

# 음원 메타데이터와 사용자 플레이리스트를 활용한 음악 추천 시스템

남경민<sup>1</sup>, 박유림<sup>2</sup>, 정지영<sup>3</sup>, 김도현<sup>4</sup>, 김현희<sup>3</sup>

<sup>1</sup> 동덕여자대학교 컴퓨터학과

<sup>2</sup> 동덕여자대학교 영어과

<sup>3</sup> 동덕여자대학교 정보통계학과

<sup>4</sup> 동덕여자대학교 경제학과

rudalswhdk12@naver.com, 20190140@dongduk.ac.kr, groundj00@gmail.com, deg7065@gmail.com, heekim@dongduk.ac.kr

## Music Recommendation System Using Audio Metadata and User Playlists

Kyoung Min Nam<sup>1</sup>, Yu Rim Park<sup>2</sup>, Ji Young Jung<sup>3</sup>, Do Hyeon Kim<sup>4</sup>, Hyon Hee Kim<sup>3</sup>

<sup>1</sup>Dept. of Computer Science, Dongduk Women's University

<sup>2</sup>Dept. of English, Dongduk Women's University

<sup>3</sup>Dept. of Information and Statistics, Dongduk Women's University

<sup>4</sup>Dept. of Economics, Dongduk Women's University

### 요 약

본 논문은 음원 메타데이터 임베딩 방법론을 기반으로 새로운 음원 추천 방법을 제안한다. 사용자 행동 데이터를 활용한 개인 맞춤형 음악 추천 모델은 신규 사용자의 데이터가 부족할 경우, 적절한 추천이 어려운 콜드스타트 현상을 초래할 수 있다. 본 연구에서는 플레이리스트의 음원 메타데이터를 Song sentence 로 구성하고, 고차원 벡터 공간에 임베딩하여 유사도를 계산한 추천 알고리즘을 구축한다. 사용자 행동 데이터가 아닌 음원의 자체적인 정보에 근거하기 때문에 콜드 스타트 현상을 보완하여 사용자에게 편리한 음악 감상 경험을 제공할 수 있을 것으로 기대된다.

### 1. 서론

온라인 스트리밍 서비스의 등장으로 사용자의 음악 취향에 정확히 부합하는 음원을 선별하기 위해 개인 맞춤형 추천 알고리즘이 적극적으로 연구되고 있다. 그러나 신규 사용자의 경우, 과거 검색 기록/좋아요 수/사용자 프로필/리뷰 등 사용자의 선호도를 정의할 수 있는 행동 데이터를 확보하기 어려워 여전히 콜드 스타트 현상[1]을 해결하지 못한다는 문제점이 존재한다.

본 연구에서는 이러한 문제점을 극복하기 위해 사용자 행동 데이터인 사용자 플레이리스트를 활용하고, 가수명, 장르, 작곡가, 작사가뿐만 아니라 음악의 '감정'과 같은 잠재적 속성까지 고려하여 음원 메타데이터를 구성하였다. 언어학 기반 모델인 SGNS (Skip-gram Negative Sampling) 알고리즘 [2]으로 메타데이터를 임베딩하고, 음원 간 유사도를 계산하여 음원 추천 알고리즘을 구축하였다.

음원의 특성을 음악 추천에 적용한 Song2Vec[3]은 사용자 행동데이터에 의존하지 않고 음악이 가진 자체적 특성, 즉 멜로디, 리듬, 화음, 화성 등을 기준으로 음원 고유 벡터를 임베딩하여 유사성을 계산한

다. 본 연구에서는 아티스트, 장르, 태그와 같은 표면적 속성 뿐만 아니라, 가사를 분석하여 음원에 내재된 감정까지 추천을 위한 정보로 활용하였다.

제안하는 추천 알고리즘은 음원이 자체적으로 보유한 고유 특성을 강조함으로써 신규 사용자들은 정확한 추천을 제공받을 수 있을 것이며, 새로운 음악으로 더욱 다양하고 풍부한 청취 경험을 할 수 있을 것으로 기대된다.

### 2. 데이터 수집 및 전처리

멜론 음원 플랫폼에서 다양한 태그들을 포함한 100개의 사용자 플레이리스트를 크롤링하였다. 총 2873개의 곡으로 Song Database 를 구성하였으며, [음원 ID, 제목, 순위, 가수명, 앨범명, 좋아요수, 발매일, 댓글수, 가사, 장르, 작곡가, 작사가, 편곡가, 태그리스트] 속성으로 구체화하여 데이터프레임으로 구축하였다. 발매일 칼럼은 '계절(Season)'과 '시기(Era)'로 분리하였다. 또한 '공포', '놀람', '분노', '슬픔', '중립', '행복', '혐오' 레이블로 사전 학습된 koBERT 모델에 소절 단위로 분리한 가사를 학습시켜 두 가지의 감정을 추출하고, 새로운 '감정(Emotion)' 칼럼으로 구축하였다.

### 3. 음악 추천 알고리즘

#### 3.1. Song sentence 벡터 임베딩

Song Database 의 새롭게 생성된 칼럼들을 활용하여 음원을 표현하는 문장인 Song sentence 를 구성하였다. 사용자 플레이리스트에 포함된 하나의 음원은 그 음원이 가진 메타데이터들 중 [가수명, 장르, 시기(Era), 계절(Season), 감정(Emotion), 태그리스트] 조합으로 하나의 Song sentence 가 생성된다.

Song sentence 는 SGNS (Skip-Gram with Negative Sampling) 알고리즘으로 고차원 벡터 공간에 임베딩하여 하나의 음원을 표현하는 고유 벡터를 생성하였다. SGNS 는 주어진 단어로 주변 단어를 예측하기 때문에 단어 간 의미적 관계를 포착할 수 있는 언어학적 기법이다[3]. window=5, negative=100 으로 조정하여 주어진 단어 주변의 문맥을 충분히 파악함과 동시에 특정 데이터에 과도하게 의존하는 데이터 불균형 문제를 예방하였다.

#### 3.2. Song sentence 기반 추천 알고리즘

생성된 Song sentence 는 3.1 에서 구축한 SGNS 모델을 학습시켜 벡터값을 추출하였다. 이 과정에서, [장르]에 해당하는 단어에 가중치 4 를 할당하여 음원의 특징을 결정적으로 구분하는 속성을 고려하고자 하였다. Song sentence 를 구성하는 단어 벡터들의 가중 평균값을 계산하고 한 음원에 해당하는 특징 벡터를 정의하였다. Song Database 의 음원 또한 위와 동일한 방법으로 특징 벡터를 생성하였다. 이후 사용자 플레이리스트의 음원 벡터와 Song Database 의 음원 벡터 사이의 코사인 유사도를 계산하고, 상위 5 개의 유사한 노래들을 추천 목록으로 포함시켰다. 추천 목록에 누락된 음원 중 가장 많이 추천된 음원들을 기준으로 내림차순 정렬하여 20 개의 음원으로 구성된 최종 추천 플레이리스트를 반환하였다.

#### 3.3. 성능 평가

n 개의 곡으로 구성된 사용자 플레이리스트에서 하나의 음원을 랜덤으로 제외하고, 총 n-1 개의 곡을 추천 시스템에 적용하였다. 최종 추천 플레이리스트에서, 앞서 제외한 음원이 포함(hit)되었는지의 여부를 파악하기 위해 전체 플레이리스트의 개수 대비 hit 한 플레이리스트 수의 비율을 정의하여 적중 점수를 도출하였다. 100 개의 모든 플레이리스트는 동일한 방식으로 성능평가를 진행하였으며, 최종적으로 0.87 의 높은 성능을 기록하였다.

### 4. 모바일 앱 구현

#### 4.1 모바일 애플리케이션 인터페이스

그림 1.은 모바일 애플리케이션에서 구현된 추천 알고리즘의 인터페이스를 보여준다. 이 애플리케이션은 사용자가 노래를 직접 검색하고 선택하여 플레이리스트를 생성할 수 있는 사용자 친화적 인터페이스를 제

공한다. 사용자는 이 인터페이스를 통해 자신의 음악 취향에 맞는 플레이리스트를 임의로 구성할 수 있다.



(그림 1.) 모바일 앱 구현 화면

#### 4.2 플레이리스트 처리 및 백엔드 통신

사용자가 생성한 플레이리스트는 음원 ID 목록 형태로 Google Cloud Platform 에 배포된 백엔드 모델로 전송된다. 이 통신은 API 를 통해 이루어지며, 백엔드 모델은 이 목록을 받아 추천을 수행한다. 백엔드에서는 Song sentence 기반의 추천 알고리즘을 사용하여 사용자의 취향의 적합한 새로운 음원 ID 목록을 생성한다.

추천 엔진은 처리된 결과를 음원 ID 목록으로 애플리케이션에 전달한다. 이 결과는 사용자의 화면에 동적으로 업데이트 되어 표시된다.

### 5. 결론

본 연구에서는 한 음원에 해당하는 메타데이터를 Song sentence 로 정의하여 벡터 공간에 임베딩하는 방법론을 제시하였다. 이 과정에서 메타데이터를 적극적으로 활용하여 음원을 표현할 수 있는 최적의 벡터값을 생성하였으며, 이를 기반으로 코사인 유사도를 측정해 음원을 추천해주는 알고리즘을 제안할 수 있었다. 사용자 행동 데이터가 아닌 사용자가 직접 구성한 플레이리스트 음원을 활용함으로써 음원의 자체적인 특성을 정확하고 다양하게 표현하고자 하였다.

이로써 새로 유입되어 데이터가 부족한 유저들도 정확한 추천을 제공받을 수 있을 것이며, 사용자가 청취했던 곡에 한정되지 않고 새로운 음악을 접할 기회를 제공하여 정보 편식 현상 문제를 보완할 수 있을 것으로 기대된다.

#### 참고문헌

- [1] Gope, Jyotirmoy & Jain, Sanjay. (2017). A survey on solving cold start problem in recommender systems. 133-138. 10.1109/CCAA.2017.8229786.
- [2] Mikolov, T., Chen, K., Corrado, G.S., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. International Conference on Learning Representations.
- [3] Ramakrishnan, P. (2017). song 2 vec : Determining Song Similarity using Deep Unsupervised Learning.