International Journal of Engineering Sciences & Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164

IJESRT

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

FACTORS AFFECTING ANALYSIS of THE SEVERITY of AUTOMOBILE AND ELECTRIC BICYCLE ACCIDENTS USING RANDOM FOREST MODEL

Long Chen [1], Shangshang Hu [2], Yaping Wang [1], Xue Wen [3] & Fulu Wei[*1]
[1]School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255000, Shandong, China
[2]Transportation Research Department, Changzhou City Planning and Design Institute, Shandong University of Technology, Changzhou, 213002, Jiangsu, China
[3]Traffic Accident Division, Traffic Police Detachment of Zibo Public Security Bureau, Zibo, Shandong, 255000, China

DOI: 10.29121/ijesrt.v11.i3.2022.4

ABSTRACT
The frequency of automobile and electric bicycle accidents has shown a rising trend. The occurrence of such accidents has caused great harm to the safety of electric bicycle drivers and passengers. In order to analyze the affecting factors of the of the severity of automobile and electric bicycle accidents, the data of automobile and electric bicycle accidents in a city from 2010 to 2019 were collected, and the severity of automobile and electric bicycle accidents was predicted by random forest model, and the importance of relevant factors was ranked. The results show that visibility, drivers' age and driving age, road cross section location, accident time and other factors have significant effects on the severity of electric bicycle drivers. The drivers' age of electric bicycle and automobile, the time of accident, the responsibility of accident and the severity of electric bicycle driver have significant effects on the severity of automobile bicycle driver. The research is of positive significance to reduce the severity of urban automobile and electric bicycle traffic accidents.

KEYWORDS: Traffic safety; Affecting factors; Random forest; Automobile and electric bicycle; Severity.

1. INTRODUCTION
Road traffic accidents is a worldwide public hazard and cause up to 1.25 million deaths worldwide every year. It has brought loss of life and property to people all over the world [1,2]. As the largest developing country, China's traffic situation is particularly complicated and various traffic accidents occur frequently[3,4]. Through the statistics of the means of transportation used by Chinese citizens, it is found that the use of electric bicycle accounts for 22%, second only to bicycle. Electric bicycle is one of the main means of transportation for Chinese citizens. The experts predict that the proportion of using electric bicycles will continue to rise in the future. The analysis of the National Bureau of statistics shows that there were 265,204 road traffic accidents in China in 2019. Road traffic accidents resulting in 73,484 deaths, 304,919 injuries and direct property loss of 1.01 billion yuan [5]. In all vehicles, automobiles account for 47.70%, motorcycles account for 10.98%, electric bicycles account for 20.91%, and other vehicles account for 20.41%. Accidents between automobiles and electric bicycles account for one-third of China's traffic accidents, these accidents account for more than 50% in some areas, so the study of accidents between automobiles and electric bicycles has important significance.

Experts and scholars have done a lot of research on electric bicycle accidents, but most of them tend to analyze the injured parts. It includes for vehicle and pedestrian accident reconstruction, the accident severity is analyzed from collision angle as well as to assessment the between automobile and pedestrian accident risk, research shows that through the optimization of the structure of city road, improve road traffic design can significantly reduce automobiles and pedestrian traffic accident happened [6-8]. Soria analyzed the injury characteristics by studying the cases of fatal traffic accidents between motorcycles and pedestrians. The study was conducted by calculating the injury scores (MAIS) of the three most seriously injured parts of the victim, the results showed that head, chest

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

and lower limbs are the main injuries suffered by pedestrians in fatal accidents with motorcycles. It provides a direction for the study of the mechanism of fatal accidents between motorcycles and pedestrians [9]. Yuan studied the risk factors of head injury in accidents between electric two-wheeled vehicles and bicycles. He found that the risk degree of head injury increases with the increase of vehicle speed. The influence of the speed of two-wheeled vehicle and the relative position of pedestrian and vehicle on the motion response is highly nonlinear. Larger passengers were more likely to miss their heads. The study revealed the mechanism that different risk of head injury in electric two-wheelers compared with bicycles [10]. Thajudeen Hassan studied the influence between the age of electric two-wheeled vehicle drivers and dangerous driving behavior with safety attitude as medium. The study showed that the application of safety strategies could reduce dangerous driving tendency by shaping drivers' safety awareness [11].

The research on automobile and electric vehicle accidents in China mostly focuses on the single accident of electric vehicle. Some scholars use uncertainty analysis theory to study the head injury of electric two wheeled vehicle drivers [12,13]. In addition, some scholars have studied the factors affecting the severity of electric bicycle accidents [14-16] and the injuries of cyclists [17,18]. Wang studied the risky driving behaviors and accident affecting factors of electric bicycles and studied systematically structure and occurrence mechanism of significant factors and potential factors affecting electric bicycle accidents. He established a combined accident prediction model under the joint influence of human factors and external environment [19]. Jiang obtained the risky driving behaviors of electric two-wheeler drivers according to exploratory factor analysis (EFA) of questionnaire survey and constructed a traffic accident logistic regression prediction model. He verified the model through confirmatory factor analysis and obtained the main parameters affecting traffic accidents **Error! Reference source not found.**.

2. DATA ANALYSIS

2.1. Data Description

This paper selects 3090 effective accident data between automobile and electric bicycles in a city of Shandong Province from 2010 to 2019 to study the affecting factors of automobile and electric bicycle accidents. The data related driver characteristics, road and environment characteristics, spatio-temporal characteristics, driving behavior characteristics and other characteristics, including severity, driving age, gender, accident pattern, road cross section type, visibility, road section type and other 24 variables. For automobile and electric bicycle crashes (AE), AE accident data are automobile data (A_type), electric bicycle data (E_type), automobile and electric bicycle data (AE_type).

2.2. Variable Description

Statistical analysis was conducted on variable data, as shown in Table 1. With the severity of driver injury as the dependent variable, driving age, gender, age and other 24 variables as independent variables. In order to evaluate the degree of injury caused by affecting factors, three classification methods are used to study the severity of accident, including no injury, minor injury, serious injury and death. The severity of AE accidents is classified, including two target variables, the degree of injury of automobile drivers (A_severity) and the degree of injury of electric bicycle drivers (E_severity), as shown in Table 2.

Table 1 Description of independent variables

Variable	Classification	Assignment	Number	Variable	Classification	Assignment	Number
Severity	No harm	1	A=2634	Accident mode	Collision stationary vehicle	1	329
			E=394		Collision moving vehicle	2	2761
	Minor injuries	2	A=333	Type of road cross section	A motor vehicle,	1	2232
			E=1914		Non-motorized road	2	427
	Serious injury and death	3	A=123		Machine non-mixing zone	3	431
			E=783		good	1	2897

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Driving age	1-3 years	1	565	Route table status	bad	2	193
	3-15 years	2	1535	Road type	Grade highway	1	1728
	More than 15 years	3	990		Urban road	2	1362
Gender	male	1	A=2851	Road section type	Ordinary road	1	2075
			E=2924		Multi-fork section	2	1015
	female	2	A=239	Time of the accident	When [22, 6)	1	592
			E=164		[6, 14)	2	1062
Age	Under the age of 30,	1	A=682		[14, 22)	3	1436
			E=655	Week	Working days	1	2220
	30 to 50 years old	2	A=2014		The working day	2	870
			E=1573	Weather	good	1	2652
	50 years of age or older	3	A=394		bad	2	438
			E=862	Visibility	The < 50 m	1	546
Occupation	farmers	1	A=1546		50-200m	2	842
			E=1707		> 200 m	3	1702
	The farmers	2	A=1544	Blood alcohol level	No alcohol and no examination	1	A=2315
			E=1383				E=2194
Collision type	Frontal and rear-end collisions	1	991		Drink driving	2	A=775
	Side impact	2	1959				E=876
	Scratch and other types	3	140	Use of safety devices	Use of safety devices	1	A=1284
Accident liability	Full and primary responsibility	1	A=1394				E=938
			E=932		No safety device was used	2	A=1806
	Equal responsibility	2	A=603				E=2152
			E=696	Types of vehicle-to-vehicle accidents	Electric bicycles and passenger cars	1	2001
	Secondary liability and no liability	3	A=1093		Electric bicycles and heavy goods vehicles	2	274
			E=1462		Electric bicycles and small and medium trucks	3	815

Table 2 AE_severity distribution

Category	The degree of damage	Frequency	The percentage / %	Effective percentage / %	Cumulative percentage / %
A_ Severity	No harm	2634	85.24	85.24	85.24
	Minor injuries	333	10.78	10.78	96.02
	Serious injury and death	123	3.98	3.98	100.00
A combined		3090	100.00	100.00	
	No harm	394	12.75	12.75	12.75

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

E_ Severity	Minor injuries	1914	61.91	61.91	74.66
	Serious injury and death	783	25.34	25.34	100.00
A combined		3090	100.00	100.00	

3. MODEL CONSTRUCTION AND RESULT ANALYSIS

3.1. Model construction theory

As an ensemble learning method, random forest is a classifier containing multiple decision trees and whose output categories are jointly determined by the modes of individual output categories. The random forest algorithm takes each decision tree as a classifier and the mode output by multiple classifiers as the output result. The random forest algorithm combines the advantages of the decision tree algorithm. It adopts bagging algorithm which makes the model more accurate and stable and avoids the complicated parameter tuning process of SVM classifier.

In the process of constructing the random forest model by IBM SPSS Modeler, the C&R Tree (Classification and Regression Tree) is adopted. The Classification Tree is suitable for the Classification of target variables. The Regression Tree is used for the continuity of target variables. In the C&R Tree algorithm, the Gini coefficient is used to confirm the segmentation point when the target variable is a classification variable. Gini coefficient can reflect the degree of difference between groups of target variables. The smaller the coefficient, the greater the difference between groups. Gini calculation formula as shown in Eq. 1.

$$G_{(t)}=1-\left(\frac{t_1}{T}\right)^2-\left(\frac{t_2}{T}\right)^2-\cdots-\left(\frac{t_n}{T}\right)^2 \quad (1)$$

Where, T is the total number of records, t_1, t_2, $\cdots$, t_n are the number of records of each category of the output variable respectively.

When the target variable is numerical, the regression tree algorithm is adopted. The strategy for determining the best grouping variables in regression trees is the same as in classification trees. The main difference is the index to test the heterogeneity of the output variable. Regression tree is used to evaluate the reduction level of heterogeneity through the reduction index of variance. Its mathematical expression as shown in Eq. 2.

$$\Delta R(t) = R(t) - \frac{N_{t1}}{N} R_{t1} - \frac{N_{t2}}{N} R_{t2} \quad (2)$$

Where, $R(t)$ and N are respectively the variance and sample size of output variables before grouping; R_{t1}, N_{t1}, R_{t2}, N_{t2}, are respectively the variance and sample size of left and right subtrees after grouping.

The variable that reaches the maximum is the current optimal grouping variable $\Delta R(t)$.
Random forest takes decision tree as the basic unit and adopts Bagging algorithm to correct the shortcomings of decision tree easily causing model overfitting. Random forest makes the model more accurate and more stable. Bagging adopts Bootstrap sampling with random placing and uniform sampling. Each sample has the same weight, so that the sample utilization rate is more sufficient and the error of model parameters is reduced. Bagging algorithm constructs decision tree in random forest. Its working process is as follows:

- N training samples were randomly selected from the original sample set using the Bootstraping Method. A total of k training sets were obtained (k training sets were independent of each other and elements could be repeated).
- For k training sets, we train k decision trees;
- For classification problem: the classification result is generated by the mode output of decision tree.

3.2. Model construction theory

The random forest model was constructed to study the factors influencing the severity of automobile and electric bicycle accidents based on IBM SPSS Modeler. The severity of AE accident was predicted by using the random forest model, and the importance of related factors was ranked. The modeling process is shown in Figure 1.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Random forest model training target is A_severity and E_severity and screening through experiment. It tries every branch of the variable is set to 30, has established 500 classification trees. The tree depth is set to 10 and the boy node size is set to 5, missing value maximum percentage is set to 70%. The form is selected to analyses effectiveness of nodes towards the model.

3.3. Model Analysis

A. Model information

A_severity and E_severity were taken as the target variables respectively to analyze the injury degree of automobile and electric bicycle drivers in the accident. The model information is shown in Table 3. There are 25 predictive variables in the model. The accuracy rate of the model is 73.7% and 67.4% respectively. The misclassification rate is 26.3% and 32.6% respectively. The misclassification rate is relatively high because there are 25 selected input characteristic variables. The classification error rate of the model can be effectively reduced by reducing the number of selected characteristic variables. Overall, the model accuracy rate is about 70% higher.

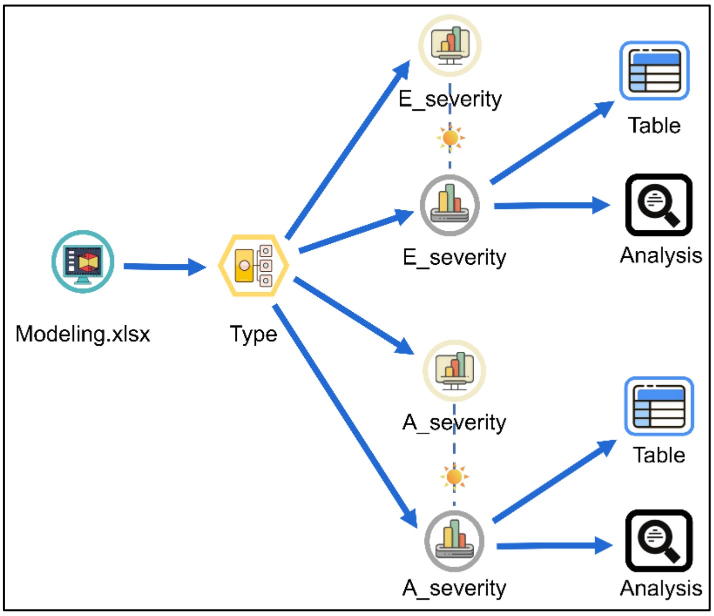

FIG. 1 Random forest model construction process

Table 3 Model information

The target field	A_Severity	E_Severity
Model building method	Random Trees Regression	Random Trees Regression
The predictive variable of the input	25	25
Model accuracy	0.737	0.674
Misclassification rate	0.263	0.326

B. Result analysis

Through the analysis of the random forest model, the significance factors affecting the severity of between automobile and electric bicycle accidents were obtained. The importance of variables was ranked according to the decrease of the average accuracy of the model, as shown in Figure 2 and Figure 3 respectively, the importance of A_severity and E_severity influencing variables were ranked.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

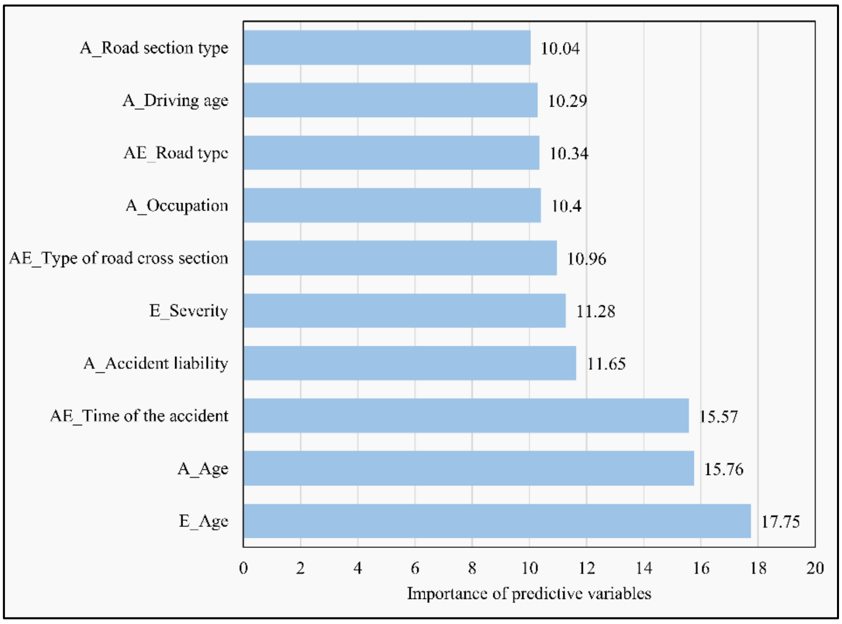

FIG. 2 Order of importance of affecting factors of A_severity

As shown in Figure 2, the importance of factors affecting the A_severity is ranked as E_age, A_age, AE_accident occurrence time, A_accident responsibility, E_severity, AE_road cross section position, A_person type, AE_road type, A_driving age, and AE_road section type.

Through the analysis of the significant factors affecting the A_severity, it is found that the age of electric bicycle driver has a great influence on the A_severity. The second is the age of the driver. The older the driver, the slower the perception of the accident situation. Too young is prone to excessive driving behavior. Therefore, the age has a higher impact on the accident. Then the occurrence time of the accident has a higher impact on the severity of the accident. These three factors are the most important factors causing the severity of the accident.

According to the analysis results, in order to reduce the A_severity when accidents occur, it is necessary to strengthen traffic control measures, optimize road cross sections and road types, improve traffic safety education, and standardize road sections and intersections, so as to avoid accidents and reduce the severity of accidents.

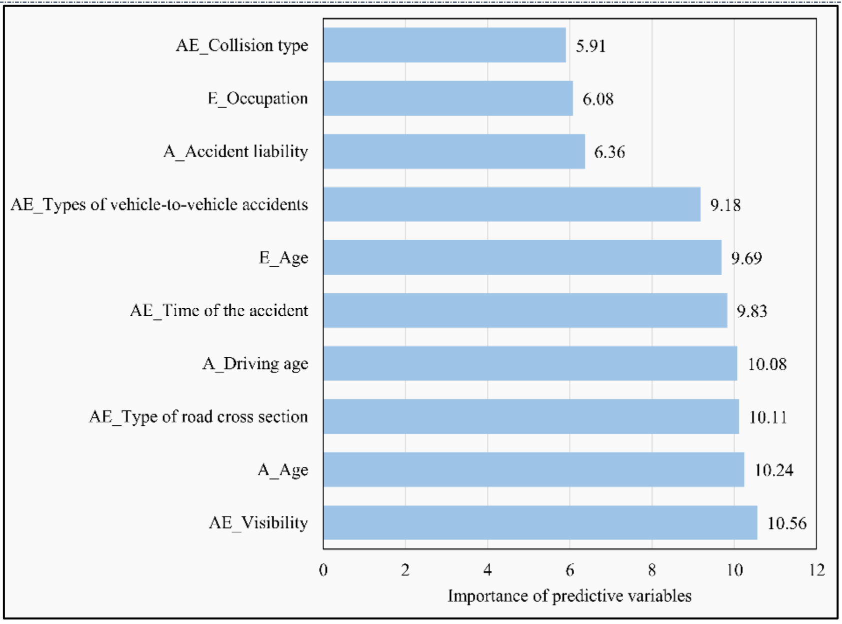

FIG. 3 Order of importance of affecting factors of E_severity

As shown in Figure 3, the importance of factors affecting the A_severity is ranked as AE_visibility, A_age, AE_cross-sectional position of road, A_driving age, AE_accident occurrence time, AE_age, AE_accident type between vehicles, AE_accident responsibility, E_person type and AE_collision type.

Through to analyze the factors influencing the E_severity significance, it is considered that the visibility is important for E_severity has the greatest impact, visibility seriously affects the visual perception of electric vehicle drivers. Due to the weak protection of electric vehicles to drivers, serious accidents are more likely to occur. Secondly, the age of the driver has a great influence on the driving mode and behavior of drivers. The measures taken by car drivers in the event of an accident with electric vehicles are directly related to the severity of the injury of electric bicycle drivers. Finally, the cross-sectional position of the road, whether motor vehicles and non-motor vehicles are separated is directly related to the occurrence of automobile and electric bicycle accidents. Optimizing the cross-sectional design of the road can avoid the occurrence of accidents and reduce the severity of the accident. The three factors are the most important factors causing the severity of the accident.

For the results of the analysis, in order to reduce E_accident severity, the departments concerned need to improve the operating conditions of electric vehicles in poor visibility environment, optimize the road cross section and road type, improve degree of attention to the traffic safety education, maintain good sight distance conditions by standardizing road sections and intersections to avoid the happening of accidents and reduce the severity of the accident.

C. Model Verification

As shown in Table 4, the prediction accuracy of A_severity and E_severity is 81.3% and 75.2%. The standard deviation is 0.296 and 0.217. The results show that the prediction effect of A_severity and E_severity of the model is good and the accuracy is high. The significance affecting factors can reflect the influence effect on the severity of the accident to a large extent. The occurrence of such accidents can be better reduced by controlling the screened significance affecting factors.

Table 4. Fuzzy quantitative values of comfort indicators.

Type	Comparison of A_severity with predicted A_severity	Comparison of E_severity with predicted E_severity
The average error	0.004	0.001
Mean absolute error	0.163	0.317
The standard deviation	0.296	0.217

Linear correlation	0.813	0.752
The incidence of	3090	3090

4. CONCLUSIONS

Based on the database of automobile and electric bicycle accidents in a city from 2010 to 2019, the random forest model was established to analyze the affecting factors of automobile and electric bicycle accidents. The significant affecting factors of the severity of automobile and electric bicycle accidents were obtained, respectively. The results show that: (1) the drivers' age of automobile and electric bicycle, driving age of automobile drivers, accident occurrence time, cross section and occupation are all the significant factors influencing the severity of automobile and electric bicycle accidents; (2) Electric bicycle drivers are belong to vulnerable groups. To ensure good visibility under the operation environment, improve the traffic awareness level of occupations, optimize the type of road cross section can help to avoid such accidents. According to the significant factors, the improvement suggestions are helpful to prevent the occurrence of such accidents and reduce the severity of accidents.

Acknowledgments

The authors are thankful for the support of the Center for Accident Research in Zibo (CARZ).

REFERENCES

[1] Z. Cai, F. Wei, Z. Wang, Y. Guo, L. Chen, X. Li, "Modeling of Low Visibility-Related Rural Single-Vehicle Crashes Considering Unobserved Heterogeneity and Spatial Correlation" in Sustainability, 13(13), 7438, 2021.
[2] F. Wei, Z. Cai, Z. Yan, P. Li, Y. Qing, L. Chen, "Analysis of the incidence of death accidents based on Bayesian Logit model" in Journal of Guangxi University (Natural Science Edition), 46(4): 1054-1062, 2021.
[3] F. Wei, Z. Cai, P. Liu, Y. Guo, X. Li, Q. Li, "Exploring Driver Injury Severity in Single-Vehicle Crashes under Foggy Weather and Clear Weather" in Journal of Advanced Transportation, 2021, 9939800, 2021.
[4] F. Wei, Z. Cai, Y. Guo, P. Liu, Z. Wang, Z. Li, "Analysis of Roadside Accident Severity on Rural and Urban Roadways" in Intelligent Automation & Soft Computing, 20(3): 754-767, 2021.
[5] National Bureau of Statistics, China Statistical Yearbook, China Statistics Press, 2020.
[6] Y. Weng, X. Jin, Z. Zhao, X. Zhang, "Car-to-pedestrian collision reconstruction with injury as an evaluation index" in Accident Analysis and Prevention, 42: 1320-1325, 2010.
[7] F. Wei, Z. Cai, Z. Wang, Y. Guo, X. Li, X. Wu, "Investigating Rural Single-Vehicle Crash Severity by Vehicle Types Using Full Bayesian Spatial Random Parameters Logit Model" in Applied Sciences, 11, 7819, 2021.
[8] N. Clabaux, J. Fournier, J. Michel, "Powered two-wheeler drivers' risk of hitting a pedestrian in towns" in Journal of Safety Research, 21(51): 1-5, 2014.
[9] M. Carmen, A. Carlos, S. David, V. Juan, G. Ignasi, "Injury pattern in lethal motorbikes-pedestrian collisions, in the area of Barcelona, Spain" in Journal of Forensic and Legal Medicine, 43: 80-84, 2016.
[10] Y. Huang, Q. Zhou, K. Caroline, Q. Li, B. Nie, "Are riders of electric two-wheelers safer than bicyclists in collisions with motor vehicles?" in Accident Analysis and Prevention, 134: 105336, 2020.
[11] H. Thajudeen, M. Vinodkumar, V. Neethu, "Role of sensation seeking and attitudes as mediators between age of driver and risky driving of Powered Two Wheelers" in Journal of Safety Research, 62: 209-215, 2017.
[12] Y. Han, J. Xu, L. Shi, X. Gao, Y. Qian, Z. Yang, "Uncertainty analysis of head injury via reconstruction of electric two-wheeler accidents" in China Journal of Highway and Transport, 33(1): 172-180+190, 2020.
[13] J. Xu, "Uncertainty analysis and head injury study of electric two-wheelers accidents" Xiamen University Of Technology, 2019.
[14] X. Xiao, "Analysis on The Affecting factors of The Severity of Electric Bicycle Road Traffic Accidents" in Highways & Automotive Applications, 2020(6): 32-36, 2020.
[15] C. Xu, M. Guan, "Analysis on affecting factors of electric bicycle traffic accident severity: A case study of Hangzhou city" in Journal of People's Public Security University of China (Science and Technology), 24(3): 75-80, 2018.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[16] T. Wang, W. Li, W. Li, "Influence factors and injury severity in electric bicycles traffic crashes", in Journal of Guangxi University (Natural Science Edition), 42(6): 2080-2088, 2017.
[17] C. Li, Y. Li, Q. He, "Factors affecting injury severity of electric bike riders in road crashes" in Journal of Transportation Engineering, 19(4): 37-43, 2019.
[18] W. Wang, X. Shen, G. Wang, Q. Fang, "Analysis of factors affecting injury to electric bicycle rider in crash" in China Safety Science Journal, 29(2): 20-25, 2019.
[19] T. Wang, "The mechanism study on the risky driving behavior and accident of electric bike riders" Southeast University, 2017.
[20] L. Jiang, Y. He, "Risky driving behavior and affecting factors analysis for electric two-wheeler" in Journal of Jilin University (Engineering and Technology Edition), 49(4): 1107-1113, 2019.