



2024-03-17[월 단위 계획]

☰ 태그	
👤 참석자	기흥 기흥 김

회의 이슈

1. stt 결과 공유
2. 성능 지표 내기
3. 로드맵

목표

- 연구실 내부적으로는 paper 1, sw 1
- 기업 R&D
 - 중간에 작업이 추가될 가능성 있음.
- 한 학기 :

이 프로젝트의 로드맵 수립을 위해, 전체적인 월 단위 계획과 최종 목표 및 결과물, 목표 달성을 위해 필요한 요구사항, 그리고 팀 구성원 각각의 역할과 책임을 아래와 같이 정리합니다.

최종 목표 및 결과물:

학기별 계획

[1학기]

- 실시간 자막 변환 소프트웨어(SW) 개발(1학기)
- 실시간 자막 변환 소프트웨어 결과 보고서(1학기)

[하계방학]

- 기존 오픈 소스 분석에 대한 논문(하계 방학) - 보류

- 방언 처리가 가능한 실시간 자막 변환 기술 개발 착수(하계 방학)

[2학기]

- 방언 처리가 가능한 실시간 자막 변환 기술 개발(2학기)
- 방언 처리가 가능한 실시간 자막 변환 기술 개발 보고서(2학기)
- 방언 처리가 가능한 실시간 자막 변환 기술 논문(2학기)

전체적인 월 단위 계획:

1. 3월~4월: 프로젝트 준비 및 조사

a. 1주차(3/11 ~ 3/17)

- 팀 구성, 역할 분담
- 기존 stt 성능 분석(google, aws)
- 프로젝트 로드맵 작성

b. 2주차~4주차(3/18 ~ 4/7)

- 기존 솔루션 조사 및 기술 리서치
 - paper로 나와 있지만 상용화되지 않은 open source 찾아보기
- 개발 도구 및 라이브러리 설치
- 개발 환경 설정 (Python, C/C++, Docker, Kubernetes)

2. 4월~5월: 음성 신호 처리 및 STT 기본 엔진 개발

a. 5주차(4/8~4/14)

- 오디오 스트림 분리 및 wav 변환
- 기본 STT 엔진 구현 시작
- [2024-04-01] - 계획 변경
 - Solution 및 SOTA Model Evaluation Metric 분석 (WER 등)
 - Audio Separation (RAW → Speech + other Noise)
 - STT Engine Baseline Code 구성

b. 6주차~10주차(4/15~5/19)

- STT 엔진 구현

- STT 엔진 최적화(구현과 일부 병행)

3. 5-6월: 성능 최적화 및 보고서 작성

a. 11주차~14주차(5/20~6/16)

- STT 성능 최적화
- 결과보고서 작성

4. 6월: 성능 개선

a. 15~17주차(6/17~6/30)

- STT 성능 최적화

5. 7월~8월: 특화 기능 개발 및 기존 오픈 소스 분석에 대한 논문

a. 18주차~24주차(7/1~8/18)

- Closed Caption 생성 알고리즘 개발
- 기존 오픈 소스 분석에 대한 논문

b. 25주차~26주차(8/19~9/1)

- 방언 처리 소프트웨어 개발 착수

6. 9월~10월: 방언 처리 소프트웨어 개발

a. 27주차~31주차(9/2~10/6)

- 방언 처리 소프트웨어 개발

7. 10월~11월: 문서 작성 및 KCI 논문 작성

- 문서화 및 사용자 매뉴얼 작성
- 논문 작성

8. 12월: 프로젝트 마무리 및 발표

- 결과 검토 및 최종 발표 준비
- 결과 발표 및 논문 제출

목표 달성을 위해 필요한 요구사항:

- 하드웨어 요구사항:
 - 고성능 GPU (예: NVIDIA RTX 3080 이상) - 딥러닝 모델 훈련 및 Fine-tuning에 필요
 - 충분한 저장 공간과 메모리 (예: SSD 1TB, RAM 32GB 이상)

- **소프트웨어 요구사항:**

- 개발 언어: Python, C/C++
- 프레임워크 및 라이브러리: TensorFlow or PyTorch, Kaldi (STT), Docker, Kubernetes
- 데이터 처리 및 분석 도구

- **기타 요구사항:**

- 방송 콘텐츠 샘플 데이터
- 클라우드 서비스 계정 (모델 훈련 및 테스트를 위한)

팀 구성원 각각의 Role & Responsibility:

1. 프로젝트 리더: 박성준

- 전체적인 프로젝트 관리 및 진행 상황 모니터링
- 주요 알고리즘 개발 및 최적화

2. 데이터 엔지니어: 하은지

- 데이터 수집, 전처리, 관리
- 음성 신호 분리 알고리즘 개발 지원

3. 머신러닝 엔지니어: 김기흥

- STT 모델
- 훈련 및 Fine-tuning
- 모델 성능 평가 및 개선