# 김삼문 1주차 발표

김가희 김예빈 김현진 문수빈 전수아

#### 주제



# 인디 공연 정보 플랫폼(웹)



전국 공연장·공연 정보를 한 곳에

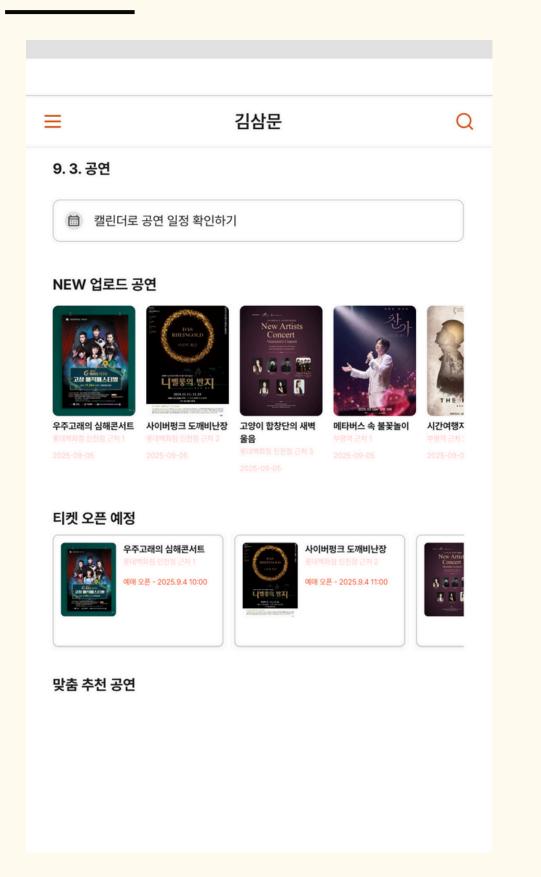


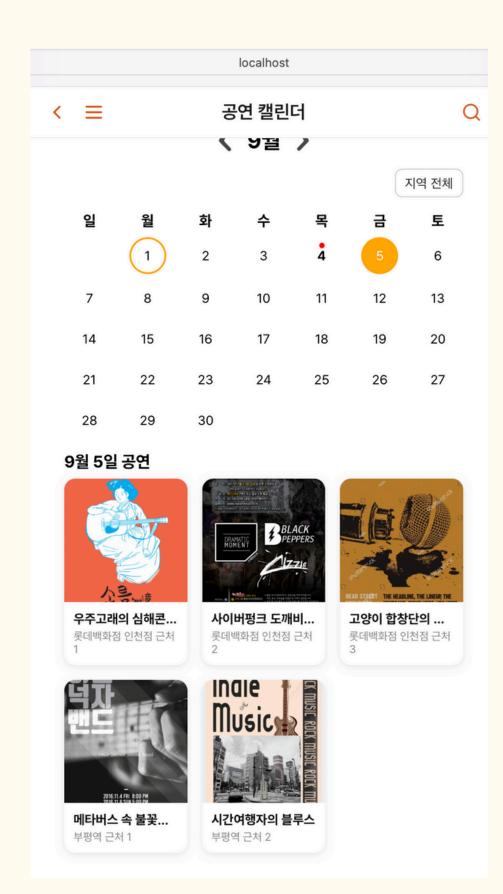
필터·검색·캘린더로 쉽게 탐색



SNS 크롤링으로 최신 정보 자동 업데이트

#### 주제







#### 오후 3시 28분 이후 공연







CLUB FF 오후 4시 10분



**언더플레이** 오후 5시



#### 여름, 파랑

오후 4시

#### 언플러그드 홍대점

서울 마포구 와우산로33길 26 1층 🏻 🗖







#### 인디 공연 공지글 판별

```
※ 8:30 '유지희 '@jihui_1011
💓 9:00 ' 화백 ' @band_ hwabaek
√공연정보
- 공연일시 : 2024. 11. 1(금) 오후 7:30
- 공연장소 : 홍대우주정거장
(서울특별시 마포구 와우산로 29길 16, 지하 1층)
✓티켓가격 : 예매-20,000원 현매-25,000원

✓ 예 매 처 (SPACE CULTURE)

프로필 링크 -> 집집라이브 예매 클릭!
#공연 당일 오후 7시부터 입장 가능합니다.
#선착순 입장이며, 비지정석입니다.
final score=6.30 / threshold=2.5
label=인디공연
→ ('인디공연', 6.3)
```

```
기본적인 주류지식 및 메뉴 설명은 함께 공유해드릴
지원 방법
👉 지원서 작성하러 가기
https://forms.gle/oxkhAKXy3DNmakqA9
또는
△인스타그램 하이라이트 '스탭 지원하기.' 클릭!
_poster by @muyahojung
UZUPA 우주파
제주시 탑동로 81
81, Tapdong-ro, Jeju
final score=-1.50 / threshold=2.5
label=비인디공연
→ ('비인디공연', -1.5)
```

# • 문제점

- 긴 게시물인 경우
  - 단어 多, + 와 -가 다양하게 섞이면서 오판단 발생



- 가중치 조절
- 공연 공지글임을 확실하게 판단하는 조건 추가

# • 공연 정보 추출 모델 훈련

- 바이오 태깅(BIO TAGGING)을 통한 학습 데이터 수집
- CRF (순차적인 데이터에서 라벨 예측하는 머신러닝 알 고리즘) 모델

279	Tristes	B-TITLE	
280	Tropiques	I-TITLE	
281	Concert	I-TITLE	
282	;	0	
283	Seoul	0	
284	Univ.	0	
285	Anthropol	0	
286	2025	0	
287	Summer	0	
288	<b>✓</b>	0	
289	6월	B-DATE	
290	27일	I-DATE	
291	금요일	I-DATE	
292	오후	B-TIME	
293	7시	I-TIME	
294	O	0	
295	라디오가기	B-VENUE	
296	@radl0_ga	0	
297	&	0	

#### 공연 정보 추출 모델 훈련

🎇 3호선 버터플라이가 긴 수면 상태에서 깨어나 새 앨범[환희보라바깥]으로 날개짓을 펼?

펜타포트 락 페스티벌 무대를 시작으로 2025년 여름 싱글<너의 속삭임>을 발매 하고 8월 남상아의 몽환적인 보컬과 느릿하면서도 섹시한 리듬, 9년 만에 복귀한 성기완의 사이키덜

#### <예매 정보>

- 공연명: 3호선 버터플라이 [환희보라바깥] 발매 기념 콘서트
- 공연 일시: 2025년 9월 13일 7pm
- 공연 장소: 무신사 개러지
- 좌석: 스탠딩
- 예매: NOL 티켓
- => 프로필 링크
- 티켓 가격: 예매 55,000원, 현매 66,000원
- 관람 등급: 만 7세 이상
- 관람 시간: 약 90분
- \* 티켓오픈: 2025.08.12(화) 12시(정오)\*

#3호선버터플라이 #3호선\_버터플라이 #단독공연 #무신사개러지 #놀티켓 #NOL\_TICKET #알

예매: NOL티켓

티켓오픈: 8월12일(화) 12시(정오)

#3호선버터플라이 #3rdlinebutterfly #3rd\_line\_butterfly #live #단독공연 #무신사개

TITLE: 3호선 버터플라이 [환희보라바깥] 발매 기념 콘서트

ARTIST:

VENUE: 무신사 개러지 DATE: 2025년 9월 13일

TIME: 7pm

PRICE\_ON: 55,000원, PRICE\_OFF: 66,000원

TICKET\_OPEN: 2025.08.12(화) 12시(정오)\* / 8월 12일(화) 12시(정오)

ACCOUNT:

#### 공연 정보 추출 모델 훈련

### • 문제점

- 정확한 공연 TITLE 추출
  - 앞뒤 맥락 없음
  - 타이틀 없는 공연 많음



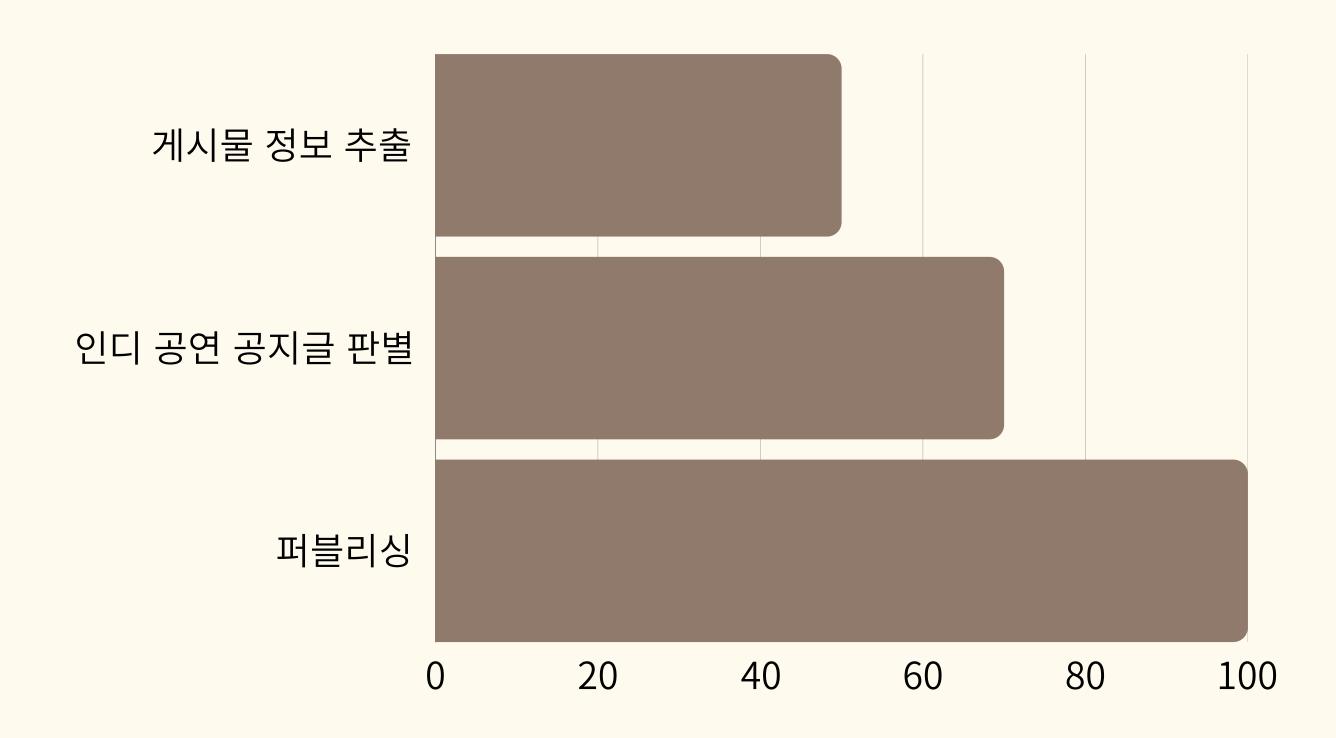
- 타이틀이 포함된 학습 데이터 추가
- 자주 나타나는 타이틀 표기 형식 후처리 강화

- 정확한 ARTIST 추출
  - DJ, 기획자 이름 등도 추출
  - 제목이나 문장 안에서만 언급



- 문장 학습 데이터 추가
- 잘못 뽑은 케이스만 따로 모아서 재학습

# 진행상황



# 평가 요소



### 주간 웹 방문자 수 1000명 이상

Google Analytics 4(GA4)로 1주일간 수집한 사용자 데이터를 기준으로 계산

모델 정확도

### 게시물 정보 추출 알고리즘 모델 정확도 > 90%

모든 공연장 계정의 게시물에서 모델이 추출한 값 vs 사람이 추출한 값(정답 데이터) → 전체 라벨 중 90%이 일치하면 성공

사용자 만족도

## 사용자 300명의 웹 만족도 4점 이상 (5점 만점)

웹사이트 접속자 구글폼 설문조사 + 사용자 인터뷰 진행

# 평가 표 정리

구분	현재 상황	평가 기준	달성여부
주간 방문자 수	없음	1000명 이상	X
모델 정확도	60 %	90% 이상	진행중
사용자 웹 만족도	없음	5점 만점 4점 이상	X