Sleep Prediction Algorithm Based on Machine Learning Technology

Kyungmee Park[1], Sangeun Lee[2], SeungMin Lee[3], Songlee Cho[2], Soohyun Wang[4], Seoungkyun kim[4], Eun Lee[1*]

1 Institute of Behavioral Science in Medicine and Department of Psychiatry, Yonsei University College of Medicine, Seoul, South Korea
2 Yonsei University Health System, Yonsei University College of Medicine, Seoul, South Korea
3 Division of information and computer science, Dankook University, Yongin, Kyungki-do, South Korea
4 Future Information & Network Architecture Lab, Korea University, Seoul, South Korea

*Correspond to leeeun@yuhs.ac

Poster #: P749

INTRODUCTION

- Insomnia is common and important psychiatric condition. Clinicians believe that maladaptive sleep-related factors such as inactivity during daytime, excessive time spent in bed, poor exposure to light, and high stress level contribute development of chronic insomnia. However, there's paucity of evidences what factors affect sleep quality.
- For knowing better of relationship between sleep and sleep related factors, we need to careful observation of them. Common ways of acquiring information about them are polysomnography, sleep diary, and actigraphy.
- One of the most crucial limitations of them is that prediction of sleep quality is wholly dependent on researchers. Traditional assessment methods cannot integrate sleep outcome with sleep-related factors by itself. In addition, there are no ways to assess stress-related index which is believed to affect sleep quality.
- From this study, we want to find the relationship between sleep quality and sleep-related factors including daytime activity, exposure to light, and stress-related index. Also, we tried to develop algorithm that can predict sleep quality by using machine learning technology.

METHODS

Participants

- Sixty-nine healthy participants were enrolled to our study.

Assessments

- We measured their sleep and sleep-related factors including daytime activity, exposure to light, and heart rate variability by using actigraphy and heart rate sensor (ActiGraph GT3X® and Polar H7®) for 2 weeks. We also gathered information of subjective sleep quality and alcohol or caffeine consumption by using sleep diary.
- Daytime activity was indicated by the acceleration of the x-, y-, and z axis. As a unit of momentum, metabolic equivalent of task (MET) was used and sorted as following: vigorous (MET≥6), moderate (3≤MET<6), and light (MET<3) activity. Exposure to light was divided into outside light and inside light to analyze (outside ≥1000 lux).
- For the stress-related index, we used heart rate variability, which is easy to assess and proven its relatedness to stress. RMSSD(root mean square successive difference of inter-beat of heart interval) and pNN50(percentage of successive normal sinus intervals of heart > 50ms.) was used for heart rate variability.
- Data was subdivided into 3 periods: wake-up to noon (P1), noon to 18:00 (P2), and 18:00 to bedtime (P3).
- For the machine learning analysis, we randomly divided data into 2 subset: training set(80%) and test set(20%).

Statistical analysis

- We developed 3 prediction model for sleep quality by using machine learning techniques. First, we performed logistic regression analysis which was optimized enough to maximize accuracy. We also used random forest (RF) model and multi-input one-dimensional convolutional neural network (1D-CNN) model for develop sleep prediction algorithm.
- Sleep quality was determined by sleep efficiency, divided as good or bad sleep. (good sleep indicates sleep efficiency ≥90%).

Table 1. Demographic and clinical characteristics

Variable	Total (n=69)
Age, years (mean ± SD[1])	30.8 ± 8.4
Sex, female (n, %)	43 (62.3)
BMI[2] (mean ± SD)	22.8 ± 3.5
Education, years (mean ± SD)	17.10 ± 1.9
Regular Alcohol consumption[3] (n, %)	41 (59.4)
Current smoking (n, %)	5 (7.2)
Regular caffeine consumtion[4] (n, %)	62 (89.9)
Sleep with partner (n, %)	43 (62.3)
PSQI[5] (mean ± SD)	6.7 ± 2.3
ISI[6] (mean ± SD)	5.9 ± 4.1

[1]SD : Standard deviation. [2]BMI: Body mass index. [3]Alcohol consumption more than 300cc/week
[4]Caffein consumption more than 1 cup/day [5]Pittsburgh Sleep Questionnaire Index [6]Insomnia severity index

RESULTS

Factors associated with sleep quality

- By ten times of optimized logistic regression, vigorous activity during P1, total activity during P1, exposure to light during all periods showed significant correlations with sleep quality.
- Among exposure to light, exposure to outside light showed significant correlations with sleep quality.
- No difference was found between the results with and without sleep diary.

Table 2. Logistic regression analysis of factors affecting sleep quality

	Weight[1]	
Variables	With sleep diary	Without sleep diary
Sex	0.35	0.21
BMI	0.72	None
Alcohol	0.11	None
Caffeine	0.59	None
Recuperation	0.74	None
Nap	0.49	None
Moderate activity (P1[2])	1.19	1.34
Vigorous activity (P1)	**7.31**	**6.94**
Total activity (P1)	**4.76**	**4.62**
Outside light (P1)	**15.27**	**14.84**
Total light (P1)	**21.37**	**22.64**
Moderate activity (P2[3])	0.08	0.40
Vigorous activity (P2)	2.05	2.57
Total activity (P2)	2.91	3.30
Outside light (P2)	**12.49**	**10.56**
Total light (P2)	**13.74**	**11.60**
Moderate activity (P3[4])	3.77	3.19
Vigorous activity (P3)	1.03	2.50
Total activity (P3)	2.86	3.86
Outside light (P3)	**7.04**	**6.72**
Total light (P3)	**5.97**	**5.41**
Average heart rate	0.20	0.51
RMSSD[5]	2.53	2.42
pNN50[6]	1.44	1.10

[1]Presented in absolute value, rounded at the third decimal place. [2]P1: wake-up to noon. [3]P2: noon to 18:00. [4]P3: 18:00 to bedtime. [5]RMSSD: root mean square successive difference of inter-beat of heart(RR) interval. [6]pNN50: percentage of successive normal sinus intervals of heart(NN intervals) > 50ms.

Accuracy rate of 3 sleep prediction models

- With sleep diary, the accuracy rate was 77.5% in logistic regression model, 83.1% in RF model, and 87.2% in 1D-CNN model (94.0% in training set).
- Without sleep diary, the accuracy rate was 79.2% in logistic regression model, 83.6% in RF model, 87.2% in 1D-CNN model (93.4% in training set).

Figure 1. Accuracy rate of 3 models without sleep diary

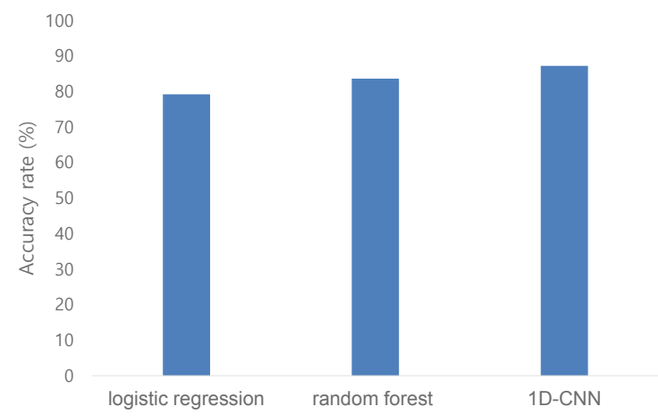
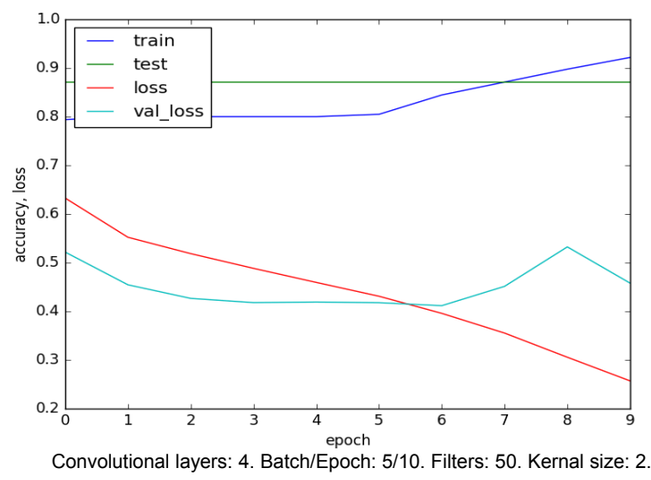


Figure 2. Loss and accuracy curves of 1D-CNN model predicting sleep quality without sleep diary



Convolutional layers: 4. Batch/Epoch: 5/10. Filters: 50. Kernal size: 2.

CONCLUSIONS

- The results suggest that morning activity (especially vigorous) , and exposure to total and outside light during daytime are important to sleep quality. The algorithm made by 1D-CNN model predict sleep quality accurately, which is comparable to the random forest or logistic regression.
- This result can enable more objective and accurate sleep prediction by automation of integration and interpretation of the data.

This research was supported by the Basic Science Research Program (No. 2017R1A2B3008214) through the National Research Foundation of Korea funded by the Ministry of Science, ICT, and Future Planning and the Korean Mental Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea [HM15C0995].