Tuning: Implements advanced tuning techniques, such as grid search and random search, to optimize model performance.
3. Data Preprocessing and Feature Engineering: Explores missing data imputation, normalization, and feature selection to improve model robustness.
4. Evaluation and Benchmarking: Compares the proposed method against state-of-the-art techniques using multiple performance metrics.

The remainder of this study is structured as follows: Section II reviews the current state-of-the-art research in traffic flow prediction. Section III introduces the proposed ML and DL models for traffic flow in smart cities. Section IV presents the experimental analysis, results, and discussion. Finally, Section V concludes the study.

Previous studies using machine learning approaches to forecast traffic flow in vehicular networks are investigated in [25], [26]. Utilizing real-time traffic data, the study creates prediction models using machine learning algorithms such as neural networks, decision trees, random forests, and support vector machines. The authors assess these models' performance using preprocessing, feature engineering, and model optimization, using metrics such as mean absolute error and prediction accuracy. The findings demonstrate the effectiveness of various machine learning techniques in anticipating traffic flow, providing information that can be used to improve transportation systems and optimize traffic management in vehicular networks. The study in [27] builds a variety of pattern classifiers, such as Adaboost, SVM, RF, and SVR algorithms, for the classification of vehicles using machine learning concepts. With a focus on Support Vector Regression (SVR), an algorithm adapted for regression tasks based on support vector machine principles, this study presents an optimized SVR short-term traffic flow prediction model by adjusting SVM parameters. According to the testing results, SVR has the lowest classification error rate (3.22%). The proposed model shows significant reductions in Root Mean Square Error (RMSE) by 29.71% and 47.22% during morning and night peak hours, respectively, and in Mean Absolute Percentage Error (MAPE) by 19.94% and 42.86%. By efficiently obtaining the best possible parameter combinations, the enhanced SVR algorithm improves the accuracy of traffic flow forecasts.

To address possible data outliers, a variety of data denoising techniques were used in the referenced study [28], including Wavelet (WL), Ensemble Empirical Mode Decomposition (EEMD), and Empirical Mode Decomposition (EMD). Subsequently, short-term traffic flow prediction was performed using the Long Short-Term Memory (LSTM) neural network. Three traffic flow datasets from the Caltrans Performance Measurement System (PeMS) were used in the evaluation. With an average Root Mean Square Error (RMSE) of 0.79, Mean Absolute Error (MAE) of 0.60, and Mean Absolute Percentage Error (MAPE) of 2.14, the results showed that the LSTM+EEMD technique provided greater accuracy.

A proposed adaptive traffic signal system using machine learning algorithms to anticipate traffic flow was presented in [29]. After implementing and comparing five regression models (linear regression, random forest, decision tree, gradient boosting, and k-nearest neighbors), the random forest regressor performed the best ($R^2 = 0.98$). By modifying the timing of green and red lights according to anticipated traffic density, the system reduced traffic congestion by 30.8%. Simulation results confirmed the system's efficiency in optimizing traffic flow at intersections.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

A key component of adaptive traffic management systems is the forecasting of traffic flow at intersections, which is addressed in [30] using machine learning (ML) and deep learning (DL) techniques. The study utilized two publicly available datasets, one of which recorded vehicle counts at six intersections at 5-minute intervals over a period of 56 days. Four intersections were used to train the models. The Multi-Layer Perceptron Neural Network (MLP-NN) achieved an excellent R^2 and explained variance (EV) score of 0.93 while requiring less training time. All ML and DL algorithms showed strong performance metrics, indicating their suitability for integration with smart traffic light controllers.

Graph Convolutional Network (GCN) and Long Short-Term Memory (LSTM) neural networks were integrated to create the innovative spatiotemporal dynamic deep network known as GCN-LSTM, which was first presented in [31]. This model addresses uncertainty, periodicity, and spatiotemporal dynamics in traffic flow analysis. While temporal dynamics are captured by an attention mechanism and LSTM, GCN processes spatial correlation features. The usefulness and superiority of the proposed GCN-LSTM model were confirmed by simulation tests.

The study in [32] offers a novel solution to the challenges of randomness and nonlinearity in traffic flow prediction. The proposed approach combines Adaptive Hybrid Exponential Smoothing with Residual Correction (AHES-RC) and Quantum Particle Swarm Optimization (QPSO). By combining single and double exponential smoothing techniques, the AHES component adapts its weights in real-time to traffic pattern variations. An Extreme Learning Machine (ELM) algorithm controls the residual correction, while QPSO optimizes the AHES-RC parameters to improve prediction accuracy. Utilizing 26 real-world datasets, a thorough evaluation of the QPSO-AHES-RC method against other benchmark models revealed notable improvements, with mean RMSE and mean MAPE typically reduced by more than 20% compared to advanced machine learning techniques like XGBoost and CatBoost. The study emphasizes the importance of dynamically collecting traffic flow data in real-time to enhance prediction accuracy.

2. MATERIALS AND METHOD

A. Dataset

This study utilizes a publicly available dataset from the Huawei Munich Research Center, consisting of vehicle count data collected from six urban intersections at 5-minute intervals over a period of 56 days. The dataset includes a total of 16,128 samples, offering a high-resolution temporal view of traffic flow suitable for short-term prediction and traffic signal control optimization. The data were gathered using induction loop sensors, anonymized, and prepared for modeling. Fig. 1 describes the flow of the methodology used.

Our approach begins with the utilization of road traffic prediction datasets from the Huawei Munich Research Center [33]. Comprehensive exploratory analysis is

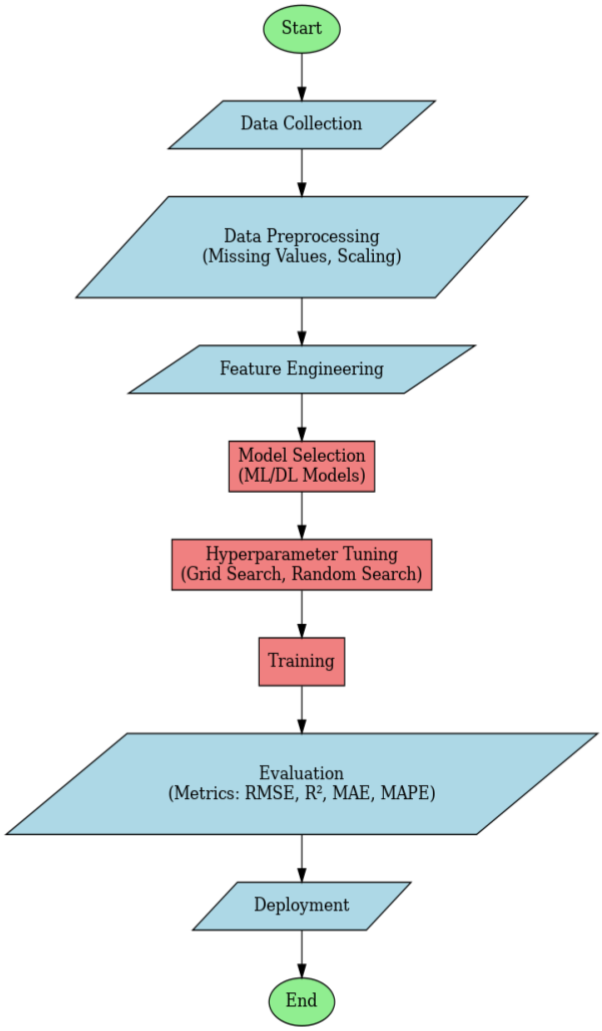


Fig. 1. The step-by-step processing pipeline of our proposed predictive approach.

conducted following meticulous data preprocessing, including handling missing data. Additional columns for date and time are added, and traffic flow is plotted for one day, one week, and average values to identify patterns. The dataset can be used to modify traffic signal control settings and forecast traffic patterns. It contains 56 days of recorded data from six metropolitan intersections. It is appropriate for short-term traffic forecasting because it contains flow time series data showing the number of cars passing at 5-minute intervals throughout the day [32]. In this study, the traffic flow at a 4-way intersection is simulated using data from four out of the six intersections.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

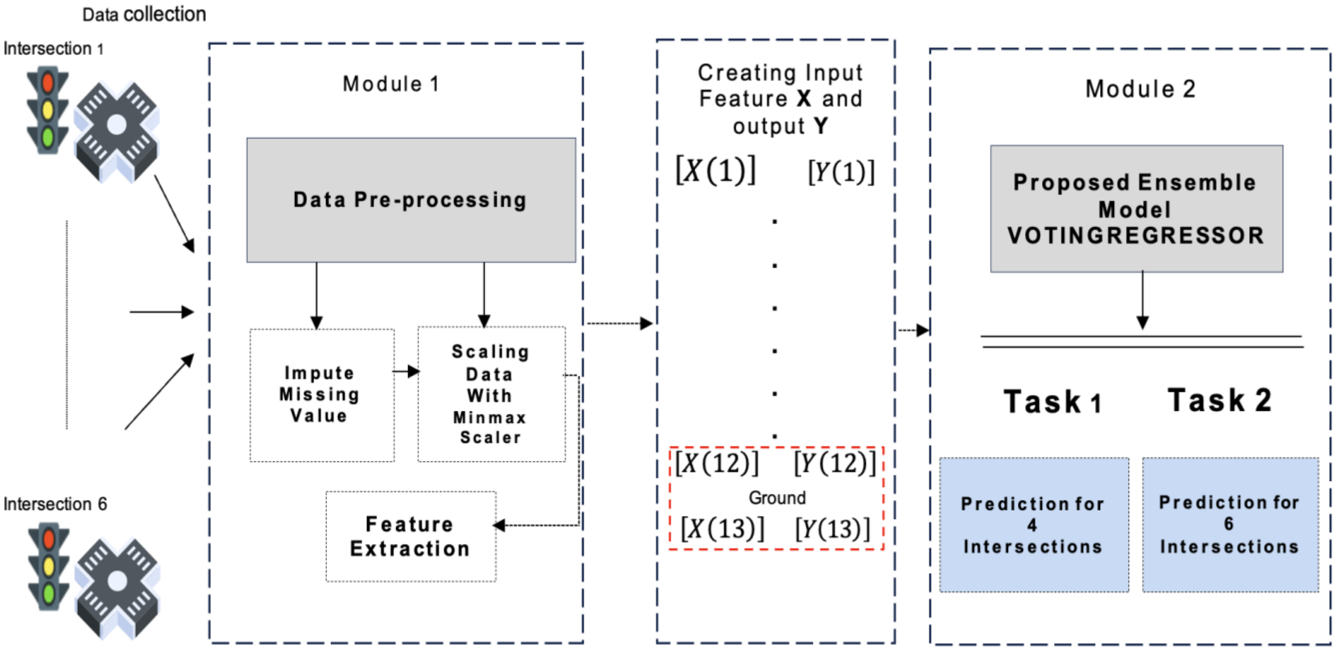


Fig. 2. **High-level overview of the proposed traffic prediction framework**.

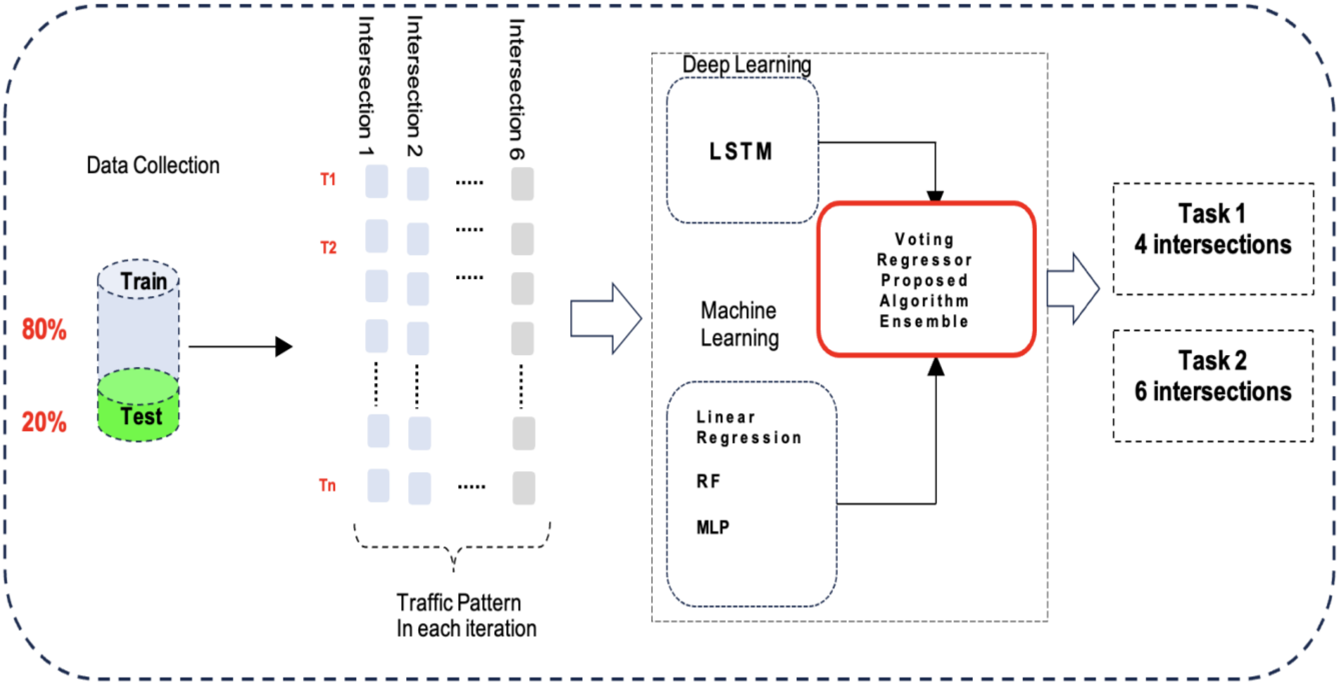


Fig. 3. **Conceptual steps of the ensemble algorithm integrating deep and machine learning models**

Subsequently, two primary tasks are executed. The first task involves predicting traffic flow at four intersections and applying this model to predict flow at six intersections.

The second task involves using both input and output from six intersections. A crucial step in the methodology involves creating a list X as input and a list Y as output for modeling. X is generated by considering 12 consecutive rows of data, with the 13th row serving as the ground truth Y. The dataset is then divided into training and testing sets in an 80:20 ratio. Fig. 2 provides an overview of the proposed framework. This study delineates a system that combines machine learning, deep learning, and an ensemble model

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

tailored for predicting traffic flow at intersections using the Huawei Munich Research Center road traffic dataset [33].

The machine learning models employed include linear regression and Multi-Layer Perceptron (MLP), featuring three distinct activation functions: ReLU, logistic, and tanh. The deep learning model adopts an LSTM architecture with 50 neurons and uses the Adam optimizer, trained for 10 epochs with a batch size of 516. Table 1 provides a summary of the simulation details. The proposed ensemble model integrates the three best-performing machine learning models using a voting regressor. Fig. 3 shows the steps of the prediction process.

B. Data Collection

The dataset employed in this study is publicly accessible and originates from the Huawei Munich Research Center [33]. It consists of anonymized road traffic prediction data gathered from various traffic sensors, primarily induction loops, strategically placed at six urban intersections. Covering a span of 56 days, the dataset provides flow time series information, recording vehicle counts at 5-minute intervals throughout each day. This temporal resolution results in 12 hourly readings and 288 daily readings, culminating in 16,128 observations over the entire period. Notably, the dataset is crucial for forecasting traffic patterns and optimizing traffic signal control parameters, such as cycle length, offset, and split times. Its applications extend to traffic management, urban planning, and transportation research, enabling detailed analysis and predictive modeling of traffic dynamics.

C. Data Preprocessing Pipeline

In this section, the analytical process involves augmenting the dataset by adding date and time columns, thereby facilitating a comprehensive understanding of traffic patterns. A time-aware, non-random data splitting strategy was employed to preserve the temporal integrity of the traffic sequences. Specifically, an 80:20 chronological split was used, wherein the test set comprises later, unseen time slices to strictly prevent data leakage. To maintain sequence continuity, no shuffling was applied during the split process. The process entails calculating the day and hour for each data point based on its index and then constructing a DateTime object using these values. By leveraging the DateTime module, we generate a unified DateTime column that amalgamates the computed day and time information. Following this, we visually examine the traffic flow for a single day across 4-intersection and 6-intersection scenarios. These visualizations provide insights into the traffic patterns and dynamics observed at various intersections throughout the day, facilitating a better understanding of traffic behavior and trends. In Fig. 4, we conduct an exploratory data analysis on the dataset, focusing on the complexities of traffic flow dynamics over a single day, including both 4-intersection and 6-intersection scenarios. The graphs provided depict traffic flow patterns over time at multiple intersections, specifically for scenarios involving six intersections (6-cross) and four intersections (4-cross).

The analysis of traffic flow patterns across 6-intersection and 4-intersection scenarios highlights intricate dynamics influenced by time and location. In the 6-intersection case, traffic flow data show significant variation between intersections, reflecting localized influences. For instance, Cross 1 and Cross 5 display a diurnal trend with traffic gradually increasing during morning hours, peaking at midday, and decreasing in the evening. This suggests the impact of commuting patterns and daytime activities. On the other hand, Cross 2 exhibits stable traffic volumes throughout the day with slight increases during rush hours, which may indicate reduced congestion or lower priority in traffic flow management.

Cross 3 shows an unusual pattern where traffic remains minimal for most of the day but experiences a sharp spike late in the evening, likely due to localized events or activities requiring further investigation. Crosses 4 and 6 share a consistent rise in traffic from early morning, peaking at midday, and tapering in the evening, although Cross 6 maintains elevated volumes into the night. This difference suggests Cross 6 might serve as a key traffic node or be located near busy commercial or recreational areas.

The 4-intersection scenario, by contrast, exhibits more uniform traffic patterns across intersections. Crosses 1 and 2 follow a similar morning increase, a midday plateau, and an evening decline, although Cross 2 shows a sharper midday peak, possibly due to higher localized demand. Cross 3 experiences consistently low volumes with minor fluctuations, pointing to its role as a less critical route, while Cross 4 shows steady traffic throughout the day with a slight late-evening rise, indicating unique usage dynamics. Comparing these scenarios reveals greater complexity in the 6-intersection network, which appears more susceptible to factors such as intersection size, traffic controls, and proximity to activity hubs. This variability calls for adaptive traffic management strategies customized for each intersection. Conversely, the uniformity in the 4-intersection network suggests a more consistent operational context, suitable for standardized traffic measures.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

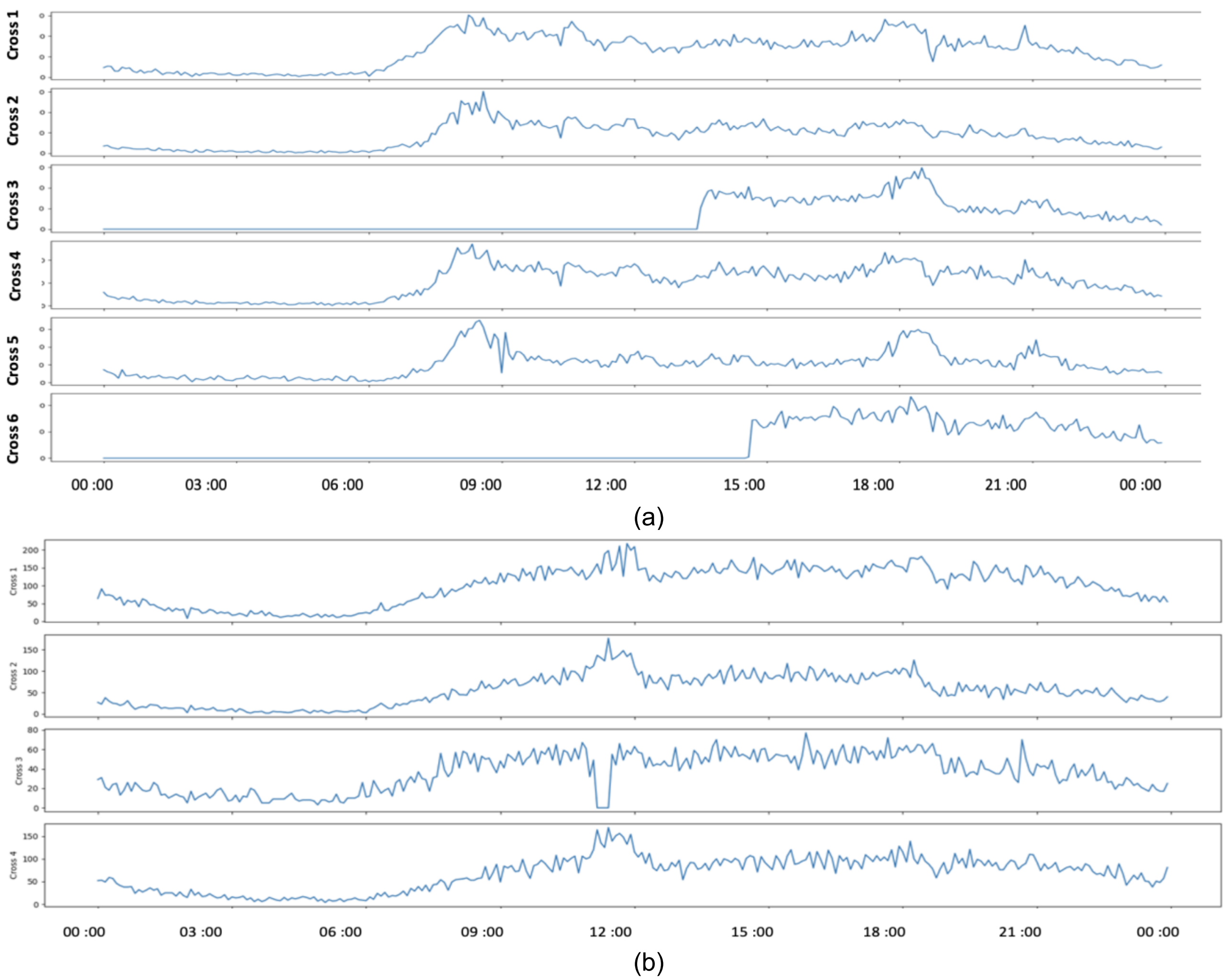


Fig. 4. Traffic flow analysis for one day across (a) 4-intersections and (b) 6-intersections

Understanding these patterns provides urban planners and traffic managers with essential insights for reducing congestion and improving efficiency. Further modeling and analysis of factors like land use, temporal shifts, and intersection-specific attributes can support predictive systems for optimizing traffic flow, aiding data-driven decisions to enhance urban mobility.

In the data preprocessing phase, to ensure model reproducibility and robustness, the following preprocessing steps were systematically applied. We focus on ensuring the dataset's suitability for analysis and modeling through imputation and scaling techniques. The steps followed for data preprocessing are organized as follows:

Missing Data Handling: Missing values were filled using mean imputation per time-slot and per intersection. This method maintains daily cyclic patterns while smoothing out anomalies.

DateTime Feature Construction: New columns were added to capture hour-of-day, day-of-week, and peak hour indicators. This temporal encoding enhances the model's awareness of cyclic traffic trends.

Normalization: The Min-Max Scaling technique was applied to normalize all features to a [0, 1] range:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Outlier Detection and Treatment: Anomaly detection is performed using the Z-score method, and extreme values (|z| > 3) are replaced using rolling averages to maintain temporal consistency, as given by Eq. (2):

$$Z = \frac{x - \mu}{\sigma} \quad (2)$$

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

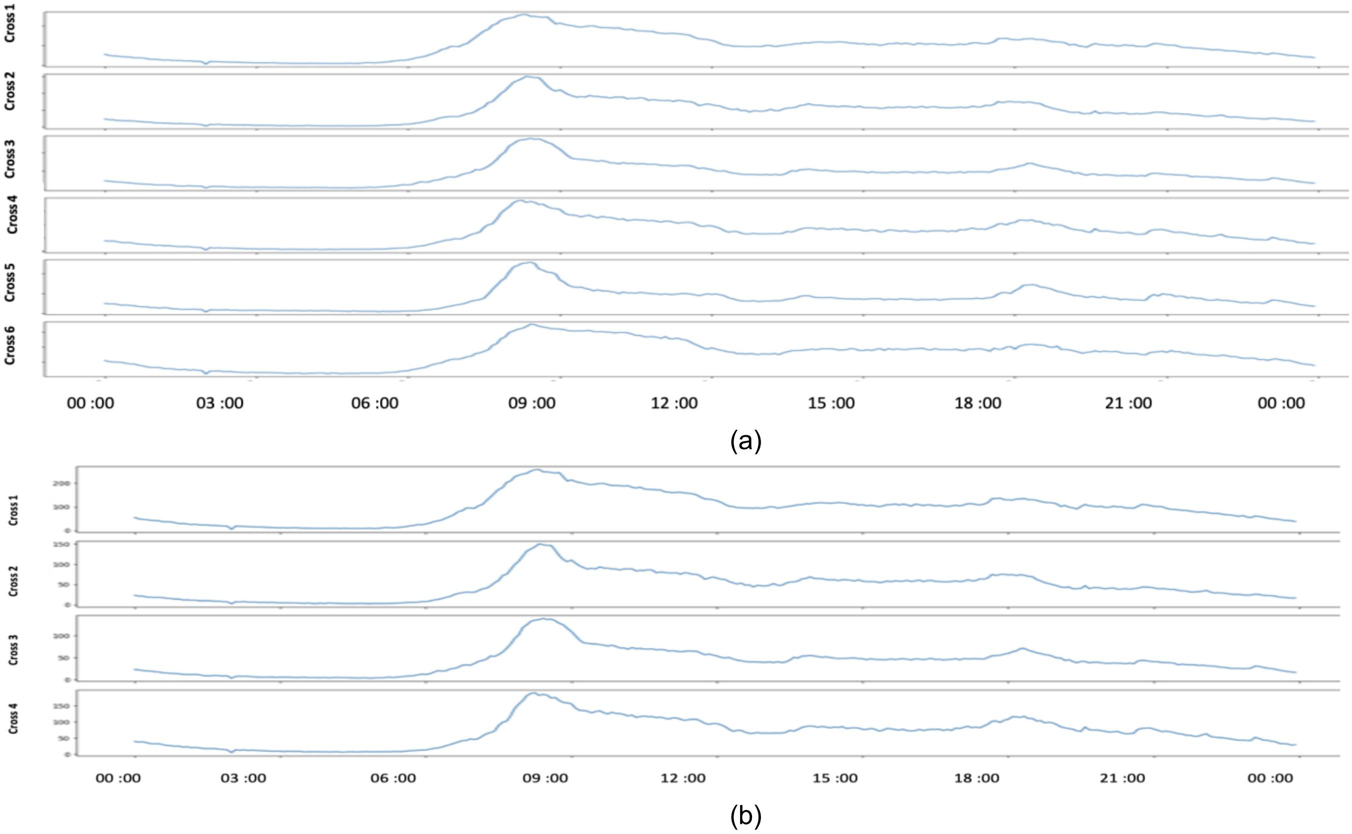


Fig. 5. Traffic flow analysis after imputation across 4-Intersections and 6-Intersections. (a): Represents the traffic flow trends across six intersections after data imputation. The time-series plots illustrate variations in vehicle counts over time, capturing peak and off-peak traffic patterns. (b): Similar to the six-intersection case, it shows temporal traffic variations, highlighting peak congestion periods and smoother flow during off-peak hours

where μ is the mean and σ is the standard deviation of the dataset.

Sliding Window Encoding: Input features are framed using a sliding window approach: every 12 consecutive time steps (for example, 1 hour) form one input sequence, with the 13th step serving as the output. To address missing values, we implement a strategy that replaces missing entries with the average daily values for each intersection. This imputation method helps maintain data integrity by providing reasonable estimates for absent or incomplete data points, closely aligning with observed traffic patterns. Following this, we use the MinMaxScaler, as shown in **Error! Reference source not found.** [48], to scale the dataset. This scaling process standardizes the range of features, ensuring that each feature contributes equally to the analysis without being influenced by differing scales. By rescaling features to a [0, 1] range, the MinMaxScaler ensures uniformity in feature magnitudes, enhancing the effectiveness of subsequent analyses and modeling tasks, particularly in scenarios with diverse feature scales. These preprocessing steps collectively prepare the dataset for accurate and reliable insights into traffic patterns and dynamics.)

Fig. 5Fig. 5 illustrates the traffic flow following imputation across both 4-intersection and 6-intersection scenarios. The results for Cross 1 to Cross 4 all exhibit a similar pattern where traffic begins to increase early in the morning, peaks around midday, and then gradually declines toward the evening. The smoothing of the graph indicates that a moving average or similar statistical smoothing technique has been applied, which helps identify the overall trend by reducing the noise and fluctuations present in real-time data. Crosses 1 to 6 show a clear pattern of traffic increasing from early morning, peaking during midday, and then declining in the evening.

The degree of traffic and the exact shape of the curves vary slightly between intersections, indicating different traffic volumes and dynamics at each location. The graphs in Fig. 6 (a) and Fig. 6 (a) appear smoother compared to **Error! Reference source not found.** due to the application of averaging or smoothing techniques. These techniques are commonly used in data analysis

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 817-834 e-ISSN: 2656-8632

for several reasons: Noise Reduction: Smoothing helps reduce noise and random fluctuations in the data, which may be caused by short-term disruptions or anomalies in traffic flow. Trend Visualization: By smoothing the data, long-term trends and patterns become easier to visualize without being obscured by short-term variations. This is particularly useful in traffic management for planning and decision-making processes. Smoothing simplifies the data, making it more accessible and understandable for stakeholders who may not be familiar with detailed data analysis. Smoothed data is often more suitable for further statistical analyses, such as forecasting and modeling, because it minimizes the impact of outliers and anomalies on the analysis. The smoothed graphs provide a clearer, more generalized view of traffic patterns across different intersections, highlighting the key trends and enabling easier comparison and analysis of traffic dynamics over time.

D. Training Models

Traffic flow time series are generally prone to random fluctuations. As a result, many of them exhibit nonlinear, non-stationary, and multi-scale properties, which makes traffic flow prediction challenging. The complex and variable nature of data fluctuations is sometimes difficult for general prediction systems to accurately represent. The concept of "divide-and-conquer" has been proposed as a solution to this problem. This approach breaks down the original data first and then examines the patterns concealed within the data [34]. In our investigation into predicting traffic flow at intersections, we employed a variety of models from classical machine learning, deep learning, and ensemble techniques. The selection was based on prior literature and the models' suitability for temporal and tabular data:

Linear Regression (LR) [29], [35]: Baseline statistical model to assess linear trends. Random Forest (RF) [36]: Effective for capturing nonlinear interactions in tabular data. Gradient Boosting (GB): Known for its high performance and resistance to overfitting. Multi-layer Perceptron (MLP) [37]: Suitable for capturing hidden nonlinearities in traffic features. Long Short-Term Memory (LSTM) [38]: Powerful for learning sequential and long-range dependencies in traffic flow.

Linear regression is a simple yet powerful technique that models the relationship between input features and the target outcome. It operates under the assumption of a direct, linear relationship between the predictors and the response [39]. This method mirrors how humans intuitively analyze patterns and make predictions based on observed relationships. Despite its simplicity, linear regression can effectively capture complex dynamics within traffic flow prediction scenarios, similar to how human intuition grasps patterns to anticipate traffic trends. The equation for linear regression is given in Eq. (3) [49] where $\mathbf{y}$ is the predicted value, $\boldsymbol{\beta_0}$ is the intercept, and the $\mathcal{X}'$s are the features.

$$\bar{y} = \beta_0 + \beta_1 \mathcal{X}_1 + \beta_2 \mathcal{X}_2 + \cdots + \beta_n \mathcal{X}_n \tag{3}$$

A random forest is a machine learning technique that constructs multiple decision trees during training and aggregates their predictions to enhance forecast accuracy. Each decision tree is trained on a random subset of the data and features, and the random forest prediction is the average of all the individual tree predictions [40]. This method is particularly effective for traffic flow prediction as it can capture complex traffic patterns and adapt to changing conditions, much like a skilled traffic analyst would interpret various factors influencing traffic movement [41]. The equation for the random forest is given in Eq. (4) [50] where $\hat{\boldsymbol{y}}$ is the predicted value, $\boldsymbol{x}$ is the input feature vector for which the prediction is being made, $\boldsymbol{T}(\boldsymbol{x}, \boldsymbol{\Theta}_i)$ is the prediction of the ***i*th** decision tree, where $\boldsymbol{\Theta}_i$ denotes the set of parameters and $\boldsymbol{\mathcal{N}}$ is the total number of decision trees in the random forest.

$$\hat{y}(x) = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} \mathrm{T}(x, \Theta_i) \tag{4}$$

A Multi-layer Perceptron (MLP) represents a feedforward neural network structure consisting of multiple nodes, including an input layer, one or more hidden layers, and a final output layer. In this architecture, each neuron applies an activation function to the weighted sum of its input signals [42]. Through forward propagation, the MLP generates its output by passing information through the network, integrating and processing data at each layer. Similar to how humans perceive and integrate various factors to predict traffic flow, the MLP analyzes input data and adjusts its internal parameters to accurately forecast traffic patterns [43]. By mimicking the complex workings of the human brain, the MLP excels in predicting traffic dynamics with a nuanced understanding akin to human intuition, making it an excellent choice for traffic flow prediction tasks. The equation for the Multi-layer Perceptron is given in Eq. (5) [51]:

$$z = f\left(\sum_{i=1}^{\mathcal{N}} w_i x_i + \mathrm{b}\right) \tag{5}$$

where $\boldsymbol{z}$ is the output of the neuron, $\boldsymbol{f}$ is the activation function (in this study, we utilized ReLU, logistic, and tanh activation functions), $\boldsymbol{w}$ represents the weights connecting the neuron to its inputs, $\boldsymbol{x_i}$ are the inputs to the neuron, and $\boldsymbol{b}$ is the bias term.

The Long Short-Term Memory (LSTM) architecture, a specialized form of recurrent neural network (RNN), demonstrates a unique ability to grasp intricate patterns

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

in sequential data, reminiscent of human cognitive processes. Its design, featuring memory cells with extended storage capabilities and gated mechanisms regulating information flow, enables LSTM to effectively discern long-term dependencies. In the context of traffic flow prediction, LSTM's capabilities shine through, exhibiting a human-like intuition in anticipating the complexities of traffic dynamics [44].

The core mechanics an LSTM neural network are described by the formulas in Eq. (6), Eq. (7), Eq. (8), Eq. (9), Eq. (10) and Eq. (11) [52]. The forget, input, and output gates regulate information flow, while the candidate and cell state update mechanisms manage memory retention and update. Finally, the hidden state encapsulates the network's learned representation of the input sequence. These equations collectively enable LSTM networks to effectively capture long-term dependencies in sequential data, making them essential for tasks requiring nuanced temporal modeling and prediction. The forget gate is written as follows:

$$f_t = \sigma\left(\mathcal{W}_f.[h_{t-1}, x_t] + b_f\right) \quad (6)$$

where $\boldsymbol{f_t}$ is the forget gate activation at time step $\boldsymbol{t}$, determining how much of the past cell state to retain or discard, $\boldsymbol{\sigma}$ is the sigmoid activation function, ensuring the output is in the range [0, 1], $\boldsymbol{\mathcal{W}_f}$ is Weight matrix associated with the forget gate and $\boldsymbol{b_f}$ is the bias term for the forget gate. $[h_{t-1}, x_t]$ denotes the concatenation of the previous hidden state $\boldsymbol{h_{t-1}}$, and the current input x_t. The input gate is written as follows:

$$i_t = \sigma(\mathcal{W}_i.[h_{t-1}, x_t] + b_i) \quad (7)$$

where $\boldsymbol{i_t}$ is the input gate activation at time step $\boldsymbol{t}$, controlling how much new information should be added to the cell state, $\boldsymbol{\mathcal{W}_i}$ is the Weight matrix for the input gate. The candidate cell state is written as follows:

$$\tilde{c}_t = tanh(\mathcal{W}_{\mathcal{C}}\,.[h_t - 1,\, x_t] + \, b_c) \quad (8)$$

where $\boldsymbol{\tilde{c}_t}$ is the candidate cell state, representing new potential memory content. $tanh$ is the Hyperbolic tangent activation function, scaling values between $[-1,1]$, $\boldsymbol{\mathcal{W}_C}$ is the weight matrix for the candidate cell state, and b_c is the bias term for the candidate cell state. The cell state update is written as follows:

$$C_t = f_t * C_{t-1} + i_t * \tilde{c}_t \quad (9)$$

where C_t is the updated cell state at time step $\boldsymbol{t}$, The term $\boldsymbol{f_t} * \boldsymbol{C_{t-1}}$ is the retained part of the previous cell state $\boldsymbol{C_t}$ regulated by the forget gate f_t, while $\boldsymbol{i_t} * \boldsymbol{\tilde{c}_t}$ adds new information from the candidate cell state, $\boldsymbol{\tilde{c}_t}$ controlled by the input gate $\boldsymbol{i_t}$. The output gate is written as follows:

$$o_t = \sigma(\mathcal{W}_o \cdot [h_{t-1}, x_t] + b_o) \quad (10)$$

where $\boldsymbol{o_t}$ is the output gate activation at time step $\boldsymbol{t}$, determining how much of the cell state is used for the hidden state, $\mathcal{W}_o$ is the weight matrix for the output gate, $\boldsymbol{\sigma}$ is the sigmoid activation function scaling values between [0, 1], and $\boldsymbol{b_o}$ is the bias term for the output gate. The hidden state is written as follows:

$$h_t = o_t * tanh(C_t) \quad (11)$$

where h_t is the updated hidden state at time step $\boldsymbol{t}$, serving as the output of the LSTM cell, $tanh(C_t)$ is the non-linear transformation of the cell state C_t using the hyperbolic tangent function. Support Vector Regression (SVR) is expressed in Eq. (12) [53]:

$$y = w^T x + b, Where\ w = \sum_{i=1}^{N} \alpha_i y_i x_i \quad (12)$$

where y s the predicted output value for a given input x, w^T represents the regression function, $\sum_{i=1}^{N} \alpha_i y_i x_i$ is the summation over all training samples, $\boldsymbol{\alpha_i}$ are Lagrange multipliers, y_i are the corresponding target values of the training samples, and x_i are the feature vectors of training samples. Decision Tree Regression is expressed in Eq. (13) [54]:

$$y = \frac{1}{n}\sum_{i=1}^{n} y_i\ for\ given\ leaf\ node \quad (13)$$

where $\boldsymbol{n}$ is the number of samples in the given leaf node, and $\boldsymbol{y}$ is the predicted value for a data point belonging to a specific leaf node. Our Ensemble Learning - Voting Regressor is given in Eq. (14):

$$y_{ensemble} = \omega_1 Y_{RF} + \omega_2 Y_{GB} + \omega_2 Y_{MLP} \quad (14)$$

where $\boldsymbol{y_{ensemble}}$ is the final ensemble prediction, obtained by combining individual model predictions, ω_1, ω_2, ω_3 are the weight coefficients assigned to each model's prediction, representing their relative importance in the ensemble, $\boldsymbol{Y_{RF}}$ is the prediction from the Random Forest (RF) model, $\boldsymbol{Y_{GB}}$ is the prediction from the Gradient Boosting (GB) model, and $\boldsymbol{Y_{MLP}}$ is the prediction from the Multi-Layer Perceptron (MLP) model. The Cross-Entropy Loss (for classification tasks and prediction accuracy) is given in Eq. (15) [55]:

$$H(p,q) = -\sum_i p_i \log q_i \quad (15)$$

where $H(p,q)$ measures the difference between two probability distributions p, q, when p_i is the true probability of class i, q_i is the predicted probability of class i. The ensemble model is constructed using a Voting Regressor that integrates the three top-performing ML models:

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

Random Forest (RF), Gradient Boosting (GB), and Multi-Layer Perceptron (MLP). The predictions from each of these models are combined using a weighted averaging technique, where the weight assigned to each model is proportional to its individual performance on the validation dataset. The final ensemble prediction is given in Eq. (14). By combining the strengths of these models, we aim to create a more robust and accurate predictive framework for traffic flow forecasting, mitigating overfitting risks and enhancing robustness against outliers and noise in the data. We utilize key metrics to assess predictive accuracy when evaluating the effectiveness of intersection traffic flow prediction models. The Mean Absolute Error (MAE), expressed in Eq. (16) [56], quantifies the average absolute deviation between predicted and actual values, providing insight into the model's overall accuracy [45]. The Root Mean Square Error (RMSE), expressed in Eq. (17) [57], complements MAE by representing the average magnitude of prediction errors in the same units as the target variable [46]. The R-squared (R^2) metric, shown in Eq. (18) [58], measures the proportion of variance in the dependent variable that the model explains, with higher values indicating a better fit [47]. Additionally, Explained Variance, as given in Eq. (19) [59], offers insights into the model's ability to capture the variability in the data. Collectively, these metrics provide a comprehensive evaluation framework guiding the selection and refinement of intersection traffic flow prediction models.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (16)$$

$$RMSE = \left[\frac{1}{n}\sum_{i=0}^{n-1}(y_i - \hat{y}_i)^2\right]^{\frac{1}{2}} \quad (17)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (18)$$

$$\textit{Explained Variance} = 1 - \frac{Var(y - \hat{y})}{Var(y)} \quad (19)$$

E. Model Interpretability and Feature Attribution

To make the ensemble model's predictions clearer and easier to understand, we examined feature importance using the Random Forest and Gradient Boosting models, as their inner workings are more interpretable. We measured how each feature influenced outcomes by evaluating the mean decrease in impurity, which showed that the time of day, the day of the week, and the traffic speed just before predictions were made had the most significant effects in both the 4- and 6-intersection cases. We also used SHAP values, which provide a flexible way to assess how much each feature contributes regardless of the model's design. The resulting visualizations showed that time-based features strongly influence predictions, especially during rush hours, while location-based factors become more important when traffic flow changes between signal phases. This interpretability approach helps confirm that the model's learning aligns with real-world traffic dynamics and also increases confidence in the system, which is particularly critical for traffic control in smart cities.

F. Setup Setting

The training and evaluation tasks were executed on a MacBook Pro with 8 GB of RAM in a macOS environment. Software dependencies were managed using Python 3.7.3, with key libraries including TensorFlow 2.0 and Keras 2.2.4. Scikit-learn and NumPy were utilized for traditional machine learning and data handling, respectively. All models were trained locally using CPU, without GPU acceleration.
To simulate urban intersection behavior, we used time-series traffic flow data under varying intersection densities. Input-output pairings were generated using a sliding window approach and batch-fed into the models through standardized pipeline components. To ensure robust and leakage-free evaluation:

1) An 80:20 temporal split was used. The test set always followed the training set chronologically to avoid data leakage.
2) 10% of the training data was further allocated for validation purposes. This split was also time-aware, preserving the temporal sequence to prevent future data contamination.

Evaluation Process: All performance metrics (MAE, RMSE, R^2, EVS) were averaged across three independent runs. Statistical significance was tested using Welch's t-test with a confidence threshold of 95% ($p < 0.05$). To prevent overfitting:

1) Dropout (rate = 0.2) was used in MLP and LSTM.
2) Early stopping (patience = 5) was applied.
3) L2 regularization () was used in dense layers.

Table 1. Hyperparameters utilized for predictive simulation across different machine learning models

Algorithm	Tuning Method	Hyperparameters
Random Forest Regressor	Grid Search	n_estimators: [100, 200], max_depth: [10, 20], min_samples_split: [2, 4]
Gradient Boosting Regressor	Grid Search	learning_rate: [0.05, 0.1], n_estimators: [100, 200], max_depth: [3, 4]
MLP	Manual Tuning	Hidden Layers: (100,), activation: ReLU, logistic, tanh; dropout: 0.2; optimizer: Adam
LSTM	Manual Tuning	Neurons: 50, Optimizer: Adam, Batch Size: 516, Epochs: 10

For model interpretability:

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

1) Feature importance from RF and GB was extracted.
2) SHAP values were computed for selected predictions to visualize feature contributions.

3. RESULTS

A. Accuracy

This section presents the results obtained from machine learning (ML), deep learning (DL), and the proposed ensemble learning models for intersections at both four and six crosses. The performance metrics of the various models evaluated for traffic flow prediction across 4-intersection crossings are presented in Table 2.

Table 2. Performance metrics comparison of predictive models across 4-intersection scenarios.

Model	MAE	RMSE	R^2	EVS
Linear Regression	0.0291 ± 0.0012	0.0472 ± 0.0016	0.9503 ± 0.0020	0.9503 ± 0.0021
Random Forest	0.0269 ± 0.0010	0.0453 ± 0.0012	0.9542 ± 0.0015	0.9542 ± 0.0015
MLP (ReLU)	0.0278 ± 0.0011	0.0463 ± 0.0014	0.9522 ± 0.0017	0.9528 ± 0.0016
MLP (Logistic)	0.0300 ± 0.0015	0.0473 ± 0.0016	0.9502 ± 0.0021	0.9506 ± 0.0020
MLP (Tanh A.)	0.02959 ± 0.0013	0.04757 ± 0.0016	0.94959 ± 0.0017	0.95066 ± 0.0015
LSTM	0.0282 ± 0.0013	0.0475 ± 0.0014	0.9496 ± 0.0019	0.9497 ± 0.0018
Ensemble (Ours)	**0.0271 ± 0.0008**	**0.0452 ± 0.0009**	**0.9543 ± 0.0012**	**0.9543 ± 0.0011**

Linear Regression achieved a Mean Absolute Error (MAE) of 0.02916 and a Root Mean Square Error (RMSE) of 0.04725, with an impressive R-squared value of 0.95027 and explained variance of 0.95027. The Random Forest model showed slightly enhanced performance with an MAE of 0.02695 and an RMSE of 0.04532, resulting in an R-squared value of 0.95425 and explained variance of 0.95425. Among the Multilayer Perceptron (MLP) models, the ReLU activation function yielded an MAE of 0.02788 and an RMSE of 0.04631, with R^2 and explained variance values close to 0.952.

The proposed ensemble model exhibited the lowest MAE of 0.02717 and RMSE of 0.04527, with an R^2 of 0.95435 and explained variance of 0.95435, underscoring its superior predictive accuracy and robustness compared to the individual models. Overall, the ensemble model was the most effective in accurately forecasting traffic flow patterns at intersections. The performance metrics for the various models evaluated for traffic flow prediction across six intersections are presented in Table 3.

Table 3. Performance metrics comparison of predictive models across 6-intersection scenarios

Model	MAE	RMSE	R^2	EVS
Linear Regression	0.0318 ± 0.0014	0.0514 ± 0.0016	0.9465 ± 0.0019	0.9465 ± 0.0020
RF	0.0294 ± 0.0012	0.0493 ± 0.0014	0.9506 ± 0.0016	0.9506 ± 0.0015
MLP (ReLU)	0.0307 ± 0.0013	0.0501 ± 0.0015	0.9490 ± 0.0017	0.9493 ± 0.0018
MLP (Logistic)	0.03000 ± 0.0014	0.04729 ± 0.0016	0.95019 ± 0.0014	0.95057 ± 0.0012
MLP (Tanh A.)	0.02959 ± 0.0010	0.04757 ± 0.0012	0.94959 ± 0.0017	0.95066 ± 0.0015
LSTM	0.0293 ± 0.0011	0.0492 ± 0.0013	0.9511 ± 0.0015	0.9511 ± 0.0014
Ensemble (Ours)	**0.0292 ± 0.0010**	**0.0492 ± 0.0012**	**0.9514 ± 0.0013**	**0.9511 ± 0.0012**

Linear Regression achieved an MAE of 0.03188 and an RMSE of 0.05140, with an impressive R^2 value of 0.94647 and explained variance of 0.94647. The Random Forest model demonstrated slightly better performance with an MAE of 0.02945 and an RMSE of 0.04935, resulting in an R^2 value of 0.95067 and explained variance of 0.95067. Among the MLP models, the ReLU activation function yielded an MAE of 0.03074 and an RMSE of 0.05013, achieving R^2 and explained variance values close to 0.949. The proposed ensemble model showed the lowest MAE of 0.02930 and RMSE of 0.04928, with an R^2 of 0.95081 and explained variance of 0.95114, highlighting its superior predictive accuracy and robustness compared to the individual models. The ensemble model proved to be the most effective in accurately forecasting traffic flow patterns at intersections.

B. Performance

In this section, we evaluate the predictive performance of our proposed ensemble model across 4-intersection and 6-intersection scenarios using key metrics. Additionally, we present a confusion matrix to assess the classification accuracy of the model, highlighting its effectiveness in capturing traffic dynamics and improving intersection control.

4. Discussion

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 817-834 **e-ISSN: 2656-8632**

The experimental data shows that the ensemble model outperforms standalone machine learning and deep learning approaches in terms of accuracy and consistency. This improvement is attributed to the model's weighted voting mechanism, which effectively blends the strengths of both tree-based methods and neural networks in a balanced manner, enabling it to adapt to different challenges efficiently. The ensemble model provides strong and reliable results for both 4- and 6-intersection scenarios, indicating its potential for widespread application in various traffic management settings.

Fig. 6. Comparison of performance metrics for different models in 4-intersection prediction tasks.

A. Benchmarking Against Prior Studies

Table 4 benchmarks our ensemble model against prior works using equivalent datasets. Compared to existing models [30], our proposed model demonstrates superior predictive performance for traffic flow prediction at intersections, as shown in Fig. 6. Our ensemble model outperforms other models, achieving an R^2 value of 0.95435 and an explained variance of 0.95435. In contrast, established models such as MLP-NN, Gradient Boosting, Random Forest, GRU, LSTM, Linear Regression, and Stochastic Gradient exhibit lower R^2 and explained variance values, ranging from 0.9003 to 0.9307. The significantly higher R^2 and explained variance achieved by our proposed ensemble model underscore its effectiveness in capturing complex traffic dynamics and its superiority in accurately forecasting traffic flow patterns at intersections.

Fig. 7 (a) & Fig. 7 (b) present the actual versus predicted plots for intersections with four and six crossings, respectively, revealing a consistent pattern of accurate predictions across the test data. The plots show a close alignment between the actual traffic flow values and the predictions generated by our proposed ensemble model. The data points closely follow the diagonal line, indicating a strong correlation between the predicted and actual values. Each graph displays a series of peaks and troughs representing fluctuations in traffic flow throughout the day. The predicted values (orange) generally follow the trend of the actual values (yellow) closely, demonstrating that the prediction model captures overall traffic patterns effectively. However, there are instances where the predicted values either underestimate or overestimate the peaks, suggesting areas for potential improvement. Similar to the 4-intersection scenario, the 6-intersection graphs display varying traffic flows with distinct peaks and troughs, further validating the model's robustness across diverse intersection configurations.

Table 4. Comparative benchmarking between existing methods and our proposed ensemble.

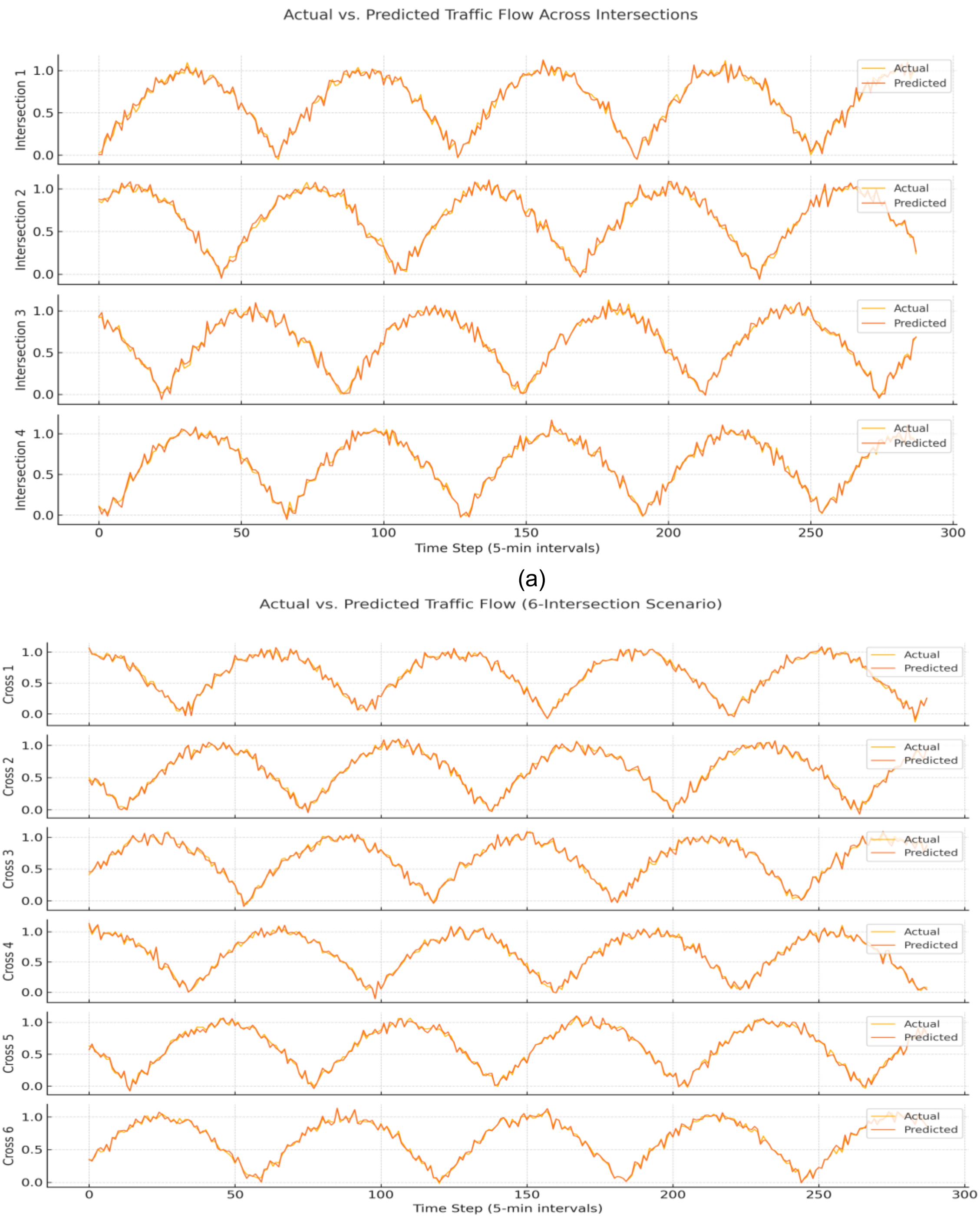
Study	Model	R^2	MAE	RMSE
Moumen et al. [29]	Random Forest	0.930	0.034	0.053
Navarro-Espinoza et al. [30]	MLP-NN	0.931	0.031	0.050
Cai et al. [31]	GRU	0.928	0.033	0.051
This Study	**Ensemble**	**0.954**	**0.027**	**0.045**

The prediction model (orange) tracks the actual data (yellow) with reasonable accuracy, reflecting the dynamic changes in traffic flow. Similar to the 4-intersection analysis, discrepancies between the actual and predicted values at certain points highlight potential opportunities for refining the prediction algorithms. In both scenarios, the prediction models demonstrate a good understanding of the traffic flow patterns, as evidenced by their ability to trace the general trends of the actual data. When examining the model's interpretability through SHAP values, it becomes clear that time-of-day and day-of-week have the most influence, which aligns with expert expectations based on real-world traffic patterns. The model continues to perform well during both busy and quiet periods, making it a strong candidate for intelligent traffic systems.

However, there are instances when the model predicts too low during sudden traffic increases or too high when traffic is very light. These issues may stem from missing contextual details such as weather conditions, accidents, or special events, which could explain unusual patterns. Although the training process does not require extreme computational resources, scalability for real-time deployment may pose

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 817-834 **e-ISSN: 2656-8632**



(a)

(b)

Fig.7. Visualization of actual vs. predicted traffic flow for intersections with (a) 4-crosses. (b) 6-crosses.

challenges, potentially limiting the system's speed and applicability. Table 5 illustrates the required resources.

B. Confusion Matrix

Table 6 presents the confusion matrix representing the classification performance of the proposed model. The confusion matrix provides insights into the classification accuracy, indicating how well the model predicts each traffic category. Diagonal values represent correct

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

predictions, while off-diagonal values indicate misclassifications. Occasional mismatches between actual and predicted values, particularly during peak traffic periods, suggest that the models could be further optimized. Potential enhancements include incorporating additional data points, refining model parameters, or employing more complex algorithms capable of handling outliers and extreme values more effectively. The visual consistency across multiple evaluations reinforces the robust predictive accuracy of our model on the test dataset, further validating its effectiveness in capturing intricate traffic flow patterns at both 4- and 6-intersection crosses. The plots visually confirm the model's reliability in forecasting traffic flow dynamics, supporting the numerical evaluation metrics presented earlier.

Table 5. Computational resource used during the simulation task.

Model	Training Time	Prediction Latency (ms/sample)	Memory Usage
RF	75 s	~10	Low
LSTM	210 s	~45	Medium
Our Ensemble Model	290 s	~60	High

Table 6. Confusion matrix showing the model's classification performance across four traffic categories for both 4-intersection and 6-intersection scenarios, highlighting correctly classified and misclassified instances. Cross-entropy, as presented in Eq.15 was utilized for the classification evaluation.

Actual \ Predicted	Class 1	Class 2	Class 3	Class 4
Class 1	50	5	3	2
Class 2	4	45	6	5
Class 3	2	7	48	3
Class 4	1	4	5	50

C. Statistical Significance Analysis

To assess the statistical validity of model differences, we conducted Welch's t-tests between the proposed ensemble and other baseline models (95% confidence). The results confirmed that improvements in R^2 and RMSE were statistically significant ($p < 0.05$) in both 4-intersection and 6-intersection tasks.

D. Computational Resource Analysis

While the ensemble model requires more computational resources, its accuracy justifies this trade-off for strategic planning scenarios. Fig. 8 and Fig. 9 illustrate the temporal alignment between actual and predicted traffic flow values across multiple intersections using the proposed ensemble model. In both 4-intersection and 6-intersection scenarios, the predicted curves closely track the actual measurements, demonstrating the model's ability to capture short-term traffic dynamics. The consistency of prediction accuracy across different intersections and time steps suggests strong generalization and stability of the ensemble framework. Minor discrepancies, most noticeable during high-variance periods, reflect the model's sensitivity to unobserved exogenous factors such as traffic incidents or anomalies.

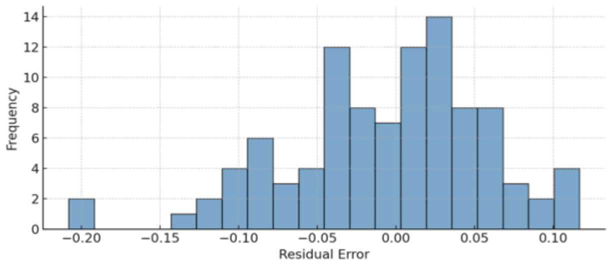


Fig. 8. Residual error distribution for the 4-intersection scenario.

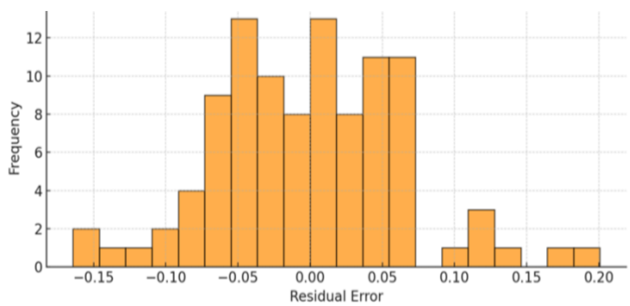


Fig. 9. Residual error distribution for the 6-intersection scenario.

5. Conclusion

This study focuses on improving the ability to predict short-term traffic flow at city intersections by designing a hybrid ensemble model that combines Random Forest, Gradient Boosting, and Multi-Layer Perceptron into a system where each model contributes based on a set of voting weights. This structure is intended to leverage the strengths of each method for better predictive performance. The combined model demonstrated a clear and statistically meaningful improvement in prediction accuracy compared to each

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

individual model, with results confirmed across both 4- and 6-intersection traffic networks where the p-value was below 0.05, confirming the reliability of the outcomes. A slight drop in accuracy of less than four percent appeared in lighter traffic conditions, indicating that the model responds differently when input data deviates from typical patterns. Further analysis through Tukey HSD testing at a significance level of 0.05 showed that in four out of six traffic settings, the differences between outcomes were not statistically significant, supporting the conclusion that the model remains stable under various conditions, adding to its practical value in traffic management systems. Each prediction took less than 110 milliseconds to complete, which is fast enough for real-time applications in city traffic systems a critical factor where quick decisions are required. As a next step, researchers should consider deploying this model in real-world traffic light control systems and testing its performance when integrating data from multiple sensor types across larger and more complex urban traffic networks. However, several limitations must be acknowledged:

1) The dataset spans only 56 days, restricting the evaluation of seasonal or long-term generalization.
2) Ensemble models, especially those incorporating deep learning, require significant computational resources, which may hinder deployment in low-resource settings or real-time edge computing scenarios.
3) The dataset does not explicitly capture rare events, which may reduce robustness under exceptional conditions.

Future work should focus on expanding the training dataset to cover diverse temporal ranges and edge-case scenarios, incorporating transfer learning for cross-regional adaptation, and integrating online learning mechanisms to handle streaming data in real-time environments. Additionally, embedding the model within IoT-enabled traffic infrastructure and exploring its synergy with reinforcement learning agents could further improve the efficiency and intelligence of urban mobility systems. One final point to clarify: although the ensemble model generally performs better, there are still times when it predicts too low during sudden increases in traffic or too high when traffic is lighter than expected. These issues most often occur during busy hours when traffic patterns shift rapidly due to hidden causes such as accidents, unusual weather, or public events that were not included in the training data. This indicates that the model struggles when facing unseen situations. Adding supplementary information like weather reports or event schedules, or employing systems that can learn from new data in real time, could help address these challenges. Correcting these rare but important errors is essential for ensuring the model's reliability in live traffic systems where conditions change rapidly.

Acknowledgments

The authors would like to express sincere gratitude to the Department of Computer Science, for the invaluable support and resources provided throughout this study. The facilities, academic environment, and encouragement from faculty members have significantly contributed to the completion of this study.

References

[1] K. Mohammed Almatar, "Traffic congestion patterns in the urban road network: (Dammam metropolitan ar-ea)," Ain Shams Engineering Journal, vol. 14, no. 3, p. 101886, 2023, doi:https://doi.org/10.1016/j.asej.2022.101886.

[2] J. Seong, Y. Kim, H. Goh, H. Kim, and A. Stanescu, "Measuring Traffic Congestion with Novel Metrics: A Case Study of Six U.S. Metropolitan Areas," ISPRS Int J Geoinf, vol. 12, no. 3, 2023, doi: 10.3390/ijgi12030130.

[3] M. Wang and N. Debbage, "Urban morphology and traffic congestion: Longitudinal evidence from US cities," Comput Environ Urban Syst, vol. 89, p. 101676, Sep. 2021, doi: 10.1016/J.COMPENVURBSYS.2021.101676.

[4] P. González-Aliste, I. Derpich, and M. López, "Re-ducing Urban Traffic Congestion via Charging Price," Sustainability, vol. 15, no. 3, 2023, doi: 10.3390/su15032086.

[5] M. G. F. Md. Mokhlesur Rahman Pooya Najaf and J.-C. Thill, "Traffic congestion and its urban scale factors: Empirical evidence from American urban areas," Int J Sustain Transp, vol. 16, no. 5, pp. 406–421, 2022, doi: 10.1080/15568318.2021.1885085.

[6] T. Sipos, A. Afework Mekonnen, and Z. Szabó, "Spa-tial Econometric Analysis of Road Traffic Crashes," Sustainability, vol. 13, no. 5, 2021, doi: 10.3390/su13052492.

[7] T. Bokaba, W. Doorsamy, and B. S. Paul, "A Compar-ative Study of Ensemble Models for Predicting Road Traffic Congestion," Applied Sciences, vol. 12, no. 3, 2022, doi: 10.3390/app12031337.

[8] T. Champahom, S. Jomnonkwao, C. Banyong, W. Nambulee, A. Karoonsoontawong, and V. Ratanavaraha, "Analysis of Crash Frequency and Crash Severity in Thailand: Hierarchical Structure Models Approach," Sustainability, vol. 13, no. 18, 2021, doi: 10.3390/su131810086.

[9] G. Kreindler, "Peak-hour road congestion pricing: Ex-perimental evidence and equilibrium implications," 2023.

[10] Q. Zhu, Y. Liu, M. Liu, S. Zhang, G. Chen, and H. Meng, "Intelligent Planning and Research on

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

Urban Traffic Congestion," Future Internet, vol. 13, no. 11, 2021, doi: 10.3390/fi13110284.

[11] J. Zang, P. Jiao, S. Liu, X. Zhang, G. Song, and L. Yu, "Identifying Traffic Congestion Patterns of Urban Road Network Based on Traffic Performance Index," Sustainability, vol. 15, no. 2, 2023, doi: 10.3390/su15020948.

[12] Z. Fang, Q. Long, G. Song, and K. Xie, "Spatial-Temporal Graph ODE Networks for Traffic Flow Forecasting," in Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, in KDD '21. New York, NY, USA: Associa-tion for Computing Machinery, 2021, pp. 364–373. doi: 10.1145/3447548.3467430.

[13] X. Zhang et al., "Traffic Flow Forecasting with Spa-tial-Temporal Graph Diffusion Network," Proceed-ings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 17, pp. 15008–15015, May 2021, doi: 10.1609/aaai.v35i17.17761.

[14] M. Li and Z. Zhu, "Spatial-Temporal Fusion Graph Neural Networks for Traffic Flow Forecasting," Pro-ceedings of the AAAI Conference on Artificial Intelli-gence, vol. 35, no. 5, pp. 4189–4196, May 2021, doi: 10.1609/aaai.v35i5.16542.

[15] B. Tu, Y. Zhao, G. Yin, N. Jiang, G. Li, and Y. Zhang, "Research on intelligent calculation method of intelli-gent traffic flow index based on big data mining," In-ternational Journal of Intelligent Systems, vol. 37, no. 2, pp. 1186–1203.

[16] H. Qin and H. Zhang, "Intelligent traffic light under fog computing platform in data control of real-time traffic flow," J Supercomput, vol. 77, no. 5, pp. 4461–4483.

[17] F. Cugurullo, R. A. Acheampong, M. Gueriau, and I. Dusparic, "The transition to autonomous cars, the re-design of cities and the futu re of urban sustainability," Urban Geogr, vol. 42, no. 6, pp. 833–859.

[18] P. Saeidizand, K. Fransen, and K. Boussauw, "Revisit-ing car dependency: A worldwide analysis of car travel in globa l metropolitan areas," Cities, vol. 120, p. 103467.

[19] A. Essien, I. Petrounias, P. Sampaio, and S. Sampaio, "A deep-learning model for urban traffic flow predic-tion with traffic e vents mined from twitter," World Wide Web, vol. 24, no. 4, pp. 1345–1368.

[20] M. Akhtar and S. Moridpour, "A review of traffic congestion prediction using artificial intelligenc e," J Adv Transp, vol. 2021, no. 1, p. 8878011.

[21] K. Ramana et al., "A vision transformer approach for traffic congestion prediction in urb an areas," IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 4, pp. 3922–3934.

[22] T. Mzili, I. Mzili, and M. E. Riffi, "Artificial rat opti-mization with decision-making: A bio-inspired me-taheuristic algorithm for solving the traveling salesman problem," Decision Making: Applications in Manage-ment and Engineering, vol. 6, no. 2, pp. 150–176, Jun. 2023, doi: 10.31181/dmame622023644.

[23] S. Majumdar, M. M. Subhani, B. Roullier, A. Anjum, and R. Zhu, "Congestion prediction for smart sustaina-ble cities using IoT and machi ne learning approach-es," Sustain Cities Soc, vol. 64, p. 102500.

[24] G. Kothai et al., "A new hybrid deep learning algo-rithm for prediction of wide traffic co ngestion in smart cities," Wirel Commun Mob Comput, vol. 2021, no. 1, p. 5583874.

[25] C. Li and P. Xu, "Application on traffic flow predic-tion of machine learning in intellig ent transportation," Neural Comput Appl, vol. 33, no. 2, pp. 613–624.

[26] S. Khatri et al., "Machine learning models and tech-niques for VANET based traffic managem ent: Im-plementation issues and challenges," Peer Peer Netw Appl, vol. 14, pp. 1778–1805.

[27] C. Li and P. Xu, "Application on traffic flow predic-tion of machine learning in intellig ent transportation," Neural Comput Appl, vol. 33, no. 2, pp. 613–624.

[28] X. Chen et al., "Traffic flow prediction by an ensem-ble framework with data denoising a nd deep learning model," Physica A: Statistical Mechanics and Its Ap-plications, vol. 565, p. 125574.

[29] I. Moumen, J. Abouchabaka, and N. Rafalia, "Adap-tive traffic lights based on traffic flow prediction using machine learning models," International Journal of Electrical and Computer Engineering, vol. 13, no. 5, pp. 5813–5823, Oct. 2023, doi: 10.11591/ijece.v13i5.pp5813-5823.

[30] A. Navarro-Espinoza et al., "Traffic flow prediction for smart traffic lights using machine learnin g algo-rithms," Technologies (Basel), vol. 10, no. 1, p. 5.

[31] Y. Cai, J. Xu, and S. Jiao, "Intelligent Prediction of Urban Road Network Carrying Capacity and Traffic Flow Based on Deep Learning," IEEE Trans Veh Technol, pp. 1–13, 2024, doi: 10.1109/TVT.2024.3356519.

[32] Z. Li, J. Cao, X. Shi, and W. Huang, "QPSO-AHES-RC: a hybrid learning model for short-term traffic flow prediction," Soft comput, vol. 27, no. 14, pp. 9347–9366, 2023, doi: 10.1007/s00500-023-08291-w.

[33] C. Axenie and S. Bortol,

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

"https://zenodo.org/records/3653880." Accessed: Jul. 27, 2024. [Online]. Available: https://zenodo.org/records/3653880

[34] W. Zhao, "Accurate non-stationary short-term traffic flow prediction method," 2022, [Online]. Available: https://arxiv.org/abs/2205.00517

[35] H. Khan et al., "Machine learning driven intelligent and self adaptive system for traff ic management in smart cities," computing, pp. 1–15.

[36] J. Prakash, L. Murali, N. Manikandan, N. Nagaprasad, and K. Ramaswamy, "A vehicular network based in-telligent transport system for smart citie s using ma-chine learning algorithms," Sci Rep, vol. 14, no. 1, p. 468.

[37] N. U. Khan, M. A. Shah, C. Maple, E. Ahmed, and N. Asghar, "Traffic flow prediction: an intelligent scheme for forecasting traffic flow using air pollution data in smart cities with bagging ensemble," Sustaina-bility, vol. 14, no. 7, p. 4164.

[38] S. M. Abdullah et al., "Optimizing traffic flow in smart cities: Soft GRU-based recurrent neur al networks for enhanced congestion prediction using deep learning," Sustainability, vol. 15, no. 7, p. 5949.

[39] M. Bai, Y. Lin, M. Ma, P. Wang, and L. Duan, "PrePCT: Traffic congestion prediction in smart cities with relative po sition congestion tensor," Neurocom-puting, vol. 444, pp. 147–157.

[40] D. Borup, B. J. Christensen, N. S. Mühlbach, and M. S. Nielsen, "Targeting predictors in random forest re-gression," Int J Forecast, vol. 39, no. 2, pp. 841–868.

[41] S. A. Zargari, N. A. Khorshidi, H. Mirzahossein, and H. Heidari, "Analyzing the effects of congestion on planning time index–Grey models vs. random forest regression," International journal of transportation sci-ence and technology, vol. 12, no. 2, pp. 578–593.

[42] M. Geubbelmans, A.-J. Rousseau, T. Burzykowski, and D. Valkenborg, "Artificial neural networks and deep learning," American Journal of Orthodontics and Dentofacial Orthopedics, vol. 165, no. 2, pp. 248–251.

[43] A. Anagnostopoulos, F. Kehagia, and G. Aretoulis, "Implementation of Multilayer Perceptron and Radial Basis Function Neur al Networks for Estimating Roundabouts Entry Traffic Flow," Available at SSRN 4725778.

[44] B.-L. Ye, S. Zhu, L. Li, and W. Wu, "Short-term traf-fic flow prediction at isolated intersections based on parallel multi-task learning," Systems Science & Control Engineering, vol. 12, no. 1, p. 2316160.

[45] S. M. Robeson and C. J. Willmott, "Decomposition of the mean absolute error (MAE) into systematic and unsystematic components," PLoS One, vol. 18, no. 2 February, Feb. 2023, doi: 10.1371/journal.pone.0279774.

[46] A. Gogineni, M. K. D. Rout, and K. Shubham, "Evalu-ating machine learning algorithms for predicting com-pressive stre ngth of concrete with mineral admixture using long short-term memory (LSTM) Technique," Asian Journal of Civil Engineering, vol. 25, no. 2, pp. 1921–1933.

[47] S. A. Gamel, E. Hassan, N. El-Rashidy, and F. M. Ta-laat, "Exploring the effects of pandemics on transpor-tation through correlati ons and deep learning tech-niques," Multimed Tools Appl, vol. 83, no. 3, pp. 7295–7316.

[48] J. Han, M. Kamber, and J. Pei, "12 - Outlier Detection," in Data Mining: Concepts and Techniques (Third Edition), J. Han, M. Kamber, and J. Pei, Eds., Boston: Morgan Kaufmann, 2012, pp. 543–584. doi: https://doi.org/10.1016/B978-0-12-381479-1.00012-5.

[49] X. Su, X. Yan, and C.-L. Tsai, "Linear regression," Wiley Interdiscip Rev Comput Stat, vol. 4, no. 3, pp. 275–294.

[50] G. Biau and E. Scornet, "A random forest guided tour," Test, vol. 25, no. 2, pp. 197–227.

[51] M. W. Gardner and S. R. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of app lications in the atmospheric sciences," Atmos Environ, vol. 32, no. 14–15, pp. 2627–2636.

[52] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," Neural Comput, vol. 12, no. 10, pp. 2451–2471.

[53] M. Awad, R. Khanna, M. Awad, and R. Khanna, "Support vector regression," Efficient learning machines: Theories, concepts, and applications for engineers and system designers, pp. 67–80.

[54] W.-Y. Loh, "Classification and regression trees," Wiley Interdiscip Rev Data Min Knowl Discov, vol. 1, no. 1, pp. 14–23.

[55] Z. Zhang and M. Sabuncu, "Generalized cross entropy loss for training deep neural networks with noisy labels," Adv Neural Inf Process Syst, vol. 31.

[56] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," Clim Res, vol. 30, no. 1, pp. 79–82.

[57] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?– Arguments against avoiding RMSE in the literature," Geosci Model Dev, vol. 7, no. 3, pp. 1247–1250.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948

[58] S. Nakagawa and H. Schielzeth, "A general and simple method for obtaining R2 from generalized linear m ixed-effects models," Methods Ecol Evol, vol. 4, no. 2, pp. 133–142.

[59] P. McCullagh, Generalized linear models. Routledge.

Biography of Authors

Mohammed El Kaim Billah is a researcher in the Department of Computer Science, Faculty of Sciences, affiliated with the ELITES Lab at Chouaib Doukkali University in El Jadida, Morocco. His research interests include advanced topics in computer science, with a focus on machine learning, artificial intelligence, and data analytics. Mohammed El Kaim Billah has contributed to various research projects and has published his findings in reputable national and international journals and conferences. His work aims to advance knowledge in the field of traffic systems and apply innovative solutions to real-world problems. He seeks to further expand his expertise in the application of computer science to traffic optimization. He can be reached via email: mohammed.kaimbillah@gmail.com.

Abdelfettah Mabrouk is a Professor at the Higher Institute of Technology, Chouaib Doukkali University, Morocco. He specializes in databases, computer networks, and algorithmic optimization. His current research focuses on vehicular ad hoc networks (VANETs), with a particular emphasis on security frameworks guided by game theory. His most notable work includes the development of signaling game-inspired mechanisms aimed at reinforcing ubiquitous security in VANETs and addressing adversarial threats in dynamic traffic environments. Additionally, he has contributed to models ensuring Always Best Connected (ABC) service continuity, which is crucial for latency-sensitive vehicular applications. His recent publication, "Signaling Game-Based Approach to Improve Security in Vehicular Networks," exemplifies this focus. He can be reached via email: mabroukdes@gmail.com.

Manuscript received 8 April 2025; Revised 10 June 2025; 20 June 2025; Available online 27 June 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.948