Machine Learning Approach for Modeling the Lateral Movement Decisions of Vehicles in Heterogeneous Traffic Conditions

Prashant Baviskar[1], Shriniwas Arkatkar[2] and Anshuman Sharma[3]

[1] M. Tech Scholar, Sardar Vallabhbhai National Institute of Technology, Surat, India
[2] Associate Professor, Sardar Vallabhbhai National Institute of Technology, Surat 39, India
[3]Centre for Infrastructure, Sustainable Transportation and Urban Planning (CiSTUP), Indian Institute of Science, Bangalore

Abstract. The article describes the modeling of lateral movement decisions of motorized passenger vehicles like Cars, Motorised Three Wheelers(3W), and Motorised Two Wheelers(2W) under heterogeneous traffic conditions. The supervised machine learning approach was used to predict the lateral movement decision by treating the decision of lateral movement as a multi-class classification problem. Based on surrounding vehicles' information, a set of parameters was identified that potentially affect the decision-making process of drivers to change their lateral position. With the help of these parameters, the prediction ability of machine learning algorithms was compared. It was identified that these algorithms could predict the lateral movement decision of vehicles with an acceptable accuracy range. The results revealed that the random forest method outperformed all other algorithms and appeared to be a potential contender for modeling lateral movement decisions. Real-time position information about nearby vehicles may be gathered using advanced sensors and analyzed using developed models, allowing for the provision of safety features linked to lateral movement.

Keywords: Lateral movement, Surrounding vehicles, Machine learning.

1 Introduction

In developing nations such as India, the traffic conditions are highly heterogeneous, which differs from homogeneous traffic due to the vehicles' vast range of operational and performance characteristics. Under such traffic conditions, the driver always looks for a sufficient gap to move forward at the desired speed. As a result of this tendency, the vehicle's lateral position frequently varies as it moves longitudinally. Furthermore, drivers' proclivity to go laterally is influenced by traffic flow conditions. The presence of other vehicles on the road does not affect the driving behavior of the subject vehicle in free-flow conditions. However, when traffic volume increases, adjacent vehicles influence driving behavior of the subject vehicle.

Also, in the era of connected and autonomous vehicles, this technology is expected to have exceptional driving comfort with maximal safety and minimum effect on the environment [1].These vehicles will soon become an efficient transportation mode. But the transition from traditional vehicles to autonomous vehicles will take time; hence, in this transition period, it is reasonable to assume that automated vehicles will have to

*Corresponding Author: Prashant Baviskar

prashantbaviskar81@gmail.com

K. R. Rao et al. (eds.), *Traffic and Granular Flow '22* , Lecture Notes in Civil Engineering 443, https://doi.org/10.1007/978-981-99-7976-9_42

co-exist with human-driven vehicles. This approaching transition scenario necessitates extensive study to have a more profound knowledge of human interactions with various types of vehicles and their decision-making processes. In this regard, it is critical to incorporate critical human variables into traffic models. Also, the deployment of Intelligent Transport Systems (ITS) policies is required to accumulate different driving behavior of the drivers of different vehicle classes.

One of the key elements representing the driving behavior is the lateral movement decision of the subject vehicle (SV) present in the traffic stream, which is influenced by surrounding vehicles. Also, the traffic composition in heterogeneous traffic conditions mainly consists of different vehicle classes. To model the decision-making process for lateral movements of these vehicle types with varying volume conditions has received significantly less attention. A detailed grasp of this element can serve as a foundation for prediction models that explain complex vehicle maneuvers in real-time. With this motivation, the present study proposes a systematic framework to model the lateral movement decision of Cars, Motorized Three Wheelers, and Motorized Two Wheelers.

2 Literature Review

The lateral movement studies in the case of homogeneous traffic conditions include modeling lane change decisions, duration, and execution. Mandatory lane changing (MLC) and Discretionary Lane changing (DLC) are two broad categories of lane change. Former is performed to avoid obstacles such as work zone, lane drop, tapered lane, or situations of turning at an intersection, or taking an exit to the ramp. The latter is performed to achieve the desired speed by changing the current lane to the target lane based on feasibility. Many researchers have attempted to model such lane change behavior based on decision rules [2] or discrete choice models [3].

But in the case of heterogeneous traffic conditions, drivers perform a series of lateral movements to achieve the desired speed based on the availability of the gap between the vehicles. Such driving behavior gives rise to progressive and continuous lateral shifting within or between the lane rather than a discrete lateral movement. Hence traditional lane-changing models developed for homogeneous traffic are unable to account for such lateral movement of vehicles in heterogeneous traffic conditions. Many researchers have attempted to model the lateral gaps [4], lateral placement of vehicles [5], and duration of lateral shifts [6]. For a detailed review of lateral movements studies readers can refer to [7].

Further from the literature, it can be concluded that for heterogeneous traffic modeling, the use of data-driven techniques is limited. The driving behavior of all the vehicle types is different due to differences in their size and shape and acceleration/deceleration capabilities. From this motivation, the current study attempts to predict the lateral movement decision of Cars, M2W, and M3W separately at different volume levels using machine learning.

3 Methodology

When a fast-moving follower approaches a slow-moving leader, it has two options either slow down its speed and follow the lead vehicle or look for the opportunity in the adjacent lanes and change lane. This phenomenon is known as Discretionary Lateral Movement. The scope of this study is limited to modeling the discretionary lateral movement decision of motorized passenger vehicles. Fig. 1 depicts a methodological framework to achieve the objectives of the present study.

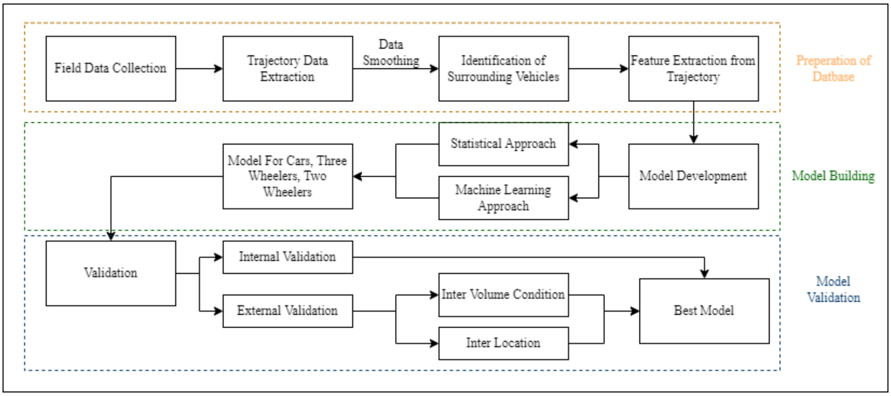


Fig. 1. Modeling Framework

4 Data Collection

The lateral movement decision of drivers gets affected by the presence of surrounding vehicles which can be studied with the help of trajectory data. In this research trajectory data developed by [9] on an urban midblock section of Western Expressway (Mumbai), India is considered as shown in Fig 2. The study section is 10-lane divided (5 lanes in each direction) having a width of 17.5 meters. The trap length of 120 meters was considered. From the video recordings collected, a total 40 min video was selected which had wide variation in traffic flow ranging from free flow to stop-and-go conditions, and the trajectories were extracted at a time resolution of 0.5 s using a semiautomated traffic data extractor tool [10]. The details of trajectory data are given in Table 1. With reference to the literature [9], smoothing techniques were applied to remove the noise in trajectory data. Based on a videographic survey broadly six types of vehicle categories were found in the selected roadway study section: Motorized three-wheelers, Motorized two-wheelers, Buses, Cars, Trucks, and Light commercial vehicles (LCV).

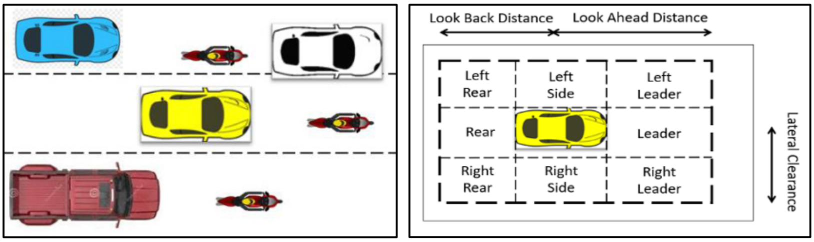
Table 1. Details of trajectory sets over the study section

Traffic Flow Level	Traffic Composition[a] (%)	Avg. Speed (kmph)	Avg. Flow (pcu/h)	Volume to capacity ratio	No. of vehicles tracked	Duration (min)
Flow 1	15/35/5/40/2/3	65	4800	0.35	1080	15
Flow 2	20/29/2/45/1/3	42	10120	0.71	1715	15
Flow 3	17/25/5/45/3/4	20	3500	<1	660	10

[a]Traffic composition: sequence 3W,2W, Bus, Cars, LCV, HCV

5 Surrounding Vehicle Identification

To study the lateral movement decision of vehicles at each time-stamp, it is required to identify the variables governing the decision. Thus, it is required to define the 'influence 'zone' of a subject vehicle. After defining the dimensions of the influence zone, the parameters of surrounding vehicles can be found. Further influence zone is divided into nine different compartments considering the subject vehicle at the centre of a three-lane road, as shown in Fig. 2. The look-ahead distance was estimated based on stopping sight distance (SSD) as given in equation 1. The look-back distance was estimated based on the Time to Collision (TTC) values as given in equation 2.



(a) (b)

Fig. 2. Influence zone to define surrounding vehicles

$$d_{front} = v_{vc} t_{vc}^{\alpha} \tag{1}$$

$$d_{back} = v_{stream} t_{\beta} \tag{2}$$

where d_{front} is the look-ahead distance, v_{vc} is the average longitudinal speed of a particular vehicle class and t_{vc}^{α} is the perception reaction time (PRT). Where d_{back} is the look back distance, v_{stream} is the average longitudinal stream speed and $t_{\beta}$is the threshold Time to Collision time (TTC). Table 2 shows the different look ahead and look back distances for all three flow levels by considering PRT of 2.5 seconds and TTC of 2 seconds [1].

Table 2. Look ahead and look back distance

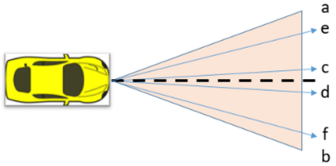
Flow Level	Average speed (m/s)			Average Stream Speed(m/s)	Look Ahead Distance (m)	Look back distance (m)
	Car	3W	2W			
Flow 1	16.83	13.88	15.47	14.61	40	30
Flow 2	11.25	9.28	12.62	8.95	30	20
Flow 3	4.667	3.463	6.322	4.35	15	10

A lateral distance of 5.5 m from the center position of the subject vehicle to the center position of the surrounding vehicles, including the total width of the subject vehicle (with an overlap of width), is considered over the entire road space (in longitudinal and lateral directions over time)

6 Defining Lateral Movement

When drivers wish to seek lateral movement opportunities for speed advantage, they can scan the peripheral area of vision. Considering the direction of motion along with traffic flow as a reference, this space is divided into N number of radial cones to define the choice of drivers at each time stamp, as shown in Fig 3. The choices were defined i.e., as straight, left, and right. Initial analysis was done to observe the angular deviation of the vehicles at a next time interval. The angle for the straight choice is considered close to human central vision and the threshold for this choice is -1^0 to 1^0 [11]. The threshold for the left and right choices according to vehicle classes is given in Table 3. If the deviation is between ‘c' and 'e’, then the vehicle is considered as shifting left and right for the deviation between 'd' and 'f'.'

Table 3. The threshold for angular deviation



	a	b	C	d	e	f
2W	-11	15	-1	1	-8	8
Car	-9	9	-1	1	-5	5
3W	-10	11	-1	1	-7	7

*All Numbers are in Degrees

Initially, the surrounding vehicles are identified using developed trajectory data and MATLAB code, considering the dimensions of the influence zone for different flow conditions. The parameters influencing the drivers' decision of lateral shift were identified. The spearman correlation test is performed between drivers' lateral movement decisions and the parameters influencing them. A threshold value of 0.70 was considered. Few variables were found highly correlated according to this threshold value; such variables were dropped as they were not providing any extra information about the dependent variables. Table 4 shows the final set of variables considered to predict the driver's decision of lateral movement.

Table 4. Set of variables considered

Surrounding Vehicle	Dummy variable for presence (0/1)	Vehicle Category	Relative speed with SV (m/s)	Longitudinal Gap (m)	Lateral Gap (m)
Leader	✓	✓	✓	✓	
Left Leader	✓	✓	✓	✓	✓
Right Leader	✓	✓	✓	✓	✓
Left Side	✓		✓		✓
Right Side	✓		✓		✓
Left Rear	✓	✓		✓	
Right Rear	✓	✓		✓	
SV	Lateral position(m), Longitudinal speed(m/s), No. of surrounding vehicles.				

7 Model Development

The drivers' decisions of lateral movements, i.e., to move straight, left, or right, were treated as a discrete choice for successive time intervals. For this, five models based on Machine Learning algorithms for classification were developed. The ML models include (i) Random Forest (RF) model, (ii) SVM model, (iii) Extreme Gradient Boosting (XGBoost), (iv) K-Nearest Neighbour (KNN) model, (v) Artificial Neural Networks (ANN) model.

8 Implementation and Results

The above-discussed models were developed separately for each flow level, and each considered vehicle class (Cars, M3W, M2W). To check the overall prediction accuracy of the models, internal validation was done by dividing 70% of the total data for the training of models and 30% of data for the testing of models. Further, to check the model's transferability and feasibility, external validation was done by

1. Testing of trained algorithms of one volume level to other volume levels
2. Testing of trained algorithms for the Mumbai dataset on the Chennai dataset.

The performance of all the algorithms was quantified with the help of the following indicators

True Positive (TP): The actual and predicted decisions are correct for all alternatives.
True Negative (TN): The actual decision and predicted decision are not in a specific alternative.
False Positive (FP): The actual decision is not in a specific alternative, but the predicted decision is in that specific alternative.
False Negative (FN): The actual decision is in a specific alternative, but the predicted decision is not in that specific alternative.

Accuracy: The ratio of correctly classified samples to the total number of samples, calculated as per Equation (3)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

Sensitivity (Recall): The ratio of correctly classified positive samples to the total number of actual positive samples of a particular alternative, calculated as per Equation (4).

$$Sensitivity = \frac{TP}{TP+FN} \tag{4}$$

Specificity: The ratio of correctly classified negative samples to the total actual number of negative samples, calculated as per Equation (5).

$$Specificity = \frac{TN}{TN+FP} \tag{5}$$

Positive Predictive Value (Precision): Ratio of correctly classified positive samples to the total number of positive predictions, calculated as per Equation (6).

$$Precision = \frac{TP}{TP+FP} \tag{6}$$

8.1 Validation of Trained Algorithms

Based on the confusion matrix obtained after developing the models, four matrices i.e., Accuracy, Sensitivity, Specificity, and Precision were compared to find the best model for a given classification problem. From the results based on four matrices, it was found that the models developed for medium volume levels showed better transferability in low and high-volume levels. Table 5 shows the performance matrices of RF model for different volume levels in the Mumbai dataset and the external validation on the Chennai dataset.

Table 5. Performance of models on test data

			Straight			Left			Right		
	Veh Type	Accuracy	Sensitivity	Specificity	Precision	Sensitivity	Specificity	Precision	Sensitivity	Specificity	Precision
Mumbai Medium Volume	Car	70.97	86.25	63.88	73.47	68.25	90.17	67.95	63.48	93.66	64.82
	M3W	72.11	90.24	71.98	72.11	72.58	84.89	72.78	68.169	95.84	70.08
	M2W	81.23	92.66	65.13	81.67	77.38	94.23	81.2	68.18	98.29	76.47
Mumbai High Volume	Car	76.12	81.05	71.39	71.93	71.59	75.12	73.02	60.74	82.49	74.63
	M3W	79.79	82.09	84.84	76.68	85.33	82.73	79.94	72.77	84.8	81.56
	M2W	78.01	88.73	80.91	79.01	79.74	72.18	76.94	65.71	83.89	80.7
Mumbai Low Volume	Car	90.56	94.88	45.45	93.65	51.25	88.93	72.5	53.57	96.29	60
	M3W	81.67	93.48	42.86	84.31	65.56	89.09	83.33	46.84	95.05	68.33
	M2W	79.58	97.22	26.67	79.91	59.63	99.53	88.88	54.24	98.06	66.67
Chennai	Car	70.36	84.2	75.36	71.64	63.16	76.74	66.36	61.24	91.36	65.79
	M3W	68.09	91.19	63.51	73.49	64.19	81.37	71.33	52.76	84.25	69.47
	M2W	74.33	85.25	70.15	78.47	66.49	79.37	74.12	61.24	84.36	70.13

Summary and Conclusions

A lateral movement decision behavior of Car, M3W, and M2W drivers in different volume levels is briefly investigated in this study. Five different machine learning algorithms for classification is trained to model the driver's lateral movement decision. A performance comparison of all these models were done and the best model to predict the decision was found. Results revealed that the RF classifier is best suited with respect to the rest of the considered ML models. The present study demonstrates the application of machine learning in the field of traffic and transportation engineering. From the results, it can be concluded that the lateral movement decision of drivers can be predicted using spatial information of surrounding vehicles. In real-time, this information can be gathered using advanced sensors, processed through these trained algorithms and provided as standard safety features in vehicles.

References

1. Amrutsamanvar, R.: Modeling lateral movement decisions of powered two wheelers in disordered heterogeneous traffic conditions. Transp. Lett. 14, 195–214 (2022). https://doi.org/10.1080/19427867.2020.1839718
2. Gipps, P.G.: A model for the structure of lane-changing decisions. Transp. Res. Part B. 20, 403–414 (1986). https://doi.org/10.1016/0191-2615(86)90012-3
3. Ahmed, K.I.: Modeling Drivers ' Acceleration and Lane Changing Behavior. Transportation (Amst). Ph.D, 189 (1999)
4. Budhkar, A.K., Maurya, A.K.: Characteristics of lateral vehicular interactions in heterogeneous traffic with weak lane discipline. J. Mod. Transp. 25, 74–89 (2017). https://doi.org/10.1007/s40534-017-0130-1
5. Kanagaraj, V., Asaithambi, G., Toledo, T., Lee, T.C.: Trajectory data and flow characteristics of mixed traffic. Transp. Res. Rec. 2491, 1–11 (2015). https://doi.org/10.3141/2491-01
6. Asaithambi, G., Joseph, J.: Modeling Duration of Lateral Shifts in Mixed Traffic Conditions. J. Transp. Eng. Part A Syst. 144, 04018055 (2018). https://doi.org/10.1061/jtepbs.0000170
7. Azam, M.S., Bhaskar, A., Haque, M.M.: Driving behaviour modelling in the context of heterogeneous traffic and poor lane discipline conditions: the state of the art and beyond. Transp. A Transp. Sci. (2021). https://doi.org/10.1080/23249935.2020.1870581
8. Kumar, P., Perrollaz, M., Lefevre, S., Laugier, C.: Learning-based approach for online lane change intention prediction. IEEE Intell. Veh. Symp. Proc. 797–802 (2013). https://doi.org/10.1109/IVS.2013.6629564
9. Raju, N., Kumar, P., Chepuri, A., Arkatkar, S.S., Joshi, G.J.: Calibration of Vehicle Following Models Using Trajectory Data Under Heterogeneous Traffic Conditions. 96th Tranportation Annu. Meet. (2017)
10. Munigety, C.R., Vicraman, V., Mathew, T. V.: Semiautomated tool for extraction of microlevel traffic data from videographic survey. Transp. Res. Rec. 2443, 88–95 (2014). https://doi.org/10.3141/2443-10
11. Sarkar, N.C., Bhaskar, A., Zheng, Z., Miska, M.P.: Microscopic modelling of area-based heterogeneous traffic flow: Area selection and vehicle movement. Transp. Res. Part C Emerg. Technol. 111, 373–396 (2020). https://doi.org/10.1016/j.trc.2019.12.013