Journal of Engineering, Project, and Production Management
2026, 16(6), 2025-258

Short-Term Traffic Flow Prediction Model for Smart Cities Based on Gradient Boosting Decision Trees

Lili Zhang
Lecturer, Department of Traffic Management Engineering, Henan Police College, Zhengzhou, 450046, China,
E-mail: zll1883893@163.com

Project Management
Received November 5, 2025; revised December 26, 2026; January 6, 2026; accepted June 21, 2026
Available online July 2, 2026

Abstract: Short-Term Traffic Flow (STTF) prediction is a pivotal technology in intelligent transportation management. This technology is crucial to the development of smart cities. However, traditional prediction models have shortcomings in modeling spatiotemporal dependencies, resisting overfitting, and adapting to external factors. Therefore, a smart city short-term traffic flow prediction model that fuses Bidirectional Long Short-Term Memory (Bi-LSTM) networks with gradient-boosted decision trees is proposed to improve prediction accuracy and robustness. The research utilizes Bi-LSTM networks to capture the bidirectional temporal characteristics of traffic flow, introduces an extreme gradient boosting algorithm framework, and combines radio frequency identification technology to achieve high-precision data collection. The experiment findings indicate that the model's forecasts closely align with the real-world data, with a maximum Mean Absolute Percentage Error (MAPE) of only 3.07%, far superior to traditional methods. The model can accurately capture the influence of weather factors: flow on rainy days is about 20% lower than on sunny days, and on rainstorm days, it is reduced by an additional 15%. After optimization, the model's inference speed reaches 1-2 milliseconds per sample (on an NVIDIA V100 GPU, a high-performance data center GPU designed for AI and deep learning), which is more than 50% faster than the traditional model. The introduced model exhibits remarkable precision and broad applicability for forecasting traffic flow in smart city environments, effectively addressing the lack of accuracy and reliability in current prediction methods under recurrent traffic patterns. It provides an efficient solution for short-term traffic flow prediction and demonstrates potential for future extension to abnormal event response and multi-city application verification.

Keywords: Bidirectional LSTM; gradient boosting decision tree; intelligent transportation systems; smart cities; traffic flow prediction.

DOI 10.32738/JEPPM-2025-258

1. Introduction

The ongoing rise in motor vehicle numbers and rapid urbanization has intensified traffic congestion, posing a major obstacle to cities sustainable growth (Tiwari, 2024). A key objective of the smart city traffic management system is to accomplish precise Short-Term Traffic Flow (STTF) forecasting, thereby providing data support for traffic signal optimization, dynamic path planning, and congestion warning (Wei et al., 2023). The validity of traffic state prediction directly depends on the accuracy of STTF prediction. This technology can accurately predict traffic conditions for the next 15 minutes to 2 hours by analyzing real-time monitoring data and historical traffic information. This predictive ability has a decisive impact on the formulation of traffic control measures (Yu et al., 2023). However, traffic flow exhibits high nonlinearity, spatiotemporal dependence, and susceptibility to external factors, making traditional prediction models often inadequate to meet actual needs (Hasanvand et al., 2023). At present, the main methods for STTF prediction in cities include neural deep learning and machine learning models (Preethi and Mamatha, 2023). However, these methods often have limitations in the comprehensive use of information. For instance, standalone LSTM models primarily capture forward temporal dependencies while underutilizing future context, resulting in insufficient bidirectional information exchange (Li et al., 2024). Furthermore, many models process temporal patterns and complex feature interactions in isolation, lacking effective integration between deep sequential features and higher-order feature combinations. This decoupling limits models ability to leverage complementary characteristics of data. To address this, we propose a novel serial fusion architecture integrating Bidirectional Long Short-Term Memory (Bi-LSTM) and Gradient Boosting Decision Tree (GBDT). Unlike parallel ensembles, our model establishes a directed feature-enhancement pipeline in which a Bi-LSTM serves as a dedicated temporal feature extractor, generating high-level bidirectional representations that are explicitly fed to GBDT as engineered inputs.

This creates a synergistic division: Bi-LSTM compensates for GBDT's limitations in sequence processing, while GBDT leverages enriched temporal features through its powerful nonlinear modeling capabilities. The hierarchical learning process enhances prediction accuracy and robustness against overfitting and traffic fluctuations through GBDT's regularization and Bi-LSTM's contextual modeling.

2. Related Works

STTF prediction in smart cities is a typical spatiotemporal data mining problem that predicts the traffic status of urban road networks or key nodes over a specific future time period using historical traffic data and external environmental factors. Domestic and foreign researchers have explored this topic. For instance, Wang et al. (2024) developed a hybrid framework that integrates Long Short-Term Memory (LSTM) and Bayesian Neural Networks (BNNs) to quantify uncertainty in STTF predictions. During the process, LSTM was used to capture spatiotemporal interrelations within traffic sequences, and BNN was used to model probabilistic prediction intervals. The experimental outcomes revealed that the proposed framework significantly outperformed traditional approaches in regard to prediction accuracy. Xu et al. (2023) proposed a combined autoregressive fractional-integral moving-average model to address long-term memory and dynamic-feature modeling issues in STTF prediction. During the process, the advantages of linear and nonlinear modeling were combined, and a two-stage prediction strategy was adopted. The experimental results showed that the proposed method was particularly well-suited to periods with significant fluctuations in traffic flow. Li et al. (2024) proposed a Bi-LSTM model that integrates an Attention Mechanism (AM) to address the limitations of traditional models, which struggle with large fluctuations and achieve insufficient prediction accuracy in STTF prediction. By combining bidirectional sequence modeling with AM, the model's ability to focus on critical time steps was enhanced. The experimental results showed that the proposed model significantly reduced prediction error. Yang and Wang (2023) introduced a hybrid model grounded in time-separated AM to address the challenges of multi-parameter collaborative modeling and the dynamic capture of spatiotemporal features in STTF prediction. During the process, multi-source traffic parameters and spatiotemporal feature extraction techniques were integrated to enhance the model's ability to perceive key spatiotemporal nodes. The experimental outcomes revealed that the newly introduced model outperformed conventional models in prediction precision. Tan et al. (2024) introduced a multimodal model based on noise immunity and AM to address the issues of insufficient data, implicit pattern mining and noise sensitivity in STTF prediction. The process innovatively combined the latest advances in computer vision and deep learning, introducing a dynamic, noise-immune Loss Function (LF) to effectively mitigate the effects of outliers and noise during model training. The experimental outcomes revealed that the introduced model significantly improved the prediction accuracy.

Since the development of ensemble learning algorithms, their theoretical foundations and practical applications have matured, and scholars from many countries have conducted thorough investigations and applied them to supervised learning problems. For example, Ouiba et al. (2023) introduced a novel system to deal with the growing security threats in the Internet of Things environment. During the process, gradient boosting and decision trees were combined to improve the system, and the open-source Catboost framework was used to integrate machine learning and deep learning algorithms to enhance the device's security. The experimental outcomes revealed that the system performed excellently in regard to detection speed and computational efficiency. Wu et al. (2023) introduced a novel approach to address privacy concerns for machine learning models in real-world applications. During the process, GBDT was combined with a robust GBDT-like model to reduce inter-tree dependencies through improved training algorithms, thereby supporting efficient and accurate data deletion. The experimental results showed that the proposed method provided a feasible solution for protecting privacy. Wang et al. (2023) designed a new predictive model to address the low accuracy of rolling force prediction in hot-rolled strip production. By using the gradient boosting decision tree method to optimize input parameters and developing a model self-training function in the control system, continuous optimization of online models was achieved. The experimental outcomes revealed that compared to traditional models, the designed prediction accuracy was notably enhanced. Jing et al. (2023) introduced a novel driving approach to address rotor temperature escalation caused by increased power density in permanent magnet synchronous motors. The research aimed to predict temperature using integrated GBDT data to avoid the risk of permanent magnet demagnetization caused by inaccurate temperature estimation. The experimental outcomes revealed that the proposed method was tested and validated for its superiority on high-power motors. Yang et al. (2024) presented a novel approach to overcome the limitation where conventional 3Ni weathering steel fails to adequately satisfy the evolving demands of ocean engineering. During the process, big data methods for corrosion were used in combination with GBDT machine learning to explore the mechanism by which structural factors affect corrosion. The experimental outcomes revealed that the proposed method effectively examined how structural elements affect the corrosion progression of 3Ni steel.

While the aforementioned studies have significantly advanced the field of STTF prediction, distinct limitations persist in various model architectures. For instance, models based solely on Bi-LSTM with Attention excel at capturing bidirectional temporal dependencies and focusing on critical time steps (Li et al., 2024), but they may lack the powerful capability to model complex, non-linear feature interactions present in heterogeneous urban data. On the other hand, standalone ensemble models like XGBoost demonstrate strong performance in handling tabular data and feature engineering (Demir and Sahin, 2023), yet they typically fall short in natively modeling the intricate long-term temporal dependencies inherent in traffic flow sequences.

Although existing hybrid models have made certain progress in traffic flow prediction, there are still obvious deficiencies in the collaborative integration of time series and feature learning, the full fusion of high-resolution multi-source data, and the customized optimization of hybrid architectures. To this end, this study proposes a novel serial fusion architecture that combines Bi-LSTM with GBDT/XGBoost and integrates RFID data acquisition technology. Its main contributions are reflected in three aspects: First, a sequence feature enhancement pipeline was designed, with Bi-LSTM as the dedicated temporal feature extractor, and its output was explicitly used as the structured input of GBDT, achieving complementary learning of long-term dependence and high-order feature interaction; Second, a multi-factor traffic modeling framework based on RFID was constructed, integrating multi-source data such as vehicle trajectories, weather, and spatio-temporal data, to support fine-grained and scene-aware predictions. Third, a customized optimization framework was established, which utilized the regularization and parallel computing capabilities of XGBoost for targeted adjustments. This not only effectively prevented overfitting but also increased the inference speed to 1-2 milliseconds per sample. This structured method not only significantly improves the prediction accuracy but also provides an expandable and interpretable solution for the smart city traffic management system.

3. Methods

3.1. Construction of STTF Prediction Model for Smart Cities

The traffic volume of motor vehicles in specific areas of urban roads (such as road sections or intersections) over a given time period is usually defined as Urban Traffic Flow (UTF), a core indicator that reflects the operational status of road traffic. It mainly includes four key parameters: flow, speed, density, and occupancy rate (Demir and Sahin, 2023). UTF is affected by numerous elements, including time, space, weather, and traffic control measures. Its changing pattern provides an important basis for traffic signal optimization, congestion warning, intelligent navigation, and urban planning, and is the basic data support for smart transportation management and urban road network optimization (Zhang et al., 2023). UTF is formed by people using transportation for different purposes and at different times. Through long-term monitoring of traffic flow at a specific road section in a certain area of C city, a continuous traffic flow change map for the road section over one week was obtained, as shown in Fig. 1.

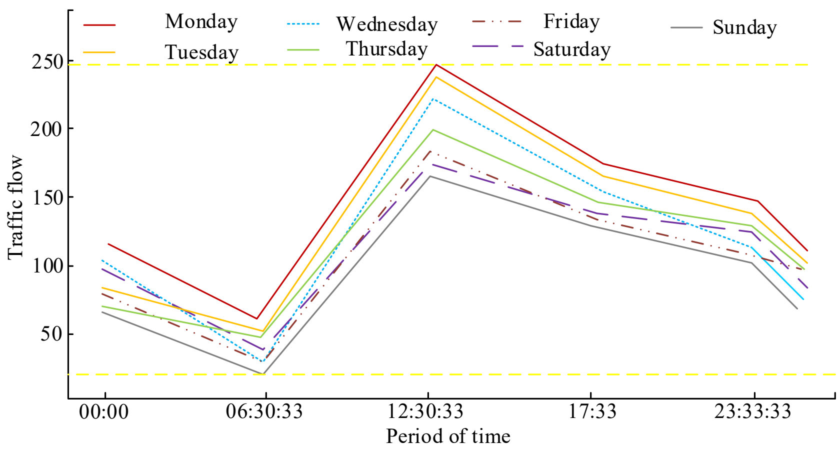


Fig. 1. Distribution of traffic flow in each time period of a week

Fig. 1 reveals distinct weekly traffic patterns with consistent weekday flows and varied weekend volumes, showing characteristic periodicity and correlation. Peak flows reach 150 units during the morning hours, then drop to near zero overnight. STTF prediction uses 5-15-minute intervals, with 5-minute granularity providing optimal accuracy. These patterns form the basis for constructing the STTF prediction model illustrated in Fig. 2.

Fig. 2 outlines the STTF prediction pipeline through key stages: target parameter definition (flow, speed, occupancy), data collection via monitoring equipment, data cleaning and feature extraction, dataset splitting, model selection and training, evaluation, and final deployment. This standardized workflow emphasizes data preprocessing and validation to ensure reliability, but faces challenges in capturing long-term temporal dependencies. To address this, our implementation employs Bi-LSTM, which enhances long-range dependency modeling by processing sequences bidirectionally and simultaneously learning from historical and future contexts through its gated mechanisms (e.g., the forget gate in Eq. (1)). This approach significantly improves prediction accuracy by comprehensively capturing temporal dynamics in traffic data.

$$F = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{1}$$

Eq. (1), F is the output of the FG. σ is the activation function. W_f is the weight vector. h_{t-1} indicates the hidden state of the previous moment. x_t represents output data. b_f indicates FG offset.

The Input Gate (IG) processing procedure is shown in Eq. (2).

$$i = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{2}$$

Eq. (2), i represents the IG output. i indicates the learning weight of the IG. The Output Gate (OG) processing procedure is shown in Eq. (3).

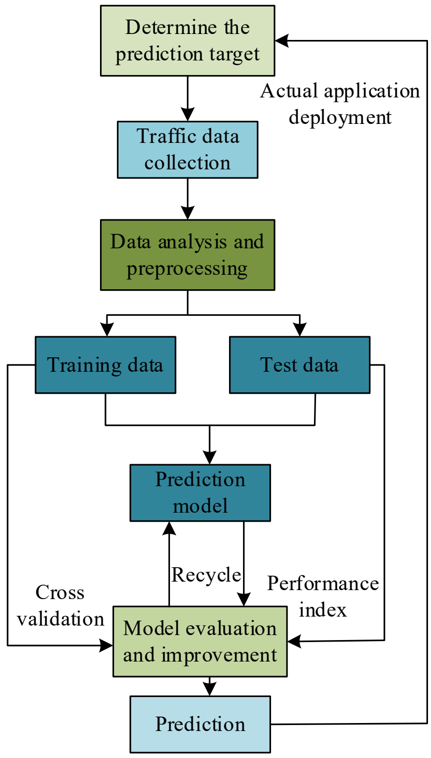


Fig. 2. STTF forecasting process

$$o = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{3}$$

Eq. (3), i is the output of the OG. b_o indicates the OG bias. The study used preprocessed traffic parameter datasets as training samples and fed them into a Bi-LSTM to train the model, ultimately achieving high-precision STTF time-series prediction. Its expression is shown in Eq. (4).

$$H = \tan\theta \cdot (F + i + o) \cdot C_t \tag{4}$$

Eq. (4), H is the output of the traffic flow time series prediction results. $\tan\theta$ represents the hyperbolic tangent activation function. C_t represents a matrix element. A Fully Connected Layer (FCL) is appended to the Bi-LSTM network as the regression output layer. It nonlinearly transforms the extracted temporal features into quantifiable traffic flow predictions, enabling end-to-end modeling from feature learning to predictive output, as formulated in Eq. (5).

$$y(t) = O\mu H + Y \tag{5}$$

Eq. (5), O is the number of training cycles for the FCL. y represents the predicted traffic volume at a specific time period at a road intersection. Y indicates the number of units in the Dense layer. O represents the topological depth of the Dense layer. Accurate STTF prediction relies heavily on real-time, comprehensive data collection. While traditional methods suffer from data fragmentation, RFID technology overcomes this limitation through its non-contact recognition capability. By automatically identifying targets and retrieving relevant data via radio frequency signals, RFID provides high-quality data for subsequent modeling. Its operational mechanism is illustrated in Fig. 3.

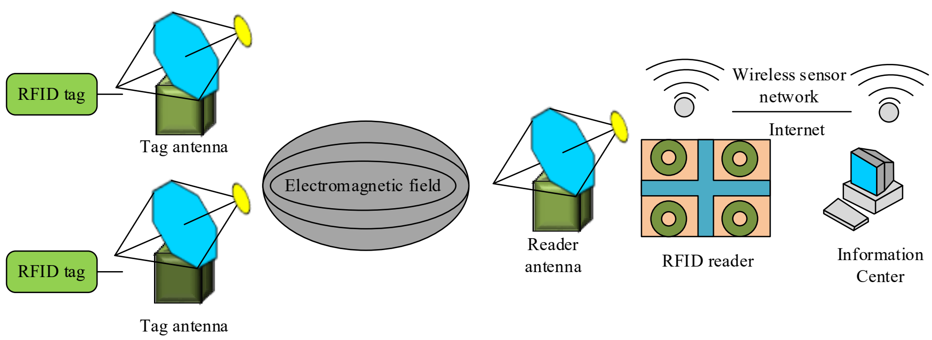


Fig. 3. RFID operational principle

Fig. 3 illustrates RFID technology's operational principle: the tag (left) activates and responds via the reader's electromagnetic signals; the central reader communicates wirelessly with tags; collected data is transmitted through wireless networks to information centers (right). This integrated process covers energy transmission, data exchange, and information processing. Raw RFID data (vehicle tag IDs and timestamps) undergo critical conversion into four quantitative traffic features. 1) Flow means unique vehicle counts per time interval. 2) Speed denotes calculated from dual-reader time differences. 3) Temporal features are time-series alignment with periodicity indicators. 4) Occupancy represents inferred density from vehicle counts relative to capacity. These processed features include flow, speed, density, and temporal indicators, forming a multi-dimensional input vector for subsequent Bi-LSTM and GBDT models, thereby enabling learning of spatiotemporal traffic dynamics. RFID tags power themselves through the reader's electromagnetic field, with received power defined in Eq. (6).

$$P_{tag} = P_{reader} \cdot G_{reader} \cdot G_{tag} \cdot \left(\frac{\lambda}{4\pi d}\right)^2 \cdot \eta \tag{6}$$

Eq. (6), P_{reader} represents the reader's transmission power. G_{reader} and G_{tag} represent the gains of the reader and the tag antenna. λ represents the wavelength of electromagnetic waves. d indicates the distance between the reader and the tag. η indicates the polarization matching efficiency factor. The minimum activation power P_{min} is required for the working distance label, and its maximum distance is shown in Eq. (7).

$$d_{max} = \frac{\lambda}{4\pi}\sqrt{\frac{P_{reader} \cdot G_{reader} \cdot G_{tag} \cdot \eta}{P_{min}}} \tag{7}$$

Eq. (7), d_{max} represents the maximum distance. Overall, STTF prediction in smart cities is achieved by integrating multiple data sources and using advanced algorithms for short-term forecasting. Meanwhile, the Bi-LSTM algorithm is used to process time series data, and RFID data collection technology is used to collect STTF data.

3.2. GBDT Optimized STTF Prediction Model for Smart Cities

The constructed STTF prediction model has improved its ability to process time series data and collect short-term data, but its learning and performance capabilities remain limited, and overfitting on the training set is often pronounced. Therefore, the GBDT model is further introduced into the research. The GBDT algorithm constructs a decision tree through gradient-based iterative updates, modeling the prediction residuals at each step and gradually reducing the LF in the negative gradient direction. When utilizing the LF as an evaluation criterion, a lower value signifies greater model accuracy. By continuously optimizing parameters, the model's performance gradually improves, exhibiting a decreasing error trend (Kilinc et al., 2023). The training diagram of the GBDT algorithm is shown in Fig. 4.

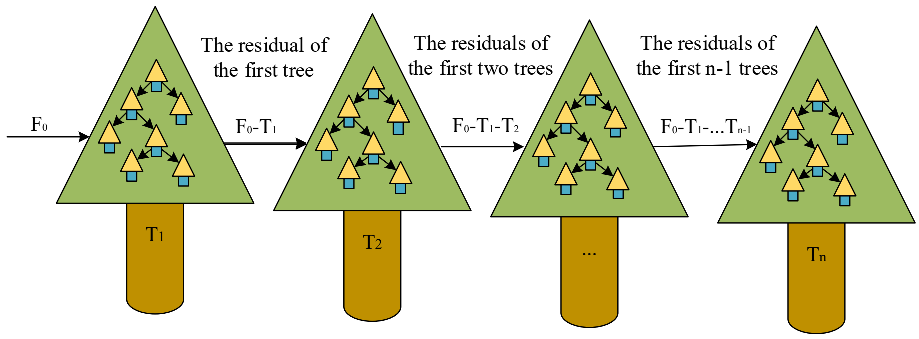


Fig. 4. Gradient lift decision tree process

Fig. 4 outlines GBDT's iterative training mechanism: starting from initial predictions, it sequentially constructs decision trees (T1 to Tn). Each new tree learns the residuals from previous predictions, progressively enhancing accuracy by error correction. This serial process culminates in a powerful ensemble model through a weighted combination of all trees, as formalized in Eq. (8).

$$r_{ni} = -\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]_{f(x)^* f_{n-1}(x)} \tag{8}$$

Eq. (8), r_{ni} represents pseudo residual. $\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}$ represents the gradient of the loss function L on the current model $f(x_i)$. $f(x)^* f_{n-1}(x)$ indicates that the "reference model" for gradient calculation is the model obtained from the previous round of training. The regression tree is presented in Eq. (9), which represents the final output of its LSTM unit.

$$\hat{f}(x) = f_M(x) = \sum_{m=1}^{M} T_m(x) \tag{9}$$

Eq. (9), $\hat{f}(x)$ represents the final prediction function. $f_M(x)$ represents the model after the M th iteration. $T_m(x)$ represents the weak model after the M th iteration. XGBOOST is an enhancement of traditional GBDT that introduces regularization mechanisms to mitigate overfitting, supports user-defined LFs, integrates parallel computing, and efficiently handles missing data. It finds extensive applications across diverse areas, including classification, forecasting, regression analysis, and sorting. Therefore, in the GBDT enhancement method, the XGBOOST method is selected for fusion, as shown in Fig. 5.

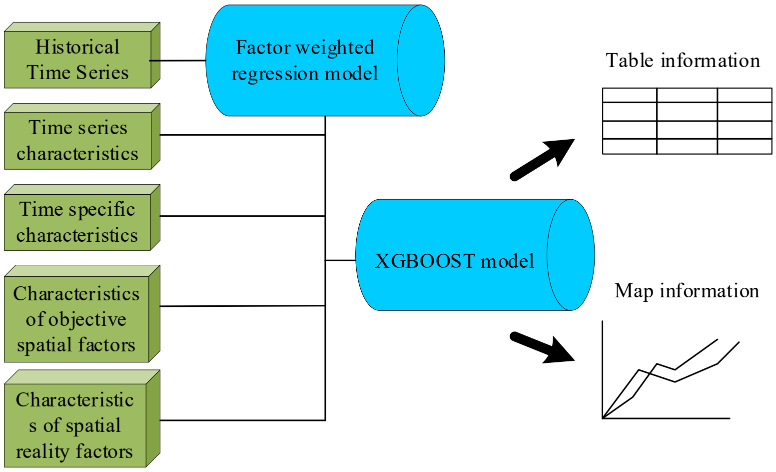


Fig. 5. XGBOOST integration process

As shown in Fig. 5, the left input contains five types of features: historical time series data, time series characteristics, time specificity characteristics, spatial objective factors, and spatial reality factors. These features are fused and learned through models such as XGBOOST. The table on the right shows that the model can output customized predictions for different scenarios (Scenario 1/2/3), reflecting the full process design from multi-dimensional feature extraction to scenario-based application, and is suitable for intelligent prediction tasks that require comprehensive spatiotemporal characteristics. The objective function of XGBOOST is presented in Eq. (10).

$$J^t = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) + constant \tag{10}$$

Eq. (10), J^t is the objective function value of the t th iteration. y_i represents the true value. $\Omega(f_t)$ represents regularization. $constant$ represents a constant term. The Taylor expansion second-order equation is shown in Eq. (11).

$$f(x+\Delta x) \cong f(x) + f'(x)\Delta x + \frac{1}{2} f''(x)\Delta x^2 \tag{11}$$

Eq. (11), $f(x+\Delta x)$ represents the true value of the function f at $x+\Delta x$. $f(x)$ indicates the function value of function f at x. $f'(x)$ represents the first derivative of function f at x. Δx indicates the increment of the independent variable. The first and second derivatives of the LF are defined as shown in Eq. (12).

$$\partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) = g_i \tag{12}$$

Eq. (12), $\partial_{\hat{y}^{(t-1)}}$ represents the first derivative partial derivative operator. g_i is the first derivative of the LF with respect to the previous round of prediction $\hat{y}_i^{(t-1)}$ of the i th sample. The second derivative is shown in Eq. (13).

$$\partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) = h_i \tag{13}$$

Eq. (13), $\partial^2_{\hat{y}^{(t-1)}} l$ represents the second-order derivative partial derivative operator. h_i indicates the calculation result of the second derivative. It ultimately reduces the objective function to the form presented in Eq. (14).

$$J^t = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j}{H_j + \lambda} + \gamma T \tag{14}$$

Eq. (14), G_j denotes the initial derivative of the LF concerning the model's predicted value. H_j represents the second derivative of the LF with respect to the model's predicted value. T represents the quantity of leaf nodes in the decision tree. The splitting gain of the objective function is shown in Eq. (15).

$$Gain = \frac{1}{2}\left\{\frac{G_R}{H_R+\lambda}+\frac{G_L}{H_L+\lambda}-(\frac{G_R+G_L}{H_R+H_L+\lambda})\right\}-\gamma \tag{15}$$

Eq. (15), $Gain$ represents the splitting gain. G_R and G_L represent the sum of the first derivatives of the right and left child nodes. H_R and H_L represent the sum of the second derivatives of the right and left child nodes. λ represents the regularization coefficient. γ represents the penalty coefficient for leaf node complexity. The performance of both GBDT and XGBoost models is highly dependent on their hyperparameter configuration. After preliminary experimentation and tuning, the following key hyperparameters were set for model training in this study: the number of estimators (trees) was set to 100, the learning rate to 0.1, and the maximum tree depth to 6. For the XGBoost model specifically, a subsample ratio of 0.8 was used to introduce randomness and prevent overfitting, and a colsample_bytree value of 0.8 was applied to control the fraction of features used for training each tree.

These parameters were optimized via a grid search to balance model complexity and generalization, minimizing MAE on the validation set. The final GBDT-based STTF prediction framework is illustrated in Fig. 6.

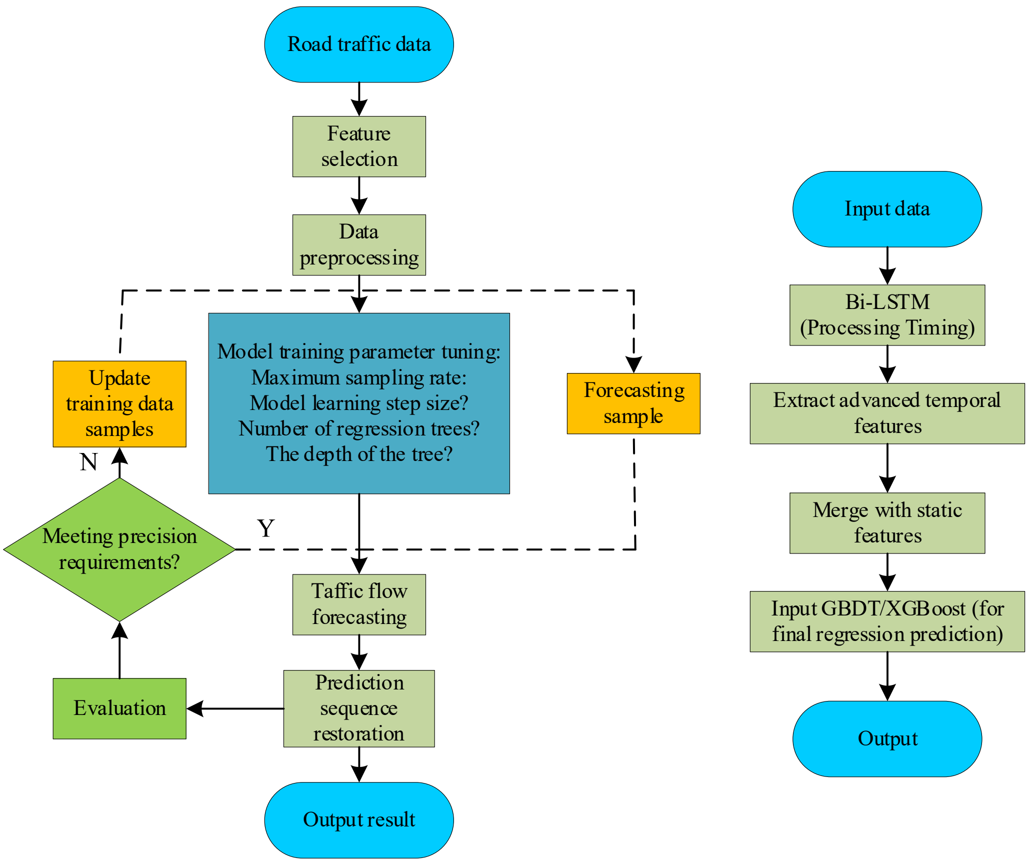


Fig. 6. Framework of the proposed GBDT-based STTF prediction model for smart cities

As shown in Fig. 6, the proposed smart city STTF prediction model follows a structured workflow: It begins with feature selection and data preprocessing, including tuning key parameters such as sampling rate, learning rate, number of trees, and tree depth. This is followed by an iterative training phase in which the model is continuously updated with training data until the accuracy requirements are met. The process then proceeds to traffic flow prediction, model evaluation, and sequence reconstruction, ultimately generating the final prediction outputs. This end-to-end pipeline represents a complete closed-loop machine learning process from data preparation to prediction. The core innovation lies in the serial fusion of Bi-LSTM and GBDT/XGBoost. In this architecture, Raw time-series data is first processed by Bi-LSTM, which serves as a temporal feature extractor to capture complex bidirectional long-term dependencies. The high-level sequential features from the Bi-LSTM are concatenated with non-sequential features (weather and time metadata). This enriched feature vector is fed into GBDT/XGBoost, which acts as a powerful non-linear regressor to model complex interactions and produce final predictions. This synergistic design allows Bi-LSTM to compensate for GBDT's limitations in temporal modeling, while GBDT enhances the model's ability to handle diverse feature interactions and improve generalization. The model uses

RFID technology for data acquisition and incorporates XGBoost enhancements, including regularization and parallel computing, to mitigate overfitting and optimize performance. The resulting GBDT-based STTF prediction framework effectively addresses the challenges of smart-city traffic-flow forecasting.

4. Validation of GBDT Smart City STTF Prediction Model

4.1. Performance Validation of STTF Prediction Model for Smart Cities

To assess the validity of the GBDT-based STTF prediction model for smart cities, the study compared it with traditional prediction models and real values. The research experimental data were sourced from the ZB laboratory. The experiments were conducted using the publicly available ("city C Urban Traffic Flow (CUTF-2025)") dataset, provided by the ZB Laboratory. The data was collected from RFID readers installed at major intersections and arterial roads in the D area of city C (as shown in Fig. 7). The raw data includes anonymized vehicle tag IDs and precise timestamps. The data sampling frequency is 1 reading per vehicle per reader, which is then aggregated into 5-minute intervals to form the time series for traffic flow (vehicle count) and average speed. The dataset spans from June 1, 2025, to June 15, 2025, providing a continuous two-week period for training and evaluation. The preset parameters of the experimental model are presented in Table 1.

Table 1. Training conditions and technical parameters

Project	Parameter
Input layer size	60 × 5 (e.g., 60 time steps × 5 features)
Activation function	RELU
Pooling layer	Max pooling
Size of pooling layer	2×2
Number of FCLs	2
Learning rate	0.01

Based on the above model parameters, this study assessed the model's predictive performance by comparing predicted values with true values, evaluating various weather prediction metrics, and assessing prediction accuracy and error rates. First, the study collected experimental data and selected area D in city C, as shown in Fig. 7.

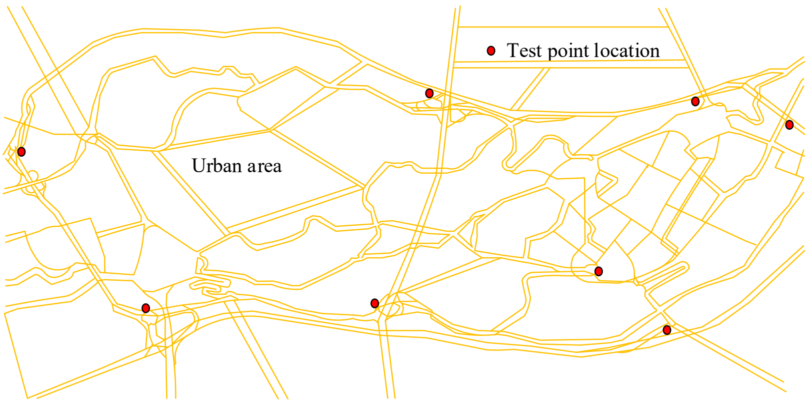


Fig. 7. Road sections in area D of City C

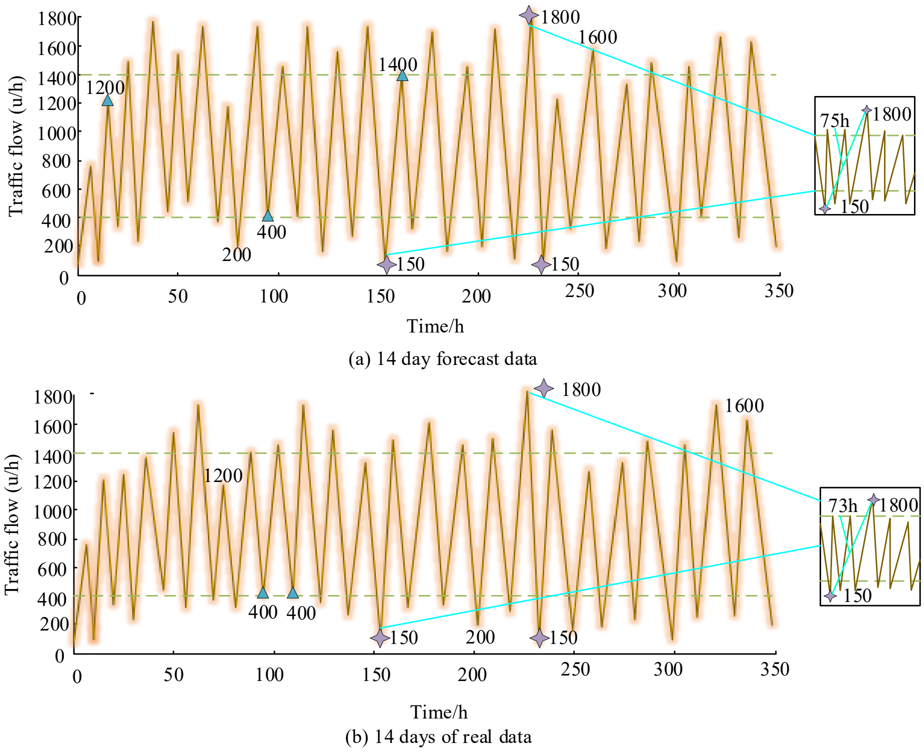
As shown in Fig. 7, the road segments within the D area road network of city C were selected. In the experiment, RFID data on vehicles on the road network were collected over 14 days, from June 1, 2025, to June 15, 2025. Subsequently, the RFID data for each road section were summarized and processed, and a complete time series was constructed from 0:00 to 24:00, with 5-minute intervals. Then, the model's predictive performance was evaluated by comparing its predictions with real traffic values. The results are shown in Fig. 8.

Fig. 8 demonstrates the model's strong predictive performance over a 14-day period, with forecasts (peak: 1800, trough: 150) closely matching actual values. Quantitative evaluation confirms high accuracy, with an MAE of 28.5, an RMSE of 45.2, and a MAPE of 3.07%. As shown in Table 2, the hybrid Bi-LSTM-GBDT model consistently outperforms all baseline methods, validating the effectiveness of the integrated approach.

To comprehensively evaluate the model's performance, a comparison was made between the study and the current mainstream deep learning models (CNN-LSTM, Transformer). The experimental results showed that the proposed Bi-LSTM-GBDT hybrid model outperformed the comparison model across the three key indicators: MAE, RMSE and MAPE. Although the Transformer model had advantages in handling long-range dependencies, it was vulnerable to noise in short-term traffic flow prediction and had relatively long training times. The CNN-LSTM model performed well at extracting spatial features but failed to fully capture interactions between temporal and high-dimensional features. The model extracted bidirectional temporal series features using a Bi-LSTM, then performed high-dimensional feature fusion and regularization with GBDT, achieving a balance between accuracy and efficiency.

4.2. Practical Application of STTF Prediction Model for Smart Cities Based on Gradient Boosting Decision Tree

Building on the verification of the GBDT-based smart city STTF prediction model's predictive prowess, this research aimed to further substantiate its practical value in application. Experiments revealed that meteorological conditions exerted a considerable influence on residents' travel choices, making them a pivotal factor in forecasting traffic flow. To this end, the study conducted comparative experiments under varying weather conditions and model settings, evaluating the performance of the GBDT-based smart city STTF prediction model across four distinct weather scenarios. The results of the four different weather predictions are shown in Fig. 9.



(a) 14 day forecast data

(b) 14 days of real data

Fig. 8. Comparison of predicted value and true value

Table 2. Comparative performance of proposed model against baselines

Model	MAE	RMSE	MAPE (%)	Training Time (s)
Proposed (Bi-LSTM+GBDT)	22.4	38.7	2.85	312
Bi-LSTM Only	35.1	58.2	4.51	285
CNN-LSTM	29.8	49.3	3.62	298
Transformer	32.4	53.1	3.94	410
XGBoost Only	41.6	65.3	5.34	120
ARIMA	52.8	79.1	6.77	5

Fig. 9 demonstrates the model's effective response to weather conditions: rainy road traffic flow was significantly lower than sunny conditions, with storm conditions showing further reduction. Critically, predictions remained stable without anomalies despite weather variations. Statistical validation through a two-sample t-test confirmed the observed 20% flow reduction during rain is statistically significant ($p < 0.001$). Subsequently, the study evaluated the model's prediction speed and efficiency through a comparative analysis with traditional methods, as shown in Fig. 10.

As shown in Fig. 10(a), the improved model maintained a stable prediction speed in the range of 2-1m/h, with the fastest predicted value reaching 1m/h, and the overall performance was stable. As shown in Fig. 10(a), the prediction results of traditional methods not only had low velocities (3-2m/h) but also fluctuated violently, with multiple instances of zeroing. Overall, the improved model was faster and more stable than traditional models in terms of prediction speed, and time-series data showed that its advantages were sustained. Finally, to verify the prediction error rate of GBDT's STTF prediction model for smart cities, the study compared it with that of traditional prediction models, and the results are shown in Fig. 11.

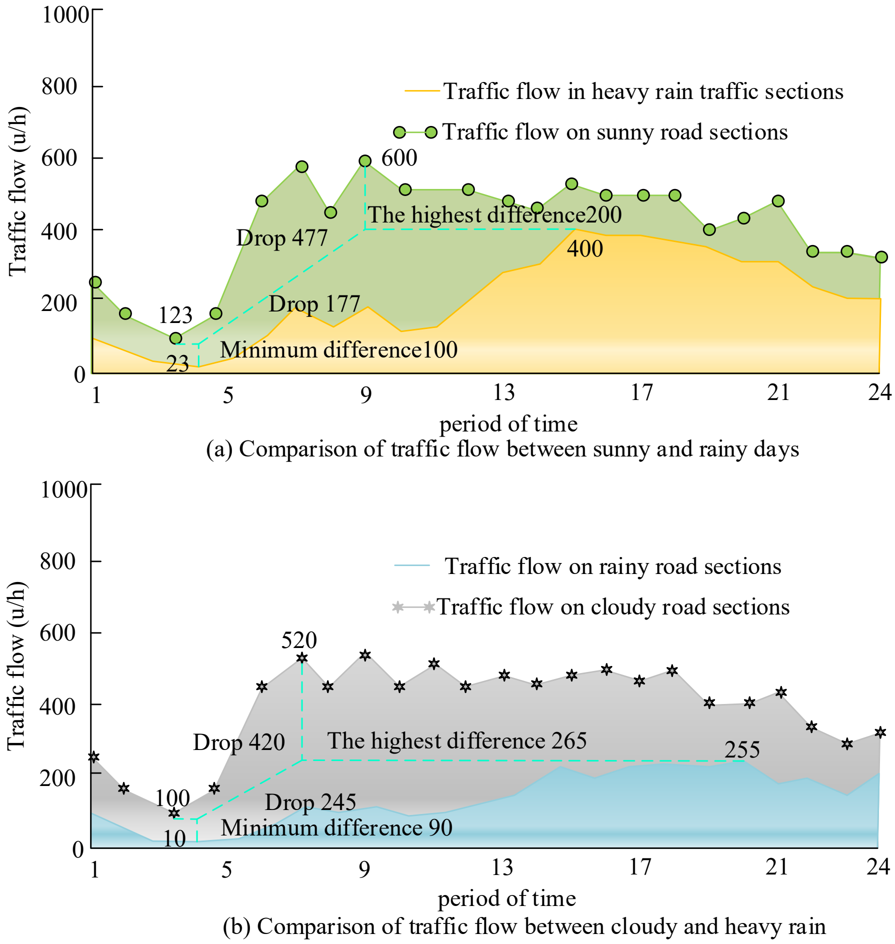


Fig. 9. Different weather forecasts

As shown in Fig. 11(a), the true value data was 95.11%. The improved model achieved a prediction accuracy of 92.13%, which was close to the true value and performed well. As shown in Fig. 11(b), the true value data was 95.11%. The prediction accuracy of traditional models was 88.89%, which was 6.22 percentage points lower than the true value, indicating a large error. Overall, the F-DNN model's actual warning performance was excellent, significantly better than that of the comparison model. The above outcomes revealed that the optimization measures for GBDT's smart city STTF prediction model significantly improved prediction accuracy and yielded performance significantly better than traditional methods.

5. Conclusion

This study developed a GBDT-based STTF prediction model that integrates Bi-LSTM for temporal processing and RFID for data acquisition, effectively addressing traditional model's limitations in long-term dependency modeling and overfitting. Experimental results demonstrated strong alignment with actual data, achieving a maximum flow of 1800 vehicles with only 3.07% error rate. The model accurately quantified weather impacts, showing 20% flow reduction during rain and an additional 15% decrease during storms. The optimized model achieved 1-2ms inference speed per sample, over 50% faster than traditional models, with 92.13% prediction accuracy. However, the model currently performs best under recurrent traffic patterns and shows limitations during non-recurrent extreme events. First, it is necessary to move from "black box" prediction to explainable, reliable decision support for proactive traffic control. Second, it is necessary to overcome the model's bottleneck in generalizing to sudden abnormal events to ensure the system's robustness. Third, the current centralized data processing model has limitations in privacy, energy consumption, and scalability, and there is an urgent need to evolve towards a privacy-protected, distributed learning framework. The model proposed in the research not only achieved excellent prediction accuracy and speed but also demonstrated strong engineering applicability. In the future, it can be connected to urban traffic control centers, navigation platforms and emergency management systems via APIs to enable real-time application of prediction results. The next step of research will focus on conducting pilot applications of the model in real traffic signal control systems and exploring its generalizability across multiple cities and scenarios.

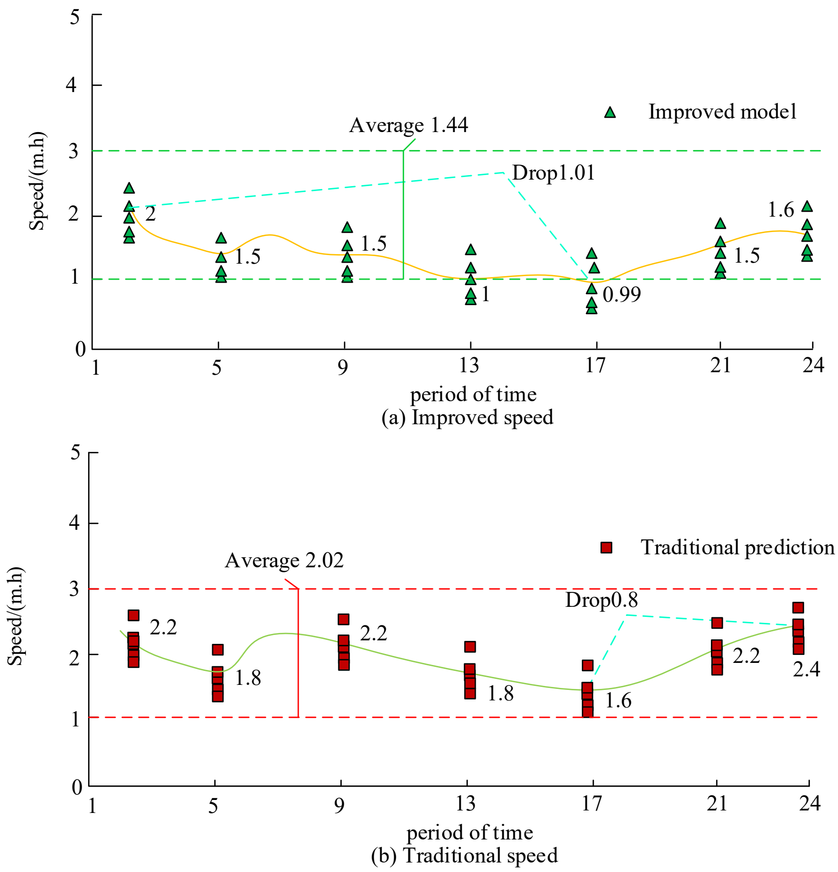


Fig. 10. Comparison between traditional and improved prediction speed

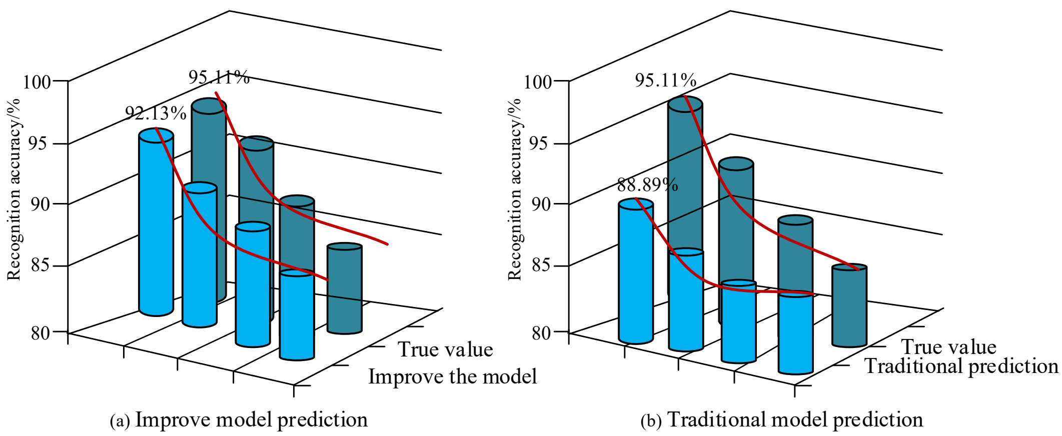


Fig. 11. Comparison of model prediction error rate

6. Practical Application Cases and Integration Potential

The proposed Bi-LSTM-GBDT hybrid model not only demonstrated high predictive accuracy but also held significant potential for real-world deployment in smart city traffic management systems. The following use cases illustrate its practical applicability: (1) Adaptive Traffic Signal Control. By providing real-time 5-15 minute traffic flow forecasts, the model can dynamically optimize signal timing at intersections. For instance, if a 20% increase in traffic is predicted for a specific intersection during the morning peak, the system can proactively extend green light duration to reduce queue length and improve throughput. (2) Dynamic Congestion Pricing and Route Guidance. The model's predictions can inform dynamic congestion pricing algorithms by identifying anticipated congestion hotspots. Simultaneously, predicted traffic conditions can be disseminated to navigation platforms (e.g., Google Maps, Baidu Maps) to enable dynamic route guidance, thereby distributing traffic load across the network. (3) Emergency Traffic Management and Incident Response. In the event of non-recurrent incidents (e.g., accidents, public events), the model can rapidly assess the impact on surrounding road networks. This supports emergency vehicle routing, temporary traffic control measures, and enhances overall urban traffic resilience. Future work will involve collaboration with municipal traffic management agencies to pilot the model in real-time traffic control platforms and validate its performance in operational environments.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The author used AI tools solely for language editing and to improve readability. The author reviewed and verified all content and takes full responsibility for the accuracy and integrity of the manuscript.

References

Bagga, P. J., Patel, K. M., Makhesana, M. A., Şirin, Ş., Khanna, N., Krolczyk, G. M., and Chauhan, K. C. (2023). Machine vision-based gradient-boosted tree and support vector regression for tool life prediction in turning. *The International Journal of Advanced Manufacturing Technology,* 126(1), 471-485.
https://doi.org/10.1007/s00170-023-11137-2

Demir, S. and Sahin, E. K. (2023). An investigation of feature selection methods for soil liquefaction prediction based on tree-based ensemble algorithms using AdaBoost, gradient boosting, and XGBoost. *Neural Computing and Applications,* 35(4), 3173-3190.
https://doi.org/10.1007/s00521-022-07856-4

Douiba, M., Benkirane, S., Guezzaz, A., and Azrour, M. (2023). An improved anomaly detection model for IoT security using decision tree and gradient boosting. *The Journal of Supercomputing,* 79(3), 3392-3411.
https://doi.org/10.1007/s11227-022-04783-y

Hasanvand, M., Nooshyar, M., Moharamkhani, E., and Selyari, A. (2023). Machine Learning Methodology for Identifying Vehicles Using Image Processing, *Artificial Intelligence Applications,* 1(3), 170-178.
https://doi.org/10.47852/bonviewAIA3202833

Jing, H., Chen, Z., Wang, X., Wang, X., Ge, L., Fang, G., and Xiao, D. (2023). Gradient boosting decision tree for rotor temperature estimation in permanent magnet synchronous motors. *IEEE Transactions on Power Electronics,* 38(9), 10617-10622.
https://doi.org/10.1109/TPEL.2023.3291464

Kilinc, H. C., Ahmadianfar, I., Demir, V., Heddam, S., Al-Areeq, A. M., Abba, S. I., and Yaseen, Z. M. (2023). Daily scale river flow forecasting using hybrid gradient boosting model with genetic algorithm optimization. *Water Resources Management,* 37(9), 3699-3714.
https://doi.org/10.1007/s11269-023-03522-z

Li, Z., Xu, H., Gao, X., Wang, Z., and Xu, W. (2024). Fusion attention mechanism bidirectional LSTM for short-term traffic flow prediction. *Journal of Intelligent Transportation Systems,* 28(4), 511-524.
https://doi.org/10.1080/15472450.2022.2142049

Mehrpour, O., Saeedi, F., Vohra, V., Abdollahi, J., Shirazi, F. M., and Goss, F. (2023). The role of decision tree and machine learning models for outcome prediction of bupropion exposure: A nationwide analysis of more than 14 000 patients in the United States. *Basic and Clinical Pharmacology and Toxicology,* 133(1), 98-110.
https://doi.org/10.1111/bcpt.13865

Preethi, P. and Mamatha, H. R. (2023). Region-Based Convolutional Neural Network for Segmenting Text in Epigraphical Images. *Artificial Intelligence Applications,* 1(2), 119-127.
https://doi.org/10.47852/bonviewAIA2202293

Tan, G., Zhou, T., Huang, B., Dou, H., Song, Y., and Lin, Z. (2024). A noise-immune and attention-based multi-modal framework for short-term traffic flow forecasting. *Soft Computing*, 28(6), 4775-4790.
https://doi.org/10.1007/s00500-023-09173-x

Tiwari, P. (2024). The machine learning framework for traffic management in smart cities. *Management of Environmental Quality: An International Journal,* 35(2), 445-462.
https://doi.org/10.1108/MEQ-08-2022-0242

Wang, Q., Song, L., Zhao, J., Wang, H., Dong, L., Wang, X., and Yang, Q. (2023). Application of the gradient boosting decision tree in the online prediction of rolling force in hot rolling. *The International Journal of Advanced Manufacturing Technology*, 125(1), 387-397.
https://doi.org/10.1007/s00170-022-10716-z

Wang, Y., Ke, S., An, C., Lu, Z., and Xia, J. (2024). A hybrid framework combining LSTM NN and BNN for short-term traffic flow prediction and uncertainty quantification. *KSCE Journal of Civil Engineering,* 28(1), 363-374.
https://doi.org/10.1007/s12205-023-2457-y

Wei, L., Yu, Z., and Ye, Y. (2023). Spatiotemporal Graph Convolutional Network with Attention for Urban Traffic Flow Forecasting. *IEEE Transactions on Intelligent Transportation Systems*, 24(8), 8525-8536.
https://doi.org/10.1109/TITS.2023.3267101

Wu, Z., Zhu, J., Li, Q., and He, B. (2023). Deltaboost: Gradient boosting decision trees with efficient machine unlearning. *Proceedings of the ACM on Management of Data,* 1(2), 1-26.
https://doi.org/10.1145/3589313

Xu, X., Jin, X., Xiao, D., Ma, C., and Wong, S. C. (2023). A hybrid autoregressive fractionally integrated moving average and nonlinear autoregressive neural network model for short-term traffic flow prediction. *Journal of Intelligent Transportation Systems,* 27(1), 1-18.
https://doi.org/10.1007/s12205-023-2457-y

Yang, X., Jia, J., Li, Q., Zhu, R., Yang, J., Liu, Z., and Li, X. (2024). Stress-assisted corrosion mechanism of 3Ni steel by using gradient boosting decision tree machining learning method. *International Journal of Minerals, Metallurgy and Materials,* 31(6), 1311-1321.
https://doi.org/10.1007/s12613-023-2661-2

Yang, Z. and Wang, C. (2023). Short-term traffic flow prediction based on AST-MTL-CNN-GRU. *IET Intelligent Transport Systems,* 17(11), 2205-2220.
https://doi.org/10.1049/itr2.12400

Yu, W., Wu, S., and Huang, M. (2023). MmgFra: A multiscale multigraph learning framework for traffic prediction in smart cities. *Earth Science Informatics*, 16(3), 2727-2739.
https://doi.org/10.1007/s12145-023-01068-7

Zhang, W., He, Y., Wang, L., Liu, S., and Meng, X. (2023). Landslide Susceptibility mapping using random forest and extreme gradient boosting: A case study of Fengjie, Chongqing. *Geological Journal,* 58(6), 2372-2387.
https://doi.org/10.1002/gj.4683

Lili Zhang earned her master's of laws from the School of Law at Henan University in 2016. Currently, she is a lecturer at Henan Police College and serves concurrently as the Secretary of the Youth League Branch of the Traffic Management Engineering Department. In recent years, she has led and completed one college-level teaching reform project, participated in the completion of one teaching reform project under the Henan Provincial Department of Education, and contributed to 26 other projects. She has won the second prize in the National First Smart Traffic Policing Innovation Skills Competition and has been recognized multiple times as an Outstanding Youth League Cadre at the Provincial-Level.