Computers in Biology and Medicine 153 (2023) 106393

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Machine learning-based injury severity prediction of level 1 trauma center enrolled patients associated with car-to-car crashes in Korea

Joon Seok Kong [a,b], Kang Hyun Lee [a,b,*], Oh Hyun Kim [a,b], Hee Young Lee [a,b], Chan Young Kang [a,b], Dooruh Choi [a,b], Sang Chul Kim [c], Hoyeon Jeong [d], Dae Ryong Kang [d], Tae-Eung Sung [e]

[a] *Center for Automotive Medical Science Institute, Yonsei University, Wonju College of Medicine, Wonju, 26426, Republic of Korea*
[b] *Department of Emergency Medicine, Yonsei University, Wonju College of Medicine, Wonju, 26426, Republic of Korea*
[c] *Department of Emergency Medicine, Chungbuk National University, Cheongju, 28646, Republic of Korea*
[d] *Department of Precision Medicine and Biostatistics, Yonsei University, Wonju College of Medicine, Wonju, 26426, Republic of Korea*
[e] *Department of Computer and Telecommunication Engineering, Yonsei University, College of Science and Technology, Wonju, 26493, Republic of Korea*

ARTICLE INFO

Keywords:
Injury severity prediction
Machine learning
Car-to-car crashes
Motor vehicle occupants
Trauma injury
Korea in-depth accident study

ABSTRACT

Injury prediction models enables to improve trauma outcomes for motor vehicle occupants in accurate decision-making and early transport to appropriate trauma centers. This study aims to investigate the injury severity prediction (ISP) capability in machine-learning analytics based on five-different regional Level 1 trauma center enrolled patients in Korea. We study car crash-related injury data of 1417 patients enrolled in the Korea In-Depth Accident Study database from January 2011 to April 2021. Severe injury classification was defined using an Injury Severity Score of 15 or greater. A planar crash was considered by excluding rollovers to compromise an accurate prediction. Furthermore, dissimilarities of the collision partner component based on vehicle segmentation were assumed for crash incompatibility. To handle class-imbalanced clinical datasets, we used four data-sampling techniques (i.e., class-weighting, resampling, synthetic minority oversampling, and adaptive synthetic sampling). Machine-learning analytics based on logistic regression, extreme gradient boosting (XGBoost), and a multilayer perceptron model were used for the evaluations. Each model was executed using five-fold cross-validation to solve overfitting consistent with the hyperparameters tuned to improve model performance. The area under the receiver operating characteristic curve of 0.896. Additionally, the present ISP model showed an under-triage rate of 6.1%. The Delta-V, age, and Principal ~ were significant predictors. The results demonstrated that the data-balanced XGBoost model achieved a reliable performance on injury severity classification of emergency department patients. This finding considers ISP model selection, which affected prediction performance based on overall predictor variables.

1. Introduction

In 2018, the World Health Organization reported that more than 1.35 million global deaths were caused by road traffic injuries [1]. Furthermore, the report claimed that 20–50 million patients sustained non-fatal injuries. Motor vehicle crashes (MVCs) are the single leading cause of mortality among traumatic injuries and are a significant cause of sudden unnatural death in the United States [2]. Although the overall incidence of road crashes has decreased worldwide, the ratio of casualties does not correspond to this decrease.

Predicting the injury severity of motor vehicle occupants (MVOs) is significantly crucial in saving trauma patients. It has been reported that patients with severe injuries transferred early to trauma centers lead to a 25% reduction in mortality [3]. During the pre-hospital stage, accurate classification of crash-related injury severity is essential for decision making for patients and their transfers to appropriate facilities [4]. Paramedics refer to various field triage recommendations to determine the injury classification of trauma patients [5,6]. Despite long clinical efforts, securing indicators (e.g., crash velocity or crash deformations) in pre-hospital trauma triage for MVOs in critical rescue circumstances has been problematic [7]. Although emergency medical services (EMS) are expected to proceed with short notice, the actual "golden hour" of

* Corresponding author. Department of Emergency Medicine, Yonsei University, Wonju College of Medicine, 20, Ilsan-ro, Wonju, 26426, Republic of Korea.
E-mail address: ed119@yonsei.ac.kr (K.H. Lee).

https://doi.org/10.1016/j.compbiomed.2022.106393
Received 27 January 2022; Received in revised form 19 November 2022; Accepted 29 November 2022
Available online 9 December 2022
0010-4825/

survival is not always observed in traditional procedures [8]. Vehicular telematics services have recently been leveraged to provide collision information to first responders, and their use is increasing in high-income countries [6,9] to overcome these issues. However, the application of these advanced technologies must be premised upon in-depth clinical research [10].

In terms of reducing fatalities, the unique characteristics of each country's high-risk crashes should be considered. In Korea, severe injuries from MVCs are aggregated in large numbers among road users overall. In particular, car-to-car (C2C) crashes account for a significant proportion of crashes leading to major traumatic injury risks. Many studies have shown that crash incompatibility between two vehicles significantly affects injury severity [11–18]. This indicates a considerable difference in vehicle design regarding mass and size, geometry, and stiffness. Consequently, in contrast to single-vehicle crashes, the significant factor affecting the presence of severe injury in C2C crashes is vehicle dissimilarity. However, most injury severity prediction (ISP) studies have focused only on the overall collide materials [19–21]. Individual crash types are significant in determining crash-related injury outcomes. However, predictive estimations of injury severity focusing on C2C crashes have not been explored.

A numerical model of ISP was provided using traditional statistics [19,21,22] in a previous study. Early predictive models (e.g., logistic regression) have the advantage of intuitive and interpretable structures [23]. The predictive performance of these algorithms depends on the sample size. Thus, it is difficult to expect good performance when there is an insufficient amount of clinical data [24]. However, the recent use of machine learning (ML) has provided an alternative that might overcome these limitations [20,24,25]. In extant works, ML models used to predict injury severity classification have reported better performance than traditional statistical models [26]. However, there is no single optimal model for predicting injury severity classifications for trauma-injured MVOs [25]. Thus, it is necessary to determine the performance of various ISP models for this purpose.

This study aims to provide ISP models using ML analytics for MVOs who have visited Level 1 trauma centers in Korea. The study suggests the noteworthy by following as 1) a primary ISP model focused on C2C crashes, 2) handling imbalanced injury severity classification based on data sampling techniques, 3) comparing the optimal model by considering an under-triage in medical point-of-view, and 4) provides the feature importance of single outperforming evaluation model. Thus, rather than simply focusing on improving the predictive performance, it is vital to represent a clinically reliable model for medical-related workers in the real world.

The remainder of the paper is organized as follows.; Section 2 describes the datasets and detailed framework methodologies; Section 3 presents the results; Sections 4 and 5 present the discussion and limitations of this study; Finally, Section 6 outlines the main conclusions and presents the scope for future research.

2. Methods

This study applied ML analytics through imbalanced clinical data processing to determine the best performing model according to binary injury classification. The overall methodological procedure is illustrated in Fig. 1. We pre-processed the class-imbalanced data using oversampling techniques to achieve results that reduced the defects of the training dataset. All models were verified using k-fold cross-validation to avoid overfitting problems. A detailed methodological description is described in the following subsections.

2.1. Data source

This retrospective study used the Korea In-Depth Accident Study (KIDAS) database of the Center for Automotive Medical Science Institute at Yonsei University. The data were collected using on-scene investigations of real-world crashes. We analyzed patients who visited five different regional trauma centers in South Korea from January 2011 to April 2020. The dataset consists of road traffic injury information related to the human, vehicle, and crash components to predict MVOs' injury severity. The patients' age was considered in adults in both males and females. The restraint system (passive safety device) was considered in 3-pointed seatbelts wearable. Furthermore, the Principal Direction of Force (PDOF) was defined as impact direction. This consists of frontal, side (including left and right), and rear-end impacts. Vehicle types were categorized into five different subjects of sedan, sports utility vehicle, light truck, van, and heavy trailers. Collision partner was defined as considering the two-vehicles similarity of mass and size considering crash incompatibility. For instance, in case of the patient's vehicle was heavier than the opponent's crash vehicle, the selection has been made

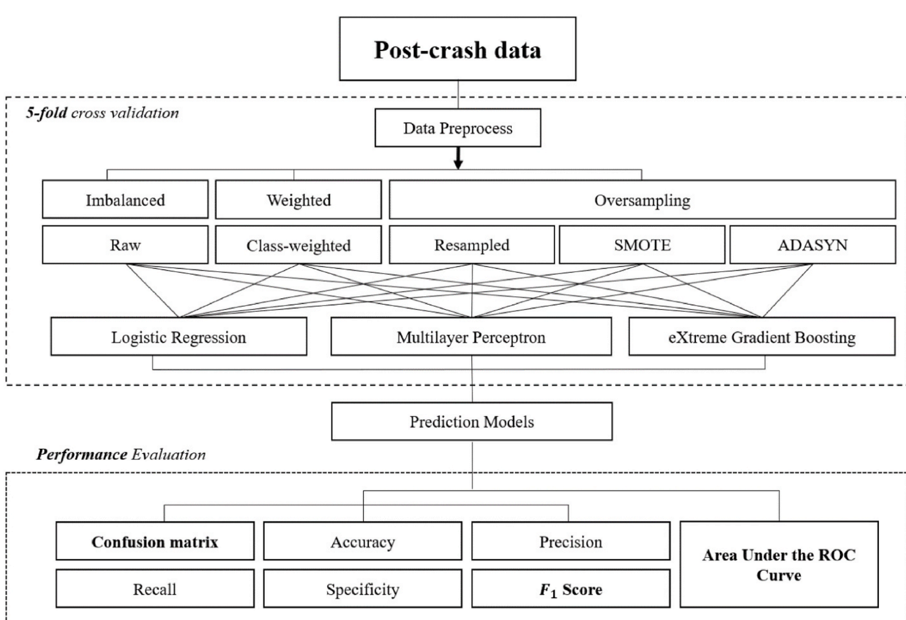

Fig. 1. Prediction model analytics of MVC occupant injuries.

as a relatively smaller component. The multiple impacts were categorized into a single and more than twice of impacts in car-to-car crashes. Delta-V is a change of velocity during pre-crash to in-crash relating to vector dynamics of MVCs. The Delta-V was obtained by crash reconstruction using PC-Crash software referring to on-scene investigated information documented by field investigators. This study was conducted following approval from the research ethics committee of the Wonju Severance Christian Hospital at Yonsei University (IRB Approval No.: CR319049).

2.2. Population

Among the 3928 occupants related to MVCs, we used the data of 1417 patients aged 18 years or older to predict severe injuries in C2C crashes. We grouped individual patients based on the classification of injury severity. In this study, simple planar crashes were considered to predict the results based on the complexity of the MVC. Rollovers were excluded from the analysis.

2.3. Injury severity classification

The Injury Severity Score (ISS) is a medical score used to assess trauma severity established by the Association for Advanced Automotive Medicine [27]. The score provides a primary anatomical diagnosis for trauma, considering the epidemiological information needed to classify injury severity and determine treatment viability. An ISS score ranges from 0 to 75 and is assigned according to the abbreviated injury scale, which addresses six anatomical body regions: head and neck, face, thorax, abdomen, extremities, and externals. ISS is used extensively as a discriminant measure for predicting severe injury in MVCs. An ISS of 1–8 is considered minor, 9–15 moderate, and more than 15+ are calculated as severe to critical trauma. In this study, patients with an ISS of 15+ were categorized as severely injured based on the criteria for injury classification.

2.4. Predictive parameters

The selection of parameters used in the prediction model requires in-depth consideration of risk factors that affect the safety of occupants. The Centers for Disease Control and Prevention (CDC) in the US provided a recommendation from an expert panel on trauma-patient classification system guidelines for advanced automatic collision notification [10]. We adopted indicators such as these and the extant national standards for predicting injury severity. This study selected seven parameters (i.e., age, restraint usage, the principal direction of force (PDOF), vehicular type, collision number, crash partner, and Delta-V) to predict patients with severe injuries in C2C crashes.

2.5. Data sampling techniques for class-imbalance data

Effective predictive analytics requires a model that uses large-scale data consisting of neutrally balanced constituents. However, we considered that the injury severity classes based on the KIDAS dataset are generally imbalanced. These clinical datasets are frequently imbalanced due to the sample count depending on the number of patients visiting trauma centers with different variances of injury severities [28]. These features a strong bias for the prediction model's performance, which causes severe errors in diagnosis. Since the class imbalance problem occurs when the majority class has more data than the minority class [29], this can calculate the imbalance ratio (ratio of majority class to minority class) [30]. When precisely balanced, the class imbalance ratio is 1:1 however, a larger ratio implies a higher imbalanced dataset. This study considered the severity of the imbalanced clinical dataset as mildly imbalanced for a ratio between 1.9 and 9 and an extreme imbalance for a ratio higher than 9 [31,32]. The approaches to handling imbalanced class datasets were selected as data sampling techniques for class balancing. Class weighting, resampling, the synthetic minority oversampling technique (SMOTE), and adaptive synthetic sampling (ADASYN) were used in the present study.

Class weighting is not an oversampling methodology. However, it could be used to assign weights to each class to calculate the model's objective function. Resampling is a method of sampling minorities by replacing as many units as the number of majorities [33]. Despite the advantage of balancing classes, the technique increases the likelihood of overfitting as it replicates random records from the minority class. SMOTE and ADASYN were used to avoid overfitting by generating a newly synthesized minority class in a relatively wider region [34,35]. This can effectively change the sparse distribution of minority-class samples. SMOTE randomly generates synthetic minority instances that contain nearby instances of the minority class. ADASYN is a similar idea that assigns a weighted distribution for different minority class samples according to the density of majority class samples around the nearest neighbor's boundary.

Overall, the imbalanced data oversampling and predictive model development was performed using the Python programming language (version 3.8.2, Python Software Foundation, Wilmington, DE, USA), and the libraries used included scikit-learn 0.24.1, Imblearn 0.7.0, TensorFlow 2.3.1, and XGBoost 1.4.0 (version SNAPSHOT) in Table 1.

2.6. Classification models

This study used three ML classification techniques to develop a model to predict injury severity in MVCs. LR was the most widely used in prediction analysis; it is a classification algorithm used to assign observations to discrete response variables. The algorithm transforms the output using the logistic sigmoid function to return a probability value. MLP is a deep learning (DL) model suitable for handling heterogeneous variables in any order. The MLP is a stacked linear model wherein the activation function is generalized similarly to the LR model [36]. XGBoost is a decision tree ML model with a boot-strapping framework [35]. XGBoost parallelly processes sequential tree buildings. This method can prevent overfitting and improve calculation speed. Among tree-based models, the performance of this method is excellent, and the importance of the features can be determined.

2.7. Model training

The dataset was divided into 80% for training and 20% for testing. In the case of the LR and XGBoost models, five-fold cross-validation was applied in training, and a grid search was used for hyperparameter tuning. Though many k-fold may be used for validation, others were conducted similarly using short-scaled datasets [28]. In the case of the MLP model, the number of hidden layers was limited to two. The optimal values of the hyperparameters were tuned for each model.

2.8. Performance evaluation

In this study, the evaluation of presented ISP models was considered

Table 1
Hyperparameters used in the prediction model.

Logistic Regression	Multilayer Perceptron	Extreme Gradient Boosting
Penalty: l1 Solver: lbfgs	Number of hidden layers: 2 Activation function: ReLU Dropout: 0.3 Loss function: Binary Crossentropy Optimizer: Adam Epochs: 100 Batch size: 32	Booster: gbtree Max depth: 10 Min child weight: 2 Gamma: 1 Colsample bytree: 0.8 Colsample bylevel: 0.9 Number of estimators: 100

for assessing the internal validity in binary injury severity. We evaluate the proposed ML models using F-measures (F1 score) which are computed based on the harmonic average of precision and recall. Also, accuracy was calculated for performance comparison with other previous studies. These are defined as the following equations (1)–(4).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1\ Score = \frac{(2 \times (Precision \times Recall))}{(Precision + Recall)} = \frac{(2 \times TP)}{((2 \times TP) + FP + FN)} \tag{4}$$

where true positive (TP), the number of actual events of severely injured patients is classified as severe injury, true negative (TN), the number of events of non-severe injured patients counted as non-severe injury, false positive (FP), the number of non-severe injured patients detected as severely injured, and false negative (FN), the number of events of severely injured presents as non-severely injury, respectively (see Table 2).

However, standard errors of false alarms represent misleading predictions, such as over-triage (false positive ratio) and under-triage (false negative ratio) classification. This study considered an under-triage levels prior to evaluating the predictive performance in clinical assessments.

Using the receiver operating characteristics curve (ROC) value, we conducted a performance evaluation for a primary classifier based on ML analytics. The curve plots the true positive rate (TPR) against the false negative rate (FPR) illustrating the predictive performance of a binary classifier. The TPR also represents equal calculating equation as recall (or sensitivity), and FPR can calculate as (1-specificty).

The AUC values ranged from 0.5 to 1. Hosmer and Lemeshow defined the evaluation of AUC as a "no discrimination" outcome when the AUC is 0.5; it is an acceptable discrimination outcome when $0.7 \leq AUC < 0.8$, and an excellent discrimination outcome occurs when $0.8 \leq AUC < 0.9$. Furthermore, an outstanding discrimination outcome occurs when the $AUC \geq 0.9$. As the AUC approaches 1.0, the response can be interpreted as a complete predictive power outcome [23].

3. Results

Scatterplots of primary continuous data (age and Delta-V) were used by each sampling technique to configure the data distribution of binary injury severities (Fig. 2). Since the oversampling was conducted only on severely injured data, the plots show an increase focusing on Resample, SMOTE, and ADASYN datasets compared to imbalance distribution. However, the class-weighted datasets demonstrate as equal to raw data due to assigning weights to each class in initial data. The degree of spread and central tendency of the sampling data was similar in cases of imbalanced and oversampled datasets. The Delta-V distribution showed a significant spread in severe injuries, whereas the central tendency of non-severe patients was focused on the low–middle range.

Table 2
Model intra-validation associated with the prediction model and traumatic clinical data abbreviation.

	Actual Positive (Severe Injury)	Actual Negative (Non-severe Injury)
Predicted Positive (Severe Injury)	True Positive (TP) (Hits)	False Positive (FP) (Over-triage)
Predicted Negative (Non-severe Injury)	False Negative (FN) (Under-triage)	True Negative (TN) (Reject)

The descriptive data were summarized as a sample for predicting severely injured C2C crash occupants (Table 3). According to the classification of injury severity, the data distribution led to performance outcomes that were nearly five times higher in the non-severe group (n = 1,181, 83.3%) than in the severe group (n = 236, 16.7%). Among the patients with the majority and minority classes, the imbalance ratio showed nearly 5:1, which is a mildly imbalanced dataset (1.9-to-9.0). Since the dataset has not satisfied an extremely imbalanced ratio (>9), it is more likely to be appropriate for predicting the majority of classes in clinical data. Also, the data indicated that young occupants were more engaged with MVCs than elderly groups.

The proportion of frequency of restrained occupants at the time of the MVC was larger. The PDOF was the largest in cases of frontal (e.g., head-on) impacts. In terms of vehicular type, the sedan met with the highest number of crashes, followed by sport-utility vehicles (SUVs) and light trucks. In this study, we classified the relative sizes of the counterparts into three categories. The incidence of impact with vehicles similar to or larger than those of the counterpart vehicle was higher.

Regarding the number of collisions, the probability of multiple impacts was lower than 10% of all MVCs. The Delta-V accounted for nearly 70% at the low and medium ranges (0–30 km/h). We developed a model to predict the severity of damage in patients based on age and Delta-V distribution.

This study assessed 15 models to predict severe injury based on the oversampling techniques of class-imbalanced MVC data. The confusion matrix of the present model was analyzed using five-fold cross-validation (Table 4). The sampling data (Resample, SMOTE, and ADASYN) oversampled nearly twice as high as the raw and weighted dataset. In addition, the Resampling and SMOTE oversampled the most sampling numbers than ADASYN. The number of samples used for ML in each dataset was identical.

A crucial role of classification problems in ML predictions may be visualized as a confusion matrix that shows the classification model being confused with the prediction. The number of correct (positive) and incorrect (negative) predictions of binary classifiers (severe or non-severe injury) is summarized with count values and broken down by each class. However, a significant error of false alarms represents misleading predictions as over-triage (false positive ratio) or under-triage (false negative ratio) in clinical outcomes. The false-negative rate (severe injury) should be considered within the lowest peak for an accurate model to avoid under-triage in MVOs classifications. This study found the best-performed model with lower bounds of the under-triage-rated model in imbalanced data (MLP = 2.5%). However, the oversampled data-enhanced prediction of severely injured patients included a good under-triage tolerance of <10%.

Table 5 shows the classification performance of injury severity results obtained from the confusion matrices for each sampled classifier in Table 4. Thanks to these matrices, it has been determined how injury severities were predicted correctly by referring to Table 2. It is clear that with the proposed method, the least incorrect injury severity estimation is made. According to the performance findings, the outperformed classifier of SMOTE-XGBoost model achieved the accuracy, precision, recall, and F1 measures as 83.1%, 81.3%, 88.4%, and 84.7%, respectively. From the obtained results, we can observe that SMOTE and ADASYN have similar performance, although the outperformed classifiers are machine learning models (especially in XGBoost) based on SMOTE sampled dataset.

The calculation of the performance parameter (accuracy, precision, recall, and F1 score) for the outperformed model SMOTE (LR, MLP, and XGB) of Table 5 have been obtained using the metrics columns in Table 4. For instance, if we analyze the same oversampled SMOTE dataset of LR from the confusion matrix parameters as given in Table 4, TP = 161, FN = 77, FP = 80, and TN = 155. Thus, the $\text{accuracy} = (161 + 155)/\ (161 + 155 + 80 + 70) = 0.668$, $\text{precision} = 161/\ (161 + 80) = 0.668$, $\text{recall} = 161/(161 + 77) = 0.676$, and the $\text{F1 score} = (2 \times 161)/\ ((2 \times 161) + 80 + 77) = 0.672$. Relatively, the matrix parameters for the

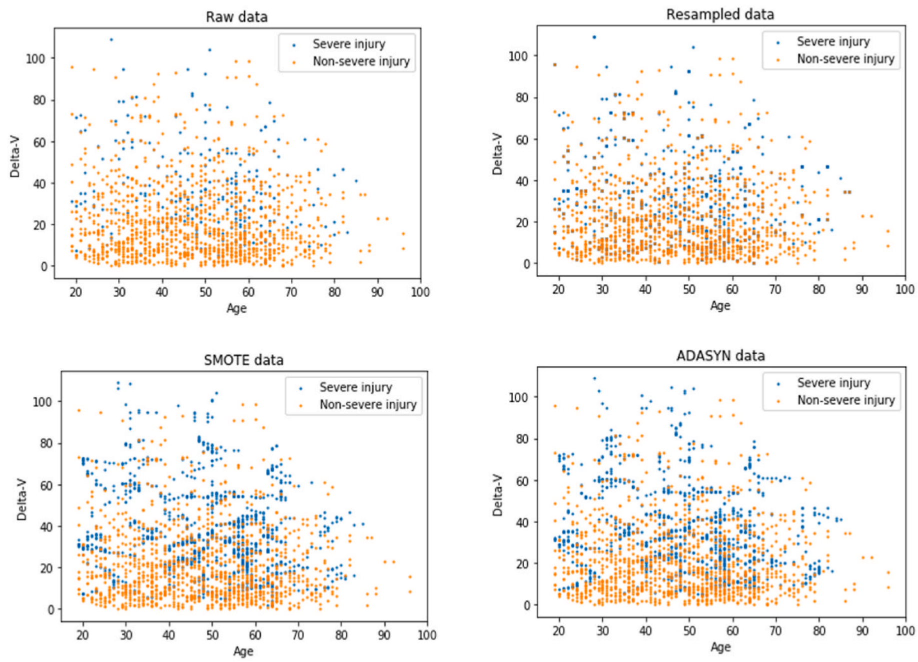

Fig. 2. Comparison of scatter plot of data obtained from tested oversampling methods.

Table 3
Demographic data of MVCs related to trauma.

Variables		Descriptions	Frequency (n = 1417)	Ratio (100%)
Dependent variables	ISS (binary)	Severe injury	236	16.7
		Non-severe injury	1181	83.3
Independent variables	Age	54 years under	907	64.0
		55–64 years	312	22.0
		65 years over	198	14.0
	Restraint usage	Restrained	930	65.6
		Unrestrained	487	34.4
	PDOF	Frontal impact	881	62.2
		Side impact	336	23.7
		Rear-end impact	200	14.1
	Vehicle type	Sedan	820	57.9
		SUV	230	16.2
		Light truck	212	15.0
		Van	122	8.6
		Heavily trailers	33	2.3
	Collision partner	Smaller	114	8.0
		Similar	931	65.7
		Larger	372	26.3
	Multiple impact	Yes	118	8.3
		No	1299	91.7
	Delta-V	0–10 km/h	470	33.2
		11–19 km/h	275	19.4
		20–29 km/h	243	17.1
		30–39 km/h	177	12.5
		40–49 km/h	97	6.8
		50 km/h over	155	10.9

Abbreviations: ISS Injury Severity Score PDOF; Principal Direction of Force, SUV; Sports Utility Vehicle.

SMOTE-based MLP classifier are presented as TP = 205, FN = 20, FP = 109, and TN = 139. This may calculated as $\text{accuracy} = (205+139)/(205+139+109+20) = 0.727$, $\text{precision} = 205/(205+109) = 0.653$, $\text{recall} = 205/(205+20) = 0.911$, and the $\text{F1 score} = (2\times205)/((2\times205)+109+20) = 0.761$, respectively. The outperformed parameter calculation is presented using the SMOTE dataset for the XGBoost classifier given as TP = 221, FN = 29, FP = 51, and TN = 172. In this case, the $\text{accuracy} = (221+172)/(221+172+29+51) = 0.831$, $\text{precision} = 221/(221+51) = 0.813$, $\text{recall} = 221/(221+29) = 0.884$, and the $\text{F1 score} = (2\times221)/((2\times221)+51+29) = 0.847$. Other classifiers may also calculate from referred by equations (1)-(4).

In the case of predicting severely injured occupants, the SMOTE-XGBoost model also yielded excellent discrimination in C2C crashes (AUC = 0.896). The comparison of prediction performance also can be visualized from the graphical plot illustrations using the ROC curve (Fig. 3). This visualizes the success rate for the classifier as quantified by calculating the curves. A higher value of evaluation metrics represents the outperforming of predictions.

This study suggests the feature importance ranking of the best performance model (SMOTE-XGBoost) indicators in predicting injury classification (Table 6). The Delta-V featured exclusive importance compared with other variables. Furthermore, the age distribution and PDOF showed nearly equal secondary importance. Though collision partners had relatively lower ranks in C2C crashes, the result has shown an advantage of importance compared to vehicle types.

4. Discussion

This study provided an ISP model using clinical data of MVOs who visited Level-1 trauma centers from January 2011 to April 2021 in South Korea. The primary outcome measurements were conducted as binary variables considering an overall ISS of 15 or greater, referring to the indicators used to evaluate trauma triage performance as recommended by the American College of Surgeon-Committee on Trauma (ACS-COT) within a limited protocol. The parameters used for prediction referred to the field triage recommendations of the CDC Expert Panel [6] and parameters of vehicle incompatibility of C2C crashes [11,12,15], including age, restraint usage (retrained or unrestrained), PDOF (frontal, side, and rear), vehicle type (sedan, SUV, light truck, van, heavy trailers), collision partner (smaller, similar, and larger-sized vehicle), multiple impacts (single or multiple), and Delta-V(kph unit).

The main findings showed that the ISP model of C2C crash-related occupants had an AUC of 0.896. This indicates the potential for improving predictive performance when considering sampling methods for imbalanced clinical data. Moreover, these results showed that the triage performance of the ML model was higher than that of traditional

Table 4
A comparison of confusion matrix used to predict injury severity classification.

Dataset		Classifier	N	Balance		Confusion matrix			
				Positive	Negative	True Positive	False Negative	False Positive	True Negative
Imbalanced	Raw	LR	284	6 (2.1)	278 (97.9)	5 (1.8)	47 (16.5)	1 (0.4)	231 (81.3)
		MLP	**284**	**163 (57.4)**	**121 (42.6)**	**40 (14.1)**	**7 (2.5)**	**123 (43.3)**	**114 (40.1)**
		XGB	284	11 (3.9)	273 (96.1)	3 (1.1)	42 (14.8)	8 (2.8)	231 (81.3)
Weighted	Class-weighted	**LR**	**284**	**89 (31.3)**	**195 (68.7)**	**26 (9.2)**	**20 (7.0)**	**63 (22.2)**	**175 (61.6)**
		MLP	**284**	**122 (43.0)**	**162 (57.0)**	**40 (14.1)**	**13 (4.6)**	**82 (28.9)**	**149 (52.5)**
		XGB	284	5 (1.8)	279 (98.2)	3 (1.1)	42 (14.8)	2 (0.7)	237 (83.5)
Over-sampled	Resampled	LR	473	218 (46.1)	255 (53.9)	132 (27.9)	106 (22.4)	86 (18.2)	149 (31.5)
		MLP	473	237 (50.1)	236 (49.9)	162 (34.2)	93 (19.7)	75 (15.9)	143 (30.2)
		XGB	**473**	**274 (57.9)**	**199 (42.1)**	**203 (42.9)**	**45 (9.5)**	**71 (15.0)**	**154 (32.6)**
	SMOTE	LR	473	241 (51.0)	232 (49.0)	161 (34.0)	77 (16.3)	80 (16.9)	155 (32.8)
		MLP	**473**	**314 (66.4)**	**159 (33.6)**	**205 (43.3)**	**20 (4.2)**	**109 (23.0)**	**139 (29.4)**
		XGB	**473**	**272 (57.5)**	**201 (42.5)**	**221 (46.7)**	**29 (6.1)**	**51 (10.8)**	**172 (36.4)**
	ADASYN	LR	459	231 (50.3)	228 (49.7)	156 (34.0)	70 (15.3)	75 (16.3)	158 (34.4)
		MLP	459	236 (51.4)	223 (48.6)	165 (35.9)	62 (13.5)	71 (15.5)	161 (35.1)
		XGB	**459**	**238 (51.9)**	**221 (48.1)**	**182 (39.7)**	**28 (6.1)**	**56 (12.2)**	**193 (42.0)**

Abbreviations: SMOTE; Synthetic Minority Oversampling Technique, ADASYN; Adaptive Synthetic Sampling, LR; Logistic Regression, MLP; Multilayer Perceptron, XGB; eXtreme Gradient Boosting.

Table 5
Predictive performance of severely injured occupants based on data sampling techniques.

Dataset		Classifier	Accuracy	Precision	Recall	F1 score	AUC
Imbalanced	Raw	LR	0.831	0.833	0.096	0.172	0.768
		MLP	**0.542**	**0.245**	**0.851**	**0.381**	**0.685**
		XGB	0.824	0.273	0.067	0.107	0.756
Weighted	Class- weighted	**LR**	**0.708**	**0.292**	**0.565**	**0.385**	**0.737**
		MLP	**0.665**	**0.328**	**0.755**	**0.457**	**0.711**
		XGB	0.845	0.600	0.067	0.120	0.806
Oversampled	Resampled	LR	0.594	0.606	0.555	0.579	0.627
		MLP	0.645	0.684	0.635	0.659	0.658
		XGB	**0.755**	**0.741**	**0.819**	**0.778**	**0.755**
	SMOTE	LR	0.668	0.668	0.676	0.672	0.735
		MLP	**0.727**	**0.653**	**0.911**	**0.761**	**0.795**
		XGB	**0.831**	**0.813**	**0.884**	**0.847**	**0.896**
	ADASYN	LR	0.684	0.675	0.690	0.683	0.748
		MLP	0.710	0.699	0.727	0.713	0.792
		XGB	**0.817**	**0.765**	**0.867**	**0.813**	**0.878**

Abbreviations: SMOTE; Synthetic Minority Oversampling Technique, ADASYN; Adaptive Synthetic Sampling, LR; Logistic Regression, MLP; Multilayer Perceptron, XGB; eXtreme Gradient Boosting, AUC; Area Under the Curve.

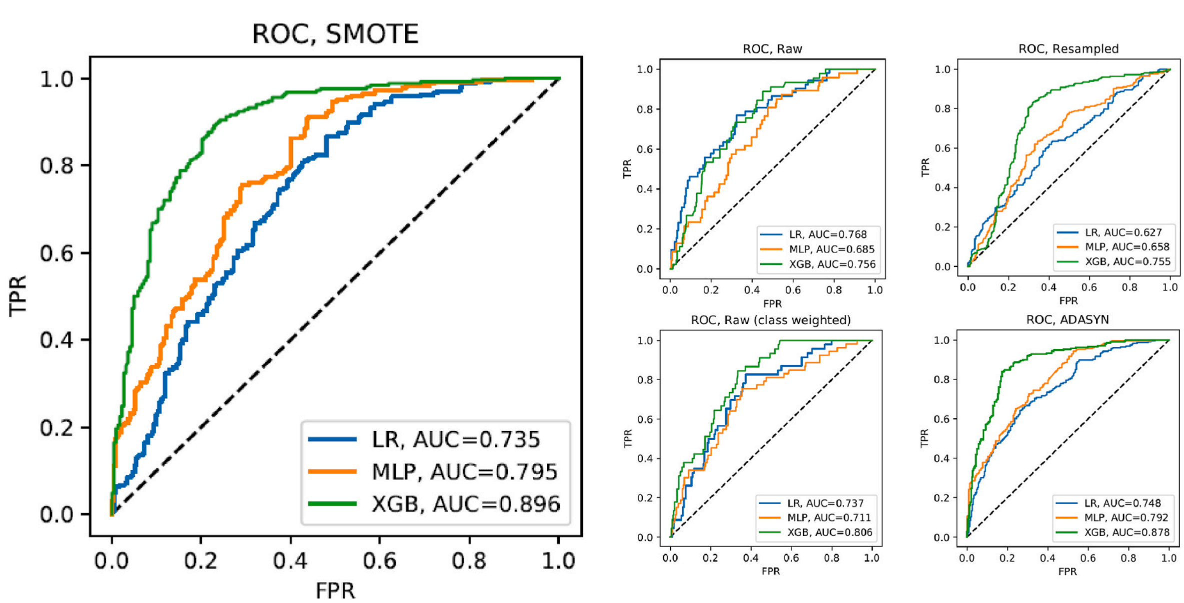

Fig. 3. Comparison of ROC curve among various sampling techniques.

Table 6
Features importance ranking of extreme gradient boost model using the synthetic minority oversampling technique dataset.

Parameters	Importance scores	Importance ratios	Features Rank
Delta-V	0.275	1.00	1
Age	0.176	0.64	2
PDOF	0.171	0.62	3
Restraint usage	0.107	0.39	4
Multiple collision	0.107	0.39	4
Collision partner	0.085	0.31	5
Vehicle type	0.079	0.29	6

statistical models (see Table 4).

This study confirmed that prediction performance improved through the data sampling technique before developing the ISP model. Most MVOs visiting trauma centers were classified as non-severely injured, resulting in a class imbalance of clinical data. Previous studies have reported that data imbalances cause prediction model bias and affect prediction performance [25,37–39]. Thus, several studies using the National Automotive Sampling System/Crashworthiness Data System (NASS-CDS) have leveraged population-weighted samples to address data bias [19,40,41]. However, a database lacking a data-weighting system has difficulty handling data under similar conditions. In contrast, data sampling techniques have recently been embraced as methodological approaches to addressing class imbalance problems. Some researchers have pointed out that data balancing should be considered to predict reliable injury outcomes [39,42–45]. This study showed similar results, with the best performance found using SMOTE-based oversampling data [25,37]. Using crash-related data, SMOTE provided an excellent prediction probability for MVO binary injuries. Meanwhile, undersampling or hybrid sampling approaches paired with different sampling techniques were not considered owing to the small sample-sized data.

Meanwhile, several studies suggest that the prediction models based on machine intelligence have improved outperformance [25,40,43,46–48]. Compared with statistical methodologies, the latest machine learning and deep learning technique enhance the predictive performance. In the previous study, various classifiers were conducted to compare the prediction performance of each model. These include decision tree [48,49], k-nearest neighbor [24], support vector machine (SVM) [50,51], tree-based model [52], neural networks [53], Naïve Bayesian classifier [54], and gradient boosting [55]. Yet, some latest methods have received the attention that is superior to conventional prediction models in the case of MVOs-related injury classification. A deep learning model, multilayer perceptron (MLP), yielded the highest accuracy as well as area under the curve (AUC) rate compared to the k-nearest neighbor, NBC, DTC, support vector machine, and logistic regression models [47]. On the other hand, the eXtreme Gradient Boosting (XGBoost) model outperformed compared to such models; K-nearest neighbor (KNN), linear SVM, radial basis function SVM (RBF SVM), Gaussian process classifier (GP), Decision tree (DT), random forest (RF), multilayer perceptron (MLP), AdaBoost, naïve Bayes (NB), and quadratic discriminant analysis (QDA) [55]. However, no study has been conducted comparing with the suggested models. Therefore, the present work has pointed out performance evaluation considering three ML models.

The results indicated that both the MLP and XGBoost models exhibited excellent discrimination for binary injury classification. In particular, the XGBoost model yielded the best predictions based on SMOTE oversampling in minority class data. The gap differences in predictive performance between XGBoost and MLP existed because most data used in the model consisted of categorical variables [56]. Because the factors affecting road traffic injuries in real-world crashes were immensely complicated, there was a tendency to categorize the data to estimate injury outcomes. For instance, it was intuitive to categorize the wearing state of seatbelts (belted or unbelted) rather than using quantitative kinematics for belt loading of MVOs to estimate injury severity. XGBoost is a gradient tree-based ML classifier with no issues encoding data with most of these categorical variables. However, predictive models based on continuous variables could expect improved MLP prediction probabilities. The implication of these findings pointed to the potential to support the selection decisions of ISP models based on different data characteristics and conditions.

In contrast to ML, statistical models have been reported to have weak ISP performance owing to their fixed assumptions [25]. ML models are flexible when capturing valuable information from nonlinear complex and heterogeneous data because they do not include pre-assured relationships between variables [45,48,57,58]. Furthermore, these methodological approaches have produced a better model fit than statistical methods [24]. Jamal et al. [58] suggested that various ML models (e.g., random forest and decision tree), including XGBoost, outperformed traditional statistical models, yielding results similar to our study. Nevertheless, regression models can classify injury severity by intuitively providing clear theoretical interpretations [59]. In previous studies, statistical models achieved acceptable discriminative predictive power using large-scale data [19,21,60,61]. However, the sample size used affected the performance of traditional statistical methods. It was difficult to expect the probability of prediction power using insufficient data acquisition at the national or regional levels. Sampling-based ML models provided effective approaches for ISP using relatively short-sized data.

Several studies proposed outperforming methods for ISP engaged with MVOs comparing various machine intelligence in binary in classification (Table 7). Most of all, they have different data collection periods for analysis in various databases. Also, there was a difference in the imbalance ratio according to the injury severity classification in each study. Although the machine learning models had superior predictive performance in related studies [46,48], others gave better results in traditionally statistical technique [40,61]. It is assumed that this may influence the performance of the model depending on the parameter selection in predicting binary class of injury outcome. In particular, Delen et al. (2017) showed that the best predictive performance in SVM, however, the under-triage results was missing to support clinical insights in real-world [48]. Therefore, this study confirmed that the logistic regression performed better than previous models (Random forest, Adaboost, Naïve Bayes, Support Vector Machine, k-nearest neighbor, Ridge Regression, Bernoulli Naïve Descent, Stochastic Gradient Descent) detecting errors in trauma classification from medical point of view [40,61]. Thus, comparing the presented methodologies in the previous studies, the XGBoost model had outperformed comparatively considering under-triage rate in medical terms of uses. However, studies applying various techniques based on the optimal parameters considering the crash injury mechanism according to the complexity of MVCs are required.

Many ISP models have been developed that consider overall crash types [19–21,61]. However, factors affecting severe MVO injuries differed depending on various crash scenarios. Unlike fixed-material collisions, vehicle incompatibilities (e.g., passenger cars versus SUVs) in C2C crashes have contributed to injury severity outcomes [16,62–64]. These vehicle mismatches of body structure increased the risk of injury severity to MVOs with disadvantageous self-protective capacities due to vehicle differences, such as mass, weight, geometry, and stiffness, based on Newtonian mechanics [11,14–16,51,54,65]. Zeng et al. (2016) reported that vans and trucks had stronger self-protection and aggressivity than passenger vehicles [16]. However, no further research has been conducted that reflects these characteristics in real-world C2C crashes. This study suggested an ISP model with collision partners that consider the crash incompatibility of two-vehicle scenarios. The collision partner was confirmed as a high discriminant feature of the best model compared to vehicle type. However, it was interpreted that these low features pointed to the distribution of vehicles with high rigidity (e.g., heavy trailers), which had insufficient numbers compared with other

Table 7
Performance comparison between the proposed models and previous studies.

Studies	Data (year) Number of crash data Number of variables	Crash injury targets	Class break-down Imbalance Ratio (%)	Data Sampling	Classification Models	Prediction Performance (%)	Under-triage (%)	Major ranked features
Kusano & Gabler [40]	• **NASS-CDS** (2002–2011) • N = 16,398 • **7**	General MVOs	• Severe injury (N/A) /Non-severe injury (N/A) • **N/A**	**Population-weighted**	**LR** (RF, AB, NB, SVM, kNN)	Accuracy: 88.3 Sensitivity: 67.5 Specificity: 88.9 **AUC: N/A**	**8.5**	• N/A
Delen et al. [48]	• **NASS-GES** (2011–2012) • N = 27,214 • **29**	General MVOs	• High level of severity (21.0) /Low level of severity (79.0) • **1:3.8**	**Under-sampling**	**SVM** (ANN, DT, LR)	Accuracy: 90.4 Sensitivity: 88.5 Specificity: 92.0 **AUC: 92.8**	N/A	• **Restraint use** • Manner of collision • Ejection
AI Mamlook et al. [46]	• **MTCF** (2010–2017) • N = 106,274 • **8**	Elderly MVOs	• Severe injury (12.4) /Non-severe injury (87.6) • **1:7.1**	**SMOTE**	**Light-GMB** (RF, DT, LR, NB)	Precision: 87.9 Recall: 81.4 F1 score: 83.7 **AUC: 87.5**	N/A	• **Age** • Traffic volume • Car age
Candefjord et al. [61]	• **NASS-CDS** (2010–2015) • N = 21,589 • **14**	General MVOs	• Severe injury (5.7) /Non-severe injury (94.3) • **1:16.5**	**Population-weighted**	**LR** (RR, BNB, SGD, ANN)	**AUC: 86.0**	**5.0–20.0**	• **Ejection** • Entrapment • Belt use
Our study	• **KIDAS** (2011–2020) • N = 1417 • **7**	C2C MVOs	• Severe injury (16.7) /Non-severe injury (83.3) • **1:5.0**	**SMOTE** (CW, Resample, ADASYN)	**XGB** (LR, MLP)	Accuracy: 83.1 Precision: 81.3 Recall: 88.4 F1 score: 84.7 **AUC: 89.6**	**6.1**	• **Delta-V** • Age • PDOF

Abbreviations: NASS-CDS: National Automotive Sampling System-Crashworthiness Data System, NASS-GES: National Automotive Sampling System-General Estimates System, MTCF: Michigan Traffic Crash Facts, MVO: Motor Vehicle Occupants, KIDAS: Korea In-Depth Accident Study, C2C: Car-to-Car crashes, SMOTE: Synthetic Minority Oversampling Technique, CW: Class-weight, ADASYN: Adaptive Synthetic Sampling, LR: Logistic Regression, RF: Random Forest, AB: AdaBoost, NB: Naïve Bayes, SVM: Support Vector Machine, kNN: k-nearest neighbor, ANN: Artificial Neural Networks, DT: Decision Trees, Light-GMB: Light-Gradient Boosting Machine, RR: Ridge Regression, BDB: Bernoulli Naïve Bayes, SGD: Stochastic Gradient Descent, XGB: eXtreme Gradient Boosting, MLP: Multilayered Perceptron.

vehicles. Thus, large-scale data might result in enhanced feature rankings for collision partners.

The application of telematics-based services (such as AACN) that can classify the injury severity of real-time crash victims through post-crash analysis is expected to be most effective for consistent golden hour [22]. It is available to transmit information to the control system through an algorithm built into the crash vehicle. Also, the dispatcher may detect the crash location automatically (i.e., GPS) and provide predicted triage to the EMS provider in real-time. Thus, patients may arrive at the trauma center quickly by minimizing the delay time compared to existing in-person responses. Therefore, advanced ISP models may potentially assist diagnosis effectively in hospital arrival time and for public use in preventing road traffic fatalities in the future.

5. Limitations

The study has several limitations. The main problem was that ML models were considered a black box, making it difficult to understand the relationships between crash inputs and injury outcomes. Meanwhile, an LR model interprets as a simple linear form. Clinically, this difference might cause problems depending on whether the structure of the model was interpretable. Therefore, ML models should be discussed in more detail before their practical application to real-world injury control, prevention, and treatment. Furthermore, compared with earlier studies, the number of data used to predict MVO injury severity was short-scaled. We used data focused on field investigations at five different regional trauma centers. In Korea, public databases (i.e., police investigations and transport-related government institutions) have not been authorized for use with ISP models. Hence, improving ISP model reliability through improved data collection was crucial. Since many hospitalized datasets have difficulties for public availability, nationalized scaled data collecting efforts collaborating from government and joint institutes are required to prevent road traffic injuries.

Additionally, it was necessary to consider the scalability of the predictor variables affecting severe injuries in C2C crashes. Although this study applied recommended variable MVC factors for CDC field triage guidelines and expert panels, advanced considerations of the characteristics of C2C crashes were limited to counterpart objects. Therefore, more detailed aspects of vehicle incompatibility (e.g., mass ratio or/and energy absorption) between two-vehicle collisions are required. However, major indicators of ISP models (e.g., ejections and entrapments) were not considered owing to a lack of prepared investigation data.

6. Conclusions

The main goal of this study was to propose an ML-based model for predicting severe injuries of C2C crash-related patients who visited Level-1 trauma centers in Korea. We evaluated the probability of the predictive performance of several ISP models (i.e., XGBoost, MLP, and LR) using a confusion matrix and F-measures. Based on the results, it was confirmed that the SMOTE-XGBoost model outperformed the other models. This demonstrated the importance of selecting an optimized ISP model while considering the variable MVC conditions. Furthermore, we confirmed that the sampling technique for class imbalanced datasets increased the prediction power. Nonetheless, it was essential to provide an interpretable algorithm for practical use in the real world through the expansion of MVO data collection. The primary features of our model were like those from a previous work. This study contributed to the literature by considering C2C-crash vehicle incompatibilities.

In a future study, external validation should be undertaken to improve the validity of the current model. Validating against different local or broad international databases is required to achieve model reliability. Additional research adopting state-of-the-art techniques (e.g., hybrid and ensemble models) using equivalent datasets should be performed. Moreover, an interpretable ISP model classifier is critical. In contrast to statistical algorithms, structural uncertainty due to the black-box phenomenon of ML models is a vital concern for medical applications. Therefore, transforming explainable artificial intelligence approaches into ML models in clinical practice is challenging. The results indicate the potential for EMS providers to improve dispatches to and field triage of MVOs while preventing emergency department overcrowding with non-severely injured patients.

Declaration of competing interest

None Declared.

Acknowledgements

This research was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 21AMDP-C162419-01) and the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (no.2021R1A2C2094669).

References

[1] WHO, Global Status Report on Road Safety, 2018, 2018, https://www.who.int/publications/i/item/9789241565684. (Accessed 1 March 2021).

[2] C.T.K. Stadtlander, CDC Health information for international travel 2016, Am. J. Trop. Med. Hyg. 95 (2016) 1219–1220, https://doi.org/10.4269/ajtmh.16-0627.

[3] E.J. MacKenzie, F.P. Rivara, G.J. Jurkovich, et al., A national evaluation of the effect of trauma-center care on mortality, N. Engl. J. Med. 354 (2006) 366–378, https://doi.org/10.1056/NEJMsa052049.

[4] H. Al-Thani, A. Mekkodathil, A.J. Hertelendy, T. Frazier, G.R. Ciottone, A. El-Menyar, Prehospital intervals and in-hospital trauma mortality: a retrospective study from a level I trauma center, Prehospital Disaster Med. 35 (2020) 508–515, https://doi.org/10.1017/S1049023X20000904.

[5] R.C. Mackersie, History of trauma field triage development andthe american college of surgeons criteria, Prehosp. Emerg. Care 10 (2006) 287–294, https://doi.org/10.1080/10903120600721636.

[6] S.M. Sasser, R.C. Hunt, M. Faul, et al., Guidelines for field triage of injured patients: recommendations of the national expert panel on field triage, MMWR Recomm. Rep. (Morb. Mortal. Wkly. Rep.) 61 (2012) (2011) 1–20.

[7] D.P. Stonko, D.C. O'Neill, B.M. Dennis, M. Smith, J. Gray, O.D. Guillamondegui, Trauma quality improvement: reducing triage errors by automating the level assignment process, J. Surg. Educ. 75 (2018) 1551–1557, https://doi.org/10.1016/j.jsurg.2018.03.014.

[8] D.E. Clark, R.J. Winchell, R.A. Betensky, Estimating the effect of emergency care on early survival after traffic crashes, Accid. Anal. Prev. 60 (2013) 141–147, https://doi.org/10.1016/j.aap.2013.08.019.

[9] H. Matsumoto, K. Mashiko, Y. Hara, et al., Dispatch of helicopter emergency medical services via advanced automatic collision notification, J. Emerg. Med. 50 (2016) 437–443, https://doi.org/10.1016/j.jemermed.2015.11.001.

[10] National Center for Injury Prevention and Control, Recommendations for the Expert Panel: Advanced Automatic Collision Notification and Triage of the Injured Patient, Centers for Disease Control and Prevention, Atlanta, GA, 2008.

[11] S. Acierno, R. Kaufman, F.P. Rivara, D.C. Grossman, C. Mock, Vehicle mismatch: injury patterns and severity, Accid. Anal. Prev. 36 (2004) 761–772, https://doi.org/10.1016/j.aap.2003.07.001.

[12] M.K. Verma, R. Nagappala, M. Murugan, Y.J. Tung, Evaluation of structural parameters for vehicle crash compatibility, Int. J. Crashworthiness 9 (2004) 577–586, https://doi.org/10.1533/ijcr.2004.0312.

[13] H. Stigson, A. Ydenius, A. Kullgren, Variation in crash severity depending on different vehicle types and objects as collision partner, Int. J. Crashworthiness 14 (2009) 613–622, https://doi.org/10.1080/13588260902920589.

[14] A. Sobhani, W. Young, D. Logan, S. Bahrololoom, A kinetic energy model of two-vehicle crash injury severity, Accid. Anal. Prev. 43 (2011) 741–754, https://doi.org/10.1016/j.aap.2010.10.021.

[15] R. Tolouei, M. Maher, H. Titheridge, Vehicle mass and injury risk in two-car crashes: a novel methodology, Accid. Anal. Prev. 50 (2013) 155–166, https://doi.org/10.1016/j.aap.2012.04.005.

[16] Q. Zeng, H. Wen, H. Huang, The interactive effect on injury severity of driver-vehicle units in two-vehicle crashes, J. Saf. Res. 59 (2016) 105–111, https://doi.org/10.1016/j.jsr.2016.10.005.

[17] H.J. Jeon, S.C. Kim, J. Shin, et al., Risk of serious injury of occupants involved in frontal crashes of cab-over-type trucks, Traffic Inj, Preview 18 (2017) 839–844, https://doi.org/10.1080/15389588.2017.1315413.

[18] L.Y. Wang, R.M. Li, C.J. Wang, Z.Y. Liu, Driver injury severity analysis of crashes in a western China's rural mountainous county: taking crash compatibility difference into consideration, J. Traffic Transport. Eng. 8 (2021) 703–714, https://doi.org/10.1016/j.jtte.2020.12.002.

[19] D.W. Kononen, C.A. Flannagan, S.C. Wang, Identification and validation of a logistic regression model for predicting serious injuries associated with motor vehicle crashes, Accid. Anal. Prev. 43 (2011) 112–122, https://doi.org/10.1016/j.aap.2010.07.018.

[20] R. Buendia, S. Candefjord, H. Fagerlind, A. Balint, B.A. Sjoqvist, On scene injury severity prediction (OSISP) algorithm for car occupants, Accid. Anal. Prev. 81 (2015) 211–217, https://doi.org/10.1016/j.aap.2015.04.032.
[21] T. Nishimoto, K. Mukaigawa, S. Tominaga, et al., Serious injury prediction algorithm based on large-scale data and under-triage control, Accid. Anal. Prev. 98 (2017) 266–276, https://doi.org/10.1016/j.aap.2016.09.028.
[22] K. He, P. Zhang, S.C. Wang, Crash telemetry-based injury severity prediction is equivalent to or out-performs field protocols in triage of planar vehicle collisions, Prehospital Disaster Med. 34 (2019) 356–362, https://doi.org/10.1017/S1049023X19004515.
[23] D.W. Hosmer, S. Lemeshow, R.X. Sturdivant, Applied Logistic Regression, third ed., Wiley, New York, NY, 2013.
[24] J. Zhang, Z. Li, Z. Pu, C. Xu, Comparing prediction performance for crash injury severity among various machine learning and statistical methods, IEEE Access 6 (2018) 60079–60087, https://doi.org/10.1109/ACCESS.2018.2874979.
[25] A. Ji, D. Levinson, Injury severity prediction from two-vehicle crash mechanisms with machine learning and ensemble models, IEEE Open. J. Intell. Transp. Syst. 1 (2020) 217–226, https://doi.org/10.1109/OJITS.2020.3033523.
[26] S. Somboon, N. Phunghassaporn, A. Tansawet, S. Lolak, Accuracy of machine learning logistic regression in death prediction in road traffic injury patients [letter], Asian J. Surg. 45 (2022) 537–538, https://doi.org/10.1016/j.asjsur.2021.09.010.
[27] T.A. Gennarelli, E. Wodzin, I.L. Barrington, The Abbreviated Injury Scale 2005-update 2008, Association for the Advancement of Automotive Medicine, Barrington, IL, 2008.
[28] S. Shilaskar, A. Ghatol, P. Chatur, Medical decision support system for extremely imbalanced datasets, Inf. Sci. 384 (2017) 205–219, https://doi.org/10.1016/j.ins.2016.08.077.
[29] D.C. Li, C.W. Liu, S.C. Hu, A learning method for the class imbalance problem with medical data sets, Comput. Biol. Med. 40 (5) (2010) 509–518, https://doi.org/10.1016/j.patrec.2020.03.004.
[30] R. Zhu, Y. Guo, J.H. Xue, Adjusting the imbalance ratio by the dimensionality of imbalanced data, Pattern Recogn. Lett. 133 (2020) (2020) 217–223, https://doi.org/10.1016/j.patrec.2020.03.004.
[31] N. Noorhalim, A. Ali, S.M. Shamsuddin, Handling imbalanced ratio for class imbalance problem using SMOTE, in: Proc Third Int Conf Computing, Mathematics and Statistics (iCMS2017), Springer, Singapore, 2019, pp. 19–30.
[32] Y.C. Wang, C.H. Cheng, A multiple combined method for rebalancing medical data with class imbalances, Comput. Biol. Med. 134 (2021), 104527, https://doi.org/10.1016/j.compbiomed.2021.104527.
[33] P. Good, Permutation Tests: a Practical Guide to Resampling Methods for Testing Hypotheses, second ed., Springer, New York, NY, 2000.
[34] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique, J. Artif. Intell. Res. 16 (2002) 321–357, https://doi.org/10.1613/jair.953.
[35] H. He, Y. Bai, E.A. Garcia, S. Li, ADASYN: adaptive synthetic sampling approach for imbalanced learning, in: 2008 IEEE International Joint Conference on Neural Networks, IEEE World Congress on Computational Intelligence), 2008, pp. 1322–1328, https://doi.org/10.1109/IJCNN.2008.4633969.
[36] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, Xgboost: extreme gradient boosting, R package version 0.4-2 (2015).
[37] M. Yahaya, X. Jiang, C. Fu, K. Bashir, W. Fan, Enhancing crash injury severity prediction on imbalanced crash data by sampling technique with variable selection, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), IEEE, Auckland, New Zealand, 2019, pp. 363–368, https://doi.org/10.1109/IJCNN.2008.4633969.
[38] C. Morris, J.J. Yang, Effectiveness of resampling methods in coping with imbalanced crash data: crash type analysis and predictive modeling, Accid. Anal. Prev. 159 (2021), 106240, https://doi.org/10.1016/j.aap.2021.106240.
[39] M. Yahaya, R. Guo, X. Jiang, K. Bashir, C. Matara, S. Xu, Ensemble-based model selection for imbalanced data to investigate the contributing factors to multiple fatality road crashes in Ghana, Accid. Anal. Prev. 151 (2021), 105851, https://doi.org/10.1016/j.aap.2020.105851.
[40] K. Kusano, H.C. Gabler, Comparison and validation of injury risk classifiers for advanced automated crash notification systems, Traffic Inj. Prev. 15 (Suppl 1) (2014) S126–S133, https://doi.org/10.1080/15389588.2014.927577.
[41] J.D. Stitzel, A.A. Weaver, J.W. Talton, et al., An injury severity-, time sensitivity-, and predictability-based advanced automatic crash notification algorithm improves motor vehicle crash occupant triage, J. Am. Coll. Surg. 222 (2016) 1211–1219, https://doi.org/10.1016/j.jamcollsurg.2016.03.028.
[42] H. Jeong, Y. Jang, P.J. Bowman, N. Masoud, Classification of motor vehicle crash injury severity: a hybrid approach for imbalanced data, Accid. Anal. Prev. 120 (2018) 250–261, https://doi.org/10.1016/j.aap.2018.08.025.
[43] N. Fiorentini, M. Losa, Handling imbalanced data in road crash severity prediction by machine learning algorithms, Infrastructure 5 (2020) 61, https://doi.org/10.3390/infrastructures5070061.
[44] S. Kim, Y. Lym, K.J. Kim, Developing crash severity model handling class imbalance and implementing ordered nature: focusing on elderly drivers, Int. J. Environ. Res. Publ. Health 18 (2021) (1966), https://doi.org/10.3390/ijerph18041966.
[45] Z.J. Ma, G. Mei, S. Cuomo, An analytic framework using deep learning for prediction of traffic accident injury severity based on contributing factors, Accid. Anal. Prev. 160 (2021), 106322, https://doi.org/10.1016/j.aap.2021.106322.
[46] R.E. Al Mamlook, T.Z. Abdulhameed, R. Hasan, H.J. Al-Shaikhli, I. Mohammed, S. Tabatabai, S, Utilizing machine learning models to predict the car crash injury severity among elderly drivers, in: 2020 IEEE Int. Conf. Electro Information Technology (EIT), 2020, pp. 105–111.
[47] A. Cigdem, C. Ozden, Predicting the severity of motor vehicle accident injuries in adana-Turkey using machine learning methods and detailed meteorological data, Int. J. Intell. Syst. Appl. Eng. 6 (2018) 72–79, https://doi.org/10.18201/ijisae.2018637934.
[48] D. Delen, L. Tomak, K. Topuz, E. Eryarsoy, Investigating injury severity risk factors in automobile crashes with predictive analytics and sensitivity analysis methods, J. Transport Health 4 (2017) 118–131, https://doi.org/10.1016/j.jth.2017.01.009.
[49] K.A. Abay, R. Paleti, C.R. Bhat, The joint analysis of injury severity of drivers in two-vehicle crashes accommodating seat belt use endogeneity, Transp. Res. Part B Methodol. 50 (2013) 74–89, https://doi.org/10.1016/j.trb.2013.01.007.
[50] S. Ferreira, M. Amorim, A. Couto, Risk factors affecting injury severity determined by the mais score, Traffic Inj. Prev. 18 (2017) 515–520, https://doi.org/10.1080/15389588.2016.1246724.
[51] N.H.T.S. Administration, Crash injury research and engineering network (ciren) program report, 2002, DOT HS 809 (2005) 564.
[52] A. Ji, D. Levinson, An energy loss-based vehicular injury severity model, Accid. Anal. Prev. 146 (2020), 105730, https://doi.org/10.1016/j.aap.2020.105730.
[53] M.I. Sameen, B. Pradhan, Severity prediction of traffic accidents with recurrent neural networks, Appl. Sci. 7 (2017) 476, https://doi.org/10.3390/app7060476.
[54] R.E. Alamlook, K.M. Kwayu, M.R. Alkasisbeh, A.A. Frefer, Comparison of machine learning algorithms for predicting traffic accident severity, in: 2019 IEEE Jordan Int. Joint Conf. Electrical Engineering and Information Technology, JEEIT), 2019, pp. 272–276.
[55] B. Pradhan, M.I. Sameen, Predicting injury severity of road traffic accidents using a hybrid extreme gradient boosting and deep neural network approach, in: Laser Scanning Systems in Highway and Safety Assessment. Advances in Science, Technology & Innovation (Ierek Interdisciplinary Series for Sustainable Development), Springer, Cham, Switzerland, 2020, pp. 119–127.
[56] S.X. Zhu, K. Wang, C.Y. Li, Crash injury severity prediction using an ordinal classification machine learning approach, Int. J. Environ. Res. Publ. Health 18 (2021), 11564, https://doi.org/10.3390/ijerph182111564.
[57] L. Wahab, H. Jiang, Severity prediction of motorcycle crashes with machine learning methods, Int. J. Crashworthiness 25 (2020) 485–492, https://doi.org/10.1080/13588265.2019.1616885.
[58] A. Jamal, M. Zahid, M.T. Rahman, et al., Injury severity prediction of traffic crashes with ensemble machine learning techniques: a comparative study, Int. J. Inj. Control Saf. Promot. 28 (2021) 408–427, https://doi.org/10.1080/17457300.2021.1928233.
[59] J. Tang, J. Liang, C. Han, Z. Li, H. Huang, Crash injury severity analysis using a two-layer Stacking framework, Accid. Anal. Prev. 122 (2019) 226–238, https://doi.org/10.1016/j.aap.2018.10.016.
[60] J. Augenstein, E. Perdeck, J. Stratton, K. Digges, G. Bahouth, Characteristics of crashes that increase the risk of serious injuries, Annu. Proc. Assoc. Adv. Automot. Med. 47 (2003) 561–576.
[61] S. Candefjord, A.S. Muhammad, P. Bangalore, R. Buendia, On Scene Injury Severity Prediction (OSISP) machine learning algorithms for motor vehicle crash occupants in US, J. Transport Health 22 (2021), 101124, https://doi.org/10.1016/j.jth.2021.101124.
[62] E.L. Toy, J.K. Hammitt, Safety impacts of SUVs, vans, and pickup trucks in two-vehicle crashes, Risk Anal. 23 (2003) 641–650, https://doi.org/10.1111/1539-6924.00343.
[63] M. Fredette, L.S. Mambu, A. Chouinard, F. Bellavance, Safety impacts due to the incompatibility of SUVs, minivans, and pickup trucks in two-vehicle collisions, Accid. Anal. Prev. 40 (2008) 1987–1995, https://doi.org/10.1016/j.aap.2008.08.026.
[64] H. Huang, C. Siddiqui, M. Abdel-Aty, Indexing crash worthiness and crash aggressivity by vehicle type, Accid. Anal. Prev. 43 (2011) 1364–1370, https://doi.org/10.1016/j.aap.2011.02.010.
[65] C. Lee, X. Li, Analysis of injury severity of drivers involved in single- and two-vehicle crashes on highways in Ontario, Accid, Anal. Prev 71 (2014) 286–295, https://doi.org/10.1016/j.aap.2014.06.008.