임상 적용을 위한 EEG 단채널 기반의 기계 학습 수면 단계 분류 모델*

이채린[1], 권기혁[2], 이소정[3], 김남재[4], 김종완[5†]
[1,3]삼육대학교 보건관리학과 학부생
[2]삼육대학교 인공지능융합학부 학부생
[4](주)글로벌비즈텍 대표
[5]삼육대학교 SW융합교육원 교수
ag0727@naver.com, 7lhyuk@gmail.com, leesojung0624@naver.com, knj7360@gbt2019.com, kimj@syu.ac.kr

Machine Learning Sleep Stage Classification Model Based on EEG Single Channel for Clinical Application

Chae-rin Lee[1], Ki-hyuk Kwon[2], So-jung Lee[3], Nam Jae Kim[4], Jongwan Kim[5]
[1,3]Dept. of Public Health, Sahmyook University
[2]Dept. of Artificial Intelligence Convergence, Sahmyook University
[4]Global Biz Tech Co., Ltd.
[5]Software Convergence Education Center, Sahmyook University

요 약

수면 장애는 성인의 약 30%가 경험하는 흔한 질환으로 수면 단계마다 다른 양상을 보인다. 현재 수면 단계 분류에 관한 연구는 활발하지만 기존 모델은 임상에서 사용할 만큼 예측 정확도가 높지 않다. 본 논문에서는 비렘수면 단계를 LS(Light Sleep)와 DS(Deep Sleep)로 분류하여 정확도를 높인 뇌파 단일 채널 기반 수면 단계 분류 모델을 제안한다. 해당 모델은 수면 단계 단순화를 통해 모델의 성능을 향상하였으므로 향후 직관적인 수면 질 파악 및 실시간 수면 모니터링 등 임상 환경에서 수면의 질 평가, 장애 진단, 치료 계획 수립에 기여할 것이다.

1. 서론

수면은 뇌, 심혈관, 위장관, 호흡, 면역, 내분비, 대사 등의 생체 기능을 안정적인 상태로 유지하도록 도와주며 수면의 질이 나쁠 경우 피로도 상승과 집중력 저하뿐 아니라 고혈압, 당뇨, 심혈관계 질환 등 각종 질병을 유발한다. 학습과 기억, 감정 조절 기능에도 관여하기 때문에 수면의 질을 향상하는 것은 중요한 문제이다.[1]

현재 수면의 질은 수면다원검사(PSG, polysomnography)에서 측정되고 미국 수면의학회(AASM, American Academy of Sleep Medicine) 매뉴얼에 따라 특정 수면 단계로 분류되는데, 이러한 분류 과정은 긴 시간이 소요되고 전문가의 주관성이 내포된다. 이와 같은 문제를 해결하기 위해 수면 단계 자동화에 관한 연구가 이루어졌지만, 예측 정확도가 균일하게 높지 않아 임상에서 사용하기 어렵다는 문제점이 있다. 따라서 본 논문에서는 임상 적용을 위해 NREM(non-rapid eye movement sleep) 단계를 얕은 수면(LS, Light Sleep)과 깊은 수면(DS, Deep Sleep)으로 분류하고 EEG(Electroencephalogram) 신호의 단일 채널(Fpz-cz)과 기계 학습 앙상블 모델인 부스팅 모델(GBM, Gradient Boosting Machine)을 사용하여 학습시킨 수면 단계 분류 모델을 제안한다.

2. 관련 연구

수면은 크게 REM(rapid eye movements) 수면과 NREM(non-rapid eye movement sleep) 수면으로 나눌 수 있고, 이는 빠른안구운동이 나타남과 나타나지 않음의 기준으로 분류된다. 미국수면의학회 가이드라인에 근거하여 NREM은 N1, N2, N3으로 분류되고 뇌파 활동과 생리적 변화에 따라 N1, N2는 얕은 수면(LS, Light Sleep), N3는 깊은 수면(DS, Deep Sleep)으로 간주된다.[2]

3. 본론

3-1. 모델 구축

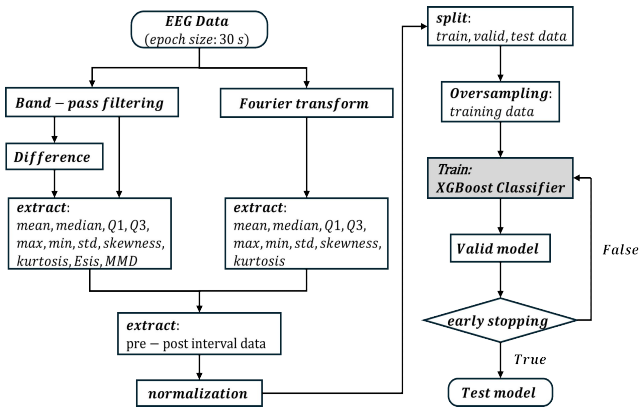
본 논문은 모델을 학습하기 위해 Sleep EDF Expanded Database의 EEG 신호를 사용하였다. 모든 신호는 100Hz로 샘플링되었으며 sleep-cassette의 Hypnogram만을 사용하였다. 제안된 모델에서는

* 본 연구는 2021년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음 (2021-0-01440).
† 교신저자(Corresponding Author),

Stage 1, 2를 묶어 LS로, Stage 3, 4를 묶어 DS로 구분하여 Sleep Stage를 [W, LS, DS, R]으로 분류하였다. 클래스 불균형 문제는 소수 클래스의 경계에 있는 샘플을 중심으로 새로운 데이터를 생성하는 방법인 BorderlineSMOTE으로 해결하였다.

모델의 학습에 들어가기에 앞서 Bandpassfiltering을 적용한 전처리를 수행했다. EEG 데이터를 주파수별로 필터링하였고, 뇌파 종류에 따라 델타파는 0.5~4 Hz, 세타파는 4~8 Hz, 알파파는 8~12 Hz, 베타파는 12~30 Hz, 감마파는 30~45 Hz로 대역을 구분했다.

모델의 학습은 GBM(Gradient Boosting Machine)을 사용하였다. GBM은 여러 개의 약한 학습기를 순차적으로 학습시켜 이전 모델의 잔여 오차를 최소화하고 예측 성능을 극대화한다.



(그림 1) 데이터 전처리 및 특징 추출

3-2. 특징 추출

3-2-1. 시간 영역

Bandpassfiltering을 적용한 EEG 데이터를 30초로 분할하여 통계적인 특징과 esis, mmd[3]를 추출하였다. 통계적인 특징에는 평균, 중앙값, IQR, 최댓값, 최솟값, 표준편차, 왜도, 첨도가 포함된다.

3-2-2. 주파수 영역

고속 푸리에 변환(FFT, fast fourier transform)을 이용하여 주파수 대역을 분리하였고 주파수 파워를 계산한 다음 통계적인 특징을 추출했다. 단계적으로 이루어지는 EEG 데이터의 특성을 고려하여 데이터의 이전 구간과 이후 구간의 일부 값을 추가하였다.

3-3. 정규화

Robust Scaling은 중앙값(median)과 사분위 범위(Interquartile Range, IQR)를 사용하여 이상값의 영향을 최소화하면서 데이터 스케일을 조정한다.

$$X_{\text{scaled}} = \frac{\text{X} - \text{median}}{\text{IQR}} \qquad (1)$$

X는 정규화되기 전 데이터 값으로, 수식 (1)에 의해 X_{scaled}로 정규화된다. median은 데이터의 중간값으로, 데이터 분포의 중심을 의미한다. IQR은 데이터의 1 사분위수 (25%)와 3 사분위수 (75%) 간의 차이로, 데이터의 중간 50%가 차지하는 범위이다. 수식 (1)에 의해 변환된 X_{scaled}의 분포는 median은 0, IQR은 1로 변환된다.

3-4. 실험 결과

수면 단계 분류 방식에 따른 모델 성능은 <표 1>, <표 2>와 같다.

<표 1> 기존 분류 방식 모델 성능표

		Precision (%)	Recall (%)	F1 – score (%)	Data Size
Sleep stage W		97.55	96.47	97.01	205,216
LS	Stage 1	73.58	71.34	72.44	67,722
	Stage 2	49.34	45.76	47.48	
DS	Stage 3	11.22	11.03	11.13	18,167
	Stage 4	4.33	6.59	5.23	
Sleep stage R		14.02	27.85	18.65	6,507

<표 2> 임상 분류 방식 모델 성능표

	Precision (%)	Recall (%)	F1 – score (%)	Data Size
Sleep stage W	97.95	97.08	92.52	205,216
LS	85.04	73.27	78.72	67,722
DS	52.69	68.28	59.48	18,167
Sleep stage R	33.60	63.61	43.97	6,507

위 성능표를 통해 NREM 단계를 임상에서 구분하는 방식인 LS와 DS의 분류가 기존 방식으로 분류했을 때에 비해 정확도가 높아졌음을 알 수 있다.

4. 결론

본 논문에서는 임상 적용을 목표로 EEG 단일 채널을 사용한 수면 단계 분류 모델을 구축하였다. 예측 정확도를 높이기 위해 데이터의 일정 구간을 추가하여 특성 개수를 늘렸고 모델이 수면 단계 경향성을 파악하게 하여 학습하는 데 어려움이 없도록 하였다. 또한 수면 단계 단순화는 모델 성능을 향상할 뿐만 아니라 직관적인 수면 질 파악 및 실시간 수면 모니터링 등 임상 환경에서 활용성을 높인다.

참고문헌

[1] National Center for Disaster and Trauma, http://www.mentalhealth.go.kr/portal/disease/diseaseDetail.do?dissId=32.

[2] 6 American Academy of Sleep Medicine (AASM) Manual for the Scoring of Sleep and Associated Events (2007).

[3] Aboalayon, Khald Ali I., Miad Faezipour, Wafaa S. Almuhammadi, and Saeid Moslehpour. 2016. "Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey and New Investigation" Entropy 18, no. 9: 272.