|  |  |  |  |
| --- | --- | --- | --- |
| **2021 혁신성장 빅데이터 분석 프로젝트 요약서** | | | |
| **팀 명** | ADAM | **팀 장** | 박소연 |
| **팀 원** | 김기영, 박소연, 송현정, 이동렬 | | |
| **주 제 명** | 광양시 재활용품 자동수거기 설치 | | |

|  |  |  |
| --- | --- | --- |
| 구분 | 프로젝트 상세내용 | |
| 프로젝트명 | 재활용품 자동수거기 설치를 위한 최적위치 선정 | |
| 구현 목표 | 주택단지별 특성을 도출해내어 재활용품 자동수거기 설치에 있어서 가장 적합한 위치를 선정하고자 함 | |
| 구현 내용 | - 최종모델링을 위한 데이터프레임을 만들기 위해 Feature값 생성 및 결측치 보간- 모델링을 통한 회귀식 산출 후, 점수화를 통해 재활용품 자동수거기 최적위치 선정 | |
| - X값(Feature) 선정을 위한 데이터 전처리  - 결측치 보간을 위한 모델링 시도(RandomForestRegressor,  DecisionTreeRegressor, ExtraTreesRegressor, LinearRegression, Ridge,  SGDRegressor, BayesianRdige, LogisticRegression, ElasticNet)  - 최종 회귀식 산출을 위한 모델링 과정(Linear Regression, Ridge)  - 최종 회귀식에 필요한 Weight값, Bias값 도출 후, 각 항목별로 점수화 진행  - 재활용품 자동수거기 최적위치 산정 | |
| 상세 설명 | **1. 연구 과제**  - 본 분석은 환경부 폐기물 발생 및 처리현황, 인문사회 데이터(인구증가율, 인구 분포도, 생활수준 등) 및 제공 데이터를 활용하여 각 지역별 재활용 쓰레기 발생량을 예측하고 예측된 발생량을 활용하여 광양시 공동주택 내 재활용품 자동수거기 50개소의 설치 위치 도출을 목적으로한다.  **2. 데이터 설명 및 모델링**  **2-1) 사용 데이터**  광양시\_공동주택\_분리수거장현황.csv  광양시\_재활용품분리수거\_장려금지급내역.csv  광양시\_금호동\_집적장별\_재할용품통계.csv  광양시\_건축물\_표제부.csv  광양시\_건축물\_전유부.csv'  광양시\_법정경계(읍면동).geojson  광양시\_도로명주소(건물).geojson  광양시\_도로명주소(건물출입구).geojson  광양시\_지적도.geojson  광양시\_인구정보(총인구).geojson  광양시\_동별인구수\_세대수.csv  광양시\_장래인구\_추계치.csv  광양시\_폐기물발생\_및\_처리현황(2017~2018).xlsx  광양시\_폐기물발생\_및\_처리현황(2019).xlsx  광양시\_건물정보.geojson  **2-2) 전처리 후 데이터 통합**  - 장려금지급내역, 표제부, 동별인구수 데이터를 하나의 DataFrame으로 통합  - 아파트 단지별로 세대수를 합하여 그룹화  - 동별 세대당인구수 데이터를 활용하여 인구 추정을 진행  - 최종 모델링에 필요한 X값 도출  **2-3) 결측치 보간**  - 장려금총액, 주요품목배출량, 기타품목배출량 3개의 값을 예측  - Ridge(solver = 'sparse\_cg', alpha = 0.0001) 모델 채택  **2-4) 회귀식 산출**  - 건물별 동주소, 세대당인구수, 추정거주인구, 기타품목배출량, 장려금총액 column을 통해 주요품목배출량을 예측하는 모델링을 구성  - LinearRegression 모델 채택  **2-5) 점수화**   * 각 Feature별 도출된 Weight값을 곱해 Points 산출 | **3. 결과 분석**  - 산출된 Point를 통해 수거기 입지에 적합한 50개소 위치 도출  - “50개소 중 12개소는 금호동 내에 설치되어야합니다” 라는 분석 요건에 따라 금호동 내의 아파트 단지들의 특성들을 반영해 금호동내에 12개소 도출  **4. 활용방안 및 기대효과**  - 본 데이터 분석을 통해 입지추천을 함으로써 수거기가 효율적으로 운영될것으로 기대된다. 이를 통해 인공지능 수거기의 기대효과(교육 효과, 수거효율 증대, 인력 감소, 환경 보전, 경제적 효율성)을 극대화 할 수 있을것이다.  - 재활용품 배출량, 특히 주요 배출품목인 패트, 플라스틱, 유리 배출량이 높은 곳에 설치하였기 때문에 자동수거기가 광양시의 생활폐기물 처리 부담을 경감시켜 줄 것이다.  **5. 프로젝트 보완점**  - 회귀식 산정시, Feature수 부족 : 데이터 탐색 기간을 늘려 외부 데이터수집을 통해 회귀모델에 반영되는 특성을 늘렸다면 좀 더 정교한 회귀예측이 가능했을 것이다.  - 모델의 예측 성능을 판별하는 score값이 낮았던 점 : 딥러닝을 시도해 배출량 예측의 정확도를 높일 수 있었을 것이다.  - 지적도나 아파트출입구, 도로데이터 등을 활용하지 못했다.  **6. 분석도구**  - Python  - 분석 모델의 종류: Linear Regression, Rdige |