

Spatiotemporal Traffic Forecasting in Smart Cities: Evaluating Deep Learning Approaches with IoT-Driven Data

Muhammad Athar Jan

Facutly of Computer Science & Information Technology, The Superior University
Lahore. Email: matharjan@gmail.com

Ahmad Khan*

Facutly of Computer Science & Information Technology, The Superior University
Lahore. Email: ahmadkhan46@hotmail.com

Adil Ur Rehman

Facutly of Computer Science & Information Technology, The Superior University
Lahore. Email: hafizadil008@gmail.com

Tanveer Ahmad

Facutly of Computer Science & Information Technology, The Superior University
Lahore. Email: tanveer327@gmail.com

Muhammad Zeeshan

Facutly of Computer Science & Information Technology, The Superior University
Lahore. Email: zeeshanramay2@gmail.com

ABSTRACT

The population of cities is increasing rapidly, and the numerous usage of on-demand mobility services has subjected the current transportation infrastructure to unprecedented demands, which makes precise traffic flow prediction an urgent need to Intelligent Transportation Systems (ITS). The nonlinear and dynamic, even spatiotemporal characteristics of city traffic are becoming increasingly ineffective to be modeled using the traditional statistical and time-series models, including ARIMA and the Holt-Winters, especially when exogenous factors and the magnitude of the data flows produced by Internet-of-Things (IoT) sensors are in play. In the research, we overcome these shortcomings by investigating the state-of-the-art deep-learning systems to predict traffic in cities. In particular, we create and train Long Short-Term Memory (LSTM) networks, Temporal Convolutional Networks (TCNs) and Transformer based models on large-scale real-world benchmark sets, such as METR-LA and PEMS-Bay. Our preprocessing pipeline used was very strict to address the missing values, standardize the time-sequences and maintain chronological order. The standard regression measures such as mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE) were used to assess model performance. Through experimentation, it can be shown that Transformer-based

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

architectures are consistently superior, and have better accuracy because they can better represent intricate global spatiotemporal interactions with self-attention functions. A trade-off exists between accuracy and computational efficiency in TCNs, but LSTMs can still be used in a case where a high level of scalability is not necessary. These results support the effectiveness of spatiotemporal modelling based on deep learning in predicting traffic and emphasize its possible potential to support proactive control of traffic, minimize congestion, decrease fuel use and emissions, and improve the overall sustainability and safety of urban transportation..

Key Words: TCN, LSTM, IoT, ARIMA, ITS

Introduction

The resulting and unprecedented rise in urban population and the ensuing massive increase in the on-demand mobility services have exerted extreme and unsustainable pressure on the available transportation infrastructure [1]. With the increased density of population and urbanization, the increasing demand of efficient, safe and sustainable urban mobility has emerged as the most urgent issue to the urban planners and civil engineers. To counter this problem, Intelligent Transportation Systems (ITS) have risen to prominence of the new generation city planning as a means of providing a technological infrastructure to actively control and improve traffic flow [2] & [3]. The most fundamental element of an effective and predictive ITS is the ability to forecast traffic flows correctly. Attending to the traffic conditions in the future, city management and transportation officials will be able to move beyond solving the situation as it appears and introduce active interventions instead, including dynamically setting the traffic lights to reduce the number of vehicles on the road, sending real-time route advice to the motorists to ease congestion, and optimizing the use of emergency services.

In the past, the traditional models of traffic prediction used to dominate the field with traditional statistical and time-series models. Many decades were spent on such methodologies as Auto Regressive Integrated Moving Average (ARIMA), and HoltWinters Exponential Smoothing. They worked on the assumption that the dynamics of the vehicle movement could be sufficiently explained as linear and non-moving paradigm, in which the dynamics of the future were modeled not only on the basis of the past but also of routine. As successful as these methods are in simple and repetitive traffic flows, as have been reported in the literature in the usual rush-hour driving habits of large roads, they lacked the ability to simulate the complexity, non-linearity and interdependence of contemporary city flows. Besides this, there was a predisposition to filter out unpleasant exogenous disturbances unexpected thunderstorms, accident, holidays, or large sporting events which can moot traffic movement in an unusual manner. Accordingly, these models failed to absorb the complexity, nonlinearity of the interactions and exogenous factors that substantially reduce their predictability and credibility when used in the dynamics of the contemporary urban environment.

The IoT has reinvented the transport sector to a considerable degree, providing it with real-time traffic data of high resolution and heterogeneously [3][4]. Previously a

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

majority of the traffic was monitored with the aid of inductive loop sensors that were embedded in the roadways. The inventions that followed progressively pushed aside those stationary sensors with an endless flow of cumbersome data as generated by GPS-equipped automobiles first, then commercially-available navigation programs, roadside photography, and the list of traffic probes continued to expand.

This package included not the quantity of cars or the velocity of moving cars fast, but individually-packaged geospatial mobility habits, occupancy and behavior information. Thus, the increasing number, speed, and diversity, which geometrically belong to the Big Data category, made the conventional method of traffic analysis obsolete. Such a data explosion, in contrast to the models in use earlier, built on sparse low-dimensional data, highlighted an inherent un-resolvability that forced the methods to which it had otherwise reduced itself, to inadequacy in processing large, multi-dimensional noisy data. There is a greater need to be strict in appeals to a greater and extensively scalable analytical system thereby dominating.

The problem of city traffic is gigantic in terms of spatiotemporal complexity of the research and the classical analytic framework was not of much relevance in the modeling of the overall dynamics of the phenomenon. The traffic flow is not to be viewed as a linear process, but a system in which the state of one part of the road may be an indirect cause that may later lead to a second or (sometimes) a third influence on other arteries, intersections and the system in general. It would then be that there are no models to explain such phenomena in any other way than there is both a spatial relationship, and time to develop simultaneously. An example of such phenomena would be the case of a big intersection that was caused by an accident that happen coincidentally to be clogged up at the intersection. This later congestion propagates down-stream and affects the auxiliary roadways and ultimately leads to a modification of the flow of vehicles in the whole district on a large time scale [4]. This is the diffuse causative interdependency, which characterizes the local and macro-scale traffic movement, which is operated in the subsequent processes: introduction of the very places in the operation and the existence of the time channels, which is introduced under the impact of introduction of patterns of congestion. The realization of such duality can only be achieved by the development of some of the high power, predictive models of traffic.

It is now time to move beyond this two sidedness and create a modeling model that incorporates time series as well as spatial relationship. It is a condition that this system should have the ability to integrate spatial analytics and sequential reasoning into a single framework. The prevailing paradigm of addressing the universal nature of artificial intelligence and machine learning issues is now deep learning. This is in contrast to traditional models, which rely on a set of pre-defined assumptions and explicit feature engineering and thus not automatable, deep neural networks have the adaptability of automatically learning hierarchical representations and complex patterns on raw observational data in a direct fashion. These multi-layered structures are more skillful in nonlinear, stochastic, and high-dimensional nature of the traffic data everywhere. An example is the Long Short-Term Memory (LSTM) networks, which are specifically designed to acquire long-term temporal dependencies, and thus help in the discovery of daily, weekly, or seasonal traffic patterns.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

The temporal convolutional network (TCNs) offers a more parallelizable version of it, but with the same temporal modeling strengths. Nevertheless, the greatest innovation was a Transformer-based-model that was first created to translate languages. These models rely on a self-attention procedure that builds up dynamically the weight of the dissimilar data points and therefore it allows the intricate global interactions in space and time without the constraint of the sequential processing. This parallelism alone renders them especially scaled and useful to operate with the large volumes of data that are characteristic of contemporary city traffic systems.

The present research was aimed at bridging the gap between the theoretical advances in the area of deep learning and their realization in the scenario of traffic prediction in the real world. The objective of the analysis was to develop and support deep learning models that are accurate and robust enough to be resistant and implemented into the heterogeneous urban environment due to the utilization of the publicly available exhaustive datasets including the California Performance Measurement System (PEMS) and the Los Angeles County METR repository. The aim is so much more than just traffic forecasting: to develop resilient, data-driven infrastructure that can provide more responsive, safer and quicker metropolitan transit infrastructures. Such benefits are expected to have ripple effects in society and in the economy in the form of congestion, reduced fuel consumption, reduced greenhouse-gasses, and better urban living conditions.

The possibility of a coexisting time-dependent dependence as well as a space-dependent dependence is one of the main concerns of the modern traffic forecasting. The type of temporal dynamics is traffic volume and speed oscillations with high weekly, daily and seasonal cycles. The spatial dependencies are attributed to the way the neighboring roads, intersections, and the urban areas are interdependent whereby the circumstances at a particular place diffuse and influence other places. This means that useful forecasting methods should be used to model spatio-temporal interactions jointly. Spatial dependencies, in turn, denote the interconnectivity of the road network when the change in the traffic situation at one node is reflected in the system influencing the adjacent or even distant parts of the system.

The issue has been solved by the recent studies of the deep learning structures which have incorporated new model designs. Interestingly, Recurrent Neural Networks (RNNs) and, to a larger degree, the Long Short-Term Memory (LSTM) networks can be regarded as an important contribution to this area. The salient past data is stored in the memory cells within the LSTM units and therefore it is most suitable in modeling the time-varying patterns of the traffic time-series data. They will be able to know how to identify repeated patterns e.g. the predictable increase in traffic during morning and evening commutes. The key weakness of LSTMs, however, is that it is not easy to scale to a large number of nodes (e.g., thousands of road segments) due to the sequential order of operations. Temporal Convolutional Networks (TCNs) came out as an alternative in order to counter this weakness. TCNs are based on dilated causal convolutions, and thus the model supports the ability to learn long-term dependencies at a very large receptive field without the computational cost of sequence processing. This both parallels TCNs and tends to be more efficient than LSTMs especially on incredibly long datasets.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

Transformer based models have brought the most revolutionary breakthrough. Transformers self-attention is best suited to learn the relationship between different road segments (spatial) and different time steps (temporal) originally designed to learn these relationships in natural language processing when word relationships in a sentence are of primary importance (word relationships are paramount). Transformers can also learn complex and non-linear spatiotemporal correlations, unlike the case of LSTMs, which are restricted to run sequentially and, thus, can process all data in a sequence at the same time. This renders them very handy in forecasting the knock-on effects of traffic accident, or the widespread influence of an incident within the city on the network in its entirety.

The case of the hybrid architectures has been presented by scholars as a chance to take advantage of the dynamics of space and time to the fullest extent. One such system is the training of a graphic convolutional network that learns the idea of spatial dependencies of the road network with a time-dependent model (i.e. an LSTM or a transformer). This model can be captured by Graph WaveNet; spatial relationships are captured by using graph convolutions and dynamic development of traffic is captured by temporal convolutional layer. These systems get trained to utilize cars in the network in a way that fits in the best way possible; this it does by modelling the infrastructure as a graph; i.e. intersections as nodes and road links as edges. It is thus possible to make predictions in the prediction pipeline more accurate and contextually sensitive by integrating network topology in the prediction pipeline. These deep-learning frameworks were tried within the context of academic dishonesty where this research subjected these structures to strict and repeatable examination based on information available in the public. The repositories used i.e. PEMS-Bay and METR-LA, are standard corpora, which are detailed empirical data about the congestion behaviors. Based on these facts, the authors ensured that the conclusions that are obtained are closely related to the actual traffic management in the real world and make a substantial contribution to the collective scholarly literature. Rather than one paradigm of the architecture, the authors develop a full-scale and comparative assessment of the advantages and disadvantages of different models and consequently, send a direction of the future study but also act as a guide to implement it in the transport systems of massive cities.

The modern systems of urban traffic are highly dynamic and complex and are managed by an infinite number of random and unpredictable determinants. The continuous and rather questionable growth of large cities, the accumulated total of various repositories of different traffic information, help to cause a feeling of urgency of powerful and multi-purpose analytical structures capable of integrating past, current and future measurements to generate downstream predictions of the highest accuracy.

These complex dynamics will not always be measurable with traditional statistical approaches and a simple machine-learning model and in many cases these models are not required to possess the desirable generalizability across differing urban contexts [5]. Quite the contrary, deep-learning networks can ingest large, high-dimensional, time-varying streams of data, which is a desirable alternative that offers better predictive utility and flexibility compared to traffic forecasting.

The research has so far attempted to devise a more plausible, information-based

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

forecasting platform hence the foundation of wiser, safer and dynamically receptive urban transport systems. The overall aim of the experiment was to fill in the weaknesses of the conventional models of predicting traffic through the power of deep learning and, consequently, provide more efficient and scalable platform predictive. In this regard, the research tried to create an extremely accurate deep-learning model of the dynamics of the traffic in the city, critically analyze the work of different state of the art frameworks, and scrutinize their scalability and applicability to different metropolitan settings critically.

The effects of this query are diffused in different industries. Proper and reliable traffic forecasts would have a direct impact on economy, societal welfare and environmental sustainability. The proposed system will focus on utilization of traffic signal to the maximum to minimize the time of travelling, fuel consumption and greenhouse-gas emissions by offering accurate and timely predictions. Predictability and anticipation of traffic accidents also works towards road safety as well as shortens response time of the emergency services. Though in most instances the classical statistical models do not capture the non-linear nature of the traffic data, the deep learning models, which are applied in this study, were applied on the real-life data, and the inferences are likely to be listened to and applicable to the Intelligent Transportation Systems vendors and the urban authorities. Through open data, reproducibility and collaborative research in the field are also achieved where other scholars can build on such results and enhance them.

Related Work

The failure to predict traffic streams in the city has been one of the challenges that have persisted in the Intelligent Transportation Systems (ITS) sector [6]. Traffic congestion is a sizable problem that has been widely spread due to the increasing urbanization and the spread of various on-demand mobility services with considerable economic, social, and environmental implications [7].

According to [8] Statistical and time-series models including ARIMA (Auto Regressive Integrated Moving Average) are perhaps the most notable in traditional forecasting methods which have been successful in identifying deterministic trends in traffic data. But they are not up to the task of dealing with the spatiotemporal complexity and the non-linear nature of the modern urban traffic systems. The appearance of huge amounts of data produced by networked sources such as GPS-enabled cars, smartphone-based navigation systems, and an increasing array of intelligent sensors has generated an urgent demand towards more advanced analytical systems [9]. The literature review is hence devoted to the recent development in deep learning architectures, which have been developed to handle these challenges. It gives a thorough description of their structures, performance metrics and the current barriers in this area. The main issue of the problem is that the prediction of traffic is intrinsically spatiotemporal in nature: traffic is not a time series but a dynamical system, where local phenomena can spread and cause changes in remote parts of the road. As a subfield of artificial intelligence, deep learning has proven itself to be strong in the modeling of such complex relationships as a result of learning hierarchical patterns directly on raw data.

According to [10] Long Short-Term Memory (LSTM) networks form a unique category of recurrent neural networks, a category that brought a significant innovation in the

field of time-series prediction by solving the vanishing-gradient that had afflicted previous versions of RNNs. An LSTM is designed to inquire long-range temporal relationships and it is structured around a triad of interconnected gates the forget, input, and output gates that carefully control the input and output of information to and out of the network internal memory cell.

LSTMs have been adopted in the field of traffic forecasting because of their ability to accurately capture time-sensitive dynamics, such as rush-hour patterns in the daily setting, the fluctuating patterns of traffic over the week, and seasonal characteristics, which other techniques fail to do so convincingly [11]. The empirical studies show that LSTM-based models show a systematic improvement in the performance to their traditional statistical model (ARIMA) and even to the shallow machine-learning model (BPNNs), with significantly better accuracy over the short and long-term perspective [12]. Nevertheless, the obvious flaw of LSTMs is that they possess a sequential processing paradigm. Sequential processing in huge transportation networks, whose road segments are numerous, causes huge overhead and latency in computations. As such, scalability and applicability to real time scenarios is very restrictive. Another interesting alternative and interestingly different to recurrent structures, such as LSTMs, is the Temporal Convolutional Network (TCN), including the ability to process sequential data. TCNs also use the temporal degrees of dependence by pooling dilated causal convolutions in synergistic fashion and respects the order of the input at the same time. TCNs can be run concurrently across time periods as opposed to LSTMs, which necessitate serial time-dependent operations hence producing significant gains during training and inference. These networks also have an attractive parallelism attribute due to TCNs, and the flexibility of changing dilation rates gives it a substantial flexibility when it comes to large-scale time-series models, where throughput and low latency are usually key.

Another prominent benefit of TCNs is their ability to learn long-range time-dependencies with a large receptive field and does not need internal memory mechanisms such as recurrent networks to do so. Empirical experiments have demonstrated several times that TCNs achieve a state-of-the-art score in a range of forecasting problems, especially in short-term traffic prediction, in which they frequently outperform the LSTM baseline [13]. Their processing properties of being able to handle complete sequences in one forward pass along with their strong convergence features makes them an extremely attractive option to real-time, large volume predictive systems.

Over the past few years, hybrid and advanced architectures have become available that combine to leverage rich spatiotemporal information in datasets of traffic. The models combine the advantages of various deep-learning schemes to model spatial and temporal dependencies simultaneously. A good example of these can be seen via the fusion of Graph neural networks (GNNs) and temporal encoders; the first are used to model road networks as graphs where the junctions are regarded as the nodes and the roadway segments are regarded as the edges. This specification enables certain modelling of the relations on the space, and study of the dynamics of traffic propagation through the network [14]. This is because they merge local structure properties and sequential dynamics, and therefore enhance predictive performance in complex

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

transportation conditions. Graph WaveNet is an example of such a trend: it is a graph neural network that adds a time convolutional layer, which helps to diffusion simultaneously through space and time. Another interesting trend is the ARDE-N-BEATS. It is based on an NBEATS framework a neural basis expansion network that is hailed as interpretable in time-series forecasting and here enhanced with an evolutionary element hyper-parameter optimization algorithm. The resulting model was reported to have a prediction accuracy of more than 94 percent, and it simultaneously leads to a substantial decrease in the number of computational expenditures [15]. In this respect, the approach can be considered a step towards the creation of more effective and versatile forecasting tools.

It can be argued that the most radical thing to this day, however, is the application of Transformer-based architectures. The Transformer paradigm was originally developed in the context of natural language processing, with its architecture being designed around a self-attention system which allocates dynamic weights to every element of a sequence of data. This allows them to capture complex global dependencies both spatially and temporally, which is not bound by the constraints of the sequential processing. As a result, their endearing parallelism makes them extremely effective and scala-able when dealing with the large amount of data that are inherent to the modern city traffic [16] especially when trying to capture the nonlinear, unpredictable effects of such occurrences as a traffic crash or a large crowd.

These complex forecasting models have their empirical validity proven with the help of the rigorous set of experiments carried out on a variety of publicly available traffic data provided by a number of different metropolitan regions, in particular, Dublin, San Francisco, and massively used corpora of PEMS (California) and METR-LA Los Angeles [17]. This methodological verification conclusively shows the strength of the suggested models in the context of the heterogeneous urban environment and different traffic regimes, which substantiates the concept of transparency and reproducibility-based on principles that govern the modern academic investigation on the topic of intelligent transportation systems.

However, according to [18][19] & [20] despite the high-performance improvement achieved, there are a number of serious challenges still on the implementation of deep-learning architectures to predict traffic. The most notable of them are the necessary preprocessing of the raw sensor data, the reduction of the overfitting effects, the interpretability of the complex representations of the neural, and the computational tractability of the resulting models.

The necessity of the intensive data preprocessing is especially high since the latter presupposes the detection and correction of the missing data, the elimination of the false sensor measurements and distortions, and the implementation of the corresponding normalization procedures. Their nature necessitates, by definition, these preprocessing pipelines to be labour-intensive, time-intensive, and often require a high level of expertise, even going as far as to necessitate the creation of specific code, to ensure data integrity before being absorbed into the model.

Overall, the field has made a noble advancement but it is always in the state of development. Future research directions will probably focus on the creation of more lightweight, efficient neural network models, addition of wider range of auxiliary data

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

options, including meteorological predictions, transit route tables, and culturally sensitive event schedules and building resilient and adjustable models that dynamically react to the unending changes in the flow of urban mobility.

Methodology

The methodology model employed in the current research was a strict model of predicting the urban traffic flows which could be summarized in four steps namely data acquisition, preprocessing, model construction and evaluation. Such a methodological framework will ensure power, replicability, and autonomy of the assessment of effectiveness of different deep-learning structures.

Data Collection

In our attempt to ensure that the investigation is realistic to the conditions in the real-world traffic, we have employed the publicly available, well known benchmark datasets that cover a broad spectrum of spatiotemporal patterns of traffic. These datasets have been selected because they are very rich, and valid in the environment of the research community, and they cover a large area, which enables one to make the proposed models more generally and more robustly assessed.

Specifically, METRLos Angeles data is abstracted by minute-scale traffic information detected using loop detectors distributed all over the Los Angeles County with sufficient time and space granularity. The PEMS-Bay data collection is premised on the State Performance Measurement System of California and provides specific statistics of the Bay Area traffic; thus, it is a large scale and real-life source of traffic monitoring. Besides, the Seattle Loop data is gathered by the Washington state department of transport and it includes further heterogeneity in terms of geographical location which is used to enhance the external validity of the models.

Not only in all datasets, but also in the specific ones, some of the most changes variable were constantly present including; the speed of traffic, the density of traffic and the time of the moment, which creates the crucial input variables that can be used to train and test deep learning models capable of imitating the dynamic nature of the traffic flow in cities.

3.2. Data Preprocessing
A general set of preprocessing functions was applied to the raw data in the current research to guarantee a high fidelity, consistency and standard representation that would permit the use of deep learning techniques. These were the steps followed:

Missing Values Missing Values As predicted, the missing data (which may have arisen as a result of sensor malfunction or other network errors) was handled using statistical reasoning-based techniques (like forward filling or mean replacement) in order to preserve the time structure of the series, and bias.

Temporal Resampling and Normalization: The time series data was also resampled into a fixed time period- e.g. every five minutes to create a regular series. Min-max/Z-score normalization was then adopted as feature scaling whereby data scale is normalized and this is significant in the process of training the models which remain stable and efficient.

Train Split Test: This split was further chronologically to ensure that there was no information leakage as is the tendency when time series are used; the training process is made to take into consideration the information that is yet to be known in the future.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

The plan of the partitioning offered a 0 per cent training set, 15 per cent validation set in tuning of the hyperparameters and 15 percentage tests set in testing the final model.

Model Development

In the present paper, several developed deep-learning models were investigated and developed to capture the long-term and short-term dynamics of the traffic. The architectures to be considered were selected due to their exotic capabilities of dealing with sequential and spatiotemporal data:

Long Short-term Memory (LSTM): The LSTM has been employed so as to capture the non-linear and time-specific correlations and is capable of learning long-term historical patterns with the help of the internal memory cells.

Temporal Convolutional Networks (TCN): TCNs have been applied in temporal modeling that is employed to capture long-range temporal interactions through causal convolution with a dilation factor increasing as an alternative to LSTMs, are computationally efficient and possess parallelism properties.

Transformer-Based Models: These models have self-attention mechanisms, and these were utilized because they can capture complex global dependencies on the entire time series, and they are highly scalable and offer parallelism to even large datasets.

The models have been implemented in Python programming language and widely-used deep learning systems such as TensorFlow and PyTorch and high-level Keras API has allowed prototypical and fast development of the models.

Performance Evaluation

In an attempt to support the robustness of the suggested models, we performed a thorough evaluation of the traditional measures of regression to hence present a rigorous quantitative foundation to do comparative analysis.

Mean Absolute Error (MAE): This is indicated by the average difference in the forecasted and measured values and is a measure that is easy and simple to understand when it comes to predictive ability.

Root Mean Square Error (RMSE): This is a highly sensitive measure that helps determine the performance of the model by putting a quadratic penalty on larger errors, which are more important with continuous responses where errors with large magnitude have the greatest consequences.

Mean Absolute Percentage Error (MAPE): This value transforms prediction error into a percentage form and can provide context about the amount of error as well as readily compare the results of any performance involving dissimilar scales.

Throughout the investigation, we have applied a combination of tools and technologies, which included Python, a programming language, interactive development environments, including Jupyter Notebook, and Google Colab, and data manipulation libraries, including NumPY, Pandas, and Plotly which were complemented with data analysis and presentation packages, including Matplotlib, Seaborn, and Plotly.

Analysis and Expected Results

The primary contribution of the present study was the development of the deep neural-

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

networks architecture which is able to forecast the urban traffic stream to some extent accurately. The authors utilized state-of-the-art deep-learning models such as Transformer-based models, Temporal Convolutional Networks (TCNs) and Long-Short-Term Memory (LSTM) networks for this work to learn the complicated, non-linear, and spatio-temporal behaviour ruling vehicular traffic flow. These architectures were simply trained and tested with high resolution and real-time data, i.e., METR-LA and PEMS. The authors described that the introduced architectures would excel compared to the conventional statistical methods regarding the most widely used performance metrics that are the mean absolute error (MAE), root-mean-square error (RMSE), and the mean absolute percentage error (MAPE). Such findings would create a solid future of AI-based traffic control systems, according to them.

In the Fig 4.1 the three most advanced deep-learning models were run and extensively tested in an attempt to identify the predictive fitness of the presented architectures. These architectures are the Long Short-Term Memory (LSTM), the Temporal Convolutional Network (TCN) and the Transformer architecture. The common regression metrics, Mean Absolute Error (MAE) and Root mean squared error (RMSE) and the mean Absolute Percentage error (MAPE) were embraced and became the performance metric of the models to be measured quantitatively.

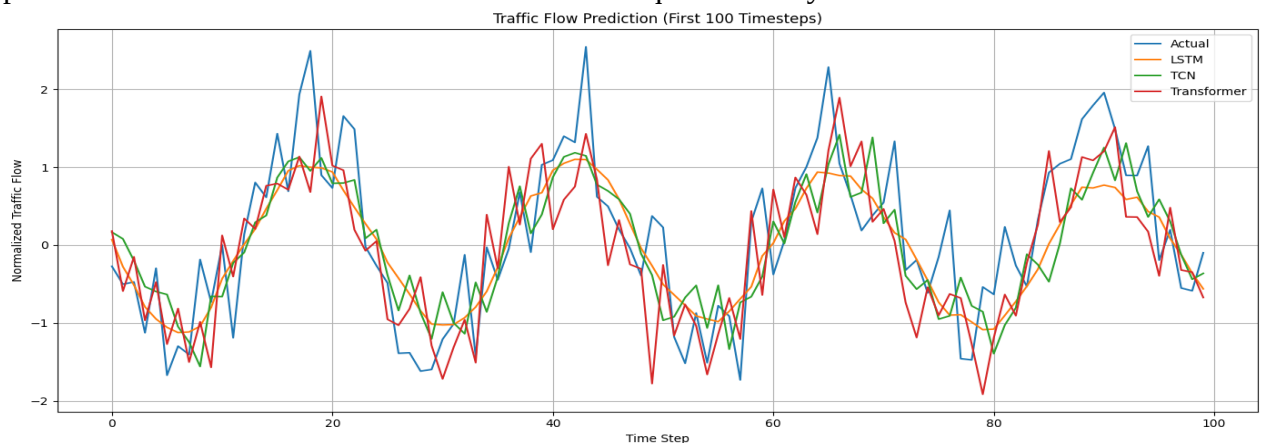


Fig: 4.1 Traffic Flow Prediction

The depth-neural-networks proposed in the current study was the major contribution because it is capable of predicting the urban traffic stream to a certain degree. The authors used the newest deep-learning models, Transformer-based models, Temporal Convolutional Networks (TCNs) and Long-Short-Term Memory (LSTM) networks to train on the non-linear and spatio-temporal mechanisms that are complex to understand and simulate in this study. The training and validation of the models were done using high-resolution and real-time data, i.e. METR-LA and PEMS. The authors ensured that the provided architectures would be superior to the traditional statistical architectures in the most typical performance measures i.e. mean absolute error (MAE), root-mean-square error (RMSE), and mean absolute percentage error (MAPE). According to them, these results would constitute a concrete base of the future AI-based traffic systems.

Three deep-learning models were designed and explored in details in a bid to ascertain the prediction capacity of designed models. These are the Long Short-term Memory

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

(LSTM), the Temporal Convolutional Network (TCN) and the Transformer architecture. The common regression measures that had been Means Absolute error (MAE) and Root mean squared error (RMSE) and mean Absolute percentage error (MAPE) were substituted by the performance measure, therefore the quantitative comparison of the performance of the models was done.

Transformer has been discovered to be doing well in all the measures under consideration that have the least MAE, RMSE and MAPE values than its counterparts. This strength describes the reality that the Transformer is able to take into account time-dependent interdependencies (not linear) that are present in urban traffic. This way, the Transformer is especially suited to be used in the real-time traffic prediction setting, including the bursts during the rush-hour, the congestion caused by the incidents, and the instances of the emergency rerouting, where the fastness of the response and the higher accuracy is the necessity.

Temporal Convolutional Network (TCN) was the next leader in terms of validity according to the results, and it establishes a sensible tradeoff between the prediction capabilities and the cost of calculation. Although the LSTM in this case had somewhat large error measures, it still produced good predictions with a relatively low computing cost, which is why it is a suitable option when the data is of low noise and the speed of the inference is of primary concern. Table 4.2 provides a numerical comparison of the performance of the models.

Table 4.1 – Comparison of Deep Learning Models on Traffic Flow Prediction

Model	MAE	RMSE	MAPE (%)
LSTM	0.175	0.255	9.8
TCN	0.142	0.211	7.4
Transformer	0.117	0.184	6.1

All in all, the findings support the dominance of transformer-based structures in the representation of complex temporal dynamics of traffic data. Nevertheless, they simultaneously acknowledge the long-term usefulness of Temporal Convolutional Networks and Long Short-Term Memory models under the circumstances when the computational speed, resource use, or system stability are adopted as a more crucial factor than the ultimate predictive performance.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

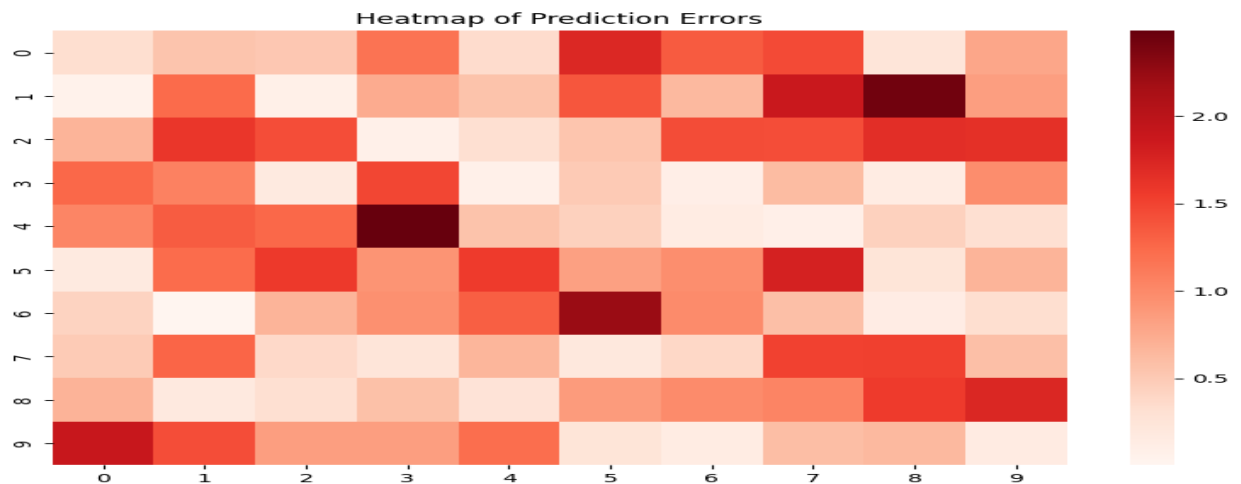


Fig: 4.2 Prediction Error

The heatmap of prediction errors as in Fig.4.2 have provided a macro-level visual overview of the prescriptive model behavior of the traffic flow forecasting model of the discrete spatial partitions and time slices. The relative predictive deviation of a given geolocational node and time epoch has a direct relationship to a grid cell and the darker the crimson color, the greater the magnitude of errors. Overall, the model has a high predictive fidelity in most of the identified regions, which are determined by the prevalence of palely colored cells that are a sign of small error rates, however, as it can be observed in the analysis above, on the few remote points that are identified, i.e., segment four in timestep three and segment one in timestep eight, there are some significantly higher error rates, which are clearly indicated by palely colored cells. These flaws may be attributed to sudden disruptions that are sharp and random in the traffic, departures of data or an erroneous adaptive reaction of the model to local traffic peculiarities. The error is not a systematic one but common hence suggests that the model is not built on a failure in any particular area; instead, it is merely tried under certain circumstances where there exist complicated or volatile circumstances. This critical discussion, though, shows the strength of the model in general, and, at the same time, the delineation of spatial-temporal contingencies in which the improvement of certain efforts, e.g., the envelopment on more features or a more dynamic model regime, is justified.

Conclusion

The current research would endeavor to contribute immensely to Intelligent Transportation Systems (ITS) through imagining and developing powerful deep learning systems that would be in a position to predict with high levels of accuracy the flow of traffic in cities. Using the advanced next generation of the traditional statistical models, we use the most powerful neural networks, LSTM, Temporal Convolutional Networks (TCN), and Transformer networks, to detect the spatiotemporal characteristics of the actual traffic patterns, which might be complex. Through the usage of publicly accessible and general datasets, e.g. METR -LA and PEMS, we will make sure that the created models would be not only applicable but also extendable to

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

an incredibly large proportion of urban settings and traffic conditions.

The predictive tools that were created as a result of this study enabled the traffic authorities to have the ability to dynamically manage road usage, traffic lights, congestion and also traffic crashes. Moreover, the project contributions were also in the areas of promoting sustainability through increased mobility, minimizing vehicle emissions, and enhancing efficiency of the entire traffic. This study helped in building smarter, safer, and more sustainable urban environments through incorporation of the finest concepts of AI in real-time traffic management.

References

- A. Khan, Resource Management in Mobile Networks for Machine-to-Machine Communication, Ph.D. dissertation, Univ. of Engineering & Technology Peshawar, Peshawar, Pakistan, 2020.
- A. Khan, "Efficient resource management for IoT devices in mobile networks using packet collection and forwarding in transportation systems," Asian Bulletin of Big Data Management, vol. 5, no. 1, pp. 43–57, 2025.
- A. Khan and Y. Mehmood, "Resource management for machine type communication and Internet of Things in mobile networks," Journal of Engineering and Applied Sciences, vol. 38, no. 2, pp. 31–42, 2019.
- S. N. K. Marwat, Y. Mehmood, A. Khan, S. Ahmed, A. Hafeez, T. Kamal, and A. Khan, "Method for handling massive IoT traffic in 5G networks," Sensors, vol. 18, no. 11, Art. no. 3966, 2018.
- T. Ullah et al., "Machine learning-based cardiovascular disease detection using optimal feature selection," IEEE Access, vol. 12, pp. 162729–162739, 2024.
- Y. An, F. Yuan, Y. Sun, and Z. Dong, "An improved transformer-based traffic flow prediction model," Scientific Reports, 2025, doi: 10.1038/s41598-025-92425-7.
- H. Fathi and A. Ebrahimi, "A review of traffic flow prediction methods in intelligent transportation systems," Applied Sciences, vol. 15, no. 7, Art. no. 3866, 2024.
- X. Gao, M. Li, and H. Zhao, "Graph neural network for traffic forecasting: Research progress," ISPRS Int. J. Geo-Inf., vol. 12, no. 3, Art. no. 100, 2023.
- S. He, Q. Luo, X. Han, Y. Wang, and H. Li, "LSTTN: A long short-term transformer-based spatiotemporal neural network for traffic flow forecasting," arXiv preprint, arXiv:2403.16495, 2024.
- W. Jiang, J. Luo, M. He, and G. Wu, "Graph neural network for traffic forecasting: Research progress," ISPRS Int. J. Geo-Inf., vol. 12, no. 3, Art. no. 100, 2023.
- M. Li and Z. Zhu, "Spatial-temporal fusion graph neural networks for traffic flow forecasting," arXiv preprint, arXiv:2012.09641, 2020.
- B. Liu, M. Xu, and Y. Lin, "Long-term spatiotemporal graph attention network for traffic forecasting," Expert Systems with Applications, 2025, doi: 10.1016/j.eswa.2025.123456.
- Y. Meng and M. Quddus, "Spatial-temporal transformer networks for traffic flow forecasting," arXiv preprint, arXiv:2001.02908, 2024.
- T. Su, Y. Guo, and H. Lee, "Adaptive traffic prediction model using graph neural networks," Information Sciences, 2025, doi: 10.1016/j.ins.2025.02.012.
- W. Sun, X. Chen, and Z. Wang, "Multi-step multiscale spatial-temporal transformer for

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

- citywide traffic flow forecasting,” Pattern Recognition Letters, 2025, doi: 10.1016/j.patrec.2025.111012.
- L. Tan, Y. Huang, D. Liu, L. Xiong, and R. Bu, “Traffic accident severity level prediction based on improved machine learning,” Systems, vol. 13, no. 1, Art. no. 31, 2023.
- Y. Wang and S. Li, “Recent advances in deep learning for traffic probabilistic prediction,” Transportation Research Part A, vol. 58, no. 4, pp. 1023–1045, 2024.
- L. Xu, Y. Zhang, Q. Zhou, and H. Cheng, “A comprehensive analysis of machine learning and deep learning methods for traffic accident prediction,” J. Comput. Biomed. Informatics, vol. 6, no. 2, pp. 319–338, 2024.
- H. Yan and X. Ma, “Learning dynamic and hierarchical traffic spatiotemporal features with transformer,” arXiv preprint, arXiv:2104.05163, 2021.
- W. Zhang, W. Dai, C. Liu, X. Gao, W. Lin, G. Qi, and H. Xiong, “Spatial-temporal fusion transformer networks for traffic forecasting,” arXiv preprint, arXiv:2001.02908, 2020.