

# 음악 분류 모델 학습을 위한 데이터셋 구성

서용석, 임동혁  
한국전자통신연구원(ETRI)

yongseok@etri.re.kr

## Dataset configuration for music classification model learning

Yongseok Seo, Dong-Hyuck Im  
Electronics and Telecommunications Research Institute

### 요 약

본 논문은 음악의 무드에 기반하여 자동 태깅, 분류 및 유사 음악 추천 등을 목적으로 하는 딥러닝 모델 개발을 위해 필요한 음악-무드 데이터셋 구성에 관한 것이다. 음악 무드 자동 태깅을 위한 딥러닝 모델의 학습을 위해서는 음악 데이터와 함께 음악별 무드가 태깅된, 즉 레이블링된 데이터셋이 준비되어야 한다. MSD, Last.FM, AllMusic 및 Jamendo 데이터셋을 관통하는 공통의 무드 태그를 선별하고 정제하여 한글-영문 단어로 구성된 총 45 종의 음악 무드 태그 데이터셋 구성을 제안한다.

### I. 서론

디지털 음악 시장이 커짐에 따라 방대한 양의 음악을 효과적으로 관리할 필요성과 함께 대용량 음악 검색이나 유사 음악 추천에 관한 기술 수요가 증가하고 있다.

본 논문에서는 음악의 자동태깅, 분류 및 유사 음악 검색/추천 등을 목적으로 하는 딥러닝 모델 개발을 위해 필수적인 음악 학습 데이터셋에서의 무드 태그(레이블) 구성에 관해 제안한다.

구분되어 연구되고 있다. 음악 장르 분류는 주로 미국 팝 음악을 대상으로 10 개 태그로 구성되며, 아래와 같다.

Blues, Jazz, Country/Western, Baroque, Classical, Romantic, Electronica, Hip-Hop, Rock, HardRock/Metal

한편, 음악의 무드 분류 기술의 성능 평가를 위해 MIREX 에서 2020 년까지 사용한 무드 태그는 표 1 과 같이 5 개의 클러스터로 분류하는 작업이었다. 데이터셋은 1,438 개의 곡으로 구성되어 있으며, 각 음악은 5 개의 클러스터 중 하나로 레이블링 되어있다.

### II. 본론

음악-무드 태그 데이터셋을 이용한 음악-무드 자동 태깅 모델은 그림 1 과 같이 구성할 수 있다. 그림과 같이 음악 무드 자동태깅을 위한 딥러닝 모델 학습을 위해서는 학습을 위한 음악 데이터와 함께 음악별 무드가 태깅된 즉, 레이블링된 데이터셋이 준비되어야 하며 레이블링된 데이터가 제한되어 있는 것은 딥러닝 모델의 성능 향상을 위해 반드시 해결되어야 할 문제이다 [1].

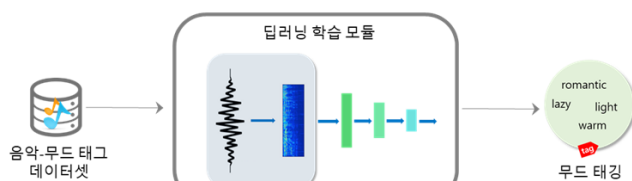


그림 1 딥러닝 기반 음악-무드 자동태깅 모델

음악 정보분석 성능평가 대회인 MIREX(Music Information Retrieval Evaluation eXchange)를 중심으로 음악 자동 분류에 대한 문제 정의와 연구가 진행되고 있으며, 음악 분류 문제는 크게 장르 분류와 무드 분류로

표 1 MIREX 무드 클러스터

클러스터	무드 형용사
클러스터 1	passionate, rousing, confident, boisterous, rowdy
클러스터 2	rollicking, cheerful, fun, sweet, amiable/good natured
클러스터 3	literate, poignant, wistful, bittersweet, autumnal, brooding
클러스터 4	humorous, silly, campy, quirky, whimsical, witty, wry
클러스터 5	aggressive, fiery, tense/anxious, intense, volatile, visceral

2021 년 MIREX[2]에서는 표 2 와 같이 Last.FM 에서 사용되는 음악 태그들을 정리하여 134 개의 무드 태그를 포함하는 18 개 그룹으로 확장하였으나, 의미가 중복되는 태그의 사용과 감성, 스타일을 포괄하는 단일 태그로 표현하기에 제한적인 부분이 있다.

약 100 만곡에 달하는 MSD(Million Song Dataset) 데이터셋에 대해 Last.FM 에서는 무드 태그를 포함한 약 50 만개의 음악 태그, AllMusic 에서는 289 개의 무드 태그를 작성하였다. 그러나 Last.FM 과 AllMusic 태그들은 의미론적으로 같지만 품사가 잘 정리되어 있지 않고 단복수 처리가 되지 않아 적절한 수의 무드 태그로 정제할 필요가 있다. 또한 음악 검색 및 추천을 위해 사용되는 다수의 공개 데이터셋을 이용함으로써 각 데이터셋의 단점들을 보완할 수 있다.

표 2 MIREX 무드 태그 데이터셋

그룹 ID	무드 태그	태그의 수	해당 곡의 수
G12	calm, comfort, quiet, serene, mellow, chill out, calm down, calming, chillout, comforting, content, cool down, mellow music, mellow rock, peace of mind, quietness, relaxation, serenity, solace, soothe, soothing, still, tranquil, tranquility	24	1,680
G15	sad, sadness, unhappy, melancholic, melancholy, feeling sad, sad song, mood: sad - slightly	8	1,178
G5	happy, happiness, happy songs, happy music, glad, mood: happy	6	749
G32	romantic, romantic music	2	619
G2	upbeat, gleeful, high spirits, zest, enthusiastic, buoyancy, elation, mood: upbeat	8	543
G16	depressed, blue, dark, depressive, dreary, gloom, darkness, depress, depression, depressing, gloomy	11	471
G28	anger, angry, choleric, fury, outraged, rage, angry music	7	254
G17	grief, heartbreak, mournful, sorrow, sorry, doleful, heartache, heartbreaking, heartsick, lachrymose, mourning, plaintive, regret, sorrowful	14	183
G14	dreamy	1	146
G6	cheerful, cheer up, festive, jolly, jovial, merry, cheer, cheering, cheery, get happy, rejoice, songs that are cheerful, sunny	13	142
G8	brooding, contemplative, meditative, reflective, broody, pensive, pondering, wistful	8	116
G29	aggression, aggressive	2	115
G25	angst, anxiety, anxious, jumpy, nervous, angsty	6	80
G9	confident, encouraging, encouragement, optimism, optimistic	5	61
G7	desire, hope, hopeful, mood: hopeful	4	45
G11	earnest, heartfelt	2	40
G31	pessimism, cynical, pessimistic, weltschmerz, cynical/sarcastic	5	38
G1	excitement, exciting, exhilarating, thrill, ardor, stimulating, thrilling, titillating	8	30
합계		134	6,490

본 논문에서는 MSD, Last.FM, AllMusic 및 Jamendo 데이터셋 전체를 관통하는 공통의 무드 태그를 선별하고 국내 온라인 음원 서비스에서의 감성/스타일/테마 태그들을 참고하여 한글-영문 단어로 구성된 총 45 종의 음악 무드 태그를 제안한다. 제안된 45 개의 정제된 무드 태그를 이용해 MSD 음악을 선별하거나 Last.FM API 를 이용하여 무드 태그에 해당하는 음악을 확보할 수 있으며, 이를 음악-무드 태깅을 위한 데이터셋으로 구축하여 딥러닝 모델 학습에 사용할 수 있다.

본 논문에서 제안하는 한글-영문의 음악 무드 태그는 표 3 과 같다.

표 3 한글-영문 음악 무드 태그

1 가벼운	Light	24 슬프고도 아름다운	Bittersweet
2 감상적인	Sentimental	25 슬픈	Sad
3 강렬한	Intense	26 신나는	Exciting
4 강한	Strong	27 신비한	Mysterious
5 고양되는	Uplifting	28 심각한	Serious
6 공격적인	Aggressive	29 외로운	Lonely
7 관능적인	Sensual	30 우울한	Gloomy
8 우울한	Melancholy	31 웅장한	Epic
9 극적인	Dramatic	32 자신만만한	Confident
10 긍정적인	Positive	33 잔잔한	Calm
11 꿈꾸는 듯한	Dreamy	34 재미있는	Fun
12 낭만적인	Romantic	35 점잖은	Gentle
13 달콤한	Sweet	36 정력적인	Energetic
14 따뜻한	Warm	37 진지한	Earnest
15 무서운	Scary	38 질주하는	Driving
16 밝은	Bright	39 차가운	Cold
17 복잡한	Complex	40 친밀한	Intimate
18 부드러운	Smooth	41 쾌활한	Cheerful
19 분위기 있는	Atmospheric	42 편안한	Relaxed
20 사나운	Fierce	43 행복한	Happy
21 사랑스러운	Lovely	44 향수어린	Nostalgic
22 서정적인	Lyrical	45 화난	Angry
23 섹시한	Sexy		

Last.FM 에서 제공하는 다양한 API 를 이용하면 해당 무드 태그에 해당하는 곡들을 확인할 수 있다. 아래는 light 태그에 대한 tag.getTopTracks API 의 예시이다.

[https://ws.audioscrobbler.com/2.0/?method=tag.gettoptracks&tag=TAG\\_NAME&api\\_key=YOUR\\_API\\_KEY&format=json](https://ws.audioscrobbler.com/2.0/?method=tag.gettoptracks&tag=TAG_NAME&api_key=YOUR_API_KEY&format=json)



그림 2 Last.FM 의 API 결과 예시

또한 아래 URL 처럼 무드 태그를 이용하여 관련 음악을 직접 검색하는 것도 가능하다.

<https://www.last.fm/tag/sentimental/tracks>

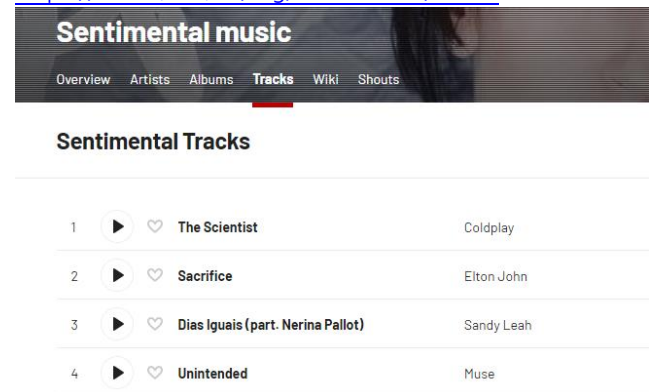


그림 3 Last.FM 태그 기반 음악 검색 결과

### III. 결론

본 논문에서는 음악 분류 목적의 딥러닝 모델 개발에 필요한 음악-무드 데이터셋 구성을 위해 다양한 공개 데이터셋에서의 공통의 무드 태그를 선별하고 정제된 한글-영문의 무드 태그를 제안하고 Last.FM 에서 무드 태그를 이용한 음악 검색 방법을 살펴보았다. 이를 이용하여 음악 무드에 기반한 자동태깅, 분류 및 유사음악 검색/추천 기술개발에 활용될 것으로 기대한다.

### ACKNOWLEDGMENT

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2022 년도 저작권기술 연구개발사업으로 수행되었음 (과제명: 교육 콘텐츠에 대한 인공지능 기반 저작권 침해 의심요소 검출 및 대체 재료 콘텐츠 추천 기술 개발, 과제번호: CR202104003)

### 참 고 문 헌

- [1] Gong, et. al., "Psla: Improving audio tagging with pretraining, sampling, labeling, and aggregation." IEEE/ACM Transaction on Audio, Speech, and Language Processing 29 (2021): 3292-3306.
- [2] [https://www.music-ir.org/mirex/wiki/2021:Audio\\_Tag\\_Classification](https://www.music-ir.org/mirex/wiki/2021:Audio_Tag_Classification)