< 퍼, 터, 커 >

**1. 개요**

**1-1. 소프트웨어 목적**

**• 전체 목적:** 마비말장애 평가

**• 퍼, 터, 커 :** 단일 음절 초당 평균 반복 개수 측정

**• 퍼터커:** 서로 다른 3음절 초당 평균 반복 개수 측정

**1-2. 소프트웨어 개발 환경 및 실행 환경**

**1-2-1. 개발 환경**

**• 프로그래밍 언어:** Python 3.10

**• 패키지 관리 및 환경 구성:** Conda 및 pip

**• 가상 환경 이름:** CLAP\_PC

**• DB :** MySQL 8.0.43 Community

**1-2-2. 주요 라이브러리**

|  |  |  |
| --- | --- | --- |
| **범주** | **라이브러리** | **버전** |
| 데이터 처리 | numpy | 2.0.2 |
| pandas | 2.3.1 |
| 딥러닝 | tensorflow | 2.19.0 |
| tensorflow-macos (macOS 전용) | 2.15.1 |
| 오디오 처리 | librosa | 0.11.0 |
| 모델 API 호출 | tensorflow-hub | 0.16.1 |
| 데이터베이스 연결 | mysql-connector-python | 9.1.0 |

**2. 소프트웨어 구현 사항**

**2-1. 소프트웨어 구성 모듈**

• audio\_preprocess()

• wav\_padding()

• x\_data\_preprocess(): 종합 전처리 수행

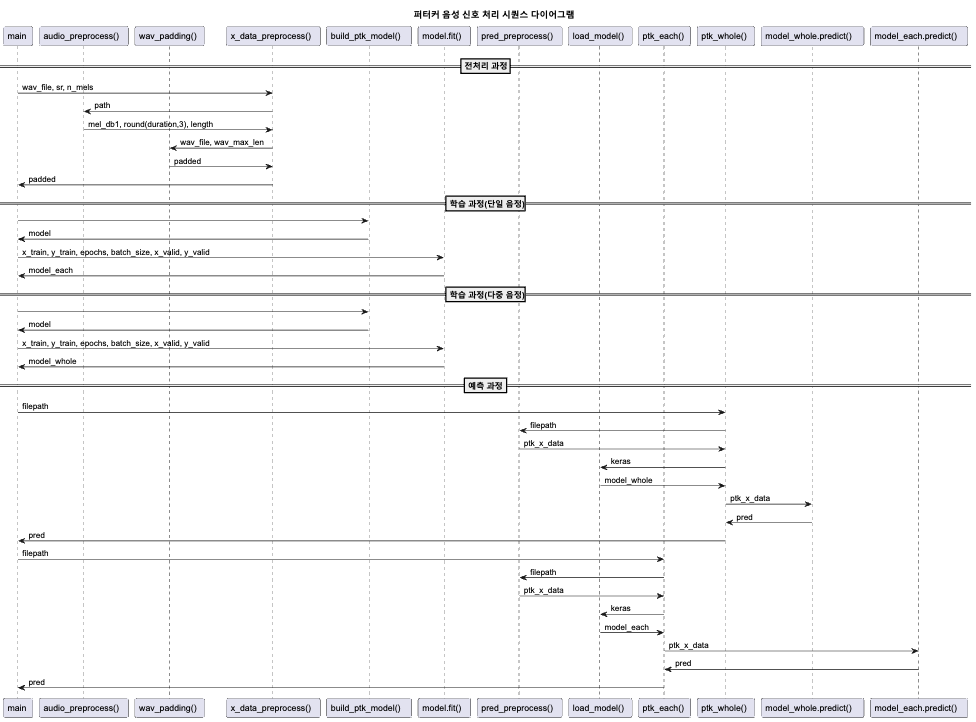
• pred\_preprocess()

• build\_ptk\_model()

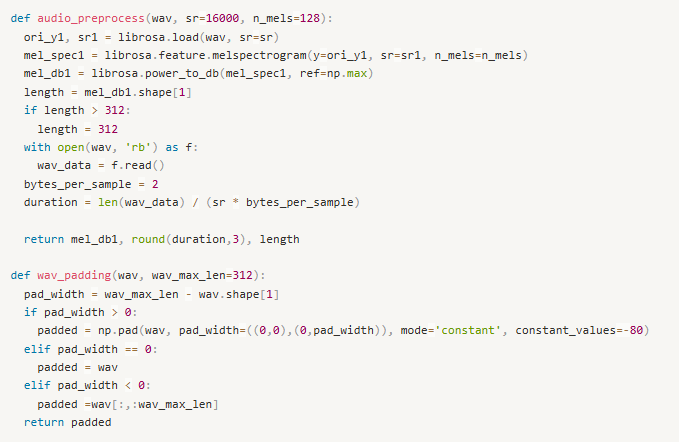
• ptk\_each()

• ptk\_whole()

**2-2. 소프트웨어 구성 모듈별 설명**



**2-2-1. 데이터 전처리 모듈**



**2-2-1-1. audio\_preprocess()**

**• 목적:** 오디오 데이터 정규화

**• 파라미터:** 파일 저장 경로(string) , sr(Sample Rate), n\_mels(차수)

**• 리턴 값:** 정규화된 파일, time, 배열 크기

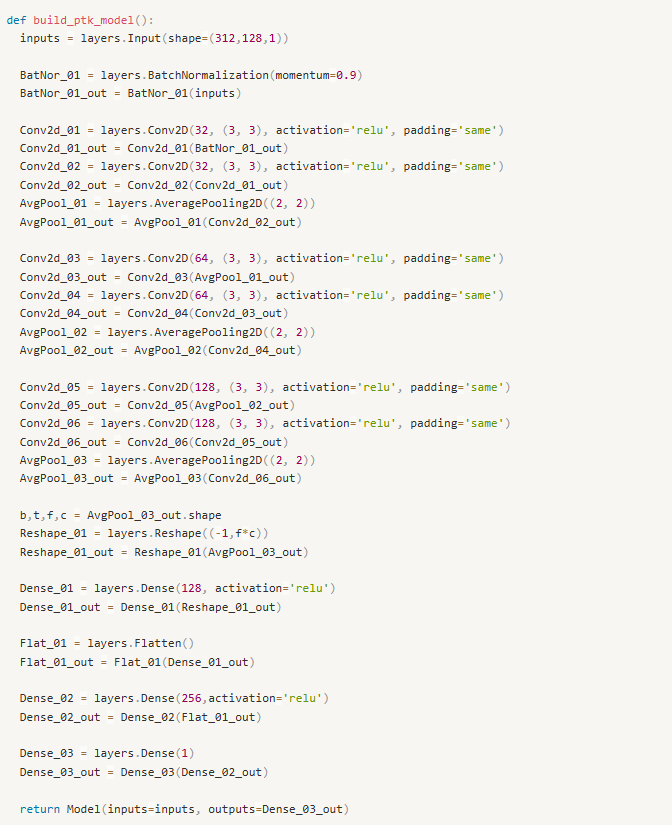
**2-2-1-2. wav\_padding()**

**• 목적:** Input data 차수 통일을 위한 패딩

**• 파라미터:** 파일 저장 경로(string), 패딩 길이(기본값=312)

**• 리턴 값:** 패딩된 파일

**2-2-2. 학습 모듈:** build\_ptk\_model()



**• 목적:** 단일/다중 음절 평균 반복 시간 측정하는 Conv2D 기반 CNN 회귀 모델 구조 구축

**2-2-2-1. 입력 데이터**

**•** **멜스펙토그램화된 음성파일**

**• shape:** (312, 128, 1) = time, freq(멜 밴드), channel(모노채널)

**2-2-2-2. 출력 데이터**

**•** **shape:** (batch, 1) = 각 입력 멜스펙토그램에 대한 회귀값

**2-2-2-3. 모델 전체 구조**

텍스트, 스크린샷, 번호, 폰트이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**2-2-2-4. 모델 레이어별 설명**

**• Batch\_normalization:** 첫 input data 처리 시 배치 정규화

**• CNN:** 특히 이미지 분류 작업에서 시퀀스 분석을 수행하여 관련이 높은 핵심 특징(feature)을 올바르게 추출하는 머신러닝 모델

**2-2-2-5. 학습방법**

**• 옵티마이저:** adam

**• 손실 함수:** mse

**• 배치 크기:** 16

**• 에포크:** 100(Early stopping 적용: patience=10, monitor=val\_loss)

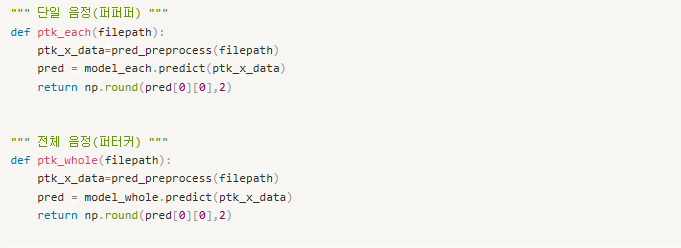
**• 비고**

**•** train, test 데이터 비율은 9:1로 설정

**•** 단일 음성 측정(퍼/터/커) 모듈과 다중 음성 측정(퍼터커) 모듈을 분리함(ptk\_each, ptk\_whole)

- 모든 shape는 동일하며 hyperparameter만 상이

**2-2-3. 예측 모듈**



**• 함수명**

**•** ptk\_each(): 단일 음정 분석(퍼/터/커)

**•** ptk\_whole(): 다중 음정 분석(퍼터커)

**• 입력 데이터:** wav 파일 경로

**• 출력 데이터:** 단위 시간당 카운트된 평균 음절 수(float32)

**• 모델 구조:** 전처리, 모델 호출, 결과 반환

**• 상관계수:** 두 모델 모두 0.97**텍스트, 폰트, 스크린샷이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.**

텍스트, 폰트, 스크린샷이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**3. 소프트웨어 개선 사항**

**3-1. 예외 사항 처리**

**3-2. 기능 구현**

분석 데이터가 총 60명 가량으로 소수이며, 이 중에서도 이상 환자의 데이터는 특히 부족함. 데이터 추가 수집 또는 생성적 적대 신경망(GAN)을 활용해 성능을 높일 수 있을 것으로 기대함.