



스마트 배너 입지 선정 및 홍보 관리 서비스 구축

스마트 홍보 시스템

8조 : 권지혜, 송예은, 문구영 | KU홍보지기 TEAM

2020 KU 스마트 캠퍼스 데이터톤

Contents

배경



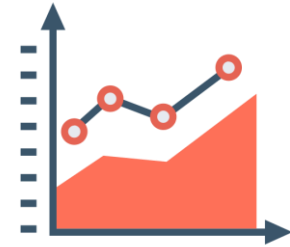
스마트 배너
입지 선정



홍보물 관리
시스템 구축



기대효과



Contents

배경



- 현재 교내 홍보 시스템의 운영 방식
- 현재 교내 홍보 시스템의 문제점 및 필요성
- 프로젝트 과정

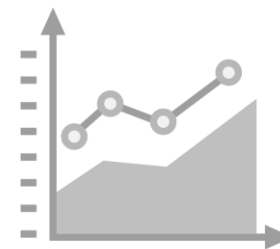
스마트 배너
입지 선정



홍보물 관리
시스템 구축



기대효과



현재 교내 홍보 시스템의 운영 방식

*출처: 고려대학교 총학생회

- 1 게시물 부착에 대한 총학생회의 허가제는 별도로 존재하지 않는다.
- 2 단, 과대 게시 방지를 위해 개강 직후에 한해 **게시물 총량제**를 적용한다
*게시물 총량제란?
총학생회의 게시 허가를 받고 일자 별 스티커가 주어지면 정해진 기간에 홍보물을 게시할 수 있는 제도
- 3 개강 직후가 아닌 경우에는 게시가 자유롭다.
 - 게시 제한 기간 X
 - 외부 단체의 게시에 대한 제약 X

배경 _ 현재 교내 홍보 시스템의 문제점 및 필요성

이로 인해 발생하는 문제점에는 어떤 것이 있을까?



시간적, 금전적 부담

→ 홍보물을 게시판에 부착하기 위해선 포스터의 인쇄 및 부착, 게시 허가를 위한 많은 시간과 돈이 수반됨



게시물 총량제 위반 다수 발생

→ 홍보허가를 받지 않은 홍보물이 부착되어 이를 단속하기 위한 인력이 필요함



홍보물 미철거

→ 모집기간이 지났음에도 부착한 홍보물을 철거하지 않음



나에게 필요한 홍보물을 찾기 어려움

→ 내가 정말 필요로 하는 홍보물을, 필요한 순간에 접하지 못함

배경 _ 현재 교내 홍보 시스템의 문제점 및 필요성



홍보물의 무분별한 탈 부착으로 인한
미관 저해

뚜렷한 기준을 통한 홍보물의 분류가
없음 (즉, 원하는 정보를 찾기 어려움)



앱을 이용한
체계적인 교내 홍보 관리 시스템,
스마트 배너의 필요성 확인



STEP1. 교내 스마트 배너 설치 최적 입지 선정

- 교내 수업 데이터와 건물 데이터를 사용해 학생들의 유동인구 파악
- 가장 많은 사람들이 홍보물을 확인할 수 있는 스마트 배너 설치 입지 제안

STEP2. 홍보물 자동화 관리 시스템 구축

- 학생들이 홍보물을 오프라인 및 온라인에 쉽게 등록하고 관리 할 수 있는 서비스 구축
1. 오프라인 홍보물 게시
 - 게시물 (홍보 포스터 등) 업로드
 - 홍보물 분석을 통한 적절한 게시 장소 제안
 - 학교로 전송, 학교측에서 검토
 - 전자배너에 게시
 2. 온라인 홍보물 게시
 - 교내 게시판이 아닌 앱 내에서 홍보물 업로드 및 관리
 - 데이터분석을 통해 앱 사용자의 조회 이력, 관심사에 부합하는 **맞춤형 홍보물 추천**

Contents

배경



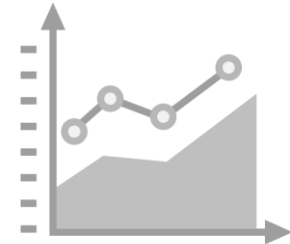
스마트 배너
입지 선정



홍보물 관리
시스템 구축



기대효과



- 스마트 배너란?
- 분석 FLOW
- Visualization
- 분석 결과
- 최종 입지 선정

스마트 배너 설치 최적 입지 선정_ 스마트 배너란?

★ 스마트 배너란?



코엑스 언더월드 파노라마

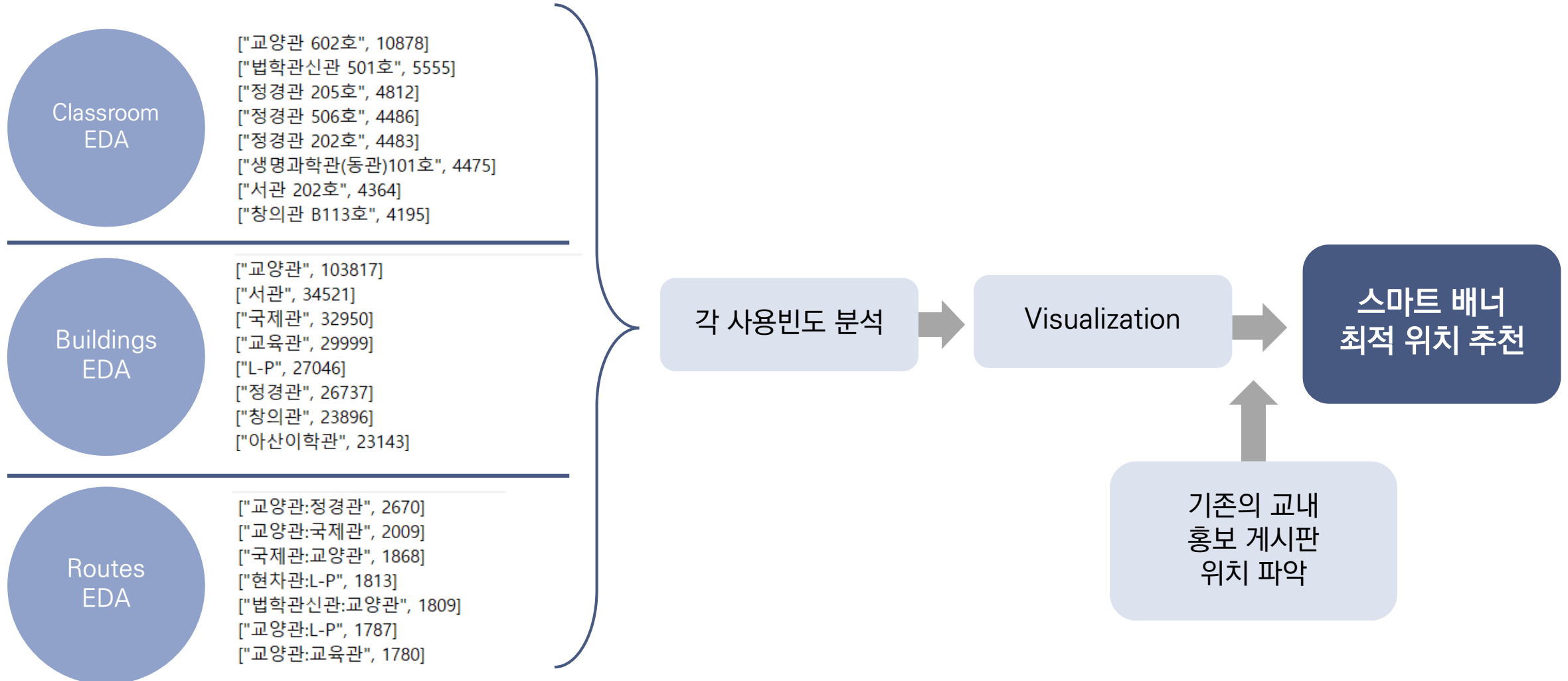


강남역 파노라마 미디어 플랫폼

- **디지털 전광판**으로, 다양한 디스플레이로 정보를 송출할 수 있는 홍보 매체이다.
- 기존의 출력 광고와 다르게 물리적인 작업이 필요 없으며 설치 이후엔 비용이 거의 들지 않아 자원 효율적이다.
- 이를 학교에 설치할 시, 최대한 많은 학생들이 스마트 배너에 노출될 수 있는 장소를 선정한다.

스마트 배너 설치 최적 입지 선정_ 분석 FLOW


제공된 2019년도 강의 시간 및 강의실별 학생 데이터 활용

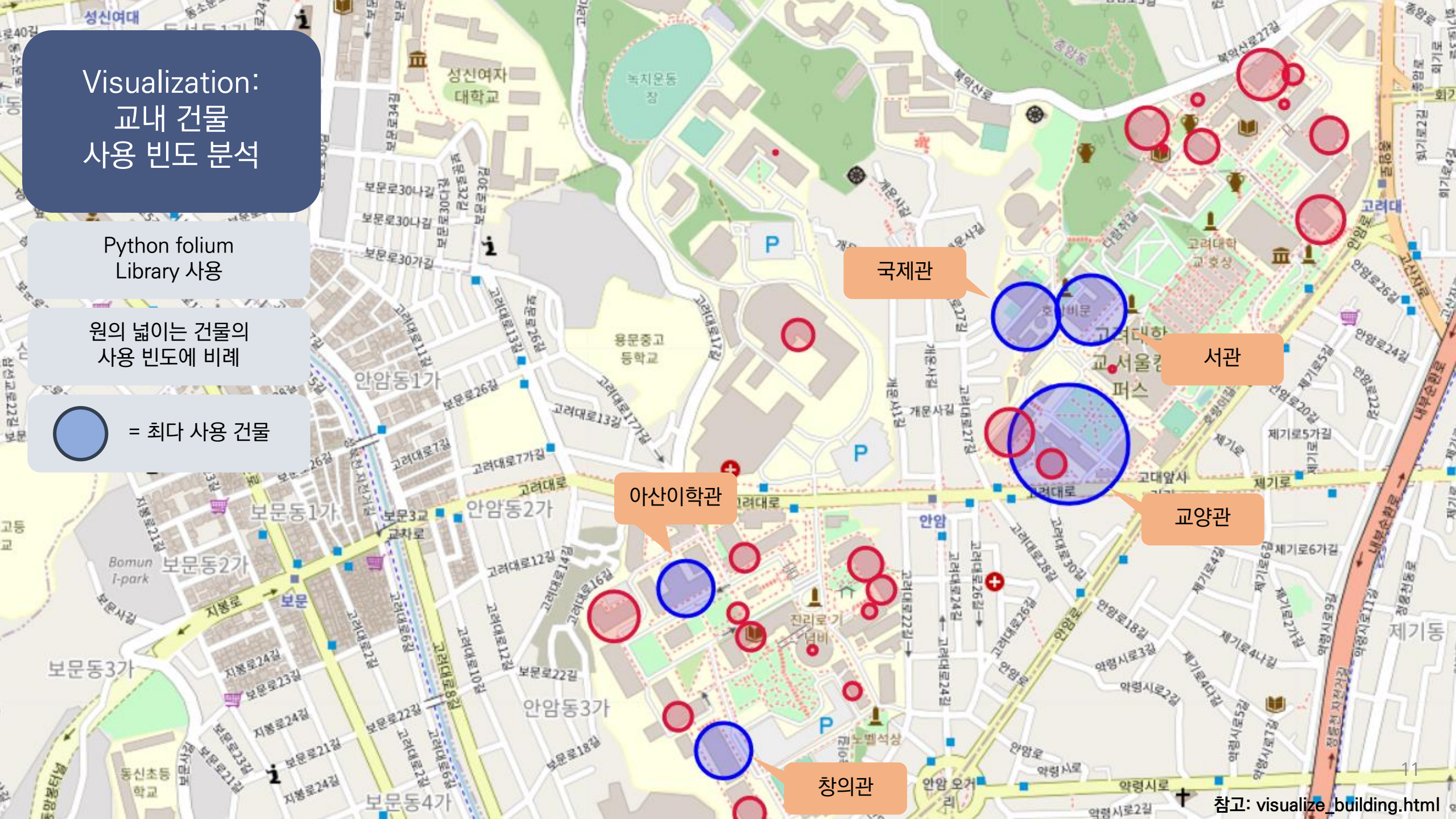


Visualization: 교내 건물 사용 빈도 분석

Python folium
Library 사용

원의 넓이는 건물의
사용 빈도에 비례

 = 최다 사용 건물

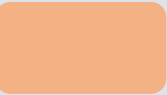


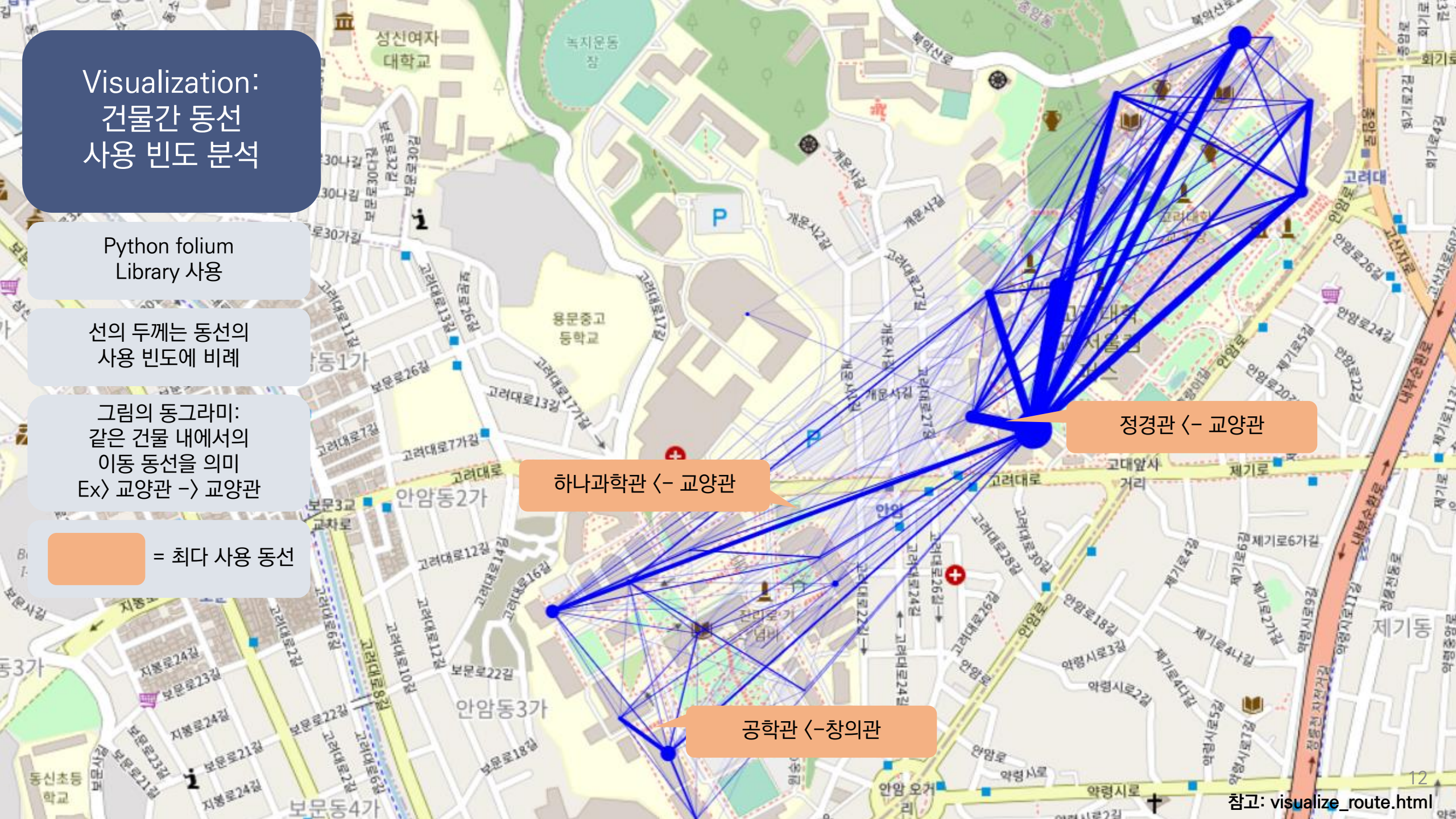
Visualization: 건물간 동선 사용 빈도 분석

Python folium
Library 사용

선의 두께는 동선의
사용 빈도에 비례

그림의 동그라미:
같은 건물 내에서의
이동 동선을 의미
Ex) 교양관 -> 교양관

 = 최다 사용 동선



정경관 <- 교양관

하나과학관 <- 교양관


공학관 <- 창의관

Visualization: 건물간 동선 사용 빈도 분석

Python folium
Library 사용

선의 두께는 동선의
사용 빈도에 비례




그림의 동그라미:
같은 건물 내에서의
이동 동선을 의미
Ex) 교양관 -> 교양관

 = 최다 사용 동선

교양관 -> 법학관신관

교양관 -> 서관

스마트 배너 최종 입지 선정 - 인문계

-  = 현재 종이 게시판이 설치되어 있으며, 스마트 배너로 교체 제안
-  = 현재 종이 게시판이 설치되어 있으며, 적은 유동인구로 배너의 필요성이 적다
-  = 게시판 자체가 설치되어 있지 않으며, 스마트 배너의 도입이 시급

현차관-LP관
사이 골목

다람쥐길

국제관 내부




교양관 - 서관
사이 골목

정경대 후문

교양관 내부



스마트 배너 최종 입지 선정 - 자연계

-  = 현재 종이 게시판이 설치되어 있으며, 스마트 배너로 교체 제안
-  = 현재 종이 게시판이 설치되어 있으며, 적은 유동인구로 배너의 필요성이 적다
-  = 게시판 자체가 설치되어 있지 않으며, 스마트 배너의 도입이 시급

하나과학관 내부

자연계 정문

하나스퀘어 내부

노벨광장

공학관 - 창의관
사이 골목

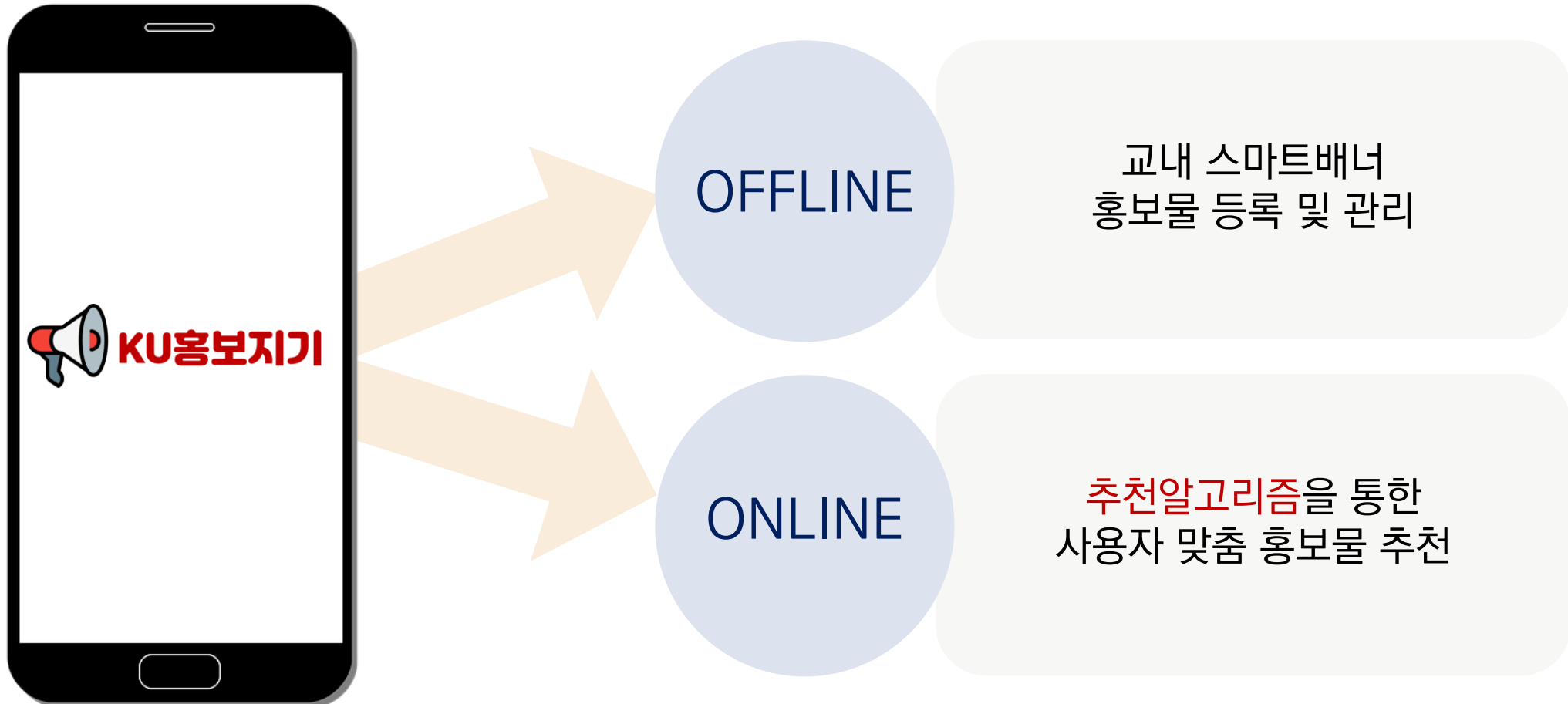
애기능학생회관 정문

Contents



홍보물 관리시스템 구축 _ 시스템 개요

교내 스마트배너 설치와 더불어 홍보물을 온/오프라인으로 쉽게 등록하고 관리할 수 있는 어플리케이션을 구현한다.



홍보물 관리시스템 구축 _ 추천 알고리즘

▶ 추천 알고리즘이란?



○○ • 2개월 전

알 수 없는 유튜브 알고리즘이 나를 꽤 괜찮은 곳으로 인도한 것 같다.

👍 8.9천 🗨️ 38

답글 38개

- 수많은 데이터의 축적, 알고리즘의 전문화에 따라 좋은 성능을 보이는 중

✓ 추천 시스템 님, 좋아하는 콘텐츠를 3개 선택하세요.

회원님이 좋아하실 만한 영화와 TV 프로그램을 더 많이 추천해 드릴 수 있습니다. 마음에 드는 콘텐츠를 선택하세요.

다음



- 이용자의 취향을 고려하는 고객 맞춤형 시스템의 발전

Objective : 앱 이용자들의 관심사, 조희 이력 등의 정보를 고려해 적절한 홍보 게시물을 추천해주는 맞춤형 시스템 구현

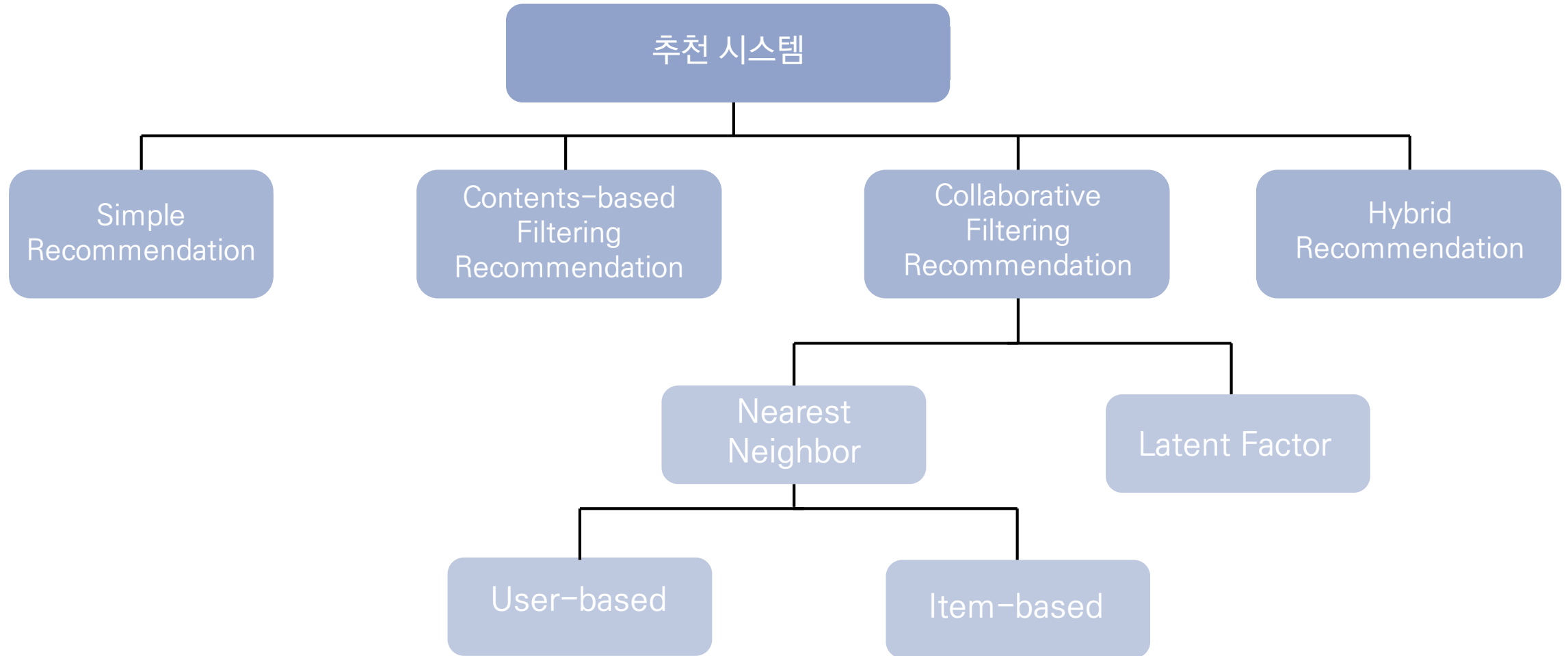


효과1) 시간 단축 : 검색 등의 중간과정 없이 direct로 콘텐츠 추천

효과2) 커버리지 확보 : 이용자가 예측하지 못한 콘텐츠를 접하게 함 (콘텐츠 폭 확장)

홍보물 관리시스템 구축 _ 추천 알고리즘 유형

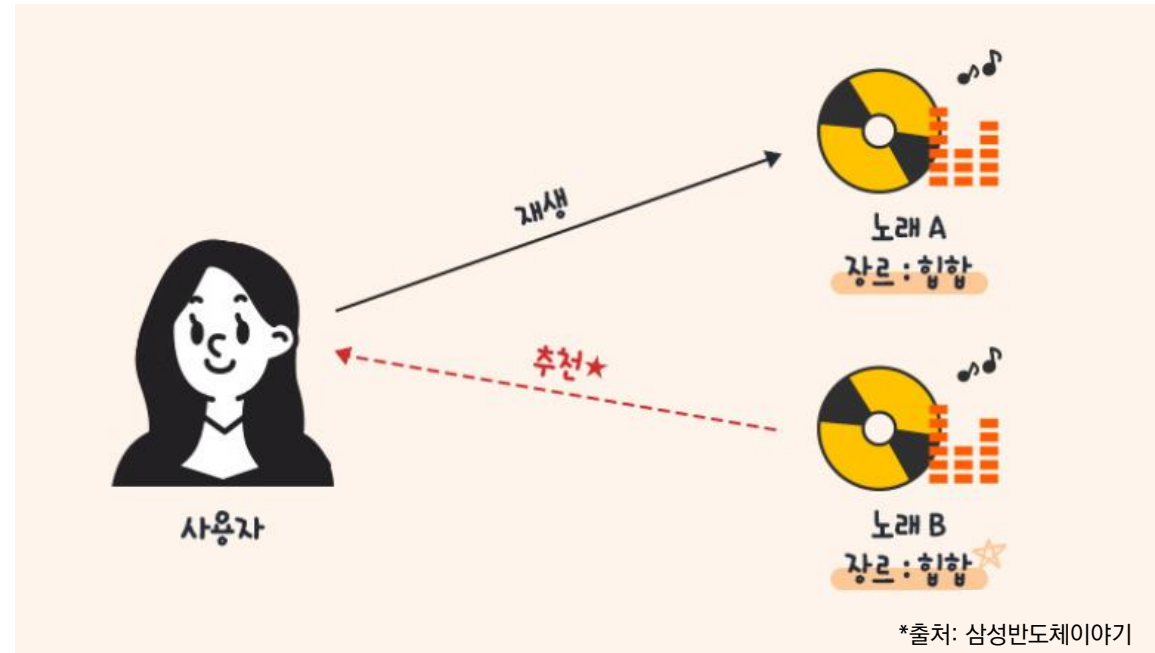
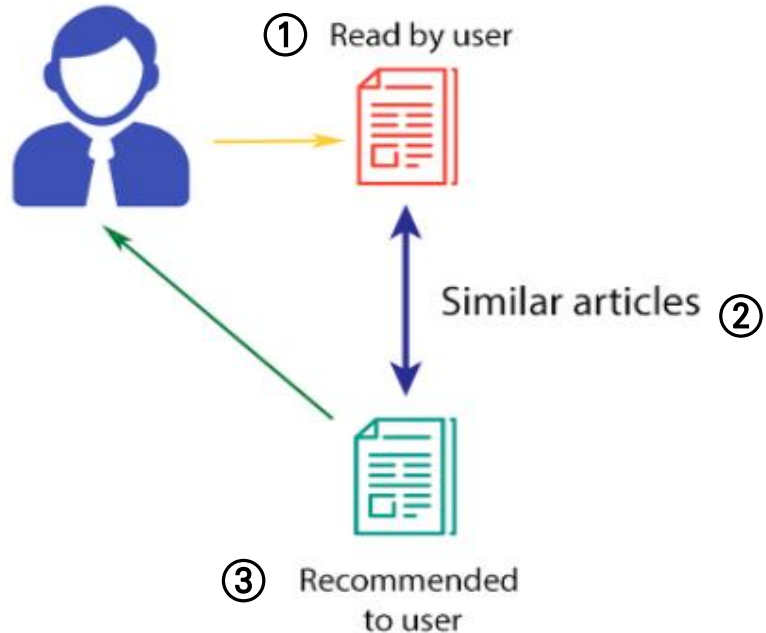
▶ 추천 알고리즘 유형



홍보물 관리시스템 구축 _ 추천 알고리즘 유형

▶ Contents-based Filtering Recommendation (내용 기반 추천시스템)

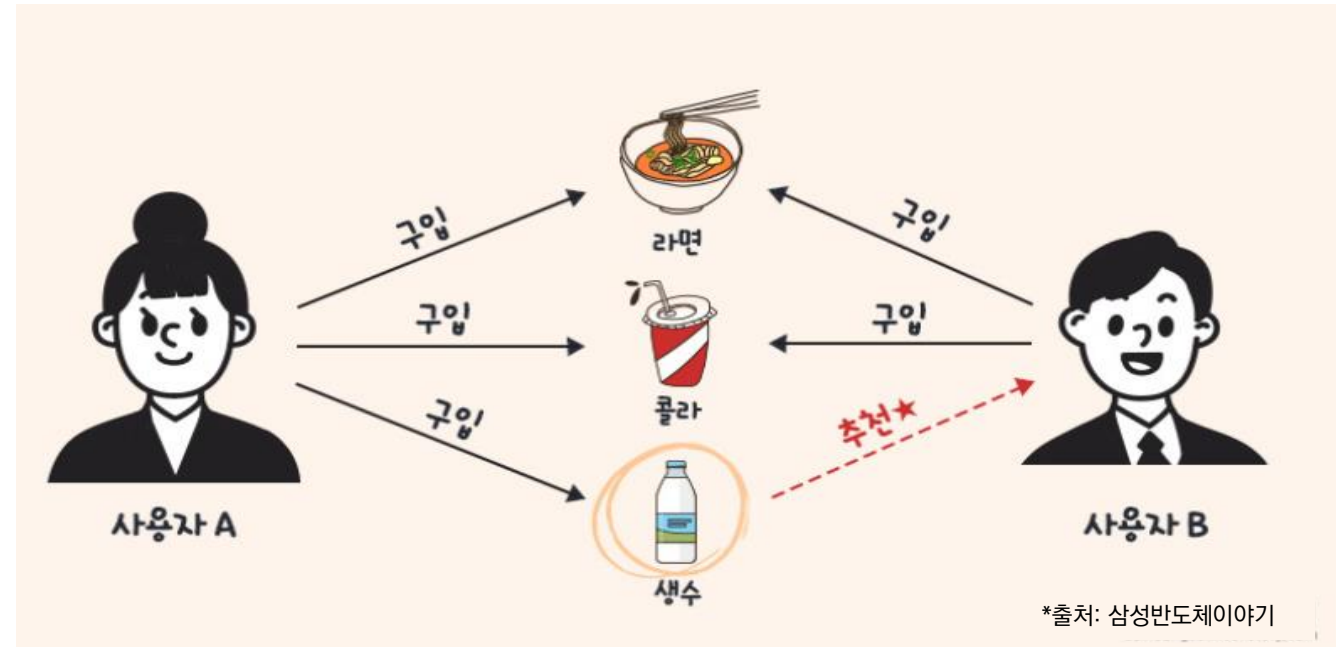
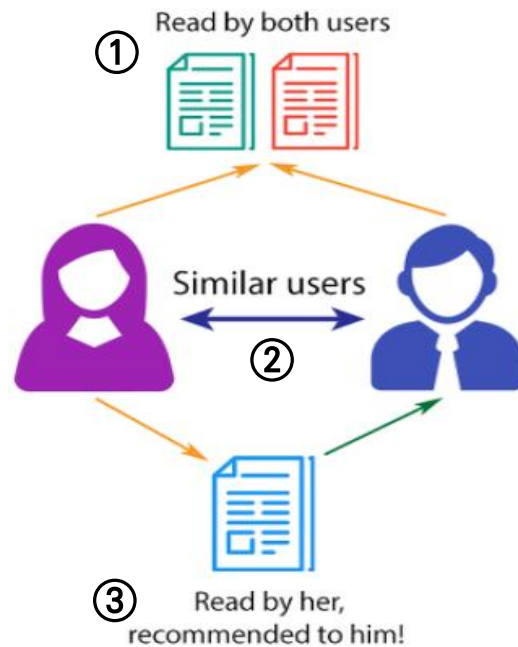
- 콘텐츠 내용의 유사성을 바탕으로 추천을 하는 알고리즘
- 이용자가 조회하거나 구매한 콘텐츠 자체의 특징을 분석하여 그와 유사한 콘텐츠 추천
- ex) 클래식 음악을 듣는 이용자에게 최근 발매된 클래식 음악이나, 평소 듣는 음악가의 다른 음반 추천



홍보물 관리시스템 구축 _ 추천 알고리즘 유형

▶ Collaborative Filtering Recommendation (협업 필터링 추천시스템)

- **이용자와 콘텐츠 사이의 유사성을** 기준으로 **비슷한 성향의 이용자들이 선호하는 콘텐츠 추천**
- 나와 가장 유사한 사람을 찾아주거나(사용자 기반) / 나의 관심사와 가장 유사한 콘텐츠를 찾음(아이템 기반)
- ex) 모차르트 음악을 듣는 이용자가 베토벤의 음악도 들었다면, 모차르트 음악을 듣는 다른 이용자에게도 베토벤 추천



홍보물 관리시스템 구축 _ 추천 알고리즘 유형

▶ Latent Factor Collaborative Filtering Recommendation

- 2009년 넷플릭스 추천 시스템 경진대회에서 해당 알고리즘이 우승하면서 주목 받음
- **행렬 분해**(Matrix Factorization)를 활용해 user-item 간 **잠재 요인**을 파악하는 알고리즘

예시) 이용자-홍보물 조회수 행렬에 잠재되어 있는 어떠한 요인 검출 → 이를 기준으로 다른 홍보물 추천

	홍보물1	홍보물2	홍보물3	홍보물4	홍보물5
이용자1	4			2	
이용자2		5		3	1
이용자3			3	4	4
이용자4	5	2	1	2	

=

	요인1	요인2	요인3
이용자1	0.96	0.47	-0.76
이용자2	-0.03	0.84	-2.47
이용자3	2.38	0.11	-1.20
이용자4	0.59	1.10	-1.06

×

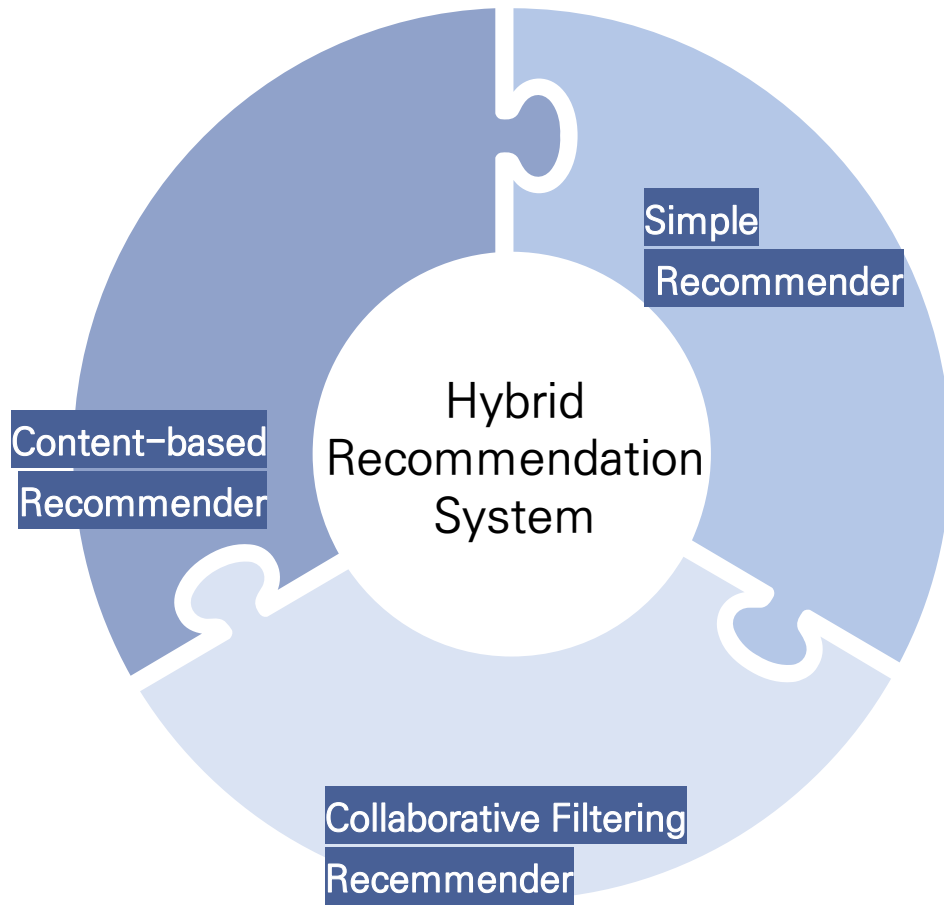
	홍보물1	홍보물2	홍보물3	홍보물4	홍보물5
요인1	1.62	-0.79	1.04	1.07	1.43
요인2	1.51	0.45	-0.06	0.12	-0.21
요인3	-2.22	-1.85	0.43	1.18	-0.50

- 각 이용자별 홍보물들에 대한 조회수
- 빈 칸 = 아직 조회되지 않은 홍보물

- (이용자, 홍보물) = (이용자, 잠재요인) × (잠재요인, 홍보물)
- **잠재요인** : blackbox element로 구체적인 정의는 어려우나, 추천의 근거가 됨
- 분해된 두 행렬의 곱은 원 행렬의 조회수 예측치 제공

홍보물 관리시스템 구축 _ 추천 알고리즘 유형

▶ 추천 알고리즘 유형 -Hybrid Recommendation



Simple Recommendation(고객유형/인기도를 기준으로 아이템 분류 및 추천)
+
Contents-based Filtering Recommendation,
+
Collaborative Filtering Recommendation

의 조합으로 이루어진 알고리즘

- 유저의 상황에 따라 각 알고리즘에 적절한 가중평균을 취함
- 성격이 다른 추천 알고리즘을 조합함으로써 **다양한 추천 시스템들의 장점 극대화, 단점 보완**

홍보물 관리시스템 구축 _ 추천 알고리즘 구축

▶ 홍보물 데이터 수집

공모전, 대외활동, 동아리, 학회, 취업활동 등 다양한 분야에서의 홍보물을 크롤링

KOREAPAS^o

쿠 카 이 브
KU chive

링커리어
LINKareer

KOREA UNIVERSITY
고려대학교
경력개발센터

BeautifulSoup,
Selenium

Crawling

articleid	title	host	duedate	description	image	tag
1	한국야쿠르트 마케팅 공모전	한국 야쿠르트	2020-08-31	2020년 한국야쿠르트 마케팅 공모전 ■ 공모주제 · 언택트 시대에 걸맞은 차별...	https://s3.ap-northeast-2.amazonaws.com/media....	['공모전', '경영']
2	행정안전부 제1회 도전.한국 대국민 아이디어 공모전	행정안전부	2020-09-15	제1회 도전.한국 대국민 아이디어 공모전 > 지원자격 : 관심 있는 국민·기업·단체...	https://s3.ap-northeast-2.amazonaws.com/media....	['공모전', '봉사']
3	K-water 2020년 대국민 환경 빅데이터 공모전	k-water	2020-08-31	2020년 K-water 대국민 환경 빅데이터 공모전 > 담당자 전화번호 :042-6...	https://s3.ap-northeast-2.amazonaws.com/media....	['공모전', 'IT', '과학']
4	기획재정부 국민참여예산 대학생 공모전	기획재정부	2020-08-31	[기획재정부] 국민참여예산 대학생 공모전 (~8/31) ■ 신청기간 2020년 6월 2...	https://s3.ap-northeast-2.amazonaws.com/media....	['공모전', '사회']

참고: 2-1. 고파스 크롤링.ipynb / 2-2. 링커리어+쿠카이브 크롤링.ipynb / 2-3. 경력개발센터 크롤링.ipynb

본래 유저들이 홍보물을 자발적으로 등록하는 방식으로 시스템이 운영되지만, 초기 알고리즘 구축 및 앱 세팅을 위해 외부에서 게시된 홍보물을 임시로 가져오는 것을 택함

➔ 따라서 데이터 수집은 오로지 알고리즘 학습용이며, 유저가입 이후 따로 데이터를 수집할 필요 없음

홍보물 관리시스템 구축 _ 추천 알고리즘 구축

▶ 홍보물 데이터 전처리

데이터 내 결측치 제거
특수문자 제거
불용어(Stopwords) 제거

1단계 Data Cleansing

한글 형태소 분석기인
KoNLPy의 Komoran을
활용해
Tokenizing /
Stemming /
Normalizing /
Lemmatization 진행

2단계 Tokenizing

Scikit-Learn. TfidfVectorizer
를 활용해
데이터 토큰을 TFIDF 벡터로 변환

*TFIDF : 문서 내 frequency와
rarity를 모두 고려하여
각 단어에 가중치 부여하는 알고리즘

3단계 Embedding

홍보물 관리시스템 구축 _ 추천 알고리즘 구축

▶ 유저 데이터 수집

설문조사를 통해 유저정보 & 해당 유저의 홍보물에 대한 관심도를 수집하여 조회테이블과 유저테이블을 생성

조회 테이블

조회 ID, 유저ID, 홍보물 ID, 조회수(=관심도에 비례)

	lookupid	userid	articleid	viewed
0	0	0	487	2
1	1	0	665	2
2	2	0	97	2
3	3	0	590	2
4	4	0	810	2

유저 테이블

유저 ID, 학년, 성별, 소속대학, 유저별 총 조회수

	userid	grade	college	sex	viewed
0	0	4학년	사범대학	M	12
1	1	2학년	경영대학	F	30
2	2	4학년	생명과학대학	M	26
3	3	1학년	국제학부	F	33
4	4	2학년	생명과학대학	M	8

홍보물 관리시스템 구축 _ 추천 알고리즘 구축

▶ Simple Recommendation

1. 유저의 학년, 성별, 소속대학을 기준으로 홍보물 카테고리를 mapping

```
{ '1학년' : [ '동아리', '친목', '연합' ],
  '2학년' : [ '동아리', '대외활동', '봉사' ],
  '3학년' : [ '대외활동', '봉사' ],
  '4학년' : [ '학회', '스터디', '공모전', '취업' ] }
```

```
{ '정보대학' : [ 'IT' ],
  '경영대학' : [ '경영', '대외활동' ],
  '디자인조형학부' : [ '예술' ],
  '미디어학부' : [ '경영', '대외활동', '예술', '음악' ], }
```

```
{ 'F' : [ '음악', '예술', '언어' ],
  'M' : [ '운동', '음악', 'IT' ] }
```

2. 추천 시스템이 이용자 ID를 입력으로 받으면, 사용자 정보(성별, 소속대학, 학년)에 매핑

→ 매핑된 사용자 정보에 속하는 홍보물 추출

3. 조회수(인기도) 별로 정렬

- 멜론 TOP100, 교보문고 베스트 셀러와 같이 “대다수 학생들이 조회한 홍보물은 나도 좋아할 것이다”라는 개념을 기반

추출 결과)

<학년:1학년, 성별:M, 단과대학:경영대학> 인 당신에게 어울리는 홍보물을 조회순으로 나열
1 Young Tigers 2 엔더스 3 코기토 4 예술비평연구회 5 고란도란 ...

→ 실제로 남자에게 인기가 많은 운동, 음악 동아리가 다수 추천됨을 알 수 있음

홍보물 관리시스템 구축 _ 추천 알고리즘 구축

▶ Content-based Filtering Recommendation

1. 데이터 전처리를 통해 추출한 게시물 TFIDF Matrix를 활용해 홍보물 간의 유사도를 측정 (Cosine Similarity 사용)

```
pd.Series(cosine_sim[20]).sort_values(ascending=False)[1:]
```

```
121    0.180002
174    0.147396
33     0.146640
110    0.139188
114    0.132134
...
```



“20번 홍보물과 가장 유사한 홍보물의 cosine similarity 출력”

2. 추천 시스템이 이용자 ID를 입력으로 받으면, 사용자가 조회했던 홍보물에 해당하는 코사인 유사도 를 추출
3. 홍보물과 유사도가 높은 순으로 정렬

추출 결과)

당신이 조회했던 홍보물과 유사한 홍보물을 추천

```
109          명상, 어떻게 할까? 밥 먹기보다도 쉽다!
39          포근하고 편안한 원불교학생회로 오세요! (누구나 환영!)
86          “편안해지다”
50          [원불교퀘스트] 신앙은 은혜! 사은(四恩)에 대해 알아보시다
27          명상? mindfulness? 어떻게 할까?
```

➔ 실제로 해당 사용자는 종교 동아리의 홍보물을 다수 조회하였으며, 조회했던 홍보물과 관련된 홍보물이 다수 추천되는 것을 알 수 있음

홍보물 관리시스템 구축 _ 추천 알고리즘 구축

▶ Collaborative Filtering Recommendation

1. 조회테이블을 기반으로 user-item 간 **pivot table**을 생성

→ 집계할 값=조회수, 행=유저, 열=홍보물로 구성된 행렬

articleid	2	10	15	19	26	28	31	46	56	57	...	853	854	860	862	864	867	869	870	872	873
userid																					
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	3.0	3.0	0.0	0.0	0.0	0.0	0.0

(0 : 미 조회)

2. Matrix Factorization의 종류인 **SVD(Singular Vector Decomposition)** 을 이용해 Pivot table 분해

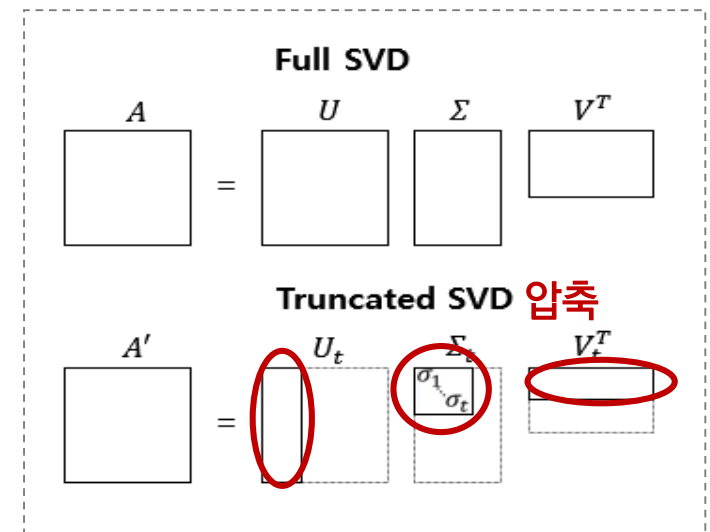
- (유저 x 홍보물) = (유저 x 잠재요인) x (잠재요인 x 홍보물)
= 유저 x (잠재요인 x 잠재요인) x 홍보물

- 총 3가지 행렬이 추출됨

→ (우측 이미지 참고) U, V : 분해 행렬 / Σ : 잠재요인 대각행렬

3. 잠재요인 행렬의 **대각원소 중 상위 원소만 추출** (데이터 압축, Truncated SVD)

- 상위 원소일수록 데이터의 대부분을 설명함 → 따라서 데이터가 압축되어도 원 데이터의 설명력을 보존



홍보물 관리시스템 구축 _ 추천 알고리즘 구축

▶ Collaborative Filtering Recommendation

4. 앞서 구한 SVD로 분해된 행렬들의 축소(압축, Truncated)버전을 다시 곱 함 → **user-item 간의 예측 조회수 계산**

- 원 pivot table 행렬과 값이 다름 (원 데이터를 압축하여, 일부분만 사용했기 때문)

원 행렬													복구된 원 행렬 (예측치)							
articleid	2	10	15	19	26	28	31	46	56	57	...	853	0	1	2	3	4	5	6	7
userid													0	1	2	3	4	5	6	7
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	-6.956458e-03	2.145354e-03	2.145354e-03	-8.257389e-03	-6.956458e-03	-2.149288e-02	-2.149288e-02	-1.962897e-02
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	-5.027598e-03	1.091612e-03	1.091612e-03	-4.714519e-03	-5.027598e-03	-1.401978e-02	-1.401978e-02	-8.572646e-03
2	0.0	0.0	0.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	...	0.0	-1.754851e-03	5.290506e-04	5.290506e-04	-6.142286e-03	-1.754851e-03	1.985429e+00	1.985429e+00	1.843759e-03
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	-1.317898e-03	2.915258e-04	2.915258e-04	-1.098941e-03	-1.317898e-03	1.641643e-03	1.641643e-03	-4.067049e-04

5. 추천 시스템이 이용자 ID를 입력으로 받으면, userID 별로 SVD 예측값을 내림차순으로 정렬

6. 상위 예측값의 인덱스 (홍보물 ID)중 이전에 조회하지 않은 것만 추출

추출 결과) what you have seen : ['쿠리에이터', '엔더스', '고대농악대', ...'KU HOPE', 'EBS on Campus']
 Recommended for you : ['Young Tigers', 'KULSOM', 'KUHS', ...'와일드아이즈', 'KUBT', 'Talk Through Piano']

→ {{해당 유저가 조회한 홍보물}을 조회한 다른 유저}가 조회한 홍보물을 추천

홍보물 관리시스템 구축 _ 추천 알고리즘 구축

▶ Hybrid Recommendation

앞서 세 가지의 추천 알고리즘을 종합한 알고리즘 구축

➔ 목적 : 유저에 따라 **최적의 유저 맞춤형 추천시스템** 제공

→ 유저의 상황을 3단계의 USER PHASE로 분류, 각 단계마다 다른 하이브리드 추천 알고리즘 적용

USER PHASE1) 신입 이용자 (조회수 10 미만) ← Simple Recommendation

USER PHASE2) 중간 이용자 (조회수 10 이상) ← $\text{Simple} \times 0.1 + \text{Content-based} \times 0.6 + \text{Collaborative} \times 0.3$

USER PHASE3) 구 이용자 (조회수 20 이상) ← $\text{Simple} \times 0.1 + \text{Content-based} \times 0.4 + \text{Collaborative} \times 0.5$



이는 가장 널리 통용되는 하이브리드 추천시스템의 가중치를 사용하였으며,
추후 충분한 유저 데이터와 조회 데이터가 축적된 뒤 회귀분석을 통해 가중치를 추정할 예정

홍보물 관리시스템 구축 _ 태그 제안 모델 구축

사용자가 등록한 홍보물의 카테고리를 제안하는 Tag Predictor 모델을 구축

홍보물 제목 : [KU_BIG] 고려대학교 빅데이터 연구회 KU-BIG에서 12기 신입부원을 모집합니다!

Host : KU-BIG

홍보물 내용 : KUBIG은 고려대 내 유일한 데이터 분석 학회로 빅데이터 시대에 필요한 통계학, 머신러닝, 딥러닝 등에 대한 지식을 함께 공부하고 실제 데이터에 적용하는 프로젝트를 진행하는 학회입니다. 학회 활동 내용으로는...

예측

'학회', '스터디', 'IT', '취업'

- App 내에서 사용자는 제안된 카테고리를 수락하거나, 수동으로 선택할 수도 있음
 - ➔ 홍보물 카테고리를 추천함으로써 사용자는 일일이 카테고리를 선택하지 않아도 됨
- 홍보물의 카테고리는 추후 Simple Recommendation, 오프라인 게시 장소 제안에 사용됨

사용 모델 : LSTM

홍보물 관리시스템 구축 _ 태그 제안 모델 구축

LSTM (Long Short Term Memory)이란?

- 텍스트, 시계열 데이터와 같이 **순서가 중요한 데이터의 분석**에 유용한 **딥러닝 모델**
- 텍스트 문서의 Tag Prediction에 자주 쓰임
- Objective : 홍보물의 내용(Input, 2D tensor)이 주어졌을 때, 해당 홍보물의 태그(Target) 예측하는 모델 학습



학습과정

Data Preprocessing

- 홍보물 내용을 X(설명변수), 태그를 Y(반응변수)로 Set
- 홍보물 내용 Tokenize / Normalization
- 서로 다른 길이의 입력(홍보물 내용)을 고정된 길이의 벡터로 변환
→ 최종 Input : (홍보물 개수 x 고정길이) 형태의 2D Matrix
- Target Data의 경우 One-hot Encoding 진행
- Data Holdout => Train, Test Data Split

Model Train

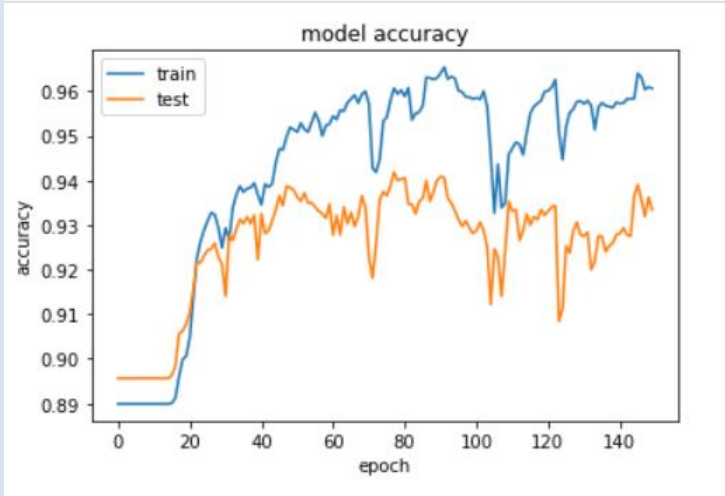
- Tensorflow.Keras.LSTM 사용
- [Embedding Layer + LSTM Layer + Fully Connected Layer]로 구성된 모델 Build
- Train Data에 대해 모델 fit
- Hyperparameter Tuning을 통해 적절한 초모수 탐색

Model Test

- 학습한 모델을 Test Data 에 적용시켜 Overfitting 여부 확인
- 모델 구조 & weight를 별도의 파일에 저장
- 추후 앱 구축 시 해당 모델 탑재

홍보물 관리시스템 구축 _ 태그 제안 모델 구축

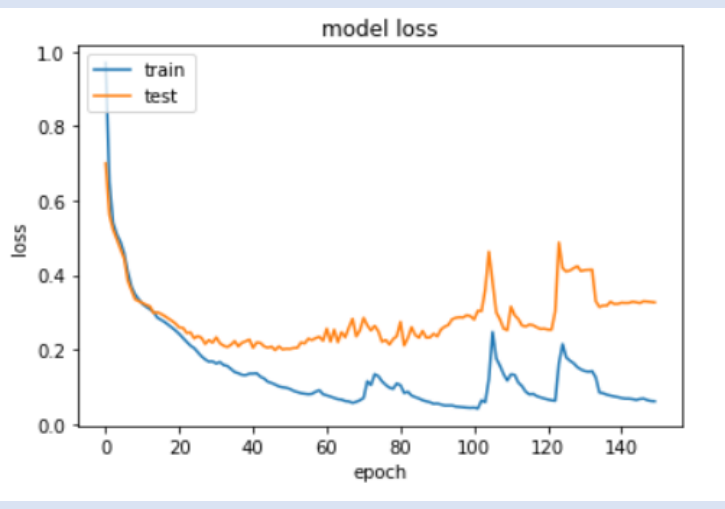
Model Accuracy



LSTM Model Analysis

1. Train Data model Accuracy: 0.96
 2. Test Data model Accuracy: 0.93
- 93%의 model accuracy로, 홍보물의 태그 예측 정확도가 높음을 확인

Model Loss



LSTM 적용 결과

1. REAL_TAGS: 실제 홍보물의 태그
 2. PRED_1: LSTM이 첫 번째로 예측한 홍보물의 태그
 3. PRED_2: LSTM이 두 번째로 예측한 홍보물의 태그
- 홍보물의 실제 태그를 높은 확률로 예측했음을 확인

REAL_TAGS	PRED_1	PRED_2
[운동, 동아리]	[동아리]	[운동]
[봉사, 미술, 대외활동]	[대외활동]	[동아리]
[동아리, 친목, 음악]	[공모전]	[언어]
[취업]	[취업]	[봉사]
[공모전, 경영, 미술]	[공모전]	[예술]
[운동, 동아리]	[동아리]	[사회]
[미술, 대외활동]	[공모전]	[IT]
[경영, 대외활동]	[동아리]	[공모전]
[공모전, 미술]	[대외활동]	[경영]
[경영, 취업, 스터디, 학회]	[스터디]	[과학]
[취업]	[취업]	[미술]
[사회, 스터디, 학회]	[대외활동]	[대외활동]
[경영, 미술, 대외활동]	[사회]	[기타]
[공모전]	[공모전]	[운동]
[공모전, 경영, 미술]	[공모전]	[동아리]
[동아리, 친목, 음악]	[동아리]	[언어]
[동아리, 언어, 스터디, 연합]	[동아리]	[스터디]
[운동, 동아리, 친목]	[동아리]	[연합]

홍보물 관리시스템 구축 _ 앱 구현

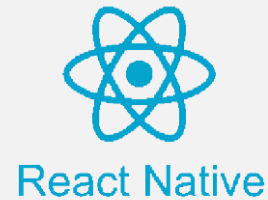
최종 목표 :

학생들이 홍보물을 편리하게 관리하고 조회할 수 있는

Mobile Application 생성



〈Front-End〉:



〈Back-End〉:



Flask



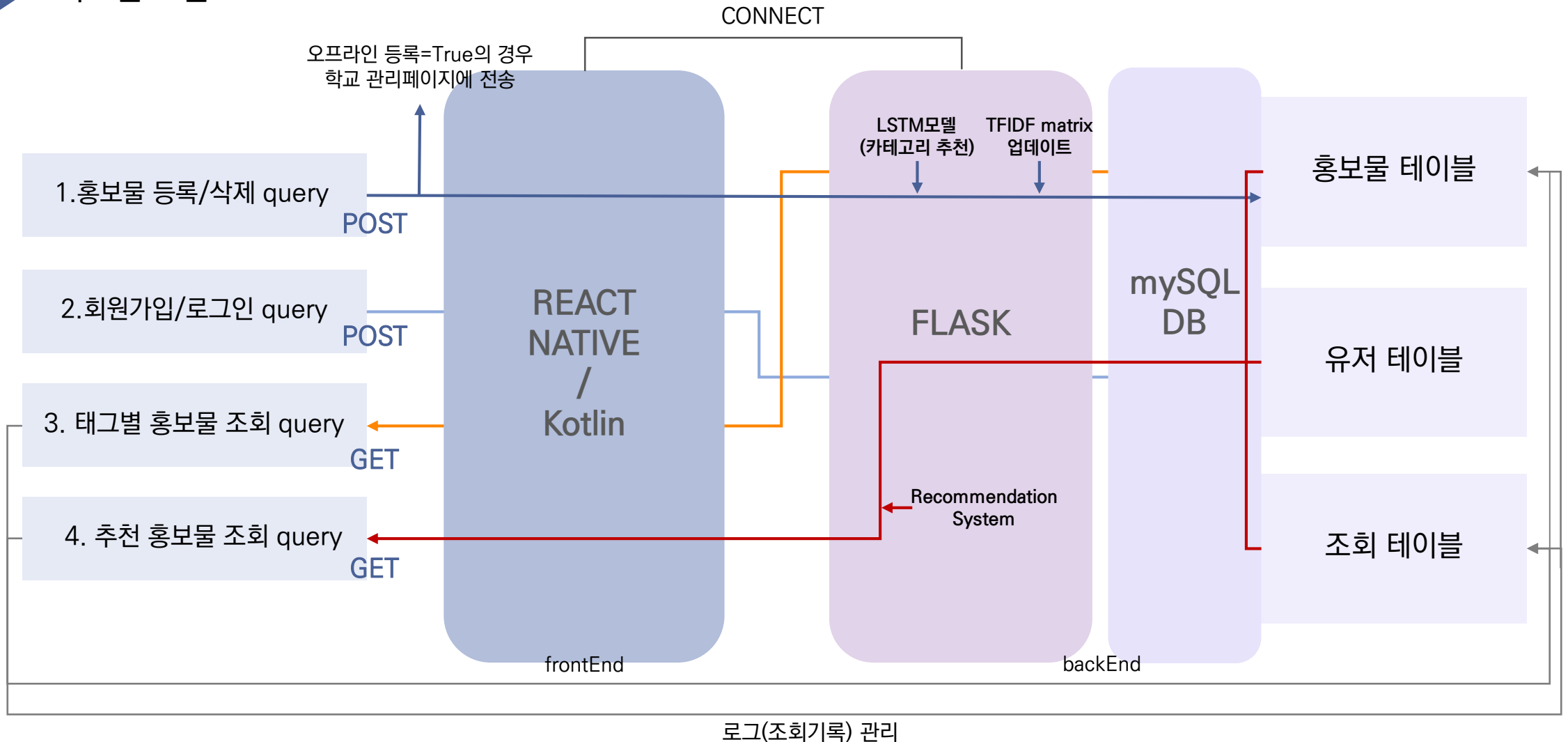
REST API형태로
서비스 제공



Azure DB for
MySQL

홍보물 관리시스템 구축 _ 앱 구현

▶ 시스템 흐름도



홍보물 관리시스템 구축 _ 앱 구현

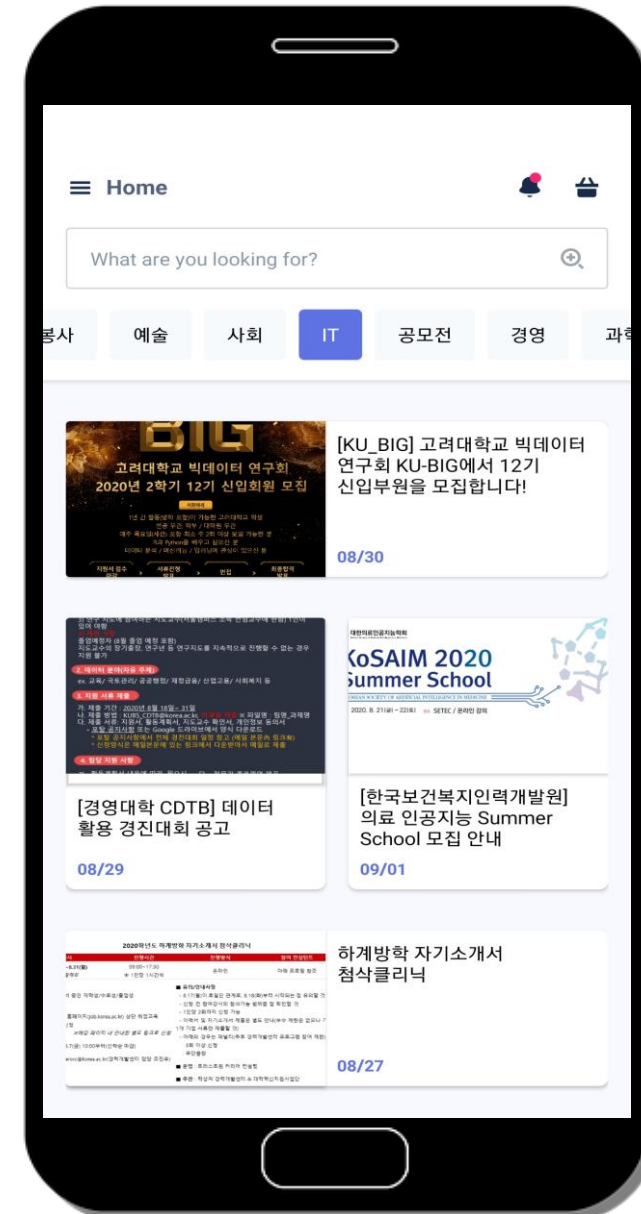
▶ Database E-R modeling



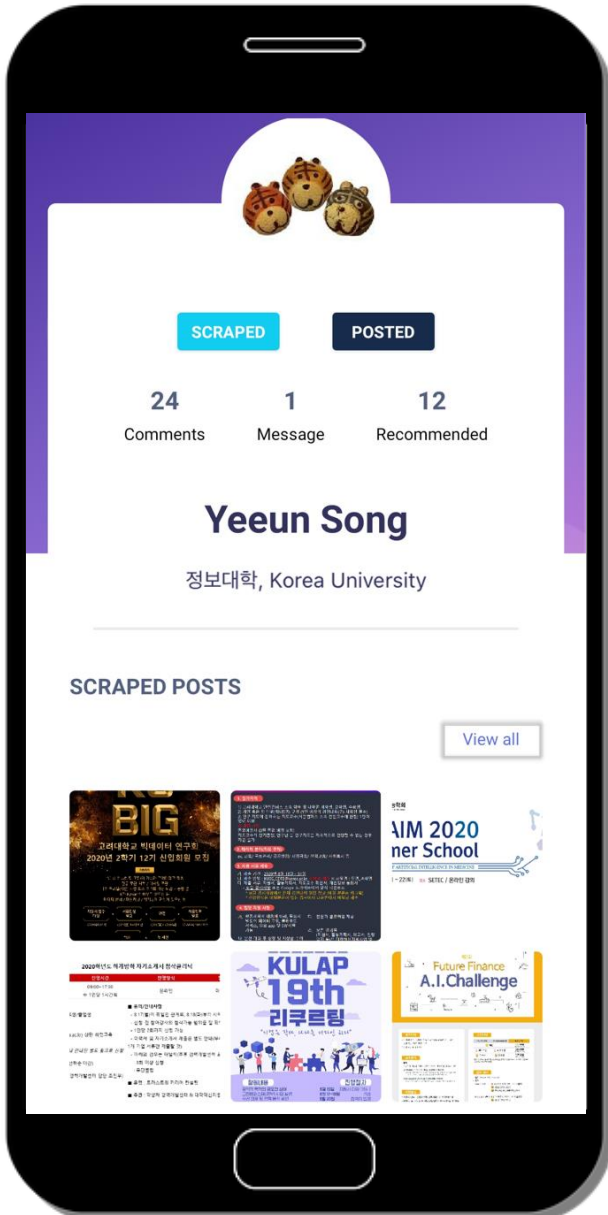
홍보물 관리시스템 구축 _ 앱 구현

Home Screen

- Hybrid Recommendation Algorithm을 기반으로 사용자의 조회 이력/프로필 정보를 사용해 **사용자 맞춤 홍보물** 표시
- 키워드를 중심으로 관심 있는 분야 검색 가능
 - 위 예시는 4학년/정보대학/여학생인 유저가 {IT}를 키워드로 검색하였을 때 필터링 된 추천 홍보물



홍보물 관리시스템 구축 _ 앱 구현

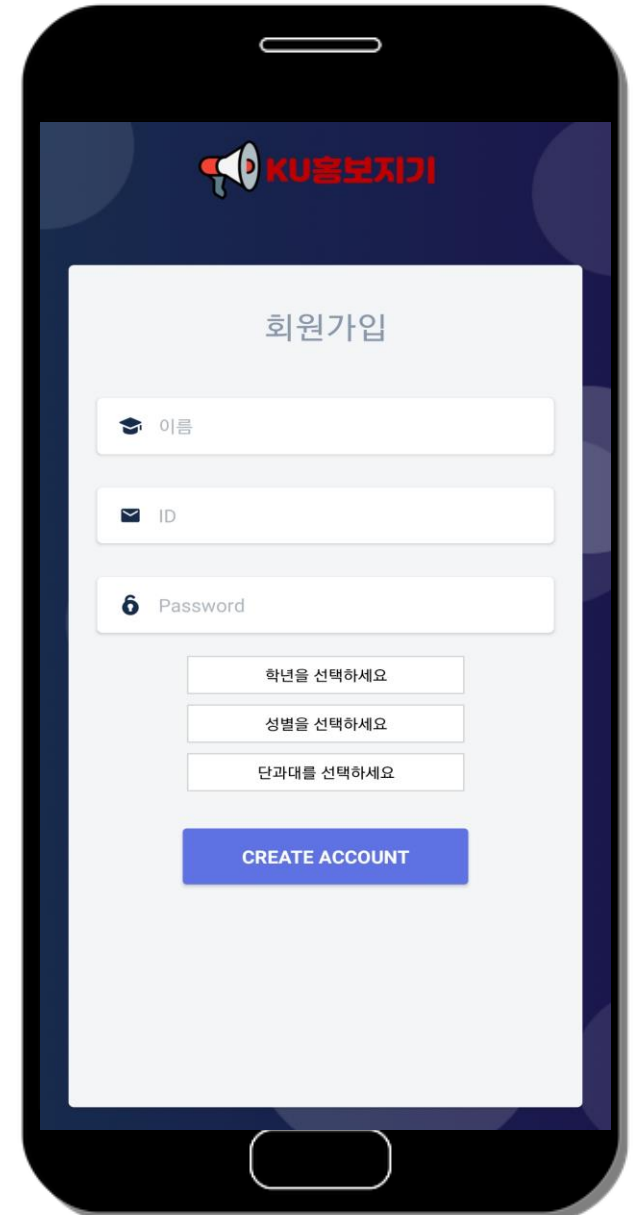


Profile Screen

- 앱에서의 나의 활동 내역 확인
- 스크랩한 홍보물을 모아서 볼 수 있음 (SCRAPED POSTS)
- 게시한 홍보물 또한 확인 가능

Register Screen

- Creating Accounts
- 유저 맞춤형 추천시스템을 구현하기 위한 {학년/성별/단과대} 가입 버튼 존재
- 특정 단과대에서 많이 추천되는 홍보물이 해당건물 근처 스마트 배너에서 홍보 빈도가 높아지도록



홍보물 관리시스템 구축 _ 앱 구현

Upload Screen

- 홍보 주최가 홍보물을 모바일로 업로드 가능
- “학교 게시판에 전송하기” 체크 시 홍보물 인쇄 없이 교내에 홍보물 게시 가능

Login Screen

- 앱을 사용하기 위해 로그인

Contents

배경



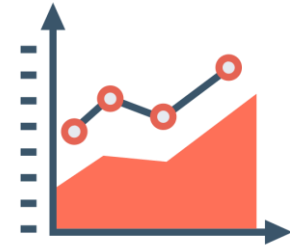
스마트 배너
입지 선정



홍보물 관리
시스템 구축



기대효과



기대효과

돈, 시간 낭비 down



1. 홍보 비용 절감
 - 다수의 홍보물 인쇄 제작 필요 x
2. 홍보물 게시 허가 절차의 자동화
 - 관리의 편의성을 향상시키고, 현장 허가의 시간 비용을 절감시킴으로써 홍보 주체들의 편의성을 보장

효과적인 홍보



1. Application을 통한 이용자 맞춤 홍보물 추천
2. 홍보물 내용을 통해 교내 적절한 홍보 입지 추천
 - 관심 있는 분야의 홍보물을 지나칠 가능성을 줄일 수 있음

깔끔한 교내 미관



1. 인쇄 홍보물의 필요성이 사라지며 무분별한 홍보물 부착으로 인한 미관 침해 X

기대효과

쉬운 관리



1. 게시 유효기간이 지난 홍보물의 철거 자동화
2. 자동화 앱의 체계적 관리를 통해 수많은 온/오프라인 홍보물의 관리 용이
3. 수작업이 필요했던 절차들의 자동화 → **스마트 캠퍼스**로의 도약이 가능

참여형 어플리케이션



1. “모두가 만들어가는 어플리케이션”
 - 홍보 주최측은 어플리케이션을 통해 손쉬운 게시 가능
 - 홍보 주최자와 홍보물의 수요자의 니즈를 모두 충족
2. 유저가 **참여할수록 추천 알고리즘의 질이 향상**되어 더 풍부한 콘텐츠 제공

지속 가능성



1. 새로운 홍보물의 꾸준한 업데이트에도 지속 가능한 Dynamic Recommendation System
2. **일회성이 아닌 지속적인** 추천관리시스템
 - 유저의 성향, 검색 내용 등에 따라 꾸준하게 유저에게 fitting & update 되는 추천시스템

참고문헌 및 사용 데이터

참고문헌

〈Kim Falk〉〈Practical Recommender Systems〉, Kim Falk, Manning Publications

Creating a mobile chat app from scratch using React Native and Flask
<https://www.youtube.com/watch?v=AlbmVWWTv4I>

추천시스템 Content Based Filtering(CBF) Python 기반 구현
<https://simonezz.tistory.com/19>

RESTful에 대해서 설명해주세요.(REST, RESTful, RESTful API 개념 정리)
<https://jeong-pro.tistory.com/180>

React Native Official Documentation
<https://reactnative.dev/docs/getting-started>

Android Studio 가상 기기 만들기 및 관리
<https://developer.android.com/studio/workflow>

Keras for Multi-Label Text Classification
<https://medium.com/towards-artificial-intelligence/keras-for-multi-label-text-classification-86d194311d0e>

Keras API Reference/Layers API/Recurrent Layers/LSTM Layer
https://keras.io/api/layers/recurrent_layers/lstm/

활용데이터

고파스 동아리 게시판 홍보물
 Linkcareer 대외활동, 공모전 게시판 홍보물
 KuChive 홍보물

고려대학교 경력개발센터 취업교육 게시판 홍보물
 2019년도 강의시간 및 강의실별 학생 데이터

사용 도구

