Critical Review: Real Time Traffic flow prediction with Time Series Models

Nivedita Tiwari [a], Dr. Lalji Prasad [b]
[a] Research Scholar, Sage University, Indore, India, nivedita.8584@gmail.com
[b] Professor, Sage University, Indore, India, hoi.iac@sageuniversity.in

Abstract
Traffic prediction using time series models is important for forecasting the volume and density of traffic flow, usually for the purpose of managing vehicle movement, reducing congestion, and generating the optimal route. This work will analyze the performances of different predictive time-series models (ARIMA, SARIMA, LSTM) for predicting traffic flow. The performance of each model is evaluated based on the error functions incurred by each approach. It is necessary to understand how to choose the right combination algorithms and the dataset approachable. So, in this research we are going to explore how to analyze time-series models and algorithms and how machine learning helps to forecast congestion and plan optimal routes. This work attempts to highlight the usefulness of Time Series analysis in traffic forecasting by using multivariate and univariate analysis to understand the structure of the data to choose the right modeling technique. The result showed which technique is the best to model the time-series traffic conditions.

Keyword: **Predictive Modeling; Time-series Analysis; Traffic Flow; LSTM; ARIMA; SARIMA.**
DOINumber:10.48047/nq.2022.20.22.NQ10277 **NeuroQuantology2022;20(22):2838-2850**

1. Introduction

Prediction is the process of predicting the outcomes of unknown facts. Forecasting, on the other hand, is a sub- discipline of prediction in which we make predictions about the future based on time-series data. The sole distinction between prediction and forecasting is that the temporal component is taken into account. Forecasting deals with estimation of future demand for a product.

Time series data changes according to time. Real-world data [1] may be non-stationary and have seasonality. Organizations can utilize time series analysis to figure out what's causing trends or systemic patterns across time.

Organizations can use time series forecasting to predict the likelihood of future events when they study data at regular Intervals [5]. Predictive analytics includes time series forecasting. It can reveal likely data changes, such as seasonality or cyclic behavior, allowing for a better understanding of data factors and better forecasting. Predictive modeling is a process of creating, testing and validating a model and methods for doing predictive modeling is available in predictive analysis. Time-series is a predictive analytics methodology.

The accurate prediction of the parameters is important for future traffic condition assessment. Accurate and timely short-term traffic prediction is important for Intelligent Transportation System (ITS) to solve the traffic problem. The key to the ITS lies in the accurate forecast of traffic flow [6]. Short-term traffic prediction allows Intelligent Transport Systems to proactively respond to events before they happen. With the rapid increase in the amount, quality, and detail of traffic data, new techniques are required that can exploit the information in the data in order to provide better results while being able to scale and cope with increasing amounts of data and growing cities [11].

2. Related Work

Multivariate time series models and univariate time series models are two types of time series models. Univariate time series models are models used when the dependent variable is a single time series. Only one variable is varying over time. You will have only a one-dimensional value, which is the temperature.

When there are several dependent variables, multivariate time series models are used. In addition to depending on their own past values, each series may depend on past and present values of the other series. Multiple variables are varying over time. A time series forecast can be classified into two categories.

If you use only the previous values of the time series to predict its future values, it is called Univariate Time Series Forecasting. And if you use predictors other than the series to forecast it is called Multivariate Time Series Forecasting.

To solve the problem of complexity and stochasticity arise in short term traffic prediction, Saiqun Lu, et.al., proposed a combined prediction method for short term traffic flow based on the autoregressive integrated moving average (ARIMA) model and long short-term memory (LSTM). MAE, MSE RMSE and MAPE used as evaluation Indicators and three real highway data sets were used and compared with the three comparative baselines of ARIMA and LSTM two single methods and equal weight combination [1].

Kong Yan et.al. [2], addresses the problem of Traffic congestion in smart and major cities due to the fast economic growth and the highly increasing number of vehicles. To successfully predict accurate traffic flow information, applied three different kinds of recurrent neural network architecture such as simple RNN, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) by considering different time intervals. The data was gathered from the California Department of Transportation in the years of 2018 and 2019. LSTM model is proposed in this study on both short and long-time intervals. Two popular metrics, including Mean Absolute Percentage Errors (MAPE) and Root Mean Squared Error (RMSE), have been used to evaluate the prediction efficiency.

LSTM is a special version of RNN which solves short-term memory problems and, in this research, we will explain LSTM in a very simple manner using real life examples [2], [4].

Mizanur Rahman and Mashrur Chowdhury [4], worked on traffic flow parameters, such as average travel speed of a roadway segment and average space-headway between vehicles from the connected vehicles (CV's) broadcasted data. They developed a real-time traffic-data prediction model that combines LSTM with Kalman filter-based Rauch–Tung–Striebel (RTS) noise reduction model that improves the prediction accuracy using noisy traffic flow parameters. LSTM/RTS model reduced the mean absolute percentage error (MAPE) from 19% to 5% for speed prediction and from 27% to 9% for space-headway prediction.

Jianhu Zheng and Mingfang Huang [5], discussed the problem of Traffic congestion posing a serious threat in many large and medium-sized cities, to sustainable urban development. Based on the long short-term memory (LSTM) network, the authors created a traffic flow forecast model. Through long-term traffic flow forecast tests using an actual traffic flow time series from OpenITS, the proposed model was compared to two traditional forecast methods, namely the autoregressive integrated moving average (ARIMA) model and the backpropagation neural network (BPNN) model. The suggested LSTM network beat the classic models in prediction accuracy, according to the results. The research discloses the dynamic evolution law of traffic flow, and facilitates the decision-making of traffic management.

Yudi Xu and Yang Yang [6], presents a hybrid model called SpAE-LSTM that considers the temporal and spatial features of traffic flow and it consists of sparse autoencoder that extracts the spatial features within the spatial-temporal matrix via full connected layers and long short-term memory (LSTM) network based on memory units that cooperates with the LSTM network to capture the spatial-temporal features of traffic flow evolution and make prediction. As a result, proposed model SpAE-LSTM effectively captures the spatial-temporal features of the traffic flow and achieves promising results.

Zainab Abbas et. al., compared three models for short-term road traffic density prediction based on Long Short- Term Memory (LSTM) neural networks. In order to deal with the challenge of scale and to improve prediction accuracy, they proposed to partition the road network into road stretches and junctions, and to model each of the partitions with one or more LSTM neural networks. Also shows reduced complexity of LSTM network by limiting the number of input sensors, on average to 35% of the original number, without compromising the prediction accuracy [11].

3. Objective:

Predicting traffic flow in real time is a very challenging task. We are using a time series model to forecast accurate traffic predictions with efficiency. This work attempts to analyze the performances of different predictive time-series models for predicting traffic flow. It is important for forecasting the volume and density of traffic flow for managing the traffic by reducing vehicle movement and generating the optimal route. For this objectives are:

1. To identify and analyze temporal dependency models to predict traffic flow with respect to static and dynamic behavior of traffic systems.
2. Analyze the different time series models in context of different parameters to identify and evaluate better prediction results.

The results of the analyzed time series model for predicting traffic flow have been compared for relative levels of accuracy with the help of different evaluation matrices. This is the initial step in developing an insight for choosing the best suitable model to estimate future traffic flow which is imperative from many aspects. This study proves the Time series analysis is a better alternative to traditional methods of predicting traffic flow. In general, this study attempts to estimate the appropriateness of Time Series forecasting techniques for traffic flow prediction.

4. Discussion

4.1 Data Analysis:

Effective time-series models require quality training and testing data to make accurate predictions.

Training data. ML algorithms require training data to achieve an objective. The algorithm will analyze this training dataset, classify the inputs and outputs, and then analyze it again. Prepared sufficient, a calculation will basically memorize all of the inputs and yields in a preparing dataset - this gets to be an issue when it should consider information from other sources, such as real-world clients [26].

Validation data. During preparing, approval information implants unused information into the model that it hasn't assessed some time recently. Approval information gives the primary test against validation data, allowing information researchers to assess how well the model makes forecasts based on the unused information [26]. Not all information researchers utilize approval information, but it can give a few supportive data to optimize hyper parameters, which impact how the show surveys information.

Validation data gives an introductory check that the demonstrator can return valuable forecasts in a real-world setting, which preparing information cannot do. The ML model can evaluate preparing information and approval information at the same time.

Test data. After the model is built, testing data once again approves that it can make precise expectations. In case training and validation data include labels to monitor performance metrics of the model, the testing information ought to be unlabeled [26]. Test data provides a final, real-world check of an unseen dataset to confirm that the ML algorithm was trained effectively.

High-quality datasets are essential for accurate traffic forecasting. In this section, we summarize some of the public data information used for the prediction task. TABLE I represents the list of commonly used public and large-scale real-world datasets used in traffic prediction [25].

Table1. List of Public Datasets [25]

Data Set	Description	Source

PeMS	California Transportation Agency Performance Measurement System (PeMS), which is displayed on the map and collected in real-time by more than 39000 independent detectors.	http://pems.dot.ca.gov/.
METR-LA	It records four months of statistics on traffic speed, ranging from 3/1/2012 to 6/30/2012, including 207 sensors on the highways of Los Angeles County.	https://github.com/liyaguang/DCRNN
Los-loop	This dataset is collected in the highway of Los Angeles County in real time by loop detectors. It includes 207 sensors and its traffic speed is collected from 3/1/2012 to 3/7/2012. These traffic speed data is aggregated every 5 minutes includes 207 sensors and its traffic speed is collected from 3/1/2012 to 3/7/2012. These traffic speed data is aggregated every 5 minutes.	https://github.com/lehaifeng/T-GCN/tree/master/data
TaxiBJ	Trajectory data is the taxicab GPS data and meteorology data in Beijing from four time intervals: 1st Jul. 2013 - 30th Otc. 2013, 1st Mar. 2014 -30th Jun. 2014, 1st Mar. 2015 - 30th Jun. 2015, 1st Nov. 2015 - 10th Apr. 2016.	https://github.com/lucktroy/DeepST/tree/master/data/TaxiBJ
SZ-taxi	This is the taxi trajectory of Shenzhen from Jan.1 to Jan.31, 2015. It contains 156 major roads of Luohu District as the study area. The speed of traffic on each road is calculated every 15 minutes.	https://github.com/lehaifeng/TGCN/tree/master/data
NYC Bike	The bike trajectories are collected from NYC CitiBike system. There are about 13000 bikes and 800 stations in total.	https://github.com/lucktroy/DeepST/tree/master/data/Bi keNYC
NYC Taxi	The trajectory data is taxi GPS data for New York City from 2009 to 2018.	https://www1.nyc.gov/site/tlc/about/tlc_triprecord-data.page
Q-Traffic	It consists of three sub-datasets: query sub-dataset, traffic speed sub-dataset and road network sub-dataset. These data are collected in Beijing, China between April 1, 2017 and May 31, 2017, from the Baidu Map.	https://github.com/JingqingZ/BaiduTraffic#Dataset
ENG-HW	It contains traffic flow information from inter-city highways between three cities, recorded by British Government, with a time range of 2006 to 2014.	http://tris.highwaysengland.co.uk/detail/trafficflowdata
DiDi chuxing	DiDi gaia data open program provides real and free desensitization data resources to the academic community. It mainly includes travel time index, travel and trajectory datasets of multiple cities.	https://gaia.didichuxing.com

2841

4.2 Real-Time Traffic Data:

Real-Time Traffic comprises highly accurate data from multiple sources, including connected car probes, roadway sensors, and live operations centers [23]. Real-Time Traffic contains three types of traffic data:

Flow: information about the speed of travel space headway, location, travel time, occupancy, demand and congestion along a segment of a roadway. Flow data is updated every minute.

Incidents: information about events includes various holidays, traffic control, traffic accidents, sports events, concerts and other activities, affecting the flow of traffic or that may be important for drivers to know. Incident data is updated every two minutes.

Seasonal: information about climate includes temperature, humidity, wind speed, visibility and weather state (sunny/rainy/windy/cloudy etc.). Seasonal data is updated according to weather conditions.

4.3 Experimental Data Analysis:

In this section we will have to do data analysis of time series data. For this purpose we have chosen the EDA technique to find out the pattern and structure of the data. Factors related with the Patterns and structure are data with some seasonality, linearity of data, increasing and decreasing trend in the data, data is stationary or non-stationary Figure1. All these factors are important for the modeling purpose. In case of TS data it is important to understand the structure of the data, so that we can impute the right modeling technique. For this we have to do multivariate and univariate analysis. We have to plot Autocorrelation (ACF) plot and partial Autocorrelation (PC) plot Figure 2. Visualization of dataset in represent in Figure 3 with respect to date and number of vehicles at different junctions.

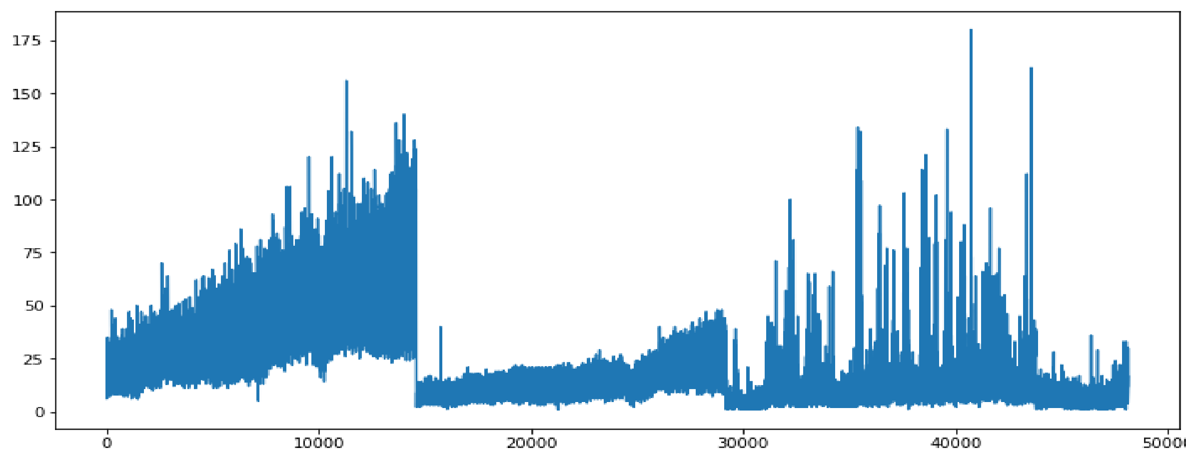
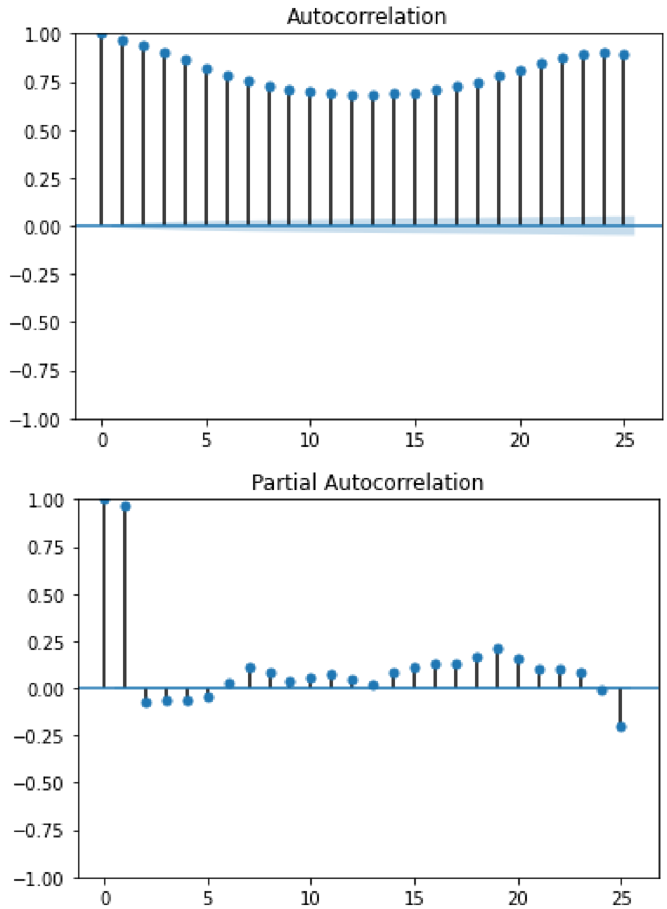


Figure1. Stationary and Seasonality



(a) (b)

Figure2. (a) ACF plot and (b) PACF plot.

Figure3. Visual Representation of Dataset

4.4 Methodology Adopted:

It is a contextual framework that we need to process the research work. Figure 4 shows the process to be used for
the research consist of 3 main stages:

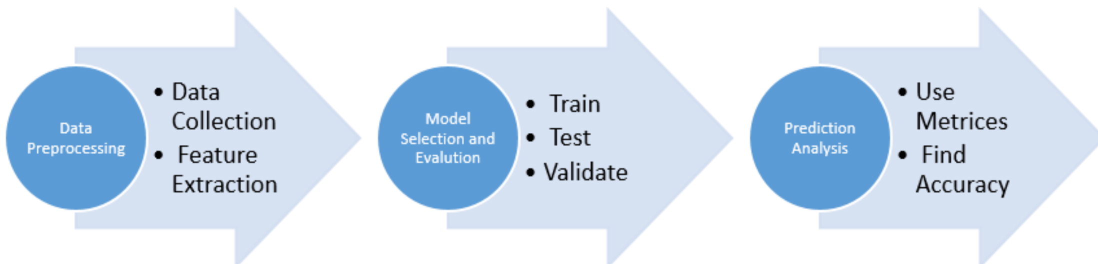


Figure4. Research Methodology Analysis

Data Pre-Processing:
Data Collection from different sources is must for processing the data. Data that has been well-prepared for your model can help it run more efficiently. It can aid in decreasing the model's blind areas, resulting in improved forecast accuracy. As a result, it's a good idea to think about and examine your data sets so that they can be fine-tuned to• produce more accurate and useful findings.

- Wrangle data and prepare it for training lean that which may require it (remove duplicates, correct errors, deal with missing values, normalization, data type conversions, etc.)
- Randomize data, which erases the effects of the particular order in which we collected and/or otherwise prepared our data
- Visualize data to help detect relevant relationships between variables or class imbalances, or perform other exploratory analysis
- Split into training and evaluation sets

Feature Extraction is a technique for reducing the number of features in a dataset by generating new ones from existing ones (and then discarding the original features). The original set of features should then be able to summarize the majority of the information in the new reduced set of features. In this way, a summarized version of the original features can be created from a combination of the original set.

2844

Model Selection and Evaluation:
There are three types of temporal dependency models most commonly used in epidemic time series forecasting and prediction, which includes:
Long-Short Term Memory (LSTM),
Autoregressive Integrated Moving Average (ARIMA), and

- Seasonal Autoregressive Integrated Moving Average (SARIMA) model.

Train and Test the selected model using machine learning techniques. After training the model we will evaluate the result for different parameters.
Prediction Analysis:
After evaluation we will predict the traffic forecasting for real time traffic in terms of evaluation metrics (RMSE, MAE, MAPE) to find accuracy.

4.5 Index of Performance/Errors Metrics:
The commonly used accuracy metrics for Time Series Forecast to judge forecasts are:

- Mean Absolute Percentage Error (MAPE)
- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)

Typically, if you are comparing forecasts of different series, the MAPE (Mean Absolute Percentage Error), the MAE (Mean Absolute Error) and the RMSE (Root Mean Square Error) can be employed to measure the performance of different methods.

4.6 Different Modeling techniques:

There are various methods of doing Time series Analysis and Forecasting.

LSTM: Most popular model in time-series domain. It is a class of RNN. RNN is a class of NN tailored to deal with temporal data, Useful to process and predict events with time series. Difficult to solve exceedingly long-term dependencies because LSTM errors increase as the sequence length increases. It is designed to recognize pattern in sequences of data. It takes time and sequence into account; they have a temporal dimension [19]. LSTM works on Multivariate time-series. LSTM takes time and sequence into account; they have temporal dimensions. It possesses a certain type of memory and memory is a part of the human condition, we will make repeated analogies to memory in the brain. LSTM help preserve the error that can be back propagation through time and layers [7]. LSTM contains information outside the normal flow of the RN in a gated cell. Information can be stored in written to or need to read from a cell, much like data in a computer's memory.

- Use to find out total percentage reduction in prediction error.
- Percentage reduction in error during morning rush hour.
- Percentage reduction in error during evening rush hour.

Advantages of LSTM

1. No prerequisites (stationary, no level shifts).
2. Can model non-linear functions with neural networks.
3. Need a lot of data.

ARIMA: AR predicts the value of future time period as a function of values at previous time period, but does not account for seasonality. ARIMA works on univariate time-series. It is a time-domain approach, models future values as a function of past values and present values the time series regression of present values of a time series on its own past values and past values of other variables is the cornerstone of this approach. These regressions' estimations are frequently utilized for forecasting, and this method is widely employed in time series econometrics [8].

ARIMA stands for 'Auto Regressive Integrated Moving Average,' and it is a forecasting technique based on the premise that past values of a time series can be utilized alone to predict future values.

ARIMA models can be used to model any 'non-seasonal' time series that has patterns and isn't random white noise.

An ARIMA model is characterized by 3 terms: p, d, q Where, p is the order of the AR term, q is the order of the MA term, d is to make the time series stationary, for the number of differencing is required.

- Use ARIMA to predict per road segment future speeds based on previously observed values.
- Can model hour-of-week and day-of- week patterns.
- Cannot handle non-periodic incidents.

Advantages of ARIMA

1. Simple to implement, no parameter tuning.
2. Easier to handle multivariate data.
3. Quick to run.

SARIMA: SARIMA is also capable of working with univariate time series. If a time series has seasonal trends, seasonal terms must be added, and the time series becomes SARIMA, short for 'Seasonal ARIMA.'

Seasonality in Time Series: It's critical to recognize the presence of seasonality in time series. Incorrect models and interpretations might result from failing to detect the regular and predictable patterns of seasonality in time series data [15].

How to Identify Seasonality: Seasonality in time series data must be identified in order to construct a meaningful time series model. There are many tools that are useful for detecting seasonality in time series data: Background theory and knowledge of the data can provide insight into the presence and frequency of seasonality [12]. Seasonal subseries plots, autocorrelation plots, and spectral plots are examples of time series plots that can help uncover clear seasonal trends in data [23]. Seasonality can be detected using statistical analysis and tests such as the autocorrelation function, periodograms, and power spectrums.

- Support time series with a seasonal component.
- Used on univariate data containing trends and seasonality.
- Designed to predict errors at different times of the day round the year.

Dealing with Seasonality in Time Series Data: Once seasonality is identified, to deal with its presence, the appropriate procedures must be done. Seasonality in time series data can be addressed in a number of ways:

Choose a model that incorporates seasonality, like the Seasonal Autoregressive Integrated Moving Average (SARIMA) models.

1. Remove the seasonality by seasonally detrending the data or smoothing the data using an appropriate filter.

If the model is going to be used for forecasting, the seasonal component must be included in the forecast.

2. Use a seasonally adjusted version of the data. The Bureau of Labor Statistics, for example, provides labor and employment data for the United States, with many series available in both seasonally adjusted and non- seasonally adjusted versions.

TABLE 2 and TABLE 3 represent the comparison between selected models. One major difference between these models is that ARIMA could only perform well on stationary time series where ARIMA concentrates on seasonality and trend. LSTM is useful for dealing with huge amounts of data, while ARIMA is better for smaller datasets.

Table2. Comparison between ARIMA, SARIMA, LSTM on the basis of Basic Parameters

Basic Parameters / Models	Type	Time	Matrices	Data	Variables
ARIMA	Use ARIMA to predict per road segment future speeds based on previously observed values.	Can model hour-of-week and day-of- week patterns.	Cannot handle non-periodic incidents.	ARIMA requires a series of parameters which must be calculated based on data.	Support Multivariate
SARIMA	Support time series with a seasonal component.	Used on univariate data containing trends and seasonality.	Designed to predict errors at different times of the day round the year.	SARIMA used the data contains s	Univariate: only one variable varying over time. Or you will have only a one-dimensional value, which is the temperature.

LSTM	Use to find out total percentage reduction in prediction error.	Percentage reduction in error during morning rush hour.	Percentage reduction in error during evening rush hour.	This model is useful for dealing with huge amounts of data and for training data.	Support Multivariate i.e.multiple variables are varying over time.

Table3. Comparison between ARIMA, SARIMA, LSTM on the basis of Implementation Parameters

Implementation Parameters / **Models**	**Prediction Method**	**Model- Based**	**Support nature of Time-Series**	**Hyper parameters**
ARIMA	Short-term	Traditional algorithms model	Stationarity	(p, d, q): A tuple p, d, and q parameters for the modeling of the trend.
SARIMA	Long-term	Variant of ARIMA has seasonality.	Stationarity, Seasonality	(P, D, Q and m): A tuple of P, D, Q, and m parameters for modeling the seasonality.
LSTM	Both	Deep learning-based algorithms model	Non-linear and Non-stationary	Number Of Nodes and Hidden Layers, Number Of Units In A Dense Layer, Dropout.

ARIMA and SARIMA require a series of parameters which must be calculated based on data, while LSTM does not require setting such parameters. However, there are some hyper parameters we need to tune for LSTM. TABLE 4 summarized available predictive performances of ARIMA and LSTM for traffic flow by different authors. Same is not available for SARIMA.

Table4. Available Predictive performances of different algorithms for traffic flow

Paper Reference	**ARIMA**			**LSTM**		
	RMSE	**MAE**	**MAPE (%)**	**RMSE**	**MAE**	**MAPE (%)**
[1]	24.151	10.502	0.158	16.798	8.698	0.108
[2]	149.52	-	13.28	-	-	-
[6]	1.910	-	9.539	1.398	-	6 .841
[7]	30.86	20.59	-	0.044	0.04	-

[11]	7.67	6.74	-	5.49	2.41	-
[24]	7.16	4.22	7.62	7.73	4.79	8.51

5. Conclusion

In this research paper we discussed the terms related with traffic forecasting and prediction and performed experimental data analysis (EDA) for checking compatibility of data sets with selected models. For doing analysis and taking Future decisions Time Series forecasting is really convenient. The problem is dealing with predicting time series values. So far, we have come across three models ARIMA, SARIMA and LSTM from which we can quickly do the time series forecasting and also predict future values.

After analyzing these models it has been found that LSTM is useful to process and predict events with time series. And by using LSTM it is difficult to solve exceedingly long term dependencies. LSTM errors increase as the sequence length increases. Further, notified that SARIMA performs time series analysis with seasonal parameters this can not be performed using ARIMA.

REFERENCES

[1] Saiqun Lu, Qiyan Zhang, Guangsen Chen, Dewen Seng, A combined method for short-term traffic flow prediction based on recurrent neural networks. Received 21 April 2020; revised 1 June 2020; accepted 4 June 2020.

[2] Umuhoza Kibogo Aimee Vanessas, Kong Yan, Evaluating the Performance of LSTM in Traffic Flow Prediction at Different Time Scales. International Journal of Engineering Research & Technology (IJERT), Vol. 10 Issue 03, March-2021.

[3] Roberto Casado-Vara, Angel Martin del Rey, Daniel Pérez-Palau and Luis de-la-Fuente-Valentín and Juan M. Corchado, Web Traffic Time Series Forecasting Using LSTM Neural Networks with Distributed Asynchronous Training. Mathematics 2021, 9, 421. https://doi.org/10.3390/math904042.

[4] Mizanur Rahman, Mashrur Chowdhury, Real-time Traffic Flow Parameters Prediction with Basic Safety Messages at Low Penetration of Connected Vehicles. IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. Manuscript received February 20, 2019.

[5] JIANHU ZHENG AND MINGFANG HUANG, Traffic Flow Forecast Through Time Series Analysis Based on Deep Learning. Received April 10, 2020, accepted April 19, 2020, date of publication April 27, 2020, date of current version May 15, 2020.

[6] Fei Lin , Yudi Xu , Yang Yang, and Hong Ma, A Spatial-Temporal Hybrid Model for Short-Term Traffic Prediction. Mathematical Problems in Engineering Volume 2019, Article ID 4858546, 12 pages https://doi.org/10.1155/2019/4858546.

[7] Bailin Yanga, Shulin Suna, Jianyuan Li, Xianxuan Lina, Yan Tiana, Traffic flow prediction using LSTM with feature enhancement. © 2018 Elsevier B.V. All rights reserved.

[8] Soheila Mehrmolaei, Mohammad Reza Keyvanpour, Time series forecasting using improved ARIMA. 978-1- 5090-2169-7/16/$3l.00 ©2016 IEEE.

[9] Sumanta Bhattacharyya Arindrajit Seal Arindam Mukherjee, Real-Time Traffic Incidence dataset. 978-1-7281- 0137-8/19/$31.00 ©2019 IEEE.

[10] Xianglong Luo, Danyang Li, Yu Yang, and Shengrui Zhang, Spatiotemporal Traffic Flow Prediction with KNN and LSTM. Journal of Advanced Transportation Volume 2019, Article

ID 4145353, 10 pages https://doi.org/10.1155/2019/4145353.

[11] Zainab Abbas, Ahmad Al-Shishtawy, Sarunas Girdzijauskas, Vladimir Vlassov, Short-Term Traffic Prediction Using Long Short-Term Memory Neural Networks. 978-1-5386-7232-7/18/$31.00 ©2018 IEEE. DOI 10.1109/BigDataCongress.2018.00015.

[12] Ling Zhao, Yujiao Song, Chao Zhang, Yu Liu , Pu Wang, T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. 1524-9050 © 2019 IEEE.

[13] Huaxiu Yao, Yanwei Yu, Modeling Spatial-Temporal Dynamics for Traffic Prediction. © Copyright held by the owner/author(s). ACM ISBN 123-4567-24-567/08/06. https://doi.org/10.475/123_4. March 2018.

[14] Vinayakumar R, Soman KP and Prabaharan Poornachandran, Applying Deep Learning Approaches for Network Traffic Prediction. 978-1-5090-6367-3/17/$31.00 ©2017 IEEE.

[15] Xianglei Zhu and Shuai, Zhao Yaodong Yang, Hongyao Tang, Zan Wang and Jianye Hao, A Real-time Ensemble Classification Algorithm for Time Series Data. 2017 IEEE International Conference on Agents (ICA).

[16] Mizanur Rahman, Mashrur Chowdhury, Senior Member, IEEE, and Jerome McClendon, Real-time Traffic Flow Parameters Prediction with Basic Safety Messages at Low Penetration of Connected Vehicles. IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. Manuscript received February 20, 2019.

[17] Bhaskar Chowdhury, Maithili Kinhikar, Mrs. N. Noor Alleema, Road Traffic Prediction using Machine Learning. International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 06 Issue: 04 | Apr 2019.

[18] Shen Fang, Qi Zhang, Gaofeng Meng, Shiming Xiang and Chunhong Pan, GSTNet: Global Spatial-Temporal Network for Traffic Flow Prediction. Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19).

[19] Rui Fu, Zuo Zhang, and Li Li, Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction. Research gate November 2016. DOI: 10.1109/YAC.2016.7804912.

[20] Wasim Hashmi Syeda, Ansar Yasarb, Davy Janssensb, Geert Wets, Analyzing the Real Time Factors: Which Causing the Traffic Congestions and Proposing the Solution for Pakistani City. ScienceDirect Procedia Computer Science 32 (2014) 413 – 420.

[21] Aniekan Essien & Ilias Petrounias & Pedro Sampaio & Sandra Sampaio, A deep-learning model for urban traffic flow prediction with traffic events mined from twitter. World Wide Web (2021) 24:1345–1368.

[22] Noor Afza Mat Razali , Nuraini Shamsaimon , Khairul Khalil Ishak, Suzaimah Ramli, Mohd Fahmi Mohamad Amran and Sazali Sukard, Gap, techniques and evaluation: traffic flow prediction using machine learning and deep learning. Razali et al. Journal of Big Data (2021) 8:152. https://doi.org/10.1186/s40537-021-00542-7.

[23] Wei Zhou, Wei Wang, Xuedong Hua and Yi Zhang, Real-Time Traffic Flow Forecasting via a Novel Method Combining Periodic-Trend Decomposition. Sustainability 2020, 12, 5891; doi:10.3390/su12155891.

[24] BASHARAT HUSSAIN, MUHAMMAD KHALIL AFZAL, SHAFIQ AHMAD AND ALMETWALLY M. MOSTAFA, Intelligent Traffic Flow Prediction Using Optimized GRU Model. Received June 8, 2021, accepted July 10, 2021, date of publication July 14, 2021, date of current version July 22, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3097141.

[25] Xueyan Yin, Genze Wu, Jinze Wei, Yanming Shen, Heng Qi, and Baocai Yin (2020). A Comprehensive Survey on Traffic Prediction arXiv: 2004.08555v2.

[26] Machine Learning – Training, Validation & Test Data Set June 13, 2021 by Ajitesh Kumar. https://vitalflux.com/machine-learning-training-validation-test-data-set/.