



# 음성인식을 활용한 보안 시스템

20518035 박성재 | 20518053 심성관

# 01

## 지난 발표 후 변경 사항

### 1. 주제 변경 및 응용분야 결정



음성 보안을 활용한 스마트 홈



음성 보안 시스템

#### 응용분야

- 스마트 홈
- 스마트 도어락
- 차량 도난 방지 잠금
- 개인 금고 잠금 장치

- 온라인 시험 부정행위 방지
- 게임에서 화자의 음성별 상황 구현

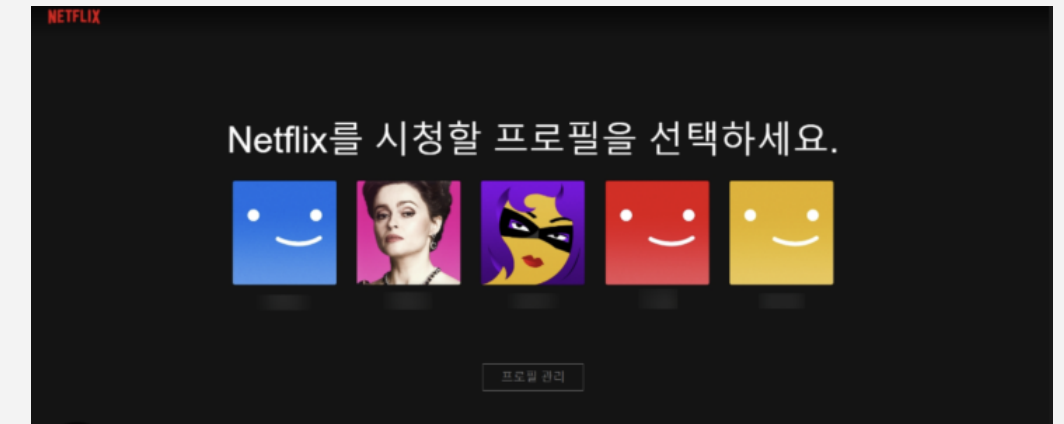
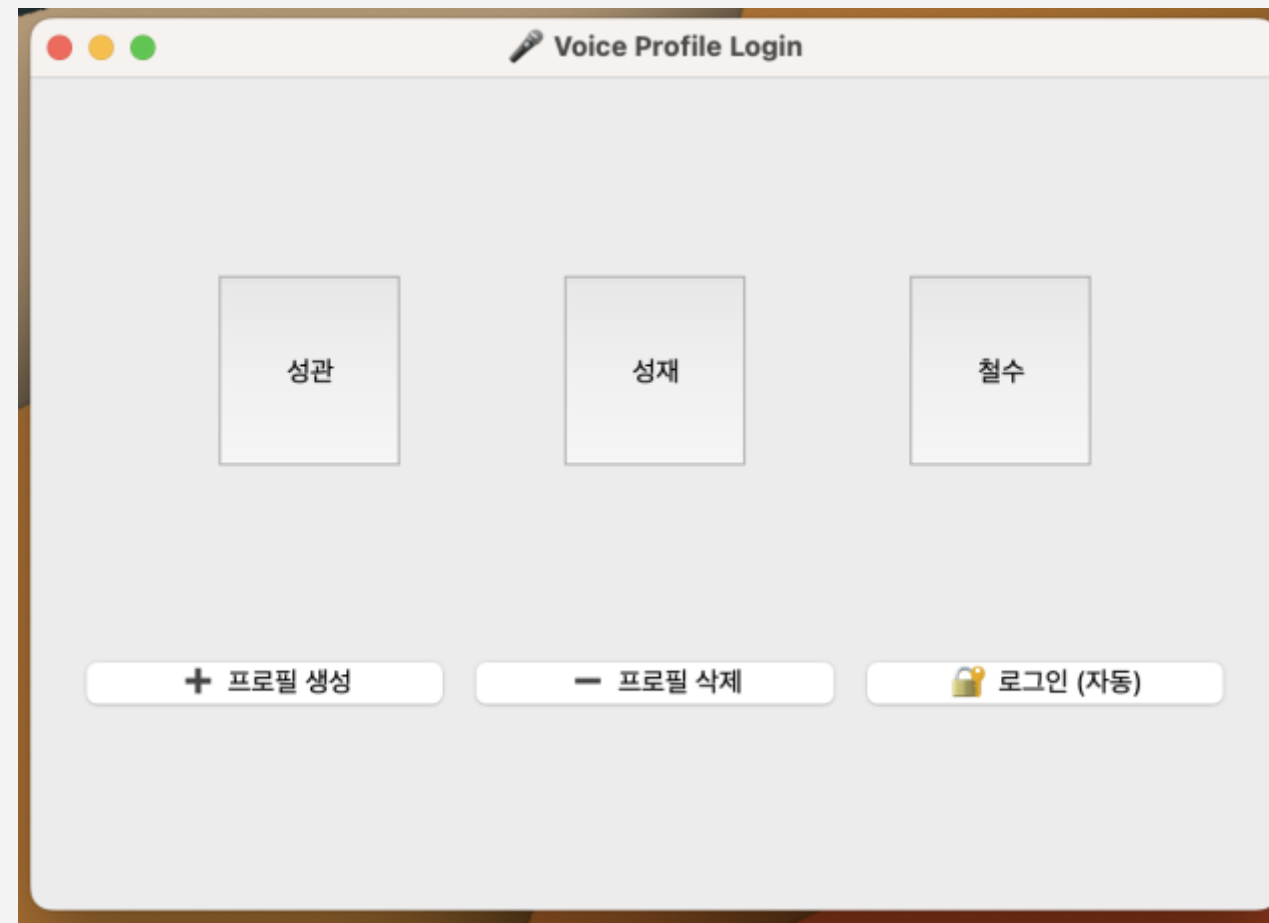
# 01

## 지난 발표 후 변경 사항

### 2. 디스플레이 활용

디스플레이를 활용한 UI는 음성 보안 시스템의 직관성을 크게 향상시킬 수 있음

간단하게 Python으로 구현한 UI

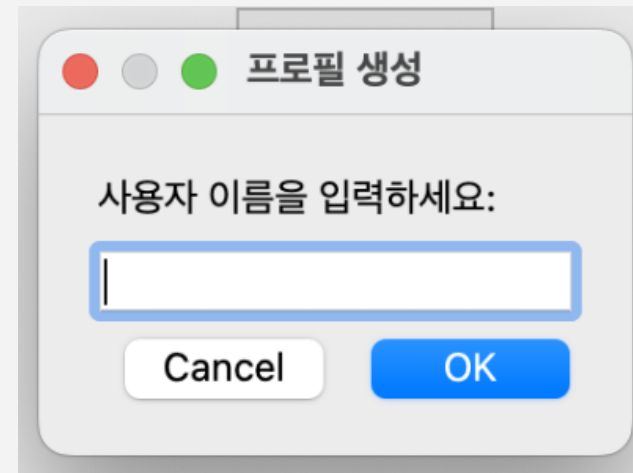


예) 넷플릭스 로그인 화면

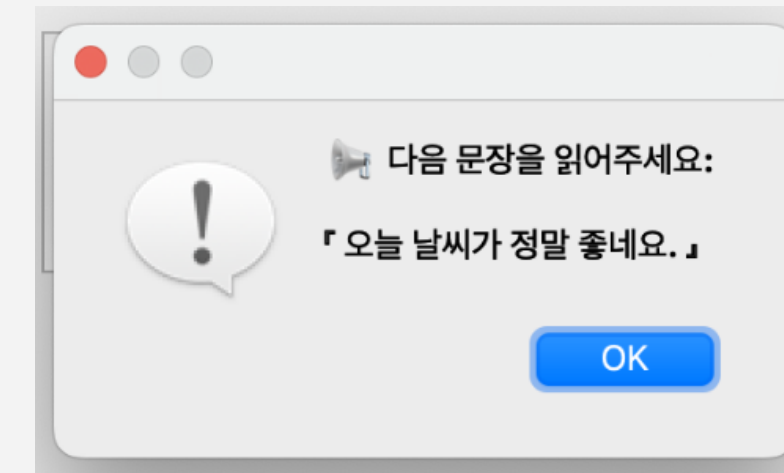
# 02

## UI 설명

### 1. 프로필 생성



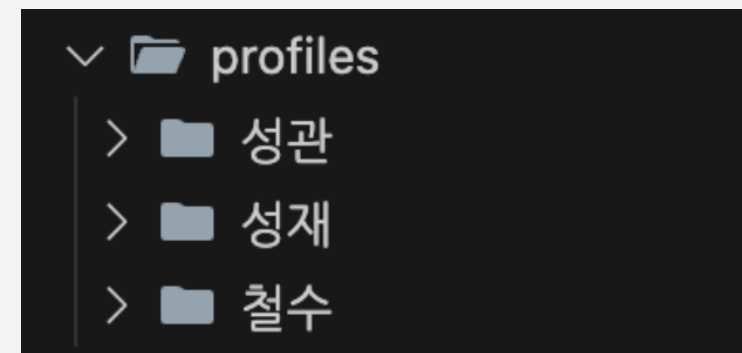
프로필 이름 설정



프로필 음성 녹음

```
ENROLL_SENTENCES = [  
    "오늘 날씨가 정말 좋네요.",  
    "인공지능은 우리의 삶을 바꾸고 있어요.",  
    "파이썬은 정말 유용한 언어입니다.",  
    "이 문장을 또박또박 읽어주세요.",  
    "커피 한 잔의 여유를 즐기세요."  
]
```

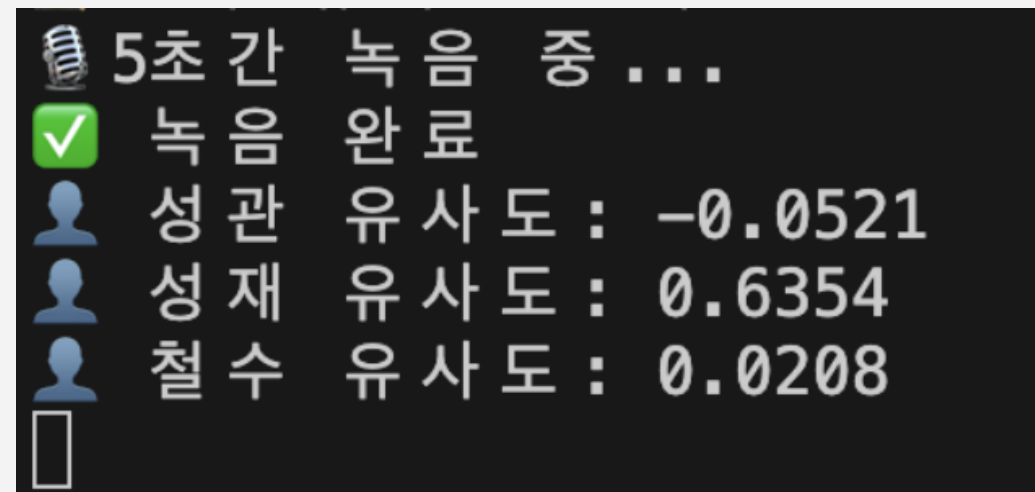
사용자에게 보여줄 녹음 문구



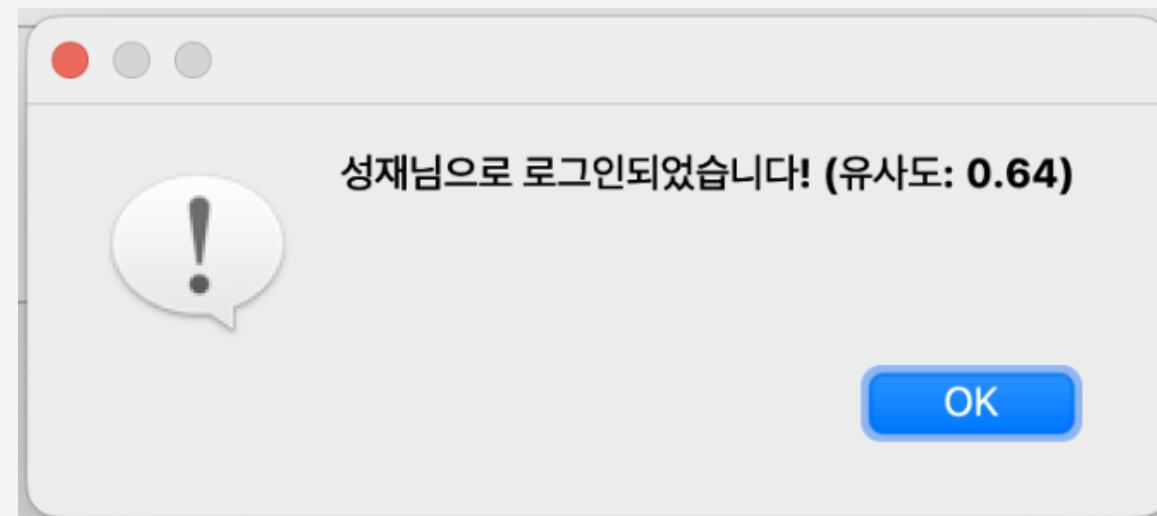
# 02

## UI 설명

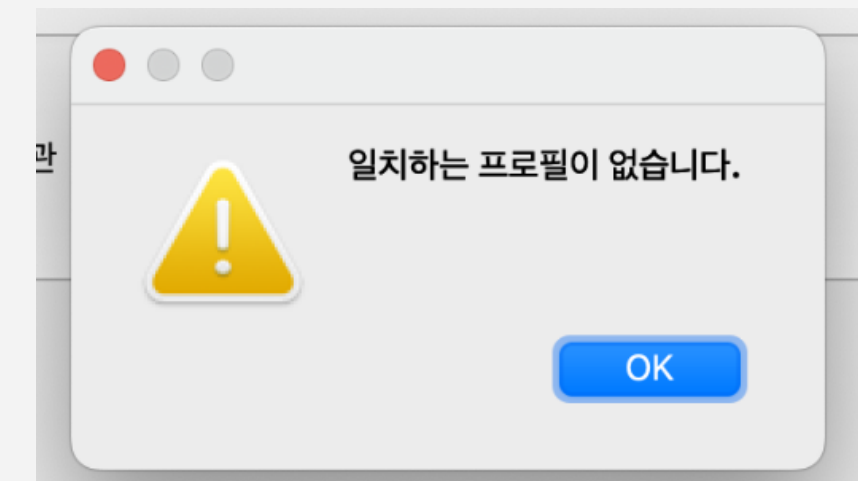
### 2. 로그인 - 1차 인증



로그인에 필요한 음성 5초 녹음 요구



프로필에 저장되어있는 녹음 문구와 같은 문구로 로그인

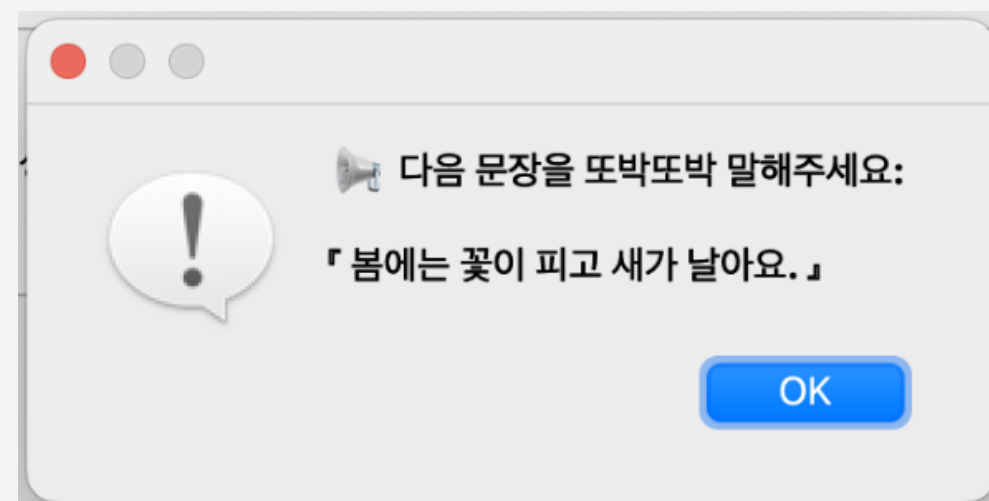


프로필에 일치하는 음성 없음

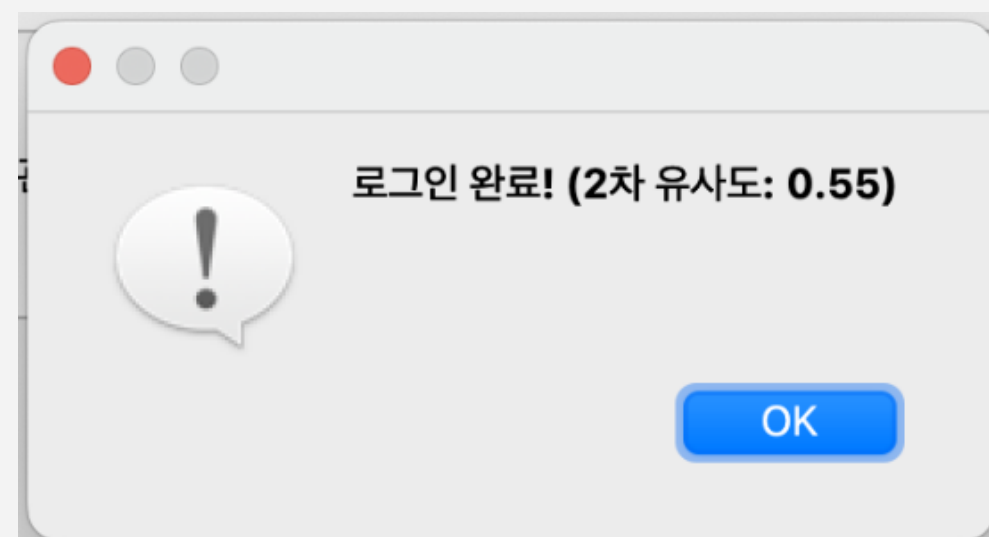
# 02

## UI 설명

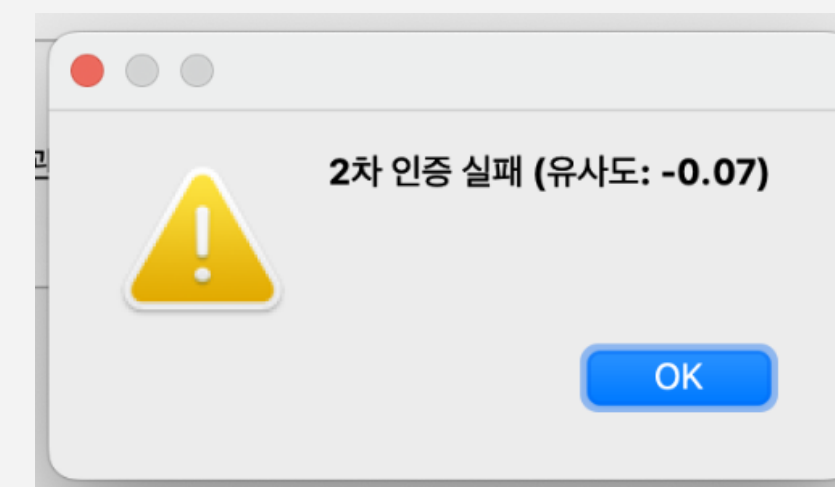
### 3. 로그인 - 2차 인증



2차 인증에 필요한 음성 5초 녹음 요구



프로필에 저장되어있는 녹음과 2차로 녹음한 음성을 비교

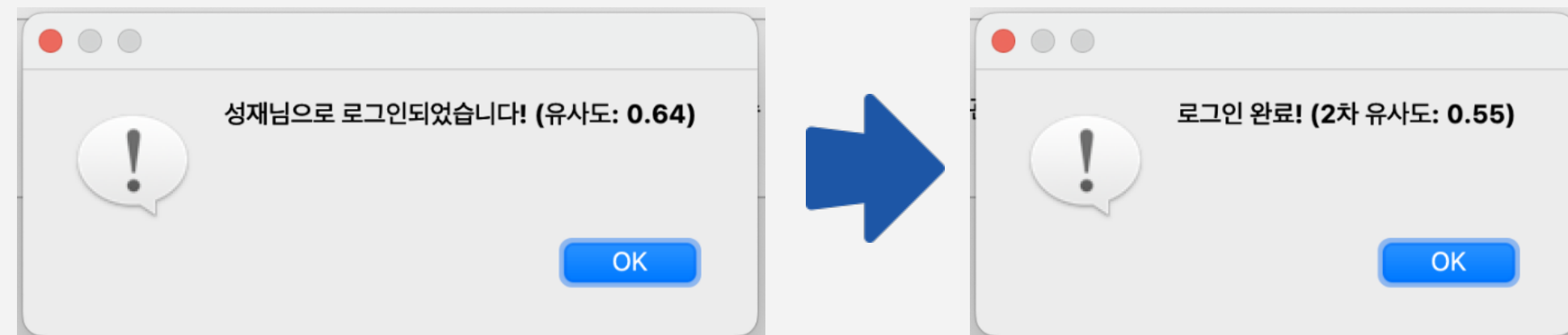


2차 인증 실패 팝업창

# 03

## 문제점 발견 및 해결 방안

2차 인증 시 유사도가 떨어짐



1차 유사도 0.64 -> 2차 유사도 0.55

# 03

## 문제점 발견 및 해결 방안

### 해결 방안

1. 녹음 시간 늘리기 (5초 -> 10~15초)
2. 여러 문장 평균 임베딩 사용

### "임베딩" 이란?

- 사람의 음성은 파형이지만, 인공지능 모델은 파형 자체로는 의미를 잘 파악하지 못함.
- 그래서 AI 모델은 이 음성을 "숫자 벡터(임베딩, embedding)"로 바꿈

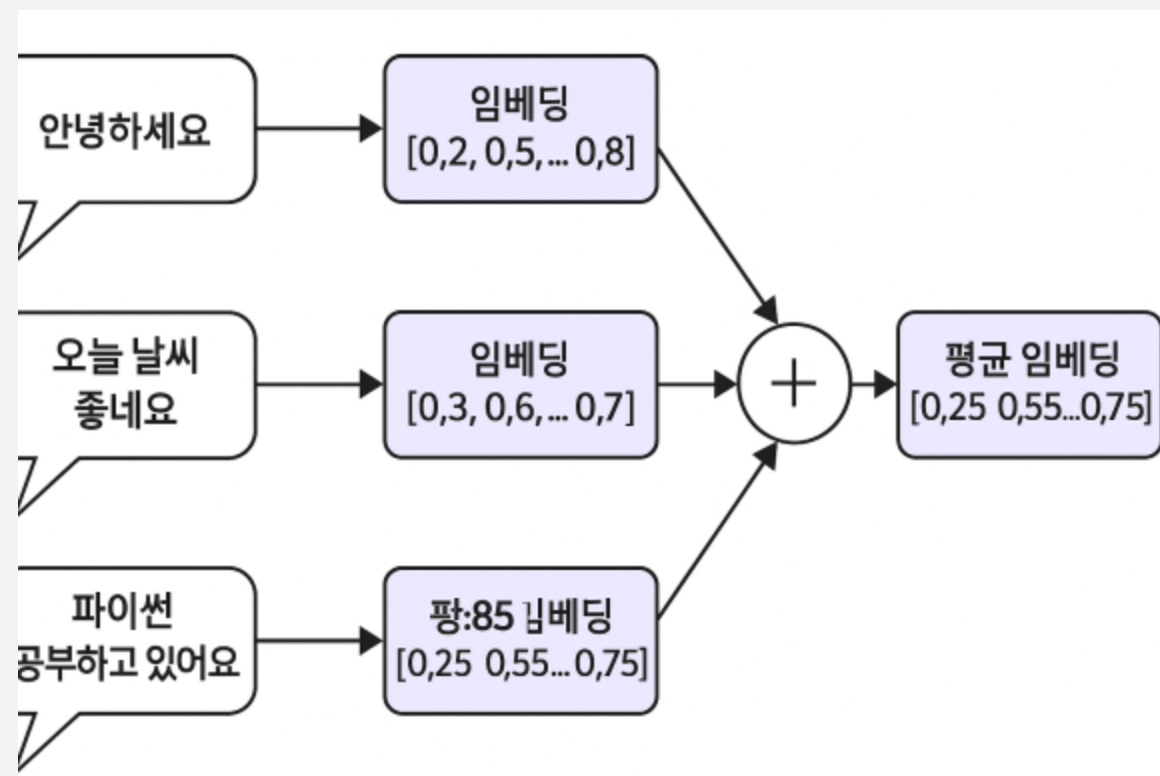
문장	생성된 임베딩
"안녕하세요"	[0.2, 0.5, ..., 0.8]
"오늘 날씨 좋네요"	[0.3, 0.6, ..., 0.7]
"파이썬 공부하고 있어요"	[0.25, 0.55, ..., 0.75]



# 03

## 문제점 발견 및 해결 방안

### 해결 방안



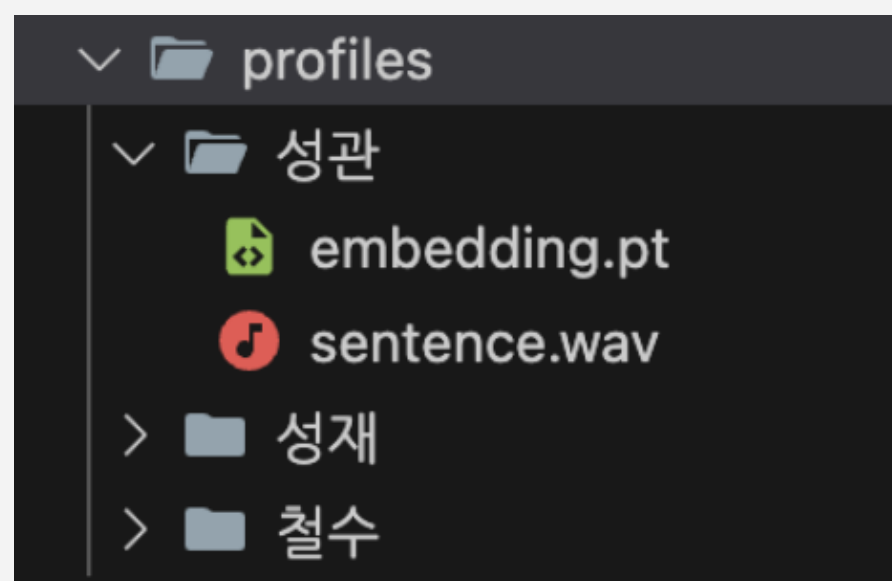
여러 번의 음성 녹음을 통해 나온 여러 임베딩 벡터들을 평균 내면,  
화자의 본질적인 특징만 남게 되어 인식 정확도가 높아진다.

# 03

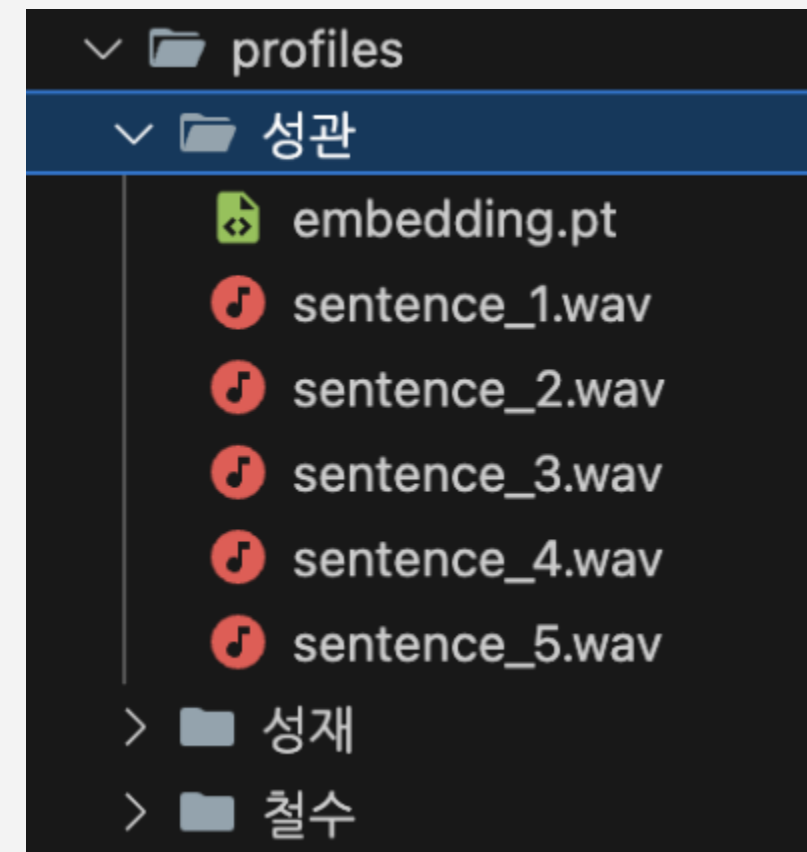
## 문제점 발견 및 해결 방안

### 해결 방안

5개의 문장을 녹음한 후 임베딩 벡터들의 평균을 내 embedding.pt 파일을 생성해 비교한다.



기존 프로필 생성 시 1개의 문장을 읽어 프로필 생성



프로필 생성 시 5개의 문장을 읽어 프로필 생성

# 04

다음 발표까지의 목표



## 응용 분야 선택

음성 인식 보안 기술을 적용할 분야를 선택



## 라즈베리파이 사용

라즈베리파이에서 지금까지 구현한 코드를 올려 테스트



## 유사도 점수 올리기

여러 문장을 녹음해 평균 임베딩을 높이고 녹음할 5개의 문장 선택

Q&A