

Research Article

Exploring the Mechanism of Crashes with Autonomous Vehicles Using Machine Learning

Hengrui Chen , Hong Chen , Ruiyu Zhou , Zhizhen Liu , and Xiaoke Sun 

College of Transportation Engineering, Chang'an University, Xi'an 710000, China

Correspondence should be addressed to Hong Chen; glch@chd.edu.cn

Received 9 January 2021; Accepted 19 February 2021; Published 27 February 2021

Academic Editor: Mostafa M. A. Khater; mostafa.khater2024@yahoo.com

Copyright © 2021 Hengrui Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The safety issue has become a critical obstacle that cannot be ignored in the marketization of autonomous vehicles (AVs). The objective of this study is to explore the mechanism of AV-involved crashes and analyze the impact of each feature on crash severity. We use the Apriori algorithm to explore the causal relationship between multiple factors to explore the mechanism of crashes. We use various machine learning models, including support vector machine (SVM), classification and regression tree (CART), and eXtreme Gradient Boosting (XGBoost), to analyze the crash severity. Besides, we apply the Shapley Additive Explanations (SHAP) to interpret the importance of each factor. The results indicate that XGBoost obtains the best result (recall = 75%; G-mean = 67.82%). Both XGBoost and Apriori algorithm effectively provided meaningful insights about AV-involved crash characteristics and their relationship. Among all these features, vehicle damage, weather conditions, accident location, and driving mode are the most critical features. We found that most rear-end crashes are conventional vehicles bumping into the rear of AVs. Drivers should be extremely cautious when driving in fog, snow, and insufficient light. Besides, drivers should be careful when driving near intersections, especially in the autonomous driving mode.

1. Introduction

In recent years, academia and industry have invested enormous human and material resources in the research and development of AVs. One of their original intentions is to reduce the rate of traffic accidents. Every year, the economic loss caused by traffic accidents is 277 billion dollars, which is twice as much as traffic congestion [1]. Drivers' errors cause 90% of traffic accidents. More than 40% of fatal crashes are related to alcohol, distraction, drug addiction, and fatigued driving [2]. Even when nonhuman factors primarily cause crashes, they usually include some human factors such as distractions or unfamiliar driving skills. With improved AV technology, fatal accident rates are likely to fall by at least 40%, and human factors may disappear. Many studies have been carried out to explore the features that influence the severity of conventional vehicle crashes [3, 4]. However, the literature review shows that, relatively, few studies have focused on factors influencing the severity of AV-involved crashes. The driving system of AVs is different from

conventional vehicles. Therefore, clarifying the mechanism of AV-related crashes is of great importance to improve the safety of AVs.

In the previous literature on the contributing factors of AV-related crashes [5–8], less attention was paid to the correlation between different factors. These research studies have proposed various methods (such as classification and regression trees and neural networks) to explore the features and mechanism of the AV-involved crash. However, these methods are based on the hypotheses that the factors are independent of each other, and it is easy to ignore the causal relationship between factors. Association rule mining is a crucial technology among numerous data mining technologies, especially in analyzing the cause of traffic crashes, because it does not rely on any assumptions and can discover meaningful relationships hidden in large datasets [9]. Therefore, to interpret the interrelationships between factors and further explore the mechanism of AV-involved crashes, the association rule mining algorithms need to be adopted.

The interpretability of model results is another valuable issue. There are many machine learning methods (such as random forest [10], classification and regression tree [11], and gradient boosting model [12]) that have been utilized to study the severity of traffic crashes. These models are more complex and data-driven and have higher accuracy than traditional calculation models [13]. However, these models are usually regarded as a “black box” because the complicated and nonlinear effects of features on the prediction results cannot be explained [14]. The current study uses SHAP (Shapley Additive Explanations) to interpret how a variable affects model prediction results. SHAP, proposed by Lundberg and Lee [15], originated from cooperative game theory. The prediction results are explained by calculating the contribution of individual variables to the results.

The objective of this study is to explore the mechanism of AV-involved crashes and investigate the impact of each feature on accident severity. We adopted 131 AV-involved crash reports received in California from 2019 to October 2020. We used synthetic minority oversampling technique (SMOTE) to balance the dataset. The Apriori algorithm and classification models are used for accident mechanism analysis. We apply the SHAP to interpret how a variable affects model classification results.

2. Literature Review

The method of improving traffic safety through advanced driving assistance system (such as antilock brake system (ABS), electronic stability program (ESP), autonomous emergency braking (AEB), and lane-keeping assist system (LKA)) has been used in the automotive industry for many years. Statistics show that these systems have effectively improved vehicle safety and reduced the rate of traffic accidents [2]. As the level of AV technology increases, the driving tasks gradually transferred from the driver to the autonomous driving system. Since AVs have a strong environmental perception (such as the vehicle to everything (V2X)), data processing, and rapid response capabilities, they can make up for the driver’s inherent shortcomings to a certain extent [13]. It is foreseeable that, with the development of AV technology, traffic safety problems will be further alleviated. In the 1970s, Haddon proposed a theory from the perspective of human-vehicle-environment, which divided the crash into three stages: precrash, crash, and postcrash [16]. The three factors at each stage of the crash process are arranged and combined to form the famous Haddon matrix, shown in Table 1.

With the development of AV technology, the causes of traffic accidents are significantly changed. As shown in Figure 1, drivers make driving decisions based on complex environmental information and vehicle status for conventional vehicles. Therefore, if the vehicle is in danger of collision, the driver needs to quickly make a decision based on the current driving scene and combined with his own experience. However, the situation for AVs is different. As shown in Figure 1, the AV technology eliminates the driver’s unstable factor. In the preset operation design domain (ODD), AVs rely on their sensors to perceive all

environmental information (including traffic information, environmental conditions, and road conditions) and make driving decisions [16]. Simultaneously, through human-computer interaction techniques, V2X, and other technologies, AVs share some of their driving status with other traffic participants. If an accident occurs in the ODD, the primary responsibility is the AV. Therefore, to protect passengers’ safety and avoid accident liability, the AV manufacturer will carry out comprehensive testing and verification before the AV launch. Thus, based on the perception of the environment and the infrastructure’s support, the traffic accidents caused by human drivers may disappear under ideal circumstances. The Haddon matrix will change accordingly. However, limited by current AV technology and transportation facilities, this goal cannot be achieved in the short term.

The previous studies on the safety of AV technology were carried out in analyzing driver behavior in the driving simulator and testing the autonomous driving system’s stability in closed environments. To avoid potential collisions and improve traffic efficiency, some research studies concentrate on trajectory optimization of AVs. Omidvar et al. [17] designed an AV trajectory optimization algorithm for closed road signalized intersections. The algorithm can optimize the signal control scheme before the AVs arrive at the intersection, control the vehicle’s speed, and ensure that the vehicles can quickly pass the intersection. Li et al. [18] developed an integrated local trajectory optimization scheme and tracking control framework to avoid obstacles in time. With safety and comfort as the objective function, the best trajectory plan is selected. Zhu et al. [19] proposed a speed control method for AVs based on vehicle speed prediction, which improves vehicle operating efficiency and comfort. Many studies use driving simulators to analyze the characteristics of drivers. In the environment of automatic driving, attention should be paid to drivers’ physiological and psychological reactions. Winter et al. [20] found that drivers do not need to monitor the vehicle’s automation process for a long time when driving a highly automated vehicle. They can shift their attention to nondriving-related tasks without affecting the safety of the vehicle. However, the insufficient sample size is one of the limitations of AV safety testing. Both field and driving simulator research attempt to analyze the safety problems of AVs from the perspective of vehicle control and human factors. However, we also need to fully explore the mechanism of AV-involved crashes and analyze the impact of each feature on crash severity. Table 2 lists the studies conducted on AV-involved crashes. This table presents a list of variables used in the studies, analysis methods, and the significant factors obtained from the study.

Determining the cause of a traffic crash is the most critical process for taking preventive measures to reduce the severity and traffic crashes. The Apriori algorithm proposed by Agrawal et al. [21] is the most commonly used association rule mining algorithm. It has been widely used in traffic safety analysis [22–24]. Xu et al. [9] used the Apriori algorithm to explore the causes of traffic crashes with heavy casualties and their interdependent relationship in China. The results indicated that serious casualty crashes resulted

TABLE 1: Haddon matrix.

Phase	Human factors	Vehicle factors	Environment factors
Precrash	Information	Road worthiness	Road design and road layout
	Attitudes	Lighting	Speed limits
	Impairment	Braking	Pedestrian facilities
Crash	Police enforcement	Speed management	Other safety devices
	Use of restraints	Occupant restraints	Crash-protective roadside objects
	Impairments	Other safety devices' crash-protective design	
Postcrash	First-aid skills	Ease of access	Rescue facilities
	Access to medics	Fire risk	Congestion

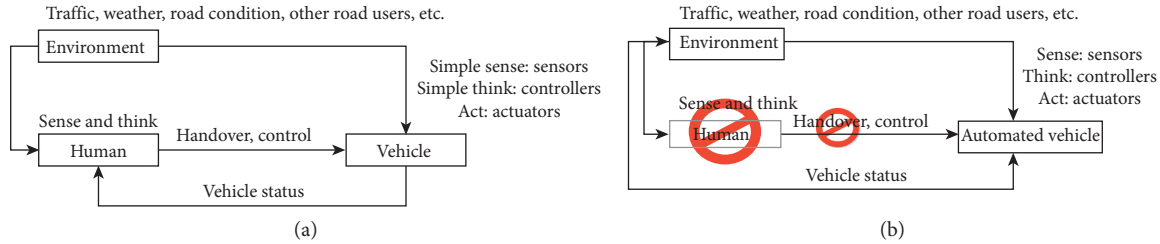


FIGURE 1: (a) Conventional and (b) autonomous vehicles' human-vehicle-environment interaction diagram.

TABLE 2: Methods and significant factors in previous studies.

Study	Variables used in the studies	Analysis methods	Significant factors
Wang and Li [6]	Crash severity, collision type, roadway characteristics	CART models	Location (highway), accident liability
Xu et al. [5]	Type of intersection, type of collision, roadway characteristics, weather	Binary logistic regression model	Driving mode, collision location, roadside parking
Boggs et al. [7]	County, test year, time of the day, vehicle state, type of crash, crash severity	Hierarchical Bayesian heterogeneity-based model	Weather condition
Chen et al. [8]	Crash severity, collision type, weather, vehicle damage, accident location, weather, vehicle state, land-used data	XGBoost model	Weather condition, location (intersection)

from complex interactions between traffic participants, vehicle, road, and environmental conditions. Montella et al. [25] applied the Apriori algorithm to analyze the Italian PTW collision to find the interdependencies and differences between the collision features. Yu et al. [23] adopted the Apriori algorithm to recognize risk factors that are prominently linked to the severity of crash accidents. Many studies have begun to use SHAP for model interpretation. In terms of traffic safety, Mihaita et al. [26] applied the SHAP to study the influence of various characteristics on crash duration. Parsa et al. [1] adopted the SHAP to explain the individual features' importance on accident detection. Zhou et al. [14] applied the SHAP to interpret the influence factors of the severity of car and truck driver injury in the car-truck collision.

By combining the Apriori algorithm and classification models to explore the mechanism of AV-involved crashes, this study provides some useful information for taking preventive measures to promote the safety of AVs.

3. Data Analysis and Feature Extraction

3.1. *Data Sources.* The Department of Motor Vehicles (DMV) was required to provide AV-related crash reports

within ten business days of an accident. This study adopted 131 crash reports received in California from 2019 to October 2020. We extract AV-related information from crash reports such as type of collision, crash severity, vehicle information, and weather. According to the heat map (shown in Figure 2) of the accident location, we find that the crash mainly occurred in Northeast San Francisco because it is the primary test site for AVs.

3.2. *Imbalanced Data Treatment.* By nature, accident data are unbalanced because most crashes cause property damage, and a few lead to injuries. The techniques applied to cope with imbalanced data can be divided into two groups: oversampling and undersampling. The undersampling technique often leads to the loss of a large amount of data, leading to decreased model accuracy. Therefore, oversampling is usually preferred. In this study, we adopted SMOTE to cope with imbalanced data.

3.3. *Variable Collinearity Analysis.* Multicollinearity refers to the distortion or difficulty of estimation due to the high correlation between explanatory variables in a linear regression model. Variance inflation factor (VIF) is a common

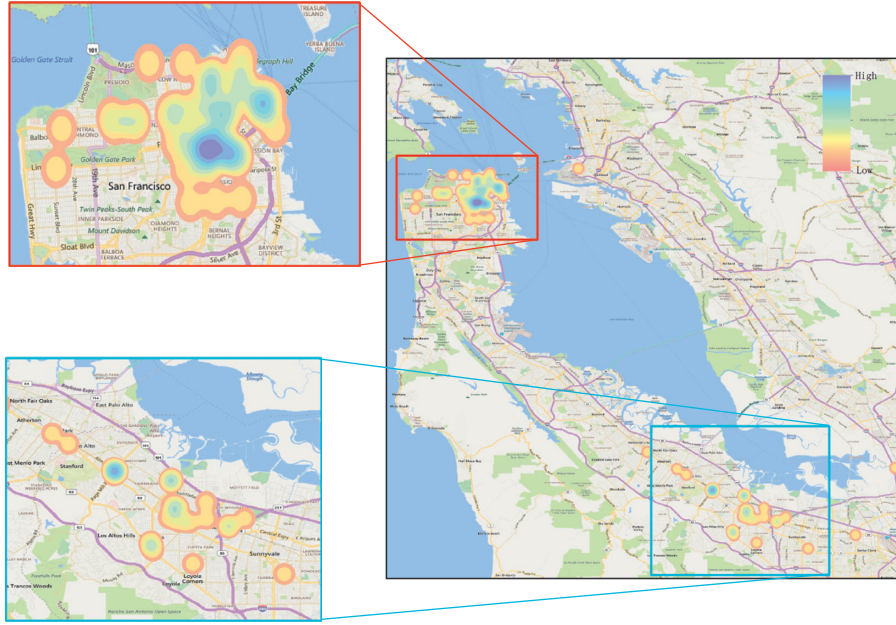


FIGURE 2: Heat map of the accident location.

indicator to measure variables' multicollinearity [27–29]. We used SPSS 26.0 to calculate the VIF value of each variable. When the VIF value is greater than 10, there is serious collinearity between the two variables. It is recommended to eliminate one of the variables [30]. Finally, we selected nine categorical variables. The variables' description and distribution are shown in Table 3.

4. Methodology

4.1. Association Rule Mining. The Apriori algorithm identifies sets of items (i.e., the crash patterns in our study) that occur together in a given event (i.e., a crash in our research). It is the most basic and widely used algorithm in association rule mining. It uses layer-by-layer iterative search to calculate frequent itemsets in the database and determine strong association rules. The association rules can be expressed as $X \Rightarrow Y$ (where X means antecedent and Y means consequent). That is to say, when X appears in the dataset, Y may also occur. In this study, the Apriori algorithm interprets the interrelationships between factors and further explores the mechanism of AV-involved crash.

In the Apriori algorithm, lift (L), support (S), and confidence (C) are three essential indicators for discovering association rules. S means the percentage of occurrence of several related data in the complete dataset, which can be calculated as

$$\text{support}(X, Y) = P(XY) = \frac{\text{number}(XY)}{N}, \quad (1)$$

where $\text{number}(XY)$ is the frequency of X and Y appearing in the dataset at the same time and N is the total number of samples.

C can be interpreted as the conditional probability $P(Y|X)$ (the probability of finding itemset Y in crashes given that the dataset already contains X). It can be calculated as

$$\text{confidence}(X \Rightarrow Y) = P(Y|X) = \frac{P(XY)}{P(X)}. \quad (2)$$

We used S and C to exclude meaningless rules. L means the effect of antecedent on the probability of consequent and is used to determine whether the rule has actual value. It can be calculated as

$$\text{lift}(X \Rightarrow Y) = \frac{P(Y|X)}{P(Y)} = \frac{\text{confidence}(X \Rightarrow Y)}{P(Y)}. \quad (3)$$

When the value of L is greater than 1, it indicates that X 's occurrence increases the probability of Y 's occurrence; otherwise, it is an invalid rule.

In general, the rationality of a rule is judged by support, confidence, and lift of the rule, but in the crash dataset, the same support has different values for the classification factors under different variables. Therefore, when mining accident association rules, the reference value of the support indicator is small. The confidence is essentially a conditional probability, representing the probability of the following itemset appears when the previous itemset occurs. Thus, the confidence indicator will be used as the primary indicator for later analysis and screening. We used Python 3.6 to code the Apriori algorithm.

4.2. Classification Models. The mining of association rules for traffic accidents can explore the causal relationship between multiple factors to explore the mechanism of AV-involved crashes. Besides, we use classification models to investigate the primary factors affecting the severity of the accident. This section briefly introduces each classification model and classification performance evaluation indicators. We used the Scikit-learn (sklearn) library in Python 3.6 to code the following models.

TABLE 3: Variables' description and distribution.

Variable	No. of crashes	Distribution (%)
<i>TOC (type of collision)</i>		
Head-on = 0	11	8.60
Side swipe = 1	20	15.05
Rear end = 2	85	64.52
Broadside = 3	15	11.83
<i>PMAV (the precrash movement of AVs)</i>		
Stopped = 0	51	38.71
Moving = 1	80	61.29
<i>PMCV (the precrash movement of conventional vehicles)</i>		
Stopped = 0	12	9.16
Moving = 1	119	90.84
<i>VD (vehicle damage)</i>		
None = 0	10	7.52
Minor = 1	80	60.07
Mod = 2	35	26.89
Major = 3	5	3.82
<i>DM (precrash driving mode)</i>		
Conventional = 0	69	52.69
Autonomous = 1	62	47.31
<i>AL (accident location)</i>		
Intersection = 1	62	47.31
Street = 2	46	35.48
Highway = 3	18	13.98
Parking lot = 4	4	3.23
<i>W (weather conditions)</i>		
Clear = 1	99	75.57
Cloudy = 2	22	16.79
Raining = 3	6	4.58
Fog/visibility = 4	4	3.05
<i>L (lighting conditions)</i>		
Daylight = 1	85	64.52
Dusk-dawn = 2	3	2.15
Dark-street lights = 3	44	33.33
<i>CS (crash severity)</i>		
Uninjured = 0	106	80.65
Injured = 1	25	19.35

4.2.1. XGBoost Model. The XGBoost model is a boosting algorithm, which generates multiple weak learning classifiers through fitting residual and finally accumulates the generated weak learning classifier to obtain the strong learning classifiers [31]. Chen and Guestrin made some improvements based on the gradient boosting [32] and presented XGBoost in 2016. XGBoost expands second-order Taylor of the loss function in the optimization process and introduces the second-order derivative information, which makes the model converge faster in the training process. In addition, XGBoost also adds a regularization term to the loss function to suppress model complexity and prevent overfitting. The regularized objective L_k for the k^{th} iteration can be expressed in the following equation:

$$L_k = \sum_{i=1}^n l(y^{(i)}, \hat{y}_k^{(i)}) + \sum_{j=1}^k \Omega(f_j), \quad (4)$$

where n is the sample number, $\hat{y}_k^{(i)}$ is the prediction value of sample i at iteration k , and l is the original loss function.

Ω represents the regularization term, as shown in the following equation:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2. \quad (5)$$

Here, T is the number of leaf nodes, and γ and λ are two constants employed to constrain the degree of regularization.

Another development of XGBoost is the application of an additive learning approach that combines the most reliable tree model $f_k(x^i)$ into the current classification model to provide the m^{th} iteration prediction result. Therefore, equation (5) can be expressed further as follows:

$$L_k = \sum_{i=1}^n l(y^{(i)}, \hat{y}_{k-1}^{(i)} + f_k(x^{(i)})) + \Omega(f_k) + \sum_{j=1}^{k-1} \Omega(f_j). \quad (6)$$

Additionally, XGBoost utilizes the second-order Taylor expansion to the objective function, and equation (6) can be expressed further as

$$L_k = \sum_{i=1}^n \left[l(y^{(i)}, \hat{y}_{k-1}^{(i)} + g_i * f_k(x^{(i)}) + \frac{1}{2} h_i * f_k(x^{(i)})) \right] + \Omega(f_k) + C. \quad (7)$$

Here, $g_i = \frac{\partial}{\partial \hat{y}_{k-1}^{(i)}} l(y^{(i)}, \hat{y}_{k-1}^{(i)})$ and $h_i = \frac{\partial^2}{\partial \hat{y}_{k-1}^{(i)2}} l(y^{(i)}, \hat{y}_{k-1}^{(i)})$ are the first and second derivatives of the loss function, respectively, and C represents the constant.

4.2.2. CART Model. The CART model is a decision tree learning algorithm that uses the Gini index (as shown in equations (8) and (9)) to select and classify attributes [4]. The smaller the Gini value, the higher the purity of the dataset and the better the classification effect.

$$\text{Gini}(D) = 1 - \sum p_i^2, \quad (8)$$

$$\text{Gini_index}(D) = \sum \frac{|D_V|}{D} \text{Gini}(D_V). \quad (9)$$

Here, D is the dataset, p_i is the probability that category i appears in D , a is the attribute to be divided; v is the desirable value of attribute a , and D_V is the sample with the value of V in the dataset. The decision tree generated by the CART algorithm is a binary tree, and the segmentation steps are as follows:

- (1) Select the attribute with the minimum Gini coefficient as the segmentation point for each binary tree node
- (2) Select the optimal segmentation point for this node from all the optimal attribute values as this node's segmentation rule
- (3) Repeat the above steps to continue the left and right nodes' segmentation until all samples belong to the same category and stop the segmentation

4.2.3. SVM Model. Among various machine learning algorithms, SVM has been widely used in classification and regression research. The core idea of the algorithms is to construct the optimal hyperplane to improve classification precision. For a linear classification problem, two parallel hyperplanes are chosen to maximize the distance between different classes. For nonlinear classification problems, the classification problem is transformed into higher-dimensional space by using three basic kernel functions involving linear, polynomials, and Gaussians. Take the binary classification as an example, and there is a training set as the following:

$$(x_i, y_i), \quad i = 1, 2, \dots, n, \quad x_i \in R^d, \quad y_i = \pm 1, \quad (10)$$

where y_i is the category label to which the sample belongs to, d is the dimension of the sample, and n is the number of training samples. An optimal separating hyperplane is calculated by the following formula:

$$H: w^T * x + b = 0, \quad (11)$$

where w is an n -dimensional vector and b is the offset. The label of a sample can be represented by the following equation:

$$f(\bar{x}) = \begin{cases} 1 & \text{if } w^T * x + b \geq 1, \\ -1 & \text{if } w^T * x + b \leq -1. \end{cases} \quad (12)$$

4.2.4. Evaluation Indicator. The confusion matrix, also called the error matrix, is a multidimensional measurement indicator representing accuracy evaluation. It is mainly used to compare the classification result with the actual measured value and display the classification's accuracy in a confusion matrix (see Table 4) [33].

The overall accuracy is measured as follows:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}}. \quad (13)$$

Regrettably, this indicator may not apply to unbalanced data. Because the number of injuries in the current study is significantly lower than the number of noninjuries, overall accuracy can be high even if all minority instances are misclassified. We used the G -mean to evaluate the classification accuracy of unbalanced data. It is considered a reasonable indicator for assessing unbalanced data by balancing the classification accuracy of minority and majority instances [34]. The G -mean is calculated as follows:

$$G - \text{mean} = \sqrt{\frac{\text{TP}}{\text{TP} + \text{FN}} * \frac{\text{TN}}{\text{TN} + \text{FP}}}. \quad (14)$$

The recall rate indicates the classification accuracy of minority instances, as shown in the following equation:

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (15)$$

Finally, we choose the G -mean and recall rate as indicators to measure the model performance because we need to identify injury accidents as much as possible.

TABLE 4: Confusion matrix.

	Predicted positive	Predicted negative
Actual positive	True positive (TP)	False negative (FN)
Actual negative	False positive (FP)	True negative (TN)

4.3. Model Results' Interpretation. The objective of establishing the crash severity classification model is to reveal the relationship between each factor and the crash severity. Subsequently, corresponding countermeasures can be implemented to reduce the crash severity. Therefore, the interpretability of the model output is as important as its accuracy.

We apply the SHAP to interpret how a specific variable influences model classification results. It was proposed by Lundberg and Lee and originated from the cooperative game theory [15]. It produces a predicted value for each sample, and the SHAP value is allocated to each feature in the sample. The importance of each feature on the model output $v(N)$ (ϕ_i is the importance of feature i) is assigned based on its boundary contributions. The following equation calculates the Shapley values:

$$\Phi_i = \sum_{S \subseteq N(i)} \frac{|S|!(n-|S|-1)!}{n!} [v(S \cup \{i\}) - v(S)]. \quad (16)$$

The linear function of binary feature g is determined according to the following formula:

$$g(z') = \phi_0 + \sum_{i=1}^M \phi_i z'_i. \quad (17)$$

Here, $z' \in \{0, 1\}^M$, when the feature is detected, it is equal to 1; otherwise, it is equal to 0, and M is the number of input features.

5. Results and Discussion

5.1. Association Rule Analysis. In previous studies, the threshold of L was usually set at 1, and the threshold of S and C was usually set at 10–20% [24]. In the current study, association rules were generated by setting the min-support equal to 0.2, min-confidence equal to 0.2, and min-lift equal to 1. A total of 90 association rules were generated. As shown in Figure 3, it is a bubble chart of all association rules. The following will be an excavation analysis of uninjured and injured crashes, respectively.

5.1.1. Association Rules for Uninjured Crashes. There were 106 uninjured crashes in the dataset, accounting for 80.65%. We sorted the obtained rules according to the confidence index and deleted the rules with no obvious value. Finally, we obtained the top 10 strong association rules, as shown in Table 5.

The above results show that the ten strong association rules for uninjured accidents have a high overlapping rate of factors and are the most common factors under related variables. This rule conforms to objective facts and proves the rationality of the results of association rules from the

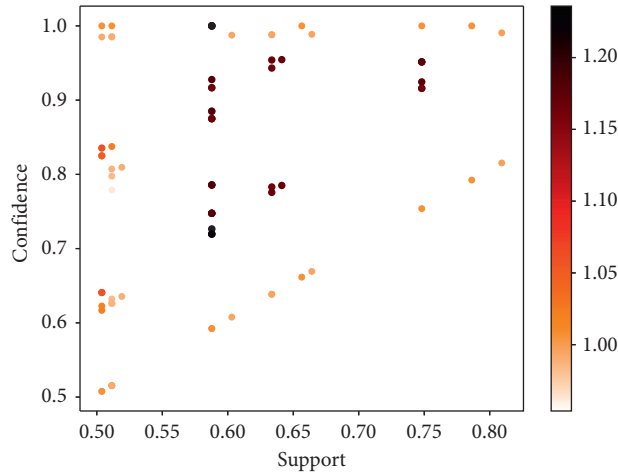


FIGURE 3: Bubble chart for association rules.

TABLE 5: Strong association rules of uninjured crashes.

Rule ID	Association rules		S	C	L
	Antecedents	Consequents			
1	W = 1 + TOC = 2	PMCV = 1	0.570	1.000	1.009
2	L = 1 + W = 1	PMCV = 1	0.589	1.000	1.009
3	DM = 0 + W = 1	PMCV = 1	0.514	1.000	1.009
4	TOC = 2 + VD = 1	PMCV = 1	0.523	1.000	1.009
5	PMAV = 0 + TOC = 2	PMCV = 1	0.598	1.000	1.009
6	DM = 0	PMCV = 1	0.533	0.983	0.992
7	PMAV = 0 + PMCV = 1	TOC = 2	0.598	0.970	1.059
8	DM = 0 + PMCV = 1	W = 1	0.514	0.965	1.054
9	DM = 0	PMCV = 1 + W = 1	0.514	0.948	1.035
10	L = 1 + PMCV = 1	W = 1	0.589	0.940	1.027

side. Among the factors involved in the ten strong association rules for uninjured crashes, “TOC = 2,” “PMCV = 1,” “L = 1,” and “W = 1” are the most frequent itemsets. It shows that most uninjured crashes occurred in ordinary situations. According to rules 5 and 7, the precrash movement of AVs has stopped (“PMAV = 0”), and then the conventional vehicle is still moving forward (“PMCV = 1”), and the type of collision is rear-ended (“TOC = 2”). It shows that most rear-end crashes are conventional vehicles bumping into the rear of AVs, consistent with previous research results [35].

5.1.2. Association Rules for Injury Crashes. There were 25 injury crashes in the dataset, accounting for 19.35%. The same as the previous section, the top 10 strong association rules are shown in Table 6.

The top ten strong association rules for injury crashes involve more factors than uninjured crashes from the above results. There are some unconventional factors (such as “W = 2,” “VD = 2,” and “DM = 1”), indicating that unfavorable conditions often accompany the occurrence of injury accidents. This rule conforms to objective facts and proves the rationality of the results of association rules from the side. Compared with Table 5, injury crashes are more likely to occur on cloudy days (“W = 2”). According to rules

1 and 6, the AVs were still moving before the collision on a cloudy day indicating that the autonomous driving system failed to apply emergency measures. It could be caused by the detector failing to find anomalies in time when there is insufficient light. Besides, the driver may distract their attention to secondary tasks when the drive mode is automatic, so it is more likely to be injured in an accident (rules 7 and 9).

5.2. Classification Model Results. We divided the test set and training set according to the ratio of 7 : 3. We used the grid search to decide the best combination of parameters to prevent overfitting of the model. The performances for each classification model are reported in Table 7.

Table 7 shows that the XGBoost model represented better than the other models because it has the highest G-mean and recall ($G\text{-mean} = 67.82\%$; $\text{recall} = 75.00\%$). Although the overall accuracy is low (overall accuracy = 61.10%), the G-mean and recall are the main indicators for unbalanced data because we want to recognize more injury crashes. In summary, the XGBoost model can more accurately identify injured accidents. Then, we used the SHAP to explain the results of the XGBoost model. Figure 4 shows the impact of features on AV-involved crashes. The ordinate represents the features and is sorted

TABLE 6: Strong association rules for injury crashes.

Rule ID	Association rules		S	C	L
	Antecedents	Consequents			
1	$W = 2 + PMAV = 1$	$CS = 1$	0.792	1	1
2	$W = 2 + CS = 1$	$PMCV = 1$	0.792	1	1
3	$VD = 2 + TOC = 2$	$CS = 1$	0.667	1	1
4	$VD = 2 + W = 2$	$CS = 1$	0.625	1	1
5	$TOC = 2$	$PMCV = 1 + CS = 1$	0.792	1	1
6	$L = 2 + PMAV = 1 + W = 2$	$CS = 1$	0.583	0.680	1
7	$DM = 1 + TOC = 2$	$CS = 1$	0.542	0.627	1
8	$AL = 1$	$PMCV = 1 + CS = 1$	0.542	0.533	1
9	$DM = 1 + PMCV = 1$	$CS = 1$	0.583	0.507	1
10	$W = 2 + L = 1 + CS = 1$	$PMCV = 1$	0.583	0.507	1

TABLE 7: Model testing results.

Models	Evaluation indicator		
	Overall accuracy (%)	G-mean (%)	Recall (%)
XGBoost	64.10	67.82	75.00
CART model	66.67	65.07	62.50
SVM model	71.79	55.00	30.75

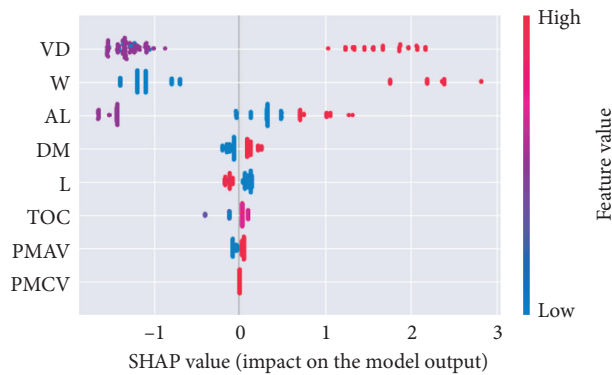


FIGURE 4: Impact of features on AV-involved crashes.

according to the importance, and the abscissa is the Shapley value. Each point in the figure represents a sample. Color represents the size of the feature value.

Obviously, the greater the degree of vehicle damage is, the more likely it is to cause a severe accident. The interpretation of the model results suggests that weather is another key feature. Particularly, in low-visibility conditions, such as fog and snow, injury accidents are more likely to occur [36]. This is probably due to the sensors' worse perception performance in extreme weather [37]. According to Hasirlioglu et al. [38], the reflector can only detect a short distance in foggy weather, and crashes are more likely to occur in this situation.

The next most important features are accident location and driving mode. The possibility of a crash is higher at intersections [39]. This is probably due to the complex and changeable traffic environment at intersections because vehicles, nonmotor vehicles, and pedestrians are highly mixed [40, 41]. According to crash reports, AVs usually switch to the conventional driving mode when they arrive at intersections because intelligent transportation facilities are

not perfect enough now. These infrastructures can increase vehicle stability during driving and improve the safety of all traffic participants [42]. In terms of driving mode, Figure 4 illustrates that the automatic driving mode will increase the risk of injury, surprising. This is probably because the driver diverts their attention to secondary tasks (such as playing mobile phones) in the automatic driving mode, so it is more likely to be injured in an accident.

6. Conclusion

The objective of this study is to explore the mechanism of AV-involved crashes and analyze the impact of each feature on crash severity. We employ 131 accident reports involving AVs received in California from 2019 to October 2020. We use the Apriori algorithm to explore the causal relationship between multiple factors to explore the mechanism of crashes. Given the imbalanced crash severity distribution, we apply the SMOTE to balance the dataset. Three different classification models are used to compare the classification performance: XGBoost, CART, and SVM. The result shows

that the XGBoost model can better recognize the injured crashes involving AVs. We apply the SHAP (Shapley Additive Explanations) to interpret how a specific variable influences model classification results. Both the XGBoost and the Apriori algorithm effectively provided meaningful insights about AV-involved crash characteristics and their relationship.

For the analysis of crash mechanisms involving AVs, we use the Apriori algorithm to mine association rules for uninjured and injured crashes, respectively. Among the top ten strong association rules for uninjured crashes, we can find that most rear-end collisions are conventional vehicles bumping into the rear of AVs. It is probably because the AVs have stopped before the collision, while the conventional vehicles are still moving forward (“PMAV = 0 + TOC = 2” → “PMCV = 1”). Among the top ten strong association rules for injured crashes, we can find that the AVs were still moving before the collision on a cloudy day. It could be caused by the detector failing to find anomalies in time when there is insufficient light (“L = 2 + PMAV = 1 + W = 2” → “CS = 1”). Besides, the driver may distract their attention to secondary tasks in the automatic driving mode, so it is more likely to be injured in an accident (“DM = 1 + TOC = 2” → “CS = 1”). For the crash severity analysis, XGBoost generates the best result (overall accuracy = 64.10%, G-mean = 67.82%, and recall = 75%). To make the results of the XGBoost model more informative, we apply the SHAP to analyze the impact of each feature on crash severity. Among all these features, vehicle damage, weather conditions, accident location, and driving mode are the most critical features. The greater the degree of vehicle damage is, the more likely it is to cause a severe accident. Injured accidents are more possible to occur in low-visibility conditions (such as fog and snow). Intersections are more prone to injury accidents, especially in the automatic driving mode. This study may provide some help in reducing the severity of AV-involved crashes. For example, autonomous vehicle drivers should be extremely cautious when driving in low-visibility conditions (e.g., fog and snow). They should be more careful when driving near intersections, especially in the autonomous driving mode. It is recommended to use vehicle sensors with strong stability and high sensitivity.

However, the current study has certain limitations. Firstly, this study’s sample size and variables could be extended to increase the model result’s reliability. In the future, we will collect driver characteristics, traffic flow information before the crash, and vehicle speed to understand the mechanism of AV-involved crashes. Secondly, this study uses only crash reports received in California for modeling. Future research should continue to collect accident data from other countries and regions because driving habits and traffic laws in different countries and regions may be completely different.

Data Availability

The data used to support the findings of this study are publicly available at https://www.dmv.ca.gov/portal/dmv/detail/vr/autonomous/autonomousveh_0l316.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the 111 Project of Sustainable Transportation for Urban Agglomeration in Western China (no. B20035).

References

- [1] A. B. Parsa, A. Movahedi, H. Taghipour, S. Derrible, and A. Mohammadian, “Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis,” *Accident Analysis & Prevention*, vol. 136, p. 105405, 2020.
- [2] J. Stilgoe, “Machine learning, social learning and the governance of self-driving cars,” *Social Studies of Science*, vol. 48, no. 1, pp. 25–56, 2018.
- [3] L.-Y. Chang and J.-T. Chien, “Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model,” *Safety Science*, vol. 51, no. 1, pp. 17–22, 2013.
- [4] H. Jeong, Y. Jang, P. J. Bowman, and N. Masoud, “Classification of motor vehicle crash injury severity: a hybrid approach for imbalanced data,” *Accident Analysis & Prevention*, vol. 120, pp. 250–261, 2018.
- [5] C. Xu, Z. Ding, C. Wang, and Z. Li, “Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes,” *Journal of Safety Research*, vol. 71, pp. 41–47, 2019.
- [6] S. Wang and Z. Li, “Exploring the mechanism of crashes with automated vehicles using statistical modeling approaches,” *PLoS One*, vol. 14, pp. 1–16, 2019.
- [7] A. M. Boggs, B. Wali, and A. J. Khatkhat, “Exploratory analysis of automated vehicle crashes in California: a text analytics & hierarchical Bayesian heterogeneity-based approach,” *Accident Analysis & Prevention*, vol. 135, Article ID 105354, 2020.
- [8] H. Chen, H. Chen, Z. Liu, X. Sun, and R. Zhou, “Analysis of factors affecting the severity of automated vehicle crashes using XGBoost model combining POI data,” *Journal of Advanced Transportation*, vol. 2020, Article ID 8881545, 12 pages, 2020.
- [9] C. Xu, J. Bao, C. Wang, and P. Liu, “Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China,” *Journal of Safety Research*, vol. 67, pp. 65–75, 2018.
- [10] Z. Liu, H. Chen, Y. Li, and Q. Zhang, “Taxi demand prediction based on a combination forecasting model in hotspots,” *Journal of Advanced Transportation*, vol. 2020, Article ID 1302586, 13 pages, 2020.
- [11] B. Zhou, Z. Li, S. Zhang, X. Zhang, X. Liu, and Q. Ma, “Analysis of factors affecting hit-and-run and non-hit-and-run in vehicle-bicycle crashes: a non-parametric approach incorporating data imbalance treatment,” *Sustain*, vol. 11, pp. 1–14, 2019.
- [12] S. R. Mousa, P. R. Bakhit, and S. Ishak, “An extreme gradient boosting method for identifying the factors contributing to crash/near-crash events: a naturalistic driving study,” *Canadian Journal of Civil Engineering*, vol. 46, no. 8, pp. 712–721, 2019.

- [13] F. Mannering, "Temporal instability and the analysis of highway accident data," *Analytic Methods in Accident Research*, vol. 17, pp. 1–13, 2018.
- [14] B. Zhou, X. Wang, S. Zhang et al., "Comparing factors affecting injury severity of passenger car and truck drivers," *IEEE Access*, vol. 8, pp. 153849–153861, 2020.
- [15] K. Zhang, Y. Zhang, and M. Wang, "A unified approach to interpreting model predictions scott," *Nips*, vol. 16, pp. 426–430, 2012.
- [16] W. Haddon, "The changing approach to the epidemiology, prevention, and amelioration of trauma: the transition to approaches etiologically rather than descriptively based," *Injury Prevention*, vol. 5, no. 3, pp. 231–235, 1999.
- [17] A. Omidvar, M. Pourmehr, P. Emami et al., "Deployment and testing of optimized autonomous and connected vehicle trajectories at a closed-course signalized intersection," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 19, pp. 45–54, 2018.
- [18] X. Li, Z. Sun, D. Cao, D. Liu, and H. He, "Development of a new integrated local trajectory planning and tracking control framework for autonomous ground vehicles," *Mechanical Systems and Signal Processing*, vol. 87, pp. 118–137, 2017.
- [19] M. Zhu, H. Chen, and G. Xiong, "A model predictive speed tracking control approach for autonomous ground vehicles," *Mechanical Systems and Signal Processing*, vol. 87, pp. 138–152, 2017.
- [20] J. C. F. Winter, R. Happee, M. H. Martens, and N. A. Stanton, "Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 27, pp. 196–217, 2014.
- [21] A. K. Agarwal, S. Wadhwa, and S. Chandra, "Diagnosis of tuberculosis--newer tests," *Journal of the Association of Physicians of India*, vol. 42, no. 8, p. 665, 1994.
- [22] P. Wu, X. Meng, L. Song, and W. Zuo, "Crash risk evaluation and crash severity pattern analysis for different types of urban junctions: fault tree analysis and association rules approaches," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 1, pp. 403–416, 2019.
- [23] S. Yu, Y. Jia, and D. Sun, "Identifying factors that influence the patterns of road crashes using association rules: a case study from Wisconsin, United States," *Sustain*, vol. 11, pp. 1–14, 2019.
- [24] J. Weng and G. Li, "Exploring shipping accident contributory factors using association rules," *Journal of Transportation Safety & Security*, vol. 11, no. 1, pp. 36–57, 2019.
- [25] A. Montella, M. Aria, A. D'Ambrosio, and F. Mauriello, "Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery," *Accident Analysis & Prevention*, vol. 49, pp. 58–72, 2012.
- [26] A. S. Mihăiță, Z. Liu, C. Cai, and M. A. Rizoiu, "Arterial incident duration prediction using a bi-level framework of extreme gradient-tree boosting," 2019, <http://arxiv.org/abs/1905.12254>.
- [27] D. A. Belsley, "A Guide to using the collinearity diagnostics," *Computer Science in Economics and Management*, vol. 4, pp. 33–50, 1991.
- [28] C. H. Mason and W. D. Perreault, "Collinearity, power, and interpretation of multiple regression analysis," *Journal of Marketing Research*, vol. 28, no. 3, p. 268, 1991.
- [29] T. Næs and B. H. Mevik, "Understanding the collinearity problem in regression and discriminant analysis," *Journal of Chemometrics*, vol. 15, pp. 413–426, 2001.
- [30] Y. Pan, S. Chen, T. Li, S. Niu, and K. Tang, "Exploring spatial variation of the bus stop influence zone with multi-source data: a case study in Zhenjiang, China," *Journal of Transport Geography*, vol. 76, pp. 166–177, 2019.
- [31] M. Petrere, "Pesque-solte," *Ciência Hoje*, vol. 53, pp. 1189–1232, 2014.
- [32] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367–378, 2002.
- [33] X. Li, J. Tang, X. Hu, and W. Wang, "Assessing intercity multimodal choice behavior in a Touristy City: a factor analysis," *Journal of Transport Geography*, vol. 86, Article ID 102776, 2020.
- [34] A. Bradley, "The use OF the area under the roc curve IN the evaluation OF machine learning algorithms," *Transportation Research Part A: Policy and Practice*, vol. 44, pp. 291–305, 1997.
- [35] S. Wang and Z. Li, "Exploring causes and effects of automated vehicle disengagement using statistical modeling and classification tree based on field test data," *Accident Analysis & Prevention*, vol. 129, pp. 44–54, 2019b.
- [36] X. Yu and M. Marinov, "A study on recent developments and issues with obstacle detection systems for automated vehicles," *Sustain*, vol. 12, 2020.
- [37] R. H. Rasshofer, M. Spies, and H. Spies, "Influences of weather phenomena on automotive laser radar systems," *Advances in Radio Science*, vol. 9, pp. 49–60, 2011.
- [38] S. Hasirlioglu, A. Riener, W. Ruber, and P. Wintersberger, "Effects of exhaust gases on laser scanner data quality at low ambient temperatures," in *Proceedings of the IEEE Transactions on Intelligent Transportation Systems*, pp. 1708–1713, Los Angeles, CA, USA, 2017.
- [39] V. Gadepally, A. Krishnamurthy, and Ü. Özgüner, "A framework for estimating long term driver behavior," *Journal of Advanced Transportation*, vol. 2017, Article ID 3080859, 11 pages, 2017.
- [40] L. Hu, J. Ou, J. Huang, Y. Chen, and D. Cao, "A review of research on traffic conflicts based on intelligent vehicles," *IEEE Access*, vol. 8, pp. 24471–24483, 2020.
- [41] M. M. Morando, Q. Tian, L. T. Truong, and H. L. Vu, "Studying the safety impact of autonomous vehicles using simulation-based surrogate safety measures," *Journal of Advanced Transportation*, vol. 2018, Article ID 6135183, 11 pages, 2018.
- [42] O. Grembek, A. Kurzhanskiy, A. Medury, P. Varaiya, and M. Yu, "Making intersections safer with I2V communication," *Transportation Research Part C: Emerging Technologies*, vol. 102, pp. 396–410, 2019.