

GiTi-4



김연수(팀장)
신유진
김석영
이태훈

목차

1. 팀원 소개
2. 프로젝트 소개
3. 결과물
4. 아키텍처
5. 구현 & 한계점
6. QnA

팀원 소개



김연수(팀장)

- 팀장업무
- TTS 구현
- Voice Cloning
모델 구현



신유진

- 욕설탐지모델 구현
- XAI 구현



김석영

- 영문 탐지 프롬프트
엔지니어링
- 영문 변환 프롬프트
엔지니어링



이태훈

- GCP 환경 구성
- Streamlit 구현
- STT 구현
- 욕설 빼소리 기능 구현
- 욕설 리스트를 활용한 욕설 탐지

프로젝트 소개

선정배경, 목적&필요성

프로젝트 소개

Task 1: clean contents 제작

선정배경

유튜브 및 개인 인터넷 방송에서 수위 조절 변경으로 인해 직접적인 욕설 및 폭언이 노출되고 있습니다.
이로 인해 아이들이 불쾌한 정서적 경험을 할 수 있습니다.

목적 & 필요성

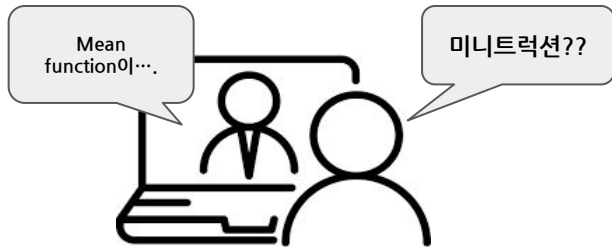
시청자들에게는 비폭력적인 장면만을 제공하며,
영상 제작자들은 의도치 않은 폭력적인 장면을 피해 불쾌한 상황을 방지합니다.

프로젝트 소개

Task 2: Enhancing YouTube technical jargon

선정배경

유튜브 자동 자막 사용시, 영어표현 및 문장이 잘못 번역되는 불편함이 있습니다.

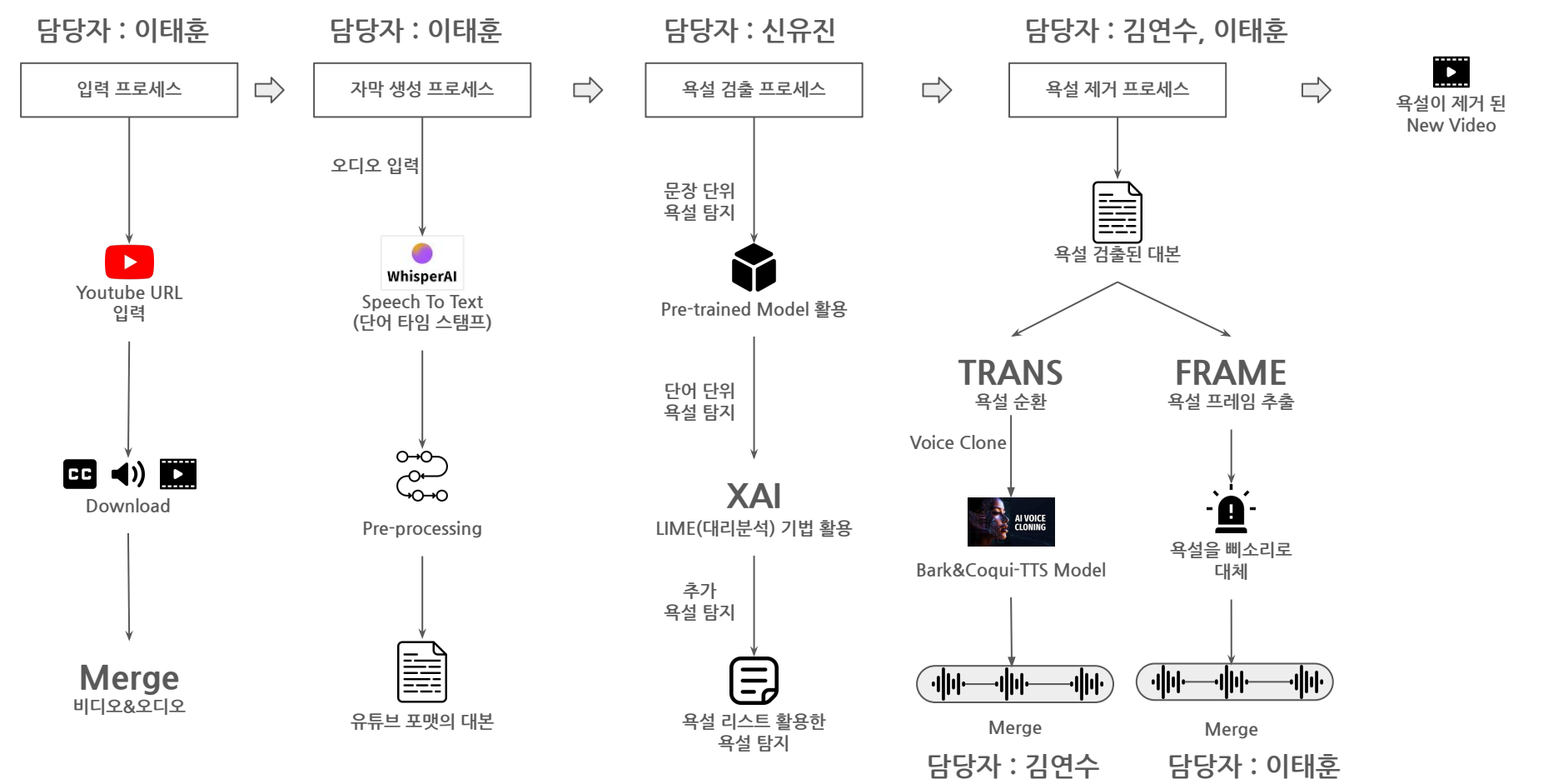


목적 & 필요성

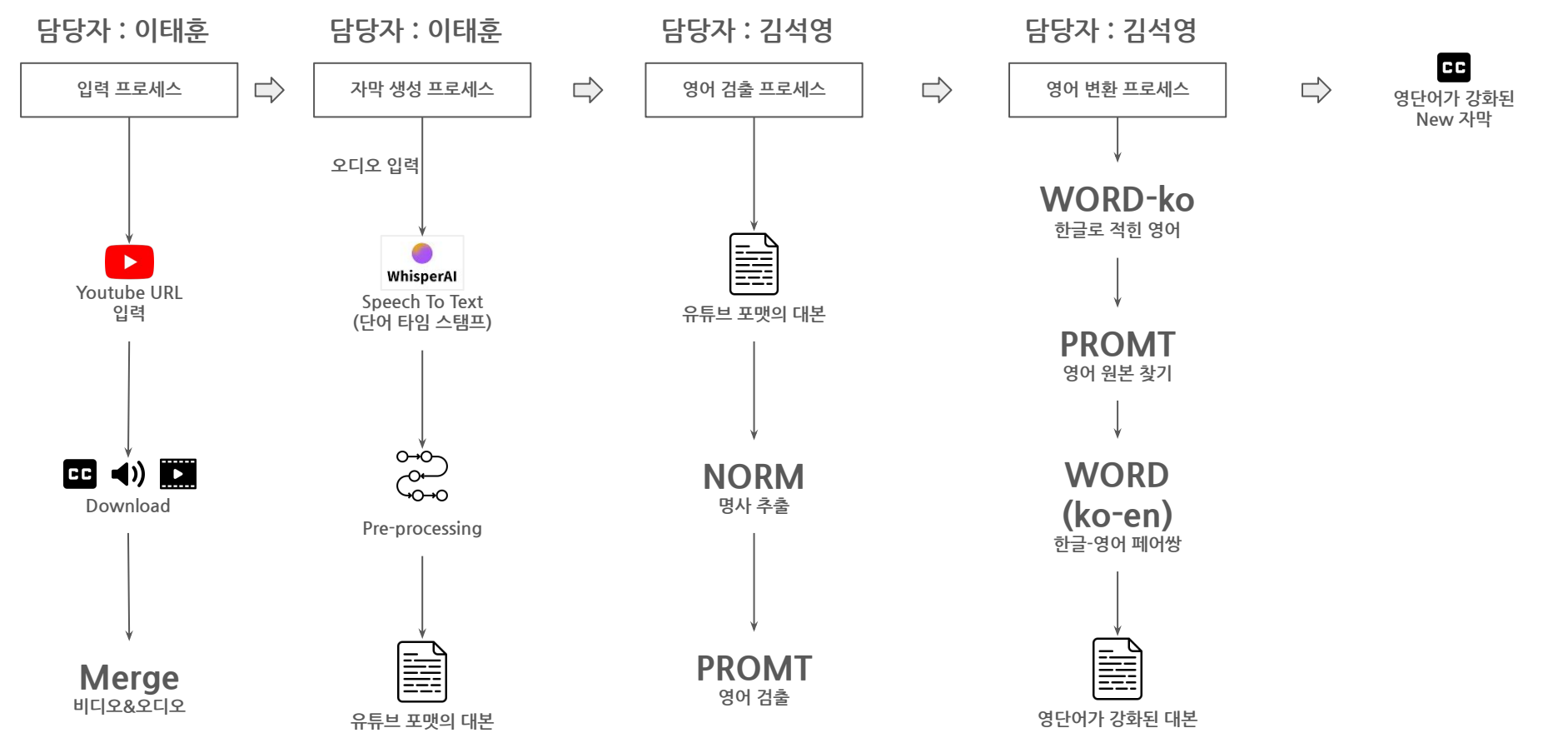
잘못된 영어표현 감지 및 교정을 통해 학습에 도움을 줍니다.

아키텍처



아키텍처 1



아키텍처 2



결과물

 ※주의) 영상에 욕설이 포함되어 있습니다. 

Youtube

Scripte-Enhancer & Voice-Cloner

Sample

- ▾

Youtube URL

<https://www.youtube.com/~>

Play

<http://34.30.79.104:8501/> <- 외부 접속 불가능(GCP 반납으로 접속 불가)

결과물 링크 공유

-

보이스피싱(1분)

더글로리(40초)

멜로가체질(55초)

카지노(37초)

샘플 중 ‘더글로리’, ‘멜로가체질’, ‘카지노’을 선택하시면 결과물을 바로 보실 수 있습니다

결과물

1. 유튜브 기본 제공 Script(자막)보다 성능 향상

자막만으로는 의도 파악 어려움

*Youtube 자막

[0:07] 나이 어린거
[0:10] 언제까지 어려 내년에도 어려
[0:12] 훈연이도 어릴 거니 이제 아무리

훈연이? 흐름 파악이 어려운 단어가 있음

*Ours 자막

[0:08] 너 나이 어린거?
[0:09] 너 뭐 언제까지 어려? 내년에도 어려?
[0:12] 후년에도 어릴거니?

자막 성능 향상

*더글로리 중 한 장면

결과물

1. 유튜브 기본 제공 Script(자막)보다 성능 향상



https://youtube.com/shorts/bGFkdfYPPAo?si=xwQB3GSjVs_zuD9T

결과물

1. 유튜브 기본 제공 Script(자막)보다 성능 향상

*Youtube 자막

[0:02] 나 여기서 결하고 내 형이야 엄마가
[0:06] 달라
[0:07] **엄마 아버지가 얘기 안 했을 거야**
[0:09] 아버지 성격 알잖아

자막만으로 보면
'엄마와 대화하는 장면'으로 보임

*Ours 자막

[0:00] 어, 니가 무식이냐?
[0:02] 나 영석이라고, 니 형이야.
[0:04] 어, 엄마가 달라.
[0:07] **아마 아버지가 얘기 안 했을 거야.**

실제론
'엄마와 대화하는 장면'이 아님

*카지노 중 한 장면

결과물

1. 유튜브 기본 제공 Script(자막)보다 성능 향상



<https://youtube.com/shorts/5-7NPzsgAag?si=DFZLVKEDzA7AkihE>

1. 유튜브 기본 제공 Script(자막)보다 성능 향상

*Youtube 자막

| | |
|--------|----------------|
| [0:20] | 걸리적거리는 거야 으 |
| [0:24] | tuck 축축하게 |
| [0:26] | 으 |
| [0:29] | 기 0 |
| [0:32] | 니가 멋지 가나 |
| [0:34] | 음 으 |
| [0:36] | bc 9 900,000 5 |
| [0:39] | pocket 마 |
| [0:42] | 으 |
| [0:44] | 으 |
| [0:47] | 으 |
| [0:50] | 와 어 유기 el |
| [0:54] | 내가 이겼다 으 |
| [0:55] | [음악] |

자막만으로 보면
'랩하는 장면'으로 보임

*Ours 자막

| | |
|--------|--------------------------|
| [0:20] | 니가 걸리적거린 거라고. |
| [0:23] | 어디다 들고 남의 귀한 자식한테 욕짓거리야? |
| [0:26] | 어디다 들고 사람을 괴로워한다고? |
| [0:29] | 사람이야! |
| [0:30] | 귀한 사람이야! |
| [0:32] | 니가 뭔데 지랄이야! |
| [0:36] | 이씨... |
| [0:38] | 너! |
| [0:40] | 이 개새끼... |
| [0:41] | 개새끼 뭐! |
| [0:50] | 뭐! |
| [0:52] | 유빈 |
| [0:54] | 네가 이겼다. |

실제론 '두 사람이 말싸움 하는 장면'

결과물

1. 유튜브 기본 제공 Script(자막)보다 성능 향상



https://youtube.com/shorts/t7aJ1WvVCCI?si=Jwj8kr_NjR-Z0h4N

2. 욕설 검출 및 자막표시 기능

*Youtube 자막

이 자막에는 욕설이 몇 개 있을까요?

스크립트

0:05

예 1주에서 친한 제 친구 제 신상을 담은글 1대 등입니다

0:10

그래서 올해 작구요 신안 대책 임박을 법령 단어입니다 범
박동 아등 품고

0:16

음 근데 이렇게 한국말 잘 보세요

0:19

요새 근데 왜 한국말 잘 못 하나구요 으

0:24

한국말 못하지만 사람이 무슨 멈춘 게 오이 지금 그쪽 행
하고 말하는

0:28

완공이 그제 서식하는데 처진 하긴 한데 한번 말이죠 까치
들이 잤아요

0:33

뒤질래 에어 뭐어 dc 나는 어디

0:36

노인은 뭐야 위에 있지

0:39

228 없 으 10시간 목표 신세 껴야 이 밝혀 그렇다면 이 씹
새끼야

0:45

누가 이 신발 엄 있겠냐 이중 신재 깨지고 저축 저축은행
너 이름 뭐야

0:50

드 이라고 예 미실 타려 잇 아이템이 될 거야

0:54

애야 소개 필요하신 새끼야 어디 시발농아

한국어 (자동 생성됨)

*Youtube 자막

자막 성능 및 양이 많은 문제로
한 눈에 **욕설 찾기 어려움**

스크립트

⋮

✕

0:05 예 1주에서 친한 제 친구 제 신상을 담는글 1대 등입니다

0:10 그래서 올해 작구요 신안 대축 임방을 범랑 단어입니다 범
박동 아등 품고

0:16 음 근데 이렇게 한국말 잘 보세요

0:19 요새 근데 왜 한국말 잘 못 하나구요 으

0:24 한국말 못하지만 사람이 무슨 멈춘 게 오이 지금 그쪽 행
하고 말하는

0:28 완공이 그게 서식하는데 쳐진 하긴 한데 한번 말이죠 까지
들이 잡아요

0:33 뒤질래 예어 뭐어 dc 나는 어디

0:36 노인은 뭐야 위에 있지

0:39 228 업으 10시간 목표 신세 켜야 이 밝혀 그렇다면 이 썰
새끼야

0:45 누가 이 신발 얻 있겠냐 이중 신재 깨지고 저축 저축은행
너 이름 뭐야

0:50 드 이라고 예 미실 타려 잇 아이템이 될 거야

0:54 애야 소개 필요하신 새끼야 어디 시발놓아

한국어 (자동 생성됨)

2. 욕설 검출 및 자막표시 기능

*Our 자막

욕설부분에 하이라이트(형광펜) 효과 기능

[0:00] 여보세요.
[0:02] 안녕하세요. 심규환 검사님 아세요?
[0:05] 네.
[0:07] 1일 저 신한제축은행 심사부 담당자 한다성입니다.
[0:10] 뭐라고? 누구야?
[0:12] 신한제축은행 심사부 담당자 한다성입니다.
[0:15] 잠깐 통화 괜찮으세요?
[0:17] 근데 왜 이렇게 한국말 잘 못해요?
[0:19] 여보세요.
[0:21] 근데 왜 한국말 잘 못하냐고요?
[0:24] 한국말 못하시다는 말은 무슨 말인지요?
[0:26] 그게 표준어긴 한데 한국말이 쯔들리잖아요.
[0:33] 죽을래요? 어디? 신한은행 어디?
[0:36] 너 이름 뭐야?
[0:38] 미애들 개새키야.
[0:40] 이 사보놈아.
[0:42] 미친 새끼야.
[0:44] 니 말투 그렇게 하면 누가 사보놈아 믿겠냐.
[0:48] 저축은행? 너 이름 뭐야?
[0:50] 니 애미다.
[0:51] 니 애미 사보놈이다.
[0:53] 개새키야.
[0:54] 니 애미야.
[0:55] 어디 사보놈아.

결과물

3. 영상 내 욕설이 있는 프레임 뺄소리로 대체 기능



처음 보여준 영상으로 대체

결과물

*우리 모델은 왜 Youtube보다 우수한가?

최신 모델 Whisper을 사용
(현재 Large-v2 사용, 최근 Large-v3도 출시)

구현 & 한계점

Streamlit 구현



1. Youtube URL을 이용한 유튜브 영상, 오디오, 자막 다운로드 기능
2. 다운로드한 영상, 오디오 Merge 기능
3. 샘플 영상 제공 기능
4. 욕설 하이라이트(형광펜) 기능
5. 자막에서 욕설 감지 기능
6. 영상에서 욕설있는 프레임 감지 기능
7. 영상에서 욕설을 삐소리 대체 기능
8. 외부 접속 가능

욕설감지 파트

딥러닝을 활용한 욕설 검출 프로세스 개발 및 XAI로 욕설구분에 미치는 결정적인 단어 구분

beomi/KcBERT-Base(Pre-trained Model) 활용

훈련된 데이터: **Smilegate Ai**에서 공개하는 혐오표현 “🙄 UnSmile” 이라는 자체적으로 수집한 각종 커뮤니티의 댓글 데이터셋

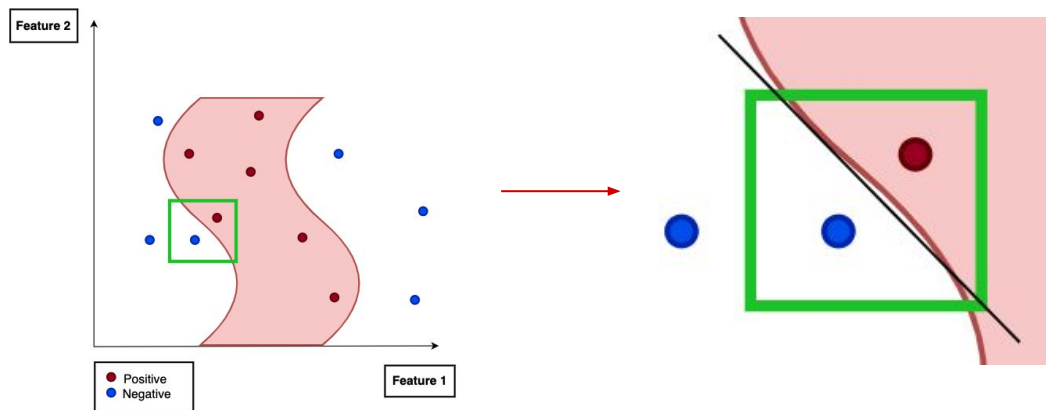


데이터셋 규모:

| Dataset | train | test |
|----------|---|------|
| features | text, label (0: 혐오표현, 1: 비혐오표현) | |
| num_rows | 15005 | 3737 |

Lime 패키지 - Explainable AI (XAI)

Local Interpretable Model-agnostic Explanation



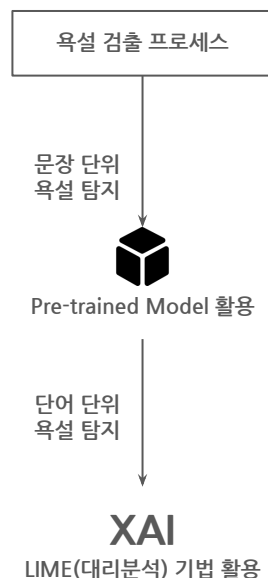
이 결정경계 가운데에 지역적인 선형회귀 모델(검정색 선)을 적합해서,
이 회귀식의 계수를 사용해 모델의 판단 메커니즘을 분석하는 것이 바로 LIME의 아이디어라고 할 수 있습니다.

원본

[0:00] 이 쌍년이
[0:02] 쌍년짓을 할려면 제대로 해
[0:03] 그래야 쌍년 안되고 난년 되는거야
[0:06] 씨발
[0:08] 너 나이 어린거?
[0:09] 너 뭐 언제까지 어려? 내년에도 어려?
[0:12] 후년에도 어릴거니?
[0:14] 잊어보지 않겠어 아악
[0:15] 아무리 콜라겐을 쳐먹고 쳐바르고 용을 써
도 내 자린 어림도 없단 뜻이란다
[0:22] 이 어리기만한 년아
[0:23] 그래서 뭐? 뭔데?
[0:26] 원하는게 뭐냐고 시발 년아
[0:28] 야 이 미친 년아
[0:29] 이 시발 년아
[0:30] 이 미친 년이
[0:31] 이 좇밥이단 년이
[0:33] 대가리에 국영수를 채웠다고 아주 시발 년
이 됐네
[0:36] 먼저 일어날게
[0:38] 시발 년이 다음 일정이 있어서
[0:40] 들았네 이 시발 년이

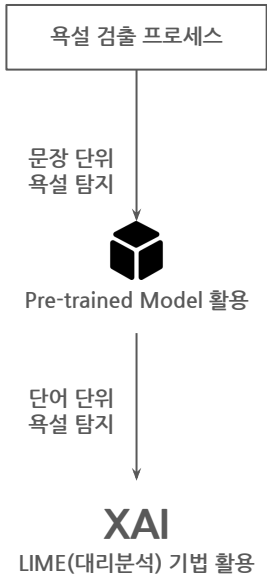
원본

[0:00] 이 쌍년이
[0:02] 쌍년짓을 할려면 제대로 해
[0:03] 그래야 쌍년 안되고 난년 되는거야
[0:06] 씨발
[0:08] 너 나이 어린거?
[0:09] 너 뭐 언제까지 어려? 내년에도 어려?
[0:12] 후년에도 어릴거니?
[0:14] 잊어보지 않겠어 아악
[0:15] 아무리 콜라겐을 쳐먹고 쳐바르고 용을 써
도 내 자린 어림도 없단 뜻이란다
[0:22] 이 어리기만한 년아
[0:23] 그래서 뭐? 뭔데?
[0:26] 원하는게 뭐냐고 시발 년아
[0:28] 야 이 미친 년아
[0:29] 이 시발 년아
[0:30] 이 미친 년이
[0:31] 이 좇밥이단 년이
[0:33] 대가리에 국영수를 채웠다고 아주 시발 년
이 됐네
[0:36] 먼저 일어날게
[0:38] 시발 년이 다음 일정이 있어서
[0:40] 들았네 이 시발 년이



원본

[0:00] 이 쌍년이
[0:02] 쌍년짓을 할려면 제대로 해
[0:03] 그래야 쌍년 안되고 난년 되는거야
[0:06] 씨발
[0:08] 너 나이 어린거?
[0:09] 너 뭐 언제까지 어려? 내년에도 어려?
[0:12] 후년에도 어릴거니?
[0:14] 잊어보지 않겠어 아악
[0:15] 아무리 콜라겐을 쳐먹고 쳐바르고 용을 써
도 내 자린 어림도 없단 뜻이란다
[0:22] 이 어리기만한 년아
[0:23] 그래서 뭐? 뭔데?
[0:26] 원하는게 뭐냐고 시발 년아
[0:28] 야 이 미친 년아
[0:29] 이 시발 년아
[0:30] 이 미친 년이
[0:31] 이 좆밥이단 년이
[0:33] 대가리에 국영수를 채웠다고 아주 시발 년
이 됐네
[0:36] 먼저 일어날게
[0:38] 시발 년이 다음 일정이 있어서
[0:40] 돌아왔네 이 시발 년이



Text with highlighted words

[0:00] 이 쌍년이
[0:02] 쌍년짓을 할려면 제대로 해
[0:03] 그래야 쌍년 안되고 난년 되는거야
[0:06] 씨발
[0:08] 너 나이 어린거?
[0:09] 너 뭐 언제까지 어려? 내년에도 어려?
[0:12] 후년에도 어릴거니?
[0:14] 잊어보지 않겠어 아악
[0:15] 아무리 콜라겐을 쳐먹고 쳐바르고 용을 써
도 내 자린 어림도 없단 뜻이란다
[0:22] 이 어리기만한 년아
[0:23] 그래서 뭐? 뭔데?
[0:26] 원하는게 뭐냐고 시발 년아
[0:28] 야 이 미친 년아
[0:29] 이 시발 년아
[0:30] 이 미친 년이
[0:31] 이 좆밥이단 년이
[0:33] 대가리에 국영수를 채웠다고 아주 시발 년
이 됐네
[0:36] 먼저 일어날게
[0:38] 시발 년이 다음 일정이 있어서
[0:40] 돌아왔네 이 시발 년이

Text with highlighted words

못잡아낸 단어

긍정으로 판단된 무의미한 단어

욕설에 가깝다고 인식한 단어

[0:00] 이 **쌍년**이
[0:02] **쌍년**이를 할려면 제대로 해
[0:03] 그래야 **쌍년** 안되고 **난년** 되는거야
[0:06] **씨발**
[0:08] 너 나이 어린거?
[0:09] 너 뭐 언제까지 **어려?** 내 **년**에도 **어려?**
[0:12] **후년**에도 어릴거니?
[0:14] 잊어보지 않겠어 아악
[0:15] 아무리 콜라겐을 쳐먹고 쳐바르고 용을 써
도 내 자린 어림도 없단 뜻이란다
[0:22] 이 어리기만한 **년**아
[0:23] 그래서 뭐? 뭔데?
[0:26] 원하는게 뭐냐고 **시발** **년**아
[0:28] 야 이 **미친** **년**아
[0:29] 이 **시발** **년**아
[0:30] 이 **미친** **년**이
[0:31] 이 **죽밥**이단 **년**이
[0:33] 대가리에 국영수를 채웠다고 아주 **시발** **년**
이 됐네
[0:36] 먼저 일어날게
[0:38] **시발** **년**이 다음 일정이 있어서
[0:40] 들었네 이 **시발** **년**이

*결론 : XAI의 성능 개선 필요

- 개선 방법
 - 욕설 리스트를 활용한 추가 필터링
 - 데이터셋 재구축

TTS & Voice Cloning



감정을 표현하는 AI 보이스

Voice Studio는 오디오북, 영화 해설을 포함해 다양한 분야에 활용할 수 있는 목소리를 제공합니다.

캐릭터 걸스 키즈 외국인



줄리
30대 느릿한 매력적인
· 우아한



카이지
20대 활기찬 진지한 정직한



순저 지나
30대 단호한 열정적인
· 진지한



상구
20대 개진 당당한 열정적인



Text



TTS & Voice Cloning

➡  **Bark** &  **Coqui-TTS**  [Bark - TTS 0.20.5 documentation](#)

<모델 선정 기준>

✓ 최신 모델:  2023.04.20 released&updated,  since 2021

✓ 한국어로 훈련된 모델이어야 함. 

✓ TTS + Voice Cloning  + 

? 성능이 좋지 않으면 수요 및 활용도가 떨어질 것 -> 후처리 고려

TTS & Voice Cloning

실험 변수 및 결과

- noise ↓
- sentence length ↓
- Ton →
- Language(English)

<raw.wav>



<cloned.wav>



영어검출 및 변환 파트

목표

Before

하지만 이렇게 할 경우 이전 값을 계속해서 이어서 쓰게 되니
백 프라퍼게이션 과정에서 배니싱 그레이디언트 문제가 발생하게 됩니다.



After

하지만 이렇게 할 경우 이전 값을 계속해서 이어서 쓰게 되니
백 프라퍼게이션(back propagation) 과정에서 배니싱 그레이디언트(Vanishing Gradient) 문제가 발생하게 됩니다.

영발음으로 적혀진 자막은 종종 원본 단어를 혼돈 가능성이 있다.

영어검출 및 변환 파트

입력 데이터

그래서 보죠. 그러면 gp라는 게 뭔지 알았으니까 이제 gp를 가지고 리그레션 어떻게 될까 아웃션 프로세스라는 게 뭔지 아는 것과 gp랑 gpr은 달라요. gp는 가우셰어 프로세스 자체로 gpr은 아우셰어 프로세스를 이용해서 리그레션 문제 클핑스 로그인에서 인프 아웃풋이 주어졌을 때 새로운 입력에 대한 프레이트 아웃풋을 계산하는 문제를 푸는 걸 의미합니다. 그 자리에서 우리 어떤 꺼는 코드를 찾게 되면 내가 찾고자 하는 출력, 어떤 함수, 형태 프로모션이 이런 프라이어를 만든다는 걸 의미하고 이거를 우리가 앞에서 시피 프라이어로 프라우드를 했었어요. 그랬는데 프로젝팅을 어떻게 할 수 있냐면 어떤 컨디셔널 디스플레이션을 표현할 수 있는 거예요. 사실 내가 x랑 y라는 이스터 아웃이 좋아졌을 때 새로운 입력 x 스타를 플레인 데이터의 컨디션한 컨디션드 gvc를 표현할 수 있는 거죠.

프롬프트

위 내용은 STT 기술을 이용해 음성을 텍스트로 변환한 거야.
그런데 한국어 문장에 영어 용어가 섞여 있다보니 발음 등의 문제로 영어 용어가 제대로 변환이 되지 않은 것들이 있어.

위 내용에 대한 지식과 아래 관련 키워드들을 참고해서
위 내용에서 부정확하게 변환이 된 부분에 올바른 용어를 괄호로 추가해
위 문장/문단을 수정해줘.

관련 단어 추가

- gaussian process
- regression
- continuous domain
- predicted output
- current function
- prior
- conditional distribution
- input
- ouput

영어검출 및 변환 파트

생성 내용

그래서 보조, 그러면 가우시안 프로세스(Gaussian Process)라는 게 뭔지 알았으니까 이제 가우시안 프로세스를 가지고 회귀(Regression) 어떻게 될까요? 아웃 프로세스라는 것은 뭘까요? 아우션 프로세스를 이용해서 리그레션 문제를 클리핑스 로그틱(클릭스 로지스틱)에서 인풋 아웃풋이 주어졌을 때 새로운 입력에 대한 프레딕티드 아웃풋을 계산하는 문제를 푸는 걸 의미합니다. 그 자리에서 우리가 어떤 함수, 형태, 프라이어(Prior)를 만든다는 것을 의미하고 이것을 우리가 앞에서 시피 프라이어로 프로우드(Proof)를 했었어요. 그랬는데 프로젝션을 어떻게 할 수 있을까요? 어떤 컨디셔널 디스트리뷰션(Conditional Distribution)을 표현할 수 있는 거예요. 사실 내가 x 와 y 라는 이스테이트 아웃이 주어졌을 때 새로운 입력 x 스타를 플레인 데이터의 컨디션한 컨디션드 GPC를 표현할 수 있는 거죠.

영어검출 및 변환 파트

입력 데이터

그래서 보죠. 그러면 gp라는 게 뭔지 알았으니까 이제 gp를 가지고 리그레션 어떻게 될까 아웃션 프로세스라는 게 뭔지 아는 것과 gp랑 gpr은 달라요. gp는 가우셰어 프로세스 자체로 gpr은 아우셰어 프로세스를 이용해서 리그레션 문제 클핑스 로그인에서 인프 아웃풋이 주어졌을 때 새로운 입력에 대한 프레이트 아웃풋을 계산하는 문제를 푸는 걸 의미합니다. 그 자리에서 우리 어떤 꺼는 코드를 찾게 되면 내가 찾고자 하는 출력, 어떤 함수, 형태 프로모션이 이런 프라이어를 만든다는 걸 의미하고 이거를 우리가 앞에서 시피 프라이어로 프라우드를 했었어요. 그랬는데 프로젝팅을 어떻게 할 수 있냐면 어떤 컨디셔널 디스플레이션을 표현할 수 있는 거예요. 사실 내가 x랑 y라는 이스터 아웃이 좋아졌을 때 새로운 입력 x 스타를 플레인 데이터의 컨디션한 컨디션드 gvc를 표현할 수 있는 거죠.

프롬프트

위 내용은 STT 기술을 이용해 음성을 텍스트로 변환한 거야.
그런데 한국어 문장에 영어 용어가 섞여 있다보니 발음 등의 문제로 영어 용어가 제대로 변환이 되지 않은 것들이 있어.

위 내용에 대한 지식과 아래 관련 키워드들을 참고해서
위 내용에서 부정확하게 변환이 된 부분에 올바른 용어를 괄호로 추가해
위 문장/문단을 수정해줘.

위키피디아, OCR 추가

1) 교안
Function-space View
The specification of the kernel function implies a prior distribution over functions.
 $f^* \sim \mathcal{N}$
(0, K(X*X))
Prediction can be done using the conditional distribution. Suppose that we are given n training data
=

2) 위키피디아
From Wikipedia, the free encyclopedia
In probability theory and statistics, a Gaussian process is a stochastic process (a collection of random variables indexed by time or space), such that every finite collection of those random variables has a multivariate normal distribution, i.e. every finite linear combination of them is normally distributed. The distribution of a Gaussian process is the joint distribution of all those (infinitely many) random variables, and as such, it is a distribution over functions with a continuous domain, e.g. time or space.

생성 내용

"그래서 보죠. 그러면 Gaussian Process (가우시안 프로세스)라는 게 뭔지 알았으니까 이제 GP를 가지고 회귀 (Regression) 어떻게 할까요? 아웃풋 프로세스 (Gaussian Process)라는 게 뭔지 아는 것과 GP랑 GPR은 달라요. GP는 가우시안 프로세스 자체이고, GPR은 가우시안 프로세스를 이용해서 회귀 문제를 의미합니다. 즉, 클립스 로그인 (Continuous Log-in)에서 인풋 아웃풋 (Input과 Output)이 주어졌을 때 새로운 입력에 대한 프리딕트 아웃풋 (Predicted Output)을 계산하는 문제를 푸는 걸 의미합니다. 그 자리에서 우리가 어떤 코드를 찾게 되면 내가 찾고자 하는 아웃풋이, 어떤 함수, 형태를 프로모트 (Promote)한다는 걸 의미하고 이것을 우리가 앞에서 시피 프라이어 (CP Prior)로 프라우드를 했었어요. 그랬는데 프로젝팅 (Projecting)을 어떻게 할 수 있냐면 어떤 콘디셔널 디스트리뷰션 (Conditional Distribution)을 표현할 수 있는 거예요. 사실 내가 x랑 y라는 이스터 아웃이 주어졌을 때 새로운 입력 x 스타를 플레인 데이터의 콘디셔널 커버리언스 (Conditional Covariance)를 표현할 수 있는 거죠. 그래서 우리가 N개의 데이터가 있고 그리고 한 개의 테스트 이력이 있으면 n개의 y_1, \dots, y_n 이 있고 1개의 y^* 라는 게 있으면 애가 어떤 분포를 따를 건데 이거는 다 0이고 여차피 미니트렉션 (Mean Function)으로 써놔지 않습니까? 포먼트렉션 (Mean Function)이 그렇게 생겼을 거냐면 여기에 n바이 n 짜리와 1xN 이렇게 생겼을 거예요. 이 큰 블록 매트릭스가 N개의 데이터에 대한 커널 매트릭스이고 나머지는 테스트 데이터와 메인 데이터 사이의 플레이스 텍터 (Placeholder)를 정의할 수 있어요. 그래서 이거는 트레이닝 데이터의 조인트 디스트리뷰션 (Joint Distribution)과 어떤 플레이트의 아웃풋 한계 (Output Limit)를 정의할 수가 있습니다. 그러면 우리가 GP의 성질을 알 수 있는 게 뭐냐면 이런 식으로 가우시안 디스트리뷰션 (Gaussian Distribution)이라는 게 있으면."

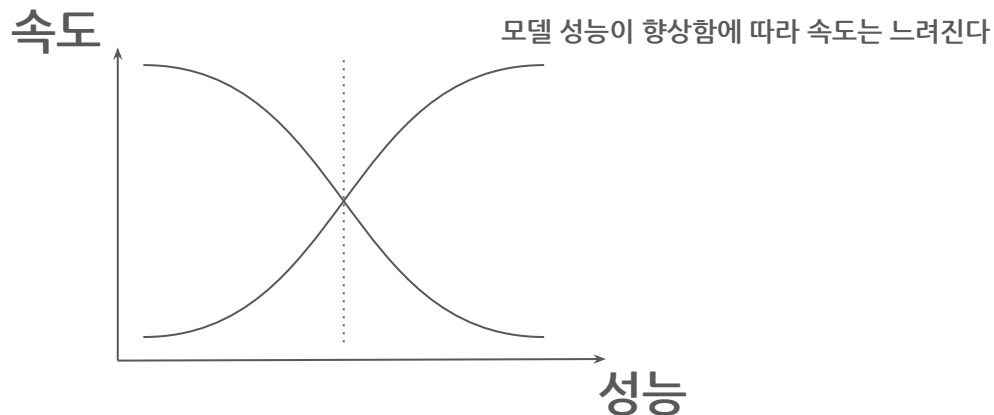
심화된 프롬프트 엔지니어링 필요

한계점

STT 한계점



(GCP 기준)
Large-v2 모델 사용



해결방법

1. Whisper.cpp 모델 사용
2. Faster-Whisper 모델 사용
3. 하드웨어 GPU 추가(현재 사용 방법)

QnA

Appendix

- Whisper : <https://github.com/openai/whisper>
- Whisper.cpp : <https://github.com/ggerganov/whisper.cpp>
- Faster-Whisper : <https://github.com/SYSTRAN/faster-whisper>
- Streamlit Docs : <https://docs.streamlit.io/>
- Pytube : <https://github.com/pytube/pytube>
- youtube-transcript-api : <https://github.com/jdepoix/youtube-transcript-api>
- Bark & Coqui TTS : <https://tts.readthedocs.io/en/dev/models/bark.html>
- suno-bark : <https://huggingface.co/suno/bark/tree/main>
- TTS : <https://github.com/coqui-ai/TTS>
- Pre-trained Model: <https://github.com/hegelty/hate-classifier>
- Unsmile dataset: https://github.com/smilegate-ai/korean_unsmile_dataset
- Other models: <https://github.com/2runo/Curse-detection>
- Lime: <https://github.com/marcotcr/lime>

아키텍처

