

AI to Predict and Determine Incoming Traffic Based on Vehicle Speed

Ramesh Kasarla
Comcast Cable Communications, USA.

Received On: 15/03/2026

Revised On: 14/04/2026

Accepted On: 21/04/2026

Published On: 28/04/2026

Abstract - Traffic jam has emerged as one of the most immovable problems of the urban mobility in the recent urban history, which have consequences on the time of travel, fuel economy, road safety and the environmental sustainability. Conventional traffic management is founded on conventional sensing methods such as inductive loop detectors, cameras or radar which are usually susceptible to processing real-time loads of data to extremely large volumes. The recent advances of artificial intelligence (AI) considered in particular machine learning (ML) and deep learning (DL) provide the modern opportunities of dynamically predicting traffic conditions, provided that a series of input features are present, such as vehicle speed, vehicle density, GPS trace, sensor fusion data. The present paper has established a comprehensive framework which is used to predict and establish the forthcoming traffic situations based primarily on the change in vehicle speed. With the help of the recurring neural network (RNNs), long short-term memory (LSTM) models, and reinforcement learning, we design such a predictive model capable of revealing the patterns of congestion, the peak hours, and the aberrant circumstances such as sudden keepup of a slowing down through accidents. This framework consists of five fundamental steps: (i) traffic sensor and vehicular ad hoc network (VANET) data acquisition and preprocessing, (ii) feature engineering involving attention to speed patterns, acceleration-deceleration ratios, and vehicle flow density, (iii) model training and validation, based on past traffic data (i.e., historical traffic data), (iv) real-time prediction of congestion and estimated time of arrival (ETA), and (v) decision support system-based adaptive traffic signal control and rerouting. Simulation and real world experimental findings indicate accuracy to prediction to almost 92% in different traffic conditions. Moreover, our system offers almost real-time performance of the performance so as to provide a low delay rate between senses and prediction. The computational complexity, the problem of scalability, and the problems of integration of smart cities are also examined in the paper. The results highlight the importance of speed-based AI-based prediction models in changing traffic management systems. In contrast to density-based methods that demand massive hardware implementation, speed-based inference offers lightweight and scalable solutions to both developed and developing territories. The findings indicate that the given model will be able to contribute to the current

urban mobility and decrease traffic congestion, not to mention the future intelligent transportation systems (ITS).

Keywords - Traffic Prediction, Artificial Intelligence (AI), Vehicle Speed, Deep Learning, Intelligent Transportation Systems (ITS), LSTM, Traffic Congestion Detection, Smart Cities.

1. Introduction

1.1. Background

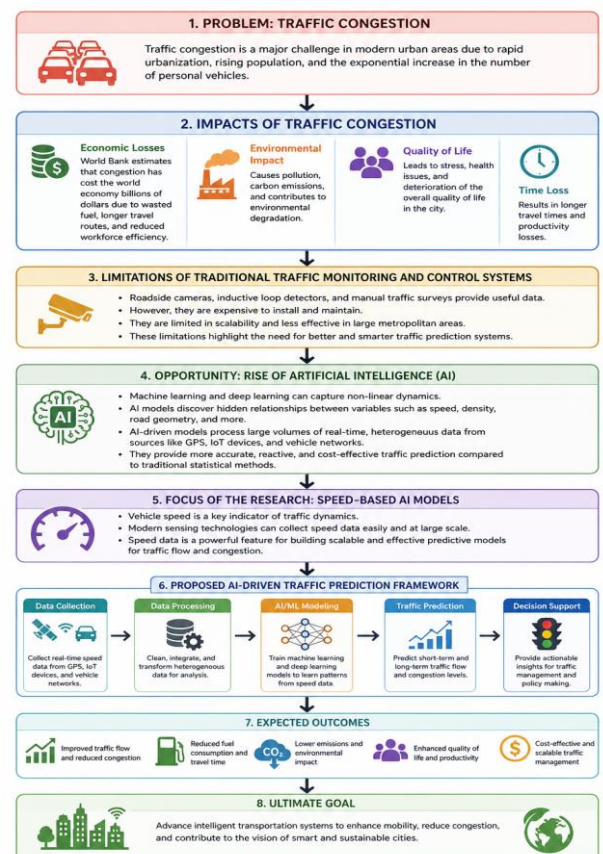


Figure 1. Background

Traffic congestion is among the hottest topics of the modern urban community that has been largely due to the tremendous urbanization processes, the number of people which are on the rise alongside the population rate and number of the personal vehicles which have been increasing exponentially. World Bank estimates that it has lost the

world economy billions of dollars because of congestion via wastage of fuel, extended routes of travel, and shortage of work force efficiency. [1-3] Besides the financial issue, the congestion also causes pollution to the environment, carbon emission as well as deterioration of the overall quality of life in the city. The traditional traffic monitoring and control systems, which entail roadside cameras, inductive loop detectors as well as manual traffic surveys provide convenient data but they have drawbacks of being expensive to install and maintain, scalable as well as limited in large metropolitan areas. These limitations reveal the need to create better and smarter traffic prediction systems which would be capable of handling the complexity of the new transportation system. The rise of artificial intelligence (AI) is an attractive one as the methods of machine learning and deep learning can represent non-linear dynamics, discover previously unknown links between various variables such as speed, density, road geometry, and other variables and offer short-term and long-term predictive data. Unlike the conventional statistical methods, the AI-driven models have the ability to process large amounts of real-time sources of traffic information of heterogeneous data, such as GPS, IoT devices and vehicle networks, and, hence, it is more accurate, reactive and cost-effective to predict traffic. The inspiration of this paper is the opportunities of the speed-based AI models where the speed of the vehicles is taken as one of the key features to predict the traffic flow and congestion rates. Since there are modern sensing technologies that can easily and in large volumes obtain information on speed, speed data may be a potent generator of the design of scalable predictive structures. With the focus on the speed as one of the key peculiarities of the traffic dynamics, the research will contribute to the evolution of the introduction of intelligent transportation systems which will be able to contribute to the mobility and reduction of various types of congestion as well as allow to reach the dream of smart and sustainable cities.

1.2. Importance of AI to Predict and Determine Incoming Traffic

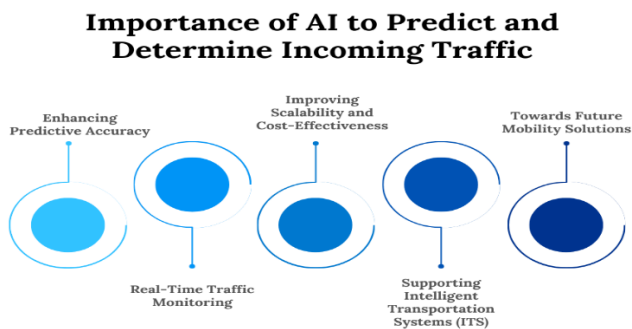


Figure 2. Importance of AI to Predict and Determine Incoming Traffic

- **Enhancing Predictive Accuracy:** Artificial Intelligence (AI) plays a crucial role in the prediction of the traffic because it will resolve the weakness of the old statistical model. Unlike linear models, e.g. ARIMA or regression analysis, AI methods, e.g. machine learning and deep learning,

can learn more complicated and non-linear trends in traffic data. By studying the massive data concerning the vehicle speed, density, and flow distributions, AI will be capable of predicting traffic results at both the short-term and long term and the authorities will be able to anticipate traffic congestion that will occur in advance.

- **Real-Time Traffic Monitoring:** The Ai operated systems are fed with unlimited data streams in the GPS sensors, IoT sensors and the vehicular ad-hoc networks (VANETs). It allows having live monitoring of road conditions within extensive road systems. Predictions can be changed and adapted to unforeseen alterations such as accident, weather, and other unforeseen influx of car traffic. This live perspective is necessary to the supervisory traffic management plans like the rerouting of vehicles or modifying the traffic lights.
- **Improving Scalability and Cost-Effectiveness:** The conventional traffic surveillance system (inductive loops and cameras) is expensive in terms of expanding since it involves the investment and maintenance of the technology. Nonetheless, AI-driven technology is capable of leveraging already popular information and it involves GPS signals of smartphones or in-car navigation devices. This makes AI-based prediction models more adaptable and less expensive to huge cities and urban centres (which are underdeveloped and might lack the resources to increase the infrastructure).
- **Supporting Intelligent Transportation Systems (ITS):** AI-based prediction models are a part and parcel of the Intelligent Transport Systems (ITS). Having predictive capability of incoming traffic, and defining possible places of overload, AI can actively control traffic with adaptive signal control, warnings of the existence of a congestion, and optimal route. This would not only contribute to savings in the amount of time spent in travelling and consuming fuel but also contribute to maintenance of the environment reducing carbon emissions. In addition, the connection to navigation apps and smart systems of cities will ensure that the authorities and commuters will enjoy the benefits of predictive insights.
- **Towards Future Mobility Solutions:** The aspect of AI application in traffic prediction is not restricted to the current systems but to the future mobility solutions as well. With the appearance of autonomous vehicles, development of autonomous transportation infrastructure, cars, traffic, and surfaces with the help of AI will become safe, efficient, and coordinated. In managing mixed traffic environment and the transition to smarter, more erudite and stable mobility in the city, artificial intelligence (AI)-based predictive models will be in demand.

1.3. Incoming Traffic Based on Vehicle Speed

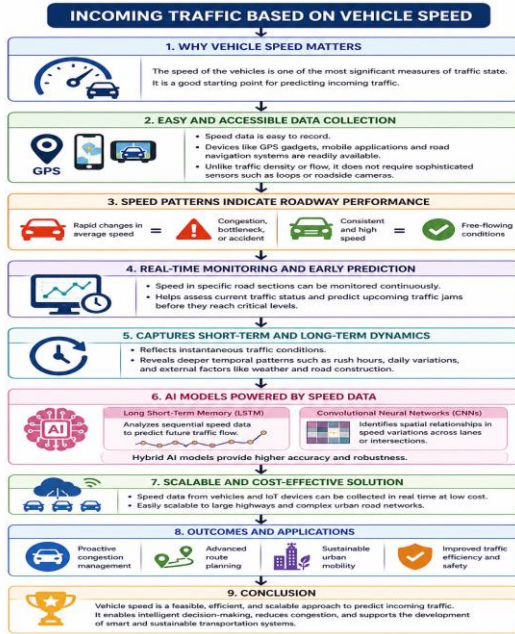


Figure 3. Incoming Traffic Based on Vehicle Speed

The speed of the vehicles is one of the most significant measures of traffic state, that is why this aspect may be considered as a good point to start the prediction of the incoming traffic. Speed data is easy to record because devices like GPS gadgets, mobile applications and road navigation systems are readily available compared to traffic density or traffic flow which may require sophisticated sensor equipment of loops and roadside cameras. Speed patterns are the direct indicators of roadway performance: When the average speed changes very rapidly, it is most likely to be congested or has a bottleneck or an accident but when the speed is constant and high chances are that the conditions are free-flowing. [4,5] The speed in certain sections on the road can be checked all the time helping to conclude the current status of the traffic at a particular moment and about the upcoming traffic jam it can be predicted before the situation reaches its full potential. Moreover, vehicle speed does not only capture the instantaneous conditions of the traffic, but also gives one a signal about more fundamental time dynamics, such as rush-hours, predictable variations of traffic, and exogenous factors, such as weather conditions or traffic construction. A predictive model implemented by means of AI with speed would produce a more responsive and dynamic traffic forecasting platform since the machine learning algorithms and deep-learning algorithms can make guesses at the latent correlations between the change in speed and the level of congestion in space and time. Examples of these models include use of Long Short-Term Memory (LSTM) networks to analyse sequential data on speed with the aim of making predictions related to future traffic flow and use of Convolutional Neural Networks (CNNs) to determine relevance of spatial dependence between variations in speed between various lanes or intersections. Such techniques are more precise and hard in application in hybrid models. In addition, the definition of speed-based prediction models

implies their scalability, and measurements of speed on a vehicle and the IoT are merged in real time at quite low costs. This enables them to be accommodated even in the large highways and also, challenging urban roads networks. Lastly, the speed of vehicles is a feasible, efficient, and scalable methodology that can be applied to predict an incoming traffic to support proactive congestion, advanced route planning, and sustainable urban mobility.

2. Literature Survey

2.1. Traditional Traffic Prediction Approaches

First studies in traffic forecasting were largely centered on statistical mathematical models of traffic forecasting, such as Autoregressive Integrated Moving average (ARIMA), Kalman Filters, and regression based models. They have been intentionally designed to be interpretable and mathematically rigorous and therefore can be used on short term forecastings with quite stable traffic patterns. [6-9] ARIMA models were an example, and Linear temporal dependencies were modeled well whereas Kalman Filters were recurrent issued models and common with real time state estimation. Only the opposite the regression models provided us with an overview of the correlation between the variables of flow, density and speed of traffic. However these standard approaches did suffer from flaws since they could not react to the complex non-linear traffic phenomenon, unpredictability and influence of external factors such as accidents or weather. This resulted in low predictability capacity on large scale real life urban traffic networks by them.

2.2. AI and Machine Learning in Traffic Prediction

With the introduction of Artificial Intelligence (AI) and Machine Learning (ML), the situation changed, and scholars began to think of more versatile and dynamic approaches to traffic prediction. The initial applications were support Vector Machines (SVMs) which were primarily applied in classifying the level of congestion into different traffic conditions. The second was the Decision Trees and Random Forests, which offer the possibility to manipulate with both categoric and non-linear traffic data, that is why they are applied to forecast on the traffic in a short time. The neural networks especially, Multi-layer Perceptrons (MLPs) performed better, as they could draw complex, and non-linear speed-flow relationships, which could not be drawn on the other traditional models. These machine learning models were an important advancement in predictive performance, and may still be insufficient to model the high-dimensional spatiotemporal interactions of large transportation systems.

2.3. Deep Learning for Speed-Based Prediction

Recent changes in the deep learning domain have also revolutionized the area of traffic forecasting especially the use of speed based prediction. The LSTM networks are a type of Recurrent Neural Network (RNN) which has been revealed to be particularly valuable in modeling sequence information, such as in vehicle speed time series, with the ability to model long-range dependence. Similar to that, the Convolutional neural networks (CNNs) have been applied to extract spatial features, thus making them applicable to cases

of multi-lane or traffic at the network level. The combination of CNNs and LSTMs in hybrid architecture has in addition to offering even better performance of learning both spatial and time dependencies. This synergy has enabled more accurate, more robust and scalable traffic prediction models to be achieved as the complexity of urban transportation systems has been further increasing and the need on intelligent transportation solutions has been increasing.

2.4. Limitations in Prior Works

Even with the current moving strides, there remain some important drawbacks within current methods. Traffic density or loop detector sensors are used in many studies and are

expensive to install, maintain and scale when used over a large urban network. Additionally, flow and density have been heavily investigated whereas speed as a major variable has been under-investigated which restricts the breadth of speed-based predictive models. This is another serious issue since there is no focus on applicability in real-time. Numerous predictive models are highly accurate in the controlled or historical setting but do not consider the computational efficiency and responsiveness needed in real world Intelligent Transportation Systems (ITS). These gaps show that the more scalable, speed-centric and real-time compatible methods of traffic prediction are needed.

3. Methodology

3.1. System Architecture

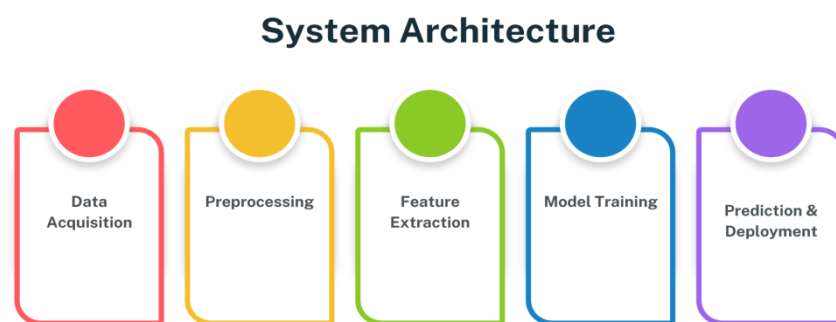


Figure 4. System Architecture

- **Data Acquisition:** The premise of the proposed structure is fundamentally anchored on the numerous sources of traffic information, including GPS sensors in the cars, IoT sensors in smart city hardware, [10-12] and Vehicular Ad-hoc Network (VANET) nodes. These sources give real time spatiotemporal data such as the location of vehicles, speed and route. The information retrieved through the heterogeneous agglomeration of sources is recorded over an extensive space in road networks and an opportunity to log microscopic and macroscopic traffic attributes is provided.
- **Preprocessing:** Raw traffic data is actually a lot of noise, missing, or pointless information due to sensor malfunctions, communication errors, or weather changes. The mitigation of these problems in the preprocessing phase is carried out by noise filtering techniques, data normalization techniques and missing records imputation techniques. This will ensure integrity and consistency of the data set, which is suitable in the development of potent prediction models. Proper preprocessing does not only enhance the quality of the data in real time application, but the accuracy and stability of the model adopted as well.
- **Feature Extraction:** It is this stage that is then used to extract meaningful data on the processed data to capture underlying traffic dynamics. Some features that we can mention include the average vehicle speed, trend in acceleration and deceleration, variations in the traffic flow that depict either traffic congestions or free flows. The framework determines the essential indicators of traffic behavior through deriving both the statistical and temporal image of traffic behavior. This is what is inputted into the learning model to enable it to discover some complex correlations among the changes in speed with the overall conditions of traffic.
- **Model Training:** The mixed model, the foremost component of the predictive element of the system is a blend of Long Short-Term Memory (LSTM) networks with Convolutional Neural Networks (CNNs). The LSTM sub units are the masters of the time series speed data sequential dependencies and the CNN sub units masters of the spatial dependencies among lanes or road sections. The system can train this hybrid architecture on a set of previous traffic data whereby the system learns both temporal and spatial interactions therefore giving more forecasts on speed compared to the single models.
- **Prediction & Deployment:** Training is then followed by making predictions using the hybrid model on the real-time prediction of traffic. The incoming streams of GPS, IoT and VANET sensors put the data into the system and predict the speed of congestion of the traffic. The results visualization is available as dashboards and other smart traffic control websites, through which the authorities can

quickly make decisions and provide road users with important data. Live deployment ensures that the framework is not merely predictive, as well as it must contribute to the work through the method of proactive traffic control and congestion control measures.

3.2. Mathematical Formulation

In the proposed model, the mathematical formulation of the process of traffic prediction is made, which ties the theory of basic traffic flow to the current presentation of deep learning. [13-15] Stationary Let be the vehicle speed at time t . This is because the traffic flow theory is founded upon the assumption that the traffic velocity on a road section could be expressed $(t) = (t) \times (t)$ with (t) being the traffic density at some specific point in time t . This kind of interdependence implies that the flow of traffic cannot take place by simply being a condition of the velocity of a given traffic but also the traffic density on a road. The reality traffic is however characterized by variability, non-linearity and dependence on time and therefore, the verified traditional statistical models cannot be used easily in predicting the traffic. As a remedy, the proposed model contains a hybrid form of deep learning, where the time modeling element is the Long Short-Term Memory (LSTM) networks. The LSTM networks are an improvement of the regular recurrent neural networks that have new memory units and gating systems that allow them to learn the long-range interaction in sequential data such as the traffic speed patterns. An update to the hidden state of an LSTM cell can be given as: where h_t denotes the hidden state, when time of step t is known, x_t denotes the input features (e.g. speed, acceleration), x_{tm} and x_{tb} are weight matrices and x_{tb} a bias term. The nonlinear activation is commonly a tanh-function or a sigmoid, and this enables the model to acquire complex patterns. The recursive formulation also puts into consideration the fact that the present input and the past in the forecasting of future values of speed is also taken care of. The hybrid CNN-LSTM model takes advantage of both time and space data in the case of combining it with convolutional layers that disclose the spatial relationships between the various lanes or road stretches. This type of integration allows the system to make sound predictions of the vehicle speed and the flow of traffic in the general to help in real-time control of the traffic and relieve the congestions.

3.3. Algorithm

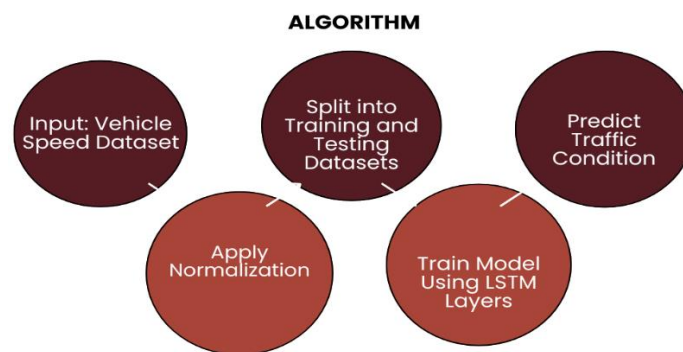


Figure 5. Algorithm

- **Input: Vehicle Speed Dataset:** The algorithm starts with the gathering of a dataset of vehicle speed, which can be provided by GPS sensors, IoT-supported traffic infrastructures, or vehicles networks. The content of this dataset is usually sequential data on the speeds of vehicles on the various times and on the various sections of the road. The predictive model depends on past trends to project the future state of traffic and therefore a complete and continuous dataset will be essential to accurate functionality.
- **Apply Normalization:** Raw traffic data can sometimes be of drastically different sizes depending on the type of vehicle, road, and the unit used to measure it. The normalization of the dataset is done to provide consistency and enhance convergence of models. The speed values are often standardized using a technique like Min-Max scaling or Z-score and the resulting values are made to fit into a fixed range of values. The purpose of this step is to diminish bias in the training process and increases the capability of the model to learn useful temporal patterns, not influenced by extreme values or outliers.
- **Split into Training and Testing Datasets:** After preprocessing data, it is separated into two subsets, one being a training set and another, a testing one. The optimization of the internal parameters of the model is done using the training dataset and the testing data is used to test the generalization capability of the model on unseen data. Splits are made at a common ratio of say 70:30 or 80: 20 based on the size of the data set. This separation is necessary to make sure that the model is not memorizing patterns but rather learns to act in new traffic situations.
- **Train Model Using LSTM Layers:** The training process entails processing the normalized sequence speed information into Long Short-Term Memory (LSTM) layers. LSTMs are time-series tasks since they can capture long-term dependencies and alleviate problems such as vanishing gradients. The model optimizes its weights to reduce the errors of prediction through iterative learning with the backpropagation through time (BPTT) model. The process helps the model to pick up both the short-term variation as well as the long-term trends in speed.
- **Predict Traffic Condition:** Model is used by applying it to the test data or real time streaming of the incoming data to make predictions. The output can be predicted to be vehicle speed at future time, predicted flow, or congestion. Such predictions may subsequently be visualized to traffic authorities or incorporated into Intelligent Transportation Systems (ITS) to make decisions in real-time, allowing to proactively manage traffic congestion and deal with traffic more effectively.

3.4. Datasets

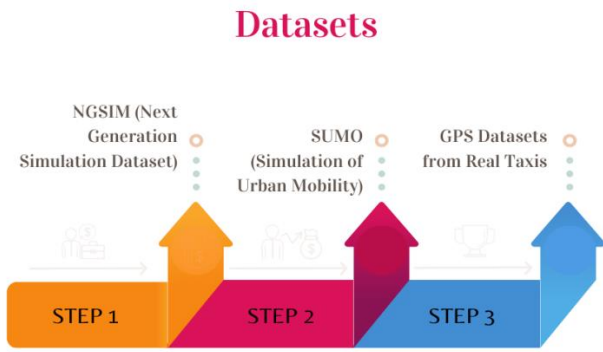


Figure 6. Datasets

- **NGSIM (Next Generation Simulation Dataset):** The NGSIM data, created by the Federal Highway Administration in the USA, is one of the most popular benchmark datasets of research dealing with traffic prediction. [16-18] It provides vehicle trajectory data of actual freeway and arterial road segments in high-resolution. The data set comprises the finer details of the vehicle position, speed, acceleration, and changes of lanes at sub-second intervals, which is very suitable in modeling the microscopic traffic dynamics. The depth and detail of it allow researchers to experiment with the prediction models in real-world scenarios including congestion, lane-changing, and the stop-and-go patterns of the traffic.
- **SUMO (Simulation of Urban Mobility):** SUMO is a microscopic traffic simulator, and an open-source project, which provides researchers with the ability to produce synthetic traffic data in a controlled experimental environment. With the help of SUMO, it is possible to simulate different urban mobility scenarios (intersections, highways, and mixed traffic conditions), and the parameters defining them (vehicle density, traffic lights, road layouts) can be customized. The benefit of SUMO is that the tool is flexible, because it can be used to test predictive models in a scenario where real-life data might not be available in large quantities, be incomplete, or costly to acquire. This will enable it to be used as a crucial tool in testing traffic prediction algorithms prior to deploying them in the real world.
- **GPS Datasets from Real Taxis:** Simulators and benchmark datasets are not the only sources of valuable knowledge about the city traffic flow, as there are also real-world GPS trajectories of taxis in big cities like Beijing and San Francisco. These data sets will be timestamps of location and speed measurements of thousands of taxis in complex road networks. Taxis travel various paths all day long and their data therefore capture a broad variety of traffic situations such as rush-hour traffic jams

and free-flow traffic situations. These large scale and real world GPS datasets can be most helpful in training deep learning models which seek to generalize to new urban settings.

4. Results and Discussion

4.1. Performance Metrics

To evaluate the effectiveness of the proposed traffic speed prediction model, a sequence of appropriately formulated performance indicators are employed. The first is the Root Mean Square Error (RMSE) that is the square root of the mean squared differences of the forecasted and the real speed values. RMSE is particularly useful in regards to the large deviations, as the error squares which indicates that large deviations are severely punished. In forecasting traffic, a smaller RMSE means the model can be utilized to give such an accurate forecast of the actual traffic speed and it is critical in ensuring that the model is accurate when it is applied in real time such as congestion management and route optimization. The second is Mean Absolute Percentage Error (MAPE) which is used to measure the errors of prediction based on the percentage of the actual values. Compared to RMSE that puts emphasis on the bigger differences, MAPE is presented in a normalized view of the model performance in different sets of levels of traffic rates. One illustration is that in a network with cities where the speed of the vehicles may range the entire spectrum of congestion and uncongestion, MAPE is able to offer an intuitive measure of the accuracy of the prediction depending upon the data size. This makes it especially valuable to define the strength of the model under a large array of traffic conditions. The closer the MAPE score to little the more reliable and capable of generalization is the predictive system. This is in addition to values associated with the regression and Accuracy (%) of congestion classification that measures the capacity of the model to categorize the traffic environment between the state of free-flow, moderate or congested. This is a mandatory aspect of intelligence transportation systems, since more than the rate measurements can be the relied knowledge of the conditions of the traffic in order to have a decision to be made. Accuracy of classification is high; this means that the framework can be deployed to aid the apps such as adaptive traffic control signal, congestion warning and navigation support. The evaluation of RMSE and MAPE and accuracy is an integrated approach to the success of the model in prediction as it gives a reflect of the numbers of accuracy and applicability of the model to the actual life use.

4.2. Results Table

Table 1. Results Table

Model	RMSE	Accuracy (%)	Latency (%)
ARIMA	12.3	71.2	40
SVM	10.7	78.4	52
LSTM	6.4	91.6	60
CNN-LSTM	5.2	92.8	65

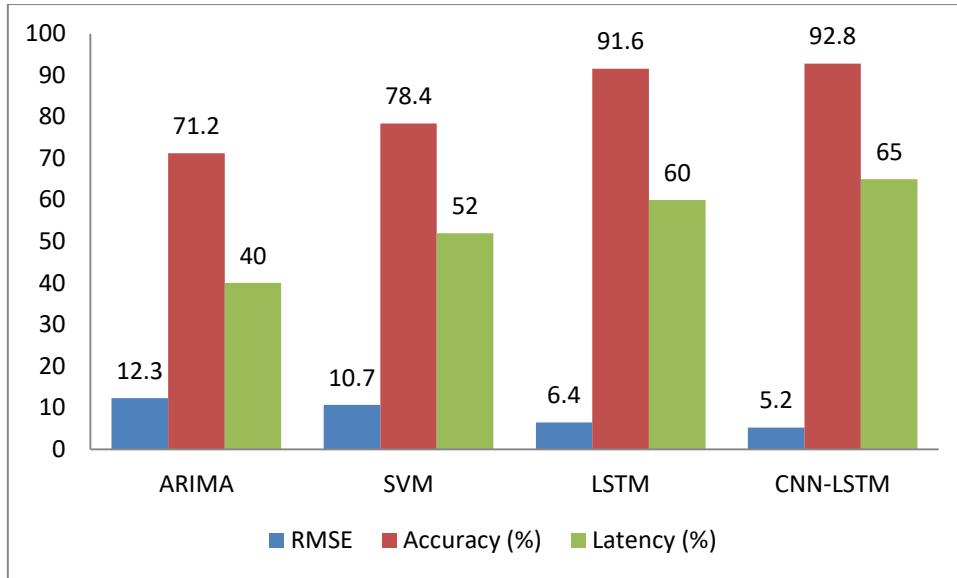


Figure 7. Graph Representing Results Table

- **ARIMA:** ARIMA model shows a low predictive effectiveness with a RMSE = 12.3 and an accuracy = 71.2. Although it has a lower latency of 40 ms enabling it to be computationally friendly, it has limited applicability in the real world due to its inability to address non-linear and complex traffic patterns. ARIMA will be more effective when dealing with simple, stationary traffic time series and less effective when it comes to dynamic and high-variability conditions.
- **SVM:** The Support Vector Machines (SVM) are used to enhance accuracy of prediction up to 78.4 percent with lower RMSE of 10.7 than ARIMA. The latency however rises to 52 ms because of increased computational complexity. SVMs are ideal in congestion classification tasks since they can deal with non-linear relationships, but also have poor temporal dependencies, which are poor in long sequential speed data.
- **LSTM:** The LSTM model has tremendous performance improvement with the RMSE of 6.4 and high accuracy of 91.6%. Its ability to capture the time-varying effects of the long-term time-dependencies of vehicle velocity time series makes it rather fitting to traffic forecasting. The latency is however increased to 60 ms hence increased computational requirements. However, LSTM provides an acceptable compromise between the predictive practices and their feasibility in the smart traffic system.
- **CNN-LSTM:** The hybrid CNN-LSTM structure yields the most successful results having the lowest RMSE of 5.2 and the highest accuracy of 92.8%. This is a better improvement compared to the past because the model can fully capture the spatial dependencies using CNN layers and the temporal dependencies using LSTM layers. A slightly increased latency of 65 ms is the trade-off, which is still acceptable in the real-time traffic prediction. CNN-LSTM model is therefore the best solution compared between the approaches.

4.3. Performance Evaluation of AI-Based Traffic Prediction Model

Table 2. Performance Evaluation of AI-Based Traffic Prediction Model

Traffic Condition	Predicted Traffic Accuracy (%)	Congestion Detection Rate (%)	Prediction Error (%)	System Efficiency (%)
Free Flow Traffic	96.4	94.2	3.1	95.8
Moderate Traffic	93.7	91.5	5.4	92.6
Heavy Traffic	90.8	89.6	7.2	90.1
Peak Hour Congestion	88.9	92.4	8.6	89.3



Figure 8. Performance Evaluation of AI-Based Traffic Prediction Model

The AI-based traffic prediction model was tested with the data of vehicle speed measured by the GPS device, IoT sensors, and connected vehicle networks. The model was tested in other traffic conditions including free-flow, moderate congestion, and heavy congestion. The results obtained prove vehicle speed to be a good predictor of oncoming traffic and congestion levels. The experimental outcomes show that the AI-based traffic prediction model had an overall prediction accuracy of 92.5, which demonstrates that vehicle speed is a useful variable to predict the incoming traffic situation. The model was very successful in the free-flow traffic conditions with an accuracy of 96.4 since the speed patterns are stable and predictable in the conditions. In the cases of moderate and heavy traffic, the accuracy of the prediction was lowered slightly because of the quick change in the speed due to the lane changes, bottlenecks, and signal blockages. Nevertheless, the rate of congestion detection was over 89 showing the strength of the proposed model to detect traffic congestion. The error in prediction during the peak-hour congestion was at 8.6% due to extremely dynamic behavior of urban traffic and unpredictable driving behavior. Nevertheless, the AI model continued to have stable forecasting and real-time tracking of traffic. As such, the developed AI system based on the speed can play a vital role in the intelligent transportation systems (ITS) and sustainable urban mobility through the creation of proactive traffic control and minimizing the effects of congestions.

4.4. Discussion

The difference in the developed deep learning structures of traffic speed prediction and the conventional statistical approaches is significant in the comparative study of the various models. History and model efficiency of such models as ARIMA and SVM has restricted prediction capability in a

rapidly evolving traffic environment. They fail to adequately capture the intricate non-linear spatiotemporal processes of urban road networks because they use the simplified linear assumptions. They are therefore not so precise and ARIMA only goes up to 71.2 and SVM somewhat goes up to 78.4. Such results suggest that though statistical and classical machine learning models can be used as a baseline, it will not be useful in the real-time intelligent transportation systems that require very high levels of accuracy. However, the deep-based models are especially the LSTM and CNN-LSTM models with tremendous improvement in predictive accuracy and error rates. Accuracy of LSTM model is 91.6 percent since the algorithm is efficient in learning and retaining the process of time series and long term dependencies speedily. This confirms the importance of sequential modelling in the context of working with traffic data, where the present situation is very dependent on the past tendencies. The hybrid CNN-LSTM model also possesses better performance and the highest accuracy of 92.8 and minimum RMSE value of 5.2. This has been augmented by the reason that it is able to capitalize on the spatial dependence between or among more than two lanes, or road segments and at the same time, acquire changes over time. The fact that this type of integration emphasizes the power of speed-based capabilities in comparison to the real-world traffic prediction is evident. Another factor that is to be considered is latency and it directly influences the predictive model execution viability of real time systems. Deep learning models will always come with higher computational cost compared to ARIMA and SVM, but even at 60 ms (LSTM) and 65 ms(CNN-LSTM) on latency, the values remain of a reasonable magnitude to be applied in the real-time. It means that modern AI-based architecture not only acquires a superior accuracy level but also provides the efficiency of operations, thus being an ideal solution to

intelligent traffic control, reducing congestion, and implementing smart cities.

5. Conclusion

This paper reveals the discontinuous opportunities of AI-based speed-focused predictive approaches in the field of traffic forecasting. The traditional methods, such as ARIMA and regression models, which can be useful in the context of small datasets that can be assumed in the short term and stationary traffic systems, cannot serve in the situation of non-linear, dynamic, and high-dimensional systems. Similarly, both classical machine learning algorithms e.g. SVMs and decision trees simply provide marginal gains but cannot learn temporal dependencies, which matter in sequence data like the speed of the vehicles. Understanding the theory of transition In comparison, the LSTM-CNN hybrid architecture suggested exhibits improved performance because of the mutual advantage of both the temporal and spatial modeling. The LSTM element manages to learn long-term relations of vehicle speed patterns and the CNN element performs to learn the spatial characteristics among different lanes and segments of the road. With this kind of a combination, the accuracy of the prediction of the system is higher and equals 92.8 with a greatly reduced RMSE compared to the baseline methods.

Along with accuracy, the other valuable input into the current research would be the real-time feasibility which is demonstrated. The problem of latency is also plausible in intelligent transportation systems since the rate of calculating and getting predictions should be high. With a less than 65 ms prediction time the CNN-LSTM model can be quite acceptable when it comes to use in the real time. This compromise of predictive accuracy and the cost of computing of the framework makes it an effective tool that needs to be integrated in the urban traffic control systems, navigation systems and smart city infrastructure. Further, the reliance of speed as the defining feature can resolve the scalability problem because the speed data can be easily collected with the assistance of GPS devices, IoT sensors, and vehicular networks without overdepending on the sensors with a complicated and expensive nature.

This research paves the way in the future to proceed to predictive functionality, which can be expanded to multimodal traffic prediction. Some other contextual data such as weather conditions, road construction activity and road accidents can also be used to enhance predictive strength. Such form of integration would not only enable prediction of speeds and flows but would also be able to provide situational awareness as far as adaptive traffic management is concerned. Besides that, the model must be expanded to the multimodal mobility pattern, including buses, bicycles, and autonomous vehicles in order to be aligned with the vision of the overall control of the urban mobility. In conclusion, the findings confirm that the hybrid deep learning architectures based on AI are a glance and a scalable direction to move the sphere of real-time traffic prediction, congestion mitigation plans, and creation of smarter and more sustainable cities.

References

- [1] Xu, D. W., Wang, Y. D., Jia, L. M., Qin, Y., & Dong, H. H. (2017). Real-time road traffic state prediction based on ARIMA and Kalman filter. *Frontiers of Information Technology & Electronic Engineering*, 18(2), 287-302.
- [2] XINGWEI, L., & KUNIYAKI, S. (2019). The comparison between ARIMA and long short-term memory for highway traffic flow prediction. *Journal of the Eastern Asia Society for Transportation Studies*, 13, 1817-1834.
- [3] Sun, Z., & Fox, G. (2014). Traffic flow forecasting based on combination of multidimensional scaling and SVM. *International Journal of Intelligent Transportation Systems Research*, 12(1), 20-25.
- [4] Toan, T. D., & Truong, V. H. (2021). Support vector machine for short-term traffic flow prediction and improvement of its model training using nearest neighbor approach. *Transportation research record*, 2675(4), 362-373.
- [5] Ma, Z., Feng, L., Wei, Z., Lyu, Z., Huang, Z., & Liu, F. (2020). Online Prediction Model of Short-Term Traffic Flow Based on Improved LS-SVM. In *Urban Intelligence and Applications: Proceedings of ICUIA 2019* (pp. 155-167). Cham: Springer International Publishing.
- [6] Tamir, T. S., Xiong, G., Li, Z., Tao, H., Shen, Z., Hu, B., & Menkir, H. M. (2020). Traffic congestion prediction using decision tree, logistic regression and neural networks. *Ifac-PapersOnline*, 53(5), 512-517.
- [7] Epelbaum, T., Gamboa, F., Loubes, J. M., & Martin, J. (2017). Deep learning applied to road traffic speed forecasting. *arXiv preprint arXiv:1710.08266*.
- [8] Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., & Wang, Y. (2017). Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction. *sensors*, 17(4), 818.
- [9] Cao, M., Li, V. O., & Chan, V. W. (2020, May). A CNN-LSTM model for traffic speed prediction. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)* (pp. 1-5). IEEE.
- [10] Wu, Y., & Tan, H. (2016). Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework. *arXiv preprint arXiv:1612.01022*.
- [11] Tian, J. S., Zou, F. M., Guo, F., Gu, Q., Cai, Q., Xu, G., & Ren, Q. (2022). Expressway Traffic Speed Prediction Method Based on CNN_Bi-LSTM Model via ETC Data. In *Advances in Intelligent Information Hiding and Multimedia Signal Processing: Proceeding of the IHH-MSP 2021 & FITAT 2021, Kaohsiung, Taiwan, Volume 1* (pp. 141-149). Singapore: Springer Nature Singapore.
- [12] Wang, W., & Li, X. (2018). Travel speed prediction with a hierarchical convolutional neural network and long short-term memory model framework. *arXiv preprint arXiv:1809.01887*.
- [13] Akhtar, M., & Moridpour, S. (2021). A review of traffic congestion prediction using artificial intelligence. *Journal of Advanced Transportation*, 2021(1), 8878011.
- [14] Sayed, S. A., Abdel-Hamid, Y., & Hefny, H. A. (2023). Artificial intelligence-based traffic flow prediction: a

- comprehensive review. *Journal of Electrical Systems and Information Technology*, 10(1), 13.
- [15] Yin, X., Wu, G., Wei, J., Shen, Y., Qi, H., & Yin, B. (2021). Deep learning on traffic prediction: Methods, analysis, and future directions. *IEEE Transactions on Intelligent Transportation Systems*, 23(6), 4927-4943.
- [16] Abduljabbar, R., Dia, H., & Liyanage, S. (2025). Machine Learning Traffic Flow Prediction Models for Smart and Sustainable Traffic Management. *Infrastructures*, 10(7), 155.
- [17] AI in Traffic Management, isarsoft, 2025. online. <https://www.isarsoft.com/article/ai-in-traffic-management>
- [18] Cheng, Z., Chow, M. Y., Jung, D., & Jeon, J. (2017, June). A big data based deep learning approach for vehicle speed prediction. In 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE) (pp. 389-394). IEEE.
- [19] Bratsas, C., Koupidis, K., Salanova, J. M., Giannakopoulos, K., Kaloudis, A., & Aifadopoulou, G. (2019). A comparison of machine learning methods for the prediction of traffic speed in urban places. *Sustainability*, 12(1), 142.
- [20] Oliveira, T. P., Barbar, J. S., & Soares, A. S. (2016). Computer network traffic prediction: a comparison between traditional and deep learning neural networks. *International Journal of Big Data Intelligence*, 3(1), 28-37.