



특집논문 (Special Paper)

방송공학회논문지 제31권 제2호, 2026년 3월 (JBE Vol.31, No.2, March 2026)

<https://doi.org/10.5909/JBE.2026.31.2.218>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

키워드 기반 음악 추천 시스템의 설계 및 구현

봉 지 훈^{a)}, 김 화 정^{a)†}

Design and Implementation of a Keyword-based Music Recommendation System

Jihoon Bong^{a)} and Hwajung Kim^{a)†}

요 약

본 논문에서는 차트 및 과거 청취 기록 중심이었던 추천 시스템이 현재 사용자의 기분과 상황을 반영하지 못한다는 한계를 개선하고자, 키워드 기반 음악 추천 시스템을 제안한다. 제안하는 시스템은 사용자가 입력한 키워드를 곡의 가사와 세부 단어 단위 유사도, 문맥 의미 유사도를 각각 60%, 40%로 가중 결합한 뒤, 순위를 산출하여 추천한다. 제안한 음악 추천 시스템과 기존의 대표적인 3가지 음악 추천 시스템인 콘텐츠, 협업 필터링, 하이브리드 추천 시스템을 개별 곡 평가 및 리스트 평가를 통해 비교하였으며, 키워드 외에 연도 정보를 적용한 성능 평가를 수행하였다. 실험 결과, 추천 시스템 평가 실험에서 제안한 시스템이 기존 시스템 대비 개별 곡 평가에서 최대 53%, 리스트 평가에서 최대 51%로 높은 만족도를 보였다. 또한, 연도를 적용한 평가 결과 역시 개별 곡 평가 0.62점, 리스트 평가 7.73점으로 가장 높은 만족도를 보였다.

Abstract

In this paper, we propose a keyword-based music recommendation system to address the limitations of existing recommendation systems that are primarily centered on charts and historical listening records and thus fail to reflect users' current moods and situational contexts. The proposed system generates ranked recommendations by combining word-level similarity between user-input keywords and song lyrics with contextual semantic similarity, weighted at 60% and 40%, respectively. The proposed music recommendation system is compared with three representative existing approaches content-based, collaborative filtering, and hybrid recommendation systems through both item-level and list-level evaluations, in addition, a performance evaluation incorporating release-year information alongside keyword-based recommendation is conducted. Experimental results show that, compared to existing systems, the proposed approach achieves improvements of up to 53% in item-level evaluation and up to 51% in list-level evaluation. Furthermore, the evaluation incorporating the release-year information demonstrates the highest level of satisfaction, with scores of 0.62 for individual song evaluation and 7.73 for recommendation list evaluation.

Keyword : Music recommendation system, Keyword-based music recommendation, User-centric evaluation, Context-aware music recommender, Cosine similarity

Copyright © 2026 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

1. 서론

최근 음악 소비의 흐름은 과거의 차트 중심적 구조에서 벗어나, 개인의 취향과 상황을 중시하는 방향으로 변화하고 있다^[1]. 스트리밍 플랫폼의 확산으로 사용자는 언제든 원하는 음악을 들을 수 있게 되었으며, 이에 따라 음악 추천 시스템은 단순한 편의 기능을 넘어 개인화된 청취 경험을 제공하는 핵심 기술로 자리 잡고 있다^[2]. 이러한 배경 속에서 음악 추천 시스템 분야에서는 콘텐츠 기반 협업 필터링, 그리고 두 방식을 결합한 하이브리드 추천 시스템이 주로 활용되고 있다^[2]. 콘텐츠 기반 추천 시스템은 가사와 메타데이터를 활용해 사용자의 선호 곡과 유사한 곡을 추천하고, 협업 필터링은 사용자 간 선호 곡의 교집합을 이용해 추천을 생성하며, 하이브리드 방식은 두 접근법을 가중 결합하여 추천을 생성한다^[3]. 기존의 음악 추천 시스템은 주로 사용자의 과거 청취 이력이나 선호 데이터를 기반으로 작동한다. 그러나 이와 같은 방식은 장기적인 취향을 반영하는 데에는 효과적이지만, 현재의 기분이나 일시적인 상황과 같은 맥락적 요소를 반영하기에는 한계가 있다^[4].

이러한 한계를 보완하기 위해, 본 연구는 사용자의 현재 상태를 직접적으로 반영할 수 있는 키워드 기반 음악 추천 시스템을 제안한다. 제안하는 시스템은 사용자가 입력한 키워드를 세부 단어 단위 유사도와 문맥 의미 유사도를 기반으로 함께 고려하여 음악을 추천한다. 결과적으로 사용자는 장기 선호로 학습된 프로필 위에 현재 상황을 잘 드러내는 키워드에 대한 가중치를 부여함으로써, 사용자가 평소 좋아하는 음악이 아닌 지금 이 순간 듣고 싶은 음악을

추천받을 수 있도록 한다.

제안하는 키워드 기반 추천 시스템은 콘텐츠 기반, 협업 필터링, 하이브리드 추천 시스템과 비교하는 평가 실험과 연도 정보를 적용한 실험을 통해 성능을 검증하였다. 평가 지표로는 오프라인 평가 방식이 아닌 실제 청취 경험에 기반한 사용자 평가 지표인 Precision@5로 측정하는 개별 곡 평가와 추천 리스트의 전반적인 분위기 만족도를 측정하는 10 Scale Score인 리스트 평가를 사용하였다. 평가 실험 결과, 제안한 키워드 기반 추천 시스템은 개별 곡 평가에서 평균 0.577점, 리스트 평가에서 평균 7.2점으로 네 가지 추천 방식 중 가장 높은 성능을 기록하였으며, 기존 추천 시스템 대비 개별 곡 평가에서는 최대 53%, 리스트 평가에서는 최대 51%의 성능 향상을 보였다. 연도 정보를 추가 적용한 실험에서도 제안한 시스템은 개별 곡 평가에서 0.62점, 리스트 평가에서 7.73점을 기록하였다. 이러한 결과는 키워드 기반 추천 시스템이 사용자의 현재 기분과 상황을 직접적으로 반영함으로써, 기존 추천 방식이 가지는 한계를 보완할 수 있음을 보여준다.

II. 연구 배경 및 동기

1. 콘텐츠 기반 추천 시스템

콘텐츠 기반 추천 시스템은 곡의 메타데이터인 장르, 아티스트, 발매 연도와 가사를 Word2Vec^[5]을 이용하여 학습하고, 사용자 선호 곡들의 가사 평균 벡터와 후보 곡 가사 평균 벡터를 코사인 유사도^[6]로 계산한다. 코사인 유사도는 두 벡터 방향이 얼마나 비슷한지를 측정하는 지표로, 식 (1)과 같이 정의된다.

$$\cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (1)$$

또한, 장르 정보를 반영하기 위해 One-hot encoding을 적용하였다^[7]. One-hot encoding은 각 장르를 0과 1로 구분하는 방식으로 코사인 유사도 계산값에 추가 점수 형

a) 서울과학기술대학교 스마트ICT융합공학과(The Department of Smart ICT Convergence Engineering, Seoul National University of Science and Technology)

‡ Corresponding Author : 김화정(Hwajung Kim)

E-mail: hwajung.kim@seoultech.ac.kr

Tel: +82-2-970-6878

ORCID: <https://orcid.org/0000-0001-7134-823X>

※ 이 성과는 2025년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원받아 수행된 결과입니다(RS-2025-2453332).

※ 이 논문의 연구 결과 중 일부는 한국방송·미디어공학회 2025년 추계 학술대회에서 발표한 바 있음.

· Manuscript January 6, 2026; Revised February 2, 2026; Accepted February 9, 2026.

태로 반영되어, 가사와 장르를 함께 고려한 추천이 이루어진다. 예를 들어 사용자가 선호하는 곡의 장르가 발라드인 경우, 동일한 가사를 가지더라도 발라드 장르로 분류된 곡이 다른 장르의 곡보다 더 높은 추천 점수를 부여 받는다. 콘텐츠 추천 시스템 방식은 신규 아이템 대응에 유리하지만, 본질적으로 장기 선호 프로필에 의존하기 때문에 현재의 기분이나 상황의 변화에는 민감하게 반응하기 어렵다.

2. 협업 필터링 추천 시스템

협업 필터링 추천 시스템은 사용자-사용자 간 상호작용에 기반하여, 선호 곡이 유사한 다른 사용자의 아이템을 추천해 주는 방식이다^[8]. 본 연구에서는 사용자 간 선호 곡의 교집합을 이용하여 추천 후보를 생성하고, 교집합이 여러 사용자와 동시에 존재하는 경우 각 사용자로부터의 추천 후보 곡을 번갈아 가며 순차적으로 배치하였다. 해당 방식은 추천 곡이 과도하게 집중되는 현상을 방지하고, 여러 사용자의 취향이 추천 리스트에 균형 있게 반영되도록 하는 장점이 있다. 이를 통해 추천 다양성을 일정 수준 확보하고자 하였다. 그러나 이러한 방식에도 불구하고, 사용자가 선호한 곡의 수가 적은 경우와 같은 데이터 희소성 문제나 사용자 간 교집합이 존재하지 않는 구간에서는 추천이 불안정하거나 생성되지 않는 한계를 지닌다^[9].

3. 하이브리드 추천 시스템

하이브리드 추천 시스템은 그림 1과 같이 콘텐츠 기반 추천 시스템과 협업 필터링 추천 시스템을 가중 결합하여 추천의 정확도와 다양성을 높이는 방식이다^[10]. 이러한 하이브리드 추천 시스템에서는 두 추천 방식의 기여도를 조절하기 위한 가중치 설정이 중요한 요소로 작용하며, 해당 가중치에 대한 최적 값은 존재하지 않고 적용 환경과 목적에 따라 휴리스틱하게 결정되는 경우가 많다^[11]. 본 연구에서는 최적 가중치 설정을 위해 콘텐츠 기반과 협업 필터링 추천 방식의 가중치를 각각 5:5와 6:4로 설정한 추천 시스템의 만족도 평가를 사전 실험으로 진

행하였다. 각 실험의 평균 만족도가 37% 및 48%로 측정되어, 콘텐츠 기반과 협업 필터링 결과를 각각 60%와 40%로 가중 결합하여 추천 안정성과 다양성을 균형적으로 반영하고자 하였다. 다만, 이러한 하이브리드 방식 또한 사용자의 현재 기분이나 상황적 맥락을 즉시 반영하기 어렵다는 한계가 있다.

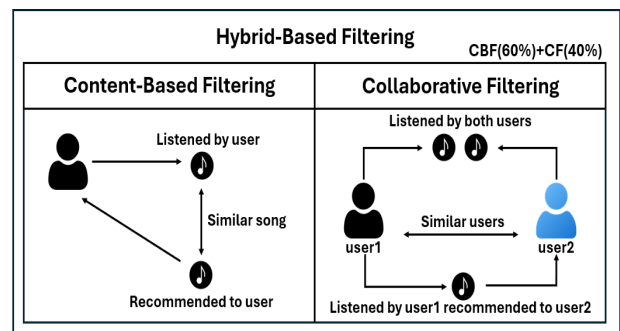


그림 1. 기존 음악 추천 시스템의 전체 매커니즘
Fig. 1. Overall mechanism of the conventional music recommendation system

4. 연구 동기

본 연구는 기존 음악 추천 시스템이 주로 사용자의 과거 청취 기록과 선호 프로필에 기반하여 작동함으로써, 사용자가 지금 이 순간 느끼는 기분이나 처한 상황과 같은 맥락적 요소를 즉시 반영하는 데 한계를 가진다는 문제를 해결하고자 한다. 콘텐츠 기반 추천은 가사와 메타데이터 유사도를 통해 추천을 제공하지만, 현재 맥락의 변화에 둔감하고, 협업 필터링은 사용자 간 교집합에 의존하여 데이터 희소성에 취약하며 추천이 생성되지 않는 경우가 발생할 수 있다. 하이브리드 방식은 두 접근을 결합해 성능을 개선하려 하지만, 기본적으로 과거 선호 중심의 구조를 벗어나기 어렵기 때문에 “사용자가 평소 좋아하는 음악”은 잘 맞추더라도 “지금 듣고 싶은 음악”을 직접적으로 반영하기에는 부족하다. 이에 본 연구는 사용자의 현재 상태를 키워드로 입력받아, 키워드와 가사 간 세부 단어 단위 유사도와 문맥 의미 유사도를 결합하여 추천을 생성함으로써 맥락 반영의 실효성을 높이고자 한다.

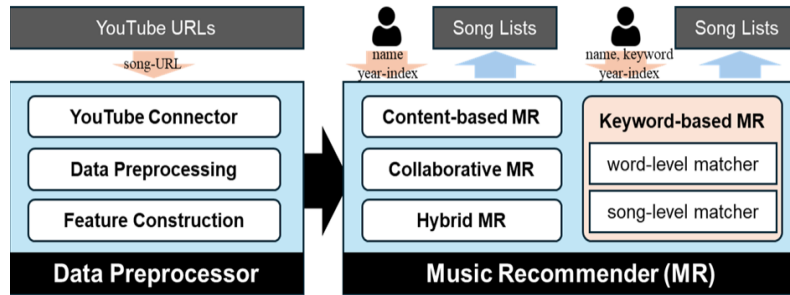


그림 2. 음악 추천 시스템의 전체 아키텍처
 Fig. 2. Overall architecture of the music recommendation system

III. 시스템 설계 및 구현

1. 시스템 아키텍처

그림 2는 본 연구에서 사용자 중심 평가를 위해 구성된 음악 추천 시스템의 전체 구조를 나타낸다. 해당 추천 시스템은 데이터 전처리기(Data Preprocessor)와 음악 추천기(Music Recommender, MR)로 구성된다. 데이터 전처리기에서는 사용자의 음악 접근성 향상을 위해 YouTube Data API를 사용하여 곡의 URL을 수집하고, 데이터 전처리 단계(Data Preprocessing)에서는 가사 텍스트와 메타데이터를 저장한다. 이후, Word2Vec 기반 가사 임베딩을 통해 특징 구성(Feature Construction) 단계를 수행한 뒤, 음악 추천을 수행하게 된다. 음악 추천기에서는 사용자 명, 연도 정보를 바탕으로 코사인 유사도 기반 콘텐츠 추천, 사용자-사용자 기반 협업 필터링, 가중 결합 방식의 하이브리드 추천 결과를 제공한다. 연도 정보는 5년 단위의 구간으로 구분하여, 사용자는 하나 이상의 구간을 중복하거나 비연속적으로 선택할 수 있다. 제안하는 추천 시스템은 사용자에게 키워드를 추가로 입력받아 각 음악의 가사 내 단어 간 word-level matcher, 각 음악 가사 평균 벡터 간 song-level matcher를 각각 계산하여 추천 결과를 제공한다. 각 추천 시스템은 입력된 정보를 바탕으로 점수가 높은 상위 5곡(Top-5)을 순위화하여 추천 리스트로 제공한다. 음악 추천 결과는 그림 3과 같이 제공된다.

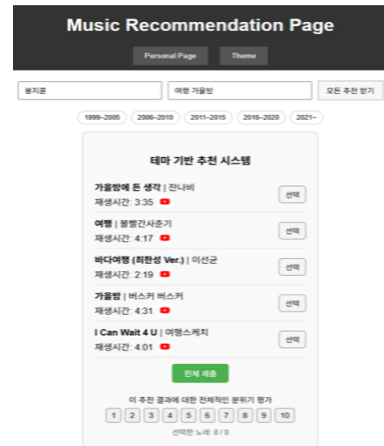


그림 3. 음악 추천 결과
 Fig. 3. Music Recommendation Results

2. 키워드 기반 추천 시스템

제안하는 키워드 기반 추천 시스템은 단순한 키워드 포함 여부를 확인하는 기존의 텍스트 검색 방식과 달리, 고차원 벡터 공간에서의 세부 단어 단위 유사도와 문맥 의미 유사도를 산출하여 추천을 수행한다. 이는 사용자가 입력한 키워드와 가사의 단어가 문맥적으로 유사한 의미를 지닌 곡을 찾아낼 수 있다는 점에서 기존의 텍스트 검색 방식과 차별점을 가진다. 세부 단어 단위 유사도는 각 키워드와 가사 토큰 400개의 단어 벡터 간 코사인 유사도를 통해 점수를 반환하고, 문맥 의미 유사도는 키워드 평균을 벡터화하여 가사 평균 벡터와 코사인 유사도를 계산한다. 이때, 키워드 평균 벡터는 개별 키워드 점수들의 평균을 내어 하

나의 벡터로 만든 것이다. 제안하는 시스템은 사용자가 입력한 키워드의 의미가 추천 결과에 우선 반영되도록 세부 단어 단위 유사도에 60%, 전체 가사의 흐름을 파악하는 문맥 의미 유사도에 40%의 가중치를 부여하였다. 가중치를 5:5로 설정할 경우, 키워드 본연의 의미가 희석되고, 세부 단어 단위 유사도에 70% 이상의 가중치를 부여할 경우, 특정 단어의 포함 여부에 치중하여 추천 품질이 저하되는 것을 확인하였다. 결과적으로 6:4의 비율은 키워드의 의미를 전달하면서도 곡 전체의 분위기를 반영하는 가중치로 확인되어 이를 식 (2)와 같이 설정하였다.

$$S_{keyword} = 0.6_{word} + 0.4_{avg} \quad (2)$$

제안하는 시스템은 입력된 모든 키워드를 조합한 벡터를 생성하여 추천에 활용한다. 예를 들어, 동일한 슬픔 키워드라도 ‘슬픔, 행복’과 ‘슬픔, 우울’은 서로 다른 벡터를 생성한다. 제안하는 시스템은 이 차이를 코사인 유사도 계산을 통해 다른 감정의 곡을 추천한다. 곡의 가사를 평균 벡터로 만들기 이전, 가사에 불필요한 기호를 제거하고, 한글과 영문만 남긴 후 소문자로 변환하여 전처리한다. 이후 각 가사를 Word2Vec으로 벡터화한 뒤 평균을 계산해 곡별 사전 임베딩 벡터를 생성하며, 이는 키워드 - 곡 간 유사도 계산 시간을 줄이기 위함이다. Word2Vec은 `vector size=300`, `window=10`, `min_count=10`, `workers=5`, `epochs=50`으로 설정하여 학습하였다^[12]. `Vector size=300`은 단어 간 의미 관계를 안정적으로 표현하면서 계산 복잡도를 과도하게 증가시키지 않는다^[13]. `Window=10`은 한 단어를 학습할 때 양쪽으로 최대 10개의 단어를 함께 고려하는 값으로, 가사 한 줄 정도의 문맥을 반영하기에 적절하다^[13]. `Min count=10`은 출현 빈도가 매우 낮은 단어를 제외하여 노이즈를 줄이기 위한 설정이며, `epochs=50`은 모델이 수렴하도록 반복 학습을 진행하기 위한 값이다.

이렇게 전처리와 Word2Vec 과정을 걸쳐 세부 단어 단위 유사도와 문맥 의미 유사도를 점수화한 키워드 점수에 콘텐츠 기반 추천 시스템과 유사한 사용자별 개인화 점수를 합친다. 개인화 점수는 해당 사용자가 선호한 10곡의 평균 임베딩과 후보 곡 임베딩의 코사인 유사도를 이용하여 계

산한다. 이는 동일한 키워드임에도 불구하고 각 사용자의 기본적인 음악 취향을 반영하기 위해 가중치 20%를 부여하였다. 결과적으로, 최종 점수는 식 (3)과 같이 키워드 점수와 개인화 점수를 각각 80%와 20% 가중치로 결합하고, 키워드가 제목에 포함되면 추가 점수인 `boost`를 2%(0.02)만큼 부여하여 더 밀접하게 추천해 준다. 이후 점수를 내림차순으로 정렬하여 최종 추천 리스트를 반환한다.

$$S_{final} = (0.8S_{keyword} + 0.2S_{user}) + 0.02boost \quad (3)$$

IV. 실험 및 평가

1. 실험 환경

본 연구의 실험은 총 60명의 참여자를 두 개의 그룹으로 나누어 수행하였다. 첫 번째 그룹인 전체 60명은 기존 3가지 추천 시스템과 제안하는 키워드 기반 추천 시스템의 성능 평가에 참여하였으며, 이 중에서 30명은 동일한 평가 실험에 참여한 후, 연도 정보를 추가로 적용한 경우의 추천 성능 평가를 진행하여, 추가 정보에 의한 추천 성능 평가에 참여하였다. 실험에 사용된 음악 데이터셋은 2000년부터 2024년까지의 멜론 Top 100 차트 곡을 기반으로 하며, 중복 곡을 제외한 총 7,720곡으로 구성되었다. 실험 전, 각 참여자는 본인이 선호하는 10곡을 입력하도록 하였으며, 이를 통해 총 600곡의 개인 선호 데이터가 사전에 구축되었다. 해당 데이터는 이후 기존 3가지 추천 시스템과 제안하는 키워드 기반 추천 시스템에서 개인화 추천의 기준으로 활용되었다.

2. 실험 평가 지표

추천 결과에 대한 평가는 두 가지 정량적 지표를 활용한다. 첫 번째로 사용자가 개별 곡을 평가하는 `Precision@5`는 선호하는 곡을 기반으로 각 추천 시스템을 통해 사용자에게 추천된 전체 곡 중에서 사용자가 실제로 선호한 곡으로 선택하였는지에 관한 지표이다. 예를 들어, 사용자에게 추천한 5곡 중 사용자가 3곡을 선호 곡으로 선택한 경우, `Precision@5`는 0.6으로 계산된다. 두 번째 지표는 추천된

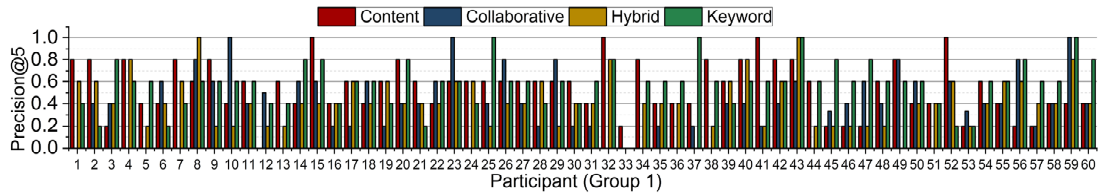


그림 4. 각 추천 방식별 개별 곡 평가(Precision@5) 비교 결과

Fig. 4. Comparison of item-level evaluation results(Precision@5) across recommendation approaches

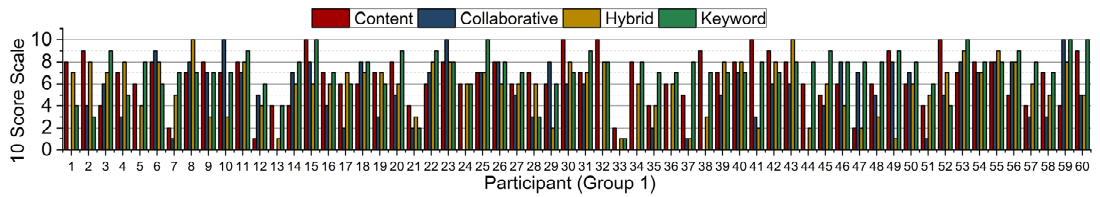


그림 5. 각 추천 방식별 리스트 평가(10 Scale Score) 비교 결과

Fig. 5. Comparison of list-level evaluation results(10 Scale Score) across recommendation approaches

음악 리스트에 대한 평가 결과로, 추천된 음악 리스트에 대해 1에서 10 사이의 점수를 부여하여 각 추천 리스트에 대한 평가를 진행한다. 음악 추천 시스템은 정량적 지표로 성능을 평가하는 것도 중요하지만, 음악은 결국 사람이 느끼는 경험이 핵심이기 때문에 본 연구에서는 정답 레이블 기반의 실험이 아닌 실제 사용자의 체감 평가를 중심으로 진행하였다^[14]. 이러한 이유로 Recall과 F1-score는 정답 레이블이 존재하는 실험에 적합한 지표이므로 본 연구에서는 사용하지 않았다.

3. 실험 결과 및 분석

3.1 평가 실험 결과 및 분석

그림 4는 추천 방식별 개별 곡 평가 결과를, 그림 5는 리스트 평가 결과를 각각 참여자 단위로 나타낸 것이다. 그림 6 (a)와 그림 6 (b)는 이러한 결과를 추천 시스템별 평균값으로 요약한 그래프이다. 개별 곡 평가인 그림 6 (a)에서 키워드 기반 추천 시스템은 평균 0.577점으로 가장 높은 성능을 보였으며, 콘텐츠 기반(0.547점), 협업 기반(0.376점), 하이브리드(0.42점) 대비 각각 6%, 53%, 37% 높은 점수를 기록하였다. 그림 6 (b) 리스트 평가에서도 평균 7.2점으로 최고 성능을 보였으며, 콘텐츠, 협업, 하이브리드 대비 각각 10%, 51%, 25% 높은 수치를 나타냈다.

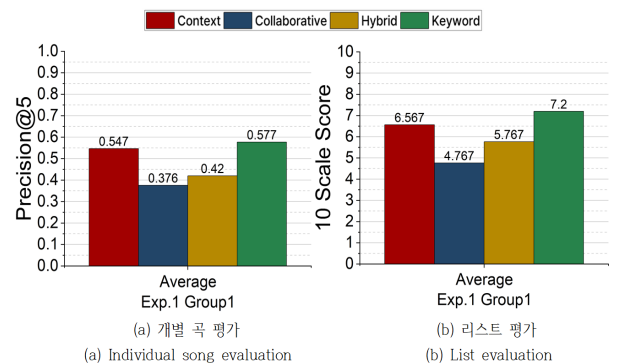


그림 6. 평가 실험 결과

Fig. 6. Evaluation experiment results

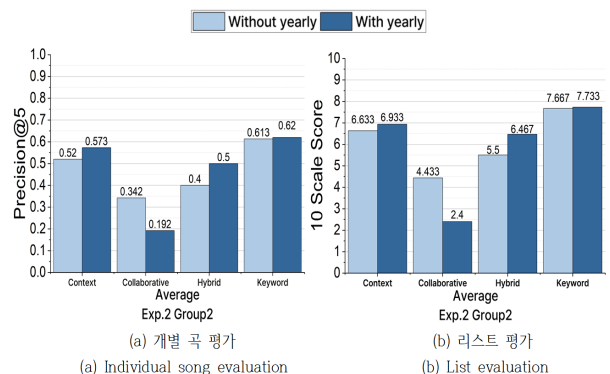


그림 7. 연도 정보 적용 실험 결과

Fig. 7. Experiment results with year information

콘텐츠 기반 추천 시스템은 개별 곡 평가에서는 평균 0.547점을 기록하였으며, 리스트 평가에서도 평균 6.567점을 기록하였지만, 키워드 기반 추천 시스템은 9.6% 향상된 평균 7.2점을 기록하였다. 이는 콘텐츠 기반 추천이 가사 및 장르 중심으로 이루어져, 키워드가 반영하는 맥락 및 상황 정보가 고려되지 않아 사용자 기대와의 불일치로 만족도가 낮아진 것으로 해석된다. 협업 필터링 추천 시스템은 사용자 간 교집합에 기반하나, 그림 4와 5에서 확인할 수 있듯이 실험 참여자 60명 중 10명(16.7%)이 교집합이 없어 추천이 생성되지 않았고, 이로 인해 개별 곡 및 리스트 평가 모두 가장 낮은 성능을 보였다. 이는 협업 필터링 추천 시스템이 사용자 간 교집합에 의존하는 특성으로 데이터 희소성 문제에 취약함을 보여준다⁸⁾. 하이브리드 추천 시스템은 개별 곡 평균 점수에서 협업 기반보다 높고 콘텐츠 기반보다는 낮은 성능을 보였다. 리스트 평가에서는 콘텐츠 기반과 유사한 수준을 유지했으나, 키워드 기반 추천 시스템은 이에 비해 약 25% 향상된 성능을 나타냈다. 하이브리드 추천 시스템은 협업 필터링의 사용자 교집합 부족 문제가 하이브리드 추천 시스템에도 반영되어 한계를 보였다.

3.2 연도 정보 추가 적용 실험 결과 및 분석

그림 7의 결과에 따르면, 연도 정보를 적용한 콘텐츠 기반 추천은 개별 곡 및 리스트 평가에서 성능이 향상되었으나, 여전히 키워드 기반 추천 시스템보다 낮은 성능을 보였다. 협업 필터링 추천 시스템은 개별 곡 평가와 리스트 평가에서 연도 정보 적용 시 오히려 성능이 감소하는 경향을 보였으며, 네 가지 추천 방식 중 가장 낮은 성능을 기록하였다. 이는 연도 정보 추가로 추천 후보가 제한되면서 사용자 교집합이 감소해 데이터 희소성 문제가 심화되고, 그 결과 추천 생성 실패 및 전반적인 성능 저하가 발생했기 때문이다. 하이브리드 추천 시스템은 개별 곡 평가에서 협업 필터링보다 높고 키워드 기반보다는 낮은 성능을 보였으며, 리스트 평가에서는 콘텐츠 기반과 유사한 수준을 나타냈다. 연도 정보 적용으로 성능은 향상되었으나, 최종적으로는 키워드 기반 추천 시스템에 미치지 못했다.

연도 정보를 추가로 입력받아 추천에 적용한 결과, 협업 필터링 추천 시스템은 연도 정보 적용 시 추천 후보 곡의 범위가 더욱 제한되어 데이터 희소성 문제가 심화되어 성

능이 오히려 감소하는 경향을 보였다. 반면 콘텐츠 기반 및 하이브리드 추천 시스템에 연도 정보를 적용함으로써 개별 곡 평가와 리스트 평가 모두에서 성능 향상이 관찰되었다. 그러나 연도 정보를 적용하지 않은 키워드 기반 추천 시스템이 개별 곡 평가에서 0.61점, 리스트 평가에서 7.67점으로 연도 정보를 적용한 콘텐츠 기반 및 하이브리드 추천 시스템보다 높은 성능을 기록하였다. 이를 통해, 키워드뿐만 아니라 연도 정보를 사용자로부터 입력받는 경우, 평가 시점에서의 기본 및 상황적 맥락을 사용자 입력을 통해 추천에 반영함으로써, 기존 추천 시스템이 가지는 한계를 보완함을 확인할 수 있다.

V. 결론 및 향후 연구

본 연구에서는 사용자의 현재 기분이나 상황을 반영하는 키워드 기반 추천 시스템을 제안하고, 기존의 콘텐츠 기반, 협업 필터링 기반, 하이브리드 기반의 3가지 음악 추천 시스템과 비교 분석 실험을 수행하였다. 평가 결과, 제안한 시스템은 개별 곡 평가에서 콘텐츠 기반 대비 6%, 하이브리드 기반 대비 37%, 협업 기반 대비 최대 53%로 높은 성능을 보였다. 또한, 리스트 평가에서도 콘텐츠 기반 대비 10%, 하이브리드 기반 대비 25%, 협업 기반 대비 최대 51%로 높은 만족도를 기록하였다. 또한, 연도 정보를 추가로 입력한 경우, 제안하는 추천 시스템이 개별 곡 평가에서 0.62점, 리스트 평가에서 7.73점으로 기존 시스템 대비 전반적으로 우수한 품질의 추천 결과를 확인하였다.

본 연구에서 실시한 평가는 사용자 중심의 만족도를 확인하였다는 점에서 유의미하나, 객관적 지표가 부족하다는 한계가 있다. 향후 연구에서는 사용자 선호도 정보가 포함된 Last.fm 등의 공개 데이터를 활용하고 추천 시스템의 주요 평가지표인 Recall, F1-score를 도입하여 시스템의 성능을 정량적으로 평가하여 검증할 계획이다¹⁵⁾. 또한, BPM 등의 음악적 요소를 반영하여 정서적 기반을 반영한 세부적인 추천이 가능하도록 발전시킬 예정이다. 이후, 사용자 피드백을 활용한 개인화 가중치 학습 기법을 적용해, 상황과 취향에 더욱 정밀하게 대응하는 추천 시스템으로 확장할 계획이다.

참 고 문 헌 (References)

- [1] J. Webster, "The promise of personalisation: Exploring how music streaming platforms are shaping the performance of class identities and distinction.", *New Media & Society*, Vol.25, No.8, pp.2140-2162, 2021.
doi: <https://doi.org/10.1177/14614448211027863>
- [2] Z. Zeng, Y. Wang, Y. Zhao, & W. Shi, "A survey of music recommendation systems.", In *Proceedings of the 5th International Conference on Computer Information and Big Data Applications*, Wuhan, China, pp. 507-519, April 2024.
doi: <https://doi.org/10.1145/3671151.3671243>
- [3] S. Bostandjiev, J. O'Donovan, & T. Höllerer, "TasteWeights: a visual interactive hybrid recommender system.", In *Proceedings of the sixth ACM conference on Recommender systems*, Dublin, Ireland, pp. 35-42, September 2012.
doi: <https://doi.org/10.1145/2365952.2365964>
- [4] X. Meng, Y. Du, Y. Zhang and X. Han, "A Survey of Context-Aware Recommender Systems: From an Evaluation Perspective," *IEEE Transactions on Knowledge and Data Engineering*, Vol.35, No.7, pp. 6575-6594, July 2023.
doi: <https://doi.org/10.1109/TKDE.2022.3187434>
- [5] T. Mikolov, K. Chen, G. Corrado, & J. Dean, "Efficient estimation of word representations in vector space", *arXiv preprint arXiv:1301.3781*, 2013.
doi: <https://doi.org/10.48550/arXiv.1301.3781>
- [6] G. Salton, & C. Buckley, "Term-weighting approaches in automatic text retrieval." *Information processing and management*, Vol.24, No.5, pp. 513-523, 1988.
doi: [https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0)
- [7] Y. Qu, H. Cai, K. Ren, W. Zhang, Y. Yu, Y. Wen, & J. Wang, "Product-based neural networks for user response prediction.", In *2016 IEEE 16th international conference on data mining (ICDM)*, Barcelona, Spain, pp. 1149-1154, December 2016.
doi: <https://doi.org/10.1109/ICDM.2016.0151>
- [8] U. Shardanand, & P. Maes, "Social information filtering: Algorithms for automating "word of mouth".", In *Proceedings of the SIGCHI conference on Human factors in computing systems*, Denver, Colorado USA, pp. 210-217, May 1995.
doi: <https://doi.org/10.1145/223904.223931>
- [9] A. Popescul, L. H. Ungar, D. M. Pennock, & S. Lawrence, "Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments." *arXiv preprint arXiv:1301.2303*, 2013.
doi: <https://doi.org/10.48550/arXiv.1301.2303>
- [10] R. Burke, "Hybrid Recommender Systems: Survey and Experiments.", *User modeling and user-adapted interaction*, Vol.12, No.4, pp. 331 - 370, 2002.
doi: <https://doi.org/10.1023/A:1021240730564>
- [11] G. Adomavicius, & A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions." *IEEE transactions on knowledge and data engineering*, Vol.17, No.6, pp. 734-749, 2005.
doi: <https://doi.org/10.1109/TKDE.2005.99>
- [12] Gensim: Word2Vec model documentation, Official Gensim Documentation, 2024, <https://radimrehurek.com/gensim/models/word2vec.html> (accessed Dec, 18, 2025)
- [13] D. Jatnika, M. A. Bijaksana, & A. A. Suryani, "Word2vec model analysis for semantic similarities in english words", *Procedia Computer Science*, pp. 157, 160-167, 2019.
doi: <https://doi.org/10.1016/j.procs.2019.08.153>
- [14] R. Ungruh, K. Dinnissen, A. Volk, M. S. Pera, & H. Hauptmann, "Putting popularity bias mitigation to the test: A user-centric evaluation in music recommenders.", In *Proceedings of the 18th ACM Conference on Recommender Systems*, Bari, Italy, pp. 169-178, October 2024.
doi: <https://doi.org/10.1145/3640457.3688102>
- [15] D. Roy, & M. Dutta, "A systematic review and research perspective on recommender systems.", *Journal of Big Data*, Vol.9, No.59, May 2022.
doi: <https://doi.org/10.1186/s40537-022-00592-5>

저 자 소 개



봉 지 훈

- 현재 : 서울과학기술대학교 스마트ICT융합공학과 학사과정
- ORCID : <https://orcid.org/0009-0008-2866-8957>
- 주관심분야 : 추천시스템, 데이터베이스 최적화



김 화 정

- 2023년 2월 : 서울대학교 컴퓨터공학과 박사
- 2023년 8월 ~ 현재 : 서울과학기술대학교 스마트ICT융합공학과 조교수
- ORCID : <https://orcid.org/0000-0001-7134-823X>
- 주관심분야 : 운영체제, 분산시스템, 데이터베이스 시스템