Sleep Duration and Quality Assessment using Smartphones

by

Yuxiao Peng

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the Degree of Master of Science

in

Bioinformatics and Computational Biology

May 2021

APPROVED:

Professor Emmanuel O. Agu, Major Thesis Advisor

Professor Dmitry Korkin, Head, BCB Program

Abstract

Sleep problems affect millions of people globally, presenting a tremendous burden on society and people. Sleep monitoring is important to understand patient-specific problems and to guide treatment. However, the traditional methods of sleep quality detection are inconvenient, costly, and time-consuming. In this research, we conducted a user study to investigate if sensors built into smartphones can be utilized for passive sleep duration and quality monitoring. Smartphone sensor and 'ActiGraph' ground truth data were simultaneously collected from 6 subjects. The data were preprocessed data to remove errors and inconsistencies. The preprocessed data was extracted into seven types of features that were then used to train machine learning classification models. In rigorous evaluation, gradient boosting machine learning algorithm performed best, achieving an accuracy of 88.08%. The final model was used to create a tool that will extract and analyze patient's sleep features from smartphone sensors data. In the future, this model can be used to develop a smartphone App to monitor users' sleep quality automatically.

Table of Content

ACKNOWLEDGMENTS

I would like to offer my special thanks to Prof. Emmanuel O. Agu for his guidance and support when my thesis reached a challenging situation. His broad experience in research helped me overcome many difficulties.

I would also like to thank all the volunteers who take their time to wear the ActiGraph GT9X watch in our study and help me collect the sleeping data. Without their assistance, my work would not have been able to come to an end.

I would like to express my gratitude to the Bioinformatics and Computational Biology program at WPI for allowing me to finish my Master of Science degree here.

Last, but not least, my special thanks are extended to my parents for their support and encouragement throughout my study.

LIST OF TABLES

Table **Page**

LIST OF FIGURES

1. Introduction

1.1 Background and Motivation

The quality of sleep is an important factor in human physical and mental health. Millions of young people are suffering from chronic partial sleep deprivation or sleep disorder around the world [1]. According to the American Academy of Sleep Medicine(AASM), sleeping disorder is classified into four main parts [18 (See Table 1)]:

Table 1. Four Main Classification of Sleeping Disorders				
Sleeping Disorder	Examples			
DYSSOMNIAS	Hypersomnia	Insomnia	Obstructive sleep apnea syndrome	
PARASOMNIAS	Confusional arousals	Sleep-walking	Sleep-talking	Nightmares
MENTAL, NEUROLOGIC, OR OTHER MEDICAL DISORDERS	Psychoses	Dementia	Nocturnal cardiac ischemia	
PROPOSED SLEEP DISORDERS	Subwakefulness syndrome	Short sleeper	Long sleeper	

Table 1. The 4 main classifications of sleeping disorders. The causes of sleeping disorders vary differently, such as environments, psychological stress, eating habits, age, and medicines.

Sleep as a symptom of serious ailments: These sleep problems are often associated with or symptoms of many serious diseases, such as diabetes, cardiovascular disease, obesity, and depression [2][3][4]. More seriously, these even affect memory and cognitive functions [5]. Sleep disorders are also always associated with Traumatic Brain Injury (TBI) population [21]. 50% of people suffered from some form of sleep disturbance after a TBI and 25–29% had a diagnosed sleep disorder (insomnia, hypersomnia, apnea) [21].

Prevalence of sleep disorders: Insomnia is the most common sleep disorder in the world. Insomnia is defined as difficulty of falling asleep or staying asleep, waking up too early, and not being able to get back to sleep. Insomnia affects 40% people worldwide, and is a significant cause of morbidity and mortality [20]. Therefore, sleep monitoring has broad market prospects and potential economic benefits.

The need for sleep monitoring: As mentioned above, sleep disorders have become a worldwide health problem and are often a sign of serious diseases. Sleep monitoring can be used to understand patient-specific problems and remind those who are not aware of their own sleep problems. Furthermore, early detection of sleeping disorders can also help in the discovery of serious diseases related to sleeping disorders in advance.

Challenges in sleep monitoring: Although sleeping monitoring is important, it also faces some challenges. First, the physical state of a person lying in bed before sleeping and falling asleep is similar. It's difficult to tell when a person really fell asleep. Second, people will wake up few times during sleep, especially those with sleep disorders. How to accurately distinguish these awakenings through external mobile phone sensors is very difficult because people always awake unconsciously during two sleep periods, and their physical state does not change significantly.

1.2 Measuring Sleep

The Pittsburgh Sleep Quality Index(PSQI): is a self-rated questionnaire that assesses sleep quality and disturbances over a one-month time interval [11]. It asks a set of 19 obligatorily-fill questions and 5 selectively-fill questions determined by the condition that if you have a roommate/bed partner (see Table 2). PSQI set a metric to measure these questions and give a long-term sleep scoring. The metric scores 7 components of the subject's sleep: 1) sleep quality, 2) sleep latency, 3) sleep duration, 4) habitual sleep efficiency, 5) sleep disturbances, 6) use of sleeping medication, and 7) daytime dysfunction. And the summary of these 7 scores is the Global PSQI score. But when looking through the questionnaire, we can find that the answers to these questions are too subjective and lack accurate measurements of some questions, such as question 2 (how long (in minutes) does it usually take you to fall asleep each night), because no one would fall asleep while looking at his watch to count the minutes. In my opinion, PSQI is only a rough self-check method which guides you to calculate your sleep; it is far from being precise.

Table2. The Pittsburgh Sleep Quality Index [11]				
1. During the past month, what time have you usually gone to bed at night?				
2. During the past month, how long (in minutes) has it usually taken you to fall asleep each night?				
3. During the past month, what time have you usually gotten up in the morning?				
4. During the past month, how many hours of actual sleep did you get at night?				
5. During the past month, how often have you had trouble sleeping because you...	Not during the past month	Less than once a week	Once or twice a week	Three or more times a week
a. Cannot get to sleep within 30 minutes				
b. Wake up in the middle of the night or early morning				
c. Have to get up to use the bathroom				
d. Cannot breathe comfortably				
e. Cough or snore loudly				
f. Feel too cold				
g. Feel too hot				
h. Have bad dreams				
i. Have pain				
j. Other reason(s), please describe:				
6. During the past month, how often have you taken medicine to help you sleep (prescribed or "over the counter")?				
7. During the past month, how often have you had trouble staying awake while driving, eating meals, or engaging in social activity?				
	No problem at all	Only a very slight problem	Somewhat of a problem	A very big problem
8. During the past month, how much of a problem has it been for you to keep up enough enthusiasm to get things done?				
	Very good	Fairly good	Fairly bad	Very bad
9. During the past month, how would you rate your sleep quality overall?				
	No bed partner or room mate	Partner/room mate in other room	Partner in same room but not same bed	Partner in same bed
10. Do you have a bed partner or roommate?				

Polysomnogram (PSG): PSG is a set of sensors that are used in diagnostics of sleep disorders. A PSG typically requires the recording of multiple sensors including electroencephalography (EEG), electromyography (EMG), electrocardiography (ECG), heart rate, respiratory effort, airflow, and oxygen saturation [10] (See figure 1). The sleep summary report of PSG often consists of onset of sleep, sleep onset efficiency, sleep stages, any breathing irregularities, Arousals, and cardiac rhythm abnormalities. Sleep Onset Latency(SOL) is the onset of sleep from the time the lights were turned off. Sleep Efficiency(SE) is the number of minutes of sleep divided by the number of minutes in bed. Any breathing irregularities are mainly apneas and hypopneas. Apnea is a complete or near-complete cessation of airflow for at least 10 seconds followed by an arousal and/or 3% [22] oxygen desaturation; hypopnea is a 30% or greater decrease in airflow for at least 10 seconds followed by an arousal and/or 4% oxygen desaturation.[23] "Arousals" are sudden shifts in brain wave activity [23].

Figure 1: Patient monitoring using a Polysomnogram (PSG)

ActiGraph: is a wearable accelerometer-based biosensor designed for Scientific research use. ActiGraph watch uses its 3-axis accelerometer to precisely detect movements of the user. It can detect sleep accurately based on the movement of its monitors. Two built-in algorithms Cole [13] and Sadeh [14] are available in actigraphy to score sleep. The average accuracy of predicting sleep/wake state is around 90%

(reported 88% in [13] and 91-93% in [14]). Actigraphy watch can measures Sleep Latency, Total Sleep Time, Wake After Sleep Onset, Arousals and Sleep Efficiency. In this thesis, we use ‘ActiGraph GT9X Link’ watch showed here.

Figure 2: Actigraph GT9X Link smartwatch for activity and sleep monitoring

1.3 Limitations of Prior Sleep Monitoring Approaches:

As mentioned above, sleep disorder is a common illness all around the world. It is also a threat to many serious diseases like diabetes, cardiovascular disease, obesity, and depression [2][3][4]. Moreover, insomnia brings a tremendous burden on society and economy [20]. Therefore, early recognition of the sleep disorder can reduce the costs associated with the condition, as well as possibly prevent other illnesses [20].

But nowadays, a personal, convenient, and economic sleep detection approach is not available to people. Polysomnography (PSG) is the primary clinical method for sleep monitoring [6]. Due to its use of multiple sensors, utilizing of PSG is usually limited to clinical. Actigraphy is an alternative to assess sleep and wakefulness based on body movement [7]. It is generally a watch worn on the participant’s non-dominant wrist, recording people’s movements. The data can later be read to a computer and analyzed to get the sleep quality of the people. Actigraphy allows the patient to be movable and to continue her or his normal routines while the required data are being recorded in his or her natural sleep environment. However, both of the methods require additional sensors to be worn on people and they may be potentially invasive to the human being.

1.4 Opportunity: Smartphone sensing for passive health monitoring

The mobile phone has become an inseparable part of our lives. Google's consumer barometer shows that 83% of people in America use a mobile phone/smartphone in 2020 [8]. It is no longer a tool of connecting people, but an intelligent apparatus which has gone through an extremely fast evolution in the last decade. It can work as a butler to monitor our heart rate, count our movement, and manage our schedules. A recently listed smartphone on the market has several sensors which can monitor every aspects of our daily life:

1.Accelerometer: The accelerometer detects acceleration, vibration and tilt to determine the movement and precise direction along the three axes. The app uses this smartphone sensor to determine whether your phone is portrait or landscape. By using this sensor's data we can detect whether the person is moving or static. It is a very important factor of judging a person is sleep or not.

2.Gyroscope: The gyroscope also provides azimuth details and directions, such as up/down and left/right, but the accuracy is higher. It can accurately detect the tilt angle of the device. Same as the accelerometer, gyroscope sensor contributes a lot in detecting person's sleep.

3.Magnetometer: It can detect magnetic fields around the smartphone, which can help the app to find out the northward. Recent study [27] showed direct evidence that human's brain can detect magnetic field and take sensitive neural response to it.

4.GPS: Global Positioning System (GPS) units in smartphones connect with the satellites to determine the precise location of the smartphone. With this sensor we can get the location of the people and his movement across the map.

5.Proximity Sensor: A proximity sensor can find out how close the phone is to an outside object. In this study it can help detect the situation of the smartphone.

6.Ambient Light Sensor: The light sensor detects the lighting levels around the environment. We can also use it to detect the lightness around the person to help judge the person's sleeping or not. The lightness is also a factor that can influence a person's sleep quality.

7.Microphone: The microphone is a sound sensor that detects and measures the loudness of sound. Acoustic data is a very important factor that can help smartphone to detect a person's sleep. Obviously, acoustic data is also a factor which can influence a person's sleep.

8. Barometer: The barometer measures the air pressure. Its data can numerically reflect changes in the weather, which is also an important factor affecting human's sleep.

1.5 Thesis Goals

The goal of this research is to investigate machine learning classification of sleep that could be used in the future to develop a mobile app to passively assess the sleep of smartphone users. Using a smartphone to gather sleep data and score the sleep quality is a convenient and ideal way to help people know their sleeping condition.

To achieve the goal, simultaneously collected both mobile phone sensor data and Actigraphy watch data on 6 participants. Sleep features were extracted from the mobile phone data and used the Actigraphy data as labels. Thereafter, the extracted smartphone sensor features and ActiGraph labels were used to train a machine learning model to predict the sleep attributes and overall sleep quality of users.
Finally, the sleep machine learning model will be programmed into the smartphone sleep app in the future, which can predict the user's sleep quality using the collected data by smartphone. Users of this passive smartphone app will not be required to wear an Actigraphy watch to have their sleep analyzed.

The sleep measures that will be classified and investigated based on Actigraphy ground truth data are showed in Table 3. below.

Table 3. ActiGraph Watch Variables and Their Definition	
Variable	**Definition**
Sleep Onset Latency (SOL)	Time elapsed between when the subject lies down and goes to sleep
Time Awake prior to Rising (TWAK)	Time elapsed between when the subject wakes up and gets up from bed
Number of Awakenings (NWAK)	The number of times subject woke up within a sleep period
Wake time After Sleep Onset (WASO)	Total duration of time awake within sleep period
Total Sleep Time (TST)	Total amount of time spent asleep across entire sleep period
Final Awakening Time	Time of day subject finishes sleep period
Time in bed (TIB)	Total time spent in bed
Sleep Efficiency (SE)	Percent of time subject spent asleep while in bed

2. Related Work:

2.1 Using smartphone acoustic features to give sleep quality

Tian et al. [15] created a smartphone app called iSleep that uses the built-in microphone of the smartphone to detect the events related to sleep quality. They focused on acoustic features extracted from microphone data in the smartphone. They built a noise detection system classify the sound segments into noise-only segments and events segments. Noise-only segments are fed back to upgrade the noise detection model to make sure it is flexible to different environments. In the meanwhile, statistical acoustic features (such as root mean square, variance) are extracted from the events segments.

They integrated statistical acoustic features and noise detection model into an event detection model, this machine learning model can detect sleep-related events. And the model was trained using a lightweight decision-tree-based algorithm to classify the events. After events classification, they involved ActiGraph sleep quality estimation as standard to judge one-night sleep quality(sleep time/in-bed time), and PSQI to estimate multiple-nights sleep quality based on its scoring rules.

The study involved 7 volunteers and detected 51 nights of sleep per person. The results showed 90% accuracy for the event detection model. The study indicate that the sound data is a potential source of detecting sleep.

2.2 Using smartphone sleep diary as a label to get sleep quality

Jun-Ki et al.[16] collected microphone, accelerator, light intensity& screen proximity, running apps, battery states, screen states data from smartphone sensors to generate sleep-related features. And they make a daily sleep diary originated from PSQI which need participances to fill by themselves everyday as ground truth(label).

They recruited 27 participants and collect data for one month. After feature extraction, they use decision tree and Bayesian network to train a model to detect sleep and wake states with 93.06% accuracy, daily sleep quality with 83.97% accuracy, and global sleep quality with 81.48% accuracy.

2.3 Using mobile sensor data to predict sleep duration.

Zhenyu et al. [17] presented a radically different approach for measuring sleep duration based on a novel best effort sleep (BES) model. They classify smartphone sensors' features into 4 categories: light feature, phone usage feature, stationary feature, and silence feature. And they assume these kinds of features follow a linear combination with different weight to calculate sleep duration. The authors recruited 8 participants for one week to train the BES model. They asked participants to keep a daily diary to record their sleep duration. The record is set as ground truth. Finally, they find the BES model can estimate sleep duration within ± 49 minutes.

3. Methodology:

3.1 Research Process

The following main research steps were taken to accomplish the goals of the thesis:

1. Collect data from the ActiGraph watch to gather data on the sleep quality features as ground truth(label).

    a. Trouble falling asleep features:
    Sleep Onset Latency (SOL)
    Time Awake prior to Rising (TAWK)
    b. Trouble staying asleep features:
    Number of Awakenings (NWAK)
    c. Sleeping too much
    Total Sleep Time
    Final Awakening Time
    Time in bed
    Sleep Efficiency

2. Simultaneously collect sensor data from the following mobile phone sensors.
    The data that we can get from a mobile phone:
    - Accelerometer
    - Magnetometer
    - Audio
    - Weather, temperature, humidity
    - light
    - phone usage
    - application usage

3. Using mobile phone data as features and Actigraphy watch sleep data as label to train a model to detect sleep.

4. Extract the Sleep Quality features with the help of the sleep detecting model and Compare them with the Actigraphy Sleep features to adjust the extracting tools.

Figure 3 presents an overview of our research approach.

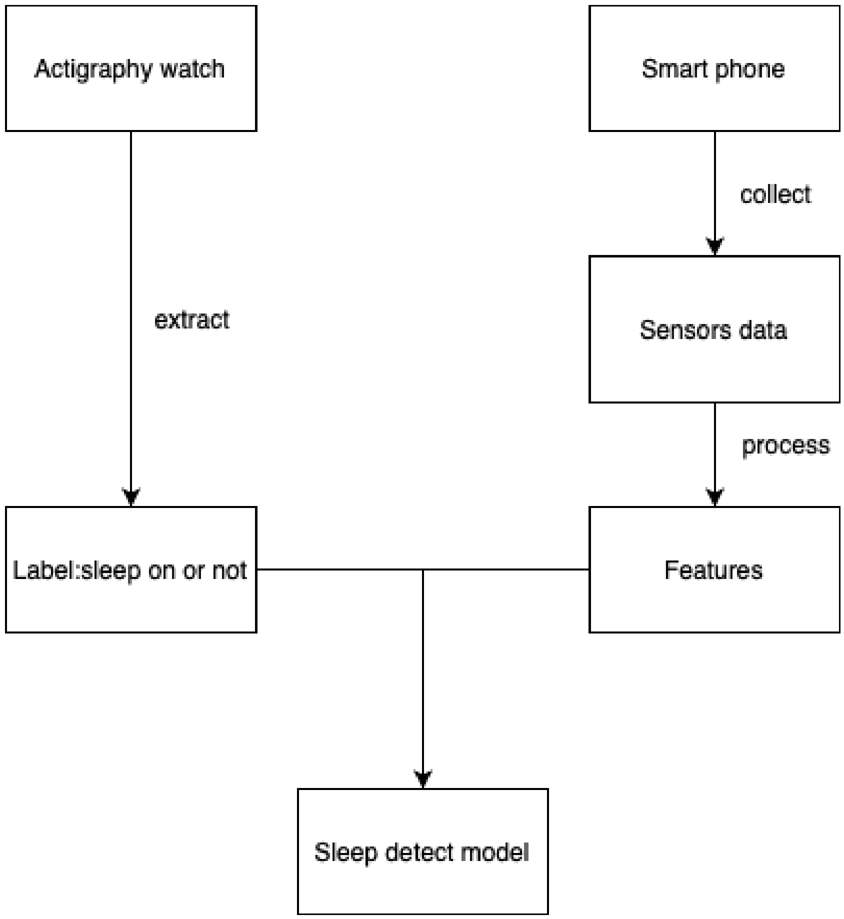

Figure 3. Data collection and processing, sleep feature extraction and machine learning model development. Mobile Sensor data and Actigraphy data are simultaneously collected for each participant. The raw mobile sensor data is preprocessed into sensor features. The ActiGraph data of sleep state is extracted as label(ground truth). The features and labels are integrated together to train a sleep detection model.

3.2 Data Gathering Study

We recruited 6 students in WPI as subjects. Each volunteer would wear one Actigraphy GT9X coupled with an Android phone for 4 days. The ActiGraph data is fed back to the

Actigraph’s platform ‘Centrepoint’ which can load and analyses the data in the Actigraph watch and measures the wearer’s sleeping. The actigraphy data is used to be set as ground truth to show the sleeping state at that time. At the same time, the smartphone collects sensor data as the potential features to be trained as a sleep detection model.

3.2.1 Phone Data Collection

For Android phones, we installed a modified ExtraSensory APP to collect sensors’ data and continually upload it to our online server. The ExtraSensory App was developed for research purposes. It provides capabilities for collecting sensor-measurements (from the phone and other possible mobile devices) and for collecting labels describing the user's context through various methods of self-reporting [24]. Consequently, the ExtraSensory App can be used for different purposes:

Collection of labeled-data:
Sensor-measurements and their assigned context-labels can both be collected, in order to train classifiers for automatic context-recognition. The sensors are recorded automatically in the background, and the labels are collected by self-reporting through the rich user-interface. Self-reporting methods include "active feedback" (where the user initiates to report their immediate future context), app-initiated notifications (which the user can respond to), and a rich history-page (real-time recognized contexts help the user organize their day and recall their actual context - they can confirm or edit the automatically recognized context-labels).

Offline context-recognition:
The ExtraSensory App can run in the background, collecting only sensor-measurements (without self-reported labels). This can be beneficial in the framework of some long-term behavioral monitoring (e.g. for patients recovering from surgery). Automatic recognition can be done either in real-time or offline (whenever the app is connected to the internet and can get context predictions from the server).

Real-time context-recognition:

The ExtraSensory App can run in the background, while connected to the internet. It can serve as another context-aware application by reacting to the recognized context of the user. For example, a music streaming app can be built around the ExtraSensory App, and it changes the streaming music throughout the day, to be relevant to the user's current environment or activity.

Self-behavioral logging:

The ExtraSensory App can be used for the sole purpose of self-logging behavior. The user-interface and the automatically recognized contexts can make it easier for the user to log their behavior. In our study, we only used ExtraSensory as a tool to collect mobile phone sensors' data. The data is collected and uploaded to the server every 1 minute with the details of phone sensors. The epoch of the data is 1 minute.

3.2.2 Actigraph watch Data Introduction

As for Actigraphy watch GT9X, we select the Sadeh algorithm[14] to detect sleep.

The Sadeh algorithm uses an 11-minute window that includes the five previous and five future epochs. And it gets 4 parameters from the watch:

AVG: The arithmetic mean (average) of the activity counts for the window

NATS: Number of epochs that have counts >= 50 and < 100

SD: The standard deviation for the first 6 epochs of the window

LG: Natural (base e) logarithm of a current epoch. Note: If the epoch count is 0, we make this value 0 to avoid infinity problems

Those calculations are put through the following algorithm:

(7.601 - (0.065 * AVG) - (1.08 * NATS) - (0.056 * SD) - (0.703 * LG))

If the result of that algorithm is GREATER than -4, then the current epoch is considered asleep.

3.2.3 Extracting Sleep-Wake Patterns

Using sleep detection algorithms such as Sadeh and Cole-Kripkee on the accelerometer data collected from the actigraphy watch was an important part of preprocessing. These algorithms help detect sleep/wake cycles within the data, which in turn helps generate the actigraphy variables used as labels for training artificial intelligence prediction models. The algorithms were implemented on the dataset using a software tool developed by Actigraphy LLC called ActiLife (See figure 4).

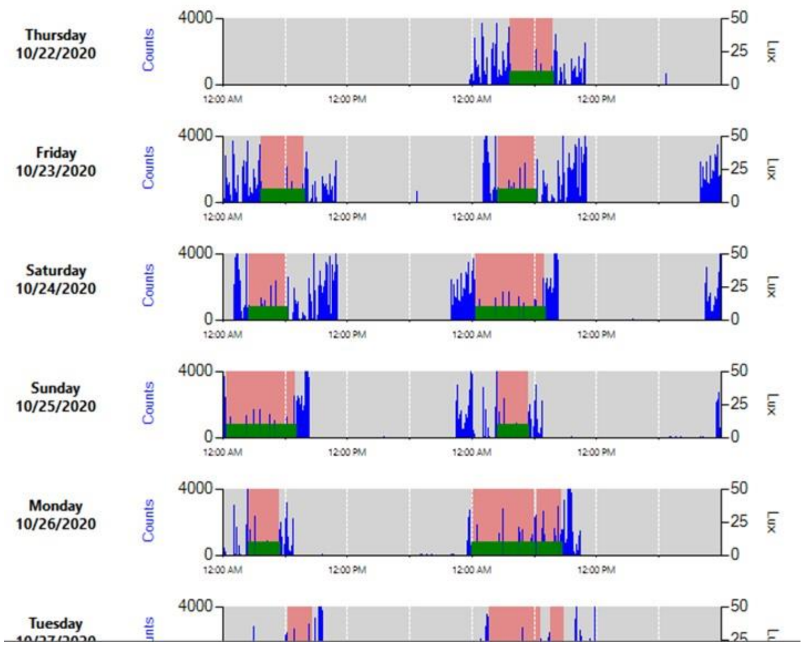

Figure 4: Sample of ActiLife Data Analyses

ActiLife's robust screening and analysis toolkit allows users to process and score collected data using a comprehensive selection of independently developed and validated algorithms. The software allows an easy implementation of scoring using these algorithms as well as allowing users to change parameters of the algorithms in order to get the most accurate fit to the dataset collected. Sadeh and Cole-Kripkee algorithms were used to detect sleep/wake activity cycles. Then the Tudor-Locke algorithm took 60 seconds segments of sleep/wake cycles as input and scores the segmented data to generate the actigraphy variables (SOL, NWAK, WASO etc...)

3.3 Machine Learning Classification Pipeline

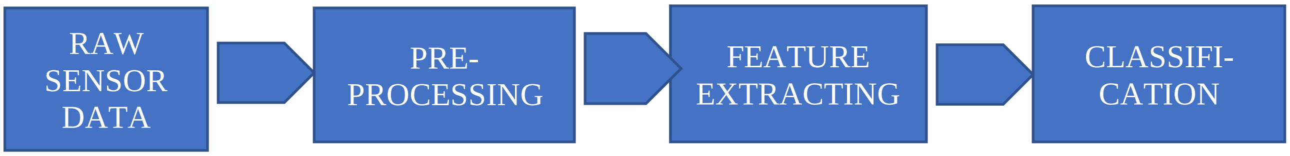

Figure 5. Overview of our Machine Learning Data Analysis Pipeline

3.3.1 Preprocessing

Preprocessing consists of identifying errors or inconsistencies within the data, extracting more features, modifying the data to make it compatible to analysis techniques and formatting the data into input for artificial intelligence algorithms.

3.3.2 Feature Extraction

We extract smartphone features from raw data, and the features can be classified into 6 main parts. Every feature is calculated or loaded every minute.

1. Accelerate features (26 features): These features contain the statistical features of accelerate magnitude mean, median, standard deviation, percentiles, etc. These features show the acceleration state of the smartphone when it is taken by a person.

2. Gyroscope features (26 features): These features contain the gyroscope magnitude mean, median, standard deviation, percentiles, etc. These features show the rotation state of the smartphone when it is taken by a person.

3. Location features (17 features): These features contain longitude, altitude and speed. These features show the location state of the smartphone.

4. Microphone features (26 features): The microphone collected and saved sound in the MFCC format. It is transferred into mean and standard deviation in 13 sound frequency intervals.

5. Discrete state features (34 features): These are Boolean features show the app state, battery state, ringer mode, WIFI state and time period of a day, etc.

6 Low-frequency measurements: light, pressure, proximity, relative humidity, battery level, screen brightness, temperature ambient. This can be used to detect how much light is present when the participant sleeps. This is important as light is important for the body to regulate its circadian rhythm.

The details of the features are showed in Table 4. below

Table 4. Feature sets for sleep detection

Modality	**Features**
Accelerometer	{Min, Q1, Median, Q3, Max, Mean, Std} of accelerometer and its 3 dimensions (X, Y, Z)
Gyroscope	{Min, Q1, Median, Q3, Max, Mean, Std} of gyroscope and its 3 dimensions (X, Y, Z)
Location	Number of valid updates, {min, max} of speed, {min, max} of altitude, log of latitude, log of longitude, diameter, best horizontal accuracy, best vertical accuracy, log diameter.
Microphone	{Mean, Std} of mfcc0 - mfcc12 (from 0 to 12 are different entries), RLH, FFT
Discrete state	Application state, battery plug, battery state, ringer mode, WIFI status, time period of the day
Low-frequency measurements	light, pressure, proximity, relative humidity, battery level, screen brightness, temperature ambient

ActiGraph watch raw data is processed using Sadeh algorithm mentioned above and it is implemented by the ActiGraph platform. By using R package 'read.gt3x', we can extract the Boolean feature sleep state (1 for asleep and 0 for awake), number feature total sleep time, number of awakenings, total awake time, average awakenings and sleep efficiency from ActiGraph watch raw output .gt3x file. The time window of the asleep feature is 1 minute, same as smartphone data we collected. The sleep state feature is set as label in the machine learning model.

3.4 Machine Learning Model Training

3.4.1. Sleeping Detecting Model:

We recruited 6 WPI students as participants in our study. The data contains 129 features from smartphone and 1 label from Actigraphy watch. To find out the best performing machine-learning algorithm to detect sleep, all popular Machine Learning algorithms in the sklearn library were tested by 6-fold cross validation. These included:

Ensemble Methods:
- AdaBoost Classifier,
- Bagging Classifier,
- Extra Trees Classifier,
- Gradient Boosting Classifier,
- Random Forest Classifier.

Gaussian Processes:
- Gaussian Process Classifier.

Generalized Linear Models
- Logistic Regression,
- Passive Aggressive Classifier,
- Ridge Classifier,
- SGD Classifier,
- Perceptron.

Naives Bayes:
- Bernoulli Naive Bayes,
- Gaussian Naive Bayes.

Nearest Neighbor:

K Neighbors Classifier.

Trees:

Decision Tree Classifier,

Extra Tree Classifier.

Discriminant Analysis:

Linear Discriminant Analysis,

Quadratic Discriminant Analysis.

3.4.2. Sleeping Quality Variables Extraction:

After selecting the best algorithm to train the sleep detecting model. We predict whether a participant was asleep or not for every minute using the smartphone. After straightforward calculation, we can get these sleep features below:

Duration: Total sleep time.
Formula: Sum of the asleep minutes during the asleep period.

Number of awakenings: the number of wake periods in the sleep time.
Formula: Sum of the continuous awakening minute clusters.

Total awake time: the time awakening during the sleep.
Formula: Sum of the awakening minutes during the asleep period.

Average awakening time: total awake time/ number of awakenings

Sleep efficiency: duration/ (duration + total awake time)

4. Evaluation and Results

4.1 Sleeping Detecting Machine Learning Model

Table5. The 7 most accurate algorithm in Model Selecting

Algorithm Name	Test Accuracy	Training Accuracy	Algorithm Time
Gradient Boosting	**0.896753**	**0.97235**	9.2533
AdaBoost	0.885429	0.95645	5.2106
KNN	0.820832	0.93562	**0.19974**
Bagging	0.790456	0.86734	2.86422
Random Forest	0.623974	0.76432	2.64039
Extra Tree	0.613984	0.75982	0.6386
Decision Tree	0.590238	0.79283	0.50811

Table 5. shows the most accurate machine learning algorithm of. It shows Gradient Boosting has the best training and test accuracy. So we choose Gradient Boosting as our sleep detecting model algorithm.

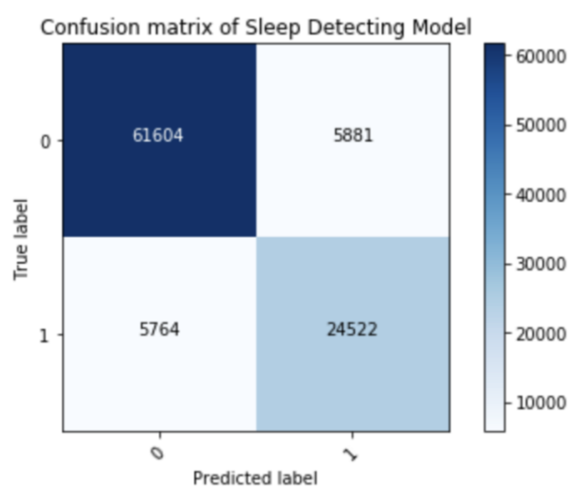

Figure 6. Confusion Matrix of Sleeping Detecting Model

Figure 6. shows the confusion matrix of our Gradient boost Sleeping Detecting Model. The accuracy of it is 88.08%. The True Positive Rate (Sensitivity) is 80.97% and the True Negative Rate (Specificity) is 91.29%. Relatively high specificity refers to the model that can detect non-sleep time more accurately.

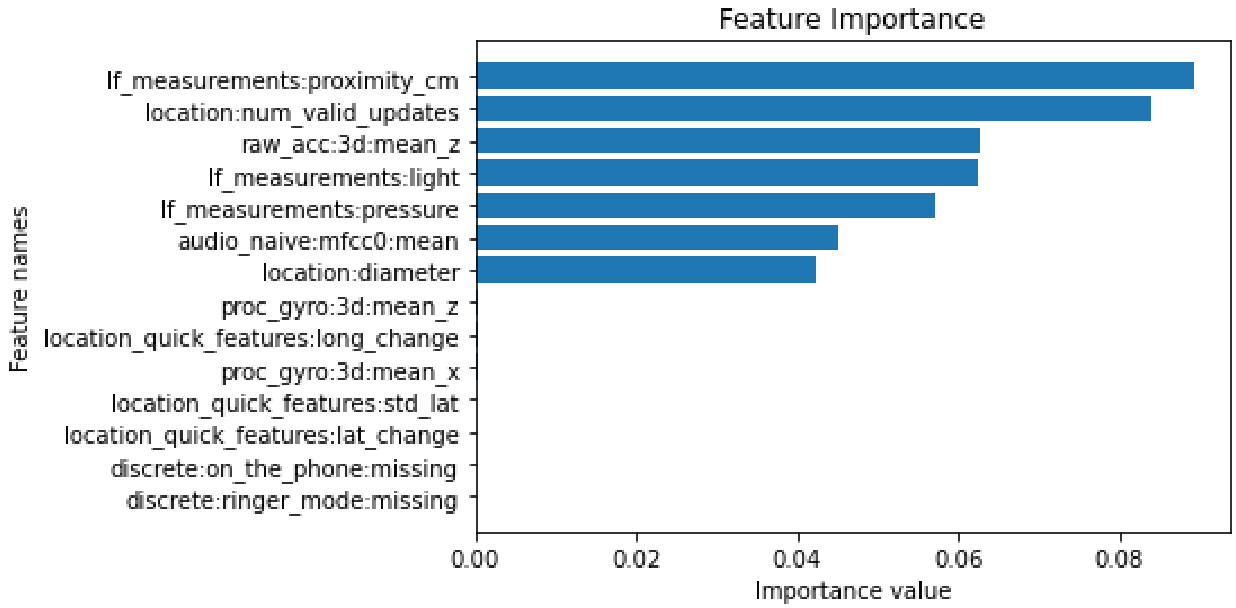

Figure 7. The 7 most important features and the 7 least important

Figure 7. The 7 most important features and the 7 least important features are showed here. We can see that the feature of proximity sensor contributes most in the sleep detection model. While there are many unimportant features that contribute only a few in the model. We can exclude these features off the model for more accurate detection.

4.2 Sleeping Quality Variables Performance

We then compared the sleep quality variables extracted by our model with the sleep quality variables extracted from the ActiGraph watch. Table 6 shows the performance of our model for the various sleep variables that our model predicted.

Sleep features	Accuracy
Total_sleep_time	82.3%
Number of awakenings	60.7%
Total_awake_time	72.5%
Average_awakenings	65.2%
Sleep_efficiency	75.2%

Table 6. Accuracy of the sleep model in predicting ground truth sleep attributes generated by the Actigraph using features extracted from a smartphone.

5. Discussion

The results of our study indicate that our sleep model that uses features extracted from smartphone data can predict ground truth sleep labels gathered using an Actigraph with reasonable accuracy. Our models generally performed well for sleep duration variables such as total sleep time and total time awake as well as sleep efficiency. However, it did not perform well in predicting the number and length of awakenings. This might be attributable to three reasons:

1. *Distance between smartphone and ActiGraph:* Unlike the Actigraphy watch that is worn on the participant's wrist, the smartphone was located far from their body. This introduces prediction errors especially for detecting awakenings.

2. *Varied placements of smartphone across participants:* When sleeping, the distance between the smartphone varied for different participants. For instance, some participants placed the smartphone near their ear, others placed the smartphone on the other side of the bed while sleeping. These placement differences may cause errors between predicting sleep variables of participants.

3. *Insufficient number of participants:* Due to logistics reasons including COVID-19, we were only able to recruit 6 participants for our study. With such a low number of participants, variations in sleep patterns may have magnified errors.

6. Conclusion and Future Work:

This thesis demonstrated that the feasibility of using the smartphone as a viable tool for passively detecting and monitoring sleep quality. A small data gathering study with 6 participants was conducted, in which sleep quality data were gathered simultaneously. These data were preprocessed, and features extracted before classification of sleep quality variables were extracted. Various machine learning variables were investigated and compared for the task of predicting ground truth Actigraph sleep quality labels from smartphone sensor features. The Gradient Boosted Machine Learning classifier had the best accuracy of 0.896753 for the task of predicting whether the subject was awake or asleep. Our machine learning models generally performed well for sleep duration variables such as total sleep time and total time awake as well as sleep efficiency. However, it did not perform well in predicting the number and length of awakenings.

Our study had several limitations including inadequate data due to the COVID-19 pandemic. In the future, we plan to collect more data from Actigraphy watch and smartphone sensors to improve our sleep prediction models.

References:

[1] M. H. Bonnet and D. L. Arand, “We are chronically sleep-deprived,” SLEEP-NEW YORK-, vol. 18, pp. 908–911, 1995.

[2] Newman A B, Nieto F J, Guidry U, et l. Relation of sleep-disordered breathing to cardiovascular disease risk factors: the Sleep Heart Health Study[J]. American journal of epidemiology, 2001, 154(1): 50-59.

[3] A. NT, W. DP, M. JE, and et al, “A prospective study of sleep duration and coronary heart disease in women,” Archives of Internal Medicine, vol. 163, no. 2, pp. 205–209, 2003.

[4] G. DJ, P. NM, N. AB, and et al, “Association of sleep time with diabetes mellitus and impaired glucose tolerance,” Archives of Internal Medicine, vol. 165, no. 8, pp. 863– 867, 2005.

[5] P. Maquet, “The role of sleep in learning and memory,” Science, vol. 294, no. 5544, pp. 1048–1052, 2001.

[6] A. Chesson et al. Practice parameters for the indications for polysomnography and related procedures. In Sleep, volume 20, pages 406–422, 1997.

[7] S. Ancoli-Israel et al. The role of actigraphy in the study of sleep and circadian rhythms. In Sleep, volume 26, pages 342–392, 2003.

[8] Data from The Google Consumer Barometer.

[9] Iber C, Ancoli-Israel S, Chesson A L, et al. The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specifications[M]. Westchester, IL: American Academy of Sleep Medicine, 2007.

[10] A. Chesson et al. Practice parameters for the indications for polysomnography and related procedures. In Sleep, volume 20, pages 406–422, 1997.

[11] Buysse D J, Reynolds C F, Monk T H, et al. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research[J]. Psychiatry res, 1989, 28(2): 193-213.

[12] S. Ancoli-Israel et al. The role of actigraphy in the study of sleep and circadian rhythms. In Sleep, volume 26, pages 342–392, 2003.

[13] Cole R J, Kripke D F, Gruen W, et al. Automatic sleep/wake identification from wrist activity[J]. Sleep, 1992, 15(5): 461-469.

[14] Sadeh A, Sharkey M, Carskadon M A. Activity-based sleep-wake identification: an empirical test of methodological issues[J]. Sleep, 1994, 17(3): 201-207.

[15] Hao T, Xing G, Zhou G. iSleep: unobtrusive sleep quality monitoring using smartphones[C]//Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems. 2013: 1-14.

[16] Min J K, Doryab A, Wiese J, et al. Toss'n'turn: smartphone as sleep and sleep quality detector[C]//Proceedings of the SIGCHI conference on human factors in computing systems. 2014: 477-486.

[17] Chen Z, Lin M, Chen F, et al. Unobtrusive sleep monitoring using smartphones[C]//2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops. IEEE, 2013: 145-152.]

[18] American Sleep Disorders Association. International classification of sleep disorders. Revised: Diagnostic and coding manual[J].

[19] Cao X L, Wang S B, Zhong B L, et al. The prevalence of insomnia in the general population in China: A meta-analysis[J]. PloS one, 2017, 12(2).

[20] Stoller M K. Economic effects of insomnia[J]. Clinical Therapeutics: The International Peer-Reviewed Journal of Drug Therapy, 1994.

[21] Castriotta R J, Lai J M. Sleep disorders associated with traumatic brain injury[J]. Archives of physical medicine and rehabilitation, 2001, 82(10): 1403-1406.

[22] "Current Definitions for Sleep Disordered Breathing in Adults". FDA.

[23] Berry, Richard et al. (2012). A The AASM Manual for the scoring of Sleep and Associated Events: Rules Terminology and Technical Specifications, Version 2.0. Darien, IL: American Academy of Sleep Medicine

[24] Vaizman Y, Ellis K, Lanckriet G, et al. Extrasensory app: Data collection in-the-wild with rich user interface to self-report behavior[C]//Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018: 1-12.

[25] H. Lu, J. Yang, Z. Liu, N. Lane, T. Choudhury, A. T. Campbell. The Jigsaw Continuous Sensing Engine for Mobile Phone Applications, In SenSys ’10.

[26] H. Lu, W. Pan, N. D. Lane, T. Choudhury, A. T. Campbell. Soundsense: Scalable Sound Sensing for People-centric Applications on Mobile Phones, In MobiSys ’09.

[27] Connie X. Wang, Isaac A. Hilburn, et al. Transduction of the Geomagnetic Field as Evidenced from alpha-Band Activity in the Human Brain. eNeuro 18 March 2019, 6 (2)