

산학협력 프로젝트 중간보고서

| | | | | |
|-------|--|-----------|--------|---------|
| 과제명 | 빅데이터와 인공지능을 활용한 상권분석 서비스 | | | |
| 수행기간 | 2020.08.24.~2020.11.30. | | 책임교수명 | 육동철 (인) |
| 협력기관명 | 와이폴라리스(YPolaris) | | 기관담당자명 | 김승연 |
| 참여학생 | 이름 | 학번 | 팀명 | AI STAR |
| | 팀장: 허정훈 | 201511152 | | |
| | 팀원: 권기호 | 201542104 | | |
| | 팀원: 이가영 | 201611129 | | |
| | 팀원: 강다미 | 201645801 | | |
| 참여인원 | 참여교수 1명(육동철), 학부생 4명(허정훈, 권기호, 이가영, 강다미), 협력기관 1명(이현수 CTO) | | | |
| 수행현황 | 상권 데이터 결과 지도 시각화 및 상권 예측 | | | |

1. 과제 요구 사항 변경내용

개발팀과 분석팀으로 나누어 작업을 진행하며 팀 간의 협업과 업체와의 정보 교류는 구글드라이브를 사용했다.

분석팀은 회귀분석을 이용해 상권분석을 위한 변수를 지정하고 평균선을 내어 좋은 상권과 그렇지 못한 상권을 예측하는 식을 만들고 시각화했다.

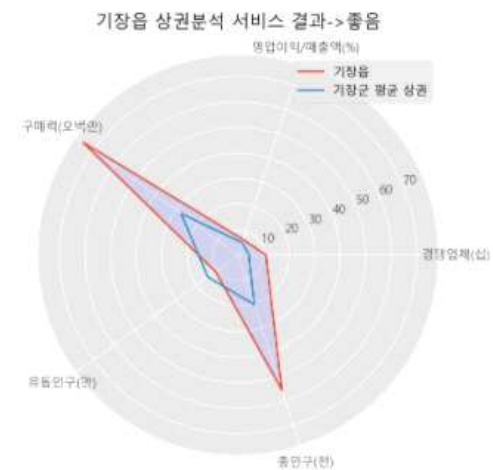
개발팀은 메인페이지 시스템상의 오류를 수정하고 검색 시 나올 수 있는 경우의 수를 나열해 페이지를 재구성하며 해당 페이지에 검색바와 카테고리를 생성했다. 인구수는 지도를 확대해 보여주는 식으로 구현했다.

2. 수행 내용 및 중간 결과

분석팀은 상권분석을 위한 변수로 유동인구, 경쟁업체 수, 구매력, 총인구, 매출대비 영업이익으로 구성된 뒤 임의의 가중치를 두어 평균선을 중심으로 상권분석 초안을 만들어 방사형 그래프로 만들었다. 상권분석 시각화를 위한 등급을 매길 때는, 먼저 원하는 상권과 그 상권이 속한 구의 평균 상권의 차이를 구해서 각 항목의 합을 구했다. 단위 차이에 대한 방지책으로 속성별로 Z-normalization을 해주고 가중치를 조절하였다. 임의로 유동인구, 영업이익/매출액(%), 경쟁업체 순으로 영향을 미친다고 가정하여 그에 따른 가중치를 추가하고 모든 동의 결과의 기준을 $\max=11.9$, $\min=-6.04$, $\text{mean}=0.0$, 상위 25%=1.71 하위 25%=-1.79 라고 두고, 하위 25% : 나쁨, 하위 25%-평균 : 고려, 평균-상위 25% : 양호, 상위 25% : 좋음이라고 평가하였다.

회귀분석을 이용한 상권분석의 결과와 방사형 그래프는 아래와 같다.

['우1동', '우2동', '중1동', '중2동', '좌1동', '좌2동', '좌3동', '좌4동', '송정동', '반여1동', '반여2동', '반여3동', '반여4동', '반송1동', '반송2동', '재송1동', '재송2동']
 사하구
 ['괴정1동', '괴정2동', '괴정3동', '괴정4동', '당리동', '하단1동', '하단2동', '신평1동', '신평2동', '장림1동', '장림2동', '다대1동', '다대2동', '구평동', '감천1동', '감천2동']
 금정구
 ['서1동', '서2동', '서3동', '부곡1동', '부곡2동', '부곡3동', '부곡4동', '장전1동', '장전2동', '선두구동', '청룡노포동', '남산동', '구서1동', '구서2동', '금성동']
 강서구
 ['대저1동', '대저2동', '강동동', '가락동', '녹산동', '가덕도동']
 연제구
 ['거제1동', '거제2동', '거제3동', '거제4동', '연산1동', '연산2동', '연산3동', '연산4동', '연산5동', '연산6동', '연산8동', '연산9동']
 수영구
 ['남천1동', '남천2동', '수영동', '망미1동', '망미2동', '광안1동', '광안2동', '광안3동', '광안4동', '민락동']
 사상구
 ['삼락동', '모라1동', '모라3동', '덕포1동', '덕포2동', '괘법동', '감전동', '주례1동', '주례2동', '주례3동', '학장동', '엄궁동']
 기장군
 ['기장읍', '장안읍', '일광면', '철마면']
 기장읍



상권 분석 시각화

상권의 등급 매기기

- 원하는 상권과 그 상권이 속한 구의 평균 상권의 차이를 구해서 각 항목의 합을 구함
 - 단위가 차이 나기 때문에 개별 컬럼별로 Z-normalization을 해주고 가중치를 조절함
 - 유동인구, 영업이익/매출액(%), 경쟁업체 순으로 영향을 미친다고 생각하여 그에 따른 가중치 추가
 - 모든 등의 결과 max=8.17, min=-5.81, mean=0, 상위 25%=2.34 하위 25%=-2.24
 - 하위 25% : 나쁨, 하위 25%-평균 : 고려, 평균-상위 25% : 양호, 상위 25% : 좋음이라고 평가

그림1 [회귀분석을 통한 방사형 그래프]

개발팀은 카카오 API를 이용해 지도를 구성했다. 지도는 요구사항에서 요청한 각각의 경우의 수에 대한 메인페이지 에러를 구현하고 각 동의 중심을 동사무소로 설정해 마커를 띄웠으며 지도 확대시 인구수를 표출하는 방식으로 구현했다. 지도에 표출시키기 위한 데이터(csv)들은 데이터베이스화 해서 서버에 삽입시켰다. 데이터들은 상권의 정보가 담긴 파일들을 포함해 인구수 및 필요에 따라 편집한 파일들로 이루어져 있다.

상황별 페이지와 데이터베이스는 아래와 같다.



그림2 [메인 페이지1]

사이트에 처음 접속했을 때 볼 수 있는 메인 페이지이다. 기본으로 장전2동의 위치와 인구 수가 출력된다.



그림3 [메인 페이지2]

업종을 선택하지 않으면 업종을 선택해달라는 알림창이 뜬다.



그림4 [메인 페이지3]

부산에 존재하지 않는 동을 입력했을 시 검색결과가 존재하지 않는다는 알림창이 뜬다.



그림5 [메인 페이지4]

업종은 대형마트, 편의점, 기타 슈퍼마켓으로 분류되어 선택이 가능하다.



그림6 [인구수 출력 화면]

문현1동의 편의점을 검색했을 때 처음으로 나오는 화면이 문현동의 인구수임을 알 수 있다.



그림7 [zoom in 시 편의점 정보 출력 화면]

ZOOM IN을 하면 문현1동의 편의점 위치와 주소 그리고 상호명을 확인할 수 있다.

```
mysql> show tables;
+-----+
| Tables_in_retail_info |
+-----+
| busan_admin_population |
| busan_map_analysis     |
| busan_population_analysis |
| busan_store            |
| busan_store_select     |
| busan_store_select2    |
+-----+
6 rows in set (0.03 sec)
```

그림8 [데이터베이스 화면]

현재 데이터베이스의 테이블은 총 6개로 구성되어 있다.

busan_admin_population 테이블은 부산시의 구와 행정동명 그리고 행정동별의 인구수가 저장되어 있다.

busan_map_analysis 테이블은 지도에 표출되는 상권 예측 및 추천 프로그램에 필요한 데이터들이 저장되어 있다.

busan_population_analysis 테이블은 부산시의 인구수와 유동인구 수 그리고 경쟁업체 및 상권을 분석하기 위한 데이터와 분석의 기본이 되는 분석 데이터가 저장되어 있다.

busan_store 테이블은 부산시의 모든 상점 정보가 저장되어 있다.

busan_store_select 테이블과 busan_store_select2 테이블은 busan_store 테이블의 열들을 필요에 따라 편집한 정보들이 저장되어 있다.

3. 추진 계획 및 진척도

분석팀은 상권분석 초안에 관한 수정을 거치고 방사형 그래프를 이용해 상권분석 결과를 나타낸 다음 5개의 변수를 이용해 일정한 기준을 정하고 그것을 바탕으로 상권이 좋은지와 그렇지 않은지를 판단하게 하는 것을 목표로 한다.

개발팀은 상권분석의 결과에 해당하는 방사형 그래프를 지도에 표출시키고 분석팀의 데이터를 바탕으로 해당 상권이 타 상권들과 비교했을 때 경쟁력이 있는지를 표현하는 것을 목표로 한다.

상기와 같이 산학협력 프로젝트 중간보고서를 제출합니다.

2020 년 10 월 30 일

팀장 허정훈