



# 딥러닝을 활용한 노래 가사 분석 및 추천 시스템

팀 E 김은비 유창현 이은경 이인석

# 발표 내용

이번 주차 진행 사항

프로젝트 변경 사항

다음 주차에 진행해야 할 사항

# 크롤링 방식


멜론 음원 스트리밍 서비스 사이트 사용

‘멜론DJ’ 메뉴의 해시태그 기능을 검색

검색 결과 나온 플레이리스트들의 곡을 크롤링

#사랑 태그를 포함하는 10627개의 플레이리스트가 있습니다.

인기순 • 제작일순 • 수정일순




봄날, 벚꽃 엔딩부터 시작하는 완벽한 봄 노래 Playlist

DJ foreven.

♡ 1,368 | 99곡

#봄날 #벚꽃엔딩




90년대생들 모여라! 그 시절, 아이리버로 즐겨들던 추억의 노래 모음

DJ melle

♡ 3,630 | 279곡

#추억 #분위기

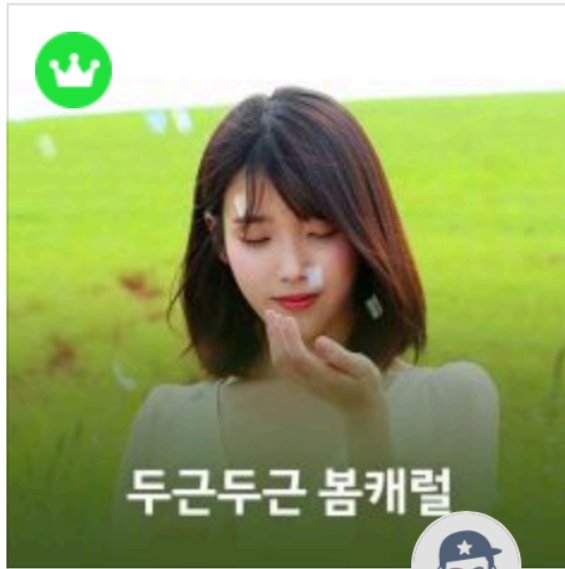


도입부터 봄 냄새 가득한 로맨틱 POP

DJ Hipster

♡ 1,352 | 73곡

#로맨스 #연인




두근두근 봄캐럴

봄이 오면 찾게 돼! 두근두근 기분 좋은 봄캐럴

DJ 멜군

♡ 7,055 | 152곡

#봄 #샬렘




드라마 속 남주인공이 애뜻히 바라볼 때 나오는 OST Part.2

DJ 별이빛남

♡ 191 | 53곡

#OST #샬렘




내 하루가 즐거워지는 기적 같은 플레이리스트

DJ Roopretelcham

♡ 2,190 | 60곡

#기분좋은 #행복




과몰입 준비 완료! 듣는 순간, 드라마 속으로 빨려 들어가는 OST ...

DJ melle

♡ 489 | 155곡

#드라마 #OST



8090 추억 소환! 그 시절 봄의 향수가 느껴지는 노래들

DJ Hipster

♡ 407 | 63곡

#봄 #90년대

# 기존 방식

기존 논문 분석을 통한 사용 기법 논의

특정 키워드를 활용하여 노래 가사 크롤링



Word2Vec, Doc2Vec 등을 활용한 가사 문장 단위의 분석, 유사도 측정



사용자에게 추천

# 의문점

## 기존 논문 분석을 통한 사용 기법 논의

노래 가사의 단순 문장 분석 시 결과가 잘 나올까?

결과가 나오더라도 가사가 말하는 감정에 대한 결과가 올바르게 나올 것인가?



# 새로운 방식

## 기존 논문 분석을 통한 사용 기법 논의

‘사랑’ 키워드를 활용하여 노래 가사 크롤링



‘사랑’이라는 키워드에서 느낄 수 있는 적절한 세부 감정들을 자료와 회의를 통해 결정



크롤링한 노래 가사를 분석하여  
세부 감정과 유사도가 높은 단어들로 감정 사전을 구축



세부 감정으로 부터 뽑은 단어들을 데이터 라벨링하여 딥러닝 모델 (예 : LSTM)들을 활용하여  
분류하고, 분류된 그룹 내 유사도 측정



사용자에게 추천

# 감정 사전 간단 예시

## 기존 논문 분석을 통한 사용 기법 논의

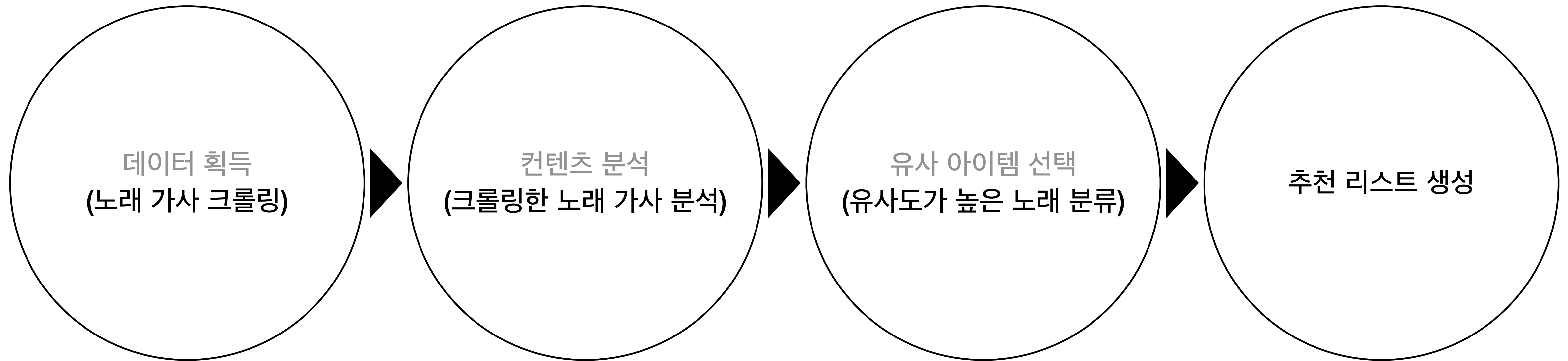
|     |        |    |     |
|-----|--------|----|-----|
| 슬픔  | 부정     | 분노 | 무관심 |
| 그리움 | 오해     | 한심 | 반응  |
| 외로움 | 질투     | 덤덤 | 차분  |
| 욕심  | 미워하(다) | 씩씩 | 느슨  |
| 삶   | 욕하(다)  | 허무 | 막연  |



# 추천 시스템 결정

## 컨텐츠 기반 필터링

음악 간에 측정한 유사도를 바탕으로 사용자가 듣거나 선택한 음악과 유사도가 높은 음악을 추천





# 다음 주차에 진행해야 할 사항

## - 노래 가사 크롤링

여러 패키지들과 코드를 활용하여 멜론 사이트의 #사랑 키워드를 포함하는 음악 크롤링하여 데이터 수집

## - 데이터 전처리와 세부 감정 결정

수집한 가사 데이터에서 영어, 특수문자와 같은 불용어를 제거

단어를 토큰화하고, 품사를 파악하는 등의 전처리 과정

‘사랑’ 키워드로 부터 느낄 수 있는 세부 감정들을 관련 자료 수집과 회의를 통해 결정