Proceedings of the 3rd International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/135/2025.21214

Traffic Flow Prediction Using Stacked Ensemble Models for Intelligent Traffic Management

Shouming Li[1,a], Hanfeng Xiong[2,b,*]

[1]*College of Computer Science, Beijing University of Technology, Beijing, China*
[2]*School of Power and Mechanical Engineering, Wuhan University, Wuhan, Hubei, China*
a. 2314791397@qq.com, b. 2022302191667@whu.edu.cn
**corresponding author*

Abstract: As urbanization accelerates, the importance of traffic prediction in optimizing intelligent traffic management systems becomes increasingly evident. However, the process of accurate traffic prediction faces challenges such as the complexity of traffic data, the dynamic nature of traffic patterns, and the inconsistency of model performance across different conditions. To address these challenges, this paper proposes an innovative prediction method based on ensemble models. Specifically, the approach begins by creating a voting classifier that filters out highly accurate models, ensuring that only the best-performing models are selected for further use. Following this, a stacked ensemble model is constructed, which combines the predictions of these high-accuracy models to improve the overall prediction performance. The experimental results, evaluated based on prediction accuracy, show that this method significantly outperforms traditional approaches regarding precision and reliability. This research is highly significant, as it not only enhances the accuracy of traffic flow prediction but also contributes to the optimization of traffic management systems. Additionally, it promotes further exploration and development of more robust and efficient prediction models, potentially benefiting other fields that rely on predictive analytics.

Keywords: Ensemble models, Traffic prediction, Stacking, Accuracy Rate.

1. Introduction

In the world today, to build sustainable cities and improve the living quality of the people, pushing forward the construction of the Smart City is of great importance [1, 2]. Among these, the Intelligent Transportation System is one of the important sectors [3]. Researchers in the field have made a lot of groundbreaking achievements. Among the nonparametric methods, the most well-known ones are machine learning (ML) based methods [4]. This paper adopts the ensemble model to predict the traffic flows.

The traditional traffic prediction methods which are based on statistics do not have many parameters, most of them are designed by experts rather than data mining [5]. Dai combined the spatial-temporal analysis and gated recurrent units to predict short-term traffic flows, the model shows higher accuracy and stability than that of the convolutional neural network [6]. In recent years, a graph multi-attention network (GMAN) has been brought forward, by modeling the relationship between different historical and future time steps to alleviate the spread of mistakes, and to predict

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/135/2025.21214

the traffic situation in different time steps [7]. Long Short-term Graph Convolutional Networks (LSGCN) can not only predict the traffic situation in different time steps but also capture the complex spatial-temporal features [8]. Li proposed a Dynamic Graph Convolutional Recurrent Network (DGCRN), the model combines the dynamic graph with the pre-defined graph and skip connection, and possesses the ability to describe the dynamic characteristics of road networks [5]. The Spatial-Temporal Graph Synchronous Aggregation Model (STGSA) can extract spatial-temporal relationships in the process of graph convolution, the method improves the extracting efficiency and breaks the limitation of the inability of interaction between nodes [9]. A novel Spatio-Temporal Self-Supervised Learning (ST-SSL) traffic prediction framework performs adaptive enhancement on flow diagrams, promoting the perception of spatial-temporal heterogeneity [10].

The accuracy rate of the single models still has a lot of room for improvement while the data set is insufficient. The key method the study uses is using the VotingClassifier to select basic models and meta-models which have relatively high accuracy. With the basic models and the meta-models, the StackingClassifier can be constructed.

The paper contains four major parts. The first chapter Introduction, includes the research background, the newly emerging major prediction models, and the research methods of this paper. The second chapter Data and Method, introduces the data set, the basic models, and the ensemble model used in this study. The third chapter Result and Discussion, concentrates on analyzing the experiment results and concluding the advantages and disadvantages of the ensemble model. The fourth chapter the Conclusion, generalizes this paper and points out the limitations of the research as well, providing directions for future studies.

2. Data and Method

2.1. Data Collection and Description

The data set used in this study was obtained from the Kaggle website, shown as Table 1. The dataset contains 2960 sets of data. Each count node respectively gauges the numbers of cars, bikes, buses, trucks, and the total amounts. There are nine columns: Time, Date, Day of the week, CarCount, BikeCount, BusCount, TruckCount, Total, and Traffic Situation. The data set updates every fifteen minutes, and it covers the traffic records over two months. The researchers train the ensemble model with part of this data set, enabling it to make accurate predictions of future traffic situations.

Table 1: Part of the Dataset.

	Time	Date	Day of the week	CarCount	BikeCount	BusCount	TruckCount	Total	Traffic Situation
0	12:00 AM	10	Tuesday	31	0	4	4	39	low
1	12:15 AM	10	Tuesday	49	0	3	3	55	low
2	12:30 AM	10	Tuesday	46	0	3	6	55	low
3	12:45 AM	10	Tuesday	51	0	2	5	58	low
4	1:00 AM	10	Tuesday	57	6	15	16	94	normal

DOI: 10.54254/2755-2721/135/2025.21214

2.2. Data Pre-processing

The data set contains three columns of text data: Time, Day of the week, and Traffic Situation. Firstly, the researchers transformed the Day of the week into numbers 1 to 7 in order. Then conduct transformation on the Traffic Situation as follows: Low-0, Normal-1.High-2, Heavy-3. Lastly, the study introduces LabelEncoder to encode the Time as consecutive integers.

2.3. Basic Models

AdaBoost (ADA), XGBClassifier (XGB), GradientBoostingClassifier (GBC), and LGBMClassifier (LGBM) constitute a strong classifier by training multiple weak classifiers iteratively, ADA assigns them different weights. RandonForestClassifier (RF) integrates several weak decision trees and takes the mode or the average of which as the final result. The difference between RF and ExtraTreesClassifier (ET) is that ET selects features completely random on every node for dividing. CatBoostClassifier (CAT) is an algorithm based on gradient gradient-boosting decision tree, it is mainly used for classification problems. KneighborsClassifier (KN) finds K data points closest to the input point and then classifies the input point according to the classification of these K points. Logistic regression (LG) is a general linear regression model, always being used for binary classification. RidgeClassifier (RC) introduces L2 regularization terms to solve the multicollinearity problems that probably exist, generalization ability is improved. HistGradientBoostingCLassifier (HBC) differs from traditional methods of gradient boosting trees, it brings in histograms to accelerate the training and prediction process. BaggingClassifier (BC) is an integrated classification model, it builds independent models on random subsets of the data set. GaussianNB (GNB) supposes the features are mutually independent and obey Gaussian distribution. DummyClassifier (DC) makes predictions following the pre-defined simple rules and takes the prediction as the benchmark to assess the performance of other models. SupportVectorClassifier (SVC) constructs the best classification boundaries to maximize the gap between different categories.

2.4. Stacking Model

The StackingClassifier this study built is inherited from BaseEstimator, ClassifierMixin, and TransformerMixin. BaseEstimator is the base class of all kinds of estimators. It defines the basic interface and other measures implemented in the subclasses. ClassifierMixin and TransformerMixin are abstract classes, used separately for classification and data conversion. The outputs of the basic models serve as the inputs of the meta-model. Specifically, the StackingClassifier takes the average prediction result of each basic model as meta-features, which the meta-model predicts. In this way, the prediction accuracy of the model is improved.

3. Result and Discussion

3.1. Voting Classifier

This experiment selected 15 types of classifiers as voters, performed 5-fold cross-validation on each classifier, calculated their accuracy on the training set, and then obtained the average accuracy and standard deviation, as shown in Table 2.

Table 2: Accuracy of Classifiers on the Training Set

Model	Accuracy Mean	Accuracy Std
ADA	0.7920	0.0931
XGB	0.9988	0.0011

DOI: 10.54254/2755-2721/135/2025.21214

Table 2: (continued).

GBC	0.9991	0.0012
LGBM	0.9988	0.0017
RF	0.9958	0.0039
ET	0.9820	0.0019
CAT	0.9994	0.0007
KN	0.9400	0.0093
LG	0.8935	0.0055
RC	0.7686	0.0119
HBC	0.9985	0.0019
BC	0.9985	0.0013
GNB	0.8229	0.0100
DC	0.6089	0.0009
SVC	0.9502	0.0056

The following eight classifiers have higher accuracy on the training set: XGB, RF, ET, CAT, HBC, BC, GBC, and LGBM, with CAT having the highest accuracy. In the next step, these classifiers were selected as base models and meta-models.

3.2. Stacking Classifier

The stacking classifier adopts a two-tier architecture, as shown in Figure 1. In the first tier, multiple base models are constructed, each of which is trained through k-fold cross-validation and generates prediction results on the validation set. The second tier is a meta-model, which uses the prediction results from each base model in the first tier as features for training. When employing this stacking classifier for traffic flow prediction, each base model will output a predicted value for traffic flow. These predictions are then averaged to form a new feature set. Finally, the meta-model will make predictions based on this new feature set and output the final prediction results.

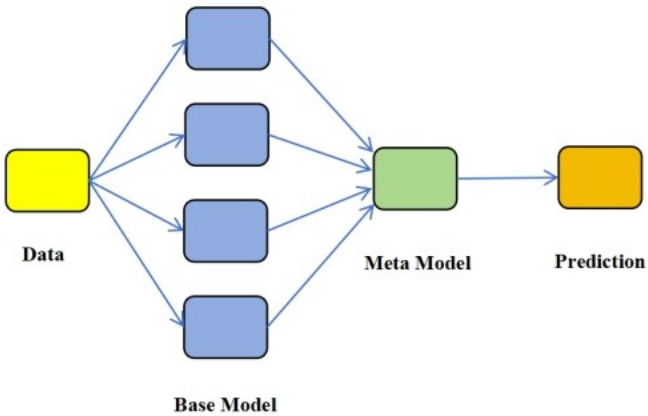

Figure 1: Structure of the Stacking Classifier. (Picture credit: Original).

In the previous step, among the eight models with higher accuracy rates, CAT was chosen as the meta-model, while XGB, RF, ET, CAT, HBC, BC, GBC, and LGBM were selected as the base models to construct the ensemble model. The experimental results of this study are visualized in Figure 2. The proposed ensemble model demonstrated superior performance on the test set, achieving an accuracy rate of 99.944%.

DOI: 10.54254/2755-2721/135/2025.21214

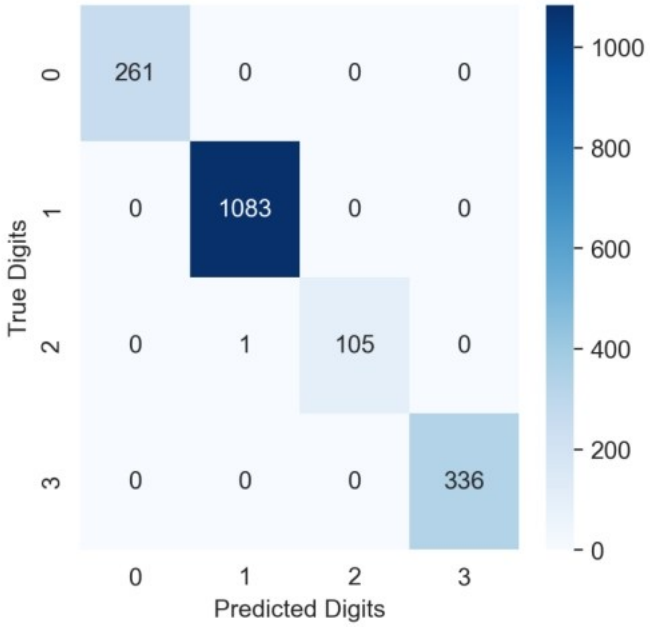

Figure 2: Confusion Matrix of the Integrated Model's Stacking Classifier. (Picture credit: Original).

Multiple experiments were conducted to compare the stacking classifier's performance with that of the better-performing base models, and the average accuracy rates were calculated, as shown in Table 3.

Table 3: Prediction Accuracy of the Models

Model	Accuracy Mean (%)
Stacking	99.535
XGB	98.018
RF	99.507
ET	97.844
CAT	99.608
HGB	99.602
BC	99.490
GBC	99.574
LGBM	99.580

In this experiment, the performance of the stacking classifier was satisfactory, with its accuracy rate on par with that of the best-performing base model. However, no significant performance improvement was observed in the current dataset. This phenomenon may be attributed to the dataset's lower complexity, which failed to fully reveal the potential advantages of the stacking classifier over single models.

In light of this, this paper believes that in future research, it is necessary to further optimize the model combination and explore more effective integration strategies to enhance the model's generalization capability and prediction accuracy. At the same time, to more comprehensively evaluate the effectiveness of the ensemble model, it is recommended to test on datasets that are more challenging and complex, to better reveal the performance improvement that ensemble learning technology can bring when dealing with complex issues.

DOI: 10.54254/2755-2721/135/2025.21214

3.3. Discussion

This experiment verified the following disadvantages and advantages of the stacking classifier.

In terms of disadvantages, the stacking classifier has several issues. Firstly, its computational cost is significantly higher than that of a single model. The stacking classifier involves the training of multiple base models and the construction of a meta-model, which undoubtedly leads to a significant increase in training time and places higher demands on hardware resources. Secondly, there is a certain complexity in the selection of base models and the tuning of parameters for the stacking classifier. This process requires researchers to invest a considerable amount of time and effort in repeated experimentation and optimization.

In terms of advantages, the stacking classifier exhibits the following characteristics. Firstly, while ensuring performance, the stacking classifier has a more robust performance compared to a single model. Thanks to the presence of a voting mechanism, even if there are performance differences among the base classifiers, the ensemble model can still maintain a high prediction accuracy rate. Secondly, the ensemble model can reduce the risk of overfitting. For example, some base models such as XGB may perform well on the training set but poorly on the test set, indicating that the model may have a certain degree of overfitting. By integrating multiple base models, the ensemble model reduces the negative impact of individual models.

4. Conclusions

Traffic flow prediction plays a significant role in intelligent transportation applications such as traffic management, congestion mitigation, and public risk assessment. To construct a reliable traffic flow prediction system, this paper investigates the use of ensemble models for predicting traffic conditions. This paper employs a voting mechanism to select the base models and then constructs an ensemble model, achieving better prediction results. Compared to the base models, the ensemble model exhibits more stable performance and a lower risk of overfitting. However, it also faces the challenges of increased computational costs and more complex parameter tuning. Actual traffic prediction problems are often more complex, influenced by external factors such as weather and holidays, and challenged by incomplete data and high requirements for real-time performance. These are all directions that future research needs to explore. In conclusion, while ensemble models improve prediction accuracy and stability, challenges like computational complexity, external factors, and real-time performance need further exploration. Future research should focus on optimizing these aspects for more practical and efficient traffic flow prediction systems.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] *Firmansyah, H., S., & Supangkat, S., H. (2022). Smart World Living Lab: A Living Lab Approach to Improve Smart City Implementation (Introducing the DDG areas as an integrated Living Lab). IEEE International Smart Cities Conference (ISC2), pp. 1-6.*

[2] *Yan, H. (2023). A Novel Cloud-IPSO-DBN Algorithm for Evaluating Digital Smart City Construction. In IEEE Systems Journal (Vol. 17, no. 4, pp. 5109-5117).*

[3] *Chen, Y., Guo, J., Xu, H., Huang, J., & Su, L. (2023). Improved Long Short-Term Memory-Based Periodic Traffic Volume Prediction Method. In IEEE Access (Vol. 11, pp. 103502-103510).*

[4] *Boukerche, A., & Wang, J. (2020). Machine learning-based traffic prediction models for intelligent transportation systems. Computer Networks, 181, 107530.*

DOI: 10.54254/2755-2721/135/2025.21214

[5] *Li, F., Feng, J., Yan, H., Jin, G., Yang, F., Sun, F., ... & Li, Y. (2023). Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution. ACM Transactions on Knowledge Discovery from Data, 17(1), 1-21.*

[6] *Dai, G., & Xu, X. (2019). Short-Term Traffic Flow Prediction Method for Urban Road Sections Based on Space-Time Analysis and GRU. In IEEE Access (Vol. 7, pp. 143025-143035).*

[7] *Zheng, C., Fan, X., Wang, C., & Qi, J. (2020, April). Gman: A graph multi-attention network for traffic prediction. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 01, pp. 1234-1241).*

[8] *Huang, R., Huang, C., Liu, Y., Dai, G., & Kong, W. (2020, July). LSGCN: Long short-term traffic prediction with graph convolutional networks. In IJCAI (Vol. 7, pp. 2355-2361).*

[9] *Wei, Z., Zhao, H., Li, Z., Bu, X., Chen, Y., Zhang, X., ... & Wang, F. Y. (2023). STGSA: A novel spatial-temporal graph synchronous aggregation model for traffic prediction. IEEE/CAA Journal of Automatica Sinica, 10(1), 226-238.*

[10] *Ji, J., Wang, J., Huang, C., Wu, J., Xu, B., Wu, Z., ... & Zheng, Y. (2023, June). Spatio-temporal self-supervised learning for traffic flow prediction. In Proceedings of the AAAI conference on artificial intelligence (Vol. 37, No. 4, pp. 4356-4364).*