



2025학년도 1학기 SW 캡스톤디자인 경진대회

부안군에 대한 온라인 감성 분석을 통한 장기 정주 유도 전략

오픈소스 URL : https://github.com/locotive/stay_buan

팀 명 머무름
지도교수 김윤경

팀 원 남재민(컴퓨터공학부,4), 경태현(컴퓨터공학부,4), 신철언(컴퓨터공학부,4)
산업체 한국국토정보공사

개발 동기 및 목적



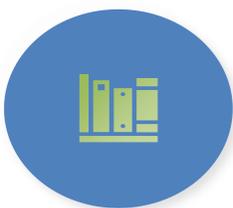
지방 소멸 위기

전라북도 부안군은 2024년 기준 주민등록인구 5만 명 이하로 감소하였으며, 65세 이상 고령층 비율은 37%를 초과하여 전라북도 평균 24%보다 현저히 높은 수준입니다.



데이터 기반 접근 필요성

온라인에서 형성되는 감성 데이터를 실시간으로 분석하고, 이를 정책 결정 과정에 반영하는 데이터 기반 접근 방식이 요구됩니다.



관광 산업 정체

변산반도 국립공원, 채석강, 내소사 등 다양한 관광 자원을 보유하고 있음에도 불구하고, 연간 관광객 수는 2019년 700만 명에서 2023년 500만 명 이하로 감소했습니다.

주요 기술

구분	주요 내용
데이터 크롤링	네이버 블로그/뉴스/카페, 유튜브, 구글, 부안군청 등 6개 이상 플랫폼 자동 수집 Selenium + BeautifulSoup, 키워드 기반 AND 검색, 중복/날짜 필터링)
감성 분석 모델	KoBERT, KcELECTRA-base-v2022, KcELECTRA, KcBERT-large, KoSentenceBERT 5개 모델 앙상블 문장 단위 Positive/Negative/Neutral 분류 및 Confidence Score 산출
시각화 대시보드	Streamlit 기반 웹 대시보드 감성 분포, 키워드/플랫폼/시계열별 통계 실시간 시각화 Docker 배포 지원 Ngrok 배포 지원
데이터 전처리	불용어 제거, 중복/광고/스팸 필터링, 텍스트 정제 및 토큰화, 키워드 추출 등 데이터 품질 향상 처리
Git 버전 관리/협업	Git 및 GitHub를 활용한 코드 버전 관리, 이슈 트래킹, 팀원 간 협업 및 오픈소스 공개

개발 내용

단계	단계
데이터 수집	다양한 플랫폼(네이버 블로그/뉴스/카페, 유튜브, 구글, 부안군청 등)에서 키워드 기반 자동 크롤링 및 데이터 축적
데이터 전처리	불용어/중복/광고/스팸 필터링, 텍스트 정제, 키워드 추출 등 데이터 품질 향상 작업
감성 분석	5개 딥러닝 모델(앙상블)로 문장별 긍정/부정/중립 분류 및 신뢰도 산출, 댓글 등 다양한 텍스트 분석
시각화 대시보드	Streamlit 기반 웹 대시보드 구현, 감성 분포·키워드·시계열·플랫폼별 통계 실시간 시각화, Docker 배포 지원, ngrok 배포 지원
협업 및 관리	GitHub 기반 버전 관리, 코드 공유 및 팀원 협업, 오픈소스 공개

결과 및 분석

단계	주요 내용
수집 데이터	1,000,000건 이상의 온라인 텍스트(후기, 댓글 등) 수집 및 분석
주요 키워드	부안, 채석강, 변산반도, 내소사, 숙소, 교통, 맛집 등
감성 분석 결과	긍정 감정은 전체적으로 감소, 부정 감정은 상대적으로 증가 추세
불만 요인	숙소 부족, 교통 불편, 편의시설 미흡 등 주요 불만 키워드 도출
시계열 분석	기간별 긍정/부정 감성 변화 시각화, 특정 시기 불만 급증 등 트렌드 파악 가능
정주 관련 인사이트	실제 후기를 기반으로 일자리 부족, 정주 매력 미흡 등 정주 관련 문제점 일부 포착
활용 및 기대효과	정책 수립, 관광 인프라 개선, 민간 마케팅, 지역 브랜드 관리 등 다양한 분야에 데이터 활용 가능