05.17

빈도수랑 평점간 상관분석

선크림의 데이터를 가져와 경쟁사들과의 특징을 비교하여

크로스 조인

결론이 많을 수도 있지 않을까

리뷰 데이터가 있다면 사용자 이용데이터를 못구하는 경우가 있었고,

이용데이터가 있다면 리뷰데이터가 없는 경우가 많다..

사용자 데이터가 있는 데이터

여행 데이터를 가지고 여행객의 사용자 데이터를 분석해서,

-> 장소가 그룹으로 묶이면 각 그룹의 특성 파악, 특성을 토대로 추천?

줌부

잘 분류되지 않아도

우리 나라에 오는 외국인 데이터를 보고

영어권, 일어권에서 원하는 관광지와 내국인이 원하는 관광지가 다를 수도 있지

관광지가 가진 특성에 따라 분류하여 최적의 여행지 코스

컨텐츠 기반 필터링

-> 숫자로 나오는 형태가 있어야 함..

주제 참고

관광쪽 논문에

평점데이터 분석,

해외 외국인 관광객이 자연 ? 어디에 ???  
인기 관광지를 알기 위해

평점과 댓글의 유사도를 파악

cf 방식?

몇번 이상 나오면 5점

청결의 기준 점수가 있어야 하는데

tf - idf??

추천시스템

평점을 주는 야놀자를 가지고 오는거

내 생각이 아니라 통계자료가 있어야함.

몇번 언급된거 -? 텍스트 마이닝

정형데이터 분석이 가능해야함!!

다음주에는 저 특성데이터를 기준에 맞게, 수치로 변화하는 모습을 변경해야함

cf

lda

예제를 만들어 사이트 비교

요인분석…………키워드 중심………………..

통계를 정규분포를 그려서 …….

외국인 관광지 데이터 수치화를 이용하여 순위로 한다음에 내국인

결론 -> 인기 많은 관광지를 추천

신라호텔 9.7 (청결, 시설 만족도와의 상관관계)

관광 관련 데이터가 많다는 점

관광지에 따른 인프라 분석

여행자들이 가지는 특성?

* 여행지를 나누기

-> 각 여행지의 특성을 분석하고 분류

* 2030을 관객 분석 특성

-> 소비패턴 등을 분석