Prediction Models for Severely Injured Occupants using Machine Learning Analytics Based on Oversampling Class Imbalanced Data

Joon Seok Kong

The Graduate School
Yonsei University
Department of Medicine

Prediction Models for Severely Injured Occupants using Machine Learning Analytics Based on Oversampling Class Imbalanced Data

A Dissertation
Submitted to the Department of Medicine
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Joon Seok Kong

December 2022

This certifies that the dissertation of Joon Seok Kong is approved.

____
Thesis Supervisor: Kang Hyun Lee

____
Thesis Committee Member: Oh Hyun Kim

____
Thesis Committee Member: Sang Chul Kim

____
Thesis Committee Member: Dae Ryong Kang

____
Thesis Committee Member: Tae-Eung Sung

The Graduate School

Yonsei University

December 2022

Acknowledgements

먼저 연구자로서 가져야 하는 진실된 마음과 넘치는 열정을 바탕으로 아낌없는 조언과 성심을 다해 지도해주신 이강현 교수님께 진심으로 감사를 드립니다. 국내 유일한 자동차의학 분야에서 제가 이바지할 수 있는 기회를 주시고, 연구에 매진할 수 있도록 큰 버팀목이 되어 주셔서 여기까지 올 수 있었습니다.

박사학위 과정을 진행하며 도움주신 분들이 너무나 많습니다. 머나먼 미국에서 자동차의학 분야의 최전선에서 해외연수를 겸하시며 아낌없이 조언주신 김오현 교수님, 누구보다 실사고 분야에서 많은 연구와 관심을 쏟아 주신 김상철 교수님, 연구에 임하는데 있어 자신의 소신과 전문성을 알려주신 강대용 교수님, 그리고 따스한 조언과 연구의 객관성을 조언주신 성태응 교수님께 깊은 감사를 드립니다. 또한 인공지능 모델링 분석에 어려움을 겪을 때마다 친절히 명쾌한 설명을 보태주신 정호연 선생님께도 감사의 말씀을 드립니다.

본 응급의학교실에서 환자를 최우선으로 생각하시고 연구에 최선을 다하는 모습을 몸소 보여주신 황성오 교수님, 김현 교수님, 차경철 교수님, 차용성 교수님, 정우진 교수님, 이윤석 교수님, 박경혜 교수님, 노영일 교수님께도 감사와 헌신에 존경을 표합니다. 자동차의과학연수소 설립부터 함께 지낸 연구원 선생님들과 의국원, 그리고 응급구조사 선생님들께도 감사의 마음을 전달 드립니다.

마지막으로, 언제나 저를 믿고 따스한 마음으로 응원해 주시는 아버지와 어머니께 감사드립니다. 늘 긍정적인 사고와 겸손한 태도로 인생을 마주하라는 말씀 간직하겠습니다. 또한 언제나 저에게 크나큰 관심을 주시고 사랑으로 베풀어 주신 우리 가족과 하늘에서 축복으로 저를 지켜보고 계신 할아버지께 이 논문을 바칩니다.

2022년 12월

공 준 석 올림

Table of Contents

List of Figures iii

List of Tables v

Abstract viii

Chapter 1. Introduction 1

1.1 Research background 1

1.2 Purpose of the research 4

Chapter 2. Materials and Methods 5

2.1 Korea In-Depth Accident Study (KIDAS) database 5

2.2 Data source 23

2.3 Study population 24

2.4 Sample size calculation 25

2.5 Injury Severity classification 26

2.6 Study design 27

2.7 Predictive parameters 28

2.8 Data sampling techniques for class imbalanced data 29

2.9 Classification model 31

2.10 Model training and testing 32

2.11 Performance evaluation 32

Chapter 3. Results 35

3.1 Data sampling techniques for class imbalance clinical outcomes 35

3.2 Baseline characteristics of crash data distributions enrolled in truama centers 37

3.3 Confusion matrices for triage controls for injury severity classifications 39

3.4 Evaluating predictive performances on ISP classifiers 41

3.5 Rankings of features important to outperforming SMOTE-XGBoost 45

Chapter 4. Discussion 46

Chapter 5. Limitation 54

Chapter 6. Conclusion 56

References 58

Abstract in Korean 67

List of Figures

Figure 2.1. Regional investigations for MVCs based on the KIDAS database............... 6
Figure 2.2. Flowchart for investive procedures based on the emergency department..... 7
Figure 2.3. Initial stage investigation report forms collected at truama centers.............. 9
Figure 2.4. On-scene investigation of crash vehicles.. 10
Figure 2.5. Crash vehicle data extraction using Event Data Recorder......................... 10
Figure 2.6. Real-world flight mapping sequence based on cloud-based photo scenes.. 11
Figure 2.7. Generating geometric snapshots on real world crash scene....................... 12
Figure 2.8. 3-Diemensionalized image rendering of crash scene................................ 12
Figure 2.9. Perspective view shots of crash scene rendering using Pix4D................... 13
Figure 2.10. Collision optimizer to assess trajcetory errors and confidence intervals ... 14
Figure 2.11. Correlation analysis of in-vehicle recording and crash reconstruction 16
Figure 2.12. 3D reconstruct crash scene based on car-to-car crashes.......................... 17
Figure 2.13. A crash scene with the pre-hospital care units 18
Figure 2.14. Funtional road photography on crash environments 18
Figure 2.15. Vehicle exterior deformation in frontal crashes 19
Figure 2.16. Vehicle interiors and passive safety device status investigation 19
Figure 2.17. Opponent vehicle's exterior deformations 20
Figure 2.18. Opponent vehicle's interiors with the activated airbags.......................... 20
Figure 2.19. Patient's records and truama scores in car-to-car crash........................... 21

Figure 2.20. Diagnostic imaging using computed tomography 22

Figure 2.21. In-hospital surgical findings and injury surveillance 23

Figure 2.22. Flowchart of data selections of overall crash-related KIDAS dataset....... 25

Figure 2.23. Prediction model analytics of MVCs occupant injuries 28

Figure 3.1. Comparison of scatter plots of data obtained from tested oversampling..... 36

Figure 3.2. Comparison of ROC curve using various sampling techniques 44

List of Tables

Table 2.1. An example of Injury Severity Score calculation 27

Table 2.2. Hyperparameters used in the prediction model 31

Table 2.3. Model intra-validation associated with the prediction model and traumatic clinical data abbreviation 33

Table 3.1. Demographic data of MVCs related to traumas 38

Table 3.2. A comparison of confusion matrix used to predict injury severity classification 40

Table 3.3. Predictive performance of severely injured MVOs based on sampling techniques....................42

Table 3.4. Feature importance ranking of outperformed classifier 45

Table 3.5. Performance comparison between the proposed model and previous studies 51

ABBREVIATIONS

Abbreviations	Full name
AACN	Advanced Automatic Collision Notification
AUC	Area Under the Curve
ADASYN	Adaptive Synthetic Sampling
AIS	Abbreviated Injury Scale
C2C	Car-to-Car crashes
CAMS	Center for Automotive Medical Science Institute
CI	Class Imbalance
CW	Class-weighted
DL	Deep Learning
ED	Emergency Department
FN	False Negative
FP	False Positive
IR	Imbalance Ratio
ISP	Injury Severity Prediction
ISS	Injury Severity Score
KCD	Korean Standard Classification of Disease

KIDAS	Korea In-Depth Accident Study
LR	Logistic Regression
ML	Machine Learning
MVC	Motor Vehicle Crashes
MVOs	Motor Vehicle Occupants
MLP	Multilayer Perceptron
NASS-CDS	National Automotive Sampling System -Crashworthiness Data System
NASS-GES	National Automotive Sampling System -General Estimates System
OS	Over Sampling
PDOF	Principle Direction of Force
SMOTE	Synthetic Minority Oversampling Technique
TN	True Negative
TP	True Positive
UT	Under-triage
XGB	eXtreme Gradient Boosting

ABSTRACT

Prediction Models for Severely Injured Occupants using Machine Learning Analytics Based on Oversampling Class Imbalanced Data

Joon Seok Kong

Department of Medicine

The Graduate School

Yonsei University

Injury prediction models improve trauma outcomes for motor vehicle occupants with accurate decision-making and early transport to appropriate trauma centers. This study aimed to investigate the injury severity prediction (ISP) capability of machine-learning analytics based on five-different regional Level 1 trauma center-enrolled patients in Korea.

We studied car crash-related injury data from 1,417 patients enrolled in the Korea In-Depth Accident Study database from January 2011 to April 2021. Severe injury classification was defined as an Injury Severity Score ≥ 15. Planar collisions were considered by excluding rollovers which would 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 was 0.896. Additionally, the present ISP model showed an under-triage rate of 6.1%. The Delta-V, age, and Principal Direction of Force (PDOF) 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.

Keywords: Injury severity prediction, Machine learning, Motor vehicle occupants, Trauma center, Injury severity, Delta-V, Under-triage, Class-imbalance, Oversampling technique, Korea In-Depth Accident Study (KIDAS)

Chapter 1

Introduction

1.1. Research background

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 traumatic injury-related mortality 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 crucial to saving trauma patients. It has been reported that of patients with severe injuries are transferred to trauma centers early, it leads 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-standing 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 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. In particular, car-to-car (C2C) crashes account for a significant proportion of collisions 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, size, geometry, and stiffness. Consequently, in contrast to single-vehicle crashes, vehicle dissimilarity significantly affects severe injury in C2C crashes. However, most injury severity prediction (ISP) studies have focused only on the overall collision 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]. Meanwhile, the predictive performance of these algorithms depends on the sample size. Thus, it is difficult to expect good performance when there is insufficient 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.

1.2. Purpose of the research

This study aims to provide ISP models using ML analytics for MVOs who have visited Level 1 trauma centers in Korea. The study suggests that the noteworthy by address the 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 from a medical point-of-view, and 4) providing the feature importance of a single outperforming evaluation model. Thus, rather than simply focusing on improving predictive performance, it is vital to represent a clinically reliable model for medical workers in the real world.

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

Chapter 2

Materials and Methods

2.1. Korea In-Depth Accident Study (KIDAS) database

2.1.1. Hospital centered accidental injury data

The Korea In-Depth Accident Study (KIDAS) is a research unit established by emergency medical institutes in Korea to conduct research by forming a consultative body with various related organizations to prevent bodily injury and reduce the mortality of road traffic users. This is to promote injury prevention that is differentiated from other countries according to the unique road traffic environments and vehicular distributions in South Korea. In addition, based on Haddon's matrix, various indicators are being collected from the perspective of humans, vehicles, and the environmental index by crash stages (i.e., Pre-crash, In-crash, and Post-crash) in real-world crashes. Centering on the Center for Automotive Medical Science Institute (CAMS) established at Wonju Severance Christian Hospital, other regional trauma centers in Korea are collecting in-depth data associated with road traffic injuries (Figure 2.1).

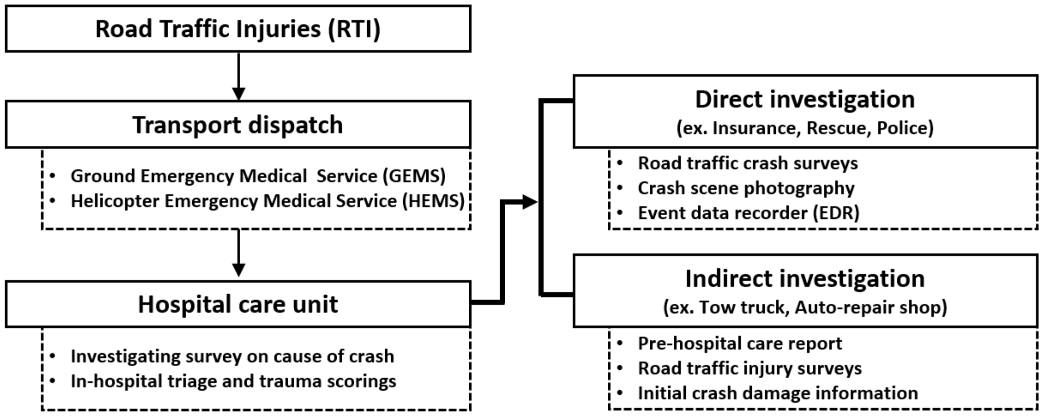


Figure 2.1. Regional investigations for motor vehicle crashes based on the KIDAS database

The collected information consists of a large group of different types of information; classification index, patients index, vehicle information, vehicular safety features, road-environmental characteristics, accident and collision types, vehicular deformations, prehospital information, medical records and trauma registry, Korean Standard Classification of Disease (KCD7), Abbreviated Injury Scale (AIS), Injury Severity Score (ISS), and crash dynamics. These include a total of 290 indicators related to Korean road traffic injuries to analyze and develop an injury prevention model.

2.1.2. Team-makers and action roles for data collections

The professionals involved with the investigation and management of road traffic injuries are led by regional trauma centers and are subdivided into medical staff, crash investigators, and reconstruction engineers. Initial information sharing is implemented by disseminating preliminary investigation results from emergency room. Each division performs it duties through a mutual information exchanged for traffic injury patients who have visited regional trauma centers. All stakeholders participate in the overall process until every record related to the incidents is collected to finalize the data recordings (Figure 2.2).

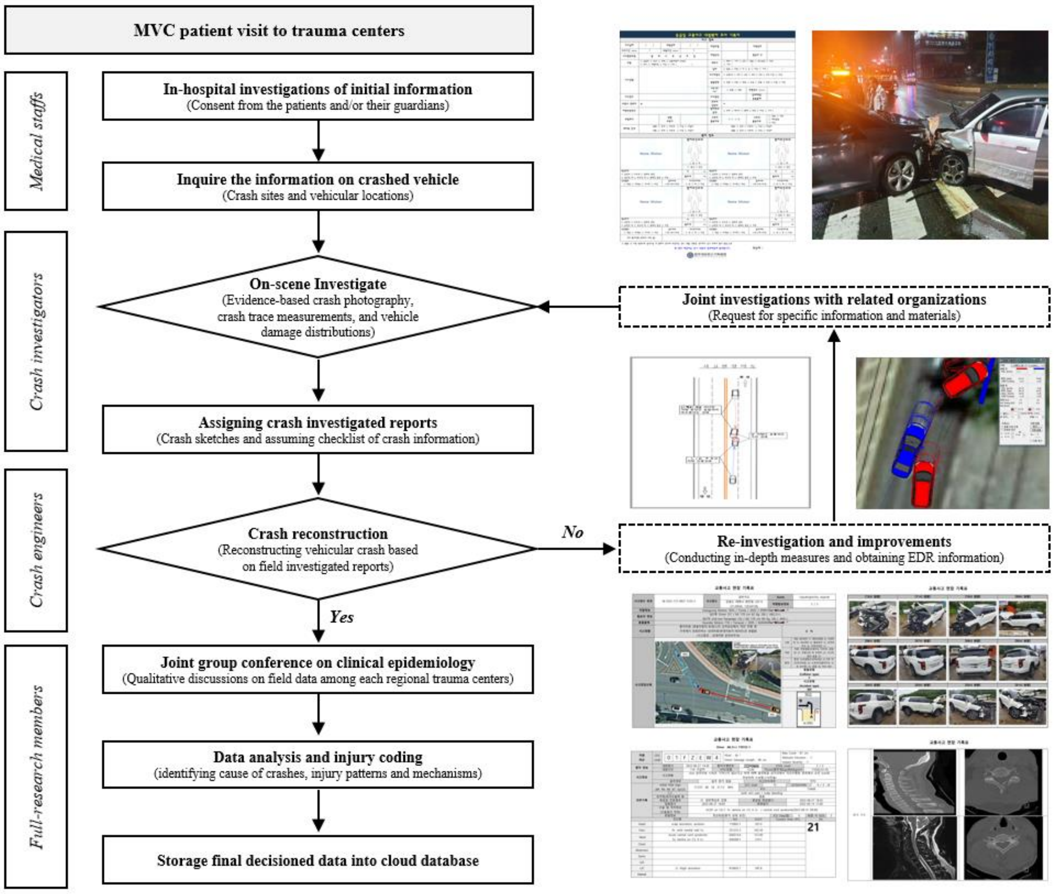


Figure 2.2. Flowchart for investigative procedures based on the emergency department

2.1.3. Hospitalized Team-call investigation systems

In the pre-hospital stage, the medical staff collects the initial information on the road traffic collision from the patient or their guardians who have visited the emergency department. Depending on necessity, the fire department requests quality on-site photographic evidence of the crash scene. The information is disseminated to the crash investigators through team calls. After the patient is given an in-hospital triage classification, the final injury outcomes are recorded using the Emergency Medical Record (EMR) system and trauma registry.

2.1.4. On-scene crash investigation protocols

Based on the survey report, the established information is used to complete a crash eport before being entered into the database (Figure 2.3). After being notified of the situation by the hospital medical staff, the crash investigators immediately attempt to record the crash site information and the vehicular damage distribution (Figure 2.4). In addition, the Event Data Recorder (EDR) is used to secure the vehicle's safety device and dynamic information from the time of the crash (Figure 2.5).

In the field stage, the crashed vehicle’s crash location and momentum trajectory are recorded as sketches by investigators. Moreover, an aerial drone scans the crash site and surrounding the roads with scale-based measurements to improve the qualitative information. High-precision 3-Diemensionalized road objects are generated through Pix4D mappers (Pix4D) by point cloud photogrammetry (Figure 2.6-2.8). Through the calibration of geolocated images during flight, detailed information such as the elevation difference of the road and the inclination of the road curvature is measured (Figure 2.9).

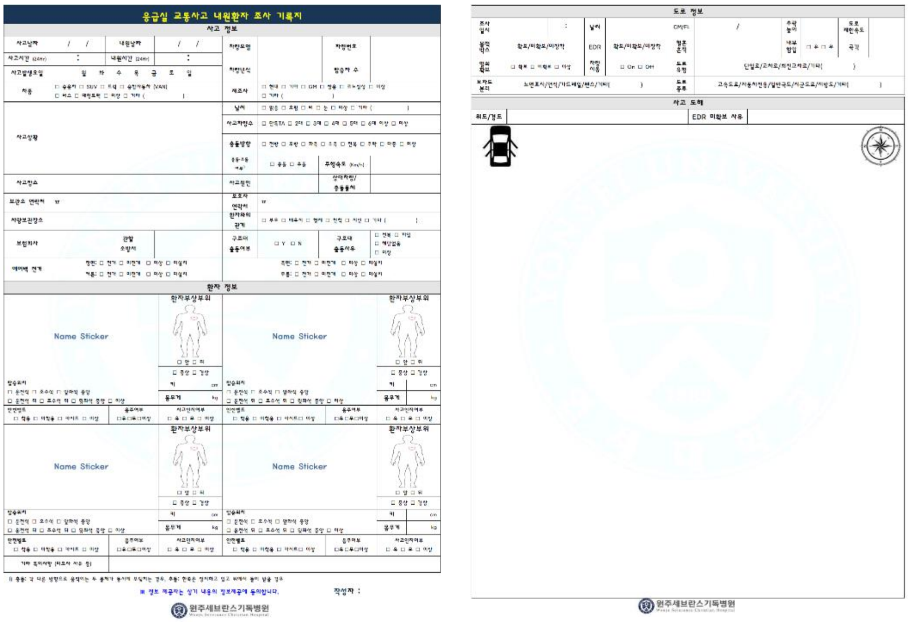


Figure 2.3. Initial stage investigation report forms collected at trauma centers

Figure 2.4. On-scene investigation of crash vehicles

Figure 2.5. Crash vehicle data extraction using Event Data Recorder

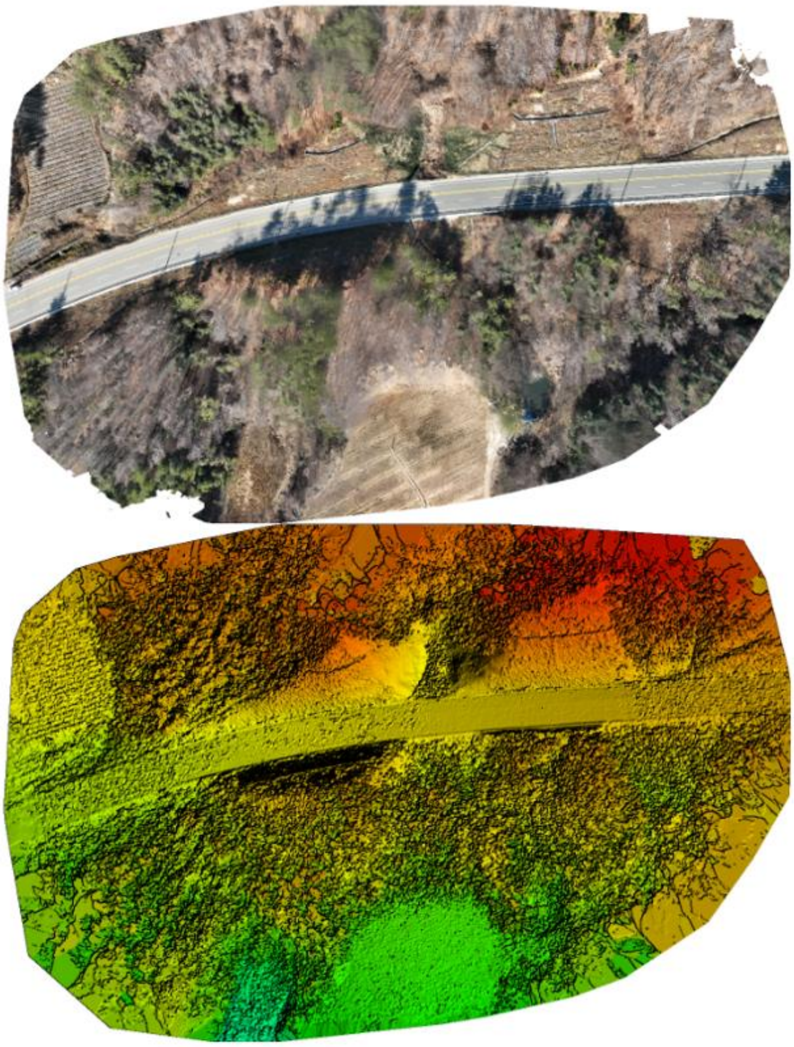


Figure 2.6. Real-world flight mapping sequence based on cloud-based photo scenes

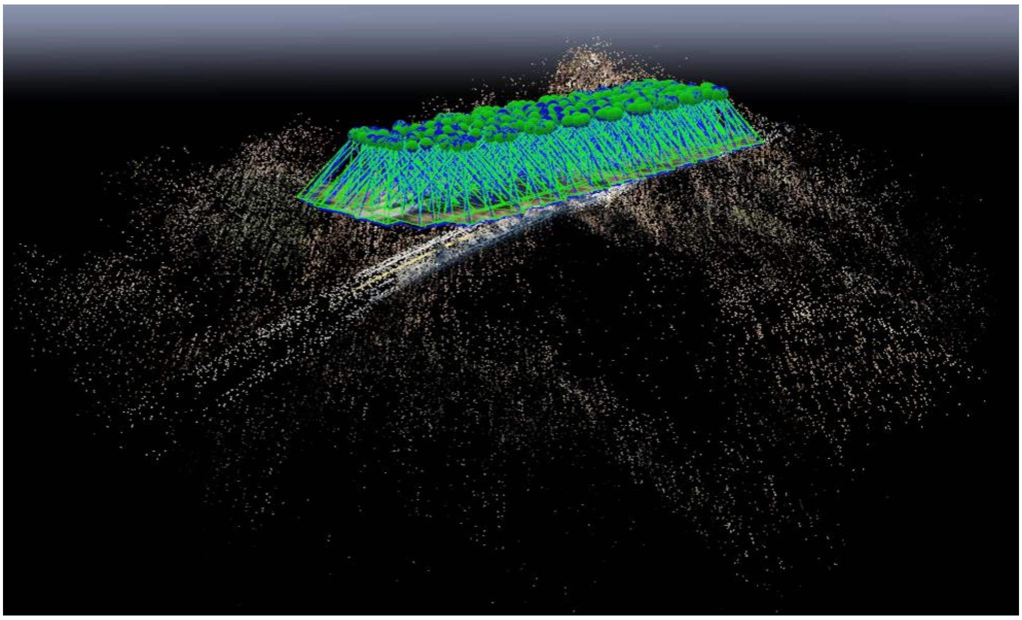


Figure 2.7. Generating geometric snapshots of real-world crash scene

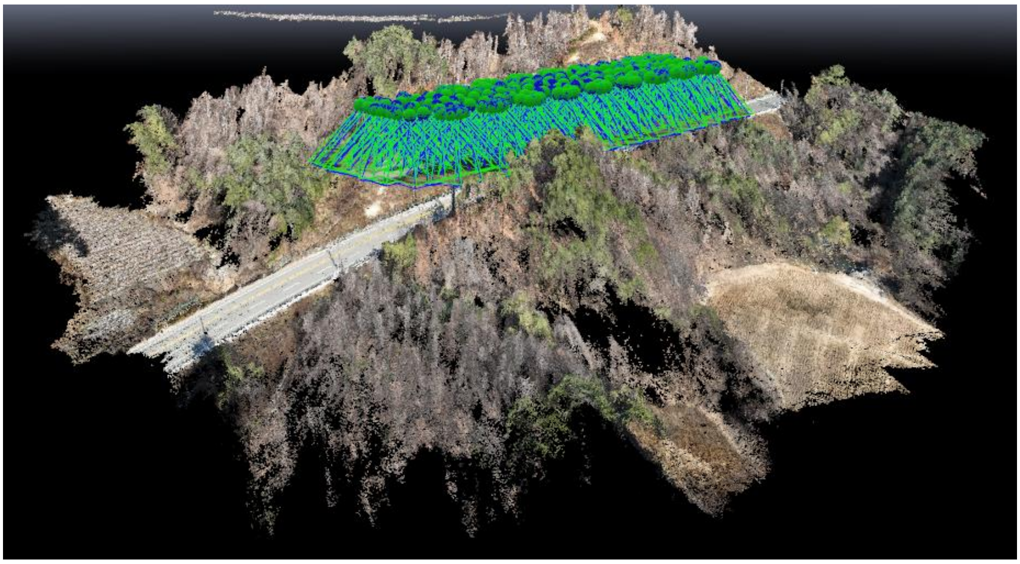


Figure 2.8. 3-Diemensionalized image rendering of crash scene

Figure 2.9. Perspective view shots of crash scene rendering using Pix4D

2.1.5. Crash reconstruction for real-world crashes

Crash reconstruction promotes an intuitive understanding of the actual collision scene and enables easy analysis of the dynamic behaviors of the crash vehicles. To ensure that the reconstruction of a motor vehicle crashes is precise, it is necessary to optimize the on-scene investigative information considering every detail. Based on the in-depth crash investigation report, the reconstruction is performed using the PC-Crash (DSD, Dr. Steffan Datentechnik GmbH Linz, Austria) software. This allows simple vehicle dynamics and kinematic responses to be reconstruct based on real world crashes. The collision optimizer evaluated the accuracy of reconstructed data and show a trajectory error rate of less than 5% within the 95% confidence intervals (Figure 2.10). Through this, there is a higher potential for injury to the occupants. Delta-V can be simply defined as the difference between the initial speed before the collision and the lowest peak speeds after the collision.

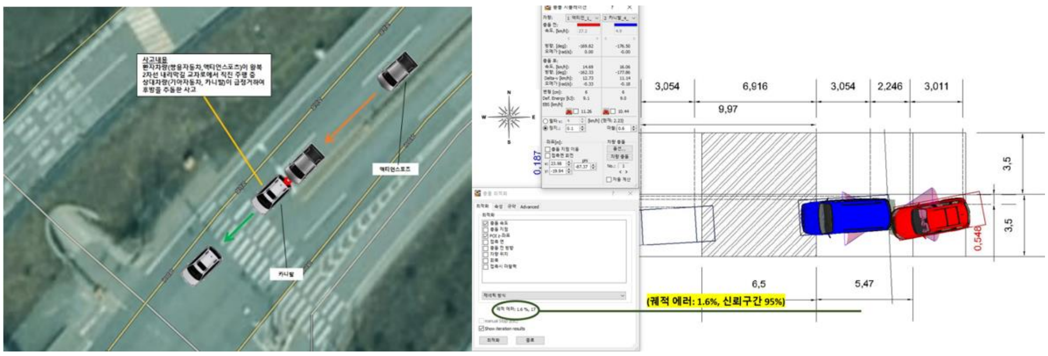


Figure 2.10. Collision optimizer to assess trajectory errors and confidence intervals

According to Newton's law of motion, Delta-V can be calculated using the mass of the two vehicles and their velocities before and after the collision. Based on the hypothesized case for the frontal crashes of two vehicles, the equations may address simplified as follows.

Equation 2.1 is the formula for the law of conservation momentum, where m_1 and m_2 is the mass of two vehicles, v_1 and v_2 is the initial velocity of the crash vehicle, v_1' and v_2' is the velocities of after crash of the both vehicles, respectively.

$$m_1 v_1 + m_2 v_2 = m_1 v_1' + m_2 v_2{}' \tag{2.1}$$

Equation 2.2 shows the coefficient of restitution (e) calculation formula, and each of the other variables is the same as in Equation 2.1.

$$\mathrm{e} = \frac{v_2' - v_1{}'}{v_1 - v_2} \tag{2.2}$$

Equation 2.3 is a combination of Equations 2.1 and 2.2, and using this equation, the post-collision speed of the two vehicles can be obtained by applying the mass and initial collision speed of the two vehicles.

$$v_2' = v_2 + \frac{m_1(1+e)}{m_1+m_2}(v_1 - v_2) \tag{2.3}$$

However, the accuracy of these mathematical calculations may be insufficient depending on the complexity of the collision (i.e., rollovers, multiple collisions) in real-world crashes. Therefore, the calculations are performed using the Delta-V of the vector unit calculated in the rigid body-based universal crash reconstruction software (Figure 2.11-2.12).

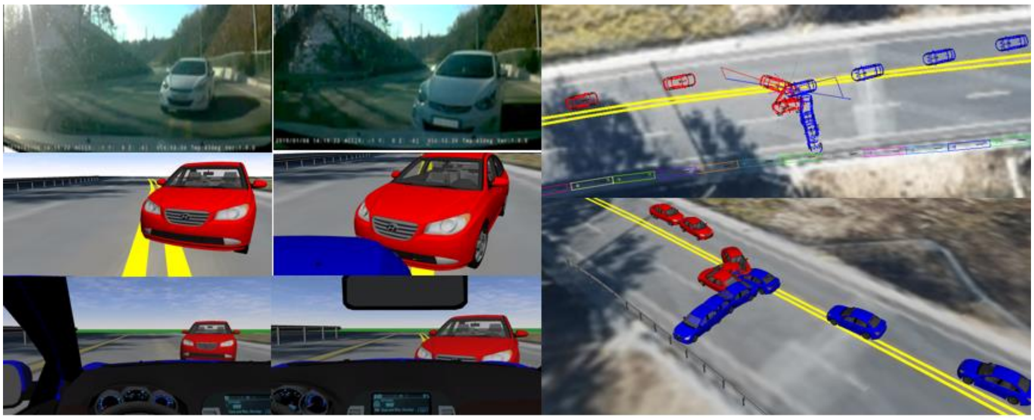


Figure 2.11. Correlation analysis of in-vehicle recording and crash reconstruction

Figure 2.12. 3D reconstructed crash scene based on car-to-car crashes

2.1.6. In-depth data review and decision-making process

The comprehensive collection of information is reviewed by an expert meeting where the decision to enter it into the final database is made. An in-depth reviewing process for decision-making is held every week with participants including medical doctors, trauma coordinators, crash investigators, reconstruction engineers, and related researchers. Experts discuss their field crash analysis (Figure 2.13-2.18) and injury outcomes (Figure 2.19-2.21) with full consideration for the mechanism of injury in vehicular crashes. Incorrect input and unclear information will be re-reviewed after reinvestigating the related feature. Finally, the preservation of the final documents is carried out, and the data are entered into the database through the database management system.

Figure 2.13. A crash scene with the pre-hospital care units

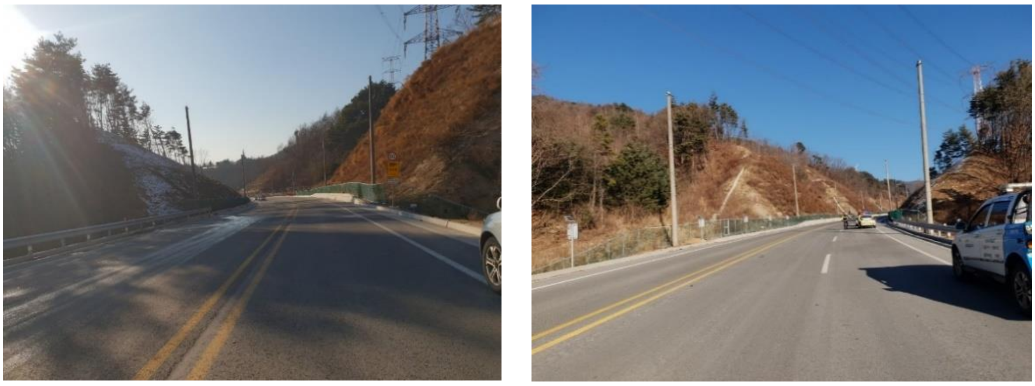


Figure 2.14. Functional road photography of crash environment

Figure 2.15. Vehicle exterior deformation in frontal crashes

Figure 2.16. Vehicle interiors and passive safety device status investigation

Figure 2.17. Opponent vehicle's exterior deformations

Figure 2.18. Opponent vehicle's interior with the activated airbags

교통사고 현장 기록표

환자 정보	내원일시	2019-01-06 15:29	환자식별번호	[redacted]	KTAS Level	Level Ⅱ
	내원수단	119 구급차	나이/성별	20/F	키(cm)/몸무게(kg)/BMI(kg/m^2)	156/49/20.1
의무기록	Initial Vital Sign	(100/67) (102) (18) (36.0℃) (100%)		Initial GCS		Total : 12 (E: 3 V: 5 M: 4)
	과거력 및 과거수술력	특이 내과적 과거력 없음 맹장-2013년도 용인local병원				
	응급실 진료결과	수술 후 중환자실로 입원				
	입원정보	2019-01-06 ~ 2019-03-01 (ICU : 6)				
	수술 및 처치정보	**2019-01-06 외과** Semental Resection of small bowel and end to end anstomosis Primary repair of colon Partial Omentectomy **2019-01-14 성형외과** Orbital floor reconstruction with Medpor (Blow out fx, Lt.) Closed reduction (Nasal bone fractre) **2019-01-21 정형외과** Flection extention injury on L2,3,4 **2019-01-28 정형외과** CR & Intrramedulary Depusy Synthes 6.5mm headless screw fixation for olecranon elbow Lt.				
	퇴원정보	전원 (이후의 지속적인 재활치료 위해 아주대병원으로 전원합니다.)				

진단명		AIS	ISS
Head	**T-SAH** (small amount of SAH at Rt. cerebral cortical sulci and falx) **Cerebral contusion** (hemorrhagic cerebral contusions, Rt. posterior temporal lobe) **Skull fracture, closed** (Rt. temporal bone, mastoid)	140694.2 140605.2 150402.2	**17**
Face	**Maxilla fracture including sinus, closed** lateral maxillary sinus wall to pterygoid plate. Lt. **Mutiple oribit fracture, closed** (inferior, medial and lateral wall of Lt. orbit) Injury of optic nerve and pathways **Lt. nasal bone** **Faical laceration**	250800.2 251205.2 230203.2 251000.1 210602.1	
Neck			
Chest	**Lt. lung contusion, minor** (Lung contusion at left upper lobe anterior.)	441407.2	
Abdomen	**Mesentery laceration, major** Colon laceration, no perforation (T-colon) Omentum laceration, minor **L-spine fracture** (L2,3,4 body fracture, L3 transverse process fx) **Disc herniation on L4** (no nerve root damage)	542024.3 540822.2 542222.2 650617.2 650602.2	
Spine			
U/E	Fracture, olecranon, closed	752113.2	
L/E			
External			

Figure 2.19. Patient's records and trauma scores in car-to-car crash

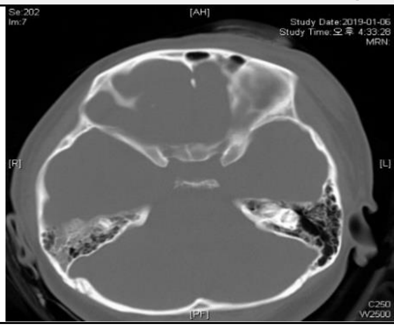
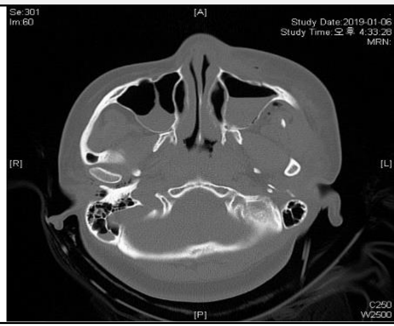
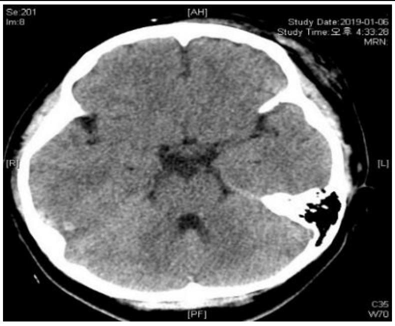
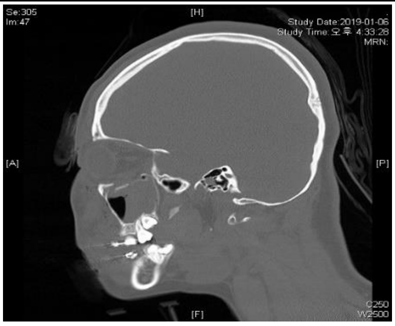
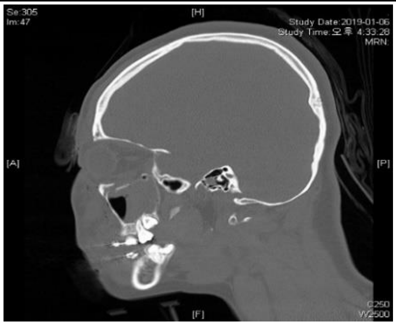
Injury regions	Computed tomography of bodily injuries	
Head (Skull fracture)		
Brain (T-SAH, Cerebral contusion)		
Facial (Maxilla fracture, Multiple orbit fracture, Injury of optic nerve and pathways, Lt. nasal bone, Facial laceration)		
Abdomen (Mesentery laceration, Colon laceration, Omentum laceration, L-spine fracture, Disk herniation on L4)		

Figure 2.20. Diagnostic imaging using Computed Tomography

수술 중 특이사항 ○무 ⊙유

Seat belt sign (+)
하복부에 seat belt line을 따라 subcutaenous layer가 tunneling되어있었다.

Operative findings ☐R0 ☐R1 ☐R2

개복하였을때 다량의 혈액이 확인되었다.
four quadrant packing을 시행한후 whole bowel을 tracing 하였다.
T-colon의 중간부분에 serosa tearing된 부분이 세군데 확인되어 repair하였으며
partial omentectomy를 시행하였다.
lesser sac이 열려있었으며 pancreas 는 특이소견 없음을 확인하였다.
treitz ligament로 부터 40cm 하방에 small bowel mesentery가 두군데 심하게 손상되어 모두 포함하여 약 30cm가량의 소장을 segmental resection with end-to-end anastomosis시행하였다.
Ascending & descending colon에도 serosa tearing된 부분을 repair하였다.

Figure 2.21. In-hospital surgical findings and injury surveillance

2.2. Data source

This retrospective study used the Korea In-Depth Accident Study (KIDAS) database from the Center for Automotive Medical Science Institute (CAMS) at Yonsei University. The data were collected using on-scene investigations of real-world crashes. We analyzed patients from five different regional trauma centers in South Korea from January 2011 to April 2020.

The dataset consisted of road traffic injury information related to the human, vehicle, and crash components and is used to predict the severity of the MVOs' injury. The patients age was recorded for both male and female genders. The restraint system (passive safety device) was assessed, and evidence of wear or fault was recorded. Furthermore, the Principal

Direction of Force (PDOF) was defined as the impact direction. This consists of frontal, side (left and right), and rear-end impacts. Vehicle types were grouped into five categories; sedan, sports utility vehicle, light truck, van, and heavy trailer. The collision partner was defined by considering the two vehicle's mass and size, and assessing them for crash incompatibility. For instance, if the patient's vehicle was heavier than the opponent's crash vehicle, the collision partner would have only been a relatively small component of the overall force sustained by the patient. Finally, the number of impacts between the two vehicles in car-to-car crashes was categorized into single and two or more. Delta-V is a change in velocity from pre-crash to in-crash relating to the vector dynamics of MVCs. The Delta-V was obtained by crash reconstruction using PC-Crash software referring to on-scene 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).

2.3. Study population

Among the 3,928 occupants related to MVCs, we used the data of 1,417 patients aged ≥ 18 to predict severe injuries in C2C crashes (Figure 2.22). We grouped individual patients based on the classification on severity of their injuries. In this study, simple planar crashes were assessed to predict the results based on the complexity of MVCs. Rollovers were excluded from the analysis.

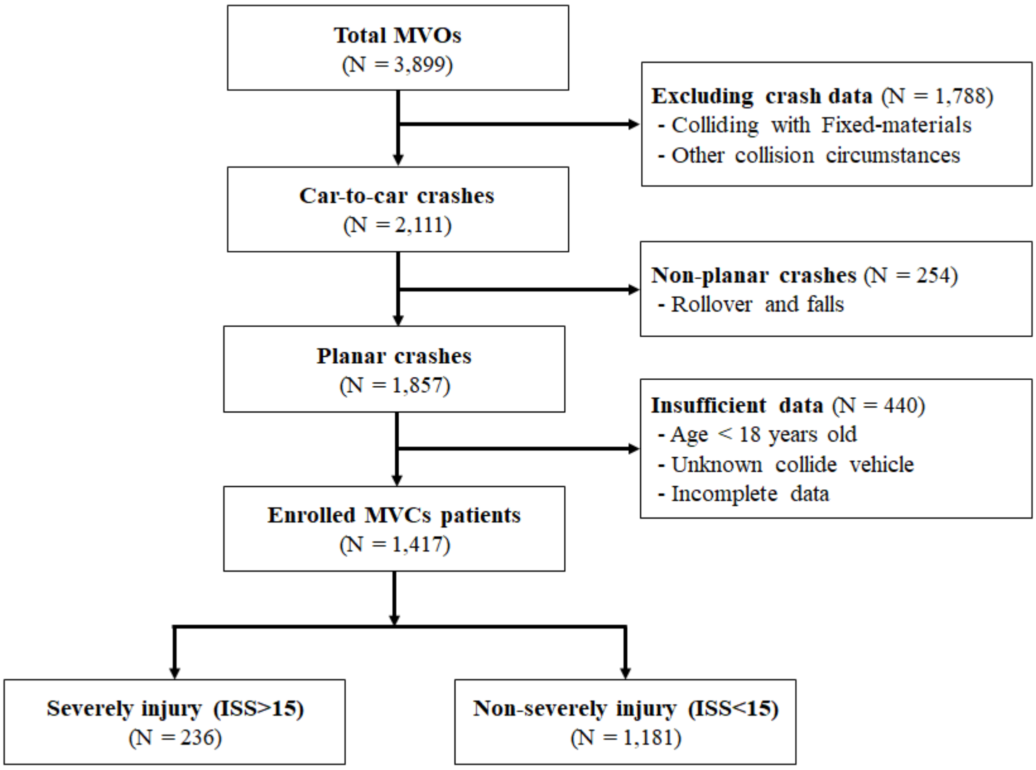


Figure 2.22. Flowchart of data selections of overall crash-related KIDAS dataset

2.4. Sample size estimation

In this study, the sample size required for the MVOs serious injury prediction model was calculated using G-power 3.1 software (Heinrich-Heine- Universität, Düsseldorf). It was calculated as the sample size (F test) for multiple logistic regression (deviation of R^2 from zero), which is the most widely used statistical model.

We set the input parameters as 0.02 for effect size f-square, 0.05 for α error probability, and 0.80 for Power (1-β error probability), and set seven predictors (i.e., age, seat belts usage, the principal direction of force, vehicle type, collision partner, multiple impact, and Delta-V) for the model. Based on the assumption above, the desired sample size is 725 patients. The 1,417 patients used in this study, therefore, more than adequate.

2.5. Injury severity classification

The Injury Severity Score (ISS) is a medical score used to assess trauma severity and was 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 (See Table 2.1). The 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+ 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.

Table 2.1. An example of Injury Severity Score calculation

Region	Description of Injury	AIS	Square top three
Head & Neck	Scalp laceration	110602.1	
Face	Lt. medial orbital wall fracture	251231.2	4
Chest	Rt. 4-9th rib fractures	450203.3	9
Abdomen	Adrenal grand contusion	540212.1	
Extremity	Combined fracture at Lt. radius and ulna	853171.3	9
External	Multiple abrasion	910200.1	
ISS			22

AIS, abbreviated injury scale; ISS, injury severity score

2.6. Study design

This study applied ML analytics through imbalanced clinical data processing to determine the best-performing model according to a binary injury classification. The overall methodological procedure is illustrated in Figure 2.23. 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.

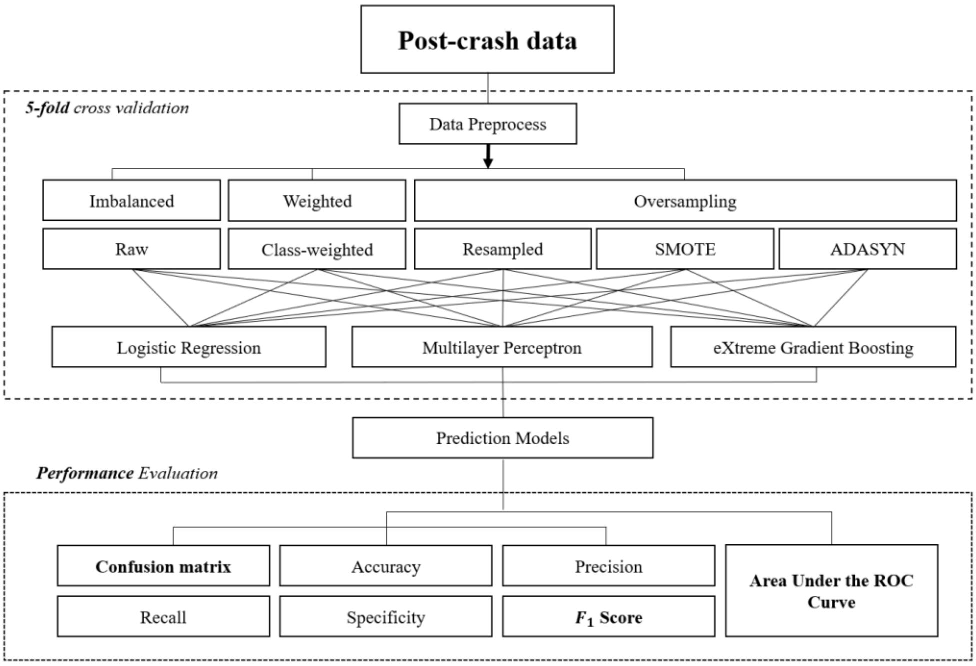


Figure 2.23. Prediction model analytics of MVC occupant injuries

2.7. 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 United States provided recommendations 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, seat belts usage, the principal direction of force (PDOF), vehicular type, collision number, crash partner, and Delta-V) to predict the severity of the patient's injuries in C2C crashes.

2.8. 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 degrees of injury severity [28]. This creates a strong bias for the prediction model's performance, which then 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 then facilitate the calculation of the imbalance ratio (ratio of majority class to minority class) [30]. The calculations of imbalance ratio can be simplified as follows:

$$\text{Imbalance Ratio (IR)} = \frac{N_{maj}}{N_{min}} \tag{2.4}$$

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 extremely imbalance for a ratio higher than 9 [31,32]. The approach to handling an imbalanced class dataset is to select a data sampling techniques that will balance the class. 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 2.2.

Table 2.2. Hyperparameters used in the prediction model

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

2.9. 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 processes sequential tree buildings in parallel. This method can prevent overfitting and improve calculation speed. Among the tree-based models, the performance of this method is excellent, and the importance of its features is still to be determined.

2.10. Model training and testing

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.

2.11. Performance evaluation

In this study, we evaluated the presented ISP models for their internal validity in binary injury severity. We evaluated the proposed ML models using F-measures (F1 score) which were computed based on the harmonic average of precision and recall. Also, accuracy was calculated for performance comparison with other previous studies. These are defined in the following equations.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{2.5}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2.6}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{2.7}$$

$$\text{F1 score} = 2 \times \frac{Precision \times Recall}{Precision+Recall} = \frac{2 \times TP}{2 \times TP+FP+FN} \tag{2.8}$$

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 (Table 2.3).

Table 2.3. 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)*

However, standard errors of false alarms represent misleading predictions, such as over-triage (false positive ratio) and under-triage (false negative ratio) classification. Therefore, this study considered an under-triage level before 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 an equal calculating equation as recall (or sensitivity), and FPR 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 was 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 $AUC \geq 0.9$. Finally, as the AUC approaches 1.0, the response can be interpreted as a complete predictive power outcome [23].

Chapter 3

Results

3.1. Data sampling techniques for class imbalance clinical outcomes

Scatterplots of primary continuous data (age and Delta-V) were used by each sampling technique to configure the data distribution of binary injury severities (Figure 3.1). Since the oversampling was conducted only on severely injured data, the plots show an increased focus on Resample, SMOTE, and ADASYN datasets compared to the imbalance distribution. However, the class-weighted datasets are demonstrated to be equal to the raw data due to the assignment of weights to each class in the initial data. The degree of spread and central tendency of the sampling data was similar in cases of imbalance 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.

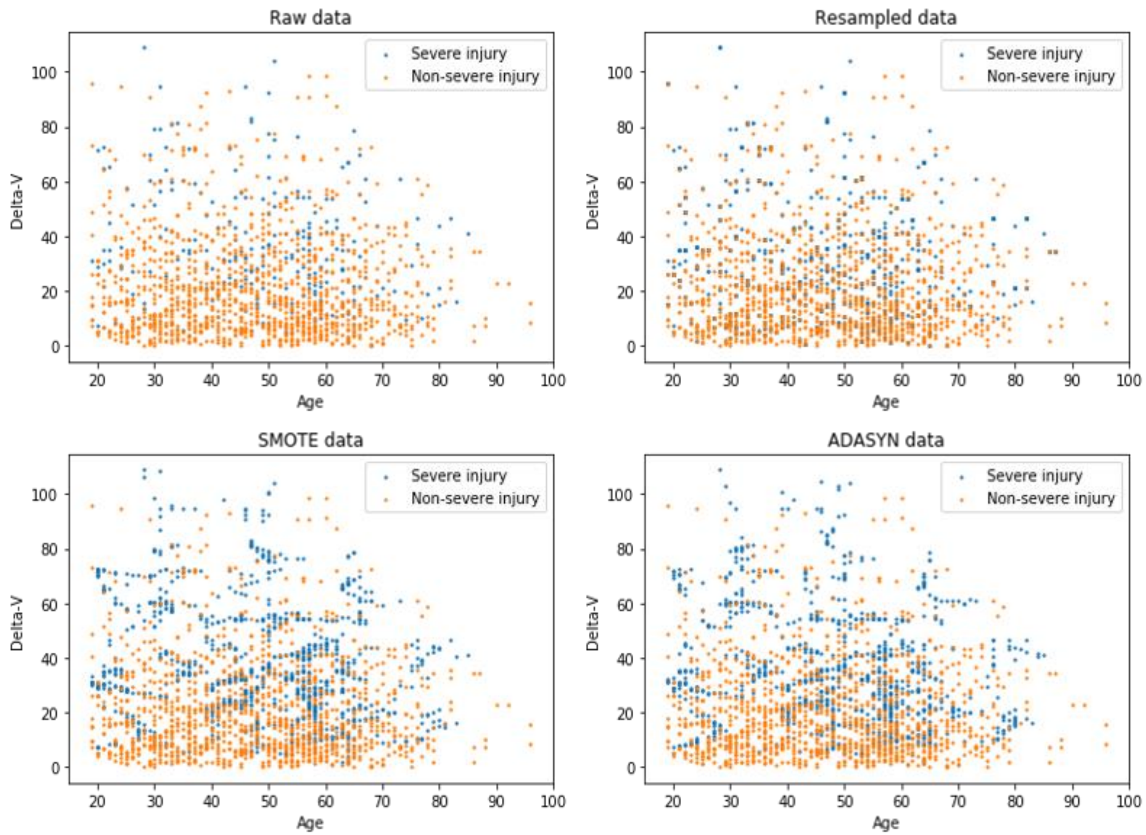


Figure 3.1. Comparison of scatter plots of data obtained from tested oversampling

3.2. Baseline characteristics of crash data distributions enrolled in regional trauma centers

The descriptive data were summarized as a sample for predicting severely injured C2C crash occupants (Table 3.1). According to the classification of injury severity, the data distribution led to performance outcomes 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 the clinical data. Also, the data indicated that young occupants were more engaged with MVCs than older groups.

The proportion 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 vehicle 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 impact incidence with vehicles similar to or larger than the counterpart vehicle was higher. Regarding the number of collisions, the probability of multiple impacts was lower than 10% in 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 to patients based on age and Delta-V distribution.

Table 3.1. Demographic data of MVCs related to trauma

Variables		Descriptions	Frequency (n=1,417)	Ratio (100%)
Dependent variables	ISS (binary)	Severe injury	236	16.7
		Non-severe injury	1,181	83.3
Independent variables	Age	54 years under	907	64.0
		55-64 years	312	22.0
		65 years over	198	14.0
	Seat belts usage	Fastened	930	65.6
		Unfastened	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	1,299	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

ISS, injury severity score; PDOF, principal direction of force; SUV, sports utility vehicle

3.3. Confusion matrices for triage controls on injury severity classifications

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 3.2). 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 more sampling numbers than ADASYN. However, 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 model. 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 reporting 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-performing 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 3.2. A comparison of the confusion matrix used to predict injury severity classification

Dataset		Classifier	N	Balance		Confusion matrix			
				Positive	Negative	TP (TPR)	FN (FNR)	FP (FPR)	TN (TNR)
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)**

SMOTE, synthetic minority oversampling technique; ADASYN, adaptive synthetic sampling; LR, logistic regression; MLP, multilayer perceptron; XGB, extreme gradient boosting; TP, true positive; TRP, true positive ratio; FN, false negative; FNR, false negative ratio; FP, false positive; FPR, false positive ratio; TN, true negative; TNR, true negative ratio

3.4. Evaluating predictive performances on injury severity prediction classifiers

Table 3.3 shows the classification performance of injury severity results obtained from the confusion matrices for each sampled classifier in Table 3.2. Thanks to these matrices, it has been determined how injury severities were predicted correctly by referring to Table 2.3. 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 the 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 perform similarly, although the outperformed classifiers are machine learning models (especially in XGBoost) based on the 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 3.3 have been obtained using the metrics columns in Table 3.2.

Table 3.3. Predictive performance of severely injured MVOs based on 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**

SMOTE, synthetic minority oversampling technique; ADASYN, adaptive synthetic sampling; LR, logistic regression; MLP, multilayer perceptron; XGB, extreme gradient boosting; AUC, area under the curve

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 accuracy = (161+155)/(161+155+80+70) =0.668, precision = 161/(161+80) = 0.668, recall = 161/(161+77) = 0.676, and the F1 score = (2×161)/((2×161)+80+77) = 0.672.

The matrix parameters for the SMOTE-based MLP classifier are presented as TP = 205, FN = 20, FP = 109, and TN = 139. This may be calculated as accuracy = (205+139)/(205+139+109+20) = 0.727, precision = 205/(205+109) = 0.653, recall = 205/(205+20) = 0.911, and the F1 score = (2×205)/((2×205)+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 accuracy = (221+172)/(221+172+29+51) = 0.831, precision = 221/(221+51) = 0.813, recall = 221/(221+29) = 0.884, and the F1 score = (2×221)/((2×221)+51+29) = 0.847. Other classifiers may also be calculated by referring to equation 2.5-2.8.

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 graphic plot illustrations using the ROC curve (Figure 3.2). 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.

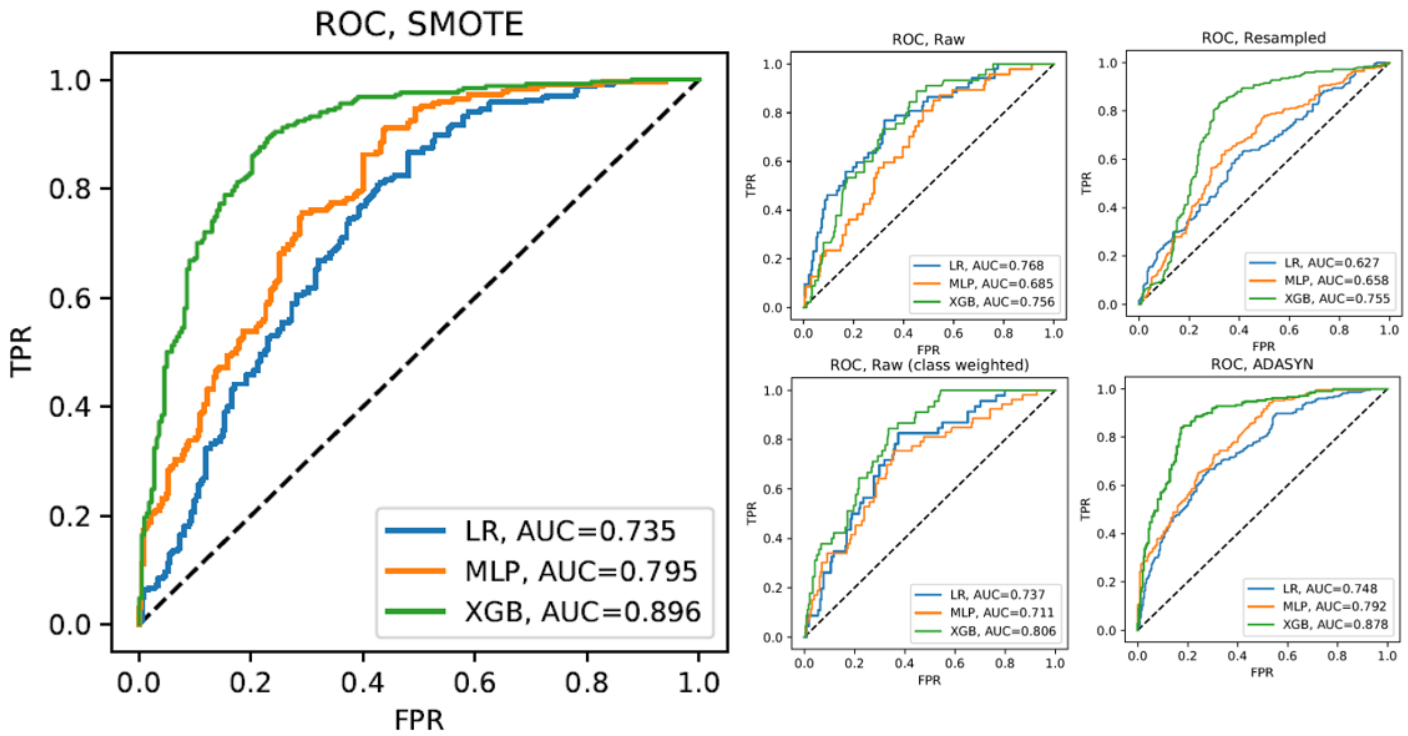


Figure 3.2. Comparison of ROC curve using various sampling techniques

3.5. Rankings of features important to outperforming SMOTE-XGBoost

This study suggests which indicators are essential in ensuring the best performance model (SMOTE-XGBoost) when predicting a patient's injury classification (Table 3.4). 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.

Table 3.4. Features importance ranking of outperformed classifier

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
Seat belts 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

PDOF, principal direction of force

Chapter 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, seat belts usage (fastened or unfastened), 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 statistical models (see Table 3.2).

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 clinical data class imbalance. 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 performance [25,40,43,46-48]. Compared with statistical methodologies, the latest machine learning and deep learning techniques enhance the predictive performance. In the previous study, various classifiers were used to compare the prediction performance of each model. These include a 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 of the latest methods have received the attention that implies that they are superior to the 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 the suggested models. Therefore, the present work is evaluating the performance of 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 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 an ISP using relatively small datasets.

Several studies proposed outperforming methods for ISP engaged with MVOs comparing various machine intelligence in binary classification (Table 3.5). 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 traditional statistical techniques [40, 61]. It is assumed that this may influence the performance of the model's performance depending on the parameter selection in predicting the binary class of injury outcome. In particular, Delen et al [48] showed the best predictive performance in SVM; however, the under-triage results were missing, so they could not support clinical insights in the real-world. 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 outperformed all of others especially considering the under-triage rate in medical terms. However, additional studies are required to apply techniques based on the optimal parameters when considering complex crash injury mechanisms.

Table 3.5. Performance comparison between the proposed models and previous studies

Studies	Data (year) Crash data Variables	Crash injury targets	Class break-down (%) Imbalance Ratio	Data sampling	Classification Models	Performance evaluation (%)	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=1,417 • **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

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-weighted, 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; BNB, Bernoulli Naïve Bayes; SGD, stochastic gradient descent; XGB, extreme gradient boosting; MLP, multilayered perceptron

Many ISP models have been developed to consider all 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 body structure mismatches increased the risk of injury severity to MVOs with disadvantageous self-protective capacities due to the 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 considers the crash incompatibility of two-vehicle scenarios. The collision partner was confirmed as a highly 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 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 a consistent golden hour [22]. It can 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.

Chapter 5

Limitation

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 is 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.

Chapter 6

Conclusion

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.

References

[1] WHO, Global status report on road safety 2018. https://www.who.int/publications/i/item/9789241565684, 2018. (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, Prehosp. Disaster Med. 35 (2020) 508-515. https://doi.org/10.1017/S1049023X20000904.

[5] R.C. Mackersie, History of trauma field triage development and the 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, 2011, MMWR Recomm. Rep. 61 (2012) 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. Atlanta, GA: Centers for Disease Control and Prevention; 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. Safety 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. Prev. 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 Transp 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, Prehosp. Disaster Med. 34 (2019) 356-362. https://doi.org/10.1017/S1049023X19004515.

[23] D.W. Hosmer, S. Lemeshow, R.X. Sturdivant, Applied logistic regression. 3rd ed. New York, NY: Wiley; 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. Barrington, IL: Association for the Advancement of Automotive Medicine; 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, J. H. (2020). Adjusting the imbalance ratio by the dimensionality of imbalanced data. Pattern Recognit Lett, 133(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. 2nd ed. New York, NY: Springer; 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. 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence); 2008. p. 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. 2019 IEEE Intelligent Transportation Systems Conference (ITSC). Auckland, New Zealand: IEEE; 2019. p. 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) S126-133.

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, Infrastructures (Basel) 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. Public 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, Accident 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. 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. Transp. Health 2017, 4, 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, doi:10.1080/15389588.2016.1246724.

[51] Administration, N.H.T.S. Crash injury research and engineering network (ciren) program report, 2002. DOT HS 2005, 809, 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. Applied Sciences 7(2017), 476, doi: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. Public 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. Contr. 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. Transp. 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.

Abstract in Korean

클래스 불균형 데이터의 오버샘플링 기반 기계학습 기법을 적용한 교통사고 탑승자의 중증외상 예측 모델

공준석
의학과
연세대학교 대학원

병원 전 단계에서 교통사고 환자의 인체상해 예측은 환자의 중증도분류에 대한 정확한 의사결정과 적절한 이송체계를 통해 인명피해를 경감시키는 효과가 있다. 최근 사고 현장에서 즉각적인 상해유형 판별을 위해 텔레메틱스를 기반한 자동검출 시스템의 법제화가 각 국에서 도입되고 있으며, 이를 위한 외상환자의 상해예측 모델에 대한 요구가 부각되고 있다. 그러나 환자의 상해예측 모델은 데이터의 클래스 불균형(Class imbalance)에 따라 실제 왜곡된 예측과 성능저하를 초래할 수 있다. 또한, 아직까지 교통사고 환자의 상해를 판별하기 위한 임상자료의 균등화(balancing)를 통한 최적화된 모델의 부재로 다양한 모델간의 성능 비교가 필요하다. 본 연구는 국내 5 개 지역의 응급의료센터에 내원한 차대차 탑승자 교통사고 환자를 대상으로 상해중증도 판별을 개선하기 위해 최신의 기계학습 모델의 성능을 평가하고자 한다.

본 연구는 2011 년 1 월부터 2021 년 4 월까지 한국형 자동차사고-인체상해 구축(Korea In-Depth Accident Study, KIDAS) 데이터베이스에 등록된 1,417 명의 교통사고 환자를 대상으로 선정하였다. 상해중증도에 대한 분류는 손상중증도점수(Injury Severity Score, ISS) 기준 15 점 이상을 중상해 환자로

선별하였다. 다양한 사고유형에 따라 보다 정밀한 예측성능 확보를 위해 전복사고를 제외한 평면충돌 사고를 고려하였다. 또한 차대차 사고에서 두 차량 간의 충돌 부조화(crash incompatibility)을 고려하여 서로 다른 차량 세그먼트 구성을 분류하였다. 탑승환자의 중증도분류 결과에 따른 데이터 불균형성을 극복하기 위해 네 가지의 데이터 샘플링 기법(i.e., class-weighting, resampling, synthetic minority oversampling, and adaptive synthetic sampling)을 사용하였다. 교통사고 환자의 상해예측 판별을 위한 기계학습 모델은 logistic regression, extreme gradient boosting (XGBoost), 그리고 multilayer perceptron (MLP)로 선정하였다. 모델 성능을 향상시키기 위해 하이퍼파리미터를 조정하고 5 겹 교차검증을 통해 각 모델의 과적합을 방지하였다. 외상환자의 상해예측은 과소분류 10% 이하의 수준을 지닌 모델을 기반으로 모델의 성능을 평가하였다.

본 연구에서 데이터 샘플링 기법을 적용한 SMOTE 와 ADASYN 모델이 클래스 불균형 데이터 보다 예측 성능이 높았다. 특히 SMOTE 기반 XGBoost 모델에서 가장 우수한 예측 성능을 보였다. 해당 모델을 활용한 특성중요도에서 두 차량간의 속도변화량(Delta-V)이 교통사고 탑승자의 상해 예측에 기여한 주요 요인으로 확인되었다.

이러한 결과는 환자의 중증도분류에 따른 클래스 불균형을 데이터 샘플링 기법을 구현하여 상해 심각도에 대한 개선된 예측 성능을 기대할 수 있다. 따라서, 교통사고 환자의 상해 예측을 위해 활용되는 변수의 유형에 따른 샘플링 기법과 학습모델 선정이 고려되어야 한다.

핵심되는 말 : 상해예측모델, 기계학습, 차량 탑승자, 권역외상센터, 손상중증도, 속도변화량, 과소분류, 클래스 불균형, 오버샘플링, 한국형 자동차사고-인체상해 심층분석 자료