

# 컨설팅 논의

## 1. 첫 번째 연구 목표 (초기 제안)

- **목표:** 편마비 환자나 특히 실어증 환자의 재활 훈련 시 **SEMG (Surface Electromyography)** 신호를 이용하여 주관적인 피로도를 추론하는 모델을 만드는 것입니다.
- **배경:** 실어증 환자의 경우 훈련 중 힘든 정도를 말로 표현하기 어렵기 때문에.
- **방법:** SEMG 센서가 받아들이는 근육의 전기적 신호(전압값)를 시계열 데이터로 변환하고, 이 객관적인 SEMG 신호와 **Borg CL 10**이라는 주관적인 피로도 점수 사이의 상관관계를 학습하여, SEMG 신호를 통해 보고 점수를 회귀(Regression) 예측하고자 합니다.
- **난제:** 이 주제를 진행할 경우 환자 섭외가 매우 어렵고, 환자의 데이터는 프라이빗한 데이터이므로 연구 서약 및 동의를 받는 것이 필수적이며 걱정되는 부분입니다.

## 2. 두 번째 연구 목표 (대안)

- **목표:** **IMU (Inertial Measurement Unit)** 센서 데이터를 사용하여 근육의 활성도를 예측하는 프로그램을 만드는 것입니다.
- **배경:** SEMG 센서는 민감하고 부착이 예민하여 측정이 어렵기 때문에, 부착이 간편한 IMU만으로 활성화 정도를 예측하고자 합니다.
- **활용 데이터:** IMU 센서에서 나오는 시간에 따른 데이터와 추가적으로 이 사람의 체중, 키, 체지방률을 고려하여 근육 활성도를 예측합니다. IMU 센서는 자이로 센서로, 위치나 가속도를 파악합니다.
- **활성도 정의:** 운동 시 각도가 중요한 것이 아니라 근육이 얼마나 활성화되었는지가 중요하기 때문에 근육 활성도를 예측하고자 합니다.

### 핵심 기술적 논의: 라벨링 및 컨버전

IMU 기반 모델(두 번째 목표)을 학습시키기 위해서는 정답 라벨(근육 활성도)이 필요한데, 이 라벨을 SEMG 신호로부터 어떻게 신뢰성 있게 추출할 것인가가 가장 큰 논점이었습니다.

#### 1. SEMG를 이용한 라벨 추출의 어려움:

- SEMG 신호는 근육 활성도를 측정하기 위한 목적의 센서 데이터입니다.
- 초기 제안은 SEMG 신호의 **RMS (Root Mean Square)** 값을 계산하여 활성도를 얻겠다는 것이었으나, SEMG 볼트값과 문헌에 나오는 근육 활성도값은 스케일이 다르므로 이를 변환(컨버전)하는 과정이 필요합니다.

- 시계열 데이터인 SEMG에서 RMS를 통해 라벨 값을 꽂는 것이 적절한 근거가 있는지에 대한 의문이 제기되었습니다.

## 2. 신뢰성 있는 라벨(Ground Truth) 확보 방안:

- 자기들이 측정한 데이터(SEMG RMS 값)가 정답이라고 판단할 근거가 없으므로, **신뢰성 있는 장비나 외부에 공인된 측정기를** 사용하여 근육 활성도를 측정하고, 이를 정답 라벨(활성도 등급 또는 스칼라 값)로 삼아야 합니다.
- 이 라벨 값이 있어야 RMS로 컨버전하는 방법이 맞는지, 아니면 뉴럴 네트워크로 컨버전해야 하는지를 확인할 수 있습니다.

## 3. 뉴럴 네트워크 활용 방안 논의:

- **기본 구조:** IMU 데이터 (인풋)를 네트워크에 통과시켜 활성도 (라벨)를 구하는 방식.
- **OTPA 기법 제안:** 박상욱 조교는 OTP(경로 다른 애를 맵핑) 기법을 활용하여, IMU 신호를 SEMG 신호로 맵핑하는 뉴럴 네트워크를 만들고 (IMU → SEMG), 이후 SEMG에서 활성도 라벨을 추출하는 방식이 더 나은 방법론일 수 있다고 제안했습니다.
- **결론:** 뉴럴 네트워크를 활성도 값을 직접 출력할 때 쓸 것인지, 아니면 **SEMG 신호로 컨버전할 때 쓸 것인지** (두 가지 옵션)를 결정해야 합니다. 최소한 SEMG에서 활성도 라벨을 추출하는 과정(컨버전)을 먼저 명확히 정의하는 것이 필요합니다.

## 데이터 수집 및 실험 계획의 난제

연구를 위해 실제로 데이터를 확보하는 것(1번)이 가장 중요하고 시급한 일입니다.

### 1. 멀티모달 데이터 수집의 어려움:

- IMU와 SEMG 데이터를 동시에, 같은 사람, 같은 시간대에 뽑아야 하므로 데이터 게더링이 꽤 까다롭습니다.
- 센서 종류가 다르기 때문에 멀티모달 데이터라고 할 수 있으며.
- **샘플링 속도 차이 (Hz):** IMU는 약 200Hz, SEMG는 약 50Hz로 데이터 헤르츠가 다르므로 이를 맞춰주는 작업(동기화)이 필요합니다.

### 2. 실험 규모 및 환경:

- 실험의 보편성을 확보하려면 최소한 5명 정도의 피실험자가, 근육 활성화 정도(예: 5개 등급)당 **최소 세 텐씩** 데이터를 뽑아야 할 것으로 추정됩니다.
- 실험 장비는 화정 체육관의 풀업(턱걸이) 기구를 사용할 계획입니다.

### 3. 전처리 (Pre-processing) 문제:

- 타임 시리즈 데이터는 처리가 지저분한 데이터입니다.
- **전처리**는 데이터 종류나 목적에 따라 정답이 없으며, 어떤 전처리 기법을 적용할지 결정하고 테스트하는 것 자체가 많은 시간이 소요되는 노동 집약적인 일입니다.

이러한 논의를 통해, **데이터 확보와 신뢰성 있는 라벨링(Ground Truth)을 정의하는 작업**이 현재 뉴럴 네트워크 구축 고민보다 선행되어야 함이 강조되었습니다.