TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Article

Transportation Research Record
1–13

DOI: 10.1177/0361198118797212
journals.sagepub.com/home/trr
SAGE

Using Deep Learning in Severity Analysis of At-Fault Motorcycle Rider Crashes

Subasish Das[1], Anandi Dutta[2], Karen Dixon[1], Lisa Minjares-Kyle[3], and George Gillette[4]

Abstract
Motorcyclists are vulnerable highway users. Unlike passenger vehicle occupants, motorcycle riders do not have either protective structural surrounding or the advanced restraints that are mandatory safety features in cars and light trucks. Per vehicle mile traveled, motorcyclist fatalities occurred 27 times more frequently than passenger car occupant fatalities in traffic crashes. In addition, there were 4,976 motorcycle crash-related fatalities in the U.S. in 2014—more than twice the number of motorcycle rider fatalities that occurred in 1997. It shows that, in addition to current efforts, research needs to be conducted with additional resources and in newer directions. This paper investigated five years (2010–2014) of Louisiana at-fault motorcycle rider-involved crashes by using deep learning, which is a competent tool for mapping a high-multidimensional input into a smaller multidimensional output. The current study contributes to the existing injury severity modeling literature by developing a deep learning framework, named as *DeepScooter*, to predict motorcycle-involved crash severities. The final deep learning model can predict severity types with 100% accuracy with training data, and with 94% accuracy with test data, which is not attainable by using a statistical method or machine learning algorithm. The intensity of severities was found to be more likely associated with rider ejection, two-way roadways with no physical separation, curved aligned roadways, and weekends. It is anticipated that the *DeepScooter* framework and the findings will provide significant contributions to the area of motorcycle safety.

Typically identified as "vulnerable roadway users," motorcyclists are an at-risk group of roadway users whose death rates have consistently remained higher than other vehicle-based roadway users. In 2014 alone, motorcyclist fatalities "occurred 27 times more frequently than fatalities in other vehicles" with 4,976 motorcycle crash-related fatalities in the U.S. This is more than twice the number of motorcycle rider fatalities that occurred in 1997, and contrasts the 27% reduction in the number of fatalities involving passenger cars and light trucks (*1*).

Motorcycle crashes represent 5% of all fatal crashes in Louisiana each year, and yet only represent 1% of all total crashes. In 2016, 98 fatal crashes (30% higher than 2010 statistics) involved motorcycles (*2*). The conventional approach to crash-severity analysis has been to establish associations between traffic and driver characteristics, roadway and environment conditions, and crash occurrence. The shortcoming of most of the models developed using this approach is that they rely on aggregate measures and general inferences. In addition, identifying contributing factors using observational data comprises a wide variety of associations because of the assumptions considered during different modeling techniques. Thus, it is essential to determine the influence of patterns of associated factors with crash severity. The 2010–2014 traffic crash data associated with at-fault motorcycle riders for Louisiana were used in this study. This study restricted the deep learning framework development by analyzing only at-fault motorcycle rider crashes because not-at-fault motorcycle rider crashes involve other vehicle drivers' driving traits, which is outside the scope of the current study. A deep learning framework "*DeepScooter*" was developed to better predict injury outcomes from the significant factors associated with motorcycle crashes.

[1]Texas A&M Transportation Institute, Texas A&M University System, College Station, TX
[2]Computer Science and Engineering Department, Texas A&M University, College Station, TX
[3]Texas A&M Transportation Institute, Texas A&M University System, Houston, TX
[4]Department of Civil and Environmental Engineering, University of California, Berkeley, CA

Corresponding Author:
Address correspondence to Subasish Das: s-das@tti.tamu.edu

Earlier Work and Research Context

The most prominent area of transportation safety analysis are crash frequency analysis (count data problem), and injury-severity analysis (classification problem). Lord and Mannering provided a detailed research synthesis on crash count data and related methods and limitations for examining such data (*3*). Savolainen et al. presented a similar assessment on injury-severity analysis (*4*). Recently, Mannering and Bhat extended and bridged both of these studies (*5*). Mannering et al.'s study presented a detailed discussion of unobserved heterogeneity in crash data analysis along with their strengths and weaknesses (*6*).

Research and analysis of motorcycle crashes throughout the years has provided an abundance of valuable and important information to assist with the overall need to reduce motorcyclists' injuries and deaths caused by crashes. Rider variables such as age, gender, and impairment, for example, have consistently been identified in the literature to influence the likelihood of increased crash severity. Researchers have identified significant relationships between the age of a rider and increased crash-severity outcomes in which older riders (over the age of 25), although less likely to be at fault in a collision, are more likely to be severely injured in crashes compared with younger riders (*7–13*). Male riders have lower probabilities of being severely injured in a crash compared with females, but on average have a higher likelihood of being fatally injured in a crash (*7*, *8*, *14*).

Intoxicated driving has been identified as a significant contributor to motorcycle crash severity. Crashes involving impaired riders have been found to be more severe with a higher risk of fatal or major injury outcomes than crashes involving non-impaired riders (*10*, *15–18*). Research has shown that impaired motorcyclists are more likely to be involved in single-vehicle crashes, be at fault, and less likely to be wearing a helmet at the time of the crash, which can affect crash severity (*9–10*, *13*, *15*).

Helmet use has been identified throughout several studies as a significant variable in motorcycle crash severity, in which the likelihood of a fatality or the severity of a crash is higher when a helmet is not worn (*8–10*, *13*, *14*). Crashes in which a motorcyclist is wearing a helmet have been associated with reduced chances of severe or fatal injuries in both rural or urban environments, decreases in injuries sustained in both single and multi-vehicle crashes, and reduced frequency of severe injuries (*7*, *12–14*, *16*, *19*).

Temporal factors such as season, day of the week (weekend vs. weekday), and time of day at which motorcycle crashes occurred have also been linked to motorcycle crash-severity outcomes in the literature. Motorcycle crashes that occur at night (between the hours of 8 p.m. and 6 p.m.) or early morning hours, on weekends, during the early months of the rider "season" or summer months have been shown to have higher fatal and severe injury probabilities (*10*, *12–14*, *20*, *21*). For example, Savolainen and Mannering found that crashes that occurred in April and July had a 111% and 98% greater probability of being fatal. Later months also showed a lower likelihood of protective gear use because of higher temperatures (*14*).

Several geometric variables have been identified in the literature to contribute to motorcycle crash-severity outcomes including roadway type, light conditions, posted speeds, roadway features such as curves or T-junctions, and weather. According to the literature, roadways that have the following characteristics significantly influence motorcycle crash severity: two-lane roads, farm to market roads, curved roads, looped or "lollipop" designed roads, and T-junction roads controlled by "stop, give-way signs or markings," (*7*, *9*, *10*, *16*, *19–22*). Motorcycle crashes that occurred on highways, near driveways, intersections, or signalized intersections were also found to influence crash severity, as it is estimated that these factors negatively affect a driver's ability to see a motorcyclist (*9*, *23*). Crashes that occurred in rural localities were also significantly linked to motorcycle crash severity, which may be because of overall higher posted speeds, roadway geometry, the absence of streetlights, and a higher propensity for two-lane roadways (*16*, *18*).

Motorcycle injury severity has also been significantly linked to speed. Roadways with higher posted speeds or motorcycle crashes that occur at "unsafe speeds," where speed is listed as the primary contributing factor or posted speed limits are higher than 55 mph, have been shown to increase the probability of a fatality and more severe injuries (*10*, *18*, *20*, *22*). For example, Savolainen and Mannering found a 212% increase in the probability of a fatality when the crash citation involved the factor "unsafe speed" (*6*). This is true for both single-vehicle crashes and multi-vehicle crashes (*13*, *14*, *18*).

Motorcycle crashes have been found to be more severe during daytime weather versus wet or rainy weather (*7*, *10*, *13*). Savolainen and Mannering posit this may be as a result of lower speeds maintained by riders while driving through wet or rainy conditions, in which riders may exhibit more caution and lower speeds as they adjust to "perceived higher risk" on wet roadways versus dry (*7*).

Overview of Models Used to Analyze Crash Severity

Several modeling approaches have been applied throughout the literature to analyze crash severity by examining injury-severity levels related to the various factors discussed in the previous paragraphs. This analysis will provide a snapshot of models used in the literature review and background of the research works in this paper.

Many studies have used multinomial ordered probit and logit models for crash-severity analysis, which analyzes various categories of injury-severity levels in order of severity ranging from no injury to a fatality (*11*, *12*, *17*, *21–23*). Research has identified potential limitations to this model because of constrained effects resulting from ordered modeling and underreporting of crashes that have minor or no severe injuries (*7*).

Multinomial logit models (MNL) have been used by researchers to address the above issues as MNL models "consider three or more outcomes and do not explicitly consider the ordering," (*23*). Researchers have used this model to analyze motorcycle rider accident severity in single-vehicle crashes (*15*), develop probabilistic models of motorcyclists' injury severities for single and multi-vehicle crashes (*7*), compare severity of motorcycle injury by crash types (*19*), analyze differences in factors that affect severity of motorcyclist's crash injuries (*16*), and predict the probability of crash severities (*10*).

Mixed logit analysis models allow for "heterogeneous effects and correlation in unobserved factors" (*19*) and have been used to examine crash-specific factors (crash factors, roadway and environment conditions, and vehicle attributes) on two-vehicle crash-severity outcomes involving motorcycles (*14*).

Log linear modeling provides "measures of the magnitude, direction and statistical significations of main effects" and "interactions among a set of categorical variables." Haque et al. utilized log linear modeling to investigate effects of traffic, environmental, and roadway factors on roadway crash in Singapore to establish influential factors in various location types (*24*). Other methodological approaches include Empirical Bayesian analysis and stepwise logic regression (*8*, *25*, *26*). It is important to note that conventional statistical models are good at statistical inference. The common limitation of these methods is poor prediction accuracy. In addition, the models are based on assumptions. Violations of any of the assumption will produce biased results.

Overview of Deep Learning Method Analysis

Within the field of data analysis, there are two significantly differing opinions regarding its appropriate treatment—roughly categorized as data or statistical modeling and algorithmic modeling or machine and deep learning (*27*). The data modeling approach transcribes to the belief that an underlying stochastic process has generated the data, such that the response variables can be related to a set of predictor variables. In order to evaluate the model's success, this approach would apply goodness-of-fit tests based around established and acceptable margins of certainty. In contradiction, algorithmic modeling focuses on the process of understanding the unknown by minimizing the error rate through a black box algorithmic model.

Although deep learning has not been widely used in past studies, many studies have used machine learning and data mining algorithms in transportation safety analysis (*28–40*). Deep learning is a branch of artificial intelligence that attempts to model complex information through a series of processing layers. Although deep learning has recently become the new primary focus within the artificial intelligence community, the models and ideas behind this technique have been discussed for over half a century.

In the field of transportation engineering, the majority of statistical work could be shown to fall into the category of data modeling. Although there are not similar papers in deep learning among the established literature for motorcycle safety, there are examples of how deep learning can be applied to other transportation engineering problems through traffic data imputation, short-term traffic flow prediction, vehicle classification, and sustainable guideline development. Duan et al. highlight the importance of clean and complete traffic datasets, as the growing sensor infrastructure is generating ubiquitous data for analysis (*41*). While Duan et al. were interested in the imputation of missing data, Polson et al. focused on the prediction given short-term conditions (*42*). Yu et al. developed a fine-grained vehicle classification approach that applied a convolutional neural network with a joint Bayesian network to classify a vehicle similar to the methods applied for classifying a face (*43*).

Data Processing

The master database created for this analysis includes 10,099 motorcycle-involved crashes from the police-reported crash data in 2010–2014, Louisiana. This study mainly focused on at-fault motorcycle riders. Louisiana crash data contains a variable named "*Vehicle Number*," in which 1 denotes at-fault riders. The final dataset of at-fault motorcycle crashes included 6,853 crashes. This study analyzes only at-fault motorcycle rider-involved crashes because not-at-fault rider crashes involve other vehicle drivers' driving traits, which is outside the scope of the current study. The severity of crashes is recorded as five injury levels (commonly known as KABCO injury scale): fatality (K), incapacitating injury (A), non-incapacitating injury (B), possible/complaint injury (C), and no injury (O). The fatal injury category includes crashes that result in death within 30 days of the crash. The incapacitating injury prevents the injured person from normal daily work. The non-incapacitating injury includes evidence of significant injury during police reporting. The possible injury indicates complaints of pains or stresses with no physical evidence. No-injury crashes, also known as Property Damage Only (PDO), do not involve any injury.

To determine important variables, past studies have been consulted. A wider list of variables has been selected in the preliminary analysis. As a result of lack of adequate information (for example, blood alcohol content level of motorcycle riders), some of the variables were not explored for the final analysis. The final dataset contains 16 predictors of the dataset with 79 levels, of which three are numerical, and 13 are categorical (76 levels). The general findings from the numerical variables are following:

- The median value of crash hour is 4 p.m., which may be because of the peak-hour traffic volume increase at the end of the workday.
- The median number of vehicles was two, indicating that motorcycle crashes more often are not single-vehicle crashes.
- Lastly, the median rider age was 39, which may be because of the higher number of middle-aged motorcycle riders compared with younger riders. Regardless, young riders have been found to have a higher crash likelihood compared with older riders.

Table 1 lists descriptive statistics of the categorical variables used in the final analysis. As this study focuses on at-fault motorcycle rider crashes, the number of crashes and the number of at-fault motorcycle riders are same. Of the total 6853 crashes, 4.6% crashes were classified as fatal crashes, 74.1% are classified as injury crashes and the rest of these crashes are PDO. Sixty-two percent of crashes occurred more frequently on two-lane rural roadways that did not have a dividing barrier, which is similar to previous findings in the literature. Motorcycle crashes were also more likely to occur on straight level roadways (74.9%), but it is important to note that 15.6% occurred on curve level roadways, a feature which has been shown to impact on crash severity. A majority of crashes occurred during daytime and clear weather conditions, which correlates with previous findings that suggest higher speeds are more common during clear weather. Motorcycle crashes were more likely to occur over the weekend (Friday–Sunday) and over half of crashes analyzed, 52.3%, involved a rider being thrown from the motorcycle, which is likely because a motorcycle offers a rider no protection. Localities with business entities show high likelihood of motorcycle crash involvements. Lastly, nearly 27% of crashes involved "driver inattention" compared with only 3.6% of crashes that involved alcohol-impaired driving.

Table 1. Descriptive Statistics of Categorical Variables

Category	Count	%
Road condition		
No abnormalities	6374	93.0%
Animal in roadway	110	1.6%
Construction, repair	75	1.1%
Loose surface material	68	1.0%
Object in roadway	48	0.7%
Other types	178	2.6%
Roadway type		
Two-way road with no physical separation	4247	62.0%
Two-way road with a physical separation	1667	24.3%
One-way road	687	10.0%
Two-way road with a physical barrier	218	3.2%
Other types	24	0.5%
Highway type		
State hwy	2751	40.1%
City street	1504	21.9%
U.S. hwy	1092	15.9%
Parish road	887	12.9%
Interstate	615	9.0%
Toll road	4	0.1%
Locality type		
Business continuous	1931	28.2%
Business, mixed residential	1854	27.1%
Residential district	958	14.0%
Residential scattered	950	13.9%
Open country	844	12.3%
Manufacturing or industrial	145	2.1%
Other types	171	2.5%
Alignment type		
Straight-level	5131	74.9%
Curve-level	1066	15.6%
Straight-level-elevated	142	2.1%
On grade-curve	137	2.0%
Curve-level-elevated	132	1.9%
On grade-straight	122	1.8%
Other types	123	1.8%
Lighting condition		
Daylight	4733	69.1%
Dark—continuous street light	1025	15.0%
Dark—no street lights	687	10.0%
Dark—street light at intersection only	194	2.8%
Dusk	151	2.2%
Dawn	51	0.7%
Other types	12	0.2%
Weather condition		
Clear	5762	84.1%
Cloudy	820	12.0%
Rain	206	3.0%
Fog/smoke	34	0.5%
Other types	31	0.4%
Collision type		
Single vehicle	2578	37.6%
Rear end	1317	19.2%
Right angle	908	13.2%
Left turn—opposite direction	445	6.5%
Sideswipe—same direction	442	6.4%
Other types	1163	17.0%
Day of the week		
Saturday	1352	19.7%
Sunday	1114	16.3%

(continued)

Deep Learning

History

The history of artificial neural networks dates back to 1950. Rosenblatt's Perceptron algorithm was the earliest example of an artificial neural network (ANN). In the

Table 1. (continued)

Category	Count	%
Friday	1099	16.0%
Thursday	980	14.3%
Wednesday	801	11.7%
Tuesday	768	11.2%
Monday	739	10.8%
First harmful event		
Motor vehicle in transport	4260	62.2%
Ran off road right	680	9.9%
Overturn/rollover	405	5.9%
Other non-collision	309	4.5%
Crossed median/centerline	241	3.5%
Ran off road left	239	3.5%
Other types	719	10.5%
Rider ejection		
Totally ejected	3584	52.3%
Not ejected	2665	38.9%
Partially ejected	451	6.6%
Unknown	153	2.2%
Rider condition		
Normal	3968	57.9%
Inattentive	1785	26.0%
Unknown	588	8.6%
Drinking alcohol—impaired	244	3.6%
Distracted	119	1.7%
(Other)	149	2.2%
Rider severity		
Possible/complaint	2319	33.8%
Non-incapacitating/moderate	2280	33.3%
No injury	1460	21.3%
Incapacitating/severe	478	7.0%
Fatal	316	4.6%

earlier stages, Perceptron failed to approximate many nonlinear decision functions. A solution was developed later by stacking multiple layers of linear classifiers, known as multilayer perceptron, to approximate nonlinear decision functions. Because of a lack of computational power, ANN faced a slower pace of development during 1990–2000. Since late 2000, ANN has seen explosive progress as a result of parallel processing. ANN processes many parameters and approximates nonlinear functions. For a very deep and complex problem, ANN can provide a reasonable estimate. This new branch of complex problem solving is known as "deep learning" Many researchers consider learning to be deep if ANN has more than two layers. There are other interpretations of the word "deep," including: (1) deep learning can provide solutions for unlabeled data, and (2) deep means autonomous (*44–47*). Figure 1*a* shows the timeline of the scientific evolution of deep learning. This study used the 'R with H2O.ai' platform to perform the analysis (*48*).

Theory

Interested readers can consult these references for a better understanding of deep learning architecture (*43–46*). For the sake of easy interpretation, only a brief overview of deep learning concept is provided here. Consider a set of data points $\{x^{(1)}, x^{(2)}, \ldots, x^{(m)}\}$ in which each data point has many dimensions. These data can be mapped to another set of data points $\{z^{(1)}, z^{(2)}, \ldots, z^{(m)}\}$, where zs have lower dimensionality than xs. In place of using high-dimensional x, low-dimensional z can reconstruct x. To map data back and forth, a relationship can be developed:

$$z^{(i)} = W_1 x^{(i)} + b_1 \tag{1}$$

$$\tilde{x}^{(i)} = W_2 x^{(i)} + b_2 \tag{2}$$

If x^i is a two-dimensional vector, it is possible to visualize the data to find W_1, b_1 and W_2, b_2 analytically as the experiment above suggested. For high-dimensional data, the visualization is not possible. As the target is to attain $\tilde{x}^{(i)}$ to estimate $x^{(i)}$, an objective function can be used:

$$\begin{aligned} J(W_1, b_1, W_2, b_2) &= \sum_{i=1}^{m} \left(\tilde{x}^{(i)} - x^{(i)}\right)^2 \\ &= \sum_{i=1}^{m} \left(W_2 x^{(i)} + b_2 - x^{(i)}\right)^2 \\ &= \sum_{i=1}^{m} \left(W_2\left(W_1 x^{(i)} + b_1\right) + b_2 - x^{(i)}\right)^2 \end{aligned} \tag{3}$$

which can be minimized using stochastic gradient descent. This concept is also known as a linear autoencoder (as shown in Figure 1*b*).

An example of a deep network with two hidden layers (W_1, *and* W_2) is shown in Figure 1*b*. To train the neurons, an autoencoder (with parameters W_1 and W_1') can be trained. Later W_1 will be used to compute the values for the neurons for all data, which will then be used as input data to the subsequent autoencoder. This autoencoder uses the values for the neurons as inputs, and trains an autoencoder to predict those values by adding a decoding layer with parameters W_2'.

Model Development

The current study has developed a deep learning framework named "*DeepScooter*." This framework is based on five major steps: (1) finalize the dataset by selected contributing factors, (2) divide the dataset into training, validation, and text data, (3) use a basic deep learning model and check model accuracies for training, validation, and text data, (4) perform tuning to get better estimates with higher accuracies, and (5) select the final model. Figure 2 illustrates the general framework of *DeepScooter*.

A detailed step-by-step procedure is described below:

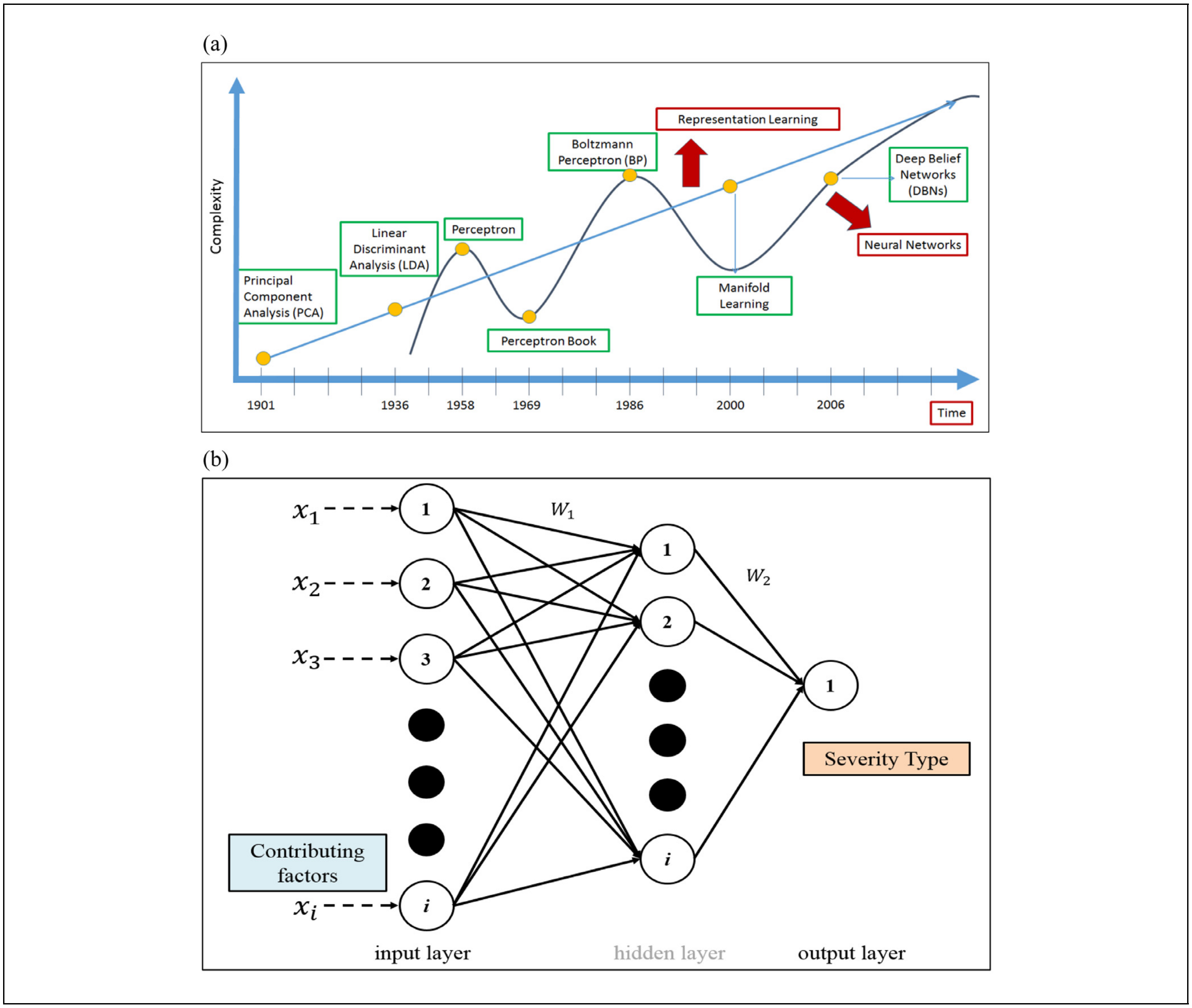

Figure 1. (*a*) Timeline of deep learning (reproduced from *44*) and (*b*) deep learning algorithmic concept.

- The dataset has been divided into three random subsets: 50% of the data was used for training, 25% was used for validation, and 25% was used for a test. All data categories (numerical, categorical, and ordinal) can be used as explanatory variables. The final matrix incorporates 6,853 rows with 79 attribute levels. The deep learning will consider training data to develop the learning framework.
- This study developed an initial model (Model 1) by using training data with one pass (known as epoch) over the training data. For Model 1, the size of hidden layer is 2.
- For Model 2, stopping criteria was applied. A larger number of epochs (100,000) was used to refine the model. For Model 2, the size of hidden layer is 32. It was found that the precision accuracy was not improving after 100 epochs.
- For the final model (Model 3), the deep learning algorithm was tuned with an adaptive learning algorithm. Two tuning parameters (ρ *and* ε) balance the global and local search efficiencies. The parameter ρ is the similarity to prior weight updates, and ε is a parameter that makes optimization work beyond local optima. Adaptive learning rate algorithm uses stochastic descent optimization. For the final model, these parameters were used:
    - Epochs: 100
    - Hidden layers: 128
    - Early stopping: enabled
    - Annealing rate: 2×10^{-6}
    - Samples = 350,000

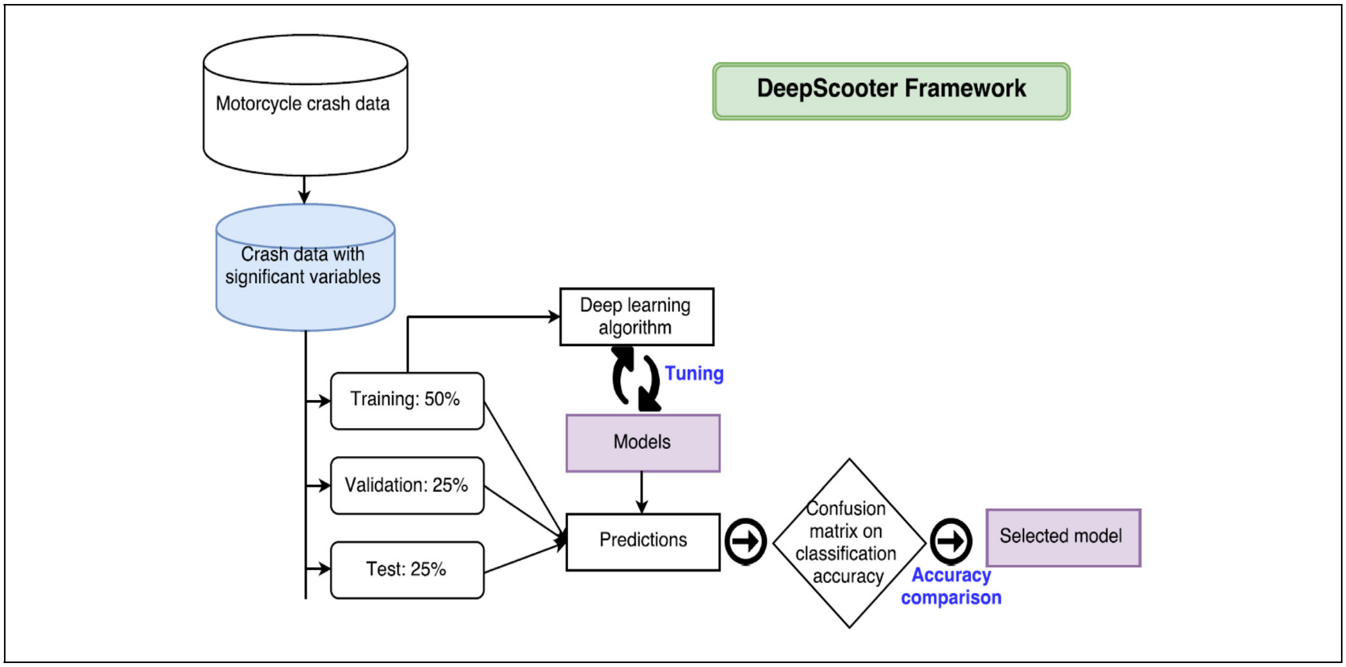

Figure 2. Framework of *DeepScooter* tool.

- Once the classifier has been trained (i.e., the parameters of the different layers of the model have been fixed), the quality of the classification outputs predicted by the model are compared against the correct "true" values stored in a labeled dataset. The confusion matrix describes the prediction accuracies and misclassification (error) rate. The outputs of the confusion matrix are reported in Tables 2 and 3.

Table 2 summarizes the results of Model 1 in confusion matrix format. For training data, the overall accuracies are 92.5%. Higher inaccuracies are found in identifying non-incapacitating injuries. For the validation dataset, the accuracy is lower than training data (around 90%). Both "*no injury*" and "*non-incapacitating injury*" show higher misclassification rates.

Table 2 also lists the results of Model 2 in confusion matrix format. For training data, the overall accuracies are improved by 3.5% (Model 1: 92.5% vs. Model 2: 96%). "*No injury*" showed a higher misclassification rate (around 8%). For the validation dataset, the accuracy was lower than training data (around 94%). "*No injury*" showed a higher misclassification rate (around 12%). For Model 1 and Model 2, results of text data are not shown.

Table 3 summarizes the results of Model 3 (training, validation, and test) in confusion matrix format. The findings include:

- For training data, the estimation accuracy of the training data was nearly 100%. Out of a sample size of 3440, the model can accurately classify 3437 crash severities.
- For validation set, the estimation accuracy was not improved from the accuracy derived in Model 2. "*Non-incapacitating injury*" showed higher misclassification.
- For the test set, the misclassification rates are 8% and 7% for Model 1, and Model 2, respectively. Model 3 results showed the lowest misclassification rate for test data. "*No injury*" showed higher misclassification for text data.

Figure 3*a* shows the overall misclassification rates of three models used for three sets of data. The training set showed that misclassification rate decreases sharply from 7% to 0%. For the validation set, the decrement is slower (from 9% to 6%). Test data also observed a slower decrease (from 8% to 6%). Overall, Model 3 performed better than the other two models. However, the validation set showed slightly a lower misclassification rate (5.79%) in Model 2 than the misclassification rate of Model 3 (6.25%).

Figure 3, *b–e*, illustrates comparison plots (Model 2 vs. Model 3) of the classification error and root mean square error (RMSE) over all epochs and samples. The values closer to zero indicate better model fit.

This study has applied two methods, multinomial logistic regression (statistical model) and support vector machine algorithm (machine learning method), on the same dataset. The highest prediction accuracy was found

Table 2. Confusion Matrix for Model 1 and Model 2

Severity level	Fatal	Incapacitating injury	Non-incapacitating injury	Possible injury	No injury	Column total	Error	Error rate	Metrics
Training (Model 1)									
Fatal	157	0	0	0	4	161	4	2.48	MSE: 0.0625
Incapacitating injury	1	245	0	0	2	248	3	1.21	RMSE: 0.250
Non-incapacitating injury	6	8	963	13	140	1130	167	14.78	logloss: 0.250
Possible injury	2	3	16	1129	9	1159	30	2.59	Mean per-class
No injury	1	2	9	41	689	742	53	7.14	Error: 0.056
Total	167	258	988	1183	844	3440	257	7.47	
Validation (Model 1)									
Fatal	69	0	0	0	3	72	3	4.17	MSE: 0.0795
Incapacitating injury	1	122	0	0	3	126	4	3.17	RMSE: 0.282
Non-incapacitating injury	4	6	520	7	75	612	92	15.03	logloss: 0.344
Possible injury	0	5	14	525	2	546	21	3.85	Mean per-class
No injury	0	3	10	27	315	355	40	11.27	Error: 0.075
Total	74	136	544	559	398	1711	160	9.35	
Training (Model 2)									
Fatal	161	0	0	0	0	161	0	0	MSE: 0.035
Incapacitating injury	1	246	0	0	1	248	2	0.81	RMSE: 0.189
Non-incapacitating injury	5	7	1085	7	26	1130	45	3.98	logloss: 0.146
Possible injury	3	4	21	1125	6	1159	34	2.93	Mean per-class
No injury	1	2	14	39	686	742	56	7.55	Error: 0.031
Total	171	259	1120	1171	719	3440	137	3.98	
Validation (Model 2)									
Fatal	71	0	1	0	0	72	1	1.39	MSE: 0.0532
Incapacitating injury	1	123	0	0	2	126	3	2.38	RMSE: 0.231
Non-incapacitating injury	4	6	584	5	13	612	28	4.58	logloss: 0.243
Possible injury	0	5	17	521	3	546	25	4.58	Mean per-class
No injury	0	2	14	26	313	355	42	11.83	Error: 0.049
Total	76	136	616	552	331	1711	99	5.79	

Table 3. Confusion Matrix for Model 3

Severity level	Fatal	Incapacitating injury	Non-incapacitating injury	Possible injury	No injury	Column total	Error	Error rate	Metrics
Training									
Fatal	161	0	0	0	0	161	0	0.00	MSE: 0.000
Incapacitating injury	0	248	0	0	0	248	0	0.00	RMSE: 0.025
Non-incapacitating injury	0	0	1128	0	2	1130	2	0.18	logloss: 0.003
Possible injury	0	0	0	1158	1	1159	1	0.09	Mean per-class
No injury	0	0	0	0	742	742	0	0.00	Error: 0.000
Total	161	248	1128	1158	745	3440	3	0.09	
Validation									
Fatal	70	1	0	0	1	72	2	2.78	MSE: 0.058
Incapacitating injury	1	124	1	0	0	126	2	1.59	RMSE: 0.241
Non-incapacitating injury	0	2	315	13	25	355	30	8.45	logloss: 0.544
Possible injury	4	6	13	582	7	612	33	5.39	Mean per-class
No injury	1	5	15	12	513	546	40	7.33	Error: 0.053
Total	76	138	344	607	546	1711	107	6.25	
Test									
Fatal	79	0	3	1	0	83	4	4.82	MSE: 0.058
Incapacitating injury	0	99	2	1	2	104	5	4.81	RMSE: 0.241
Non-incapacitating injury	3	4	511	8	12	538	27	5.02	logloss: 0.515
Possible injury	1	1	19	583	10	614	31	5.05	Mean per-class
No injury	1	3	9	29	321	363	42	11.57	Error: 0.063
Total	84	107	544	622	345	1702	109	6.40	

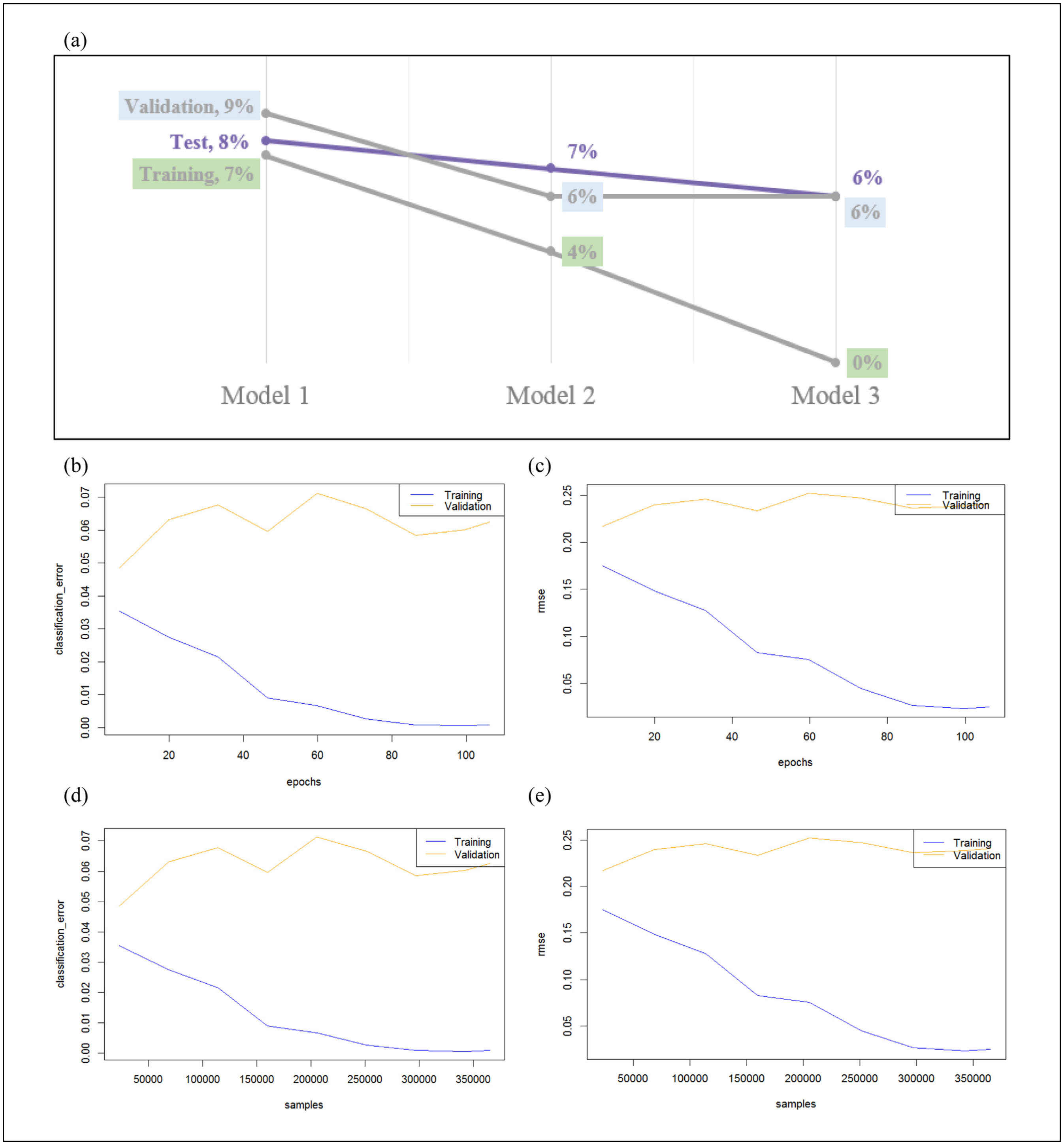

Figure 3. (*a*) Misclassification (error) rate in three models; (*b*) Model 2 classification error; (*c*) Model 2 RMSE; (*d*) Model 3 classification error; and (*e*) Model 3 RMSE.

as 78%, which is far below the attained accuracy from the deep learning method. The main contribution of this paper is the development of a deep learning framework to perform severity analysis of motorcycle crashes. Although the present study is focused on the improvement of the classification accuracy, a sound understanding of the explanatory variables would be beneficial. The heat chart shows (see Table 4) the conditional probability of significant variables from final deep learning model (Model 3). The color intensity (from red to white) varies along each

Table 4. Conditional Probabilities of Top Ten Attributes

Rider ejected	89.9	84.3	67.2	48.4	16.5
Two-way with no physical separation	62.3	62.3	63.2	61.6	60.4
State highway	48.4	29.7	38.7	43.9	38.1
Single vehicle	42.1	32.8	45.5	39.4	32.2
Two-way with physical separation	25.6	24.1	24.3	24	24.7
Business and mixed residential	24.1	28.2	26.4	27.7	35.5
Residential scattered	23.1	27.4	26.4	27.2	27.2
Saturday	22.5	21.3	18.4	20.9	18.8
Curve alignment	20.6	17.4	17.6	16.2	9.7
Sunday	19.9	19	17.6	14.5	15.2
	K	A	B	C	O

row. For example, around 90% of cases in fatal motorcycle crashes involve rider ejection. Fatal crashes are highly associated with rider ejection, a two-way road with no physical separation, single vehicle, curve aligned roadways, and weekends. Residential areas were not highly skewed in crash fatalities; this may be because of lower posted speeds in residential neighborhoods. Two-way roadways (both divided and undivided) are over-representative in severe and fatal crashes.

Conclusion

Many studies on motorcycle crash data have been conducted to understand the contributing factors that influence the severity of crashes. In 2014, the U.S. experienced 976 motorcycle crash-related fatalities. This statistics was more than twice the number of motorcycle rider fatalities that occurred in 1997. This unacceptably high number of motorcycle crashes calls for research to be conducted with additional resources and in newer directions. The 2010–2014 at-fault motorcycle rider crash data for Louisiana were used in this study. The findings include:

- The descriptive statistics show that motorcycle crashes generally feature high proportions in weekends, two-lane rural roads with no physical barriers, rider ejection, daylight, and clear weather conditions.
- The study confirms that, in modeling crash severity, the developed deep learning framework *DeepScooter* can estimate accurately up to 100%. The accuracy rate for test data ranges from 92% to 94%. The framework has sufficient reproducibility for use with a larger set of motorcycle crash data. For example, it can work as a suitable framework for identifying significant factors from FHWA Motorcycle Crash Causation Study (MCCS) (*49*).
- The conditional probability chart from *DeepScooter* shows that fatal crashes are more likely to be associated with rider ejection, a two-way road with no physical separation, single vehicle, curve aligned roadways, and weekends. Residential areas were not highly skewed in fatal crashes, as a result of lower posted speeds. In addition, younger riders are the vulnerable group in fatal motorcycle crashes.

The advantage of using a deep learning tool is its efficiency and high prediction accuracy. As this method does not require holding any statistical assumption, there is no consequence of biased results because of the violation of assumptions. However, this method has several limitations. One limitation of deep learning models is their low explanatory command because of the black box approach. In transportation safety research, interpretability is considered as one of the major constraints in adapting sophisticated deep learning models in real life. The current developed framework has the flexibility of reproduction that can be used by other researchers. It is anticipated that the *DeepScooter* framework and the findings will provide noteworthy contributions to the reduction of motorcycle crashes and crash-involved severities.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Subasish Das, Anandi Dutta; data collection: Subasish Das; analysis and interpretation of results: Subasish Das, Anandi Dutta; draft manuscript preparation: Subasish Das, Anandi Dutta, Lisa Minjares-Kyle, Karen Dixon, George Gillette. All authors reviewed the results and approved the final version of the manuscript.

References

1. National Highway Traffic Safety Administration. *Motorcycles*. https://www.nhtsa.gov/road-safety/motorcycles. Accessed July 3, 2017.
2. Louisiana Crash Data Reports. http://datareports.lsu.edu/reports.aspx?p = ci&yr = 2016&rpt = H1. Accessed August 1, 2017.
3. Lord, D., and F. Mannering. The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives. *Transportation Research Part A: Policy and Practice*, Vol. 44, No. 5, 2010, pp. 291–305.
4. Savolainen, P., F. Mannering, D. Lord, and M. Quddus. The Statistical Analysis of Highway Crash-Injury Severities: A Review and Assessment of Methodological Alternatives. *Accident Analysis & Prevention*, Vol. 43, No. 5, 2011, pp. 1666–1676.
5. Mannering, F., and C. Bhat. Analytic Methods in Accident Research: Methodological Frontier and Future Directions. *Analytic Methods in Accident Research*, Vol. 1, 2014, pp. 1–22.
6. Mannering, F., V. Sarkar, and C. Bhat. Unobserved Heterogeneity and the Statistical Analysis of Highway Accident Data. *Analytic Methods in Accident Research*, Vol. 11, 2016, pp. 1–16.
7. Savolainen, P., and F. Mannering. Probalistic Models of Motorcyclists' Injury Severities in Single-and Multi-Vehicle Crashes. *Accident Analysis and Prevention*, Vol. 39, 2007, pp. 955–963.
8. Lapparent, M. Empirical Bayesian Analysis of Accident Severity for Motorcyclists in Large French Urban Areas. *Accident Analysis and Prevention*, Vol. 38, 2006, pp. 260–268.
9. Schneider, W. H., P. T. Savolainen, D. V. Boxel, and R. Beverly. Examination of Factors Determining Fault in Two-Vehicle Motorcycle Crashes. *Accident Analysis and Prevention*, Vol. 45, 2012, pp. 669–676.
10. Turner, P., L. Higgins, and S. Geedipally. *Development of Statewide Motorcycle Safety Plan for Texas*. Technical Report, FHWA/TX-12-0-6712-1, 2013.
11. Quddus, M. A., R. B. Noland, and H. C. Chin. An Analysis of Motorcycle Injury and Vehicle Damage Severity Using Ordered Probit Models. *Journal of Safety Research*, Vol. 33, 2002, pp. 445–462.
12. Cunto, F. J. C., and S. Ferreira. An Analysis of the Injury Severity of Motorcycle Crashes in Brazil Using Mixed Ordered Response Models. *Journal of Transportation Safety & Security*, Vol. 9, No. S1, 2017, pp. 33–46.
13. Shaheed, M., and K. Gkritza. A Latent Class Analysis of Single-Vehicle Motorcycle Crash Severity Outcomes. *Analytic Methods in Accident Research*, Vol. 2, 2014, pp. 30–38.
14. Shaheed, M., K. Gkritza, W. Zhang, and Z. Hans. A Mixed Logit Analysis of Two-Vehicle Crash Severities Involving a Motorcycle. *Accident Analysis and Prevention*, Vol. 61, 2013, pp. 119–128.
15. Shankar, V., and F. Mannering. An Exploratory Multinomial Logit Analysis of Single-Vehicle Motorcycle Accident Severity. *Journal of Safety Research*, Vol. 27, No. 3, 1996, pp. 183–194.
16. Geedipally, S. R., P. T. Turner, and S. Patil. Analysis of Motorcycle Crashes in Texas with Multinomial Logit Model. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2265: 62–69.
17. Chung, Y., T. J. Song, and B. J. Yoon. Injury Severity in Delivery-Motorcycle to Vehicle Crashes in the Seoul Metropolitan Area. *Accident Analysis and Prevention*, Vol. 62, 2014, pp. 79–86.
18. Kim, K., J. Boski, and E. Yamshita. Typology of Motorcycle Crashes: Rider Characteristics, Environmental Factors and Spatial Patterns. *Transportation Research Record: Journal of the Transportation Research Board*, 2002. 1818: 47–53.
19. Schneider, W. H., and P. T. Savolainen. Comparison of Severity of Motorcyclist Injury by Crash Types. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2265: 70–80.
20. Blackman, R. A., and N. L. Haworth. Comparison of Moped, Scooter and Motorcycle Crash Risk and Crash Severity. *Accident Analysis and Prevention*, Vol. 57, 2013, pp. 1–9.
21. Pai, C., and W. Saleh. Exploring Motorcyclist Injury Severity Resulting from Various Crash Configurations at T-Junctions in the United Kingdom: An Application of the Ordered Probit Models. *Traffic Injury Prevention*, Vol. 8, 2007, pp. 62–68.
22. Rifaat, S. M., R. Tay, and A. de Barros. Severity of Motorcycle Crashes in Calgary. *Accident Analysis and Prevention*, Vol. 49, 2012, pp. 44–49.
23. Savolainen, P. T., F. L. Mannering, D. Lord, and M. A. Quddus. The Statistical Analysis of Highway Crash-Injury Severities: A Review and Assessment of Methodological Alternatives. *Accident Analysis and Prevention*, Vol. 43, 2011, pp. 1666–1676.
24. Haque, M., H. C. Chin, and A. K. Debnath. An Investigation on Multi-Vehicle Motorcycle Crashes Using Log-Linear Models. *Safety Science*, Vol. 50, 2012, pp. 352–362.

25. Allen, T., S. Newstead, M. G. Lenne, R. McClure, P. Hillard, M. Symmons, and L. Day. Contributing Factors to Motorcycle Injury Crashes in Victoria, Australia. *Transportation Research Part F: Traffic Psychology and Behavior*, Vol. 45, 2017, pp. 157–168.
26. Chen, C., G. Zhang, X. Liu, Y. Ci, H. Huang, J. Ma, Y. Chen, and H. Guan. Driver Injury Severity Outcome Analysis in Rural Interstate Highway Crashes: A Two-Level Bayesian Logistic Regression Interpretation. *Accident Analysis and Prevention*, Vol. 97, 2016, pp. 69–78.
27. Breiman, L. Statistical Modeling: The Two Cultures (With Comments and a Rejoinder by the Author). *Statistical Science*, Vol. 16, No. 3, 2001, pp. 199–231.
28. Das, S., and X. Sun. Association Knowledge for Fatal Run-off-Road Crashes by Multiple Correspondence Analysis. *IATSS Research*, Vol. 39, No. 2, 2016, pp. 146–155.
29. Chen, C., and Y. Xie. Machine Learning for Recognizing Driving Patterns of Drivers of Large Commercial Trucks. *Transportation Research Record: Journal of the Transportation Research Board*, 2015. 2517: 18–27.
30. Das, S., A. Dutta, R. Avelar, K. Dixon, X. Sun, and M. Jalayer. Supervised Association Rules Mining on Pedestrian Crashes in Urban Areas: Identifying Patterns for Appropriate Countermeasures. *International Journal of Urban Sciences*, 2018. https://doi.org/10.1080/12265934.2018.1431146
31. Sun, J., J. Sun, and P. Chen. Use of Support Vector Machine Models for Real-Time Prediction of Crash Risk on Urban Expressways. *Transportation Research Record: Journal of the Transportation Research Board*, 2014. 2432: 91–98.
32. Iranitalab, A., and A. Khattak. Comparison of Four Statistical and Machine Learning Methods for Crash Severity Prediction. *Accident Analysis and Prevention*, Vol. 108, 2017, pp. 27–36.
33. Das, S., L. Minjares-Kyle, R. Avelar, K. Dixon, and B. Bommanayakanahalli. Improper Passing-Related Crashes on Rural Roadways: Using Association Rules Negative Binomial Miner. Presented at 96th Annual Meeting of the Transportation Research Board, Washington, D.C., 2017.
34. Dong, N., H. Huang, and L. Zheng. Support Vector Machine in Crash Prediction at the Level of Traffic Analysis Zones: Assessing the Spatial Proximity Effects. *Accident Analysis & Prevention*, Vol. 82, 2015, pp. 192–198.
35. Das, S., and X. Sun. Factor Association with Multiple Correspondence Analysis in Vehicle-Pedestrian Crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2015. 2519: 95–103.
36. Ghazan Khan, G., A. Bill, and D. Noyce. Exploring the Feasibility of Classification Trees Versus Ordinal Discrete Choice Models for Analyzing Crash Severity. *Transportation Research Part C: Emerging Technologies*, Vol. 50, 2015, pp. 86–96.
37. Das, S., R. Avelar, K. Dixon, and X. Sun. Investigation on the Wrong Way Driving Crash Patterns Using Multiple Correspondence Analysis. *Accident Analysis and Prevention*, Vol. 111, 2018, pp. 43–55.
38. Mishra, D., eds. Information and Communication Technology. *Advances in Intelligent Systems and Computing*, Vol. 625, 2018.
39. Das, S., A. Mudgal, A. Dutta, and S. Geedipally. Vehicle Consumer Complaint Reports Involving Severe Incidents: Mining Large Contingency Tables. *Transportation Research Record: Journal of the Transportation Research Board*, 2018.
40. Das, S., A. Dutta, M. Jalayer, A. Bibeka, and L. Wu. Factors Influencing the Patterns of Wrong Way Driving Crashes on Freeway Exit Ramps and Median Crossovers: Exploration Using 'Eclat' Association Rules to Promote Safety. *International Journal of Transportation Science and Technology*, 2018.
41. Duan, Y., Y. Lv, Y. L. Liu, and F. Y. Wang. An Efficient Realization of Deep Learning for Traffic Data Imputation. *Transportation Research Part C: Emerging Technologies*, Vol. 72, 2016, pp. 168–181.
42. Polson, N. G., and V. O. Sokolov. Deep Learning for Short-Term Traffic Flow Prediction. *Transportation Research Part C: Emerging Technologies*, Vol. 79, 2017, pp. 1–17.
43. Yu, S., Y. Wu, W. Li, Z. Song, and W. Zeng. A Model for Fine-Grained Vehicle Classification Based on Deep Learning. *Neurocomputing*, Vol. 257, 2017, pp. 97–103.
44. Rosenblatt, F. *The Perceptron: A Perceiving and Recognizing Automaton*. Report No. 85-460-1. Cornell Aeronautical Laboratory, Ithaca, NY, 1957.
45. Zhong, G., L. Wang, X. Ling, and J. Dong. An Overview on Data Representation Learning: From Traditional Feature Learning to Recent Deep Learning. *The Journal of Finance and Data Science*, Vol. 2, 2016, pp. 265–278.
46. Le, Q. A Tutorial on Deep Learning Part 2: Autoencoders, Convolution Neural Networks and Recurrent Neural Networks. http://robotics.standford.edu/~quocle/tutorial2.pdf. Accessed July 3, 2017.
47. Goodfellow, I., Y. Bengio, and A. Courville. Deep Learning. The MIT Press, Cambridge, MA, 2016.
48. H2O.ai Team. *H_2o: R Interface for H2O*. R Package Version 3.1.0, 2015.
49. Nazemetz, J., F. Bents, J. Perry, C. Thor, and C. Tan. *Motorcycle Crash Causation Study: Final Report*. FHWA, 2016.

The Standing Committee on Motorcycles and Mopeds (ANF30) peer-reviewed this paper (18-04840).