Prediction of Heart Rate Using Phone Sensors

Yixue Wendy Feng, Nutta Homdee and Natalie Kramer

***Abstract*— Continuous, ubiquitous monitoring through wearable sensors has the potential to detect physiologic measures, such as heart rate. However, wearable sensor data can have voids in the data stream. The ability to use other sensors that a person may have on them (i.e. smartphone) to predict what the wearable sensor would have collected has useful applications for the medical field. In this paper, we use data collected from both a smart watch and a smartphone to test the idea that smartphone sensor features alone have the ability to predict the heart rate collected. With our collected data, we were able to construct both individual and group-level models using regression, SVM, and random forest. Prediction shows promising results for consistent data streams, but are limited in generalizing to all participants.**

I. INTRODUCTION

Sensor data from mobile phones and wearable devices, such as smart watches, have the ability to capture continuous data in an ”in-situ” environment. An issue with this data is the how to fill voids in the data stream from either times when the device is not on the body or lapses in capture due to settings in the operating software. The central theme in this project was to consider if there is a relationship between sensor data from the mobile phone and smart watch, respectively. More specifically, the main research question is can phone sensor data predict heart rate as collected by a wrist-worn wearable device.

II. RELATED WORK

A. Section 1: Physiologic Markers

A feature collected by a smart watch that is not passively captured by phone sensors is heart rate. Heart rate is a physiologic measurement that can be useful in monitoring both physical and mental activity [1], [2]. In turn, knowledge on these states can be useful in tracking medical conditions, such as anxiety [3], [4] and stress. It can also provide insight into aspects of disease processes like loss of mobility or decreased physical activity.

A popular health metric is heart rate variability (HRV), which is a measure of the variation in time between two consecutive heart beats [5]. Numerous studies have found HRV to be a valid measure of reactivity to acute stressors [6], [7], [8]. While we only captured heart rate, the ability to predict this physiologic measure could be a stepping stone to predicting more sought after metrics, such as HRV, in the future.

B. Section 2: Sensors and Prediction

Sensor datastreams (e.g. accelerometers, microphones, heart rate, light, and location) provide continuous, unobtrusive measurements that can be a glimpse into a participant's daily activities. Previous works have researched the ability of phone sensors to predict heart rate. In work by Herandez, McDuff and Picard [9], they were able to predict heart rate using a wrist-worn accelerometer and gyroscope during sleep [9]. The same group found the 3-axis accelerometer in a smart phone was also able to predict heart rate in varying positions and phone location with mean absolute error of 1.16 beats per minute (STD: 3). The varying phone participant positions included standing up, sitting down, and lying down, which were repeated before and after exercising, followed by watching a video, listening on the phone, and browsing the Internet while sitting down [10]. Their results are shown in Figure 1.

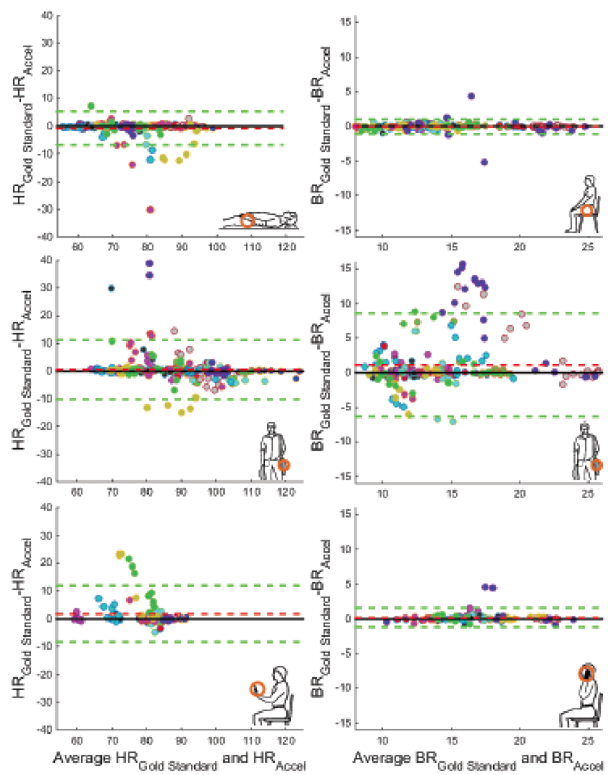


Fig. 1. Results from BioPhone[10] Bland-Altman plots for heart (left) and breathing rates (right) of the conditions that yielded the best mean absolute error when the phone was inside the pocket (top), inside the bag (middle), and on the hand (bottom). Mean error is depicted with slashed red and 95 percent limits are depicted with slashed green lines. (HR: Heart Rate in beats per minute, BR: Breathing Rate in breaths per minute, Accel: Accelerometer). N = 216 for top and middle graphs, and N = 108 for bottom graphs.

Passive monitoring of both heart rate and phone sensors is not a novel concept, however, we were unable to find another study that worked to create a heart rate prediction model from data collected in the wild. Many groups have created prediction models using passively sensed data. In work by Hao et al, they utilized a smartwatch with a PPG sensor to collect HRV to predict stressful situations for the participants. Their model, StressHacker, reliably captured daily stress dynamics with a precision of 86.1 percent and a recall of 91.2 percent. Other studies have shown relationships between accelerometery and location sensing and physiologic states. Notably, Boukhechba et al found preliminary results that passively sensed physiologic data can be used to detect changes following cognitive bias modification interpretation training. Ameko et al also used location and accelerometry to determine negative affect at the group level. Although these studies did not directly predict heart rate, passively sensed physiologic measures were collected; we used these papers to help guide our data preprocessing and model selection.

III. STUDY DESIGN

Twelve university students were recruited for a seven-day study period to collect both wrist-worn sensor data and phone sensor data. The selected group provided a relatively homogeneous sample mitigating the impact of confounding factors. Both Android and iOS phones were being used among the group. A customized mobile app (Sensus) [11] was installed on participants personal smartphones and was programmed to collect GPS coordinates every 1 second and accelerometer data at 1 Hz. All data were transmitted wirelessly to a secure Amazon Web Services server, where data was stored until further analysis (see Figure 2).

Fig. 2. Passively collected data

For the study period, participants were also given a Huawei Smart Watch 2 to wear. The watch recorded optical heart rate (beats per minute) at a sampling rate of 1 Hz. Participants were asked to wear the watch as much as possible during the 7 days. Data was stored on-board and extracted at the end of the collection period for further analysis.

IV. METHODS

A. Data Preprocessing

We first examined the raw data and selected four participants with the most robust and complete data sets. For comparison we deliberately chose two iOS users and two Android users, as the ability to capture continuous data streams varies between the two platforms. For accelerometer data, we first calculated a magnitude of acceleration $(a = \sqrt{(x^2+y^2+z^2)})$ to have an orientation free feature since the phones were use in the participants' natural environment. Accelerometry data is used to asses physical activity level of the participant and in using the magnitude value, the value of 1.0 meant the phone was still; to make this value more meaningful for prediction models, we decided to create features for each of the accelerometer metrics that represented the deviation from 1.0. Figure 3 shows HR and magnitude from one data for each participant over the seven-day study; data was resampled using means over 30 minutes time periods. In the interest of maintaining as much data as possible, all accelerometer metrics (x, y, z, and magnitude) were used for analysis.

GPS data was preprocessed into cluster using the ST-DBSCAN, an algorithm developed based on DBSCAN using spatial temporal data [12] (see Figure 4). This algorithm clusters data based on both location and temporal threshold. Datetime was parsed into hour, minute and day and extracted as features. For individual prediction models, raw heart rate was used. To compare heart rates across participants, it made the most sense to utilize a z-score as heart rate is a very individualized health metric. The heart rate z-score was utilized in the group-level model using all four participants' data.

B. Predictive Models

1) Individual Models: We used three algorithms to test the predictability of heart rate: ridge regression, support vector machine (SVM), and random forest decision tree. We chose to start with a regression as a natural first step in analyzing the data and ultimately chose to use the ridge regression analysis in order to diminish the mulitcollinearity of the variables. In addition to the regression model, the SVM regression model was chosen due to its ability to handle non-linear data. Lastly, to better understand the results of the SVM model, a random forest was utilized. The included features for the predictive models are listed in Table 1.

TABLE I
VARIABLES USED IN PREDICTION

Feature	Individual Models (target: raw HR)			General Models (target: z-score HR)	
	Ridge Regression	SVM	Random Forest	Ridge Regression	Random Forest
Time Features					
hour	x	x	x	x	x
minute	x	x	x	x	x
second	x	x	x	x	x
Location Features					
latitude	x	x	x	x	x
longitude	x	x	x	x	x
location cluster	x	x	x		
Accelerometer Features					
X	x	x	x	x	x
Y	x	x	x	x	x
Z	x	x	x	x	x
magnitude from 1	x	x	x	x	x

Fig. 3. HR and Accelerometer Magnitude Data

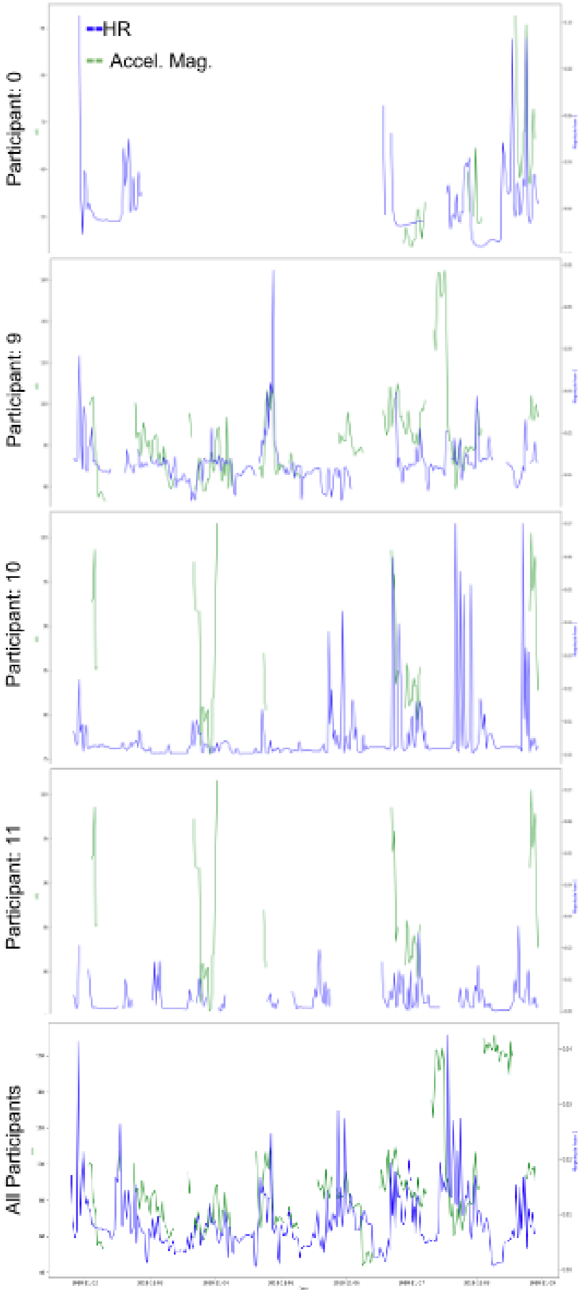
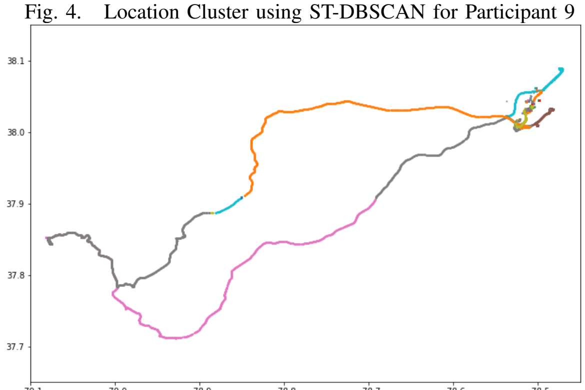


Fig. 4. Location Cluster using ST-DBSCAN for Participant 9



V. RESULTS

In this section, the performance of each regression models are shown with 1.) Coefficient of determination or the R squared value and 2.) Mean-squared error (MSE). We started by computing the simplest regression using thirty minutes prior heart rate to interpolate the next thirty minute. This thirty minute heart rate interpolation is used as a baseline to compare with other heart rate regression models. The performance of this baseline regression is shown in Table 2.

TABLE II

INDIVIDUAL BASELINE REGRESSION RESULTS

ID	Data Size	R^2	MSE
Participant 0	27911	0.2072	188.3461
Participant 9	265605	0.2739	735.9945
Participant 10	3552	0.243	108.8668
Participant 11	43036	0.4107	882.5675

A. Ridge Regression

For prediction models, we started with the ridge regression using actual heart as the outcome variable. The data was not randomized when split into train-test data (80/20 split). The decision to not shuffle the data was made to ensure the model was predicting durations of time versus random time points. The results of the individual models are listed in Table 3.

TABLE III

INDIVIDUAL RIDGE REGRESSION RESULTS

ID	Data Size	R^2		MSE	
		Train	Test	Train	Test
Participant 0	27911	0.59	0.42	102.21	77.84
Participant 9	265605	0.25	-0.13	258.56	1648.28
Participant 10	3552	0.39	-0.20	39.82	75.27
Participant 11	43036	0.71	-0.05	449.62	514.19

B. Support Vector Machine

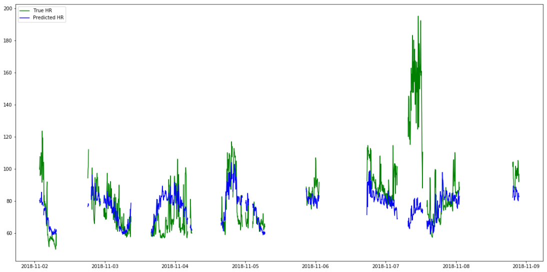
For our individual SVM models, we used the same features and outcome variable as in the ridge regression. We used the svm.SVR method in the sklearn package for training and predicting (C = 1 and epsilon = 0.2). The R^2 and MSE results are presented in Table 4 and graphical representation of the predicted and actual heart rate for participant 9 is shown on Figure 5 (resampled using mean for every 5 minutes).

TABLE IV

INDIVIDUAL RESULTS OF SVM MODEL

ID	Data Size	R^2		MSE	
		Train	Test	Train	Test
Participant 0	27911	0.49	-0.00	125.54	136.37
Participant 9	265605	0.70	0.17	104.86	1213.64
Participant 10	3552	0.62	0.69	24.67	106.35
Participant 11	43036	0.75	-0.06	383.93	762.78

Fig. 5. SVM Heart Rate Prediction results for Participant 9



Many of our participants had large intervals of missing data; to see how that impacted the SVM model, we condensed the data to approximately a three day time-frame with the most consistent datastream for participant 9. With this condensed time-frame, the ridge regression had an R^2 (MSE) of 0.45 (139.44) and 0.39 (121.67) for train and test data respectively.

C. *Random Forest*

Lastly, Table 5 gives the results of the random forest; for the individual models the tree parameter was set to 200 with a train-test split of 70/30. Charts visualizing the actual versus predicted heart are included in Figures 6-9.

TABLE V

INDIVIDUAL RESULTS OF RANDOM FOREST

ID	Data Size	R^2		MSE	
		Train	Test	Train	Test
Participant 0	27911	0.92	0.92	19.49	18.15
Participant 9	265605	0.96	0.97	22.32	21.20
Participant 10	3552	0.86	0.86	0.96	10.73
Participant 11	43036	0.96	0.96	53.43	52.72

1) General Models: In addition to our individual models, we also created general models using all four participants' data. To accommodate for the different ranges of heart rate baseline of each participant, the models are targeted the heart rate's z-score instead of the actual beat-per-minute heart rate. The same parameters (e.g. alpha, epsilon, C, number of tree) as the individual models were used. The results of the general regression, SVM models, and random forest are shown in Table 6. Figure 10 is graphical representation of the SVM model prediction versus the actual mean heart rate. Figure 11 shown the result of the random forest prediction vs. the actual heart rate z-score.

TABLE VI

GENERAL MODEL RESULTS

Model	Data Size	R^2		MSE	
		train	test	train	test
Ridge Regression	340104	0.56	-1.36	0.53	1.46
SVM	340104	0.06	-0.05	1.11	0.66
Random forest	340104	0.95	0.96	0.04	0.04

Fig. 6. Random Forest Results for Participant 0

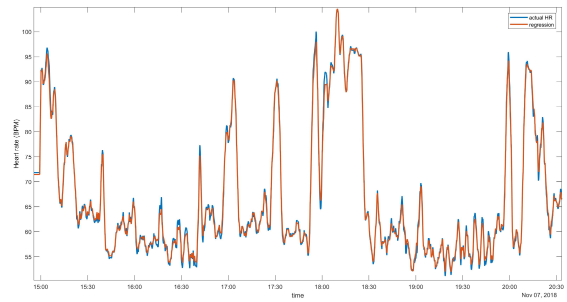


Fig. 7. Random Forest Results for Participant 9

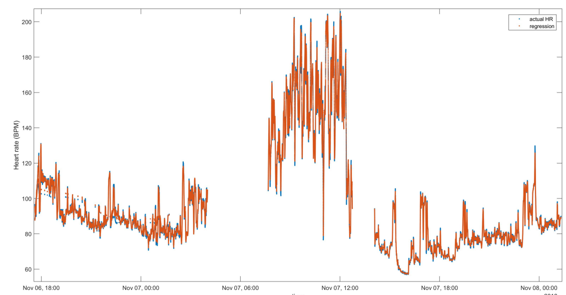


Fig. 8. Random Forest Results for Participant 10

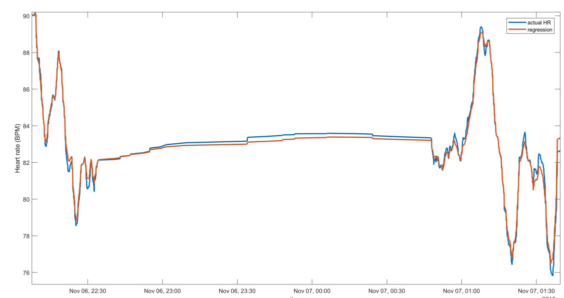


Fig. 9. Random Forest Results for Participant 11

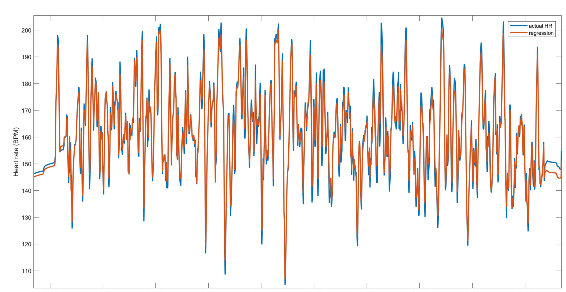


Fig. 10. SVM General Model Results

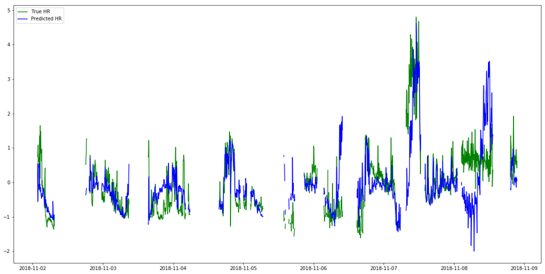
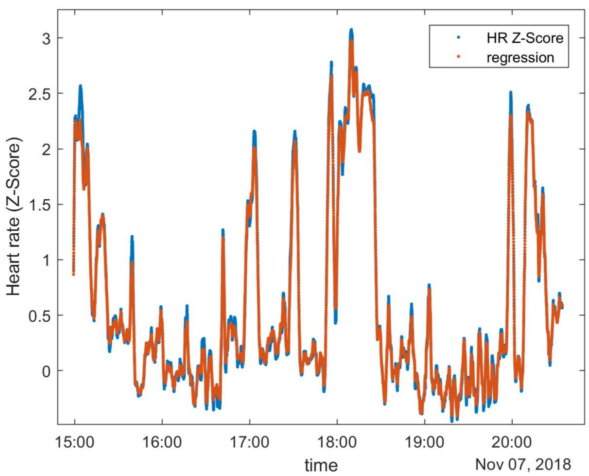


Fig. 11. Random Forest General Model Results



VI. CONCLUSIONS

In this study, we were able to show heart prediction using smart phone sensors. Although there are many limitation and confounders to the models, we believe it is a positive first step. The low R-squared values in our beginning regression models are most likely due to two reasons; first, being the non-linear nature of our data. To combat this in the future, we would consider a Gaussian model with the heart rate being classified into deviations away from the mean heart rate. Gaussian models have been recommended for healthcare researcher [13] due to the model's ability to handle variable dependence and non-linear data, while giving the researcher the ability to add knowledge about smoothness or periodicity using the covariance functions [14]. Secondly, when doing human research and real-world data collection, expected R-squared values can be lower than a controlled labratory experiment.

As we expected, our SVM and random forest regression demonstrated the best results in heart rate prediction for individuals. The SVM model tended to be better at predicting higher heart rates than heart rates that were lower the individual mean. We think this could be due limited phone sensor data when heart rate would be at its lowest (e.g. phone is still and participant is sleeping). Adding additional features such as light and audio to the model could help predict heart rate when the phone is not moving.

The random forest had the best results, however, the graphical representation of the predicted versus actual heart rate are so close we are concerned about over fitting. Typically random forests are utilized due to their robustness against overfitting [15]. In the future, the parameters of our random forest may need to be evaluated.

While we were able to create a heart rate prediction model, there are limitations to our model. First, we only used a small sample of participants who all occupied the same university setting. Also, data capture varied between participants; in the future we consider ways, such as protocol changes, to streamline collection and ensure the maximum amount of data is recorded. In addition, there were many collected variables we did not use for analysis, such as audio and light, that could be processed and used to improve the models.

The focus of this investigation was to provide a evaluate predicting heart rate given smartphone sensors. We developed both individual and group-level models with the individual models having a better prediction results. There are many ways to improve on these models as outlined above; future work will study more participants with more complete data that can continue to be applicable to more populations and evaluate uses for heart rate prediction through ubiquitous monitoring .

REFERENCES

[1] M. C. E. A. Haskell, W. L.and Yee and P. J. Irby, "Simultaneous measurement of heart rate and body motion to quantitate physical activity," *Med Sci Sports Exerc*, vol. 25, no. 1, pp. 109–115, Jan 1993.

[2] K. Madden and G. K. Savard, "Effects of mental state on heart rate and blood pressure variability in men and women," *Clin Physiol*, vol. 15, no. 6, pp. 557–569, Nov 1995.

[3] J. M. Gorman and R. P. Sloan, "Heart rate variability in depressive and anxiety disorders," *Am. Heart J.*, vol. 140, no. 4 Suppl, pp. 77–83, Oct 2000.

[4] C. Schiweck, D. Piette, D. Berckmans, S. Claes, and E. Vrieze, "Heart rate and high frequency heart rate variability during stress as biomarker for clinical depression. A systematic review," *Psychol Med*, pp. 1–12, Aug 2018.

[5] G. Theile, V. Klaas, G. Troster, and M. Guckenberger, "mHealth Technologies for Palliative Care Patients at the Interface of In-Patient to Outpatient Care: Protocol of Feasibility Study Aiming to Early Predict Deterioration of Patient's Health Status," *JMIR Res Protoc*, vol. 6, no. 8, p. e142, Aug 2017.

[6] J. W. Hughes and C. M. Stoney, "Depressed mood is related to high-frequency heart rate variability during stressors," *Psychosom Med*, vol. 62, no. 6, pp. 796–803, 2000.

[7] J. Healey and R. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 6, pp. 156 – 166, July 2005.

[8] H. Ruediger, R. Seibt, K. Scheuch, M. Krause, and S. Alam, "Sympathetic and parasympathetic activation in heart rate variability in male hypertensive patients under mental stress," *J Hum Hypertens*, vol. 18, no. 5, pp. 307–315, May 2004.

[9] J. Hernandez, D. McDuff, and R. W. Picard, "Biowatch: Estimation of heart and breathing rates from wrist motions," in *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, May 2015, pp. 169–176.

[10] J. Hernandez, D. J. McDuff, and R. W. Picard, "Biophone: Physiology monitoring from peripheral smartphone motions," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2015, pp. 7180–7183.

[11] H. Xiong, Y. Huang, L. E. Barnes, and M. S. Gerber, "Sensus: A cross-platform, general-purpose system for mobile crowdsensing in human-subject studies," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp

'16. New York, NY, USA: ACM, 2016, pp. 415–426. [Online]. Available: http://doi.acm.org/10.1145/2971648.2971711

[12] D. Birant and A. Kut, "St-dbscan: An algorithm for clustering spatial-temporal data," *Data Knowl. Eng.*, vol. 60, no. 1, pp. 208–221, Jan. 2007. [Online]. Available: http://dx.doi.org/10.1016/j.datak.2006.01.013

[13] L. Clifton, D. A. Clifton, M. A. Pimentel, P. J. Watkinson, and L. Tarassenko, "Gaussian processes for personalized e-health monitoring with wearable sensors," *IEEE Trans Biomed Eng*, vol. 60, no. 1, pp. 193–197, Jan 2013.

[14] M. Ameko, L. Cai, M. Boukhechba, A. Daros, P. Chow, B. Teachman, M. Gerber, and L. Barnes, "Cluster-based approach to improve affect recognition from passively sensed data," March 2018.

[15] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct 2001. [Online]. Available: https://doi.org/10.1023/A:1010933404324