







# Using Data-Mining Techniques for the Prediction of the Severity of Road Crashes in Cartagena, Colombia

Holman Ospina-Mateus<sup>1,2</sup> ,  
Leonardo Augusto Quintana Jiménez<sup>2</sup>,  
Francisco José López-Valdés<sup>3</sup> , Natalie Morales-Londoño<sup>4</sup> ,  
and Katherinne Salas-Navarro<sup>5</sup> 

<sup>1</sup> Department of Industrial Engineering, Universidad Tecnológica de Bolívar,  
Cartagena, Colombia

Hospina@utb.edu.co

<sup>2</sup> Department of Industrial Engineering, Pontificia Universidad Javeriana,  
Carrera 7 # 40-62, Bogotá, Colombia

lquin@javeriana.edu.co

<sup>3</sup> Instituto de Investigacion Tecnológica (IIT), ICAI Engineering School,  
Universidad Pontificia de Comillas, c/Alberto Aguilera 25, 28250 Madrid, Spain

fjlvaldes@comillas.edu

<sup>4</sup> Projects Construction and Effective Management - PCEM SAS, Cartagena,  
Colombia

nmorales@pcem.com.co

<sup>5</sup> Department of Industrial Management, Agroindustry and Operations,  
Universidad de la Costa, Barranquilla, Colombia

ksalas2@cuc.edu.co

**Abstract.** *Objective:* Analyze the road crashes in Cartagena (Colombia) and the factors associated with the collision and severity. The aim is to establish a set of rules for defining countermeasures to improve road safety. *Methods:* Data mining and machine learning techniques were used in 7894 traffic accidents from 2016 to 2017. The severity was determined between low (84%) and high (16%). Five classification algorithms to predict the accident severity were applied with WEKA Software (Waikato Environment for Knowledge Analysis). Including Decision Tree (DT-J48), Rule Induction (PART), Support Vector Machines (SVMs), Naïve Bayes (NB), and Multilayer Perceptron (MLP). The effectiveness of each algorithm was implemented using cross-validation with 10-fold. Decision rules were defined from the results of the different methods. *Results:* The methods applied are consistent and similar in the overall results of precision, accuracy, recall, and area under the ROC curve. *Conclusions:* 12 decision rules were defined based on the methods applied. The rules defined show motorcyclists, cyclists, including pedestrians, as the most vulnerable road users. Men and women motorcyclists between 20–39 years are prone in accidents with high severity. When a motorcycle or cyclist is not involved in the accident, the probable severity is low.

**Keywords:** Road crashes · Prediction · Data mining · Severity

## 1 Introduction

According to the World Health Organization WHO [1], approximately 1.4 million people died in traffic accidents in 2016, and it is estimated that more than 50 million people suffered severe injuries. Traffic accidents in 2016 were the eighth cause of death, and the main causes of death for people between the ages of 15 and 29 years old. In addition, such accidents affect pedestrians, cyclists, and motorcyclists. Half of the road deaths occur among motorcyclists (28%), pedestrians (23%), and cyclists (3%). Mortality rates in low-income countries are 3 times higher than in high-income countries. Although only 1% of motor vehicles are in emerging countries, 13% of deaths occur in these nations [1]. Colombia in 2016 obtained a rate of 18.5 fatalities per 100,000 inhabitants. This figure is close to the global average (18.2) and the average middle-income countries (18.8). Between 2012–2018 in relation to the road users involved, the motorcyclists correspond to 50% of the victims. Pedestrians with 24%. The users of vehicles with 17%. The cyclists with 5% of accidents. The objective of this research was to analyze the road crashes in Cartagena (Colombia), and the factors associated with the collision and severity. Cartagena has 1.2 million inhabitants and more than 120,000 motor vehicles. Cartagena in the last two years (2017–2018) has remained in the top positions for fatal accidents in capital cities. In the last 8 years, Cartagena has been considered the fifth most dangerous city in road safety after Medellin, Cali, Bogota, and Barranquilla.

## 2 Method

The method is based on official information from the control entities, which allowed for the application of data mining and machine learning techniques. The method is based on the application of Decision Tree, Rule Induction, Support Vector Machines, Naïve Bayes, and Multilayer Perceptron with WEKA software. The decision tree constructs classification models in the form of trees. Rule Induction is an iterative process that follows a divide-and-conquer approach. Naïve Bayes is a classification algorithm based on Bayesian theorem. Multilayer Perceptron creates a feed-forward artificial neural network. Support vector machines are learning algorithms for classification and regression analysis. In the prediction of road accidents, data mining techniques have been implemented, such as: regressive models [2], neural networks [3], artificial intelligence [4], decision trees [5], Bayesian networks [6], SVM [7], and combined methods. The aim is to establish a set of decision rules for defining countermeasures to improve road safety. The effectiveness of each algorithm was implemented using cross-validation with 10-fold. The methodological process of this investigation is presented in four steps: (a) Pre-processing of accident dataset; (b) application of data mining techniques through Weka Software; (c) analysis of results and metrics; (d) definition of decision rules and analysis of associated factors.

### 2.1 Data Sources (Include Sample Size)

The registration of traffic accidents corresponds to the Cartagena database from January 2016 to December 2017. The dataset corresponds to the reports by the Administrative Department of Traffic and Transport (DAT). In total, 10,053 traffic accidents were

reported by agents and police. The data records information about temporality, road users, gender and age. For the pre-processing of the data set, 27 categorical variables were defined (see Table 1). The variables were classified into four categories: (1) road actors involved in the crash, (2) individuals involved, (3) weather conditions and timing, (4) accident characteristics. The levels of injury severity were determined as low-level-of-injury (material damages, minor, non-incapacitating injury) or high-level-of-injury (injured victims and fatality).

**Table 1.** Definitions and values of traffic accident variables from four categories in the dataset.

Categories	Description	Label	Definition [Code] (Value)
Road actors	Number of cars	NLV	[0] Without vehicle; [1] one vehicle and [2] more than two vehicles
	Number of buses	NB	[0] Without vehicle; [1] one vehicle and [2] more than two vehicles
	Number of motorcyclists	NM	[0] Without motorcycle, [1] One motorcycle and [2] more than two motorcycles
	Number of cyclists	NC	[0] Without bicycles; [1] one bicycle; and [2] more than two bicycles
	Number of heavy vehicles	NHV	[0] Without vehicle; [1] one vehicle and [2] more than two vehicles
	Number of others (pedestrians, Non-register)	NOT	[0] Without; [1] one road actors; and [2] more than two road actors
	Relationship of road actors	RAV	[1] Solo motorcycle; [2] motorcycle-car; [3] motorcycle-heavy-vehicle; [4] motorcycle-bicycle; [5] Solo Cars; [6] Car-heavy-vehicle; [7] Solo-heavy-vehicle; [8] Solo bicycle; [9] Car-bicycle; [10] heavy-vehicle-bicycle; [11] others non-register
	Relationship type of service	RTS	[1] Particular; [2] particular-public; [3] public; [4] official-particular; [5] official-public; [6] official and [7] unidentified
Individuals	Gender of driver cars	GC	[1] Male; [2] female; [3] UNISEX; and [4] not defined
	Gender of the motorcyclist	GM	
	Gender of the bus driver	GB	
	Gender of the cyclists	GC	
	Gender of drivers of heavy vehicles	GHV	
	Driver's age	YC	
	Motorcyclist's age	YM	
	Bus driver's age	YB	
	Cyclist's age	YC	
	Heavy vehicle driver's age	YHV	
	Victim's age injured	YV	

(continued)

**Table 1.** (continued)

Categories	Description	Label	Definition [Code] (Value)
Weather/timing	Month	MON	[1] January–March; [2] April–June; [3] July–September; and [4] October–December
	Day of the week	DW	[1] During the week and [2] weekend
	Type of day	TD	[1] Holiday and [2] working day
	Type of commercial day	TCD	[1] Normal, [2] biweekly salary and [3] monthly salary
	Weather conditions (Precipitation, unit: mm)	WC	[1] (0–15); [2] (15–30); [3] (30–60); [4] (60–90); and [5] (< 90)
Accident characteristic	Accident class	CLA	[1] Collision; [2] falling from the vehicle; [3] run over; [4] dump Vehicle; [5] fire and [6] Others
	Injured or dead victim	VID	[0] Not recognized; [1] one injured/dead; and [2] two injured/dead or more
	Severity	SEV	[1] Low-level-of-injury and [2] high-level-of-injury

### 3 Results

The software WEKA contributed to the purification of the information by means of the Remove Misclassified method and duplicated instances. The dataset was reduced to 7,894 instances. Table 2 summarizes the variables used and their relationship to the severity. The analyzed data was divided into 16% of low, and 84% high-level-of-injury. In the descriptive analysis of the data, the greatest number of accidents occurs between cars (45%), followed by cars-heavy vehicles (28%), and finally between cars-motorcycles (14%). Accidents between private and public vehicles are prevalent (44%). Accidents involving motorcyclists (76%) and bicycles (88%) are more severe. The most frequent type of crash is the collision (99%), and the most severe are being run over (100%) and falling off the vehicle (93%).

**Table 2.** Statistics of road crashes for Cartagena in 2016–2017.

Label	Definition [Code] (Value)	% Total	Severity [1]	Severity [2]
NLV	[0] Without vehicle	12%	57%	43%
	[1] One vehicle	49%	81%	19%
	[2] More than two vehicles	39%	96%	4%
NB	[0] Without vehicle	77%	82%	18%
	[1] One vehicle	21%	90%	10%
	[2] More than two vehicles	2%	95%	5%
NM	[0] Without motorcycle	81%	94%	6%
	[1] One motorcycle	17%	42%	58%
	[2] More than two motorcycles	2%	12%	88%
NC	[0] Without bicycles	99%	84%	16%
	[1] One bicycle	1%	17%	83%

(continued)

**Table 2.** (continued)

Label	Definition [Code] (Value)	% Total	Severity [1]	Severity [2]
NHV	[0] Without vehicle	85%	82%	18%
	[1] One vehicle	14%	92%	8%
	[2] More than two vehicles	1%	97%	3%
NOT	[0] Without	91%	85%	15%
	[1] One road actors	9%	77%	23%
	[2] More than two road actors	0%	71%	29%
RAV	[1] Solo motorcycle	3%	14%	86%
	[2] Motorcycle-car	13%	47%	53%
	[3] Motorcycle-heavy-vehicle	2%	30%	70%
	[4] Motorcycle-bicycle	0%	4%	96%
	[5] Solo Cars	45%	95%	5%
	[6] Car-heavy-vehicle	28%	96%	4%
	[7] Solo-heavy-vehicle	6%	93%	7%
	[8] Solo bicycle	0%	0%	100%
	[9] Car-bicycle	0%	27%	73%
	[10] Heavy-vehicle-bicycle	0%	19%	81%
	[11] Others non-register	0%	79%	21%
RTS	[1] Particular	37%	75%	25%
	[2] Particular-public	44%	88%	12%
	[3] Public	17%	92%	8%
	[4] Official-particular	1%	83%	17%
	[5] Official-public	1%	85%	15%
	[6] Official	0%	58%	42%
	[7] Unidentified	0%	85%	15%
GC	[1] Male	74%	87%	13%
	[2] Female	6%	88%	12%
	[3] UNISEX	7%	97%	3%
	[4] Not defined	13%	58%	42%
GM	[1] Male	18%	39%	61%
	[2] Female	0%	18%	82%
	[3] UNISEX	0%	0%	100%
	[4] Not defined	81%	94%	6%
GB	[1] Male	22%	90%	10%
	[2] Female	0%	0%	0%
	[4] Not defined	77%	82%	18%
GC	[1] Male	1%	18%	82%
	[2] Female	0%	10%	90%
	[4] Not defined	99%	84%	16%
GHV	[1] Male	15%	93%	7%
	[2] Female	0%	100%	0%
	[3] UNISEX	0%	100%	0%
	[4] Not defined	85%	82%	18%

(continued)

**Table 2.** (continued)

Label	Definition [Code] (Value)	% Total	Severity [1]	Severity [2]
YC	[0] Without registration	14%	60%	40%
	[1] (0–17)	1%	96%	4%
	[2] (18–35)	26%	85%	15%
	[3] (36–45)	28%	90%	10%
	[4] (46–55)	20%	89%	11%
	[5] (56–69)	9%	87%	13%
	[6] (<70)	1%	85%	15%
YM	[0] Without registration	82%	94%	6%
	[1] (0–17)	0%	16%	84%
	[2] (18–35)	11%	37%	63%
	[3] (36–45)	4%	42%	58%
	[4] (46–55)	2%	45%	55%
	[5] (56–69)	1%	37%	63%
	[6] (<70)	0%	50%	50%
YB	[0] Without registration	78%	82%	18%
	[1] (0–17)	0%	100%	0%
	[2] (18–35)	10%	89%	11%
	[3] (36–45)	7%	90%	10%
	[4] (46–55)	4%	92%	8%
	[5] (56–69)	2%	91%	9%
	[6] (<70)	0%	100%	0%
YC	[0] Without registration	99%	84%	16%
	[1] (0–17)	0%	0%	100%
	[2] (18–35)	0%	19%	81%
	[3] (36–45)	0%	10%	90%
	[4] (46–55)	0%	9%	91%
	[5] (56–69)	0%	33%	67%
	[6] (<70)	0%	0%	100%
YHV	[0] Without registration	85%	82%	18%
	[1] (0–17)	0%	50%	50%
	[2] (18–35)	4%	92%	8%
	[3] (36–45)	5%	94%	6%
	[4] (46–55)	4%	94%	6%
	[5] (56–69)	2%	93%	7%
	[6] (<70)	0%	94%	6%
YV	[0] Without registration	99%	84%	16%
	[1] (0–17)	0%	0%	100%
	[2] (18–35)	0%	0%	100%
	[3] (36–45)	0%	0%	100%
	[4] (46–55)	0%	0%	100%
	[5] (56–69)	0%	0%	100%
	[6] (<70)	0%	0%	100%

(continued)

**Table 2.** (continued)

Label	Definition [Code] (Value)	% Total	Severity [1]	Severity [2]
MON	[1] January–March	22%	89%	11%
	[2] April–June	28%	82%	18%
	[3] July–September	30%	82%	18%
	[4] October–December	20%	84%	16%
DW	[1] During the week	74%	85%	15%
	[2] weekend	26%	81%	19%
TD	[1] Holiday	4%	84%	16%
	[2] working day	96%	84%	16%
TCD	[1] Normal	52%	84%	16%
	[2] Biweekly salary	16%	82%	18%
	[3] Monthly salary	33%	84%	16%
WC	[1] (0–15)	37%	86%	14%
	[2] (15–30)	9%	81%	19%
	[3] (30–60)	12%	82%	18%
	[4] (60–90)	8%	82%	18%
	[5] (<90)	33%	83%	17%
CLA	[1] Collision	99%	84%	16%
	[2] Falling from the vehicle	0%	7%	93%
	[3] Run over	0%	0%	100%
	[4] Dump Vehicle	0%	73%	27%
	[5] Fire	0%	100%	0%
	[6] Others	0%	89%	11%
VID	[0] Not recognized	66%	100%	0%
	[1] One injured/dead; and [2] two injured/dead or more	17%	0%	100%
	[2] Two injured/dead or more	18%	0%	100%

After descriptive statistical analysis, an inferential and correlational analysis of the variables proposed in the prediction of severity was proposed. The proposed analyzes were the Spearman and Friedman ANOVA correlation. Table 3 summarizes the results.

**Table 3.** Statistical analysis of the results for the prediction of the severity.

	Spearman RC		Friedman ANOVA		
	R2	P-Value	Mean square	F-Ratio	P-Value
NLV	-0.3123	0	110.886	1384.45	0
NB	-0.094	0	74.247	927	0
NM	0.5924	0	197.696	2468.29	0
NC	0.1837	0	26.7644	334.16	0
NMH	-0.1053	0	7.59372	94.81	0
NOT	0.0608	0	0.000874166	0.01	0.9168
RAV	-0.4131	0	2.15764	26.94	0
RTS	-0.1788	0	0.546751	6.83	0.009
YD	-0.1843	0	0.674626	8.42	0.0037

(continued)

**Table 3.** (continued)

	Spearman RC		Friedman ANOVA		
	R2	P-Value	Mean square	F-Ratio	P-Value
YM	0.5697	0	0.315334	3.94	0.0472
YB	-0.0946	0	0.100419	1.25	0.2628
YHV	-0.1042	0	0.015438	0.19	0.6606
YC	0.1765	0	0.0846726	1.06	0.3039
YVC	0.1989	0	13.7402	171.55	0
GC	0.173	0	0.479928	5.99	0.0144
GM	-0.5852	0	2.69317	33.63	0
GB	0.0942	0	0.507256	6.33	0.0118
GHV	0.1045	0	0.104282	1.3	0.2538
GC	-0.1801	0	0.286762	3.58	0.0585
MON	0.048	0	1.91877	23.96	0
DW	0.0451	0.0001	0.0695317	0.87	0.3515
TD	-0.0006	0.9591	0.18272	2.28	0.1309
TCD	0.0104	0.3555	0.142195	1.78	0.1827
WC	0.0369	0.0011	0.264161	3.3	0.0494
CLA	0.1109	0	0.797542	9.96	0.0016

The variables NOT, YB, YHV, YC, GHV, DW, TD, and TCD do not evidence a significant statistical association with the direct prediction of the severity of the accident ( $p$ -value  $> 0.05$ ). The variables with a statistical association on the prediction will be represented in the definition of the rules ( $p$ -value  $< 0.05$ ).

After the data pre-processing, the selected data mining techniques are applied and parameterized (See Table 4) with the 10-fold cross-validation technique. The results are compared with the metrics: precision, accuracy, recall, and area under the ROC curve (See Table 5). The results show a high consistency and similarity in the prediction metrics in the applied techniques.

**Table 4.** Parameters settings for all classifiers.

Method	Parametrization
DT-J48	weka.classifiers.trees.J48 -C 0.25 -M 2
PART	weka.classifiers.rules.PART -M 2 -C 0.25 -Q 1
NB	weka.classifiers.bayes.NaiveBayes;Use kernel estimator F, Use supervised; discretization
SVM	weka.classifiers.functions.SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
MLP	weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a



**Table 5.** Comparison of all prediction results and performance metrics based on 10-fold cross-validation

Method	Classified instances		Accuracy	Precision	Recall	ROC Area	Time (seg)
	Correctly	Incorrectly					
DT-J48	7059	835	89.4%	89.1%	89.4%	83.0%	0.11
PART	6972	922	88.3%	87.6%	86.9%	86.4%	0.21
NB	6969	925	88.3%	89.5%	88.3%	86.8%	0.02
SVM	7002	892	88.7%	89.7%	88.7%	83.1%	46
MLP	6976	915	88.4%	88.2%	88.4%	86.2%	386

From the best results of each of the techniques, 12 priority decision rules for road safety were defined (See Table 6).

**Table 6.** List of rules identified with all methods applied.

Number	Identified rule (decoded)
1	If (NM $\geq$ 1) then Severity = High
2	If (NM: 0 & NC:0) then Severity = Low
3	If (NM > 1 & GM: Male, NVL:1, MON: July–December) then Severity = High
4	If (NM > 1 & GM: Male, NVL:1, MON: April–June, WC > 30 mm) then Severity = Low
5	If (NC > 1) then Severity = High
6	If (NM > 1 & GM: Male, NVL:1, MON: 2, WC > 90 mm, TD: Working day, YM: 20–39 years, TD: Weekend, GC: Male, TDC: biweekly and monthly salary) then Severity = High
7	If (GM: Male or Female, YM: 20–39 years) then Severity = High
8	If (NM $\geq$ 1, NVL:1, GM: Males, YM: 40–79 years) then Severity = Low
9	If (CLA: Collision, NM = 0, NC:0) then Severity = Low
10	If (CLA: Collision, NLV = 2, WC < 30 mm) then Severity = Low
11	If (CLA: Runover) then Severity = High
12	If (RAV: motorcycle-heavy-vehicle, WC > 60 mm) then Severity = High

## 4 Discussion

The results show cyclists and motorcyclists as the most vulnerable road users. Motorcyclists men and women between 20 and 39 years are predictive of high severity accidents. When there are no motorcycles or cyclists involved in the accident, the probable severity is low. Also, the collision between two motorcycles is considered of high severity. If the crash is a runover the severity is high, and it is inferred that the victim is a pedestrian. Finally, if the crash is between vehicles with rigid protection systems such as cars, buses or carriages, the accident decreases its severity.

This investigation allowed analyzing the accident records of Cartagena (Colombia) with data mining techniques. The rules contribute to the definition of strategies for the reduction of the severity of accidents. The presence of vulnerable road users (motorcyclists, cyclists, and pedestrians) were predictive variables on the severity of the accidents, as well as the results of [4, 8].

Rules 1, 2, 3, 4, 6, 8 and 12 show that more than 50% of the rules are related to the users of motorcycles. Motorcyclists are a population with significant growth due to the conditions for mobility, transportation, sports and other economic activities [9–11]. Subsequently, some countermeasures are exemplified by the defined rules. These rules are based on findings in the United States and the European Union on vulnerable users that can be replicated to reduce and eliminate road accidents. Among the policies, strategies and countermeasures to improve the road safety of the motorcyclist and the recommendations of WHO [12], are: Promote culture and education in road safety [13]. Analysis and monitoring of accident reports [14]. Road safety campaigns on the most vulnerable users in the ages of 15 to 30 years [15]. Promote the use of protective elements [16]. Restrict and punish driving under the influence of drugs and alcohol [17]. Control the speed according to the road type [18]. Improve the quality of roads, or design exclusive lanes [19]. Improve mechanical conditions and maintenance [20]. Improve the visibility of the motorcyclist [21]. Improve road safety conditions such as lighting, and infrastructure [22]. Forbid the transport and exposition of children on motorcycles [23]. Penalize violations and risky behaviors [24]. Restrict the manipulation of electronic devices while driving [25].

Finally, some additional countermeasures based on the rules are: Define road safety control plans according to the season. Rules 3 and 4 contrast that there are more demanding months in road control. Define speed limits conditioned by the intensity of rainfall. Rules 6 and 12 show that moderate and intense rains increase the possibility of accidents. Motorcyclist accidents can be avoided if interaction with other users is reduced. This is achieved from single circulation lanes. Rules 1, 3, 4, 6, 8 and 12 relate cars when the motorcycle interacts with another vehicle (e.g. Trucks and Cars). Finally, rules 7 and 8 confirm that the age of the motorcyclist influences the severity of the accident. These rules allow you to define license plans according to age. For example, in Europe, there are restrictions on circulation, speed, displacement, and violations according to age. Rule 11 is closely related to the high severity of the abuses. To define effective countermeasures from these rules it would be good to include new variables. These should be focused on additional aspects such as accident location, interception, location of the road, type of road, signaling, lighting, time, among others.

## 5 Conclusion

In this study, motorcyclists at young adult ages are related to predictive factors of severity at a level of injury or fatality. Global records in 2016 placed Colombia in the tenth position worldwide, the third in the region and second in South America in motorcyclist accident. In 2018 Colombia has 8.3 million motorcycles registered. In the last 7 years (2012–2018) the proportion of dead and injured motorcyclists is close to 50% of road users. Analyzing the accident rate of motorcyclists in Colombia and their

causality are future investigations essential to improve road safety. The limitations of the study are the sub-registration of data by traffic control entities. If more causal variables are included, the creation of significantly more strategic rules for road safety can be achieved (for example, state and road conditions).

This investigation allowed analyzing the accident records of Cartagena (Colombia) with data mining techniques. The rules contribute to the definition of countermeasures, focused on vulnerable users for the reduction of the severity of accidents. The definition of rules from data mining is more effective than analyzing information with a simple descriptive statistical analysis. Because the analysis of the information is done in a correlational way, this contributes to obtain results that are easier to understand and apply. Techniques such as multivariate analysis or black-box techniques require additional steps for the analysis of information.

**Acknowledgements.** Funding for first author was covered by (CEIBA)—Gobernación de Bolívar (Colombia). We thank the Administrative Department of Traffic and Transportation (DAT) in the accompaniment and support of the information required for this investigation.

## References

1. World Health Organization (WHO): Global status report on road safety 2018 (2018). <https://apps.who.int/iris/bitstream/handle/10665/276462/9789241565684-eng.pdf?ua=1>
2. Savolainen, P., Mannering, F.: Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes (in English). *Accid. Anal. Prev.* **39**(5), 955–963 (2007)
3. Abdelwahab, H., Abdel-Aty, M.: Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. *Transp. Res. Rec.: J. Transp. Res. Board* **1746**, 6–13 (2001)
4. Hashmienejad, S.H.-A., Hasheminejad, S.M.H.: Traffic accident severity prediction using a novel multi-objective genetic algorithm. *Int. J. Crashworthiness* **22**(4), 425–440 (2017)
5. Sohn, S., Shin, H.: Data mining for road traffic accident type classification. *Ergonomics* **44**, 107–117 (2001)
6. Huang, H., Abdel-Aty, M.: Multilevel data and Bayesian analysis in traffic safety. *Accid. Anal. Prev.* **42**(6), 1556–1565 (2010)
7. Li, Z., Liu, P., Wang, W., Xu, C.: Using support vector machine models for crash injury severity analysis. *Accid. Anal. Prev.* **45**, 478–486 (2012)
8. Delen, D., Tomak, L., Topuz, K., Eryarsoy, E.: Investigating injury severity risk factors in automobile crashes with predictive analytics and sensitivity analysis methods. *J. Transp. Health* **4**, 118–131 (2017)
9. Balasubramanian, V., Jagannath, M.: Detecting motorcycle rider local physical fatigue and discomfort using surface electromyography and seat interface pressure. *Transp. Res. Part F* **22**, 150–158 (2014)
10. Shafiei, U.K.M., Karuppiyah, K., Tmrin, S.B.M., Meng, G.Y., Rasdi, I., Alias, A.N.: The effectiveness of new model of motorcycle seat with built-in lumbar support (in English). *Jurnal Teknologi* **77**(27), 97–103 (2015)
11. Ospina-Mateus, H., Jiménez, L.A.Q.: Understanding the impact of physical fatigue and postural comfort experienced during motorcycling: a systematic review. *J. Transp. Health* **12**, 290–318 (2019)

12. World Health Organization (WHO): Seguridad de los vehículos de motor de dos y tres ruedas: manual de seguridad vial para decisores y profesionales (2017). <https://apps.who.int/iris/bitstream/handle/10665/272757/9789243511924-spa.pdf?sequence=1&isAllowed=y>
13. Seguí-Gomez, M., Lopez-Valdes, F.J.: Recognizing the importance of injury in other policy forums: the case of motorcycle licensing policy in Spain (in English). *Inj. Prev. Short Surv.* **13**(6), 429–430 (2007)
14. Schneider Iv, W.H., Savolainen, P.T., Van Boxel, D., Beverley, R.: Examination of factors determining fault in two-vehicle motorcycle crashes (in English). *Accid. Anal. Prev.* **45**, 669–676 (2012)
15. Ivers, R.Q., et al.: Does an on-road motorcycle coaching program reduce crashes in novice riders? A randomised control trial (in English). *Accid. Anal. Prev.* **86**, 40–46 (2016)
16. Donate-López, C., Espigares-Rodríguez, E., Jiménez-Moleón, J.J., de Dios Luna-del-Castillo, J., Bueno-Cavanillas, A., Lardelli-Claret, P.: The association of age, sex and helmet use with the risk of death for occupants of two-wheeled motor vehicles involved in traffic crashes in Spain. *Accid. Anal. Prev.* **42**(1), 297–306 (2010)
17. Albalate, D., Fernández-Villadangos, L.: Motorcycle injury severity in Barcelona: the role of vehicle type and congestion (in English). *Traffic Inj. Prev.* **11**(6), 623–631 (2010)
18. Clabaux, N., Brenac, T., Perrin, C., Magnin, J., Canu, B., Van Elslande, P.: Motorcyclists' speed and "looked-but-failed-to-see" accidents (in English). *Accid. Anal. Prev.* **49**, 73–77 (2012)
19. Sager, B., Yanko, M.R., Spalek, T.M., Froc, D.J., Bernstein, D.M., Dastur, F.N.: Motorcyclist's lane position as a factor in right-of-way violation collisions: a driving simulator study (in English). *Accid. Anal. Prev.* **72**, 325–329 (2014)
20. Rizzi, M., Strandroth, J., Holst, J., Tingvall, C.: Does the improved stability offered by motorcycle antilock brakes (ABS) make sliding crashes less common? In-depth analysis of fatal crashes involving motorcycles fitted with ABS (in English). *Traffic Inj. Prev.* **17**(6), 625–632 (2016)
21. Clarke, D.D., Ward, P., Bartle, C., Truman, W.: The role of motorcyclist and other driver behaviour in two types of serious accident in the UK (in English). *Accid. Anal. Prev.* **39**(5), 974–981 (2007)
22. López-Valdés, F.J., García, D., Pedrero, D., Moreno, J.L.: Accidents of motorcyclists against roadside infrastructure. In: *IUTAM Symposium on Impact Biomechanics: From Fundamental Insights to Applications*, vol. 124, pp. 163–170, Dublin (2005)
23. Brown, J., Schonstein, L., Ivers, R., Keay, L.: Children and motorcycles: a systematic review of risk factors and interventions. *Inj. Prev.* **24**(2), 166–175 (2018)
24. Elliott, M.A., Baughan, C.J., Sexton, B.F.: Errors and violations in relation to motorcyclists' crash risk (in English). *Accid. Anal. Prev.* **39**(3), 491–499 (2007)
25. Truong, L.T., Nguyen, H.T., De Gruyter, C.: Mobile phone use while riding a motorcycle and crashes among university students. *Traffic Inj. Prev.* **20**, 1–7 (2019)