



경로 기반 맛집 추천 서비스를 위한 맛집크롤링, 맵크롤링

김미정, 김준성, 이정려, 최재철



목차

1. 목적
2. 데이터 수집처 및 수집 방법
3. 프로젝트 진행 순서 및 구조
4. 현재 진행상황
5. 향후 진행과제



데이터 수집처, 방법

1. 맛집 평점

- a. 네이버 - json
- b. 블루리본서베이 - json, html
- c. 망고플레이트 - json
- d. 다이닝코드 - selenium
- e. 메뉴판닷컴 - scrapy

2. 경로

- a. 자동차경로 (네이버 플랫폼)
- b. 대중교통 OPEN API (<https://lab.odsay.com/>)

프로젝트 진행 순서 및 구조

1. 맛집 크롤링 및 업데이트



NAVER

MANGO
PLATE

DININGCODE



매일매일맛있게
Menupan

< 사이트별 평점에 가중치 선정 종합평점 생성안 >

1안				
플랫폼명	가중치(만점)	계산식	null값 처리	비고
망고플레이트	2	평점(100점 만점) * 0.02	Q2	월간이용자(MAU) 200만명
다이닝코드	2	평점(5점 만점) * 0.4	Q2	월간이용자(MAU) 180만명
블루리본 서버이	3	기본점수 1.5 + (0.5 * 리본수)	0	
네이버 리뷰&평점	3	평점(5점만점)*0.6 + 리뷰수.MinMaxScaler	Q2	
메뉴판닷컴	0.2	0.2점 보너스 점수(10점 초과시 10점 처리)	0	
2안				
플랫폼명	가중치(만점)	계산식	null값 처리	비고
망고플레이트	2	평점.MinMaxScaler * 2	Q2	월간이용자(MAU) 200만명
다이닝코드	2	평점.MinMaxScaler * 2	Q2	월간이용자(MAU) 180만명
블루리본 서버이	3	기본점수 1.5 + (0.5 * 리본수)	0	
네이버 리뷰&평점	3	평점.MinMaxScaler * 3	Q2	
메뉴판닷컴	0.2	0.2점 보너스 점수(10점 초과시 10점 처리)	0	

append + 중복값 삭제

최종 db mongodb 생성

프로젝트 진행 순서 및 구조

2. 경로내 waypoint 설정

지하철 경로

버스 경로

자동차 경로

경로 내 waypoint 및 precision 설정





프로젝트 진행 순서 및 구조

3. 1+2

경로에서 **waypoint** 설정
반경 500~1km 이내 선택 메뉴 관련 맛집 표출 (서비스 구현)



현재 진행상황

- 총 4개 맛집 평점 사이트 크롤링, -> 2만여개의 데이터 수집 및 정제
- 크롤링한 4개의 DataFrame을 통일된 format으로 db로 구축 완료

향후 진행과제

- 주기적 업데이트를 위한 서버 및 db 구축
- 경로 정보에 맛집 정보를 입히는 기능 구현 - Haversine package를 이용하여 위치별 반경설정
- visualization on map
- 웹인터페이스 구현