WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Traffic Flow Prediction for Intelligent Transportation System Using Machine Learning

V Chandra Sekhar Reddy[1], Srikanth Ganji[2], Mudavath Mohan Nayak[3], Medaboina Manish Yadav[4], Guddeti Deepak Reddy[5]

[1]Associate professor, Department of Computer Science and Engineering, ACE Engineering College, Hyderabad, Telangana, India. E-mail: vcsreddy2003@gmail.com

[2,3,4,5] IV B.Tech Students, Department of Computer Science and Engineering, ACE Engineering College, Hyderabad, Telangana, India.

E-mail: ganjisrikanth0@gmail.com[2], mudavathmohannayak300@gmail.com[3], manishyadavmedaboina@gmail.com[4] , deepakreddyguddeti@gmail.com[5]

ABSTRACT

For a very longtime, traffic control has been a challenge. It is necessary in the era of technology. Thevehicle is one of the main ways that technology is advanced nowadays. Communication and information technologies are used by intelligent transportation systems, also known as intelligent traffic systems, to handle traffic control problems. The intelligent transportation system is the main problem in transportation.It's a program, ITS. It is utilized to improve the efficiency of transportation with cutting-edge technology by leveraging sensors and connection. The most contemporary traffic management technologies may be used to alleviate several problems, including clogged roads and poor safety. ITS is improved by using information, control, and electrical technologies based on wireless and wired connectivity. forecast oftraffic flow

INTRODUCTION

Urban regions frequently experience traffic congestion, which increases travel times, fuel consumption, and environmental pollution. The need for effective traffic management grows as cities expand and transportation needs increase. In this setting, traffic flow prediction becomes an important tool for streamlining traffic management plans, boosting transportation systems, and increasing commuter experiences.The estimating and forecasting of future traffic conditions, such as traffic volume, speed, and congestion levels, on a given road network or places is known as traffic flow prediction. Predictive models can offer insights into anticipated traffic patterns and enable proactive decision-making by using historical traffic data, weather conditions, road network features, and other pertinent aspects.

An accurate and trustworthy model or system that can anticipate traffic conditions with precision is what a project to predict traffic flow is aiming to create. These forecasts may be used for a number of things, such as traffic management, route planning, navigational aids, infrastructure planning, and the creation of intelligent transportation systems.Numerous advantages may be derived from precise traffic flow projections. Based on projected traffic circumstances, transportation authorities can efficiently deploy resources, optimize the timing of traffic signals, and conduct traffic control measures. Commuters may choose their routes wisely, avoiding busy regions and cutting down on travel time. Additionally, traffic flow projections may be used by infrastructure planners to locate congested locations, plan road improvements, and create transportation systems that can handle future development.

In this project, we will investigate the field of traffic flow prediction with the goal of creating a solid and reliable model. We want to build a system that can anticipate traffic in real-time or very close to real-time using historical traffic data, sophisticated analytics, and machine learning methods. High forecast accuracy,scalability, and interaction with current traffic management systems will be the project's main goals.

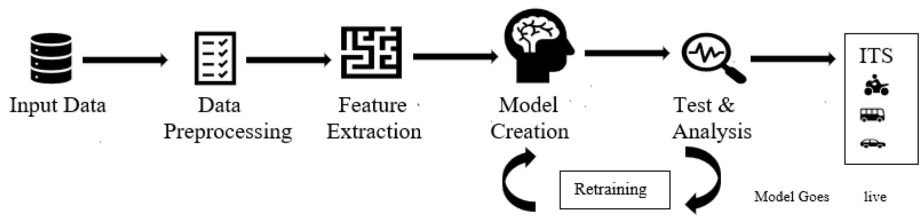
EXISTING SYSTEM

The use of tiny sample data in prediction research is a result of the difficulty in obtaining early traffic data as well as the limited and poor quality of the data environment. As a result, a lot of prediction studies have adopted the complicated model of pure mathematics theory, which overlooks the fundamental traits and evolutionary process of traffic flow. However, too sophisticated models, including neural networks and combination models, involve intricate computations and complex processes, which make them unsuitable for use in actual short-term traffic forecast. The main limitation of the existing system is human and natural factors, such as communication problems, equipment failures are some of them

PROPOSED SYSTEM

To achieve greater efficiency and more precise outcomes, we deployed and evaluated several machine algorithms. The Decision Tree Algorithm (DT) was used to distinguish between classification and regression. Predicting the value of the target variables is the aim of this approach. A decision tree learning function is one that accepts a vector of attribute values as input and outputs a single value as "Decision."It is considered a supervised learning method. Both regression and classification problems may be solved with it. A series of tests on the training dataset are used by DT to determine its findings. Another crucial stage for an accurate output is outliers' detection, and we have employed Support vector machines (SVMs),a collection of supervised learning algorithms, for this.

Database design:



Machine learning Pipelining

PERFORMANCE EVALUATION

Random forest algorithm

The Random Forest algorithm, which belongs to the ensemble learning family, is a versatile and powerful machine learning method. It combines several decision trees to produce a robust and accurate model. Each Random Forest decision tree is built independently using a fraction of the training data and a random selection of characteristics.

During the training phase, the algorithm generates an ensemble of decision trees by selecting a random subset of the training data with replacement and repeating the process. This method, known as bootstrap aggregation or "bagging," aids in the introduction of variation into individual trees. Furthermore, only a random subset of features are examined for splitting at each node of the decision tree, increasing the model's variety.

To make predictions, the Random Forest algorithm aggregates the predictions of all individual trees in the ensemble. For classification tasks, it uses a majority voting scheme, where the class with the most votes across the trees is selected as the final prediction. For regression tasks, it averages the predicted values from all the trees.

Random Forest's capacity to handle high-dimensional datasets with a large number of features is one of its primary advantages. It is capable of capturing complicated interactions as well as non-linear correlations between variables. Furthermore, the technique is resistant to overfitting since the averaging of several trees reduces the influence of outliers and noisy data.

Formula :

$n_{i} = N_{t} / N_{(impurity - N_{t(right)} / N_{t} * right\ impurity) - N_{t(left)} / Nt * left\ impurity)}$

where N t is the number of rows that particular node has

N is th etotoal number if rows present in data

impurity is our gini index value

$N_{t(right)}$ is the number of nodes in right mode

$N_{t(left)}$ is number of nodes in left mode

Decision tree algorithm

Because of its simplicity and interpretability, the decision tree method is a common choice for traffic flow prediction. A decision tree can be trained to forecast future traffic conditions using past traffic data and pertinent features such as time of day, weather conditions, and road infrastructure. The algorithm partitions the data recursively depending on the selected features, selecting the optimum feature at each step to minimise information gain or maximise the separation of traffic flow classes.

Following training, the decision tree model is assessed using a test set to determine its performance. Depending on the task, metrics such as accuracy, precision, recall, or mean squared error can be utilised. Once the model has been validated, it can be used to forecast traffic flow for new instances by traversing the decision tree and calculating the projected traffic flow based on the path followed.

It's crucial to understand that decision trees have limitations. They can overfit the training data and struggle to capture complex feature interactions. As a result, it is recommended that the model be refined by modifying hyperparameters such as the maximum depth of the tree or the minimum number of samples required to split a node. Additionally, merging decision trees with other methodologies or employing more complex machine learning algorithms can improve traffic flow prediction accuracy.

Overall, the decision tree method provides a simple solution to traffic flow prediction, providing insights into congestion sources and assisting in successful traffic management.

Formula :

Entropy (H):

Formula: $H(X) = -\Sigma p(x)*\log2(p(x))$

Information Gain (IG)Formula: $IG(X, A) = H(X) -\Sigma(|X\_v|/|X|)*H(X\_v)$

Gini Index:Formula: $G(X) = 1 -\Sigma p(x)\verb|^|2$

Gini Gain (GG)Formula: $G(X, A) = G(X)-\Sigma(|X\_v|/|X|)*G(X\_v)$

Support vector manchine algorithm

The Support Vector Machine (SVM) is a strong machine learning technique that is often used to forecast traffic flow. It has attracted a lot of attention in transportation research because of its capacity to handle complex data patterns and make accurate predictions. The SVM method works by translating the input data into a high-dimensional feature space and searching for the best hyperplane that distinguishes various classes or predicts continuous values. In the context of traffic flow prediction, SVM uses historical traffic data as input features, such as traffic volume, speed, and occupancy, to forecast future traffic conditions.

A labelled dataset including historical traffic data and matching traffic flow values is required to train an SVM model for traffic flow prediction. The data is then analysed by the SVM algorithm in order to uncover patterns and correlations between the input features and traffic flow outcomes. SVM is able to generalise successfully to unseen data and produce accurate predictions by determining the best hyperplane that maximises the margin between distinct traffic flow classes.

SVM's capacity to handle high-dimensional feature spaces and accommodate a wide range of input features is one of its advantages for traffic flow prediction. This enables the system to capture complex interactions and dependencies between different traffic factors, resulting in more accurate predictions. Furthermore, SVM is resistant to outliers, guaranteeing that extreme values or anomalies in the data have little impact on the model's performance.

Overall, the Support Vector Machine technique is a flexible and useful tool for predicting traffic flow. Its capacity to handle high-dimensional feature spaces, resistance to outliers, and ability to predict non-linear interactions make it well-suited to forecasting traffic conditions effectively. SVM can aid in traffic management and planning by utilising historical traffic data, allowing authorities to make informed decisions and implement effective strategies to optimise traffic flow and alleviate congestion.

Formula :

$G(x) = W^t x + b$

Maximize k such that;

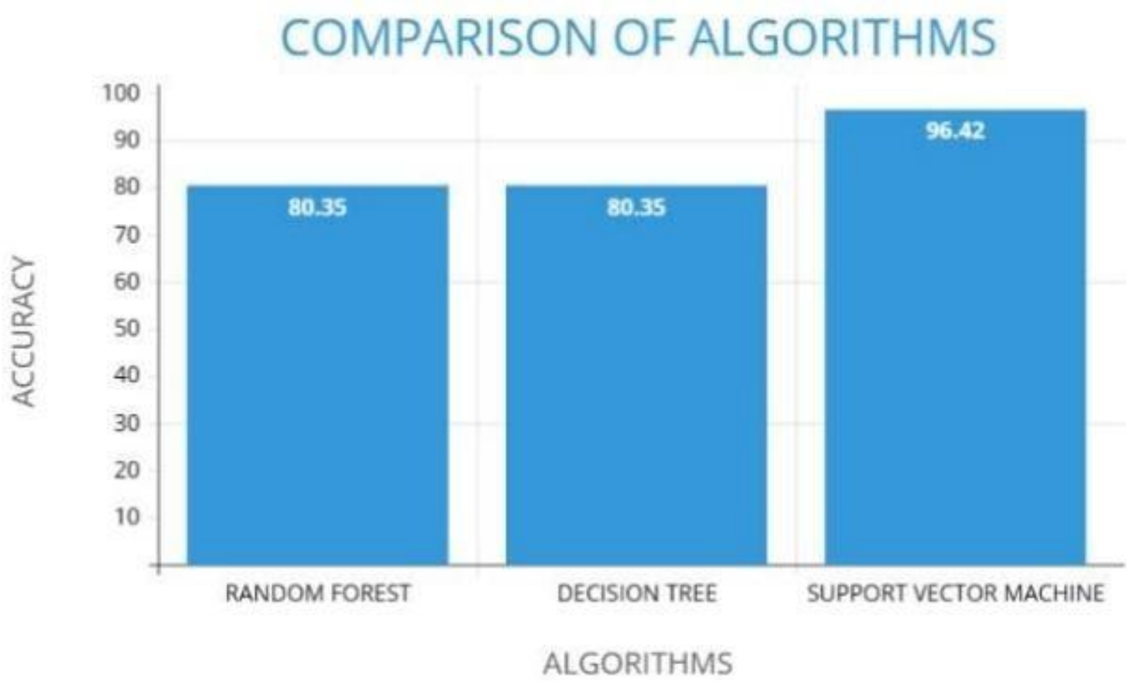
$-w^t x+b \geq k$ for $d\ i == 1$

$-w^t x+b \leq k$ for $d\ i == -1$

Value of g(x)depends upon ||w||:

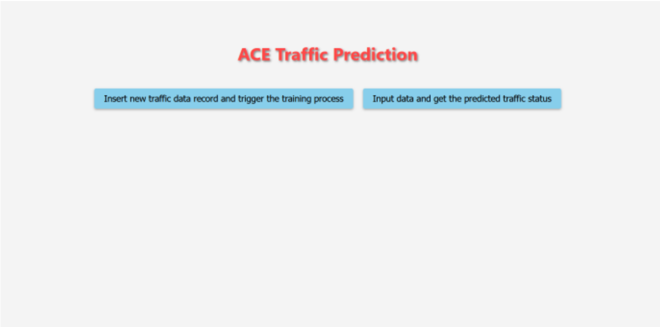
1) Keep ||w|| = 1 and maximize g(x) or

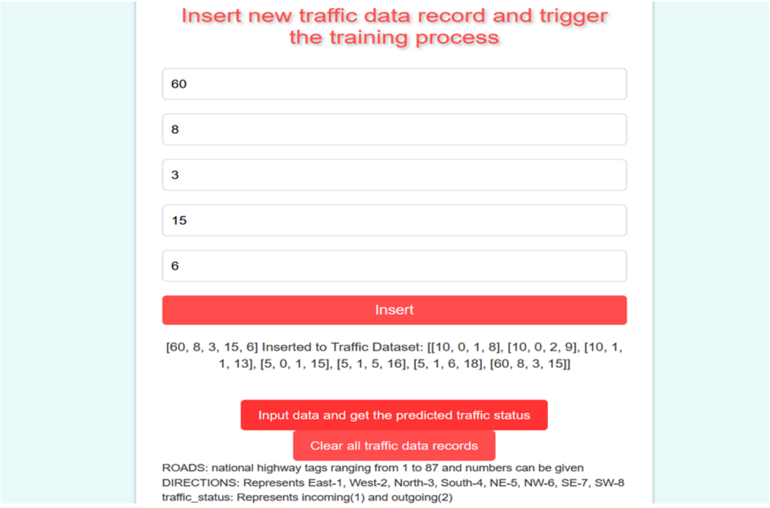
2) g(x) >1 and maximize ||w||



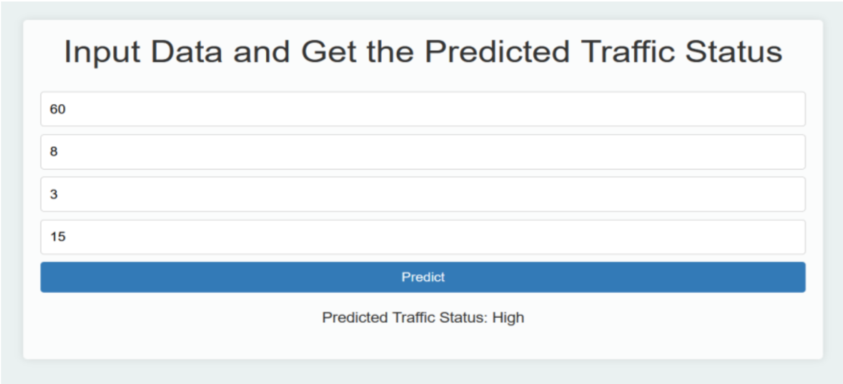
EXPERIMENTAL RESULTS

INPUT :





OUTPUT:



ACKNOWLEDGEMENT

We are grateful to our mentor, **Mr. V Chandra Sekhar Reddy**, for his constant guidance and assistance. Additionally, we appreciate the help and priceless time provided by **Dr. M. V. VIJAYA SARADHI**, Head of the Department of Computer Science and Engineering at Ace Engineering College, and **Mrs. Soppari Kavitha**, our project coordinator.

CONCLUSION

In conclusion, traffic flow prediction algorithms have shown tremendous promise in enhancing transportation systems and alleviating traffic-related difficulties. Machine learning methods, such as Random Forests, have emerged as useful tools for forecasting traffic flow.

These algorithms can create accurate and timely forecasts by analysing previous data and taking into account many elements such as time, weather conditions, transportation infrastructure, and events. These forecasts help traffic management authorities to make educated judgements, optimise traffic flow, and apply proactive congestion-relief measures.

The adoption of machine learning algorithms has various advantages. They can manage big and complicated datasets, discover subtle patterns and correlations, and adapt to changing traffic dynamics. Furthermore, they may include real-time data and continually learn from fresh information, enhancing the accuracy of their predictions over time.

Implementing traffic flow prediction algorithms can provide several benefits. It enables better traffic planning, signal timing optimisation, effective routing and navigation suggestions, and better resource allocation. By anticipating traffic conditions in advance, commuters may make educated judgements, resulting in shorter travel times, increased safety, and less environmental impact.

However, it is important to acknowledge the challenges associated with traffic flow prediction algorithms. These challenges include the need for reliable and comprehensive data sources, potential biases in the training data, and the requirement for regular model updates and retraining to account for evolving traffic patterns.

In summary, traffic flow prediction algorithms offer tremendous potential for enhancing transportation systems. By leveraging machine learning techniques, these algorithms can provide valuable insights into traffic conditions and facilitate more efficient and sustainable traffic management strategies. Continued advancements in data collection, algorithm development, and model refinement will further enhance the accuracy and effectiveness of traffic flow predictions in the future.

Machine learning algorithms may provide accurate forecasts about future traffic flow based on these learnt patterns, allowing traffic management systems and transportation authorities to optimise traffic flow, prepare for congestion, and make educated decisions about infrastructure investments.

Models based on machine learning for traffic flow prediction have various benefits over older techniques. They can handle huge and complicated information, assess several factors at the same time, and react to changing patterns and trends. Furthermore, these models can capture non-linear correlations and interactions between variables that manual analysis may miss.

References

[1] Wang, J., Xu, X., Zhang, K., and Lin, W. H. (2011). Data-driven ITS: a survey. IEEE Transactions on ITS, 12(4), 1624–1639.

[2] Hong, J. S., and Kim, Y. J. (2015). Using a multifactor pattern recognition paradigm, the urban TFP system. 16(5), 2744–2755, IEEE Transactions on ITS.

[3] Hong, J. S., and Kim, Y. J. (2015). Using a multifactor pattern recognition paradigm, the urban TFP system. 16(5), 2744–2755, ITS,IEEE Transactions

[4] Olabarrieta, I. I, Lana, I., and Del Ser, J.,(2016).. Analyzing traffic flow to understand everyday mobility trends in urban road networks. IEEE/IFIP Network Operations and Management Symposium, NOMS 2016-2016, 1(1), 1157-1162. i

[5] Jiang, Z. Wu, Y., Tan, H., Qin, L., and Ran, B., (2018). A hybrid TFP approach built on deep learning and its comprehension. Emerging Technologies in Transportation Research, 90(1), 166–180.

[6] Yan, P., and Jia, T. (2020). use deep spatiotemporal neural networks to forecast citywide traffic flow. Intelligent Transportation Systems,IEEE Transactions ,22(5), 3101-3111.

[7] Zhai, X., Kong, Q. J., Xu, Y., and Liu, Y. (2016). An urban TFP approach based on spatiotemporal variable selection. 50(4), 489–506 Journal of Advanced Transportation.

[8] Zhao, Q., Liu, S., Yang, L., Li, D., and Qi, Y. Qu, W., Li, J., and others (2020). Temporal intersection anticipating the flow of traffic.Sustainability, 12(19), 1-13.