

Recommender System

#오늘#듣기
#좋은#노래

Team | 멜로니어리
20기 김유진 김재훈 박준희

KUBIG
DATA SCIENCE & AI

CONTENTS

01

프로젝트 개요

#오늘 #듣기 #좋은 #노래

02

데이터

Melon 데이터

데이터 크롤링

03

모델링

추천 과정

태깅

벡터화

04

시연

오늘의 하루를
입력해주세요.





01. 프로젝트 개요

01. 프로젝트 개요

#오늘 #듣기 #좋은 #노래

: 오늘의 일기를 입력하면 키워드와 분위기에 맞는 곡을 추천해드립니다!

사용자가 입력한 일기의 [감정 태그와 키워드 태그]를 분석하고,
멜론 플레이리스트 데이터에서 수집한 (1)플레이리스트 태그와
곡의 가사에서 수집한 (2)가사 태그, (3)나머지 곡 정보 (가수, 좋아요수, 장르 등)을
활용하여 곡을 추천하는 알고리즘을 구현하고자 함



첫 구현 목표

오늘 집에 오는 길에 소나기가 내렸는데 빗소리가 참 좋더라. 나는 비를 좋아해서 기분이 좋았어.
→ [기쁨 / 비, 소나기, 집에 오는 길] 추출 → 김뜻돌의 '비 오는 거리에서 춤을 추자'는 어떠신가요?

01. 프로젝트 개요

Motivation

일반적인 추천시스템 : 유저, 아이템 사이의 관계 데이터 활용

BUT, 유저(사용자)/아이템(노래) 데이터 구할 수 x

기존 추천시스템

USER

USER FEATURE

ITEM FEATURE

멜로니ერი

PLAYLIST

PLAYLIST HASHTAGS

LYRICS, GENRE, # OF LIKES...



취향저격! 감성 DJ

러블리쓰_

취향저격 탕탕탕! 여러분의 감성을 자극할 테마들로 구성된 참신한 플레이리스트를 공개합니다. 개인적인 취향이지만 세심하고 꼼꼼하게 선곡했으...

👤 16,344 | 🎵 133



음악 기상캐스터

YevelyK

음악만 들어도 날씨가 보여요! 날씨 또는 계절에 따라 달라지는 음악 플레이리스트들! 제가 책임져 드리겠습니다.)

👤 6,932 | 🎵 159



늘 새로운 음악을 찾아 다니는 DJ

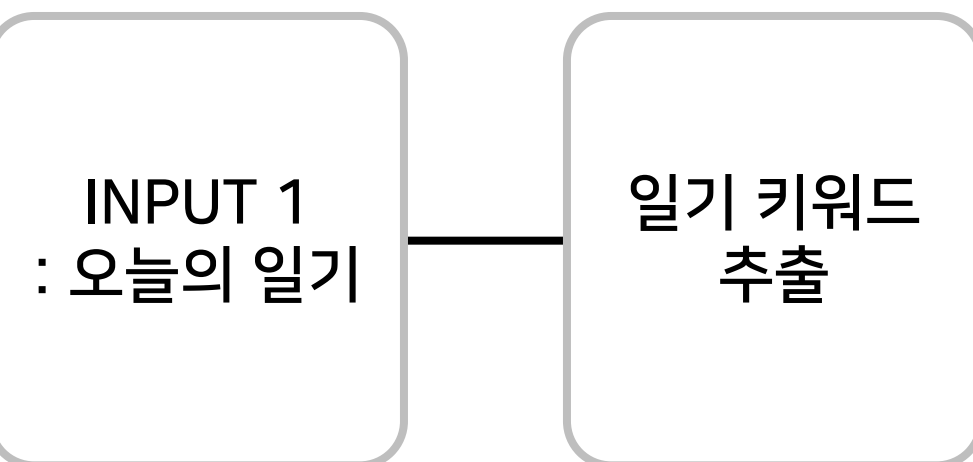
Mogu_

음악만큼은 깐깐한 DJ가 정성으로 채운 플레이리스트

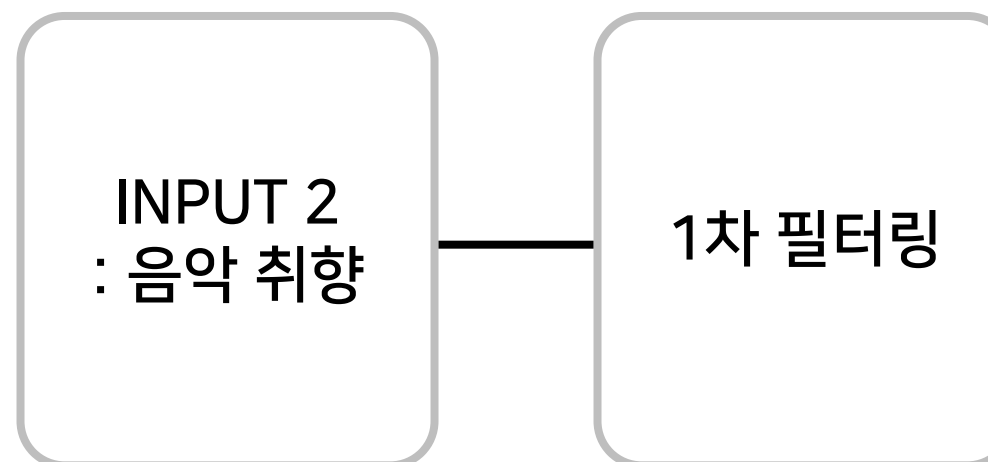
👤 9,356 | 🎵 50

01. 프로젝트 개요

Key Idea



+



ex. "시험이 끝나
서 완전 후련해!
크리스마스의
몽글몽글함을
실컷 즐길거야!"

ex. [후련, 크리스
마스, 몽글몽글함]

ex. 최신 노래,
유명한 곡, POP

입력된 취향에 맞
는 노래들 필터링

1. 음악 취향 1차 필터링,
후보곡에 추가
 2. 일기 키워드와 유사한
플레이리스트 필터링,
후보곡에 추가
- => 두 가지 종합하여
[가사 태그] 기반 추천



02. 데이터

02. 데이터

o Melon

1) 인기차트 - 1980년대부터 / 한국 & 해외 [총 6853곡]

2) 장르음악 - 한국 & 해외 [총 2960곡]

3) DJ 플레이리스트 - 파워 DJ 52명의 플레이리스트 [총 39190곡]

: 1과 2는 수합하여 취향 필터링 후 사용


: 3은 플리 태그를 가지고 있기 때문에 유사도 기반 필터링 후 사용

02. 데이터

o Melon

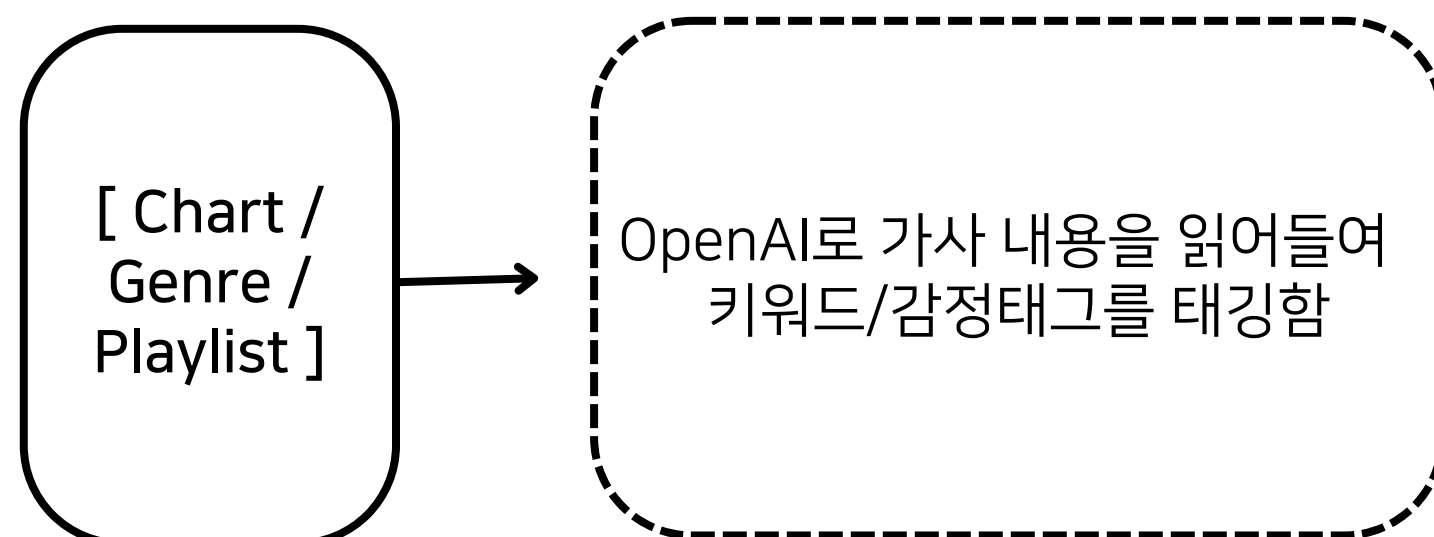
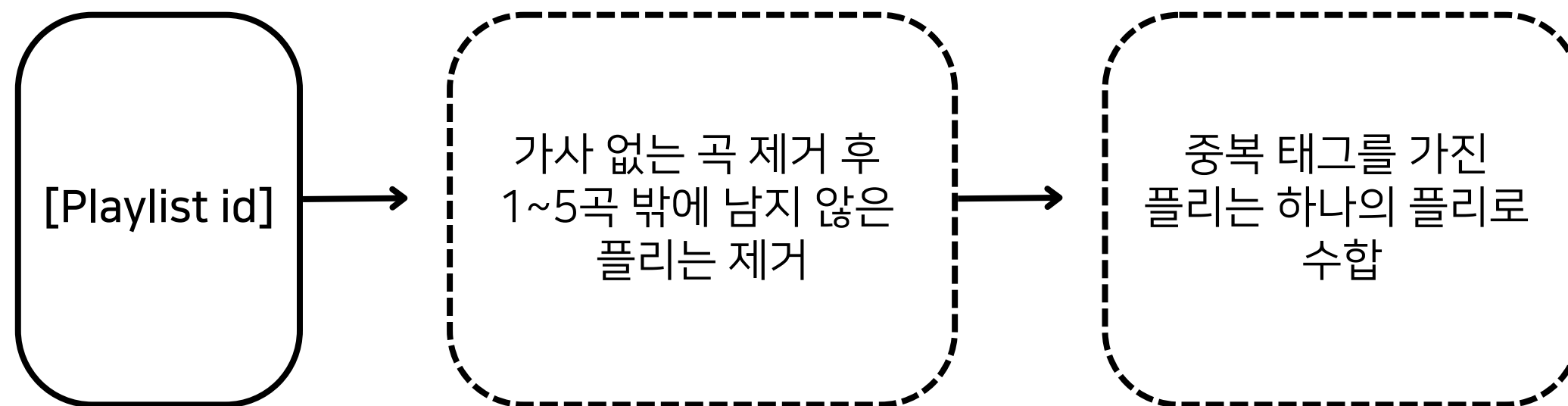
- 연도, 노래 제목, 가수, 가사, 발매년도, 장르, 좋아요 수 크롤링
- 총 49003곡

	연도	제목	가수	가사	발매년도	장르	좋아요 수
0	2023	Ditto	NewJeans	Woo woo woo woo oohWnWoo woo woo wooWnStay in ...	2022	댄스	295,491
1	2023	Hype Boy	NewJeans	(1,2,3,4)WnBaby, got me looking so crazyWn빠져버리...	2022	댄스	304,682
2	2023	I AM	IVE (아이브)	다른 문을 열어 따라 갈 필요는 없어Wn넌 너의 길로 난 나의 길로 음Wn하루하루마...	2023	댄스	221,848
3	2023	OMG	NewJeans	이 노래는 it's about you babyWnOnly youWnYou you y...	2023	댄스	221,321
4	2023	사건의 지평선	윤하 (YOUNHA)	생각이 많은 건 말이야Wn당연히 해야 할 일이야Wn나에겐 우리가 지금 1순위야Wn안...	2022	록/메탈	330,876
5	2023	Kitsch	IVE (아이브)	It's our time 우린 달라 특별한 게 좋아WnOh What a good t...	2023	댄스	161,428
6	2023	퀸카 (Queencard)	(여자)아이들	Hey youWn월 보니? 내가 좀 Sexy Sexy 반했니WnYe youWn뭐 하...	2023	댄스	148,625
7	2023	Attention	NewJeans	You and meWn내 맘이 보이지Wn한참을 쳐다봐Wn가까이 다가가WnYou se...	2022	댄스	253,852
8	2023	헤어지자 말해요	박재정	헤어지자고 말하려 오늘Wn너에게 가다가 우리 추억 생각해 봤어Wn처음 본 네 얼굴W...	2023	발라드	179,930
9	2023	After LIKE	IVE (아이브)	또 모르지 내 마음이Wn저 날씨처럼 바뀔지Wn날 나조차 다 알 수 없으니Wn그게 뭐...	2022	댄스	244,885



03. 모델링

03. Overview - 전처리



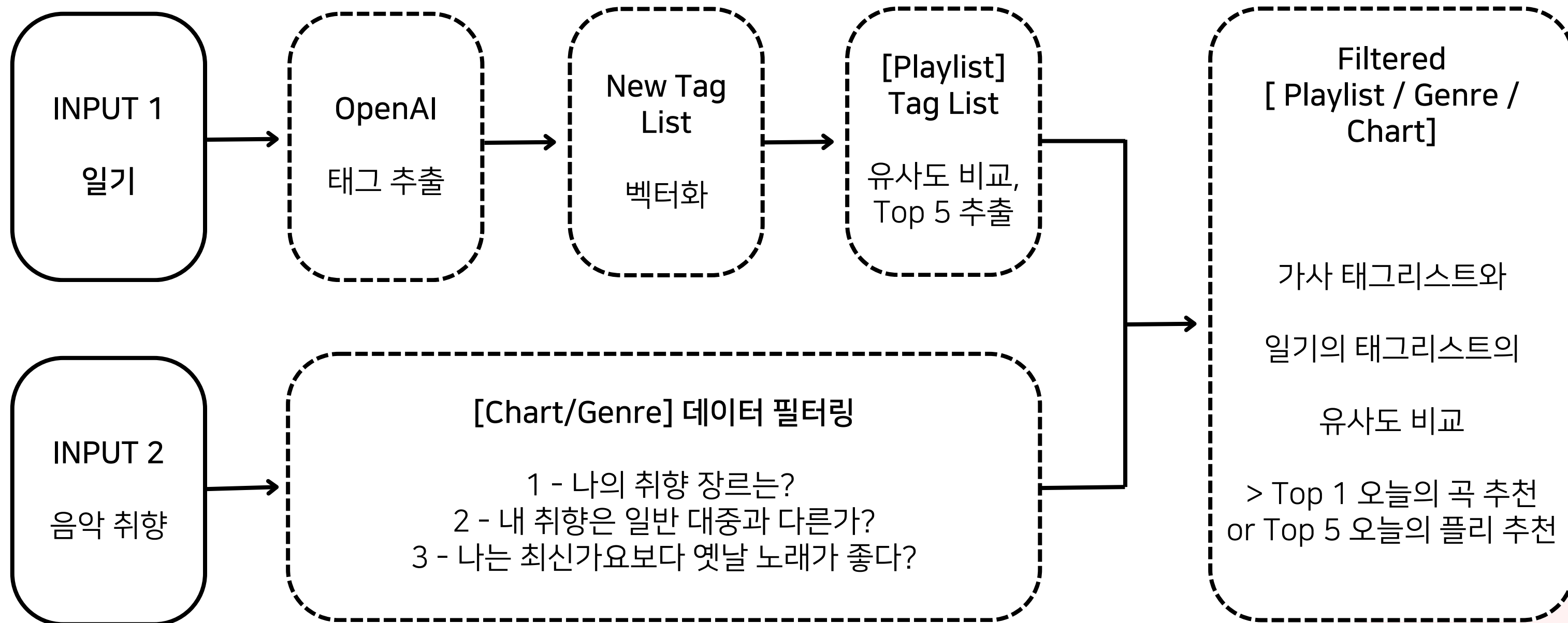
제목	가수	키워드	감정
비행기	거북이	[비행기, 거북이, 하늘]	[설렘, 두려움, 그리움]
미리 메리 크리스마스 (Feat. 천둥 Of MBLAQ)	아이유	[너, 눈꽃, 선물]	[사랑, 행복, 불안]
좋은 날	아이유	[하늘, 오빠, 눈물]	[설렘, 슬픔, 사랑]
I Don't Care	2NE1	[남자친구, 사랑, 속임수]	[분노, 슬픔, 해방]
너랑 나	아이유	[비밀, 마음, 시간]	[두려움, 벽참, 서러움]

03. 일기 / 가사 Tagging

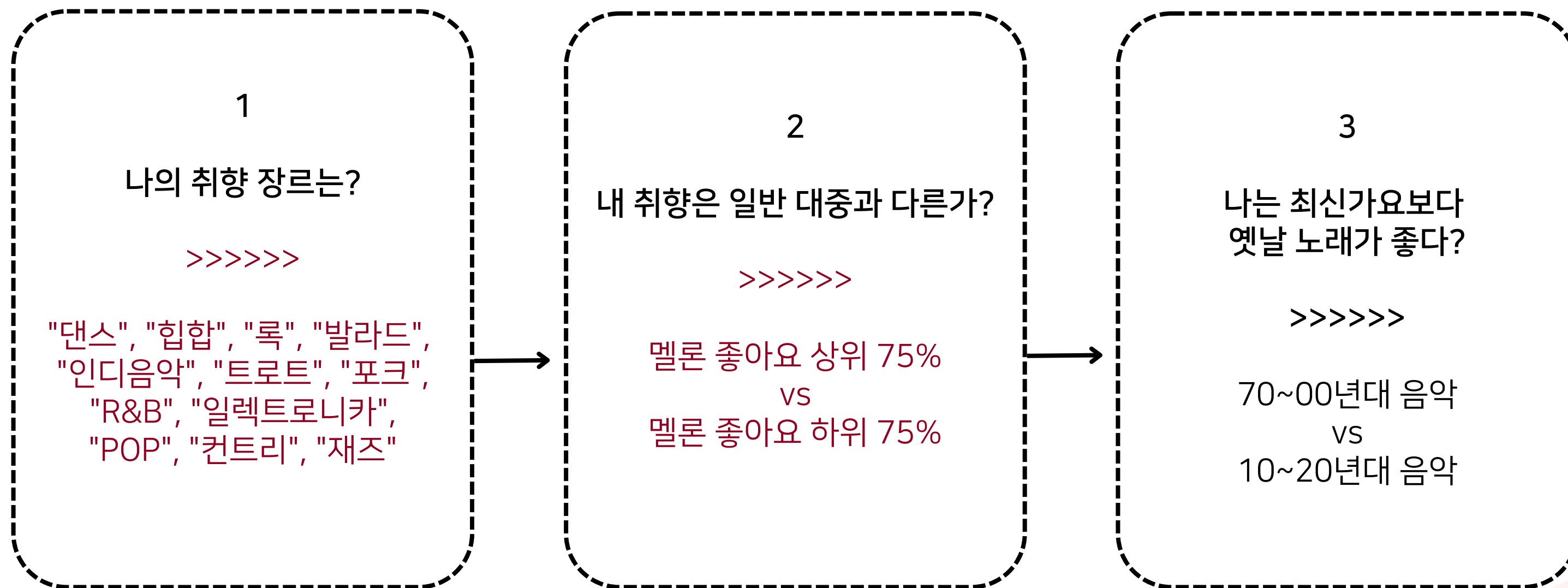


- OpenAI API 활용
- INPUT : input_text (일기)
- PROMPT : 감정 3개, 키워드 3개 추출
 - Extract **three emotions** and **three keyword** in **Korean** from the following text : {input_text}.
 - Make sure all the answers are **nouns**.
- OUTPUT : 단어 6개를 담은 리스트

03. Overview - Recommend



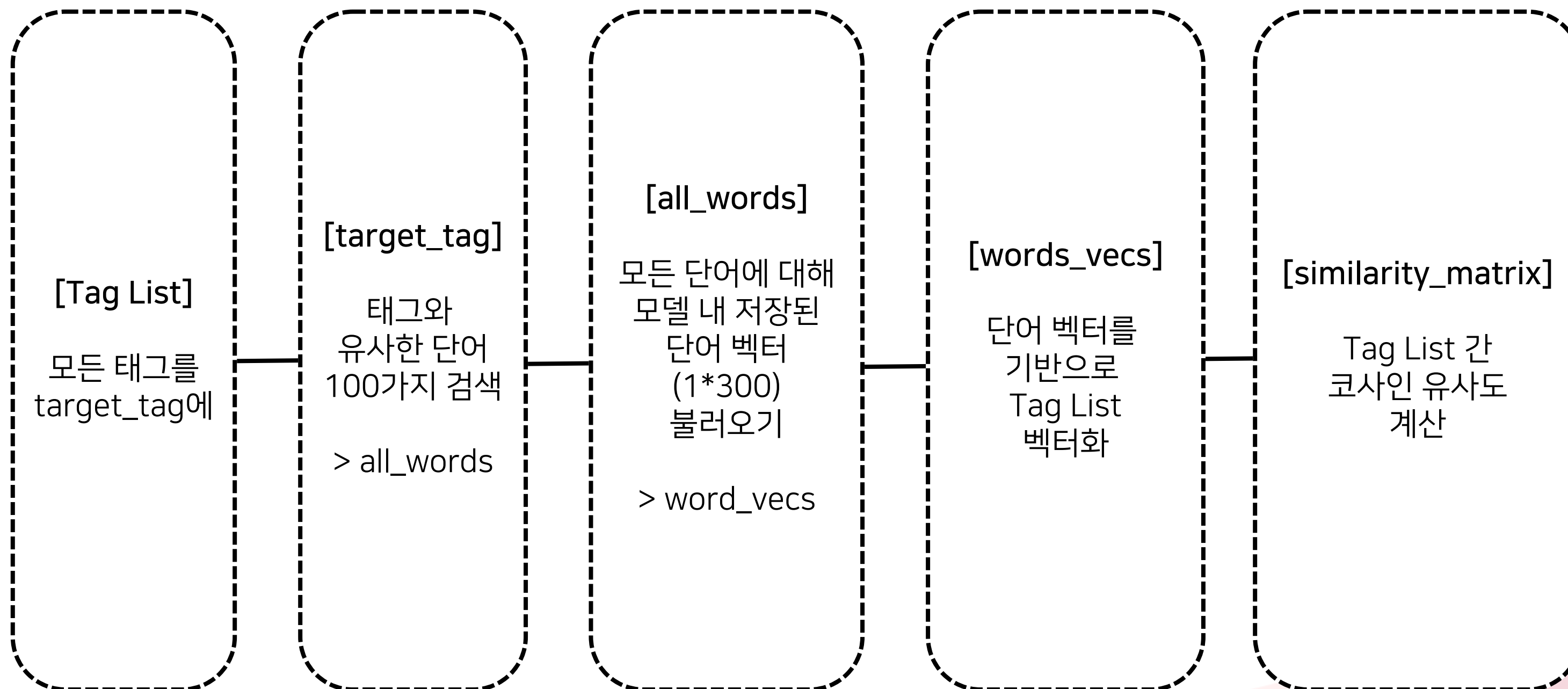
03. 취향 필터링



*** Q 그럼 선택한 범위 이외의 곡들은 추천 받지 못하는 건가요?

A 일기 키워드와 유사한 플레이리스트에서 선택한 범위를 벗어난 곡들도 후보곡들로 들어갈 수 있음

03. Overview - Vectorization



03. Vectorization

*fast*Text

fastText의 Pretrained 한국어 모델을 사용

fastText는 common crawler와 wikipedia의 데이터를 학습한 Pre-trained model을 제공
추가로 학습이 가능한 cc.ko.300.bin 대신 cc.ko.300.vec를 사용하여 이미 학습된 단어 벡터를 가져옴

장점

- fastText는 인터넷에서 수집한 단어를 포함. 신조어 대응에 좋은 결과를 보임
- 한국어 모델에도 영어 데이터가 존재하며 유사한 단어를 검색할 수 있는 장점

한계

- 직접 학습시키는 것이 아니기 때문에 입력하는 단어가 missing key일 수 있음
- '우울'은 존재하지만 '우울할때'와 같은 단어는 존재하지 않음
- 유사한 단어에 이상한 한자어, 영어가 나올 때가 있음



04. 시연

04. 시연

오늘의 일기

일기 키워드 생성

“시험이 끝나서
완전 후련해!

이제 연말이니까,
크리스마스의
몽글몽글함을
실컷 즐길거야!

다음 학기 시작 전까지
푹 자고 푹 쉬어야겠어”

['후련함', '즐거움',
'휴식', '시험', '연말',
'크리스마스']

```
#####
```

```
input_text = ""
```

```
시험이 끝나서 완전 후련해!  
이제 연말이니까, 크리스마스의 몽글몽글함을 실컷 즐길거야!  
다음 학기 시작 전까지 푹 자고 푹 쉬어야겠어  
""
```

```
#####
```

```
keywords = get_keywords(input_text)
```

```
diary_tag = ast.literal_eval(keywords)  
print(diary_tag)
```

```
✓ 0.9s
```

```
['후련함', '즐거움', '휴식', '시험', '연말', '크리스마스']
```

04. 시연

INPUT 2
: 음악 취향

필터링

힙합,
유명한 노래,
최신 노래

입력된 취향에
맞는 노래들
필터링

힙합

장르를 입력하세요 (예: 힙합, 발라드 등): (확인하려면 'Enter' 키를 누르고, 취소하려면 'Escape' 키를 누름)

0

좋아요 수에 대한 조건을 입력하세요 (O: 상위 75%, X: 하위 75% 이하): (확인하려면 'Enter' 키를 누르고, 취소하려면 'Escape' 키를 누름)

X

옛날 노래가 좋으신가요? (O, X): (확인하려면 'Enter' 키를 누르고, 취소하려면 'Escape' 키를 누름)

	장르	제목
416	랩/힙합	어땠을까 (Feat. 박정현) \
418	랩/힙합	한여름밤의 꿈
419	랩/힙합	자니 (Feat. Dynamic Duo)
420	랩/힙합	Officially Missing You
427	랩/힙합	삐딱하게 (Crooked)
...
6272	랩/힙합	문득
6276	랩/힙합	만남은 쉽고 이별은 어려워 (Feat. Leellamarz) (Prod. TOIL)
6291	랩/힙합	Shut Down
6305	랩/힙합	Yet To Come
6388	랩/힙합	Summer (Feat. BE'O (비오))

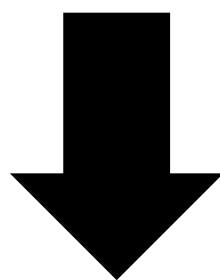
04. 시연

“시험이 끝나서 완전 행복해!

즐거운 노래를 들으며 신나게 춤도 추고,

연말의 분위기도 즐길거야!”

=> ['행복', '즐거움', '신나', '시험', '춤', '노래']



제목

Singing in the Rain (진솔)
How We Do (Feat. 마크 of NCT)

내가 제일 잘 나가
아무노래
그런 밤

Alone (Feat. GSoul)

기대꽃

쏘리 쏘리 (Sorry, Sorry)

Uptown Funk (Feat. Bruno Mars)

수채화

SUPER JUNIOR (슈퍼주니어)

Mark Ronson

가수

이달의 소녀 \
시우민 (XIUMIN)

2NE1

지코 (ZICO)

어반자카파

CHEEZE (치즈)

정키

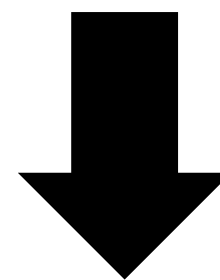
싸이 (PSY)

“오늘 하루 종일 내리는 비를 바라보니 우울하면서 센치하다.

해가 안 뜨는 것을 보니 이제 장마가 시작되는건가보다.

장마가 끝나면 더운 여름이 시작될텐데 벌써 무섭다.”

=> ['우울', '센치', '무서움', '비', '장마', '여름']



제목

추억속의 재회
다 알아요
참새와 허수아비
순애보

오늘부터 우리는 (Me Gustas Tu)

침묵의 기록

이별살이

비 내리면 (Feat. sogumm) (Prod. BOYCOLD) 릴러말즈 (Leellamarz)

내가 아는 한가지

준비 없는 이별

가수

조용필 \
양파

조정희

유리상자

여자친구 (GFRIEND)

이승환

린

이덕진

녹색지대

04. 시연

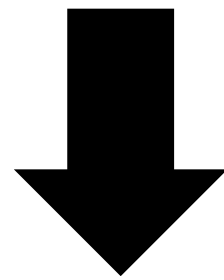
“오늘은 오랫동안 준비해 온 시험을 망쳐서 기분이 안 좋았다.

공부한 적 없는 부분만 문제로 내시는 교수님이 미웠다.

그대신 맛있는 거라도 먹어서 기분을 풀려고 카페에 가서 케이크를 먹었다.

역시나 맛있었고 친구와 대화하면서 많은 응원을 받았다.”

=> ['기쁨', '분노', '평온', '시험', '카페', '케이크']



제목	가수
Africa	Toto \
i wanna be your girlfriend	girl in red
Boulevard of Broken Dreams	Green Day
Take A Picture (Andy Platts' Demo Ver Of John ...	Mamas Gun
Memories	Conan Gray
네온사인	버스커 버
Modern Times	아이유
Couch Potato	Jakubi
귀여운 남자	JUNIEL
Falling	WNRDA

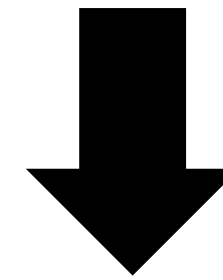
“오늘은 빙판길에 넘어졌다.

무릎이 까지고 온 몸에 눈과 흙이 묻은 채로 덜덜 떨며 집에 돌아가야 했다.

집에 와보니 오늘 처음 입은 옷이 찢어져 있는 것을 보고 마음이 아팠다.

최악의 하루였다.”

=> ['아픔', '답답함', '실망', '빙판길', '무릎', '옷']



제목	가수
이번 겨울	권진아 \
서울의 달	김건모
동경	박효신
내가 너의 곁에 잠시 살았다는 걸	토이 (Toy), 김연우
첫눈에... (Snowy Wish)	소녀시대 (GIRLS' GENERATION)
Beautiful Life	V.O.S
내사랑 내곁에 (응답하라 1988 삽입곡)	김현식
영화 같던 날	CHEEZE (치즈)
서리	이제
제자리 걸음	김종국



Thank You