



# Performance Evaluation and Comparison of a Stacking - Based Ensemble Model for Traffic Speed Prediction

Yuanzhe Cheng<sup>1</sup>, Haoyang Lv<sup>2</sup>, Hanrui Chen<sup>3,\*</sup>, Chengjie Ni<sup>4</sup>, Yuyang Hu<sup>5</sup>

<sup>1</sup>School of Yingyang Financial Technology, Zhejiang University of Finance and Economics, Hangzhou, China

<sup>2</sup>Jingying High School, Shijiazhuang, China

<sup>3</sup>Office of Scientific Research and Foreign Cooperation, Zhejiang-California International NanoSystems Institute, Hangzhou, China

<sup>4</sup>School of Information, Zhejiang University of Finance and Economics, Hangzhou, China

<sup>5</sup>School of Mechatronic Engineering, Sanming University, Sanming, China

## Email address:

[cyz1467927689@zufe.edu.cn](mailto:cyz1467927689@zufe.edu.cn) (Yuanzhe Cheng), [3038290281@qq.com](mailto:3038290281@qq.com) (Haoyang Lv), [1059808669@qq.com](mailto:1059808669@qq.com) (Hanrui Chen), [1374706274@zufe.edu.cn](mailto:1374706274@zufe.edu.cn) (Chengjie Ni), [3538686130@qq.com](mailto:3538686130@qq.com) (Yuyang Hu)

\*Corresponding author

## To cite this article:

Yuanzhe Cheng, Haoyang Lv, Hanrui Chen, Chengjie Ni, Yuyang Hu. Performance Evaluation and Comparison of a Stacking - Based Ensemble Model for Traffic Speed Prediction. *International Journal of Economics, Finance and Management Sciences*.

Vol. 11, No. 5, 2023, pp. 255-260. doi: 10.11648/j.ijefm.20231105.15

**Received:** August 31, 2023; **Accepted:** October 23, 2023; **Published:** October 31, 2023

---

**Abstract:** Accurate prediction of traffic speed plays a key role in easing traffic congestion and improving road utilization efficiency. However, traditional traffic analysis methods often fail to capture complex traffic patterns. With the rapid development of artificial intelligence, traffic prediction using machine learning models has become a focal point of research. This study aims to explore the application of machine learning models in traffic speed analysis and prediction, by constructing a multi-model fusion method through stacking-based ensemble. Initially, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) networks, Multilayer Perceptron (MLP), Linear Regression (LR), K-Nearest Neighbors (KNN), and Support Vector Regression (SVR) were selected as base models to predict the traffic speed. Then, their predictive performance was improved by optimizing the model parameters through Bayesian optimization algorithm. After contrast experiments, LR was adopted as a meta-regressor to merge the predictive factors of the optimized base models into a stacking-based ensemble model, improving the performance of traffic speed prediction further. Finally, the proposed ensemble model was evaluated using multiple traffic datasets. The experimental validation demonstrates that the ensemble model achieves outstanding performance in predicting traffic speed. The findings of this study highlight the potential of machine learning models, particularly the stacking-based ensemble method, in predicting the traffic speed.

**Keywords:** Ensemble Model, Data Mining, Traffic Speed Prediction, Machine Learning, Bayesian Optimization

---

## 1. Introduction

In recent years, the rapid development of artificial intelligence and machine learning has catalyzed a big shift in the domain of intelligent transportation systems (ITS). Powered by advancements in Internet and Internet of Things technologies, these modern ITS have shown significant promise in alleviating traffic congestion and optimizing road network utilization. Traffic speed prediction, as underscored by recent studies [1], directly influences the alleviation of traffic congestion, thereby enhancing the efficiency of road

network utilization. This is mainly due to the influence of many complex factors such as weather conditions, driver behaviors, varying road conditions, holidays, and special events. By predicting traffic speed, future traffic information can be provided to facilitate transportation decision-making, e.g., enhancing the current transportation efficiency and improving the future traffic network layout.

The historical reliance on statistical models, exemplified by the usage of Autoregressive Integrated Moving Average (ARIMA) and Seasonal Decomposition of Time Series (STL), has garnered some successes [2]. However, the nonlinear and

highly volatile nature of traffic speed data often surpasses the predictive capacities of these traditional models [3]. Recent work have explored the application of advanced data mining technologies, including machine learning and deep learning, for traffic speed analysis. By using historical data, these modern techniques can unearth traffic speed patterns that traditional statistical models may overlook [4].

Guided by these insights, this study embarks on an exploration of an ensemble learning paradigm, specifically tailored to address different situation of traffic speed prediction. An in-depth analysis was first conducted on the traffic speed data, wherein promising foundational predictive models such as Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) networks, Support Vector Regression (SVR), Linear Regression (LR), Multi-Layer Perceptron (MLP), and the k-Nearest Neighbors algorithm (KNN) were trained and compared. Subsequent to this, parameter optimization method, in particular, Bayesian optimization algorithm searches for the most optimal parameter configuration, enhancing the model performance. It also improves the predictive accuracy and generalization capability of the stacking-based ensemble models that consists of optimized base models. Building on this, several competent base models were pitted against each other in comparative experiments to ascertain the most appropriate meta-regressor. Ultimately, by stacking diverse optimized base models, an ensemble learning model, with LR used as its meta-regressor, was constructed. Experimental validations, as delineated in subsequent sections, reveal the ensemble model's superior performance in the domain of traffic speed predictions. Capitalizing on these advancements, this study also underscores the potential of stacking-based ensemble models in other fields.

The remainder of this paper is organized as follows: Section 2 reviews previous research efforts in data mining, ensemble learning, and traffic prediction. Section 3 explores the data preprocessing process, the proposed ensemble learning model, and the underlying base models. Section 4 demonstrates the experimental design and results. Section 5 concludes the primary findings of this study and discusses future research directions and potential applications.

## 2. Related Work

### 2.1. Application of Data Mining

Data mining delves into the intricate process of unearthing valuable, implicit insights from vast and often incomplete datasets, translating them into comprehensible principles and patterns. It is an interdisciplinary science, applicable across various domains, and encompassing technologies from artificial intelligence, mathematical statistics, pattern recognition, and databases. Not only can data mining unlock the latent value of abundant idle data, but it also enables businesses and institutions to predict and strategize for the future based on known data.

In the realms of social sciences and natural sciences, data mining techniques have been widely deployed. Baesens et al.

[5] evaluated a myriad of classification algorithms in the context of genuine credit scoring datasets, spanning from classical algorithms like logistic regression and decision trees to avant-garde kernel-based classifiers such as support vector machines and least squares support vector machines (LS-SVMs). Ngai et al. [6] gave a systematic, recognizable, and comprehensive academic literature review of six categories of data mining techniques (including classification, regression, clustering, forecasting, anomaly detection, and visualization) applied to Financial Fraud Detection (FFD).

In the natural sciences, Bui et al. [7] devised and validated five spatial prediction models of shallow landslide hazard using SVM, MLP Neural Nets, radial basis function neural networks (RBF Neural Nets), kernel logistic regression (KLR), and logistic model trees (LMTs). Russ et al. [8] concentrated on predicting wheat yield using neural networks within data mining. Komi et al. [8] explored early predictions for diabetes through five different data mining methods, including GMM (Gaussian Mixture Model), SVM, LR, ELM (Extreme Learning Machine), and ANN (Artificial Neural Network).

With the thriving development of advanced information technology and the entrance of intelligent transportation paradigm into the mainstream vision, data mining techniques have also contributed to the construction of intelligent transportation systems. Wu et al. [10] developed a vehicle accident prediction model using classification-based data mining techniques, including steps such as data preparation, mining mechanism implementation, and validation. Huang et al. [11] proposed a multi-view dynamic graph convolutional network to capture different levels of spatiotemporal dependencies for traffic flow prediction.

### 2.2. Application of Ensemble Methods

Ensemble learning is a distributed machine learning framework that resolves problems unattainable by a single model by combining multiple base models. It has been proven to broadly enhance the performance and robustness of machine learning algorithms and models. For example, He et al. [12] employed two tree-based classifiers, including random forests and extreme gradient boosting, as basic classifiers to form an ensemble model for credit scoring, boasting exceptional performance and high adaptability across diverse imbalanced ratio datasets.

Stacking is a popular ensemble method, which usually consists of a meta-regressor and several base models. Through appropriate adjustments to the dataset, quantity of classifiers, types of classifiers, and related model parameters, stacking-based ensemble learning can achieve enhanced performance.

Stacking-based ensemble learning has also shown strong performance in prediction. In natural disaster prediction, Cui et al. [13] put forth a stacking-based ensemble learning method for earthquake casualty prediction. In the field of medical diagnosis, Khoei et al. [14] applied a stacking-based ensemble learning model to detect early stages of Alzheimer's disease concerning biological genetics. Wang et al. [15] used decision trees in conjunction with

stacking-based ensemble learning methods for the interpretation of prostate cancer detection.

### 2.3. Application in Traffic Prediction

Traffic speed prediction is a vital part of intelligent transportation system. Examples of traditional traffic speed prediction methods include spatiotemporal correlative k-nearest neighbor models for short-term traffic multi-step prediction [16], online-SVR for short-term traffic flow prediction under typical and atypical conditions [17], and time-series models for predicting main road traffic volumes in cities [18]. With the development of machine learning, various new data mining methods have been applied to traffic prediction. Wei et al. [19] established empirical mode decomposition and neural networks for short-term subway passenger flow prediction. Wang et al. [20] proposed a meta-learning-based spatiotemporal graph attention network for traffic signal control. Ma et al. [21] utilized neural network methods for traffic speed and traffic volume predictions.

Accurate prediction of traffic speed can furnish future

information for traffic-related system decision-making, enhancing current traffic efficiency and facilitating future transportation network planning. However, few researches have applied stacking-based ensemble learning methods to traffic speed prediction studies. This study proposes a new traffic speed prediction model through stacking-based ensemble learning methods, and optimizes it through various techniques such as Bayesian optimization algorithm.

## 3. Method

### 3.1. Data Description

As shown in Table 1, this study employed two datasets from PeMS (Performance Measurement System), spanning from 01/06/2017 to 30/06/2017. PeMS is a database utilized for monitoring and analyzing the traffic speed. By recording the rate of transportation tools through sensors, the measurement and computation of traffic speed data are accomplished.

**Table 1.** Statistics of the two datasets used in this study.

Dataset	Area	#Sensor	#Records	Time Span
PeMSD4	Bay Area	400038	8640	June 1, 2017 to June 30, 2017
PeMSD7	Los Angeles	717431	8640	June 1, 2017 to June 30, 2017

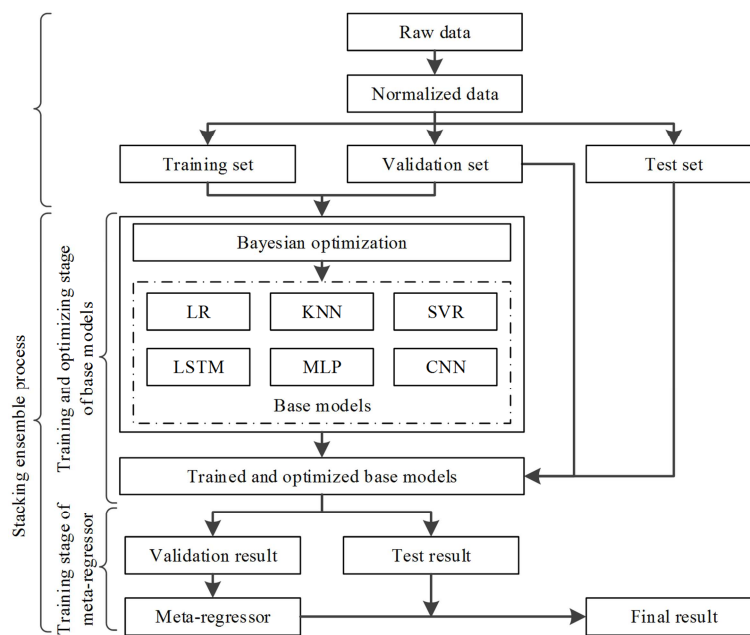
### 3.2. Data Normalization

Data normalization transforms data with various scales and ranges into a unified standard, avoiding biases stemming from data disparities. By mapping the data into a common scale, the learning efficacy of the model can be enhanced, and its sensitivity to input data is reduced. Applying the following normalization formula, the data will be scaled to a range from 0 to 1.

$$X(\text{normalized}) = \frac{x - \min(x)}{\max(x) - \min(x)}$$

### 3.3. Ensemble Model

The workflow of this section is shown in Figure 1.



**Figure 1.** Schematic overview of the stacking-based ensemble model.

### (1) Data Preprocessing and Partitioning

The collected traffic speed data requires preprocessing steps, such as normalization, to facilitate more effective training of the base models. To evaluate the performance of the models, the dataset was partitioned into three segments: 64% as the training set, 16% as the validation set and 20% as the test set. The training set is used to construct base models and validation set helps to optimize model parameters, while the test set is employed to assess the ultimate performance of the proposed ensemble models.

### (2) Application of Bayesian Optimization

This study employed Bayesian optimization algorithm as a sequential model-based optimization technique to enhance the performance of base predictive models. Bayesian optimization algorithm utilizes probability models to tactically explore and exploit the search space for optimal hyper-parameters [22]. Each hyper-parameter is assigned a range of values for exploration. For each iteration of the Bayesian optimization algorithm, a set of hyper-parameters is evaluated from the current probability model. This iterative process persists until convergence is reached or a predetermined number of iterations is exhausted.

### (3) Training and Optimizing the Base Predictive Models

Considering the characteristics of traffic speed data, the promising base models chosen for this study, such as CNN, LSTM, SVR, LR, MLP, and KNN are trained and optimized through Bayesian optimization algorithm.

### (4) Stacking-based Ensemble Modeling Framework

In the first layer of the stacking-based ensemble model, predictions from all optimized base models are concatenated. In the second layer, a meta-regressor is employed to integrate all the prediction results from the first layer. Subsequently, the meta-regressor is trained using predictions from the optimized base models as inputs and the actual outputs from the training data as targets, followed by final evaluation on the test set. Finally, the predicted results are compared with the actual traffic speeds, and the model's error is computed using specific evaluation metrics, such as Mean Squared Error (MSE), etc.

The stacking-based ensemble method integrates optimized base models into a unified ensemble model, enhancing the model's generalization capability and contributing to the alleviation of overfitting.

## 4. Experiments

This section delineates the statistical measures used for evaluating the base models and ensemble learning models. All models and methods were implemented using the Python programming language.

### 4.1. Evaluation Metrics

To assess the performance of the proposed stacking-based ensemble model and its baseline models, several statistical measures were employed.

These metrics collectively represent the predictive capability of the models, capturing various aspects of performance, and are calculated using the following equations. In the following equations,  $Y_i$  represents the actual observation of the  $i$ -th sample point and  $\hat{Y}_i$  represents the prediction to the  $i$ -th sample point.

#### (1) Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) represents the mean distance between predicted and actual values, capturing the average level of error in the model's predictions. It is calculated in Equation (1).

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i| \quad (1)$$

#### (2) Root Mean Square Error (RMSE)

The Mean Absolute Error (MAE) represents the mean distance between predicted and actual values, capturing the average level of error in the model's predictions. It is calculated in Equation (2).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2} \quad (2)$$

#### (3) Symmetric Mean Absolute Percentage Error (SMAPE)

SMAPE evaluates model performance by calculating the mean of the absolute percentage errors between predictions and actual values. A small SMAPE means the enhanced predictive accuracy, as illustrated in Equation (3).

$$SMAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|\hat{Y}_i - Y_i|}{(|\hat{Y}_i| + |Y_i|) / 2} \quad (3)$$

### 4.2. Experimental Results and Discussion

Using the above evaluation metrics, Table 2 illustrates in detail the comparative results on the test set for ten machine learning models, including six individual models (i.e., CNN, LSTM, MLP, LR, KNN, and SVR) and four stacking-based ensemble models with different meta-regressor (i.e., Stacking+LSVR (Linear SVR), Stacking+LR, Stacking+SVR, and Stacking+Ridge). The bold font indicates the significant values.

**Table 2.** The comparative results among various models.

Metrics	Dataset A (PeMSD4)			Dataset B (PeMSD7)		
	MAE	RMSE	SMAPE	MAE	RMSE	SMAPE
LR	6.0079	8.7618	11.9922	1.6330	2.5808	2.6036
KNN	6.1184	9.5707	12.4354	1.1624	2.2065	1.8961
SVR	5.7126	8.9567	11.4722	1.4573	2.4959	2.3436

Metrics	Dataset A (PeMSD4)			Dataset B (PeMSD7)		
	MAE	RMSE	SMAPE	MAE	RMSE	SMAPE
MLP	5.3370	8.1561	11.0128	1.5872	2.5538	2.5414
CNN	5.6086	8.3330	11.4945	1.6029	2.6405	2.5821
LSTM	5.9026	8.7110	11.8369	1.5815	2.5648	2.5323
Stacking+LSVR	5.3774	8.4995	10.9033	1.1589	2.1777	1.8872
Stacking+LR	5.1679	8.1042	10.5787	1.2015	2.1483	1.9437
Stacking+SVR	5.2403	8.0542	10.6135	1.4731	2.1195	2.3345
Stacking+Ridge	5.2185	7.9851	10.7259	1.2865	2.2105	2.0684

Experimental results demonstrate the following average rankings to various models on different evaluation metrics:

For MAE, the models rank in the order: Stacking+LR, Stacking+Ridge, Stacking+LSVR, and Stacking+SVR.

For RMSE, the models rank in the order: Stacking+Ridge, Stacking+SVR, Stacking+LR, and Stacking+LSVR.

For SMAPE, the models rank in the order: Stacking+LR, Stacking+LSVR, Stacking+Ridge, and Stacking+SVR.

The above comparative results show that the ensemble models outperform the individual models, and LR as meta-regressor performs the best in all the stacking-based ensemble models. Therefore, LR was finally adopted as a meta-regressor to merge the predictive factors of the optimized base models into a stacking-based ensemble model.

## 5. Conclusion

Accurate traffic speed prediction can provide intelligent support for the efficient operation of modern urban transportation systems. This study proposes a stacking-based ensemble model optimized by Bayesian optimization algorithm. Experimental validation on various traffic speed data sets through multiple evaluation metrics has demonstrated the proposed model's superiority over the base models in terms of predictive performance for traffic speed.

However, the current study also highlights the areas that require further investigation and refinement. Firstly, the integration of other advanced ensemble methods, such as ensemble boosting and bagging, may be explored to further improve the model performance. Secondly, the incorporation of other hyper-parameter optimization techniques could be considered to further enhance the model's robustness and adaptability. Moreover, the integration of some population-based optimization algorithms, such as genetic algorithms, could be considered to further enhance the model's generalization capabilities.

## Author Contribution

Study conception and design: Yuanzhe Cheng, Haoyang Lv, Hanrui Chen.

Data collection: Chengjie Ni, Hanrui Chen.

Analysis and interpretation of results: Yuanzhe Cheng, Haoyang Lv, Yuyang Hu.

Draft manuscript preparation: Yuanzhe Cheng, Hanrui Chen.

All authors reviewed the results and approved the final version of the manuscript.

## References

- [1] Wang, J. Y., Gu, Q., Wu, J. J., Liu, G. N. and Xiong, Z. (2016). Traffic speed prediction and congestion source exploration: A deep learning method. In *Proceedings of 2016 IEEE 16th International Conference on Data Mining (ICDM)*, Barcelona, Spain, December 12-15, pp. 499-508.
- [2] Ding, Q. Y., Wang, X. F., Zhang, X. Y. and Sun, Z. Q. (2011). Forecasting traffic volume with space-time ARIMA model. *Advanced Materials Research*, 156, 979-983.
- [3] Jia, Y. H., Wu, J. P. and Du, Y. M. (2016). Traffic speed prediction using deep learning method. In *Proceedings of 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Rio de Janeiro, Brazil, November 1-4, pp. 1217-1222.
- [4] Sheng, D. M., Zheng, X. X. and Yi, X. T. (2019). Intelligent traffic control system based on cloud computing and big data mining. *IEEE Transactions on Industrial Informatics*, 15 (12), 6583-6592.
- [5] Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J. and Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society*, 54(6), 627-635.
- [6] Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y. J. and Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), 559-569.
- [7] Bui, D. T., Tuan, T. A., Klempe, H., Pradhan, B. and Revhaug, I. (2016). Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides*, 13 (2), 361-378.
- [8] Russ, G., Kruse, R., Schneider, M. and Wagner, P. (2008). Data mining with neural networks for wheat yield prediction. In *Proceedings of the 8th Industrial Conference on Advances in Data Mining: Medical Applications, E-Commerce, Marketing, and Theoretical Aspects*, Leipzig, Germany, July 16-18, pp. 47-56.
- [9] Komi, M., Li, J., Zhai, Y. and Zhang, X. G. (2017). Application of data mining methods in diabetes prediction. In *Proceedings of 2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, Chengdu, China, June 2-4, pp. 1006-1010.
- [10] Wu, C. H., Kuo, S. W. and Kao, S. C. (2019). Classification-based data mining applied in vehicle accident prediction. *Fuzzy Systems and Data Mining*, 320, 218-223.

- [11] Huang, X. H., Ye, Y. M., Yang, X. F. and Xiong, L. Y. (2023). Multi-view dynamic graph convolution neural network for traffic flow prediction. *Expert Systems with Applications*, 222, 119779.
- [12] He, H. L., Zhang, W. Y. and Zhang, S. (2018). A novel ensemble method for credit scoring: Adaption of different imbalance ratios. *Expert Systems with Applications*, 98, 105-117.
- [13] Cui, S. Z., Yin, Y. Q., Wang, D. J., Li, Z. W. and Wang, Y. Z. (2021). A stacking-based ensemble learning method for earthquake casualty prediction. *Applied Soft Computing*, 101, 107038.
- [14] Khoei, T. T., Labuhn, M. C., Caleb, T. D., Hu, W. C. and Kaabouch, N. (2021). A stacking-based ensemble learning model with Genetic Algorithm for detecting early stages of Alzheimer's disease. In *Proceedings of 2021 IEEE International Conference on Electro/Information Technology*, Mount Pleasant, USA, May 14-15, pp. 215-222.
- [15] Wang, Y. Y., Wang, D. J., Geng, N., Wang, Y. Z., Yin, Y. Q. and Jin, Y. C. (2019). Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection. *Applied Soft Computing*, 77, 188-204.
- [16] Cai, P. L., Wang, Y. P., Lu, G. Q., Chen, P., Ding, C. and Sun, J. P. (2016). A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. *Transportation Research Part C: Emerging Technologies*, 62, 21-34.
- [17] Castro-Neto, M., Jeong, Y. S., Jeong, M. K. and Han, L. (2009). Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert Systems with Applications*, 36 (3), 6164-6173.
- [18] Hamed, M. M., Almasaeid, H. R. and Said, Z. M. B. (1995). Short-term prediction of traffic volume in urban arterials. *Journal of Transportation Engineering-asce*, 121 (3), 249-254.
- [19] Wei, Y., Chen, M. C. (2012). Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. *Transportation Research Part C: Emerging Technologies*, 21 (1), 148-162.
- [20] Wang, M., Wu, L. B., Li, M., Wu, D., Shi, X. C. and Ma, C. (2022). Meta-learning based spatial-temporal graph attention network for traffic signal control. *Knowledge-based Systems*, 250, 109166.
- [21] Ma, X. L., Tao, Z. M., Wang, Y. H., Yu, H. Y. and Wang, Y. P. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, 54, 187-197.
- [22] Du, L., Gao, R., Suganthan, P. N. and Wang, D. Z. W. (2022). Bayesian optimization based dynamic ensemble for time series forecasting. *Information Sciences*, 591, 155-175.