

주간프로젝트 기획안

기획안 작성일자 : 2024년 5월 20일

조 명	4조 : E1i4
조 원	조장: 김지우 / 부조장 : 김성원 / 조원: 이가희, 성수린, 박태영
프로젝트 주제 및 개요	<p>주제 : 방문자 리뷰 유사도 분석을 통한 서울 관광지 추천 시스템</p> <p>개요 :</p> <ul style="list-style-type: none"> - 배경 <ul style="list-style-type: none"> ● 관광객들은 각기 다른 취향과 필요를 가지고 여행지를 방문 ● 다양한 요구를 충족시키기 위해 맞춤형 관광지 추천 시스템이 필요 ● 사용자가 제공하는 키워드(예: 청결, 주차, 가격)과 원하는 관광지를 기반으로 가장 적합한 관광지를 추천할 수 있는 시스템이 유용할 것이라 판단 - 목적 <ul style="list-style-type: none"> ● 개인의 선호도에 맞는 관광지를 추천함으로써 여행 경험을 향상시키는 것을 목표 ● 사용자들은 시간과 노력을 절약하고, 만족스러운 여행 경험을 함 - 타겟 <ul style="list-style-type: none"> ● 여행지를 처음 방문하는 국내 관광객 ● 특정 요구를 가지고 있는 관광객 ● 짧은 시간 내에 맞춤형 관광지 정보를 얻고자 하는 사람들 <p>기대효과</p> <ul style="list-style-type: none"> ● 효율적인 여행 계획 ● 사용자의 만족도 향상
프로젝트 수행 방향 및 내용 (수행 방법/도구)	<ul style="list-style-type: none"> ● 해결하고자하는 문제, 최종 산출물의 청사진 <p>1. 데이터 수집</p> <ul style="list-style-type: none"> - 한국 관광 공사_구석구석 웹크롤링 https://korean.visitkorea.or.kr/list/travelinfo.do?service=ms - 네이버 맵 크롤링을 통한 리뷰 키워드와 참여자 수 수집 https://map.naver.com/p?c=15.00,0,0,0,dh - 네이버 API를 통한 위도, 경도 데이터 수집 - 네이버 API를 통한 이미지 데이터 수집 https://openapi.naver.com/v1/search/..

2. 텍스트 전처리

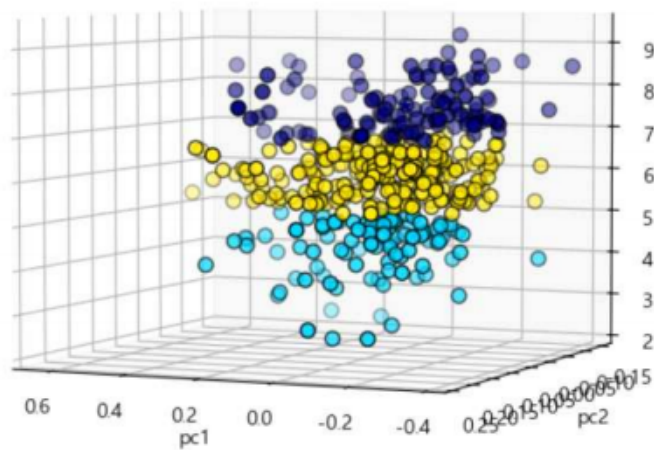
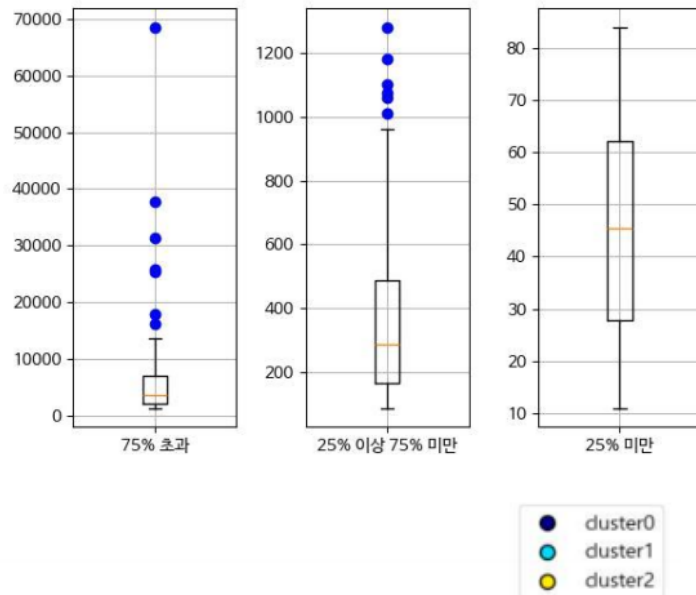
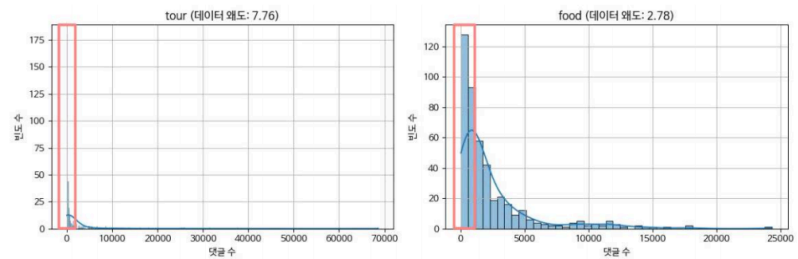
- 불용어 처리, 기호 등을 제거
- 문장 및 단어 토큰화

3. 키워드 추출 및 분류

- 키워드 정의, 매핑 및 빈도 계산
- 속성에 따른 분류

4. 시각화

- Matplotlib 사용하여 boxplot, barchart 등 시각화



	<p>5. 모델링</p> <ul style="list-style-type: none"> - 테이블 레이블링, 특징추출(TF-IDF, LDA등)을 통한 단어의 가중치 부여 <p>5. 추천 시스템 구현</p> <ul style="list-style-type: none"> - Streamlit 활용한 추천 시스템 구현 (사용자 정보 선택 → 유사도 계산 → 관광지 추천) <p>● 수행도구</p> <ul style="list-style-type: none"> - 분석도구 : VScode(python), Matplotlib, Konlpy, Sklearn, gensim - 일정관리 : JIRA - 코드공유 : Google Drive
<p>프로젝트 조직 (구성원 및 역할)</p>	<p>● 역할분담</p> <p>김지우: (조장) 기획안 작성, 일정관리, 데이터 분석(자료수집/ 전처리/분석/시각화), 발표자료 제작, 발표</p> <p>김성원: (부조장) 자료 취합 및 검토, 데이터 분석(자료수집/ 전처리/분석/시각화), 발표자료 제작, 발표</p> <p>이가희: 데이터 분석(자료수집/ 전처리/분석/시각화), 발표자료 제작, 발표</p> <p>성수란: 데이터 분석(자료수집/ 전처리/분석/시각화), 발표자료 제작, 발표</p> <p>박태영: 데이터 분석(자료수집/ 전처리/분석/시각화), 발표자료 제작, 발표</p>
<p>프로젝트 추진 일정</p>	<p>● 일정 (5/10~6/12)</p> <p>5/10 ~16: 프로젝트 기획</p> <ol style="list-style-type: none"> 1. 주제 선정 (5/10~5/13) 2. 추진 일정 계획 (5/14) 3. 주제 관련 자료 수집 (~5/16) 4. 기획안 작성 및 피드백 (~5/16) <p>5/17 ~ 6/7: 프로젝트 수행</p> <ol style="list-style-type: none"> 1. 데이터 수집 (5/17~5/18) 2. 데이터 이해 (5/18) 3. 데이터 전처리(5/18~6/7) 4. 데이터 시각화 및 대시보드 화면(6/8 ~) 5. 제품 개발 (6/8 ~) <p>6/7 ~ 6/12: 프로젝트 발표</p> <ol style="list-style-type: none"> 1. 발표 자료 취합 및 구성 (6/7~6/9) 2. 발표 대본 작성 (6/9~)

