

책을 듣다

Name 이남기

Date 2019.03 ~ 2019.06

Git <https://github.com/brainmining-for-bigdata>



Index



Intro



Service



Model



Conclusion

Intro



Background

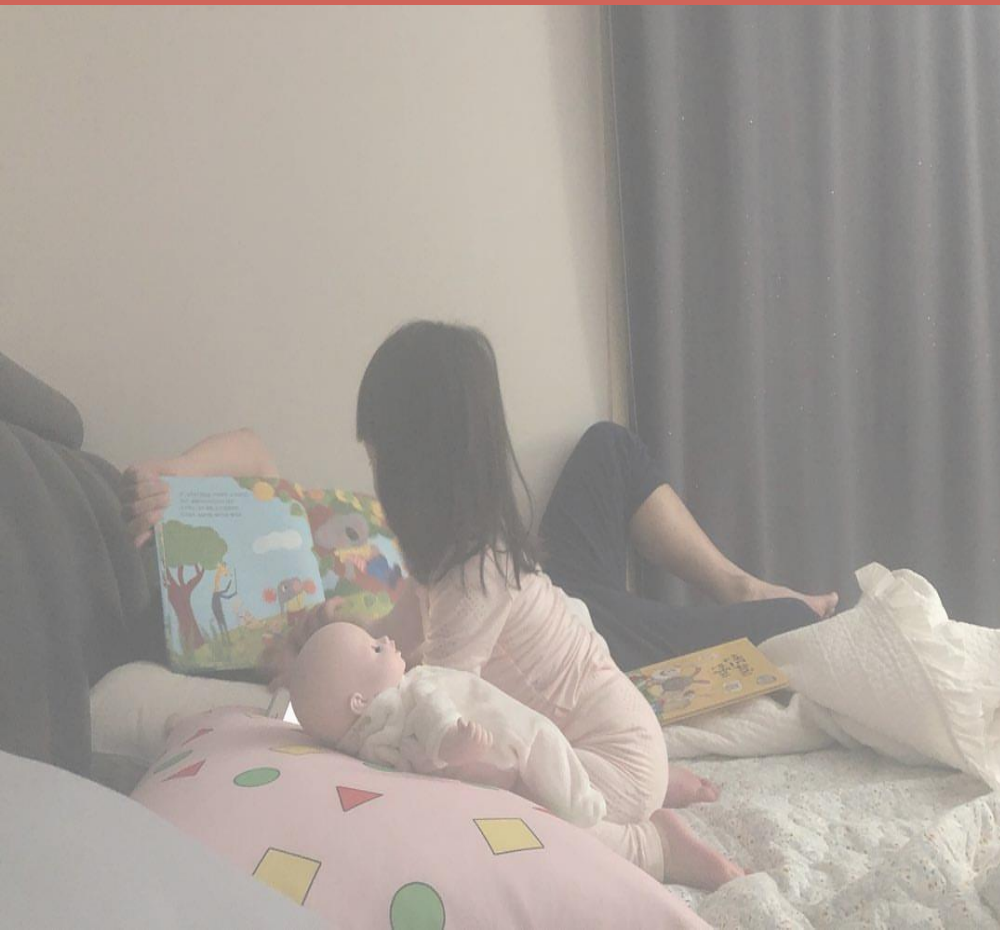
발디딜 틈도 없는 지하철.
간신히 이어폰을 꺼내어 귀에 꽂습니다.

들려오는 목소리.....

「내가 읽고 싶은 책을 누군가
아름다운 목소리로 읽어준다면
얼마나 좋을까요?」



Background



퇴근 후 매일 밤 아이가 잠들기 전 책을 읽어주고 싶지만.

쏟아지는 피곤에 정작 먼저 잠드는건 아빠네요...

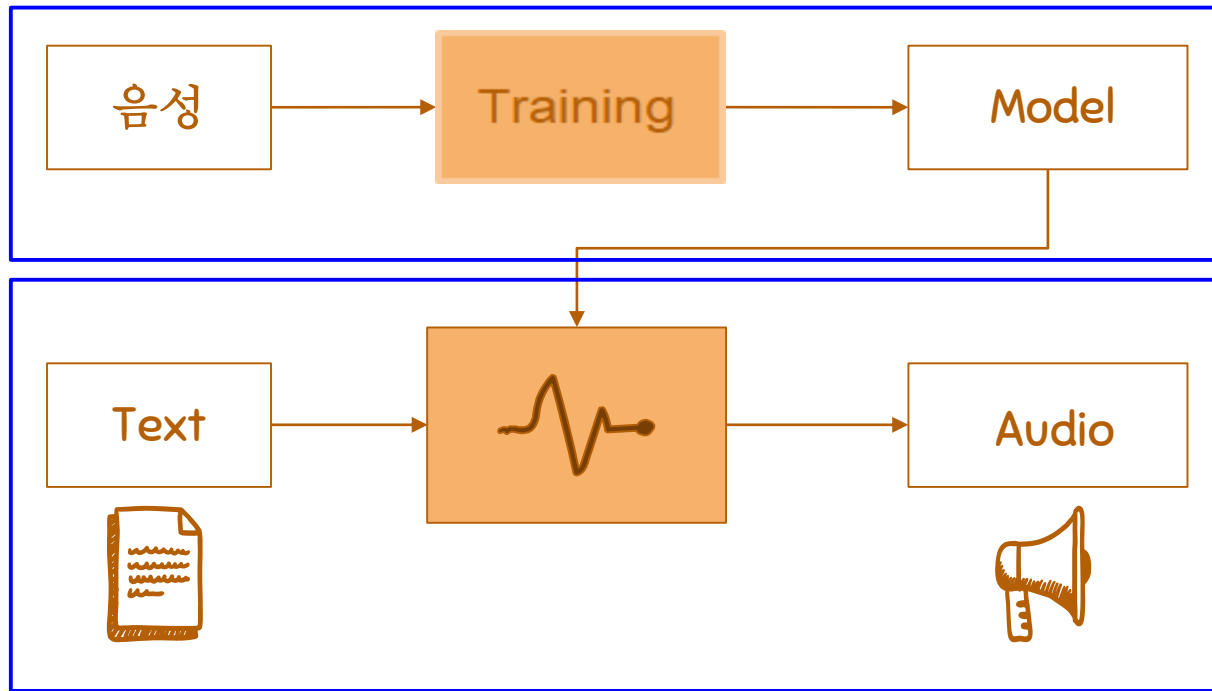
「아이에게 엄마아빠의 목소리로 책을 읽어 준다면?」

Summary

- 프로젝트명 : MARATRON
- 서비스명 : 책을 듣다.
- 프로젝트 정의 : 개인화 음성 합성 기술 TTS(Text to Speech)를 기반으로 목소리를 분석해 학습된 목소리로 글을 읽어주는 웹 서비스
- 활동기간 : 2019.03.12 ~ 2019.06.22

Concept

충분한 시간의 음성만
있다면 텍스트를
원하는 음성으로 합성



Technology Stack

Collaboration



Front-End



Back-End



Database



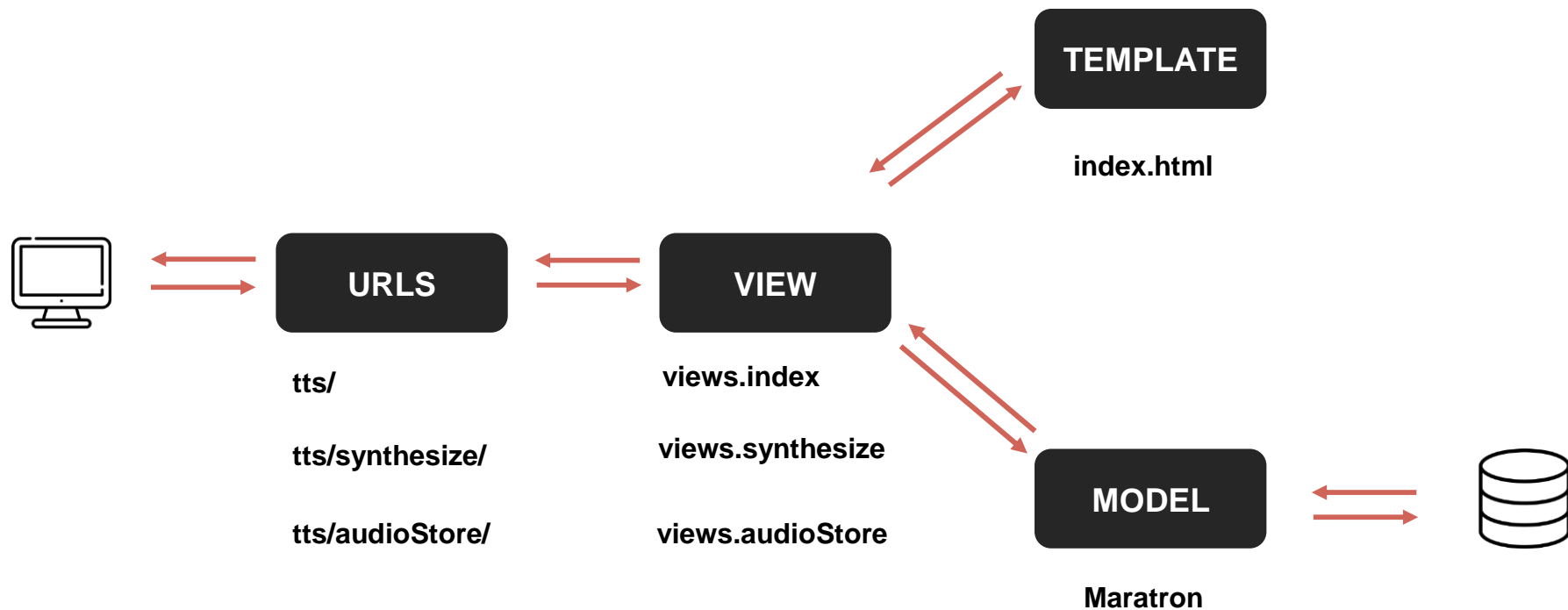
Infra



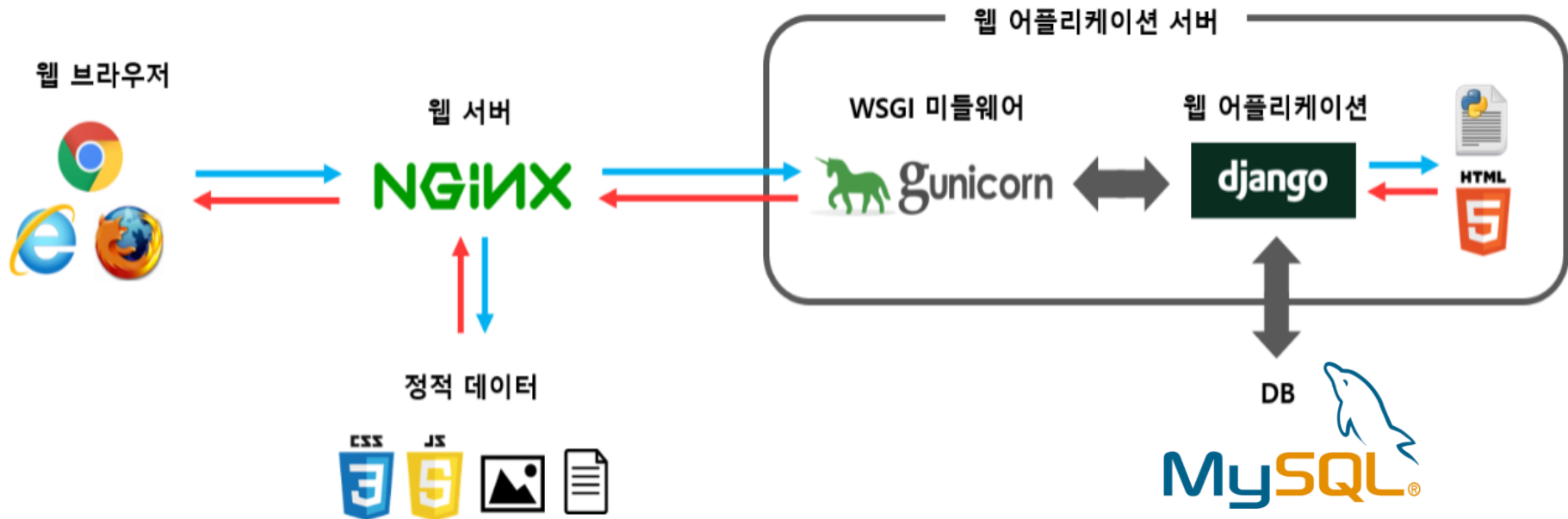
Development



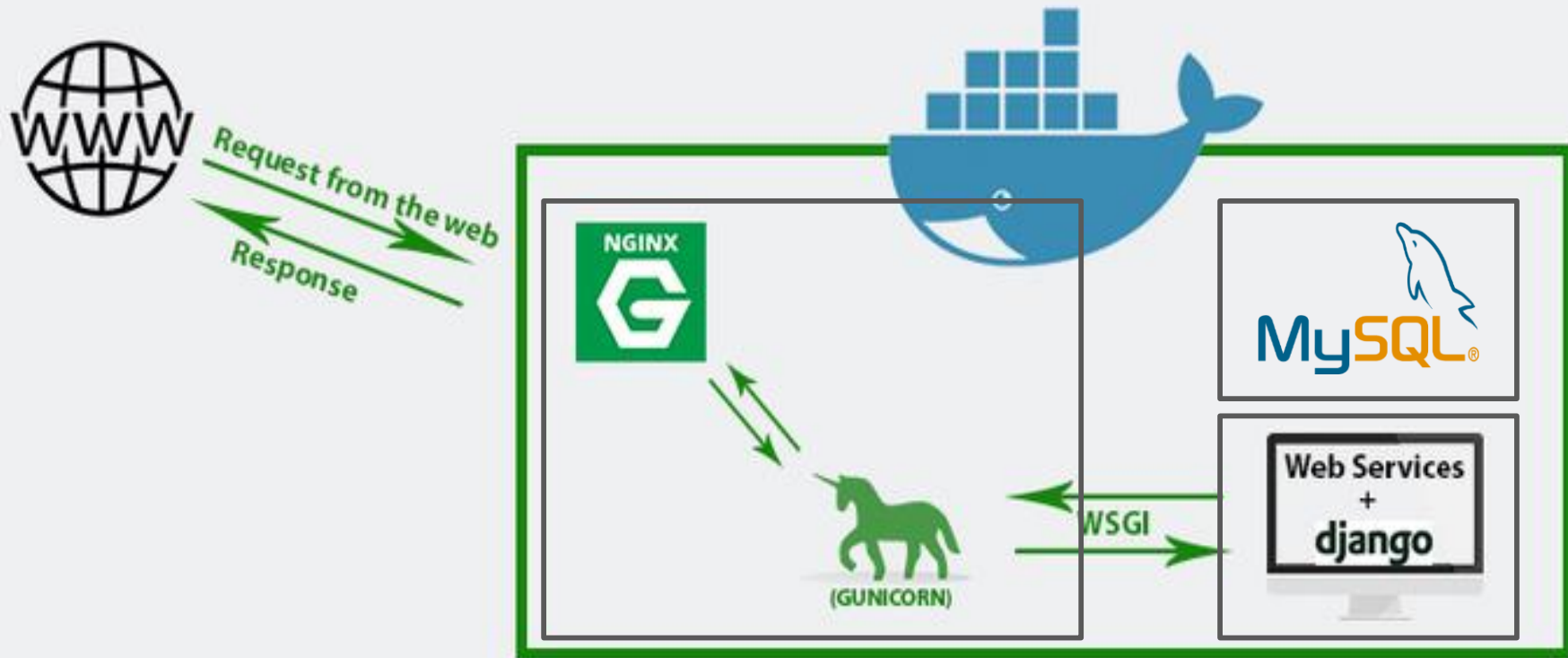
Django MVT Pattern



Deployment : Django Server Architecture



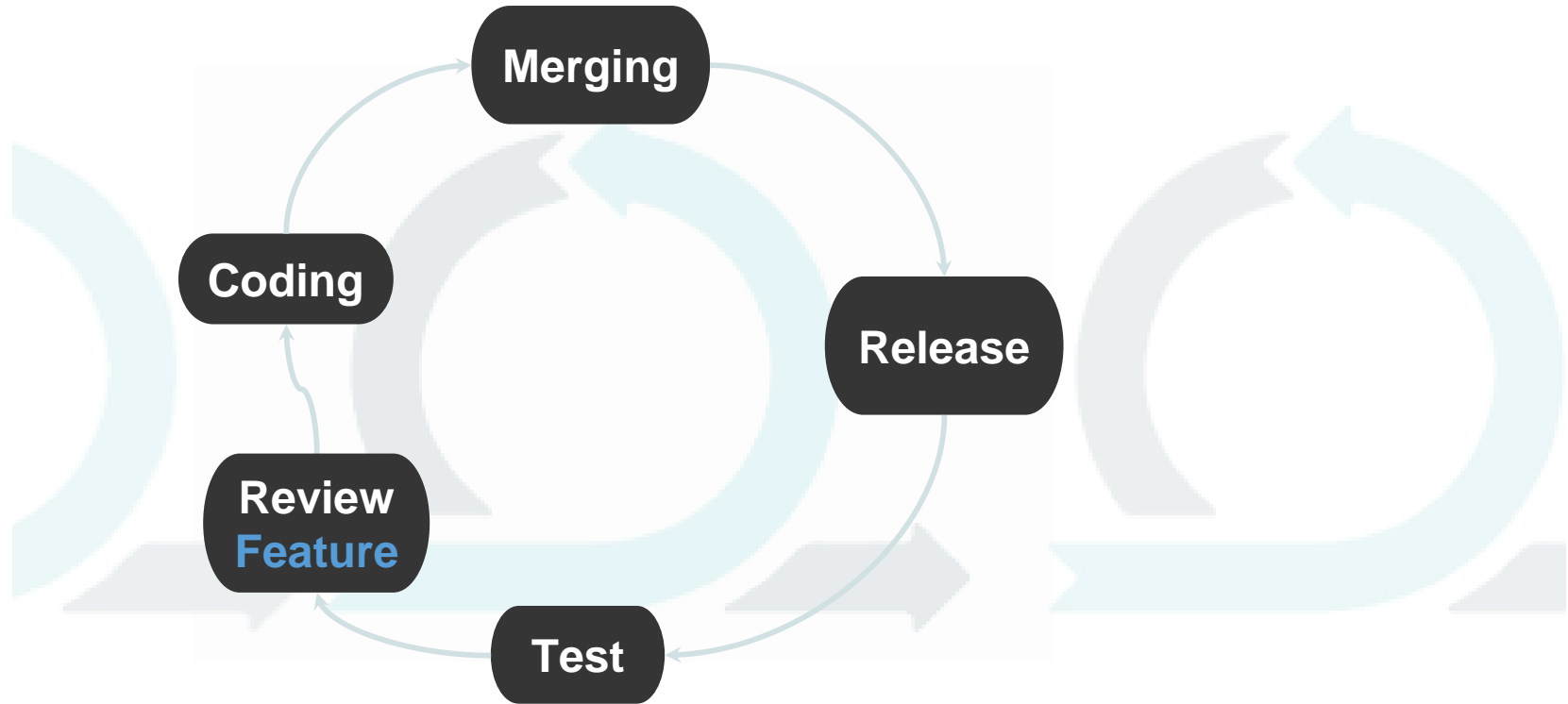
Deployment : Docker 사용 - 3개의 container로 구성




Service Release : 배포 플랫폼 선정



Management : Agile



Management : GitHub Organization account



BrainMining

Repositories 2

People 6

Teams 1

Projects 1

Settings

Type: All ▾



Language: All ▾

Customize pins







New

maratron
Text2Speech
Python ★ 1 Updated 2 days ago

project-text2speech-blog
Project Team Blog
Python 🍷 1 Updated 3 days ago

Top languages
Python

People 6 >



Invite someone

Project Blog

- Overview
- Study
- Technical Docs
- ToDoProgress : 회의록

The screenshot shows the GitHub interface for the repository 'brainmining-for-bigdata / project-text2speech-blog'. The repository has 81 commits, 1 branch, 0 releases, and 5 contributors. The 'Code' tab is selected, showing a list of files and their commit history. The files listed are: Overview, Study, TechnicalDocs, ToDoProgress, .gitignore, Linux Cent OS에서 Jupyter Notebook 서버 외부 접속.pdf, and README.md. The README.md file is selected, showing its content. The content of the README.md file includes the title 'project-text2speech-blog' and a section titled 'Papers' with a list of papers, including 'Google paper (published in April 2017): Tacotron: Towards End-to-End Speech Synthesis'.

Search or jump to... Pull requests Issues Marketplace Explore

brainmining-for-bigdata / project-text2speech-blog Watch 0 Star 0 Fork 1

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

Project Team Blog Edit

Manage topics

81 commits 1 branch 0 releases 5 contributors

Branch: master New pull request Create new file Upload files Find File Clone or download

srshin Update README.md Latest commit f356f62 4 minutes ago

| | | |
|---|------------------------------|----------------|
| Overview | Delete 2019-04-22_Dataset.md | 25 days ago |
| Study | Update README.md | 19 days ago |
| TechnicalDocs | Update README.md | 4 minutes ago |
| ToDoProgress | Update 2019-05-13_Todo.md | 6 days ago |
| .gitignore | initial commit | 2 months ago |
| Linux Cent OS에서 Jupyter Notebook 서버 외부 접속.pdf | Add files via upload | 2 months ago |
| README.md | Update README.md | 14 minutes ago |

README.md

project-text2speech-blog

Papers

- Google paper (published in April 2017): Tacotron: Towards End-to-End Speech Synthesis

Management: Project Management

Kanban System

- Todo
- In progress
- Done

The screenshot displays a GitHub Kanban board for the 'maratron' project. The board is organized into three columns: 'To do', 'In progress', and 'Done'. Each column contains a list of issues, each represented by a card with a title, issue number, and status. The 'To do' column has two cards, 'In progress' has one, and 'Done' has nine. Each card includes a status label (e.g., 'enhancement', 'bug') and a small profile picture of the user who opened the issue. The board is titled 'Maratron' and shows it was updated 2 days ago. The top navigation bar includes links for 'Pull requests', 'Issues', 'Marketplace', and 'Explore'.

Search or jump to... Pull requests Issues Marketplace Explore

brainmining-for-bigdata / maratron

<> Code ① Issues 4 Pull requests 0 Projects 1 Wiki Insights Settings

Maratron Updated 2 days ago

2 To do

- [fix] tokenizer pickle load path 변경 #66 opened by iamnamki enhancement
- [feature request] Google Cloud Platform 배포 #15 opened by iamnamki enhancement

1 In progress

- [feature request] Eng Tokenizer 변경 #43 opened by iamnamki enhancement

29 Done

- [fix] db field update #68 opened by jjung0121
- [fix] index 중복코드 수정 #64 opened by jjung0121
- tokenizer pickle load path 변경 #67 opened by iamnamki enhancement
- [fix] model index field명 수정 및 UI 1자 완성 #57 opened by jjung0121
- [fix] textToSpeech section update #55 opened by jjung0121
- [fix] audiobook 썸네일 업데이트 #48 opened by bhy304 enhancement
- [fix] audiostore / textToVoice 색션 로직 update #46 opened by jjung0121
- [feature request] Eng Tokenizer 변경 #40 opened by iamnamki enhancement
- [feature request] Docker version

Management : Issue Tracking

Issue Type

- [Feature Request]
- [Bug Fix]

Issue Status

- Open
- Closed

Integration

- Kanban 과 연동
- Commit과 연계

The screenshot shows the GitHub interface for the repository `brainmining-for-bigdata / maratron`. The top navigation bar includes links for Pull requests, Issues, Marketplace, and Explore. The repository name is displayed with icons for Watch (0), Star (1), and Fork (0). Below the repository name, there are tabs for Code, Issues (4), Pull requests (0), Projects (1), Wiki, Insights, and Settings. A search bar with the text "is:open" is visible, along with filters for Labels (8) and Milestones (1). A green "New issue" button is on the right. Below the search bar, there is a link to "Clear current search query, filters, and sorts". The main content area displays a list of issues with the following details:

| Issue Type | Issue Title | Issue Number | Opened | By | Labels | Assignee |
|-------------------|---|--------------|--------------------|--------------|-------------|----------|
| [bug fix] | docker- media is not loaded in nginx server | #70 | opened 2 days ago | srshin | | |
| [fix] | tokenizer pickle load path 변경 | #66 | opened 3 days ago | iamnamki | enhancement | |
| [feature request] | Google Kubernetes Engine 배포 | #42 | opened 6 days ago | elephamaximu | enhancement | |
| [feature request] | Google Cloud Platform 배포 | #15 | opened 20 days ago | iamnamki | enhancement | |

ProTip! Exclude everything labeled bug with `-label:bug`.

Management : Issue Tracking

+

↓

↑

↺

↻

↺↻

↺↻

↺↻

커밋

물

무시

폐지

브랜치

병합

상태시

🌐

📄

🔍

원격 보기

파인더에서 보기

터미널

📁 워크스페이스

📄 파일 상태

📜 히스토리

🔍 검색

🔗 브랜치

📌 태그

📁 원격

📁 > origin

📁 치워두기

📁 부모물

📁 SUBTREES

모든 브랜치

원격 브랜치 표시

원형 순

그래프

설명

↗️ origin/SangrimShin#10

python eval.py --checkpoint ./california-12-logs/model.ckpt-112000 동작안하는 현상 해결

↗️ Merge pull request #17 from brainmining-for-bigdata/inchangJUNG#16

↗️ origin/inchangJUNG#16

recognition 모듈 google.py 추가

Update README.md

↗️ Merge pull request #14 from brainmining-for-bigdata/SangrimShin#12_모델선택기능구현

↗️ origin/SangrimShin#12_모델선택기능구현

선택된 목소리에 따른 모델 선택 기능 구현

↗️ Merge pull request #9 from brainmining-for-bigdata/JungeunPark#6

↗️ origin/JungeunPark#6

목소리 선택 기능 추가 생성

↗️ Merge branch 'master' into JungeunPark#6

↗️ Merge pull request #8 from brainmining-for-bigdata/inchangJUNG#7

↗️ origin/inchangJUNG#7

오디오 모듈 추가 및 유틸 모듈 수정

↗️ Merge pull request #5 from brainmining-for-bigdata/HayeonBaek#4

AudioStore 업데이트 및 기본 템플릿 재구현

↗️ Merge pull request #3 from brainmining-for-bigdata/SangrimShin#2_dataset

↗️ origin/SangrimShin#2_dataset

dataset folder추가

↗️ Merge pull request #1 from brainmining-for-bigdata/srshin_tacotron_merge

↗️ origin/srshin_tacotron_merge

tacotron code integration

파일 정리 및 기본 템플릿 재구현!

Django 기본 로직 구현 완료

커밋

작성자

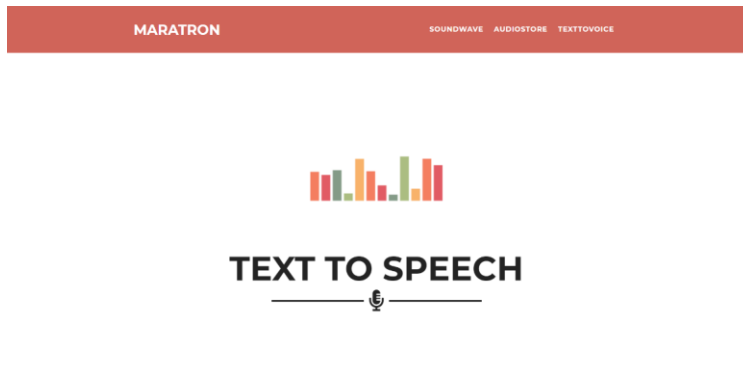
날짜

| | | |
|---------|---------------------|--------------------|
| a52e8ac | srshin <sangrims... | 2019. 4. 30. 오전... |
| 23b11bb | 신상림 <sangrims... | 2019. 5. 2. 오전... |
| a14b127 | 정인창 <elephama... | 2019. 4. 30. 오전... |
| e398b2e | 신상림 <sangrims... | 2019. 4. 30. 오후... |
| 0b8769e | 신상림 <sangrims... | 2019. 4. 30. 오전... |
| 6dc0c01 | srshin <sangrims... | 2019. 4. 29. 오후... |
| bffc4b0 | jung0121 <4566... | 2019. 4. 26. 오후... |
| e5abec2 | jung0121 <4566... | 2019. 4. 26. 오후... |
| 10cdc50 | jung0121 <4566... | 2019. 4. 26. 오후... |
| f1a29ea | elephamaximu <... | 2019. 4. 25. 오전... |
| 58bb8d7 | 정인창 <elephama... | 2019. 4. 25. 오전... |
| 0467d2b | HayeonBaek <43... | 2019. 4. 24. 오후... |
| c41a379 | bhy304 <bhy051... | 2019. 4. 24. 오후... |
| 73b6cdf | 신상림 <sangrims... | 2019. 4. 24. 오후... |
| 083b1ac | srshin <sangrims... | 2019. 4. 24. 오후... |
| 04d5964 | 신상림 <sangrims... | 2019. 4. 24. 오전... |
| c967df2 | srshin <sangrims... | 2019. 4. 24. 오전... |
| f2d9344 | bhy304 <bhy051... | 2019. 4. 24. 오전... |
| 3aa1fa6 | bhy304 <bhy051... | 2019. 4. 23. 오후... |

Service

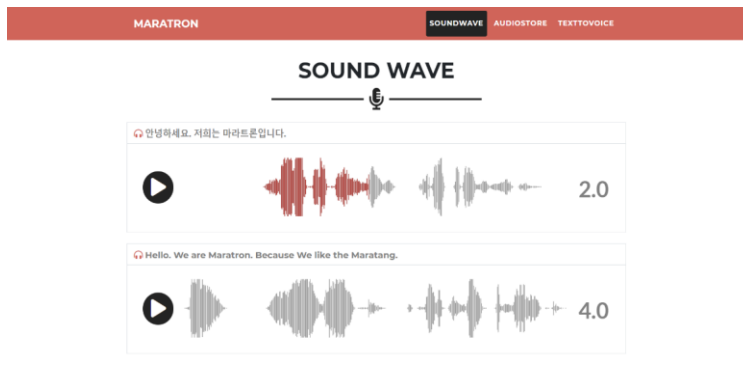


Implement



MARATRON

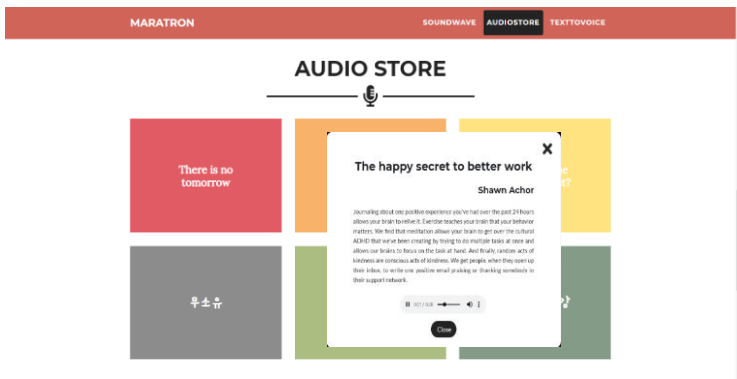
AUDIO STORE , SOUND WAVE , TEXT TO VOICE 메뉴 선택 가능.



SOUND WAVE

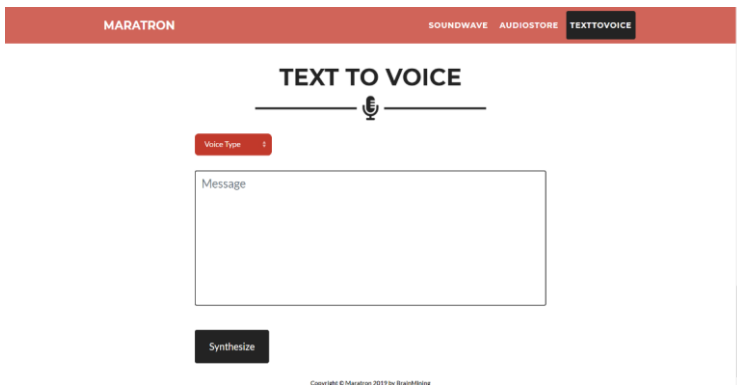
Training되어 있는 목소리의 샘플 오디오를 통해 목소리를 확인할 수 있다.

Implement



AUDIO STORE

미리 등록되어 있는 오디오 콘텐츠를 들을 수 있다.



TEXT TO VOICE

Voice Type을 선택 - 읽어주길 원하는 Text 입력 - Synthesize - 오디오제공

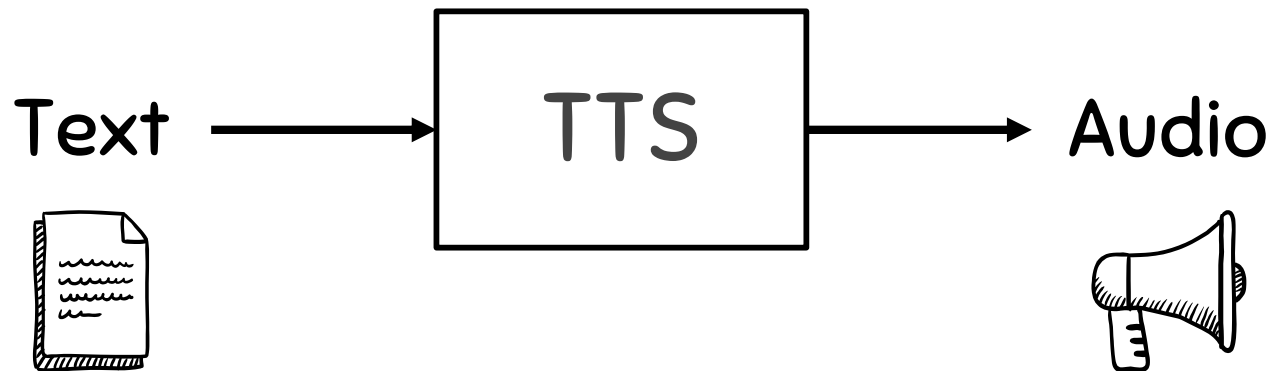
Model



Model : Dataset

| Voice Type | English_Female | English_Male | Korean_Female |
|-----------------------|----------------|--------------|---------------|
| Total Size | 2.6 GB | 567 MB (zip) | 3 GB |
| Total Length | 24 hours | 20 hours | 12 hours |
| Number of audio-clips | 13,100 | 1 audio file | 12,853 |
| Training Check-Point | 40,000 | 112,000 | 64,000 |

TEXT-TO-SPEECH : Concept



TACOTRON: TOWARDS END-TO-END SPEECH SYNTHESIS

Google paper (published in April 2017)

TACOTRON: TOWARDS END-TO-END SPEECH SYNTHESIS

Yuxuan Wang*, RJ Skerry-Ryan*, Daisy Stanton, Yonghui Wu, Ron J. Weiss[†], Navdeep Jaitly,

Zongheng Yang, Ying Xiao*, Zhifeng Chen, Samy Bengio[†], Quoc Le, Yannis Agiomyriannakis,

Rob Clark, Rif A. Saurous*

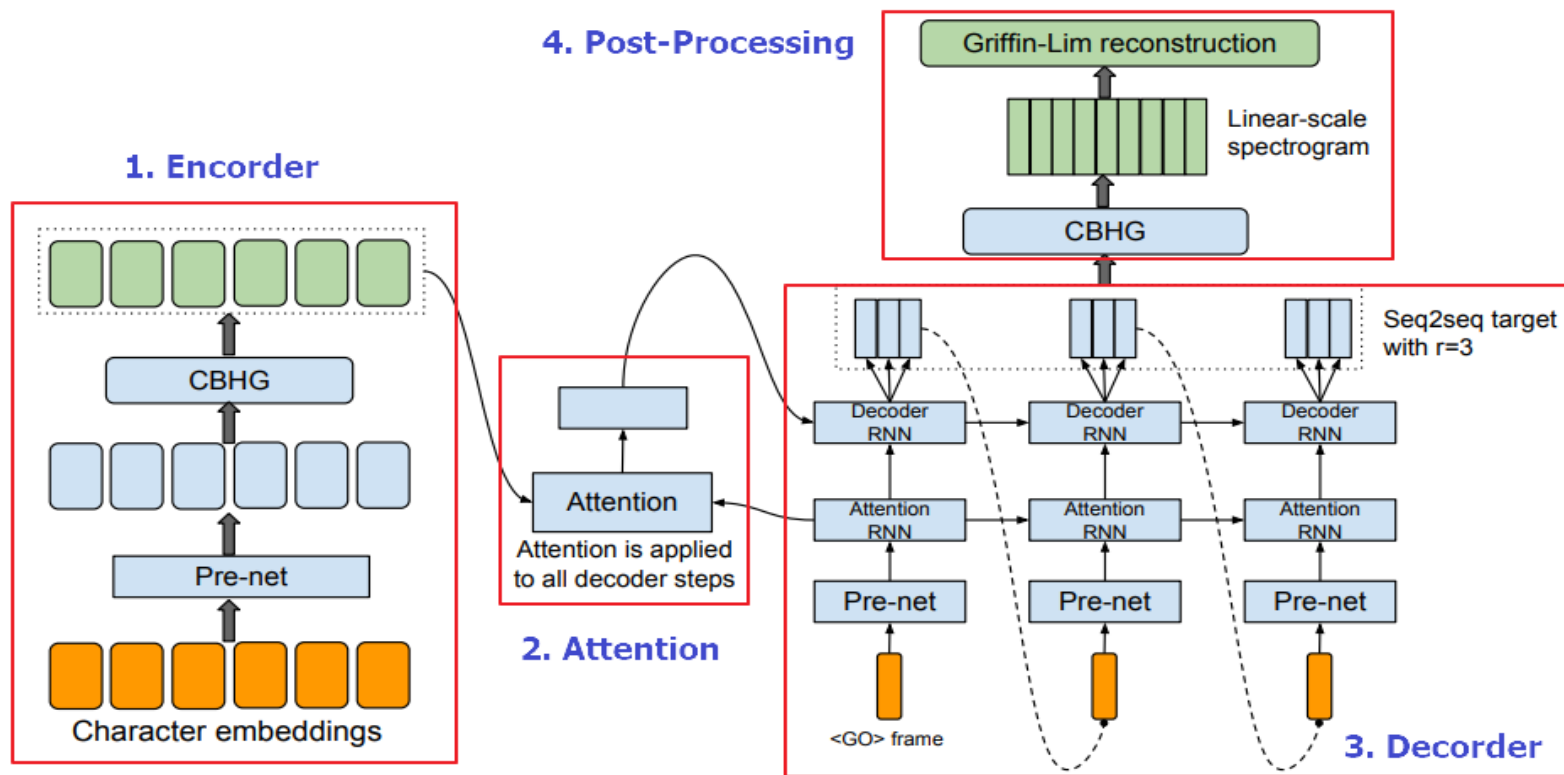
Google, Inc.

{yxwang, rjryan, rif}@google.com

ABSTRACT

A text-to-speech synthesis system typically consists of multiple stages, such as a text analysis frontend, an acoustic model and an audio synthesis module. Building these components often requires extensive domain expertise and may contain

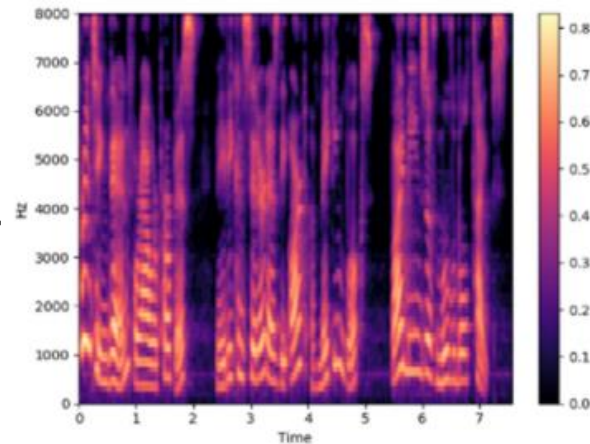
Model : Tacotron



Model : Training Model

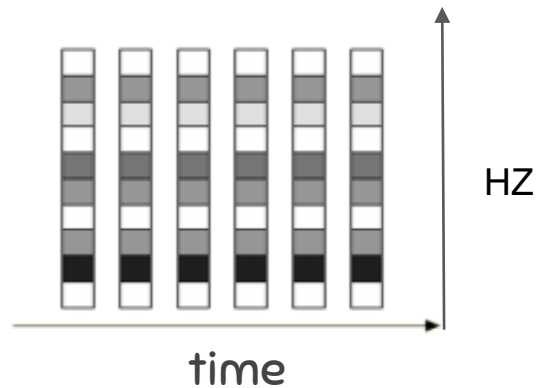
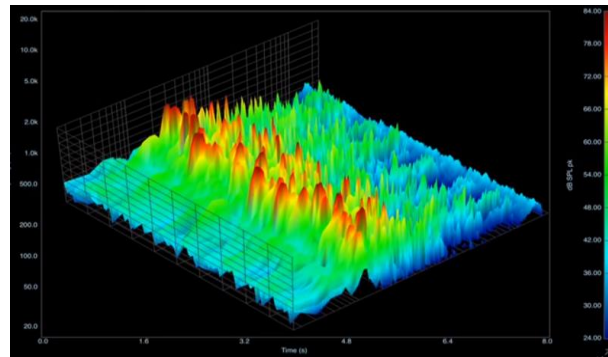
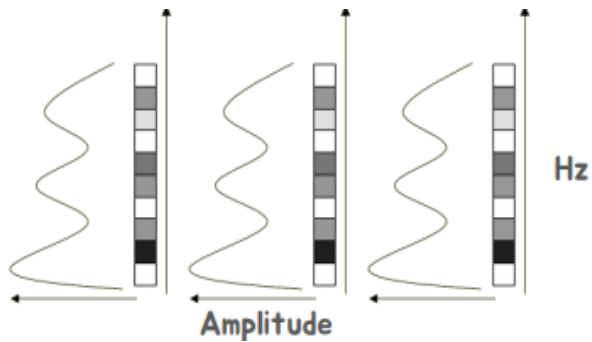
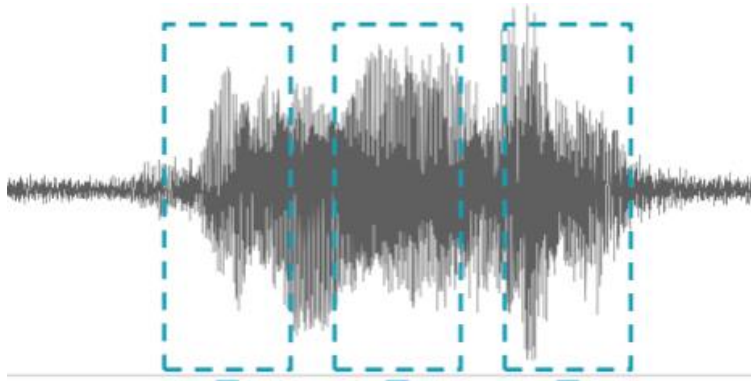
Over the MIRACLE

<Text>

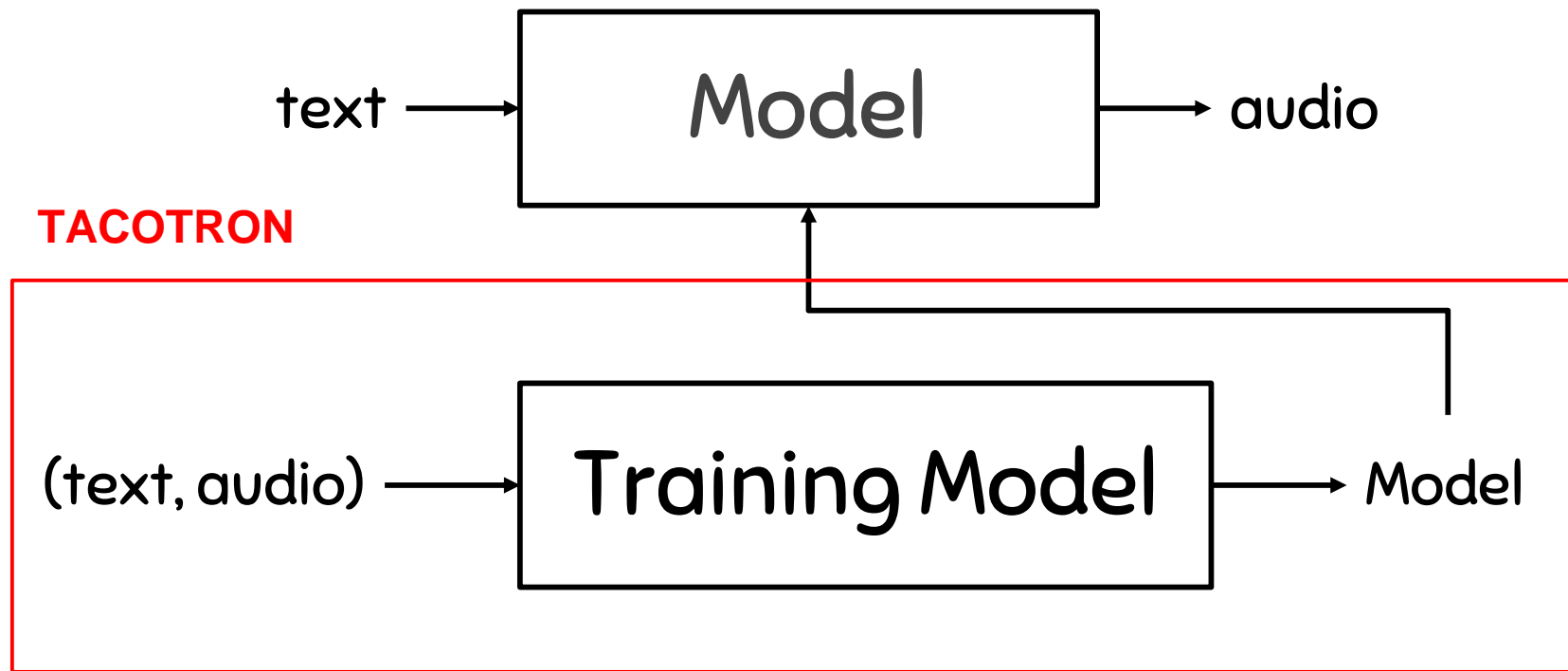


<spectrogram>

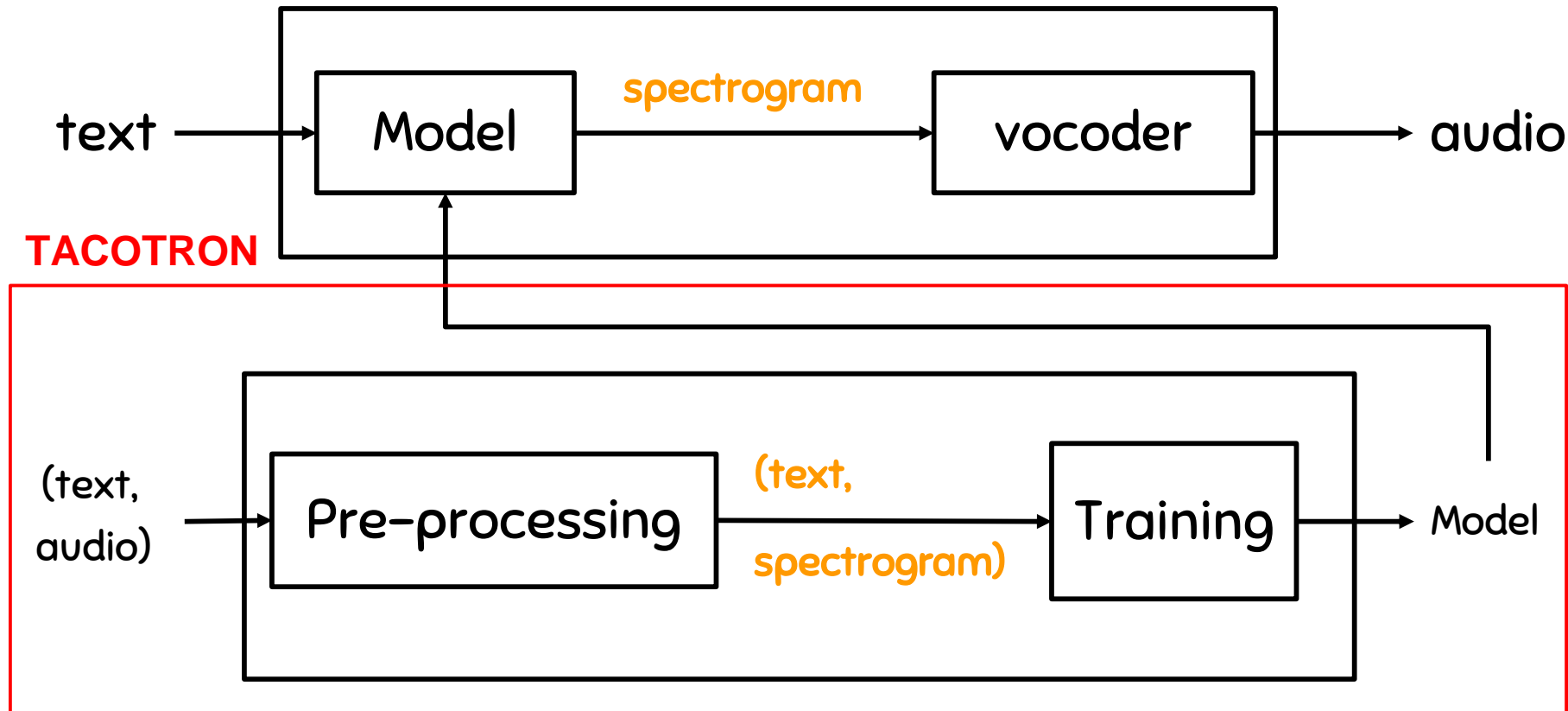
Model : Spectrogram - 소리를 형상화



Model : structure

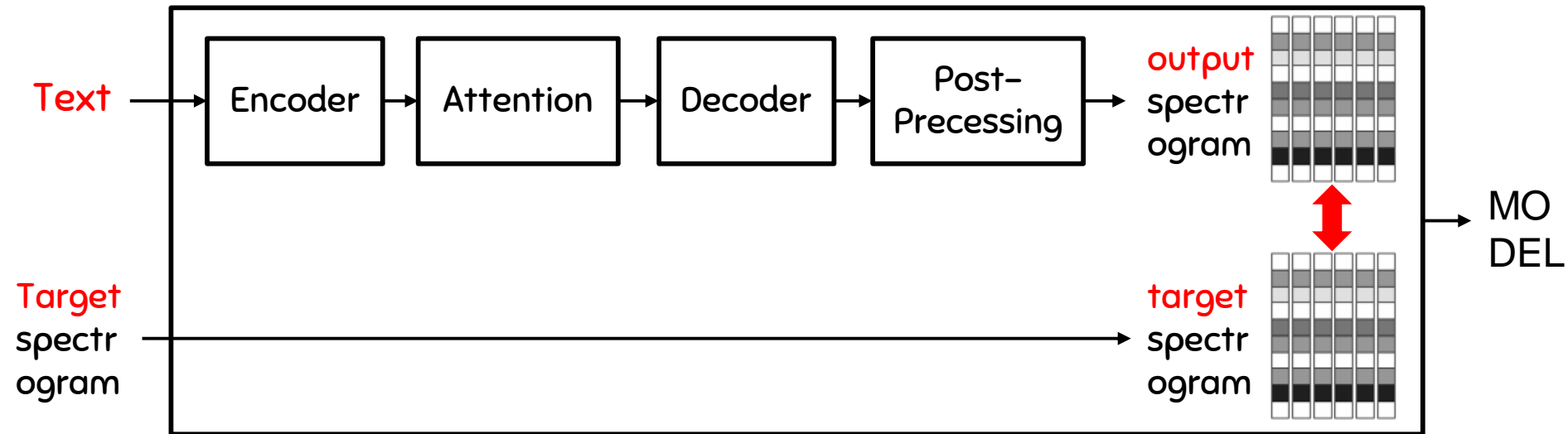


Model : structure

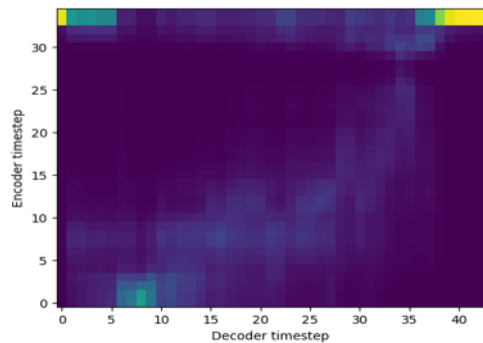


Model : Tacotron

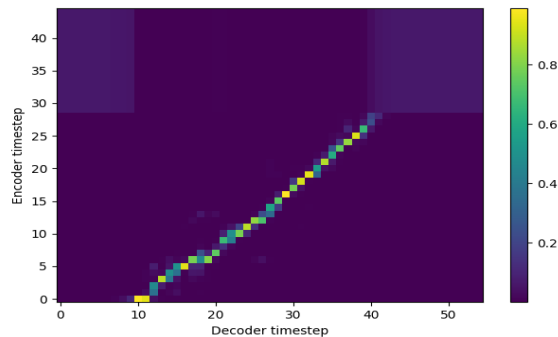
TACOTRON



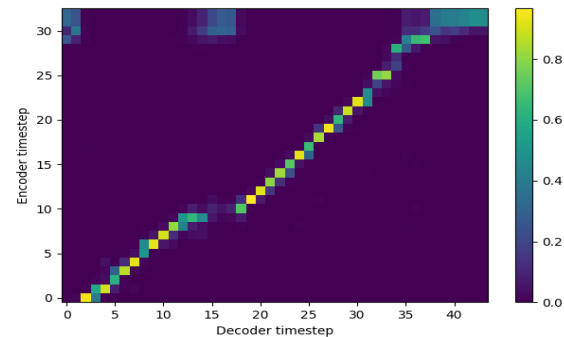
Training Model : linear Spectrogram



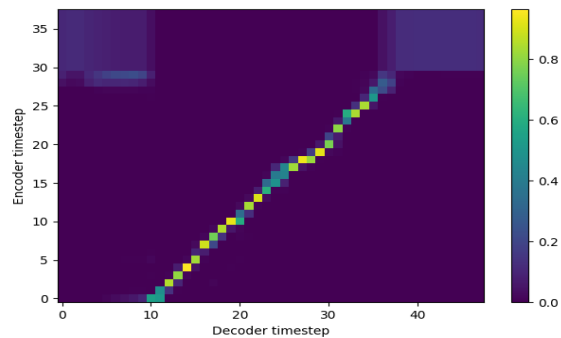
tacotron, None, 2019-05-03 18:51, step=5000, loss=0.12108



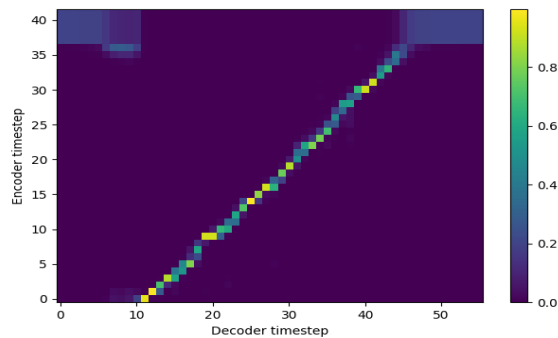
tacotron, None, 2019-05-04 01:44, step=10000, loss=0.10240



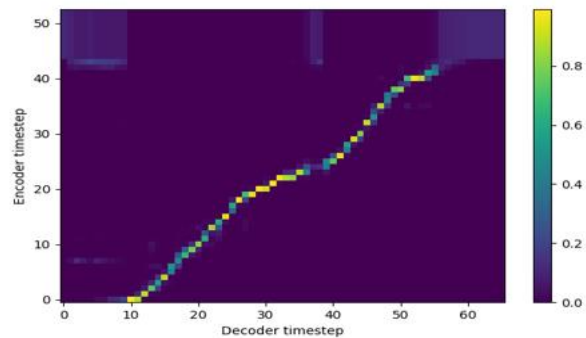
tacotron, None, 2019-05-05 05:49, step=30000, loss=0.09098



tacotron, None, 2019-05-05 20:09, step=40000, loss=0.08470



tacotron, None, 2019-05-06 10:42, step=50000, loss=0.08184

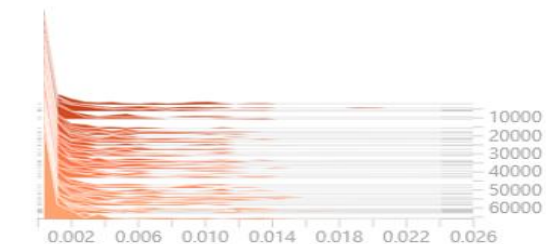


tacotron, None, 2019-05-07 09:17, step=65000, loss=0.08697

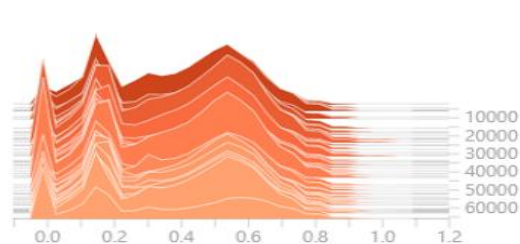
Tensorboard

model

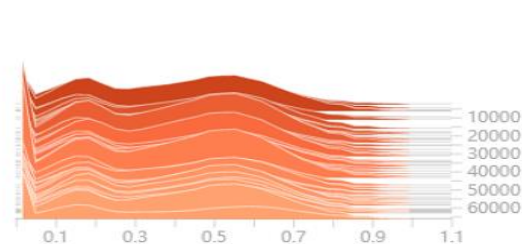
model/stats/gradient_norm



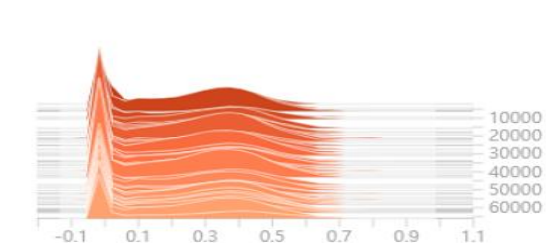
model/stats/linear_outputs



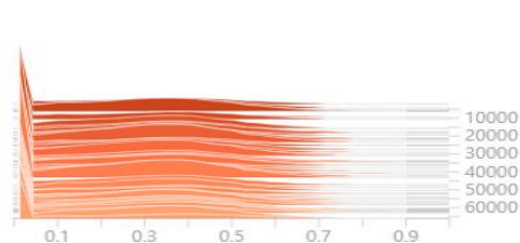
model/stats/linear_targets



model/stats/mel_outputs



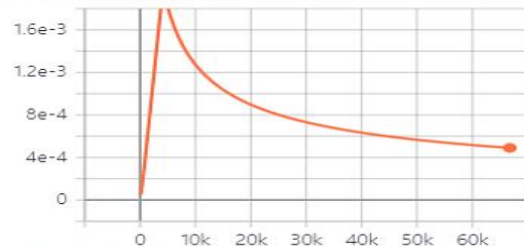
model/stats/mel_targets



Tensorboard

model

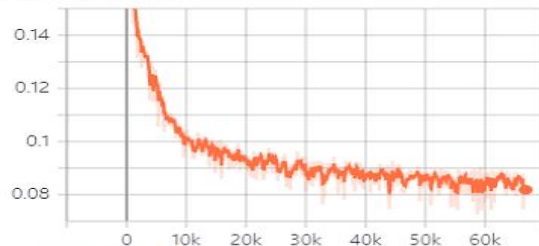
stats/learning_rate
tag: model/stats/learning_rate



run to download

CSV JSON

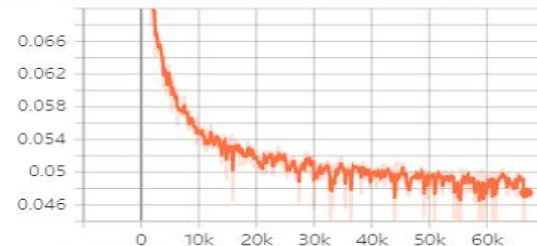
stats/loss
tag: model/stats/loss



run to download

CSV JSON

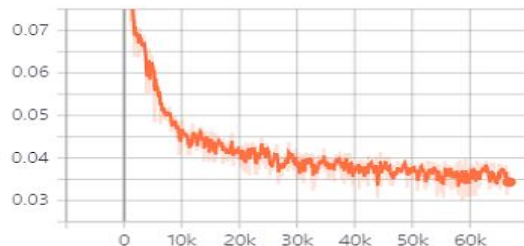
stats/loss_linear
tag: model/stats/loss_linear



run to download

CSV JSON

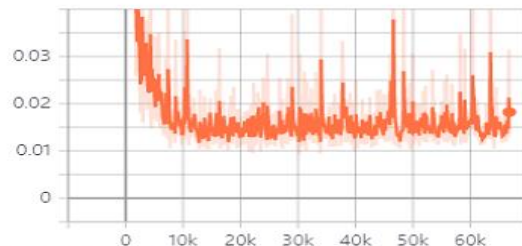
stats/loss_mel
tag: model/stats/loss_mel



run to download

CSV JSON

stats/max_gradient_norm
tag: model/stats/max_gradient_norm



run to download

CSV JSON

감사합니다.