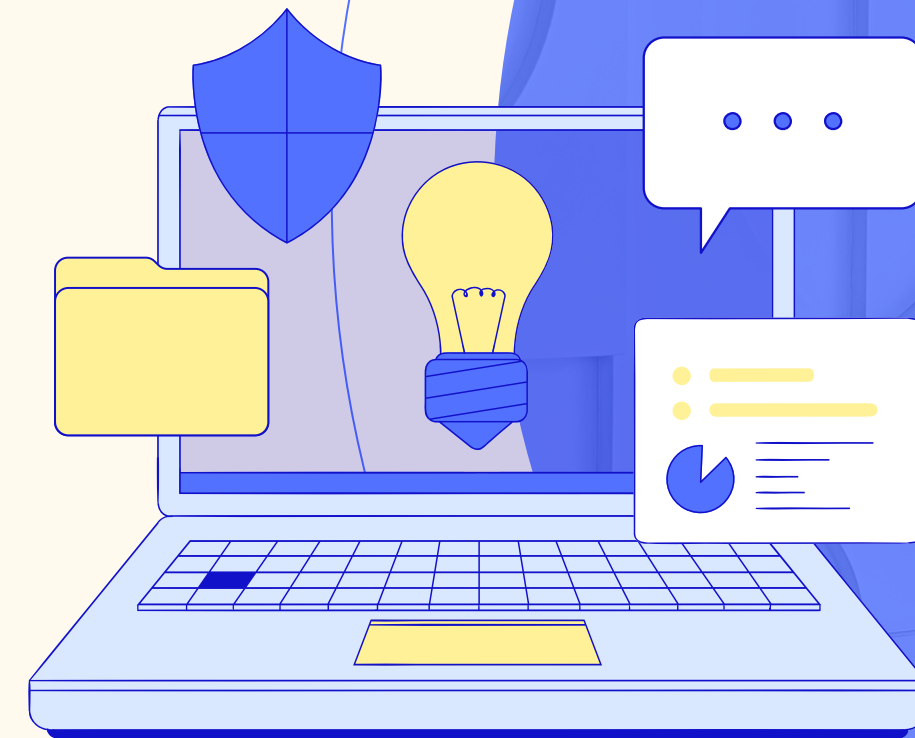


EST BOOTCAMP

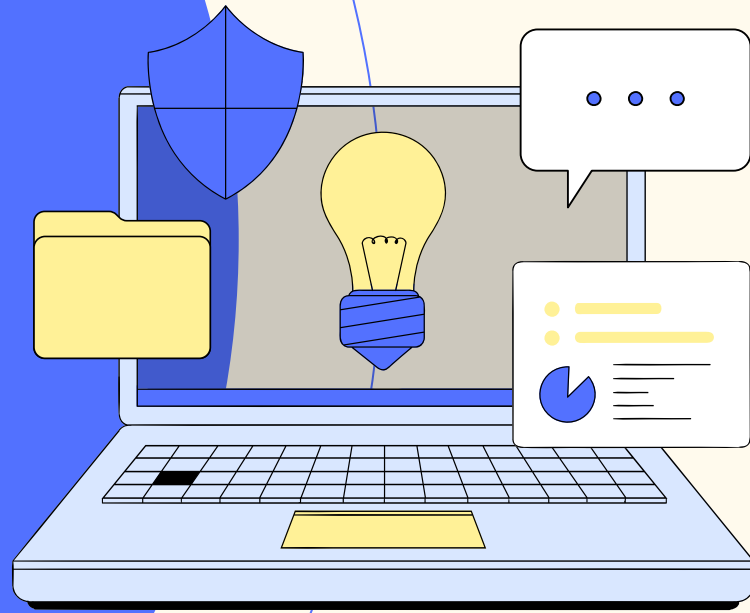
# 시계열 프로젝트

1조 김호곤팀(김호곤, 김정철, 신민경)



# 목차

---



01

**프로젝트 소개**

02

**유사서비스 비교 및  
차별점**

03

**설계**

04

**개발내용**

05

**시연**

06

**기대효과·활용방안**

# 소개

## 프로젝트 소개

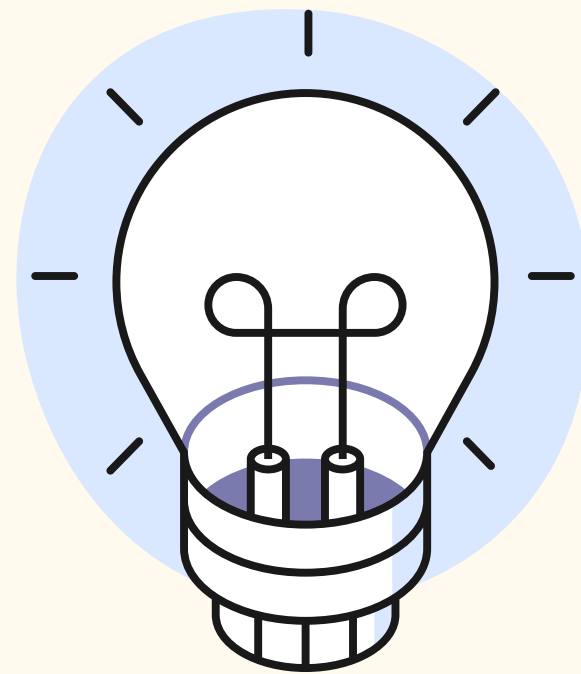
### 유사서비스 비교 및 차별점

### 설계

### 개발내용

### 시연

### 기대효과·활용방안



## 음성을 내가 원하는 텍스트 말투로 STT

음성 언어 입력 → 새로운 텍스트 문체

- 평소 음성인식에 어려움을 겪는 청각장애인들
- 평소 글을 작성할 때 문체에 어려움을 겪는 사람들

프로젝트 소개

유사서비스 비교 및  
차별점

설계

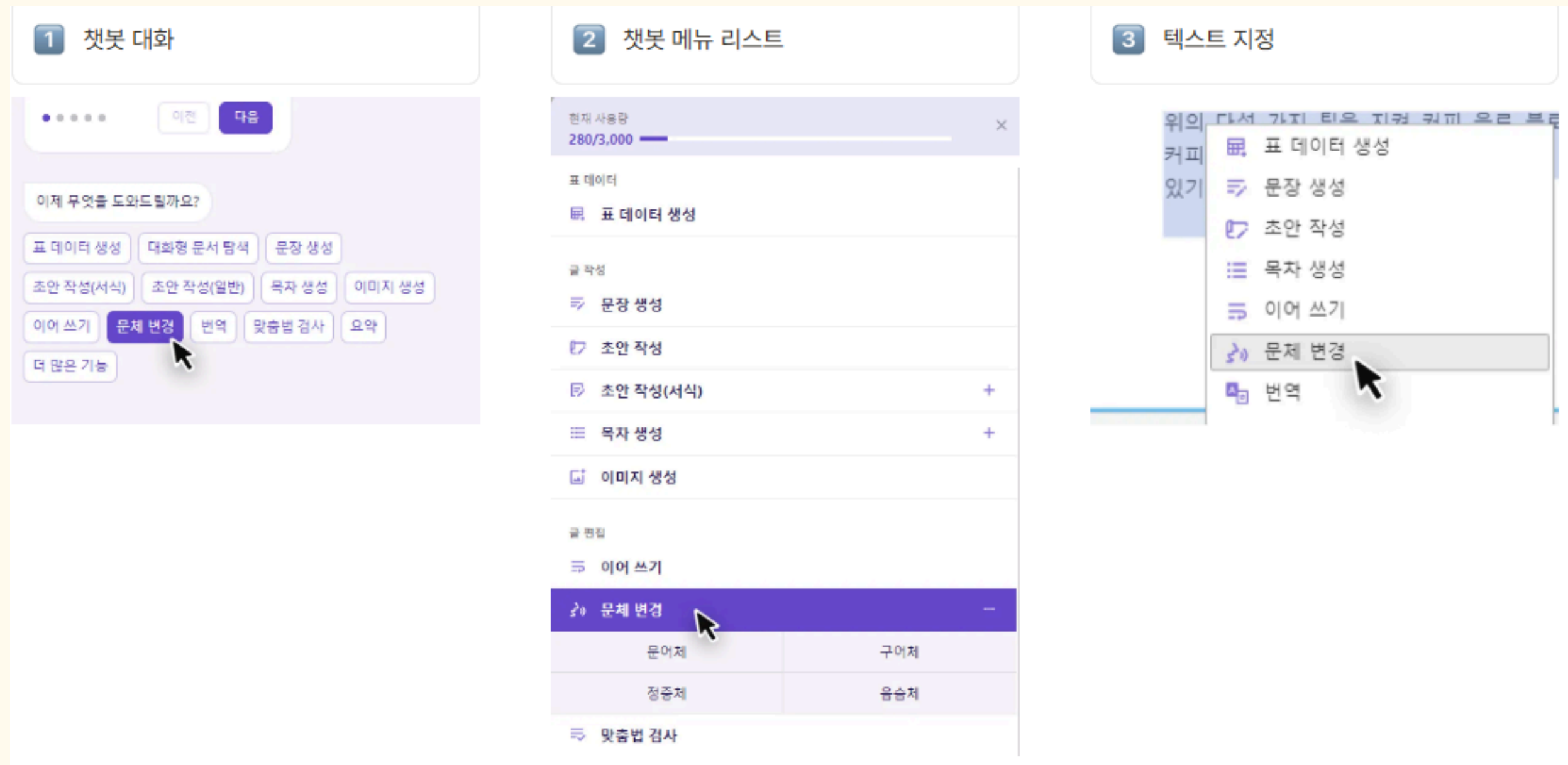
개발내용

시연

기대효과·활용방안

# 유사서비스 비교 및 차별점

## 유사서비스 : 한컴 AI 문체 변경 기능



프로젝트 소개

유사서비스 비교 및  
차별점

설계

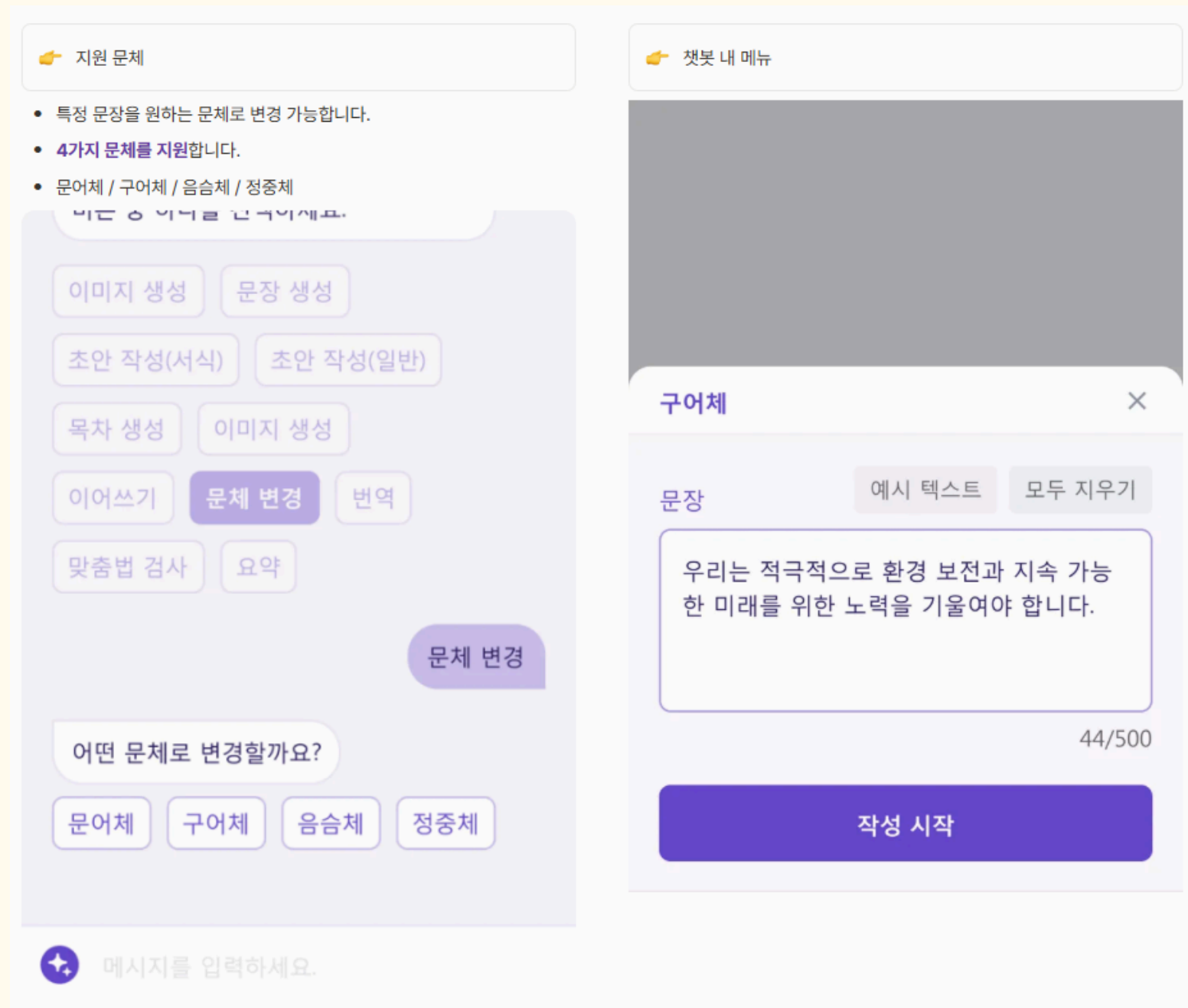
개발내용

시연

기대효과·활용방안

# 유사서비스 비교 및 차별점

## 유사서비스 : 한컴 AI 문체 변경 기능



프로젝트 소개

유사서비스 비교 및  
차별점

설계

개발내용

시연

기대효과·활용방안

# 유사서비스 비교 및 차별점

## 차별점

1. 해당 모델을 사용하면 17가지 문체를 사용할 수 있습니다.
2. 음성을 텍스트로 변환하고 그 텍스트의 문체를 변환하여 사용할 수 있습니다.

입력 문장 :

어쩌다 마주친 그대 모습에  
내 마음을 빼앗겨 버렸네  
어쩌다 마주친 그대 두 눈이  
내 마음을 사로잡아 버렸네  
그대에게 할 말이 있는데  
왜 이리 용기가 없을까  
음 말을 하고 싶지만 자신이 없어  
내 가슴만 두근두근 답답한  
이 내 마음 바람 속에 날려 보내리

formal(존댓말) :

어쩌다 마주친 사람의 두 눈에 마음이  
휘말려버렸네요.  
어쩌다 마주친 사람의 두 눈이 내 마음  
을 휘저어 버렸네요.

informal(반말) :

어쩌다 마주친 사람의 눈에 마음이 휘  
말려버렸어.  
어쩌다 마주친 사람의 두 눈이 내 마음  
을 휘저어 버렸어.  
왜?

choding(초등학생) :

어쩌다 마주친 두 눈에 마음이 휘둘러  
버림,,  
어쩌다 마주친 두 눈이 내 마음을 휘어  
감,, 왜?

enfp :

어쩌다 마주친 두 눈에 마음이 휘말려  
버렸어  
어쩌다 마주친 두 눈이 내 마음을 휘어  
잡아 버렸어 왜??

프로젝트 소개

유사서비스 비교 및  
차별점

설계

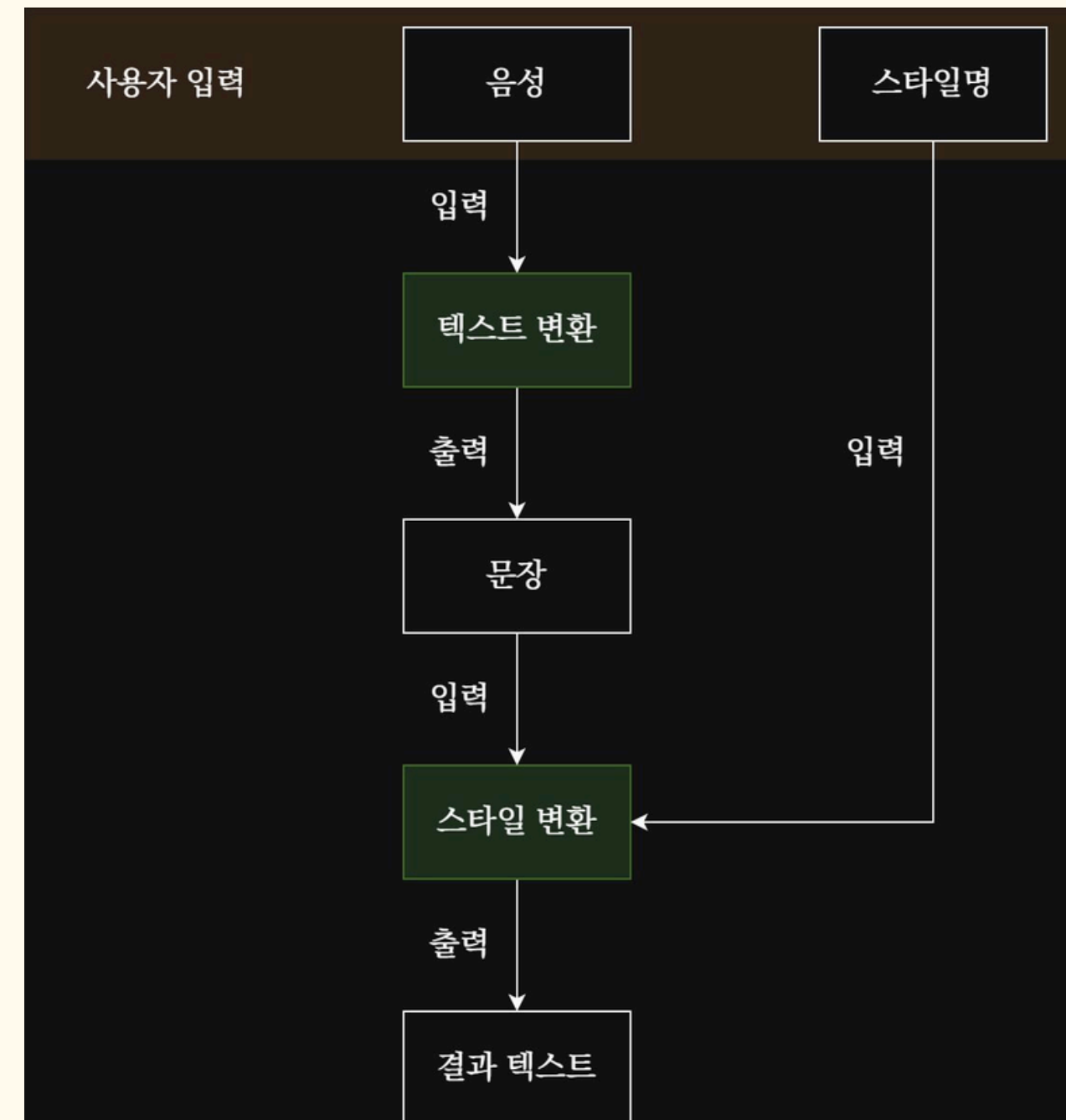
개발내용

시연

기대효과·활용방안

# 설계

## 마인드 맵 및 내용



### 1. 주요기능

- 1) 사용자가 음성 입력합니다.
- 2) 음성-텍스트 변환 성공 후 사용자가 변환하고자 하는 스타일명 입력합니다.
- 3) 텍스트 문체 변환 후 결과 출력됩니다.

### 2. 사용 기술 및 도구

- 1) 개발도구 : Google Colab
- 2) 개발언어 : Python
- 3) 음성-텍스트 변환 모델 : Whisper
- 4) 문체 변환 :
  - (1) 모델 : KoBart-Base-V2
  - (2) 도구 : Transformers
  - (3) 데이터 셋 : SmileGate 한국어 스타일 데이터 셋

프로젝트 소개

유사서비스 비교 및  
차별점

설계

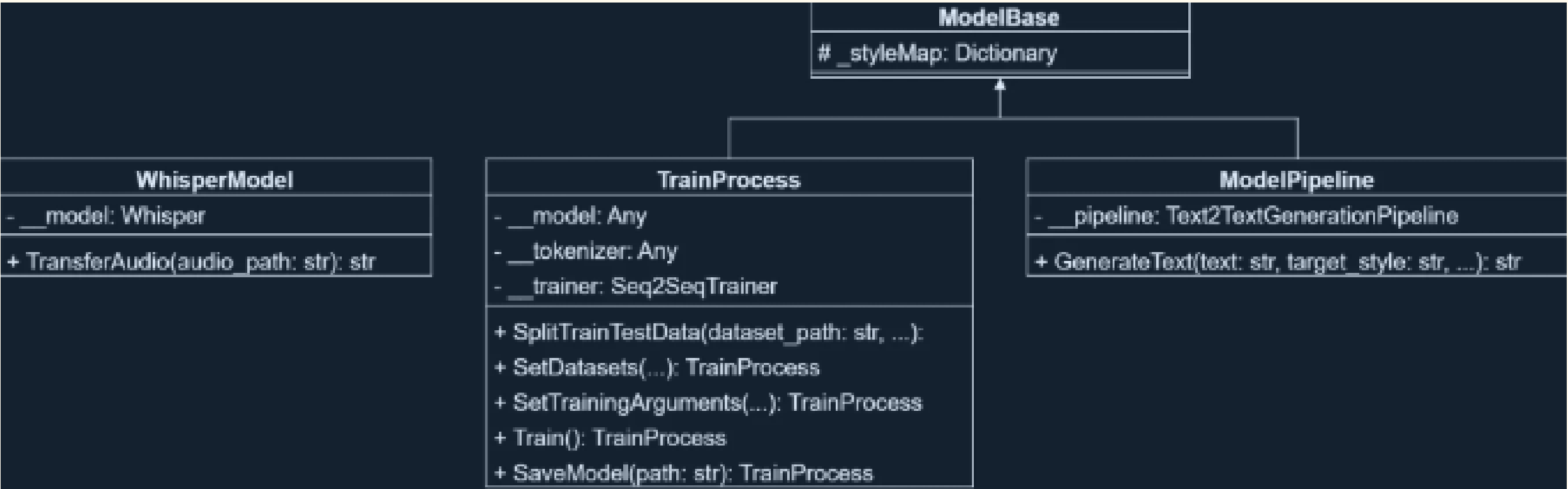
개발내용

시연

기대효과·활용방안

# 개발내용

## 1. 모듈화



- (1) 모델 학습과 학습된 모델의 사용을 모듈화합니다.
- (2) Main 스크립트에서는 최소한의 모듈을 임포트하여 실행 가능합니다.
- (3) 학습 과정, 모델 사용 코드를 은닉화하여 수정 불필요한 코드와 데이터를 보호하고 Main 스크립트의 가독성을 증가시킵니다.
- (4) 정의된 함수를 호출하여 간단한 코드로 동작 구현 가능합니다.



프로젝트 소개

유사서비스 비교 및  
차별점

설계

개발내용

시연

기대효과·활용방안

# 개발내용

## 2. 학습

```
from TextStyleTransfer import TrainProcess

# 모델 학습

model = TrainProcess(

    ).SplitTrainTestData(dataset_path = dataset_path

    ).SetDatasets(

    ).SetTrainingArguments(output_dir = f"{model_path}dummy/"

    ).Train(

    ).SaveModel(path = model_path

)
```

(1) 모든 함수는 자기 자신의  
참조를 반환하기 때문에  
줄줄이 매서드 접근 가능합니다.  
(2) 함수에 추가적인 인자를  
전달하여 세부적인 학습 방법  
조절할 수 있습니다.

프로젝트 소개

유사서비스 비교 및  
차별점

설계

개발내용

시연

기대효과·활용방안

# 개발내용

## 3. 최종 평가

(1) 오디오-텍스트 변환 평가

1) 평가 지표

- WER: (삽입 오류 + 삭제 오류 + 치환 오류) / 전체 정답 단어 수
  - CER: (삽입 + 삭제 + 치환) / 전체 정답 문자 수
- 숫자가 낮을 수록 높은 정확도를 보입니다.

2) 평가 내용 예시

- |                          |                          |                           |
|--------------------------|--------------------------|---------------------------|
| • 정답: 지금 개봉했어?           | • 정답: 그러면 그 모자는?         | • 정답: 나 그때 뭐냐 뭐지? 무슨 영화지? |
| • 예측: 지금 개봉했어?           | • 예측: 그러면 그 모자는?         | • 예측: 나 그때 뭐냐 뭐지 무슨 영화지   |
| • WER: 0.000, CER: 0.000 | • WER: 0.000, CER: 0.000 | • WER: 0.333, CER: 0.105  |

3) 최종 평가

- 평균 WER: 0.321
- 평균 CER: 0.137

프로젝트 소개

유사서비스 비교 및  
차별점

설계

개발내용

시연

기대효과·활용방안

# 개발내용

## 3. 최종 평가

(3)문체 스타일 변환 평가

- 평가 지표
  - Validation Loss
- 최종 평가
  - 에포크 16, 스텝 간격 500으로 2500 스텝에서 Validation Loss가 약1.29로 가장 낮았습니다.

[3136/3136 09:39, Epoch 16/16]

Step	Training Loss	Validation Loss
500	No log	1.828878
1000	No log	1.557056
1500	No log	1.528180
2000	No log	1.438569
2500	No log	1.286781
3000	No log	1.429533

프로젝트 소개

유사서비스 비교 및  
차별점

설계

개발내용

시연

기대효과·활용방안

# 개발내용

## 모델 과정



프로젝트 소개

유사서비스 비교 및  
차별점

설계

개발내용

시연

기대효과·활용방안

# 시연영상

```
오디오 경로 >> ./data/GSF03031LJH0_한국어/SW201105KTRNGSF03031LJH0001.pcm  
pcm 파일을 wav 파일로 변환.  
저장 경로: ./data/GSF03031LJH0_한국어/SW201105KTRNGSF03031LJH0001.wav  
스타일 입력 >> king
```

원본 문장: 더 좋은 모습 기대할게요. 주소 광고에서 너무 귀여우시던데요.

변환 문장: 더 좋은 활약을 기대하오. 주소가 광고에서 귀여운 듯 하오.

종료하려면 end 입력. 계속하려면 아무 키 입력 >> end

1. 사용자는 오디오 경로에 음성을 입력하거나, 변환하고자 하는 문체 스타일을 선택합니다.
2. 모델이 원본 문장과 출력 문장을 반환합니다.
3. pcm 파일 입력 시 자동으로 wav 파일로 변환됩니다.

프로젝트 소개

유사서비스 비교 및  
차별점

설계

개발내용

시연영상

기대효과·활용방안

# 기대효과·활용방안

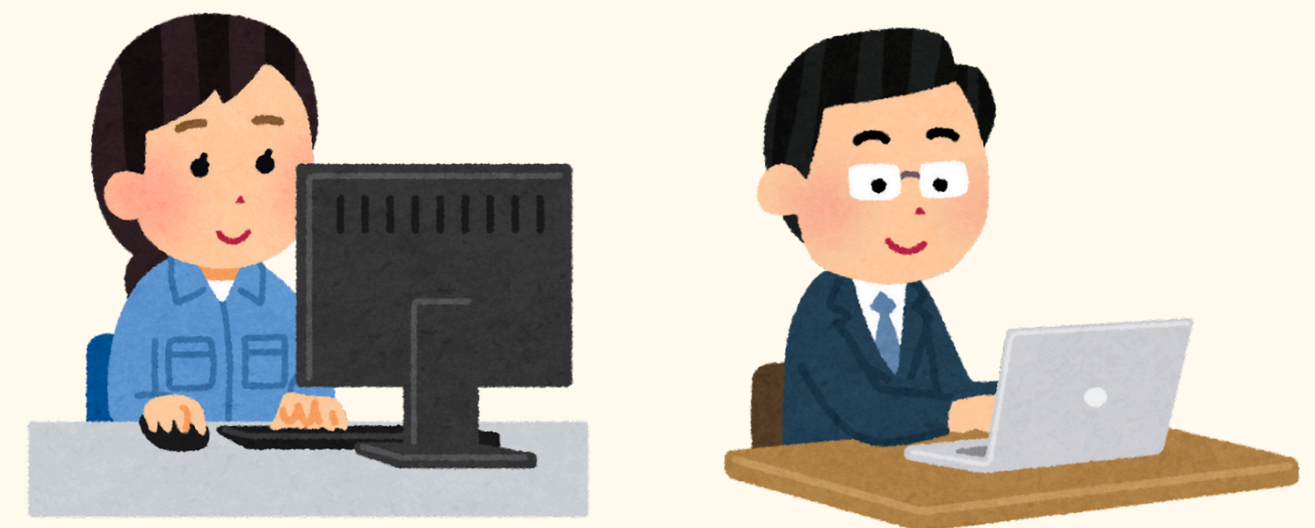
## 기대효과

1. 청각장애인을 도울 수 있습니다.  
해당 모델을 사용하면 청각장애인에게 음성을 텍스트로 보여줄 수 있습니다. 그래서 청각장애인이 원하면 언제, 어디서든지 음성을 텍스트화 시킬 수 있습니다.
2. 고령자의 편의를 향상시킵니다.  
청각적으로 어려움을 겪는 고령자는 힘들게 들을 필요 없이 글자로 상대방의 음성을 파악할 수 있습니다.



## 활용방안

1. 비즈니스 이메일 자동화  
사용자가 편하게 텍스트를 입력하면 해당 모델은 문체를 비즈니스 상황에 맞게 변환시켜 줍니다.
2. 안내음성 텍스트화  
청각적으로 어려움을 겪는 사람들은 안내음성을 듣지 못할 수도 있습니다. 하지만 해당 모델을 사용하면 음성을 텍스트로 바꿀 수 있기에 그 어려움을 해결할 수 있습니다.





경청해 주셔서 감사합니다

---

1조 김호곤팀(김호곤, 김정철, 신민경)