



ToBig's

14th

Conference

Recommendation



Music Recommendation by Image

김종우, 나다경, 염제윤, 이지수



Recommendation





ToBig's

14th

Conference

Recommendation



01

Motivation



MRI

02

Data



Melon Playlist
Image&Caption Data

03

Modeling



Music Recommendation
Image Captioning

04

Result



https://github.com/emforce77/tobigs_conference_14_rec



Recommendation





이미지의 분위기와 유사한 노래 추천



	부산 바캉스 레게 강 같은 평화	⋮
	바다보러갈래 (SEE SEA) 효린	⋮
	여름여름해 (Sunny Summer) 여자친구 (GFRIEND)	⋮
	여름 안에서 by 싸쓰리 (Feat. 황광희) 싸쓰리 (유두래곤, 린다G, 비룡)	⋮





Kakao Arena Melon playlist continuation

- 약 20만개의 Tag, playlist dataset
- 약 60만개의 Song dataset
- song_meta, mel_data, playlist_meta





Motivation

Data

Modeling

Result



Song_meta

곡 제목, 앨범 제목, 아티스트명, 장르, 발매일

Arena_mel

곡에 대한 mel-spectrogram을 담고있는 파일

Playlist_meta

제목, 플레이리스트에 수록된 곡, 태그 목록

Used Data





Motivation

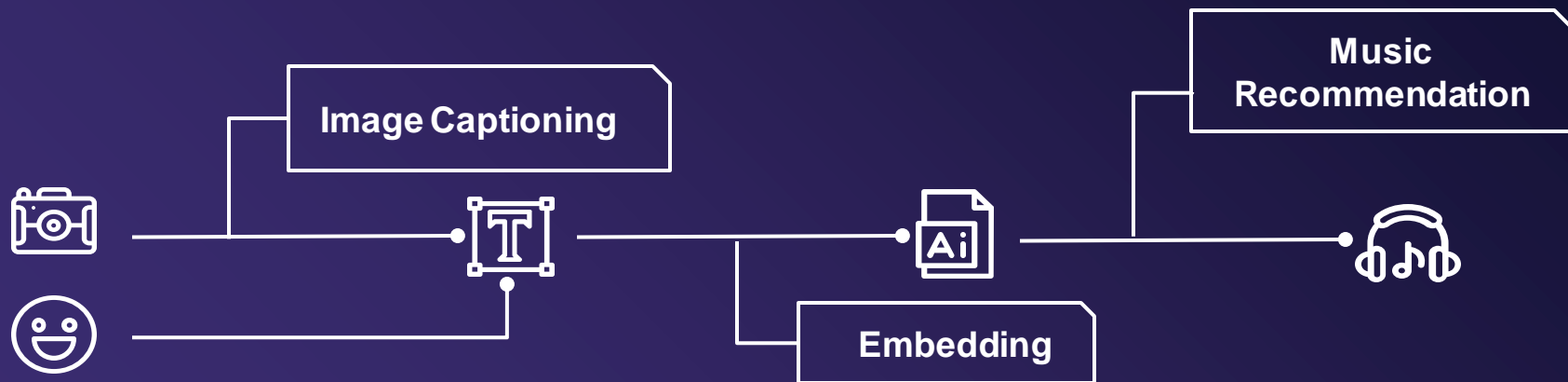
Data

Modeling

Result



Overall Framework



Service Flow





Motivation

Data

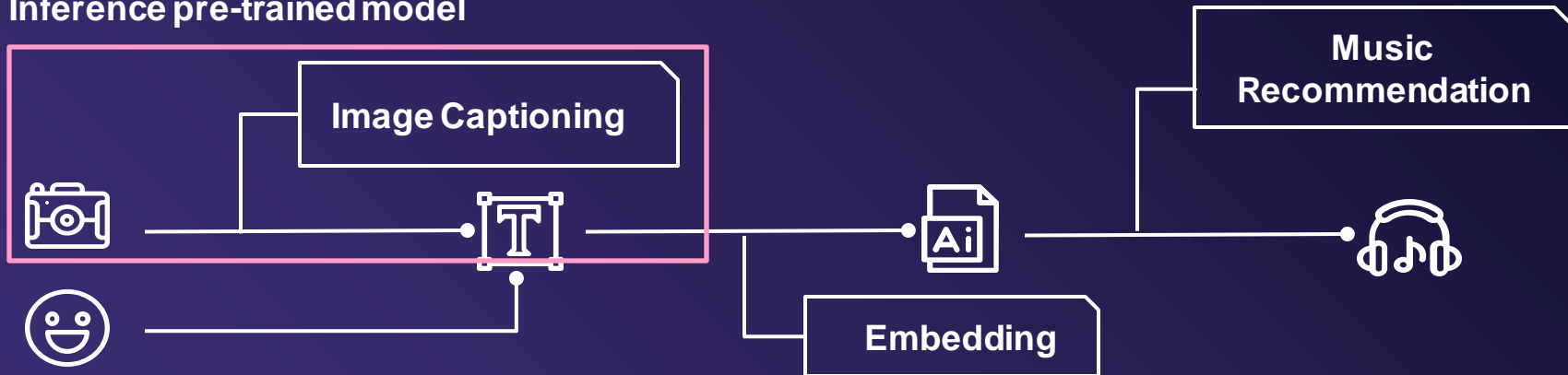
Modeling

Result



Overall Framework

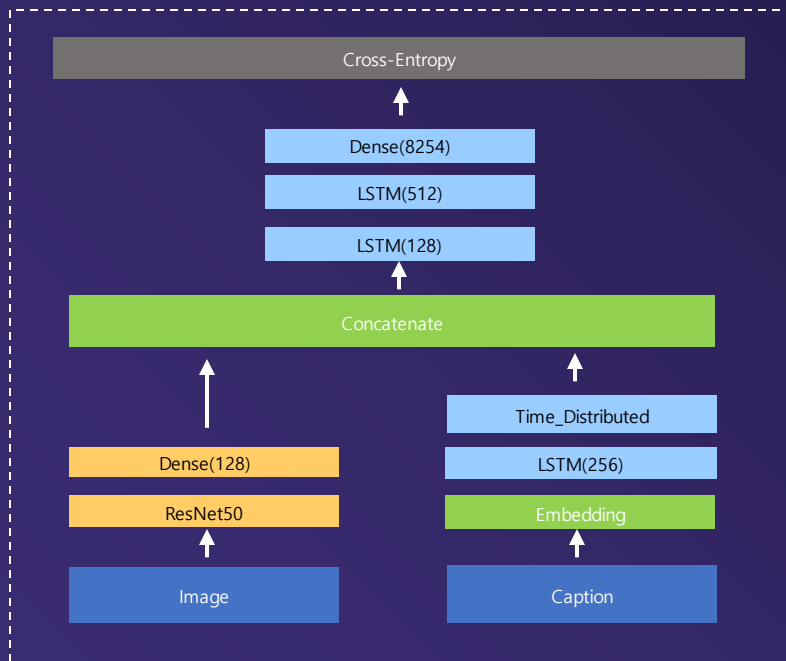
Inference pre-trained model



Service Flow

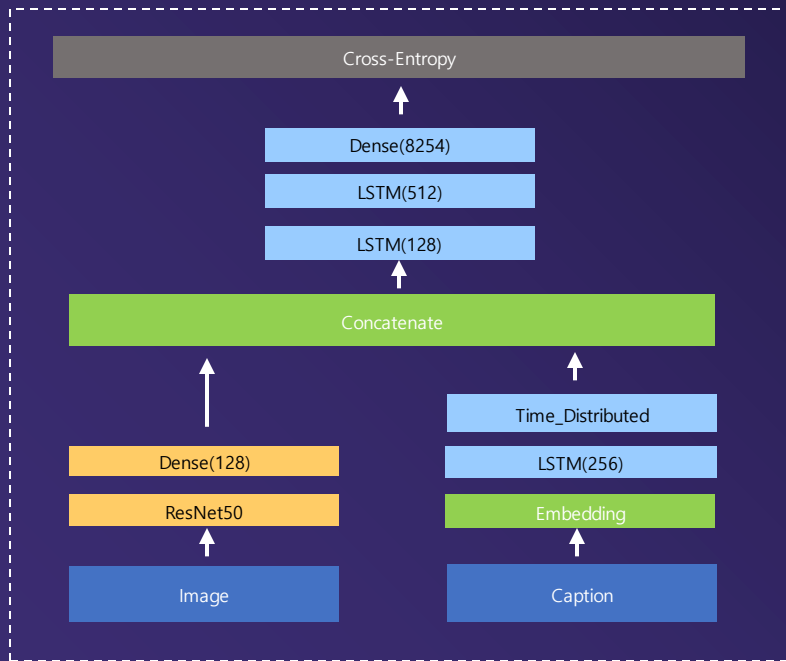


Image captioning model



A man in an orange hat starring at something.
A man wears an orange hat and glasses.
A man with gauges and glasses is wearing a Blitz hat.
A man with glasses is wearing a beer can crocheted hat.
The man with pierced ears is wearing glasses and an orange hat.

Image captioning model



predict



A man and a woman are standing on the beach .

↓
남자와 여자가 해변에 서있다.

↓
남자, 여자, 해변



Motivation

Data

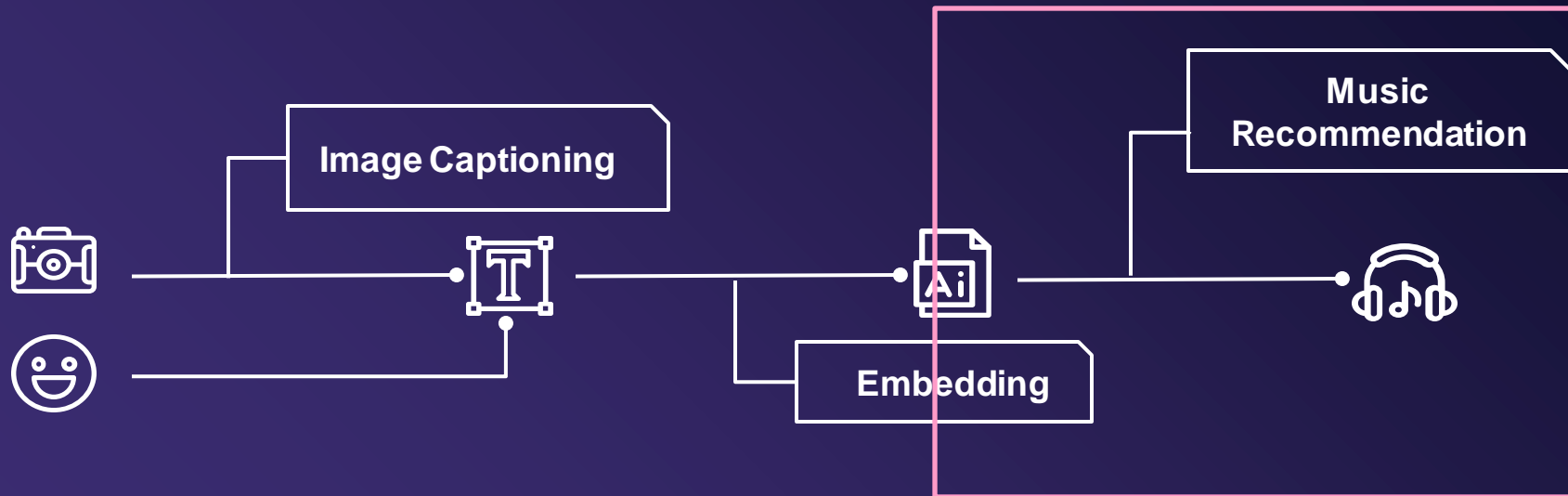
Modeling

Result



Overall Framework

Train Recommendation model

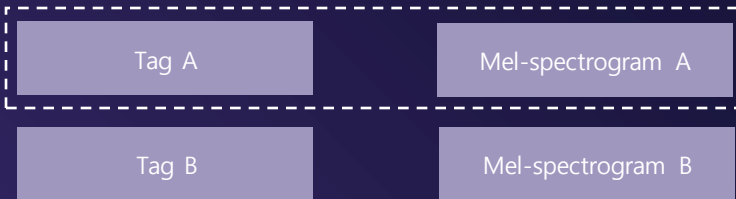


Service Flow

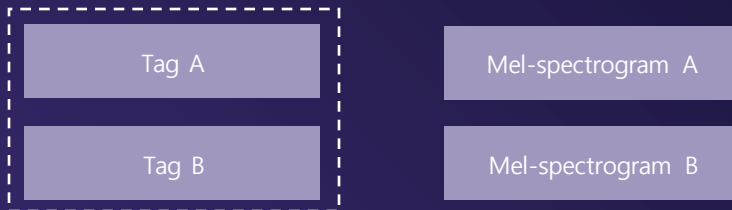


Input Data (define label)

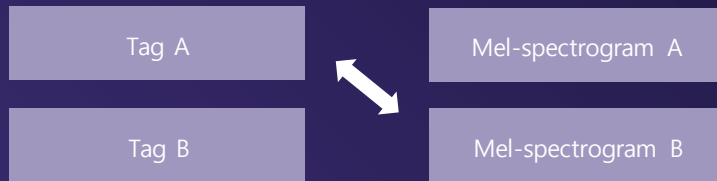
100%의 similarity를 가지는 것으로 가정



Tag A와 Tag B의 similarity를 계산



Tag A와 Mel-spectrogram B의 similarity로 가정





Motivation

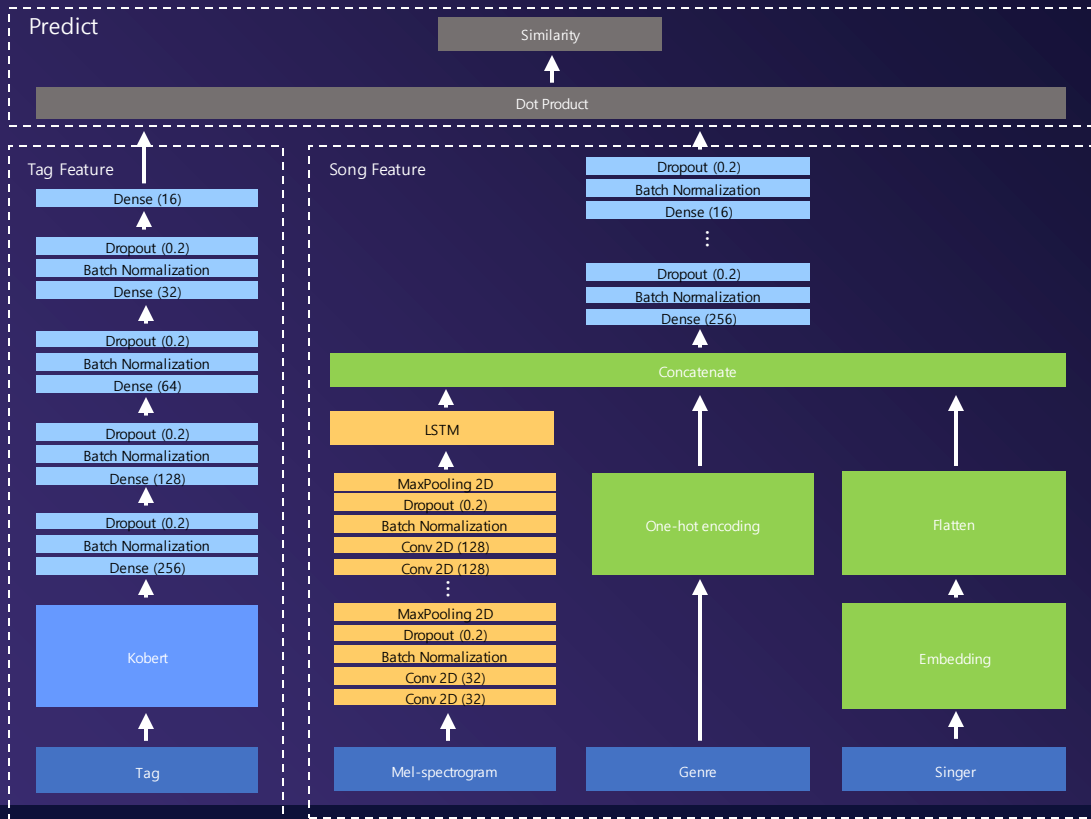
Data

Modeling

Result



Recommendation model



Modeling





Motivation

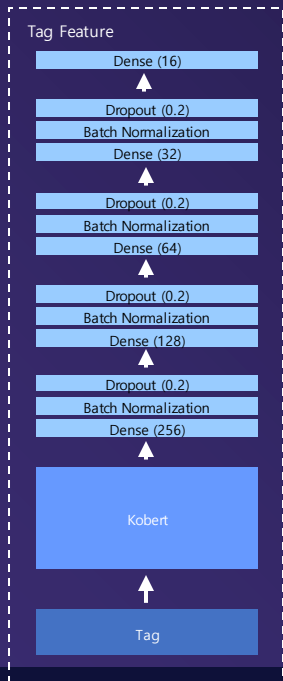
Data

Modeling

Result



Recommendation model



Input data

	tags	songs
	[락]	[525514, 129701, 383374, 562083, 297861, 13954...
	[추억, 회상]	[432406, 675945, 497066, 120377, 389529, 24427...
	[카페, 잔잔한]	[83116, 276692, 166267, 186301, 354465, 256598...
	[연말, 눈오는날, 캐럴, 분위기, 따뜻한, 크리스마스캐럴, 겨울노래, 크리스마스,...]	[394031, 195524, 540149, 287984, 440773, 10033...
	[댄스]	[159327, 553610, 5130, 645103, 294435, 100657,...

	[록메탈, 밴드사운드, 록, 락메탈, 메탈, 락, extreme]	[429629, 441511, 612106, 516359, 691768, 38714...
	[일렉]	[321330, 216057, 534472, 240306, 331098, 23288...
	[담시, 가죽, 눈물, 그리움, 주인공, 나의_이야기, 사랑, 친구]	[50512, 249024, 250608, 371171, 229942, 694943...
	[잔잔한, 버스, 퇴근버스, Pop, 풍경, 퇴근길]	[533534, 608114, 343608, 417140, 609009, 30217...
	[노래추천, 팝송추천, 팝송, 팝송모음]	[26008, 456354, 324105, 89871, 135272, 143548,...

TAG

- English to Korean
- Extract Noun and Adjective

Processing

- 노래A가 포함된 Playlist들의 Tag의 embedding 추출 후 average를 Tag A embedding으로 정의

Embedding

- Kobert-small (256-d)를 사용

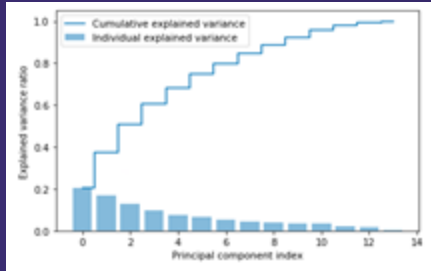
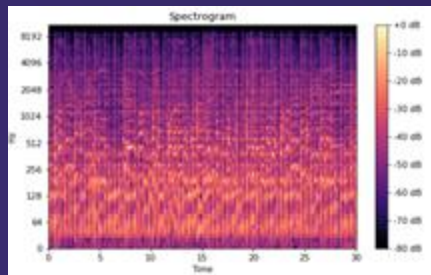
Modeling





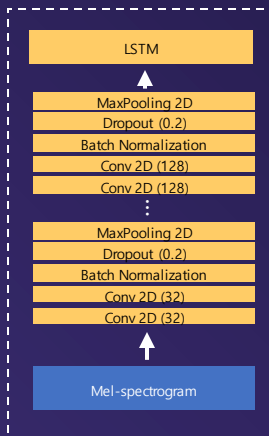
Recommendation model

Input data

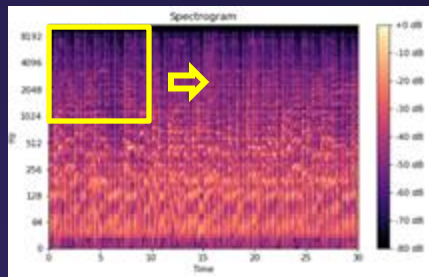


48x1892
240GB \Rightarrow 48x20
2GB

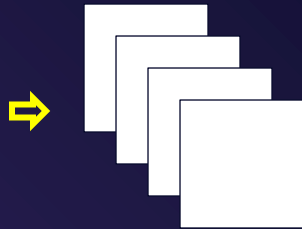
About 85%
Explained Variance



Mel spectrogram



CNN을 통해 feature extract





Motivation

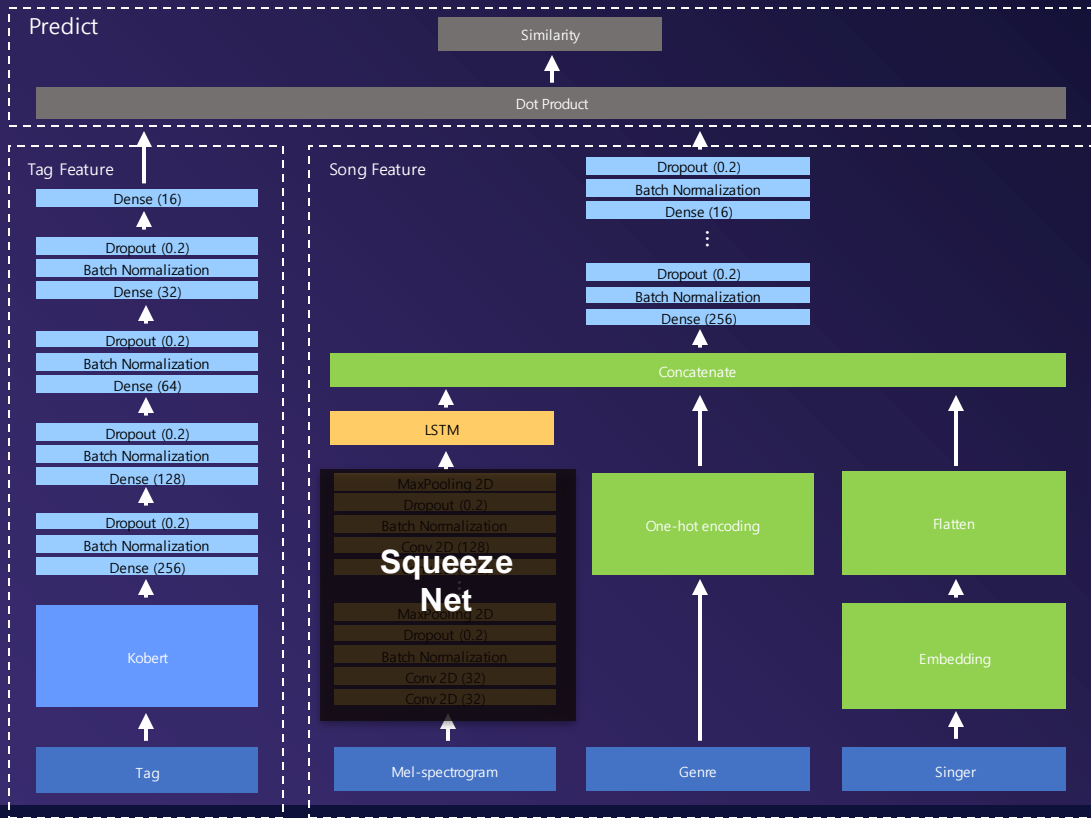
Data

Modeling

Result



Recommendation model



Modeling





Motivation

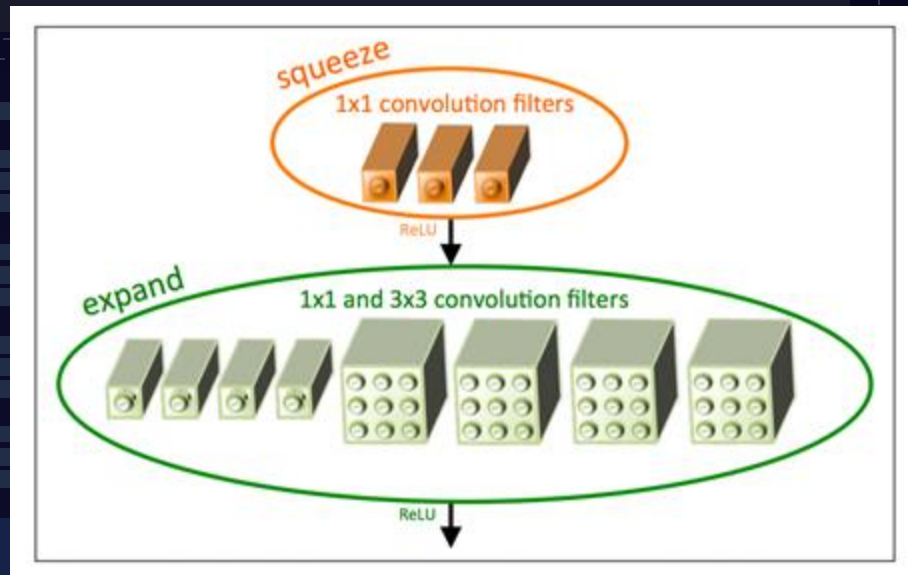
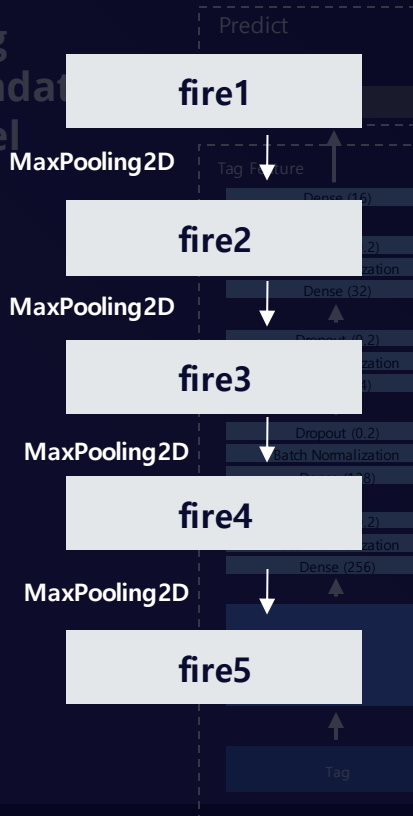
Data

Modeling

Result



Song Recommendation Model



Modeling





Motivation

Data

Modeling

Result



Song Recommendation Model



MSE

nDCG

Kendall's τ

0.0083

0.99

0.4075

Modeling





Motivation

Data

Modeling

Result



해변, 남자, 여자, ...



로맨틱, 밝음

Song_name	Artist_name
Over The Rainbow	Karel Boehlee Trio
다시 여기 바닷가	썩쓰리
Summer Feelings	Lennon stella
Sanctify	Years & Years, Alma
Lie About Us	Avant



Result





Motivation

Data

Modeling

Result



Limitation



Incomplete Mel Data

Mel-spectrogram이 곡의
일부만을 갖고 있음.



Lack of Captioning Adjective

캡셔닝 결과 감정형용사의 부족



Incomplete Metric

불완전한 성능지표



Lack of Infra

전반적인 자원 부족

Restriction





Motivation

Data

Modeling

Result



Improvement



Complete Mel-DATA

전체 곡 데이터를 적용



Adding More Data

곡 관련 데이터 추가



Improvement





Reference

Image Captioning DataSet : Flickr8ksau

<https://www.kaggle.com/datasets/waelboussbat/flickr8ksau>

Kakao Arena Melon-data

<https://arena.kakao.com/c/7>

Deep Convolutional and LSTM Neural Networks for Acoustic Modelling in Automatic Speech Recognition

<http://cs231n.stanford.edu/reports/2017/pdfs/804.pdf>

SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size, Forreset N. Iandola et al, ICLR2017

<https://arxiv.org/abs/1602.07360>





Motivation

Data

Modeling

Result



17기 엄제윤

KINGKINGKINGKING
KINGKINGKING
KING



16기 김종우

내 mbti는 C.U.T.E



17기 나다경

최고의 엔딩 고양이
집사



17기 이지수

음악이랑 거리두기
3단계



Our Team





Motivation



Data



Modeling



Result



THANK YOU



Thank U

