PLOS ONE

RESEARCH ARTICLE

Posttraumatic stress disorder hyperarousal event detection using smartwatch physiological and activity data

Mahnoosh Sadeghi[ID]**, Anthony D. McDonald**[ID]**, Farzan Sasangohar**[ID]*

Department of Industrial and / Systems Engineering, Texas A&M University, College Station, Texas, United States of America

* sasangohar@tamu.edu

Abstract

Posttraumatic Stress Disorder (PTSD) is a psychiatric condition affecting nearly a quarter of the United States war veterans who return from war zones. Treatment for PTSD typically consists of a combination of in-session therapy and medication. However; patients often experience their most severe PTSD symptoms outside of therapy sessions. Mobile health applications may address this gap, but their effectiveness is limited by the current gap in continuous monitoring and detection capabilities enabling timely intervention. The goal of this article is to develop a novel method to detect hyperarousal events using physiological and activity-based machine learning algorithms. Physiological data including heart rate and body acceleration as well as self-reported hyperarousal events were collected using a tool developed for commercial off-the-shelf wearable devices from 99 United States veterans diagnosed with PTSD over several days. The data were used to develop four machine learning algorithms: Random Forest, Support Vector Machine, Logistic Regression and XGBoost. The XGBoost model had the best performance in detecting onset of PTSD symptoms with over 83% accuracy and an AUC of 0.70. Post-hoc SHapley Additive exPlanations (SHAP) additive explanation analysis showed that algorithm predictions were correlated with average heart rate, minimum heart rate and average body acceleration. Findings show promise in detecting onset of PTSD symptoms which could be the basis for developing remote and continuous monitoring systems for PTSD. Such systems may address a vital gap in just-in-time interventions for PTSD self-management outside of scheduled clinical appointments.

Introduction

PTSD is a psychiatric condition experienced by individuals after exposure to life-threatening events, such as physical assault, sexual abuse, and combat exposure [1]. PTSD symptomology includes avoidance, hyperarousal, and reexperiencing trauma through dreams and recollections [1]. Avoidance symptoms include circumventing activities or thoughts associated with the traumatic event, decreased interest in daily life, and an overall feeling of detachment from

one's surroundings. Hyperarousal symptoms include hypervigilance, feelings of irritability, and an exaggerated response following a startling event. Other symptoms of PTSD include anxiety, insomnia, fatigue, anger, and aggression [2, 3].

Over 70% of the U.S. population will experience a traumatic event in their lifetime of whom 20% will subsequently develop Post-Traumatic Stress Disorder (PTSD) [4]. Combat veterans are particularly prone to PTSD [5], with recent estimates of prevalence as high as 20% [4]. Veterans with PTSD are also at a greater risk of suicide [6], and suicidal thoughts [7]—according to a 2019 U.S. Department of Veteran Affairs annual report, an average of 20 veterans per day commit suicide, with a majority of cases linked to PTSD [8]. Beyond personal costs, PTSD has an enormous societal cost associated with healthcare service utilization. The costs of caring for war veterans usually peaks 30–40 years following a major conflict. For example, for Iraq/Afghanistan conflicts, it is estimated that costs of veterans' care will peak around 2035 [9, 10]. This increased healthcare use is estimated to cost the United States over $60 billion each year and the costs are also expected to increase due to secondary and tertiary comorbidities including depression, substance abuse, smoking, heart disease, obesity, diabetes, chronic fatigue, and increased dementia [9, 11].

PTSD is typically managed by a combination of therapeutic and pharmaceutical treatments, although many cases go undiagnosed or untreated potentially due to mental illness stigma and care shortages [12]. Therapeutic methods include eye movement desensitization and reprocessing; exposure therapy; cognitive therapy; cognitive restructuring therapy; cognitive processing therapy; stress inoculation therapy as part of stress management therapy [13]; and the trauma-focused cognitive behavioral therapy [14]. While these methods are effective [15, 16], there are several barriers to care access including geographical, financial, and cultural constraints, and limited care delivery resources [17, 18]. In addition, most intense symptoms of PTSD are often experienced outside clinical environments and in-between therapeutic sessions [19]. Therefore, there is a critical need for tools and methods for real-time monitoring and detection of PTSD signs and symptoms, as well as mechanisms to support self-management of such symptoms. Recent advances in wearable physiological sensors and mobile health (mHealth) technologies may provide a viable alternative to address this need.

Prior work has shown that PTSD is correlated with several physiological measures including heart rate, heart rate variability, blood pressure, respiratory rate and skin conductance [20]. Among these, heart rate has shown promise as a reliable correlate of PTSD [21]. Sadeghi et al. [22] used the Autoregressive Integrated Moving Average (ARIMA) analysis to model veterans' heart rate patterns during PTSD hyperarousal events. Their results indicated strong correlation between heart rate characteristics (e.g., autocorrelation and fluctuation) and hyperarousal events. Recent efforts have utilized supervised machine learning tools to detect perceived stress using heart rate, other physiological metrics, and self-reported measures with reported accuracies ranging between 67%-92% [23–28], however, to our knowledge, only one study [23] used machine learning algorithms to predict the onset of PTSD symptoms among veterans based on heart rate data.

In their study, McDonald et al. [23] investigated self-reported periods of hyperarousal to extract heart rate time-dependent features and developed five machine learning algorithms: a Conventional Neural Network, Neural Network, Support Vector Machine (SVM), Random Forest, and Decision Tree. Among these methods, the SVM showed the highest accuracy (over 70%). While this study provided preliminary evidence supporting the efficacy of using heart rate to detect hyperarousal events, using heart data alone may be subject to significant noise associated with movement or physical activity [29, 30]. In line with previous research [e.g., 25, 26, 31], McDonald et al. [23] suggested that body acceleration data might improve the accuracy of machine learning algorithms and enable algorithms to distinguish between heart rate
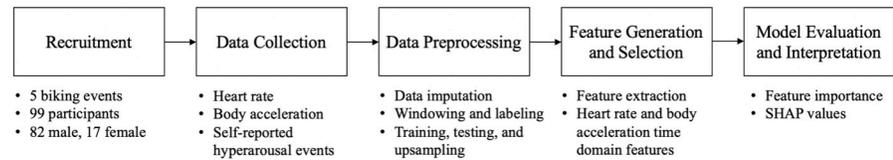
Fig 1.

https://doi.org/10.1371/journal.pone.0267749.g001

fluctuations due to physical activity and heart rate fluctuations due to mental stress. Therefore, the objective of this article is to expand McDonald et al.'s [23] study to machine learning algorithms that uses body acceleration and heart rate data to predict PTSD hyperarousal events in veterans. In addition, in an effort to improve the interpretation of the algorithm, we further analyze the developed model to investigate significant factors contributing to model's detection output.

Materials and methods

Four machine learning algorithms were trained using self-reported data collected naturalistically from veterans to predict PTSD hyperarousal events: Random Forest, XGBoost, Logistic Regression and non-linear SVM. Fig 1 provides an overview of the methods used in this study. This study was approved by the Institutional Review Board (IRB) at Texas A&M University (IRB2017-0210D) and all participants completed an informed consent prior to data collection. Fig 1 provides an overview of the methods used in this study.

Participants

Participants were recruited from seven Project Hero's United Healthcare Ride 2 Recovery (R2R) challenges. Project Hero is a non-profit organization dedicated to help veterans and first responders diagnosed with PTSD. In each challenge, veterans rode for an average of 7 days between key destinations in California, Washington DC, Minneapolis, Texas, and Nevada. Each day of the challenge involved approximately 8 hours of biking with the remaining time for resting and socializing. The research team joined a total of 5 rides in 2017, 2018 and 2019.

Data from 99 veteran participants (82 male; 17 female) were used in this study. Participants' age ranged from 22 to 75 years old (M = 45.5, SD = 10). Majority of participants reported Veterans Affairs disability rating of over 90% related to PTSD. Table 1 summarizes other relevant demographics.

Data collection

The data collection application (app) for smart wearable devices utilized in [23] was used. Participants were asked to wear smart watches (MOTO 360 Gen 1 or Gen 2, Apple Watch series 3 or 4) with the app installed on them for the duration of the study. The app ran continuously in the background and connected to participants' phones for the purpose of data transfer. The app had the ability to continuously and remotely collect physiological data including heart rate and acceleration from participants at the frequency of 1 Hz. The app included functionality which allowed the user to report a hyperarousal event (symptomatic of PTSD) through a simple 'double tap' anywhere on the watch face which created a time-stamped self-reported event. These events were used for training the machine learning algorithm.

Table 1. Participants' demographics; the numbers show the number of veterans per each group.

Ethnicity		Branch		VA Disability Ratings for PTSD	
American Indian or Alaska Native	4	Air Force	5	40%	2
Asian	1	Army	68	50%	4
Black/African American	15	40%	2	70%	6
Hispanic/Latino	26	50%	4	80%	13
Native Hawaiian	1	70%	6	$\geq$ 90%	74
White	44	80%	13		
Other	8	$\geq$ 90%	74		

Data preprocessing

All data analysis including data preprocessing and machine learning were conducted in Python 3.8.2 and R 3.6.2. The data preprocessing included four main steps: (1) imputation, (2) windowing and labeling, (3) dividing the data into training and testing, and (4) resampling the training dataset.

Further, four machine learning algorithms including Random Forest, XGBoost, Logistic Regression and non-linear SVM were used to detect hyperarousal events. These four algorithms have been used and shown promise for stress detection in previous research [32–40] and have several strengths that makes them suitable for this specific research. For example, Random Forest has been widely used for large physiological datasets due to its robustness in dealing with missing values [41, 42]. In addition, Logistic Regression are computationally efficient algorithms commonly used for predicating binary output [43]. Lastly, XGBoost was chosen due to its efficiency, accuracy, interpretability, and ease of integration with mobile applications [42].

Data imputation. Kalman filter imputation was used to impute missing acceleration and heart rate data. Kalman filter imputation is an established method for timeseries data imputation [44], especially for heart rate data [45]. To determine the cut off range, we calculated the average Mean Square Error (MSE) of the imputed data and corresponding actual values. A cut off range of 15 MSE for estimating the randomly dropped values is suggested by Gui et al. [45]. Based on Kalman filter imputation analysis, we chose 5 as the maximum imputation range because it was the greatest value among a set of successive values to have the highest MSE less than 15 [23].

Windowing and labeling. To investigate the patterns of hyperarousal events, the data was divided into 60-second sliding windows with 30 seconds overlap, chosen based on prior work [24] to predict stress severity based on physiological reactions. Each window was assigned a label based on the presence or absence of reported hyperarousal events. If a hyperarousal event occurred anywhere in the window, it was labeled as hyperarousal event; otherwise, it was labeled as non-hyperarousal event. All windows with over 80% missing values were dropped from the dataset. The final dataset included 530 and 13,554 instances of hyperarousal and non-hyperarousal events, respectively.

Training, testing, and upsampling. To validate the algorithm, the data was separated into training (70%) and testing (30%) sets by participant. Participants were assigned to only the testing or training set to ensure generalizability of the results. Table 2 shows the initial dataset classifications. One of the challenges of training the algorithm to detect PTSD hyperarousal events was the imbalanced dataset—96.2% of the windows were labeled non-hyperarousal events. To address this issue, we upsampled the training data. Upsampling was used because it

Table 2. Training and testing datasets after and before resampling.

	Label	Training set	Test set
Training and testing datasets before resampling	Non-hyperarousal events	9486	4068
	Hyperarousal events	372	158
Training and testing datasets after resampling	Non-hyperarousal events	9486	4068
	Hyperarousal events	7114	158

https://doi.org/10.1371/journal.pone.0267749.t002

decreases the information lost in the quantification process, thereby reducing the noise and increasing the resolution of the results [46, 47]. Upsampling has several advantages over down-sampling. The main advantage of upsampling is its ability to use the entire information in the data rather than omitting part of the data [48]. In addition, for noisy datasets, which is often the case when collecting data in naturalistic settings, oversampling is more robust to the noise in the data than undersampling and performs better for prediction [49]. Based on a sensitivity analysis comparing different resampling ratio including 1–1, 2–1, 3–1, 3–2, and 4–3, a ratio of 4 (non-hyperarousal events) to 3 (hyperarousal events) windows was used for upsampling (Table 2).

Feature generation and selection

Heart rate data. Previous research has shown that that time domain features of heart rate are strongly correlated with PTSD [50]. We extracted time domain features of heart rate including maximum heart rate (bpm), minimum heart rate (bpm), heart rate standard deviation (bpm), heart rate range (max-min) (bpm), and average heart rate (bpm) from each window of time to use for PTSD hyperarousal prediction. Key features were extracted based on recommendations from a review article on detecting psychological stress using bio signals [51].

Acceleration data. Research on stress prediction have used scalars of body acceleration to estimate body activity and to remove noise from the data [25, 51–53]. Garcia-Ceja et al. [52] used time domain and frequency domain features of acceleration to predict stress in participants in real work environments. In line with these approaches, we calculated the vector of body acceleration for each moment using the following widely used formula:

$$\text{body acceleration} = \sqrt{a_x^2 + a_y^2 + a_z^2}$$

Where a_x is the body acceleration in X direction, a_y is body acceleration in Y direction, and a_z is body acceleration in Z direction.

Further, based on previous research [e.g., 25, 52], time domain features of body acceleration including average body acceleration (m/s^2), maximum body acceleration (m/s^2), minimum body acceleration (m/s^2), and range of body acceleration (m/s^2) were extracted to feed machine learning algorithms.

Model assessment. Four models including Random Forest (size of feature set = 10, number of trees = 50) XGBoost (Maximum depth = 37, number of trees = 50, learning rate = 0.3), Logistic Regression (nLambda = 100) and non-linear SVM (degree = 3, C-Classification, Radial Basis Function) were trained to predict hyperarousal events. We generated a confusion matrix to assess the performance of each model. Model comparisons were conducted with a 5x2 cross validation test following the recommendations in Dietterich [54] to minimize the type 1 error. This method uses p values and t statistics to compare the algorithms. The null

hypothesis indicates that there is no significant difference between the algorithms in terms of performance (average accuracy) where the alternative hypothesis shows that one algorithm is more accurate than the others. Algorithms were further assessed with the Area Under the receiver operating characteristic (ROC) Curve (AUC).

Feature importance and model interpretation. The complexity of black box machine learning models and the need to make these models explainable necessitate an evaluation of the influence of algorithm's features on algorithm predictions. In this study, we used Shapley Additive exPlanations (SHAP) to address this. SHAP uses game theoretic concepts to allocate values to features in a model based on their importance in prediction [55]. SHAP values indicate how much each feature contributes to the prediction of the machine learning algorithm. SHAP value summary plots generate a feature importance list along with the distribution of each feature and shows how each value affects the output of the model. This method is computationally efficient, consistence with human intuition, and interpretable for explaining class differences [56]. Using SHAP values to interpret machine learning algorithms has several advantages over using more traditional methods such as dependency plots [56]. For example, dependence plots do not usually show features' distributions, which may lead to misinterpreting regions with significant missing data. Conversely, SHAP plots show feature distributions. SHAP values also indicate how much a feature affects the output of the prediction by considering interaction effects, whereas partial dependence plots do not account for interactions between features.

Results

Model performance and comparison

Fig 2 shows the ROC curves for each algorithm along with the AUC values. As shown in the plots, XGBoost had the highest AUC (0.70). Random Forest, Logistic Regression, and non-linear Kernel SVM had AUC of 0.63, 0.62, and 0.61, respectively. Table 3 shows confusion matrices for developed algorithms at three probability cutoffs. The first confusion matrix prioritizes hyperarousal detection, the second confusion matrix balances the true positives and false positives rate, and the third matrix prioritizes minimizing false positive rates. The pairwise 5*2 cross validation test results showed that XGBoost significantly outperformed the Random Forest (t = -13.25, p < 0.001), SVM (t = -13.02, p < 0.001), and Logistic Regression (t = -11.97,
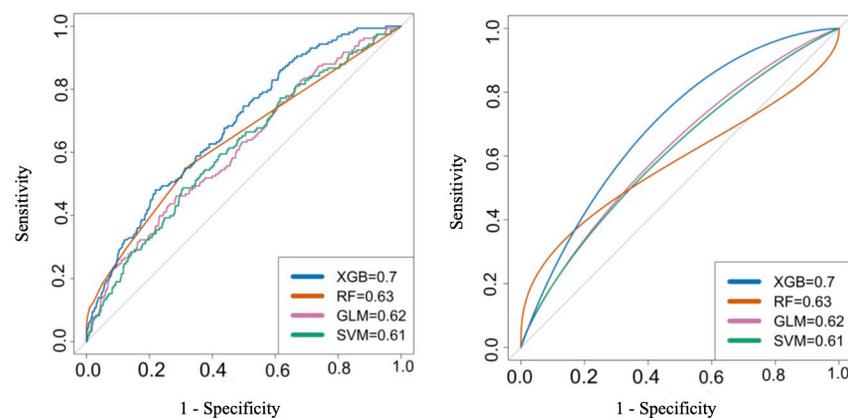
Fig 2.

https://doi.org/10.1371/journal.pone.0267749.g002

Table 3. Confusion matrices for all models at different probability cut offs.

Design Performance	Algorithm	TP	FN	FP	TN	TPR	FPR
Prioritize hyperarousal detection (TPR = 1)	XGB	158	0	3821	247	1	0.94
	RF	158	0	3853	205	1	0.95
	GLM	158	0	4066	2	1	0.99
	SVM	158	0	4065	3	1	0.99
Balanced priorities (TPR = 0.5)	XGB	79	79	1064	3004	0.5	0.26
	RF	88	70	1322	2746	0.55	0.33
	GLM	79	79	1467	2601	0.5	0.36
	SVM	79	79	1401	2667	0.5	0.34
Prioritize false positive minimization (FPR = 0.1)	XGB	46	112	420	3648	0.29	0.1
	RF	29	129	205	3863	0.18	0.1
	GLM	38	120	410	3658	0.24	0.1
	SVM	30	128	409	3659	0.19	0.1

p < 0.001). Based on the results from Table 3, XGBoost showed the best performance in detecting PTSD hyperarousal events.

Model interpretation

Fig 3 shows the SHAP summary plot for the XGBoost model. In Fig 3, the Y-axis shows the feature as well as the mean SHAP values ordered from top to bottom, the color shows the

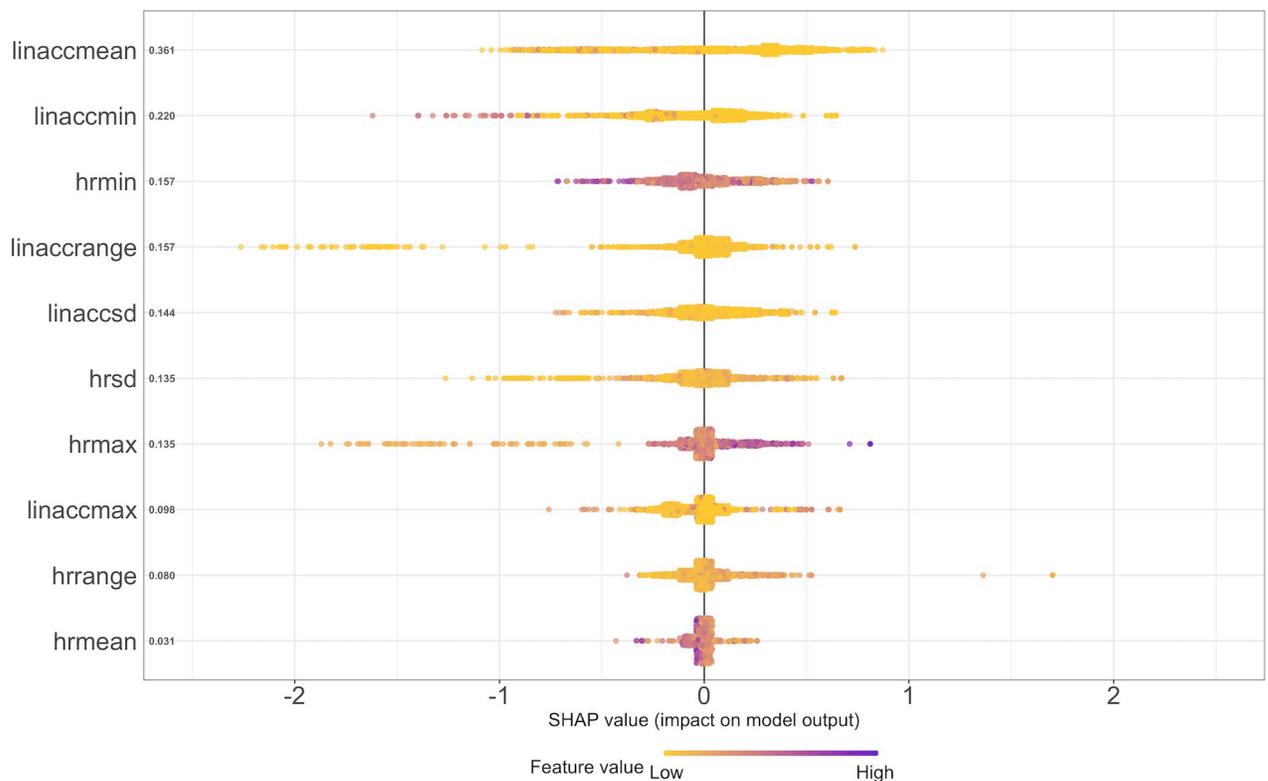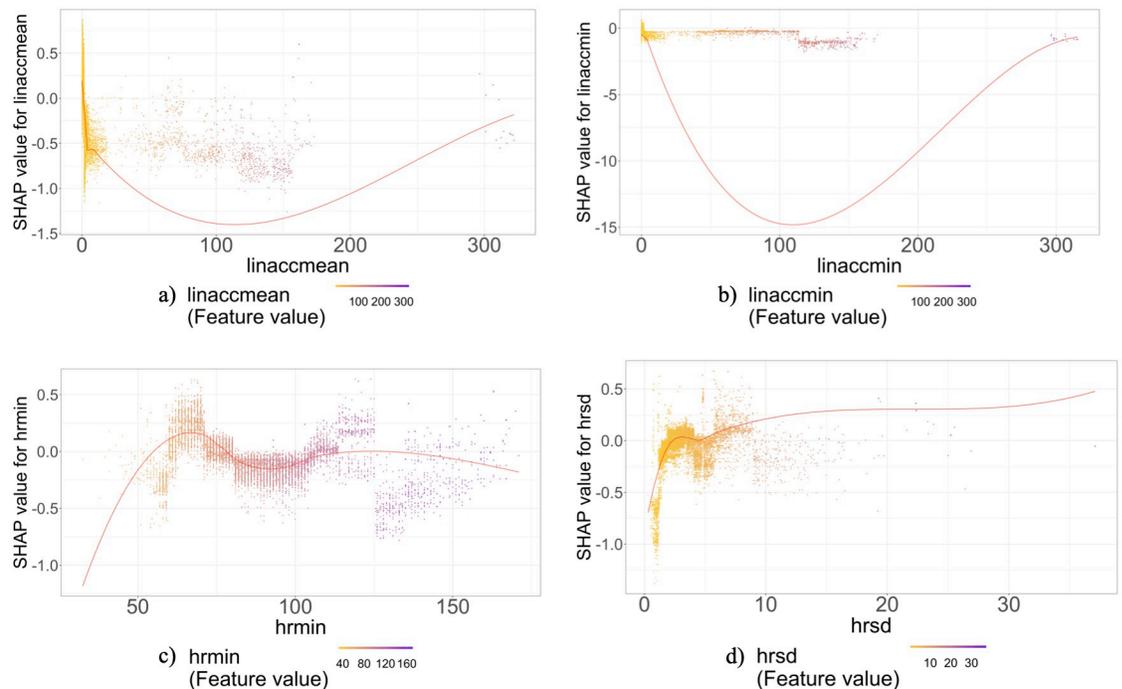**Fig 3.**

Fig 4.

significance of the feature's value in predicting the output, and the X-axis indicates how the feature affects the output of the model (whether that feature with that specific value is contributing to experiencing a hyperarousal event or not). The X axis further indicates log-odds of perceiving a PTSD hyperarousal event. According to the SHAP analysis, the most important body acceleration features are average body acceleration (linaccmean) and minimum body acceleration (linaccmin). The most important heart rate time-domain features for predicting PTSD hyperarousal events are minimum heart rate (hrmin) and heart rate standard deviation (hrsd).

Fig 4 shows the SHAP dependence plots for the two most important acceleration and heart rate features. SHAP dependence plots show contribution of a specific feature to a model based on the feature's distribution. In this plot each point shows an observation from the dataset, the X-axis line shows the value of the feature in that row, and the Y-axis shows the SHAP value for that feature that indicates the effect of that feature with that specific value on the prediction. The unit of X-axis is the same as the unit of the feature (for instance for heart rate measures it is beats per minute), and the unit of the Y-axis is log-odds of perceiving a PTSD hyperarousal event.

As shown in Fig 4, hyperarousal events are more likely to be observed with higher minimum heart rate values over the window. When the minimum heart rate is over 140 the risk of perceiving hyperarousal events increases. Also, as the average body acceleration and minimum body acceleration increase, the odds of the detecting PTSD hyperarousal events decrease. Finally, higher heart rate standard deviation, i.e. higher heart rate fluctuation, increases the risk of hyperarousal events.

Discussion

This study developed, evaluated, and explicated machine learning algorithms to predict PTSD hyperarousal events among veterans using smartwatch based naturalistic heart rate and

accelerometer data. The ground truth was subjectively-reported PTSD hyperarousal events. After preprocessing the data, we trained four different algorithms including Random Forest, SVM, Logistic Regression and XGBoost. Among the developed algorithms, the XGBoost was the most robust algorithm which yielded an AUC of 0.70 and over 81% accuracy. We sorted the most important features in the prediction process. The top three body acceleration features included average body acceleration, minimum body acceleration and range of body acceleration. The top three heart rate time domain features were minimum heart rate, standard deviation of heart rate and maximum heart rate. The initial analysis from the SHAP summary plot and SHAP dependence plots show that heart rate and body acceleration features have nonlinear relationships with PTSD episodes. A deeper look into SHAP plots indicate that as the body acceleration increases, indicating more activity from the participant, the algorithm is less likely to predict a PTSD hyperarousal events. This result is consistent with prior studies demonstrating a significant relationship between increased physical activity and a reduction in PTSD hyperarousal events [57–59].

The SHAP dependence plot for the average heart rate data corroborates that when the heart rate is between 60–70 bpm, PTSD hyperarousal events are more likely to happen (cf. [60]). The SHAP summary plot indicates that heart rate standard deviation was one of the most important features contributing to the odds that the algorithm will predict PTSD hyperarousal event manifestation. In particular, our findings suggest that as the heart rate standard deviation increases, i.e., as heart rate fluctuates more and in higher ranges, the odds of detecting a PTSD hyperarousal event increases. This result supports previous findings [22, 61, 62] which showed that during PTSD hyperarousal events, participants experience increased heart rate acceleration and fluctuation.

While previous research on objective assessment of PTSD exists, most of such research has been conducted in controlled lab settings by inducing external stimuli or so-called triggers (e.g., [62–64]). Naturalistic studies investigating common internally generated stimuli such as thoughts or flashbacks [65] are largely absent. In addition, while several studies have attempted to detect stress using physiological data, most of these studies have used Heart Rate Variability (HRV) features (e.g., [24, 34, 66, 67]). Given the availability and non-intrusiveness of heart rate (or pulse rate) sensors, it is timely to investigate the efficacy of using heart rate features to monitor or predict stress and more specifically PTSD hyperarousal. However, to our knowledge, only one study [23] has used heart rate features to predict PTSD hyperarousal events using data collected in naturalistic settings.

This study builds on a prior study conducted by McDonald et al. [23] but there are several key differences that both enhance the algorithm and improve our understanding of algorithm performance. First, we expanded McDonald et al.'s dataset by conducting two additional field studies. Second, McDonald et al. indicated that one of the main limitations of their work is utilizing downsampling for data preprocessing. We addressed this limitation by using an upsampling method for preprocessing the data. Third, McDonald et al. used frequency domain features of heart rate such as coefficients of Fourier decomposition of heart rate. In our analysis we used time domain features of heart rate which may provide more tangible and interpretable results for this specific context [50, 68]. This change may be beneficial for integrating the detection algorithm into a treatment device because explaining predictions to veterans based on time domain features (e.g., standard deviation of heart rate) will be more understandable than frequency domain features (e.g., the phase of the 5th Fourier component). Fourth, in our analysis we added body acceleration features to decrease the noise in data by differentiating heart rate changes due to physical activity versus fluctuations due to stress (similar to [27]) and used a new XGBoost algorithms to train the machine learning model. We believe these changes have contributed to the significant improvement in the machine learning algorithm

performance compared to the findings from McDonald et al. (81% in this study compared to 70% in [23]). Finally, McDonald et al. explain that one of their study limitations is the limited insight into the performance-shaping factors for the algorithm predictions. We addressed this issue by adding feature importance and SHAP analysis to improve the interpretability of the results.

Several limitations of this study should be addressed in future work. First, stress and PTSD hyperarousal events are highly idiosyncratic. A stimulus that triggers one individual may or may not trigger someone else. Because of the subjective and sustained characteristics of stress, defining the start, end, duration, and intensity of a hyperarousal event is an uncertain task [69]. As a result, it is significantly complex and difficult to define and measure a ground truth for stress. Hyperarousal events might have been over or under reported due to the subjectivity of the perceived events. Individual differences such as gender, age, lifestyle and other factors can affect PTSD hyperarousal events; therefore, personalizing machine learning algorithms might boost their performance. Another issue in this study was the high number of heart rate missing values due to the naturalistic nature of the study. While non-intrusiveness of smart watches makes them suitable for naturalistic data collection, like most wrist-based sensors, smart watches use optical technology the accuracy of which is affected by skin tone and proximity to skin [70]. Future work may validate the findings presented here using more accurate sensors such as chest straps that use electrical pulse. Lastly, although machine learning algorithms work in theory, external validation of these algorithms in naturalistic settings are necessary to evaluate the accuracy and applicability of these algorithms in the real world settings.

This article provides preliminary evidence of efficacy for data-driven real-time PTSD hyperarousal detection tools that can be used beyond clinic walls to remotely and continuously monitor veterans suffering from PTSD. In addition to the promise shown by the machine learning algorithm, in this article we utilized analytical techniques to which identifies most important features contributing to such detection, hence, improving the interpretation of the outcomes and moving towards explainable ML tools for PTSD monitoring. Although other machine learning algorithms exist to detect stress, to the best of our knowledge, the algorithm documented in this paper is one of very few algorithms that is specific to PTSD. The work is in progress to validate this algorithm in longitudinal home studies using smart watches and smart phones.

Author Contributions

Conceptualization: Anthony D. McDonald, Farzan Sasangohar.

Data curation: Mahnoosh Sadeghi.

Formal analysis: Mahnoosh Sadeghi, Anthony D. McDonald.

Funding acquisition: Farzan Sasangohar.

Methodology: Anthony D. McDonald, Farzan Sasangohar.

Project administration: Anthony D. McDonald, Farzan Sasangohar.

Resources: Farzan Sasangohar.

Supervision: Anthony D. McDonald, Farzan Sasangohar.

Writing – original draft: Mahnoosh Sadeghi.

Writing – review & editing: Anthony D. McDonald, Farzan Sasangohar.

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub; 2013.

2. Carlson KF, Kehle SM, Meis LA, Greer N, MacDonald R, Rutks I, et al. Prevalence, assessment, and treatment of mild traumatic brain injury and posttraumatic stress disorder: a systematic review of the evidence. The Journal of head trauma rehabilitation. 2011; 26: 103–115. https://doi.org/10.1097/HTR.0b013e3181e50ef1 PMID: 20631631

3. Xiong K, Zhang Y, Qiu M, Zhang J, Sang L, Wang L, et al. Negative emotion regulation in patients with posttraumatic stress disorder. PLoS One. 2013; 8: e81957. https://doi.org/10.1371/journal.pone.0081957 PMID: 24349161

4. Stefanovics EA, Potenza MN, Pietrzak RH. PTSD and obesity in US military veterans: Prevalence, health burden, and suicidality. Psychiatry research. 2020; 291: 113242. https://doi.org/10.1016/j.psychres.2020.113242 PMID: 32763532

5. Xue C, Ge Y, Tang B, Liu Y, Kang P, Wang M, et al. A meta-analysis of risk factors for combat-related PTSD among military personnel and veterans. PloS one. 2015; 10: e0120270. https://doi.org/10.1371/journal.pone.0120270 PMID: 25793582

6. Brake CA, Rojas SM, Badour CL, Dutton CE, Feldner MT. Self-disgust as a potential mechanism underlying the association between PTSD and suicide risk. Journal of anxiety disorders. 2017; 47: 1–9. https://doi.org/10.1016/j.janxdis.2017.01.003 PMID: 28126678

7. McKinney JM, Hirsch JK, Britton PC. PTSD symptoms and suicide risk in veterans: Serial indirect effects via depression and anger. Journal of affective disorders. 2017; 214: 100–107. https://doi.org/10.1016/j.jad.2017.03.008 PMID: 28288403

8. U.S. Department of Veteran Affairs. National Veteran Suicide Prevention Annual Report. 2019; 32.

9. Geiling J, Rosen JM, Edwards RD. Medical costs of war in 2035: long-term care challenges for veterans of Iraq and Afghanistan. Military medicine. 2012; 177: 1235–1244. https://doi.org/10.7205/milmed-d-12-00031 PMID: 23198496

10. Mavranezouli I, Megnin-Viggars O, Grey N, Bhutani G, Leach J, Daly C, et al. Cost-effectiveness of psychological treatments for post-traumatic stress disorder in adults. PloS one. 2020; 15: e0232245. https://doi.org/10.1371/journal.pone.0232245 PMID: 32353011

11. Allabadi H, Alkaiyat A, Zahdeh T, Assadi A, Ghanayim A, Hasan S, et al. Posttraumatic stress disorder predicts poor health-related quality of life in cardiac patients in Palestine. Plos one. 2021; 16: e0255077. https://doi.org/10.1371/journal.pone.0255077 PMID: 34314470

12. Mittal D, Drummond KL, Blevins D, Curran G, Corrigan P, Sullivan G. Stigma associated with PTSD: Perceptions of treatment seeking combat veterans. Psychiatric rehabilitation journal. 2013; 36: 86. https://doi.org/10.1037/h0094976 PMID: 23750758

13. Post-Traumatic-Stress-Disorder-Fact-Sheet-.pdf. https://www.sidran.org/wp-content/uploads/2018/11/Post-Traumatic-Stress-Disorder-Fact-Sheet-.pdf

14. Haagen JF, Smid GE, Knipscheer JW, Kleber RJ. The efficacy of recommended treatments for veterans with PTSD: A metaregression analysis. Clinical Psychology Review. 2015; 40: 184–194. https://doi.org/10.1016/j.cpr.2015.06.008 PMID: 26164548

15. Foa EB, Rothbaum BO. Treating the trauma of rape: Cognitive-behavioral therapy for PTSD. Guilford Press; 2001.

16. Rauch SA, Eftekhari A, Ruzek JI. Review of exposure therapy: a gold standard for PTSD treatment. Journal of rehabilitation research and development. 2012; 49: 679–688. https://doi.org/10.1682/jrrd.2011.08.0152 PMID: 23015579

17. Kazlauskas E. Challenges for providing health care in traumatized populations: barriers for PTSD treatments and the need for new developments. Global health action. 2017; 10: 1322399. https://doi.org/10.1080/16549716.2017.1322399 PMID: 28562198

18. Rodriguez-Paras C, Tippey K, Brown E, Sasangohar F, Creech S, Kum H-C, et al. Posttraumatic stress disorder and mobile health: App investigation and scoping literature review. JMIR mHealth and uHealth. 2017; 5: e156. https://doi.org/10.2196/mhealth.7318 PMID: 29074470

19. Moon J, Smith A, Sasangohar F, Benzer JK, Kum H-C. A descriptive model of the current PTSD care system: Identifying opportunities for improvement. Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care. SAGE Publications Sage CA: Los Angeles, CA; 2017. pp. 251–251.

20. Zoladz PR, Diamond DM. Current status on behavioral and biological markers of PTSD: a search for clarity in a conflicting literature. Neuroscience & Biobehavioral Reviews. 2013; 37: 860–895.

21. Pitman RK, Lanes DM, Williston SK, Guillaume JL, Metzger LJ, Gehr GM, et al. Psychophysiologic assessment of posttraumatic stress disorder in breast cancer patients. Psychosomatics. 2001; 42: 133–140. https://doi.org/10.1176/appi.psy.42.2.133 PMID: 11239126

22. Sadeghi M, Sasangohar F, McDonald AD, Hegde S. Understanding Heart Rate Reactions to Post-Traumatic Stress Disorder (PTSD) Among Veterans: A Naturalistic Study. Human Factors. 2021; 00187208211034024.

23. McDonald AD, Sasangohar F, Jatav A, Rao AH. Continuous monitoring and detection of post-traumatic stress disorder (PTSD) triggers among veterans: a supervised machine learning approach. IISE Transactions on Healthcare Systems Engineering. 2019; 9: 201–211.

24. McDuff D, Gontarek S, Picard R. Remote measurement of cognitive stress via heart rate variability. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2014. pp. 2957–2960.

25. Sano A, Phillips AJ, Amy ZY, McHill AW, Taylor S, Jaques N, et al. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN). IEEE; 2015. pp. 1–6.

26. Sano A, Picard RW. Stress recognition using wearable sensors and mobile phones. 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction. IEEE; 2013. pp. 671–676.

27. Smets E, De Raedt W, Van Hoof C. Into the wild: the challenges of physiological stress detection in laboratory and ambulatory settings. IEEE journal of biomedical and health informatics. 2018; 23: 463–473. https://doi.org/10.1109/JBHI.2018.2883751 PMID: 30507517

28. Pavlova I, Zikrach D, Mosler D, Ortenburger D, Góra T, Wąsik J. Determinants of anxiety levels among young males in a threat of experiencing military conflict–Applying a machine-learning algorithm in a psychosociological study. PloS one. 2020; 15: e0239749. https://doi.org/10.1371/journal.pone.0239749 PMID: 33027278

29. Comtois G, Mendelson Y, Ramuka P. A comparative evaluation of adaptive noise cancellation algorithms for minimizing motion artifacts in a forehead-mounted wearable pulse oximeter. 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2007. pp. 1528–1531.

30. Matsumura T, Chemmalil VT, Gray ML, Keating JE, Kieselbach RL, Latta SB, et al. Device for measuring real-time energy expenditure by heart rate and acceleration for diabetic patients. 2009 IEEE 35th Annual Northeast Bioengineering Conference. IEEE; 2009. pp. 1–2.

31. Gjoreski M, Gjoreski H, Luštrek M, Gams M. Continuous stress detection using a wrist device: in laboratory and real life. proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: Adjunct. 2016. pp. 1185–1193.

32. Han L, Zhang Q, Chen X, Zhan Q, Yang T, Zhao Z. Detecting work-related stress with a wearable device. Computers in Industry. 2017; 90: 42–49.

33. Sun SL. HRVBased Stress Recognizing by Random Forest. Fuzzy Systems and Data MiningII: Proceedings of FSDM 2016. 2016;293: 444.

34. Al-Shargie F, Tang TB, Badruddin N, Kiguchi M. Towards multilevel mental stress assessment using SVM with ECOC: an EEG approach. Medical & biological engineering & computing. 2018; 56: 125–136. https://doi.org/10.1007/s11517-017-1733-8 PMID: 29043535

35. El Haouij N, Poggi J-M, Ghozi R, Sevestre-Ghalila S, Jaïdane M. Random forest-based approach for physiological functional variable selection for driver's stress level classification. Statistical Methods & Applications. 2019; 28: 157–185.

36. Al-Shargie FM, Tang TB, Badruddin N, Kiguchi M. Mental stress quantification using EEG signals. International Conference for Innovation in Biomedical Engineering and Life Sciences. Springer; 2015. pp. 15–19.

37. Vanitha L, Suresh GR. Hybrid SVM classification technique to detect mental stress in human beings using ECG signals. 2013 International Conference on Advanced Computing and Communication Systems. IEEE; 2013. pp. 1–6.

38. Bozorgmehr A, Thielmann A, Weltermann B. Chronic stress in practice assistants: An analytic approach comparing four machine learning classifiers with a standard logistic regression model. Plos one. 2021; 16: e0250842. https://doi.org/10.1371/journal.pone.0250842 PMID: 33945572

39. Srividya M, Mohanavalli S, Bhalaji N. Behavioral modeling for mental health using machine learning algorithms. Journal of medical systems. 2018; 42: 1–12. https://doi.org/10.1007/s10916-018-0934-5 PMID: 29610979

40. Lu Y, Fu X, Guo E, Tang F. XGBoost algorithm-based monitoring model for urban driving stress: Combining driving behaviour, driving environment, and route familiarity. IEEE Access. 2021; 9: 21921–21938.

41. Harrington P. Machine learning in action. Manning Publications Co.; 2012.

42. Ayyadevara VK. Pro machine learning algorithms. Apress: Berkeley, CA, USA. 2018.

43. Hsieh FY, Bloch DA, Larsen MD. A simple method of sample size calculation for linear and logistic regression. Statistics in medicine. 1998; 17: 1623–1634. PMID: 9699234

44. Welch G, Bishop G. An introduction to the Kalman filter. Citeseer; 1995.

45. Gui Q, Jin Z, Xu W. Exploring missing data prediction in medical monitoring: A performance analysis approach. 2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). IEEE; 2014. pp. 1–6.

46. Beferull-Lozano B, Ortega A. Coding techniques for oversampled steerable transforms. Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers (Cat No CH37020). IEEE; 1999. pp. 1198–1202.

47. Stewart RW, Pfann E. Oversampling and sigma-delta strategies for data conversion. Electronics & Communication Engineering Journal. 1998; 10: 37–47.

48. Mohammed R, Rawashdeh J, Abdullah M. Machine learning with oversampling and undersampling techniques: overview study and experimental results. 2020 11th International Conference on Information and Communication Systems (ICICS). IEEE; 2020. pp. 243–248.

49. Kaur P, Gosain A. Comparing the behavior of oversampling and undersampling approach of class imbalance learning by combining class imbalance problem with noise. ICT Based Innovations. Springer; 2018. pp. 23–30.

50. Nagpal M, Gleichauf K, Ginsberg J. Meta-analysis of heart rate variability as a psychophysiological indicator of posttraumatic stress disorder. Journal of Trauma & Treatment. 2013; 3: 2167–1222.

51. Giannakakis G, Grigoriadis D, Giannakaki K, Simantiraki O, Roniotis A, Tsiknakis M. Review on psychological stress detection using biosignals. IEEE Transactions on Affective Computing. 2019.

52. Garcia-Ceja E, Osmani V, Mayora O. Automatic stress detection in working environments from smartphones' accelerometer data: a first step. IEEE journal of biomedical and health informatics. 2015; 20: 1053–1060. https://doi.org/10.1109/JBHI.2015.2446195 PMID: 26087509

53. Porumb M, Stranges S, Pescapè A, Pecchia L. Precision medicine and artificial intelligence: A pilot study on deep learning for hypoglycemic events detection based on ECG. Scientific reports. 2020; 10: 1–16.

54. Dietterich TG. Approximate statistical tests for comparing supervised classification learning algorithms. Neural computation. 1998; 10: 1895–1923. https://doi.org/10.1162/089976698300017197 PMID: 9744903

55. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Advances in neural information processing systems. 2017. pp. 4765–4774.

56. Molnar C. Interpretable Machine Learning. Lulu.com; 2020.

57. Asmundson GJ, Fetzner MG, DeBoer LB, Powers MB, Otto MW, Smits JA. Let's get physical: a contemporary review of the anxiolytic effects of exercise for anxiety and its disorders. Depression and anxiety. 2013; 30: 362–373. https://doi.org/10.1002/da.22043 PMID: 23300122

58. Fetzner MG, Asmundson GJ. Aerobic exercise reduces symptoms of posttraumatic stress disorder: A randomized controlled trial. Cognitive behaviour therapy. 2015; 44: 301–313. https://doi.org/10.1080/16506073.2014.916745 PMID: 24911173

59. Oppizzi LM, Umberger R. The effect of physical activity on PTSD. Issues in mental health nursing. 2018; 39: 179–187. https://doi.org/10.1080/01612840.2017.1391903 PMID: 29319376

60. Woodward SH, Murburg MM, Bliwise DL. PTSD-related hyperarousal assessed during sleep. Physiology & behavior. 2000; 70: 197–203. https://doi.org/10.1016/s0031-9384(00)00271-7 PMID: 10978496

61. Sadeghi M, Sasangohar F, McDonald AD. Toward a taxonomy for analyzing the heart rate as a physiological indicator of posttraumatic stress disorder: systematic review and development of a framework. JMIR Mental Health. 2020; 7: e16654. https://doi.org/10.2196/16654 PMID: 32706710

62. Elsesser K, Sartory G, Tackenberg A. Attention, heart rate, and startle response during exposure to trauma-relevant pictures: a comparison of recent trauma victims and patients with posttraumatic stress disorder. Journal of abnormal psychology. 2004; 113: 289. https://doi.org/10.1037/0021-843X.113.2.289 PMID: 15122949

63. Elsesser K, Sartory G, Tackenberg A. Initial symptoms and reactions to trauma-related stimuli and the development of posttraumatic stress disorder. Depression and Anxiety. 2005; 21: 61–70. https://doi.org/10.1002/da.20047 PMID: 15782424

64. Adenauer H, Catani C, Keil J, Aichinger H, Neuner F. Is freezing an adaptive reaction to threat? Evidence from heart rate reactivity to emotional pictures in victims of war and torture. Psychophysiology. 2010; 47: 315–322. https://doi.org/10.1111/j.1469-8986.2009.00940.x PMID: 20030756

65. Buswell G, Haime Z, Lloyd-Evans B, Billings J. A systematic review of PTSD to the experience of psychosis: prevalence and associated factors. BMC psychiatry. 2021; 21: 1–13.

66. Zubair M, Yoon C, Kim H, Kim J, Kim J. Smart wearable band for stress detection. 2015 5th International Conference on IT Convergence and Security (ICITCS). IEEE; 2015. pp. 1–4.

67. Melillo P, Bracale M, Pecchia L. Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination. Biomedical engineering online. 2011; 10: 1–13.

68. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Frontiers in public health. 2017; 5: 258. https://doi.org/10.3389/fpubh.2017.00258 PMID: 29034226

69. Gjoreski M, Luštrek M, Gams M, Gjoreski H. Monitoring stress with a wrist device using context. Journal of biomedical informatics. 2017; 73: 159–170. https://doi.org/10.1016/j.jbi.2017.08.006 PMID: 28803947

70. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ digital medicine. 2020; 3: 1–9.