

프로젝트 기획서

과정명	퍼블릭 클라우드 서비스를 활용한 파이썬 기반 AI-ML 인재 양성 과정
팀명	드랍 더 비트
팀원	윤상하(조장), 박성훈, 김동규, 정선우

프로젝트명	서울시 범죄데이터 분석	기간	2024.01.08 ~ 2024.02.07
프로젝트 목표	서울특별시의 5대 범죄 발생률과 음주운전범죄를 바탕으로 각 범죄유형 및 서울시 행정동 별 범죄율에 대해 분석, 서울의 지역별 사회적 특성과 범죄 발생 간의 상관관계 확인 범죄 예방을 위해 시행된 제도 및 법령의 효과를 확인하고, 실효성을 검증, 예측해, 자체 안전 지표 설립		
시나리오 내용			
<p>지난해 발표된 경찰청 체감안전도 조사는 2022년 국민 체감안전도 조사 결과'에 따르면 작년 국민의 범죄 체감안전도는 전년보다 0.4점 내린 83점이었습니다. 이 중 여성 상대 범죄는 81.4점, 강도·살인은 84.6점으로 각각 0.3점, 0.6점 내리는 등 전 분야에서 체감안전도가 낮아졌습니다. 이어 지난해 8월을 기점으로 이상동기범죄를 시작으로 범행 예고·협박글이 줄을 이었고 전문가들은 시민들이 재난 수준의 불안을 느낄 것이라고 진단했습니다. 이러한 결과를 보고 우리 사회 치안에 대한 의문을 가지기 시작했고, 가장 인구 밀도가 높은 서울 특정지역 범죄 데이터를 분석하기로 결정했습니다.</p> <p>가설</p> <p>1. 지역별 가설 : 서울시의 특정 지역에서는 다른 지역에 비해 범죄 발생률이 높을 것으로 가정</p> <p>2. 인구밀도 별 가설 : 거주인구보다 생활인구가 범죄율과 더 상관관계가 있을것으로 가정</p> <p>3. 범죄 유형 간 상관 가설: 특정 범죄 유형이 다른 범죄 유형과 상관 관계가 있을 것으로 가정하고, 범죄 간의 유의미한 패턴 확인</p>			

구상도

1. 데이터 수집

경찰청, 도로교통공단, 공공데이터포털, 서울 열린데이터광장 에서 서울시 5대 범죄 발생현황, 서울시 구별 인구밀도, 생활인구 밀도, 음주운전 추이와 같은 범죄 데이터와 사회 특성 데이터를 수집한다.

2. 데이터 전처리

결측치와 이상치를 제거해주고, 다양한 데이터 셋을 복합적으로 사용하므로, 동일한 특성을 가지는 열은 칼럼명을 변경해준다.

3. 데이터 모델링

전처리된 데이터를 바탕으로 다양한 데이터 셋을 조합해 시나리오에서 세운 가설에 대해 확인해보고, 앞으로의 범죄율, 음주운전 다발 구역 지역 예측을 진행한다.

4. 시각화

Seaborn, Tableau, QGIS 등을 활용해 분석하고 예측한 데이터들을 시각화 해준다.

서비스 및 기능

사용자의 실시간 위치 정보를 활용하여 주변 반경 내 공공기관, 안전알리미벨, 방범용 CCTV, 24시 영업 편의점과 같은 안전시설과 단란주점, 유흥업소등과 같은 위험시설을 식별하고, 사용자에게 안내하는 서비스를 통해 사용자의 경로 설정과 안전에 대한 불안도를 낮출 수 있을 것으로 가정합니다. 이를 통해 안전한 경로 선택이 가능하고, 자체적인 안전지표를 통해 사용자의 안전도를 시각적으로 제공하여 사용자가 인근의 안전인프라에 대해 효과적으로 파악할 수 있게 될 것입니다.