Proceedings of 10th IOE Graduate Conference
Peer Reviewed
ISSN: 2350-8914 (Online), 2350-8906 (Print)
Year: 2021 Month: October Volume: 10

Study on Drivers' Yielding Behavior at Pedestrian Crossings

Sumona Bajracharya [a], Anil Marsani [b]

[a, b] *Department of Civil Engineering, Pulchowk Campus, IOE, Tribhuvan University, Nepal*
Corresponding Email: [a] SUMONA.752560@puc.tu.edu.np, [a] Kathmandu.bajra@gmail.com, [b] anilmarsani@ioe.edu.np

Abstract
Driver yielding at pedestrian crossings is essential for the safety of all road users. While driver yielding behavior may depend upon various road and road-user characteristics, this study investigates driver yielding behavior for twelve variables. Data is extracted from six study locations and analyzed by using binary logistic regression. The final model, Model III indicates that driver yielding depends upon ten significant parameters: no. of pedestrians crossing the road at the same time (N), pedestrian age category, pedestrian gender, pedestrian speed, driver age category, driver gender, vehicle type, vehicle speed, presence of a median strip and zebra-crossing marking condition. The study did not find any increase in yielding for older age groups (more than 60 years) or for females carrying an infant.

Keywords
vulnerable road users, zebra-crossing, road safety

1. Introduction

1.1 Background

Pedestrian and driver interaction is one of the major sources of road crashes. This relationship between the two groups of road users is most prominent at unsignalized pedestrian crossings where drivers are required to give the right of way to pedestrians as per traffic rules although the reality is questionable.

A pedestrian crossing is a specified part of a road in which pedestrians have the right of way to cross. Pedestrian crossings can generally be categorized into two types i.e., signalized and unsignalized crossings. Different types of markings and installation criteria are provided in order to distinguish the crossing area for increased safety of road users. Nepal Road Standard 2070 states a minimum of 2.5m width should be provided as walkway for pedestrian separation. [1]

According to Kumar (2010), road traffic injuries were the most common type of injuries in Nepal, most of which (42.5%) involved motorcycles and pedestrians were the most vulnerable group.[2]

Anup Ojha (2019) in his article states that Kathmandu has 107 road crossings. The article also mentions about a study conducted in 2018 which showed that 60% of the road markings in the capital had faded away and 80% of the roads without any zebra crossings at all. [3]

Drivers are unwilling to give way to pedestrians on road crossings. The drivers travelling at high speed have difficulty in sufficiently reducing their speed when reaching the road crossings. Driver behavior may depend on many factors like road congestion level, time of the day, vehicle characteristics, road condition, pedestrian behavior etc. For this purpose, the study is carried out in order to evaluate the impact of different variables that influence drivers' yielding at unsignalized pedestrian crossings.

1.2 Research Objective

The primary objective of the study is to determine the extent of road and road user characteristics which encourage yielding behavior in drivers. This can be summarized into two points given below:

- To determine the most significant parameters that influence driver yielding behavior.
- To develop a model that could establish the relationship between driver yielding behavior with significant parameters.

2. Literature Review

Andra's Va'rhelyi (1998) used Space Mean Speed and Time to Zebra (TTZ) as indicators for driver's willingness to yield in road crossings and found that out of the 790 cases with pedestrian presence, the pedestrian passes first in just 42 of the cases. It found three out of four drivers maintaining the same speed or accelerating and only one out of four slowing down or braking during an encounter. [4]

Fuller (1984) developed a 'Threat Avoidance Model' which implies that when confronted with a discriminative stimulus for a possible aversive event, what a driver does depends specifically on the rewards and punishments for alternative responses. Two possible responses:

- The driver considers the pedestrian to be a threat and slows down allowing the pedestrian to pass first so the driver is punished with a loss of time.
- The driver maintains the same speed because he considers the pedestrian to be a threat but chooses a "non- avoidance response", signalling to the pedestrian that he has no intention to yield. [5]

Francesco Bella, Manuel Silvestri (2015) made the following driver's speed behavior observations due to safety measures at pedestrian crossings :

- The lowest value of non-yielding behaviour was reached when curb extensions were present, likely due to improvement in pedestrian visibility.
- The highest value of non-yielding behaviour was found with parking restrictions near the crossing. [6]

Anciaes et al. (2010) found that all four design elements of courtesy crossings considered (stripes, colored or textured surfacing, visual narrowing of the road, and ramps) significantly increased yielding behavior. The study also found that yielding rates were higher when pedestrians were crossing from and to a median strip. Weak or no evidence was found supporting that yielding behavior was higher with the no. of pedestrians or with pedestrians including children, older adults or people with mobility restrictions. [7]

Bastian J. Schroeder and Nagui M. Rouphail (2011) have suggested that driver yielding increases with assertive behavior demonstrated by the pedestrians who walk briskly on the crosswalks. Additionally, the study also concluded that the yield probability reduced with higher driving speeds and vehicles travelling in platoons. [8]

Curtis M. Craig et al. (2019) found that crosswalk markings increased yielding rates. However, presence of crosswalks is not sufficent to ensure safety in multilane roads and the markings should be provided in addition to other enchancements. [9]

Cambon de Lavalette (2009) asserted the road crossing environment, like its topographical features such as the no. of traffic lanes to cross, its infrastructure such as the presence of a central traffic island, its control systems such as traffic signals for vehicles and pedestrians, and the pedestrian's primary motive for crossing the road are important aspects in determining pedestrian behavior. [10]

In essence, the literature suggests that road characteristics such as curb extensions, presence of a median strip, crosswalk markings etc significantly increased yielding behavior. The studies also indicated that pedestrian speed and vehicle speed influence driver yielding. However, no significant increase in yielding behavior has been identified by the increase in the no. of pedestrians or for the pedestrians of older age groups.

3. Methodology

3.1 Study Area

The methodology can be divided into two parts after the selection of study areas. Data from six sites was obtained via on site study and from videography. An hour-long video recording was taken for every site.

Study locations were selected to fulfill the following criteria:

- Unsignalized crossing
- Straight section of road
- Marked cross-walk

Following site selection, an elevated, safe location was chosen for video recording at each site. After setting the video camera, a suitable trap length was designated in order to find the time taken by vehicles to cross the trap length and determine vehicle speed. Zebra crossing length was measured at site for the

Table 1: Variable Definition

S.N.	Variable	Type of Variable	Description
1	Pedestrian Number (N)	continuous	no. of pedestrians crossing at the same time
2	Pedestrian Gender (Pedgen)	categorical	male, female or male & female
3	Pedestrian Age (Pedage)	categorical	0 - 20 (category "1"), 20 - 60 (category "2"), above 60 (category "3")
4	Pedestrian Speed (Pedspeed)	continuous	crossing speed in meter per sec
5	Vehicle Type (Vehtyp)	categorical	bike, bus, car, cycle, jeep, microbus, platoon, scooter, taxi, van
6	Vehicle Speed (Vehspeed)	continuous	vehicle speed at the crossing in meter per sec
7	Driver Gender (Drigen)	categorical	male, female or male & female
8	Driver Age (Driage)	categorical	upto 30 years (category "1"), above 30 years (category "2")
9	Crossing length	continuous	length of zebra-crossing
10	Median-strip (Median)	categorical	presence or absence of a median strip
11	Road Marking Visibility (Marking Condition)	categorical	good (bright in color, evenly painted, free from discontinuities), fair (faded but free from discontinuities or bright at the edges but faded at the center), poor (faded, uneven color, discontinuous throughout the zebra crossing length)
12	Trafffic Direction	categorical	one-way or two-way traffic

calculation of pedestrian speed. The locality at each site was evaluated for the crossing marking condition.

The twelve independent variables used in the study are listed in Table 1. Four of the independent variables are continuous whereas the remaining variables are categorical in nature.

Figure 1: Pulchowk Campus Main Gate

Features of the first study area are:

- Trap length: 29.3m
- Length of zebra crossing: 9.8m
- Marking visibility: good
- Median Strip: absent
- Traffic flow: two-way
- No. of lanes: two

The first study area "Pulchowk Campus Main Gate" is shown in Figure 1. A total of 105 pedestrian-vehicle interactions were observed consisting of 139 pedestrians. Of the total observations, 86 interactions had only male, 18 had only female and 1 had both male and female pedestrians. The average speed of the pedestrians crossing the road was 1.2 m/s whereas the average speed of the interacting vehicles was 7.7 m/s.

Figure 2: Pulchowk Campus Back Gate

Features of the second study area are:

- Trap length: 38.7m
- Length of Zebra crossing: 15.32m
- Marking Visibility: good
- Median Strip: present

- Traffic flow: two-way
- No. of lanes: four

The second study area "Pulchowk Campus Back Gate" is shown in Figure 2. A total of 138 pedestrian-vehicle interactions were observed consisting of 247 pedestrians. Categorized by gender, 74 interactions had only male, 38 had only female, 20 had both male and female, 2 had a female carrying an infant, 2 had a female with a toddler and 2 has male, female with a toddler. The average speed of the pedestrians crossing the road was 0.99 m/s whereas the average speed of the interacting vehicles was 9.63 m/s.

Figure 3: Durbar Marga

Features of the third study area are:

- Trap length: 15.45m
- Length of Zebra crossing: 11.7m
- Marking Visibility: good
- Median Strip: present
- Traffic flow: two-way
- No. of lanes: four

The third study area "Durbar Marga" is shown in Figure 3. A total of 237 pedestrian-vehicle interactions were observed consisting of 383 pedestrians. Of the total observations, 139 interactions had only male, 55 had only female and 43 had both male and female pedestrians. The average speed of the pedestrians crossing the road was 1.09 m/s whereas the average speed of the interacting vehicles was 5.78 m/s.

Figure 4: Rabibhawan

Features of the fourth study area are:

- Trap length: 44m
- Length of Zebra crossing: 16.7m
- Marking Visibility: fair
- Median Strip: absent
- Traffic flow: two-way
- No. of lanes: four

The fourth study area "Rabibhawan" is shown in Figure 4. A total of 87 pedestrian-vehicle interactions were observed consisting of 105 pedestrians. Of the total observations, 59 interactions had only male and 28 had only female pedestrians. The average speed of the pedestrians crossing the road was 1.01 m/s whereas the average speed of the interacting vehicles was 11.4 m/s.

Figure 5: Kalimati

Features of the fifth study area are:

- Trap length: 38.7m
- Length of Zebra crossing: 11.3m
- Marking Visibility: poor
- Median Strip: absent
- Traffic flow: two-way
- No. of lanes: four

The fifth study area "Kalimati" is shown in Figure 5. A total of 268 pedestrian-vehicle interactions were observed consisting of 406 pedestrians. Of the total observations, 124 interactions had only male, 96 had only female, 24 had both male ans female pedestrians, 5 had a male with a cycle, 5 had a female carrying an infant, 9 had a female with a toddler, 2 had male, female, toddler and 3 had a female with children. The average speed of the pedestrians crossing the road was 0.946 m/s whereas the average speed of the interacting vehicles was 7.35 m/s.

Features of the sixth study area are:

- Trap length: 21.2m
- Length of Zebra crossing: 8.44m
- Marking Visibility: poor
- Median Strip: absent
- Traffic flow: one-way

- No. of lanes: four

Figure 6: New Road

The sixth study area "New Road" is shown in Figure 6. A total of 378 pedestrian-vehicle interactions were observed consisting of 678 pedestrians. Of the total observations, 205 interactions had only male, 104 had only female, 61 had both male, female and 8 had a female with a toddler. The average speed of the pedestrians crossing the road was 0.50 m/s whereas the average speed of the interacting vehicles was 4.00 m/s.

3.2 Video-graphic Study and Data Extraction

For data extraction, each pedestrian vehicle interaction was taken as an observation. Many pedestrians crossing the zebra had more than one interaction. The video footage was used to count the time taken by the pedestrians to cross the zebra in order to calculate pedestrian speeds. Further, the time taken by the vehicles to cross the trap length was also found out in order to calculate the vehicle speeds.

The video recording was also utilized to find the pedestrian and driver gender as well as age categories. All the data was integrated to generate a logistic regression model in the form of an equation. This equation is the major output of the study. It is anticipated that the equation will be valuable to make necessary changes at zebra crossings for increased safety of road users. Additionally, the equation may be used to predict driver yielding behavior for different values of the independent variables.

Data extraction was done following the procedure listed below:

- Data count was taken in terms of the number of observations.
- An observation was taken each time one or more pedestrians crossing the road made an interaction with a vehicle.

The data obtained from the six study locations is shown in Fig 7. The highest number of data was obtained from "New Road". The largest proportion of male pedestrians was at "Pulchowk Campus Main Gate" whereas the largest proportion of female pedestrians was at "Kalimati".

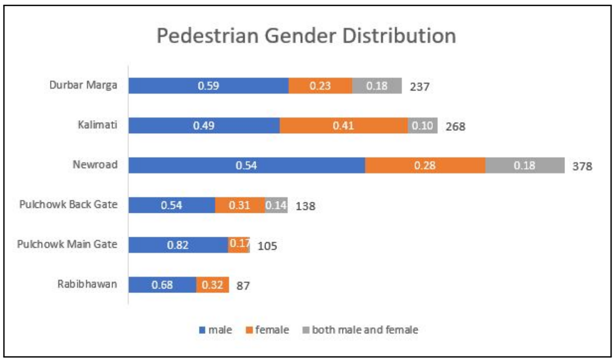

Figure 7: Location-wise Data Collection Categorized by Gender

3.3 Data Analysis

In this study, statistical analysis software Rstudio [11] was to determine the most influential factor in driver yielding behavior at different road crossings. Video observations were taken at various road sections to obtain the data.

For data analysis, logistic regression was performed by which yielding (Y) was taken as the dependent variable and the twelve variables mentioned in Table 1 were taken as the independent variables. Following the initial analysis, step-wise regression was performed whereby only the significant variables were taken for the next regression analysis in order to build a parsimonious model. Seventy percent of the data was used to build the model whereas thirty percent was used for testing. Testing of the model was done with the help of sensitivity and specificity analysis as well as area under ROC (Receiver Operating Characteristic) curve.

The area under the ROC (Receiver Operating Characteristics) curve, called AUC, is a single scalar value which measures the overall performance of a binary classifier. It is necessary because in logistic regression, the outcome is binary. This means for any value of probability; the outcome is required to be classified as 0 or 1. A value for cut-off needs to be stated so that beyond that value of, say 0.5, the outcome is 1 and the outcome for a probability up to a value of 0.5 is 0. An optimal value of cut-off is necessary to make the most accurate model possible.

To do so, ROC curve is plotted in which the y-axis represents the values for sensitivity and the x-axis represents values for (1 – specificity) for different cut-off points. The AUC value ranges from 0.5 to 1. An AUC of 0.5 shows no discrimination i.e., the model is unable to distinguish the observations with and without yielding.

3.4 Model Testing

A confusion matrix is created as shown in Table 2. The confusion or error matrix is a specific table layout which is used to describe the performance of a classification model on a set of test data, of which the true values are known.

Table 2: Confusion Matrix

	Actual Negative (N)	*Actual Positive (Y)*
Predicted Negative (0)	True Negative	False Negative
Predicted Positive (1)	False Positive	True Positive

True Positive (TP) is the number of outcomes in which the predicted values are positive and correctly predicted because the actual test data observations are also positive. True Negative (TP) is the number of outcomes in which the predicted values are negative and correctly predicted because the actual test data observations are also negative. False Positive (FP), also called Type I error is the number of outcomes incorrectly classified as positive by the model. False Negative (FN), also called Type II error is the number of outcomes incorrectly classified as negative by the model.

$$\text{Sensitivity} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Negative (FN)}}$$

$$\text{Specificity} = \frac{\text{True Negative (TN)}}{\text{True Negative (TN)} + \text{False Positive (FP)}}$$

Sensitivity is the ability of a model to correctly identify the cases in which the outcome is positive whereas specificity is the ability of a model to correctly identify the cases in which the outcome is negative. In this case, sensitivity tells us what percentage of the observations with yielding were correctly identified and specificity tells us what percentage of the observations with non-yielding were correctly identified.

3.5 Data Oversampling

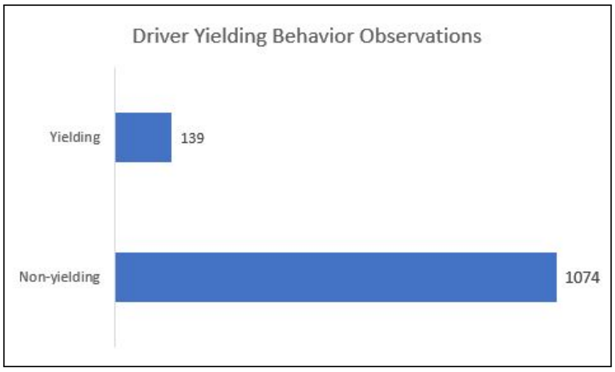

Figure 8: Total Yielding and Non-Yielding Observations

The data obtained has a class imbalance problem as shown in Fig. 8. The majority of the outcomes are negative (non-yielding) i.e 1074 while a small proportion is positive (yielding) i.e 139. The result of the imbalance is that the model performs well at predicting non-yielding behavior as given by the specificity value but the developed model has a low sensitivity value which means that the model is weak at predicting yielding behavior from a testing dataset. To improve the model, data oversampling is performed by reproducing the yielding observations to the number of observations accounted by the non-yielding ones. [12]

3.6 Outcome

3.6.1 Model Development

Model I

In this model, five of the variables were found to be significant at 95% confidence interval, which are: N, Pedestrian speed, vehicle speed, vehicle type (bus, car, jeep, taxi, van) and presence of median strip.

The categorical variables are listed alphabetically with the first category taken as the reference level from which the estimates for all the other categories are calculated. For instance, Vehtyp is a significant categorical variable of which "bike" is the first category listed alphabetically and is taken as the reference level. So, an estimate of 2.36 for "bus" means that "bus" increases the log of the odds of yielding by 2.36 times as compared to "bike".

- The outcome indicates that the log of the odds of yielding increases: 0.43 times per unit increase in N, 2.29 times per unit increase in pedestrian speed, 2.36 times in the presence of vehicle type "bus", 1.57 times in the presence of vehicle type "car", 1.54 times in the presence of vehicle type "jeep", 1.21 times in the presence of vehicle type "taxi", 1.53 times in the presence of vehicle type "van" and decreases by: 0.46 times per unit increase in vehicle speed, 2.58 times in the presence of a median strip.

- Null Deviance: 625.3 on 855 degrees of freedom
  Residual Deviance: 432.83 on 820 degrees of freedom
  AIC: 504.83
  No. of Fisher Scoring Iterations: 14
- Since the residual deviance is less than the null deviance, the model developed with the intercept and independent varaibles performs better than a model with only the intercept.
  AIC stands for Akaike Information Criteron which helps to compare different models of the same dataset a lower value of which indicates a better fit.

Model II

The second time, only significant variables were taken for model development.

- The regression equation is:
  y* = -3.03 + 0.33*N + 2.68*Pedspeed + a*Vehtyp - 0.36*Vehspeed + b*Median strip
- The values of "a" and "b" for categorical variables vehtyp and median strip can be found in Table 3. The estimates with a postitive value suggest that the presence of that variable increases the log of the odds of yielding whereas the presence of categories with a negative estimate suggest a decrease in the log of the odds of yielding.
- Null Deviance: 625.3 on 855 degrees of freedom
  Residual Deviance: 461.73 on 842 degrees of freedom
  AIC: 489.73
  No. of Fisher Scoring Iterations: 7
  Pseudo R-squared: 0.65 and corresponding P-value: 0
  Goodness of Fit: p-value= 3.73e^-28
  Sensitivity = $\frac{8}{8+29}$ = 21.6%
  Specificity = $\frac{316}{316+4}$ = 98.75%

Here, the value of residual deviance is less than of null deviance. Hence, the model with the intercept and the variables is better than the model with only the intercept. The intercept term is the expected value of the model when all the independent variables are zero.

Table 3: Estimate for categorical variables

Variable	Category	Estimate
Vehtyp	bike	0.00
	bus	2.02
	car	1.36
	cycle	1.49
	jeep	1.49
	microbus	0.00
	platoon	-2.26
	scooter	0.00
	taxi	0.04
	van	0.00
Median Strip	absent	0.00
	present	-0.98

Table 4: Confusion Matrix from Test Dataset

	N	*Y*
0	316	29
1	4	8

The pseudo R-squared is the overall effect size and its value from 0.2 ro 0.4 indicates an excellent model fit. It is used to compare different models. Here, the pseudo R-squared value of 0.65 indicates that 65% variation in yielding behavior can be explained by the independent variables.

For the goodness of fit, chi-square test is performed by which the p-value for the event to happen by random chance is 4.91e^-11 so the probability that the event would occur out of random chance is extrememly low.

The low value of Sensitivity indicates that the model is unable to accurately predict yielding behaviour. This is because of the less number of data showing yielding. However, the specificity value is high which means that the model is able to accurately predict non-yielding behavior.

In order to compare the accuracy of the various developed models, ROC curve is plotted and the Area Under Curve (AUC) of the ROC curve is calculated as shown in Figure 9. The value of AUC ranges from 0.5 to 1. Here, the AUC of 0.85 is an acceptable value.

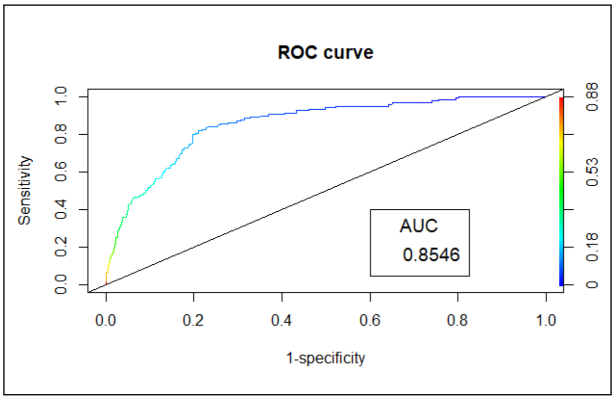

Figure 9: Area under ROC

Table 5: Estimate for Categorical Variables

Variable	Category	Estimate
Pedgen	male and female	-0.82
	others	0.00
Pedage	2	-0.55
	others	0.00
Vehtyp	bike	0.00
	bus	2.02
	car	1.36
	cycle	1.49
	jeep	1.49
	microbus	0.00
	platoon	-2.26
	scooter	0.00
	taxi	0.04
	van	0.00
Drigen	male	-0.89
	others	0.00
Driage	1	0
	1 and 2	-2.35
	2	-0.95
Median Strip	absent	0.00
	present	-3.64
Marking Condition	good	2.99
	others	0.00

3.6.2 Model III

In order to increase the Sensitivity i.e the ability of the model to correctly predict yielding, data oversampling is performed whereby the yielding observations are reproduced to the number of non-yielding observations. The oversampled data is analysed using all the variables and then using only the significant variables. The model is tested using the oversampled test data set. 70% of the data is used to train the model and the remaining 30% is used to test the model.

- The regression equation is:
  y* =1.31 + 0.3*N + a*Pedgen + b*Pedage +2.33*Pedspeed + c*Vehtyp - 0.49*Vehspeed + d*Drigen - e*Driage + f*Median + g*Marking Condition
- The values of "a", "b", "c", "d", "e", "f", "g" of the categorical variables can be found in Table 5.
- Null Deviance: 2096.1 on 1511 degrees of freedom
  Residual Deviance: 1299.5 on 1478 degrees of freedom
  AIC: 1367.5
  No. of Fisher Scoring Iterations: 13
  Pseudo R-squared: 0.89 and corresponding P-value: 0
  Goodness of Fit: p-value= 4.2e^-146
  Sensitivity = $\frac{271}{271+45}$ = 85.76%
  Specificity = $\frac{246}{246+74}$ = 76.87%

Table 6: Confusion Matrix from Over-sampled Test Dataset

	N	***Y***
0	246	45
1	74	271

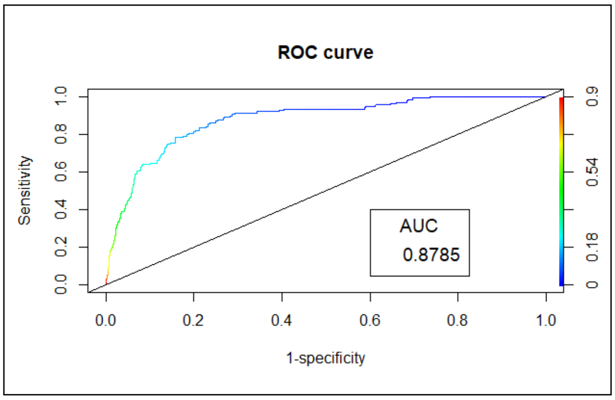

Figure 10: Area under ROC for Oversampled Dataset

From the final model, Model III it has been observed that the log of the odds of yielding increases by: 0.3 times per unit increase in N, 2.33 times per unit increase in pedestrian speed, 3.18 times in the pressence of "bus", 1.95 times in the presence of "car", 1.69 times in the presence of "jeep", 0.97 times in the presence of "microbus", 1.87 times in the presence of "taxi", 1.92 times in the presence of "van", 2.99 times in the presence of "good" road marking condition and decreases by : 0.81 times in the presence of pedestrians of "male and female" gender category, 0.55 times in the presence of pedestrian age category "2", 0.99 times in the presence of "platoon", 0.49 times per unit increase in vehicle speed, 0.89 times in the presence of "male" driver, 2.35 times in the presence of driver age category "1 and 2", 0.95 times in the presence of driver age cateogry "2" and 3.64 times in the presence of a medians strip.

Y* represents the log of the odds of yielding. In order to find the probability of yield, the following formula needs to be used:

Probability of yield = $\frac{exp(Y*)}{1+exp(Y*)}$

After the probability of yield is obtained, a threshold value of probability is taken with the help of which the data can be classified as a yield or a non-yield.

The higher sensitivity in the oversampled dataset shows that oversampling the yielding observations has resulted in a model which is a lot better at correctly predicting yielding behavior as compared to the previous model.

Also, the higher value of AUC i.e 0.87 in Fig.10 compared to 0.85 in Fig. 9 suggests that the oversampled dataset has produced a model with a greater predictive capability.

4. Conclusion

From the proposed Model III, the most significant variables in driver yielding behavior are ten in number. They are: no. of pedestrians crossing the road at the same time (N), pedestrian gender (male and female crossing), pedestrian age category "2", pedestrian speed, vehicle type (bus, car, jeep, microbus, platoon, taxi, van), vehicle speed, driver gender (male), driver age category (1 and 2, 2), presence of a median strip and good road marking condition.

The study did not find significant increase in yielding for females carrying an infant or for pedestrians of age category "3" (above 60 years).

The zebra crossing length or the traffic direction did not have any significant impact on yielding behavior.

5. Recommendation

From the research, it is clear that road safety can be improved by controlling a few of the road parameters which are: introducing a speed limit which is controlling the vehicle speed, prohibiting zebra-crossings on roads with a median strip and instead encouraging the use of road bridges and finally by improving the road marking conditions. Since changing the marking condition from "poor" to "fair" did not significantly improve yielding behavior, it is recommended to maintain "good" crosswalk markings.

In order to develop a more robust model, a larger data set could be utilized. The analysis can also be performed using alternative approaches and by evaluating the outcome from different methods, the most suitable one can be chosen.

References

[1] Department of Roads. *Nepal Road Standard 2070.*

[2] J S Kumar and S Suvash. Injuries and violence in nepal: A review. 2010.

[3] Anup Ojha. With 60 percent zebra crossings faded, pedestrians at risk of being run over https://kathmandupost.com/valley/2019/01/20/with-60-percent-zebra-crossings-faded-pedestrians-at-risk-of-being-run-over. 2019.

[4] Andra's Va'rhelyi. Driver's speed behaviour at zebra crossing: A case study. 1998.

[5] Ray Fuller. A conceptualization of driving behaviour as threat avoidance. 1984.

[6] Francesco Bella and Manuel Silvestri. Effects of safety measures on driver's speed behavior at pedestrian crossings. 2015.

[7] Paulo Anciaes, Giovanni Di Guardo, and Peter Jones. Factors explaining driver yielding behaviour towards pedestrians at courtesy crossings. 2010.

[8] Bastian J. Schroeder and Nagui M. Rouphail. Event-based modeling of driver yielding behavior at unsignalized crosswalks. 2011.

[9] Curtis M. Craig, Nichole L. Morris, and Yubin Hong. A case study on the impact of crosswalk marking on driver yielding to pedestrians. 2019.

[10] Brigitte Cambon de Lavalette, Charles Tijus, Sébastien Poitrenaud, Christine Leproux, Jacques Bergeron, and Jean-Paul Thouez. Pedestrian crossing decision-making: A situational and behavioral approach. 2009.

[11] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, PBC., Boston, MA, 2020.

[12] Weihong Han, Zizhong Huang, Shudong Li1, and Yan Jia. Distribution-sensitive unbalanced data oversampling method for medical diagnosis. 2019.