Fourth International Conference on Electronics, Communication and Aerospace Technology (ICECA-2020)
IEEE Xplore Part Number: CFP20J88-ART; ISBN: 978-1-7281-6387-1

Short Term Predictions of Traffic Flow Characteristics using ML Techniques

B. Rajendra Kumar
PG student
Department of Civil Engineering
VNR Vignana Jyothi Institute of Engineering and Technology
Hyderabad, India
rajendra9391@gmail.com

Dr. Naveen Kumar Chikkakrishna
Associate Professor
Department of Civil Engineering
VNR Vignana Jyothi Institute of Engineering and Technology
Hyderabad, India
doccnk@gmail.com

Teja Tallam
Assistant Professor
Department of Civil Engineering
VNR Vignana Jyothi Institute of Engineering and Technology
Hyderabad, India
tnteja@gmail.com

***Abstract*— This research article proposes a new model for traffic volume prediction, where it can be effectively used for transportation domain particularly in safety planning, management and assessment at any time. Various prediction methods are proposed for predicting the traffic flow, by including historical method, real-time method, time series analysis, but the precision and efficiency of time in forecasting are a couple of difficult contradictions. Real-time traffic information prediction with ANN and SVR are applied for developing an effective and efficient traffic prediction. This study develops the model for prediction of traffic volume for Nizampet road stretch, an urban area by analyzing the measured data in city of Hyderabad. In this study the artificial neural network model is best suited to Nizampet road stretch as the R-square value is 0.89 and performance measures are less compared with SVR model.**

Keywords— *Traffic volume, SVR, Regression analysis, ANN, Short term traffic prediction, Traffic flow prediction.*

I. INTRODUCTION

Jams (traffic) occur daily in most of the active urban centers around the world. The urban arteries carry a significant part of the daily traffic. The increase in demand for transport services and facilities become more complex and larger with the increase in social and economic activities. This led to delays and huge congestion of arteries and city roads. Therefore, we can see too much research on control measures of traffic in transportation facilities. The traffic control system includes the study of traffic speed, volume and traffic density. Generally, these characteristics are considered as measures of performance of roads. The information about the temporal and spatial distribution of traffic is important factor requirement in almost all transport planning and design strategies. In a road network traffic control system, it is desired to predict the characteristics of future traffic to consider the appropriate control and management strategy. The establishment of a road traffic control system requires a means to acquire a temporal and spatial distribution of traffic.

II. LITERATURE REVIEW

Many researchers studied the short-term traffic predictions. An attempt has been made by [1] using kernel methods and support vector regression model, which minimizes the generalization errors linked to get a general performance. Using autoregressive model and neural network model it was found that the forecast in the short term and 24 hour forecast are importantly lower in the base forecast. It was also found that Short-term forecasts have negligible errors when compared to long term predictions [2]. The proposed method [3] using Neural network and Markov chain model predicts the speed study accurately on metropolitan roads and override two basic speed study created by the specified speed limit and traffic model. [4] studied about the Bayesian support vector regression BSRV model. The detection performance of Bayesian support vector regression is analyzed with the false positive rate kept 20% below. The value of the error bar is greater than the average rate and the corresponding instance is recognized by as a positive event Bayesian support vector regression. A higher resolution of data classes is expected to result in higher prediction performance using artificial neural network ANN model [5]. The model absorbing Markov chain model and fundamental matrix was able to predict the rate at which traffic will decrease in the event of congestion. The results indicated that AMC model is accurate and that the well-functioning is obtained with a short-term prediction approach presented in RMSE and MAPE format. It was also observed that the on comparing seasonal and neural network ARIMA models, the AMC model also shows advanced prediction accuracy [6]. Hybrid model and base line model as described by [7], a method of combining data based and model based approaches which improves forecast performance by using the real time traffic data with traffic models from historic traffic data. The model used symbolic regression and genetic programming using Pareto front method, makes it possible to predict the traffic speed under all traffic conditions. The proposed model for predicting traffic speed outperforms SVR and ARIMA [8, 9]. Time series model to predict the volume of traffic in urban arteries was carried out by [10] using Autoregressive integrated moving average order model and found out that the model has proven to be the most appropriate model to replicate original time series. The developed model is easy to understand and implement. In addition, the model is computable and only requires the

storage of the last expected error and the observation of the current traffic. [11] worked on short term volume forecast on the incorporation of spatial and temporal features in the forecasting procedure, the results show that the proposed procedure provides more accurate predictions than the standard statistical method, as well as neural networks static forms. Support vector machine model (SVM), one of the artificial intelligence method for predicting the traffic flow and SVM-HPT model outperformed the other three models for prediction [12]. Traffic flow prediction among k-nearest neighbor (KNN), support vector machine (SVM), and artificial neural network (ANN), KNN gives more accurate results than SVM and ANN [13]. Forecasting urban traffic was analysed using spatiotemporal and video tools with reduced processing time and improved datasets for video processing was studied by [14]. A cost effective automatic traffic control system for monitoring the railway gates that were unmanned in India was developed by [15].

Based on the literatures it is clearly understood that the existing models are not capable of estimating the effect of different parameters in traffic volume, inspite they only predict the traffic volume based on historical data. Therefore, there is a need in shift from conventional models to machine learning models like Support vector regression model and Artificial Neural Network models for traffic prediction on the road section.

III. METHODOLOGY

Literature study was conducted on the Short-Term Predictions of Traffic and different models used, to know the effect of the parameters on Traffic Flow Characteristics. Different statistical and soft computing techniques were used to develop the models like SVR and ANN model. These models were developed to find the effect of parameters. SVR model was developed using R-Studios and ANN model was developed using MATLAB. Traffic Volume was considered as dependent variable and traffic speed, travel time, distress rating and road width were considered as independent variables. An equation was developed such that it could be used further for prediction of traffic volume on the road section at any time with independent variables.

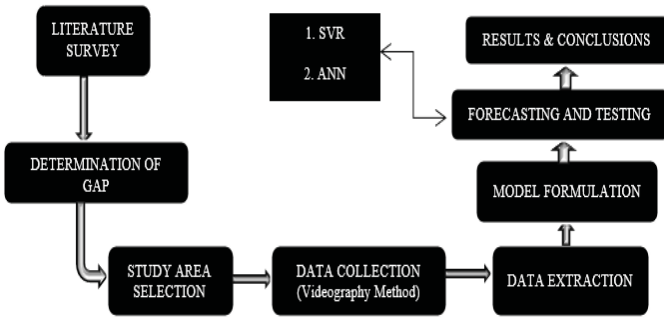

Figure 1: Methodology

The models like ANN and SVR are used in this study such that they will increase the utility of the predicted models as they are capable of giving different parameter effects on traffic volume.

IV. STUDY AREA AND DATA COLLECTION

The study area refers to the specific domain in which the research area was considered. Specific locations were chosen to study the forecasting analysis influencing the road stretch. The location stretching from Nizampet X roads to Bachupally X roads located at Hyderabad, India. The intermediate locations for road stretching between Nizampet X roads to Bachupally X roads are vijetha super market x road, GRIET X road, Hill county X road, SLG hospital X road. All the locations were the combination of the commercial and residential area.

At Nizampet X road 24 hrs data was collected to find out the peak hours. Vijetha super market x road located at Nizampet X road, peak hour was observed to be at 8 am to 11 am and 4.30 pm to 8.30 pm, based on the peak hour data the intermediate locations peak hours were observed to be as follows. At GRIET X roads it was 8am to 11 am and 4.30 pm to 8.30 pm, at Hill county X road peak hour was 8 am to 10.30 am and 4.30 pm to 7.30 pm and at SLG hospital X road peak was 8 am to 10.30 am and 4.30 pm to 7.30 pm. At Bachupally x road the peak hour was observed at 8 am to 12 am and 4.30 pm to 8.30 pm on a week day.

The map view of the study area selected is shown in the figure 2.

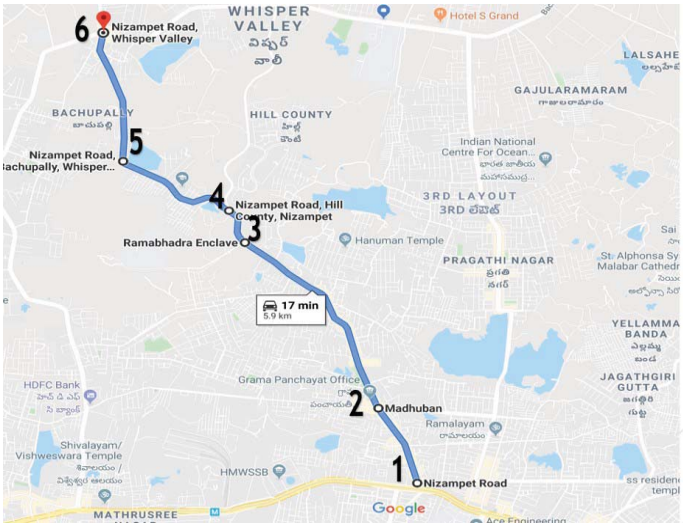

Figure 2: Map of study area

A. Data Collection

Data is collected through videography survey along the Nizampet Stretch for 24 hours at major junction like Nizampet X roads and Bachupally X roads for traffic volume. In spite of

traffic volume, off peak speed and peak speed along with road condition data was collected along the 8.4 km Nizampet stretch with 100m interval on both sides. That is from Nizampet X roads to Bachupally X roads.

B. Data Extraction:

From the videography survey and field survey following parameters are extracted.

1). Traffic volume (5 minutes interval)

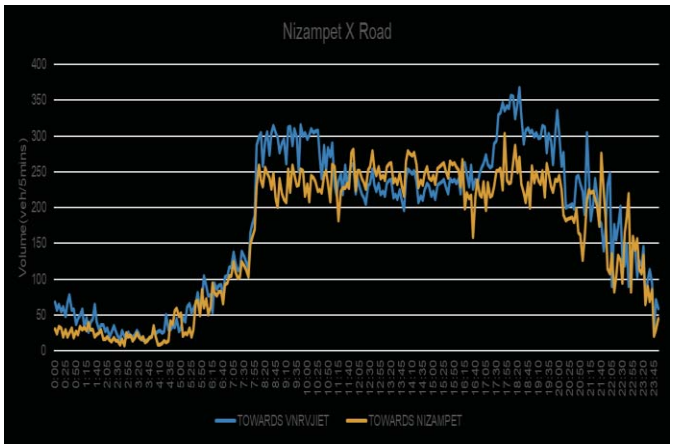

Figure 3: Nizampet X Road

From the figure 3, we can conclude that the peak traffic hours for Nizampet location were found to be 8 a.m. to 11 a.m. and 4.40 p.m. to 9 p.m. in the evening.

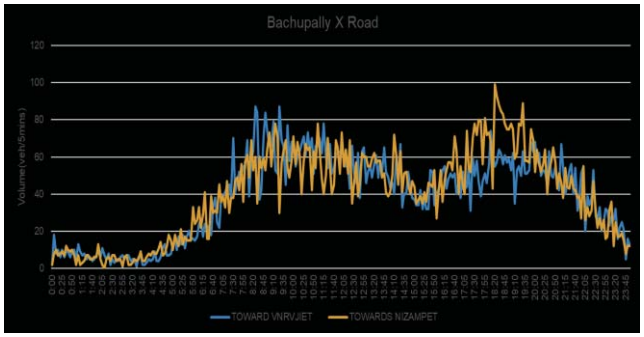

Figure 4: Bachupally X Road.

From the figure 3, we can conclude that the peak traffic hours for Bachupally location were found to be 7.50 am to 10.45 am in the morning and 4.30 pm to 9 pm in the evenings.

2). Road Condition data (Potholes, Rutting, Undulation in mm3)

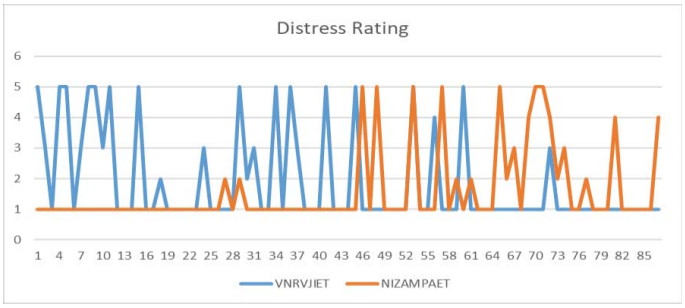

Figure 5: Distress Rating

The distress data for both the locations shown in figure 5 are taken in the form of ratings from 1 to 5 where 1 refers to good and 5 refers to worst condition.

3). Speed data (Peak and Off-Peak in Kmph)

The speed profiles are mention in the figure 6 and 7 in terms of km per hour (KMPH)

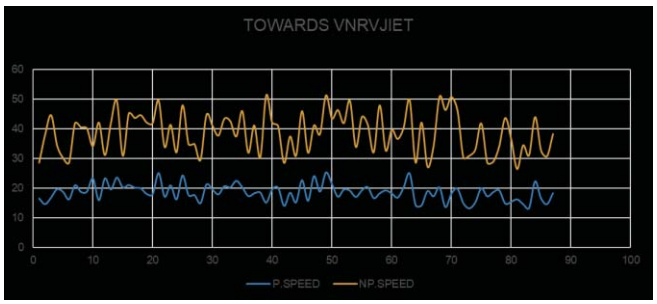

Figure 6: Speed profile along Jntu to Vnrvjiet.

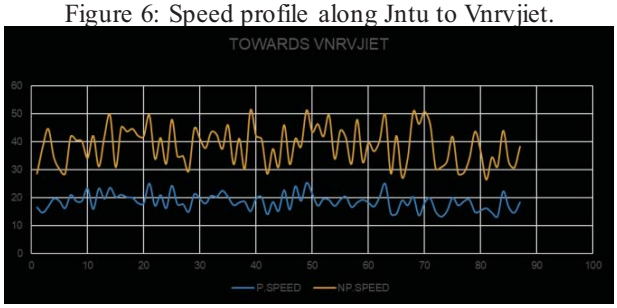

Figure 7: Speed profile along Vnrvjiet to Jntu

V. MODEL DEVELOPMENT

SVR and ANN models are developed for traffic volume prediction which is formulated using R-Software and MATLAB.

A. Correlation Matrix:

The below table represents the correlation between different parameters like speed and distress in to the traffic volume.

Table 1: Correlation Matrix

	VOLUME	*TRAVEL TIME*	*SPEED*	*DISTRESS RATING*	*Road Width*
VOLUME	1.000				
TRAVEL TIME	0.747	1.000			
SPEED	-0.872	-0.697	1.000		
DISTRESS RATING	0.020	0.135	-0.176	1.000	
Road Width	0.003	-0.014	-0.097	0.050	1.000

The correlation matrix describes the relation between each and every variable. The speed parameter has inversely proportional relation with volume, travel time, distress rating and road with. Similarly, the relations between other parameters can be identified based on the negative or positive signs shown in the table 2.

B. Support Vector Regression:

Support Vector Regression (SVR) are learning machines that implement the inductive principle of minimizing structural risks to obtain a good generalization in a limited number of learning models. The SVR model results are tabulate in table

Table 2: Results from SVR.

SVR MODEL PARAMETER	NIZAMPET
R Square	0.83
MAE	0.0005
MAPE	0.87
MSE	0.000287
RMSE	0.01695
Chi Square	0.07
Critical Chi Square	391.43

*RMSE: Root Mean Square Error *MAD: Mean Absolute Deviation
*MSPE: Mean squared prediction error *MAPE: Mean Absolute Percentage Error

From the above table 2 we can observe that the SVR model is having R-Square as 0.83 with RMSE and 0.016 and having the Chi-Square value less than Critical Chi-Square.

Equation formulation: The below equation is formed from the SVR model developed in R-software with predicted the traffic volume, when speed and distress of the road is given as input

$$LOG.TRAFFIC\ VOLUME = 3.13 - 0.47 * LTS - 0.05 * TT - 0.002 * DR - 0.03 * LRW \quad (1)$$

Here,
LTS – Log Traffic Speed (Kmph) , TT- Travel Time(mins), DR- Distress Rating (Number), LRW- Log road width(meters).

C. Artificial Neural Network:

A mathematical model that attempts to simulate the structure and functionality of biological neural networks. This model has three sets of simple rules: multiplication, addition, and activation. The inputs are weighted against the input of artificial neurons, which means that all input values are multiplied by a single weight. The middle section is the summary function, which summarizes all weighted input and bias. Many trails were conducted with different training function, number of layers, number of neurons, epchos etc. Among all the trials, the trail which is having least errors with highest R-Square value is stated suitable for the traffic prediction accurately.

Table 3: Results from Artificial Neural Network (ANN)

	Name of Network	Trial1	Trial2	Trial3
Network Properties	**Training Function**	TRAINBR	TRAINBR	TRAINBR
	Layer Numbers	2	10	7
	Neurons Numbers	10	10	10
	Transfer Function	TANSIG	TANSIG	TANSIG
In Training Parameters	**Epochs**	10000	5000	5000
Regression (R Value)	**Training**	0.56	0.73	0.87
	Validation	0.78	0.77	0.85
	Test	0.69	0.76	0.78
	All	0.65	0.71	0.82
Performance	**Chi-Square**	156.23	148.92	144.56
	MSE	0.0089	0.0052	0.000776

	RMSE	0.29	0.098	0.02785
	MAPE	15.26	2.56	1.46
	MAE	0.09	0.06	0.02
	R2	**0.65**	**0.78**	**0.89**
	SBC	4.98	4.37	3.21

From the above table 3, we can conclude that Bayesian Regression training function with 10 layers and 10 neurons having 10000 Epochs have resulted to be best fitted model for the Pragathi Nagar data with R-square as 0.89 and least SBC value.

VI CONCLUSIONS

- The peak traffic hours for Bachupally location were found to be 7.50 am to 10.45 am in the morning and 4.30 pm to 9 pm in the evenings.
- The peak traffic hours for Nizampet location were found to be 8 am to 11 am in the morning and 4.40 pm to 9 pm in the evenings.
- The article contains a traffic forecasting method based on ANN and SVR.
- In ANN model the R-square value was found to be 0.89 where as in SVR 0.83. Based on the performance measures obtained, ANN model was found to be more accurate than SVR.
- The traffic volume prediction equation was also developed using Support Vector Regression to estimate the traffic volume based on traffic speed and distress on road.

REFERENCES

[1] Debasish Basak, Srimanta Pal, and Dipak Chandra Patranabis., 2007, Oct. Support vector regression. Neural Information Processing, Vol. 11, No. 10.

[2] Tom Thomas, Wendy Weijermars and Eric Cor van Berkum (2010) "Predictions of Urban Volumes in Single Time Series" in IEEE Transactions on Intelligent Transportation Systems 11(1):71 – 80

[3] Jungme Park, Yi Lu Murphey, and Johannes Kristinsson, Ryan McGee., 2013, August. Intelligent speed profile prediction on urban traffic networks with machine learning. In Proceedings of the 2013 IEEE International Joint Conference on Neural Networks (pp. 4673-6129). IEEE.

[4] Gopi, Gaurav, Justin Dauwels, and Muhammad Tayyab Asif, Sridhar Ashwin., 2014, Oct. Bayesian support vector regression for traffic speed prediction with error bars. 16th International IEEE Conference on Intelligent Transportation Systems. IEEE.

[5] Alfred Csikos, Zsolt Janos Viharos, and Tamas Tettamanti, Istvan Varga., 2015, June. Traffic speed prediction method for urban networks-an ANN Approach. 2015 Models and Technologies for Intelligent Transportation Systems (MT-ITS), pp.3-5.

[6] Haibo Mei, Athen Ma, Stefan Poslad, and Thomas O. Oshin., 2015. Short-term traffic volume prediction for sustainable transportation in an urban area. Journal of Computing in Civil Engineering, 29(2), pp.0401-4036.

[7] Jiaming Xie, Yi-King Choi., 2017, Hybrid traffic prediction scheme for intelligent transportation systems based on historical and real time data. In International Journal of Distributed Sensor Networks (Vol. 13(II)).

[8] Li Linchao, Tomislav Fratrovic, and Zhang Jian, Ran., 2017, June. Traffic speed prediction for highway operations based on a symbolic regression algorithm. Swarm Intelligence in Transportation Engineering, Vol.29, pp.433-441.

[9] Fratrovic T, Bin R, Jian Z, Linchao L., 2017 Traffic Speed Prediction For Highway Operations Based On A Symbolic Regression Algorithm. Promet – Traffc&Transportation, Vol. 29, 2017, No. 4, 433-441.

[10] Hamed, M.M., Al-Masaeid, H.R. and Said, Z.M.B., 1995. Short-term prediction of traffic volume in urban arterials. Journal of Transportation Engineering, 121(3), 249-254.

[11] Vlahogianni, E.I., Karlaftis, M.G. and Golias, J.C., 2014. Short-term traffic forecasting: Where we are and where we're going. Transportation Research Part C: Emerging Technologies, 43, 3-19.

[12] Mingheng, Z., Yaobao, Z., Ganglong, H. and Gang, C., 2013. Accurate multisteps traffic flow prediction based on SVM. Mathematical Problems in Engineering, 2013.

[13] Rahman, F.I., 2020. Short Term Traffic Flow Prediction Using Machine Learning-Knn, Svm And Ann With Weather Information. International Journal for Traffic & Transport Engineering, 10(3).

[14] Kumar, T.S., 2020. Video based Traffic Forecasting using Convolution Neural Network Model and Transfer Learning Techniques. Journal of Innovative Image Processing (JIIP), 2(03), 128-134.

[15] Anand, J.V., 2020. Automatic Traffic Control Technologies for Remote Monitoring Of Unmanned Railway Gates. Journal of Electronics, 2(01), 30-37.