

추천알고리즘 설계방향 _ ver2

API + 음성 => 최종 감정 결과 => (추천알고리즘) => 노래

추천알고리즘은 어떻게 ‘분류’하느냐가 관건이다. 8 개의 감정 카테고리 별로 노래를 분류하는게 제일 중요하다. 이때, 분류기준이 납득가능해야한다!

1. 지금은 사용자에 대한 데이터가 없으므로 랜덤추출 방식을 사용하기로 한다. 후에 사용자 기반의 알고리즘을 활용한다면, 개개인의 노래 취향과 만족도 조사 등으로 알고리즘을 발전시킬 수 있다.

2. 1 위,2 위,3 위로 노래 랭킹이 나온다. 이는 곧, 한 감정에 대해 꼭 한 가지 장르만 나올 필요가 없다는 말이기도 하다. 최종결과에서 판단을 내리기 힘든 애매한 결과가 나온다면(한 장르의 노래로 추천하기 힘든 분석결과 등이 해당할 것이다) 이를 활용할 수 있을 것이다.

3. 비율이 다양하게/골고루 나올 경우, 어떤 카테고리를 선택할 것인지도 고려해야 한다. 예를들어, 슬픔의 감정이 60%정도가 나오고 나머지 감정이 골고루 나왔을 때, 어두운 노래를 틀어도 될지가 해당한다.

4. 슬플 때 어둡고 슬픈 분위기의 노래만 틀어주는게 좋을지 혹은 기분이 업될 만한 노래를 틀어주어 기분전환을 이끄는게 좋을지 등은 연구결과나 논문을 더 조사해봐야한다. 물론 사용자마다 상황에 따른 노래 선택은 매우 달라질 수 있겠지만 말이다. 지금은 3 가지 노래를 추천함으로써 보완한다. 아니면 설문조사를 통해 낸 통계결과를 활용할 수도 있다.

5. 저작권이 없는 노래, NCS 가 있다. 현재 600 여 개의 노래가 있다. 이를 사용하면, 저작권 문제에 걸리는 것이 없으므로 자유롭게 스트리밍 할 수 있다. 또한 저작권이 있는 노래는 매일같이 쏟아져서 너무 많기 때문에 거기에서 선정하는게 아니라 범위를 좁히고 시작한다(노래 선정의 객관성 문제에 연관이 된다.).

6. 발전 방향(데이터가 쌓이면서 활용가능한 방안)

(1) 개발 초기 단계에서는 사용자 개개인에 대한 데이터 확보가 불가능하므로 사용자 기반의 음악 추천 시스템은 구현하기 힘들다. 그러므로 데이터가 확보되면 다음과 같은 사항을 도입할 수 있다.

(2) 사용자 개개인에 대한 데이터를 쌓아 기계학습을 활용해 추천 알고리즘을 발전시킨다.

(3) 사용자 개개인이 아닌 전체에 대해, 감정 상태가 거의 비슷한 경우에 어떤 류의 노래를 주로 듣는지를 조건부 확률(이런 상황에서 이럴 가능성이 높다)을 사용할 수 있다. 기존의 알고리즘에서, 확률에 대한 가중치로 합산하는 방향으로도 발전시킬 수 있다.