기획

목적

차량 수 대비 사고율을 회귀하는 것의 의의 : 일반적이지 않은 변수가 있음. -> 구체적 정의 필요

문제인식

문제정의

솔루션제시

데이터수집

어떤 데이터셋 활용

차량 수 회귀 -> 택시데이터

지역별 사고 -> TASS

관심속성은 무엇인가?

택시데이터 -> 시간, 위경도, 주행속도

TASS -> 시간, 위경도, 사고유형, 기상상태(고려)

결측값, 이상원소 -> 전처리계획

택시데이터 -> 주행속도 0인거 제외

빵구 처리

택시데이터의 위경도가 너무 다양함 -> 군집화 과정으로 어느정도 그룹으로 묶어야 함 -> TASS의 위경도가 어느 군집에 속했는지를 판단하는 쪽으로 모델 훈련 계획

모델개발

회귀모델

시간대별, (군집화된위도)지역 별 차량 대수 대비 사고율 회귀모델

* 1. 지역 별 차량 대수를 주행속도로 회귀해야 함
* 2. 차량 대수 -> 주행속도 / 전체 주행 평균속도로 차량많음대수 지표
* 3. 사고위험지표 -> 사고건수(지표) / 차량대수지표
* 4. 사고건수(지표) -> 사고의 가중치 -> 이것도 차량대수처럼 희귀도로 가중치를 놓는쪽으로 생각

평가 / 시각화 전략

데이터셋 3단계 분할(Traning, Validate, Test)

교차검증으로 결과 일관성 확인 -> 회귀 모델의 검증 -> 눈으로 볼수있는 의미는 지도에 색깔 비슷하게 칠해지는지

시각화 전략 -> 지도 매핑

지도에 그린다..

* 핵심 내용
* 발표할때 강조해야할 내용들(긴글)
* 시각화 자료

출처 있으면 명시하고 자료

만약 직접 만들어야 하면

피피티 제작하면 폰트랑 사용 RGB