Intelligent Transportation System Using Multi Stream Feature

Tulasi Thanusree[1], Gunti Sai Chaitanya[2], Yama Ruth Elizabeth[3], Mrs. V. Indrani[4], Mrs. A. Prasanthi[5]
Scholar[1,2,3], Associate Professor[4], Assistant Professor[5]
Department of Computer Science and Engineering,
Nalla Narasimha Reddy Education Society's Group of Institutions, Hyderabad, India

Abstract— Traffic flow prediction accuracy is very important for intelligent transportation systems (ITS). Many studies have proposed different methods for traffic flow prediction including ARIMA, ANN and SVM. With the development of deep learning technology, the evolutionary models of RNN such as LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Units) models have been found to perform well in traffic flow prediction. This paper aims at investigating the use of the Random Forest Regressor Model, an ensemble learning algorithm, for improved and accurate traffic prediction. Random Forest algorithm is highly robust and is well suited for large datasets with many features, thus making it suitable for traffic forecasting. In this research, historical traffic data is used to train the model together with other variables such as traffic flow, weather and time. The Random Forest model performance is compared with the traditional prediction methods using Mean Squared Error. It shows that the Random Forest model is better than the conventional methods and can give better accurate forecasts of traffic flow and can be used in real time traffic management. It presents the actual and predicted vehicle count per hour.

Keywords—Traffic Prediction, Intelligent Transportation Systems, Urban Traffic Management, Traffic Volume, Historical Traffic Data, Ensemble Learning, Adaptive Traffic Control Systems.

Introduction

Modern urban mobility faces mounting challenges from rising population densities alongside evolving infrastructure requirements with advancing needs for improved operational safety and performance. Advanced technologies power Intelligent Transportation Systems to solve current transportation challenges in modern urban environments.ITS achieves advancement by integrating multi-stream features into transportation systems that use various data sources for capacity enhancement.anced safety and efficiency. To address these challenges, Intelligent Transportation Systems (ITS) have emerged as a transformative approach, leveraging advanced technologies to optimize transportation networks. A key advancement within ITS is the integration of multi-stream features, which enhances the capability of transportation systems by incorporating and analyzing data from diverse sources.

ITS includes various comprehensive systems which enhance transportation efficiency and safety while promoting network sustainability. Transportation infrastructure experiences improved management alongside superior decision-making ability through information gathering and processing systems enabled by technology platform. Sheer ITS systems operate by using individual isolated data inputs that stem from traffic cameras or vehicle sensors to track and direct roadway conditions. Multi-stream data integration serves as one of the most important advancements within ITS. The multi-stream ITS approach functions differently from single-stream techniques because it conducts simultaneous analysis of various data sources. System data integration uses traffic cameras as well as provides real-time video information about traffic situations and incidents. The system presents information from Vehicle Sensors that tracks vehicle speeds and locations and reports their movement patterns. The system incorporates GPS Data to provide vehicle location alongside travel details and Weather Data serves as meteorological information that affects driving safety conditions and Social Media Feeds capture real-time user reports and observations. The Public Transportation Systems module integrates transportation data from buses alongside trains and all other services of public transit. Multiple data streams combined in multi stream ITS systems deliver complete transportation condition visibility which enables superior analysis and smarter decision making processes.

The Intelligent Transportation System (ITS) employs information technology and communication systems to enhance transportation network efficiency and safety operation.ion networks. ITS obtains real-time processing and information collection while delivering services intended to maximize transportation management along with enhancing safety for roads and vehicles and creating best-in-class travel experiences. The Integrated Analytics of Various Data Flows across Different Origins Represents Multi-Stream Features as the Core Mechanism of the ITS Transport System. The system combines operational capabilities of Data Collection with Data Integration while performing Real-Time Analysis to deliver Decision Support alongside Adaptive Responses. Traffic cameras together with road sensors along with vehicle GPS data and social media feed information deliver real-time details about ongoing traffic conditions as well as movement patterns and roadway problems. Recording devices and data synthesis methods enable numerous scattered information sources to generate comprehensive understanding of transportation connectivity. The unified data stream enables both accurate and fast real-time measurements. Analysis of real-time combined data streams produces vital information to help track patterns and forecast traffic congestion and spot developing challenges. The resulting analysis enables adaptive traffic control systems to modify signal timing and implement incident-based route modifications. ITS generates decision-supporting insights using multi-stream data which enables traffic management centers as well as drivers and other stakeholders to make informed choices. Through navigation systems and alert functions Intelligent Transportation Systems enable drivers to receive real-time traffic data while helping traffic authorities identify potential roadway threats. The system demonstrates flexible capabilities that enable adaptive responses to changing situations by organizing cursory reroutes to minimize congestion and optimize emergency service deployment across systems. The ability of systems to adjust on demand is essential for operating smooth traffic movement and improving road safety. Through multi-stream integration in Intelligent Transportation Systems we achieve significant improvement in transportation network management and optimization capabilities. ITS utilizes diverse data sources plus sophisticated analytical methods to generate enhanced traffic management solutions which simultaneously increase safety measures and improve journey experiences.

Enhanced Traffic Flow will be done by integrating various data sources, ITS can optimize traffic signal timings and manage congestion more effectively. In addition, there are benefits from Improving Safety through real-time monitoring and analysis provides opportunities to respond more rapidly to accidents or hazardous conditions, thus reducing the risk of secondary incidents. It is a source of benefit to drivers from the more recent information on traffic conditions which enables them to plan routes and circumvent delays. Data driven insights have an efficient resource allocation, which means that resources such as traffic management strategies can be adjusted based on current conditions.

I. Related Research

A. ARIMA

ARIMA (Auto Regressive Integrated Moving Average) functions constitute an effective prediction model for time series forecasting with persistent trends or seasonal patterns of daily or weekly intervals. ARIMA is a combination of three major components; Auto Regression (AR) along with differencing (I) to tame the series and the Moving Average (MA) for past error data. It proves to be an excellent linear time series data forecasting technique that can handle a few items of trends and periods. The ARIMA models fail in modeling the non linear traffic data relationships as well as interrupted data relationships by irregular events like accidents and weather disturbances.

B. ANN

Artificial Neural Networks (ANN) is a type of deep learning which helps to transform biological neural network logic to model an irregular pattern between input and output data. It is based on multiple interconnected neural layers, each neuron computes weighted input values and outputs an output by application of his activation function. It recognizes the complex patterns through weighted adjustment based on input data.

ANNs have very good ability to extract nonlinear traffic patterns of how both time divisor and meteorological factors modify the vehicular flow rates. Users of ANN systems are obliged to supply huge amounts of training data and face extremely high computational burdens. Despite this, the constant problem of overfitting means that these models must be tuned correctly.

C. SVM

SVM, which operates as classification and regression supervised machine learning method. The support vector regression variant of support vector machines (SVR) is used to predict the traffic flow in the project. SVM's basic function is of the form of a hyperplane since it finds the region of maximum spacing of data points. SVM attempts to find boundaries composed of lines and planes located in a specified error band to original data points in a regression task. In MultiDimensional scenarios it achieves excellent performance with a populated Radial Basis Function or other kernel SVMs. SVM computational demands do increase with increasing the volume of the traffic data, and it becomes hard to adjust the kernel parameters without hyper parameters.

D. LSTM

In Recurrent Neural Network LSTM (RNN LSTM), it is specialized in dealing with sequential patterns and learning complex long term relationship and is employed through its implementation. LSTM controllers have an Inner mechanism through input forget and output gates which Information that operates inside the memory cells. Because LSTM models *have memory cells that can hold important information throughout long periods of time, long LSTM models can function reasonably well in traffic flow forecasting as the previous patterns influence the ensuing traffic conditions. LSTM is powerful in traffic data process for long term pattern, and also can be used for traffic data process by its ability to the detect long term pattern. LSTMs have high computational cost generally and intensive training data sample needs. LSTMs require intensive changes by means of fine tuning processes to achieve their best results.*

E. GRU

The stripped down version of a LSTM is referred to as GRU (Gated Recurrent Unit) which consumes sequential information with reduced gate count and faster operational efficiency. This framework is complex, but GRUs merge the forget and input gates in a single update gate, thus avoiding this extra complexity, and generally maintaining this capacity in order to process long sequential data streams.

II. METHODOLOGY

A. METHODOLOGY

Information driven techniques are merged with the ensemble learning to solve intricate estimates of traffic movement as a single defined sequence of operations. While working with large input variables as a dataset, Random Forest Regressor shows itself as an excellent solution. Results indicate that Random Forest has outstanding traffic prediction ability compared to ARIMA or deep learning method of LSTM in meeting real time management demand.

III. ARCHITECTURE

Random Forest

The algorithm that we have used in this project to predict traffic flosw is Random Forest Regressor. This is a technique that will construct several decision trees during the training time. Random Forest algorithm is widely considered to be robust and capable to deal with a large dataset with many features and it is said to be appropriate for modeling traffic forecasting. In a regression task such as traffic forecasting, we make the final prediction by averaging these trees' predictions. Random Forest is a random bunch of trees created at each split, which in turn uses random samples of the data and random subsets of features. That is why it helps to prevent overfitting and improve the robustness of the model, and is especially of interest for datasets with numerous features (like traffic volume, weather, time of day, etc). Other than the nonlinear relationship between features (meaning, how traffic volume changes with weather and time), It also Handles. This gives out actual vs predicted volume of vehicles per hour.

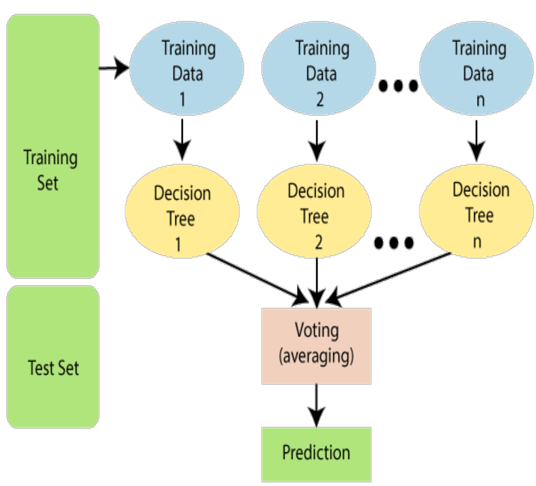

Fig: Random Forest Architecture

B. *ARIMA*

ARIMA is a traditional statistical model used for time-series forecasting. It is based on the concept of modelling the linear relationships in data over time by combining autoregression, differentiating (integration), and moving averages. ARIMA predicts future values by analysing the dependencies between current and past data points. It works best when the data shows consistent patterns or trends. ARIMA is useful for traffic flow prediction when the traffic data shows clear, linear trends over time (e.g., traffic increasing during rush hour). It is capable of modelling seasonal or cyclic patterns, which is useful when traffic flow follows predictable daily or weekly cycles.

C. ANN

ANN consists of multiple layers of nodes (neurons) that process input data. Each connection between neurons has a weight that is adjusted during training to minimize prediction errors. For traffic prediction, ANN can learn patterns from traffic volume, time-of-day, and weather data to predict future traffic conditions. ANN excels in capturing nonlinear patterns in large, complex datasets, such as those found in traffic flow prediction. It can model complex interactions between variables (e.g., how traffic changes differently depending on weather at different times).

D. SVM

SVM works by finding a hyperplane that best separates data points in the feature space, or in the case of regression, it tries to fit the data within a margin of tolerance. SVM is effective when the data is not linearly separable (e.g., traffic flow data affected by multiple complex factors). It can model both linear and nonlinear relationships, depending on the kernel used.

E. LSTM

LSTM networks have memory cells that can maintain information over long periods, which allows them to remember important traffic patterns and trends over time (e.g., rush hour peaks). LSTM excels in capturing both short-term fluctuations and long-term patterns in traffic data, making it a powerful tool for forecasting when traffic conditions change gradually or periodically. Can learn long-term dependencies between past and future traffic flow.

F. GRU

GRU also has memory capabilities, but it combines the forget and input gates into a single gate, reducing complexity while still being able to capture dependencies in traffic flow data. It is preferred when computational efficiency is more critical, as it can often provide similar results with less training time.

IV. EVALUATION

This study finds that the Random Forest model provides more precise predictions compared to traditional approaches like ARIMA and machine learning models like ANN and SVM. MSE is widely used for regression tasks because it enalizes larger errors more severely, which helps in understanding the accuracy of the prediction models. Lower MSE values indicate better predictive performance. The strength of Random Forest lies in its ability to work with big datasets and complicated (nonlinear), such as traffic volume, weather, etc, relations while ARIMA mainly deals with the linear relations. It also provides better performance than the SVM and ANN with regard to accuracy.

However, the results of the evaluation towards the Random Forest model indicate it is well enough as being able to be deployed in real time traffic management systems. It is suitable for practical applications such as traffic signal control, congestion management and route optimization, because its ability to predict traffic flow with precision. The evaluation also focuses on comparing the actual vs. the predicted traffic volume per hour. The prediction of vehicle volumes by the Random Forest model is robust and highly accurate, hence it supports the argument for Random Forest as a suitable model for real time traffic forecasting.

V. RESULT:

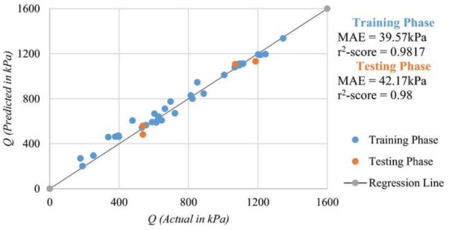

Fig:1 Random Forest Result

The first graph illustrates the performance of a regression model, likely a Random Forest Regression model. Here's a breakdown of the graph's components:

Axes:

The x-axis represents the actual values of the target variable (Q, measured in kPa), while the y-axis represents the values predicted by the model.

Data Points:

The blue dots represent data points from the training phase, where the model learns from the provided data. The orange dots represent data points from the testing phase, where the model's performance is evaluated on unseen data.

Regression Line:

The gray line represents the ideal scenario where the predicted values perfectly match the actual values (a perfect diagonal line).

Performance Metrics:

MAE (Mean Absolute Error): This metric measures the average absolute difference between the predicted and actual values. Lower MAE values indicate better model performance.

R-squared (R^2): This metric indicates the proportion of the variance in the dependent variable that is predictable from the independent variable(s). A higher R^2 value (closer to 1) indicates a better fit.

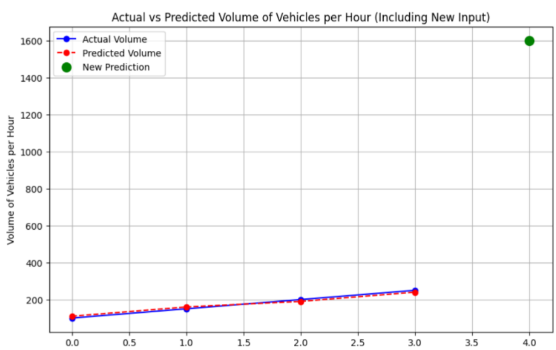

Fig:2 Random Forest Result

The second graph illustrates the comparison between actual and predicted traffic volume per hour, incorporating a new input or feature.

Actual Volume: The blue line with dot markers represents the actual recorded volume of vehicles per hour.

Predicted Volume: The red line with dot markers signifies the predicted volume of vehicles per hour based on a model (likely a machine learning model).

New Prediction: The green dot represents a prediction made using the new input or feature.

VI. CONCLUSION

Our proposed project illustrates the usefulness of the Random Forest Regressor model for highly and reliable traffic flow prediction in an intelligent transportation system (ITS). Through historical traffic data and these key variables — traffic volume, weather conditions and time of day — the Random Forest algorithm shows that it is at the same level for handling big data with lots of features. It was shown that Mean Squared Error (MSE) values of the model are significantly smaller compared to that of traditional prediction approaches such as ARIMA, ANN, and SVM, as well as producing more exact traffic predictions. This demonstrates the random forest model's superior performance for real time management of traffic. It is promising in augmenting more efficient and also dynamic decision making in transportation systems and contribute to better traffic control and congestion management.

VII. REFERENCES

[1] **Jang, H.C.; Lin, T.K.** Traffic-aware traffic signal control framework based on SDN and cloud-fog computing. In Proceedings of the 2018 IEEE 88th Vehicular Technology Conference (VTC 2018-Fall), Chicago, IL, USA, 27–30 August 2018.

[2] **Krizhevsky, A.; Sutskever, I.; Hinton, G.E.** ImageNet classification with deep convolutional neural networks. In Proceedings of the NIPS'12, 25th International Conference on Neural Information Processing Systems, **Red Hook, NY,** USA, 3–6 December 2012; Volume 1, pp. 1097–1105.

[3] **Jang, H.C.; Chang, Y.H**. Traffic flow forecast for traffic with forecastable sporadic events. In Proceedings of the 12th International Conference on Ubi-Media Computing (Ubi-Media 2019), Bali, Indonesia, 6–9 August 2019.

[4] **Du, X.; Zhang, H.; Nguyen, H.V.; Han, Z.** Stacked LSTM deep learning model for traffic prediction in vehicle-to-vehicle communication. In Proceedings of the 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall), Toronto, ON, Canada, 24–27 September 2017.

[5] **Chen, Y.Y.; Lv, Y.; Li, Z.; Wang, F.Y.** Long short-term memory model for traffic congestion prediction with online open data. In Proceedings of the IEEE 19th International Conference on Intelligent Transportation System, Rio de Janeiro, Brazil, 1–4 November 2016.

[6] **Shao, H.; Soong, B.H.** Traffic flow prediction with long short-term memory networks (LSTMs). In Proceedings of the 2016 IEEE Region 10 Conference (TENCON), Singapore, 22–25 November 2016.

[7] **Lewis, C.D.** Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting; Butterworth Scientific: London, UK, 1982.

[8] **Kang, D.; Lv, Y.; Chen, Y.Y.** Short-term traffic flow prediction with LSTM recurrent neural network. In Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017.