

Prediction of Road Traffic Congestion using Deep Neural Network Technique

Emmanuel. O. Taiwo^{1*}, Gbenga O. Ogunsanwo², Olumuyiwa B. Alaba³

¹*Department of Computer Science, Alvan Ikoku Federal University of Education, Owerri, Imo State, Nigeria.*

^{2,3}*Department of Computer Science, Tai Solarin University of Education, Ijebu-Ode, Ogun State, Nigeria.*

**Corresponding Author's e-mail: emmanuel.taiwo@alvanikoku.edu.ng*

ABSTRACT

Traffic congestion remains one of the major challenges in many parts of the world, negatively impacting urban communities in various ways. These effects include the loss of valuable time, delayed deliveries, financial losses, fuel wastage, vehicle wear and tear, and increased stress levels. The current practice of manually controlling traffic by traffic wardens or managers lacks analytical capabilities. This study focuses on developing predictive model for road traffic congestion to enhance traffic management in Lagos State. The historical dataset was collected for a period of six months (February-July, 2024) on Oshodi-Badagry expressway. The dataset consists of 4600 instances with carefully selected 13 attributes which were splitted into 80% 20%; 60% 40% for training and for testing. Deep Neural Network (DNN) was employed to formulate the predictive model. The formulation and simulation of the predictive models were carried out using Python Google Colab environment. The performance of model was assessed through validating metrics such as accuracy, F1 score, precision, specificity and recall. The results showed Accuracy (%) 98, Precision of 0.98, Recall of 0.98, F1-Score of 0.98, and Specificity of 0.97. The model is recommended to support decision-making in the transportation sector so as to reduce traffic congestion in metropolitan cities.

KEYWORDS: Prediction, Traffic Congestion, Deep Neural Network, Deep Learning Technique

I. INTRODUCTION

Traffic management is a major concern in most cities of the world due to the incessant traffic congestion. Traffic congestion takes place whenever there is an increase in the number of vehicles beyond the capacity of the road (Taiwo et al. 2023). This leads to loss of valuable time, vehicle wear and tear, high consumption of fuel and accidents (Taiwo et al. 2023). Ramchandra & Rajabhushanam (2021) described traffic congestion as the situation in road network that occurs through vehicles' slower speed. To decrease ongoing problem in real-time response, timely prediction information is necessary. Traffic data increases daily but its analysis is majorly concerned with bringing out predictions of future data. One crucial step needed in the intelligent transportation system (ITS) is an accurate traffic prediction system because it will assist traffic managers in the area of route planning and traffic regulations (Jiber et al. 2020). The increased number of vehicles each day has made traffic congestion in big cities unavoidable most especially in peak hours. However, the improvements in Artificial Intelligence (AI) techniques can solve this difficulty by analyzing and learning the traffic flow patterns and forecasting traffic situation at any point in a given time (Wafa et al. 2019). Zhu et al. (2016) noted that there is a need for drivers to know traffic information when navigating through various routes. This will assist them in the

upstream and downstream traffic flow state of the road network. Whenever congestion occur in a city, because of an increase in traffic flow it will spread through road networks. By virtue of present developments in the AI discipline and an exponential increase in the real-time transportation data, Deep Learning (DL) has particular set of techniques that could be used to predict traffic congestion. DL is a subfield of machine learning (ML) that makes use of artificial neural networks (ANNs) that are capable of self-learning and inspired by the function and brain's structure (Inzunza et al., 2020). It requires using large amount of data to train an ANN. The network learns by itself and it requires enormous data to learn in deep learning. It is a crucial tool for AI research with applications in numerous areas. Traffic prediction is needed in the transport system to handle congestion. The goal of this paper is to develop a model for predicting traffic congestion for effective traffic management on Nigerian roads.

II. RELATED WORKS

Ramchandra & Rajabhushanam (2021) developed a model to forecast traffic congestion. They applied DAN, DBN, RF and LSTM to forecast traffic flow of a particular zone. The results indicated that LSTM performed better than the other three algorithms with 94.3% accuracy. Ravi et al. (2021) conducted a study on a system that can be used to analyze traffic and predict congestion on a particular path and informing ahead vehicles planning to pass through the congested path. They applied You Only Look Once (YOLO) algorithm to implement the model. The model provided accurate prediction of the traffic congestion. Odesanya & Odesanya (2021) worked on the performance analysis of traffic congestion using DNN training algorithms, aiming to develop a model to analyze traffic flow. They employed BR, SCG and LM techniques to build models. Experimental results showed that BR outperformed SCG and LM. Majumdar et al. (2021) utilized LSTM networks to predict congestion propagation on a road network. They used both univariate and multivariate algorithms to create predictive models. The analysis of these models indicated an accuracy of 84% to 95%. Hassan et al. (2021) developed a system that will control traffic congestion in smart societies. They applied IoT-ITCHS-ML. In training and validation, the system accomplished 81% accuracy.

Rajendran & Ayyasamy (2020) conducted a study for short term traffic flow prediction on a model. They employed structure pattern and regression to build models. The proposed system outperformed the conventional systems and turns out to facilitate in conserving energy. Rahman (2020) predicted short term traffic flow using ML approaches such as KNN, SVM and ANN. For better traffic flow prediction, related parameters were applied with the algorithms in this study. The results revealed that KNN produced accurate results than ANN and SVM. Mei et al. (2020) conducted a study to forecast traffic congestion. They used fuzzy comprehensive evaluation technique to divide traffic status level (TSL) as well as five various algorithms to build predictive models. The results showed that BP neural network is more accurate than other algorithms. Inzunza et al. (2020) developed a system to forecast traffic flow in a city. They employed machine learning, computer vision, deep learning, and neuronal networks to implement the model. The proposed system works relatively fast and produced promising results. Verma & Badade (2019) worked on a system to control traffic in a smart city. They utilized Long Range Wide Area Network (LoRa) and Intelligent Traffic Management System (ITMS) for traffic prediction. The results indicated that the system would produce high performance in traffic prediction.

III. METHODOLOGY

This section presents the detail framework of the materials and methods used to accomplish objective of the research in this study. Figure 1 shows the diagram flow (DF) of the study. The process includes dataset collection, data processing, dataset used, evaluation metrics, model training, and evaluation.

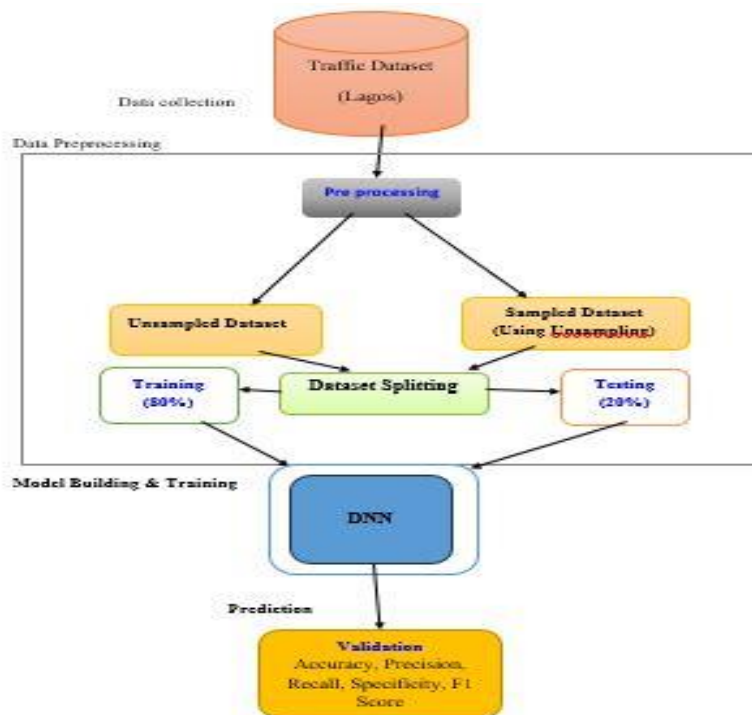


Figure 1: Diagram Flow of the Study

A. Dataset Collection

Formal interviews were conducted with Federal Road Safety Corps (FRSC) Oshodi Unit Command regarding factors causing traffic congestion in Lagos State while dataset was collected using Road Traffic Congestion Questionnaires via Google Forms from residents and road users that are plying Oshodi-Badagry expressway for a period of six months (February-July, 2024). The data collected included information such as Date, Time, Travel Time, Length of the road, road Volume, Occupancy of the road, Speed of the car, Type of road, Direction, Congestion Level, Presence of traffic control officer, Potholes, Poor drainage, Faulty vehicle, Traffic volume, Rain, Weather, Accident, Holiday, Narrow road network, Bad access roads, Presence of too many vehicles on the road, Weather description. The dataset contains 4600 instances. Attribute selection processes were performed on the collected data to identify major crucial input variables: Holiday, Narrow road network, Bad access roads, Presence of too many vehicles on the road, Accident, Weather, Weather description, Rain, Presence of traffic control officer, Potholes, Poor drainage, Faulty vehicle, Traffic volume as seen in Table 1.

Table 1: Input and Target Attributes

Input attributes	Target attributes
Holiday	Congestion
Narrow road network	
Bad access roads	
Presence of too many vehicles on the road	
Accident	
Weather	
Weather description	
Rain	
Presence of traffic control officer	
Potholes	
Poor drainage	
Faulty vehicle	
Traffic volume	

B. Data Processing

Data preprocessing techniques were performed on the dataset in order to increase the accuracy of the predictive model. They were performed using an online cloud-based Python notebook called Google CoLaboratory that was provided by Google to its users. This provided a convenient way of preprocessing the different steps in the working procedure.

C. Model Development

The dataset was splitted into 80%, 60% training to construct predictive model while 20%, 40% testing was used to validate predictive model performance.

D. Deep Neural Network (DNN)

DBN is a popular deep learning with artificial intelligence features that permits the network to learn patterns in data. DBN is an artificial neural network that helps in learning and recognizing patterns which involves different layers of restricted boltzmann machines (RBM) (Ramchandra & Rajabhushanam, 2021). The RBM layers provide a robust structure that can classify data based on different categories, and the layers of DBN follow the top-down technique, permitting communication throughout the system.

E. Evaluation Metrics

The performance evaluation metrics were used to evaluate the performance of predictive model include accuracy, precision, recall, F1 Score, specificity, execution time for each of the algorithms.

- a) **Accuracy Score:** Model accuracy measures the proportion of true positives to true negatives in relation to all observed positive and negative cases. The more the accuracy value approaching 100%, it indicates a more effective model. Essentially, accuracy reflects the frequency with which our ML model correctly predicts an outcome compared to the total number of predictions made as shown in equation 1.

$$\text{Accuracy Score} = \frac{(TP+TN)}{(TP+FN+TN+FP)} \times 100\% \quad (1)$$

Where:

TP - True Positive

TN - True Negative

FN - False Negative

FP - False Positive

- b) **Precision:** It refers to the proportion of accurate cases identified by the model. A value closer to 1 indicates the better the model.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

- c) **Recall or Sensitivity:** The number of positives identified by the model. A value closer to 1 indicates the better the model.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

- d) **Specificity:** The number of negatives returned by the model. A value closer to 1 indicates the better the model.

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

- e) **F1 Score:** F1score combines recall and precision scores of a model into a single metric to have a greater understanding of model performance.

$$F1 = \frac{2TP}{2TP+FP+FN} \quad (5)$$

- f) **Confusion Matrix:** Confusion matrix shows how many correct and incorrect predictions a classification model makes. It is useful for checking how well the model is doing by comparing what it predicts with the actual values. The confusion matrix is used to summarize how well the classification algorithm performs. Table 2 presents a Three-level classification confusion matrix that encompasses light-congestion, medium-congestion, and high congestion within the road network.

Table 2: Confusion Matrix

		Predicted Value		
Actual value	Congestion	Light	Medium	High
	Light	LL	LM	LH
	Medium	ML	MM	MH
	High	HL	HM	HH

Where: LL: Light-congestion is accurately predicted, LM: Light-congestion is predicted as Medium-congestion, LH: Light-congestion is predicted as High-congestion, ML: Medium-congestion is predicted as Light-congestion, MM: Medium-congestion is accurately predicted, MH: Medium-congestion is predicted as High-congestion, HL: High-congestion is predicted as Light-congestion, HM: High-congestion is predicted as Medium-congestion, HH: High-congestion is accurately predicted

IV. RESULTS AND DISCUSSION

DNN prediction model was created using road network datasets. As a result, it produced confusion matrix as shown in Figure 2 and Table 4 which can be used to calculate performance evaluation metrics. The confusion matrix showed that out of 728 data classified as predicted value in row 1, 725 was actually classified as light-congestion, 3 was classified as medium-congestion while 0 was classified as high-congestion. It also showed that out of 736 data classified as predicted value in column 1, 725 was actually classified as light-congestion, 11 was classified as medium-congestion while 0 was classified as high-congestion. The confusion matrix showed that out of 126 data classified as predicted value in row 2, 11 was actually classified as light-congestion, 109 was classified as medium-congestion while 6 was classified as high-congestion.

Table 3: Results of DNN Model

Validation Metric	Result
Accuracy	98%
Precision	0.98
Recall	0.98
F1-Score	0.98
Specificity	0.97

It also showed that out of 112 data classified as predicted value in column 2, 3 was actually classified as light-congestion, 109 was classified as medium-congestion while 0 was classified as high-congestion. The confusion matrix showed that out of 72 data classified as predicted value in row 3, 0 was actually classified as light-congestion, 0 was classified as medium-congestion, while 72 was classified as high-congestion. It also showed that out of 78 data classified as predicted value in column 3, 0 was actually classified as light-congestion, 6 was classified as medium-congestion, while 72 was classified as high-congestion.

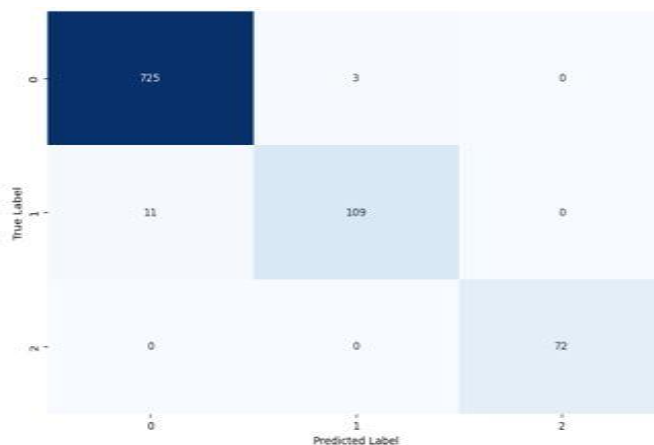


Figure 2: Python-based DNN Confusion Matrix

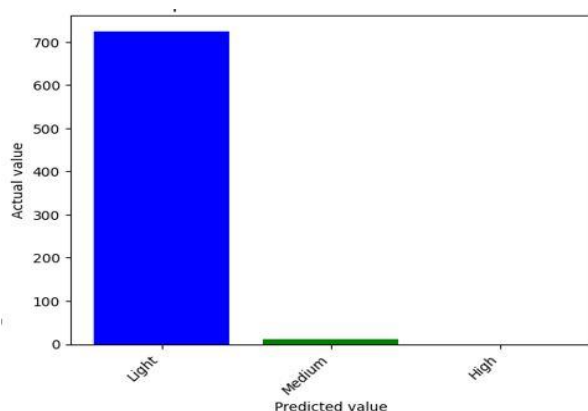
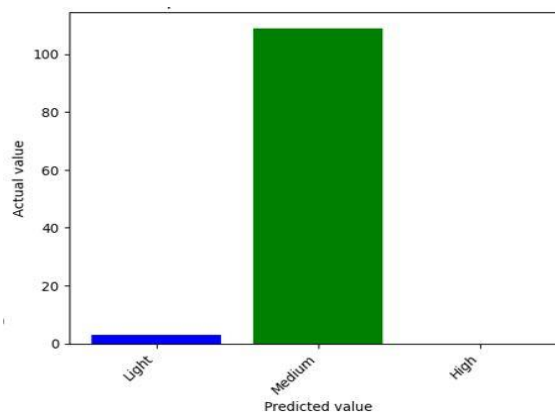
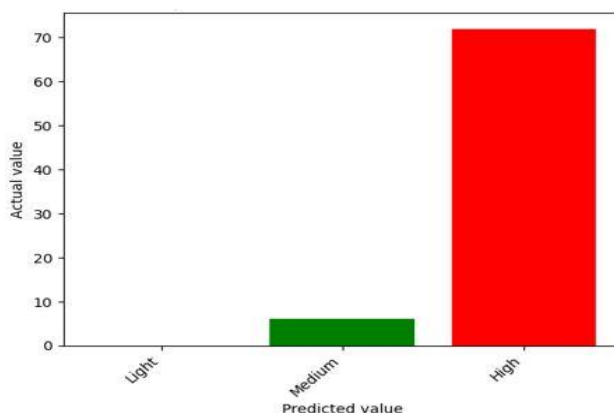
Table 4 presents the Python-based DNN confusion matrix, showing the classification performance of the model across the three congestion levels: light, medium, and high. The table provides the actual and predicted classifications, thereby showing the correctly classified instances as well as the cases where the model misclassified one congestion level as another.

Table 4: Python-based DNN Confusion Matrix

		Predicted Value		
		Light	Medium	High
Actual value	Light	725	3	0
	Medium	11	109	6
	High	0	0	72

A. Discussion on DNN Prediction Model

The results indicated that the DNN had accuracy of 98% which indicated a very high performance of the model. It also recorded a precision of 0.98, a recall of 0.98, an F1 Score of 0.98, and a specificity of 0.97. From the results above, a value closer to 1 indicates the better the model. By inputting datasets to the DNN prediction model, it obtained prediction probabilities for three levels of congestion. Additionally, the bar chart visually displayed the probabilities for 'Light-congestion', 'Medium-congestion', and 'High-congestion' are shown in Figure 3, Figure 4, and Figure 5 respectively.

**Figure 3:** Light-congestion**Figure 4:** Medium-congestion**Figure 5:** High-congestion

The probability distribution graph illustrated the three levels of congestion predictions in descending order: Medium-congestion, Light-congestion, and High-congestion. For the specific dataset tested, there was a notably higher probability of Medium-congestion while High-congestion had a lower probability.

B. DNN Accuracy and Loss Graphs

The DNN graphs for the accuracy and loss at epoch 1000 for the train and validation accuracy as shown in Figure 6 and figure 7. Figure 6 shows that the model was able to predict the congestion with high accuracy while figure 7 shows that the model is trained with many epochs to produce loss graph. The model learned not only the useful patterns in the training but it also learned specific details of the training set. As a result of that, there was a gap between training loss and validation loss.

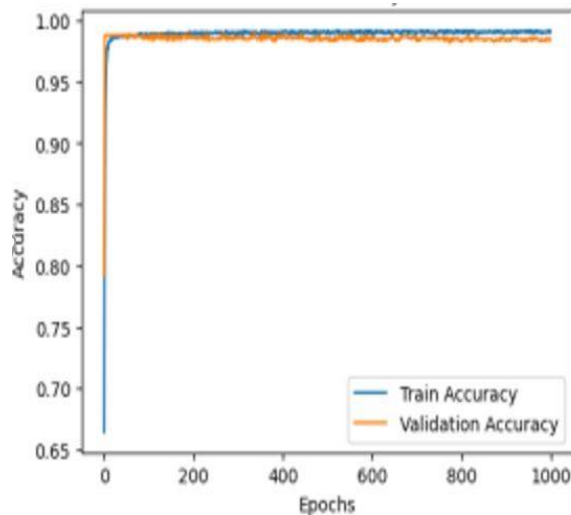


Figure 6: DNN Accuracy graph

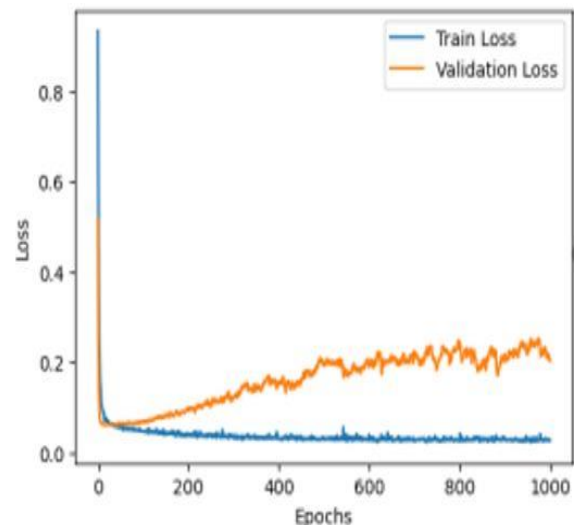


Figure 7: DNN Loss graph

The study of prediction of road traffic congestion using Deep Neural Network Technique in this paper is compared with the results of the study carried out by Majumdar et al. (2021) on congestion prediction for smart sustainable cities using IoT and machine learning approaches. The model of the study achieved accuracy 84% to 95%. This study outperforms the study with accuracy of 98%. Also compared with study carried out by Hassan et al. (2021) on Iot enabled intelligent traffic congestion handling system empowered by machine learning where the model achieved accuracy result of 81%. Our model outperforms the model with an accuracy of 98%.

V. CONCLUSION

This study analyzed the effectiveness of DNN technique on traffic dataset collected through via Google Forms from residents and road users that are plying Oshodi-Badagry expressway. 13 attributes were carefully selected with 4600 instances from the dataset which were further divided into 80%, 60% for training and 20%, 40% for testing. The dataset was used to formulate predictive model for traffic congestion. The results showed Accuracy of 98%, Precision of 0.98, Recall of 0.98, F1 Score of 0.98, and Specificity of 0.97. The model is recommended for transportation sector to assist in decision-making for metropolitan traffic congestion challenges.

REFERENCES

1. Hassan, F., Ijaz, A., Ali, M., Afzal, Z. & Arslan, F. (2021). Iot enabled intelligent traffic congestion handling system empowered by machine learning. *International Journal of Scientific & Technology Research (IJSTR)*, 10(6), 22-26.
2. Inzunza, M. C., Robles, L. H., Mancilla, M. A. & Neri, E. L. (2020). Traffic prediction architecture based on machine learning approach for smart cities. *Research in Computing Science*, 149(11), 23-33.
3. Jiber, M., Mbarek, A., Yahyaouy, A., Sabri, M. A. & Boumhidi, J. (2020). Road traffic prediction model using extreme learning machine: The case study of Tangier, Morocco. *Information*, 11, 542-549.
4. Majumdar, S., Subhani, M. M., Roullier, B., Anjum, A. & Zhu, R. (2021). Congestion prediction for smart sustainable cities using IoT and machine learning approaches. *Sustainable Cities and Society*, 64(102500), 1-10. <https://doi.org/10.1016/j.scs.2020.102500>
5. Mei, Y., Hu, T. & Yang, L. C. (2020). Research on short-term urban traffic congestion based on fuzzy comprehensive evaluation and machine learning. *Proceedings of the 5th International Conference on Data Mining and Big Data (DMBD)*, 95–107. Belgrade, Serbia, July 14–20, 2020.
6. Odesanya, I. & Odesanya, F. M. (2021). Performance analysis of traffic congestion using designated neural network training algorithms. *International Journal of Engineering and Technologies (IJET)*, 20, 23-33. <https://doi:10.18052/www.scipress.com/IJET.20.23>
7. Rahman, F. I. (2020). Short term traffic flow prediction using machine learning – KNN, SVM and ANN with weather information. *International Journal for Traffic and Transport Engineering*, 10(3), 371-389. [https://dx.doi.org/10.7708/ijtte.2020.10\(3\).08](https://dx.doi.org/10.7708/ijtte.2020.10(3).08)
8. Rajendran, S. & Ayyasamy, B. (2020). Short-term traffic prediction model for urban transportation using structure pattern and regression: An Indian context. *A Springer Nature Journal. SN Applied Sciences* 2020(2), 1159-1168. <https://doi.org/10.1007/s42452-020-2946-2>
9. Ramchandra, N. R., & Rajabhushanam, C. (2021). Traffic Prediction System Using Machine Learning Algorithms. Retrieved from <https://doi/10.4108/eai.7-6-2021.2308860>
10. Ravi, A., Nandhini, R., Bhuvaneshwari, K., Divya, J. & Janani, K. (2021). Traffic management system using machine learning algorithm. *International Journal of Innovative Research in Technology (IJIRT)*, 7(11), 303-308. ISSN: 2349-6002.
11. Taiwo, E. O., Ogunsanwo, G. O., Alaba, O. B. & Ogunbanwo, A. S. (2023). Traffic congestion prediction using supervised machine learning algorithms. *TASUED Journal of Pure and Applied Sciences (TJOPAS)*, 2(1), 110-116.
12. Verma, S. & Badade, S. (2019). Traffic Prediction Using Machine Learning. *Proceedings of National Conference on Machine Learning*, 138-141. March 26, 2019. ISBN: 978-93-5351-521-8
13. Wafa Shafqat, Sehrish Malik, Yung-cheol Byun & Do-Hyeun Kim (2019). A short term traffic flow prediction based on recurrent neural networks for road transportation control in ITS. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 8(3C), 2278-3075.
14. Zhu, G., Song, K., Zhang, P., & Wang, L. (2016). A traffic flow state transition model for urban road network based on hidden Markov model. *Neurocomputing*, 214, 567–574. <https://doi:10.1016/j.neucom.2016.06.044>