



Traffic Flow Prediction On Road Using Machine Learning

Anuja Phapale^{1*}, Sushant Shravagi²

AISSMS Institute Of Information Technology, Pune, Maharashtra, India.

Corresponding Author: Sushant Shravagi sushantshravagi@gmail.com

ARTICLE INFO

Keywords: Traffic Prognostication, Support Vector Machines (SVM), Artificial Neural Networks (ANN), Predictive Analytics, Traffic Congestion.

Received : 20, November

Revised : 23, December

Accepted: 25, January

©2024 Phapale, Shravagi:
This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/).



ABSTRACT

The Intelligent Transportation System (ITS) plays a vital role in numerous smart city applications, particularly in improving transportation and commuting processes. A primary goal of ITS is to tackle traffic-related challenges, especially the issue of traffic congestion. The prediction system for road traffic flow has significant relevance in urban transportation and area management. Many urban centers grapple with the daunting task of effective traffic management. However, the incorporation of predictive modeling that considers environmental and weather conditions, such as rainfall and thunderstorms, has proven to be remarkably effective.

In response to this challenge, we have introduced a road traffic flow prediction model specifically designed to forecast traffic conditions at hourly intervals extending up to 24 hours. Although various algorithms have been applied in previous research, there is a notable absence of accessible and user-friendly platforms dedicated to road traffic flow prediction.

INTRODUCTION

Traffic flow prediction has been a topic of extensive exploration utilizing diverse machine learning approaches. This paper offers a thorough overview of different researchers and the methodologies they have employed for predicting traffic flow. Various algorithms produce unique outcomes when applied to datasets gathered from various sources. Despite the introduction of various applications like vehicle navigation systems, congestion management, and vehicle routing to aid the public in real-time route planning, obtaining precise real-time traffic data remains a formidable challenge.

The primary obstacles to obtaining this data result from the scarcity of well-equipped traffic sensors and external factors such as accidents, public events, and adverse weather conditions. Generally, traffic flow prediction falls into two primary categories: Short-Term and Long-Term Traffic Flow Prediction. Long-term algorithms may struggle to provide highly accurate predictions, as they typically operate on an hourly basis, such as 12 or 24-hour intervals. In contrast, short-term approaches tend to yield more accurate results, providing predictions in minutes, such as 5 to 15 minutes or 30 to 50 minutes, making them more reliable for predictive purposes.

Our predictive model has been fine-tuned within a maximum one-hour time frame to deliver prediction results. In developing this model, we conducted a thorough review of research contributions in the field of machine learning, as outlined in the literature review section. Researchers have devoted substantial effort to leveraging various machine learning algorithms to enhance traffic prediction in light of these challenges.

LITERATURE REVIEW

A literature review on traffic flow prediction on roads using machine learning would typically involve an extensive exploration of research papers, articles, and studies in this field. Below, I provide a structured outline with a list of topics and potential sources you can consider for your literature review on traffic flow prediction:

1. Define the importance of traffic flow prediction in urban planning and transportation management.
2. Discuss various machine learning algorithms used in traffic flow prediction.
3. Differentiate between short-term and long-term traffic flow prediction and their respective methodologies.
4. Explore techniques for selecting relevant features and preprocessing traffic data.
5. Discuss the challenges and limitations in traffic flow prediction, such as data collection, environmental factors, and road conditions.
6. Real-Time Data Sources: Examine sources of real-time data for traffic flow prediction.

7. Explore the application of Support Vector Machines (SVM) in traffic flow prediction.
8. Investigate the role of deep learning, neural networks, and recurrent neural networks in traffic flow prediction.
9. Wavelet Neural Networks and Dynamic Models: Explore dynamic models, such as wavelet neural networks, for traffic flow forecasting.
10. Nonlinear Dynamics and Temporal Evolution: Discuss the temporal evolution of short-term urban traffic flow from a nonlinear dynamics perspective.

This structured outline will help you organize your literature review on traffic flow prediction using machine learning techniques. You can delve into each topic in more detail and cite relevant studies and sources to build a comprehensive review of the field.

METHODOLOGY

1. Problem Formulation

Define the specific problem you want to address, such as short-term traffic flow prediction in a particular location and time interval.

2. Data Collection

Gather historical traffic data, which may include information on traffic volume, speed, road conditions, weather, and any other relevant factors. Utilize sources like traffic sensors, GPS data, publicly available traffic databases, or open data platforms.

3. Data Preprocessing

Handle missing data by imputation or removal. Detect and deal with outliers and anomalies. Normalize or standardize numerical features to ensure uniform scales. Engineer relevant features, such as time of day, day of the week, holidays, and special events.

4. Data Splitting

Divide the dataset into three subsets: training, validation, and test data. Common splits include 70% for training, 15% for validation, and 15% for testing.

5. Model Selection

Choose appropriate machine learning models for traffic flow prediction, between decision trees or random forests

6. Experiment Hyperparameters

Experiment with different hyperparameters and settings to optimize model performance. Use the validation dataset to fine-tune hyperparameters.

7. Model Evaluation

Evaluate model performance using relevant metrics, such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), or others. Compare the results of different models to select the best-performing one.

8. Hyperparameter Tuning

Fine-tune the model's hyperparameters to improve prediction accuracy. You can use techniques like grid search or random search.

9. Model Validation

Validate the chosen model using the test dataset to ensure it generalizes well to unseen data.

10. Deployment

Deploy the trained model in a real-time traffic prediction system if needed.

11. Model

Update the model with new data as it becomes available to maintain accuracy.

12. Documentation

Document the entire methodology, including data sources, preprocessing steps, model architectures, hyperparameters, and results.

13. Reporting and Visualization

Present the results and insights using visualizations, reports, and dashboards for stakeholders.

14. Ethical Considerations

Ensure that the model respects ethical and legal considerations, such as privacy, fairness, and transparency.

15. Future Improvements

Identify areas for improvement and future research to enhance the accuracy of traffic flow prediction.

This methodology provides a structured approach to developing a traffic flow prediction system using machine learning. Adapt and refine the steps according to your specific project requirements and data sources.

RESEARCH RESULT

1. **Historical Traffic Data:** This category encompasses past records of traffic flow, speed, and congestion for specific road segments. Such historical data often serves as the foundation for training and validating machine learning models.
2. **Real-Time Traffic Data:** Live traffic data is sourced from various real-time providers, including GPS devices, traffic sensors, and traffic cameras. These sources offer up-to-the-minute information about current road conditions, which is vital for accurate predictions.
3. **Weather Data:** Weather conditions have a substantial impact on traffic flow. Data related to factors like temperature, precipitation, wind speed, and visibility are often integrated into prediction models to account for weather-related disruptions.
4. **Event Data:** Information about special events, road closures, construction activities, accidents, and other events that can affect traffic patterns is crucial for understanding and explaining sudden traffic changes.
5. **Geographic Data:** Geographic Information System (GIS) data provides details about road networks, intersections, and geographic features that influence traffic flow. This type of data aids in creating accurate models.
6. **Time and Date Information:** Time and date data play a pivotal role in modeling daily and weekly traffic patterns. This includes factors such as rush hours and variations related to weekends and holidays.
7. **Vehicle Count and Type:** Some prediction models consider the number and types of vehicles on the road, as these factors can significantly impact traffic flow and congestion.
8. **Public Transportation Data:** Data on public transportation schedules, routes, and disruptions can have a cascading effect on road traffic. Researchers may integrate this data into their models to provide a more comprehensive understanding of the overall transportation system.
9. **Social Media Data:** Data from social media platforms can provide real-time insights into traffic conditions and user experiences. This unstructured data can be leveraged for additional context.

10. Sensor Data: In some cases, sensors embedded in the road or traffic lights can provide data on vehicle presence, speed, and traffic density. These sensors contribute to real-time data collection.

Access to these data sources often relies on collaborations with government agencies, private companies, or open data platforms. When seeking specific datasets for your research, consider exploring the following potential sources:

- 1. Open Data Platforms:** Many cities and regions maintain open data portals where you can access a variety of traffic-related datasets.
- 2. Transportation Research Institutions:** Organizations such as the U.S. Federal Highway Administration (FHWA) and the European Space Agency (ESA) often provide datasets related to transportation.
- 3. APIs:** Some platforms and services offer Application Programming Interfaces (APIs) that provide access to real-time traffic data, weather data, or other relevant information.
- 4. Collaborations:** Collaborating with local transportation authorities or universities can provide access to proprietary data sources, which may be valuable for your research.

When working with these datasets for research, it's crucial to consider data quality, preprocessing steps, and privacy concerns. Ensuring proper data cleaning and feature engineering is essential for developing accurate traffic flow prediction models.

CONCLUSION

In this system, we have successfully created a traffic flow prediction system through the implementation of a traffic flow prediction algorithm. We have specifically integrated two well-established prediction algorithms, namely ANN and SVM, with the aim of providing highly accurate prediction results within our system.

The advantages offered by this system to the general public are significant. Users can access real-time information about current traffic flow conditions and anticipate traffic conditions one hour in advance. Furthermore, the system provides up-to-date weather information for road conditions, enabling users to make informed travel decisions. Additionally, the system offers insights into accident records, aiding users in planning safer routes for their future journeys.

Looking forward, we envision further improving this system by introducing features such as traffic congestion prediction and the utilization of various deep learning techniques to consider additional factors impacting traffic management. Users will also benefit from personalized route recommendations based on their preferences and search queries within the system.

ADVANCED RESEARCH

In writing this article the researcher realizes that there are still many shortcomings in terms of language, writing, and form of presentation considering the limited knowledge and abilities of the researchers themselves. Therefore, for the perfection of the article, the researcher expects constructive criticism and suggestions from various parties.

REFERENCES

- B. Yu, W. H. K. Lam, and M. L. Tam, "Bus arrival time prediction at bus stop with multiple routes," *Transportation Research C*, vol. 19, no. 6, pp. 1157-1170, 2011.
- B. Yu, Z. Z. Yang, and B. Z. vol. 10, no. 4, pp. 151-158, 2006.
- B. Z. Yao, C. Y. Yang, J. B. Yao, and J. Sun, "Tunnel surrounding rock displacement prediction using support vector machine," *International Journal of Computational Intelligence Systems*, vol. 3, no. 6, pp. 843-852, 2010.
- Journal of Advanced Transportation*, 2017.
- L. J. Cao and F. E. H. vol. 14, no. 6, pp. 1506-1518, 2003.
- Learning paradigm*, CEUR-WS.org, 2017.
- Mahmassani, National Academy of Sciences: Transportation Research Board 2018
- V. N. no. 5, pp. 988-999, 1999. R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, "High-speed digital-to-RF converter," U.S. Patent 5 668 842, Sept. 16, 1997.

Phapale, Shravagi

Vlahogianni, et all. *Transportation Research Part C: Emerging Technologies*,
Vol. 13, No. 3, 2005, pp. 211–234.

Vlahogianni, et all. *Computer-Aided Civil and Infrastructure Engineering*, Vol.
23, No. 7, 2008, pp. 536–548.