



광역버스 DTG를 활용한 경부고속 도로 운전 위험구간 예측

2017년 공공빅데이터분석사업

광역버스 DTG를 활용한 경부고속도로 운전 위험구간 예측

- I 모델의 개요
- II 분석 데이터
- III 분석 내용 및 절차
- IV 분석결과
- V 향후 활용계획

위험운전 발생 가능성이 높은 고속도로 구간을 예측



9월 초에 경부고속도로에서 고속버스가 승용차를 들이받아 40대 부부가 숨진 안타까운 사고가 있었습니다.

7월에는 광역버스 기사가 졸음운전을 하다가 차선을 넘어 앞서가던 차량을 추돌해서 50대 부부가 숨지기도 하였습니다.

고속도로에서의 버스 사고가 대형사고로 이어져 소중한 생명과 재산을 잃는 사고가 빈번하게 발생하고 있습니다.



고속도로에서의 교통사고로 소중한 인명과 재산 피해 증가

디지털 운행기록 3천9백만건, VDS 지점정보 1백40만건 및 추가 데이터 확보

No	데이터 명 (제공기관)	데이터 설명(항목)	데이터 범위	데이터 건수	데이터 크기
1	디지털 운행기록계 (교통안전공단)	경부고속도로를 경유하는 광역버스 등의 운행기록 데이터	2017년 1월	39,111,575	4.7 Gb
2	VDS 설치정보 (한국도로공사)	VDS_ID, VDS존유형구분명, 노선구성순번, 기점종점방향 구분코드 등	전국	7,466	1 Mb
3	VDS 지점교통량 (한국도로공사)	집계일자, 집계시분, VDS_ID, 차로유형구분코드, 교통량	2017년 1월	1,360,248	72 Mb
4	VDS 지점통행속도 (한국도로공사)	집계일자, 집계시분, VDS_ID, 차로유형구분코드, 평균속 도	2017년 1월	1,360,248	95 Mb
5	교통사고 (경찰청)	교통사고의 발생시각, 발생지점, 사망자수, 부상자수, 사고 차량종류, 법규위반내용 등	2012 ~ 2016년	4,744,645	2 Gb
6	도로중심선 (한국도로공사)	노선번호, 도로명, 이정, X좌표값, Y좌표값, GRS80X좌표 값, GRS80Y좌표값 등	경부고속도로	41,808	3 Mb
7	AWS 날씨데이터 (기상청)	지점, 일시, 기온(°C), 풍향(deg), 풍속(m/s), 강수량(mm)	경부고속도로	15,261	82 Kb
8	콘존 (한국도로공사)	IC와 IC, IC와 JC 사이의 구간, 본선영업소와 IC나 JC 사이 의 구간에 대한 기준 데이터	전국	1,371	84 Kb

고속도로 위험운전 행동별 위험구간 예측

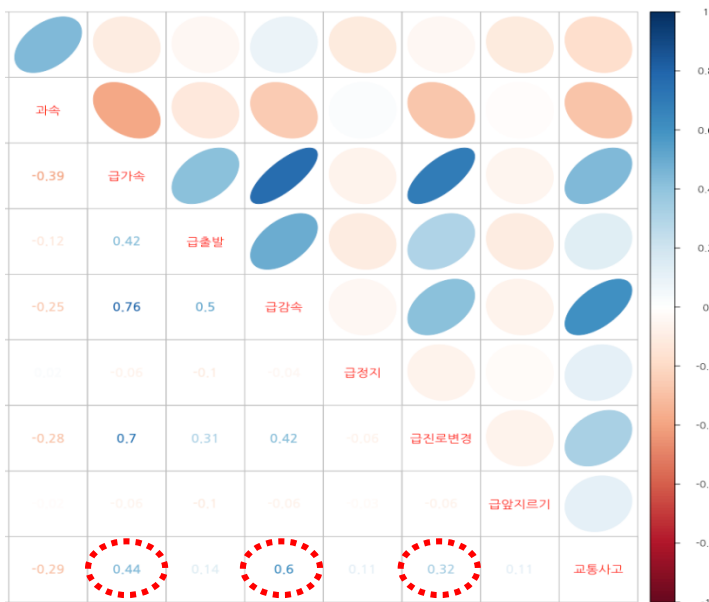
□ 분석 목적

- 교통안전공단에서 제공하는 11대 위험운전행동 기준으로 위험운전행동이 많이 발생하는 경부고속도로 구간 예측
- 위험운전행동이 많이 발생하는 구간은 교통사고 발생 가능성이 높음

□ 분석 방향

- 교통사고와 위험운전행동 간 상관관계 도출
- 교통사고와 상관성이 높은 위험운전행동을 활용하여 사고 위험구간 예측

➡ 교통사고는 위험운전행동 중 급감속, 급가속, 급진로변경과 상관성이 높음



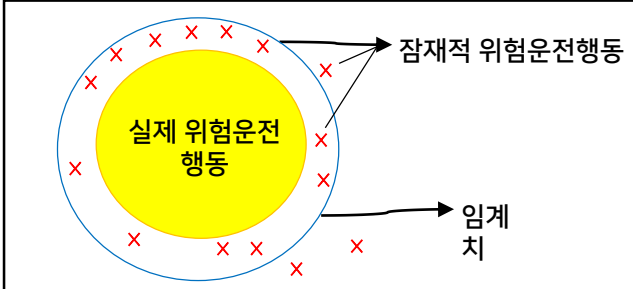
위험운전행동 범위		
구분	버스 기준	
과속 위험구간	과속	도로 제한속도보다 20km/h 초과 운행 시
	장기과속	도로 제한속도보다 20km/h 초과해서 3분 운행시
급가속 위험구간	급가속	6.0km/h 이상 속도에서 초당 6km/h 이상 가속 운행하는 경우
	급출발	5.0km/h 이하 속도에서 초당 8km/h 이상 가속 운행하는 경우
급감속 위험구간	급감속	초당 9km/h 이상 감속하여 속도가 6.0km/h 이상인 경우
	급정지	초당 9km/h 이상 감속하여 속도가 5.0km/h 이하인 경우
급차로변경 위험구간	급진로변경	속도가 30km/h 이상에서 진행방향이 좌/우측 8°sec 이상으로 차로 변경하고, 5초 동안 누적각도±2°/sec 이하, 가감속이 초당 ±2km/h 이하인 경우
	급앞지르기	속도가 30km/h 이상에서 진행방향이 좌/우측 8°sec 이상으로 차로 변경하고, 5초 동안 누적각도±2°/sec 이하, 가감속이 초당 3km/h 이상인 경우

고속도로 위험운전 행동별 위험구간 예측

□ 머신러닝(SVM)을 활용하여 위험운전행동 예측

- 실제 위험운전행동을 활용하여 위험운전행동을 예측
- 예측된 위험운전행동 중 실제 위험운전행동은 아니지만, 학습을 통해 위험운전에 가깝게 운전한 경우를 추출
- 해당 데이터들은 실제 위험운전행동의 DTG 데이터로 판단되지 않지만, 기계학습을 통해 도출된 위험운전에 준하는 행동임

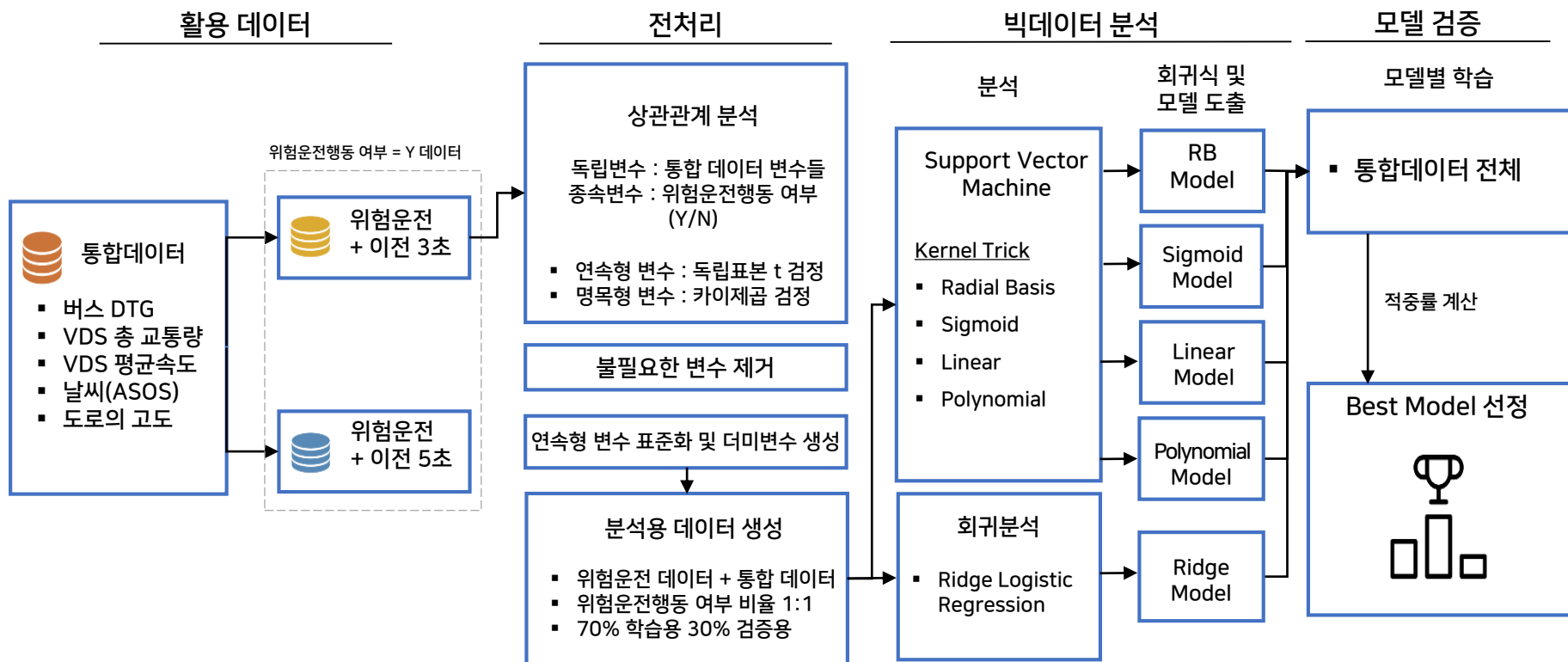
➔ 실제 위험운전행동과 예측 위험운전행동을 모두 고려하여 보다 정교한 운전 위험구간 도출 가능

단계	알고리즘	비고									
1	DTG 데이터를 활용, 실제 발생한 위험운전행동 도출	<div>전체 DTG 데이터</div> <div></div>									
2	도출된 위험운전행동 DTG 데이터 개수와 동일하게 정상 DTG 데이터랜덤 샘플링										
3	1, 2단계에서 생성된 데이터를 병합하여 학습용 데이터셋 생성										
4	학습용 데이터셋 및 SVM 알고리즘을 활용하여 위험운전행동 분류모델 도출										
5	실제 발생한 위험운전행동은 아니나, SVM 모델을 통해 위험운전행동에 가깝다고 판단된 데이터 추출(잠재적 위험운전행동)	<div>예측위험운전행동</div> <table><tr><th></th><th>YES</th><th>NO</th></tr><tr><th>YES</th><td>실제 위험운전행동</td><td>실제 위험운전행동</td></tr><tr><th>NO</th><td>잠재적 위험운전행동</td><td>비위험운전행동</td></tr></table>		YES	NO	YES	실제 위험운전행동	실제 위험운전행동	NO	잠재적 위험운전행동	비위험운전행동
	YES		NO								
YES	실제 위험운전행동		실제 위험운전행동								
NO	잠재적 위험운전행동	비위험운전행동									
6	잠재적 위험운전행동 판단 임계치는 시범 시행 과정을 거쳐 조정 <i>예) 데이터 분류 시의 일반적인 임계치:0.5(확률:50%)를 0.8, 0.9 등으로 상향조정</i>										
7	각 구역별 실제 위험운전행동 및 잠재적 위험운전행동 빈도 도출										

고속도로 위험운전 행동별 위험구간 예측

□ Support Vector Machine(SVM) 모델 활용

- 현재 위험운전행동기준으로는 실제 위험운전으로 판단되지 않지만, 기계학습을 통해 위험운전에 준하는 행동 예측
- 예측 정확도 상승 및 이전 상황을 고려하기 위해 각 DTG 데이터마다 이전 3초/5초간 속도·RPM·브레이크 데이터 병합
- 다양한 방식의 회귀분석 및 SVM을 시도하여 최적의 예측모델 도출



고속도로 위험운전 행동별 위험구간 예측

□ 예측 결과

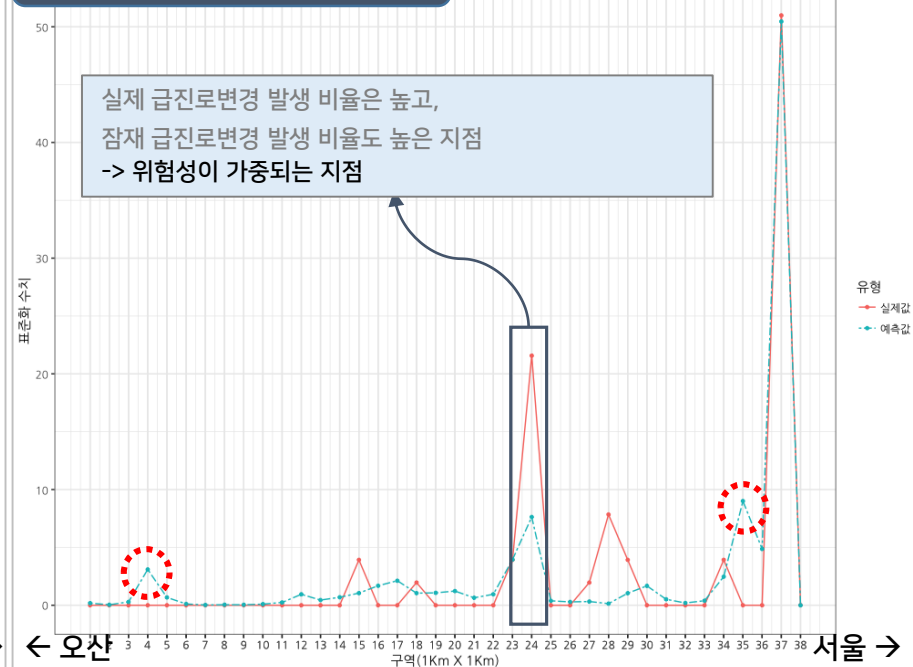
- 실제 위험운전행동을 활용해 해당 위험운전행동이 많이 발생하는 위험구간 도출
- 예측 위험운전행동을 활용해 실제 위험운전행동만으로는 위험성이 높지 않았지만 예측 위험운전행동까지 고려하면 위험성이 높아지는 구간 도출
- 실제 위험운전행동 및 예측 위험운전행동을 모두 활용하여 위험구간 정교화

✓ 급진로변경 예측 결과

상행선 잠재적 급진로변경 예측



하행선 잠재적 급진로변경 예측



고속도로 이상운전 위험구간 예측

□ 분석 목적

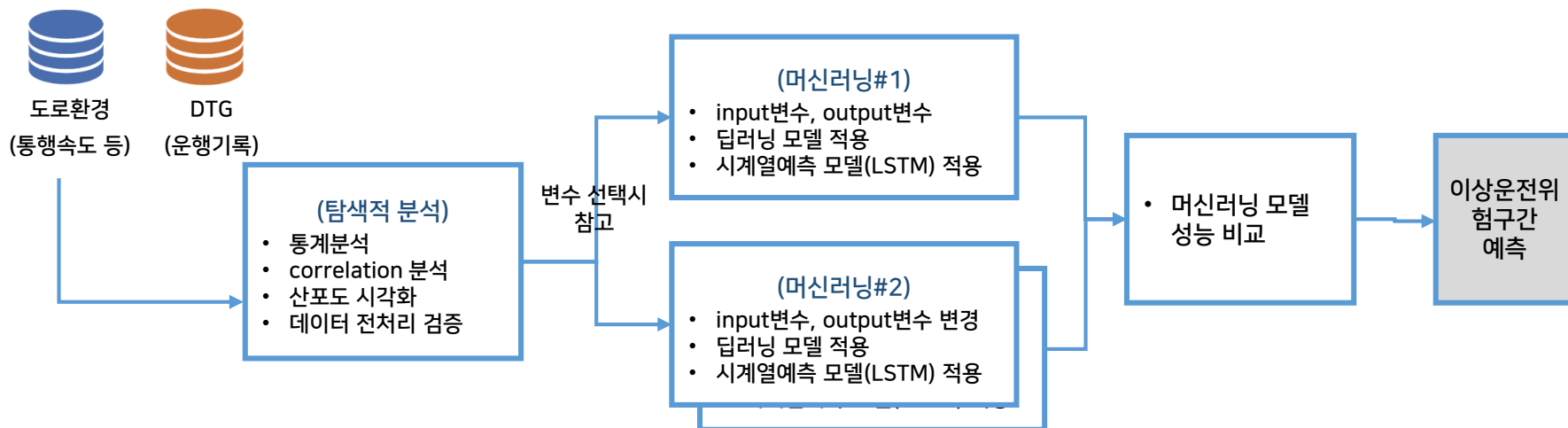
- 시계열 형식인 DTG 데이터의 특징을 활용
 - LSTM 머신러닝을 통해 평소 운전패턴을 학습시킨 후, 평소와는 다른 이상운전패턴 탐색
- ※ LSTM(Long Short-Term Memory) : 시계열 데이터와 같은 시간의 흐름에 따라 변화하는 데이터를 예측하기 위한 딥러닝 모델

□ 분석 방향

- 기존 운전패턴과 머신러닝을 통해 도출된 예측 운전패턴을 비교
- 두 패턴 간 차이값을 계산하여, 차이값이 클수록 평소 운행패턴을 벗어난 것이라고 판단

이상운전 패턴 도출을 위한 가설

- 운전 중 졸음이 발생하면 평상시와 다른 이상운전 패턴을 보일 것이다.
- 평상시와 다른 운전 패턴은 속도, RPM, 브레이크 조작에서 나타난다.
- 평상시와 다른 운전 패턴을 보이는 구간이 많은 곳은 졸음운전이 의심된다.



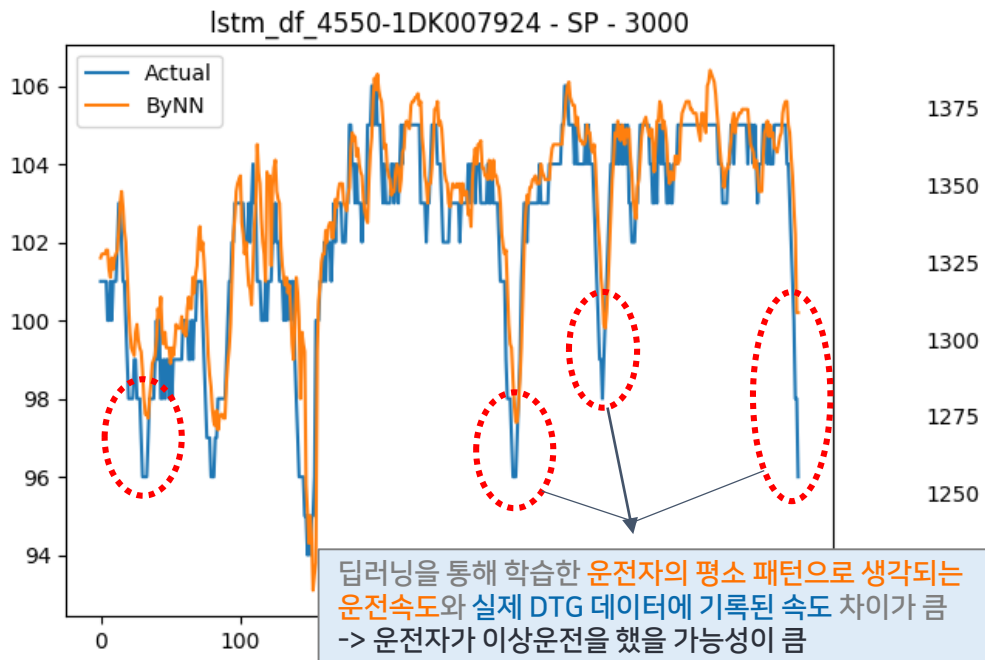
고속도로 이상운전 위험구간 예측

□ LSTM 딥러닝을 통한 운전 행태 학습 결과

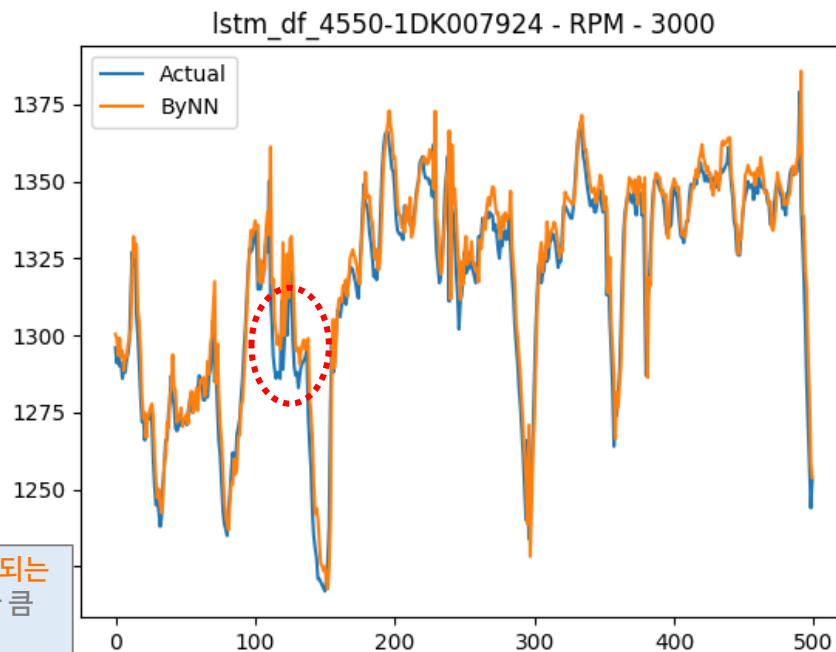
- 딥러닝을 통한 속도예측은 오차율이 상대적으로 작으며, RPM 예측은 상대적으로 높음
- 속도는 변화가 부드러우나, RPM은 외부환경, 운전습관, 브레이크 조작 등에 의한 변화가 큰 경향을 보임
- 이상운전패턴이 다발생하는 구간이 존재함

□ 실제운행데이터와 예측운행데이터 간 속도 차이 비교

속도에 대한 NN출력값과 실제운행기록값 비교

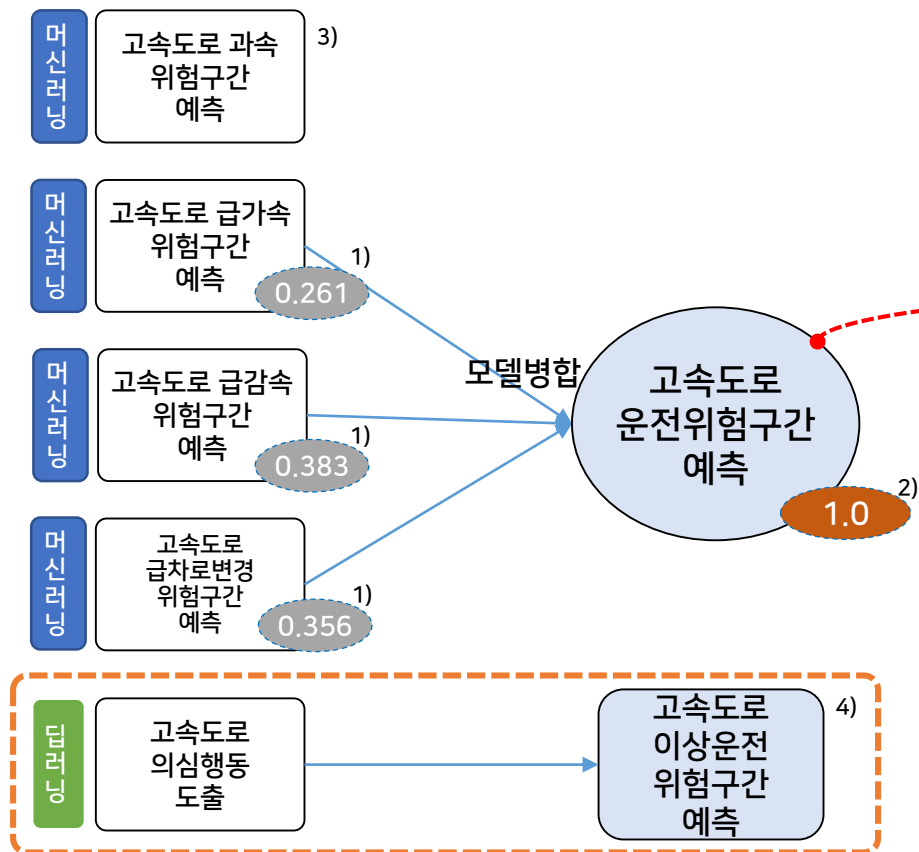


속도에 대한 NN출력값과 실제운행기록값 비교



운전위험구간 예측 알고리즘

통계분석 결과와 선행 연구자료를 참조하여 요소 모델별 가중치를 부여하였으며, 모델들을 병합하여 최종 운전위험구간 예측 산출식을 도출



운전위험구간 예측 산출식

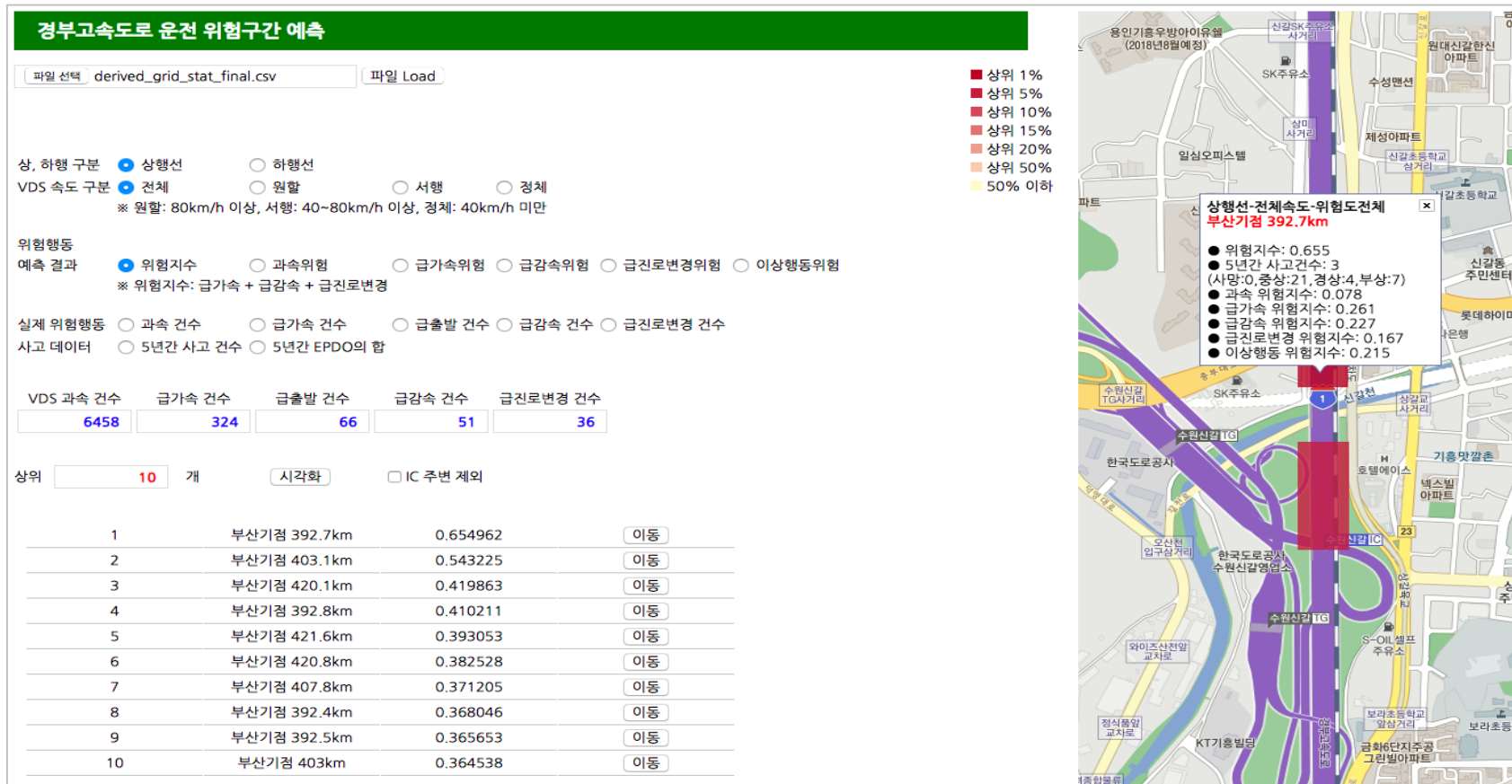
$$\begin{aligned}
 Y \text{ (위험지수)} &= \text{격자내 급가속 위험지수} * 0.261 \\
 &+ \text{격자내 급감속 위험지수} * 0.383 \\
 &+ \text{격자내 급차로변경 위험지수} * 0.356
 \end{aligned}$$

- 주 1) 사업용차량 통합단말 표준플랫폼 및 안전운전지원기술 개발 최종보고서(교통안전공단. 2017.06)에서 위험운전행동 가중치를 각각 급가속 0.18, 급감속 0.264, 급진로변경 0.245로 산출한 것을 3개의 합이 1이 되도록 가중치 조절함
- 주 2) $1.0 = 0.261 + 0.383 + 0.356$
- 주 3) 과속은 교통사고와의 상관관계가 낮으므로 예측 알고리즘에서 제외
- 주 4) 이상운전과 사고와의 관련성은 향후 많은 연구가 필요한 부분
이므로 이상운전을 운전위험구간 예측 알고리즘에서 제외

운전위험구간 예측 결과

□ 운전 위험구간 시각화

- 다음맵 API를 활용하여 100m X 100m 구간별 위험지수 시각화
- 도로방향 및 VDS 속도 구분별 세부 시각화
- 격자 선택시, 해당 격자의 통계정보 확인 가능



운전 위험구간 예측 결과

□ 운전 위험구간 예측 결과

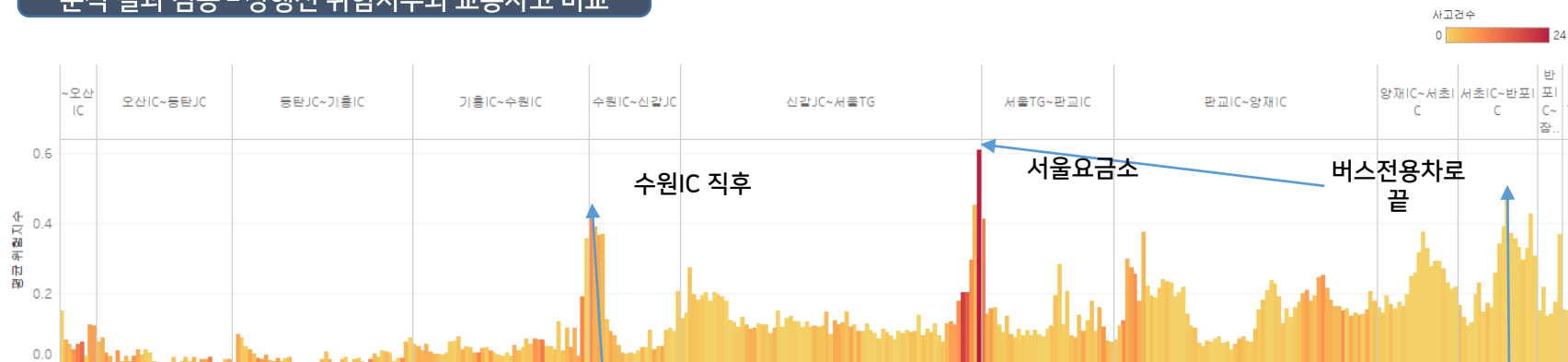
상행선 운전 위험구간				하행선 운전 위험구간		
순위	위험 지수	상세 내역	지도	위험 지수	상세 내역	지도
1	0.610	부산기점 403.1km ● 위험지수: 0.610 ● 5년간 사고건수: 24 (사망:4, 중상:124, 경상:22, 부상:116) ● 과속 위험지수: 0.003 ● 급가속 위험지수: 0.092 ● 급감속 위험지수: 0.372 ● 급진로변경 위험지수: 0.147 ● 이상행동 위험지수: 0.673		0.518	부산기점 421.6km ● 위험지수: 0.518 ● 5년간 사고건수: 0 (사망:0, 중상:0, 경상:0, 부상:0) ● 과속 위험지수: 0.000 ● 급가속 위험지수: 0.261 ● 급감속 위험지수: 0.040 ● 급진로변경 위험지수: 0.217 ● 이상행동 위험지수: 0.823	
2	0.498	부산기점 420.1km ● 위험지수: 0.498 ● 5년간 사고건수: 0 (사망:0, 중상:0, 경상:0, 부상:0) ● 과속 위험지수: 0.043 ● 급가속 위험지수: 0.206 ● 급감속 위험지수: 0.251 ● 급진로변경 위험지수: 0.041 ● 이상행동 위험지수: 0.798		0.463	부산기점 391.9km ● 위험지수: 0.463 ● 5년간 사고건수: 0 (사망:0, 중상:0, 경상:0, 부상:0) ● 과속 위험지수: 0.027 ● 급가속 위험지수: 0.015 ● 급감속 위험지수: 0.282 ● 급진로변경 위험지수: 0.167 ● 이상행동 위험지수: 0.050	
3	0.452	부산기점 403km ● 위험지수: 0.452 ● 5년간 사고건수: 3 (사망:1, 중상:13, 경상:4, 부상:9) ● 과속 위험지수: 0.005 ● 급가속 위험지수: 0.091 ● 급감속 위험지수: 0.261 ● 급진로변경 위험지수: 0.101 ● 이상행동 위험지수: 0.494		0.414	부산기점 403.2km ● 위험지수: 0.414 ● 5년간 사고건수: 7 (사망:1, 중상:30, 경상:7, 부상:22) ● 과속 위험지수: 0.000 ● 급가속 위험지수: 0.019 ● 급감속 위험지수: 0.116 ● 급진로변경 위험지수: 0.279 ● 이상행동 위험지수: 0.363	

운전 위험구간 예측 결과

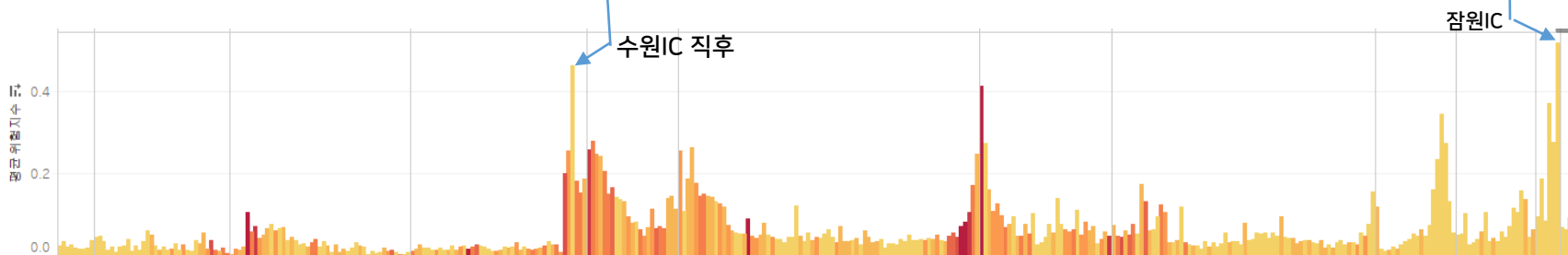
□ 운전 위험구간 예측 결과 검증

- 5년간 경부고속도로에서 발생한 교통사고와 운전 위험구간 비교검증
- 예측 운전 위험구간과 실제 교통사고 간 비교는 데이터 건수, 기간, 대상 차량 등 많은 측면에서 한계점이 존재함
- 상/하행선에서 5년간 교통사고가 가장 많이 발생한 서울요금소 부근의 교통사고와 위험지수를 비교해보면 위험구간을 충분히 예측하고 있음을 알 수 있음

분석 결과 검증 - 상행선 위험지수와 교통사고 비교



분석 결과 검증 - 하행선 위험지수와 교통사고 비교



운전위험구간 예측 결과

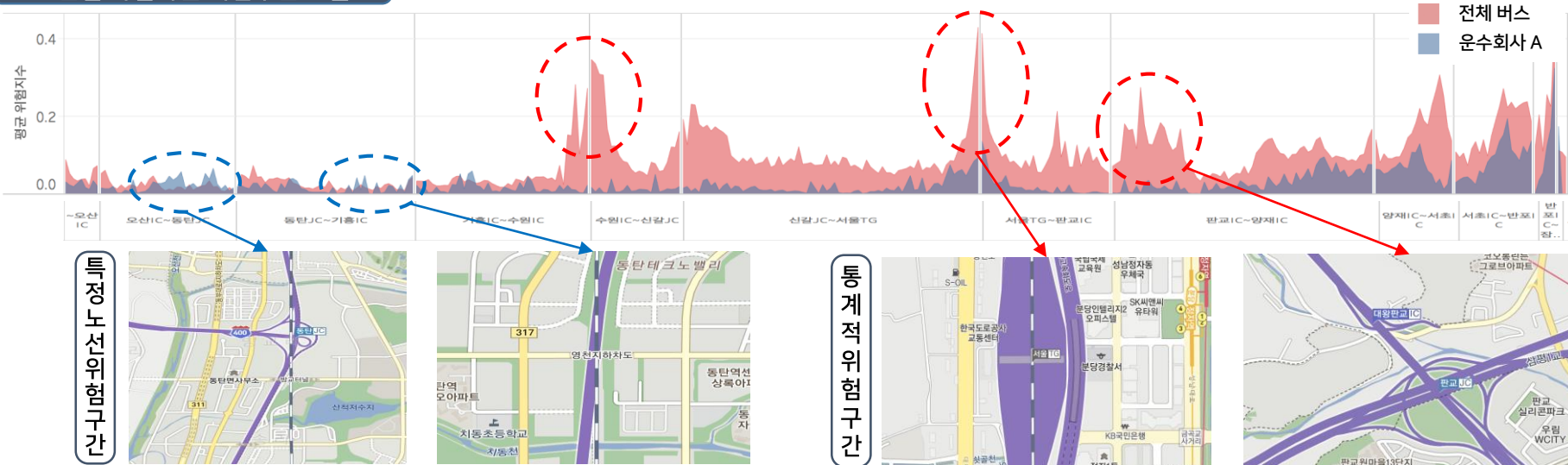
□ 통계적 위험구간과 운수회사별 위험구간 비교분석

- 비교분석을 통해 통계적으로 위험한 구간 뿐만 아니라 운수회사의 노선별 차별화된 위험구간까지 도출하여, **보다 효과적인 운수 종사자 교육 및 교통 안전관리**가 가능함

활용 데이터 특징

	전체 버스	운수회사 A
활용 데이터	모든 회사, 모든 노선의 차량 289대 DTG 데이터	운수회사 코드가 XX939 인 차량 15대
데이터 건수	39,111,575건	2,180,106건
기간	2017년 1월	
운행구간	오산IC ~ 한남IC 사이	오산IC 출발 - 한남IC 도착 OR 한남IC 출발 - 오산IC 도착

노선별 차별화된 위험구간 도출



향후 시스템 적용 및 활용방안

실시간·전국적인 교통사고 예방이 가능하며, 추가적인 교육 및 정책 수립 시에도 효과적인 활용

