DEVELOPING PEDESTRIAN FATALITY PREDICTION MODELS USING HISTORICAL CRASH DATA: APPLICATION OF BINARY LOGISTIC REGRESSION AND BOOSTED TREE MECHANISM

Malaya Mohanty[1,*], Bitati Sarkar[2], Piotr Gorzelańczyk[3], Rachita Panda[1], Srinivasa Rao Gandupalli[4], Ampereza Ankunda[1]

[1]School of Civil Engineering, KIIT Deemed to be University, Bhubaneswar, India
[2]Department of Civil Engineering, National Institute of Technology, Silchar, India
[3]Department of Engineering, Stanislaw Staszic University of Applied Sciences, Pila, Poland
[4]GITAM School of Technology, GITAM Deemed to be University, Visakhapatnam, India

*E-mail of corresponding author: malaya.mohantyfce@kiit.ac.in

Malaya Mohanty 0000-0002-6116-782X, Piotr Gorzelańczyk 0000-0001-9662-400X, Srinivasa Rao Gandupalli 0000-0002-4937-2652

Resume

Pedestrian fatality rate plays a key role in examining effectiveness of the road safety. The present study attempts to examine the effect of various categories of accused vehicles and the average 85th percentile speed at accident location on the pedestrian crash fatality. The study also attempts to develop pedestrian crash severity models using the binary logistic regression and boosted trees technique. Historical crash data, along with the video recording technique at accident sites, have been utilized for the present study. From regression equations, it is observed that when the heavy vehicle (HV) hits a pedestrian as compared to two-wheeler (2W), the average chance of death increases 2.44 times. According to the Boosted tree model, the contribution of speed is 60%, whereas the contribution of category of accused vehicle is 40% for pedestrian fatality prediction. The study should help in planning better strategies like all red time at intersections or pedestrian foot over bridge at critical locations.

Available online: https://doi.org/10.26552/com.C.2023.036

Article info

Received 25 November 2022
Accepted 28 February 2023
Online 21 March 2023

Keywords:

Pedestrian fatalities
vehicle category
85th percentile speed
binary logistic model
boosted tree model

ISSN 1335-4205 (print version)
ISSN 2585-7878 (online version)

1 Introduction

Pedestrians form an integral part of every roadway environment. The conflict between pedestrians and vehicles is a vital factor to examine the efficiency of the roadway traffic [1]. World Health Organisation (WHO) [2] has reported that more than 1.3 million deaths occur every year around the world, due to the road crashes. Among them, 54% of the global road accidental deaths comprise of pedestrians, cyclists, and motor cyclists with pedestrians outnumbering the other two categories [3]. Further, 93% of fatalities due to road crashes occur in low-and middle-income countries even though they comprise only 60% of the total number of vehicles in the world [2]. In a developing country like India, more than 150,000 people die in road crashes every year [3]. Although the number of persons killed in road crashes have slightly declined in 2019 (-0.2% as compared to 2018), however, the share of pedestrian deaths has shown increasing trend. Despite the pedestrians being involved in 14% of total road accidents in the country, their share in total deaths is alarming at 17% [3], suggesting that the fatality rate in road accidents involving pedestrians is very high. The cases of hit and run has also increased over the years [3], which concludes pedestrians being the worst sufferers in the road accidents. All these numbers press for an imperative need to study and examine the pedestrian safety on roads.

Many studies have been conducted over the years to assess pedestrian safety, as well as the overall safety on roads. While earlier studies [1, 4-5] focused mainly on studying the pedestrian behaviour from recorded videos with relation to the accepted gaps and vehicles dynamics, recent studies [6-10] are focused on development of mathematical models like MLR, Logit and Probit models, etc. to predict different characteristics of both the vehicle and a pedestrian during the conflict situation. Most of these models and simulations-based analysis are based on the data collected from the field through recorded videos, historical crash data, or questionnaire-based surveys. Few other researchers [7, 11] have also examined the pedestrian safety based on various surrogate safety measures (SSM), like Post encroachment time (PET), Time to vehicle (TTV), Deceleration time (DT), Time to accidents (TTA), etc. Similarly, few researchers [12-13] have also utilized decision trees, like CART, Random Forest, Boosted Trees, or ANN, to model the safety of vehicles and pedestrians on road. Hu and Cai (2022) [14] reported that the logistic model is first applied for relationship explanation, and then the machine learning classifiers like ANN or Fuzzy are applied for prediction modeling. Tamakloe et al. (2022) [15] opined that logistic regression helps in identifying the individual effect of factors affecting crash severity. According to Pande et al. (2010) [16], various decision trees, like CART and boosted trees, make it easier to interpret the results unlike the ANN. The CART and boosted trees also remove the problems of multicollinearity.

Although the usage of SSMs have grown popularity in last few years due to its ability to predict crashes without occurrence of the incident [17-18], still the analysis of historical reported road crashes is common due to its proof of occurrence. Moreover, with historical data, varied geometrical and driver behavior along with vehicle flow data can be considered. Safety assessment of road networks and heterogenous traffic conditions of developing countries like India have historically relied on police-reported crash data [19-20]. According to [19], the fatal crashes should be assessed in detail for developing countries like India, to understand what factors contributes to fatal crashes. Past studies reveal that the role of accused vehicles/perpetrators have not been probed much to inspect their effects on pedestrian deaths. Few studies have undertaken the effect of motorcycles and heavy vehicles [20-21] on all the road crashes. However, their effect on pedestrian deaths was not in focus much.

In the present study, historical crash data have been utilized to assess the role of accused vehicles on pedestrian deaths. Along with the historical crash data, field studies are conducted at accident locations to gather speed data, which is also used as a factor to analyze the pedestrian fatalities. Thereafter, pedestrian severity prediction model has been developed by using logistic regression and boosted tree. The developed severity prediction models can help to understand the pedestrian fatalities better and help in framing better strategies for reducing the death rate during their crashes.

2 Data collection and methodology

Data collection is a significant aspect for attaining the objectives of the present study since better prediction models can be developed with a greater number of data points. In order to pursue the study, historical road crash data has been collected from Visakhapatnam, an emerging smart city in India. A total of 2425 road crash data have been analyzed, out of which 789 crashes involved pedestrians. All these crashes have occurred in 2019. The data were extracted on spreadsheets and analyzed using the MINITAB and JMP SAS software for correlations and development of models. Since the study deals with examining the role of accused vehicle, all the road crash incidents where victims were pedestrian, the details regarding the accused vehicle category and the outcome of the crash (injury or death), are noted. Further, field studies using video recording technique have been conducted at the accident locations to collect the 85th percentile speed for the road stretch where the accident occurred. It was observed that there was no major difference between the 85th percentile speeds collected on different days at those locations for various time periods. In addition, since this research has been ongoing for a longer time period of more than 2 years, it was seen that the 85th percentile speed also remained statistically similar in different years, too. Therefore, the 85th percentile speed data used to understand the effect of operating speed on pedestrian fatality is not much different to the 85th percentile speed on the day of the road accident at that location. Figure 1 displays the camera setup to collect speed data at various locations. In order to collect the speed data, 2 lines are marked 20 m apart at certain road sections near busy streets or accident blackspots. A camera is placed at the locations strategically so that it can cover a view of 20 meters. The data are collected during both the peak and off-peak periods. Next, the recorded videos are played on monitor to extract time data of various vehicles while crossing these 20 m interval lines, which in turn is used for extraction of speed data. Further, based on the time during which the road crash has occurred, the 85th percentile speed during that time interval is extracted and used in the study.

All the above-mentioned factors have been considered and correlated with pedestrian fatality to identify causative factors for pedestrian deaths and develop pedestrian fatality prediction models.

3 Results and discussions

The road crash data used for the study had 2425 reported crashes, which included all kinds of vehicles. It was observed that out of 2425 road crashes, pedestrians were the victim in 789 instances (32.5 %). Similarly, although 550 of the total 2425 crashes were fatal (22.7 %), the fatality percentage when pedestrians are victim is 35.1% (277 out of 789 crashes). The preliminary data itself shows two major points. Firstly, how rampant is the pedestrian hit road crashes when compared to other vehicles on the road, and secondly, how the fatality rate escalates when pedestrians are hit on roads. The section has been divided into two major parts:

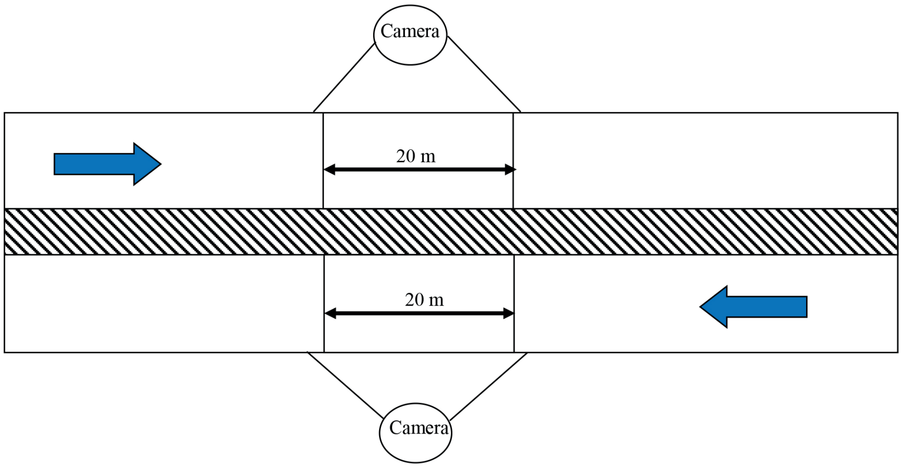

Figure 1 *Camera setup for the field data collection*

Table 1 *Pedestrian fatality rate based on various category of perpetrators*

Accused vehicle/Perpetrator	Number of crashes	Number of fatal crashes	Proportion of fatal crashes
2W	326	81	24.85
3W	51	18	35.29
4W	234	94	40.17
HV	178	84	47.19
Total number of crashes involving pedestrians	789	277	35.11

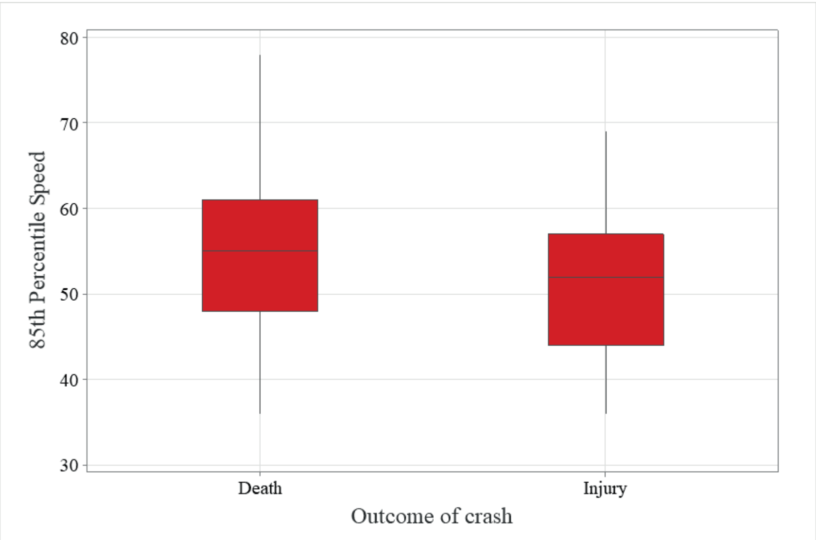

Figure 2 *Boxplot for assessing the impact of 85th percentile speed on pedestrian fatalities*

1. Relationship of accused vehicle category and speed with pedestrian fatality and
2. Development of pedestrian severity prediction models.

3.1 Relationship of accused vehicle category and speed with pedestrian fatality

The major category of vehicles, which hit pedestrians on roads are the Two-wheelers (2W), Three-wheelers (3W), cars, SUVs, jeeps, etc. (4W), and heavy vehicles (HV). Table 1 represents the number of accidents with each accused vehicle category and the percentage of fatal crashes with each category as perpetrator.

Results from Table 1 clearly show the effect of each category as perpetrator, while hitting a pedestrian. The effect of HV and 4W are particularly concerning, since the fatality rate increases to 40 and 47 %, respectively, as compared to the total average fatality rate of pedestrians (35.11 %), which itself is a high severity rate. In addition to the category of perpetrators, the operating speed (85th percentile speed at accident location is considered) has also been analyzed to study its effect on pedestrian

Table 2 *Odd's ratio for independent predictors*

Level A	Level B	Odds Ratio	95% CI
Accused category			
3W	2W	1.411	(0.7376, 2.6993)
4W	2W	1.737	(1.1892, 2.5360)
HV	2W	2.441	(1.6320, 3.6500)
4W	3W	1.231	(0.6401, 2.3665)
HV	3W	1.730	(0.8863, 3.3759)
HV	4W	1.405	(0.9308, 2.1221)
85^{th} percentile speed		1.073	(1.053 to 1.093

NB - Odds ratio for level A relative to level B

fatality rate. A box plot (as shown in Figure 2) is plotted to understand the effect of 85^{th} percentile speed on pedestrian fatality rate. Figure 2 clearly presents that with increase in 85^{th} percentile speed, the pedestrian fatality rate also increases. The average 85^{th} percentile speed when road accidents, involving pedestrians lead to fatalities, is observed to be 56.4 km/h, whereas the same average 85^{th} percentile speed reduces to 50.9 km/h when only injuries are sustained by pedestrians. 2 sample t-test is also conducted, which shows that the mean 85^{th} percentile speeds resulting in pedestrian injury and fatalities are significantly different from each other ($p < 0.05$) at 5% significance level. Next, pedestrian severity prediction modelling is conducted by two techniques - (a) Binary logistic model, and (b) Boosted tree model.

3.2 Development of pedestrian severity prediction models

In the present study, the preliminary analyses revealed clear relationship to be existing among pedestrian fatality, category of accused vehicle, and 85^{th} percentile speed. Two techniques have been utilized - (a) Binary logistic, and (b) Boosted tree to developed predictive models.

3.3 Binary logistic model

Being a nominal variable with death/fatality and injury as two outputs, at first a binary logistic regression is carried out to predict the severity of pedestrian hit crashes. The general equation of logistic regression, to calculate the probability of fatality/deaths of pedestrians, is provided in Equation (1). P(fatality) refers to probability of pedestrian fatality in the case of a crash event. Equations (2) to (5) present the Y' functions for each category of accused vehicle.

$$P(fatality) = \left[\exp(Y')/(1+\exp(Y'))\right], \tag{1}$$

$$Y'_{2W} = -4.779 + 0.0702 * 85th\ per.speed\ , \tag{2}$$

$$Y'_{3W} = -4.435 + 0.0702 * 85th\ per.speed\ , \tag{3}$$

$$Y'_{4W} = -4.227 + 0.0702 * 85th\ per.speed\ , \tag{4}$$

$$Y'_{HW} = -3.887 + 0.0702 * 85th\ per.speed\ . \tag{5}$$

Equations clearly depict that the HV shall give higher probability of pedestrian fatalities, as compared to other categories since the constant term is lowest here with a negative sign. Hence, the least number will be subtracted from the other term (85^{th} per. speed*0.0702) in the equation. That term is same for all equations. Therefore, the Y' for HV shall be highest and therefore the probability of fatality for same 85^{th} percentile speed. The proposed model predicts the outcome of 67.2% of the pedestrian crashes correctly with correct prediction for 88% of injuries and only 35% of deaths. However, the accuracy of predicting fatalities can be increased by decreasing the cut-off value for deciding fatality/injuries. Default settings suggest that if the pedestrian fatality is predicted to be more than 0.5, then the regression predicts the outcome as deaths, else injuries. However, the actual fatality rate of pedestrian as per the recorded crash data used for the study is 35.1%. This suggests that if the predicted probability for death is more than 35.1%, it is more likely to result in death of a pedestrian, since it is more than total average. If the cut-off value is decreased from 0.5 to 0.35, the prediction of accuracy for pedestrian deaths increases to 60% and for injuries, decrease to 65%. The equations (Equations (2) to (5)) reveal the higher effect of HV, which was eminent from the crash data too. However, more precise comparison can be made from the odd's ratio obtained from the logistic regression. Table 2 provides the odd's ratio table obtained for the proposed model.

The odd's ratio suggests the increase or decrease of fatality rate for each accused vehicle as compared to other vehicles, considering the 85^{th} percentile speeds. For instance, it is observed that the HV when hits a pedestrian, as compared to the 2W, the average chance of death increases 2.44 times ranges from 1.63 times to 3.65 times. Similarly, when 4W hits a pedestrian the average probability of fatality increases 1.737 (1.189 to 2.536) times as compared to when a 2W hits a pedestrian. Similarly, 1 km/h increase in 85th percentile speed increases the probability of pedestrian deaths by 1.073 times. Although the results of logistic model provide reasonable estimates of pedestrian crash outcomes, the predictions have not been very accurate. The odd's ratio, on the other hand provides important results regarding pedestrian fatalities, which is useful for development and execution of proper strategies to reduce the pedestrian accident-related deaths. The composition of traffic along with the 85th percentile speeds, at any vulnerable location, can be utilized to assess the threat to pedestrians. Next, decision tree namely, the boosted tree, was used for developing the prediction models due to its higher level of accuracy in predicting the outcome [16, 22].

3.4 Boosted tree model

Boosted tree is a type of decision tree, which uses boosting algorithm at every step to improve the accuracy of prediction of a required outcome. The boosted tree model is a more modern, versatile, and powerful predictive tool for developing the prediction models. In the present study, the model is developed using JMP SAS. The model predicts 95% of injuries and 71% of

Layer	Split	Label	Estimate
1	1	85th percentile speed<69	0.0230073206
1	2	85th percentile speed<70	-0.141280978
1	3	85th percentile speed<76	-0.789928889
1	4	85th percentile speed>=76	-0.789928889
2	1	85th percentile speed<69	0.0204074934
2	2	85th percentile speed<70	-0.125316242
2	3	85th percentile speed<72	-0.679900345
2	4	85th percentile speed>=72	-0.680794883
3	1	Accused Vehicle(3W, 4W, HV)	0.0059343076
3	2	85th percentile speed<44	0.2696723715
3	3	85th percentile speed>=44	0.1014870345
3	4	85th percentile speed>=53	-0.0327151
4	1	85th percentile speed<69	0.0160494116
4	2	85th percentile speed<70	-0.1074587
4	3	85th percentile speed<72	-0.625550795
4	4	85th percentile speed>=72	-0.626690869
5	1	Accused Vehicle(3W, 4W, HV)	0.0034501505
5	2	85th percentile speed<43	0.6235797646
5	3	85th percentile speed>=43	0.0944578748
5	4	85th percentile speed>=53	-0.025769774
6	1	85th percentile speed<69	0.0118506542
6	2	Accused Vehicle(2W)	-0.033801144
6	3	85th percentile speed<70	-0.177177805
6	4	85th percentile speed>=70	-0.59805459
7	1	Accused Vehicle(3W, 4W, HV)	-0.019553723
7	2	85th percentile speed<44	0.1976984364
7	3	85th percentile speed>=44	0.0761925523
7	4	85th percentile speed>=53	0.0012379576

Figure 3 *Sample tree details at each layer*

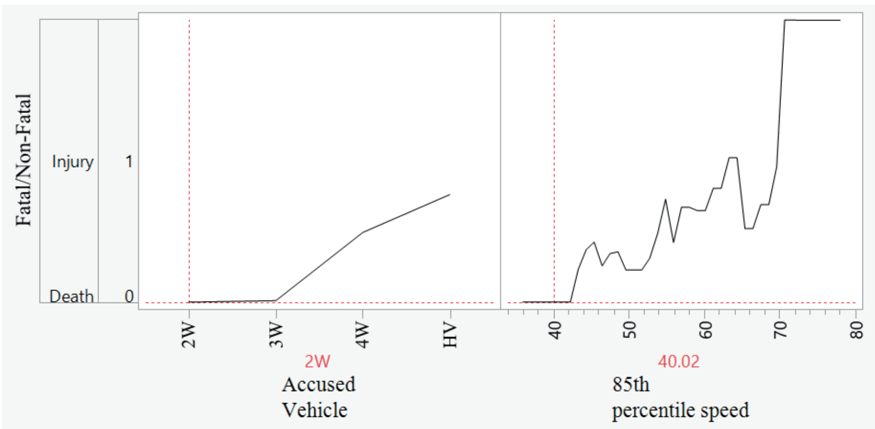

Figure 4 *Prediction profiler for boosted tree method - prediction profiler when the speed is 40 km/h*

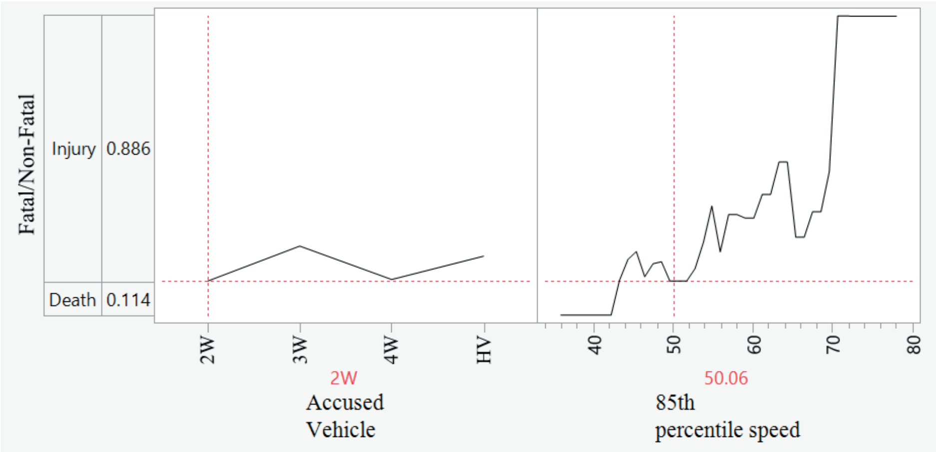

Figure 5 *Prediction profiler for boosted tree method - prediction profiler when the speed is 50 km/h*

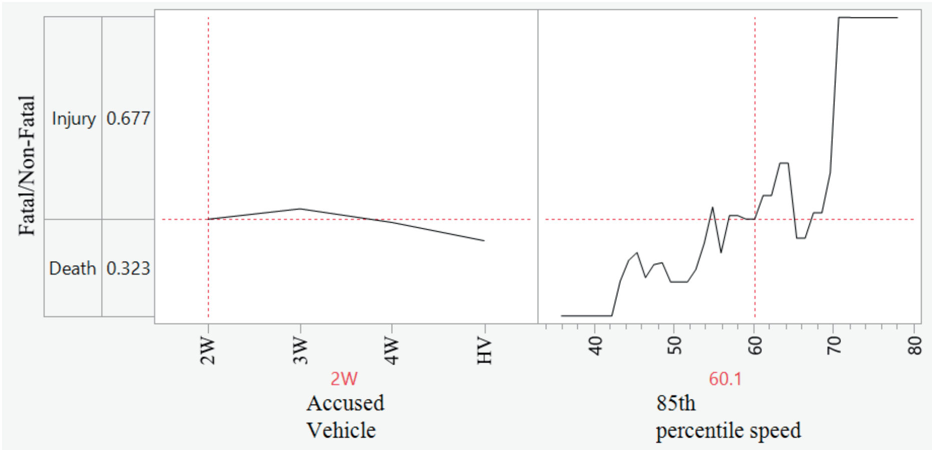

Figure 6 *Prediction profiler for boosted tree method - prediction profiler when the speed is 60 km/h*

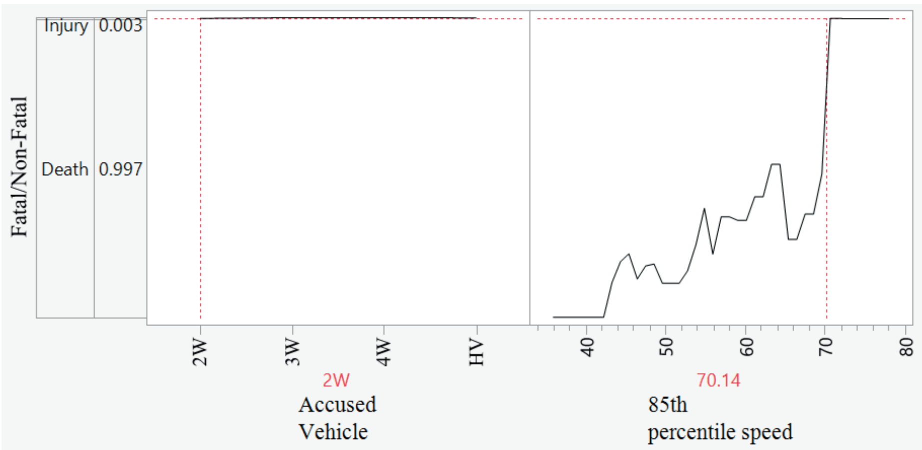

Figure 7 *Prediction profiler for boosted tree method - prediction profiler when the speed is 70 km/h*

deaths correctly. As mentioned earlier, 789 pedestrian involved incidents have been used for study, which implies that the model uses 789 layers of trees to develop the prediction model. At each layer, 3 splits have been modelled, based on data and the developed algorithm changes the estimated percentage of pedestrian fatality and makes it better. The algorithm that works at the backend of the boosted tree is highly dependent on number of data points. With increase in data points, the accuracy of prediction will increase, since that many number of trees shall be used to develop the prediction model. The sample labels of algorithm at each layer and resulting changes in predicted probabilities (tree details) are provided in Figure 3. As can be seen from Figure 3, for each layer, 3 splits or 3 conditions are being developed. For instance, in the first layer, first split describes that when 85th percentile speed decreases below 69 km/h, the probability of fatality decreases by 0.02 times. The second split denotes that when the speed is less than 70 km/h, the probability of pedestrian fatality reduces by 0.14 times. The third split describes when the speed is less than 76 km/h, the death probability decreases by 0.79 times. The reference level for the algorithm is considered as 100% probability of pedestrian deaths. Figures 4-7 present the prediction profilers for the two independent variables and their effects on pedestrian fatality. Since the plots dynamically changes with change in accused category and 85th percentile speed, therefore, 4 cases are presented in Figures 4-7, viz. the probability of deaths for all category of accused vehicles when speed increases by 10 km/h (40, 50, 60, and 70 km/h). As can be seen, when 4W and HV are hitting a pedestrian while travelling at 40 km/h, the probability of death is around 0.4-0.5. However, of a 2W or 3W hits a pedestrian while travelling at 40 km/h, the probability of death is much smaller. With increase in speed, the probability increases for each category of vehicles. At 70 km/h (Figure 7), the probability of pedestrian fatality is 0.997, irrespective of category of vehicles.

Boosted tree's prediction accuracy is high and can be increased if more data is fed into the algorithm. Prediction profiler is handy to understand the microscopic effect of each variable on pedestrian fatality. In achieving these higher levels of prediction accuracy, speed has 60% contribution, while category of accused vehicle has 40% contribution.

4 Conclusions and recommendations

The present study focuses on examining the role of accused vehicle category and 85th percentile speed of vehicles on pedestrian fatalities. To attain the objective, historical crash data along with the field study (for collecting 85th percentile speed data) have been used. Out of 2425 reported crashes, 789 incidents involved pedestrians. The average fatality rate for pedestrians is found to be 35.1%, which is much more than when other vehicles are hit (22.7 %). The category of perpetrators was classified into the 2W, 3W, 4W and HV. It was found that proportion of pedestrian fatal crashes increases when the HV is the accused vehicle. Along with category, 85th percentile speed has also been utilized in the study. It is observed that the average 85th percentile speed is 56.4 km/h and 50.9 km/h for pedestrian fatalities and injuries respectively. Then, the 2 sample t-test also shows that the mean 85th percentile speeds resulting in pedestrian injury and fatalities are significantly different from each other ($p < 0.05$) at 5% significance level.

After identification of clear relationship to be existing among pedestrian fatality, category of accused vehicle, and 85th percentile speed, pedestrian crash severity model is developed using binary logistic regression and boosted tree model. The binary logistic regression clearly depicts the higher effect of the HV of pedestrian fatalities, as compared to other categories. Besides the category, the specific regression equations, odd's ratio, helps in understanding as how each category impacts pedestrian fatalities as compared to other categories. For instance, from the odd's ratio table is observed that chance of pedestrian crashes increases 2.44 times when the HV hits a pedestrian as compared to the 2W. Similarly, when the 4W hits a pedestrian as compared to the 3W, the probability of death increases 1.23 times. Similarly, 1 km/h increase in the 85th percentile speed, increases the probability of pedestrian deaths by 1.073 times. Next boosted tree model is developed, whose accuracy of predicting injury correctly is 95% and deaths is 71%. The boosted tree employs 3 splits, based on data of each layer, to improve the accuracy denoting that more the data points, the better shall be the accuracy of the developed model. Dynamic prediction profilers, utilizing the change in accused category and the 85th percentile speed, are very helpful to understand the pedestrian fatality rates microscopically. For instance, the probability of pedestrian fatality is almost 0 for the 2W and 3W, when the speed is 40 km/h with around 0.4 for the 4W and HV. At 70 km/h the probability of pedestrian fatality is 0.997, irrespective of category of vehicles. Boosted tree's model is developed with 60% contribution of speed and 40% contribution of category of the accused vehicle. The model identifies that for a speed of around 40 km/h to 50 km/h, the pedestrian crash with the 2W and 3W have very less chance of fatality; whereas the higher speeds increase the probability of fatalities exponentially.

The predicted models can be used to predict the road crashes with great accuracy. The results of the present study should be useful for analyzing the pedestrian safety on any road. Utilizing the composition of a vehicle and the average 85th percentile speed, at any traffic facility, the probability of pedestrian fatality can be obtained using the boosted trees model. Based on the obtained pedestrian fatality probabilities, different strategies, like all red time at intersections or pedestrian foot over bridge at mid-block sections, can be planned and executed to reduce the pedestrian crashes and deaths.

Acknowledgement

The authors would like to thank the Traffic Police Authorities, Visakhapatnam City, Andhra Pradesh, India for providing the dataset used in this study. All opinions and results are those of the authors.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] GUANGXIN, L., KEPING, L. Pedestrian's psychology on gap selecting when crossing street. In: 2nd International Conference on Transportation Engineering ICTE 2009: proceedings [online]. ASCE. 2009. ISBN 9780784410394, p. 2138-2144. Available from: https://doi.org/10.1061/41039(345)354

[2] Road traffic injuries - WHO [online]. 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries/

[3] Road accidents in India - MORTH [online]. 2019. Available from: https://morth.nic.in/sites/default/files/RA_Uploading.pdf

[4] AL-AZZAWI, M., RAESIDE, R. Modeling pedestrian walking speeds on sidewalks. *Journal of Urban Planning and Development* [online]. 2007, **133**(3), p. 211-219. ISSN 0733-9488, eISSN 1943-5444. Available from: https://doi.org/10.1061/(ASCE)0733-9488(2007)133:3(211)

[5] RASTOGI, R., THANIARASU, I., CHANDRA, S. Design implications of walking speed for pedestrian facilities. *Journal of Transportation Engineering* [online]. 2011, **137**(10), p. 687-696. ISSN 2473-2907, eISSN 2473-2907. Available from: https://doi.org/10.1061/(ASCE)TE.1943-5436.0000251

[6] MUKHERJEE, D., MITRA, S. A comparative study of safe and unsafe signalized intersections from the viewpoint of pedestrian behavior and perception. *Accident Analysis and Prevention* [online]. 2019, **132**, 105218. ISSN 0001-4575, eISSN 1879-2057. Available from: https://doi.org/10.1016/j.aap.2019.06.010

[7] CHAUDHARI, A., GORE, N., ARKATKAR, S., JOSHI, G., PULUGURTHA, S. Exploring pedestrian surrogate safety measures by road geometry at midblock crosswalks: a perspective under mixed traffic conditions. *IATSS Research* [online]. 2021, **45**(1), p. 87-101. ISSN 0386-1112, eISSN 0386-1112. Available from: https://doi.org/10.1016/j.iatssr.2020.06.001

[8] KIM, D. The transportation safety of elderly pedestrians: Modeling contributing factors to elderly pedestrian collisions. *Accident Analysis and Prevention* [online]. 2019, **131**, p. 268-274. ISSN 0001-4575, eISSN 1879-2057. Available from: https://doi.org/10.1016/j.aap.2019.07.009

[9] MOHANTY, M., PANDA, R., GANDUPALLI, S. R., ARYA, R. R., LENKA, S. K. Factors propelling fatalities during road crashes: a detailed investigation and modelling of historical crash data with field studies. Heliyon [online]. 2022, 8(11), e11531. eISSN 2405-8440. Available from: https://doi.org/10.1016/j.heliyon.2022.e11531

[10] MOHANTY, M., PANDA, R., GANDUPALLI, S. R., SONOWAL, D., MUSKAN, M., CHAKRABORTY, R., DANGETI, M. R. Development of crash prediction models by assessing the role of perpetrators and victims: a comparison of ANN and logistic model using historical crash data. *International Journal of Injury Control and Safety Promotion* [online]. 2022, latest articles, p. 1-17, Available from: https://doi.org/10.1080/17457300.2022.2089899

[11] KUMAR, K. P., KUMAR, M. U., SEKHAR, V. C., KRISHNA, P. R. Analysis of pedestrian deaths in road traffic accidents-an autopsy based study in Visakhapatnam. *East African Scholars Journal of Medical Sciences* [online]. 2019, **2**(2), p. 74-79. ISSN 2617-4421, eISSN 2617-7188. Available from: https://doi.org/10.36349/easms.2019.v02i02.006

[12] LEE, C., LI, X. Predicting driver injury severity in single-vehicle and two-vehicle crashes with boosted regression trees. *Transportation Research Record* [online]. 2015, **2514**(1), p. 138-148. ISSN 0361-1981, eISSN 2169-4052. Available from: https://doi.org/10.3141/2514-15

[13] ABDEL-ATY, M., ABDELWAHAB, H. Analysis and prediction of traffic fatalities resulting from angle collisions including the effect of vehicles' configuration and compatibility. *Accident Analysis and Prevention* [online]. 2004, **36**(3), p. 457-469. ISSN 0001-4575, eISSN 1879-2057. Available from: https://doi.org/10.1016/S0001-4575(03)00041-1

[14] HU, M., CAI, W. Analysis of lifeboat embarkation efficiency for cruise passengers under multiple scenarios. *Scientific Programming* [online]. 2022, **2022**, 5491621. ISSN 1058-9244, eISSN 1875-919X. Available from: https://doi.org/10.1155/2022/5491621

[15] TAMAKLOE, R., DAS, S., AIDOO, E. N., PARK, D. Factors affecting motorcycle crash casualty severity at signalized and non-signalized intersections in Ghana: insights from a data mining and binary logit regression approach. *Accident Analysis and Prevention* [online]. 2022, **165**, 106517. ISSN 0001-4575, eISSN 1879-2057. Available from: https://doi.org/10.1016/j.aap.2021.106517

[16] PANDE, A., ABDEL-ATY, M., DAS, A. A classification tree-based modeling approach for segment related crashes on multilane highways. *Journal of Safety Research* [online]. 2010, **41**(5), p. 391-397. ISSN 0022-4375, eISSN 1879-1247. Available from: https://doi.org/10.1016/j.jsr.2010.06.004

[17] MOHANTY, M., PANDA, B., DEY, P. P. Quantification of surrogate safety measure to predict severity of road crashes at median openings. *IATSS Research* [online]. 2021, **45**(1), p. 153-159. ISSN 0386-1112, eISSN 0386-1112. Available from: https://doi.org/10.1016/j.iatssr.2020.07.003

[18] WANG, J., MA, J., LIN, P., SARVI, M., LI, R. Pedestrian single file movement on stairway: Investigating the impact of stair configuration on pedestrian ascent and descent fundamental diagram. *Safety Science* [online]. 2021, **143**, 105409. ISSN 0925-7535, eISSN 1879-1042. Available from: https://doi.org/10.1016/j.ssci.2021.105409

[19] VEDAGIRI, P., PRAGNA, T. Evaluation of traffic safety at unsignalized intersection under mixed traffic conditions. In: Transportation Research Board 92nd Annual Meeting: proceedings. 2013. 13-4000.

[20] PUTHAN, P., LUBBE, N., SHAIKH, J., SUI, B., DAVIDSSON, J. Defining crash configurations for powered two-wheelers: comparing ISO 13232 to recent in-depth crash data from Germany, India and China. *Accident Analysis and Prevention* [online]. 2021, **151**, 105957. ISSN 0001-4575, eISSN 1879-2057. Available from: https://doi.org/10.1016/j.aap.2020.105957

[21] ANDERSON, J., HERNANDEZ, S. Heavy-vehicle crash rate analysis: comparison of heterogeneity methods using Idaho crash data. *Transportation Research Record* [online]. 2017, **2637**(1), p. 56-66. ISSN 0361-1981, eISSN 2169-4052. Available from: https://doi.org/10.3141/2637-07

[22] PAN, Y., CHEN, S., QIAO, F., UKKUSURI, S. V., TANG, K. Estimation of real-driving emissions for buses fueled with liquefied natural gas based on gradient boosted regression trees. *Science of the Total Environment* [online]. 2019, **660**, p. 741-750. ISSN 0048-9697, eISSN 1879-1026. Available from: https://doi.org/10.1016/j.scitotenv.2019.01.054