Identifying NREM Sleep Stages in Consumer Wearables Proposal

Daniel J. Fasciano, BS[1], Matthew S. Shimko, BS[1], Tran N. Ton, BS[1], An T. Butchers, BS[1]
[1]Georgia Institute of Technology, Atlanta, GA, USA

Abstract

In this project, we are working with Apple Watch health data such as acceleration and heart rate gathered from June 2017 to March 2019 by the University of Michigan. The data come from 31 subjects. We are building a sleep stage classifier based on an existing work by Walch et al[4]. While the original work has a Wake vs Sleep and Wake/NREM/REM sleep classifier, we plan to develop more classification such as Wake/N1/N2/N3/REM or Wake/N1+N2/N3/REM to bring more benefits to sleep studies.

Motivation

Sleep is an important part of people's life and sleep quality can tell a great deal about brain health. Research has found a strong link between abnormal sleep and brain disorders such as brain tumors[1, 2]. Therefore, people nowadays care more and more about their personal sleep quality, hence the development of several kinds of wearable devices to track sleep data[3, 4]. Studying sleep data used to deal with big machines, many different sensors, probes, and time-consuming manual sleep staging processes[5]. These old, complicated study methods not only carry the risks of human errors but also the high chances of inaccurate input data due to uncomfortable sleep because of attached devices on human subjects. With the development of technology, collecting sleep data is becoming easier than ever. The new compact wearable devices can be electroencephalography (EEG) devices that track the electrical activity of the brain; or they may be smart fitness watches such as Apple watches that gather heart rate, movement, steps, and other health and fitness data.

We should care about utilizing wearable devices' data because this data is easy to get and becoming quite popular. However, the raw data does not yield sleep stage labels without being processed by certain algorithms. With the help of machine learning algorithms, we can categorize the sleep stages from smart wearable devices' data. This will lead to better utilization of this huge source of data for sleep studies and easier sleep analysis. As a result, brain diseases linked to sleep disorders can be prevented and cured more efficiently and quickly.

Literature Survey

During a good night sleep, our bodies typically go through multiple sleep stages. There are two basic types of sleep: rapid eye movement (REM) and non-REM sleep. Non – REM stage is further divided into three different stages N1, N2, and N3. Identification of sleep stages is a crucial step in determining sleep disorders as well as other studies to improve sleep quality. Traditionally, to identity sleep stages, researchers would first monitor participants' physiological signals such as EEG while they are sleeping. The recorded signals will then be manually observed and classified into sleep stages by well-trained sleep technicians. Since this process are often time-consuming and labor-intensive, there is an increased need to automate the sleep stages identification from sleep monitoring data. Many researchers have attempted to automate this process using machine learning techniques. Biswal et al[7] trained a deep recurrent neural network (RNN) using expert-defined features to automatically annotate EEG epochs with the five sleep stages - wake, REM, N1, N2, and N3. Tsinalis et al[1] proposed using convolutional neural network (CNN) integrates with feature extraction to effectively score sleep stage based on EEG without requiring prior domain knowledge. Alternatively, Hassan and Subasi[8] used bootstrap aggregating (Bagging) classifier to learn sleep EEG annotating. The EEG signals were first decomposed into sub-bands, from which four statistical moments - mean, variance, skewness, and kurtosis, are extracted. Zhao et al[5], on the other hand, utilized data from radio technologies, which is a less expensive and more comfortable way to monitor sleep, to determine sleep stages. They combined CNN and RNN models to learn sleep stages from radio frequency data.

While the above studies achieve promising results, they used data that are extracted from expensive and single-purpose devices. In many cases, the study might even require the subjects to spend some nights at sleep laboratories to undergo polysomnogram (PSG). These inconvenience setups limit the amount of data available to train machine learning models. Nowadays, the uses of smart watches like Apple Watch and Fitbits have become more mainstream. Beside their function of showing time and other cool apps, most of these watches are equipped with advanced suite of sensors that constantly contact users' wrist to deliver health data. These data could be advantageous in training sleep stage classification models since they are more readily accessible. In addition, since the study population using smart watches would include more diverse participants (e.g.: people wear Apple Watch could be different in ages, genders, races, with or without sleep disorders), the training examples would be more diverse which would help train supervised machine learning models more effectively. To take advantage of those data, Goldberger et al[9] proposed a method to classify sleep stages using microelectromechanical systems (MEMS) raw acceleration and photoplethysmography heart rate data derived from Apple Watch. Multiple classification algorithms such as Random Forest, Logistic Regression, k-Nearest Neighbors and Neural Net, were applied to the datasets to learn sleep staging. The results were then validated using ground truth PSG data to choose the best model. Although this is an inspiring study which shows the ability of using consumer wearable devices to determine sleep stages, the study only classifies sleep stages into Wake, REM, and Non-REM. To fill in the gap, our project will expand this research by building classification models that can use smart watch sensor data to predict the conventional five sleep stages - Wake, REM, N1, N2, and N3. We believe five-stage sleep classifications would bring more benefits to sleep studies. Apple Watch and other smart devices include three MEMS accelerometers for each axis - x, y, z, which help detect movement in any direction. We will transform these data into one of the features to be used by converting the MEMS accelerometer data into traditional actigraphy movement counts using the method described by Lindert and Van Someren[6].

Data

There are multiple datasets for this experiment. The Apple Watch produced raw sensor data. These are the data our project is trying to map to sleep stages. The other dataset is ground truth polysomnography data.

<u>Smartwatch Data</u>

First, the Apple Watch sensors used in this dataset are accelerometers, which capture motion, and photoplethysmography, which captures heart rate. The accelerometer that captures motion in three-dimension space, is measured in g, and is generally sampled at 50 Hz. Time, using Unix time, is also captured. The photoplethysmography data is heart beats per minute as sampled 10-13 times per minute. These are direct measurements to be used in calculating sleep stages.

Secondly, a subject's circadian rhythm needs to be known. One could use a cosine wave that starts at the beginning of the subject's sleep. Alternatively, one could attempt to estimate their circadian rhythm by using other data the watch produces. Ideally, the watch would sense light and the sensor data could be converted into a rhythm using a researched formula. However, the model used in the experiment does not do that. Instead, step data is collected, and a light value was estimated based on steps and time of day. Steps are reported every 10 minutes**.**

<u>Polysomnography Data</u>

While the subjects wore Apple Watches, they also underwent polysomnography recording. The PSG data were captured in accordance with American Academy of Sleep Medicine specifications and when scored is considered a very accurate classification of sleep stages. The raw PSG data will not be used in this research. Instead, we will use the resulted polysomnography data that was scored with a 0-5 sleep stage every 30 seconds.

<u>Data processing</u>

The features for this project will be heart rate, motion, and circadian rhythm. Since the polysomnography data were scored in 30 second intervals, project features will need to be converted to a consistent time interval.

Since heart rate is time-series data, we will need to perform either running averages or exponential smoothing in order to migrate them from current values to regular intervals.

Motion data is time-series data that will need to be converted into actigraphy movement counts using the method described in Lindert and Van Someren [6].

Circadian rhythm will be estimated with a cosine wave at the start of sleep. If time permits, steps may be used to perform the estimate.

Approach

Our approach is to take the Apple Watch data and process it into a set of features, which will be used to train a classifier which can take that data and predict for each time interval what stage of sleep the subject was in. As we are not sure which algorithm will be best for judging this data, we will need to split the data into test, train and validation sets in order to train many different models and compare them. Some Machine Learning algorithms we plan to train are decision trees, random forest, support vector machines, and K-nearest neighbors (KNN). However, it might be necessary to consider the time series elements of the data, in which case other types of algorithms can be used, such as KNN with dynamic time warping, time series forest classifier, or Random Interval Spectral Ensemble. These methods, instead of just treating each unique time point as a record to classify, can consider time elements as part of the data and incorporate the trends and cycles which can occur over time, and are especially relevant to sleep related data.

The classifier will return the probability that a given record belongs to each possible stage, which will allow us to use logarithmic loss as an evaluation metric.

As our main goal is to properly categorize the watch data to correctly identify the various sleep stages, we will need to use a method to determine how accurately it is able to correctly classify what the stage a subject is in during a specific time period. In order to do this, we can use logarithmic loss. Some metrics, such as AUC or accuracy will be difficult with this data set, as it is not a binary classifier so AUC cannot be used, and as the distribution of time spent in sleep stages is not very even, accuracy would not be as helpful. In order to use log loss as the method of evaluation, we need to make sure our classifier returns a probability of each stage.

Experimental Setup

Most complex and memory intensive computation will be performed in an AWS environment, using EMR and the Spark framework to perform necessary data preparation and transformation. This will take the raw data provided and generate the features that will be used in training.

We plan to use Python libraries, most likely Scikit-Learn, to train the data processed by the above step in a local environment.

Timeline

Table 1 below provides a list of our planned tasks. Generally, every step in the project execution will be carried out within a week. There will be two milestones when the project draft and the final deliverables will be turned in toward the end of April and early May.

Table 1. Timeline

Timeline	Tasks	Due	Deliverables
03/15/2021 - 03/21/2021	Define cohort, target, features + Split data into training, validation		
03/22/2021 - 03/28/2021	Clean and process data		
03/29/2021 - 04/04/2021	Develop and implement the modeling pipeline		
04/05/2021 - 04/11/2021	Evaluate the model candidates on the performance metrics + Start Project Draft		
04/12/2021 - 04/18/2021	Interpret the results from the models + Compose Project Draft	04/19/2021 8 AM EST	Project Draft
04/19/2021 - 04/25/2021	Clean-up / Wrap up the Code		
04/26/2021 - 05/02/2021	Make Presentation + Compose Final Paper	05/03/2021 8 AM EST	Code + Presentation + Final Paper

Conclusion

With the development of technology and big data, various topics in health care are being studied more efficiently. Sleep care is an important part of well-being; hence many new technical devices have been developed to track sleep as well as well-being activities. Researching the data gathered by these devices will contribute to this trendy way of studying health care. We hope this work can produce a well-performing classifier and obtain more sleep stages labeled. As a result, more detailed and accurate sleep information will be available and help in the work to prevent and treat sleep disorders.

References

1. Tsinalis O, Matthews PM, Guo Y, Zafeiriou S. Automatic sleep stage scoring with single-channel EEG using convolutional neural networks [Internet]. arXiv [stat.ML]. 2016. Available from: http://arxiv.org/abs/1610.01683
2. Wulff K, Gatti S, Wettstein JG, Foster RG. Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. Nat Rev Neurosci. 2010;11(8):589–99.
3. Samuel N, So E, Djuric U, Diamandis P. Consumer-grade electroencephalography devices as potential tools for early detection of brain tumors. BMC Med. 2021;19(1):16.
4. Walch O, Huang Y, Forger D, Goldstein C. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. Sleep [Internet]. 2019;42(12). Available from: http://dx.doi.org/10.1093/sleep/zsz180
5. Zhao M, Yue S, Katabi D, Jaakkola TS, Bianchi MT. Learning sleep stages from radio signals: A conditional adversarial architecture. In 2017. p. 4100–4109,.
6. Lindert BH, Ej VS. Sleep estimates using microelectromechanical systems (MEMS. Sleep. 2013;2013;36(5):781-789:5665 2648.
7. Biswal S, Kulas J, Sun H, Goparaju B, Westover MB, Bianchi MT, et al. SLEEPNET: Automated sleep staging system via deep learning [Internet]. arXiv [cs.LG]. 2017. Available from: http://arxiv.org/abs/1707.08262
8. Hassan AR, Subasi A. A decision support system for automated identification of sleep stages from single-channel EEG signals. Knowl Based Syst. 2017;128:115–24.
9. Goldberger A, Amaral L, Glass L, Hausdorff J, Ivanov PC, Mark R, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220.