A Comparative Analysis of Machine Learning Algorithms in Design Process of Adaptive Traffic Signal Control System

Savithramma R M¹, R Sumathi¹, H S Sudhira²
¹Department of Computer Science and Engineering
Siddaganga Institute of Technology, Tumakuru, Karnataka, India
²Director, Gubbi Labs, Tumakuru, Karnataka, India
savirmrl@sit.ac.in

Abstract. In recent decades machine learning technology has proved its efficiency in most sectors by making human life easier. With this popularity and efficiency, it is applied to design traffic signal control systems to mitigate traffic congestion and distribute waiting delays. Hence, many researchers around the world are working to address this issue. As a part of the solution, this article presents a comparative analysis of various machine learning models to come up with a suitable model for an isolated intersection. In this context, eight machine learning models including Linear Regression, Ridge, Lasso, Support Vector Regression, k-Nearest Neighbour, Decision Tree, Random Forest, and Gradient Boosting Regression Tree are selected. Shivakumara Swamiji Circle (SSC), one of the intersections in Tumakuru, Karnataka, India is selected as a case study area. Essential data is collected from SSC through videography. The selected models are developed to predict green time based on traffic classification and volume in Passenger Car Units (PCU) for each phase on the PyCharm platform. The models are evaluated based on various performance metrics. Results revealed that all the selected models predict green splits with 91% accuracy using traffic classification as input, whereas, models showed 85% accuracy with PCU as input. And also, Gradient Boosting Regression Tree is the best suitable model for the selected intersection, whereas, Decision Tree is not referred model for this application.

1. Introduction
Traffic congestion is the major issue experienced in urban areas all around the world [1]. The primary cause of congestion is urban expansion [2] in recent decades. It is observed that the volume of automobiles entering the road is increasing day-by-day due to many reasons like population growth [3], industrialization, economical growth [4], advancement in technology, etc., this in-turn lead to traffic congestion. Even with the effort of government and local administration to improve urban transport, the problem of congestion has been exacerbated by the intensive use of private vehicles [5]. Further, the congestion leads to problems like elongated travel time, additional fuel consumption, cost overhead [6], environmental pollution, and so on. As the waiting delay at intersections contributes more for overall elongated travel time [7], there is a necessity for an efficient signal planner at intersections thereby, reducing queuing delay and ensuring safety as well.

Currently, there exists a variety of Traffic Signal Controllers ranging from pre-timed to intelligent controllers. Progress in technology has supported the implementation of an advanced Traffic Signal
Control System (TSCS). Machine Learning (ML) is one of the cutting-edge technology wherein, an electronic machine learns by itself from experience gained through data. ML algorithms can estimate adaptive signal timings for each phase based on traffic demand dynamically. ML models are eliminating the necessity of traffic police at intersections by empowering TSCS with autonomous capabilities. However, the design of TSCS is location-specific as the traffic volume, traffic structure, traffic rules, and driver behaviour vary from place to place. In this regard, researchers across the world have applied various technologies to design an efficient TSCS in general and location-specific as well.

Some of the authors have surveyed available TSCS design solutions and summarized them in [8][9][10]. The ML model which is best suitable for Tier-I cities may not perform in the same manner for Tier-II/III cities or vice versa. Hence, in this article, an attempt has been made to choose the suitable model for Tumakuru city for predicting the signal plan of TSCS. The main objective of this article is to predict the adaptive traffic signal configuration for an isolated intersection in Tumakuru city under heterogeneous traffic conditions using various ML algorithms and put forward a suitable model for the city. With this objective eight ML models are selected and developed to predict the traffic signal plans for each phase with two different types of input data; (1) Traffic classification (count of each type of vehicle), and (2) Traffic volume in terms of Passenger Car Unit (PCU), where, PCU is a measure to convert unstructured traffic to structured traffic (all different type of vehicles is converted into cars). In this article, the traffic is converted into PCU as per guidelines provided by IRC-109[11].

Tumakuru is a Tier-II city in Karnataka, one of the Southern states in India. Tumakuru is a hub of industrial areas near Bangalore. Being a city of education and land of religious places, there is an increase in commuters thereby increasing the traffic. The traffic in Tumakuru is heterogeneous as shown in figure 1. Tumakuru is selected under India Smart City Project and it is in progress, as a result, seven intersections along the national highway Bangalore-Honnavar (BH) are signalized. Each signalized intersection is equipped with high-definition cameras to monitor and control the traffic flow. Adaptive TSCS is in place, but still, there is a scope of improvement to the existing system.

**Figure 1.** Heterogeneous traffic in Tumakuru (Courtesy: Tumakuru Smart City Limited)

The remaining sections of the paper are ordered as follows; the next section summarizes the existing research work concerning the application of various technologies to design TSCS. The methodology followed to carry out the study is presented in section 3. The required data for developing selected ML models and the way of data collection is briefed in section 4. Whereas,
section 5 introduces the ML models selected for comparison in the study. The obtained results are presented in section 6 and discussed in detail in section 7. Finally, the paper concludes with the findings of the analysis carried out.

2. Literature
Traffic congestion, particularly at intersections is the major problem faced across the world. Congestion leads to series of problems as discussed in the previous section. But, the congestion at intersections may be alleviated to some extent with the help of advanced TSCS. Thus, many researchers around the world are working in this direction to enhance the capabilities of TSCS with the development of technology. Some of the research works in this respect are briefly in this section.

Authors in [12] used Q-learning along with fast gradient-descent as an approximation function to decide the optimal action on signal configuration. The effectiveness of the proposed algorithm is demonstrated through simulation. An intelligent traffic signal scheduling system is introduced by [13] by applying a genetic algorithm in combination with an ML algorithm to reduce the waiting time. A linear regression algorithm is used to decide the next phase timing. Accurate traffic volume prediction is an important phase of traffic signal scheduling. In this view, the authors in [14] proposed an ensemble ML model to predict the traffic at each approaching road. The authors in [15] presented a comparative analysis of optimization algorithms in optimizing phase lengths of fixed traffic signal controllers. Authors in [16] and [17] have proposed a solution to tackle the problem of congestion at an intersection by applying a decision tree classification algorithm, whereas, [18] have used gradient descent optimization technique.

An adaptive traffic signal controller has been proposed in [19] by applying linear regression to predict the green length for each phase to reduce vehicle queuing during the red period at the stopped phase. An optimized dynamic signal timing strategy by considering pedestrian volume as the main parameter is presented in [20] by using radar data to reduce waiting delay. The support vector machine is used as a regressor to hypothesize queuing delay. The article [21] presents a new traffic signal switching system designed with object detection through ML techniques. The evolutionary algorithm is applied to optimize the signal strategy to reduce the pedestrians and vehicle waiting time.

A novel method to extract essential traffic parameters such as vehicle volume, red time length, the pedestrian volume associated with the signalized highway is proposed in [22] and involves data mining k-Nearest Neighbour classification algorithm and image processing. Then the extracted information is used to optimize the signal configuration. The authors in [23], [24], and [25] applied fuzzy logic, artificial neural networks, and genetic algorithm respectively to decide the traffic signal configuration. Researchers across the world are continuously working on signal operation strategies at signalized intersections in-order to enhance the autonomous features of the traffic controller. In this direction, ML is extensively applied. The recent research in [26][27] used The reinforcement learning technique to develop an adaptive and intelligent traffic signal controller.

3. Methods
The abstract view of the study method followed in this article is presented in figure 2. The proposed study involves various phases. The study begins with the data collection from identified study intersection followed by the development of models. The selected ML models are developed to predict adaptive traffic signal plans based on traffic demand on approaching roads at an intersection using 70% of data collected as a training set and 30% as a test set on the PyCharm platform. Then a comparative analysis is conducted on traffic signal plan predicted by each model using various performance metrics in-order to decide the suitable model for the isolated signalized intersections in the city. Finally, the study concludes with remarks on the outcomes of the empirical results obtained. The detailed discussion of each phase is presented in the following sections.
Data Collection
Data is an important component in the process of machine learning to gain efficiency concerning decision-making. There exist various types of data, but, ML models mainly rely on four types of data forms including categorical, time series, numerical, and text. In this article, numerical data is used to train the selected ML models. In this context, Shivakumara Swamiji Circle (SSC), one of the intersections in Tumakuru, is selected for the study. SSC is a five-legged (SR1, SR2, SR3, SR4, and SR5) intersection with traffic flow permitted in all directions from one approaching road in one phase. SR1 and SR5 are having dedicated free left as depicted in figure 3.
With the advancement in technology, there exist different ways to collect traffic data and some of the methods are presented in [28]. In this article, data is collected through videography. The traffic movement and signal operation recorded videos are collected from Tumakuru Smart City Limited. The data necessary for the proposed study is extracted manually for one week (05/04/2021 to 10/04/2021) from 10:00 am to 12:30 pm by watching the recorded videos. The collected data is categorized into traffic data and signal data. The traffic data includes the count of each type of vehicle accumulated at each approaching road during the red time. Whereas, signal data includes information regarding green splits, red span, and cycle length in seconds.

As the traffic is heterogeneous, the whole traffic is classified into five categories such as two-wheelers, auto-rickshaws, Cars, Light Commercial Vehicles (LCV), and buses based on the Indian scenario. A heatmap is generated using the seaborn library in python for the data collected at each approaching road (denoted as RID) as shown in figure 4. Heatmap is the graphical means of visualizing the data. It is a two-dimensional plot representing the co-relation between parameters against each other. The green time shows a higher co-relation with two-wheelers, auto-rickshaws, and cars than the RID, LCV, and buses.

**Figure 4. Data representation with heatmap**

5. **Machine Learning Models**

The application of ML techniques to the design of TSCS has eased traffic management at intersections. In line with the literature discussed in section 2, eight ML algorithms are selected to check their suitability for designing TSCS for Tumakuru city. Among eight models, Linear Regression (LR), Ridge, Lasso (Least Absolute Shrinkage and Selection Operator), and Support Vector Regression (SVR) were selected under linear models, whereas, k-Nearest Neighbour (k-NN), Decision Tree (DT), Random Forest (RF) and Gradient Boosting Regression Tree (GBRT) are selected from non-linear category. A brief introduction of all selected ML models is as follows.
**LR** is one of the well-known and simple algorithms among the statistical and ML models. It is formulated based on line theory as given in equation (1), where \( Y \) is the target (dependent) variable to be predicted, \( X_1, X_2, X_3 \) are the input (independent) variables, whereas, \( C_0, C_1, C_2, C_3 \) are the regression coefficients estimated through learning.

\[
Y = C_0 + C_1X_1 + C_2X_2 + C_3X_3 + \ldots \quad (1)
\]

**Ridge** and **Lasso** are the two variants of LR. Both are used to overcome the problem of multicollinearity in the data. Lasso uses one of the more influential independent variables to predict the target value, whereas, Ridge uses influential independent variables with varying coefficients. The mathematical model is similar for both algorithms as given in equation (2), where \( e \) is the error term with normal distribution and variance \( \sigma^2 \) as given in equation (3).

\[
Y = CX + e \quad (2)
\]
\[
e \sim N(0, \sigma^2) \quad (3)
\]

The basic concept of **SVR** is to find the best hyperplane that fits the data with more data points lying on it. The linear function to predict the target is given by equation (4), which is similar to LR, except, the line is more flat as possible.

\[
f(x) = b_0 + b_1x' \quad (4)
\]

**k-NN** is used for both classification and regression. It will keep track of all possible classes. Whenever a new data point encounters, it will be assigned to one of the classes which are nearest to it. The distance is computed based on the Euclidian distance between the mean of the particular class and the new data point. Same way, it regresses the values by averaging from the neighbours. The **DT** algorithm traverses the data in a tree form to estimate the target value by considering each dependent variable (feature) as a node along the root branches. The branching is based on control statements like if then else. Whereas, **RF** and **GBRT** are extended versions of decision trees, composed with a set of DTs at each node. Merging of decisions starts from the beginning itself in the case of GBRT, but the decisions are merged at the leaf level in RF.

6. **Results**

The experimental results obtained from each model are presented in this section. Two cases have been considered to assess the efficiency of models’ prediction ability;

**Case-1**: Green length prediction based on traffic classification at each phase.

**Case-2**: Green length prediction based on traffic volume in terms of PCU at each phase.

The traffic classification is the count of each category of a vehicle such as two-wheelers, auto-rickshaws, cars, LCVs and buses, whereas, PCU corresponds to the homogeneous traffic volume which is computed as per IRC-109[11]. The empirical results obtained are presented in the form of graphs as shown in figure 5. The scatter plot of actual versus predicted green splits is presented in figures 5(a)-(h) and 5(i)-(p) for case-1 and case-2 respectively. The actual green time in seconds is plotted along the x-axis, while, the y-axis represents the predicted green time in seconds. When the predicted values exactly coincide with the actual values it forms a straight line by subtending an angle with 45° as drawn in red colour. With this reference, it is observed that most of the predicted data points lie on and around the red line in case-1, while, most of the data points are highly scattered and relatively placed away from the red line in case-2.
Case-1

(a)

Case-2

(i)

(b)

(j)

(c)

(k)

(d)

(l)
Figure 5. Scatter plot of predicted v/s actual green splits in case-1 and case-2
(a, i) LR  (b, j) Ridge  (c, k) Lasso  (d, l) SVR  (e, m) k-NN  (f, n) DT  (g, o) RF  (h, p) GBR
7. Discussion
A comparison of ML models such as LR, Ridge, Lasso, SVR, k-NN, DT, RF, and GBRT is implemented to predict traffic signal plans for an isolated intersection in Tumakuru. The comparative analysis is applied in two folds as defined below:
   a. Comparison between case-1 and case-2
   b. Comparison between selected models

7.1. Comparison between case-1 and case-2
Case-1 and case-2 are compared based on two types of error metrics Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). The error computed for an individual model in case-1 and case-2 is presented in the form of bar charts as shown in figures 6 and 7. The MAPE and RMSE are smaller for each model in case-1 as compared to case-2. This implies that the selected models can predict green splits with less deviation from actual values in case-1 when compared to case-2. Hence, all the models selected in this article perform better with classification as input in comparison with PCU (overall traffic volume) as input.

![Figure 6. MAPE computed for various machine learning models](image)

![Figure 7. RMSE computed for various machine learning models](image)

7.2. Comparison between selected models
The selected models are compared based on performance metrics including error and r2-score. Various types of errors such as MAPE, Mean Square Error (MSE), RMSE, Median Absolute Error (MdAE) and Mean Absolute Error (MAE) are estimated and presented in Table 1 and 2 for case-1 and case-2 respectively. The predictions of LR, Ridge, and GBRT showed less error with a negligible variation. As the data pattern reflects a linear relationship between independent and dependent variables, all linear models perform better than the non-linear models except GBRT. The GBRT, an extended non-linear model can generate better predictions than the other three non-linear models.
Table 1. Models validation through various types of metrics in Case-1

| Error Type | Linear Models | Non-linear Models |
|------------|---------------|-------------------|
|            | LR | RIDGE | LASSO | SVR | k-NN | DT | RF | GBR |
| MAPE       | 0.08398 | 0.08398 | 0.08613 | 0.08464 | 0.08968 | 0.10896 | 0.09070 | 0.08464 |
| MSE        | 8.16573 | 8.16577 | 8.56190 | 8.30762 | 9.14602 | 16.07066 | 9.00121 | 8.02795 |
| RMSE       | 2.85757 | 2.85758 | 2.92607 | 2.88229 | 3.02424 | 4.00882 | 3.00020 | 2.83336 |
| MdAE       | 1.89312 | 1.89309 | 1.94895 | 1.88125 | 2.01748 | 2.57389 | 2.04535 | 1.91709 |
| MAE        | 1.07105 | 1.07076 | 1.10232 | 0.92277 | 1.20000 | 2.00000 | 1.36483 | 1.26101 |

Table 2. Models validation through various types of metrics in Case-2

| Error Type | Linear Models | Non-linear Models |
|------------|---------------|-------------------|
|            | LR | RIDGE | LASSO | SVR | k-NN | DT | RF | GBR |
| MAPE       | 0.12531 | 0.12532 | 0.13242 | 0.12692 | 0.11379 | 0.12910 | 0.11825 | 0.10412 |
| MSE        | 15.0933 | 15.0938 | 15.9923 | 15.2024 | 14.0731 | 20.1284 | 15.2386 | 12.4697 |
| RMSE       | 3.88502 | 3.88507 | 3.99904 | 3.89903 | 3.75142 | 4.48647 | 3.90367 | 3.53126 |
| MdAE       | 2.88840 | 2.88842 | 2.97692 | 2.89768 | 2.79935 | 3.24218 | 2.92861 | 2.56019 |
| MAE        | 1.95167 | 1.95222 | 2.18457 | 1.95388 | 2.20000 | 2.25000 | 2.17500 | 2.08558 |

Figure 8. r2-score computed for each model in case-1 and case-2

Also, the r2-score of an individual model is computed to check the accuracy of each model and it is depicted in figure 8. Usually, r2-score lies between 0 and 1, where, 0 indicates worst and 1 implies the best possible predictions. It is observed that the r2-score of LR, Ridge, and GBRT is higher than that of Lasso, SVR, RF and DT under both cases. However, Lasso, SVR, and RF have the next highest r2-score, whereas, DT has scored lowest. Though the LR, Ridge, and GBRT showed almost the same forecasting ability, with a slight variation in predictions the GBRT is more suitable for designing adaptive TSCS, whereas, the DT is not suitable as it showed very poor prediction ability. However, the traffic classification count estimation is a limitation of the proposed idea. In this article, the classification is extracted manually by watching the recorded videos. But edge computing, one of the cutting-edge-technology, can support the idea. The suggestion of the study can be incorporated to
implement an adaptive traffic signal planner by using edge computing. The traffic classification count can be detected automatically through image processing and this information is used by the prediction model to estimate the signal length for each phase and communicated to the controller to activate the signal lights accordingly. The complete process has to be carried out on an edge device installed at the intersection.

8. Conclusion

This article presents a study carried out to identify the suitable ML model to design an adaptive traffic signal control system for an isolated intersection of a Smart city. With this objective, eight ML models and one of the intersections in Tumakuru city namely Shivakumara Swamiji Circle are selected for the study. Among eight models, LR, Ridge, Lasso, and SVR are the linear, whereas, k-NN, DT, RF and GBRT are the non-linear models. The essential data including the count of each type of vehicle on each approaching road in every cycle along with green time is collected from the study intersection through recorded videos. The selected models are developed on the PyCharm platform to predict the signal plan as per traffic demand at each approaching road of an intersection. Then, the models are compared using performance metrics including error and r2-score. Various types of errors such as MAPE, MSE, RMSE, MdAE, and MAE for each model in case of classification and PCU as input are computed. The computed error and r2-score revealed that the selected models perform better with traffic classification as input as compared to the PCU input. Also, the GBRT, LR, and Ridge showed the best prediction ability with the least deviation from actual green splits. Therefore, the study has come up with a conclusion that among eight models developed in this article, GBRT is the best suitable ML model, whereas, DT is not suggested for signal operation implementation at signalized intersections in the city. But, the traffic classification count detection is the limitation to implement the proposed idea. However, it can be resolved by adopting edge computing technology along with image processing for traffic detection. The study presented in this article is useful for concerned city transportation management authorities.

Acknowledgment

Tumakuru Smart City Limited (TSCL), Karnataka, India, has supported this research work by providing the traffic data required to carry out the study. The authors are thankful to the respected director and other employees of TSCL for their kind and continuous co-operation in conducting the study presented in this article.

References

[1] A. Bull 2004 Traffic Congestion - The Problem and How to Deal with it?.
[2] United Nations 2018 World Urbanization Prospects 2018.
[3] H. P. Kohler 2012 Population growth. Glob. Probl. Smart Solut. Costs Benefits, pp. 510–580, doi: 10.1017/CBO9781139600484.011.
[4] C. I. Jones 2016 The Facts of Economic Growth. 1st ed., vol. 2. Elsevier B.V.
[5] Directorate of Urban Land Transport, Urban Land Development, Department of Karnataka, India 2014 Comprehensive Mobility Plan for Tumkur Final Report.
[6] A. G. L. Peixoto Neto, M. L. Galves, O. F. Lima Júnior, and D. Tacla 2008 Challenges of urban transport problems and city logistics: Sao Paulo city center case. WIT Trans. Built Environ., vol. 101, pp. 133–142, 2008, doi: 10.2495/UT080131.
[7] Washington 2005 Traffic Level of Service Calculation Methods Appendix B [Online]:http://www.ccag.ca.gov/pdf/documents/archive/cmp_2005_Appendix_B.pdf.
[8] Qadri S S S M, Gökçe M A and Öner E 2020 State-of-art review of traffic signal control methods: challenges and opportunities. Vol. 12, European Transport Research Review. Springer Science and Business Media Deutschland GmbH
[9] Savithramma R M and Sumathi R 2020 Road traffic signal control and management system: A survey. Proc 3rd Int Conf Intell Sustain Syst ICISS 2020.104–10.
[10] Astarita V, Giorfè V P, Guido G and Vitale A 2020 A review of traffic signal control methods and experiments based on Floating Car Data (FCD). Procedia Comput Sci [Internet]. 2020;175:745–51. Available from: https://doi.org/10.1016/j.procs.2020.07.110

[11] Indian Roads Congress 1990 Guidelines for Capacity of Urban Roads in Plain Areas. New Delhi. Vol. 106, IRC Code of Practice.

[12] Liu W, Liu J, Peng J and Zhu Z 2014 Cooperative multi-agent traffic signal control system using fast gradient-descent function approximation for V2I networks. 2014 IEEE International Conference. Commun ICC. 2562–7.

[13] Zhao B, Zhang C and Zhang 2015 L Real-Time Traffic Light Scheduling Algorithm Based 1:385–98.

[14] Alajali W, Zhou W, Wen S and Wang Y 2018 Intersection traffic prediction using decision tree models. Symmetry (Basel). Sep 7;10(9).

[15] Malecki K, Pietruszka P and Iwan S 2017 Comparative analysis of selected algorithms in the process of optimization of traffic lights. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 10192 LNAI(April 2018):497–506.

[16] Rithesh R N, Vignesh R and Anala M R 2018 Autonomous traffic signal control using decision tree. Int J Electr Comput Eng. 8(3):1522–9.

[17] Sadollah A, Gao K, Zhang Y, Zhang Y and Su R 2019 Management of traffic congestion in adaptive traffic signals using a novel classification-based approach. Eng Optim [Internet]. 51(9):1509–28. Available from: https://doi.org/10.1080/04209006.2018.1525708

[18] Mozejko M, Brzeski M, Madry L, Skowronek L and Gora P 2018 Traffic Signal Settings Optimization Using Gradient Descent. Schedae Informaticae.19–30.

[19] Khadilkar A, Kasodekar K S, Sharma P and Priyadarshini J 2019 Intelligent traffic light scheduling using linear regression [Internet]. Vol. 697, Advances in Intelligent Systems and Computing. Springer Singapore. 329–335 p. Available from: http://dx.doi.org/10.1007/978-981-13-1822-1_30

[20] Qu X, Guo T, Guo J, Lin Y and Ran B 2019 A real-time signal control strategy at an isolated pedestrian crossing based on radar data. Adv Mech Eng. 11(3):1–7.

[21] Ng S C and Kwok C P 2020 An intelligent traffic light system using object detection and evolutionary algorithm for alleviating traffic congestion in hong kong. Int. J. Comput. Intell. Syst. 13(1):802–9.

[22] Abdullah Yusefi, Adem Alpaslan Altun and Cemil Sungur 2020 Data Mining in A Smart Traffic Light Control System Based on Image Processing and KNN Classification Algorithm. Eur J Sci Technol. 461–5.

[23] Babhulkar G 2018 Fuzzy Based Adaptive Two Phase Traffic Signal Controller. HELIX. 31;8(5):4010–3.

[24] De Oliveira M B W and De Almeida Neto A 2014 Optimization of traffic lights timing based on Artificial Neural Networks. In: 2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014. Institute of Electrical and Electronics Engineers Inc. p. 1921–2.

[25] Teo K T K, Kow W Y and Chin Y K 2010 Optimization of traffic flow within an urban traffic light intersection with genetic algorithm. In: Proceedings - 2nd International Conference on Computational Intelligence, Modelling and Simulation, CIMSim p. 172–7.

[26] Genders W and Razavi S 2019 Asynchronous n-step Q-learning adaptive traffic signal control. J Intell Transp Syst Technol Planning. Oper. 23(4):319–31.

[27] Aslani M, Seipel S, Mesgari MS and Wiering M 2018 Traffic signal optimization through discrete and continuous reinforcement learning with robustness analysis in downtown Tehran. Adv Eng Informatics.38:639–55.

[28] Ashwini B P and Sumathi R 2020 Data sources for urban traffic prediction: A review on classification, comparison and technologies. Proceedings 3<sup>rd</sup> International Conference Intelligent Sustainable Systems ICISS. 628–35.