

데 이 터 분 석 캡 스톤 디 자 인

캡스톤 디자인 수행계획 발표

팀D 유창현 이은경 김은비 이인석

INDEX

- 01 과제 개요 및 목표
- 02 기대효과 및 활용방안
- 03 수행방법과 활용 도구
- 04 추진일정

01

과제 개요 및 목표

- ❶ 과제 주제 설명
- ❷ 과제 선정 배경 및 필요성
- ❸ 최종결과물의 목표



01

가사 유사도 측정을 기반으로 한 음악 추천 시스템



기존 음원 사이트의 가사 데이터를
수집하여 노래 간 가사의 유사도 측정

+



사용자의 플레이리스트의 노래와
가장 유사도가 높은 노래 추천

음악의 여러 구성 요소 중 **가사** 즉, 노래의 내용에 초점을 맞춘 추천 서비스 구현



음악 구성요소로서의 가사

음악을 구성하는 멜로디, 리듬, 화성 등의 음악 요소는 가사의 내용을 강조함으로써 가사가 전달하는 의미와 감성을 깊이 있게 표현



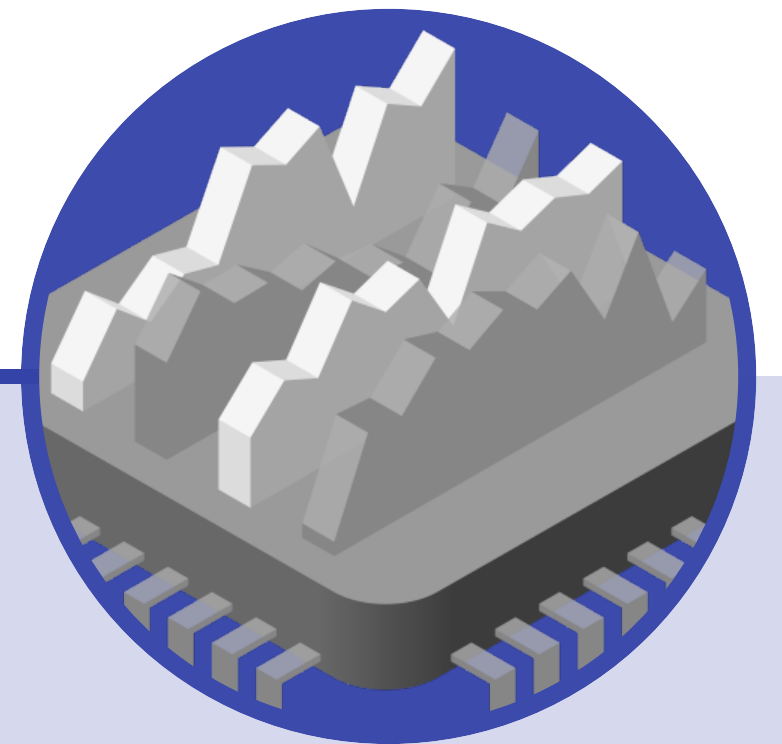
슬픈 정서적 상태

한 연구 논문에 따르면 사람은 슬픈 정서적 상태일 경우, 특히 가사가 전달하는 의미를 고려하여 음악을 선택
가사의 인지적 메시지가 감상자의 정서적 상태에 따라 선택 시 중요성을 가짐



노래 감상 시 가사

대학생을 대상으로 진행한 설문조사에서 노래 감상 시 가사 내용에 비중을 둔다는 응답은 전체의 약 90%



기존 음악 추천 시스템

기존 음악 추천 시스템은 협업 필터링, 콘텐츠 기반 필터링을 활용해 오디오 데이터, 아티스트, 장르 등에 기반한 추천 제공. 특히, 대중성을 강조하는 경우가 많음

03

최종 결과물의 목표

멜론 크롤링

특정 키워드에 해당하는 노래의 제목,
가사, 가수 데이터 수집

데이터 분석

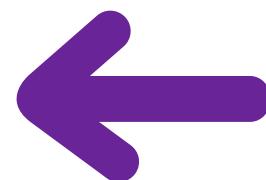
Doc2Vec 모델을 활용해 각 곡의 가
사를 하나의 문서로 벡터화하고 코사
인 유사도를 값을 계산

곡 추천 제공

가장 큰 코사인 유사도 값을 갖는 일정
개수의 곡 추천 제공

"노래 가사 분석을 통해 도출한 가사 의미의 유사
도를 기반으로 노래 간의 유사도를 수치화하여
가장 적합한 음악을 추천"

사용자A의 음원 플레이 리스트



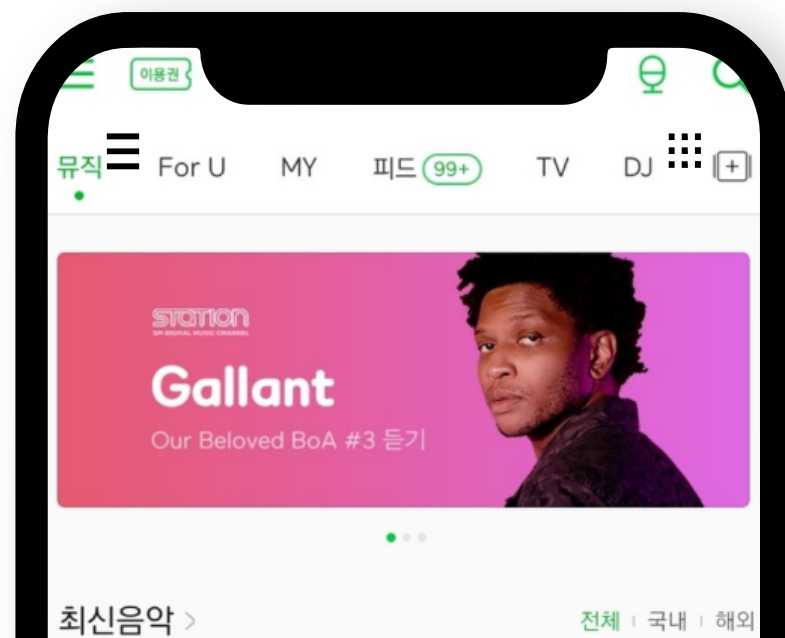
02

기대효과 및 활용방안



04

기대효과 및 활용방안



01

기존 시스템과의 차별성

멜로디, 리듬 등 음악의 다른 요소들과 가수, 장르, 인기도 등에 집중한 기존 추천 시스템과 달리 가사와 그 의미에만 집중한 차별화된 추천 가능

02

가사 기반의 정확한 추천

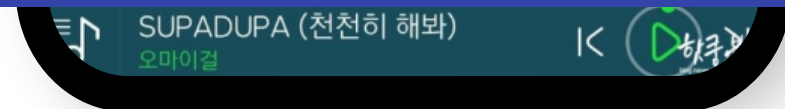
노래 가사가 담고 있는 긍정적인 메시지, 가사 자체에 아름다움, 희망과 공감적인 내용 등 가사에 집중하여 노래를 감상하는 사용자나 가사 기반의 노래 선정이 필요한 상황에 더 정확한 추천 가능

03

단순한 감정보다 가사의 의미에

집중한 분석

노래 가사는 구절, 구간마다 서로 다른 감정을 내포하고 있기 때문에 감성분석을 통해 노래에서 하나의 감정을 추출하는 것보다 가사 전체의 의미의 유사성을 기반으로 한 더 자세하고 정확한 분석 수행



03

수행방법과 활용 도구





01

분석 대상 키워드 선정
예) 이별노래, 사랑노래

02

음원 사이트 '멜론'에서
키워드에 해당하는 노래
가사 데이터 수집

03

특수문자, 영어 등 불
필요한 데이터 제거
및 단어 추출

04

문서 아이디를 할당하
고 Doc2vec모형을
활용하여 데이터 학습
및 벡터화

05

코사인 유사도 값을 계산
하여 노래 간 상호 유사도
측정

06

입력 받은 특정 사용자의
플레이 리스트의 노래 가
사와 가장 유사한 곡을 도
출하도록 구현

STEP 01

웹 크롤링 기법
BeautifulSoup

STEP 02

가사 내 단어 추출을
위한 형태소 분석기



STEP 03

Doc2Vec

STEP 04

코사인 유사도
(cosine similarity)

04

추진일정



A blue spiral-bound notebook with the text "3월" written in large black font on its white page.

3월

- 주제선정
- 중심 키워드 선정
- 웹 크롤링을 통한
가사 데이터 수집

A blue spiral-bound notebook with the text "4월" written in large black font on its white page.

4월

- 분석을 위한 데이터
전처리
- 형태소 분석기를 통
한 단어 추출
- 중간보고서 작성

A blue spiral-bound notebook with the text "5월" written in large black font on its white page.

5월

- Doc2vec 모델 설계
및 실행
- 코사인 유사도 도출
- 추천 시스템 구현

A blue spiral-bound notebook with the text "6월" written in large black font on its white page.

6월

- 추천 시스템 구현
- 결과 보고서 작성 및
발표 준비

감사합니다.
