

# 주간 활동 보고서

## 목차

### 주간 진행 사항

---

- 회의 & 활동 진행 시간 및 내용
- 데이터셋 형식의 차이

# <주간 활동 보고서 9 주차> -2024.04.29 ~ 2024.05.05-

## -주간 진행 사항

### <회의 및 활동 진행 시간>

2024.05.01 (수) 18:00 ~ 22:00 / 2024.05.02 (목) 15:00 ~ 22:00

- 총 활동 시간 : 11 시간 00 분

### -회의 및 활동 내용

>> [스마트업] 오프라인 회의 (05.01)~(05.02)

기존 계획서의 aihub 음성 데이터셋과 Kaldi와의 호환성이 맞지 않음  
이번 주차는 대학 강의용 음성 데이터셋과 Kaldi 툴킷이 서로 호환되지 않는 원인을 분석하고 해결 방안을 마련함.

AIHub 대학 강의용 음성 데이터셋과 Kaldi 툴킷의 비호환성 분석

AIHub 에서 제공하는 대학 강의용 음성 데이터셋과 Kaldi 툴킷의 호환성 문제 발생

AIHub 에서 다운로드한 음성 데이터셋과 Kaldi 툴킷 간의 호환성 문제는 다음과 같은 원인으로 분석

## 1. 데이터셋 형식의 차이

AIHub 데이터셋은 특정 형식의 오디오 파일 또는 메타데이터로 제공됩니다. 이 형식은 Kaldi 툴킷에서 바로 사용하기에 적합하지 않을 수 있습니다. Kaldi 는 일반적으로 WAV 파일 형식과 Sphinx 등의 표준 데이터셋 형식을 사용하는데, 우리가 사용하려는 데이터셋 형식이랑 호환되지 않았다

대안적으로, AIHub 데이터셋을 다른 음성 인식 시스템에 적용하기 위해 호환성을 고려한 도구를 검토했다.

따라서 AIHub 의 대학 강의용 음성 데이터셋과 Kaldi 툴킷의 호환성 문제를 극복하기 위해 Conformer 모델을 선택했다

Conformer 는 최신의 Transformer 아키텍처를 기반으로 한 음성 인식 모델로, 다양한 데이터 형식과 처리 방식에 대해 유연하게 적용할 수 있다. 이 모델은 특히 문맥을 이해하고 긴 의존성을 처리할 수 있는 Transformer 구조를 활용하여 AIHub 의 음성 데이터셋을 효과적으로 인식할 수 있고 추후에 적용될 요약 알고리즘에도 유리하다.

다음 주차는 컨포머를 기반으로 데이터 분할을 시도하여 라즈베리 파이에서 무리없이 작동할수 있도록 최적화를 시키도록 하고 요약 알고리즘을 외부 api 로 빼는 작업을 완성해야함