Multivariate Time Serious Traffic Prediction Using Long Short Term Memory Network

S.Narmadha, Dr.V.Vijayakumar

Abstract: Short term traffic prediction is essential in the modern intelligent transportation systems. Numerous algorithms were developed for time serious traffic prediction. Long short term memory network (LSTM) is a time serious prediction model which is able to integrate multiple variables such as flow, weather and precipitation. Most of the researches were carried out using the main source of data such as total flow or average speed to predict the vehicle congestion. Weather and rainfall are other factors which are gradually increase the traffic congestion in the crowded city. In this paper, LSTM network is proposed for multivariate analysis based traffic flow prediction. Traffic data contains noise and missing values due to device failures and communication problems. Missing data has been imputed and noise are removed using stacked denoise autoencoder (SDAE). Results are compared with LSTM based univariate analysis and convolutional neural network (CNN) based multivariate analysis. According to the results of LSTM based Multivariate (flow, weather, precipitation) approach without missing value reduces the RMSE error rate to 15.01 to predict the future congestion of a road.

Index Terms: Vehicle, Traffic Congestion, Flow, Weather, Precipitation, Multivariate, Deep Learning.

1 INTRODUCTION

Growing population, number of vehicle usage and space constraints of road increases the traffic congestion problem. Intelligent transportation system (ITS) is a main component in smart city to control road traffic [7]. It can effectively alleviate the traffic congestion, reduce the pollution and provide more safety road conditions [17]. As many factors such as weather, rainfall, accidents influencing the vehicle traffic, traffic modeling becomes very difficult. Traffic causing variables are highly non- linear and stochastic. Traffic congestion prediction has increases the attention of researchers because it might give solutions to travelers, traffic control authorities, traffic management systems etc. It provides the information for route planning, rescheduling and congestion prediction [14]. Advanced technology and computation power makes us to predict the congestion very effectively through the usage of large data. Speed, flow, occupancy are the three major variables used to predict the congestion. Flow is the number of vehicles used in particular segment of road [6]. Traffic flow predicts the congestion of a road at next timestamp ‘T+1’ based on current timestamp ‘T’ from the collected traffic data [3]. Spatio temporal traffic prediction or forecasting is challenging due to complex structure of a road, spatial dependency, non - linear time serious data, stochastic characteristics and uncertainties etc. Traditional and more advanced algorithms were developed to predict the traffic congestion with high accuracy [12]. However limited interest showed on integrating other factors such as rainfall, accidents, dew, wind speed which are gradually increase the congestion [6]. Machine learning algorithms have the capability to analyze large data and predict the model. Different types of machine learning algorithms and models were applied based on time series prediction. But most of the models need length of historical data which should be predefined and static [17]. Both spatial and temporal integration might be needed for efficient prediction of traffic flow. Deep learning is an emerging technology and it is acted as a novel alternative for traffic flow prediction in terms of extracting meaningful features and latent representation of data. RNN is a deep learning algorithm which integrates both spatial and temporal correlation through the use of internal memory units. LSTM network is a kind of RNN and it is a time serious algorithm efficiently used in most of the applications including speech recognition, bioinformatics and traffic prediction [6]. The rest of the paper is organized as follows: section II describes about the literature study on traffic flow prediction models. Section III contains methodology of LSTM prediction model and pre-processing technique. Section IV presents the results and comparisons with various prediction models. Section V concludes this paper.

- S.Narmadha Department of computer Science, Sri Ramakrishhna College of Arts and Science, Coimbatore. India. E-mail: narmadhas17@gmail.com
- Dr V.Vijayakumar, Department of Computer Science, Sri Ramakrishhna College of Arts and Science, Coimbatore. India. E-mail: veluvijay20@gmail.com

2 LITERATURE STUDY

Many methods were developed over the years for traffic flow prediction and forecasting. Basically traffic flow predictions are classified into three types such as short-term, mid-term and long-term which is defined based on prediction interval Δ [17]. Long term prediction usually takes the time span of months to year and midterm prediction takes hours to days whereas short term prediction takes 5 to 10 minute and short term prediction is considered as more efficient than others in terms of accurate prediction [5]. Generally traffic prediction methods are divided into two categories parametric and non parametric. Parametric approaches such as ARIMA (Auto regressive integrated moving average), Seasonal ARIMA [18], Kalman filter were used to predict the congestion. Due to stochastic and non linear nature of traffic data, parametric methods could not predict accurately. Non parametric approaches such as support vector machine (SVM) and artificial neural network predicts well with flexible model structure [17]. Artificial neural network has the learning ability and adaptability to handle high dimensional data. Stacked Denoise Autoencoder (SDAE) [8], Deep Belief Network (DBN) [13], Convolutional Neural Network (CNN) [2] are used for traffic flow prediction. But these non-parametric methods require predefined and static long term historical data [17]. Long short term memory network (LSTM) have been using in many applications including transport management [3, 17].

M Chen et al. [6] used the stacked LSTM network for congestion prediction based on traffic speed data. Recurrent neural network has the capability to handle long term dependencies but theoretical evidence shows that it suffers from gradient descent problem when time lag increases [7] and it is inadequate for process with long term dependencies [6,7]. H Shao et al. [3] introduced LSTM algorithm for short term traffic flow prediction, and it increases the accuracy with latent feature representation. J Li et al. [5] checked LSTM based traffic prediction with three different hyper parameters such as different dropout values 0, 0.1, 0.3, 0.5 were used to evaluate the prediction performance of the model, various time periods and number of hidden layers were feed into LSTM model to check the accuracy of prediction and proved LSTM model performs better than recurrent neural network (RNN). B Yang et al. [1] developed LSTM + method to enhance the prediction accuracy through utilizing long sequence data. High impact features were utilized by attention mechanism [1]. W Xiangxue et al. [14] proposed short term prediction model, in which time serious model is decomposed into trend series and residual series. Prediction is performed using LSTM-RNN based on reconstructed time series models. Prediction accuracy is improved by optimized data preprocessing and classification. Y Tian et al. [17] created short term prediction model with LSTM RNN to determine the optimal time for time sequence traffic flow data. This model generalizes well and improves the accuracy. Z Zou et al. [19] proposed city level traffic flow prediction with LSTM networks. 24 hours of history information were used to predict the future 1 hour data. Only flow data has used to predict but other factors influence the traffic information. In future more influence factors such as weather and rainfall should take into consideration. Hybrid models were developed with LSTM network to integrate both spatial and temporal correlation to predict the traffic congestion with high accuracy. X Luo et al [16] developed spatio temporal traffic prediction model with K-nearest neighbour (KNN) and LSTM network. KNN was used to choose the neighbouring stations to capture spatial features and LSTM network was used to capture temporal correlations of traffic data. R fu et al. [7] introduced Long short term memory network (LSTM) with Gated recurrent units (GRU) for short term traffic flow prediction. LSTM and GRU were proposed with forget units to forget some information which provides optimal time lags. H Zhang et al. [4] developed a multivariate short term forecasting method based on wavelet analysis and seasonal time serious (WSARIMAX). Occupancy was taken as an exogenous variable with flow to increase the prediction accuracy. Result shows that multivariate analysis based prediction gives better result than univariate analysis. D Yang et al. [2] introduced convolutional neural network for multi feature fusion based traffic flow prediction. Weather and holidays were taken as external factor with flow data to predict the vehicle congestion. High level spatio temporal features have learned and merged by convolutional neural network (CNN). It increases the efficiency and accuracy of prediction model. Based on the relative study, LSTM has proven to be superior for time serious predictions. In this paper we proposed LSTM model for multivariate time serious based traffic flow prediction.

3 METHODOLOGY

3.1 Long Short Term Memory Networks (LSTM) based Traffic Prediction

RNN (recurrent neural network) is a feed forward neural network used to extort temporal dependency from a time sequence data [5]. RNN has a "memory" to capture previous inputs to the current state for influence the output [16]. However RNN is not able to work with long time sequence data [16]. Literature shows that long term dependencies data are inadequate in training with gradient descent error [18]. LSTM is designed for model long short term dependencies and to solve sequential problems [6].

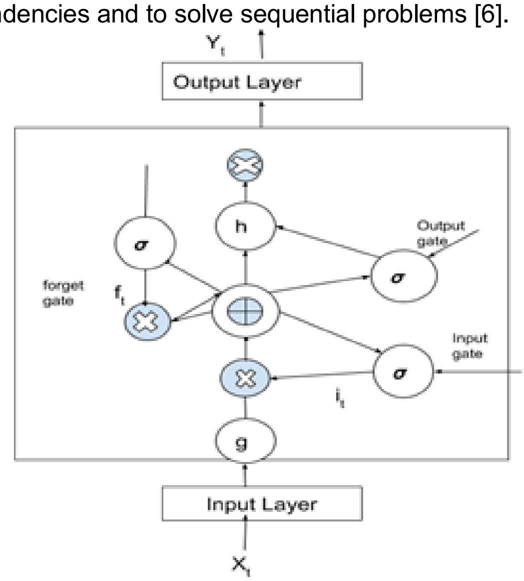

Fig 1: *Structure of LSTM block [5]*

A simple LSTM network (fig 1) contains input layer, recurrent hidden layer and output layer. Hidden layer contains memory block which is connected to layers than neurons [6, 7]. It includes gates such as 'input gate' is used to receive the new information and 'forget gate' is used to decide the discarded information. Finally 'output gate' has to store the output which is going to next high level [1, 7]. LSTM can keep required information through the input, forget and output gates [7]. Long short term memory network is mainly used for prediction with time serious data [1]. Time serious prediction works by combining the hidden states of previous time step data with the current time step.

Let us denote Input $Y_t = \{y_1, y_2, y_3, ..., y_t\}$ where t = length of input; h is a hidden state of LSTM network, $H = \{h_1, h_2, h_3 ... h_n\}$. Output $Z_t = \{z_1, z_2, z_3 ... z_n\}$.

Computation of LSTM network is as follows [6],

$$h_t = H(W_{hy} y_t + W_{hh} h_{t-1} + b_h) \quad (1)$$

$$P_t = W_{hz} Z_{t-1} + bz \quad (2)$$

'W' is the weight matrix. 'b' is a bias value.

$$i_t = \sigma(W_{iy} x_t + W_{hI} h_{t-1} + W_{ic} C_{t-1} + b_i) \quad (3)$$

$$f_t = \sigma(W_{fy}x_t + W_{hl}h_{t-1} + W_{fc}C_{t-1} + b_f) \quad (4)$$

$$C_t = f_t * C_{t-1} + i_t * g(W_{cy}y_t + W_{hh}h_{t-1} + W_{cc}C_{t-1} + b_c) \quad (5)$$

$$O_t = \sigma(W_{oy}x_t + W_{hh}h_{t-1} + W_{oc}C_{t-1} + b_o) \quad (6)$$

$$h_t = o_t * h(C_t) \quad (7)$$

σ is the sigmoid activation function [17] which is defined in Eq. (8)

$$\sigma = \frac{1}{1+e^x} \quad (8)$$

g(.) and f(.) is an exponential linear unit (Elu) activation function with range [-1, 1] and it is showed in Eq. (9) [10]

$$\{f(p) = \begin{cases} r & r > 0 \\ \alpha.(e^2 - 1, & r \le 0 \end{cases}\} \quad (9)$$

Where r = input, **α** = constant and e = exponent.

LSTM has the ability to store long term information and can determine the information which is not needed (to be forgotten) by the current state and historical information [5]. Multiple variables of time serious problem is converted into supervised learning problem for multivariate traffic prediction. Prediction helps to commuters and road users to make better travel decisions and alleviate traffic congestion and collision [18].

Algorithm 1 Multivariate LSTM Prediction

Input : Total flow $X_1 = \{x_1,x_2,x_3......x_n\}$
Weather = {Temp (X_2), dew(X_3) , humidity (X_4), wind Speed (X_5), wind direction (X_6), visibility (X_7), precipitation (X_8)}
Output : prediction Result $Z_1 = \{z_1,z_2,z_3.....z_n\}$

Step 1: Impute missing values

Initialization: Weights (w) and bias, epoch σ, no of hidden layers h;
Input : $T = \{X_1,X_2,X_3....X_8\}$
Pass raw data (Total flow & Weather) into stacked Denoise Autoencoder (SDAE);
Train model;
Output : Reconstruct the output $T_1=\{Y_1,Y_2,Y_3.....Y_8\}$ without missing values;
Step 2 : Prediction

Initialization : Weights (w) and bias (b), epoch σ , no of hidden layers h;
Input : Imputed data (total flow & weather) $T_1=\{Y_1,Y_2,Y_3.....Y_8\}$
Convert time serious data into supervised learning problem
for i = 0 to epoch do
Pass input to train LSTM model;
Do
Perform operations from equations (1) to (7).
Calculate error;
Update parameters;
end;
for i = 0 to epoch do
Test the model to generate the result;
end;
Output: Prediction result $Z = \{z_1,z_2,z_3...z_n\}$

4 EXPERIMENTAL RESULTS & DISCUSSION

4.1 **Dataset**

Traffic data collected from widely referred [17, 18] Performance Measurement System (PeMS) dataset [20]. PeMS is a california based database which contains all traffic related data such as total flow, average speed, occupancy etc. It has 44,734 detectors which are spread all over the country and traffic data is collected at every 30 seconds and aggregated into 5 min interval. District 5 (Central coast) is chosen for data collection which has 485 stations. Single station Watsonville-50003031 (cabrillo highway of santa cruz county) has taken for collecting traffic flow data for univariate and multivariate analysis.

Temp, Dew, Humidity, Wind Speed, Wind direction, Visibility, Precipitation are the weather factors which are accessible from Mesowest database [21] based on the California county. Weather data are also accessible at every 5 min interval. 4 months data from January 1 to April 30 for the year 2017 is taken for analysis of result. First 3 months of data have used for training and 1 month of data is used for testing.
Keras is used to implement LSTM model. 12 GB Tesla GPU hardware machine is used for all experiments. Proposed framework is depicted in fig 2.

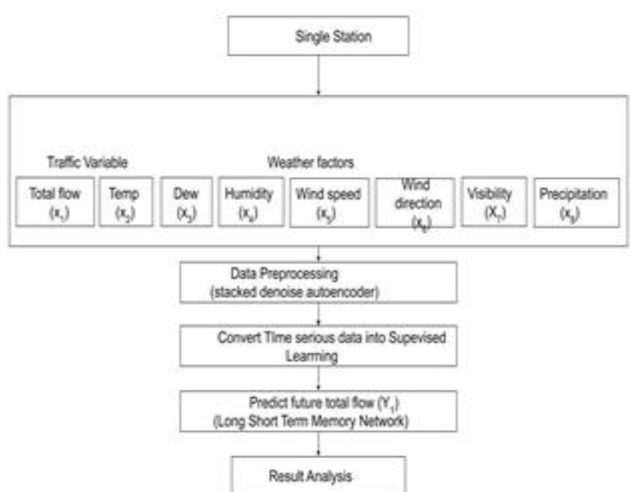

Fig 2: *Flowchart of prediction model*

4.2 Data pre-processing

Both traffic and weather data have collected using sensors, loop detectors, cameras etc. But these data may contain some missing values due to device failures, communication problems and malfunctions etc [9]. Hence it affects the quality of all upcoming traffic related applications. PEMS database also has some missing values [18]. Each and every 5 min data is crucial to predict the congestion of next time interval. Stacked denoise autoencoder deep learning algorithm used for imputing the missing values with Elu activation function [10, 11]. Weekdays, weekends and holiday data are taken for temporal analysis and single station data is taken for spatial analysis. In the year 2017, February 16^{th} to 20^{th} data is taken for analysis of result. 16^{th} and 17^{th} are weekdays, 18^{th} and 19^{th} are weekends (Saturday was a typical working day and Sunday was a holiday). At last day 20^{th} was a holiday in US federation as it is celebrated as president's day. From the plot, week days have two peaks and week end saturday flow is less than weekdays. Sunday's vehicle flow is very less and federation holiday flow was only one peak hour. Relative Error (MRE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are used to determine the prediction accuracy [18].

Mean Absolute Error (MAE = $\frac{1}{n}\sum_{i=1}^{n}|a-p|$

Mean Relative Error (MRE) = $\frac{1}{n}\sum_{i=1}^{n}\frac{|a-p|}{a}$

Root Mean Square Error (RMSE) = $\sqrt{\frac{1}{n}\sum_{i=1}^{n}|a-p|^2}$

4.3 Univariate forecasting

Traffic flow congestion has been predicted over the decades mainly using the traffic variable such as total flow or speed or combining of both them. Traffic variable is a main factor to decide the future condition of a road than other factors. Many algorithms were applied to predict or forecast the congestion [12]. Deep learning based LSTM algorithm [3] is applied to predict the congestion with 1 time step using total flow.

Table 1: *Results of univariate traffic flow prediction (4 months data)*

Prediction model	RMSE	MAE	MRE
Univariate (LSTM) with missing value	20.90	14.45	0.30
Univariate (LSTM + SDAE) without missing value	17.61	12.89	0.21

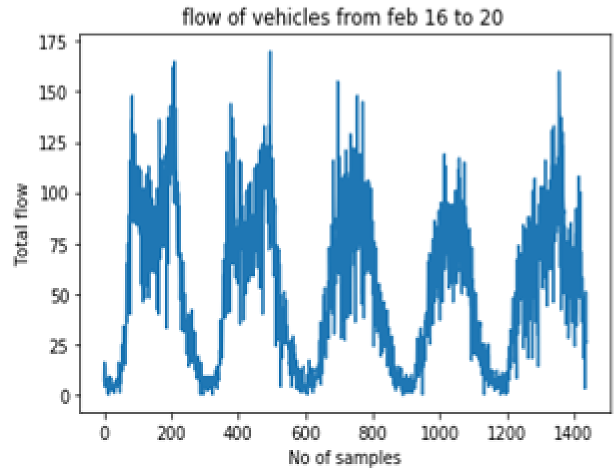

Fig 3. *Impact of the traffic volume for 2017 February* 16^{th}

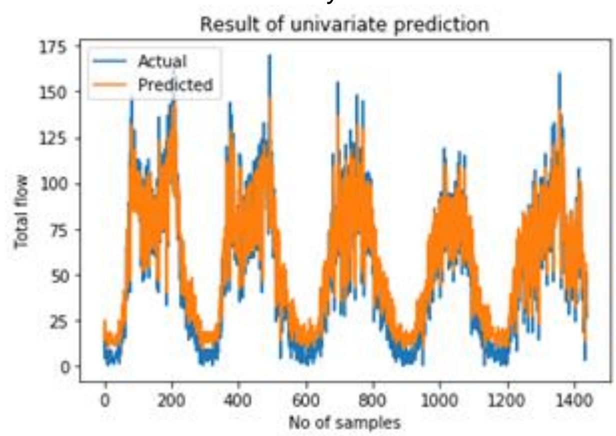

Fig: 4 *Prediction result of univariate traffic prediction (feb* 16^{th} *to* 20^{th} *)*

6

Univariate prediction without missing value reduces the error rate of RMSE to 17.61. Stacked autoencoder efficiently impute the missing values. So for single station every 5 min data is imputed based on previous occurrence of data. SDAE generalize well for large datasets with early stopping functionality. Based on result LSTM predicts well on heavy flow of traffic than usual.

4.4 Multivariate forecasting

Vehicle traffic congestion is caused by not only the flow of vehicles but other factors may reduce the flow and make the congestion live. Such factors are weather, wind speed, precipitation etc. These factors are highly non linear and stochastic. And also heavy traffic increases the chances of uncertainties like accidents, collision etc. Hence traffic variable alone is not adequate to predict the upcoming road condition. But multivariate analysis of traffic prediction is challenging due to uncertainties of data. Long short term memory network (LSTM) is proposed for multivariate analysis of traffic congestion prediction. Single station is chosen for analysis of data various over time. Single location is taken as a spatial with various temporal sequences. LSTM is a time series prediction method which effectively improves the accuracy of prediction through Elu activation function.

Table 2: *Comparison of multivariate analysis prediction (4 months data)*

Model	RMSE	MAE	MRE
LSTM (Multivariate with missing value)	19.00	12.68	0.62
LSTM + SDAE (multivariate without missing)	15.01	11.02	0.16

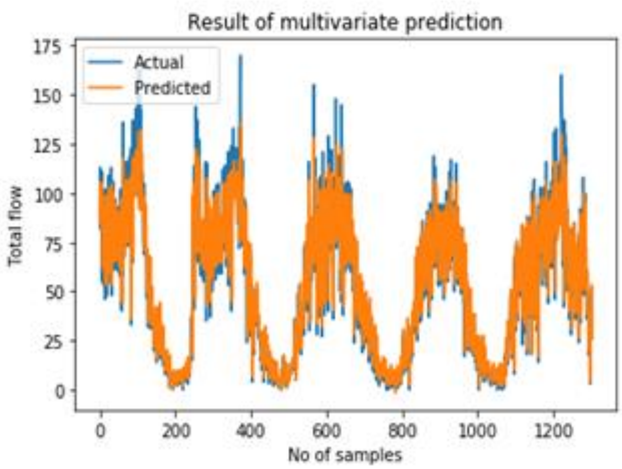

Fig: 5 *Result of multivariate prediction (feb 16th to 20th)*

From the analysis of results, noise removal of input values and capability of LSTM network yields better prediction result for multivariate traffic prediction. It ensures that traffic variable with additional factors consideration might increase the prediction accuracy. Multivariate prediction predicts the result well on both normal flow and heavy flow of vehicle.

4.5 Comparison

LSTM and Convolutional neural network (CNN) were used to predict and analyse the results. Both CNN and LSTM have the capability to integrate multiple factors.

Table 3: *Comparison of both univariate and multivariate analysis (4 months data)*

Prediction model	RMSE	MAE	MRE
Univariate forecasting (LSTM)	17.61	12.89	0.21
Multivariate forecasting (LSTM)	15.01	11.02	0.16
Multivariate forecasting with CNN	17.648	12.86	0.20

Multivariate analysis with LSTM network reduces the RMSE error rate to 15.01 and MAE to 11.02 and MRE to 0.16. Traffic prediction accuracy precision is very important to control road traffic and proactive planning. LSTM integrates other factors well with total flow based on time serious.

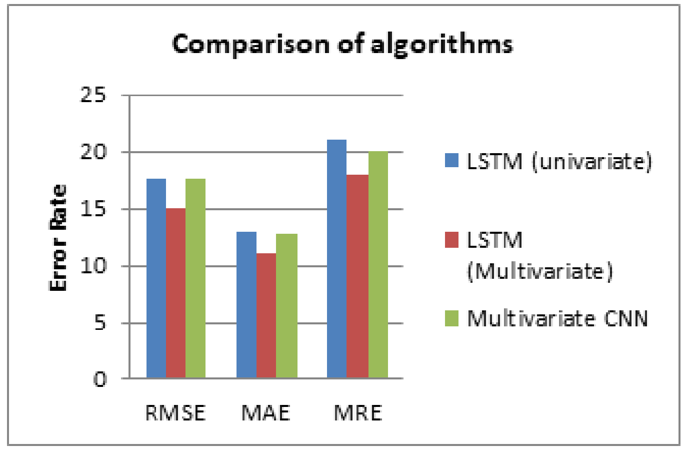

Fig: 6 *Result comparisons (4 months data)*

LSTM outperforms than CNN as it is a time serious based algorithm which accurately extracts the temporal correlation of data.

CONCLUSION

In this paper long short term memory recurrent neural network is used to examine both univariate and multivariate traffic analysis. Elu activation function makes to work the prediction model faster and provide the best result. The major development of this model is integrates all factors which are the causes of traffic congestion. Results are compared with LSTM based univariate prediction and CNN based multivariate prediction. Multivariate LSTM model reduces the RMSE error rate to 15.01 from 17.61 of univariate LSTM prediction model. Result shows that proposed model outperforms than other models. Stacked denoise autoencoder based imputation result helps to predict the congestion with high accuracy. Temporal factors and simple spatial factor only have taken into account for prediction. In future consider upstream and downstream stations as spatial locations with temporal factor to improve the accuracy of result.

ACKNOWLEDGEMENTS

We thank Sri Ramakrishna College of arts and science for giving support to doing this research work.

REFERENCES

[1] B. Yang, S. Sun, J. Li, X. Lin, Y. Tian, “Traffic Flow Prediction Using LSTM with Feature Enhancement,” Neuro computing, vol. 332, pp.320-327, 2018.
[2] D. Yang, s. Li , z. Peng , p. Wang , j. Wang , and h yang, “MF-CNN: traffic flow prediction using convolutional neural network and multi-features fusion,” IEICE trans. Inf. & syst., vol.e102–d, no.8 august 2019.
[3] H. Shao, B. Soong, “Traffic flow prediction with Long Short-Term Memory Networks (LSTMs),” 2016 IEEE Region 10 Conference (TENCON), Singapore, pp. 2986-2989, 2016.
[4] H. Zhang, X. Wang, J Cao, M Tang, Y Guo, “A multivariate short-term traffic flow forecasting method based on wavelet analysis and seasonal time series,” springer, Vol. 48, Issue 10, pp. 3827–3838, 2018.

[5] J. Li, L. Gao, L. wei and Y. Shi, “Short term traffic flow prediction based on LSTM,” Ninth international conference on Intelligent control and information processing, Wanzhou, pp.251-255, 2018.
[6] M. Chen, G. Yu, P. Chen, and Y. Wang, “Traffic Congestion Prediction Based on Long-Short Term Memory Neural Network Models,” 17th COTA International Conference of Transportation, CICTP 2017, pp.673-681, 2017.
[7] R. Fu, Z. Zhang and L. Li , “ Using LSTM and GRU Network methods for Traffic flow prediction,” 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, pp. 324-328, 2016.
[8] S. Ghosh, M. T. Asif, L. Wynter, “Denoising autoencoders for fast real-time traffic estimation on urban road networks,” 2017 IEEE 56th Annual Conference on Decision and Control (CDC), pp. 6307-6312, 2017.
[9] S. Narmadha, V. Vijayakumar, “A Study on Imputation Methods for Vehicle Traffic Data,” International Journal of Recent Technology and Engineering (IJRTE), Volume-7 Issue-5S, pp.415-420, 2019.
[10] S. Narmadha, V. Vijayakumar, “An Improved Stacked Denoise Autoencoder with Elu Activation Function for Traffic Data Imputation,” International Journal of Innovative Technology and Exploring Engineering (IJITEE), Volume-8 Issue-11, pp.3951-3954, 2019.
[11] S.Narmadha, V. Vijayakumar, “An Effective Imputation Model for Vehicle Traffic Data Using Stacked Denoise Autoencoder,” Advances in Intelligent Systems and Computing, Springer, Cham, vol 1108, pp 71-78, 2020.
[12] S. V. Kumar & L. Vanajakshi, “Short-term traffic flow prediction using seasonal ARIMA model with limited input data,” Eur. Transp. Res. Rev. (2015), 7: 21, Springer, pp. 1-9, 2015.
[13] W. Huang, H. Hong, M. Li, W. Hu, G. Song and K. Xie, “Deep Architecture for Traffic Flow Prediction,” Lecture Notes in Computer Science, vol 8347 pp.165–176, 2013.
[14] W. Xiangxue, X. Lunhui, C. Kaixun, “Data-Driven Short-Term Forecasting for Urban Road Network Traffic Based on Data Processing and LSTM-RNN,” Arabian Journal for Science and Engineering 44, 3043-3060, 2018.
[15] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang and Y. Wang, “Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction,” Sensors, pp: 1-16, 2017.
[16] X. Luo, D. Li, Y. Yang and S. Zhang, “Spatiotemporal Traffic Flow Prediction with KNN and LSTM,” Journal of advanced transportation, Vol 2019, Article ID 4145353, pp.1-10, Hindawi, 2019.
[17] Y.Tian, L. Pan, “Predicting short term traffic flow by long short term memory recurrent neural network,” 2015 IEEE International Conference on Smart City/SocialCom/SustainCom together with DataCom 2015 and SC2 2015, pp. 153-158.
[18] Y. Tian, K. Zhang, J. Li, X. Lin, B. Yang, “LSTM-based Traffic flow Prediction with Missing Data,” Neuro Computing, Vol 318, pp. 297-305, 2018.
[19] Z. Zou, P. Gao, C. Yao, “ C*ity-level traffic flow prediction via LSTM networks,” ICAIP '18 Proceedings of the 2nd international conference on Advances in Image Processing, pp. 149-153, june 2018
[20] http://pems.dot.ca.gov/
[21] http://mesowest. (The University of Utah.)