

공공 데이터를 활용한 장애인 식당 추천 서비스 구현

| 요약

- 본 연구는 식당 이용에 불편 요소를 지닌 장애인에게 편의성을 제공하려는 목적으로, **장애인 맞춤형 추천 시스템 구현**을 통해 실수요자에 맞는 서비스를 제공 하고자 함
- 편의 시설물 정보 제공 위주였던 기존 연구와 달리 본 연구에서는 **장애인들의 개인 통행 특성**을 반영한 맞춤형 추천 서비스를 제공함에 따라 차별성을 지님
- 장애인 맞춤형 추천 서비스를 구현하기 위해 **특별교통수단 데이터**의 장애 유형별 목적지, 호출지 등을 수집하였으며, 이를 ‘대구푸드’ 웹 사이트의 대구광역시 **식당 정보 데이터**와 결합하였음
- 추천시스템은 **콘텐츠 기반**과 **협업 필터링** 2가지 방법으로 분석하여 실시하였음
- 콘텐츠 기반은 식당정보 데이터를 **3개의 어절(N-gram)**로 분류하여 식당 간 **코사인 유사도**를 계산하였고, 협업 필터링은 특별교통수단 사용자별 식당방문 데이터를 이용하여 코사인 유사도, **KNN 알고리즘**을 적용하여 추천시스템을 구현함
- 분석 결과, 콘텐츠 기반의 경우 **입력한 식당과 유사한 식당이 추천**되었으며, 협업 필터링의 경우 사용자가 **이용한 식당과 통행패턴이 유사한 식당**이 추천되는 것을 확인할 수 있음
- 이용자 편의를 위해 시각화 프로그램인 Power-BI를 이용하여 **웹 형태로 추천시스템**을 구현하였으며, 이를 통해 장애인 식당 이용에 있어 고도화된 맞춤형 경험을 제공할 것으로 판단됨

주제어 : 식당 추천, 추천 시스템, 코사인 유사도, 콘텐츠 기반, 협업 필터링

1) 대구디지털산업진흥원 빅데이터활용센터 센터장, 053-655-7050, aut7767@dip.or.kr
2) 영남대학교 통계학과 연구인턴
3) 영남대학교 정보통신공학과 연구인턴
4) 영남대학교 경영학과 연구인턴
5) 경북대학교 통계학과 연구인턴
6) 경북대학교 컴퓨터학부 연구인턴

1 개 요

I 연구 필요성 및 목적

- 일반인들의 경우 개인 맞춤형 추천 서비스 활성화로 다양한 분야에서 고도화된 서비스를 제공받아 편리한 생활을 지님
- 하지만 장애인들의 경우 일반인들의 비해 상대적으로 최신 기술들을 접목한 서비스 형태가 부족하고, 정부의 장애인들의 사회적 참여 확대 정책으로 수요는 지속적으로 증가하고 있음
- 이에 동등한 기술적 혜택, 장애인 통행수요 충족, 복지정책 부합 등을 고려하여 장애인들을 위한 맞춤형 서비스 확대가 필요할 것으로 판단됨
- 이에 본 연구에서는 장애인들을 위한 식당 추천 서비스를 빅데이터, 머신러닝 기술을 통해 구현함으로써 장애인들의 식당 선택에 있어 고도화된 사용자 경험을 제공하고자 함
- 장애인 협회 담당자, 장애인들의 인터뷰를 실시한 결과 식당 선택에 있어 이동 편의시설, 보도의 높이 등과 같은 물리적 환경도 중요하지만, 식당의 친절, 맛, 마음의 장벽이 없는 식당을 선호하는 것으로 나타남
- 인터뷰 결과를 통해 현재 장애인들이 주로 방문하는 식당의 경우 해당 변수가 높게 반영되어 있음을 가정하고, 개인 맞춤형 식당 추천 서비스를 구현하고자 함



〈그림 1〉 장애인협회 담당자 인터뷰(좌), 장애인 인터뷰(우)

2 선행연구 고찰

I 선행연구 고찰 및 분석 사례

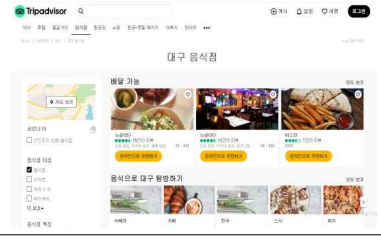
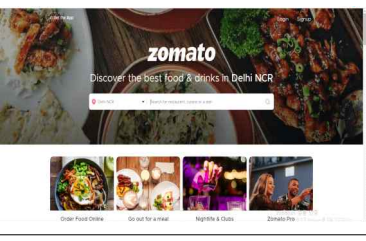

- 장애인 정보제공 관련 플랫폼 서비스를 살펴보면 <표 1>와 같이 나타남
- 대다수의 선행 연구(플랫폼)들이 편의시설, 충전소, 관광지 정보 등을 제공하고 있으며, 장애인 또는 교통약자들을 대상으로 하는 식당 추천 서비스는 부재한 상황임
- 해당 플랫폼 서비스는 장애인, 교통약자들을 대상으로 편의시설 정보를 수집, 종합하여 정보를 제공 한다는 장점은 있지만, 데이터 업데이트 및 정확도 측면에서의 한계점이 있음

구 분	대한민국 구석구석	장애in 제주	휠 누리	플랫(FLAT)
내용				
	관광약자를 대상으로 장애인 편의시설을 보유하는 관광지 정보 제공	장애인을 대상으로 제주 여행 시 장애인 편의시설 위치와 일상생활 정보 제공	교통약자(휠체어장애인) 대상으로 전국 전동보장구 충전소 정보 제공	교통약자 (장애인, 임산부) 대상으로 전국 무장애 편의시설물 정보 제공
한계	생활권 정보 미비 및 관광지 정보 제공에 국한	제주도 편의시설물에 대한 위치 정보만 기재	전동보장구 충전소 시설물 정보에 국한	편의시설물 위주의 정보를 제공해 장애인 맞춤형 정보 미비

<표 1> 장애인 관련 플랫폼 서비스

- 민간 분야에서 여행, 외식업을 영위하는 기업들 중 식당 추천 알고리즘을 비즈니스에 접목한 사례는 <표 2>와 같이 나타남
- 대다수의 기업에서 협업 필터링을 통한 개인 맞춤형 추천 서비스를 구현하고 있으며, SK 플래닛에서 운영하고 있는 Syrup Table의 경우 빅데이터 분석(블로그, 소셜 네트워크, 리뷰 등)을 통한 지역별 다양한 테마기능을 추출하여 식당을 추천하는 알고리즘을 구현함
- 또한, 외식 및 마케팅 분야의 많은 스타트업들이 개인 맞춤형 서비스 구현으로 신사업 발굴 및 비지

니스 모델을 개발 중에 있으며, 이러한 추천 서비스 개발의 활성화는 빅데이터 분석을 함에 있어 필수적인 요소로 자리매김함

구분	Trip-advisor	Zomato	Syrup Table
내용			
알고리즘	협업 필터링	협업 필터링	소셜 네트워크 및 자체 알고리즘

〈표 2〉 추천 시스템 사례

II 연구 차별성

- 기존 공공분야 서비스 플랫폼들을 살펴보면 대다수 장애인, 식당, 편의시설 등의 정보를 종합하여 제공하는데 집중하고 있으며, 이는 개인 단위의 데이터 수집의 어려움으로 인한 것으로 판단됨
- 또한, 민간 분야의 추천 서비스는 활성화 되고 있지만, 장기간 대량의 데이터를 수집해야하는 어려움으로 대규모 자본과 일부 기업에 의해서만 추천 서비스를 제공하고 있는 있음
- 전통적인 식당 추천 서비스는 방문자들의 리뷰, 유명 인플루언서들의 의견 등을 통해 식당 추천이 대다수이지만, 본 연구에서는 장애인들이 주로 이동수단으로 활용되고 있는 특별교통수단 배차이력 데이터를 활용하여 개인 맞춤형 서비스를 제공하는데 있어 차별성이 있음
- 또한, 정부에서의 공공데이터 개방 활성화 정책의 높은 정량적 성과에도 불구하고 서비스화 된 사례와 시민들의 눈높이를 맞추기에는 아직 부족한 실정이며, 이에 본 연구가 공공데이터 개방과 활성화라는 측면에서도 부합할 것으로 판단됨
- 따라서 본 연구에서는 특별교통수단(나드리콜) 배차이력, 대구 푸드 식당정보 등의 공공데이터를 활용하여 분석을 실시하였으며, 장애인을 위한 식당 추천 서비스를 제공하는데 있는 기존 연구와 비니스 모델 측면에서 차별성을 가짐

I 특별교통수단 배차이력 자료

- 장애인 식당 추천 서비스를 구현하기 위해 장애인들이 이동수단으로 활용하고 있는 특별교통수단인 나드리콜 배차이력 데이터를 활용하였으며, 자료의 구성요소는 <표 3>와 같음
- 자료의 시간적 범위는 2018년 12월부터 2020년 4월까지이며, 장애인들의 통행이 반복적이고 정기적임을 감안할 때 해당기간 이외의 자료를 추가로 포함하여도 분석결과에 큰 차이가 없을 것으로 판단함
- 공간적 범위는 대구광역시로 한정하고 특별교통수단을 이용하는 계층은 장애인, 거동이 불편한 고령자 등으로 구성됨

구분	분석 자료
데이터	1,503,679 통행 (2018.18 ~ 2020.04)
이용대상	1) 3급 이상 장애 등급으로 장애 정도가 심해 대중교통 이용이 어려운 자 2) 국가유공자 및 고령자 중 대중교통 이용이 어렵다는 진단을 받은 자
내용	이용객 ID, 날짜, 목적지, 호출장소, 장애 유형, 위도, 경도 등

<표 3> 특별교통수단 배차이력 자료의 구성 요소

- 기초 자료중 목적지, 호출장소 데이터의 경우 특별교통수단 콜센터에서 접수 시 이름, 장애유형, 목적지, 시간 등을 자연어 기반으로 담당자가 입력을 하게 되는데, 이를 데이터 전처리를 통해 장애유형, 목적지를 데이터를 추출함
- 또한, 전체 통행 데이터 1,503,679건 중에서 식당 방문과 관련된 데이터는 <표 4>와 같이 대구푸드 자료에 속해있는 식당을 추출하였으며, 이를 통해 7,153건 데이터 셋을 구축함

구분	내용
데이터	7,153 통행 (2018.18 ~ 2020.04)
장애유형	지체, 뇌병변, 신장, 시각, 노약자, 지적, 정신, 청각, 자폐, 유공자, 언어, 호흡기, 심장, 기타, 뇌전증, 요루 순
범위	대구푸드 웹사이트에서 선정된 대구광역시 식당 950여개

<표 4> 전처리된 데이터의 구성 요소

- 개인 맞춤형 추천 서비스를 구현하기 위해서 개인별 방문 목적지를 매트릭스로 변환함
- 따라서 최종 구축된 데이터셋은 <그림 2>과 같이 날짜, 장소, 위도, 경도, 장애유형 name_code가 나타남

	날짜	장소	호출장소	목적지	DB Index	call_y	call_x	target_y	target_x	장애유형	name_code	구분
0	2019-12-08	후포회수산	통덕 후포회수산(순용법 지제1급 기사메모:'17:37')	수성시장네거리 더좋은병원	7967041	35.843996	128.593767	35.851152	128.620837	지체	11374	1
1	2019-07-09	황성한우천년우정	두산 황성한우천년우정->두류(정현석 뇌병변3급 기사메모:'04:06')	두류네거리	7298593	35.836655	128.619362	35.854878	128.559655	뇌병변	19913	1
2	2019-07-08	황태성	신]새동산병원(구.365병원)(황태성 지제4급 기사메모:'07:00')	서부연합외과	7293942	35.869247	128.574751	35.834661	128.559192	지체	23415	1
3	2019-07-08	황태성	수동활]서부연합외과(황태성 지제4급 기사메모:'16:30')	신]새동산병원(구.365병원]	7297488	35.834661	128.559192	35.869247	128.574751	지체	23415	1
4	2019-07-02	황장군	평리동 황장군(황희 신장2급 기사메모:'04:54')	남산 열린내과의원	7268188	35.862260	128.557691	35.861864	128.595229	신장	23449	1
...
7147	2019-12-17	가마솔에누른밥	종구 덕곡동 564[가마솔에누른밥] >> 용산(한재성 뇌병변1급 기사메모:'13:50')	용산동 달구벌복지관	8006613	35.980872	128.622938	35.850290	128.527712	뇌병변	12534	2
7148	2019-07-11	가리산삼계탕	본리 가리산삼계탕(서진수 지제3급 기사메모:'15:06')	월성주공207동	7311580	35.834835	128.538680	35.826309	128.532324	지체	10785	2
7149	2019-07-23	가리산삼계탕	본리]달서구 와룡로 61 가리산삼계탕(김정기 뇌병변3급 기사메모:'19:57')	월성주공303동	7365082	35.834835	128.538747	35.829094	128.532442	뇌병변	5114	2
7150	2019-11-14	가리산삼계탕	본리 가리산삼계탕주차장(김정기 뇌병변3급 기사메모:'17:07')	월성주공303동	7863673	35.834835	128.538680	35.829094	128.532442	뇌병변	5114	2
7151	2019-12-05	8번식당	8번식당 (침산점)(하태훈 뇌병변3급 기사메모:'15:37')	안심주공3단지아파트314동	7955507	35.881194	128.587221	35.862559	128.706483	뇌병변	22410	2

<그림 2> 나드리콜 전처리 데이터

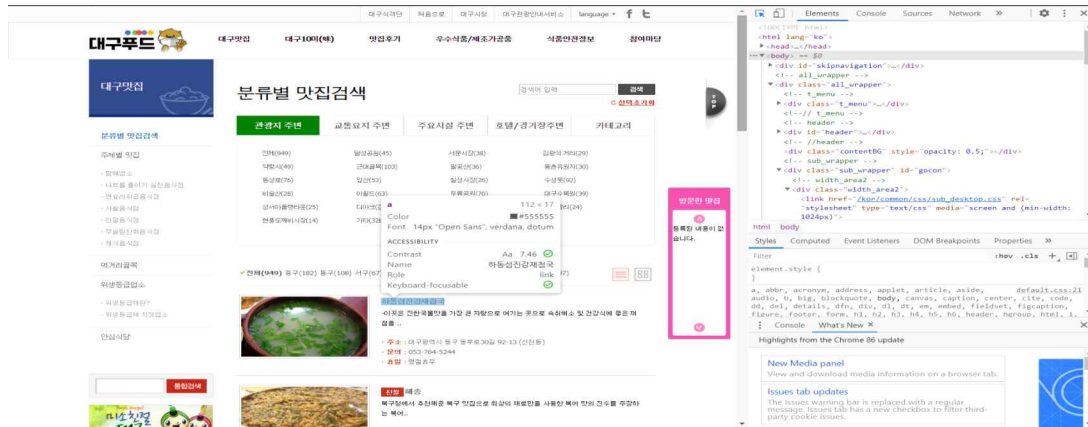
II 대구 푸드 자료

- 대구시 맛집(식당) 소개 정보는 ‘대구 푸드’ 사이트에서 제공하고 있으며, 식당 메뉴와 소개 정보를 기반으로 유사한 식당을 추천하기 위해 해당 데이터를 활용함
- 특별교통수단 회원들이 방문한 장소와 ‘대구 푸드’ 에서 제공하고 있는 식당들을 결합하여 전처리를 진행함
- 대구푸드에서 제공하는 식당정보는 <표 5>와 같이 나타나며, 식당이름, 상세소개, 주차정보 등의 내용을 포함하고 있음

구분	분석 자료
데이터	955개의 데이터
선정기준	대구시 내 식당에 대한 객관적이고 일괄적인 정보를 제공받기 위해 대구시에서 운영하는 대구 푸드 데이터를 수집
대구 푸드 자료	식당 이름, 키워드, 한줄 소개, 주소, 영업시간, 대표 메뉴, 상세 소개, 휴무일, 좌석수, 주차정보, 예약 여부, 신용카드, 가는 길

<표 5> 대구 푸드 자료의 구성 요소

- 이를 HTML 태그를 사용한 웹 크롤링으로 대구 푸드 사이트 내 식당에 대한 정보를 <그림 3>와 수집을 하였으며, 수집된 식당의 수는 최종 955건으로 나타남



<그림 3> 대구 푸드 웹 사이트 및 HTML 태그

- 수집된 데이터를 파이썬 기반의 정규식을 활용하여 영어 대/소문자 통일 및 특수문자 제거 등 텍스트 피쳐 벡터화 진행을 위한 전처리 작업을 <그림 4>와 같이 수행함

```
def main_function(start_p=1, end_p=2):
    for pagenum in range(start_p, end_p):
        url = f'https://www.daegufood.go.kr/kor/food/food.asp?gotoPage={pagenum}&snm=9&list='
        res = requests.get(url)
        bsObject = BeautifulSoup(res.content, 'html.parser') # 원 page

        ## 가게 정보만 가져오기
        #title = bsObject.find('p', class='title')
        title = bsObject.select('p[class=title]') #p[class=title]의 애플리케이션 선택
        image = bsObject.select('p[class=img]')
        # 가게 정보 및 상세 설명 링크
        store_name = []
        store_url = []
        store_imgURL = []

        for a, img in zip(title, image):
            #print(a)
            name = a.text # 가게 이름
            store_name.append(name)
            url = 'https://www.daegufood.go.kr' + a.find('a')['href'] #가게 상세 page
            store_url.append(url)
            imgURL = img.find("img")["src"] #가게 대표 이미지 url
            imgURL = ('https://www.daegufood.go.kr'+imgURL)
            store_imgURL.append(imgURL)

        for name, url, img in zip(store_name, store_url, store_imgURL):
            store_zip.append(store(name, url, img))

class store:
    store_name = ''
    store_url = ''
    store_image_url = ''
    store_dict = {}
    def __init__(self, name, detail_url, img_url=''):
        self.store_name = name
        self.store_url = detail_url
        self.store_image_url = img_url
        self.store_url_detail(self.store_url)
    def __call__(self):
        self.store_name
    def store_url_detail(self, store_detail_url):
        #가게 상세 설명 page url
        detail_url = requests.get(store_detail_url)
        detail_soup = BeautifulSoup(detail_url.content, 'html.parser')
        self.store_detail(detail_soup)
```

<그림 4> 대구 푸드 데이터 전처리 코드

- 따라서 수집 후 전처리된 데이터는 <그림 5>와 같이 나타나며, 해당 데이터 중에서 대표메뉴와 상세 소개 자료를 결합하여 별도의 데이터를 생성함

식당 이름	키워드	한글소개	주소	연락처	영업시간	휴무일	좌석수	주차정보	출입시간	외국어	예약여부	신용카드	유아시설	조식	후식	대표메뉴	상세소개	오시는길(버스)	오시는길(지하철)	이미지Url	X	Y
0	팔공산맥	북구 대원버스 주차가 능 포장 가는 데 임 독립 된 룸 주류판매 가능 한 식...	팔공산 맥은 한 우와 LA 양념갈 비 등 다 양한 숯 불고기 를 맛볼 수 있는 숯불구 이...	대구광역시 북구 칠곡동 329 (태전동)	053-311-8445	10:00 ~ 01:00	연중 무휴	354 석 (동 5 개)	50 대	없음	가능한 외국어가 없습니다.	가능	가능	불가능	가능	특제논꽃생갈비 (110g) 22,000원 안향살 (110g) 27,000원 공중물티...	칠곡 태전동에 위치한 팔공산 맥은 돼지갈비와 한우가 유명한 구이 전문점입니다. ...	버스 정류장은 '대구 보건대학3' 정류장이 가장 가깝습니다.	지하철 3호선 구암역 1번 출구 약 440m	https://www.daegufood.go.kr/data/food/small/IM...	128.546654	35.9253
1	대덕식당	남구 대원버스 주차가 능 포장 가는 데 임 독립 된 룸 주류판매 가능 한 식...	대덕식당은 대구 사람이라면 누구나 알만한 선지국밥으로 유명한 맛집입니다.	대구광역시 남구 산호동 443 (대령9동)	053-656-8111	24시간	연중 무휴	250 석	50 대	없음	영어	가능	가능	불가능	불가능	선지국 7,000원 육개장 8,000원 옛날간장 찜닭 25,000원	대덕식당은 대구 사람이라면 누구나 알만한 선지국밥으로 유명한 맛집입니다. ...	버스 정류장은 '대덕' 정류장이 가장 가깝습니다.	주변에 지하철이 없습니다.	https://www.daegufood.go.kr/data/food/small/eo...	128.573623	35.8311

〈그림 5〉 대구 푸드 식당 데이터

III 분석 개요

1) 추천 시스템 요약

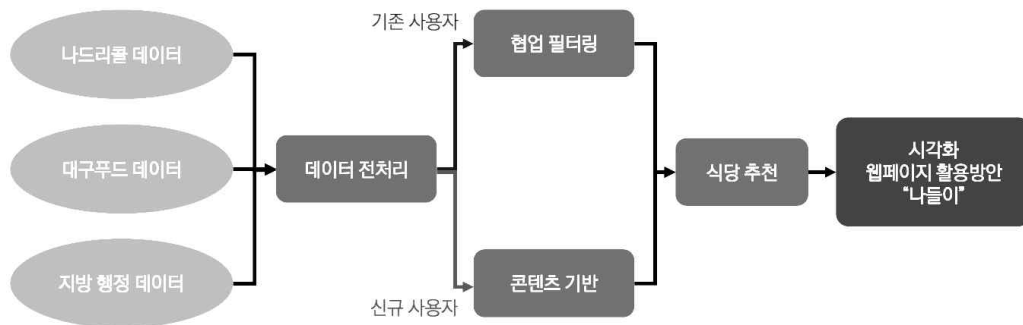
- 본 연구에서는 장애인 식당 추천 서비스를 구현하기 위해 민간 영역에 주로 활용되는 2가지 추천 시스템 알고리즘을 활용하여 구현을 하였으며, 다음과 같이 수행함
- 첫번째, 콘텐츠 기반 추천 시스템은 식당 간 대표메뉴와 상세소개의 자료를 결합하여 자연어 기반의 가장 유사한 식당을 추천해주는 기법임
- 두번째, 협업 필터링 추천 시스템은 다시 2가지로 분류되는데 해당 사용자 또는 식당과 가장 유사한 사용자 또는 식당을 추천하는 알고리즘으로 가장 널리 활용되는 알고리즘임
- 추천 시스템 구현을 위해 식당 또는 사용자 간 유사함의 정도를 코사인 유사도를 통해 산정하였으며, 협업 필터링의 경우 머신러닝 기반의 KNN 알고리즘을 적용함

2) 분석 진행 절차

- 본 연구에서 수행한 분석절차는 〈그림 6〉와 같으며, 특별교통수단(나드리콜) 배차이력, 대구푸드 식당 정보, 지방 행정 데이터들을 수집하여 결합하였으며, 이를 추천 서비스 구현이 가능하도록 이상치, 결측치 등을 제거함
- 그리고 신규 사용자에게는 특정 식당과 유사한 대표 메뉴와 소개 내용을 가지는 식당을 추천하는 콘텐츠 기반 추천 시스템을 적용하고, 식당 방문 이력 데이터가 있는 특별교통수단 이용자들에게는

협업필터링을 활용하여 해당 특정 유저와 가장 유사한 방문지를 가진 유저를 찾아 아직 방문하지 않은 식당을 추천함

- 이를 웹 사이트 형태로 시각적으로 구성하여 사용자의 편의성을 증진시키고자 함

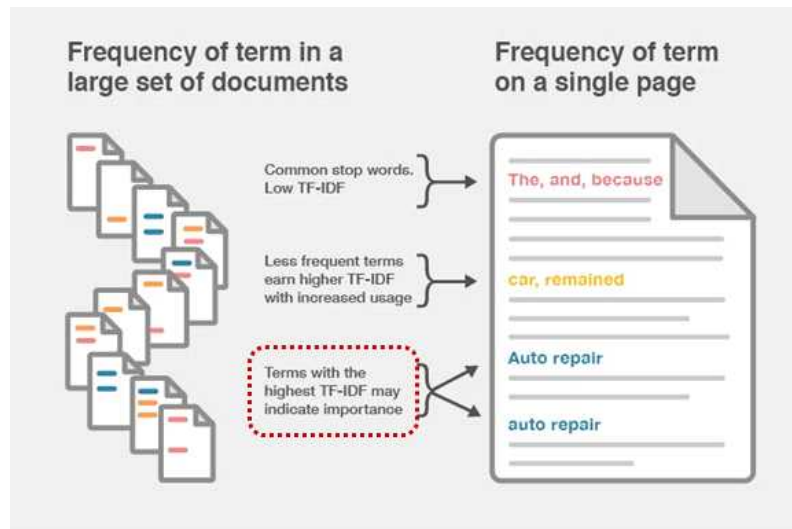


〈그림 6〉 분석 프로세스

5 콘텐츠 기반 추천 시스템

I 피쳐 벡터화

- 식당 간의 유사도를 판별하기 위하여 식당 데이터 칼럼 중 대표 메뉴와 상세 설명을 선택해 병합하여 새로운 텍스트 열 생성
- 해당 칼럼은 자료형이 텍스트이며 자연어이기 때문에 기계가 이해할 수 있도록 텍스트의 특징을 뽑아내는 별도의 전처리 작업이 필요함
- 텍스트의 특징을 추출하는 여러 기법 중 단순히 단어의 빈도수를 뽑아내는 것은 한국어의 특성상 조사가 많아 의미 없는 단어를 추출할 가능성이 있으므로 CountVectorizer 기법은 사용하지 않음
- 개별 단어의 문서에서 많이 등장하는 단어에는 가점을 주고, 여러 문서에서 많이 등장하는 단어에는 패널티를 주어 조사 등을 최대한 제거한 후 의미 있는 단어를 뽑아내는 방식인 TF-IDF 피쳐 벡터화 진행함



〈그림 7〉 TF-IDF 개념

- 피쳐 벡터화 진행 시, 문장에서 단어와 단어를 나누어 벡터화 시키는 특성으로 인하여 문장이 지닌 의미를 잃어버리게 됨
- 이를 해결하고자 N-gram을 3으로 설정하여 단어를 1~3개의 문치 단위로 끊어 하나의 토큰으로 간주하여 문장의 의미를 유지하도록 함
- N-gram을 3 이상으로 설정했을 때 결과의 성능 차이가 없는 것으로 나타났으며, 따라서 N-gram의 개수는 3으로 설정하여 분석을 실시함
- 이러한 과정을 거쳐 식당마다 의미 있는 텍스트를 추출하여 매트릭스를 형성함
- 〈그림 8〉과 같이, '어라 곱창 막창'의 메뉴인 '곱창 한 바가지'와 '돼지 김치찌개'가 의미 있는 텍스트로 추출되었음을 확인할 수 있음

식당 이름	가오리찜	가오리찜 (갈비탕)	갈비탕	어라곱창 (곱창한바가지)	장한바가지	남작만두	남작만두 (달서구)	달서구	돼지김치찌개	김치찌개
팔공산맥	0	0	0	0	0	0	0	0	0	0
대덕식당	0	0	0	0	0	0	0	0	0	0
갈비세상	0	0	0	0	0	0	0	0	0	0
별난북어	0	0	0	0	0	0	0.046076	0.046076	0	0
윤옥연할매	0	0	0	0	0	0	0	0	0	0
파크국밥	0	0	0	0	0	0	0	0	0	0
어라곱창	0	0	0	0.039472	0.039472	0	0	0	0.039472	0.039472
삼영식당	0	0	0	0	0	0	0	0	0	0
낙영찜갈비	0	0	0	0	0	0	0	0	0	0
고령촌돼지	0	0	0	0	0	0	0	0	0	0
푸른회식	0.038835	0.038835	0	0	0	0.038835	0.038835	0	0	0

〈그림 8〉 식당별 텍스트 피쳐벡터화

II 코사인 유사도(Cosine Similarity)

- 코사인 유사도는 두 벡터 간의 각도에 해당하는 코사인 값을 이용하여 구할 수 있는 두 벡터의 유사도를 의미함
- 코사인 유사도는 -1 이상 1 이하의 값을 가지며 값이 1에 가까울수록 유사도가 높다고 판단할 수 있으며, 두 벡터가 가리키는 방향이 얼마나 유사한가를 의미함
- 파이썬 기반의 Cosine_Similarity() 함수는 행을 기준으로 서로 다른 행을 비교해 유사도를 산출
- 해당 함수를 적용하면 자신을 제외하고 1번째 행의 레코드가 가장 유사도가 높음을 의미하고 가장 유사도가 낮은 것이 제일 마지막 요소임
- 문서 단어 행렬이나 TF-IDF 행렬을 통해서 문서의 유사도를 구할 때는 문서 단어 행렬이나 TF-IDF 행렬이 각각의 특징 벡터를 의미함
- TF-IDF를 통하여 만들어진 피쳐 벡터화를 사용하여 코사인 유사도를 구함
- 이를 통해서, 식당 간의 대표 메뉴 및 상세 설명을 비교하여 식당 간의 유사도를 구함
- <그림 9>과 같이 '명성숯불갈비'와 '갈비세상' 및 '큰형님숯불갈비' 식당이 타 식당에 비해 유사도가 0.022132로 높음을 확인할 수 있음
- 식당 간 유사도를 기준으로 행별로 비교 대상이 되는 행의 유사도 값이 높은 순으로 정렬된 행렬의 위치 인덱스값 추출하여, 특정 식당 검색 시 코사인 유사도가 높은 식당이 나올 수 있는 데이터 셋을 구축함

식당이름	팔공산맥	대덕식당	갈비세상	별난북어	윤옥연할매	명성숯불	큰형님숯불	돌솥당	자미성	고려가든	싱싱참치	무무스
팔공산맥	1	0.009521	0.036092	0.059082	0.006908	0.016487	0.037729	0.044931	0.035015	0.042455	0.004094	0.031113
대덕식당	0.009521	1	0.00711	0.006401	0.006542	0.005879	0.002598	0.00527	0.009199	0.006227	0.002558	0.00524
갈비세상	0.036092	0.00711	1	0.024997	0.005934	0.010632	0.130114	0.047878	0.061409	0.048539	0.004785	0.073475
별난북어	0.059082	0.006401	0.024997	1	0.006772	0.012982	0.040863	0.029048	0.024057	0.029224	0.004067	0.027395
윤옥연할매	0.006908	0.006542	0.005934	0.006772	1	0.01233	0.006115	0.009704	0.008887	0.008345	0.001519	0.007545
명성숯불갈비	0.016487	0.005879	0.010632	0.012982	0.01233	1	0.022132	0.020414	0.013942	0.022175	0.007457	0.010324
큰형님숯불	0.037729	0.002598	0.130114	0.040863	0.006115	0.022132	1	0.034141	0.07095	0.035283	0.007435	0.080127
돌솥식당	0.044931	0.00527	0.047878	0.029048	0.009704	0.020414	0.034141	1	0.041346	0.064356	0.006326	0.035877
자미성	0.035015	0.009199	0.061409	0.024057	0.008887	0.013942	0.07095	0.041346	1	0.036122	0.004876	0.066219
고려가든	0.042455	0.006227	0.048539	0.029224	0.008345	0.022175	0.035283	0.064356	0.036122	1	0.005478	0.032891
싱싱참치	0.004094	0.002558	0.004785	0.004067	0.001519	0.007457	0.007435	0.006326	0.004876	0.005478	1	0.00458
무무스	0.031113	0.00524	0.073475	0.027395	0.007545	0.010324	0.080127	0.035877	0.066219	0.032891	0.00458	1

<그림 9> 식당간 코사인 유사도

III 분석 결과

- TF-IDF 및 코사인 유사도를 사용하여 구한 식당 간의 유사도를 내림차순으로 정렬하여 식당 간 제일 유사한 식당을 차례대로 나열함
- 〈그림 10〉으로 살펴보면 입력한 식당과 메뉴가 비슷하거나 상세 설명이 비슷한 식당이 추천 됨을 확인할 수 있음

식당이름	키워드	한줄소개
180 성주숯불갈비	중구 달성공원 대형버스 주차가능 대규모 모임 독립된 룸 주류판매 한식 서문시장 동산의료원	중구 오토바이 골목에 위치. 한우 갈비살이 유명하다.
173 경주식당	달서구 포장가능 대규모 모임 독립된 룸 주류판매 wifi 가능 한식 A/W호텔 성서야...	경주식당은 한우숯불구이 전문점으로 최상급 한우만 판매하는 있는 숯불 한우 전문점임...
179 만미장	수성구 대형버스 주차가능 포장가능 대규모 모임 독립된 룸 주류판매 한식 수성관광호텔...	수성구에 위치한 대표적인 냉면 전문점으로 예전 두산오거리에 있던 고려정을 확장 이...
53 너구리	중구 한식 대구역 근대골목 대구콘서트하우스	최초영업자 정재인 는 현영업자 조인수 의 장인으로 6.25전쟁 이후 60년 초부터...
267 원조할매곰탕집	중구 달성공원 포장가능 대규모 모임 독립된 룸 주류판매 한식 노보텔 대구역 대구콘서...	원조할매곰탕집은 곰탕 전문점으로 한우만을 사용하는 34년 이상의 오래된 음식점입니다.
283 우각식육식당	중구 대형버스 주차가능 포장가능 대규모 모임 독립된 룸 주류판매 wifi 가능 한식...	우각은 한우고기 전문점으로 11 의 좋은 품질의 한우만 판매하는 한우고기 전문점...
164 안영감식육식당	수성구 대규모 모임 주류판매 한식	A 최고급 한우 사용 식육식당 상차림 3 000원 별도 한우특수부위를 골로루 선택...
308 송림식당	중구 대형버스 주차가능 포장가능 대규모 모임 독립된 룸 주류판매 wifi 가능 한식...	송림식당은 48년 이상의 전통을 지닌 북어 요리 전문점입니다.
349 함흥면옥	중구 주류판매 wifi 가능 장애인화장실 한식 김광석 거리 경대병원 mg k m g...	대백프라자 11층 불꽃냉면 불고기비빔냉면 등 신메뉴를 개발하여 고객의 입맛을 사로...
196 이모식당	중구 달성공원 포장가능 독립된 룸 주류판매 한식 서문시장 동산의료원	서성로 돼지골목에 위치한 순대집. 72년부터 운영되어 온 돼지국밥 수육 전문점.

〈그림 10〉 ‘명성숯불갈비’ 과 유사도가 높은 식당의 정보

- 앞서 인터뷰를 통해 장애인들이 식당을 선택하는 데 있어 친절 또한 중요하게 판단하고 있는 점을 감안하여 친절이라는 단어가 포함되는 경우 가점을 주어 식당 추천에 우선권을 부여함
- 해당 코드는 일반인이 활용하기에 어려우므로 사용자 편의성을 감안하여 시각화 도구인 Power-BI를 활용하여 웹 사이트 형태로 구축함
- 추천 시스템 UI는 좌측에 식당 이름을 입력하면 좌측 아래에 추천 식당이 출력되며, 우측에 추천 식당에 대한 사진, 위치, 영업시간, 장애인 화장실 정보가 나타나도록 구현함
- 단순한 맛집 추천 혹은 인플루언서들의 주관적인 추천이 아닌 식당 간의 메뉴 및 가게에 대한 상세 설명, 위치 등을 고려한 유사한 식당을 추천함으로써 기타 일반적인 추천 서비스와의 차별성을 보임
- 또한, 장애인들이 식당을 선택하는 데 있어 고려하는 요소인 친절도 및 편의 시설을 고려함으로써 장애인에게 맞춤형 추천 서비스를 제공할 수 있음



〈그림 11〉 콘텐츠기반 추천 시스템 구현(웹 사이트)

6 협업 필터링 추천 시스템

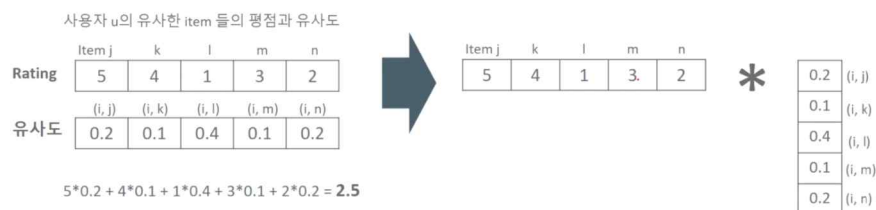
I 최근접 이웃(KNN) 알고리즘

- 최근접 이웃 알고리즘이란 새로운 데이터를 입력받았을 때 해당 데이터와 가장 근접한 데이터들의 종류를 확인 후 가장 많이 나온 종으로 분류하는 전통적인 분류예측 기법임
- 추천 시스템에서는 사용자들 간의 유사도를 바탕으로 예측 값을 계산해 높은 예측 값을 가지는 상위 n개의 추천 식당 목록을 생성하는데 사용되었으며, 다음과 같은 과정을 따름 (아이템 기반 기준)
- 첫째, 각 식당에 대한 사용자들의 ID와 방문한 식당을 기준으로 방문 횟수 매트릭스 생성
- 둘째, 두 값(방문 횟수와 코사인 유사도)을 내적하여 개인화된 방문 횟수 예측값 산출
- 셋째, 방문 횟수 예측값을 높은 순으로 정렬한 후 상위 식당 추천

Weighted Rating Sum

사용자 u의 아이템 i에 대한 평점 예측을 사용자 u가 아이템 i와 유사한 다른 아이템들(N개의 다른 아이템)의 합으로 계산하되 아이템 i와 다른 아이템들간의 유사도를 반영한 합으로 계산.

$$\hat{R}_{u,i} = \sum^N (S_{i,N} * R_{u,N}) / \sum^N (S_{i,N})$$



〈그림 12〉 최근접 이웃(KNN) 알고리즘

II 아이템 기반과 사용자 기반

- 최근접 이웃을 이용한 협업 필터링은 아이템 기반과 사용자 기반 추천 시스템으로 나뉨
- 첫째, 아이템 기반(item-based) 협업 필터링은 식당 간의 방문이 비슷한 사례를 찾아 방문하지 않은 식당을 사용자에게 추천함
- 방문 횟수를 사용해 특정 식당과 타 식당 간의 코사인 유사도를 구하여 최근접 이웃 알고리즘을 통해 사용자의 식당별 예상 방문 횟수를 구해 값이 높은 식당을 추출하여 추천하는 방식임

사용자1에게 식당1 추천	구분	사용자 1	사용자 2	사용자 3	사용자 4	사용자 5
	식당 1	0	1	1	1	0
	식당 2	1	1	1	1	0
	식당 3	0	1	0	0	1

〈그림 13〉 아이템기반 매트릭스(예시)

- 둘째, 사용자 기반(user-based) 협업 필터링은 특정 사용자와 이용 성향이 비슷한 사용자들이 방문한 다른 식당을 추천함 (기존 방문식당 제외)
- 방문 횟수를 사용해 특정 사용자와 타 사용자 간 코사인 유사도를 구하여 최근접 이웃 알고리즘을 통해 사용자와 유사도가 높은 사용자들을 선정 후, 특정 사용자가 아직 방문하지 않았지만 유사도가 높은 사용자들이 선호하는 식당을 추천함

사용자1은 사용자2와 유사도가 높음	구분	식당 1	식당 2	식당 3	식당 4	식당 5
	사용자 1	0	1	1	1	0
	사용자 2	1	1	1	1	0
	사용자 3	0	1	0	0	1

〈그림 14〉 사용자기반 매트릭스(예시)

- 두 방식은 매트릭스에서 각각의 행을 비교하여 코사인 유사도를 구해 식당에 대한 예상 선호도를 정량화된 수치로 계산하여 식당 추천의 토대로 활용함

III 분석 결과

- 분석 결과, 아이템 기반 방식이 사용자 기반 방식보다 평균제곱오차(MSE)가 낮아 정확도가 높은 것으로 나타났는데, 이는 사용자가 방문한 식당의 수가 대부분 2-3개로 낮기 때문에 보다 정확한 추천이 어려움
- 식당별 계산된 코사인 유사도 및 사용자별 계산된 코사인 유사도는 각각 <그림 15>와 같이 나타남

장소	극동구이	금강산	금강산식당	금수강산	기와집	name_code	1382	1391	1398	1401	1404
극동구이	1	0	0	0	0	1382	1	0	0	0	0
금강산	0	1	0.218218	0.065758	0	1391	0	1	0.707107	0.212598	0
금강산식당	0	0.218218	1	0	0	1398	0	0.707107	1	0.150329	0
금수강산	0	0.065758	0	1	0	1401	0	0.212598	0.150329	1	0
기와집	0	0	0	0	1	1404	0	0	0	0	1

<그림 15> 아이템 기반 코사인 유사도(좌), 사용자 기반 코사인 유사도(우)

- 첫째, 아이템 기반 추천 시스템은 KNN을 이용하여 구한 사용자별 예상 식당 방문 횟수가 높은 상위 15개 식당을 <그림 16>와 같이 예측값을 기준으로 내림차순 데이터 프레임 형태로 변환함
- 사용자 ID를 기준으로 예상 방문 횟수 데이터 프레임과 장애 유형이 있는 데이터 프레임을 결합하여 사용자별 장애 유형을 나타냄
- 식당을 추천 시 상위 15개의 가게에서 장애인들의 고려요소인 ‘친절’ 한 식당을 우선으로 하여 추천함
- 둘째, 사용자 기반 추천 시스템은 아이템 기반과 방식이 비슷하지만 코사인 유사도와 최근접 이웃을 이용할 때 식당 대신 사용자를 대입하여 분석함

user	장소	pred_score	친절	장애유형
14226	대교회식당	0.249058538	1	지체
14226	정아칼치	0.130193788	1	지체
14226	고향집칼국수,재바우	0.167574305	0	지체
14226	양반보쌈	0.833333333	0	지체
14226	본가장수촌	0.090368465	0	지체
14226	어선장	0.287765064	0	지체
14226	행복한날	0.130109189	0	지체
14226	청도동곡막걸리	0.279450021	0	지체
14226	본동복어	0.1254918	0	지체
14226	해동성	0.140681524	0	지체
14226	광장갈비	0.138018383	0	지체
14226	경산식당	0.097068163	0	지체
14226	거동구이	0.160914791	0	지체
14226	만미장	0.144664849	0	지체
14226	가야성	0.14185445	0	지체

user	장소	pred_score	친절	장애유형
14226	대교회식당	0.249058538	1	지체
14226	정아칼치	0.130193788	1	지체
14226	고향집칼국수,재바우	0.167574305	0	지체
14226	양반보쌈	0.833333333	0	지체
14226	본가장수촌	0.090368465	0	지체
14226	어선장	0.287765064	0	지체
14226	행복한날	0.130109189	0	지체
14226	청도동곡막걸리	0.279450021	0	지체
14226	본동복어	0.1254918	0	지체
14226	해동성	0.140681524	0	지체
14226	광장갈비	0.138018383	0	지체
14226	경산식당	0.097068163	0	지체
14226	거동구이	0.160914791	0	지체
14226	만미장	0.144664849	0	지체
14226	가야성	0.14185445	0	지체

<그림 16> 아이템 기반 추천 식당 리스트(좌), 사용자 기반 추천 식당 리스트(우)

- 협업 필터링 추천 시스템 UI는 아이템 기반과 사용자 기반으로 나누어 <그림 17, 18>과 같이 구성함
- 추천 시스템 UI는 좌측에 사용자 ID를 입력하면 좌측 아래에 추천 식당이 출력되며, 우측에 추천 식당에 대한 사진, 위치, 영업시간, 장애인 화장실 정보가 나타나도록 구현함
- 또한, 식당 정보를 워드 클라우드로 제공하여 식당의 특성에 대해 한눈에 볼 수 있도록 함
- 콘텐츠 기반 추천 시스템이 식당 상세내용 유사도를 기반으로 한 추천이라면, 협업 필터링 추천 시스템은 사용자의 식당 방문데이터를 이용하여 식당 또는 사용자 간 유사도를 기반으로 제공하는 추천 서비스임
- 사용자 기반 추천 시스템의 경우 아이템 기반보다 오차가 크기 때문에 실제 서비스를 제공할 때는 협업 필터링 아이템 기반 추천 서비스를 제공하되, 신규 사용자나 식당 방문 횟수가 적은 사용자는 정보가 부족하므로 콘텐츠 기반 추천 시스템을 통해 식당을 추천하는 것이 효과적이라고 판단됨



<그림 17> 아이템기반 추천 시스템 구현(웹사이트)



<그림 18> 사용자기반 추천 시스템 구현(웹사이트)

7 결 론

I 연구의 정책적 시사점

- 본 연구에서는 장애인 식당 추천 시스템을 구현하고자 특별교통수단 배차이력, 대구 푸드 데이터 등을 활용하여 콘텐츠 기반과 협업필터링 추천 서비스를 구현함
- 또한, 이를 이용하는 사용자의 편의성을 함양하기 위해 시각화 프로그램인 Power BI를 활용하여 웹 형태로 구축하였으며, 도출된 시사점은 다음과 같음
- 첫 번째, 장애 특성별 식당 수요 및 편의시설을 반영하는 추천 시스템을 통하여 장애인의 만족도 상승이 기대되며, 추후 편의시설 확충 우선 구역 선정 등 양질의 장애인 복지 서비스 개발의 기초자료로 활용될 수 있을 것으로 판단됨
- 두 번째, 특별교통수단(나드리콜) 앱과의 연계 및 외부 데이터와의 결합을 통하여 고도화된 추천 시스템 개발이 가능할 것으로 기대됨
- 마지막으로 추후 장애인 복지 정책 및 서비스 제공에 있어 본 연구가 장애인의 삶의 질 개선 및 사회적 가치 실현을 위한 기초 자료로 활용될 수 있을 것으로 판단되며, 본 연구 결과물을 통해 추가적인 인사이트 발굴이 가능할 것으로 판단됨

II 연구의 한계

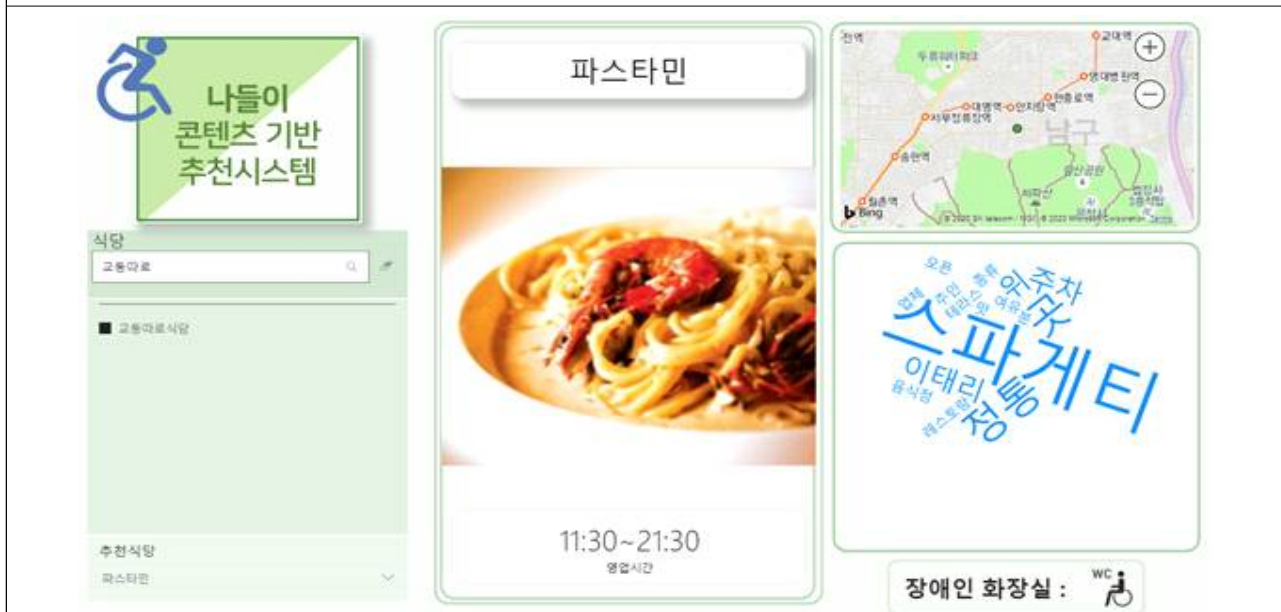
- 지체 장애인 비율이 높은 특별교통수단 배차이력 데이터를 활용하여, 특별교통수단을 이용하지 않은 계층에 대해서는 추천 서비스 구현이 어려울 것으로 판단됨
- 특별교통수단 배차이력 데이터와 대구푸드 식당 데이터 결합으로는 분석을 진행하여, 이용 식당의 이용범위가 제한됨
- 또한, 장애인 이동편의시설, 주차장, 보도폭 등의 물리적인 정보와 데이터 결합이 이루어진다면 추후 고도화된 추천 시스템 구현이 가능할 것으로 판단됨
- 마지막으로 인터뷰를 통해 장애인들이 식당을 이용함에 있어 친절이 중요하다는 의견을 반영해 알고리즘상 친절 키워드가 있을 경우 우선적으로 추천한 부분에 대해서는 객관적인 논리가 다소 부족함

I 대시보드 UI(파워BI)



대시보드 ① : 콘텐츠 기반 추천 시스템

- 좌측 식당 검색 창에 텍스트 입력 시 아래 리스트에 텍스트를 포함하는 식당 리스트 출력
- 추천식당 창에 선택한 식당과 가장 유사한 15개의 추천식당 리스트 출력



대시보드 ① : 콘텐츠 기반 추천 시스템

- 추천 식당 중 하나를 선택할 경우 좌측에 추천 식당 사진 및 영업시간, 위치, 장애인 화장실 유무 등이 표시
- 선택된 추천 식당에 관련된 상세설명 및 메뉴 정보를 Wordcloud로 제공



대시보드 ② : 협업 필터링 사용자 기반 추천 시스템

- 좌측 user-code 검색 창에 code 입력 시 아래 리스트에 code를 포함하는 user 리스트 출력
- 추천식당 창에 선택한 user와 성향이 비슷한 고객들이 방문한 식당 15개의 리스트 출력



대시보드 ② : 협업 필터링 사용자 기반 추천 시스템

- 추천 식당 중 하나를 선택할 경우 좌측에 추천 식당 사진 및 영업시간, 위치, 장애인 화장실 유무 등이 표시
- 선택된 추천 식당에 관련된 상세설명 및 메뉴 정보를 Wordcloud로 제공



대시보드 ③ : 협업 필터링 아이템 기반 추천 시스템

- 좌측 user-code 검색 창에 code 입력 시 아래 리스트에 code를 포함하는 user 리스트 출력
- 추천식당 창에 선택한 user가 방문했던 식당을 방문한 사람들의 예상 방문 횟수가 높은 식당 15개의 리스트 출력



대시보드 ③ : 협업 필터링 아이템 기반 추천 시스템

- 추천 식당 중 하나를 선택할 경우 좌측에 추천 식당 사진 및 영업시간, 위치, 장애인 화장실 유무 등이 표시
- 선택된 추천 식당에 관련된 상세설명 및 메뉴 정보를 Wordcloud로 제공

참고 문헌

1. 김건욱외 2인, 대구경북연구학회지 19(2),
빅데이터를 활용한 교통약자 특별교통수단 통행수요 분석:대구광역시를 사례로
 2. 김원호외 2인, 서울시연구. 9(2):100-119,
교통약자 유형별 이동행태분석 및 맞춤형 대중교통정보 제공방안 연구
 3. 이연숙 외 2인, 한국실내디자인학회 논문집.26(1):112-120
장애유형별 편의시설에 대한 평가 및 개선요구에 관한 연구
 4. 임완수, 월간 주민자치. 40(1):117-124
장애인 편의시설 접근성에 관한 데이터를 직접 수집해 제공 사회적 기여 성취감과 자존감 높아지다
 5. 장애인권익지원과, 보건복지부
2018년도 장애인 편의시설 실태전수조사
 6. 장애인권익지원과, 보건복지부
장애인복지법 시행규칙
-

본 보고서는 대구의 경제·사회 전망을 빅데이터로 진단하고
대응방안 제시, 다양한 분석기법들을 소개하는 분석자료입니다.

본 자료의 내용은 연구자의 견해이며 대구디지털산업진흥원의 공식 견해와
반드시 일치하는 것은 아닙니다.



대구광역시 수성구 알파 시티1로 160 2층
전화 053-655-7050 FAX 053-655-3501