



ORIGINAL

Challenges and opportunities in traffic flow prediction: review of machine learning and deep learning perspectives

Desafíos y oportunidades en la predicción del flujo de tráfico: revisión de las perspectivas de aprendizaje automático y aprendizaje profundo

Syed Aleem Uddin Gilani¹  , Murad Al-Rajab¹  , Mahmoud Bakka¹  

¹College of Engineering, Abu Dhabi University. Abu Dhabi, Emiratos Árabes Unidos.

Cite as: Uddin Gilani SM, Al-Rajab M, Bakka M. Challenges and opportunities in traffic flow prediction: review of machine learning and deep learning perspectives. Data and Metadata. 2024; 4:378. <https://doi.org/10.56294/dm2024378>

Submitted: 29-01-2024

Revised: 07-04-2024

Accepted: 28-06-2024

Published: 29-06-2024

Editor: Adrian Alejandro Vitón Castillo 

ABSTRACT

In recent days, traffic prediction has been essential for modern transportation networks. Smart cities rely on traffic management and prediction systems. This study utilizes state-of-the-art deep learning and machine learning techniques to adjust to changing traffic conditions. Modern DL models, such as LSTM and GRU, are examined here to see whether they may enhance prediction accuracy and provide valuable insights. Repairing problems and errors connected to weather requires hybrid models that integrate deep learning with machine learning. These models need top-notch training data to be precise, flexible, and able to generalize. Researchers are continuously exploring new approaches, such as hybrid models, deep learning, and machine learning, to discover traffic flow data patterns that span several places and time periods. Our current traffic flow estimates need improvement. Some expected benefits are fewer pollutants, higher-quality air, and more straightforward urban transportation. With machine learning and deep learning, this study aims to improve traffic management in urban areas. Long Short-Term Memory (LSTM) models may reliably forecast traffic patterns.

Keywords: Smart Cities; Intelligent Transportation Systems; Machine Learning (ML); Deep Learning (DL); Long Short-Term Memory (LSTM).

RESUMEN

En los últimos tiempos, la predicción del tráfico ha sido esencial para las redes de transporte modernas. Las ciudades inteligentes se basan en sistemas de gestión y predicción del tráfico. Este estudio utiliza técnicas de aprendizaje profundo y aprendizaje automático de última generación para ajustarse a las condiciones cambiantes del tráfico. Los modelos modernos de DL, como LSTM y GRU, se examinan aquí para ver si pueden mejorar la precisión de la predicción y proporcionar información valiosa. La reparación de problemas y errores relacionados con la meteorología requiere modelos híbridos que integren el aprendizaje profundo con el aprendizaje automático. Estos modelos necesitan datos de entrenamiento de primera categoría para ser precisos, flexibles y capaces de generalizar. Los investigadores están explorando continuamente nuevos enfoques, como modelos híbridos, aprendizaje profundo y aprendizaje automático, para descubrir patrones de datos de flujo de tráfico que abarcan varios lugares y períodos de tiempo. Nuestras estimaciones actuales del flujo de tráfico necesitan mejoras. Algunos de los beneficios esperados son menos contaminantes, aire de mayor calidad y un transporte urbano más sencillo. Con el aprendizaje automático y el aprendizaje profundo, este estudio pretende mejorar la gestión del tráfico en las zonas urbanas. Los modelos de memoria larga a corto plazo (LSTM) pueden predecir de forma fiable los patrones de tráfico.

Palabras clave: Ciudades Inteligentes; Sistemas de Transporte Inteligentes; Aprendizaje Automático (ML); Aprendizaje Profundo (DL); Memoria a Largo-Corto Plazo (LSTM).

INTRODUCTION

Implementing intelligent transportation systems relies on up-to-date and accurate data on traffic flows.⁽¹⁾ It is officially entered the age of big data in transportation, thanks to the explosion of traffic data in recent years. Current approaches to predicting traffic flows rely heavily on simplistic traffic prediction models,⁽²⁾ which fail to deliver in several practical contexts. Due to this circumstance, it is motivated to reconsider the traffic flow prediction issue using deep architectural models and massive traffic data.⁽³⁾

In sequence prediction tasks, LSTM captures long-term dependencies and excels. Time series, machine translation, and voice recognition⁽⁴⁾ are some of the applications that benefit from the order dependency. This article gives a comprehensive overview of long short-term memory (LSTM), discussing its concept, architecture, and foundational principles as well as its essential function in a range of application domains. Because of its one hidden state that is transmitted down over time,⁽⁵⁾ a typical RNN may struggle to learn long-term dependencies. One solution proposed by LSTMs model is the memory cell, a container with the capacity to store information for a long time.^(6,7,8)

Translation, voice recognition, and time series forecasting are just a few of the many applications that benefit from LSTM architectures' ability to learn long-term links in sequential data.

In order to achieve these goals, we need novel and innovative approaches for precisely forecasting and managing traffic,^(5,9) which might enhance traffic management and raise living standards in smart cities. Moreover, one additional objective of smart cities is to create the circumstances for future sustainability and better living conditions by lowering the emissions generated by different companies. We want to explore the remaining challenges in this area as well as the progress made in the previous few years in machine learning and deep learning to forecast traffic flow and congestion.^(5,6,9,10)

Recognizing the flaws and limitations in the present research in the topic is essential even with the developments. In this work, these issues—which vary from model-specific restrictions to more general challenges in using ML and DL techniques in real-world traffic situations—are thoroughly investigated. Furthermore emphasized throughout the paper is the significance of data for these models' training and validation. Researchers employ several datasets to enhance model performance, including information on traffic, pollution, and weather. Furthermore, as the researchers ought to have realized, feature engineering is essential for traffic flow prediction. In forecasting a known variable, it is essential. Our models have better prediction potential since we have carefully designed and chosen pertinent elements. These characteristics help us to develop precise forecasts by capturing important data. The aim is to create models that show adaptability to the unique characteristics of different urban settings and prediction accuracy.

Recent scientific research attempting to anticipate traffic flow using ML and DL models are reviewed in this paper. The primary objective is to fully explore strategies, techniques, and datasets in order to understand their effectiveness, difficulties, and shortcomings in already used techniques. Traffic prediction systems have advanced noticeably by the use of state-of-the-art ML and DL techniques. Higher forecast accuracy has been shown by a large number of studies on the effectiveness of these techniques in assessing traffic flow dynamics. Remarkable approaches that work well in many situations include ensemble methods, autoencoders, and Long Short-Term Memory (LSTM) Gated Recurrent Unit (GRU).⁽¹¹⁾

The most current advances in deep learning (DL) and machine learning (ML) models for traffic flow prediction are included in this study. For their traffic infrastructure development, it also provides academics, policymakers, and practitioners with useful information and insights. Moreover, it gives a comprehensive summary of the subject by examining methodologies, evaluation criteria, and datasets. In the end, this study facilitates and adds to the literature resource for upcoming scholars and researchers.

The paper is organized as follows: first, we introduce traffic flow, traffic flow prediction, and the significance of ML and DL algorithms in the predictions. In the second chapter, we explain our literature review and the methodology used for this research review. Chapter 3 also provides a detailed review of previous studies using ML for traffic flow predictions. Similarly, Chapter 4 discusses studies using DL for traffic flow prediction. Chapter 5 discusses the comprehensive comparison between ML and DL performance.

Furthermore, Chapter 6 briefly overviews all the datasets used in the studies reviewed. Moreover, Chapter 7 discusses previous studies' gaps and limitations. Chapter 8 briefly overviews the ML and DL models researchers employ in traffic flow prediction and shows excellent results. In addition, evaluation metrics are discussed in Chapter 9. Ultimately, we conclude the paper with a conclusion and prospects.

METHOD

The systematic review approach used in this study is considered rigorous and objective for analyzing existing research. The methodology involved the following tasks:

Comprehensive search: a thorough search was conducted methodically on reliable scholarly databases such as Google Scholar, IEEE Xplore, Elsevier, ResearchGate, and Springer. Targeting the title, abstract, or keyword list, specific search phrases like “traffic flow prediction”, “machine learning”, “deep learning”, “LSTM”, “RNN”,

“CNN” and “Congestion Prediction” were carefully used. A comprehensive search technique was implemented to ensure the retrieval of relevant journal articles and conference papers published between 2013 and 2023.

Inclusion and exclusion criteria: precise inclusion and exclusion criteria guarantee the reliability and validity of the results. The criterion centered on finding essential features in the articles that could affect how accurately traffic flow forecasts work. These included improved performance results, dataset accessibility, increased accuracy, and relevant dataset properties.

Qualitative content analysis: a qualitative content analysis technique was utilized to examine the data systematically. This technique involved gathering relevant text from the articles, summarizing the main findings, and consolidating the results. This approach made conducting a thorough and objective analysis of the literature on using deep learning and machine learning algorithms for traffic flow prediction easier.

Following a thorough literature search, we discovered an abundance of research being conducted in this field. However, after doing a comprehensive examination to identify studies conducted in the previous five years, we found that the numbers had decreased, and we selected 70 academic publications that were relevant to our investigation. The selection of literature was based on content, specifically on research involving ML algorithms like Support Vector Machines (SVM), Deep Belief Networks (DBN), and Autoencoders in the field of traffic flow predictions, as well as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU). Figure 1 presents the literature review procedure as a flowchart for better comprehension.

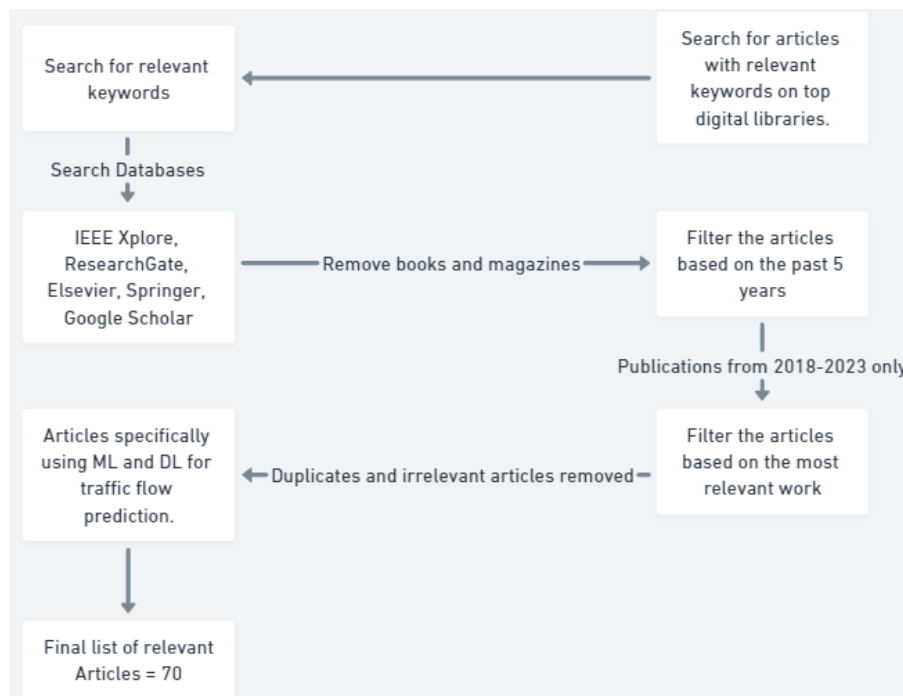


Figure 1. Research Methodology flowchart

Table 1 shows the total number of studies carried out in the timeframe and how we found the studies relevant to our research.

Table 1. Performance Comparison Results					
Searched Literature	Google Scholar	IEEE	Springer	Elsevier	Research Gate
2000 - 2023	21,400	206	18,890	5,244	75
2013 - 2023	17,900	198	17,147	4,748	65
Relevant Work	100	32	10	14	5

To improve traffic management, reduce congestion, and make travel more efficient, intelligent transportation systems (ITS) mainly depend on traffic prediction.⁽⁹⁾ Statistical models and time-series analysis, two common techniques to traffic prediction,^(12,13) are the most effective methods for capturing the complex and ever-changing nature of traffic flow. The use of deep learning (DL) and machine learning (ML) has resulted in a significant improvement in both the effectiveness and precision of traffic prediction.^(9,14)

Through the use of prior data on traffic flow, long short-term memory (LSTM) has the potential to enhance the accuracy of its short-term prediction.⁽¹⁵⁾ While more conventional approaches, such as ARIMA, cannot extract any novel insights from the data, this one can do so.

A further design option for recurrent neural networks is the combination of Autoencoders with Long Short-Term Memory (AE-LSTM) and the Gated Recurrent Unit (GRU).⁽¹⁶⁾ This is an alternative implementation of the concept. Convolutional Neural Networks (CNNs), Mixture Density Networks (MDNs), Random Forest, and Support Vector Machines (SVMs) are some of the other machine learning and deep learning approaches that are discussed in this article.^(18,19)

In the following introduction part, we will discuss the relevance of LSTM and its variations to this very important discipline. It indicates how effectively they perform and adapt to cope with challenging traffic forecast issues, particularly in urban regions, where unexpected changes to traffic patterns are widespread. This is especially true in metropolitan areas.^(9,15)

Over the period of the last five years, many machine learning models, including ensemble methods, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Autoencoders (A.E.), have been the primary focus of study in the field of traffic flow prediction. This is seen in figure 2 there has been a consistent and statistically significant rise in the amount of research that belongs into this classification. Notably, the amount of research conducted on LSTM increased by around 200 percent between the years 2018 and 2022. By comparison, throughout the same period of time, GRU and A.E. research both grew at 100 %. Furthermore, throughout the assigned time frame, research on Ensemble techniques has increased significantly, by 60 %. These findings show how popular and focused machine learning (ML) models are becoming for traffic flow prediction.⁽²⁰⁾

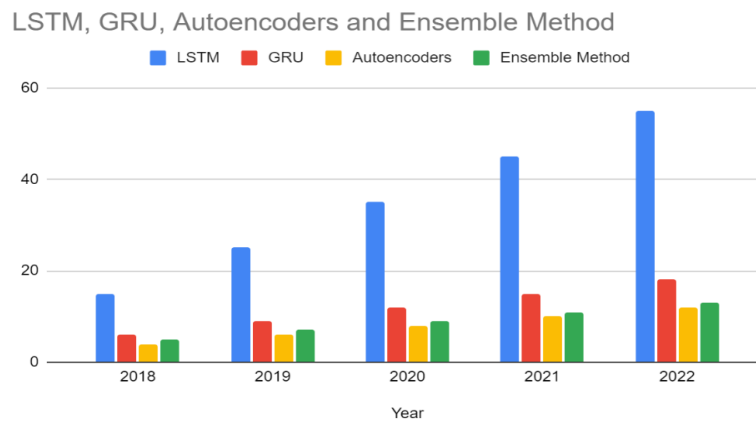


Figure 2. Statistics of LSTM vs GRU, A.E., and Ensemble⁽²⁰⁾

LSTM, CNN, and other algorithms

As powerful and useful tools for traffic flow prediction, machine learning algorithms provide experts and researchers more accuracy and efficiency when examining and predicting traffic trends. The basic citations that have helped to integrate machine learning into traffic flow prediction are examined in this extensive study of the literature, with special attention to the key publications that have fundamentally influenced this field. Figure 3 shows how much more these models are being used, and LSTM has shown to be the most accurate model.⁽²⁰⁾

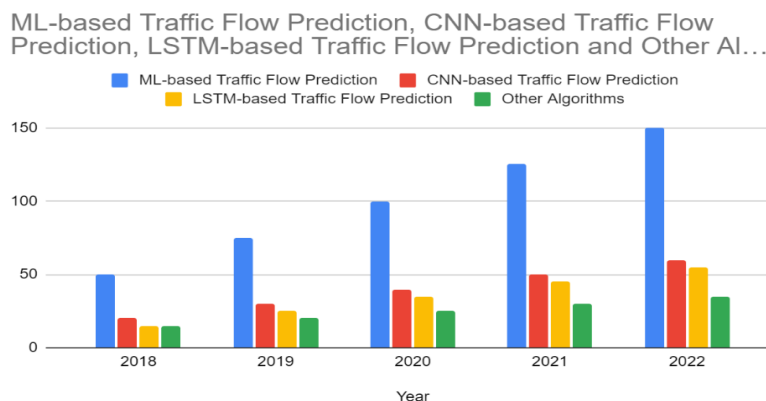


Figure 3. ML and DL models used in Traffic flow prediction over the years⁽²⁰⁾

In conclusion, long-term spatial traffic flow modelling widely employs LSTM and its variations as well as combinations with other deep learning techniques because of its capacity to precisely predict short-term trends and capture temporal correlations. Because of its recurrent structure and memory capacity, LSTM is a good choice for simulating traffic patterns, particularly in cities where traffic conditions might change suddenly. Moreover, research on traffic prediction has found that LSTM-based models perform better than more conventional machine learning techniques like ARIMA. For these reasons, they are a preferred option.

Traffic Flow

Traffic flow studies center on the interactions between various traffic participants and infrastructure in order to identify the relationship between certain traffic participants and the resulting traffic flow phenomena. Studies of traffic flow are empirical in nature and depend on precise measurements to draw their conclusions. The importance of trajectory data, particularly the Next Generation Simulation (NGSIM) trajectory dataset,^(22,23,24) has been emphasized in the analysis of traffic flow in recent years.⁽²¹⁾ Researchers may use trajectory data to look at traffic phenomena and models from many angles, and it gives them a cohesive picture of these accomplishments at the microscopic,⁽²⁵⁾ mesoscopic,⁽²⁶⁾ and macroscopic⁽²⁷⁾ levels. The research classifies traffic flow studies into three types: basic diagram and traffic wave models, studies at the mesoscopic level (such as headway/spacing distributions), and studies at the microscopic level (such as car-following and lane-changing models).⁽²⁸⁾ Asymmetric driving behavior,⁽²⁹⁾ lane-changing behavior,⁽²⁸⁾ traffic oscillations, and traffic hysteresis^(21,30) are some of the novel traffic phenomena discussed by Li et al.⁽²¹⁾ that have been confirmed or explained by trajectory data.

Intelligent Transport System

The deployment of several components, including bus schedules, traffic flow collecting, geographic information, signal control, alarm, and video monitoring systems, is the outcome of the introduction of Intelligent Transportation Systems (ITS) in 1990. Even if these systems have improved urban circumstances, problems still exist because of poor top-level design and general planning, which causes issues with numerous transit systems.⁽³¹⁾ One of these significant problems nowadays Experienced by the ITS is efficient and accurate traffic flow prediction. To alleviate this problem, researchers are carrying out a lot of work, and they are using artificial intelligence (A.I.), machine learning (ML), and deep learning (DL), which are showing promising results.

Artificial intelligence (A.I.) has expanded significantly in machine learning and deep learning, as shown in figure 4. It has demonstrated itself to be proficient at correctly predicting traffic patterns. For instance, an A.I. model may be trained using data from the day before to forecast the traffic volume on a particular highway section at a given time. To minimize congestion, this information can be utilized to change the direction of traffic, open or close lanes, and alter traffic lights.

Here are some particular instances of how traffic flow prediction is being enhanced through the use of A.I., ML, and DL: mobile surveillance cameras can track vehicle movements and identify areas of congestion.⁽³²⁾ Intelligent I.P. cameras make vehicle classification and speed identification possible.^(33,34,35) License plate data collected by LPR cameras may be utilized to follow the movements of specific cars and spot trends in traffic flow.^(36,37)



Figure 4. Structure of ITS⁽³¹⁾

A.I., ML, and DL models may be trained to forecast traffic flow in various circumstances using this data. For instance, a model might be developed to predict traffic patterns during construction, rush hour, or special occasions.

Machine Learning in Traffic Flow Predictions

Prediction of traffic flow is one of the many disciplines where machine learning (ML) has completely changed research. Intelligent transport systems (ITS) depend on traffic flow prediction. It attempts, using past and current data, to predict traffic situations in the future.

Complex pattern learning and prediction capabilities of ML approaches have shown tremendous promise in this area.^(9,38) They also manage the nonlinear and dynamic character of traffic flow, which is a problem for conventional statistical techniques.⁽⁹⁾

We reviewed a number of papers that used ML models, which perform really well. An important element of intelligent transport systems, accurate and quick traffic flow prediction is achieved by a hybrid model developed by Zhang et al.⁽³⁹⁾ that blends Deep Belief Networks (DBN) with Genetic Algorithms (G.A.). Remarkably, the FR-CG algorithm outperformed G.D. and LBFSGS and routinely produced lower MRE values. It was found, specifically, that the MRE of FR-CG was 0,5 % lower than LBFSGS and 3,4 % lower than G.D. As the amount of components increased, this performance pattern was evident throughout a variety of time frames. For models with additional weights and biases, FR-CG fared better than LBFSGS.

When predicting the flow of traffic in the near future, Hou and colleagues have offered a thorough method that accounts for the influence of weather.⁽⁴⁰⁾ Combining sequence-to-sequence organization with multivariate data analysis methods is the suggested paradigm. More than 4500 loop detectors provided traffic data, while the National Oceanic and Atmospheric Administration provided meteorological data (<https://gis.ncdc.noaa.gov/maps/ncei/lcd>).⁽⁴⁰⁾ Prior to analysis, the data was subjected to preprocessing and a plethora of procedures, such as one-hot coding and time-flow correlation expression. The goal of these methods was to improve the data's usefulness and quality so that models could be trained and evaluated more effectively. Here are the findings for the suggested model: error rate (ERR) = 10,378414 %, symmetric mean absolute percentage error (SMAPE) = 10,059485 %, mean absolute error (MAE) = 9,494741, mean squared error (MSE) = 151,153208, and root mean squared error (RMSE) = 12,294438.⁽⁴⁰⁾ As an added bonus, Xu et al.⁽⁴¹⁾ suggested a fresh approach to anticipating the flow of traffic on motorways by integrating machine learning techniques with toll data. The approach uses multilayer perceptron and Random Forest models to improve traffic flow and forecast accuracy. The study report compares the combined model's performance to that of the separate models in great detail. When compared to the separate models, the combined one performed better. Results demonstrated that the combined model's Mean Absolute Percentage Error (MAPE) was much lower than that of the random forest (2,594) and multilayer perceptron (6,490) models.

Jiang and colleagues investigated and reported a factor analysis and weighted Random Forest-based traffic incident recognition method.⁽⁴²⁾ With this method, traffic occurrences are accurately and immediately detected, therefore reducing property damage and fatalities. The work also shows how unbalanced traffic event data impacts detection performance and makes recommendations for reducing this impact.⁽⁴²⁾ The proposed model handles unbalanced data classification quite easily and exhibits superior classification performance to that of the Support Vector Machine (SVM) by combining factor analysis with weighted random forest. In this work, the detection rate (DR%) of the suggested method is 95,65 %. The figures for Factor Analysis Support Vector Machine (FA-SVM), Weighted Random Forest Automatic Incident Detection (WRF-AID), and Random Forest Automatic Incident Detection (RF-AID) are, nevertheless, 84,7 %, 90,43 %, and 89,04 %, respectively. As an additional point of interest, according to Jiang et al.⁽⁴²⁾, the accuracy rates (AR%) of RF-AID, WRF-AID, and FA-SVM are 97,5 %, 98,18 %, and 98,4 %, respectively. This technique has an AR% of 98,9 %, which is rather impressive.

On the other hand, Bratsas et al.⁽⁴³⁾ examined the ability of four different machine learning models to forecast the amount of traffic that flows through metropolitan regions. Researchers were able to collect data from probes in Thessaloniki, which they then used in combination with the road networks of Greece. Based on this study's findings, it is impossible to design or run traffic management systems or apps that use traffic signals if accurate traffic predictions are not available.

Deep Learning in Traffic Flow Predictions

Deep learning (DL) has shown to be an extremely effective method for predicting traffic flow since its implementation.^(1,44,45) Deep learning (DL) is a subfield of machine learning that enables users to model complex nonlinear connections and recognize spatial and temporal patterns in traffic data. This is made possible by Deep Learning. Additionally, deep learning is quite effective at identifying patterns in the data collected from traffic. Since they made use of DL models and achieved good findings, the following papers were selected for inclusion in this study:

They constructed and tested an LSTM-based model to anticipate traffic patterns as part of their published work. Ramchandra et al.⁽⁴⁶⁾ were responsible for this. The suggested paradigm incorporates four separate

machine learning ideas: Random Forest (R.F.), Deep Belief Network (DBN), Deep Autoencoder (DAN), and Long-Short-Term Memory (LSTM). The proposed paradigm is the result of a synthesis of many ideas. Recall, accuracy, precision, and error value were only a few of the metrics used to assess the model's efficacy. Responsible decision-making would have been on their part had they concentrated mostly on traffic pattern projections in heavily populated areas. With an accuracy rating of 95,2 %, ⁽⁴⁶⁾ the LSTM model outperformed the other three methods that were considered. Other approaches fared far worse than the ones utilized, according to the statistics.

Autoencoder (A.E.) and Long Short-Term Memory (LSTM) are combined in a novel approach proposed by Wei et al. ⁽⁴⁷⁾ for short-term traffic flow prediction. Main objectives of their research were to enhance urban living standards generally and lessen traffic congestion in urban transportation. Among the techniques looked at in the research was A.E. using LSTM, Gaussian processes, Hidden Markov model, Deep belief network, Seasonal ARIMA (SARIMA), Random Region Transmission Framework, Bayesian MARS. With historical traffic flow data and autoencoders for feature extraction, the authors investigated the characteristics of both upstream and downstream traffic flow. Performance of the algorithms was evaluated using Mean Absolute Error (MAE) and Mean Relative Error (MRE). According to the findings of the inquiry, the AE-LSTM approach consistently obtained the best likelihood of decreasing the MAE (90 %) across all 25 stations. This was the case regardless of the method that was under examination. In comparison to the other options, this strategy was superior. Given that the data presented here demonstrates that the AE-LSTM approach is capable of accurately predicting future traffic flows, ⁽⁴⁷⁾ this indicates that it has the potential to greatly improve city traffic. In light of this, it is clear that its projections about traffic flow are correct. Autoencoder Gated Recurrent Unit (AE-GRU) is an ingenious device that Chen and his colleagues developed while they were seeking for methods to estimate future traffic levels. ⁽¹¹⁾ Denoising the data was a technique that Chen and his colleagues used in order to reduce the possibility of abnormalities occurring in the traffic flow data. The results of the experiments demonstrated that the LSTM+EEMD (Ensemble Empirical Mode Decomposition) technique improved the performance of the prediction. ⁽⁴⁸⁾

Nevertheless, Chen et al. ⁽⁴⁹⁾ provided a technique for estimating traffic flow in urban road networks. This approach focused on the estimate of traffic flow. Not only does the system make use of methods from the LSTM and SAE forecasting models, but it also makes use of adaptive spectral clustering.

Wang et al. ⁽⁵⁰⁾ developed the VSTGC model by analyzing historical data on traffic flow in order to forecast future changes in transportation safety. This was done in order to predict future changes. In this instance, we make use of a method known as deep learning.

An improved Support Vector Machine (SVM) model for traffic flow forecasting was the focus of an in-depth investigation that was carried out by Li et al. ⁽⁵¹⁾ with reference number 51. During the course of this study, the settings of the Support Vector Machine (SVM) were modified with the intention of improving the accuracy of forecasts about traffic flows. The authors conducted an in-depth analysis of the inadequacies of the machine-learning algorithms that are now in use and also discussed the implications of future developments in pedestrian counting frameworks. Following their research, the scientists came to the conclusion that SVR had the highest incidence of inaccurate classifications and had the ability to improve the accuracy of traffic flow projections. This work makes a contribution to the ongoing effort to improve models and techniques for estimating traffic flows, which is a permanent aim.

Khan et al. ⁽⁵²⁾ examined data on air pollution and traffic intensity in intelligent cities using an ensemble method known as bagging with the K-nearest neighbours (KNN) model. Performance for the bagging ensemble approach was found to be 30 % higher than that of the stacked ensemble method and other options. The evaluation metrics indicated that the KNN bagging model had promising results with the lowest Root Mean Square Error (RMSE) value of 1,12 among all the models. Besides, the Mean Absolute Error (MAE) of the best ensemble model was a respectable 0,55. Reduced Root Absolute Error (RAE) values were also obtained; the model that performed the best had the lowest value of all the models, 0,4.

Furthermore, Wang et al. ⁽⁵³⁾ introduced the Selective Ensemble Method (SEL), a special ensemble method for traffic flow stage identification. For increased prediction accuracy, the proposed model combined ensemble methods with supervised and unsupervised learning techniques. They specifically developed a method called Binary Particle Swarm Optimization-Random Subspace-Selective Ensemble Learning-Support Vector Machine (BPSO-RS-SVM-SEL). The discrete traffic flow states were created from traffic flow data using fuzzy C-means clustering. The testing results demonstrated that the proposed approach reached a higher maximum accuracy of 98,68 % than other existing techniques. This work is important to improve traffic control and management systems. It progresses the area of accurate status of traffic flow detection.

A new method called the Recurrent Mixture Density Network (RMDN) was presented by Chen et al. ⁽⁵⁴⁾ in their study it combines a recurrent neural network (RNN) with a mixture density network. Positive outcomes were seen when actual traffic flow data was used. The study highlights the difficulties of making short-term traffic flow projections while also discussing the possibilities of data-driven solutions to this problem. An LSTM

network and an MDN that parameterizes mixture distributions using the neural network's outputs make up the Recurrent MDN, which may capture contextual relationships and historical data in time series. Since the expected traffic flow indices are not given deterministic values but rather produced by the prediction model, the latter uses probability distributions to arrive at the forecasts. The accuracy (Mean Relative Error, MRE) of the recurring MDN in predicting the Travel Time Index (TTI) and Average Speed (AS) was higher than 0,92 total expected timestamps.⁽⁶⁾ The efficacy of the fuzzy prediction approach was shown by the recurrent MDN's significant performance improvement over the LSTM network. With a 1,28 % improvement in TTI forecast accuracy and a 0,8 % improvement in AS prediction accuracy for the 10-minute prediction test, the recurrent MDN outperformed the ARIMA model in short-term traffic flow prediction.⁽⁵⁴⁾

Bao et al.⁽⁵⁵⁾ suggested a better Deep Belief Network (DBN) that could provide reliable traffic predictions in all weather. Due to the nonlinear nature of traffic flow and the impact of external factors like weather and traffic incidents, support vector regression (SVR) was used to improve the standard DBN. The strategy relied on SVR to forecast traffic patterns during extreme weather. Another tool used for the training of layer-by-layer DBNs was a contrastive divergence (CD-k) approach.⁽²⁵⁾ In terms of computation time, prediction accuracy, and introduced error, the enhanced DBN's performance was compared to that of the ARIMA and standard neural network methodologies. Our goal was to improve the DBN's resilience during various time periods to reduce the traffic estimate error to less than 9 %.

But the ARIMA model isn't a viable substitute for traffic forecasting since it differs substantially, by 30 %, from the actual traffic flow when the weather is unfavorable. When compared to the improved DBN, the ARIMA model required less time to process.⁽⁵⁵⁾

For proactive traffic management and traveller information services in Intelligent Transportation Systems (ITS), Chen et al.⁽⁵⁶⁾ underlined the need of precise and timely traffic forecast. The authors put up a traffic-condition-awareness ensemble method that uses graph convolution to detect spatial patterns in traffic flow in order to overcome this problem. With an eye towards their performance under certain circumstances, they created a weight matrix to efficiently integrate the predictions from basic models. The suggested approach was fully assessed using an actual traffic dataset that was collected from the Caltrans Performance Measurement System (<https://dot.ca.gov/programs/traffic-operations/mpr/pems-source>).⁽⁵⁷⁾ We carried out comparative studies against single-model and ensemble techniques of competitive models. The writers also described possible future studies to maximize the parameter settings and network architecture. These facts imply that the study had an experimental research design.⁽⁵⁶⁾

Moreover, Mohanty et al.⁽⁵⁸⁾ suggested a model to predict a congestion score depending on the ratio of vehicle accumulation to trip completion rate based on a Long Short-Term Memory neural network architecture. The model is trained and tested in the work using input signals like demand, trip times, and congestion ratings. Its performance is compared with those of reference models like the Random Forest, Holt-Winters, and 1-NN models. When using an LSTM model is less suitable, the study additionally investigates parameter selection to enhance the model's predictions. The alternative plan's performance is just little better than that of the LSTM model.

Ranjan et al.⁽⁵⁹⁾ investigated how Seoul, South Korea, may use the deep learning network PredNet to forecast traffic congestion. Convolutional neural networks (CNN), Long Short-Term Memory (LSTM), and transpose convolutional neural networks (TCNN) are included into the proposed model's Feature Extraction Network (FEN) and Recurrent Network to gather temporal and spatial information from input images. Using a training dataset for Autoencoder and ConvLSTM models, accuracy, recall, and precision were evaluated across a number of different prediction horizons, including ten, thirty, and sixty minutes. The strategy that was recommended produced better results than the models that were applicable and reduced the MAPE from 3 % to 17.⁽⁶⁰⁾

Zheng et al.⁽⁷⁾ detailed a method for predicting traffic flows across city limits that integrates deep learning and embedding learning. Built within the model's architecture are all of the embedding components, CNNs, and LSTMs that contribute to the proposed model. In order to reduce and manage traffic congestion, the article also brings up the need of creating traffic flow projections for cities. The approach that has been strongly supported is called Deep and Embedding Learning. In addition to this, the embedding component is responsible for recording structured data such as the layout of the route and the weather conditions. While the CNN component is learning more about the two-dimensional traffic flow data, the LSTM component operates in the background to preserve earlier data. Combining the three models improves the accuracy of the traffic flow forecasts. 2-D traffic flow data is the model's input; it is generated by segmenting the data based on links rather than routes. Predicting accuracy of the model has been shown to be higher than that of existing methods.⁽⁶¹⁾

Furthermore, a Path Based Deep Learning framework for accurate traffic speed prediction in urban environments was provided in the work of Wang et al.⁽⁶²⁾ part of the approach included modelling each of the main routes in the road network using a bidirectional long short-term memory neural network (Bi-LSTM NN). The efficient collection of temporal information was achieved by layering many Bi-LSTM layers, and the spatial-temporal properties obtained were fed into a fully connected layer for traffic prediction. Performance of the

model was assessed in several prediction scenarios against many benchmark methodologies, and interpretability of the model was investigated by analyzing the hidden layer output properties. Adam optimized the model's parameters to reduce the training phase's mean squared error (MSE) between the anticipated and real traffic speeds. In comparison to other methods like KNN, ANN, CNN, and LSTM NN, the results showed a significant decrease in average MSE for different prediction scenarios. The long-term performance of the recommended architecture showed improved prediction accuracy for about 70 % of the segments, which strongly supported the suggested model's applicability to urban speed prediction tasks.

Comparison of ML and DL Models

In this part, we will discuss what we found in the studies we analyzed in our literature review. Table 2 shows us how DL models are better at picking up the spatial and temporal patterns and nonlinear dependencies of traffic flow. In some cases, ML models in Hybrid or Ensemble configurations also showed outstanding results. The ones that stood out are Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Autoencoders. Among these, LSTM was the most popular, and according to the results, it performed the best when dealing with traffic data like sequences and Spatiotemporal data.

In table 2, we have consolidated the performance of algorithms in different literature. Based on the evaluation metrics provided, the highest-performing models in different categories are as follows:

1. Accuracy:
 - LSTM + SAE: 97,7 %.⁽⁴⁹⁾
 - Selective Ensemble + DL: 98,68 %.⁽⁵³⁾
 - Factor Analysis + R.F.: accuracy Rate (A.R.): 98,9 %.⁽⁴²⁾
 - LSTM: 95,2 %.⁽⁴⁶⁾
 - R.F.: 92,7 %.⁽⁴⁶⁾
 - DBN: 90,8 %.⁽⁴⁶⁾
 - DAN: 89,4 %.⁽⁴⁶⁾
2. Mean Relative Error (MRE):
 - AE-LSTM: 0,065 MRE.⁽⁴⁷⁾
 - DBN + G.A. FR-CG: FR-CG outperforms G.D. LBFSGS.⁽³⁹⁾
 - RMDN: MRE > 0,92.⁽⁵⁴⁾
3. Mean Absolute Error:
 - LSTM + EEMD: 0,19 MAE.⁽⁴⁸⁾
 - Sequence-to-Sequence Model: 9,494741 MAE.⁽⁴⁰⁾
 - The VSTGC Model outperforms other methods.⁽⁵⁰⁾
 - Bagging Ensemble + KNN: outperforms other models.⁽⁵²⁾
4. Root Mean Square Error (RMSE):
 - LSTM + EEMD: 0,25 RMSE.⁽⁴⁸⁾
 - Sequence-to-Sequence Model: 12,294438 RMSE.⁽⁴⁰⁾
 - The VSTGC Model outperforms other methods.⁽⁵⁰⁾
 - Bagging Ensemble + KNN: outperforms other models.⁽⁵²⁾
5. Mean Absolute Percentage Error (MAPE):
 - LSTM + EEMD: 3,77 MAPE.⁽⁴⁸⁾
 - LSTM + Pre-processing: improved prediction.⁽⁶⁰⁾
 - The VSTGC Model outperforms other methods.⁽⁵⁰⁾
6. Detection Rate:
 - Factor Analysis + R.F.: detection Rate (D.R.): 95,65 %.⁽⁴²⁾
7. Other Evaluation metrics:
 - Random Forest + MLP: the combined model outperforms individual models.⁽⁴¹⁾
 - Improved SVM: reduces prediction error.⁽⁵¹⁾
 - Improved DBN + SVR: within 9 % traffic prediction error.⁽⁵⁵⁾
 - PredNet: outperforms ConvLSTM and Autoencoder.⁽⁵⁹⁾
 - DELA: shows improved prediction accuracy.⁽⁶¹⁾
 - Path Based DL: outperforms KNN, ANN, CNN.⁽⁶²⁾

Highest-performing models by metric:

- Accuracy: selective Ensemble + DL (98,68 %)
- MRE: AE-LSTM (0,065)
- MAE: LSTM + EEMD (0,19)
- RMSE: LSTM + EEMD (0,25)
- MAPE: LSTM + EEMD (3,77)
- Detection Rate: factor Analysis + R.F. (95,65 %)

As illustrated in the consolidated results, the models incorporating LSTM achieve each metric’s highest results.

Table 2. Performance metrics and Algorithms used in previous studies

Reference	Algorithms	Evaluation Metrics	Results
(46)	LSTM, DBN, DAN, R.F.	Precision, Recall, Accuracy, Error	Accuracy LSTM: 95,2 %, R.F.: 92,7 %, DBN: 90,8 %, DAN: 89,4 %
(47)	A.E. + LSTM, Various Algorithms	MRE, MAE	AE-LSTM: 0,065 MRE
(11)	AE-GRU, AE-LSTM	Accuracy, Training Time	AE-GRU has similar accuracy, shorter training time
(39)	DBN + G.A., FR-CG	MRE	FR-CG outperforms G.D., LBFGS
(48)	LSTM + EEMD	RMSE, MAE, MAPE	LSTM+EEMD: RMSE 0,25, MAE 0,19, MAPE 3,77
(40)	Sequence-to-Sequence Model	MAE, RMSE	MAE 9,494741, RMSE 12,294438
(49)	LSTM + SAE	Accuracy	Accuracy: LSTM + SAE: 97,7 %
(41)	Random Forest + MLP	Various Metrics	The combined model outperforms individual models
(42)	Factor Analysis + R.F.	Detection Rate (D.R.), Accuracy Rate (A.R.)	Proposed Methodology Results: D.R.: 95,65 %, AR: 98,9 %
(43)	Various ML Models	Varies	N.N., SVM effective in different scenarios
(50)	VSTGC Model	MAE, RMSE, MAPE	VSTGC outperforms other methods
(51)	Improved SVM	Classification Error Rate	Improved SVM reduces prediction error
(52)	Bagging Ensemble + KNN	RMSE, MAE, RAE	Bagging + KNN outperforms other models
(53)	Selective Ensemble + DL	Accuracy	The proposed model achieves 98,68 % accuracy
(54)	RMDN	MRE	RMDN achieves MRE > 0,92
(55)	Improved DBN + SVR	Traffic Prediction Error	Improved DBN within 9 % traffic prediction error
(59)	PredNet	Precision, Recall, Accuracy	PredNet outperforms ConvLSTM and Autoencoder
(60)	LSTM + Pre-processing	MAPE	Improved prediction using LSTM and preprocessing
(61)	DELA	Accuracy	DELA shows improved prediction accuracy
(62)	Path Based DL	MSE	Path Based DL outperforms KNN, ANN, CNN

Table 3 displays the different machine learning and deep learning algorithms used in the studies. The frequency of the algorithms used in more than one study shows which are better for traffic flow prediction. According to the table, the model most frequently used is LSTM, which has been used ten times. After that,

SVM, an ML model, was used five times, followed by RF, which was utilized four times. LSTM is one of the best algorithms for traffic flow predictions.

Table 3. LSTM, B: RF, C: DBN, D: DAN, E: FR-CG, G: SVM, H: Ensemble Method, I: GRU, J: VSTGC, K: CNN

Reference	A	B	C	D	E	F	G	H	I	J	K
(46)	X	X	X	X							
(47)	X			X							
(11)				X					X		
(39)			X		X						
(48)	X										
(40)											
(49)	X		X								
(41)		X									
(42)		X					X				
(43)		X					X				
(50)										X	
(51)							X				
(52)								X			
(53)							X	X			
(54)	X										
(55)			X				X				
(58)	X										
(59)	X										X
(60)	X										
(61)	X										X
(62)	X										X

The long Short-Term Memory (LSTM) recurrent neural network is resistant to the vanishing gradient issue and can learn and retain information over lengthy sequences. Because of this, it serves sequential data, such as traffic flow, well.⁽⁹⁾ With the passage of time, the interdependencies of traffic flow data may be accurately captured.⁽⁶³⁾ An improved and less parameter-intensive variant of Long Short-Term Memory (LSTM) is the Gated Recurrent Unit (GRU). A new “update gate” replaced the old “input gate” and “forget gate.” Along with making changes to them, it also merges the cell and hidden states.⁽⁶⁴⁾

During training, RF builds a large number of decision trees, which is an ensemble learning technique. As a consequence, it produces a class—which can be seen as the mean of the regression predictions or the median of the classification results—from the individual trees. When compared to LSTM or GRU, RF may not be able to capture the temporal relationships in traffic flow as effectively, despite its ability to handle large dimensional spaces and without requiring feature scaling.⁽⁹⁾

Due to their unique characteristics, these models outperform others under specific conditions. For instance, LSTM and GRU perform well with sequential data, while SVM and RF are more suited for high-dimensional data. Hybrid methods like LSTM-GRU can leverage the strengths of both LSTM and GRU.

Traffic flow factors (Datasets and Features)

In the research, we cited advanced traffic flow prediction using various datasets. Ramchandra et al.⁽⁴⁶⁾ may collect data by deliberately installing sensors in highly populated areas, recording traffic flow in real time, and combining variables like weather and road capacity. Similarly, Wei et al.⁽⁴⁷⁾ used autoencoders to extract relevant characteristics for prediction from historical traffic flow data collected from monitoring systems. Chen et al.⁽⁵⁴⁾ used a detailed strategy, most likely installing sensors on different road connections to collect information for forecasts made shortly, taking network structure and upstream/downstream linkages into account. Zhang et al.⁽⁶⁵⁾ used historical traffic data and genetic algorithms for their hybrid model to optimize deep belief networks.

To investigate the effects of weather on traffic, Hou et al.⁽⁴⁰⁾ used meteorological data from NOAA and traffic data from loop detectors, applying preprocessing techniques to improve the data's relevance. In parallel, studies like Chen et al.⁽¹¹⁾ involved the Caltrans Performance Measurement System (PeMS) to develop frameworks considering multiple factors in urban traffic prediction, other works, such as Bratsas et al.⁽⁴³⁾, utilized probe data from road networks. Ranjan et al.⁽⁵⁹⁾ harnessed accurate traffic data from specific urban areas for congestion predictions. The diversity of data sources—from sensors and loop detectors to publicly available datasets, underpinned these studies analytical prowess, fostering the advancement of predictive models and insights that contribute to more effective traffic management and urban planning.

Table 4. Dataset Description and Features

References	Data Type	Data Source	Balance	Features
(46)	Traffic flow parameters	Online website	Balanced	Zone type, weather condition, day, road capacity, vehicle types
(47)	Traffic flow and model config	Not specified	Balanced	Upstream and downstream traffic flow data
(11)	Traffic flow correlation	Not specified	Balanced	Traffic flow data
(48)	Traffic flow	Caltrans Performance Measurement System	Balanced	Traffic flow direction, Traffic flow data
(40)	Traffic and weather	Various sources	Balanced	Time, traffic data, and weather data
(49)	Traffic flow	University of Minnesota	Balanced	Time, sensor data, road sections
(41)	Toll and weather data	Not specified	Balanced	Time, lane, car model, car type, weather, holiday, weekend
(50)	Video clips	Dashboard cameras on vehicles	Balanced	Driving scenarios (safe, comparative dangerous, dangerous), weather conditions, types of roads, light conditions
(51)	Traffic flow data	California Department of Transportation Caltrans Performance Measurement System (PeMS)	Balanced	Traffic flow data from sensors
(53)	Traffic flow	California Department of Transportation	Balanced	Traffic flow, speed, occupancy rate, timestamps
(54)	Traffic flow	Sensors on road networks	Balanced	Traffic flow information
(59)	Traffic congestion	Traffic Optimization & Planning Information System (TOPIS) Website	Balanced	Traffic congestion levels, timestamps
(60)	Traffic speed	Gangnam-gu, Seoul	33 % missing	Traffic speed data
(61)	Traffic flow	KDD CUP 2017 dataset	Balanced	Travel time data
(62)	Traffic speed	Automatic vehicle identification (AVI) detectors in Xuancheng, China	Balanced	Speed data, time interval
(65)	Pavement condition survey	Caltrans Pavement Program	Balanced	Traffic flow data
(52)	Pollution and traffic intensity	City pulse Aarhus, Denmark	Balanced	Air pollutants (CO, SO ₂ , O ₃ , PM), vehicle count, timestamps

Challenges in Traffic Flow Prediction (Gaps and Limitations)

Despite the significant progress in traffic flow prediction using machine learning techniques, several limitations and gaps still need to be addressed in existing research. These limitations hinder the widespread adoption and practical implementation of machine learning models for real-world traffic management applications. This section provides an illustration of the fundamental deficiencies and gaps that were present in prior research. It serves as an example of where the field may move from below. Table 5 illustrates the limitations and gaps that were present in the studies that met the criteria for inclusion in our research.

Table 5. Gaps and Limitations of previous studies

Reference	Main Findings	Limitations and Gaps
(46)	LSTM model achieved 95,2 % accuracy in traffic flow prediction, outperforming other techniques.	The system is designed mainly for densely populated areas, limiting the study to not considering other scenarios, such as special occasions. Preprocessing is used in the process, but feature engineering and a vital feature selection process were neglected.
	AE-LSTM approach consistently outperformed other methods, emphasizing the potential for improving urban traffic conditions.	The study introduces an AE-LSTM prediction model for traffic flow that combines temporal and simple spatial patterns. However, it acknowledges a limitation in only considering basic spatial correlations, raising concerns about the model's ability to handle more complex spatial relationships and their impact on prediction accuracy.
	AE-GRU achieved comparable accuracy to AE-LSTM with shorter training time.	Because it concentrated on high-relevance connections and simplified spatial interactions, the study may have overlooked complicated spatial linkages in the road network. This is because the analysis focused on high-relevance links. Because of this, it could be more difficult for the model to successfully depict intricate and variable patterns of traffic flow dynamics.
	The hybrid DBN-GA model performed better using the FR-CG algorithm, suitable for models with more parameters.	A limitation of the research is that it does not take into account the geographical connection between stations and instead concentrates only on the prediction of highway traffic. The inability to take into consideration urban traffic scenarios takes into account the fact that there is a potential of varying degrees of accuracy in projections as well as the inherent complexity of urban traffic systems.
(48)	LSTM+EEMD achieved higher prediction performance for traffic flow, particularly with Caltrans PeMS data.	The study needs to delve into potential challenges or limitations related to the real-world implementation of the proposed hybrid approach, such as issues related to the availability and quality of the data required for denoising and the practical feasibility of deploying the model in dynamic traffic environments.
	The proposed framework incorporated weather conditions for traffic flow prediction and achieved specific performance metrics.	This study proposes a combined SAE and RBF framework for short-term traffic flow prediction. It highlights data processing processes such as time periodicity representation and weather parameter fusion. Additionally, the research gives a framework for predicting short-term traffic flow. It is necessary, however, to do further research on the challenges that are linked with these treatments. There should be a comprehensive discussion of the model's ability to scale and adapt to the swings in traffic and weather that occur in the actual world.
	A framework combining LSTM and SAE achieved 97,7 % average prediction accuracy in urban road networks.	There is a possibility that the study might have benefitted from a more comprehensive analysis of complex road network scenarios; nevertheless, the research's limited focus on urban road networks makes it difficult to generalize that research. There is a deficiency in the amount of debate focused on the challenges that may be brought about by the use of novel approaches, such as reinforcement learning.
	The ensemble method combining Random Forest and KNN improved traffic flow prediction accuracy.	It is possible that noise in the weather records might have an effect on the conclusions that the models generate, and the fact that the research only used toll data from a single site restricts its application. All of these are only some of the problems that were found in the study. It is possible that the accuracy of traffic flow forecasts might be improved by removing these limits and investigating a wider variety of other data sources.
	Recurrent Mixture Density Network (RMDN) outperformed LSTM and ARIMA models for traffic prediction.	The study introduces a novel approach that combines LSTM with mixture density networks. This method is called the recurrent mixture density network, and it is implemented in order to estimate traffic flows in the near future using this method.

VSTGC model with spatiotemporal graph-based convolution outperformed existing methods for traffic safety prediction.	Due to the fact that the study relied on historical data and assumed that trends would remain consistent, it is possible that it failed to take into account rapid changes in the external environment. The model's capability to deal with complicated traffic circumstances in the actual world might be improved, even with the performance increases that have been included. Additionally, the study does not consider the difficulties encountered in the real world while attempting to apply the VSTGC model in highway networks.
Improved SVM model for traffic flow prediction emphasized better accuracy and potential enhancements.	The study introduces diverse pattern classifiers to improve vehicle detection and proposes an optimized SVM-based traffic flow prediction model. However, the limitations include the need for a more comprehensive evaluation of real-world challenges, the absence of an in-depth rationale for classifier selection, and the omission of potential implications when implementing an enhanced pedestrian counting method.
Bagging ensemble method with KNN improved air pollution prediction based on traffic intensity data.	The purpose of this research is to offer an ensemble strategy that may enhance the forecast of traffic flow. Collecting information on pollution and traffic from a variety of sources allows for the implementation of this strategy. It is important for you to be aware, however, that this technique does not come without some shortcomings. Not only does it not perform well in different locations or seasons of the year, but there is also a possibility of overfitting and a deterioration in interpretability due to the large number of base classifiers. It is necessary to do more study that makes use of satellite data to investigate the validity of the model for understanding the impacts of various seasons and other sorts of cities.
Selective Ensemble Learning (SEL) model with BPSO-RS-SVM for traffic flow state identification achieved high accuracy.	The study presents an improved SEL model for highway traffic flow state identification using clustering and ensemble techniques. However, the limitations include the reliance on specific datasets and the assumption that the proposed model will perform well across different conditions and regions. Additionally, while the model's accuracy is highlighted, potential challenges related to real-world implementation and considering various external factors still need to be thoroughly addressed.
The predNet model combined CNN, LSTM, and TCNN outperformed ConvLSTM and Autoencoder models for congestion prediction.	The study uses image data to present a hybrid deep-learning model for predicting city-wide traffic congestion. The model achieves promising accuracy and outperforms other algorithms. However, the study acknowledges the model's computational inefficiency. It suggests improving efficiency and incorporating external factors like weather information for enhanced predictions.
LSTM-based approach with preprocessing enhanced congestion prediction accuracy under various conditions.	The study proposes an LSTM-based traffic congestion prediction method. Still, it acknowledges limitations related to outliers and missing data, which can impact prediction accuracy. While the model performs better in suburban areas with less external interference, it aims to improve accuracy in low-speed regions and urban areas in the future.
DELA model integrating embedding, CNN, and LSTM improved urban traffic flow prediction.	The study proposes a Deep and Embedding Learning Approach (DELA) for accurate traffic flow prediction by combining an embedding component, CNN, and LSTM. Although DELA demonstrates superior prediction accuracy compared to existing methods, the study acknowledges the limitation of needing more explanatory power in deep learning models. It outlines plans for future enhancements without addressing specific challenges or constraints.
Path-based Deep Learning framework with Bi-LSTM NN demonstrated improved speed prediction accuracy.	The study presents an interpretable deep learning model for traffic speed forecasting, with contributions including defining critical paths and utilizing Bi-LSTM NN for spatial-temporal modeling. However, the study does not explicitly discuss the limitations of the proposed model. While it suggests future research directions related to model interpretability and critical path selection, it must address potential challenges or constraints in these areas.

In addition to the shortcomings and exceptions listed in table 5, a major shortcoming was that the majority of these studies failed to take into account the impact of feature engineering on the accuracy of traffic flow predictions, the features' effects on traffic flow, or the models' accuracy in making such predictions. The prediction process in supervised learning relies heavily on feature significance. By understanding which information directly impacts the result, we can build a more accurate and reliable model.

New developments in traffic flow prediction have shown how crucial it is to correctly detect spatiotemporal patterns and to use data-driven feature extraction and selection approaches to find complicated correlations.^(66,67) One study found that LSTM, a heuristic optimization algorithm, achieved Mean Absolute Errors (MAEs) of 4,25 and 4,38 minutes, respectively, for ATFM delay regression prediction using feature extraction modules (such as CNN, TCN, and attention modules).⁽⁶⁶⁾ In addition, present approaches for predicting traffic flows have an accuracy gap of 5 %-16 % when using a Community-Based Dandelion Algorithm for feature selection.⁽⁶⁶⁾ Data decomposition approaches, such as wavelet decomposition and empirical mode decomposition (EMD), improve the predictive performance of time series prediction models compared to standalone models.^(68,69,70)

Standard Traffic Flow ML and DL Prediction Algorithms Implemented

To gain a comprehensive understanding of traffic flow, the researchers responsible for those studies used a wide variety of deep learning and machine learning approaches. These algorithms aim to improve urban mobility and traffic management by developing prediction models that use historical data on traffic and other relevant elements. The following are examples of well-known algorithms and techniques for machine learning and deep learning:

Long Short-Term Memory LSTM

Recurrent neural networks (RNNs) take into account past contexts to make decisions based on the present input via feedback loops that allow activations to travel from one layer to another. One issue with RNNs is the vanishing gradient problem, which refers to the diminishing effect of activations from previous time steps on the network's ability to store information over time. This is further worse by the fact that activation functions like the sigmoid and hyperbolic tangent limit the range of possible input values. So, during training, error signals don't propagate very well. The inability of RNNs to capture long-term relationships is a result of this constraint. So, to fix these issues, we need more complicated solutions, such as LSTM blocks.⁽⁷¹⁾

LSTM (Long Short-Term Memory) blocks were developed in Hochreiter et al.⁽⁷²⁾ to address the problem of information loss over extended periods in RNNs. LSTM blocks, in figure 14, in contrast to conventional hidden units, include memory cells (c) and three gates (i, f, and o) to control the information flow. The cell state (ct) and output (ht) store and carry pertinent memories at each time step and layer. Using LSTM blocks, which incorporate gated linear relationships between memory cell states and peephole connections (U), allows for preserving significant information across long periods.⁽⁷¹⁾

The equations defining LSTM are shown in figure 5:

$$\begin{aligned} \mathbf{x}_t^l &= [\mathbf{h}_t^{l-1}; \mathbf{h}_{t-1}^l] \\ \mathbf{i}_t^l &= \sigma(\mathbf{W}_i^l \mathbf{x}_t^l + \mathbf{U}_i^l \mathbf{c}_{t-1}^l + \mathbf{b}_i^l) \\ \mathbf{f}_t^l &= \sigma(\mathbf{W}_f^l \mathbf{x}_t^l + \mathbf{U}_f^l \mathbf{c}_{t-1}^l + \mathbf{b}_f^l) \\ \hat{\mathbf{c}}_t^l &= \tanh(\mathbf{W}_c^l \mathbf{x}_t^l + \mathbf{U}_c^l \mathbf{c}_{t-1}^l + \mathbf{b}_c^l) \\ \mathbf{c}_t^l &= \mathbf{f}_t^l \odot \mathbf{c}_{t-1}^l + \mathbf{i}_t^l \odot \hat{\mathbf{c}}_t^l \\ \mathbf{o}_t^l &= \sigma(\mathbf{W}_o^l \mathbf{x}_t^l + \mathbf{U}_o^l \mathbf{c}_t^l + \mathbf{b}_o^l) \\ \mathbf{h}_t^l &= \mathbf{W}_{proj}^l (\mathbf{o}_t^l \odot \tanh(\mathbf{c}_t^l)) \end{aligned}$$

Figure 5. LSTM formulation⁽⁷¹⁾

The LSTM Model has proven very promising in Traffic flow predictions. Its accuracies are much higher than those of standard neural networks. Moreover, it has also shown increased performance improvements when used in a hybrid model with other ML and DL models.

Gated Recurrent Unit GRU

It is imperative to capture hidden representations of sequences when there are long-term relationships between the input signal x and the output signal y when using gated units like LSTM and GRU. These units are made to learn and store knowledge effectively over time to recognize and represent complicated sequential patterns. Click or tap here to enter text.⁽⁶⁶⁾ Figure 6 illustrates a unit formulation of the GRE model.

$$\begin{aligned} \tilde{x}(k) &= f_r [\text{net}_r(k)] \odot \hat{x}(k-1) \\ \hat{x}(k) &= f_z [\text{net}_z(k)] \odot \hat{x}(k-1) \\ &\quad + (1 - f_z [\text{net}_z(k)]) \odot f_c [\text{net}_c(k)] \end{aligned}$$

Figure 6. GRU Unit Formulation⁽⁷³⁾

The GRU (Gated Recurrent Unit) is a condensed form of the LSTM (Long Short-Term Memory) with no output gate. The GRU lacks the output gate, which increases its computational efficiency and structural simplicity in contrast to LSTM, which contains input and output gates to regulate the flow of information. GRU, unlike LSTM, cannot selectively disclose the hidden state at each time step because of the absence of an output gate. For sequential modeling jobs, the GRU can still record and save pertinent information in the hidden state.⁽⁷³⁾

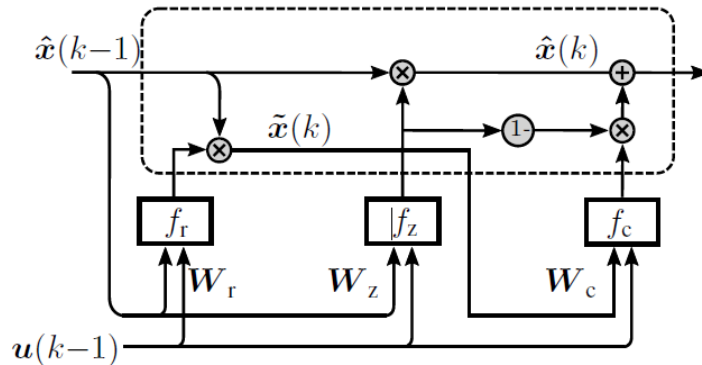


Figure 7. GRU Unit⁽⁷³⁾

Convolutional Neural Networks CNN

Convolutional neural networks (CNNs) are explicitly used for image classification tasks where the need for complex preprocessing is minimized, and the focus is on directly classifying raw images.⁽⁹⁾ CNN was first developed in 1988 by Fukushima.⁽⁷⁴⁾

In figure 8, it shows the structure of CNN. It shows a layer structure with multi-filters to allow the system for parallel processing. When dealing with enormous amounts of data, it is crucial to use deeper neural network topologies with more hidden units. The large-scale inputs may be gradually reduced by creating such a framework, which enables the underlying traits to emerge more clearly throughout the learning process. Three parts comprise a CNN architecture: the input, hidden, and output layers. The phrase “hidden layers” refers to the intermediate levels of any feedforward network, and the number of hidden layers varies depending on the kind of network design. Convolutions are carried out in hidden layers, comprising dot products between the input matrix and the convolution kernel. The feature maps produced by each convolutional layer are utilized as input by the successive layers.⁽⁷⁵⁾

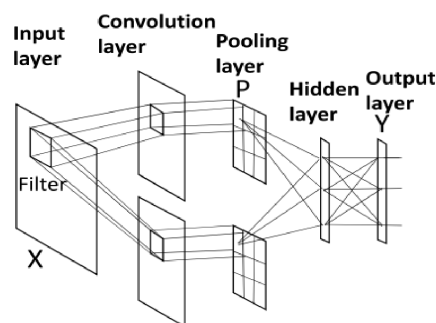


Figure 8. CNN Structure⁽⁷⁶⁾

In figure 9 the two essential parts of CNN are illustrated: feature extractor and Classifier. The network is structured with convolutional and max-pooling layers for feature extraction. There are convolutional layers (even-numbered) and max-pooling layers (odd-numbered) at the bottom and middle levels of the network. Max-pooling minimizes the dimensionality of the features, whereas convolutions extract features from the input nodes. These layers’ outputs are arranged into a 2D plane called a feature map, commonly made by

fusing different planes from the preceding levels. Every node in the convolutional layer collects features from the prior layer's linked sections.⁽⁹⁾

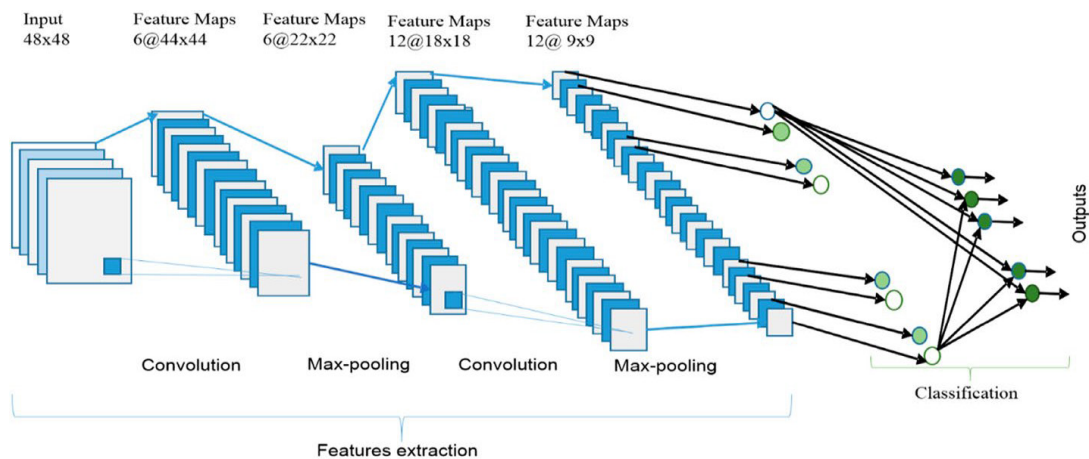


Figure 9. CNN Feature collection model⁽⁷⁷⁾

Support Vector Machine

The Support Vector Machine program allows users to perform pattern recognition and regression estimation on some given data. SVM is a new machine learning method based on statistical learning theory.⁽⁷⁸⁾ Figure 10 shows the optimal separation place of the SVM model.

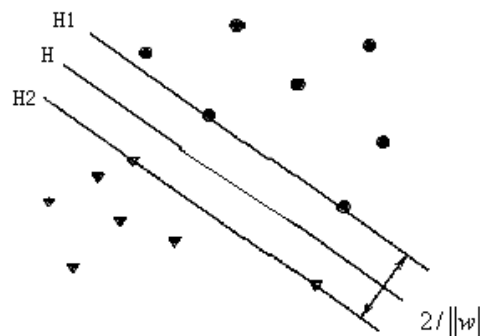


Figure 10. SVM Optimal separating surface⁽⁷⁸⁾

The fundamental idea behind SVM is margin calculation, which aims to maximize the margin-to-class distance to produce a hyperplane that divides two classes. SVM can manage scenarios with misclassifications using a soft margin and handle linearly separable data and issues with numerous features because its complexity is independent of the number of features.⁽⁹⁾

Random Forest

Random forest is a machine learning model commonly used by researchers like Ramchandra et al.⁽⁴⁶⁾ to combine the outputs of multiple decision trees to reach a single result. Xu et al.⁽⁴¹⁾, Jiang et al.⁽⁴²⁾ and Bratsas et al.⁽⁴³⁾, have used R.F. in their studies in different settings, and it is notable that in the research findings of Ramchandra et al.⁽⁴⁶⁾, R.F. was the second-best performing algorithm. R.F. is often very accurate in its predictions, and it doesn't require feature scaling or categorical feature encoding and requires minimal feature tuning.

R.F. has been effectively used in traffic flow predictions. The model is trained based on historical and real-time data, which can be collected from various sources such as cameras, sensors, GPS, social media, etc. Random Forest can predict traffic using previous data and store the expected values. One of the main advantages of R.F. is that it's highly accurate and takes less time.⁽⁸⁰⁾

Ensemble Method

Several studies utilized ensemble methods that combine multiple prediction models to achieve higher accuracy. Ensemble approaches, such as stacking and bagging, use the combined capabilities of many algorithms to improve overall prediction performance.

Machine learning is becoming a prominent area of artificial intelligence with many applications because of the abundance of data sources. Many algorithms are used in machine learning, each tested for scalability and performance. Among the noteworthy approaches in this field are deep learning and ensemble learning. Deep learning can tackle complex problems and extract features from unstructured data. Still, it requires a lot of work and expert adjustment during training, which raises the possibility of overfitting issues.

The effectiveness of ensemble learning in improving the performance of predictive models, as seen in figure 11, has attracted a great deal of attention and given rise to a growing body of research in various fields. This part uses the extensive ‘Scopus’ database to examine this tendency. This analysis focuses on the period from 2014 to 2021. It includes papers on “Ensemble Learning” and “Ensemble Deep Learning,” considering abstracts, keywords, and titles. The findings demonstrate a notable increase in papers published in the “Ensemble Learning” field, with 25 262 publications throughout the designated period.

Moreover, with an approximate total of 16 782 records, the field of computer sciences is the leader in article volume. Simultaneously, there is a clear interest in “Ensemble Deep Learning,” as seen by the 6 173 articles. The number of articles in the computer sciences sector continues to be high, indicating the growing interest in and importance of ensemble learning approaches in modern research.⁽⁸¹⁾

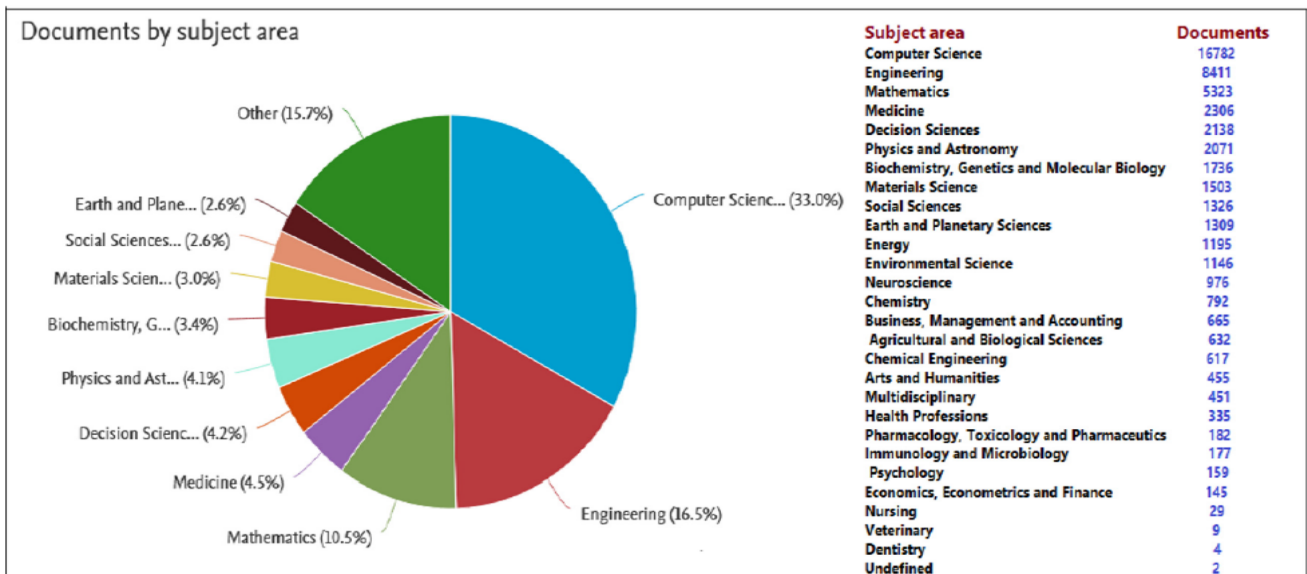


Figure 11. Ensemble learning used in different fields from 2014 - 2021⁽⁸¹⁾

Moreover, within the context of traffic flow prediction, it has been observed that the application of Ensemble methods yields exceptional outcomes when dealing with traffic-related datasets. Notably, these investigations have demonstrated that using the Ensemble method has led to a remarkable accuracy level of up to 98,68 % in predicting traffic flow patterns.

All of the ML and DL algorithms have their advantages when we are utilizing them for traffic flow predictions. ARIMA effectively captures linear temporal dependencies in traffic flow data, such as daily or weekly patterns. Random Forest can capture nonlinear relationships and interactions between features, making it suitable for complex traffic flow patterns. LSTM and GRU are particularly suited for longer-term dependencies in time series data. They are well-suited for capturing patterns that extend over hours or even days. CNNs help capture spatial patterns along a single dimension, such as road segments, where specific patterns may repeat across different segments. Ensemble Method Combining multiple models or techniques allows leveraging the strengths of each approach. For example, combining time series models with neural networks can capture linear and nonlinear dependencies.

Evaluation Metrics

The performance of the machine learning and deep learning models are analyzed using different evaluation metrics. These evaluation metrics are Accuracy, Precision, Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The equations for these metrics are as follows:

$$\text{Accuracy} = \frac{\text{Number of Correct predictions}}{\text{Total Number of Predictions}}$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True positives} + \text{False Positives}}$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_i)^2$$

$$\text{RMSE} = \sqrt{\text{MSE}}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_i - \bar{Y}_i|$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \bar{Y}_i}{Y_i} \right| \times 100$$

The equations given above show the evaluation metrics formula. In MSE, RMSE, MAE, and MAPE, the Y_i is the actual value and \bar{Y}_i is the predicted value with n being the total number of data points. Accuracy is calculated based on the number of correct predictions divided by the total number of predictions. Precision depends on the confusion matrix, which shows the total number of true positives, false positives, and false negatives. Apart from these, Mean Relative Error (MRE) was also used in many studies. MRE is the same as MAPE, except MRE is usually expressed as a fraction. In contrast, MAPE is expressed as a percentage. Overall, all these evaluation metrics are used to analyze the performance of models and compare the results when utilized for traffic flow predictions.

CONCLUSIONS

In conclusion, in this review, we examined how sophisticated AI methods like ML and DL may help better traffic flow management through accurate predictions. A total of 70 articles were selected and reviewed in this process. We discovered that various techniques, including LSTM and other models, effectively forecast the traffic flow. This is significant because it contributes to developing more innovative and livable cities. The study also showed how effective these advanced intelligent models are in identifying traffic patterns compared to previous, more straightforward techniques such as ARIMA. Some exciting methods that relied on long short-term memory (LSTM) were even more accurate than the conventional models. We were also drawn to hybrid models, a blend of many approaches, because they have the potential to be highly successful in handling complex traffic scenarios.

However, it's essential to know that these intelligent methods have some limits, and we need to be aware of them, which also brought our attention to the studies lacking the consideration of feature engineering and importance. The study suggests that future research should focus on datasets and integrate feature engineering methods to make the prediction process even better and more reliable. Also, having different information, like traffic, weather, and pollution data, can help these models work even more intelligently. Ultimately, we hope this study helps researchers, people who make decisions about cities, and those who work on traffic issues. We wish to contribute to literature and future studies by serving as a resource for other academics and researchers.

BIBLIOGRAPHIC REFERENCES

1. A. A. Kashyap, S. Raviraj, A. Devarakonda, S. R. Nayak K, S. K. V. and S. J. Bhat. Traffic flow prediction models - A review of deep learning techniques, Cogent Eng, vol. 9, no. 1, pp. 1-19. doi:10.1080/23311916.2021.2010510.
2. B. Medina-Salgado, E. Sánchez-DelaCruz, P. Pozos-Parra and J. E. Sierra. Urban traffic flow prediction techniques: A review, Sustainable Computing: Informatics and Systems, vol. 35, p. 100739, pp. 1-10. doi: 10.1016/j.suscom.2022.100739.
3. Y. Yu, W. Sun, J. Liu and C. Zhang. Traffic flow prediction based on depthwise separable convolution fusion network, J Big Data, vol. 9, no. 1, pp. 83, pp. 1-15. doi: 10.1186/s40537-022-00637-9.
4. M. Shaygan, C. Meese, W. Li, X. (George) Zhao and M. Nejad. Traffic prediction using artificial intelligence: Review of recent advances and emerging opportunities, Transp Res Part C Emerg Technol, vol. 145, p. 103921, pp. 1-21. doi: 10.1016/j.trc.2022.103921.
5. N. A. M. Razali, N. Shamsaimon, K. K. Ishak, S. Ramli, M. F. M. Amran and S. Sukardi. Gap, techniques and evaluation: traffic flow prediction using machine learning and deep learning, J Big Data, vol. 8, no. 1, pp. 1-22. doi: 10.1186/s40537-021-00542-7.

6. A. Navarro-Espinoza et al. Citation: Prediction for Smart Traffic Lights Using Machine Learning Algorithms. Traffic Flow Prediction for Smart Traffic Lights Using Machine Learning Algorithms, pp. 1-10. doi: 10.3390/technologies.
7. F. Wang, J. Xu, C. Liu, R. Zhou and P. Zhao. On prediction of traffic flows in smart cities: a multitask deep learning based approach, World Wide Web, vol. 24, no. 3, pp. 805-823. doi: 10.1007/s11280-021-00877-4.
8. S. E. Bibri and J. Krogstie. Smart sustainable cities of the future: An extensive interdisciplinary literature review, Sustain Cities Soc, vol. 31, pp. 183-212. doi: 10.1016/j.scs.2017.02.016.
9. S. A. Sayed, Y. Abdel-Hamid and H. A. Hefny. Artificial intelligence-based traffic flow prediction: a comprehensive review, Journal of Electrical Systems and Information Technology, vol. 10, no. 1, pp. 1-24. doi: 10.1186/s43067-023-00081-6.
10. R. Lakhani, D. Dalal, V. Dantwala and K. Bhowmick. Traffic Congestion Prediction Using Deep Learning Models, pp. 779-788. doi: 10.1007/978-981-19-9638-2_67.
11. D. Chen, H. Wang and M. Zhong. A Short-term Traffic Flow Prediction Model Based on AutoEncoder and GRU, in 2020 12th International Conference on Advanced Computational Intelligence (ICACI), IEEE, Aug. 2020, pp. 550-557. doi: 10.1109/ICACI49185.2020.9177506.
12. Nipun Waas. Time Series Forecasting: Predicting mobile traffic with LSTM, Medium.com.
13. Shanthababu Pandian. Time Series Analysis and Forecasting | Data-Driven Insights (Updated 2024), Analyticsvidhya.com.
14. V. Swathi, S. Yerraboina, G. Mallikarjun and M. JhansiRani. Traffic Prediction for Intelligent Transportation System Using Machine Learning, in 2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), IEEE, Apr. 2022, pp. 1-4. doi: 10.1109/ICAECT54875.2022.9807652.
15. N. Shamsaimon et al. Traffic Flow Prediction Using Long-Short Term Memory Technique for Connected Vehicles in Smart Cities, pp. 411-422. doi: 10.1007/978-3-030-90235-3_36.
16. W. Wei, H. Wu and H. Ma. An AutoEncoder and LSTM-Based Traffic Flow Prediction Method, Sensors, vol. 19, no. 13, pp. 2946. doi: 10.3390/s19132946.
17. E. P. Lens Shiang, W.-C. Chien, C.-F. Lai and H.-C. Chao. Gated Recurrent Unit Network-based Cellular Traffic Prediction, in 2020 International Conference on Information Networking (ICOIN), IEEE, Jan. 2020, pp. 471-476. doi: 10.1109/ICOIN48656.2020.9016439.
18. K. O. Ellefsen, C. P. Martin and J. Torresen. How do Mixture Density RNNs Predict the Future?, Jan. 2019, [Online]. Available: <http://arxiv.org/abs/1901.07859>
19. M. Chen, R. Chen, F. Cai, W. Li, N. Guo and G. Li. Short-Term Traffic Flow Prediction with Recurrent Mixture Density Network, Math Probl Eng, vol. 2021, pp. 1-9. doi: 10.1155/2021/6393951.
20. Google. Google Scholar Search System, <https://scholar.google.com/>.
21. L. Li, R. Jiang, Z. He, X. (Michael) Chen and X. Zhou. Trajectory data-based traffic flow studies: A revisit, Transportation Research Part C: Emerging Technologies, vol. 114, pp. 225-240. doi: 10.1016/j.trc.2020.02.016.
22. U.S Dept of Transportation. Next Generation Simulation (NGSIM) Vehicle Trajectories and Supporting Data, Data.gov.
23. L. Li, R. Jiang, Z. He, X. (Michael) Chen and X. Zhou. Trajectory data-based traffic flow studies: A revisit, Transp Res Part C Emerg Technol, vol. 114, pp. 225-240. doi: 10.1016/j.trc.2020.02.016.
24. U.S Dept of Transportation. Next Generation Simulation (NGSIM) Open Data, Transportation.gov.

25. M. Fellendorf and P. Vortisch. Microscopic Traffic Flow Simulator VISSIM, pp. 63-93. doi: 10.1007/978-1-4419-6142-6_2.
26. U. Sah. Traffic flow -Study and Modelling Levels., ResearchGate, 2019.
27. R. Mohan and G. Ramadurai. State-of-the art of macroscopic traffic flow modelling, *Int J Adv Eng Sci Appl Math*, vol. 5, no. 2-3, pp. 158-176. doi: 10.1007/s12572-013-0087-1.
28. J. Zhang, X. Wang, J. Wang and J. Wang. Decision-Making Model of Lane-Change Behavior Based on Integrated Cognitive Vehicle Cluster Situations, pp. 77-94. doi: 10.1007/978-981-10-3551-7_6.
29. W. Zhang, S. Wei, C. Wang and M. Qiu. Asymmetric Behaviour and Traffic Flow Characteristics of Expressway Merging Area in China, *Promet*, vol. 35, no. 1, pp. 12-26. doi: 10.7307/ptt.v35i1.4200.
30. H. Deng and H. M. Zhang. On Traffic Relaxation, Anticipation, and Hysteresis, *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2491, no. 1, pp. 90-97. doi: 10.3141/2491-10.
31. D. Li, L. Deng, Z. Cai, B. Franks and X. Yao. Notice of Retraction: Intelligent Transportation System in Macao Based on Deep Self-Coding Learning, *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3253-3260, Jul. 01, 2018. doi: 10.1109/TII.2018.2810291.
32. O. Messoussi et al. Vehicle Detection and Tracking from Surveillance Cameras in Urban Scenes, pp. 191-202. doi: 10.1007/978-3-030-90436-4_15.
33. A. Ghosh, Md. S. Sabuj, H. H. Sonet, S. Shatabda and D. Md. Farid. An Adaptive Video-based Vehicle Detection, Classification, Counting, and Speed-measurement System for Real-time Traffic Data Collection, in 2019 IEEE Region 10 Symposium (TENSYP), IEEE, Jun. 2019, pp. 541-546. doi: 10.1109/TENSYP46218.2019.8971196.
34. A. Gholamhosseinian and J. Seitz. Vehicle Classification in Intelligent Transport Systems: An Overview, Methods and Software Perspective, *IEEE Open Journal of Intelligent Transportation Systems*, vol. 2, pp. 173-194, 2021. doi: 10.1109/OJITS.2021.3096756.
35. V. S. R. Kosuru, A. K. Venkitaraman, V. D. Chaudhari, N. Garg, A. Rao and A. Deepak. Automatic Identification of Vehicles in Traffic using Smart Cameras, in 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), IEEE, Dec. 2022, pp. 1009-1014. doi: 10.1109/IC3I56241.2022.10072979.
36. C. An, X. Guo, R. Hong, Z. Lu and J. Xia. Lane-Based Traffic Arrival Pattern Estimation Using License Plate Recognition Data, *IEEE Intelligent Transportation Systems Magazine*, vol. 14, no. 4, pp. 133-144, Jul. 2022. doi: 10.1109/MITS.2021.3051489.
37. W. Yao, M. Zhang, S. Jin and D. Ma. Understanding vehicles commuting pattern based on license plate recognition data, *Transp Res Part C Emerg Technol*, vol. 128, p. 103142, Jul. 2021. doi: 10.1016/j.trc.2021.103142.
38. Z. Xing, M. Huang and D. Peng. Overview of machine learning-based traffic flow prediction, **Digital Transportation and Safety**, vol. 2, no. 3, pp. 164-175, 2023. doi: 10.48130/DTS-2023-0013.
39. Y. Zhang and G. Huang. Traffic flow prediction model based on deep belief network and genetic algorithm, in **IET Intelligent Transport Systems**, Institution of Engineering and Technology, Aug. 2018, pp. 533-541. doi: 10.1049/iet-its.2017.0199.
40. Y. Hou, Z. Deng and H. Cui. Short-Term Traffic Flow Prediction with Weather Conditions: Based on Deep Learning Algorithms and Data Fusion, **Complexity**, vol. 2021, 2021. doi: 10.1155/2021/6662959.
41. D. Xu and Y. Shi. A combined model of random forest and multilayer perceptron to forecast expressway traffic flow, in **2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC)**, IEEE, Jul. 2017, pp. 448-451. doi: <https://doi.org/10.1109/ICEIEC.2017.0024>.
42. H. Jiang and H. Deng. Traffic incident detection method based on factor analysis and weighted random forest, **IEEE Access**, vol. 8, pp. 168394-168404, 2020. doi: 10.1109/ACCESS.2020.3023961.

43. C. Bratsas, K. Koupidis, J.-M. Salanova, K. Giannakopoulos, A. Kaloudis and G. Aifadopoulou. A Comparison of Machine Learning Methods for the Prediction of Traffic Speed in Urban Places, **Sustainability**, vol. 12, no. 1, p. 142, Dec. 2019. doi: 10.3390/su12010142.
44. Q. Zhou, N. Chen and S. Lin. FASTNN: A Deep Learning Approach for Traffic Flow Prediction Considering Spatiotemporal Features, **Sensors**, vol. 22, no. 18, p. 6921, Sep. 2022. doi: 10.3390/s22186921.
45. S. Goswami and A. Kumar. Traffic Flow Prediction Using Deep Learning Techniques, 2022, pp. 198-213. doi: 10.1007/978-3-031-10551-7_15.
46. N. Reshma Ramchandra and C. Rajabhushanam. Machine learning algorithms performance evaluation in traffic flow prediction, in **Materials Today: Proceedings**, Elsevier Ltd, 2021, pp. 1046-1050. doi: 10.1016/j.matpr.2021.07.087.
47. W. Wei, H. Wu and H. Ma. An autoencoder and LSTM-based traffic flow prediction method, *Sensors (Switzerland)*, vol. 19, no. 13, Jul. 2019. doi: <https://doi.org/10.3390/s19132946>.
48. X. Chen et al. Traffic flow prediction by an ensemble framework with data denoising and deep learning model, *Physica A: Statistical Mechanics and its Applications*, vol. 565, Mar. 2021. doi: 10.1016/j.physa.2020.125574.
49. C. Chen, Z. Liu, S. Wan, J. Luan and Q. Pei. Traffic Flow Prediction Based on Deep Learning in Internet of Vehicles, *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3776-3789, Jun. 2021. doi: 10.1109/TITS.2020.3025856.
50. J. Wang and Q. Chen. A traffic prediction model based on multiple factors, *Journal of Supercomputing*, vol. 77, no. 3, pp. 2928-2960, Mar. 2021. doi: 10.1007/s11227-020-03373-0.
51. C. Li and P. Xu. Application on traffic flow prediction of machine learning in intelligent transportation, *Neural Comput Appl*, vol. 33, no. 2, pp. 613-624, Jan. 2021. doi: 10.1007/s00521-020-05002-6.
52. N. U. Khan, M. A. Shah, C. Maple, E. Ahmed and N. Asghar. Traffic Flow Prediction: An Intelligent Scheme for Forecasting Traffic Flow Using Air Pollution Data in Smart Cities with Bagging Ensemble, *Sustainability (Switzerland)*, vol. 14, no. 7, Apr. 2022. doi: 10.3390/su14074164.
53. Z. Wang, R. Chu, M. Zhang, X. Wang and S. Luan. An Improved Selective Ensemble Learning Method for Highway Traffic Flow State Identification, *IEEE Access*, vol. 8, pp. 212623-212634, 2020. doi: 10.1109/ACCESS.2020.3038801.
54. M. Chen, R. Chen, F. Cai, W. Li, N. Guo and G. Li. Short-Term Traffic Flow Prediction with Recurrent Mixture Density Network, *Math Probl Eng*, vol. 2021, 2021. doi: 10.1155/2021/6393951.
55. X. Bao, D. Jiang, X. Yang, and H. Wang. An improved deep belief network for traffic prediction considering weather factors, *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 413-420, Feb. 2021. doi: 10.1016/j.aej.2020.09.003.
56. Y. Chen, Y. Lv, P. Ye, and F. Zhu. Traffic-Condition-Awareness Ensemble Learning for Traffic Flow Prediction, *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 582-587. doi: 10.1016/j.ifacol.2021.04.146.
57. Caltrans. Caltrans Performance Measurement Systems, Data.gov. Available: <https://pems.dot.ca.gov/>.
58. S. Mohanty, A. Pozdnukhov, and M. Cassidy. Region-wide congestion prediction and control using deep learning, *Transp Res Part C Emerg Technol*, vol. 116, Jul. 2020. doi: 10.1016/j.trc.2020.102624.
59. N. Ranjan, S. Bhandari, H. P. Zhao, H. Kim, and P. Khan. City-wide traffic congestion prediction based on CNN, LSTM and transpose CNN, *IEEE Access*, vol. 8, pp. 81606-81620, 2020. doi: 10.1109/ACCESS.2020.2991462.
60. D. H. Shin, K. Chung, and R. C. Park. Prediction of Traffic Congestion Based on LSTM through Correction of Missing Temporal and Spatial Data, *IEEE Access*, vol. 8, pp. 150784-150796, 2020. doi: 10.1109/ACCESS.2020.3016469.

61. Z. Zheng, Y. Yang, J. Liu, H. N. Dai, and Y. Zhang. Deep and Embedded Learning Approach for Traffic Flow Prediction in Urban Informatics, *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3927-3939, Oct. 2019. doi: 10.1109/TITS.2019.2909904.
62. J. Wang, R. Chen, and Z. He. Traffic speed prediction for urban transportation network: A path based deep learning approach, *Transp Res Part C Emerg Technol*, vol. 100, pp. 372-385, Mar. 2019. doi: 10.1016/j.trc.2019.02.002.
63. H.-C. Jang and C.-A. Chen. Urban Traffic Flow Prediction Using LSTM and GRU, in *2023 IEEE 5th Eurasia Conference on Biomedical Engineering*, Basel Switzerland: MDPI, Jan. 2024, p. 86. doi: 10.3390/engproc2023055086.
64. M. Yurtsever. Unemployment rate forecasting: LSTM-GRU hybrid approach, *J Labour Mark Res*, vol. 57, no. 1, p. 18, Jun. 2023. doi: 10.1186/s12651-023-00345-8.
65. Z. Wang and T. Pyle. Implementing a pavement management system: The Caltrans experience, *International Journal of Transportation Science and Technology*, vol. 8, no. 3, pp. 251-262, Sep. 2019. doi: 10.1016/j.ijtst.2019.02.002.
66. Z. Zhao, J. Yuan, and L. Chen. Air Traffic Flow Management Delay Prediction Based on Feature Extraction and an Optimization Algorithm, *Aerospace*, vol. 11, no. 2, p. 168, Feb. 2024. doi: 10.3390/aerospace11020168.
67. L. Ma, X. Meng, and Z. Wu. Data-Driven 4D Trajectory Prediction Model Using Attention-TCN-GRU, *Aerospace*, vol. 11, no. 4, p. 313, Apr. 2024. doi: 10.3390/aerospace11040313.
68. X. Liu, X. Qin, M. Zhou, H. Sun, and S. Han. Community-Based Dandelion Algorithm-Enabled Feature Selection and Broad Learning System for Traffic Flow Prediction, *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 3, pp. 2508-2521, Mar. 2024. doi: 10.1109/TITS.2023.3321384.
69. K. Jiang, C. Wu, and Y. Chen. Revisiting the Efficacy of Signal Decomposition in AI-based Time Series Prediction.
70. X. Guo, W. Li, and J. Qiao. A modular neural network with empirical mode decomposition and multi-view learning for time series prediction, *Soft comput*, vol. 27, no. 17, pp. 12609-12624, Sep. 2023. doi: 10.1007/s00500-023-08638-3.
71. W.-N. Hsu, Y. Zhang, and J. Glass. A prioritized grid long short-term memory RNN for speech recognition, in *2016 IEEE Spoken Language Technology Workshop (SLT)*, IEEE, Dec. 2016, pp. 467-473. doi: 10.1109/SLT.2016.7846305.
72. S. Hochreiter and J. Schmidhuber. Long Short-Term Memory, *Neural Comput*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997. doi: 10.1162/neco.1997.9.8.1735.
73. A. Rehmer and A. Kroll. On Using Gated Recurrent Units for Nonlinear System Identification, in *2019 18th European Control Conference (ECC)*, IEEE, Jun. 2019, pp. 2504-2509. doi: 10.23919/ECC.2019.8795631.
74. K. Fukushima. Neocognitron: A hierarchical neural network capable of visual pattern recognition, *Neural Networks*, vol. 1, no. 2, pp. 119-130, Jan. 1988. doi: 10.1016/0893-6080(88)90014-7.
75. J. Heaton. Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning, *Genet Program Evolvable Mach*, vol. 19, no. 1-2, pp. 305-307, Jun. 2018. doi: 10.1007/s10710-017-9314-z.
76. Y. Kominami, H. Ogawa, and K. Murase. Convolutional neural networks with multi-valued neurons, in *2017 International Joint Conference on Neural Networks (IJCNN)*, IEEE, May 2017, pp. 2673-2678. doi: 10.1109/IJCNN.2017.7966183.
77. S. A. Sayed, Y. Abdel-Hamid, and H. A. Hefny. Artificial intelligence-based traffic flow prediction: a comprehensive review, *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, Mar. 2023. doi: 10.1186/s43067-023-00081-6.

78. Y. Zhang. Support Vector Machine Classification Algorithm and Its Application, 2012, pp. 179-186. doi: 10.1007/978-3-642-34041-3_27.

79. Houtao Deng. An Introduction to Random Forest, <https://towardsdatascience.com/random-forest-3a55c3aca46d>.

80. G. Meena, D. Sharma, and M. Mahrishi. Traffic Prediction for Intelligent Transportation System using Machine Learning, in Proceedings of 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things, ICETCE 2020, Institute of Electrical and Electronics Engineers Inc., Feb. 2020, pp. 145-148. doi: 10.1109/ICETCE48199.2020.9091758.

81. A. Mohammed and R. Kora. A comprehensive review on ensemble deep learning: Opportunities and challenges, Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 2, pp. 757-774, Feb. 01, 2023. doi: 10.1016/j.jksuci.2023.01.014.

FINANCING

The authors did not receive financing for the development of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Syed Aleem Uddin Gilani, Murad Al-Rajab, Mahmoud Bakka.

Data curation: Syed Aleem Uddin Gilani, Murad Al-Rajab.

Formal analysis: Mahmoud Bakka.

Research: Syed Aleem Uddin Gilani, Murad Al-Rajab.

Methodology: Syed Aleem Uddin Gilani, Murad Al-Rajab.

Drafting - original draft: Syed Aleem Uddin Gilani, Murad Al-Rajab.

Writing - proofreading and editing: Syed Aleem Uddin Gilani, Murad Al-Rajab.