**Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device**

AI 데이터 연구단 서영석

Sleep Research Society 19.08.13

## **Abstract**

수면을 추정하는 웨어러블, 다중 센서, 장치 등은 일반화되었지만 이러한 장치에서 수면score

를 매기는데 사용하는 알고리즘은 오픈 소스가 아니고 엑세스하기 쉽지 않다. 결과적으로 임상 및 연구 응용 분야에서 유용성이 제한된다. 이 논문에서는 수면다원검사를 받는 참가자와 검사 이전 보행 기간 동안 착용한 Apple Watch에서 원시 가속도 데이터와 심박수를 수집했다. 이 데이터를 사용하여 기여도를 비교했다.

이 연구는 수면 및 수면 단계 예측의 정확도 개선을 위해 가속도 및 심박수 데이터를 분석했다.

## **Introduction**

이 연구의 주요 목표는 Apple Watch에 있는 MEMS 가속도계 및 PPG에서 원시 가속도 및 심박수 데이터를 수집하고 분류 방법을 사용하여 수면과 비수면을 구별하고 PSG와 비교하여 수면 단계를 결정한다. 두 번째 목표는 “clock proxy”용어의 통합이 모든 분류기의 성능에 어떻게 영향을 미치는지 평가하는 것이다. 마지막으로는 Apple Watch 가속도계 및 PPG를 넘어 알고리즘을 일반화하기 위해, MESA로부터 수집된 데이터셋으로 훈련된 모델을 테스트했다.

## **Methods**

질환,장애 유무,몸 상태 등을 조사한 뒤 연구 진행하였다. REM 수면 행동 장애 또는 폐쇄성 수면 무호흡증을 시사하는 PSG로 입증된 대상은 제외되었다.

Apple Watch의 가속도계 및 심박수 데이터에 액세스하기 위한 OW의 코드

<https://github.com/ojwalch/sleep_accel> 에서 온라인으로 확인

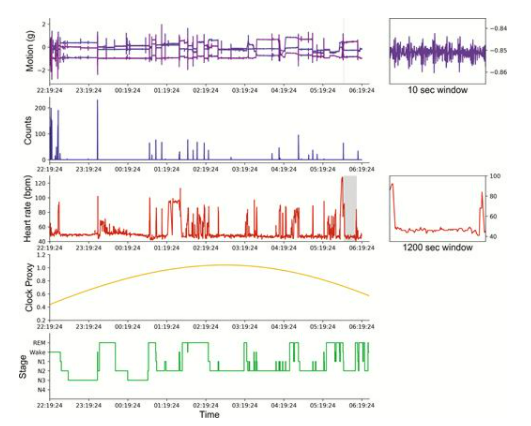
연구진행은 피실험자들은 8시간의 수면 동안 PSG로 모니터링되었고 AASM의 기술 사양에 따라 PSG를 진행했다. (이하 내용 생략)

PSG를 통한 수면 모니터링과 동시에 Apple Watch에서 원시 가속도 및 심박수를 기록하고 보안 서버로 전송했다.

Apple Watch는 x,y,z 방향의 가속도를 g(9.8m/s^2)단위로 측정하는 3축 MEMS 가속도계를 사용한다. 또한 심박수는 W의 등쪽 측면에 PPG가 있는 Apple Watch로 측정한다.

## **Analysis**

PSG매개변수(time in bed [TIB], total sleep time [TST], sleep onset latency [SOL], wake after sleep onset [WASO], sleep efficiency [SE], REM sleep minutes, and NREM sleep minutes)를 평가했다.



한 피험자의 수면에서 얻은 샘플 데이터이다.

Motion은 Apple Watch의 마이크로 전자 기계 시스템 가속도계의 원시 가속도(x,y,z)로 인해 표현

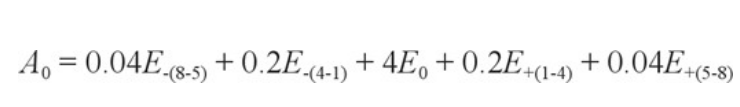
Apple Watch는 x,y,z 방향의 가속도를 g(9.8m/s^2)단위로 측정하는 3축 MEMS 가속도계를 사용한다..

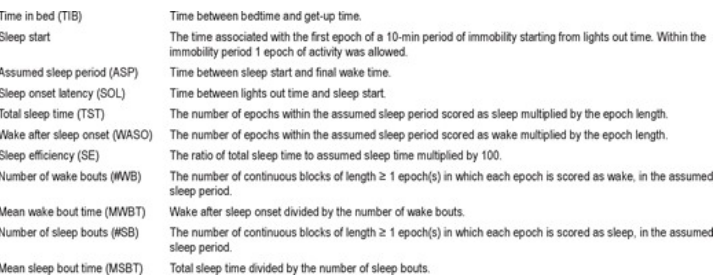
(활동 수, 원시 가속도에서 변환)

 가속도센서는 지표면을 중심으로 기울기, 가속도 등을 측정한다. 그러나 지표면에 대해 수직 방향의 가속도는 측정할 수 없다. 자이로스코프는 가속도센서로 측정할 수 없는 방위각을 측정한다

**Counts**는 아래 논문의 코드를 사용하여 활동 카운트로 처리, 여러 파라미터로 카운트하였다.

<https://academic.oup.com/sleep/article/36/5/781/2559074?login=true> – 논문 참고





**Heart rate**는 Apple Watch의 PPG(광혈류측정기)의 심박수 (심박수)

심박수는 W의 등쪽 측면에 PPG가 있는 Apple Watch로 측정한다

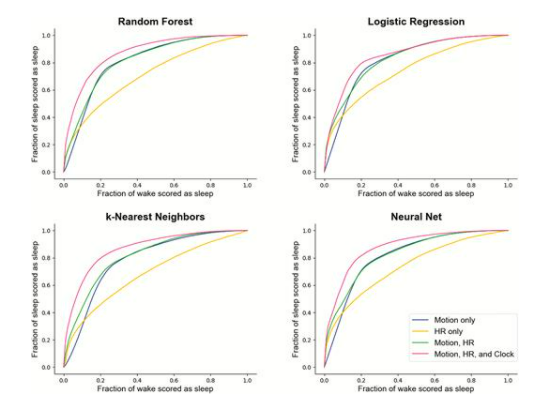
**Clock proxy**는 Apple Watch를 사용한 보행 녹음에서 예측

-두 가지 방식으로 결정된다.

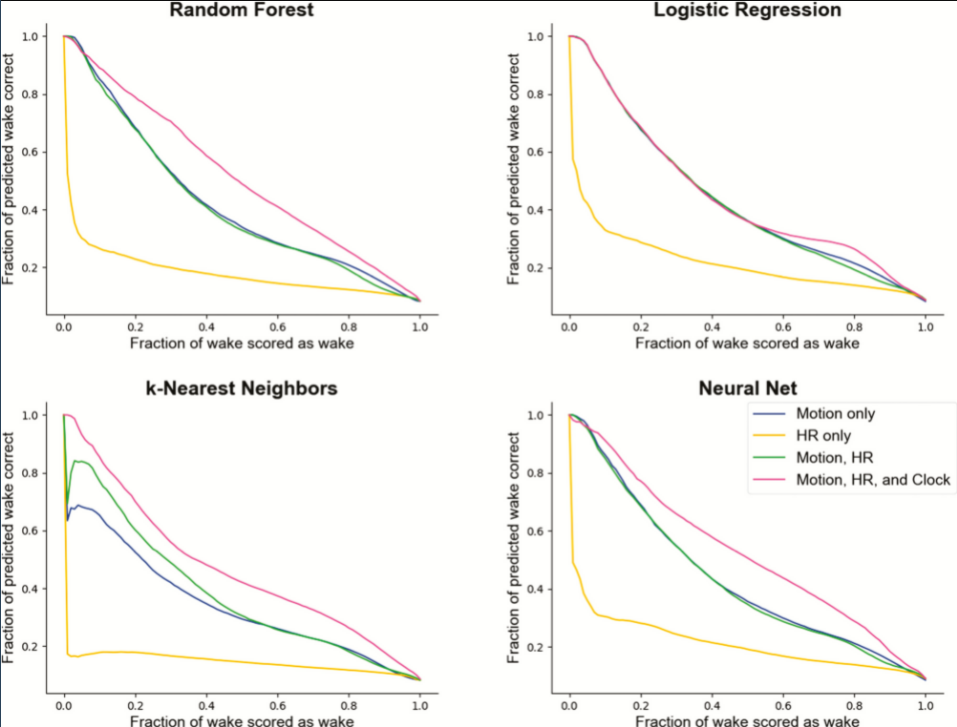
첫 번째 방법은 고정 코사인파를 사용하는 것이다. 이 파형은 녹화 시작 시간을 기준으로 이동하며 밤 동안 오르락 내리락했다. 클록 프록시 용어를 계산하는 이 방법은 입력으로 기록 시간만 필요하기 때문에. 검증된 수학적 모델을 사용해 시계 프록시 기능도 계산했다.

-healthkit(앱/프레임워크)을 통해 가져온 걸음 수 데이터로 휴식과 활동의 전형적인 일일 패턴을 추론하는데 사용되었다. 이렇게 24시간 clock proxy기능, 코사인 기능 등으로 일주기 clock proxy를 예측한다.

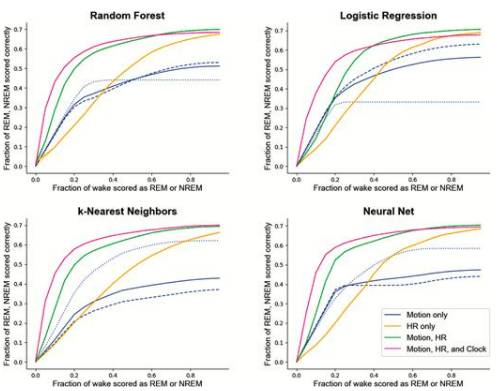
Stage는 수면 다원 검사에서의 기록이다.



수면과 비수면을 구별하기 위한 여러 분류기에 대한 ROC 곡선.



수면과 기상을 구별하기 위한 여러 분류기 및 기능에 대한 정밀 재현율 곡선



WAKE/NREM/REM 분류를 위한 여러 분류기 및 기능에 대한 ROC 곡선

motion only의 경우 점선은 각각 REM 및 NREM 정확도에 해당하고 실선은 평균이다.

다른 모든 경우 실선은 적절한 임계값을 선택하여 거의 동일하게 만들 수 있는 REM 및 NREM 정확도의 평균이다.

## **Motion feature**

가속도는 X,Y,Z방향의 가속도를 나타내는 3개의 벡터와 Apple Watch에서 나타내는 timestamp를 4번째 벡터로해서 약 50Hz에서 샘플링되었다.

## **Heart rate feature**

심박수는 Apple Watch의 PPG로 측정되었으며 몇 초마다 샘플링된 분당 비트 수로 반환되었다 이 신호를 1초마다 값을 갖도록 보안하고 평활화와 필터링으로 가우시안 필터 차이(σ 1 = 120 seconds, σ 2 = 600 seconds)로 컨볼루션하여 높은 변화의 기간을 증폭한다. 각 심박수 측정값과 수면 기간 동안의 평균 심박수 간의 차이를 정규화했다.

## **Algorithm training and selection.**

로지스틱 회귀,K-최근접 이웃,랜덤 포레스트 분류 기법 및 신경망(다층 퍼셉트론,MLP)이 서로 다른 분류 알고리즘을 비교할 때 후보 모델로 사용했다. <https://github.com/ojwalch/sleep_classifiers> (이 문서에서 분석을 수행하고 수치를 생성하는 데 사용된 모든 코드)

## **Validation against PSG**

처음에는 모든 교육 및 테스트가 Apple Watch 데이터 세트 내에서 수행되었다. 고려된 각 모델에 의한 수면 단계의 분류(수면/각성 또는 기상/NREM/REM)는 세대별 분석에서 PSG와 비교되었다. Epoch는 1970년 1월 1일(UNIX) 타임스탬프 이후 초를 사용하여 Apple Watch 기록과 정렬되었다.

모델은 Monte Carlo 교차 검증과 Leave-One-Out 교차 검증을 모두 사용하여 훈련/테스트를 했다.수면/각성 분류를 사용한 몬테카를로 교차 검증의 경우 데이터 세트를 훈련 세트(피험자의 약 70%)와 테스트 세트(약 30%)로 무작위로 50회 분할하고 기상/NREM/REM 분류를 위해 데이터 세트를 분할했다.(이하 내용 생략)

Monte Carlo 교차 검증을 사용하여 고려된 모든 기능 세트에 대한 각 알고리즘의 분류 능력은 ROC(수신기 작동 특성) 곡선과 정밀도-재현율 곡선을 사용하여 요약되었다. ROC 곡선은 임계값 매개변수를 변경하고 모든 임계값에서 서로에 대해 참 양성 및 거짓 양성 비율을 표시하여 생성 된다..

Apple Watch 데이터 세트를 사용한 수면/기상 분류에 대한 각 ROC 및 정밀도-재현율 곡선은 각 반복에서 생성된 데이터의 새로운 세분화와 함께 모든 50개 훈련 및 테스트 세트의 평균 성능을 나타낸다. 마찬가지로 WAKE/NREM/REM 분류에 대한 각 ROC 곡선은 20개의 모든 훈련 및 테스트 세트에 대한 평균 성능을 나타낸다. WAKE/NREM/REM 분류의 성능을 시각화하기 위해 1대 휴식 플롯도 생성되어 20개의 훈련/테스트 분할과 함께 보충 자료에 포함된다.

분류기 성능의 주제 변동성을 이해하기 위해 Leave-one-out 교차 검증이 사용되었다. 1개를 제외한 모든 과목에 대한 훈련과 나머지 과목에 대한 테스트 결과로부터 과목별 특이도, 민감도, 정확도에 대한 히스토그램을 구성하였다.

위 그림과 같이 조사한 모든 알고리즘에서 motion, heart rate 그리고 clock proxy 와 같은 피처들을 사용할 때 성능이 가장 좋았다.

## **Conclusion**

이 작업에서 로지스틱 회귀, k-최근접 이웃, 랜덤 포레스트 및 신경망의 네 가지 분류기를 조사했다. 분류기 방법은 기상과 수면을 구별하고 수면 단계를 구별하는 능력이 다르지만 이러한 차이는 특별히 두드러지지 않았다. 그러나 기능 포함은 성능에 상당한 영향을 미쳤다. 예를 들어, 모든 분류기의 AUC는 심박수만 사용할 때보다 심박수와 동작을 함께 취하면 크게 증가합니다. 또한 세로 데이터를 사용하여 알려진 수면의 일주기 제어를 활용하는 기능을 포함하면 성능이 측정 가능하게 향상됨을 볼 수 있다.