

축제 추천은?

大大

## TEAM MEMBER

한준희

하요한

정연종

이승윤

강진용

프로젝트 Github url :

<https://github.com/Ha-Yo-Han/metabuild-mini-project>

프로젝트 Notion url :

<https://www.notion.so/Notion-2a2e8c5a8aa180568a8adef806e31bee>

# 목차

## 1.프로젝트 소개

프로젝트 소개 .....	1
개발배경 및 필요성 .....	2
개발목표 .....	3
팀원소개 및 역할 분담 ..	4
개발 환경 .....	5
프로젝트 개발 일정 .....	7
시스템 구성 개요.....	8
데이터 흐름 .....	9
DB 설계 .....	10

## 4.웹페이지 시연

웹페이지 시연	
---------	--

## 2.머신러닝

데이터 수집 .....	11
데이터 정제 및 통합 ...	12
태그 체계 설계.....	13
단어장 .....	14
태그 예측 모델 개요....	15
모델 성능 요약.....	16
태그별 성능 및 예시....	17
추천 로직 개요.....	18
태그 · 집단 참여도 ....	19

## 5.프로젝트 마무리

프로젝트의 한계점 ....	22
향후 보완점.....	23
프로젝트 회고.....	24
참고 문헌 .....	25

## 3.웹페이지 설계

서비스 한 눈에 보기.....	11
페이지 구성 .....	12

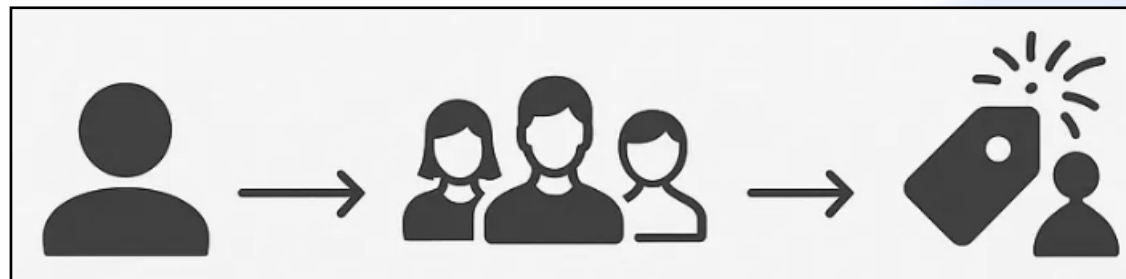
# 프로젝트 소개

축제 추천은?  
**大大**

**성별·연령 기호를 반영한  
태그 기반 전국 축제 추천 웹 서비스**

성별·연령 통계와 축제 태그를 결합해, 전국 축제를 지도에서 탐색하고 추천받을 수 있는 웹 서비스

**프로젝트 기간**: 2025.11.17- 2025.11.28



**사용자**

**통계**

**축제 추천**

# 개발배경 및 필요성

## 여행·문화 소비 트렌드 변화:

과거: **유명한 축제·관광지** 중심 이동

현재: **나이·생활패턴·취향**에 맞는 개인화된 경험 선호

## 기존 서비스의 한계:

회원의 성별·연령 등 인구통계 정보를 활용한 추천 부재

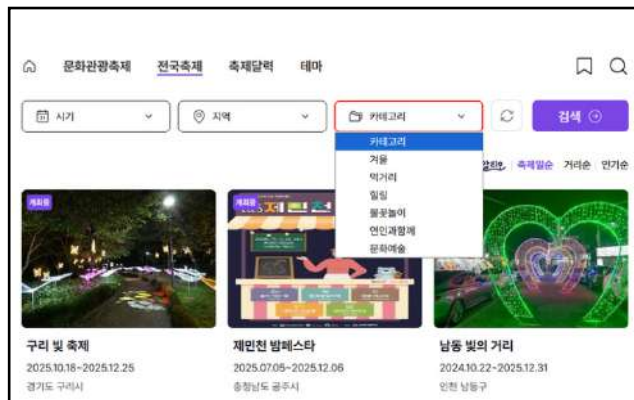
**AI 기반 추천**과 연계된 사례는 부족

## 기존 축제 정보 서비스 특징:

지역, 기간 필터 등 **기본적인 탐색 기능** 위주  
추천보다는 **정보 나열** 중심

## 프로젝트 필요성:

인구통계 정보와, 축제의 컨셉을 표현하는 태그 정보를 고려한 추천 필요  
**나와 비슷한 사람들이 선호할 축제**를 찾아주는 서비스



기존



제안

# 개발목표

## 표준화된 태그 체계 구축:

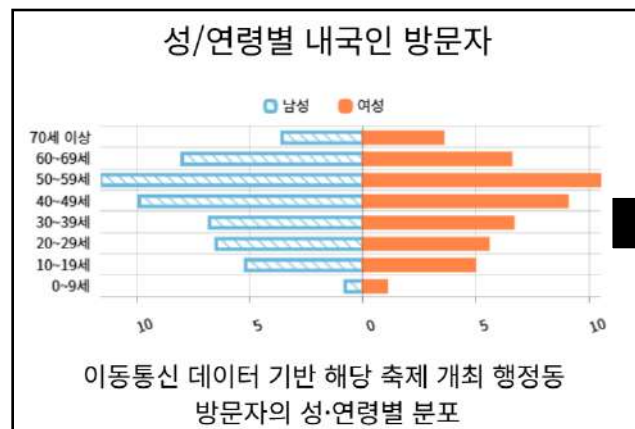
7개 태그로 축제 성격을 표현.  
텍스트 분석 + 머신러닝 활용, 축제별 핵심 태그를 자동 부여.

## 집단 특성 + 회원 선호 태그 기반 추천 설계:

방문 통계와 해당 축제의 태그를 결합하여 집단별 태그 방문 경향 지표 산출.  
사용자의 선호 태그 정보와 결합해, 그 회원이 속한 집단에서 방문도가 높고, 회원의 태그와도 잘 맞는 축제를 우선 추천.

## 추천 중심 UI 설계:

지도·캘린더·축제 리스트를 결합해 전국 축제를 한 화면에서 탐색.  
필터(지역, 기간, 태그) + AI 추천 리스트: 조건에 맞는 축제를 좁혀 나갈 수 있는 구조 지향



예시로 회원가입 시,  
체험  
선호 태그 선택하면...



# 팀원소개 및 역할 분담

## 하요한

- 프로젝트 총괄
- 공공데이터 및 외부 데이터 수집
- 데이터 전처리(ML)
- 머신러닝 모델링 및 성능 검증
- 통계 데이터 기반 지표 설계 및 산출
- 축제 태그 데이터 라벨링 및 추천 점수 로직 구현
- 풀스택 개발
  - 지도 페이지 및 축제 상세 페이지 구현
- 프로젝트 문서화

## 정연종

- 데이터 전처리(DB)
- DB 스키마 정의 및 정규화
- 웹 디자인
  - 화면 레이아웃, UI·UX 설계
- 풀스택 개발
  - 회원 가입, 이메일 인증, 로그인 기능 구현
- 사용자 관점 UX 점검
- 발표용 자료 제작 및 시각화 정리

## 이승윤

- DB 스키마 정의 및 데이터 관계 수립(ERD)
- 풀스택 개발
  - 캘린더 페이지 구현
  - 관리자 페이지 구현
- 프로젝트 발표

## 강진용

- 웹 구조 설계
- 풀스택 개발
  - 마이페이지 구현
  - 관리자페이지-리뷰관리 구현
- 클라우드 배포
  - AWS EC2 환경 구축 및 배포 자동화
  - SSL 인증 적용(HTTPS 보안 연결 지원)

## 한준희

- 웹 디자인
  - 화면 레이아웃, UI·UX 설계
- 풀스택 개발
  - 축제 리스트 페이지 및 지도 연동 화면 구현
- 발표용 자료 제작 및 시각화 정리

## 공통 작업

- 웹 서비스 QA
- Git 기반 협업 및 형상 관리
- 코드 리뷰
- 요구사항 정의 및 기능 우선순위 논의



# 개발 환경

## 백엔드



JAVA



Spring



Spring Boot



JPA



Maven



Apache Tomcat

## 데이터베이스



Oracle DB

## 데이터 분석 및 머신러닝



Python



Pandas



Numpy



Scikit Learn



soynlp

## 개발툴



Intelli J



Pycharm

## 프론트엔드



JavaScript



html5



CSS3



Tailwind



Thymeleaf



카카오맵 API



Jupyter Notebook

# 개발 환경

## 인프라



AWS



EC2



EBS



VPC



Nginx

## 협업



Notion



GitHub



Git



ZeroSSL



ERDCloud

## 웹 컬러 :

#4A80FA  
#6B96FC

#FF6A6B

#5170FF  
#FF66C4

#E6E8F2

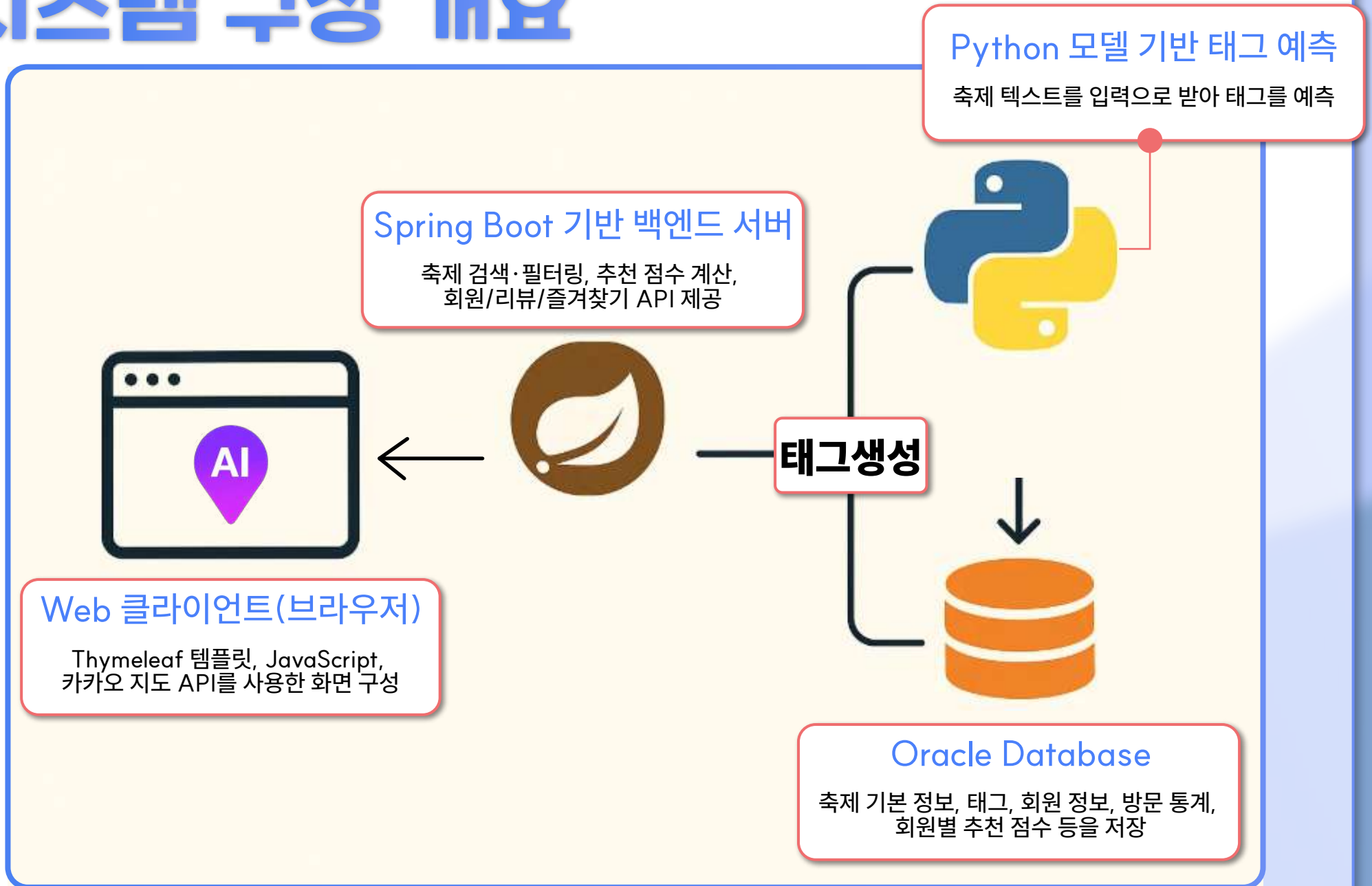
## 웹 폰트 :

**SUIT체**  
**Bagel Fat One체**



작업명	개발 일정												담당자
	17	18	19	20	21	22	23	24	25	26	27	28	
1. 계획 - 목표정의 - 계획서 작성 - 계획서 검토회의													코친눔s
2. 요구분석 - 기존 사이트 조사 - 사용자 요구 분석 - 필요 기능 정의													코친눔s
3. 설계 - 시스템 구조 설계 - 설계서 작성 - 발표자료 작성													코친눔s
4. 구현 - 기능의 구현 - 서버 및 DB 구축 - 각 부분 기능 통합													코친눔s
5. 시험 및 피드백 - 시험 및 테스트 - 매뉴얼 작성 - 최종 발표자료 작성													코친눔s

# 시스템 구성 개요



# 데이터 흐름

9

## 공공데이터 수집

축제 정보·방문 통계 등  
원천 데이터 수집



## 전처리·통합

텍스트 정제, 지역/기간 정규화,  
테이블 통합



## 태그 예측 모델 학습

라벨링 데이터로  
멀티라벨 분류 모델 학습



## 계산된 추천 결과를 API로 응답

지도/리스트/캘린더 화면에서  
추천 결과를 정렬·시각화



## 웹 요청 시 백엔드에서 추천 점수 계산

점수 계산에 필요한 정보를  
조회 후 백엔드에서  
추천 점수 로직을  
실시간으로 계산



## 태그 자동 생성 후 DB 적재

축제 소개 텍스트 >  
태그 예측 > 테이블에 저장



# DB 설계

10

password_reset_token			
id	id	Domain	INT
email	email	Domain	VARCHAR(100)
token	token	Domain	VARCHAR(200)
created_at	created_at	Domain	TIMESTAMP
expires_at	expires_at	Domain	TIMESTAMP

Review			
review_no	review_no	Domain	NUMBER
rating	rating	Domain	NUMBER(1)
content	content	Domain	VARCHAR2(1000)
created_at	created_at	Domain	DATE
updated_at	updated_at	Domain	DATE
festival_no	festival_no	Domain	NUMBER
user_no	user_no	Domain	NUMBER

Festival_detail			
festival_no	festival_no	Domain	NUMBER
infotext_1	infotext_1	Domain	CLOB
infotext_2	infotext_2	Domain	CLOB
host	host	Domain	VARCHAR2(200)
host_tel	host_tel	Domain	VARCHAR2(50)
homepage	homepage	Domain	VARCHAR2(300)
festival_fee	festival_fee	Domain	VARCHAR2(200)

Member			
user_no	user_no	Domain	NUMBER
user_id	user_id	Domain	VARCHAR2(50)
password	password	Domain	VARCHAR2(255)
name	name	Domain	VARCHAR2(50)
email	email	Domain	VARCHAR2(100)
phone	phone	Domain	VARCHAR2(20)
gender	gender	Domain	CHAR(1)
birth_date	birth_date	Domain	DATE
join_date	join_date	Domain	DATE
role	role	Domain	VARCHAR2(20)
is_active	is_active	Domain	CHAR(1)
withdraw_date	withdraw_date	Domain	DATE
favorite_tag	favorite_tag	Domain	VARCHAR2(100)

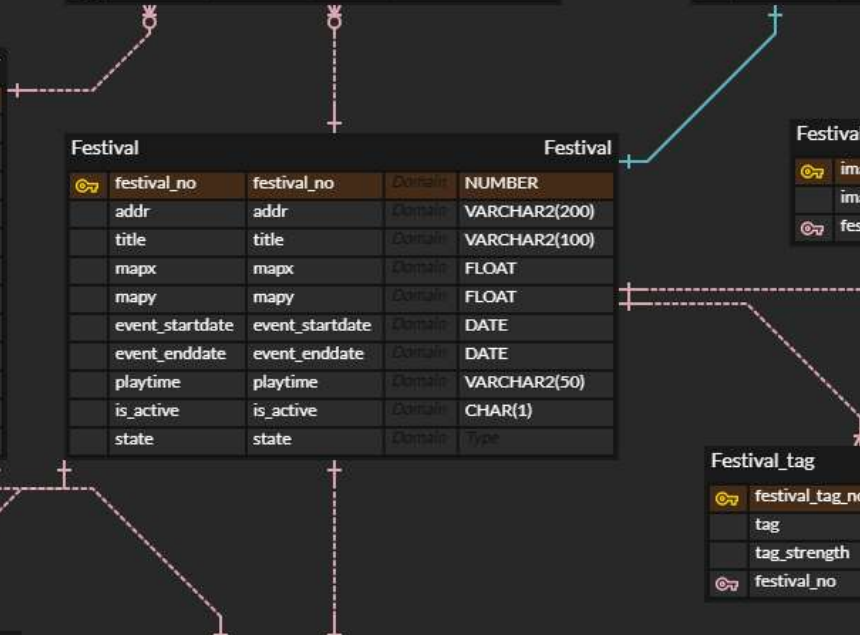
Festival			
festival_no	festival_no	Domain	NUMBER
addr	addr	Domain	VARCHAR2(200)
title	title	Domain	VARCHAR2(100)
mapx	mapx	Domain	FLOAT
mapy	mapy	Domain	FLOAT
event_startdate	event_startdate	Domain	DATE
event_enddate	event_enddate	Domain	DATE
playtime	playtime	Domain	VARCHAR2(50)
is_active	is_active	Domain	CHAR(1)
state	state	Domain	Type

Festival_image			
image_no	image_no	Domain	NUMBER
image_url	image_url	Domain	VARCHAR2(100)
festival_no	festival_no	Domain	NUMBER

Festival_tag			
festival_tag_no	festival_tag_no	Domain	NUMBER
tag	tag	Domain	VARCHAR2(30)
tag_strength	tag_strength	Domain	FLOAT
festival_no	festival_no	Domain	NUMBER

Member_AI_score			
ai_score_no	ai_score_no	Domain	NUMBER
ai_score	ai_score	Domain	FLOAT
scored_at	scored_at	Domain	DATE
festival_no	festival_no2	Domain	NUMBER
user_no	user_no2	Domain	NUMBER

Favorite			
favorite_no	favorite_no	Domain	NUMBER
favorite_date	favorite_date	Domain	DATE
user_no	user_no	Domain	NUMBER
festival_no	festival_no	Domain	NUMBER



# 머신러닝

# 데이터수집

## 축제 기본 정보

출처: 공공데이터포털

주요 항목: 축제명, 기간, 장소, 주최/주관, 요금, 소개 텍스트, 좌표 정보 등

축제 기본 정보 API  
(771개)

## 축제별 방문자 구성 통계

출처: 한국관광데이터랩(한국관광공사)

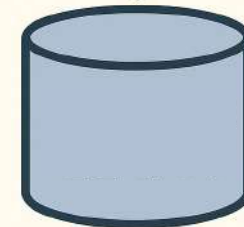
대상: 성별·연령·현지인,외지인 비율이 제공되는 일부 축제(92개)

주요 항목: 성별 비율, 연령대 비율, 현지인,외지인 비율

### 활용목적

각 축제의 태그와 결합해 집단별  
태그 방문 경향도를 추정하는 데 사용

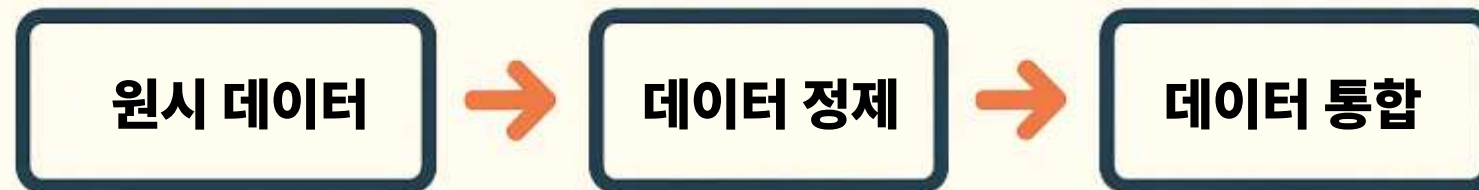
방문자 통계  
(92개)





# 데이터 정제 및 통합

## 전처리



## 통합



# 태그 체계 설계



자연

산, 바다, 강, 꽃, 생태 등  
자연 환경 중심 축제



먹거리

먹거리/특산품/푸드트럭  
중심 축제



문화

공연, 전시, 전통/예술  
중심 축제



계절

계절성(봄꽃, 단풍, 눈/  
크리스마스 등)이 강한 축제



야간

야간 조명, 불꽃놀이, 야시장  
등 밤 시간대 중심 축제



체험

참여형/체험형  
프로그램 중심 축제



아동

어린이·가족 대상  
콘텐츠 중심 축제

## 설계 원칙

- 한 축제는 최대 2개의 핵심 태그를 가질 수 있도록 제한
- 축제를 떠올렸을 때 가장 먼저 생각나는 이미지에 집중
- 서로 의미가 겹치는 태그는 최소화하고, 설명 가능하도록 유지

# 단어장

단어	태그	강도	점수
불꽃놀이	야간	강	2
벚꽃	계절	강	2
체험존	체험	강	2
야시장	먹거리	약	1
민속놀이	문화	강	2
눈썰매	아동	강	2

불꽃놀이가 열리는 야간의 다양한 먹거리를 즐길 수 있습니다.

키워드 -> 단어장

야간 -> "불꽃놀이, 야시장, 야간, 조명, 빗축제..."

먹거리 -> "먹거리, 시식, 특산품, 푸드트럭, 향토음식..."

텍스트



토크나이징

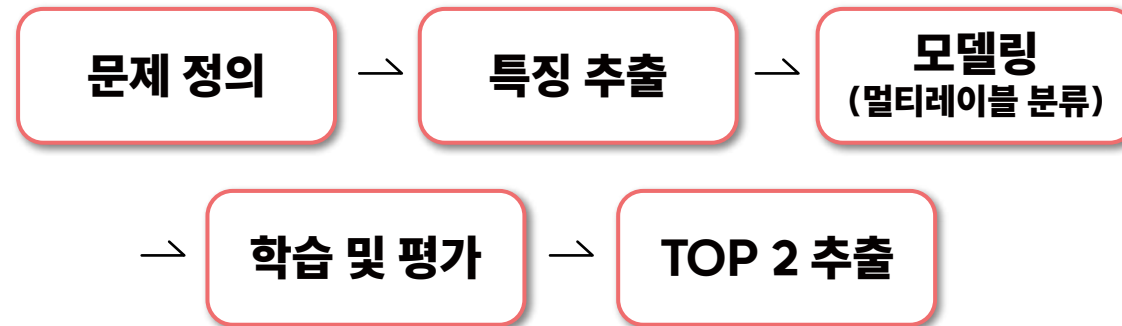


키워드매칭



TOP 2 선택

# 태그 예측 모델 개요



## 문제 정의

7개 태그에 대한 다중 라벨 예측  
각 태그가 해당 축제를 가질 확률을 산출  
상위 2개 태그를 핵심 태그로 선택

## 특징 추출

텍스트 전처리 후 TF-IDF 벡터화

## 모델 구조

로지스틱 회귀 기반 OVR 멀티레이블 분류  
태그별로 이진 분류기를 학습  
하나의 축제에 대해 태그 각각에 대한 예측

## 학습 및 평가

10-Fold 교차 검증으로 성능 평가

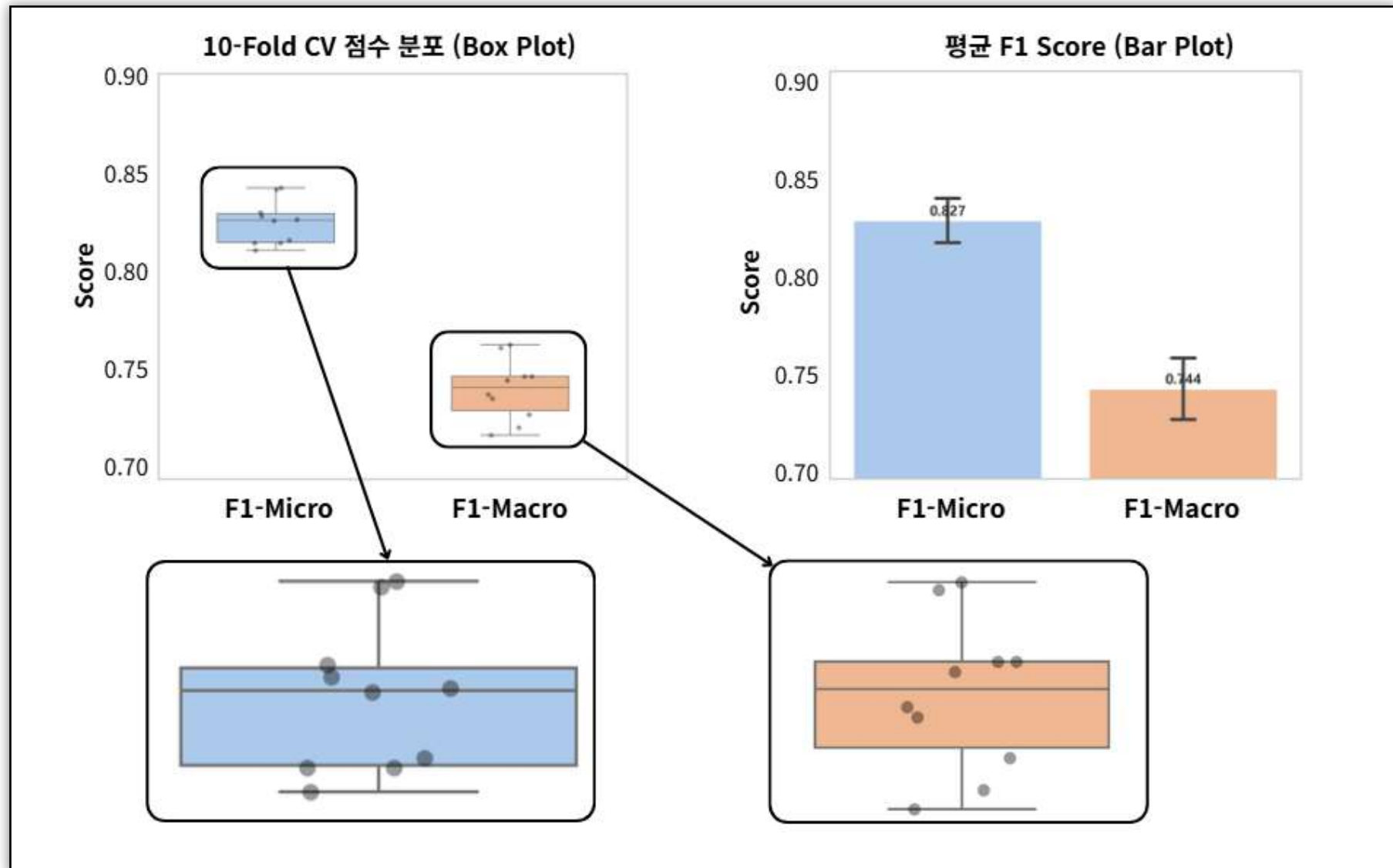
## 평가 지표

F1-micro  
F1-macro

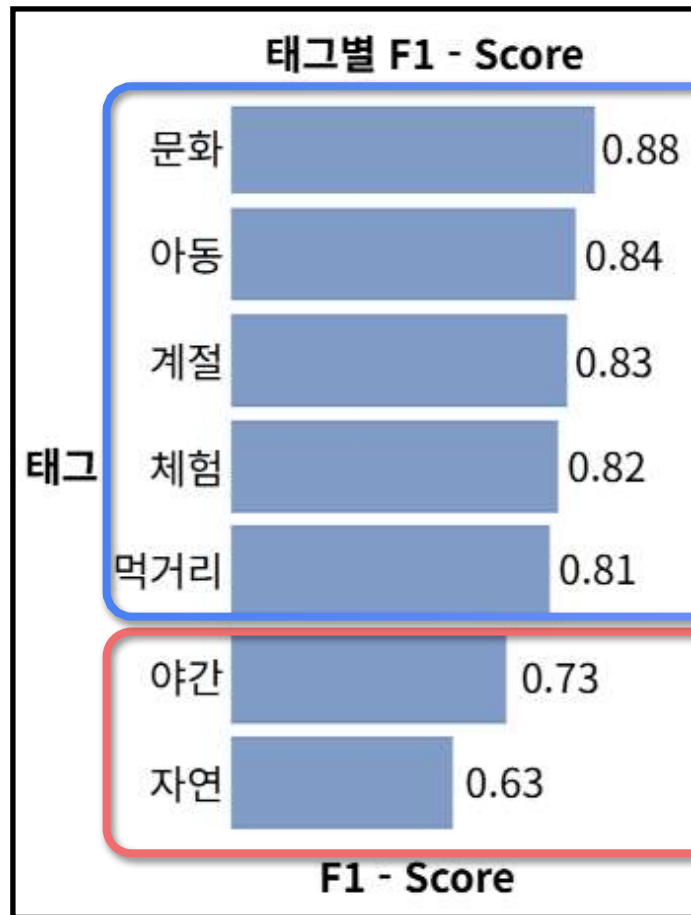
## TOP 2 추출

예측 확률이 높은 상위 2개 태그를 핵심 태그로 채택  
예측 확률을 태그의 강도로 활용

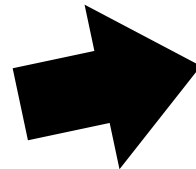
# 모델 성능 요약



# 태그별 성능 및 예시



상대적으로  
높은 SCORE



자연스러운 예측

**"제민천 밤페스타"**

야간, 먹거리(야간 먹거리 키워드 다수)

부자연스러운 예측

**"담양 대나무 축제"**

예상 라벨: 문화, 자연

모델 예측: 문화, 계절

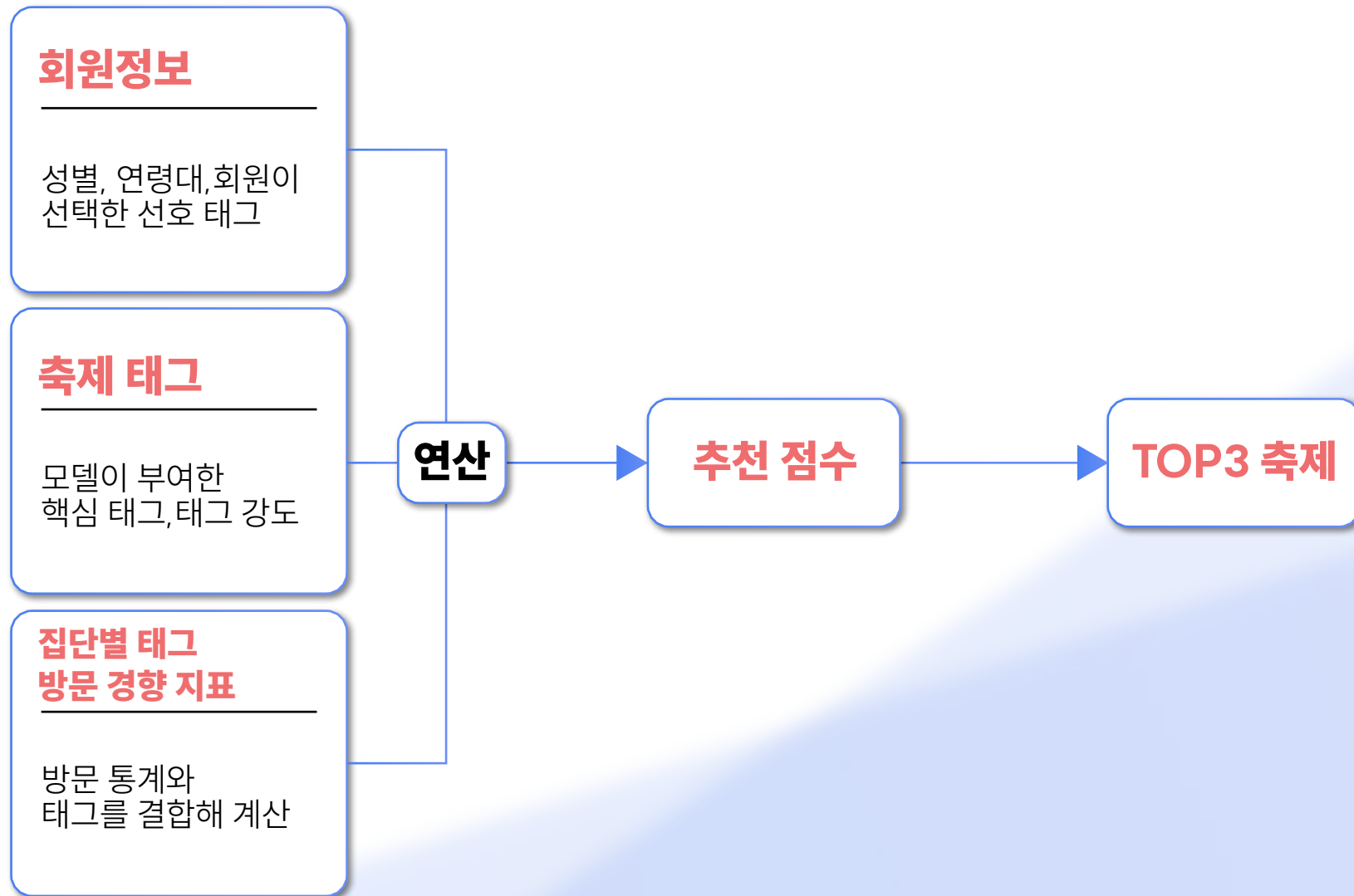
-> 자연보다는 계절 쪽으로 해석한 사례

상대적으로  
낮은 SCORE



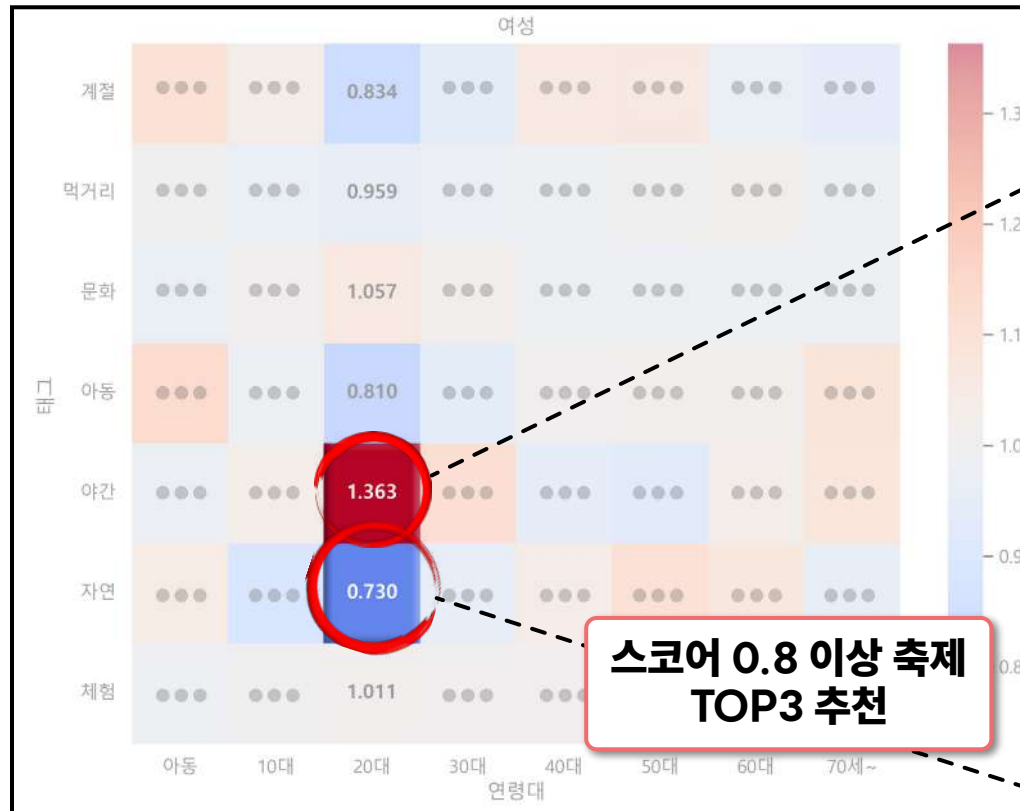


# 추천 로직 개요



# 태그 \* 집단 참여도

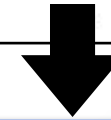
## 시스코어



### 태그 \* 집단별(성별,나이) 참여도 히트맵

### 20대 여성 추천 축제

축제명	태그명	AI 스코어
여수밤바다불꽃축제	야간, 문화	1
운정호수공원불꽃축제	야간, 문화	0.992453
한화와함께하는서울세계 불꽃축제	야간, 체험	0.990821
진주의수호자들-유등빛을 지켜라	야간, 문화	0.967039
...	...	...



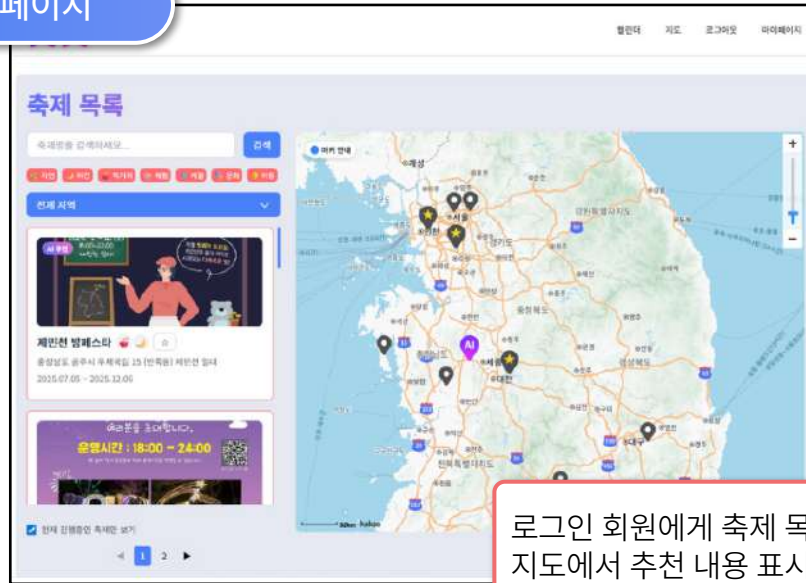
### 사용자 선호 태그 반영 결과

축제명	태그명	AI 스코어
세종낙화축제	자연, 야간	1
관악별빛산책	자연, 야간	0.995679
청도정월대보름민속한마당	자연, 야간	0.964232
만석거새빛축제	자연, 야간	0.962415
전라남도정원페스티벌	...	0.852043

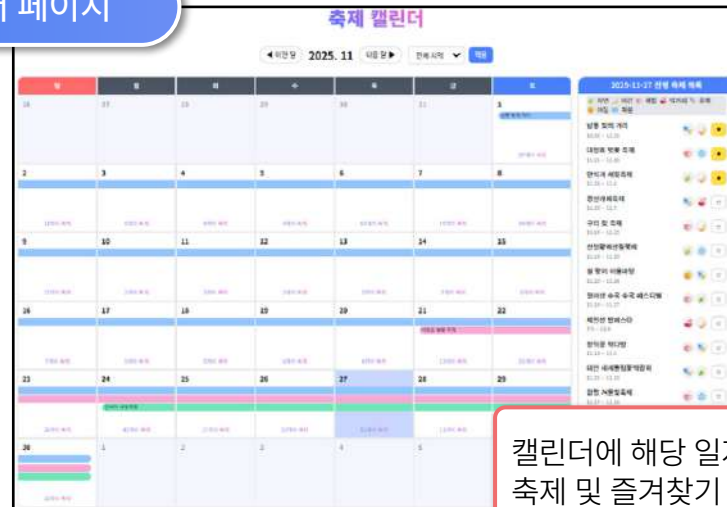
# 웹페이지 설계

# 서비스 한눈에 보기

지도 페이지



캘린더 페이지

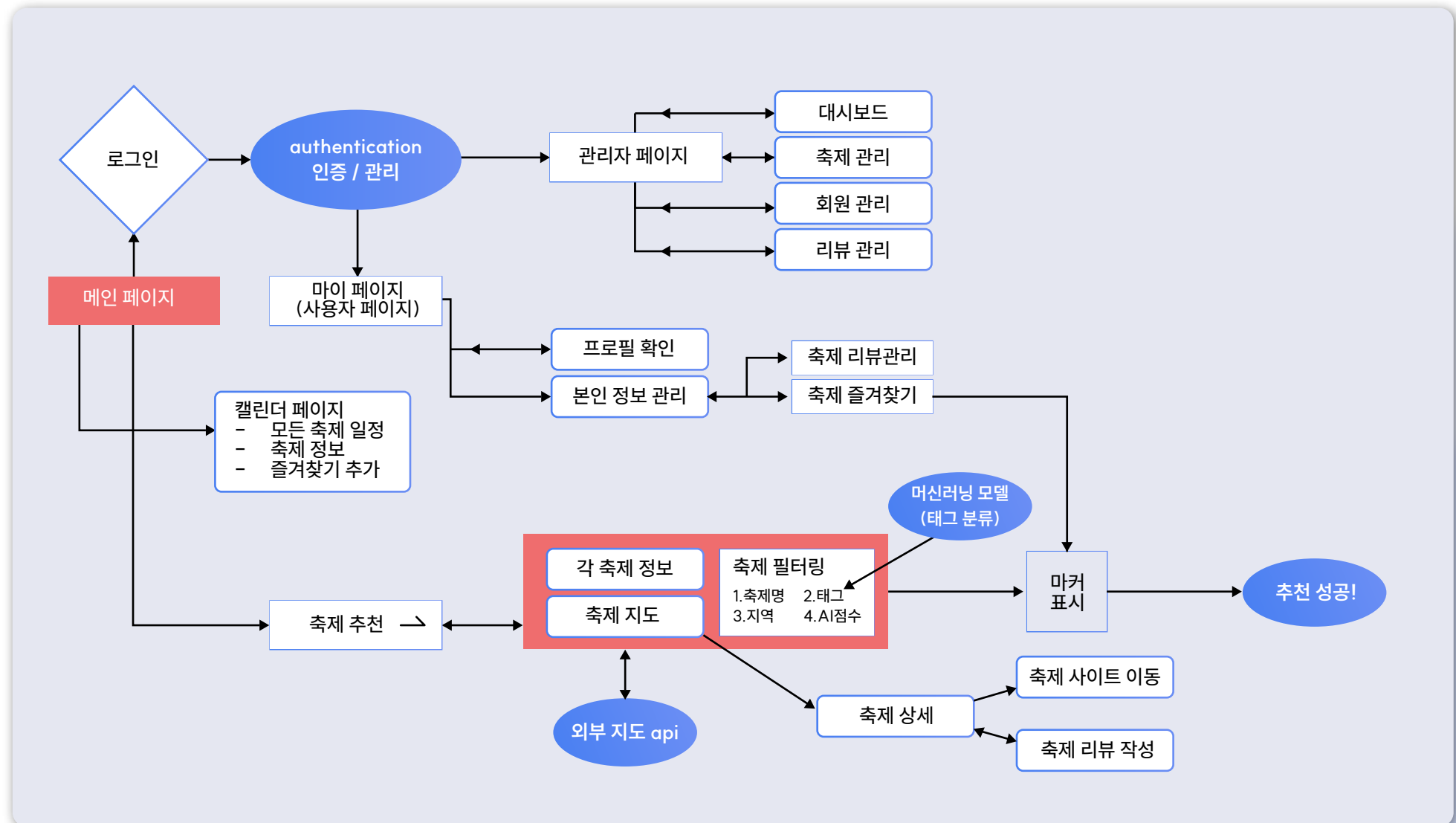


축제 상세 페이지



# 페이지 구성

→ : 데이터가 가는방향  
 ↔ : 양방향 데이터 이동



# 웹페이지 시연



# 프로젝트 마무리

# 프로젝트의 한계점

**데이터 통계 범위의 한계**

**태그·라벨 기준의 불완전성**

**단순한 모델·추천 로직**

**개인화 정보 활용의 제한**

**평가 방식의 한계**

# 향후 보완점

**데이터·통계 확장 및 보정**

**행동 기반 개인화 강화**

**태그 체계 및 라벨링 고도화**

**서비스·UX 측면 개선**

**모델·추천 알고리즘 개선**

**평가 체계 고도화**

# 프로젝트 회고

## 하요한

정답이 없는 데이터를 직접 라벨링하고 “태그”라는 기준을 합의해 가는 과정에서 데이터를 다루는 주관과 논리를 기를 수 있었고, 한정된 일정 속에서 핵심 기능에 집중하는 선택과 집중을 통해 팀원들과 함께 효율적으로 목표를 완수하는 협업 능력을 키울 수 있는 시간이었습니다.

## 정연중

프로젝트가 프론트엔드·백엔드·데이터 파트가 서로 의존성이 높은 구조였기 때문에 기능 개발전에 충분한 구조 합의 과정을 가져 프로젝트 진행함에 있어 큰 어려움 없이 마무리 할 수 있었습니다. 다시한번 협업은 기술보다 소통이 우선이라는 것을 알게 된 순간이었습니다.

## 이승운

이번 프로젝트를 통해 태그 추천, 캘린더 UI, 정렬 충돌, DB 타입 문제 등 예상치 못한 오류를 해결하며 ‘근거 기반으로 문제를 분석하는 방식’을 깊이 체득할 수 있었습니다. 데이터·UX·디버깅 모두에서 한 단계 성장했고, 앞으로도 이러한 방식으로 깊게 파고드는 개발자가 되겠다는 방향성을 확인한 시간이었습니다.

## 강진용

백엔드와 프론트를 함께 다루며 풀스택과 배포까지 직접 개발을 진행한 덕분에, 다양한 기술을 시도해 볼 수 있는 좋은 경험이었습니다. 팀원들이 각자 맡은 역할을 훌륭하게 해내 주어 더욱 원활한 팀 협업을 이룰 수 있어 좋았습니다. 다음 프로젝트에서는 이번 경험을 바탕으로 더 많은 도전을 이어가고 싶습니다.

## 한준희

디자이너로서 전체적인 로직이해와 사용자의 편의성을 고려한 UX를 매우 중요시 생각하며 프로젝트에 임했습니다. 이론적인 머신러닝 파트를 브라우저에 표현시키는 것에 성공했을 때 이루 말할 수 없는 희열을 느꼈던것 같습니다. 데이터를 다루면서 생긴 많은 트러블 슈팅 경험은 정말 값진 경험이라 생각합니다.

# 참고 문헌

1. 2024 문화관광축제 빅데이터 분석 보고서, 한국관광공사 지역관광육성팀
2. 문화체육관광부·한국문화정보원, 「전국 문화축제 표준데이터」, 공공데이터포털
3. 한국관광공사, 「TourAPI 4.0 국문 관광정보 서비스」, 공공데이터포털
4. 한국관광공사, 「관광지별 성/연령 방문 집중률」, 한국관광 데이터랩
5. 카카오, 「Kakao 지도 JavaScript API 개발가이드」, 카카오 디벨로퍼
6. scikit-learn Developers, \*scikit-learn: Machine Learning in Python\*, scikit-learn.org
7. lovit, \*soynlp: Korean NLP Library in Python\*, GitHub Repository
8. VMware, \*Spring Boot Reference Documentation\*, spring.io

**감사합니다!**