

고속도로 교통사고의 사고심각도 영향요인 분석

An Analysis of the Factors Affecting the Accident Severity of Highway Traffic Accidents

윤병조* · 이선민** · 윗위린***

Yoon, Byoung-Jo · Lee, Sun-min · WUT YEE LWIN

요약

본 연구는 2019년부터 2021년의 고속도로 교통사고 위치 좌표를 콘존 데이터와 결합한 후 파이캐럿을 활용하여 고속도로 교통사고 심각도에 영향을 끼치는 요인을 분석할 수 있는 최적 모델을 선정하고 채택된 Random Forest 기법으로 고속도로 교통사고 심각도에 영향을 끼치는 요인을 분석하고자 하였으며, 향후 전국 고속도로 교통사고에 영향을 주는 요인으로 확대하여 분석하고 사고 심각도 개선을 위한 대안 방안 마련이 가능할 것으로 판단된다.

Keywords : 고속도로 교통사고, 교통사고 심각도, Random Forest, 머신러닝

1. 서론

현재 국내의 교통사고 발생 건수는 지속적으로 감소하는 추세이며 이는 정부 및 관련 기관에서의 교통사고 심각성을 무겁게 인식하고 교통안전의 체계적인 계획과 개선대책을 수립하고 시행하는 등의 다각적인 노력을 추진함에 따른 긍정적인 결과로 판단되고 있으나, 우리나라의 도로 교통망의 중요한 축을 이루며 주요 도시를 연결하는 고속도로의 사고 건수 역시도 여전히 증감을 반복하고 있다. OECD 국가별 교통사고 건수를 살펴보았을 때, 우리나라의 교통사고 건수는 OECD 평균인 162.6건과 비교하여 약 2.5배 많은 404.5건이 발생하였고 OECD 국가별 인구 10만 명당 교통사고 건수는 회원국 평균 27.0건으로 나타난 데 비해 우리나라의 수치는 74.2건으로 OECD 회원국 평균보다 약 2.8배 많은 교통사고가 발생하여 여전히 높은 수치를 나타내고 있어 국내 교통사고의 감소를 위한 여러 가지 고민이 필요하다.

2. 본론

본 연구는 2019년부터 2021년까지 3개년의 전국 교통사고 및 고속도로 교통사고 현황 자료를 수집하였으며, 2021년 고속도로 노선을 기준으로 교통사고 자료 중 전체 사고의 71.7%를 차지하고 있는 10개의 노선을 선정하였다. 그림 1은 한국도로공사에서 제공하는 고속도로 안전성능합수 구축 단위인 콘존(Conzone) 데이터이다. 고속도로 교통사고가 발생한 위치 좌표와 콘존(Conzone) 데이터를 결합하였으며, 콘존(Conzone) 데이터를 기반으로 해당 고속도로의 차로 수, 제한속도, 곡선반경 길이 평균 및 종단경사 데이터를 결합하여 고속도로 교통사고 심각도에 영향을 미치는 요인과 상관관계를 분석하는 데 활용하고자 하였다.

머신러닝 워크 플로우를 자동화하는 오픈 소스 라이브러리인 파이캐럿(PyCaret)을 활용하여 교통사고 심각도 분석에 활용할 모형을 선정한 결과는 그림 2 와 같다.

* 정회원 · 인천대학교 도시공학과 교수 bjyoon63@inu.ac.kr

** 정회원 · 인천대학교 도시건설공학과 석사과정 sunminee2@inu.ac.kr

*** 정회원 · 인천대학교 도시건설공학과 석사과정 wutyeesoelwin@gmail.com

콘존ID	콘존길이	기점중점방향을분	시작노드 ID	종료노드 ID	차로 수	노선번호	제한속도	노선구성순번	콘존명	버스전용차로 유무	도로등급구분 코드	
0	0010C2E005	200.0	E	491	4	3.0	10	100.0	1.0	경부고속국도시흥구시IC	0	101
1	0010C2E010	1820.0	E	4	446	3.0	10	100.0	2.0	구서IC-영학IC	0	101
2	0010C2E011	1990.0	E	446	486	3.0	10	100.0	3.0	영학IC-부산TG	0	101
3	0010C2E020	1070.0	E	496	447	3.0	10	100.0	4.0	부산TG-노포	0	101
4	0010C2E030	7780.0	E	447	669	3.0	10	100.0	5.0	노포IC-노포IC	0	101
...
1569	6000C2S050	1770.0	S	668	667	2.0	6000	100.0	5.0	김해가야비-다갈IC	0	101
1570	6000C2S060	9920.0	S	669	668	2.0	6000	100.0	4.0	노포IC-김해가야비	0	101
1571	6000C2S070	1100.0	S	670	669	2.0	6000	100.0	3.0	공평IC-노포IC	0	101
1572	6000C2S080	4240.0	S	671	670	2.0	6000	100.0	2.0	기장힐마리C-공평IC	0	101
1573	6000C2S090	5580.0	S	445	671	2.0	6000	100.0	1.0	기장IC-기장힐마리C	0	101

그림 1. 콘존 마스터 예시

Model	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT (Sec)
rf	4.2897	48.0194	6.8763	0.1920	0.6173	0.8582	0.3540
lightgbm	4.6313	50.3225	7.0455	0.1550	0.6405	0.8836	1.1670
gbr	4.8078	54.4204	7.3206	0.0877	0.6482	0.9353	0.3050
lr	4.8574	56.0309	7.4268	0.0664	0.6539	0.9435	0.0350
ridge	4.8745	56.2573	7.4445	0.0606	0.6575	0.9434	0.0230
omp	4.8978	56.8055	7.4809	0.0512	0.6584	0.9548	0.0220
en	4.9447	57.6769	7.5261	0.0451	0.6622	0.9763	0.0200
lilar	4.9679	58.0690	7.5513	0.0388	0.6652	0.9828	0.0240
lasso	4.9679	58.0690	7.5513	0.0388	0.6652	0.9828	0.0200
et	4.0918	59.1592	7.6512	-0.0415	0.6492	0.8151	0.2870
dummy	5.0160	60.6834	7.7168	-0.0027	0.6744	0.9918	0.0140
huber	4.4289	61.2037	7.7490	-0.0110	0.6196	0.6706	0.0670
knm	5.0148	61.1660	7.7745	-0.0331	0.6837	0.9399	0.1060
dt	4.6504	76.7264	8.7606	-0.3475	0.7247	0.8651	0.0320
par	7.3138	112.9857	10.5648	-0.9403	0.9535	1.4034	0.0310
ada	11.6414	163.6196	12.7696	-1.9254	1.1689	3.0033	0.2000
lr	141.3988	29167374.7640	1714.4466	-366584.5136	0.6923	10.5615	0.0250
lar	304.2158	132844808.1455	3676.3185	-1669712.2987	0.8933	23.4213	0.0310

그림 2. 파이캐럿을 활용한 모형선정 결과

파이 캐럿을 활용한 결과 RandomForestRegressor 모형이 선정되었고, RandomForest 최적화 과정을 통한 하이퍼 파라미터 도출 결과는 다음 표 1과 같이 나타났으며 해당 하이퍼 파라미터를 통한 Random Forest의 최종 성능은 7.4397로 나타났다. 또한, 아래 그림 3은 최적 파라미터를 적용하여 모형을 개발한 후 SHAP 패키지를 활용한 모형 변수 중요도를 시각화한 것이다.

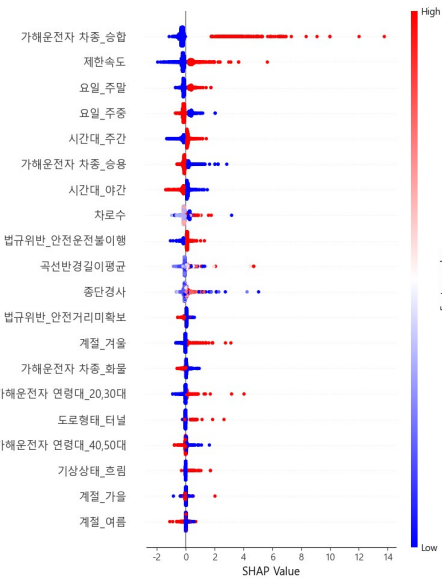


그림 3. SHAP 요약도

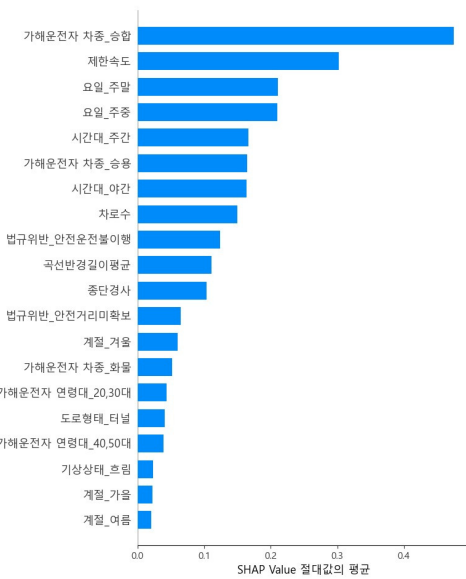


그림 4. SHAP Value 변수중요도

3. 결론

본 연구는 전국의 고속도로 교통사고 정보, 교통사고 위치 좌표를 콘존 데이터와 결합한 후 파이캐럿을 활용하여 교통사고 심각도에 영향을 끼치는 요인을 분석할 수 있는 최적 모델을 선정하였고 Random Forest 기법이 채택되었다. 2021년 개통이 완료된 고속도로 노선 중 교통사고 건수가 가장 많은 노선은 경부선, 영동선, 서해안선, 남해선, 수도권제1순환선, 중부선, 중부내륙선, 중앙선, 호남선, 서울양양선 순이며, 가해운전자 차종이 승합차일 때 고속도로 사고 심각도에 가장 큰 영향을 주는 것으로 나타났고 고속도로 제한속도가 높을수록 사고 심각도가 높은 것으로 나타났다. 분석 결과는 향후 전국 고속도로 교통사고에 영향을 주는 요인으로 확대하여 분석하고 사고 심각도 개선을 위한 대안 방안 마련이 가능할 것으로 판단된다.

참고문헌

권철우 (2021) XGBoost를 활용한 이륜자동차 교통사고 심각도 비교분석, 한국ITS학회논문지, 20(4), 1-12
 안상준 (2023) 딥 러닝을 이용한 자연어 처리 입문
 김경환, 박병호 (2011) 사고유형에 따른 교통사고 심각도 모형 개발, 한국안전학회지, 26(6), 118-123