



가사 유사도 기반 음악 추천 알고리즘

Music Recommendation Algorithm based on Lyrics Similarity

자비랩 - 한국인공지능연구소 오픈랩 8기

Pitch Date: 4 / 17 / 2021

MUSIC RECOMMENDATIONS

- 1 연구원 소개
- 2 연구의 필요성
- 3 연구의 목표
- 4 연구 방법론
- 5 연구의 과정
- 6 Demo 시연
- 7 결론 및 고찰

연구원 소개



구선민
Developer



이동인
Developer



김지훈
LAB 대표



장상근
Researcher

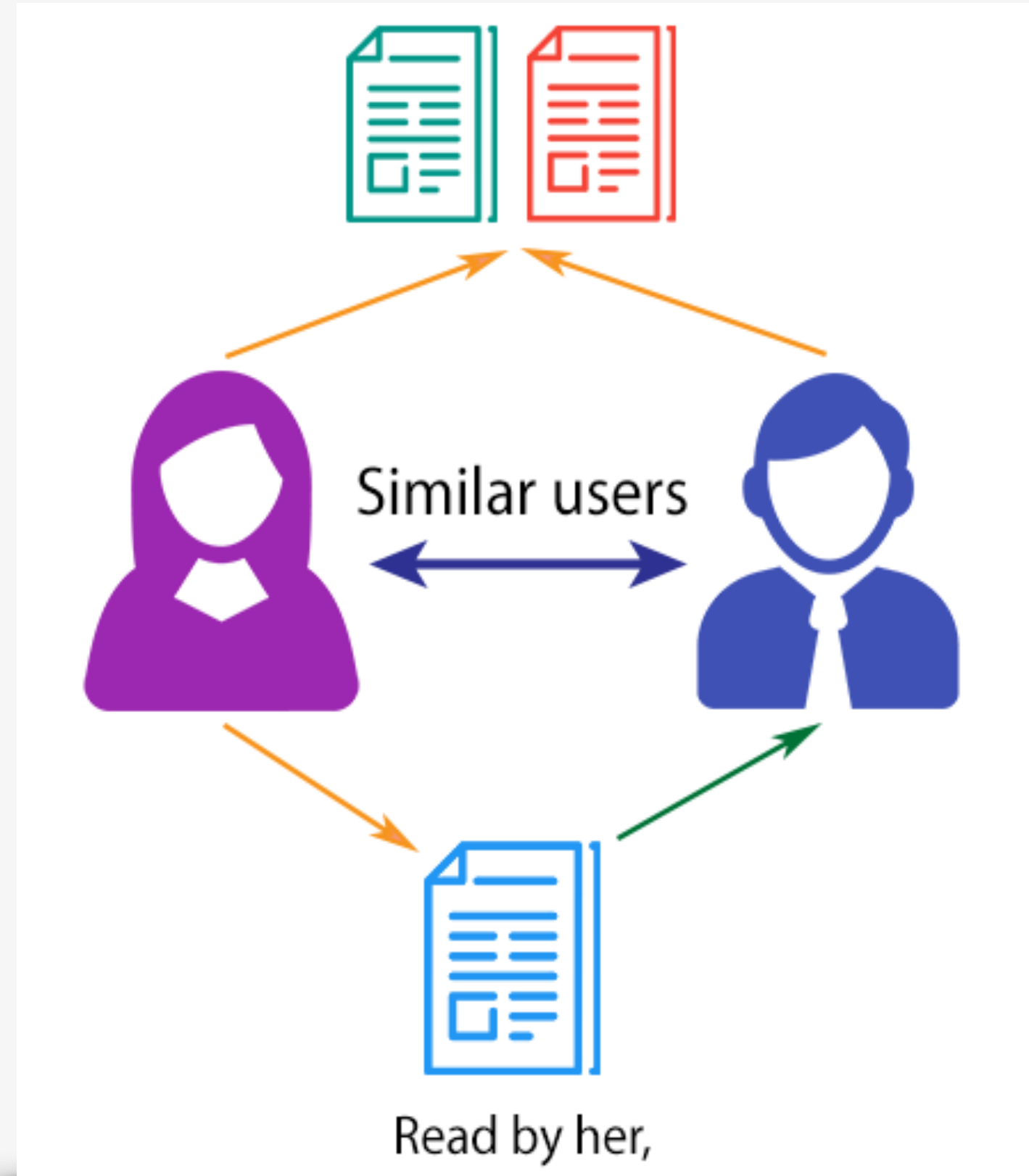


홍지완
Researcher

2. 연구의 필요성

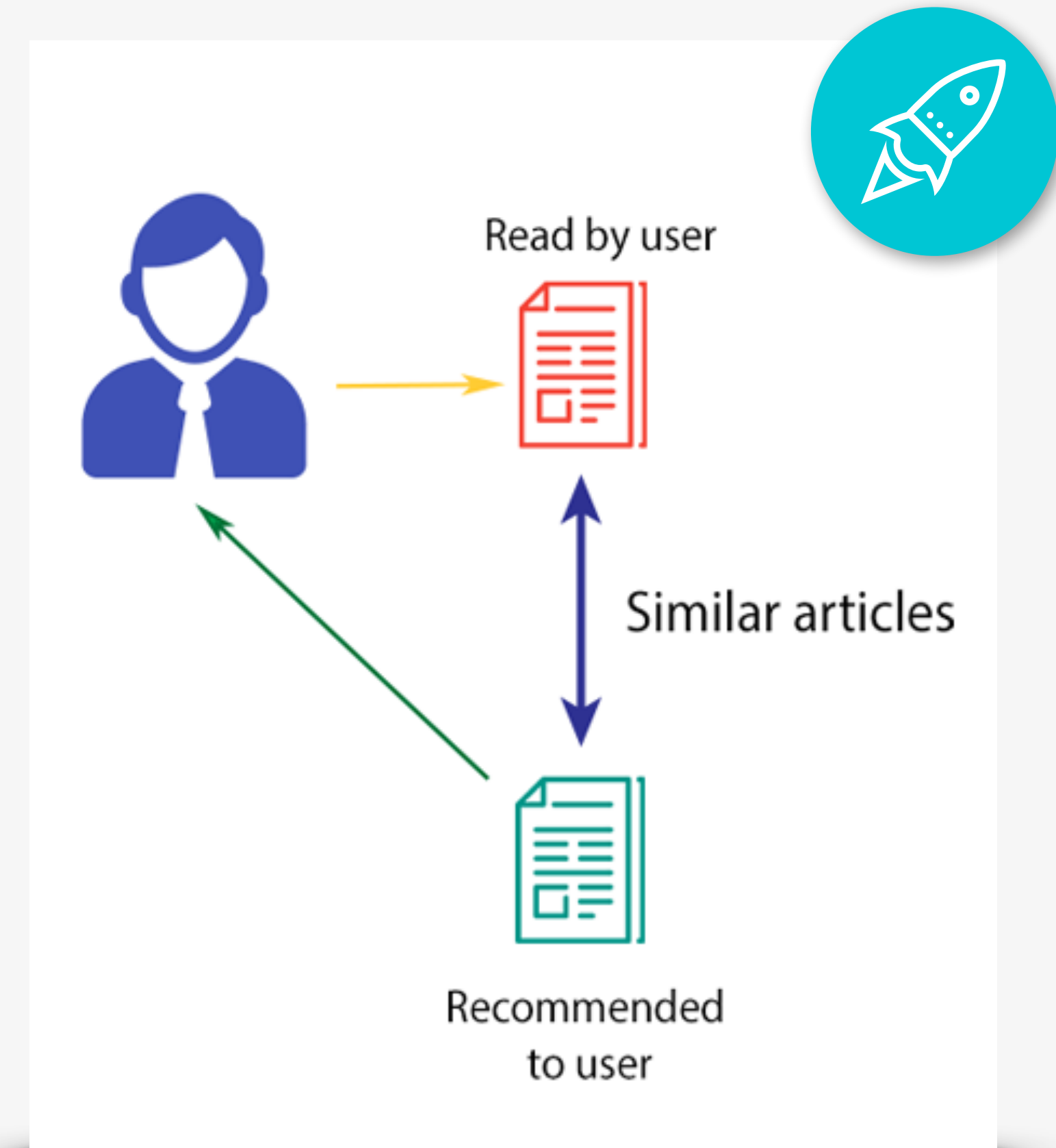
기존 음악 추천 시스템

5



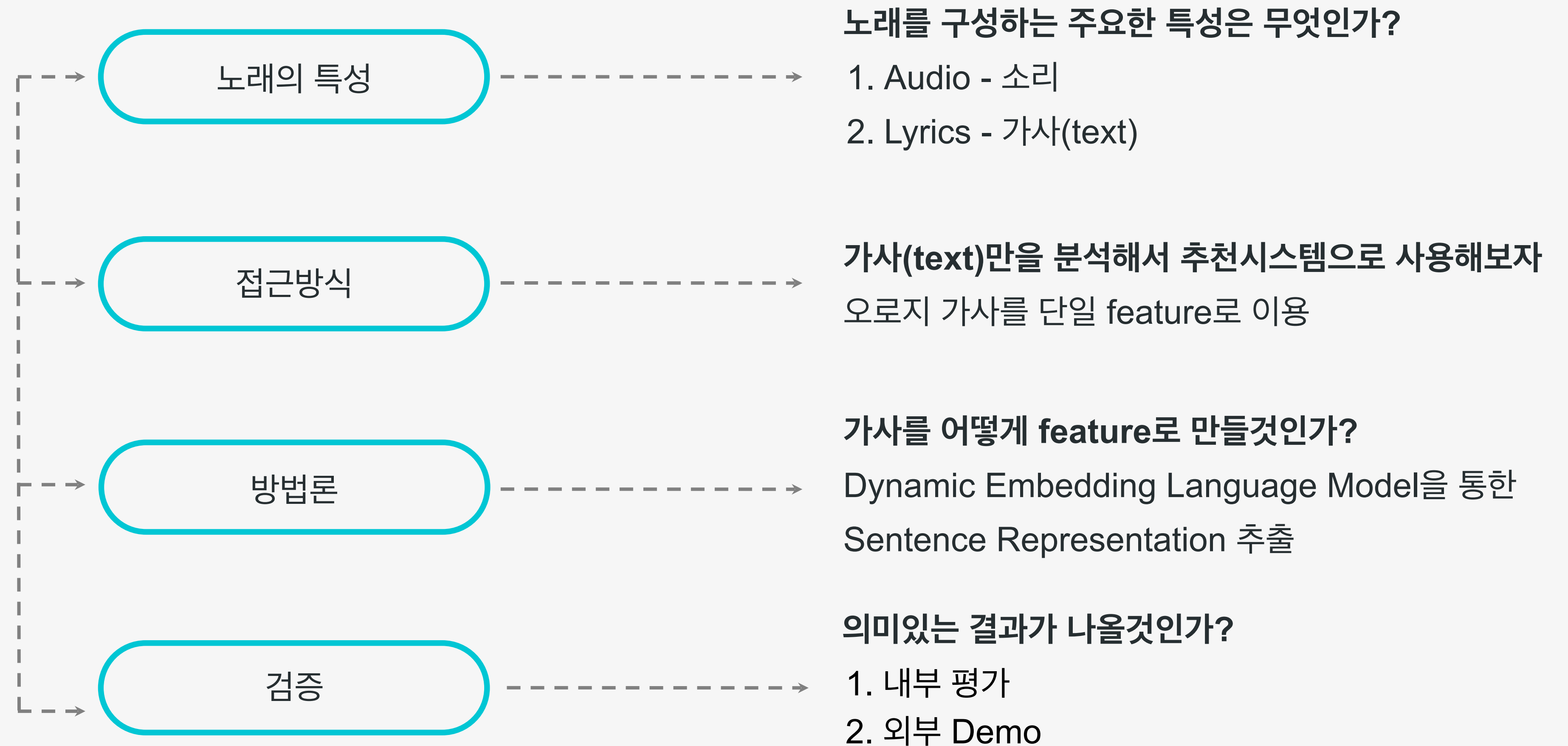
Collaborative Filtering

VS



Content-based Filtering

3. 연구의 목표



EXPLOITING SYNCHRONIZED LYRICS AND VOCAL FEATURES FOR MUSIC EMOTION DETECTION

Loreto Parisi¹

Simone Francia¹

Silvio Olivastri²

Maria Stella Tavella¹

¹ Musixmatch

² AI Labs

loreto@musixmatch.com, simone.francia@musixmatch.com

silvio@ailabs.it, stella@musixmatch.com

ABSTRACT

One of the key points in music recommendation is authoring engaging playlists according to sentiment and emotions. While previous works were mostly based on audio for music discovery and playlists generation, we take advantage of our synchronized lyrics dataset to combine text representations and music features in a novel way; we therefore introduce the Synchronized Lyrics Emotion Dataset. Unlike other approaches that randomly exploited the audio samples and the whole text, our data is split according to the temporal information provided by the synchronization between lyrics and audio. This work shows a comparison between text-based and audio-based deep learning classification models using different techniques from Natural Language Processing and Music Information Retrieval domains. From the experiments on audio we conclude that using vocals only, instead of the whole audio data improves the overall performances of the audio classifier. In the lyrics experiments we exploit the state-of-the-art word representations applied to the main Deep Learning architectures available in literature. In our

a classifier using various representations of the acoustical properties of a musical excerpt such as: timbre, rhythm and harmony [21, 26]. Support Vector Machines are employed with good results also for multilabel classification [30], more recently also Convolutional Neural Networks were used in this field [45]. Lyrics-based approaches, on the other hand, make use of Recurrent Neural Networks architectures (like LSTM [13]) for performing text classification [46, 47]. The idea of using lyrics combined with voice only audio signals is done in [29], where emotion recognition is performed by using textual and speech data, instead of visual ones. Measuring and assigning emotions to music is not a straightforward task: the sentiment/mood associated with a song can be derived by a combination of many features, moreover, emotions expressed by a musical excerpt and by its corresponding lyrics do not always match, also, the annotation of mood tags to a song turns out to be highly changing over the song duration [3] and therefore one song can be associated with more than one emotion [45]. There happens to be no unified emotion representation in the field of MER, meaning that there is no consensus upon which and how many labels are used and if emotions should be considered as categorical or contin-

EXPLOITING SYNCHRONIZED LYRICS AND VOCAL FEATURES FOR MUSIC EMOTION DETECTION

Musixmatch AI Labs



음악의 감성 인식에는 오디오 보다 가사
(텍스트)가 더 좋은 성능을 보여준다

4. 연구 방법론

연구절차

Dataset 구축

국내 음악 스트리밍 서비스 M사에서 크롤링
인기곡 순서로 정렬된 12개 장르에서 총 2만곡 준비
텍스트 전처리

Context 벡터 추출

사전 훈련 된 언어 모델인 BERT를 통해 Lyrics Representation Vector를 추출
각기 다른 사전 데이터로 훈련된 두 종류의 BERT 사용
(구글의 Multilingual BERT, SKT의 KoBERT)

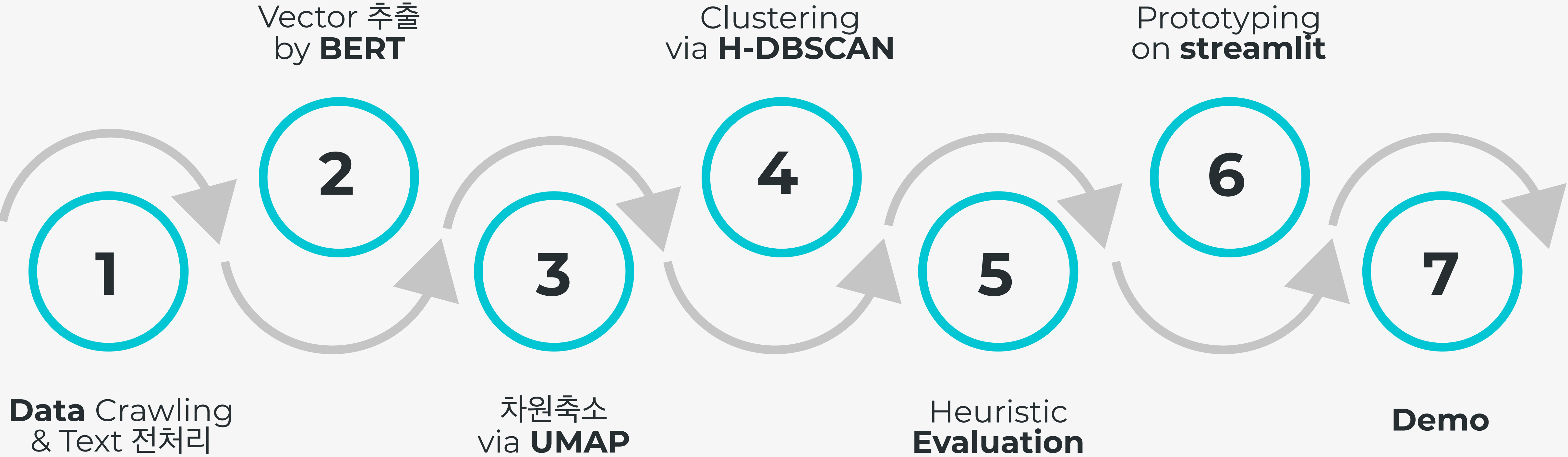
차원축소 & Clustering

768차원의 Context Vector를 UMAP을 활용하여 2차원으로 축소
H-DBSCAN 기법으로 Unsupervised Clustering

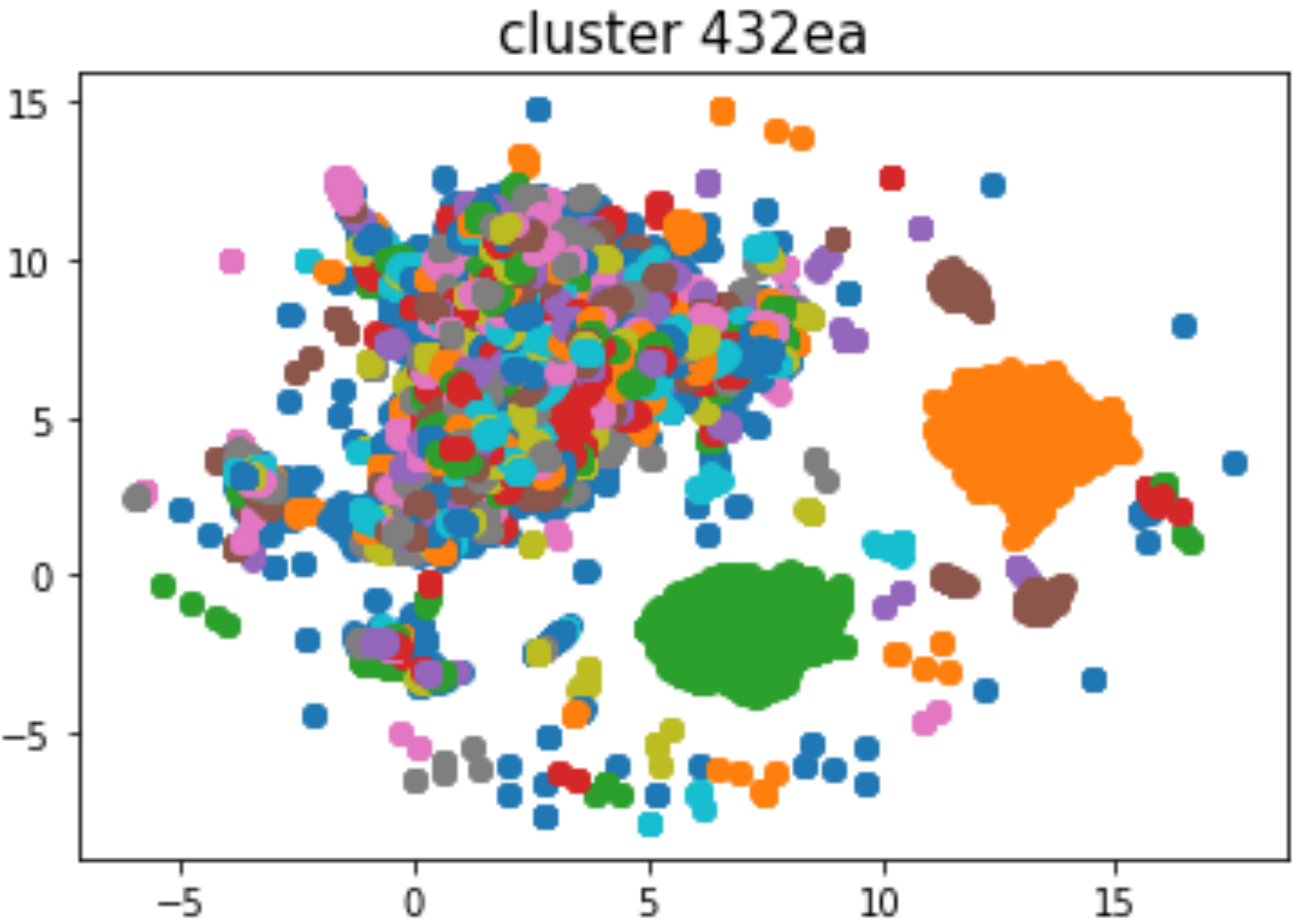
추천 알고리즘 구성 & Demo 제작

특정 음원을 키워드로 검색 하였을 경우 동일한 Cluster에서 Cosine Similarity가
높은 순으로 곡을 추천하도록 웹 기반 Demo 제작

5. 연구의 과정



Final — Result



Clustering Map(432 Cluster)

		Predicted		
		Positive	Negative	
Actual	Positive	26	16	
	Negative	24	34	
		F1 0.57	Precision 0.52	Accuracy 0.6

이진 Labeling Test(n=100)

6. Demo 시연

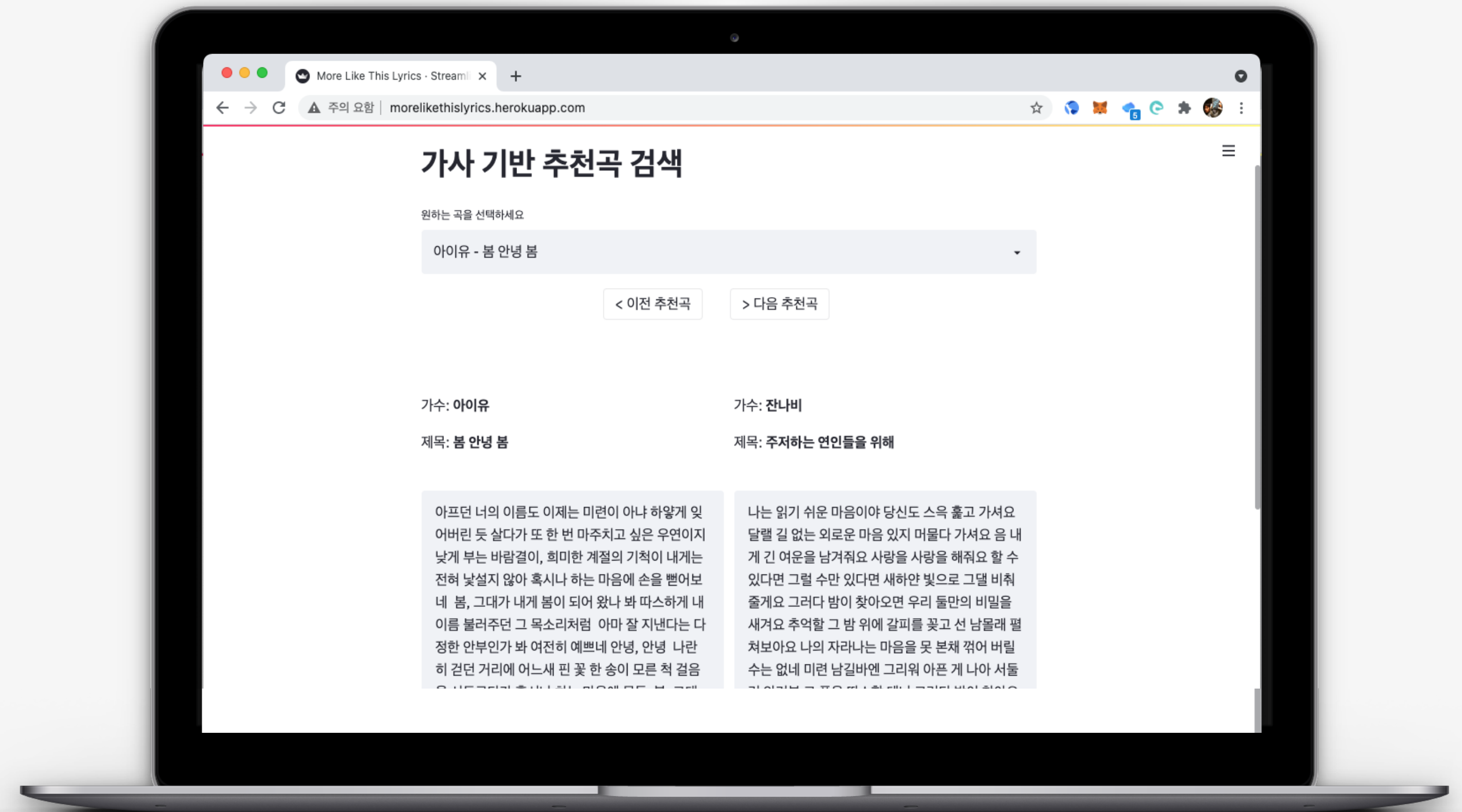
SongLike.me



Framework:
streamlit, heroku

Description:

인기곡 순서로 각 장르별 동일한 수의 노래 총 1만곡의 데이터를 바탕으로 사용자의 **검색곡을 기준으로 5곡의 노래를 추천**하여 순서대로 출력합니다.



7. 결론 및 고찰

연구결론



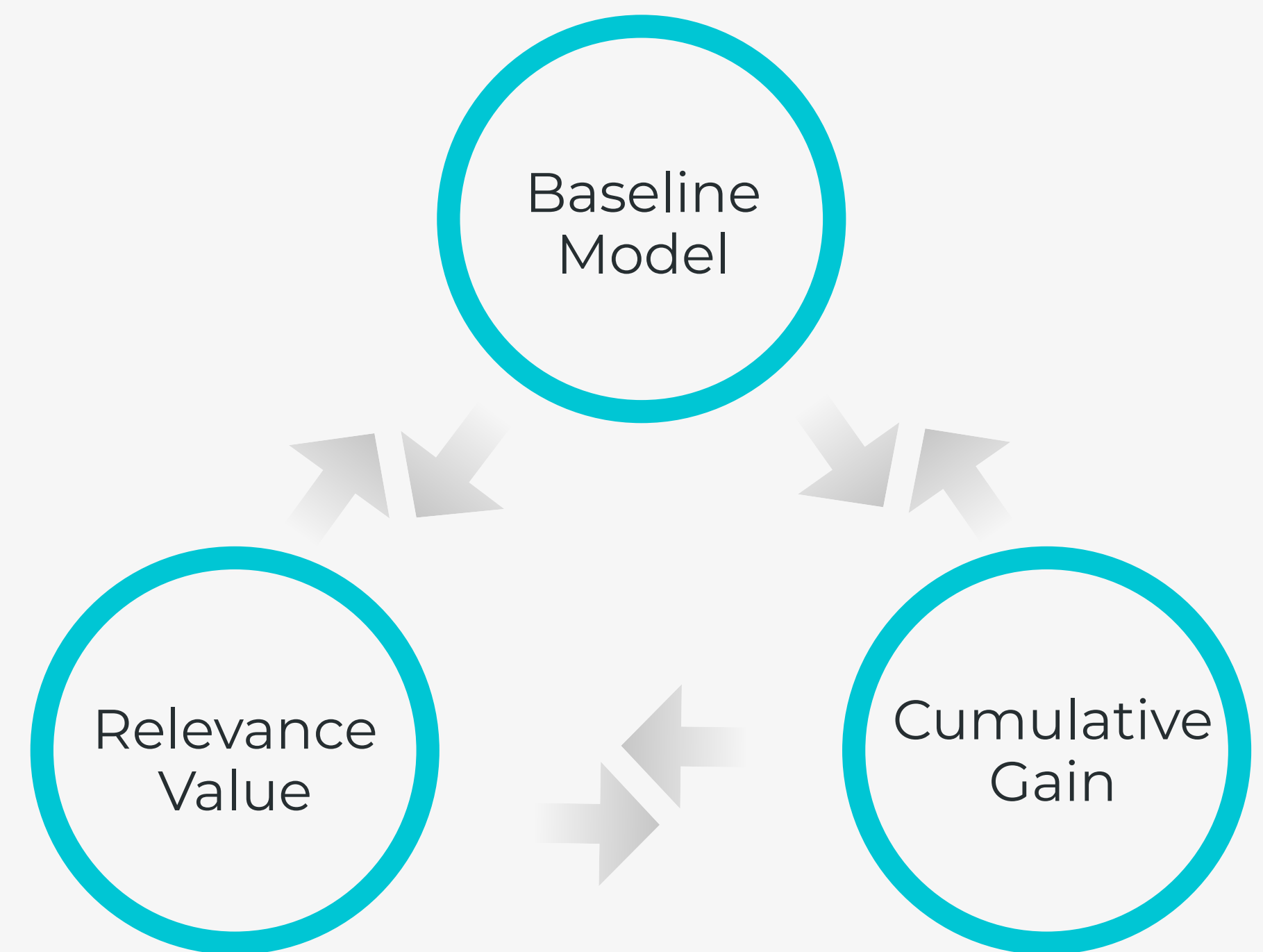
가사를 단일 Feature로 BERT를 통해 Dense하게 요약된 가사 **Embedding Vector**를 얻고, UMAP과 HDBSCAN으로 차원축소된 **Vector의 Cluster**를 형성하여 동일 Cluster에서 **Cosine Similarity**가 높은 곡을 추천해주는 알고리즘을 통해 가사 유사도 기반 **음악 추천 시스템**을 구성하였다. 실제 결과에 대해 제한적이거나 휴리스틱적으로 알고리즘의 성능을 평가 하였으며 **기존 추천 시스템에 병합되어 사용될 수 있는 Baseline Model**로의 가능성을 검토하였다.



연구 고찰



기존 자연어 처리 분야의 문서 유사도 Task를 평가하는 지표로는 본 연구의 추천 시스템의 객관적이고 **정량적인 평가가 어렵기** 때문에 어떠한 Metric을 도입할지에 대한 고민은 아쉽다. 실제 추천 시스템 평가 방식으로 자주 쓰이는 **NDGC(Normalized Discounted Cumulative Gain)**의 경우 서비스를 운영하면서 얻을 수 있는 데이터인 추천곡을 실제 사용자가 청취를 했는지에 대한 여부, 청취 횟수 등을 기반으로 **Relevance Value**를 얻고 이를 기반으로 **Cumulative Gain**을 산출하여 객관적이고 정량적인 평가가 가능하지만 **축적된 데이터가 없는 서비스 초기 단계에서 Baseline Model의 효과적인 평가는 요원**하다. 하지만 본 연구는 추천 시스템의 다양한 가능성을 제시하고 초기 추천 시스템의 Baseline Model 도입 시 고려될 수 있는 다양한 방법론의 일환으로써 향후 음악 추천 시스템이 높은 수준의 개인화를 달성하는 것에 기여하고자 한다.



Thank you :)

Speaker

김지훈 자비랩 대표

EMAIL

wanderer71718@gmail.com

SOCIAL

