Proceedings of 12th IOE Graduate Conference
Peer Reviewed
ISSN: 2350-8914 (Online), 2350-8906 (Print)
Year: 2022 Month: October Volume: 12

A Study on the Factors Influencing Red Light Violation Behavior of Pedestrians in Kathmandu Valley

Deepak Raj Shah [a], Rojee Pradhananga [b]

[a, b] *Department of Civil Engineering, Pulchowk Campus, IOE, Tribhuvan University, Nepal*
[a] dr.deeprs@gmail.com, [b] rojee.pradhananga@pcampus.edu.np

Abstract
One of the most critical reasons for pedestrian fatalities in signalized urban intersections is their risk taking behavior and tendency to run red lights. The current study is an attempt at assessing the factors influencing red light violation behavior of Nepalese pedestrians. Binary logistic regression has been conducted on a total of 600 pedestrians. The logit model results show that Age, carrying a child/heavy load, traffic volume, pedestrian speed, number of pedestrians waiting at the crosswalk and group crossing, significantly affect pedestrians' decision to commit red light violation or wait until green. Parameters such as Gender and Number of active violators were not found to be significant predictors. The accuracy of the developed logit model in predicting red light violation behavior was 90.7% while the overall accuracy of the model was 82.2%.

Keywords
Red Light Violation, Logit Model, Pedestrian Safety, Road Crossing Behavior

1. Introduction

Minimization of travel time is of the utmost importance to any road user and especially pedestrians, which might explain their willingness to opt for the shortest route (crosswalk) [1]. Thus at-grade signalized crossings seem to be the most favored by pedestrians as, at signalized crosswalks, pedestrians are sequentially separated from vehicles because of traffic signals, allowing a safe crossing experience. However, pedestrians' signal violation especially red-light running can, on unfortunate circumstances, expose them to vehicles which might lead to fatalities. Studies have also claimed that, vehicle-pedestrian crashes during the red phase are 56% more likely to be fatal than during the green phase [2].

The lack of fully operational signalized crosswalks in Kathmandu is also a major issue. According to a Kathmandu Post article, published in 2016, people who are eager to follow signals, after arriving at the crosswalk, get greeted by the "red man" every time and they don't even know if it is actually working [3]. Since then, even with the introduction of countdown timer signals, the situation has arguably gotten worse in many areas due to rise in vehicle ownership resulting in increased traffic. Out of order signals have been a constant cause of fear and frustration for pedestrians. People are frightened to cross the streets as they could easily be run over by a speeding vehicle. But as observed from the data collected during the course of this research, majority of the pedestrians don't follow the signals even when the signals are properly functioning.

As per a study conducted in China, more than 50% of pedestrian related crashes occurred at signalized intersections in 2007, of which an average 65% of the cases were related to pedestrian's indiscipline or illegal road crossing behavior [4]. In developing countries like India, of all the pedestrian road fatalities, about 60% are recorded in urban areas with about 85% of said fatalities occurring at crosswalks [5] and pedestrians' risk-taking behavior such as pedestrian red-light violation can be regarded as one of the major culprits. Similar, if not more extreme, scenarios can be expected in the Nepalese context as well.

Contemporary literature reveals that one of the most critical reasons for pedestrian crashes in signalized urban intersections is their risk taking behavior and tendency to run red lights. Hence, assessing the factors contributing to this issue is of peak curiosity to researchers. This research aims to evaluate the factors affecting pedestrians' red light violation (RLV) behavior by applying Logistic regression (Logit Model) which is one of the most popular Machine Learning algorithms, falling under the Supervised Learning technique. The key insights obtained from this research can be utilized by concerned authorities to understand pedestrian psychology as well as plan, manage or even design the facilities accordingly.

2. Relevant Literature

To analyze a pedestrian's decision to commit red light violation, existing studies have mostly focused on demographics, social and traffic parameters. Concerning demographics, a large majority of the relevant studies have concluded that males and young adults are more likely to commit red light violation than females and elderly pedestrians contending that males and younger pedestrians are prone to risk taking behavior [6, 7].

Concerning social parameters, studies have concluded that if a pedestrian perceives other pedestrians waiting for green at the crosswalk upon his/her arrival, then the chances are high that he/she will also wait for green (follow the signal) to cross the road [7, 8]. This shows the conformity behavior of pedestrians. Regarding Groups, some studies claim that fewer signal violations occur when pedestrians travel in a group (2 or more) because pedestrians in groups feel more committed to the social norms [9, 10]. However, some studies have opposed this finding by concluding that pedestrians who arrive in groups are more likely violate the signal than someone who arrives alone [4]. Concerning the number of active violators, majority of the studies seem to agree that as the number of violating pedestrians increase, the chances of a pedestrian committing red light violation also increases [11]. Studies have also analyzed the crossing speeds of pedestrians and have found crossing speed to be positively related to pedestrian red light violation behavior [8]. A trip accompanied by elders or children has also been shown to reduce a pedestrian's willingness to run a red light [12].

There seems to be a consensus among existing studies that Traffic Volume and Pedestrians' decision to violate the signal are negatively related and Traffic Volume is a major significant factor affecting red light violation behavior [8, 13]. In other words, pedestrians are considerably less likely to attempt to cross illegally on red during high traffic volume compared to low traffic. Even though these studies have used traffic volume as a parameter and found it significantly affecting the violation behavior, most of them have utilized hourly volume rates and some have utilized the total traffic volume passing through a crosswalk during each red phase. The latter one might be more useful than the former in assessing the real condition but it still does not accurately represent how each pedestrian interacts with the traffic.

3. Methodology

Since there are only a few fully operational signalized intersections in Kathmandu, two of the busiest signalized intersections—Maharajgunj and Tripureshwor—which were fully operational were chosen for this research.

3.1 Videographic Survey

The video-graphic survey was conducted using a GoPro Hero 9 Black. After obtaining ethical approval from traffic agencies, the camera was mounted inconspicuously near the selected crosswalks. The video footage was taken from 8:30 AM to 11:30 AM at both locations. This interval of time, as per the author's observation, had a high flow of pedestrians as compared to other time periods throughout the day (because of office hours) and thus was selected. Another important reason for the selection of this time period was that the signals remained fairly operational and consistent during this interval as compared to other time intervals (afternoon and evening). The pedestrian signal in Maharajgunj remained consistent (working) for the entire duration, while the signal in Tripureshwor did turn off for brief intervals. Because of this on average only about 2 hours of video footage per day could be recorded in this timeframe. Thus, about 6 hours of pedestrian flow was recorded for each site (totaling 12 hours) spanning across 3 weekdays. The camera settings were 1080p, 30fps, 16:9 aspect ratio and linear view (19-39mm).

3.2 Site and Variable Description

Sites: Maharajgunj and Tripureshwor are two of the most busy intersections in Kathmandu Valley. Regarding the pedestrian flow, it was observed by the authors that these sites had a heavier pedestrian flow as compared to others because of the presence of multiple bus stops, hospitals, malls, restaurants, colleges etc.

Figure 1: Site 1: Maharajgunj

- 4 Legged intersection (90° angle between legs)
- Crosswalk length = 24.6 metres
- 3 Lanes in both direction (6 Total)
- Median : Absent
- Two way Traffic
- Median Red Phase duration (sec): 300

Since this study aims at analyzing the road crossing behavior of pedestrians, only the pedestrians who enter the crosswalk on their own volition were considered. Thus, for the crosswalk selection, the only major criteria was minimal traffic police/personnel intervention. The crosswalk selection in Tripureshwor was a bit more tedious as compared to Maharajgunj because traffic police intervention was especially prominent in this site. Thus, among the three crosswalks in Tripureshwor, the one with the densest pedestrian flow was selected for this research.

Figure 2: Site 2: Tripureshwor

- 3 Legged intersection (Rotary)
- Crosswalk length = 18.7 metres
- 2 Lanes in both direction (4 Total)
- Median : Absent
- Two way Traffic
- Median Red Phase duration (sec): 206

Variables: Logit model requires the dependent or the outcome variable to be binary (categorical with values 0 and 1) while the independent variables can be nominal, continuous/scale or even ordinal. 4 of the 8 independent/predictor variables are categorical while the remaining 4 are continuous.

DEPENDENT VARIABLE

Ped.RLV: This is the **dependent variable**. It provides the information about a pedestrian's decision towards the red light.
0 = "Waited For Green";
1 = "Red light violation"

INDEPENDENT VARIABLES

Gender: Gender is a demographic parameter. It is categorical.
1 = "Female"; (Reference Category)
2 = "Male"

AGE: Age is a demographic parameter. It is categorical.
1 = "15 to 40";
2 = "40 to 60"
3 = "More than 60 years" (Reference Category)

Ped.Carry: Carrying a child/heavy load. Both social and movement restriction parameters are represented in this variable. It is categorical.
1 = "Not Carrying a Child/Heavy Load"; (Reference Category)
2 = "Carrying a child/heavy load"

Perceived.Vol: Perceived volume is a traffic parameter. It is continuous. As per the existing literature, majority of the studies have included the hourly traffic volume as a parameter while a few of the studies have used total red phase traffic as a parameter. These parameters are proven to be significantly affecting the red light violation behavior of pedestrians. But in the context of Nepal, those parameters don't make much sense as the pedestrian signals and traffic signals are not in sync with each other and a lot of times vehicles remain still even during the pedestrians' red phase (where they should be moving). Thus a new parameter Perceived Volume has been created which is explained below.

Perceived Volume for people who arrive at red and wait until green to cross the road (Following the traffic signal):

- Total number of vehicles passing the crosswalk (in both direction) during the total waiting time converted to per minute.

For people who violate the signal by crossing at red without waiting:

- Total number of vehicles passing the crosswalk in the field of vision (both direction) of pedestrian, while he/she is actively crossing the road (converted to vehicles per minute).

- Vehicles that don't affect the pedestrian's crossing decision (for e.g: vehicles passing the crosswalk behind the pedestrian) are not counted; only vehicles in front of the violating pedestrian and vehicles at conflict with him/her are counted.

For people who violate the signal by crossing at red after waiting a certain amount of time (showing impatience):

- Total number of vehicles passing the crosswalk (in both direction) in total waiting time plus the total number of vehicles in the field of vision after the pedestrian starts crossing (converted to vehicles per minute).

Ped.Speed: It is the uninterrupted speed of a pedestrian. It is a continuous variable and signifies the impatience/haste.

Wait.Ped: It is the total number of pedestrians waiting for green at a crosswalk just before the arrival of a pedestrian. It is a social parameter and is a continuous variable.

Group.Size: Group Size is a social parameter. It is categorical and represents the group size of a subject pedestrian at arrival.
1 = "Single Pedestrian"; (Reference Category)
2 = "Two or more pedestrians"

Violate.Ped: It is the total number of active violations perceived by a pedestrian at arrival. It is a social parameter and is a continuous variable.

3.3 Sample Size for Logistic Regression

Existing studies claim that, for logistic regression, the minimal sample size can defined as Equation 1:

$$N = 10 * k/p \tag{1}$$

Where,
N = Minimum required Sample Size
k = Number of predictor/ independent variables
p = Smallest proportion of negative or positive (binary) cases in the population, with "1" indicating that the event occurred (Positive case) and "0" indicating that the event did not occur (Negative case) [14, 15].

Population Study: The selection criteria for the pedestrians in this study, as briefly highlighted before, is as follows:

- As this study aims at assessing the red light violation behavior of pedestrians, the pedestrians who arrive at the crosswalk during the Green phase have not been included in the study; only the pedestrians who arrive at the crosswalk on red phase have been included in this research.

- Pedestrians whose road crossing decisions were influenced by an external agent such as the traffic personnel/policemen, who were out of the video frame, who crossed the street without stepping on the signalized crosswalk even once or were out of bounds were discarded from the research.

- During some cycles, through vehicles (stopping at red phase) were found temporarily halted right on top of the crosswalk, thus obstructing the pedestrians' right of way. Such instances where more than 50% of the crosswalk was covered with illegally stopped vehicles were also discarded from any further consideration.

- During some cycles, the traffic and pedestrian signals were temporarily turned off by the traffic personnel due to a myriad of reasons such as VIP traffic flow, extreme congestion etc. Pedestrians traversing the crosswalk during these periods were also not included in the current study.

The total population is illustrated in Table 1.

Table 1: Total Population

SITE	Maharajgunj	Tripureshwor	TOTAL
Population	3256	2180	**5436**
Males	1889	1416	3305
Females	1367	764	2131
Males (Followed)	466	201	667
Females (Followed)	343	110	453
Total (Followed)	**809**	**311**	**1120**
Total Followed (%)	**24.85%**	**14.27%**	**20.60%**
Total Violated (%)	75.15%	85.73%	79.39%

It is evident from the population that, the proportion of people who violated (1) (Positive Cases) is significantly greater than the proportion of people who followed (0) (Negative Cases). Thus the lower proportion p(0) or proportion of Negative cases has been chosen for the total sample size evaluation in order to ensure that a respectable sized sample which is an accurate representation of the total population is obtained. p(0) for Tripureshwor is the lowest, hence it is used for the sample size computation using equation 1.

$$N = 10 * 8/(0.1427) = 560.61 \approx 600 \tag{2}$$

The calculated required sample size of 600 was obtained using systematic sampling approach.

To avoid inclusion of predominant number of samples from a single location, the samples from each site were taken in the exact proportion to the respective site's population shown in table 1. The sampling description is clearly summarized in table 2.

Table 2: Sample Description

Site	Proportion of Total Population	Sample Size (Out of 600)	Sample Size (Per Day)
Maharajgunj	0.5989 ≅ 0.6	360	120
Tripureshwor	0.401 ≅ 0.4	240	80
TOTAL	**1**	**600**	**200*3 = 600**

Systematic sampling was conducted for Maharajgunj by extracting data for every 3rd pedestrian (as it has a comparatively large population) encountered in the video footage (Day 1, Day 2, Day 3); the sampling was stopped after total samples reached 120 for each day. Similarly, for Tripureshwor, data for every 2nd pedestrian was extracted; sampling was stopped after total samples reached 80 for each day. Systematic sampling is widely popular in pedestrian related studies due to the sheer number of pedestrians. Indian studies based on the similar premise of red light violation behavior also conducted systematic sampling by selecting the pedestrians in an odd sequence of arrival counts such as first, third, fifth, seventh and so on [8].

3.4 Logistic Regression

Logistic regression, also sometimes referred to as the logistic model or logit model, falls under the Supervised Machine Learning techniques which are simply the techniques that allow a machine to be trained using a well labeled dataset and finally on the basis of that data, the machine predicts the output. Binary logistic regression is used when the dependent variable is dichotomous or binary with only two possibilities (0 or 1, Yes or No, Failure or Success, Violated a signal or didn't violate etc.) and the independent variables are either continuous or categorical.

Odds: The odds of an event is simply a measure of how likely the event will occur or the likelihood of a particular outcome. Odds, in terms of probability, can be defined as the ratio of the probability that an event will occur (p) to the probability that it will not occur (1-p). In the current study, the odds of a pedestrian committing red light violation is given by:

$$\textbf{Odds of Ped.RLV} = \frac{P(Red\,light\,violation)}{P(wait\,till\,green)} \tag{3}$$

Log Odds, Logit Link and Probability of RLV The log odds is simply the natural log of the odds and is a linear representation of a probability scale. It converts the probability (between 0 and 1) to a scale from +∞ to -∞. It is done using the logit link function. For example, if someone's probability to violate the signal was predicted to be 0.5, the log odds, as per equation 4, would be 0 (representing he/she can either violate or follow).

$$ln(odds) = ln(\frac{p}{1-p}) \tag{4}$$

The higher the scale goes from 0 to +∞ (positive log odds), the chances of violating the signal increases and vice versa. Through logit link, the inverse is also possible, i.e. we can take a linear combination of the covariate values (which may take any value between ±∞) and convert those values to the scale of a probability, i.e., between 0 and 1.

Logistic Regression utilizes the Logit link function which is used to model the probability of 'success'—here in this case the probability of committing RLV—as a function of independent variables (covariates). The logit link function logit(Y) is defined in Equation 5.

$$Y = ln(\frac{p}{1-p}) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_n X_n \tag{5}$$

Where,
Y = *Logit or Log odds of the dependent variable (Ped.RLV)*
α = Intercept
$\beta_1, \beta_2, \beta_3, \ldots, \beta_n$ = Coefficients
$X_1, X_2, X_3 \ldots, X_n$ = Independent Variables

$$p = \frac{e^Y}{1+e^Y} \tag{6}$$

The probability of success or, in this case, the predicted probabilities of RLV (p) can be calculated by rearranging Equation 5, by talking the antilog resulting in Equation 6.

Odds Ratio: While log odds is quite effective at conveying the nature of relationship between dependent and independent variables, it is not fairly intuitive. Thus in logistic regression, a more popular and intuitive metric used by numerous disciplines is that of "Odds Ratio". Odds ratios are simply the exponential of the obtained independent variable coefficients or exp(β)

4. Results and Discussion

Stepwise Backward Elimination has been applied in order to eliminate any non significant parameters. Gender and Violate.Ped were found to be non-significant parameters (at 5% significance level) and were thus eliminated from the final model at the end of three iterations (stepwise). The final model is shown in table 3.

Table 3: Logit Model Output

	coef(β)	S.E.	Wald	df	Sig.	Exp(β)
AGE1	0.958	0.441	4.733	1	**0.030**	2.608
AGE2	1.451	0.468	9.597	1	**0.002**	4.268
Ped.Carry2	-0.860	0.340	6.382	1	**0.012**	0.423
Perceived.Vol	-0.028	0.003	84.488	1	**0.000**	0.972
Ped.Speed	1.501	0.701	4.581	1	**0.032**	4.488
Wait.Ped	-0.355	0.061	33.388	1	**0.000**	0.701
Group.Size2	0.703	0.287	5.996	1	**0.014**	2.019
Constant	2.231	1.053	4.490	1	**0.034**	9.308

Logistic regression provides us with the parameter estimates or the independent variable coefficients, [coef(β)], and the odds ratios [Exp(β)]. The results are discussed below:

- Compared to pedestrians aged *60 and above* (reference category), pedestrians between *15 to 40* and pedestrians between *40 to 60* are **2.6 times** and **4.3 times** (respectively) more likely to violate the signal.
  This is in line with existing literature; younger pedestrians tend to take more risks than the elderly.
- Pedestrians who are *carrying a child or a heavy load* are **58%** less likely to violate the signal as compared to pedestrians who are *not carrying a child or a heavy load*. This proves that pedestrians are more wary of the traffic signals when they are accompanied by a child, or when their movement has been somewhat restricted.
- Pedestrians are less likely to violate signals during *high traffic volumes* as opposed to *low traffic volumes*. The newly developed variable, *Perceived Volume*, was found to be significantly affecting a pedestrians' decision to commit violation. This shows that pedestrians don't want to take risk and commit signal violation, when a large number of vehicles are passing through the crosswalk.
- Pedestrians who are in *hurry/rush* or just generally walk faster are more likely to violate the signals than pedestrians who aren't in a rush or just generally walk slower. This might be a result of impatience or haste. As per the site study, a pedestrian has to face a median red time of 292 seconds if he/she arrives at

the crosswalk just when the red phase begins. Due to these excruciatingly long Red phases, even the pedestrians who are willing to follow the signal, are forced to take risks to cross the road. Long red duration causes impatience which leads to agitation. People engaged in risk taking behavior such as RLV tend to accept smaller gaps and also run to cross the street, which can be life-threatening.

- The more people a pedestrian perceives to be *waiting for green at a crosswalk*, the more likely he/she is to follow the signal. This proves the conformity behavior of Nepalese pedestrians.
- Pedestrians who arrive at a crosswalk, during red phase, *in groups of 2 or more people* are nearly **2 times** more likely to violate the signal than someone who arrives at a crosswalk alone. This result is in direct contrast to existing studies which claim that when pedestrians travel in a group, they feel more committed to the social order and tend to stick to social norms thus violating less [9]. The reason for the contrasting result might be because of the high conformity behavior of Nepalese pedestrians. Furthermore, due to high violation rate, when a group arrives at a crosswalk, the chances are high that the group will see someone (or another group) violating and thus will readily imitate the behavior. Also pedestrians in groups are observed to be more reckless than a pedestrian who crosses alone.
- A person's *Gender* doesn't affect the decision to violate or wait for green.
- The *total number of active violations* perceived by a pedestrian at the moment of arrival also doesn't affect the pedestrians' decision to violate or wait for green. The reason for this might be the high violation rate as the likelihood of a pedestrian perceiving an active violation–regardless of whether he/she follows or violates–is quite high.

Model Equation: The developed model equation is shown in equation 7. This model provides us the log odds of someone committing RLV, and using equation 6, we can determine the probability that someone will commit RLV. For the categorical variables, if a pedestrian falls in the reference category, the categorical variables in the model will be coded as 0. Using a cutoff value of 0.5, the obtained probability can be be used to predict if someone violates the signal *p(RLV)>0.5* or follows the signal *p(RLV)<0.5*. The model accuracy in predicting red light violation was found to be 90.7%, while the overall accuracy of the model was 82.2%.

$$\begin{aligned} Y = &\ 2.231 + 0.958 * AGE1 + 1.451 * AGE2 \\ &- 0.860 * Ped.Carry2 - 0.028 * Perceived.Vol \\ &+ 1.501 * Ped.Speed - 0.355 * Wait.Ped \\ &+ 0.703 * Group.Size2 \end{aligned} \tag{7}$$

As per table 4, the model is highly significant, and the Pseudo R-square value was 0.504 which is a normal range for studies involving the unpredictable nature of human psychology.

Table 4: Model Significance and Pseudo R^2

	Chi-square	*df*	*Sig.*	*-2 Log likelihood*	*Nagelkerke R Square*
Model	242.275	7	***0.000***	407.264	**0.502**

The sigmoid curve, helps us to easily visualize the relationship between variables and a pedestrian's decision to commit RLV. Perceived volume and Age are chosen for the graph in figure 3.

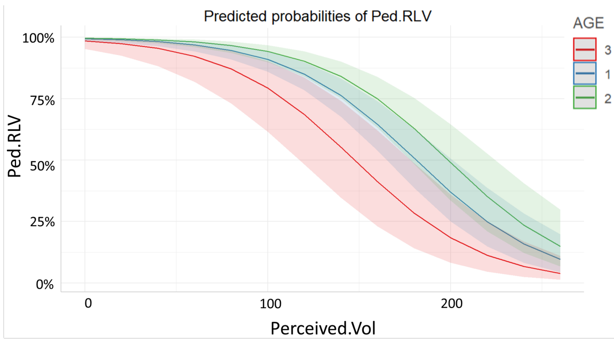

Figure 3: Sigmoid Curve: Perceived.Vol vs AGE

It is evident from the sigmoid plot that with the increase in traffic volume, a pedestrian is more likely to follow the traffic signal (wait for green) as the probability reduces below 0.5. Also we can see that age category "3" or *"pedestrians aged 60 and above"–red curve*–are the least likely to violate the signal and *"pedestrians aged 40 to 60"–green curve*–are the most active violators (regardless of traffic).

5. Conclusion & Recommendations

The overall study outcome suggested that the pedestrian signal violation at signalized intersections in Kathmandu valley is highly associated with various social as well as engineering and demographic attributes/factors. Younger pedestrians tend to take more risks than elders and pedestrians accompanied by children or those carrying a heavy load are more conscious of the traffic signal. Due to the long red phases or other personal reasons, pedestrians get impatient and violate more in haste increasing their speed exponentially. This exposes pedestrians to unwanted risks.

Concerning social attributes, when pedestrians arrive in groups of 2 or more, they are more likely to violate the signal than someone who arrives at the crosswalk alone. Furthermore, the more people a pedestrian perceives to be waiting for green at the crosswalk, the more likely he/she is to follow the signal, which explains the conformity behavior of Nepalese pedestrians. Impatience, conformity psychology, and lack of safety awareness are external and internal reasons for pedestrian violations.

Concerning the traffic/engineering parameters, in the Kathmandu Valley, because the pedestrian signals and traffic signals are not synchronized with one another, a new parameter called perceived volume (different for every pedestrian) was introduced. It was proven to be a significant factor influencing pedestrian signal violation behavior. The more vehicular volume a pedestrian perceives, the less likely he/she is to violate the signal.

A larger sample size and additional parameters could be incorporated in order to develop a more robust model. As signalized intersections are limited in the current scenario, more sites with a variety of distinct features could be included in future related studies in order to analyze pedestrians' signal violation behavior, which could be utilized to design facilities accordingly.

References

[1] Mariya Khatoon, Geetam Tiwari, and Niladri Chatterjee. Impact of grade separator on pedestrian risk taking behavior. *Accident Analysis & Prevention*, 50:861–870, jan 2013.

[2] Viola R and Roe M. New York City Pedestrian Safety Study & Action Plan. *Institute of Transportation Engineers Annual Meeting and Exhibit 2011*, (August):18–19, 2011.

[3] The Kathmandu Post. Only five out of 365 traffic lights working in Kathmandu Valley. 2016.

[4] Gang Ren, Zhuping Zhou, Weijie Wang, Yong Zhang, and Weijie Wang. Crossing Behaviors of Pedestrians at Signalized Intersections: Observational Study and Survey in China. *Transportation Research Record: Journal of the Transportation Research Board*, 2264(1):65–73, jan 2011.

[5] Ankit Bansal, Tripta Goyal, and Umesh Sharma. Pedestrian Safety on Crosswalks in India - Need of the Hour. *Journal on Today's Ideas - Tomorrow's Technologies*, 6(1):35–46, 2018.

[6] Hongwei Guo, Ziyou Gao, Xiaobao Yang, and Xiaobei Jiang. Modeling pedestrian violation behavior at signalized crosswalks in China: A hazards-based duration approach. *Traffic Injury Prevention*, 12(1):96–103, 2011.

[7] Wuhong Wang, Hongwei Guo, Ziyou Gao, and Heiner Bubb. Individual differences of pedestrian behaviour in midblock crosswalk and intersection. *International Journal of Crashworthiness*, 16(1):1–9, 2011.

[8] Rahul Raoniar and Akhilesh Kumar Maurya. Pedestrian red-light violation at signalised intersection crosswalks: Influence of social and non-social factors. *Safety Science*, 147(September 2020):105583, mar 2022.

[9] Tova Rosenbloom. Crossing at a red light: Behaviour of individuals and groups. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(5):389–394, 2009.

[10] Marilyne Brosseau, Sohail Zangenehpour, Nicolas Saunier, and Luis Miranda-Moreno. The impact of waiting time and other factors on dangerous pedestrian crossings and violations at signalized intersections: A case study in Montreal. *Transportation Research Part F: Traffic Psychology and Behaviour*, 21:159–172, 2013.

[11] P. P. Koh, Y. D. Wong, and P. Chandrasekar. Safety evaluation of pedestrian behaviour and violations at signalised pedestrian crossings. *Safety Science*, 70:143–152, 2014.

[12] Weihua Zhang, Kun Wang, Lei Wang, Zhongxiang Feng, and Yingjie Du. Exploring factors affecting pedestrians' red-light running behaviors at intersections in China. *Accident Analysis and Prevention*, 96:71–78, 2016.

[13] Abhinav Kumar and Indrajit Ghosh. Non-compliance behaviour of pedestrians and the associated conflicts at signalized intersections in India. *Safety Science*, 147(December 2021):105604, 2022.

[14] P Peduzzi, J Concato, E Kemper, T. R Holford, and A. R. Feinstein. A Simulation Study of the Number of Events per Variable in Logistic Regression Analysis. Journal of(49(12)):1373–1379, 1996.

[15] Dennis Acuna, Rajab Suliman, and Nasir Elmesmari. A Practitioner Methodology for Mitigating Electronic Data Risk Associated with Human Error. *Journal of the Midwest Association for Information Systems*, Vol. 2021(2):Article 2, 2021.