

# AI 힙합 작사가

비타민 프로젝트 2조  
고지형 김상희 배진수 황호진

>>>>  
01

주제 및 배경

02

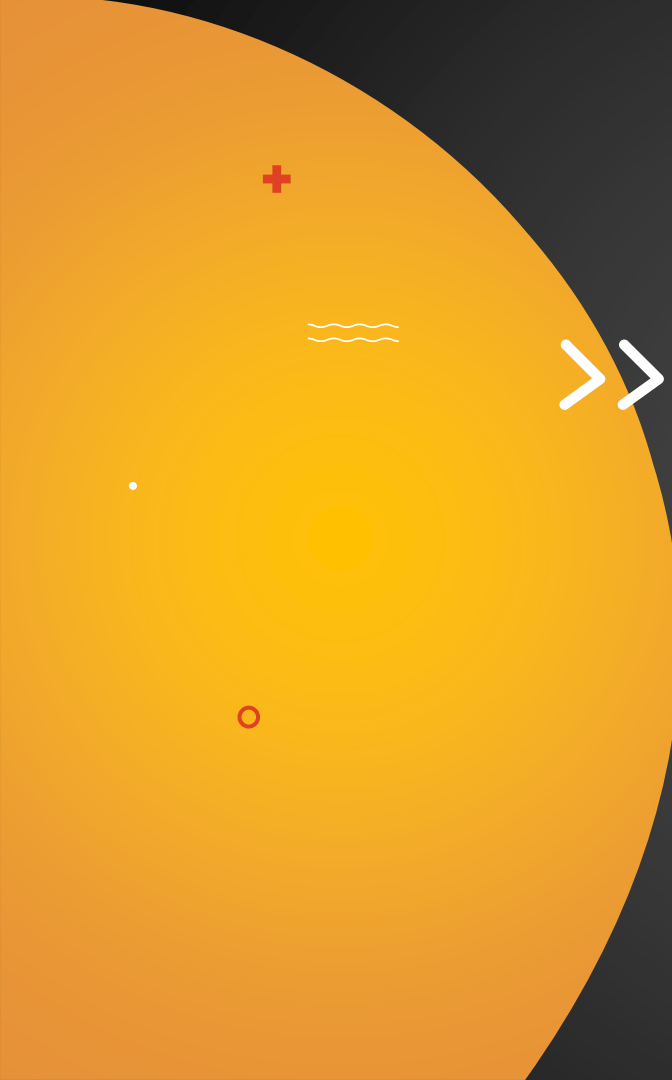
작사 시스템

03

랩 할 시간

04

기대 효과



>>>>

# 01 주제 및 배경

Topic and Background of the theme

+

# AI 힙합 작사가 Text2Hip



국내 힙합 대중화

누구나 힙합을 하고 싶다면?



A man with short dark hair, wearing large over-ear headphones and a black turtleneck sweater, is shown from the waist up. He has his eyes closed and a joyful expression, looking upwards and slightly to the left. His hands are positioned in front of him, palms facing up, as if feeling the music. The background is a solid, vibrant red. Scattered across the background are several small, white and yellow geometric shapes: a white plus sign at the top center, a white circle at the top right, a white dot on the left, a yellow plus sign near the man's right arm, a yellow dot at the bottom right, and a white circle on the man's left hand. The text 'Lets Text2Hip' is overlaid in the center in a large, white, bold, sans-serif font. The word 'Lets' is lowercase, while 'Text2Hip' is a mix of uppercase and lowercase letters with a subscript '2'.

Lets Text2Hip



빠르고



다양하고



풍부하게



Text2Hip만 있다면

힙 알 못 (배모씨) 도  
힙합 작사가 가능하다!



빠르고



다양하고

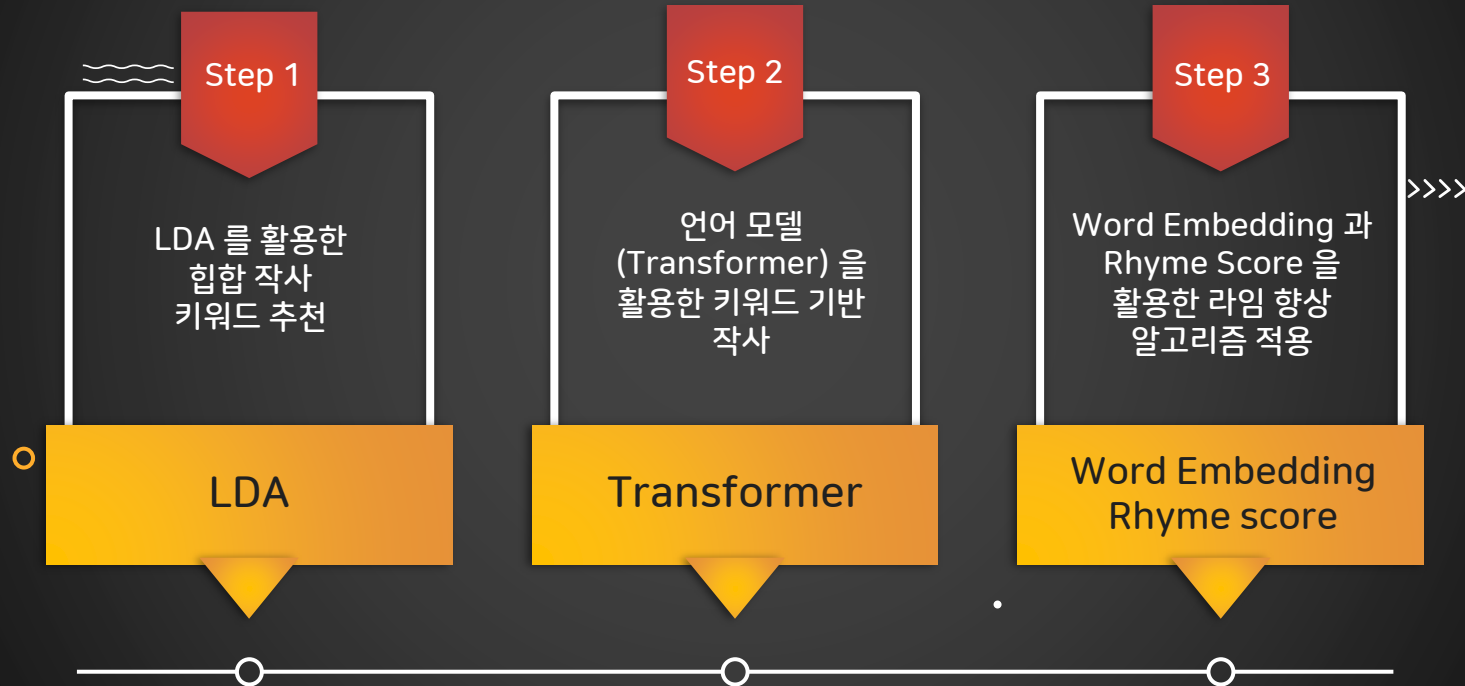


풍부하게

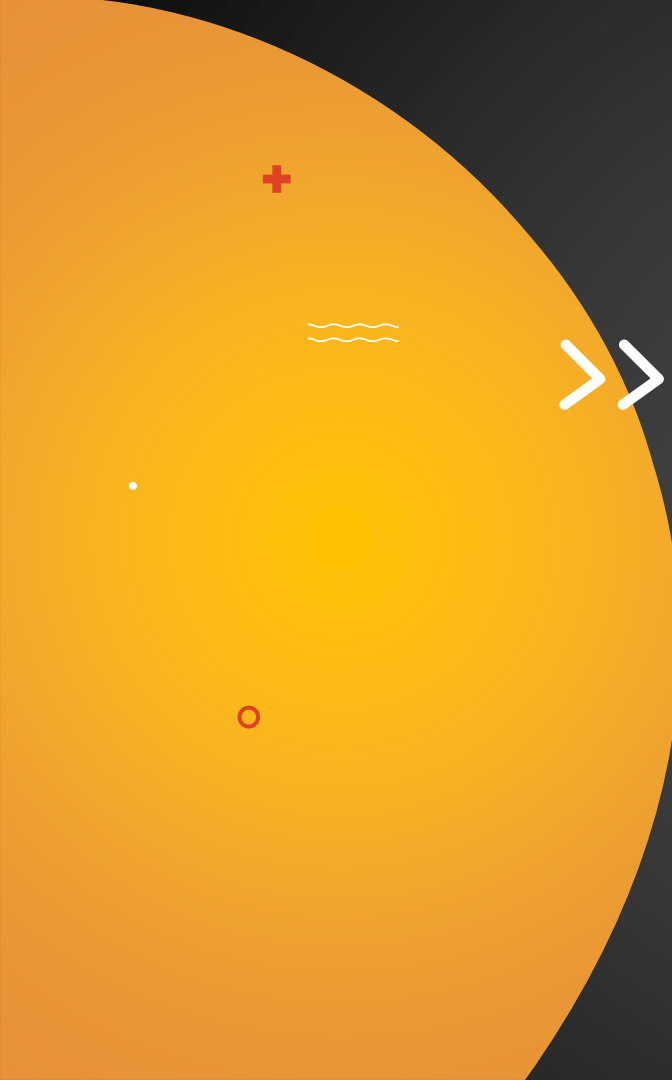


```
import texttship
```

# How ?







>>>>

>>>>

+

o

02

# 작사 시스템

Composing System

.

+

Step 1

# LDA 힙합 작사 키워드 추천

**Bugs!**

34,000 곡 데이터 구축

LDA

작사를 하고 싶지만 >>>>

키워드가 마땅히 떠오르지 않을 때

Bugs 힙합 노래 가사 데이터로

단어들을 추천해드립니다

Step 1

# LDA 힙합 작사 키워드 추천

What is LDA ?

수집된 곡

맘대로 즐겨 오늘 더 신이

Every mood is on the

난 타고 났어 Cash flow

불지를 가져다와 다아 Bank

법대로 해 법대로 생각하지 마

때론 나약해질 때도 나에게  
넌 기대라고 해 때론 실패하고  
와도 기죽지 말라고 해 뭐...

LDA

Topic 1

그대, 다시, 나를,  
우리, 너의, 내가,  
없는, 함께, ...

Topic K

Black, all, flex,  
sorry, 내가, Do,  
...

Step 1

# LDA 힙합 작사 키워드 추천

실제 결과

Topic	Key Word
이별 / 그리움	그대, 다시, 나를, 우리, 너의, 나의, 너를, 내가, 없는, 함께, 내게, 속에, 이젠, 나는, 그댈, 모든, 그대가, 이렇게, 너와, 있는
사랑 / 고백	Love, you, 사랑해, 너와, 너의, my, 나의, 너를, baby, 함께, 우리, 사랑, 싶어, 좋아, 내게, 너에게, 내가, 너만, 항상, 있어
돈 / FLEXING	내가, do, 나는, black, low, 그냥, what, sorry, it, 근데, 하는, 없어, 같은, me, 없지, 이제, all, flex, 나의, 우리

>>>>

+

Step 2

# Transformer 키워드 기반 작사



Input : 힙합 키워드 입력

꼭 , 헤어지고, 남자,  
변해, 계속, 연락했, 넌,  
번호, 바꿨

Transformer

Output : 랩 구절

>>>  
꼭 헤어지고 나면  
남자들은 변해 계속  
연락했더니 넌 번호를  
바꿨어

가족

이별

사랑



+

Step 2

# Transformer 키워드 기반 작사

Sequence data

나는 비타민 출신입니다.

나 파라미터 61만개나 있어~



Sequence data

>>>>

I am from Bitamin.

자연어 처리 분야(NLP)에서 번역 모델로 SOTA(State-of-the-ART)  
성능을 가지고 있는 모델 !

+

Step 2

# Transformer 키워드 기반 작사

Sequence data

꼭 헤어지고 남자 변해  
계속 연락했 넌 번호 바꿨



Sequence data

&gt;&gt;&gt;&gt;

꼭 헤어지고 나면  
남자들은 변해 계속  
연락했더니 넌 번호를  
바꿨어

힙합 키워드로만 구성된 문장에서 자연스러운 힙합 가사로  
복구(번역) 훈련을 진행 !



Step 2

## Transformer 키워드 기반 작사



힙합 가사



수필



시



힙합 가사, 시, 수필 등 46,000여 텍스트 수집

총 110만 문장 !



Step 3

# Rhyme Score 라임 향상 알고리즘 적용

Text2Hip



Sequence data

꼭 헤어지고 남자 뵈해  
계속 연락했 넌 번호 바꿔



Sequence data

꼭 헤어지고 나면  
남자들은 뵈해 계속  
연락했더니 넌 번호를  
바꿨어

작사가가 완료된 문장

Rhyme Score x 워드임베딩

"꼭 헤어지고 나면 남자들은 뵈해"

문장 끝 단어와 앞  
단어들 사이 Rhyme  
Score 파악

변해 차단해  
그만해 바꿨어

워드임베딩 기반  
유사도 활용

Step 3

# Rhyme Score 라임 향상 알고리즘 적용

Rhyme Score: Basic Rule

'사과'

[사과]

↓ 부분  
점수 ② + 부분  
점수 ①  
↓

- I. 소리 나는 대로 읽기
- II. 비슷할수록 높은 점수 부여

\*두 문자 간 초성, 중성, 종성의 구성에 따라 평가

Rhyme Score:

부분 점수의 가중합

'종다'

[조타]

- III. 부분 점수의 가중합을 통해 라임 점수 계산

\*뒷글자 점수 75%, 앞글자 점수 25% 반영

\*두 단어 중 한 문자로만 구성된 단어가 있을 경우, 뒷글자 점수만을 75% 반영

Step 3

## Rhyme Score 라임 향상 알고리즘 적용

Application: 단어 길이가 긴 경우

'다운돼있어'

'재미없어'

[다운돼이씨] [재미업씨]

Rhyme Score 1

[다운돼이씨] [재미업씨]

Rhyme Score 2

⋮

I. 윈도우를 움직이며 각 윈도우마다 Rhyme Score 측정

“최종 Rhyme Score:  
각 윈도우 Rhyme Score의  
가중합”

II. 구한 각 Rhyme Score의 가중합  
으로 최종 점수 계산

\*뒤쪽 윈도우의 점수일 수록 많이 반영

\*어느 하나의 윈도우가 어느 문자도 담지 못하면 필터링 종료

Step 3

# Rhyme Score 라임<sup>+</sup> 향상 알고리즘 적용

Scoring Rule

## Great

[엄] [검]

I. 모음이 모두 'ㅓ'로 일치

II. 두 문자 모두 받침이 있고,  
울림소리에 해당

## Good

[악] [향]

I. 모음이 다르지만, 발음시  
비슷한 입모양을 보임

II. 두 문자 모두 받침이  
있지만, 'ㄱ'은 안울림소리,  
'ㅇ'은 울림소리에 해당

## Bad

[하] [역]

I. 모음이 다르고, 발음시  
입모양이 다름

II. 한 문자는 받침이 없는  
반면, 다른 한 문자는 받침이  
있음

자세한 내용은 [깃허브링크]를 참조하세요



Step 3

# Rhyme Score 라임 향상 알고리즘 적용

## 모델 성능 평가 지표

단어 보존 점수

라임 향상 점수

&gt;&gt;&gt;&gt;



단어를 얼마나 보존하고  
있는지 ?



## 문장 내 단어의 라임들이 얼마나 일치하는지 ?



Step 3

# Rhyme Score 라임 향상 알고리즘 적용

단어 보존 점수



Sequence data

꼭 헤어지고 남자 변해  
계속 연락했 넌 번호 바꿨



Sequence data

~~꼭 헤어지고~~ 나면  
남자들은 변해 계속  
연락했더니 넌 번호를  
바꿨어

작사된 문장이 핵심 키워드를 얼마나 보존하고 있는지  
보여주는 평가 지표 !

Step 3

# Rhyme Score 라임<sup>+</sup>향상 알고리즘 적용

라임 향상 점수



Sequence data

“동화 같은 가사들 세상을  
아름답게 비춰”



비춰, 동화

비춰, 같은 >>>>

비춰, 가사들



비춰, 아름답게

Step 3

# Rhyme Score 라임 향상 알고리즘 적용

라임 향상 점수



Sequence data

"동화 같은 가사들 세상을  
아름답게 비춰"



비춰, 동화

0점

비춰, 같은

>>>> 2.1점

비춰, 가사들

1.4점



비춰, 아름답게

2.1점



Step 3

# Rhyme Score 라임 향상 알고리즘 적용

라임 향상 점수



Sequence data

"동화 같은 가사들 세상을  
아름답게 비춰"



비춰, **동화**

**0점**

비춰, 같은 >>>>

**2.1점**

비춰, 가사들

**1.4점**



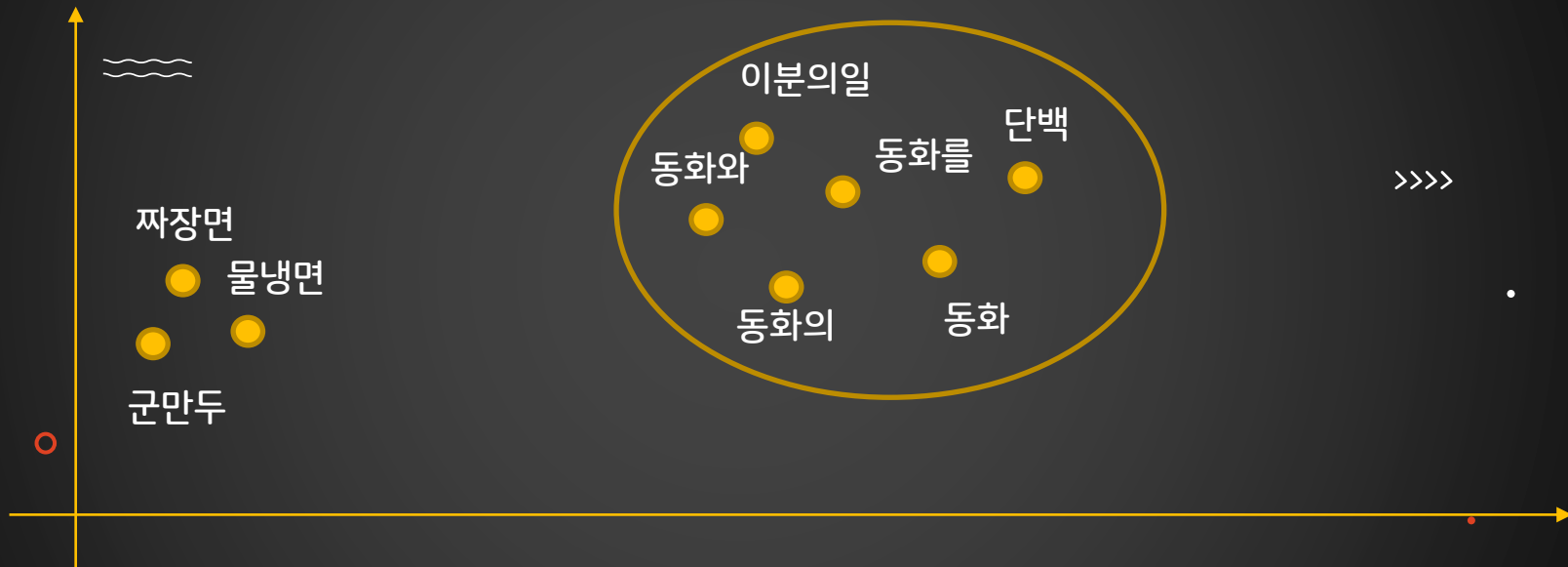
비춰, 아름답게

**2.1점**

Step 3

# Rhyme Score 라임<sup>+</sup> 향상 알고리즘 적용

워드임베딩 공간 내 유사한 단어들 파악



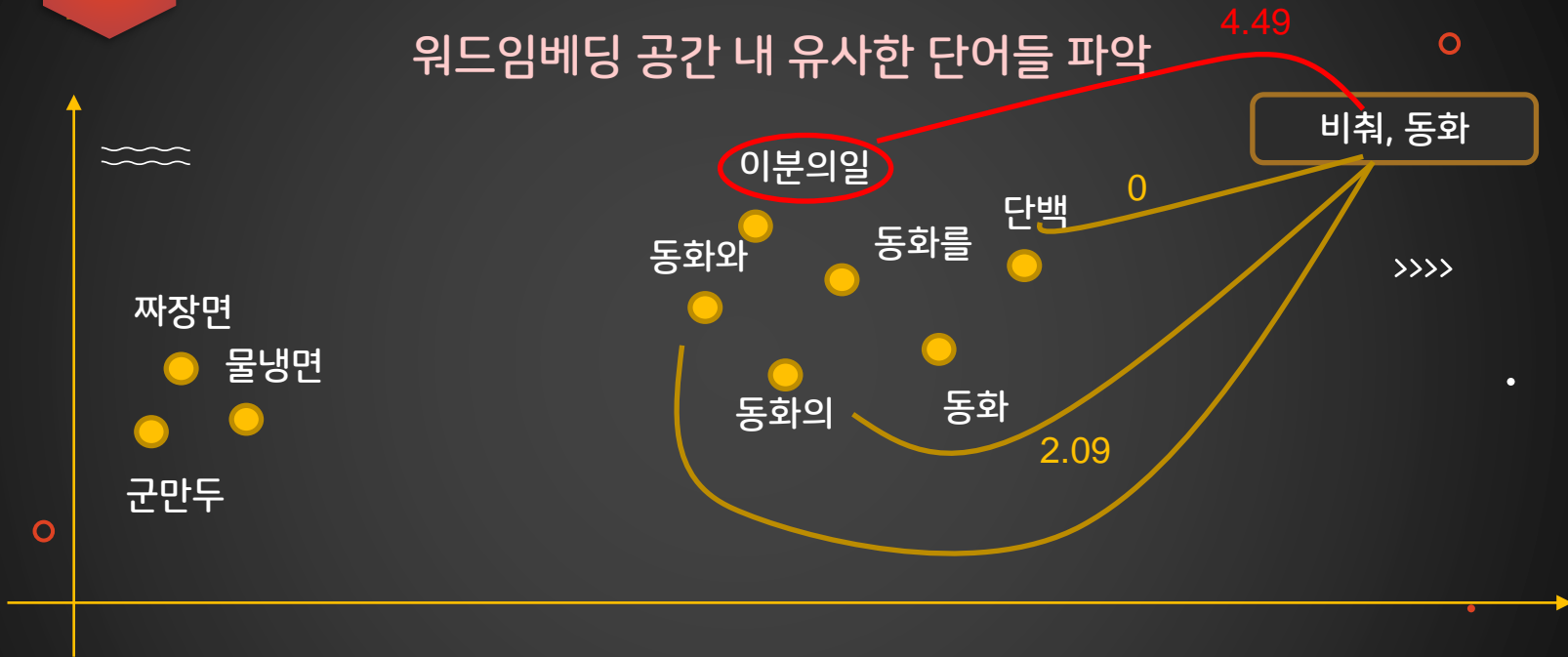
*fast*Text

위키피디아 데이터로 학습된  
pretrained<sup>+</sup> 워드임베딩 벡터!

Step 3

# Rhyme Score 라임 향상 알고리즘 적용

워드임베딩 공간 내 유사한 단어들 파악

The logo for fastText, featuring the word "fast" in red italicized font and "Text" in blue sans-serif font, set against a background of overlapping yellow and grey shapes.

# 위키피디아 데이터로 학습된 pretrained된 워드임베딩 벡터!

Step 3

## Rhyme Score 라임 향상 알고리즘 적용

워드임베딩 공간 내 유사한 단어들 파악



라임향상 전 문장

동화 같은 가사들 세상을  
아름답게 비취



라임향상 후 문장

"이분의일" 같은 가사들  
세상을 아름답게 비취

유사한 의미를 가지면서 라임을 가장 크게  
향상시킬 수 있는 단어로 대체!

Step 3

## Rhyme Score 라임 향상 알고리즘 적용

워드임베딩 공간 내 유사한 단어들 파악



라임향상 전 문장

동화 같은 가사들 세상을  
아름답게 비춰

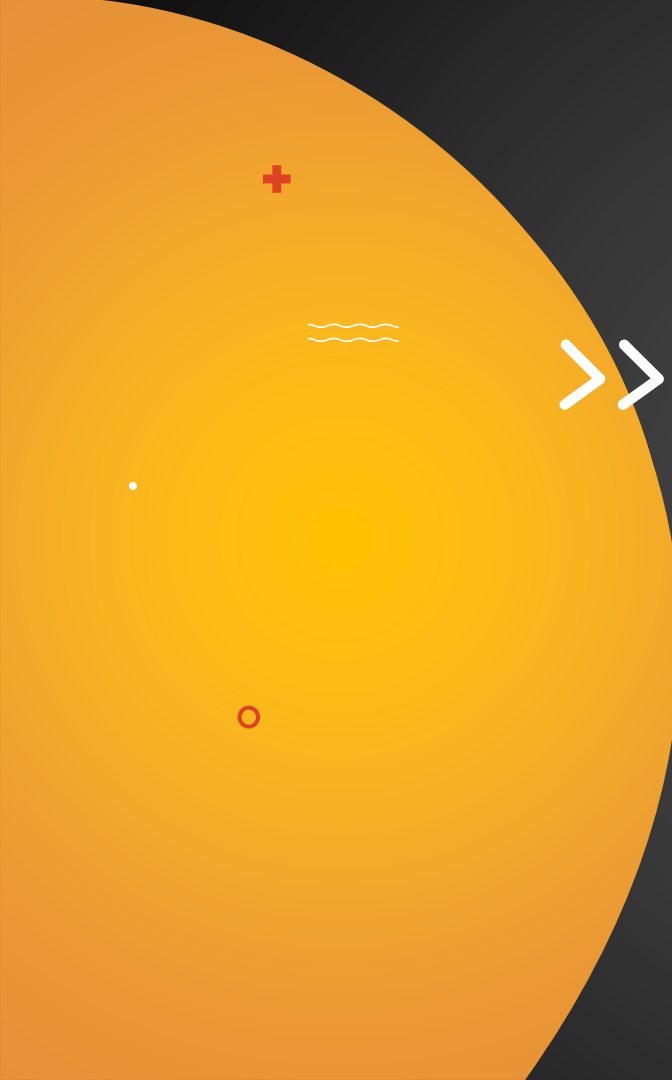


?

라임향상 후 문장

이분의일 같은 가사들  
세상을 아름답게 비춰

유사한 의미를 가지면서 라임을 가장 크게  
향상시킬 수 있는 단어로 대체!



>>>>

+

>>>>

o

03

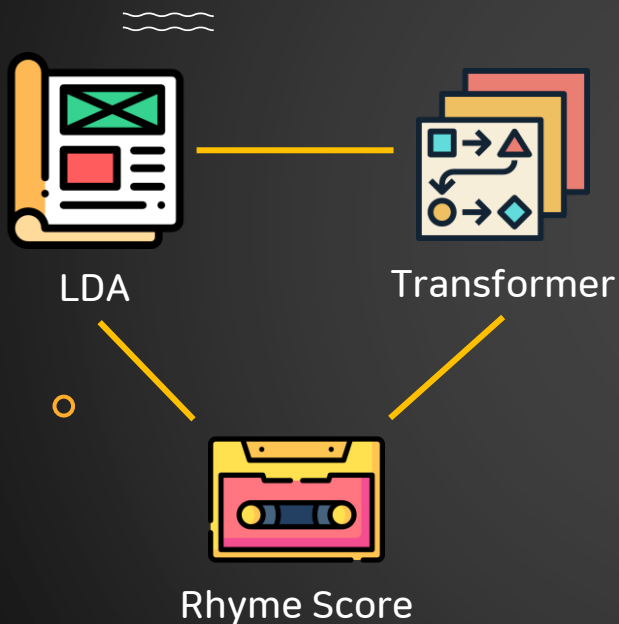
>>>>  
랩 할 시간.

It's time to RAP

.

+

# 모델 적용 최종 결과



최종 문장	Topic
내 마음이 그리워 애써 웃지 않아도 돼	이별, 그리움
매번 어떤 특별한 상황을 사고 하지 못 한 것	사랑, 후회
동화 같은 가사들 세상을 아름답게 비춰	사랑, 고백
죄인 사라진 현장 태어난 우리 어디 흘러갈지	인생
내 가슴 흔드는 잔잔한 물결 뿐 해 맑게 웃 때 나른한 걸까	사랑, 이별
알고 보면 나 다른 게 아니라 정말 요리 재능 있는 것 아닐까 하고	자기 자랑
이 하나로 나 세상 가진 건 아닐까	사랑, 고백

04

# 기대 효과.

Expectancy Effects



# Text2Hip 의 기대 효과



작사를 어려워 하던 힙썰이 (배모씨) 와 힙합 꿈나무들에게  
다양한 키워드와 주제를 보여줌으로써  
새로운 영감을 줄 수 있다

>>>>



막상 주제는 생각했지만  
가사를 쓰는 데에 걸리는 시간을 단축해 줄 수 있다



똑똑한 인공지능 "TextToHip" 를 통해  
비슷한 뜻을 가지며 라임 있는 단어를 단<sub>다</sub> 번에 파악 할 수 있다

# THANKS!

Does anyone have any questions?

비타민 프로젝트 2조  
고지형 김상희 배진수 황호진

