

# SPRT를 이용한 MI 분류 모델의 정확도에 관한 연구

조준희, 최계원

성균관대학교

jhcho9811@g.skku.edu, kaewonchoi@skku.edu

## A Study on the Accuracy of the MI Classification Model Using SPRT

Jun Hee Jo, Kae Won Choi

Sungkyunkwan Univ.

### 요약

EEG 신호를 이용한 MI 분류 기술은 학습된 모델을 이용해 입력받은 뇌파의 특징을 추출하여 MI를 분류하는 기술이며, SPRT는 상한 한계점과 하한 한계점을 설정하여 MI 분류를 할 때, 더 짧은 시간에 분류를 할 수 있는 방법이다. 본 논문에서는 오픈 데이터셋인 Lee 데이터셋을 이용해 학습된 기존 모델의 정확도를 계산하는 과정에서 SPRT의 적용 여부를 변화시켜 진행하였다. 이로써 SPRT 적용 여부에 따른 모델의 정확도 차이를 분석하였다.

### I. 서론

과거에 비해 접하기 쉬워진 뇌파 측정 장비의 도입 등으로 인해 BCI(Brain-Computer Interface)와 관련된 많은 연구들이 진행되고 있다. BCI는 인간의 뇌와 컴퓨터 간의 통신 채널을 제공하는데 그 중에서도 BCI의 연구 분야 중 하나인 MI(Motor Imagery)가 주목받고 있다. MI는 몸은 실제로 움직이지 않고 움직임을 상상하는 것을 의미한다. 예를 들어 왼손, 오른손, 발과 같은 신체의 움직임을 상상할 때 발생하는 뇌의 생체 신호에 관한 연구를 진행했다.

EEG 신호를 이용하여 MI 데이터 분류하는 기술은 상상하는 것만으로 발생한 뇌파를 입력으로 하여 분류를 하는 기술이다. 기존에는 4초 길이의 데이터를 모델의 입력으로 사용하여 분류를 진행하였다. 하지만 뇌파 신호가 연속적이라는 점을 고려하면 입력 데이터를 4초보다 짧은 시간 내에 분류하면서 정확도 또한 떨어지지 않는 것이 필요하다. 본 논문에서는 SPRT를 사용해 분류 정확도도 높고 분류에 걸리는 시간이 줄어들었음을 확인할 수 있었다.

### II. 본론

#### 가. 데이터셋 수집 과정

본 논문에서는 오픈 데이터셋인 Lee 데이터셋을 사용하였고, MOABB(mother-of-all-BCI-benchmarks)이라는 플랫폼에서 데이터셋을 다운로드하였다. 해당 데이터셋을 수집하는 과정에서 54명의 피실험자는 의자에 앉아 모니터에 나오는 무작위의 지시를 기다리고, 모니터에 MI 지시가 나타나면 그에 해당하는 상상을 한다.

그림 1을 참고하면, 피실험자가 한 번의 MI를 실행하는 4초 간격을 하나의 실험이라고 한다. 한 명당 400번의 실험을 진행하였고, 매 실험마다 무작위로 왼손, 오른손의 MI를 수행하였다. MI를 수행하는 피실험자의 뇌파는 62개의 전극으로 이루어진 장비를 통해서 수집되었다.

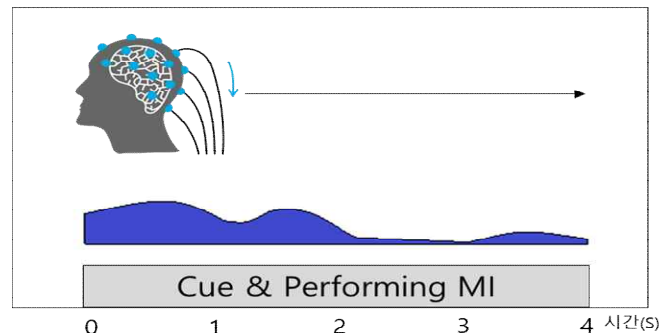


그림 1. 무작위 지시에 따른 4초 동안의 EEG 데이터 수집

#### 나. 데이터셋 구성

데이터셋으로는 54명의 피실험자에 대해 각각 400번의 실험을 수행한 뇌파 신호로 구성되어 있다. 하나의 실험은 모델의 입력 과정에서 총 64개의 토큰으로 나누어져 구성된다. 전체 데이터셋의 클래스는 왼손 또는 오른손으로 두 개의 클래스로 구성되어 있으며, 62개의 전극을 통해 수집된 이 데이터셋의 채널 수 또한 62개이다.

#### 다. 분류 모델 학습

위와 같은 데이터셋으로 트랜스포머 기반의 모델을 학습시켰으며, 이때 피실험자 의존적 학습을 위해서 10F-CV(10-Fold Cross-Validation) 방법을 사용하여 학습을 진행하였다.

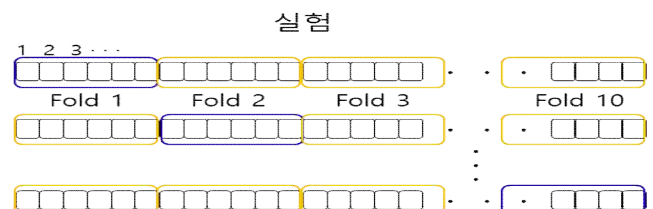


그림 2. 10F-CV 학습 방법

그림 2를 참고하면 10F-CV 학습은 하나의 실험 데이터틀 동일한 크기의 10개의 fold로 나누고, 이 중에서 1개를 테스트셋으로 사용하고, 나머지 9개를 트레이닝셋으로 사용하는 것이다. 1개의 실험에서 10F-CV를 사용해서 얻은 10개의 결과를 평균을 낸 것이 해당 실험의 결과이다.

#### 라. SPRT 개념

SPRT(Sequential Probability Ratio Test)는 가능한 빠르게 분류를 해내는 것이 목표이다. SPRT에서는 상한 한계점과 하한 한계점 설정이 필요하다. 본 논문의 경우 모델에서 출력된 결과가 상한 한계점에 도달하면 모델의 분류 결과가 오른손이고, 하한 한계점에 도달하면 왼손이다. 분류를 하고자 하는 결과가 한계점에 도달하는 시점에 분류를 종료하는 것이기 때문에 상한 한계점이 커지고, 하한 한계점이 작아지면 분류는 더 늦게 이루어지고 정확도는 더 높아진다. 본 논문에서는 상한 한계점을 1부터 0.51까지 줄이고, 하한 한계점을 0부터 0.49까지 늘리며 정확도를 분석하였다.

#### 마. MI 분류 정확도

위의 학습된 모델의 MI 분류 정확도를 구하기 위해 두 가지 방법을 사용하였다. 64토큰 길이의 데이터를 가지고 학습한 모델에 토큰의 길이를 1에서 64까지 변화시켜서 입력하였다. 한 명의 피실험자 당 400번의 실험 중에서 10F-CV로 학습하였기 때문에 40개의 실험에 대해서 평균으로 정확도를 계산하였다. 데이터셋의 라벨은 오른손이 1, 왼손이 0이기 때문에 모델의 출력이 0.5보다 크면 오른손으로 분류한 것이라 할 수 있다. 첫 번째 방법은 하나의 실험의 데이터에서 길이가 1부터 64까지인 토큰을 모델에 입력했을 때 나온 64개의 결과를 각각 0.5와 비교하여 오른손, 왼손으로 분류하였다. 이렇게 총 40번의 실험의 정확도를 평균을 구해 그래프를 얻었다.

두 번째 방법은 모델의 출력은 같지만 SPRT를 적용하여 정확도를 계산하였다. 40개의 실험에 대해서, 토큰의 길이와 분류 정확도를 평균을 구하였다. 총 50개의 한계점에 대한 분류가 끝난 시점의 토큰의 길이와 분류 정확도 그래프를 얻었다. 그림 3의 그래프는 노란색 그래프가 첫 번째 방법을 사용하여 정확도를 계산한 결과이고, 파란색 그래프는 두 번째 방법을 사용하여 정확도를 계산한 결과이다. 총 10명의 피실험자에 대해서 두 가지 방법으로 정확도를 계산한 결과, SPRT를 사용해 MI 분류 결과의 정확도를 구한 것이 더 높은 정확도를 가지는 경향이 있음을 알 수 있다.

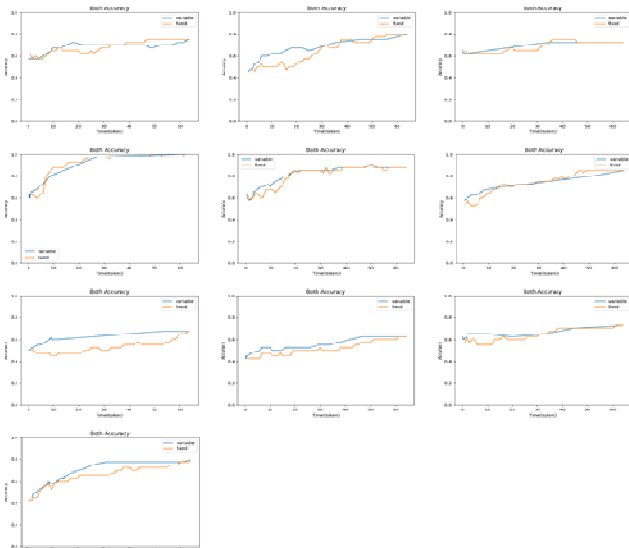


그림 3. SPRT의 적용 여부에 따른 정확도 차이

표 1을 참고하면, 입력 토큰의 길이가 5에서 25개까지로 짧을 경우, SPRT를 적용하여 정확도를 계산하였을 때가 그렇지 않았을 때보다 정확도가 높음을 확인할 수 있다.

표 1. SPRT 적용 여부에 따른 초기 분류 정확도

정확도 분석 \ 입력토큰	5	10	15	20	25
SPRT	62%	62%	73%	69%	77%
Fixed	55%	60%	63%	65%	67%

### III. 결론

본 논문에서는 SPRT의 적용 여부에 따른 MI 분류 정확도에 차이가 있었다. SPRT를 적용하면 한계점에 도달했을 때 분류가 끝나게 되었다. 이를 통해 4초가 아닌 더 짧은 길이의 데이터를 모델의 입력으로 사용할 때, 더 높은 분류 정확도를 보이는 것을 확인할 수 있었다. 본 논문에서는 피실험자 의존 실험의 경우 SPRT를 적용한 것으로, 피실험자에 따른 결과의 편차가 있었다. 더 나아가 본 논문에서는 피실험자 의존 실험의 경우 SPRT를 적용한 것이므로 피실험자 독립 실험에도 SPRT를 적용하는 것에 관한 연구를 기대할 수 있다.

### ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT명품인재양성 사업의 연구결과로 수행되었음 (IITP-2023-2020-0-01821)

### 참 고 문 헌

- [1] "Ece 830 Fall 2011 Statistical Signal Processing Lecture 9: Sequential Testing."
- [2] M.-H. Lee, O.-Y. Kwon, Y.-J. Kim, H.-K. Kim, Y.-E. Lee, J. Williamson, S. Fazli, and S.-W. Lee, "EEG dataset and OpenBMI toolbox for threeB paradigms: an investigation into BCI illiteracy," GigaScience, vol. 8, no. 5, May 2019