프로젝트 PTNM

Personal Taste Navigation Model

팀 : 찾아라 비밀의 열쇠

## 개요 ::

최근 컨텐츠 추천 방식은 Collaborative Filtering(이하 CF)과 Item Based Recommendation(아이템 기반 추천방식)을 함께 사용해서 추천을 하는 하이브리드 추천 시스템이 널리 사용되고 있다.

음원 추천에 대한 만족도로 스트리밍 시장의 강자로 성장한 ‘Spotify’의 경우 3가지 모델을 사용해서 추천하는 하이브리드 추천 방식의 서비스를 제공하며, Open API로 제공되어 ‘Naver 뮤직’에서도 Spotify의 오픈소스를 활용하여 VIBE를 출시하였다.

이러한 Spotify의 추천 방식은

- Collaborative Filtering Model

- 음원의 특징을 나타내는 태그를 통한 NLP

- 새로운 음원에 대해 CNN을 통한 분류

세가지 모델을 통해서 추천을 하고 있다. 이러한 모델이 충분히 높은 만족도를 제공하는 것을 사실이다. 하지만 사용시간이 지남에 따라 지속적으로 비슷한 음악만 추천받는 것에 대해서 사용자의 만족도가 떨어지게 된다.

따라서 우리는 사용자의 취향에 맞으면서 내가 경험한 적 없는 새로운 음악을 추천할 수 있는 모델을 만드는 것을 목표로 했다.

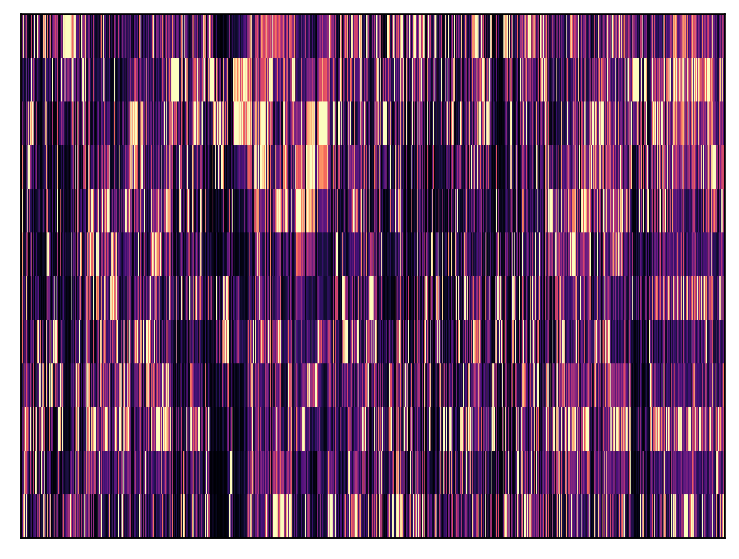
PTNM은 CNN에서 아이디어를 얻어 출발했다. CNN이란 Convolution Layer를 통해 주파수 특징을 추출하고, Neural Network를 통해 분류 or 예측을 하는 것이다.

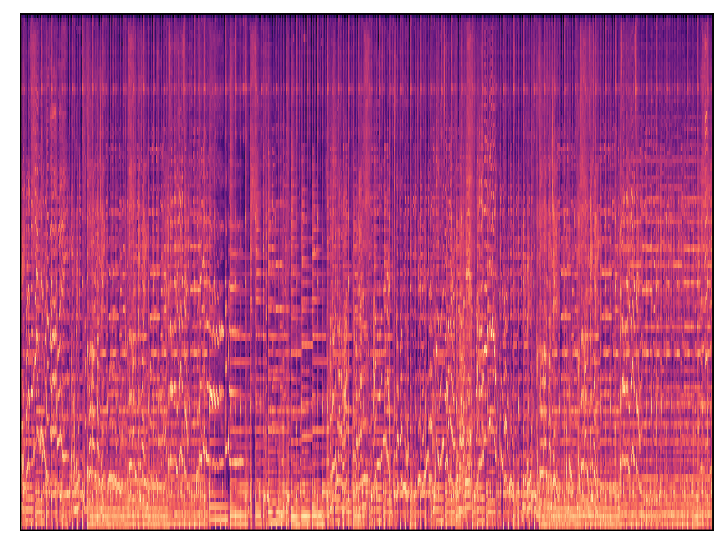
PTNM은 Convolution Layer 대신에 8가지의 특징을 추출한 벡터를 사용하였고, Neural Network를 대신하여 주파수 특징에 대한 사용자 평가를 반영할 수 있는 방법은 Page Rank Algorithm을 참고하였다. 사용자 평가를 바탕으로 주파수 특징에 대한 가중치를 구한 뒤 순위를 메겨 재 추천되는 방식이다.

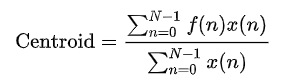
## 특징 추출 ::

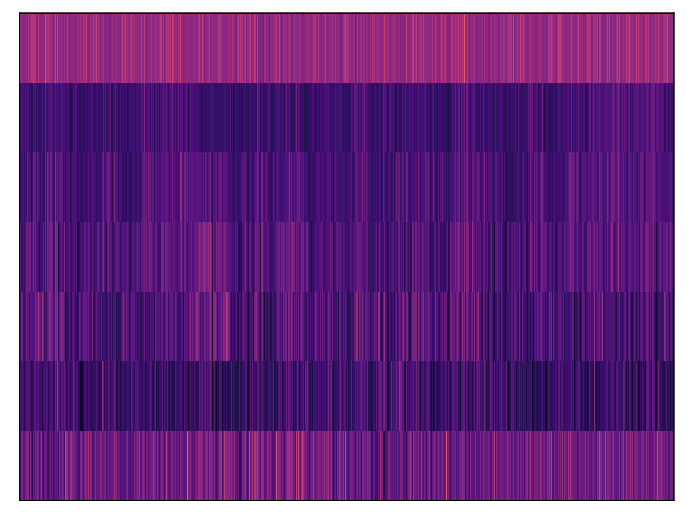
음원으로부터 8가지 특징을 추출한다. CNN에서 Convolution Layer의 역할과 비슷한 것이다. 보통은 이러한 결과 이미지를 CNN에 입력하여 정리하지만, 우리는 히스토그램 유사도 분석을 통해 샘플 그룹과의 유사도를 비교하였다. 그리고 이 결과를 각 특징별로 1개의 축으로 정리하였다.

주파수 특징 추출은 주로 ‘Librosa’ 패키지를 이용했다.

* Chromagram : 크로마그램은 주파수를 12개 bin으로 나눠서 주파수 파워를 나타내는 것이다. 음악에서 옥타브 간격으로 떨어진 음은 유사한 것으로 인식되기 때문에 절대 주파수가 없어도 Chroma 분포를 아는 것으로도 유용한 정보를 제공 받을 수 있다.   
  Funkadelic-Can’t Stand The Strain / Chromagram
* Melspectrogram : Mel-scaled spectrogram, 인간의 청각 특성상 저음역에서 민감하고 고음역에서 둔감해지는 경향을 반영하여 주파수 파워를 계산한 Spectrogram이다.

  
Funkadelic-Can’t Stand The Strain / Melspectrogram

* Spectral\_centroid : 현재 주파수 스펙트럼의 중심이 어디인지를 나타내는 것으로 소리의 밝기(Brightness) 정보와 관련이 있다.  
     
     
  Funkadelic-Can’t Stand The Strain / Spectral Centroid
* Spectral bandwidth : Spectral Centroid로부터 전체 대역폭을 구한다.  
  (sum\_k S[k, t] \* (freq[k, t] - centroid[t])\*\*p)\*\*(1/p)    
  Funkadelic-Can’t Stand The Strain / Spectral Bandwidth
* Spectral contrast : Mel\_spectrogram을 대신해서 Octave 기반의 스펙트럼 대조를 나타내는 방법이다.

   
Funkadelic-Can’t Stand The Strain / Spectral Contrast

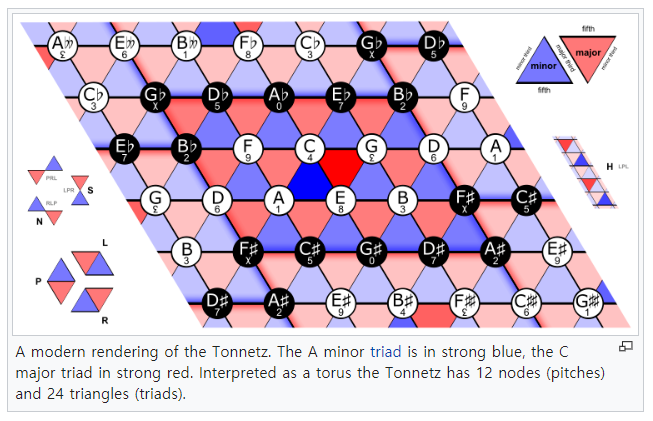
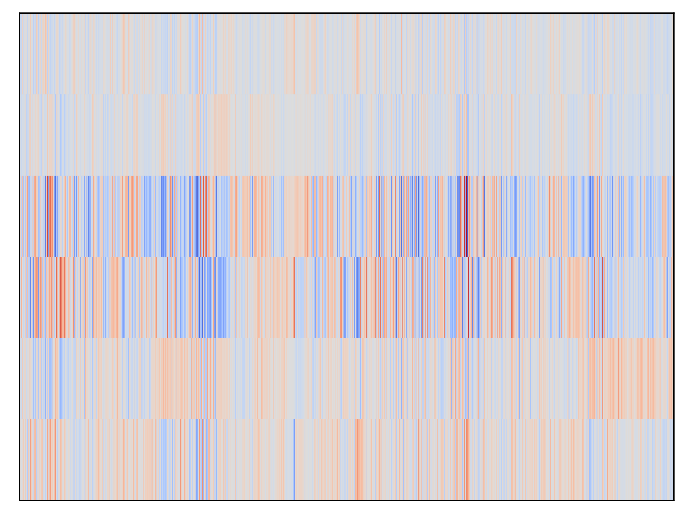
* Tonnetz : 화성 관계를 Lattice Diagram 으로 나타낸 것으로 중심으로부터 각 방향마다 3도 완전 5도 관계 등으로 구분하여 나타낸 것이다. 이러한 Tonnetz 이론을 바탕으로 구현된 Librosa Package의 tonnetz를 이용하였다.

사진 출처 :

<https://en.wikipedia.org/wiki/Tonnetz>

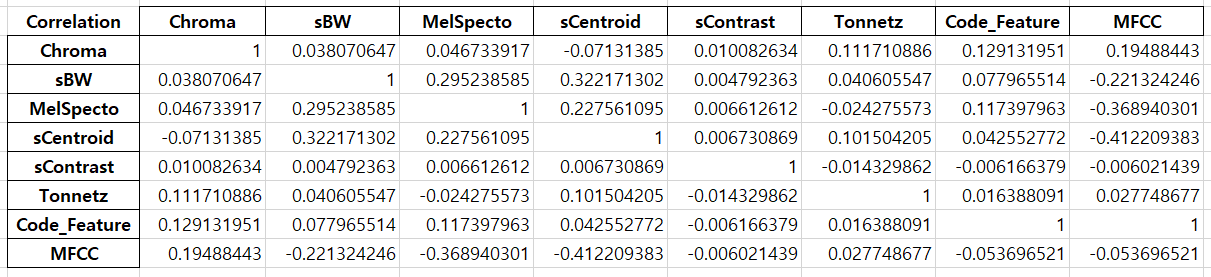
   
Funkadelic-Can’t Stand The Strain / Tonnetz

* MFCC : 주파수 정보를 특징 벡터로 만드는 것이다. 우리는 12 차원의 벡터로 만들어서 각 차원에 대한 고유 벡터와의 cos 유사도를 분석하는 것으로 정리하였다.
* 조성 분석 : Melodia를 활용해 보컬의 음정만 음악에서 분리 한 뒤 고유 주파수를 바탕으로 조성을 찾고 12개의 계이름으로 변환한 뒤 12차원 벡터화 하여 각 차원별 고유 벡터와 cos 유사도를 분석하는 것으로 음악을 정리하였다.

이미지 데이터의 경우 대표가 되는 샘플 그룹을 6개 정도 설정하여 각 그룹과의 히스토그램 유사도로 6차원 벡터화 했고, MFCC와 조성분석의 경우 12차원에 대한 정보가 존재하게 된다. 따라서 하나의 축으로 만들기 위해서는 차원을 축소하는 과정이 필요하다.

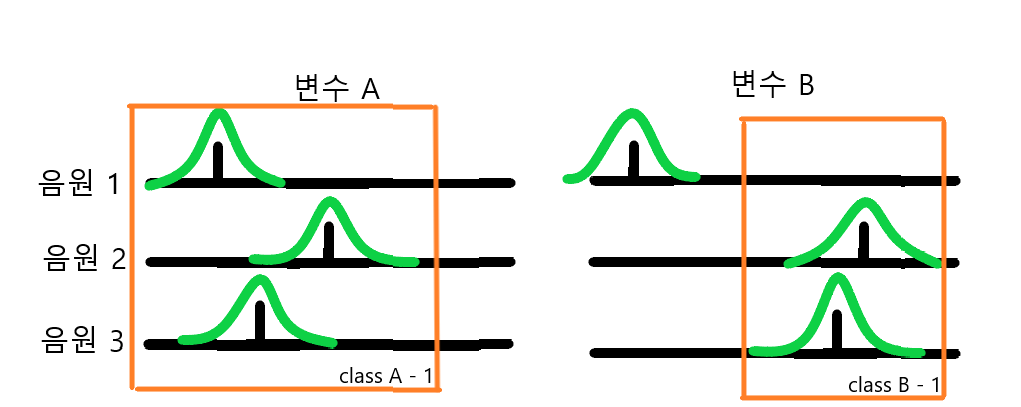
각 특징 벡터들을 각 차원 간의 비율이 중요하기 때문에, 각 그룹에서의 크기 정보를 압축할 수 있는Arithmetic Coding을 사용하였다. 결과적으로 이를 통해 각 변수들은 0~100까지의 값을 갖는 하나의 축을 만들 수 있게 되었다.

* 최종 정리된 특징들의 상관관계



## 추천 과정

1. 사용자가 기존에 선호하던 곡을 입력한다.
2. 기존 선호 곡에 대한 변수 값을 찾는다.



선택된 음원 주위로 사용자의 선호 확률이 가우시안 정규 분포를 갖는다고 가정,

80%확률 범위 안에 다른 음원의 확률 범위가 겹친다면 하나의 Class로 묶어서,

해당 변수에 대한 사용자의 선호 범위를 나타낸다.

1. 각 변수에 대한 사용자 의존도와 각 클래스에 대한 가중치를 계산한다.

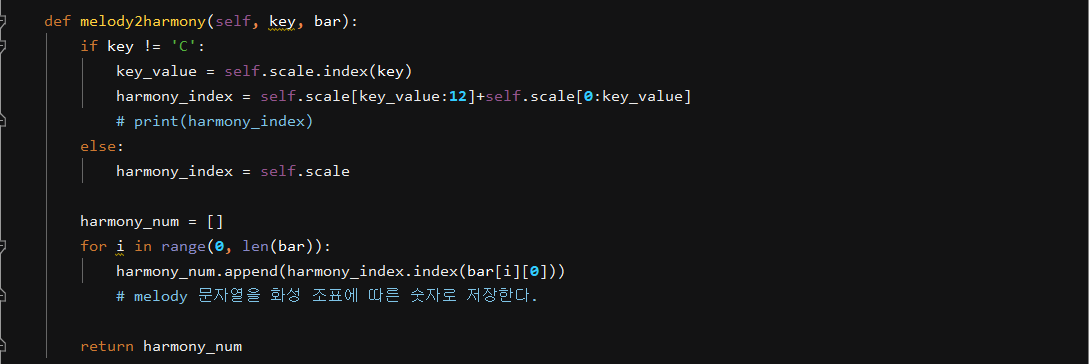
Wc : Class 가중치 , Wv : 변수 가중치, E : 클래스 내부에 포함된 사용자 평점 평균

1. 곡에 대한 가중치를 계산한다.

1. 상위 랭크 된 곡 5개를 추천하고 그에 대한 피드백을 다시 변수 축 공간에 반영한다.

## 구현 방식 :

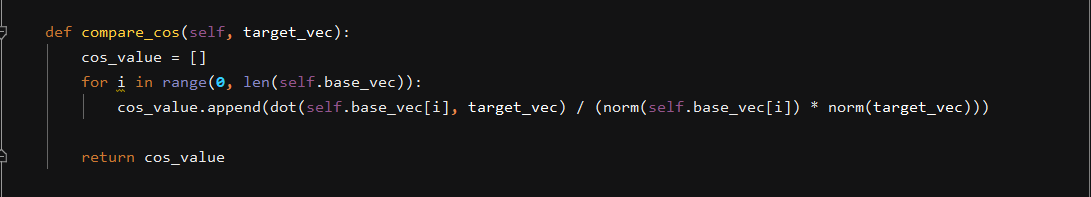
* 계이름 분석

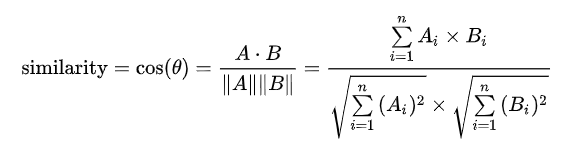


계이름이란 음과 음사이의 상대적 관계를 나타내는 방법으로 조성이 바뀔 때 마다 으뜸음이 바뀌기 때문에, 보컬 멜로디를 조성에 따라 코드를 계이름으로 변환한 뒤 계이름의 등장 횟수를 12차원 벡터화 하여 Cos 유사도 분석을 시행하였다.

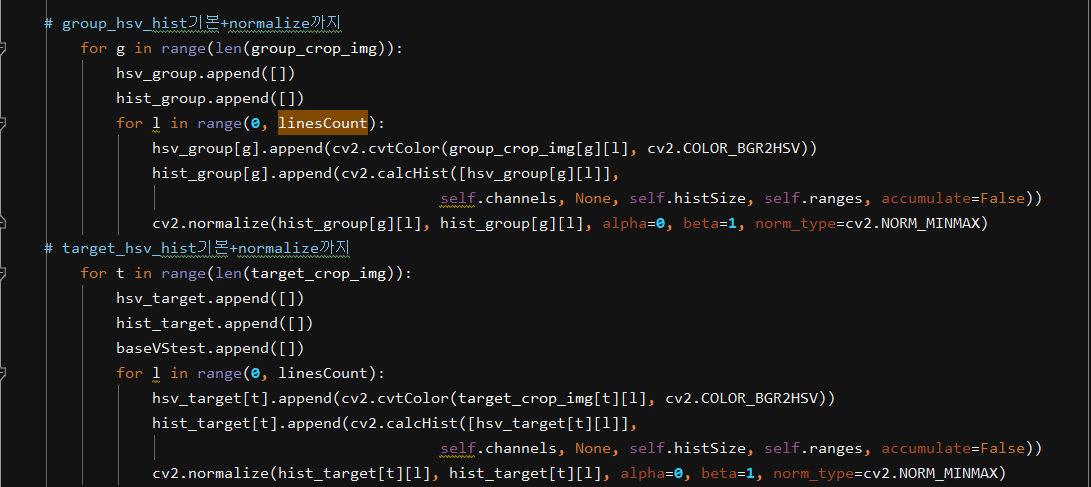
* Cos 유사도 분석

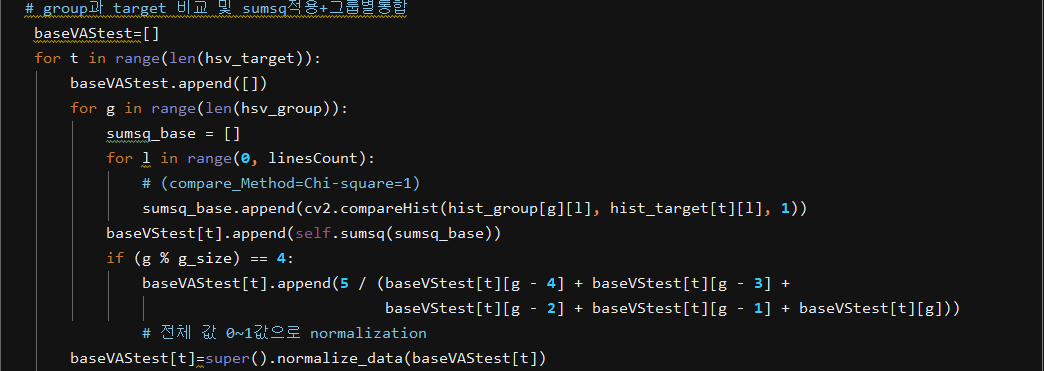


* 12개의 단위벡터  
  

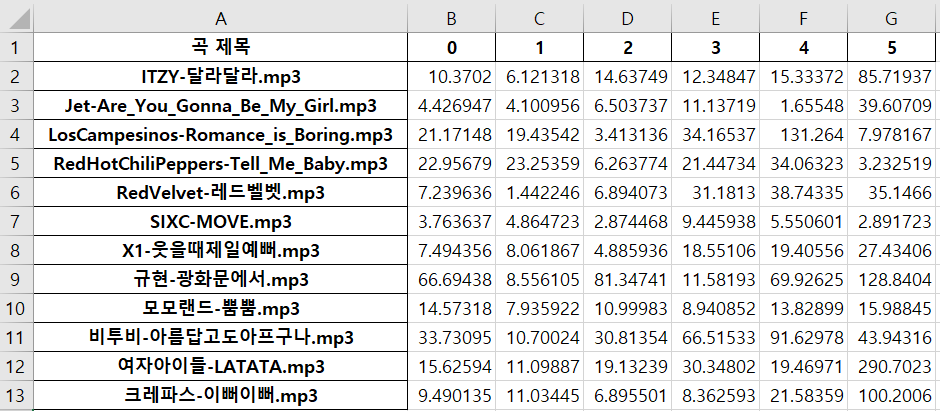


* Tonnetz의 각 영역별 Histogram 유사도 분석
* 히스토그램 유사도는 대표 샘플 그룹을 나누고 각 그룹의 히스토그램 평균과 비교할 곡의 히스토그램의 유사도를 추출하였다.
* 그룹을 구성하는 방법은 이미지의 색상 대표를 몇가지 나누거나 음악 분위기를 통해서 간략하게 분류하였다. 음악이라는 정성적 데이터를 정량화 하기 위해서 우리는 각 그룹과의 유사도를 사용하였다.   
   우리는 정답을 찾는 과정이 아닌 일정한 기준을 통해서 정렬하는 과정이 필요하기 때문에 일반적인 분류기에서 각 분류기준에 따른 확률을 출력하는 것에 대해서 아이디어를 얻어 그룹별로 유사도를 전체에 대한 비율로 정규화 한 뒤, 산술부호화를 통해서 하나의 값으로 출력하였다.
* 히스토그램 유사도 분석 코드





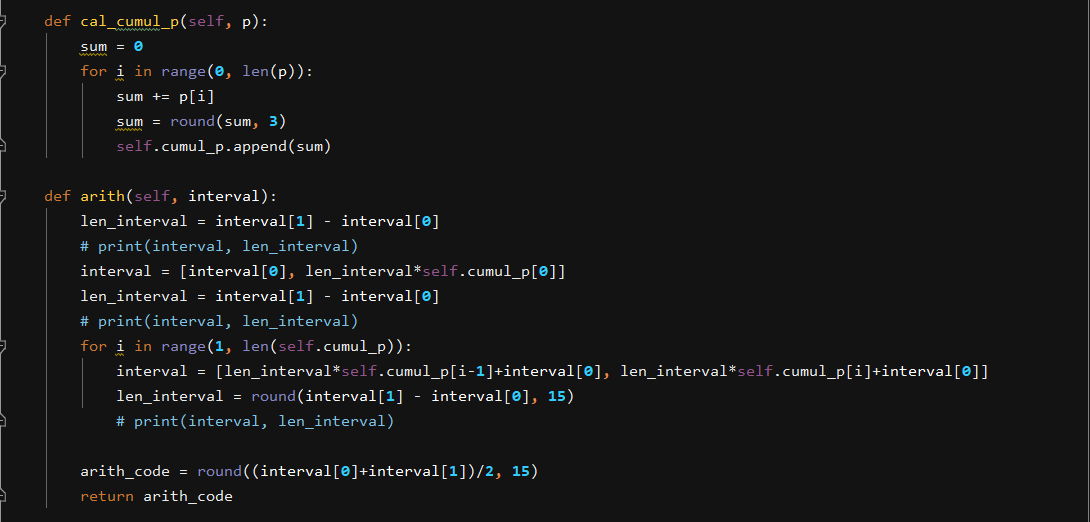
* 히스토그램 유사도 분석 결과 (Tonnetz)



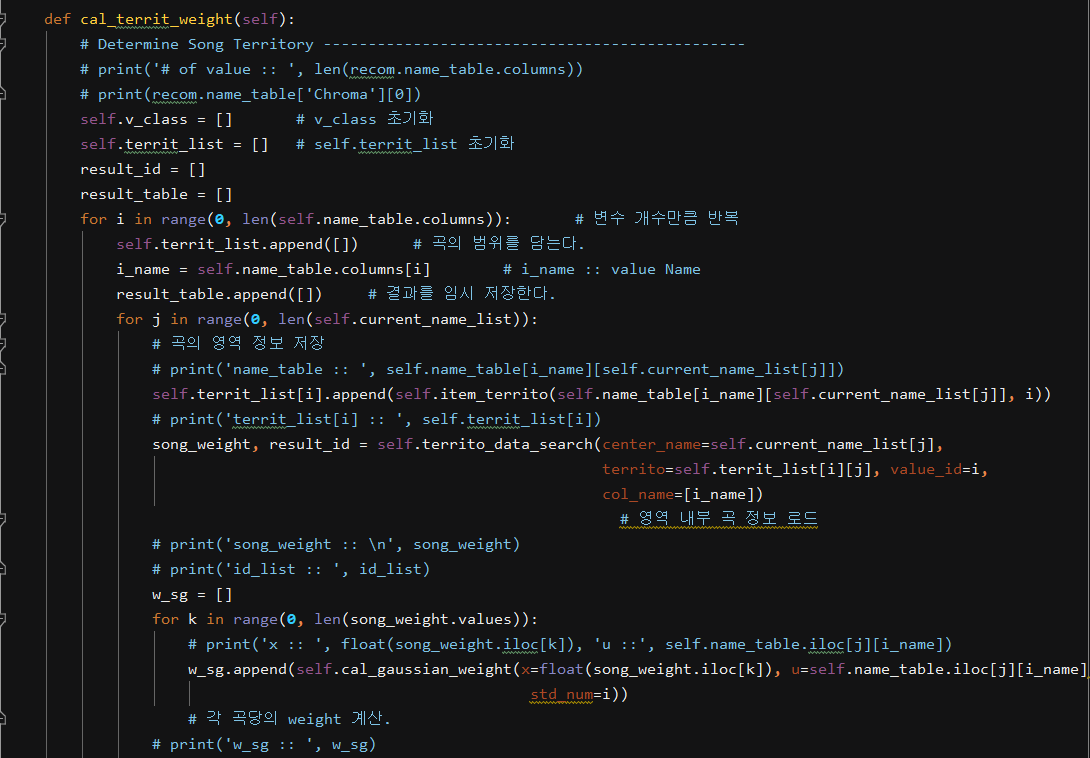
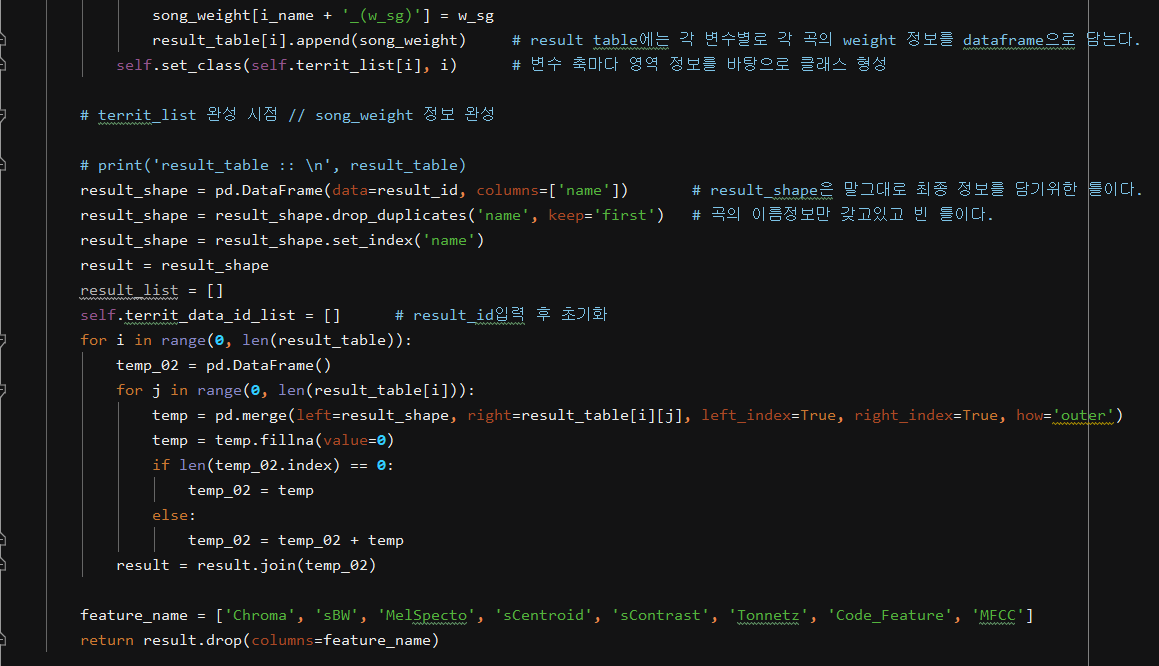
* 위 결과에서 6개의 그룹을 전체에 대한 비율로 나타내서 전체 합이 1이 되게 정규화 한뒤 산술 부호화를 시행한다.
* 산술 부호화(Arithmetic Coding)

압축 기술로 활용되는 산술 부호화는 심볼의 등장 확률에 따라 데이터를 압축하는 기술이다. 산술 부호화를 사용하면 각 심볼의 확률에 대한 정보와 심볼의 순서에 대한 정보가 하나의 숫자로 표현될 수 있다는 것을 이용하여 6~12차원의 데이터를 각각 전체에 대한 비율로 정규화 한 뒤, 산술 부호화를 통해 하나의 숫자로 만들었다.

* 산술 부호화 코드



* 곡의 가중치 계산

## 결론

테스트는 총 5602개의 곡으로 진행하였다. 전체적으로 만족스러운 결과를 얻을 수 있었다. 3곡 정도는 매우 유사한 곡이 추천되었고, 2곡 정도가 비슷한 느낌이지만 새로운 장르의 음악 등이 추천되었다. 때때로 전혀 취향과 반대인 곡이 추천되지만 피드백이 반영되면서 다시 올바른 추천 방향을 찾아가는 것까지 확인하였다.

PTNM은 음악 추천에 대해 다양한 방법으로 사용 될 수 있다. 날씨나 기분에 따른 음악 추천의 경우에도 해당 하는 상황에서 사용자가 기존에 선호했던 몇가지 곡을 입력하면 새로운 곡들을 계속 추천해 줄 수 있다. 타인의 기준으로 선정된 곡이 아닌 내가 기존에 듣던 곡들을 기반으로 날씨에 따라 내 맞춤형 서비스를 제공 받을 수 있는 것이다.

마지막으로 PTNM의 가장 큰 장점은 취향의 한계가 확장될 가능성이 높다는 것이다. 기존의 음원 분류 방식은 모두 장르를 기반으로 이루어졌는데, PTNM에서는 주파수 특징별로 음원을 나열했기 때문이다. 비슷한 음악이라도 다른 장르의 음원을 접하게 할 수 있고, 이러한 과정을 통해 사용자는 보다 다양한 음원을 접할 수 있다.

물론 모든 음원에 대해서 특징을 추출하는 과정이 필요하고, 정리된 DB로부터 효율적으로 데이터를 불러오는 방법에 대한 고민이 필요하다. 현재 추천 1회에 2~4초 가량 소모되는데, 이중 대부분이 데이터를 불러오는 과정에서 발생한다. Weight 연산 자체로 걸리는 시간이 크지 않기 때문에 데이터 검색 및 로드 과정만 보완되면 상용화 가능한 수준의 결과를 도출 할 수 있다고 생각한다.

참고문헌

‘How Does Spotify Know You So Well?’ Sohia Ciocca, Medium /

<https://medium.com/s/story/spotifys-discover-weekly-how-machine-learning-finds-your-new-music-19a41ab76efe>

Chromagram /   
Ellis, Daniel P.W. “Chroma feature analysis and synthesis” 2007/04/21

Spectral Centroid /  
<https://en.wikipedia.org/wiki/Spectral_centroid>

Spectral Contrast /  
Jiang, Dan-Ning, Lie Lu, Hong-Jiang Zhang, Jian-Hua Tao, and Lian-Hong Cai. “Music type classification by spectral contrast feature.” In Multimedia and Expo, 2002. ICME‘02. Proceedings. 2002 IEEE International Conference on, vol. 1, pp. 113-116. IEEE, 2002.

Tonnetz /  
Harte, C., Sandler, M., & Gasser, M. (2006). “Detecting Harmonic Change in Musical Audio.” In Proceedings of the 1st ACM Workshop on Audio and Music Computing Multimedia (pp. 21-26). Santa Barbara, CA, USA: ACM Press. doi:10.1145/1178723.1178727.

Librosa Documentation /

<https://librosa.github.io/librosa/feature.html>

Page Rank Algorithm /

Sergey Brin and Lawrence Page “The Anatomy of a Large-Scale Hypertextual Web Search Engine”

결정을 돕는 추천 알고리즘/

<https://blog.naver.com/happyrachy/221292844469>