

**[분석]** -자료 정리에 앞서 계획 수립(실행~1/1)

- 장르별 분석을 통해 발라드 가사를 학습 데이터로 선정한 이유 도출

-----

- \*분석 코드 동일하게 적용

- \*형태소 분석

- \*빈도수 시각화(그래프, 워드클라우드)

- \*각 장르별 빈도수 기준으로 단어 상위 10~ 20개 비교 분석 -> 장르 특성 유추 가능

- \*이유 도출

- \*결과 보고서 형태로 작성 및 발표 자료 선정, 스크립트 부분 제작

\*\*\*

특정 장르를 배제할 이유가 충분치 않음 -> 보강

발라드를 하기 위한 합리화가 되지 않도록 정리 -> 변명으로 느껴지지 않을 정도의 타당한 언어 필요

중간 발표는 왜 이 프로젝트를 진행하고자 하는지의 목적이 분명하고 이유가 타당해야 할 것 -> 있어 보이는 거

-----

## 1. 인디

-예술성에 치중한 가사(비유적, 비서사적) -> 인디 특유의 다양성과 독창성을 드러내나 문장 생성에는 적합지 않을 것으로 보임

-주제가 다양해 생성한 문장의 일관성을 찾기 어려울 것으로 보임

## 2. 국악

-예스러운 말투와 단어

-추임새 다수

## 3. 록/ 메탈

- 마이너한 장르, 다수가 공감할 수 있는 단어를 포함하고 있으나 장르적 성격이 가사에서도 드러나 문장 생성에 활용하기 적합하지 않음(하드 록, 헤비 메탈 등)

- 다른 장르와의 결합 혹은 대중적 록(록 발라드, 컨템포러리 록 등)에서는 배제할 이유를 찾기 힘들

## 4. 포크/ 블루스

- 안 되는 이유 발견하기 어려움

## 5. 댄스

-영어, 반복되는 구

## 6. 힙합(자현님)

-의미 없는 영어 단어 활용으로 인한 문장 생성에 부적합

-서사적 개연성 부족

## 7. 트로트

-트로트 특유의 말투가 문장 생성에 재치를 더해줄 수 있으나 일관성을 주기는 힘들 것

#### 8. RnB

- 이유 찾기 어려움

#### 9. 발라드(종현님)

- 사랑에 관한 내용 다수

- 계절 별로 차이를 발견하기는 쉽지 않음

- 메이저 장르의 가사를 분석하는 것이 자연스러운 문장 생성과 연결될 것이라 생각

---

#### [해야 하는 것]

=> 각각의 타당한 이유 도출 필요 (논문, 기사, 기타 자료 찾기)

=> 시각화

=> 서론 스크립트 제작

## [자료]

### (1) 한국에서는 음악보다 가사가 더 중요하게 인식된다

-Keith Howard and 양우석. (2016). 발라드 전성시대: 한국 대중음악의 변신. 音.樂.學, 24(2), 187-207.

=> 작사함에 있어 위와 같은 요소가 충분히 숙고되어 있을 것. 가사가 전반적인 사회상과 시대 흐름, 혹은 지향하고자 하는 호소를 분명히 담고 있을 것이라 판단. 문장을 생성에 있어 '동시대성'을 끌어낼 수 있을 것. 이러한 '동시대성'이 생성된 결과물을 질적으로 안정되게 해 줄 것이라 판단. (off-topic, 말이 되지 않는 구 등의 등장 축소에 기여할 것)

(2) 발라드는 대중음악의 한 부류로, 남녀간의 사랑을 주제로 한 감상적인 노래 지질, 느린 박자와 낭만적이고 감성적인 가사를 지닌 노래와 때로는 빠르면서도 반복적이고 서술적인 노래를 포함. 이야기 형태의 시나 악곡을 지칭하던 발라드에서 유래(본뜻이 스토리를 가진 노래=>담시)

역사적으로 오늘날의 신문이나 소설책의 역할을 한 것으로 보여지며 민중생활 속에 깊이 뿌리를 박고 있었음.

-위키백과. 발라드(대중음악)

=> 발라드는 서사적이라는 특징을 다른 장르보다 두드러지게 나타낸다. 단어의 본뜻이 스토리를 가진 노래라는 어원을 통해 파악할 수 있듯 서술적으로 한 곡 안에 하고자 하는 이야기를 녹아냄으로 '문장' 생성에 적합할 것.

### (3) 문화적, 개인적, 사회적 요소를 분석하기에 적합한 장르라고 판단

- 발라드를 즐겨듣는 이들이 집단주의적 성향을 띤 것으로 파악 가능
- 집단주의 성향을 지닌 사람들은 음악을 개인적, 사회적, 문화적 기능으로 고루 즐기나, 그 중 특히 음악의 사회적 기능을 주로 사용
- 집단주의 성향의 사람들이 각성 수준이 낮은 것을 추구(덜 반사회적, 덜 계몽적)
- 집단주의 성향을 지닌 사람들은 느리고 감미로운 멜로디의 음악을 좋아한다는 연구결과
- 집단주의 성향의 사람들은 집단 내의 조화와 관계 중요시, 관계를 위해 음악을 사용하는 경우 존재
- 집단주의 성향은 음악 선호의 상관관계가 유의했으나 개인주의 성향이 음악선호에 미치는 영향 유의미하지 않음

-이인영 and 박혜경. (2020). 문화를 비추는 거울: 대중음악 선호에 반영된 한국인의 문화성향을 중심으로. 한국심리학회지: 문화 및 사회문제, 26(3), 221-257.

=> 발라드는 집단주의 성향을 띄는 사람들이 선호하는 장르. 집단주의 성향을 띄는 이들은 음

악을 개인적, 사회적, 문화적 기능 모두를 즐기나 그 중에서 특히 사회적 기능을 담당하는 역할로 음악을 활용하는 경향이 있음. 공동체 문화가 사회 전반에 자리하고 있는 대한민국 사회에서는 발라드가 가지는 이러한 특징이 다수의 기준에 부합하지 않는 언어 사용을 충분히 배제할 수 있을 것이라 판단 -> 반사회적, 하위문화적 요소로 인한 불편함 감소(비속어, 성적 암시 등)

(4) 대중음악에서 사랑이야기는 흔한 주제, 그러나 오직 사랑이야기만을 다룬다는 점에서 발라드는 독특한 장르

사랑은 누구나 한 번쯤 경험해봤을 법한 광범위한 공감대 형성의 가능성을 가지고 있어 대중 영역에서 대중성을 확보하는 서사 기제로 활용되어 왔음

다른 장르와 달리 사랑 이야기 이외의 소재가 등장하는 사례 전무

1970년대 이루어진 대중음악의 단속과 정화 움직임으로 인해 장르가 '사랑'으로 더욱 한정된 것으로 보임-국가 안보와 국민 총화에 악영향을 줄 수 있는 것은 금지하나 "사랑은, 그것이 국가에 대한 것이건 부모 자식간의 것이건 남녀 간의 것이건 허용" 된다.-> 대중 음악사에서 사랑이라는 주제가 폭발적으로 증가함

-한국 대중음악에서의 발라드는 이 같은 금기와 허용이 교차 투사되어 형성된 장르라 판단

-송화숙. (2017). "이 죽일 놈의 사랑" : 발라드의 성정치학. 音.樂.學, 25(1), 7-33.

=> 대중음악 전반에 대한 단속이 행해지던 시절에 '사랑'이라는 주제에 대한 허용이 이를 주로 다루는 발라드가 주요 장르로 지속하는 것에 도움을 주었을 것이며 '사랑'이라는 주제에 대한 당연시로 이어졌을 것.

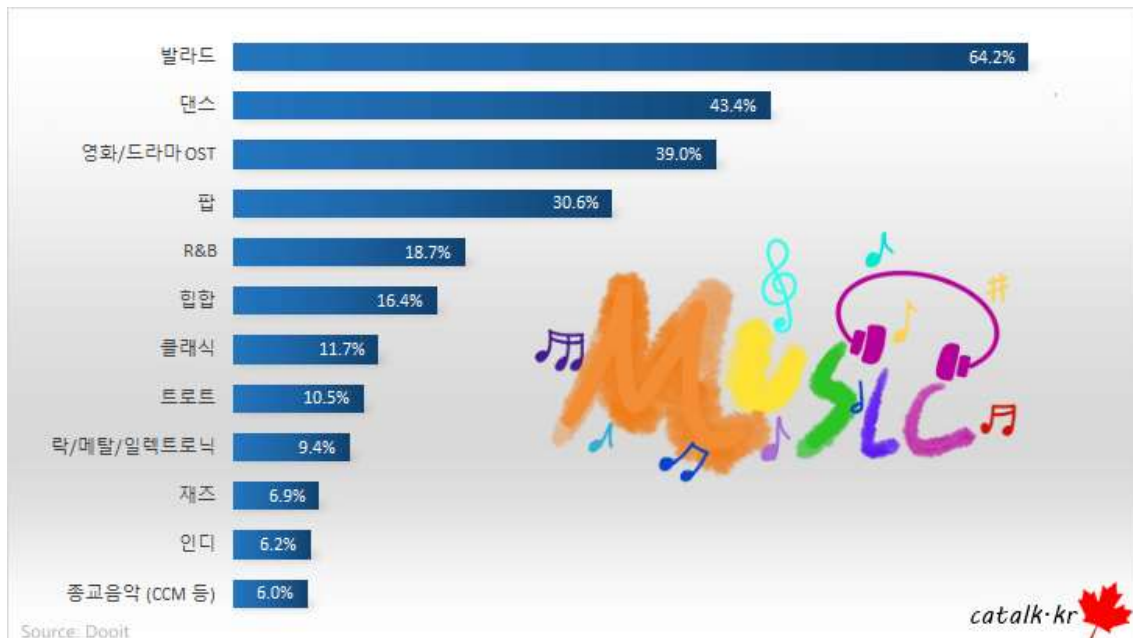
=> 누구나 한 번쯤 경험해봤을, 다수에게 공감대 형성을 할 수 있어 생성된 문장에 대해 그 문장을 접하는 이들의 언어에 대한 경계를 낮출 수 있을 것(무슨 의미인지, 무슨 말인지에 대한 의구심과 비난보다는 '이런 것도 가능하구나'하는 정도의 호기심, 흥미 고취에 기여할 친화도를 가지고 있는 주제 및 언어일 것(장르적 특성상)

(5) 한국 대중음악 명반 100의 1위가 팝'발라드'인 것

=> 발라드의 대중성

(6) 한국인이 선호하는 음악장르 TOP12

- 발라드: 64.2%



(온라인 설문 조사 두잇서베이가 2021년 10월 26일부터 11월 3일까지 전국 4,822명을 대상으로 실시한 설문조사, 표본오차는  $\pm 1.41\%P$ , 신뢰수준 95%)

-1)블로그. 한국인이 선호하는 음악 장르 TOP12

## [결과]

### \*\*\* 왜 아닌지에 대한 설득 필요

가사를 통한 문장 생성이라는 프로젝트 목표를 달성하기 위해 가사 데이터를 수집했다. 그러던 중 장르별로 가사가 상이해 문장 생성 시 맥락이 이어지지 않거나 불용어가 튀어나오는 등의 문제가 발생하는 것을 알게 되었다. 해파리 조에서는 문장 생성 결과의 질을 높이기 위해 장르를 한정하기로 결정했다.

여러 장르 중에서 발라드에 집중해 문장 생성을 하자는 의견이 나왔다. 발라드 특성상 단일한 주제를 가진 서술적 가사이므로 문장 생성 시 일관성 증진에 도움 및 비교적 매끄러운 문장 혹은 구를 생성할 것이라 판단했다. 장르를 발라드로 한정하기에 앞서 발라드 포함 9가지 장르에 대한 분석을 진행하였다.

(1) 첫 번째로 [가사의 길이 분석](#)을 진행하였다. 가사의 길이가 짧다는 것은 이를 모델에 넣고 학습했을 시 온전한 구 및 문장을 생성해내는 것에 큰 도움을 주지 않을 수 있다 판단 하여 진행하였다. 대체적으로 가사 길이는 500 전후로 나타났다. 힙합에서는 장르 특성상 가사의 길이가 굉장히 길었다. 포크/블루스와 트로트는 분석 결과 가사 길이가 인디, 록메탈, 발라드에 비해 길지 않은 것으로(300~450) 나타났다.

(2) 그 다음, [빈도수가 높은 단어를 확인](#)함으로써 장르별 특징을 파악할 수 있었다. 대부분의 장르에서 두드러지는 경향은 ‘사랑’이라는 주제가 다수를 차지하고 있다는 것이었다. 워드클라우드를 통해 시각화하여 주제 상의 유사성을 엿볼 수 있었고, 트로트/국악 과 같은 장르는 ‘님, 세월 등’ 쓰는 단어에서 장르적 특성이 드러났다. 이를 통해 우리가 일상생활에서 접하는 언어와의 유사도가 떨어지는 가사로 쓰인 장르를 선별하여 배제할 수 있었다.

(3) 가사 내 [한국어 단어 수와 영어 단어 수 비교](#)를 통해 영어의 비율이 높은 장르는 문장 생성 시 맥락적 흐름을 방해할 가능성이 높다고 판단했다. 실제로 영어의 비율이 높고 반복 단어가 많으며 다소 내용이 파편화되어 있는 댄스 장르의 경우 문장 생성의 결과가 좋지 않았기 때문이다.

- 영어 단어의 수가 높은 것은 댄스, 힙합, RnB, 인디, 록/메탈인 것으로 파악되었다.
- 국악과 트로트, 포크/블루스의 경우, 장르적 특성상 한국어의 비율이 월등히 높았다.
- 발라드 또한 한국어의 비율이 높았고 영어가 차지하는 비율이 현저히 낮았다. 이를 통해 발라드 선정 시 n행시가 문맥적으로 이어질 가능성이 높다고 판단했다.

(4) 마지막으로 발라드를 선택한 장르 특성에 대한 [자료조사](#)를 했다. 발라드는 2021년 실시한 설문조사에서 한국인이 가장 선호하는 장르 1위로 뽑힌 바 있으며, 한국 대중음악 명반 100의 1위가 발라드의 종류인 것으로 보아 한국인에게 가장 친숙한 장르이자 쉽게 받아들일 수 있는 정서와 그에 걸맞는 언어로 가사가 쓰여졌음을 알 수 있다.

이러한 발라드의 대중성과 ‘사랑’이라는 주제적 통일성, 서술적인 가사 특성으로 인해 모델 학습 데이터의 장르를 발라드로 한정하게 되었다.

