**Dominance of Artificial Intelligence and Machine Learning Algorithms in Real-Time Traffic Flow prediction and Route Optimization in Autonomous vehicles.**

**Submitted by**

**Degree Program**

**Institution  
University Name**

**University Department**

**Advisor  
Dr.**

**Date of Submission  
Month, Year**

# Abstract

The expansion of the Internet of Things has resulted in new creative solutions, such as smart cities, that have made our lives more productive, convenient, and intelligent. The core of smart cities is the Intelligent Transportation System (ITS) which has been integrated into several smart city applications that improve transportation and mobility. ITS aims to resolve many traffic issues, such as traffic congestion issues. Recently, new traffic flow prediction models and frameworks have been rapidly developed in tandem with the introduction of artificial intelligence approaches to improve the accuracy of traffic flow prediction. Traffic forecasting is a crucial duty in the transportation industry. It can significantly affect the design of road constructions and projects in addition to its importance for route planning and traffic rules. Furthermore, traffic congestion is a critical issue in urban areas and overcrowded cities. Therefore, it must be accurately evaluated and forecasted. Hence, a reliable and efficient method for predicting traffic is essential. The main objectives of this study are: First, present a comprehensive review of the most popular machine learning and deep learning techniques applied in traffic prediction. Second, identifying inherent obstacles to applying machine learning and deep learning in the domain of traffic prediction.

Table of Contents

[Abstract 2](#_Toc175669575)

[1. Introduction 4](#_Toc175669576)

[Issues for traffic congestion 4](#_Toc175669577)

[Intelligent transportation system (ITS) 5](#_Toc175669578)

[Data types, sources and TFP parameters 6](#_Toc175669579)

[Traffic theory 7](#_Toc175669580)

[Traffic flow 7](#_Toc175669581)

[Traffic density 7](#_Toc175669582)

[Free flow 7](#_Toc175669583)

[2. Literature Review 8](#_Toc175669584)

[3. Methodology 10](#_Toc175669585)

[Statistical approach 10](#_Toc175669586)

[Machine learning approach 10](#_Toc175669587)

[Deep learning approach 11](#_Toc175669588)

[Comparing the performance of our model with other novel algorithms 12](#_Toc175669589)

[LSTM 12](#_Toc175669590)

[CNN 12](#_Toc175669591)

[CNN VS LSTM 12](#_Toc175669592)

[LSTM RNN 12](#_Toc175669593)

[Data set 13](#_Toc175669594)

[The Metro Interstate dataset 13](#_Toc175669595)

[Deep learning approach (LSTM and RNN) 13](#_Toc175669596)

[4. Test Result and Analysis 14](#_Toc175669597)

[5. Conclusion 19](#_Toc175669598)

[References 20](#_Toc175669599)

# 1. Introduction

Traffic flow prediction (TFP) means predicting the volume and density of traffic flow, usually to control vehicle movement, reduce traffic jams, and create the optimal (least-time or energy-consuming) route. With the recent advancement in Artificial intelligence, Machine learning (ML), Deep learning (DL), and Big data, research in the field of predicting traffic flow has been expanded extensively.  
TFP is the key component of Intelligent Transport Systems (ITS) and can assist ITS to forecast traffic flow. Large cities have exceedingly difficult traffic regulations. many countries have adopted ITS to reduce the costs associated with traffic congestion. This study reviews the application of artificial neural network (ANN), ML, DL and other techniques and models for TFP. Finally, we will propose our own predictive model using DL, train and test it, analyze the accuracy and compare the accuracy of our model with other models.  
Comparing with conventional ML methods, DL models have the advantages such as simplifying data preprocessing procedures and outperforming other ML methods in terms of accuracy. Therefore, data-driven traffic flow prediction due to the availability of massive traffic data and DL schemes due to data preprocess procedures have received extensive attention recently in TFP.  
Furthermore, In recent years there has been a vast increase in available data with the advancement of smart cities. This modernization can have a favorable impact on transportation networks in the area of ITS, reducing travel times, boosting productivity, and minimizing the environmental impact of vehicles. ML and DL technologies are fast-growing domains for predicting traffic flow. Traffic signals, accidents, weather conditions, and road repairs are the primary causes of traffic. Since real-time traffic data are largely produced exponentially, big data principles must be used to improve data transportation. This fact motivated us to to predict the volume of traffic flow between Minneapolis and St. Paul at a specific point in Minnesota. Our aim is to build a multi-step Recurrent Nural Network (RNN) with Long Short-Term Memory (LSTM) model that makes a single prediction point of the traffic volume 2 hours into the future, given the previous 6-hour window.

## Issues for traffic congestion

In many places across the world, traffic congestion is a serious issue. It is difficult to accurately define and forecast because it is influenced by a wide range of diverse elements. Some of the main problems caused by traffic congestion are air pollution, lengthy commutes, traffic jams, unreliable public transportation, and traffic accidents.  
To create smooth traffic flow conditions on road networks, we must address the major issue of traffic congestion. Due to rising traffic demands, expanding the road network infrastructure by widening the roadways was simply insufficient to manage the smooth traffic flow conditions. To model the traffic flow circumstances, some type of traffic flow modelling methodology is needed that we will try to implement in this project.

## Intelligent transportation system (ITS)

ITS is an integrated traffic management system composed of advanced data communication, information processing and traffic management technologies. In recent years, the success of deep learning in computer vision, speed recognition, and natural language processing makes it natural to apply it to ITS. We divide the applications in ITS into visual recognition tasks, TFP, traffic speed prediction (TSP), travel time prediction (TTP) and other tasks. The term "ITS" refers to the use of systems for communication, information, transportation, and urban transportation. Two of ITS's primary objectives are efficiency and traffic safety. Reduced intersection stalls and delays, improved traffic times, improved speed control, capacity management, and incident management are all benefits of an ITS. Figures 1 depicts numerous tasks that are covered by ITS.

A diagram of a transportation system

Description automatically generated

## Data types, sources and TFP parameters

Traffic is influenced by many factors that we need to consider for accurate predictions. So, there are several main groups of data that we'll have to obtain.  
Mapping data  
First of all, we need to have a detailed map of road networks with related attributes. Connecting to such global mapping data providers as Google Maps, TomTom, HERE, or OSM is a great way to obtain complete and up-to-date information.  
Traffic information  
Then, we need to collect both historical and current traffic-related information such as the number of vehicles passing at a certain point, their speed, and type (trucks, light vehicles, etc.). Devices used to collect this data are loop detectors, cameras, weigh in motion sensors, and radars, or other sensor technologies. Fortunately, we don’t have to install these devices all over the place on our own. It’s easier to get this information from the aforementioned providers that gather data from a system of sensors, diverse third-party sources, or make use of GPS probe data. Other platforms such as Otonomo use an innovative Vehicle to Everything (V2X) technology to collect so-called connected car data from embedded modems. We can also get other important information on incidents (road closures or roadworks), places of interest, etc., from data providers.  
Weather information  
Weather data (historical, current, and forecasted) is also necessary as meteorological conditions impact the road situation and driving speed. There are lots of weather data providers such as OpenWeather or Tomorrow.io.  
Additional data on road conditions

There are external data sources that can provide important information that impacts traffic such as social media posts, local news, or even police scanners.

In short, to anticipate TFP, we consider a multi-parameters prediction approach that takes into account traffic patterns in a variety of ways.  
**a) Flow**  
The amount of traffic that passes through a particular spot on the road in a given amount of time is referred to as the flow of traffic.  
**b) Speed**  
The distance traveled per unit of time determines a vehicle's speed. In most cases, the speed of any vehicle on the road will differ from others around it due to factors such as the driver's position and the traffic conditions.  
**c) Day**  
The day can be Sunday to Saturday.  
**d) Day of type**  
The day of type is mainly described as public holiday, weekend, and working.  
**e) Clock time**  
The clock time can be divided into hours, a total of 24 hours (1- 24 hour).  
**f) Weather condition**  
Weather data such as sunny & rain can be taken to training and perdition purposes.

### Traffic theory

In this section, the variables traffic flow and density is introduced. The concept of free flow is also explained.

### Traffic flow

Traffic flow is defined as the number of vehicles per unit time in a reference point.  
n/∆T

### Traffic density

The traffic density (k) is the number of vehicles present per length of the road. For a given road this means that we have kc defined as the critical density. This is when the road has peak traffic flow. When the density goes further up the traffic flow will decrease and at some point reach jam density kj. Then max congestion have peaked.

### Free flow

When density k is less than kc the traffic is said to be in a free flow state. Traffic flow can also be viewed as flow = speed ∗ density. In that regard, the traffic state can also be understood by the mean speed of the reference point.

# 2. Literature Review

The survey of literature on TFP and ITS was conducted using several secondary sources of information such as articles, Books, and Research Reports published in various publications. The collected information is then reviewed to discover any potential TFP and ITS significant areas of concern. The prediction of traffic congestion, in particular the short-term traffic forecast, is done by evaluating various traffic parameters. To predict traffic congestion, most research uses historical data. However, several reports offered real-time predictions of traffic congestion. In city transportation and area management, TFP has a wide range of applications. The TFP issue is a time series (TS) problem that involves estimating the urban road traffic flow at a future time using information gathered from one or more observation points during prior periods. Moreover, to predict traffic flow many AI techniques in particular various ML and DL models are applied. An ITS can effectively reduce traffic congestion by forecasting short-term traffic flow.

Graph neural networks (GNNs) are deep learning models that use graph data as input and are used for a variety of applications, including molecular property prediction and traffic forecasting. Google Maps employs machine learning to combine current traffic conditions with past traffic patterns for roads all around the world in order to reliably predict future traffic. The aim of traffic forecasting is to make traffic volume predictions using past volume and speed data. Authorities are paying more attention to observing traffic congestion due to the development of the transportation industry and the collection of traffic statistics. Traffic flow data is transformed into a 2D matrix for processing in traffic forecasting due to CNN's outstanding image processing capabilities. In recent years, new traffic flow prediction models and frameworks have been quickly developed to increase TFP performance in addition to applying artificial intelligence (AI) techniques like machine learning (ML). The majority of traffic data are small-scale, concentrated on highway traffic, or lacking in auxiliary data. Traffic conditions can be predicted using both online and offline data, notably searches on map apps.

B.Karthika et al. discussed the rich mobility data and deep learning about urban traffic predictions, Deep algorithms to forecast real-world traffic data and when the traffic data becomes big data, some techniques to improve the accuracy of trafficprediction are also discussed[1]. Apurv Chandel et al. used multiple linear regression models for the dataset and obtained best results in correspondence of gradient boosting regression after performing the tunning of hyperparameters with an accuraccy of 83%[2].  
Rohit Singh et al. in his research presents several regression machine learning techniques that combine the performance of various algorithms and compares it with existing Machine Learning Algorithms using the mean square error, root mean square error and other performance metrics[6]. Renhe Jiang et al. presented a comprehensive survey of all the deep learning models used for traffic prediction and implemented the dataset using pytorch and tensor flow for better prediction[7].  
Kanokwan Khiewwan et al. described data mining techniques for the implementation and used DT, KNN and SVM for modeling and found out that DT performed better than other algorithms with an accuracy of 79.9%[8].

# 3. Methodology

A diagram of a traffic prediction

Description automatically generated

**Algorithms for generating traffic predictions (MODELS):**

Traffic prediction involves forecasting drivable speed on particular road segments, as well as jam occurrence and evolution. Let's take a look at different approaches to this task.

Statistical approach  
Using statistical techniques, we can recognize traffic patterns at various scales, such as during the day, on various days of the week, seasonally, etc. Compared to machine learning methods, they are typically quicker, cheaper, and easier to execute. However, because they can't handle as much multivariate data, they are less accurate.  
Since the 1970s, auto-regressive integrated moving average (ARIMA) models—which are simple to use and show higher accuracy than other statistical techniques—have been extensively employed to predict traffic.  
It uses a traditional statistical methodology to analyze the past and forecast the future. It gathers data from a series of regular time intervals and makes the assumption that historical patterns will continue in the future. However, traffic flow is a complex structure with many variables that can't be effectively processed with the help of the univariate ARIMA models.

Machine learning approach  
We can build predictive models using machine learning (ML) that take into account the massive amounts of heterogeneous data from many sources. The use of ML algorithms to forecast traffic has been the subject of numerous studies. Here are a few effective examples.   The random forest method builds several decision trees and combines their data to produce precise forecasts. Given enough training data, it can generate effective results quickly. In this instance, the model's input variables include the weather, time, specific road conditions, road quality, and holidays.  
Furthermore, the k-nearest neighbors (KNN) method uses the idea of feature similarity to forecast future values.

Deep learning approach  
Deep learning (DL) methods have proved highly effective in predicting road traffic in comparison to ML or statistical techniques, consistently showing about 90 percent forecasting accuracy and higher. DL algorithms are based on neural networks.  
Artificial neural networks (ANN) or neural networks (NN) are made up of interconnected nodes (neurons) organized in two or more layers and are intended to mimic the behavior of the human brain. There are many types of neural networks developed for different purposes. Here are some that were used in traffic analysis and prediction. Convolutional neural networks (CNNs) are regarded as industry pioneers in image analysis and recognition. Congestion detection utilizing images from on-road surveillance cameras is one of its natural applications to transportation issues. CNN's are not the first choice for traffic forecasts. However, attempts to develop CNN-based models for predicting transportation network speed were highly effective. To do this, researchers created a 2-dimensional picture matrix from time and space data characterizing traffic flow.

Recurrent neural networks (RNNs), as compared to convolutional neural networks (CNNs), are designed to analyse time-series data or observations gathered over specific time intervals. Such insights can be seen well in traffic patterns. RNN models have been shown to anticipate congestion evolution with excellent accuracy. The vanishing gradient problem, which is why RNNs are considered to "have a short-term memory", is its disadvantage because it causes some of the data from earlier layers to be lost. Model training is more difficult and time-consuming as a result of this "forgetfulness". Long short-term memory (LSTM) and gated recurrent unit (GRU) are variations of the RNN that address the vanishing gradient problem. A study that compared the performance of these models showed that the GRU model is more accurate in traffic flow predictions and is easier to train.

Numerous studies have recommended developing different NN models for traffic prediction, including graph neural networks, fuzzy NNs, Bayesian NNs, and others, as well as utilizing hybrid techniques that integrate two or more algorithms. As of right now, there isn't a single ideal technique that can be used in all situations to produce the most precise projections.  
There are a couple more things to mention in regards to implementing ML techniques for traffic prediction. We have to remember that ML/DL algorithms work best when there is sufficient data to train the models and fine-tune them to achieve maximum accuracy. So, the bigger datasets we manage to obtain the better results we will get.

## Comparing the performance of our model with other novel algorithms

LSTM  
Short-term traffic flow prediction is critical for intelligent transportation systems, and additional spatial and temporal traffic information might be used in future studies to accurately estimate traffic on a larger scale road network. In general, the more layers the LSTM has, the better the model's learning capacity, but it is also more susceptible to overfitting.

CNN  
CNN, or convolutional neural network, adds extra "filtering" layers where the filter weights are determined. Back propagation is still performing this task for us, but we will not make it too easy for the dependable engine that is backprop.  
A CNN has numerous parallel filters that may be adjusted to extract various aspects of interest.

CNN VS LSTM  
CNN may be used to minimize the number of parameters required for training while maintaining performance – this is the power of combining signal processing with deep learning, while LSTM requires more parameters than CNN, its benefit derives from being able to examine at lengthy sequences of inputs without expanding the network size.

### LSTM RNN

Traditional RNN architecture has the so called vanishing gradient problem. To overcome such disadvantage, certain structure of RNNs such as LSTM were proposed, which was designed to give the memory cells ability to determine when to forget certain information, thus determining the optimal time lags for time series problems. These features are particularly desirable for short-term traffic flow prediction in the transportation domain because of its long-standing memory ability .Using LSTM RNN for traffic flow prediction and showes that LSTM RNN have better performance than most of the non parametric models.

## Data set

The Metro Interstate dataset  
The Interstate Traffic Volume Dataset contains information about the hourly traffic volume on the West-bound lane of Interstate-94 (I-94) in the US. The dataset includes hourly weather and temperature reports from 2012 to 2018. The information in the dataset can be used to understand the flow of traffic on the interstate with respect to time and date and can be helpful in prediction of rush hours, weather forecasting as well as planning expansions of interstates and highways in the US. Furthermore, Hourly weather features and holidays are also included for impacts on traffic volume.

### Deep learning approach (LSTM and RNN)

The goal in this project is to build multi-step RNN with LSTM model that makes a prediction of the traffic volume 2h into the future, given 6h of history in this data set.

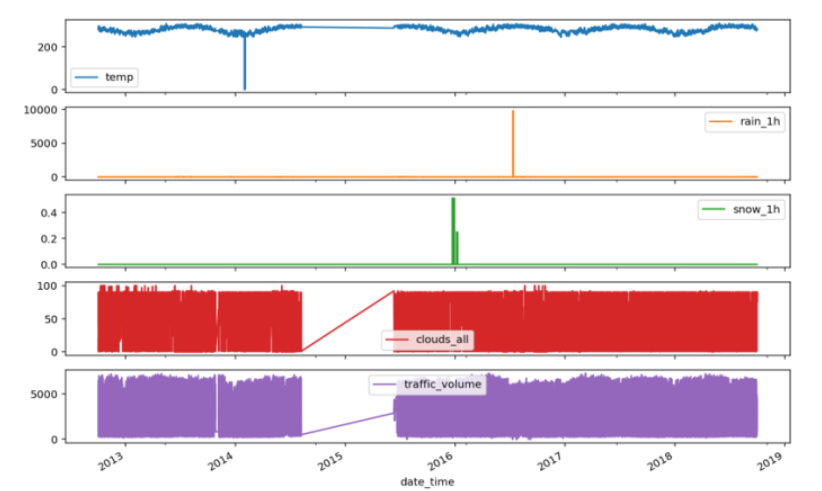
# 4. Test Result and Analysis

The Metro Interstate Traffic Volume Data Set is a hourly Interstate 94 Westbound traffic volume for MN DoT ATR station 301, roughly midway between Minneapolis and St Paul, MN. Hourly weather features and holidays were included for impacts on traffic volume. Our main aim was to build a multi-step RNN with LSTM model that makes a single prediction point of the traffic volume 2 hours into the future, given the previous 6-hour window. This can be demonstrated in (Figure), where the inputs represent the data point for a given 6 hours, and the label is the expected output 2 hours later.

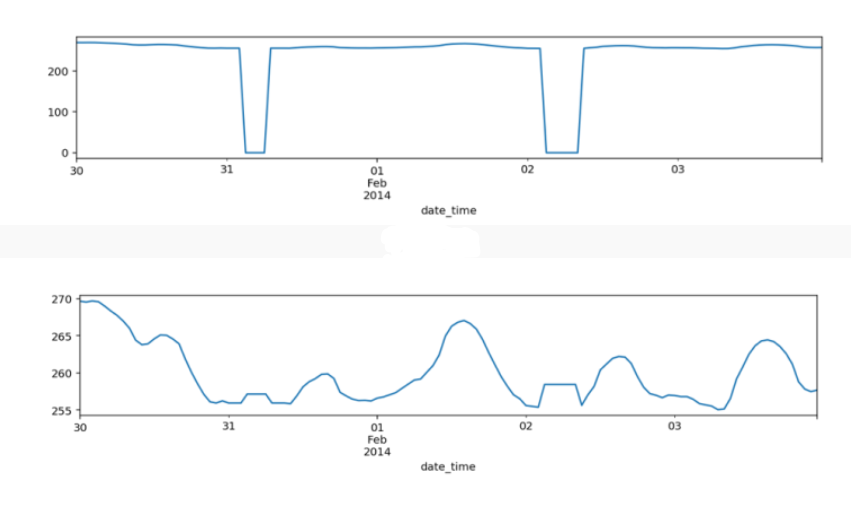
A line graph with a blue line

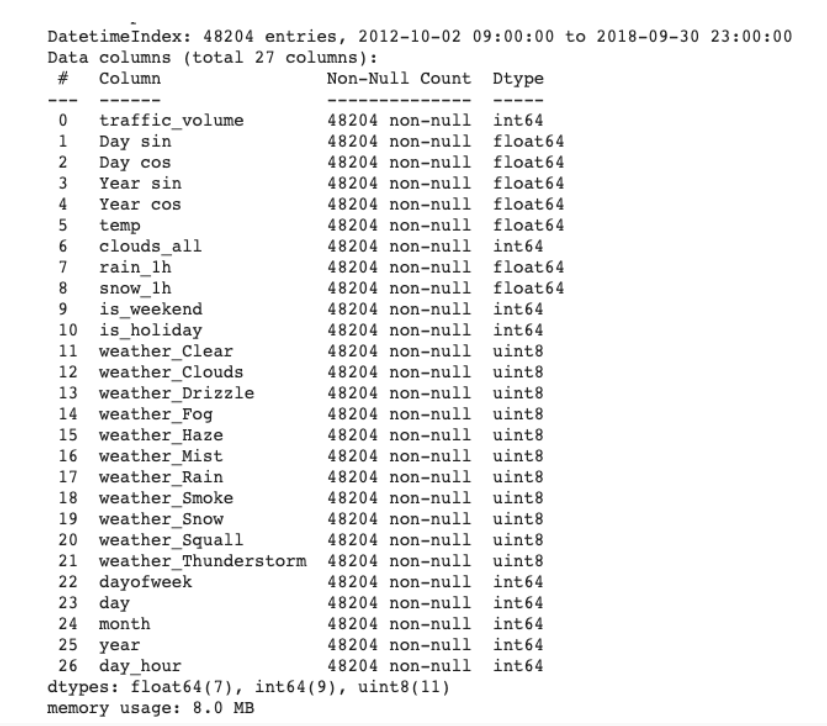
Description automatically generated

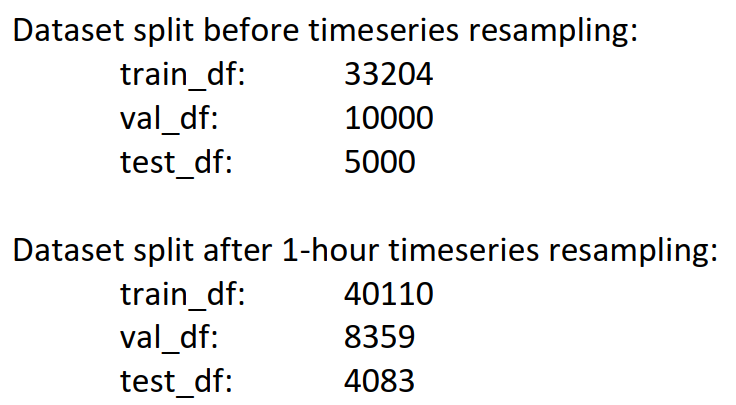
At the initial data processing almost 7629 duplicate hourly entries were found in the dataset, that means the traffic volume was repeated many times for the same day-hour. Initially, we were aware of that, but we decided to treat the dataset as an hour per record, not as actual time-series set with a time unit indexed dataset, but this led to modest validation results (mid 300s MAE). Then, We decided to preprocess the training and validation dataset properly as timeseries indexed at 1-hour intervals, and the validation results were much better (low 200s MAE). Lateron, we split the data into training, validation, and testing. To keep track of the records belonging to the right dataset. Furthermore, during data pre-processing missing data was found between 2014 and 2015 (10 months) and effected validation records because a gap could be clearly observed between the data as shown in the figure below.

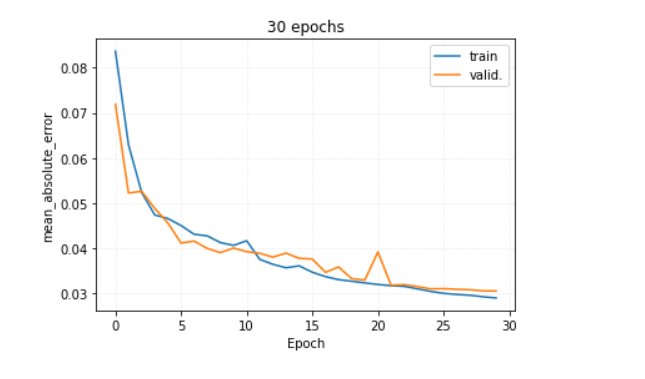


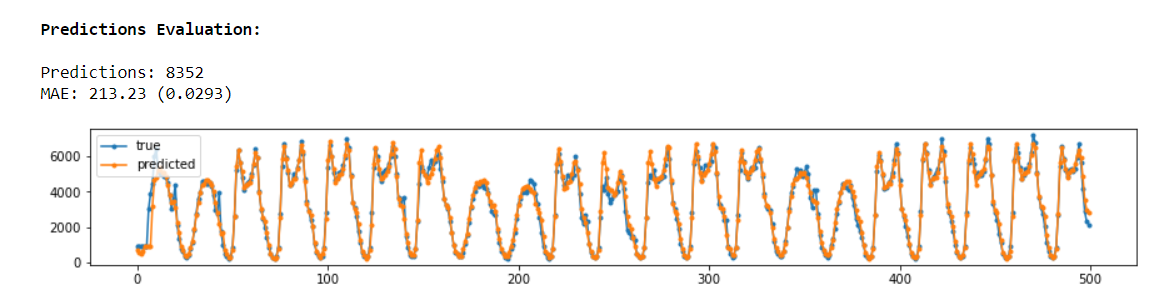
We used neural networks to deal with it and for the outliers that were obsereved in few of the records e.g two days had the temperature field set at zero, which didn't seem right, we fixed it by setting the missing values with the temperature average of each day.



Rain field also had a single extreme value, which we also set to the mean of that day. Snow had extreme values, but we could not determine if the values were to be outliers or not; it could be an exceptional winter, and as we were unaware of the weather conditions in that area, we decided to keep it as is. Moreover, during feature engineering and data transformation process, we decided to transform the weather\_main into one-hot encoded variables and to drop the weather\_description as we see that it adds a kind of redundant information with weather\_main. Also, we think the valuable information to capture is whether the day is a holiday or a weekend. We don't need to keep track of which holiday. So, a new feature is\_holiday is created, and the old feature holiday was dropped. Similarly, we don't need to keep track of which weekend it is. So, we created a new feature, is\_weekend. The date\_time field was converted into a signal using sin and cos to convert the time to clear "Time of day" and "Time of year" signals. This gives the model access to the most important frequency features. Lastly, we broke down the date\_time components into its other elemental fields, dayofweek, day, month, year, and day\_hour. We end with 27 features instead of the initial 9 features. 

In the time series index resampling section, as mentioned, there were 7629 duplicate hourly entries. And we decided to fix this by resampling the data to be on a 1-hour basis so that each record resembles only one hour. Duplicate hour records were averaged within the same hour. The transformations were done after the dataset split into training, validation, and testing. We got the best results using Min-Max Normalization vs. Standard Scaling. For most of the experiment, we aimed to evaluate the effect of tweaking each hyperparameter individually on the performance of the RNN and LSTM based models. We decided not to use automated hyperparameter search methods to understand better how each hyperparameter affects the model. Given that the data preprocessing consumed many hours, we had to apply quick and dirty methods to help us understand how the LSTM networks worked. Alternative to the hybrid automated systematic hyperparameters evaluation we applied systematic manual testing of the hyperparameters using extremely small and large values for the number of LSTM units. Then, we added other layer types such as Convolutional, Dense, GRU, and Bi-directional layers. We also experimented with different batch sizes. Then, we started mixing and matching my observations based on how the layers and the hyperparameters affected the models. 

We tried to find well-known LSTM model architectures as we previously found for the Convolutional Neural Networks. Still, we couldn't find anything beyond a couple of LSTM layers as references. The hyperparameters that were systematically tested using various values are:  
• Number of LSTM units and layers, and Dense units at the final layers (Model 1)  
• Bi-Directional LSTM (Model 2) – Furthermore, we kept a variation of three reference models provided in the TensorFlow tutorial (Dense, Conv, and LSTM). We built two other models (MyLSTM\_1, MyLSTM\_2) that performed better than the reference models. Our best model utilized bidirectional LSTMs with two custom forward and backward layers, and two dense layers with 512 units each, and a dense output layer with a single unit. The model performed the best and maintained low variance for more epochs during training (Figure). 

The predictions in (Figure) show that the model captured all the significant patterns with minor misses of some anomalies. 

# 5. Conclusion

* The LSTM Bi-Directional model performed the best with the least variance.
* We could not make LSTM networks gain better results by going deeper but having more LSTM units made a difference.
* LSTM and GRU should help with the Vanishing Gradient Descent problem in deep networks, but We did not gain any benefits in building deeper networks. Maybe I was not patient enough this time.
* We noticed that most of the models with fewer LSTM units were hardly overfitting and showed better validation scores than the training score.

# References

1. A Research of Traffic Prediction using Deep Learning Techniques. International Journal of Innovative Technology and Exploring Engineering, 2019. 8(9S2): p. 725-728.
2. Apurv Chandel, S.S., Badavath Uday Kiran, Prabhas Prasad, Nidhi Lal, An Accurate Estimation of Interstate Traffic of Metro City Using Linear Regression Model of Machine Learning.
3. Data set. Available from: [https://archive.ics.uci.edu/ml/datasets/Metro+Interstate+Traffic+Volume#](https://archive.ics.uci.edu/ml/datasets/Metro+Interstate+Traffic+Volume).
4. Ahmad, U.K. Metro Interstate Traffic Volume Time-series Forecasting Using Recurrent for Neural Networks (RNNs). Available from: <https://medium.com/@umaimakhurshidahmad/metro-interstate-traffic-volume-time-series-forecasting-using-recurrent-for-neural-networks-rnns-a73732276d1a>.
5. Garlan, E.; Available from: <https://www.kaggle.com/code/ramyahr/metro-interstate-traffic-volume/notebook>.
6. Rohit Singh, A., Mr. Sibi Amaran, Dr. K Sree Kumar, Analysis of traffic flow in different weather conditions. april, 2021.
7. Jiang, R., et al., DL-Traff: Survey and Benchmark of Deep Learning Models for Urban Traffic Prediction, in Proceedings of the 30th ACM International Conference on Information & Knowledge Management. 2021. p. 4515-4525.
8. Kanokwan Khiewwan, P.W., Khumphicha Tantisontisom, and J. Ongate, Application of Data Mining Techniques for Classification of Traffic