RESEARCH ARTICLE

SAMRIDDHI Volume 15, Issue 3, 2023 Print ISSN: 2229-7111 Online ISSN: 2454-5767

Traffic Prediction using Machine Learning

Vishwasgowda TM*, Chinnaswamy CN

Department of Information Science & Engineering, National Institute of Engineering, Mysuru, Karnataka, India

Abstract

The traffic is one of the major issues faced by people in almost every city. Predicting the traffichas always been very challenging for transportation planning as well as the city manager. The growth of population and usage of vehicles has increased the need for a reliable traffic prediction system. The system should reduce congestion, avoid accidents, and optimize the traffic flow that enhance road safety. The use of machine learning algorithms would be more suitable in handling the traffic and maintaining the flow ofvehicles. The paper focuses on the work related to handling the traffic using machine learning principles such as LSTM and Random Forest (RF) algorithm. The advanced online dataset for traffic forecasts is used. The system is tested to have improved accuracy with the help of feature engineering. The basic dataset had the following fields: Vehicle Id, Time, Date, Junction Id are the primary and significant parameters considered to anticipate the traffic at each zone. With this we have added the day and weekend component to enhance our result.

Keywords: Traffic, Artificial Intelligence, Machine learning, Traffic prediction, LSTM, Random Forest (RF), Smart traffic system.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology (2023); DOI: 10.18090/samriddhi.v15i03.10

Introduction

Our country is heading towards the advancement of urbanization with the increase in the usage of automobiles. Transportation challenges are turning out to be progressively troublesome, also accidents are common, and the traffic environment is degrading.[1] Many elated works are happening to boost the road network's capacity. One of the alternatives that comes to our mind is to add an additional roadway that deviates the traffic flow, i.e., to increase the number of lanes on the road. However, according to an academic study,increasing road capacity will result in worse traffic circumstances.[2, 3] Establishing an effective and accurate transportation system, which can help us better plan transportation resources, disperse traffic flow before it becomes overcrowded, and even provide more abundant on-road entertainment, is an efficient way to improve the traffic environment. The source of the traffic prediction could be taken from sensors adopted in roads, vehicle count, traffic cameras, parking systems and many more.

As part of Artificial Intelligence (AI), Machine Learning (ML) is one of the most important and popular growing fields these days.[4-10] Machine learning has recently emerged as an important and promising study field for transportation engineering, particularly in traffic prediction. Traffic congestion has an impact on the economy, either directly or indirectly. Every day, traffic congestion wastes people's valuable time and increases the expense of fuel.

Because traffic congestion is a big issue for individuals of all socioeconomic backgrounds, there must be small-scale traffic forecasting for people to live their lives without annoyance orstrain. The comfort of road users is essential first and foremost to ensure the country's economicprogress. This is only achievable if the traffic flow is smooth. To deal with this, we require traffic prediction so that we can estimate or predict future traffic to some degree. Figure 1 depicts the advantage of traffic prediction.

The traffic flow prediction with the use of big data could be accurately measured, the traffic information will strongly help the improvement of business in the travel sector and other government agencies. The road user can make better travel decisions, reduce traffic congestion, and contribute to carbon emissions. Traffic flow prediction aims to predicttraffic as soon as feasible for users. Traffic is extremely congested, and people on the road have no control over it. As a result, this study may be useful in predicting traffic. In summary, this paper's structures are divided into five primary components. The sections are Introduction, related work, overview, Methodology, results and discussion, and conclusion with future work.

Corresponding Author: Vishwasgowda TM, Department of Information Science & Engineering, National Institute of Engineering, Mysuru, Karnataka, India, e-mail: Vishwasgowdatm73@gmail.com

How to cite this article: Vishwasgowda, TM., Chinnaswamy CN. (2023). Traffic Prediction using Machine Learning. *SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology,* 15(3), 341-345.

Source of support: Nil

Conflict of interest: None

Objective of this paper includes

- Analyzing the traffic pattern with the time attribute
- Processing the dataset
- Increasing the accuracy and reducing the route mean square.

The major benefits of implementing the traffic prediction system are identified and illustrated in the following Figure 1.

Literature Review

In the last few years, transportation researchers have carried out a lot of research on the occurrence of traffic congestion in road transportation and the prediction of the traffic flow at various road networks.[5] has proposed a method called "scalable" to predict vehicles' traffic congestion in a grid framework. Anwar and co-workers used spectral clustering to supervise traffic congestion.[6] Considering the traffic flow density and different types of roads, Liang and his team developed a novel prediction model capable of estimating the next-time step traffic volume using a single road traffic segment to clarify traffic congestion using traffic flow variables such as the current inflow, outflow, andtraffic volume, etc. [7]

The authors of [8] attempted to handle the traffic control problem using an ML system to deal with traffic difficulties. The authors used the Q-learning RL technique to manage traffic lights and created a simulation environment called Simulation of Urban Mobility (SUMO). SUMO allows you to watch cars in motion, control the vehicle's delay time, and alter the delay time.

[9] developed a model combining artificial neural networks and root mean squared error. Both were used as a metric by applying singular point probabilities. Traffic congestion has become a global pandemic and transportation researchers are racing against time to improve the effectiveness of intelligent transportation systems. Some researchers have been able to achieve good results when it comes to traffic flow prediction. A practical traffic flow parameter prediction model was created for traffic flow conditions estimations. An autoregressive model was combined with other predictive models.[11, 12] The authors of [13] developed a unique traffic prediction approach based on DL with the least prediction error and introduced the LSTM model. The dataset for this study was real-world traffic big data from performance measurement systems (PeMS). The number of optimized parameters used in this study should be increased. In addition, the model training time must be managed.

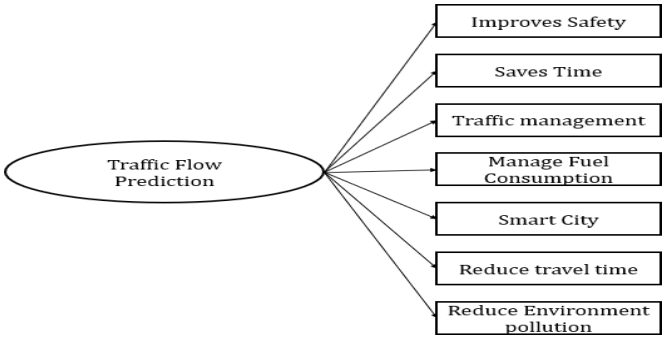

Figure 1: Benefits of implementing the traffic prediction system.

Methodology

- **Dataset:** The collection of necessary data is identified; the following are those datasets that help us in traffic prediction: This project's core dataset was a thorough traffic forecast record hosted on Kaggle.[14] This data set comprises information on 48120 traffic datasets collected between 2015 and 2017. For testing, 11808 records are used. It has five columns that collect information such as the date, time, vehicle ID, and junction ID. We generated an advanced dataset by including two attributes: days and weekends. Weekdays are denoted by the binary number 0. Weekends are defined as Saturday and Sunday, which are represented by the number 1.
- **Attributes of Dataset**: Few specific data attribute that contributes to the prediction are taken from the Kaggle dataset, they are as follows: Date, Time, Vehicle Id, Junction Id. These attributes are listed in our dataset and are the basic attributes used in analysis. We have used the featured engineering techniques to obtain accuracy. This helps add up a few more attributes that contribute to the improvement in the results. The features that are added here are the Day and weekend attribute, the day attribute specifies the day of week and weekend attribute is used to check if is a holiday.

  The traffic pattern is analyzed with these added features, if the day of week are Saturday and Sunday then it is marked as a holiday and is assigned with the value 1 and rest of the days are marked with 0.
- **Proposed System**: The proposed model's fundamental premise has two main tasks. The primary job is to accurately anticipate traffic based on historical and advanced data for the specified location. The planned work architecture is depicted in Figure 2. The augmented feature set was created using feature engineering techniques.

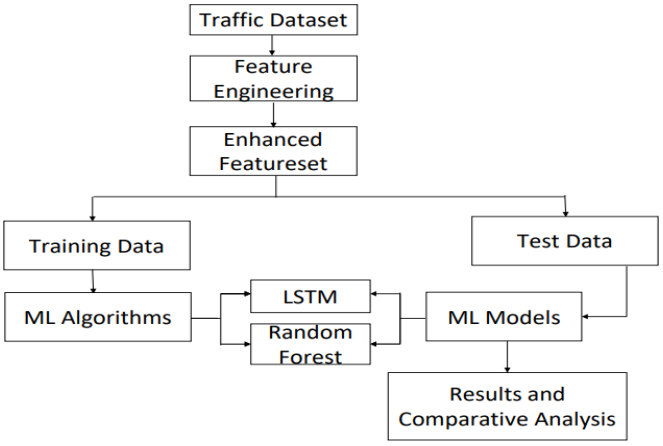

Figure 2: System Architecture

The process of changing raw data into features appropriate for machine learning models is known as feature engineering. It is the process of picking, extracting, and converting the most important features from accessible data in order to develop more accurate and efficient machine learning models.

Machine Learning Methods

Machine learning (ML) allows you to build prediction models that consider vast amounts of heterogeneous data from many sources. Numerous studies on the use of ML algorithms to forecast road traffic have been done.

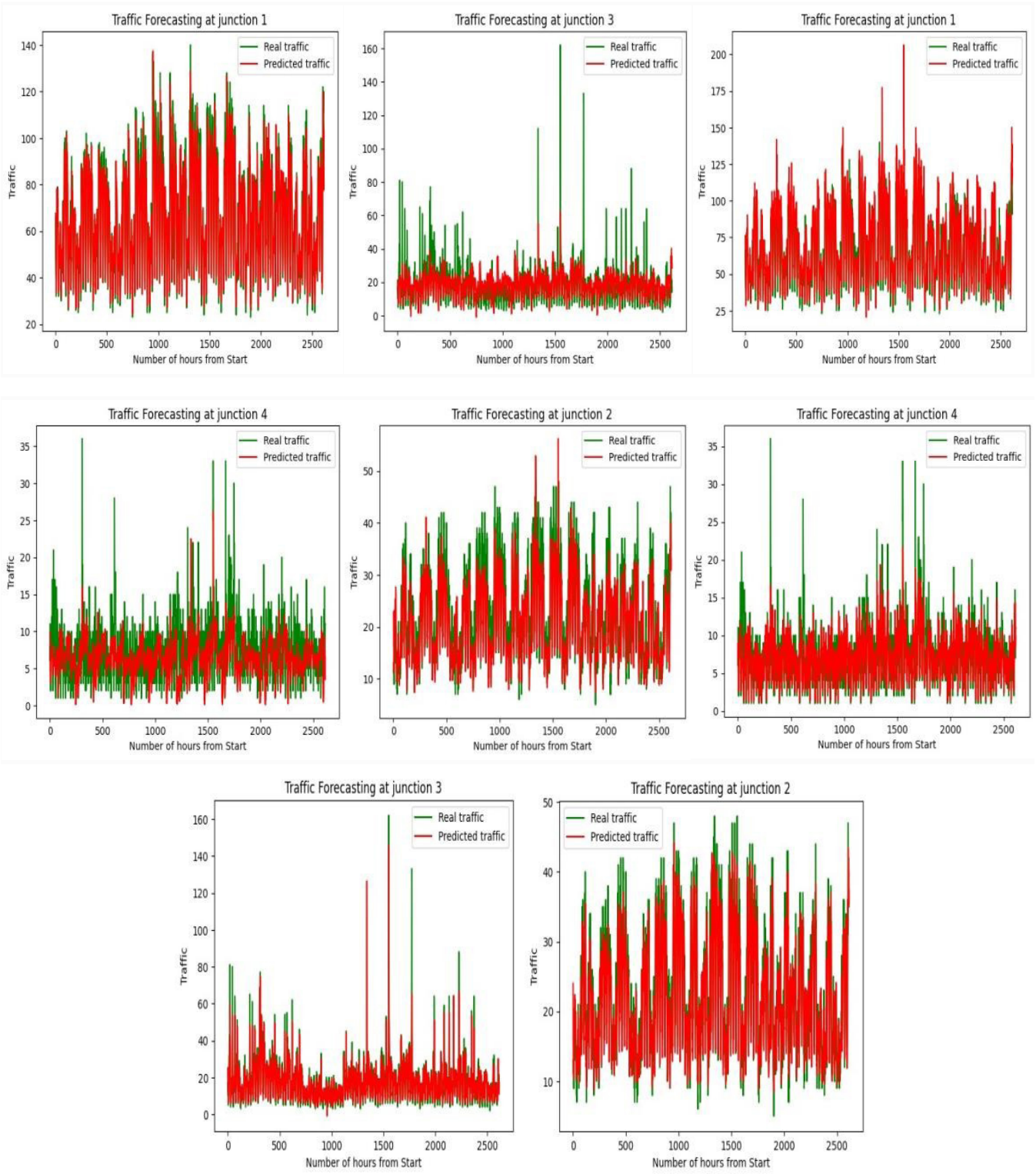

Figure 3: Traffic forecasting at junctions

- **Random Forest**: The random forest method generates numerous decision trees and integrates their data to obtain accurate predictions. Given enough training data, it is quite fast and can yield effective results. When applied to the problem of traffic congestion, this approach has an accuracy of 87.5 percent. Weather conditions, period, specific road conditions, road quality, and vacations are employed as model input variables in this scenario.
- **Long Short-term memory**: This algorithm works on time series data: it considers the data generated in the previous hour and the current hour data. With these data the analysis is performed. LSTM-RNN variants that handle the vanishing gradient problem. According to a study that examined the performance of these models, the GRU model is more accurate in predicting traffic flow and is easier to train.

  A vast variety of paper recommend developing different types of NN models for traffic prediction, such as graph neural networks, fuzzy neural networks, Bayesian neural networks, and others, as well as hybrid techniques that integrate two or more algorithms. No single best technique has been discovered that can be utilized in all instances and produce the most accurate forecasts.

Results and Discussion

Traffic congestion is regarded as one of the most serious concerns confronting urban regions, as computing technologies progress, various advanced concepts are employed to estimate traffic congestion and flow. This study employs two machine learning concepts, LSTM and RF to forecast traffic flow. The traffic data set was obtained from the internet and pre-processed by adding two attributes.

Apply machine learning concepts to forecast traffic flow after pre-processing. The proposed system's performance is assessed using the metrics accuracy, precision, and Root Mean Squared Error (RMSE), as given in Table 1.

Models have been trained for the basic and advanced datasets we developed using feature engineering. We added two additional attributes to the advanced dataset. Both the basic and advanced datasets were subjected to a comparison study. We discovered that the accuracy of the advanced dataset has improved. Figure 3 shows the traffic forecast i.e., real traffic v/s predicted traffic in all the junctions we have taken in our study. Using LSTM, we calculated Root Mean Square error for both the basic and advanced dataset. In the table we can observe error has been reduced for advanced dataset compared to basic dataset. For Random Forest model, the accuracy score has been calculated in this also accuracy has been improved for advanced dataset.

Table 1: Comparison of existing basic features approach and our enhanced features

Models	*Basic Attributes*	*Enhanced Attribute*
LSTM	6.16 RMSE	4.57 RMSE
Random Forest	77.82 Accuracy	94.96 Accuracy

Conclusion

It was determined in the system that we constructed the traffic flow forecast system employing a machine learning algorithm.

The proposed models were trained, validated, and tested using two datasets. One is a publicly available dataset, and we used feature engineering to generate a new dataset. Experiment results demonstrated that improved dataset models outperformed traditional dataset models.

The public gains benefits such as the current flow situation, the flow of traffic on the right after one hour of the situation, and they can also know how the roads are as they can know the mean of the situation, and also know how the roads are as they can know the mean of the vehicles passing through a specific junction that is 4 here. The weather has been shifting over the years. The cost of gasoline also has a significant impact on the transportation system. The high cost of petrol prevents many people from purchasing a vehicle.

As a result, there may be several differences in the traffic data. Another scenario in which people choose to travel in their own vehicle rather than carpooling is when traffic is congested. As a result, by comparing these two years of data, this projection can aid in estimating traffic flow. Forecasting or prediction can assist individuals or users in gauging road traffic more easily in advance, and they can even determine which direction to go using their navigation. This prediction will also be useful.

The proposed model with the use of LSTM and Random Forest, produced results with a reduced square mean error rate and provided a very good accuracy in traffic prediction at the junction.

References

[1] Chen, C.; Liu, B.; Wan, S.; Qiao, P.; Pei, Q. An Edge Traffic Flow Detection Scheme Based on Deep Learning in an Intelligent Transportation System. IEEE Trans. Intell. Transp. Syst. 2021, 22, 1840–1852.

[2] Boukerche, A.; Wang, J. A performance modeling and analysis of a novel vehicular traffic flow prediction system using a hybrid machine learning-based model. Ad. Hoc. Netw. 2020, 106, 102224.

[3] Ferreira, Y.; Frank, L.; Julio, E.; Henrique, F.; Dembogurski, B.; Silva, E. Applying a Multilayer Perceptron for Traffic Flow Prediction to Empower a Smart Ecosystem. In Computational Science and Its Applications; Springer: Cham, Switzerland, 2019; pp. 633– 648.

[4] Gadze, J.D.; Bamfo-Asante, A.A.; Agyemang, J.O.; Nunoo-Mensah, H.; Opare, K.A.B. An Investigation into the Application of Deep Learning in the Detection and Mitigation of DDOS Attack on SDN Controllers. Technologies 2021, 9, 14.

[5] D'Andrea, E.; Marcelloni, F. Detection of traffic congestion and incidents from GPS trace analysis. Expert Syst. Appl. 2017, 73, 43–56.

[6] Anwar, T.; Vu, H.L.; Liu, C.; Hoogendoorn, S.P. Temporal tracking of congested partitions in dynamic urban road networks. Transp. Res. Rec. 2016, 2595, 88–97.

[7] Liang, Z.; Wakahara, Y. Real-time urban traffic amount prediction models for dynamic route guidance systems. EURASIP J. Wirel. Commun. Netw. 2014, 2014, 1–13.
[8] Boukerche, A., & Wang, J. (2020). Machine learning-based traffic prediction models for intelligent transportation systems. Computer Networks, 181, 107530.
[9] Qian, Z.S.; Li, J.; Li, X.; Zhang, M.; Wang, H. Modeling heterogeneous traffic flow: A pragmatic approach. Transp. Res. Part B Methodol. 2017, 99, 183–204.
[10] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein generative adversarial networks, in: D. Precup, Y.W. Teh (Eds.), Proceedings of the 34th International Conference on Machine Learning, Vol. 70, PMLR, 2017, pp. 214–223.
[11] Kong, X.; Xu, Z.; Shen, G.; Wang, J.; Yang, Q.; Zhang, B. Urban traffic congestion estimation and prediction based on floating car trajectory data. Future Gener. Comput. Syst. 2016, 61, 97–107.
[12] Davoodi, N.; Soheili, A.R.; Hashemi, S.M. A macro-model for traffic flow with consideration of driver's reaction time and distance. Nonlinear Dyn. 2016, 83, 1621–1628.
[13] Xu, Y., Xu, W., Yin, F., Lin, J., & Cui, S. (2017, December). High-accuracy wireless traffic prediction: A GP-based machine learning approach. In GLOBECOM 2017-2017 IEEE Global Communications Conference (pp. 1-6). IEEE.
[14] Aibin, M., Chung, N., Gordon, T., Lyford, L., & Vinchoff, C. (2021, June). On short-and long-term traffic prediction in optical networks using machine learning. In 2021 International Conference on Optical Network Design and Modeling (ONDM) (pp. 1-6). IEEE.
[15] Y. Lv, Y. Duan, W. Kang, Z. Li, F.-Y. Wang, et al., Traffic flow prediction with big data: A deep learning approach, IEEE Trans. Intell. Transp. Syst. 16 (2) (2015) 865–873.
[16] Deng, D. Yu, et al., Deep learning: methods and applications, Found. Trends Signal Process. 7 (3–4) (2014) 197–387.