

# PBL 최종보고서

Track Expert 2팀 3조

멘토 : 원상혁님

멘티: 김유신, 권오민, 박관우, 송나흔, 이현주

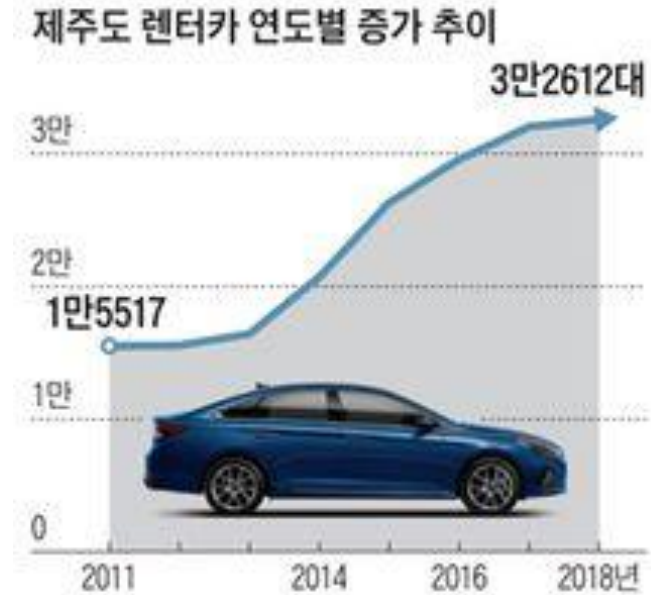
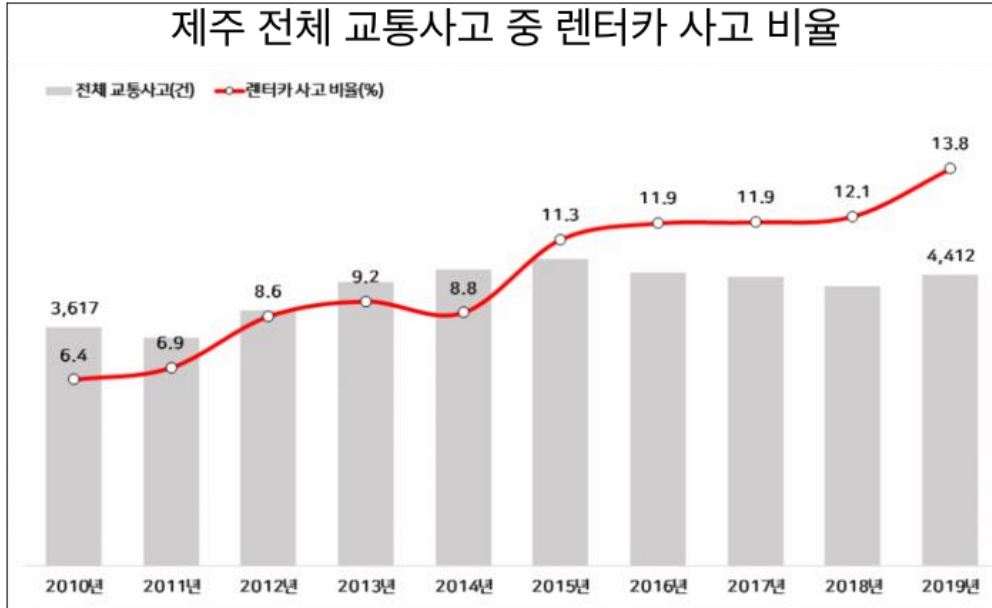
# Contents

- I. 추진 배경
- II. 문제 정의
- III. 데이터 분석 결과
- IV. 활용 방안
- V. 인턴십 소감

# Executive Summary

<b>프로젝트 명(주제)</b>	제주도 내 렌터카 교통사고의 경중 분석
<b>프로젝트 목적</b>	<ul style="list-style-type: none"> <li>• 매년 제주도에서 일어나는 교통사고 중 렌터카 비율이 10% 이상을 차지하고, 그 비율은 지속적으로 높아지고 있다.</li> <li>• 국내 다른 지역과 비교하였을 때 제주도의 렌터카 사고 발생률은 전국 평균보다 6-10배 높고, 렌터카 1만 대당 사망자 수 역시 4배 가량 높다고 한다.</li> <li>• 이렇게 렌터카 사고의 위험성이 대두되는 가운데, 사망/중상사고의 비율은 상당히 높은 편이다.</li> </ul>
<b>문제 정의</b>	<ul style="list-style-type: none"> <li>• 렌터카 사고 데이터를 사망/중상, 경상/부상으로 이진 분류하여 렌터카 사고의 경중 예측 모델을 구축한다.</li> <li>• 구축한 모델을 해석하여 변수 별 중요도와 각 특성 별 데이터들이 분류과정에 어떻게 영향을 주는지 확인한다.</li> </ul>
<b>사용 데이터</b>	<ul style="list-style-type: none"> <li>• 교통사고분석(TAAS) 홈페이지 내 연도 별 제주도 내 교통사고 데이터                         <ul style="list-style-type: none"> <li>- 파일 형식 : excel</li> <li>- 변수 : 사고번호, 사고일시, 사고위치, 사고유형, 법규위반, 노면 상태, 기상상태, 가해운전자 상해정도, 피해자 운전 상해정도</li> </ul> </li> <li>• 제주 기상 데이터                         <ul style="list-style-type: none"> <li>- 파일 형식 : excel</li> <li>- 변수 : 지점별 기상데이터 (기온, 강수량, 풍속, 습도, 현지기압, 적설, 시정, 지면온도)</li> </ul> </li> </ul>
<b>결과 및 활용방안</b>	<ul style="list-style-type: none"> <li>• Cost-sensitive 방식을 적용한 catboost로 최종 모델을 적합했다.</li> <li>• 사고유형과 발생 지역이 가장 큰 중요도를 보이고, 일부 변수들은 모델의 분류에 영향을 전혀 미치지 않는 것으로 확인되었다.</li> <li>• 기온이 높고 시정이 좋은 화창한 날일수록, 풍속이 빠를수록 발생사고의 경중이 낮은 경향성을 보인다</li> <li>• 제주 렌터카 사고의 위험도를 낮출 수 있는 신규 지자체 정책 방향 제시, 관광객 &amp; 렌터카 업체의 모델을 활용한 선제적 대응과 같은 다양한 활용도를 기대한다.</li> </ul>

# I. 추진 배경



“ 다만 제주의 경우 교통사고 100건당 사망자수를 기준으로 한 치사율에서 전국평균 1.46보다 높은 1.50을 기록, 이 부분에 대한 집중적인 보완이 필요할 것으로 보인다.

- 출처: 제주교통매거진 ”

- 매년 제주도에서 일어나는 교통사고 중 렌터카 비율이 10% 이상을 차지하고, 그 비율은 지속적으로 높아지고 있다.
- 국내 다른 지역과 비교하였을 때 제주도의 렌터카 사고 발생률은 전국 평균보다 6-10배 높고, 렌터카 1만 대당 사망자 수 역시 4배 가량 높다고 한다.
- 제주도 렌터카의 지속적인 증가추세와 맞물려 렌터카 교통사고는 제주의 안전한 교통상황을 위협하고 있는 상황이다.
- 이렇게 렌터카 사고의 발생이 잦아지는 가운데, 사망/중상사고의 비율은 상당히 높은 편이다.

## II. 문제 정의

제주 렌터카사고의 문제가 지속적으로 대두되는 가운데,  
본 프로젝트에서는 다음과 같이 문제를 정의한다.

1. 렌터카 사고 중 사망/중상, 경상/부상을 이진 분류하여  
렌터카 사고의 경중예측 모델을 구축한다.

2. 구축한 모델을 해석하여 변수 별 중요도와 각 특성 별 데이터들이  
분류과정에 어떻게 영향을 주는지 확인한다.

### 1. 연도별 제주도 내 렌터카 교통사고 데이터

가. 파일 형식 : Excel

나. 변수

- 1) 도로 요인 : 노면상태, 기상상태, 도로형태
- 2) 차량 특성 : 가해운전자 차종, 피해운전자 차종
- 3) 사고 특성 : 사고번호, 사고일시, 사고내용, 사망/중상/경상/부상자수, 사고 유형
- 4) 운전자 인적사항 : 가해운전자 성별/연령/상해정도, 피해운전자 성별/연령/상해정도
- 5) 환경 요인 : 요일, 시군구

다. 자료 출처 : 교통사고 분석시스템 (TAAS) <http://taas.koroad.or.kr/>

라. 기간 : 2017 - 2021

### 2. 기상데이터

가. 파일 형식 : Excel

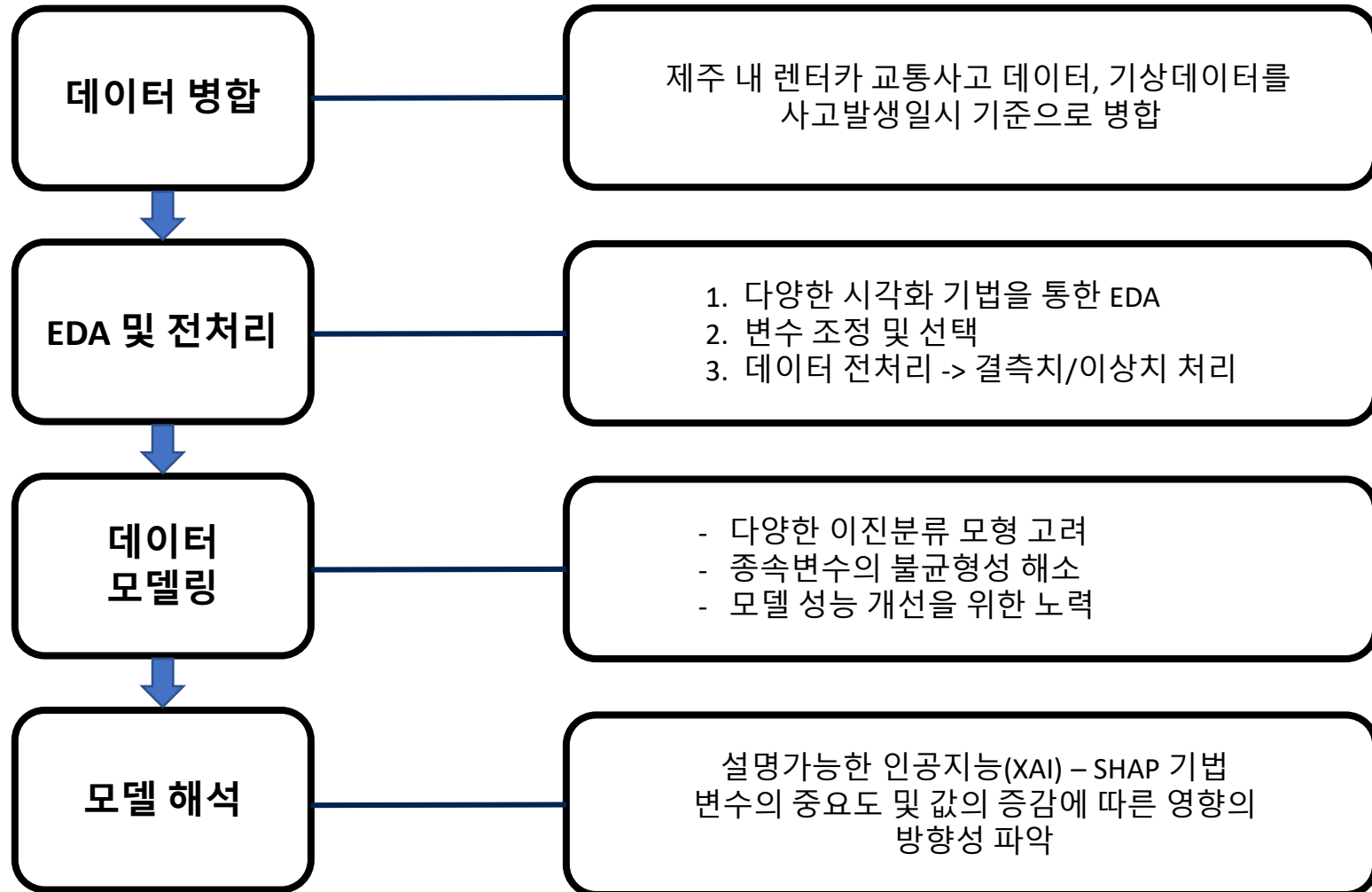
나. 변수 : 지점, 지점 명, 일시, 기온, 강수량, 풍속, 습도, 현지기압, 적설, 시정, 지면온도

다. 자료 출처 : 기상청 기상자료 개방 포털 - 종관기상관측(ASOS) 자료

<https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36>

라. 기간 : 2017-2021





## III. 데이터 분석 결과

## 1. 데이터 병합 (초기 데이터 생성)

### 가. 데이터 병합

- 제주 렌터카 교통사고 데이터와 제주 기상 데이터를 병합한다.
- 병합 기준 : 사고일시, 기상 관측 지점
- 병합 과정
  - 1) 기존 사고데이터의 string 형식의 사고일시를 datetime 형태로 변환  
ex) YYYY년 MM월 DD일 14시 -> YYYY-MM-DD HH:00:00
  - 2) 연도별 기상 데이터를 제주/서귀포로 나누어 구분 -> 각 지점별 5년의 기상데이터 생성
  - 3) 사고 데이터를 기준으로 같은 시간, 동일 지점에 해당하는 기상데이터를 병합

[제주 렌터카 교통사고]

사고번호	사고일시	요일	시군구	사고내용
2.017010e+15	2017년 1월 1일 15시	일요일	제주특별자치도 제주시 애월읍	붕상사고
2.017010e+15	2017년 1월 1일 18시	일요일	제주특별자치도 제주시 삼도이동	중상사고
2.017010e+15	2017년 1월 21시	일요일	제주특별자치도 서귀포시 남원읍	중상사고
2.017010e+15	2017년 1월 2일 10시	일요일	제주특별자치도 서귀포시 성산읍	경상사고
2.017010e+15	2017년 1월 3일 21시	화요일	제주특별자치도 서귀포시 성산읍	경상사고
...	...	...	...	...
2.021123e+15	2021년 12월 28일 12시	화요일	제주특별자치도 제주시 일도이동	경상사고
2.021123e+15	2021년 12월 29일 11시	수요일	제주특별자치도 제주시 일도이동	경상사고
2.021123e+15	2021년 12월 29일 14시	수요일	제주특별자치도 제주시 한림읍	경상사고
2.021123e+15	2021년 12월 30일 12시	목요일	제주특별자치도 서귀포시 서대동	경상사고
2.021123e+15	2021년 12월 31일 12시	금요일	제주특별자치도 제주시 애월읍	경상사고

2738 rows x 21 columns



[제주 기상 데이터]

지점명	일시	기온(* C)	풍속(m/s)	습도(%)
184.0 제주	2017-01-01 00:00:00	6.0	2.3	...
184.0 제주	2017-01-01 01:00:00	6.6	2.6	...
184.0 제주	2017-01-01 02:00:00	7.1	2.4	...
184.0 제주	2017-01-01 03:00:00	7.5	2.7	...
184.0 제주	2017-01-01 04:00:00	8.1	2.4	...
...	...	...	...	...
184.0 제주	2021-12-31 19:00:00	4.8	5.6	...
184.0 제주	2021-12-31 20:00:00	4.5	4.8	...
184.0 제주	2021-12-31 21:00:00	4.1	4.2	...
184.0 제주	2021-12-31 22:00:00	3.8	5.0	...
184.0 제주	2021-12-31 23:00:00	3.6	4.4	...

43823 rows x 10 columns



[초기 병합 데이터]

요일	사망자수	중상자수	경상자수	시	읍면동	기온(* C)	풍속(m/s)	습도(%)
일요일	0.0	0.0	0.0	제주	애월읍	11.9	3.4	63.0
일요일	0.0	1.0	0.0	제주	삼도이동	10.7	1.4	69.0
일요일	0.0	1.0	0.0	서귀포	남원읍	11.2	0.7	70.0
일요일	0.0	0.0	1.0	서귀포	성산읍	11.9	0.2	90.0
화요일	0.0	0.0	1.0	서귀포	성산읍	10.5	1.1	76.0
...	...	...	...	...	...	...	...	...
화요일	0.0	0.0	1.0	제주	일도이동	8.0	2.5	60.0
수요일	0.0	0.0	1.0	제주	일도이동	10.9	5.4	57.0
수요일	0.0	0.0	3.0	제주	한림읍	10.4	5.7	60.0
목요일	0.0	0.0	1.0	서귀포	색달동	8.2	3.7	45.0
금요일	0.0	0.0	4.0	제주	애월읍	5.9	4.9	56.0

사고 데이터

기상 데이터

## 2. 초기 데이터 EDA



### 가. 종속변수 생성

- 탐색적 데이터 분석에 앞서 종속변수를 생성 -> EDA시 종속변수 별 분포까지 고려할 수 있도록 한다

### 나. 변수별 특성 파악

- 수치형 변수의 경우 히스토그램, 범주형 변수는 막대그래프를 이용해 변수별 분포와 특성 파악한다.

- 파악 결과 종속변수(0/1의 구분)에 관계없이 비슷한 분포를 보이는 변수가 대부분이다.

위와 같은 상태에서 모델링을 할 경우 중상/사망, 부상/경상을 제대로 구분하지 못할 확률이 높다.

- eda 결과를 반영하여 설명력을 높일 수 있도록 변수를 조정하거나 선택하는 과정을 수행한다.

## 3. 데이터 전처리

### 가. 변수 선택

- 1) 미래 참조의 위험이 있는 변수 제거
  - 본 프로젝트는 렌터카 사고가 발생함에 있어 그 원인을 파악하는 것을 목표로 하는데, 사고 이후에 파악가능한 정보를 사용할 시 미래 참조를 피할 수 없으며 정확한 해석이 어렵다.
  - 이에 ‘피해 운전자 관련 변수’, ‘경상/부상/중상/사망자 수’, ‘가해운전자 상해정도’의 변수들은 분석에서 제외한다.

### 나. 변수 조정

- 1) 요일변수 -> 주말&공휴일 변수로 변경
  - 기존 요일 변수는 일곱가지의 요일을 고루 포함하고 있을 확률이 높으며 교통량과 직결되는 공휴일 관련 정보가 반영 되어있지 않다.
  - 주말&공휴일 / 평일로 구분하여 변수를 조정함으로써 더 나은 설명력을 기대한다.
- 2) 가해 운전자의 나이 -> 데이터 type을 int형으로 변환
- 3) 사고 발생지역 간략화
  - 세부 행정동으로 구성 되어있어 기존 변수로는 분석에 어려움이 있다.
  - 더 포괄적인 행정구역으로 구분하여 범주의 개수를 줄인다.
- 4) 사고 유형 간략화
  - 기존 ‘사고유형 대분류 - 사고유형 소분류’ 형식의 데이터를 분할하여 각각 사고유형 1, 유형 2 변수를 생성한다
- 5) 노면 상태 변수의 간략화
  - 기존 노면 상태 변수의 경우 ‘포장여부 - 건조상태여부’의 형식으로 구성되어 있는데, eda 결과 비포장 관련 데이터는 단 7개 존재한다.
  - 비포장 관련 데이터를 제거하고, 노면의 건조상태 여부 데이터만 사용한다.

[제주 지역 구분]

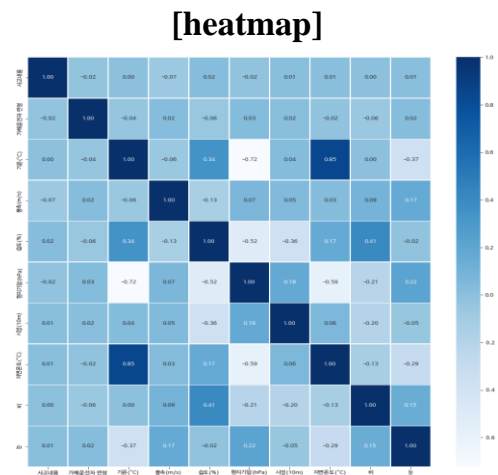


## 3. 데이터 전처리

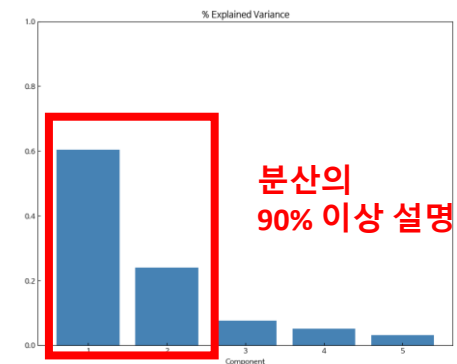
### 다. 파생변수 생성

1) 수치형 변수의 상관관계, 다중 공선성 파악

- 독립변수들 사이에 강한상관관계가 나타난다면 하나의 독립변수의 변화가 다른 독립변수에 영향을 미치므로 결과적으로 모델이 불안정해진다.
- heatmap과 VIF를 이용해 상관관계와 다중공선성을 시각화 한 후, 높은 변수들에 한해 PCA를 진행한다.



VIF Factor	features
0	눈
1	비
2	풍속(m/s)
3	가해운전자 연령
4	시정(10m)
5	지면온도(°C)
6	습도(%)
7	기온(°C)
8	현지기압(hPa)



	PC1	PC2	PC3	PC4	PC5
시정(10m)	0.036592	0.365917	0.911345	0.086266	0.163626
지면온도(°C)	0.532136	0.373487	-0.190630	-0.449258	0.582379
습도(%)	0.335815	-0.809393	0.353401	-0.316562	0.083707
기온(°C)	0.655658	0.224220	0.064779	0.004689	-0.718066
현지기압(hPa)	0.415731	0.145674	0.063410	-0.830958	-0.333818

2) PCA를 통한 파생변수 생성

- '시정', '지면온도', '습도', '기온', '현지기압' 변수들에 대해 PCA 진행한다.
- PCA 결과 pc1, pc2 가 전체 분산의 90%를 설명하므로, 두 변수를 각각 weather 1, weather 2로 명명하고 다섯 변수를 대체하는 파생변수로서 사용한다.

## 1. 최종 데이터

변수명	변수 내용	dtype	변수명	변수 내용	dtype
일시	사고 발생 일시	object	읍면동	세부 주소	object
주말/공휴일	주말&공휴일 -> 1 공휴일이 아닌 평일 -> 0	Int64	풍속	단위:m/s	Int64
사고내용	중상&사망 -> 1 경상&부상 -> 0	Int64	비	비가 온 날 -> 1 오지 않은 날 -> 0	Int64
사고유형	차대차, 차대 사람 등의 사고 주체	object	눈	눈이 온 날 -> 1 오지 않은 날 -> 0	Int64
도로형태	단일로, 교차로 등의 구분	object	노면상태	건조/젖음 등의 구분	object
가해운전자 차종	승합/승용의 구분	object	Weather1	PCA를 통한 날씨변수 (시정,지면온도,습도,기온, 현지기압)	float64
가해운전자 성별	남/여의 구분	object	Weather2	동일	float64
가해운전자 연령	가해 운전자의 연령	Int64			
시	서귀포/제주의 구분	object			

가. 수집 기간 : 2017-01-01 ~ 2021-12-31

나. 종속 변수 : 사고유형

다. 독립 변수 - 도로 요인 : 도로형태, 노면 상태

- 사고 특성 : 사고 일시, 사고 내용, 사고 유형

- 환경 요인 : 주말/공휴일, 시, 읍면동

- 차량 특성 : 가해운전자 차종

- 운전자 인적사항 : 가해운전자 차종/성별/연령

- 기상 요인 : 풍속, 비, 눈, weather1, weather 2

### 1. Catboost 모델링

#### 가. 모델선정

##### 1) catboost 선정

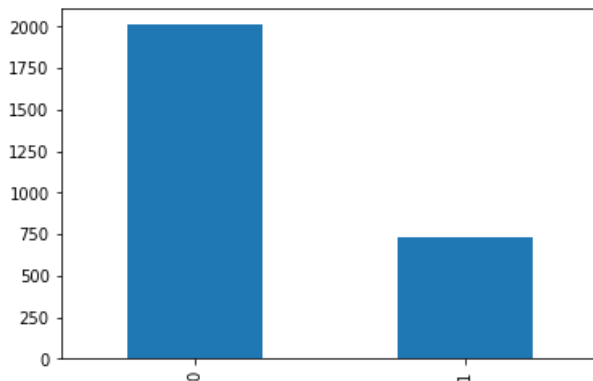
- 동일한 조건에서 모델을 적합시킨 후 모델별 accuracy와 f1 score 비교 결과 catboost로 선정했다.
- catboost는 categorical feature를 처리하는데 중점을 둔 알고리즘으로 범주형 변수가 다수 존재하는 본 프로젝트의 데이터를 효율적으로 분석 할 수 있을 것이라 기대한다.

#### 나. 최종 모델

##### 1) 불균형성 해소

- 종속변수인 ‘사고내용’은 1, 0 으로 이진분류 되어있고, 그 비중의 차이가 매우 크다.
- 별 다른 조치를 취하지 않고 학습시킨다면 0에 대한 과적합이 일어날 수 밖에 없다.
- 이에 모델에 가중치 설정(cost-sensitive)을 달리해가며 적정 가중치를 선별하여 적용함으로써 불균형을 해소한다.
- train/test를 분리할 시 증화추출을 통해 클래스의 비율이 동일하게 반영되도록 한다.

[종속변수 class 별 개수]



[가중치 별 모델의 성능]

Weights	Accuracy	F1 score
가중치 미적용	0.7339	0.0
{0:1, 1:2}	0.6771	0.2903
{0:1, 1:3}	0.6183	0.4851
{0:1, 1:4}	0.4899	0.4715

최종 선정



## 2. 결과 해석

### 가. XAI - SHAP

#### 1) XAI 알고리즘

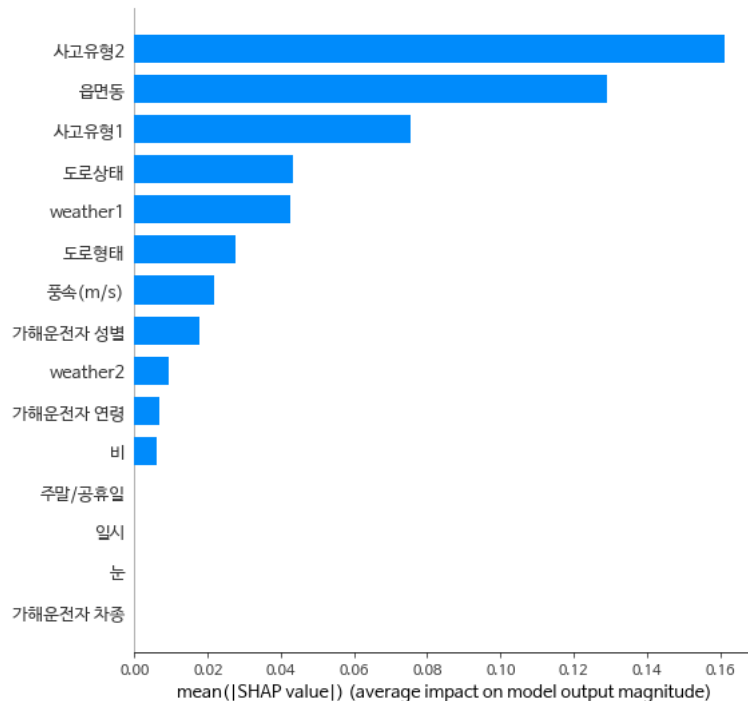
- 설명 가능한 알고리즘 (eXplainable AI)으로, 블랙박스 알고리즘의 해석을 가능케 한다.

#### 2) shap 기법

- shapley value를 이용해 변수의 중요도 및 영향의 방향을 시각화 하여 보다 직관적인 해석을 기대할 수 있다.

- 일반 트리 기반 모델의 중요도와 달리, shap의 변수 중요도는 변수간 종속성을 고려한다.

### 나. Summary plot - bar



#### 1) 중요도 시각화

- 사고유형 1&2, 읍면동이 압도적으로 높은 중요도를 보인다

- 그 외의 도로상태, weather 1&2, 풍속, 가해운전자 성별, 가해운전자 연령, 비 등의 변수들은 이진분류에 미미한 영향을 미친다

- 주말/공휴일, 일시, 눈, 가해운전자 차종은

이진분류에 있어 아무런 영향을 미치지 않는다.

#### 2) 중요도 해석

- 사고의 특성 자체(사고유형 1&2)가 사고의 경중을 좌우한다.

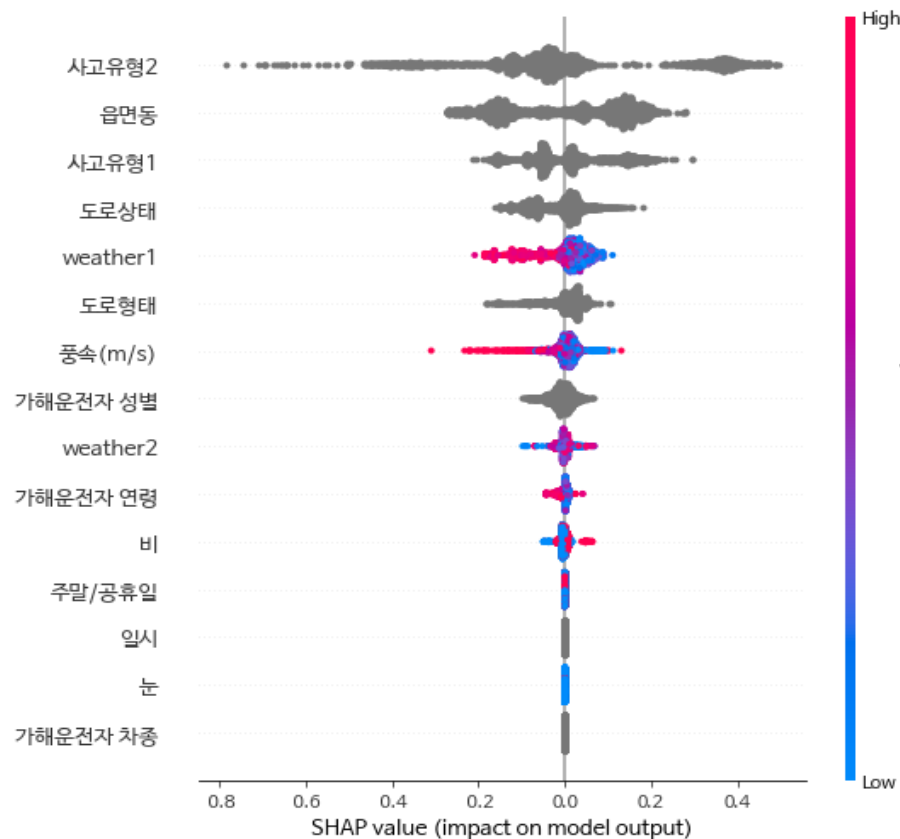
이는 사고 당시의 충돌 상황이 상해정도에 가장 큰 영향을 미친다고 해석 할 수 있다.

- 사고 발생 지역 또한 큰 영향을 미친다.

이는 타 환경변수의 중요도가 낮다는 사실과 더불어 보았을 때 도로의 상황과 주변 환경보다 지역별 특색(교통량, 관광지)이 더 큰 중요도를 가진다고 유추할 수 있다.

### 2. 결과 해석

#### 다. Summary plot – shap value



#### 1) summary plot 해석 방식

- y축 : 각 특성 / x 축 : shapley value
- 색깔 : 특성 값. 빨간색으로 갈수록 높은 값.
- x축에서 0을 기준으로 왼쪽으로 갈수록 경상/부상에, 오른쪽으로 갈수록 중상/사망에 기여함.
- 범주형 변수에 대한 해석은 불가능

#### 2) summary plot 수치형 변수 해석

##### 가) weather 1

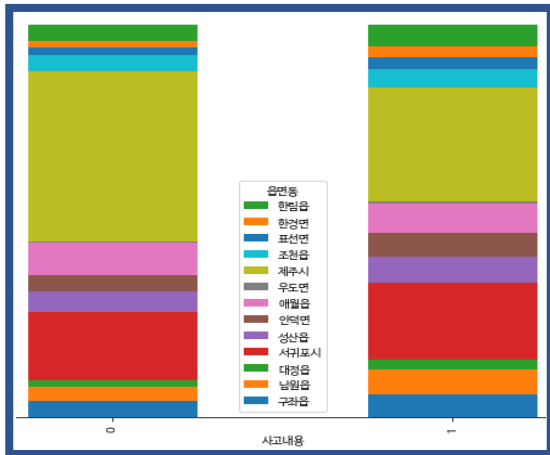
- weather 1은 시정, 지면온도, 현지기압, 습도, 기온이 모두 양의 주성분 계수를 가지는 주성분이다.
- 기온이 높고 시정이 좋은 화창한 여름에 가까울수록 사고의 경중이 낮고, 이와 반대되는 기상상태일수록 사고의 경중이 높다고 해석할 수 있다.

##### 나) 풍속

- 풍속의 경우 바람의 세기가 강할수록 경상/부상에, 약할수록 중상/사망에 기여한다.
- 이는 거센 바람에 의해 운전이 제약을 받을 시 안전에 더욱 주의를 가하는 경향성으로 해석할 수 있다.

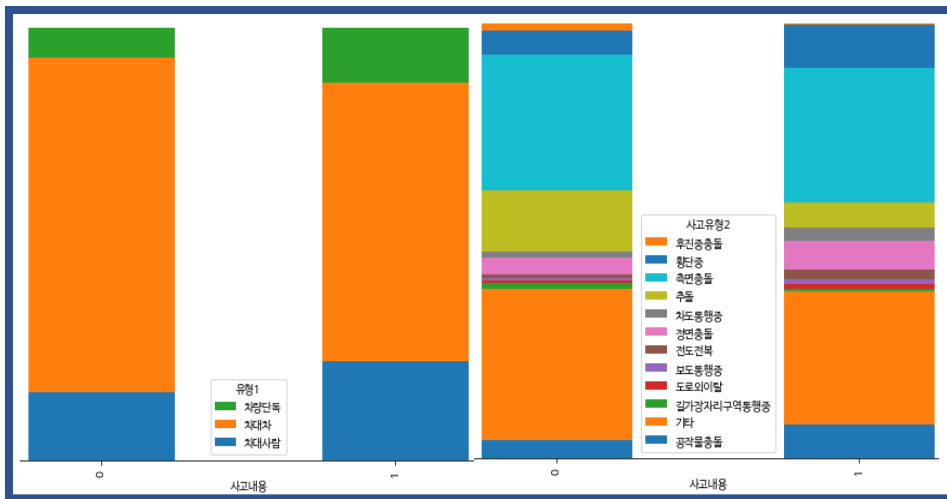
## 2. 결과 해석

### 라. 주요 범주형 변수 시각화 및 해석



#### 1) 사고 발생 지역

- 부상/경상의 경우 제주시가 압도적인 비율을 차지하지만, 중상/사망은 부상/경상에 비해 타 지역들의 비율이 모두 높다.
- 이는 압도적인 교통량과 유동인구가 가벼운 상해의 사고를 야기하는 데에는 주요한 요소이나, 오히려 중상/사망 사고에서는 그 영향의 정도가 적어진다는 사실을 알 수 있다.



#### 2) 사고 유형

- 유형1에서는 차량단독/ 차대사람, 유형2에서는 횡단중/ 정면추돌/ 공작물충돌 일수록 중상/사망 사고에 가까워진다.
- 차량단독&공작물 충돌의 경우 운전자 본인이, 차대사람&횡단중의 경우 보행자가, 정면추돌은 두 경우 모두 발생하는 상황으로서 해당 상황들의 발생을 최대한 억제해야 한다.

## IV. 활용 방안

### 1. 지방자치단체 정책 수립/수정

#### 가. 제주 렌터카 사고의 위험도를 낮출 수 있는 신규 정책 방향 제시

##### 1) 전 범위적 정비 필요

- 렌터카의 중상/사망에 있어서는 제주시와 서귀포시 같은 대량의 유동인구와 교통량을 보유한 지역보다 타 지역의 비율이 더 높아짐을 확인할 수 있었다.
- 교통이 잘 정비되어 있는 변화가에서는 사고의 강도가 낮아짐을 보았을 때, 아직 타 지역은 지자체의 정비가 미치지 않은 곳이 있다고 해석할 수 있다
- 전수 적 차원의 교통정비가 필요하다.

##### 2) 공작물 사고 방지

- 사고 유형에서 차량단독, 차대 사람, 횡단중, 정면충돌, 공작물 충돌의 유형이 사고의 강도가 높아지는 데에 영향을 미치는 것을 확인했다.
- 지자체에서 사전 대응을 할 수 있는 분야는 공작물 충돌로, 기존에 사고가 다수 발생했던 공작물의 경우 철거나 안전하게 지나칠 수 있는 구조물의 추가 설치를 고려할 수 있다.

#### 나. 관광객, 렌터카 업체의 선제적 대응

##### 1) 관광객의 사고 관련 정보 사전과약을 통한 높은 위험도의 사고 방지

- 횡단중 & 정면충돌의 경우 중상과 사망에 직결된다는 사실을 인지하고, 무단횡단과 신호준수에 유의한다
- 기온이 낮고 시정이 좋지 않은 겨울철 눈/비가 오는 날 렌터카 운전을 자제한다.

##### 2) 렌터카 업체의 보험료 책정

- 위험도가 높은 상황에 대비하여 보험료를 상이하게 책정한다.
- 기온이 낮고 시정이 좋지 않을 경우, 풍속이 약할 경우 보험료를 높게 책정한다.

## V. 인턴십 소감

권오민



데이터 분석에 많은 관심이 생겨서 이번 인턴십에 참여하게 되었습니다. 관심만 가지고 있었던 분야에 처음으로 발을 디게 되어서 많은 불안감으로 시작했습니다. 하지만, 강의나 멘토링을 사무국에서 제공해주면서 기초부터 탄탄하게 성장할 수 있도록 많은 도움을 받았습니다. 단계별로 문제를 해결해 나갈 때마다 많은 것들을 배우고 저도 성장을 했던 것 같습니다.

데이터 분석 과정이란 게 어떻게 진행하는지도 몰랐던 저는, 진행 과정부터 어떻게 분석해 나가야 할지 많은 것들을 배운 인턴십이었습니다. 또한 현직자인 멘토님께서 프로젝트 외에도 많은 얘기들을 해주셨습니다. 개발자가 되고 싶은 저에게 너무 유익한 시간이었습니다.

좋은 기회로 참가할 수 있었던 인턴십은 걱정 반, 기대 반으로 시작하였습니다. 데이터 분석은 배워보고 싶었지만 한 번도 도전해보지 않았던 분야라 내가 과연 잘 따라갈 수 있을까 하는 생각이었습니다.

그러나 약 2개월 동안 정말 많은 지식을 배울 수 있었고, 다양한 분들과 함께하며 평소에 잘 들을 수 없었던 좋은 이야기들도 많이 들을 수 있었고, 조원들과도 여러 추억을 쌓을 수 있었습니다. 팀 멘토링과 조 멘토링을 통해 멘토님께 받았던 피드백들은 인턴십에 참가하지 않았더라면 들을 수 없어 더욱 뜻깊었습니다. 또한 제공해주신 이러닝은 기초부터 차근차근 알려주셔서 처음 접하는 분야인데도 불구하고 잘 따라갈 수 있었습니다.

2개월동안 함께했던 멘토님, E2팀 분들, 그리고 3조 조원분들께 다양한 지식과 좋은 추억 남길 수 있게 해주셔서 모두 감사드립니다.

송나흔



김유신



운이 좋게 인턴십에 참여를 하게 되었는데, 평소에 배워보고 싶었지만 장벽이 있는 것 같아 망설여졌던 파이썬, 분석 기법 등을 배울 수 있게 되어 알찬 2개월이 되었습니다!

또한 매주 실행되었던 조, 팀 멘토링 때 우리 조 뿐만 아니라 다른 조의 분석 내용들을 보고 멘토님의 피드백을 받으면서 현업에서 어떤 걸 더 선호하는지, 어떤 흐름으로 가야 도움이 될 수 있을지 와 같이 평소에 궁금증을 가졌지만 해결할 수 없던 여러 의문들을 해소할 수 있게 되어 의미 있는 시간을 보낼 수 있었습니다!

데이터 사이언스 분야로 진출하고 싶은 학생으로서, 체계적으로 여러 사람들과 함께 이 분야를 공부하고 싶다는 갈증을 항상 가지고 있었습니다. 또한 현직자 분들과의 커뮤니케이션 통로가 있었으면 하는 바람도 있었는데, CJ 리모트 인턴십을 통해 제가 바라왔던 것들을 모두 해소할 수 있었습니다.

체계적으로 진행되는 강의와 프로젝트를 통해 기초부터 실전까지 하나하나 다져갈 수 있었고, 꾸준한 멘토링을 통해 즉각적인 피드백과 현업에서의 노하우들을 받을 수 있어서 뜻깊었습니다.

충분히 성장할 수 있는 발판이 되는 시간들이었습니다. 함께한 조원들과 멘토님께 진심으로 감사드립니다!

박관우



이현주



데이터 분석이라는 분야의 문외한이라고 봐도 무방했던 제가 과연 이 인턴십에 참여하는 것이 맞는지, 옳은지에 대한 걱정이 많았습니다. 이런 걱정이 무색할 정도로, 소감문을 쓰는 지금의 저는 참가하길 잘했다는 생각만이 듭니다.

무엇보다도 데이터 분석을 하면서 현직자님의 피드백을 받을 수 있다는 점이 정말 좋았고, 활동 내용을 다른 조원들과 공유하면서 새로운 시각에서 현안을 바라볼 수 있어서 흥미로웠습니다. 또 온라인 강의를 통해 기초지식을 배울 수 있어서 저 같은 비전공자도 따라갈 수 있었던 것 같습니다!

좋은 멘토님과 조원들을 만난 덕분에 성장한 것 같아 기쁩니다. 지난 8주 동안 자그만 행복과 추억을 안겨 주셔서 정말 감사합니다!



감사합니다