# Mental Stress Classification During a Motor Task in Older Adults Using an Artificial Neural Network

Apostolos Kalatzis
Montana State University
Bozeman, Montana
apostolos.kalatzis@student.montana.edu

Laura Stanley
Montana State University
Bozeman, Montana
Laura.stanley@montana.edu

Rohith Karthikeyan
Texas A&M University
College Station, Texas
rohithkarthikeyan@tamu.edu

Ranjana K. Mehta
Texas A&M University
College Station, Texas
rmehta@tamu.edu

## ABSTRACT

All people cope with mental stress from time to time. Stress can affect our emotional and physical health, which can lead to physical and/or mental health issues. Our experiment aimed to derive the stress levels of 57 older adults from the electrocardiogram (ECG) signal during a lab study that involved a hang-grip strength task. This experiment bridges the gap between previous studies by classifying the mental stress state of older adults while performing a motor task before and after the stressor was induced. In this study heart rate and heart rate variability multi-dimensional features in the time-, and frequency-domain are extracted and an optimized Artificial Neural Network (ANN) created to identify two states — stress, or no-stress. We achieved accuracy of 90.83%.

## CCS CONCEPTS

• **Human-centered computing** → *Ubiquitous and mobile computing*;

## KEYWORDS

Machine learning; Elderly population; Stress detection; ECG signal; Artificial Neural Network

**ACM Reference Format:**
Apostolos Kalatzis, Laura Stanley, Rohith Karthikeyan, and Ranjana K. Mehta. 2020. Mental Stress Classification During a Motor Task in Older Adults Using an Artificial Neural Network. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers (UbiComp/ISWC '20 Adjunct), September 12–16, 2020, Virtual Event, Mexico.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3410530.3414360


*UbiComp/ISWC '20 Adjunct, September 12–16, 2020, Virtual Event, Mexico*

ACM ISBN 978-1-4503-8076-8/20/09…$15.00
https://doi.org/10.1145/3410530.3414360

## 1 INTRODUCTION

The world's population is aging rapidly. According to the World Health Organization [1], between 2015 and 2050, the proportion of the world's older adults is estimated to almost double from about 12% to 22%. In the United States older adults often continue to work after retirement [7]. It has been noticed the labor force participation for this age group has been increased in recent years [13]. However, the prevalence of stress among elderly people is increasing, and the physical and cognitive changes some people experience as they age can make this issue even more challenging [18]. Therefore, the need to monitor mental stress especially under work conditions is crucial to avoid health issues.

The electrical activity of the heart has been widely used to predict stress. The most common stress predictors from the ECG signal is heart rate (HR) and heart rate variability (HRV) [16]. Heart activity reflects the Autonomous Nervous System (ANS) balance or the imbalance in our body. Stress plays a major role in the balance of our ANS [10]. HRV is based on the variations between heartbeats (RR intervals) and it has been proven to be a reliable indicator of the autonomic nervous system's activity [11].

In recent years, several studies have been conducted on the detection of stress. HR and HRV features were utilized to train machine learning algorithms to classify stress levels. In the work of Fan et al. (2019), ECG data were collected from 15 participants under a mentally relaxed and stress state. HRV features were extracted using 2.5 minutes ECG segments and they achieved an accuracy of 80.56% using a k-nearest neighbor classifier [5]. Castaldo et al. (2016) achieved an accuracy of 79% from HRV features extracted from 3 minutes ECG excerpts using the C4.5 tree algorithm. The ECG signal was recorded from 42 university students during an oral exam for the stress state and during vacation for the non-stress state [3]. In the work of Kashan et al. (2015) ECG data was collected from 17 participant drivers, where the authors used a random tree classifier and achieved an accuracy of 88.24% for three levels of stress: low (initial and rest state), medium (highway), high (inside the city) [9].

The focus of this study was to classify the mental stress state (stress, no stress) of the participants while performing a hand-grip strength task. HR and HRV features were used to classify these states. The analysis of HR and HRV during a motor task has inherent

[1] www.who.int/news-room/fact-sheets/detail/aging-and-health

challenges but we were able to successfully use these features to classify "social stress" – a stressor that is more prevalent in older adults. In all the previously mentioned studies the stress state was classified while the stress elicited, in this study we classified the stress state of the participant after the stressor was induced, which increases model classification utility post stressor event.

In this study, HR and HRV features extracted from 1-minutes ECG signal segments and an Artificial Neural Network (ANN) was developed to predict different levels of stress on elderly people. Previous research has shown that HRV time domain and frequency domain features can be extracted using an ultra-short HRV window. According to Li et al. (2009) HRV features can be extracted using a window of size 30 seconds [14]. Also, the work of Salahuddin et al. (2007) indicates that both time domain and frequency domain features can be extracted in a 60 seconds window [22]. Our results indicate that stress levels on the elderly can be successfully detected from ECG signals using ultra short windows.

## 2 DATA COLLECTION

This study recruited a total of 57 participants, all of whom were residents of the local Bryan-College Station community. All participants were right-hand dominant, over 65 years in age (mean = 72.76±5.69), and self-reported sedentary lifestyles with occasional recreational activity. None of the participants reported any known musculoskeletal injuries or disorders within the past year. Participants were designated as non − obese if their body mass index (BMI) was in the range of 18.5 − 25 kg/m2 while those with BMI > 30 kg/m2 were considered obese. The participants were placed into four experimental groups – non-obese males (n = 15), non-obese females (n=15), obese males (n=14), and obese females (n = 15). All participants were subject to the experiment protocols approved by Texas A&M University's Institutional Review Board on providing written informed consent prior to data collection.

The stress data presented in this investigation derives from a larger study that entailed a sequence of experiments that spanned several days. Figure 1 shows the protocol we followed in this study. Upon informed consent, participants perform a one-handed grip stress experiment. This experiment protocol consisted of instrument setup, baseline measurements for strength (maximal voluntary contraction (MVC)), salivary cortisol samples, a Trier Social Stress Test (TSST) [8], and pre/post-stressor hand-grip performance test. During MVC and subsequent hand-grip experiment procedures, participants sat upright with their dominant upper arm at their side, elbows flexed at 90°, and lower arm supported by an armrest. Prior to the motor task, ECG signal, using a 3-lead chest configuration, was captured during a baseline period in which participants were asked to relax without any movement for three minutes.

Following baseline tests, participants grasped a hand dynamometer (BIOPAC, CA, United States) to perform isometric hand-grip contractions. After a warm-up period of roughly two minutes, participants performed three isometric MVCs with two-minute rest intervals to measure their hand-grip strength. The maximum value from the three MVC trials was used to determine a target force level of 30% MVC for the subsequent motor tasks pre-/post-stressor. Participants were provided familiarization of the hand-grip trials at 30% MVC and adequate rest before starting the next motor task. The

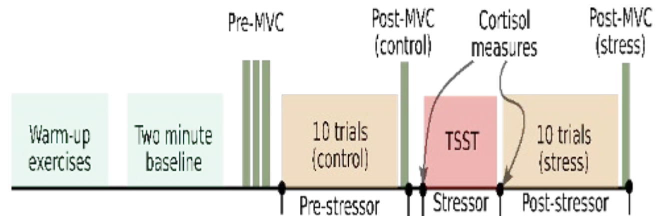


**Figure 1: Study Protocol**

pre/post-stressor tests required participants to maintain hand-grip force levels at 30% MVC for 15 s followed by 15 s rest repeatedly for ten sessions [28]. During this time participants were instructed to maintain their hand-grip force level as close to the target force level as possible utilizing real-time visual feedback on a computer screen.

After the pre-stressor hand grip sessions, salivary cortisol samples were taken from each participant. Following which participants were subject to a TSST protocol that entailed extemporaneous speech, and arithmetic tasks in front of an unfamiliar group of panelists (N=2). Following the TSST, another salivary cortisol sample was collected from the participants. After this phase participants were subject to the post-stressor experiment protocol that entailed a repetition of the 30% MVC handgrip exercise for ten sessions. Once they terminated the post-stressor task, they performed a post MVC trial to gauge levels of fatigue associated with the complete test protocol. In this protocol participant ECG data is acquired at 1000 Hz using a two lead ECG probe, and the BIOPAC ECG100C (BIOPAC Systems Inc., Aero Camino Goleta, CA, USA) amplifier interface. The salivary cortisol swabs serve as ground truth to confirm elevated stress-levels post-TSST, while the TSST serves as an accepted, proven stress induction paradigm for human factors studies.

## 3 METHODS

Figure 2 shows the steps we followed to accomplish stress classification from ECG signals. Data segmentation, data preprocessing, feature selection, feature extraction, model optimization, and model evaluation are explained in detail in the following sections.

### 3.1 Data Segmentation

In this study, the pre stressor test serves as the no-stress state and the post stressor test as the stress state. The total duration of each state for each subject was approximately five minutes. The next step was to segment each state in windows to extract the HR and HRV features for every participant. Each stress level was segmented using a window size of 60,000 (one minute) data points which corresponds to the amount of data copied from the ECG signal to our window.

### 3.2 Data Preprocessing

R–R interval time series obtained from the continuous ECG signal by detecting each QRS complex for every window we created at the segmentation process. Also, the normal-to-normal (NN) intervals

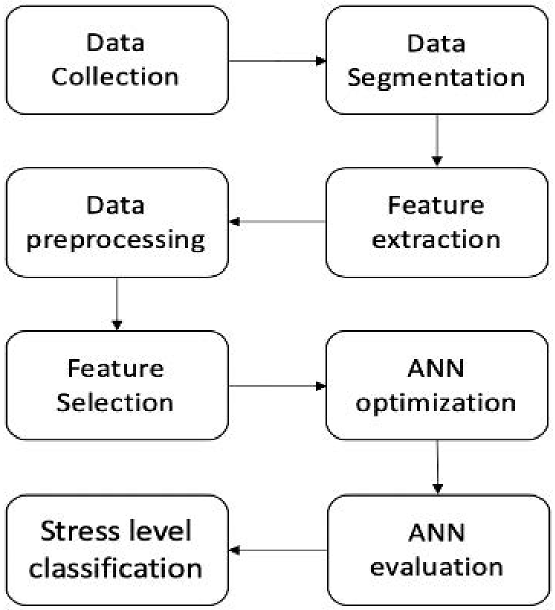


**Figure 2: Stress Level Classification Procedure**

that contained only R-R intervals computed from sinus node depolarizations was determined [17]. Furthermore, power spectrum computation was performed using a nonparametric, signal representation technique (wavelet transform). Wavelet Transform is a time-frequency analysis method. In this study, we applied Wavelet Transform to scale the decomposed ECG signal into different frequency band signals. After we eliminate the noises the ECG signal is reconstructed using only the useful parts of the original signal [1].

## 3.3 Feature Extraction

After segmented and processing the data HR and HRV features were extracted from the ECG signal for each window.

*3.3.1 Heart Rate Features.* Heart rate features. From Heart rate we extracted two features in total:

(1) HR_mean: The mean heart rate.
(2) 2. HR_std: The standard deviation of heart rate.

*3.3.2 Heart Rate Variability Features.* HRV features were extracted using two different methods: time-domain methods, frequency-domain methods [15, 20, 24, 27, 29].

*Time Domain Features.* In total, we extracted 9 features:

(1) sdNN: The standard deviation of the time interval between successive normal heartbeats.
(2) meanNN: The mean RR interval.
(3) RMSSD: The root mean square of the RR intervals during a period of time.
(4) CVSD: The coefficient of variation of successive differences, the RMSSD divided by meanNN.
(5) cvNN: The Coefficient of Variation, the ratio of sdNN divided by meanNN.
(6) medianNN: The median of the Absolute values of the successive Differences between the NN intervals.
(7) madNN: The median Absolute Deviation (MAD) of the RR intervals.
(8) mcvNN: The median-based Coefficient of Variation, the ratio of madNN divided by medianNN.
(9) pNN50: The proportion derived by dividing NN50 (The number of interval differences of successive NN intervals greater than 50 ms) by the total number of NN intervals.

*Frequency Domain Features.* In total, we extracted 9 features:

(1) VLF: The variance in HRV in the very low frequency (.003 to .04 Hz).
(2) LF: The variance in HRV in the low frequency (.04 to .15 Hz).
(3) HF: The variance in HRV in the high frequency (.15 to .40 Hz).
(4) Total Power: The total power of the density spectra.
(5) LFHF: The LF/HF ratio.
(6) LFn:The normalized LF power LFn = LF/(LF+HF).
(7) HFn: The normalized HF power HFn = HF/(LF+HF).
(8) LFp: The ratio between LF and Total Power.
(9) HFp: The ratio between HF and Total Power.

## 3.4 Feature Selection

The next step is feature selection. It is very important to choose the most informative features and at the same time features that eliminate redundant data. This will contribute to reducing the dimensionality and the complexity of the chosen predictive model [21]. Filter methods are very popular in determining the best subset of uncorrelated features. They are usually very fast, and they are not computationally intensive [21]. We applied Pearson's correlation to determine the highly correlated features. Pearson's correlation quantifies the linear dependence between two continuous variables X and Y. The correlation values vary from -1 to 1. Coefficient values close to -1 indicate a strong negative correlation. Values close to 1 indicate strong positive correlation and a value close to 0 implies weak correlation and a value of exact 0 implies no correlation [2]. Figure 3 shows the correlation degree between the features.

Highly correlated features were removed, and the final subset of features used to build our predictive model includes: HR_mean, HR_std, RMSSD, meanNN, VFL, VHL, and LFn.

## 3.5 Optimized ANN for Mental Stress Classification During a Motor Task.

Artificial Neural Networks can learn from data and generate models that receive a number of inputs and optimally map them to the desired outputs. ANNs learn and model non-linear and complex relationships [19]. Many machine learning algorithms impose restrictions on the input features, ANN is not one of them and it usually performs better with high volatility and non-constant variance data [4]. This is very important in our application because the volatility of the ECG signal is very high. After selecting the less correlated HR and HRV features we created an ANN. To determine the best hyperparameters for the neural network we applied the grid search method with 10-fold cross validation for each combination of the parameters [26]. Grid search is a brute-force approach where

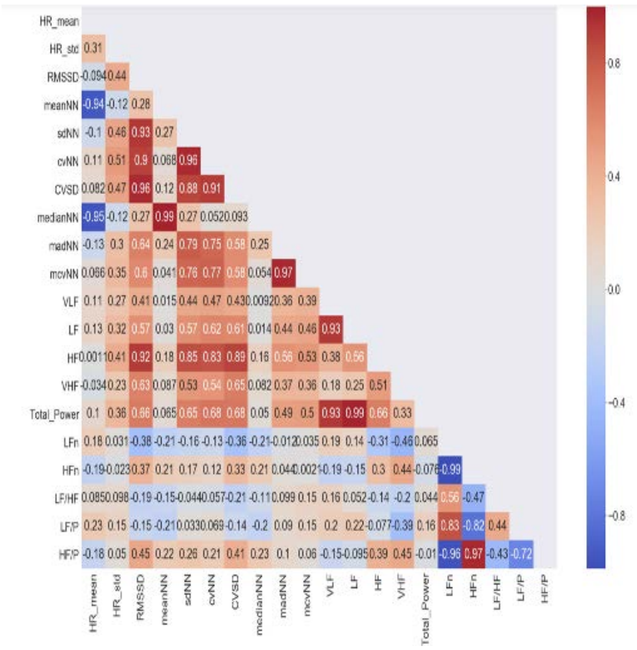

Figure 3: Correlation Between Features

every combination of a specified set of hyperparameter values is tried to find the optimal values for the neural network. In total five parameters were tuned: number of hidden neurons (range from 1 to 10), activation function for the hidden layer (identity, logistic, tanh, relu), the solver for weight optimization ( lbfgs , sgd, adam), learning rate schedule for weight updates (constant, invscaling , adaptive ), alpha penalty parameter (0.1, 0.05, 0.02, 0.01), and maximum number of iterations (100,150,200,250). We concluded that the optimal parameters for the neural network were achieved using a one hidden layer network with 4 hidden neurons. The optimal activation function was tanh, and Adam is determined as the optimal stochastic gradient optimizer. Also, according to the grid search method adaptive learning, regularization rate of alpha=0.1, maximum iteration= 100 and batch size=16 were selected as the best parameters for the neural network.

## 3.6 ANN Evaluation

The dataset that was created from the segmentation and the preprocess phase was used to test the performance of the developed ANN. The data set was split in two parts where 80% was used as the training set and the rest 20% as the test set [12]. The training and the testing set did not contain data from the same participant. We calculated the overall accuracy, the precision, recall and F1-Score of the developed model for the 80/20 training/testing split of the dataset.

$$Accuracy = (TP + TN)/(TP + FP + FN + TN) \quad (1)$$

$$Precision = (TP)/(TP + FP) \quad (2)$$

$$Recall = (TP)/(TP + FN) \quad (3)$$

$$F1 - Score = \text{ weight average of Precision and Recall} \quad (4)$$

Table 1: The Precision, Recall, F1-Score of the ANN Model.

| Classes | Precision | Recall | F1-Score |
|---|---|---|---|
| No-stress | 88% | 95% | 91% |
| stress | 95% | 87 % | 90% |
| macro avg | 91% | 91 % | 91% |

Where TP: True positive, TN: True negative, FP: False positive, and FN: False negative [25]. Also, a confusion matrix was created to provide information about the performance of the classification model and to give an insight on what the neural network misclassified [23]. Finally, we performed leave-one-subject-out cross validation [6] where, one participant was randomly selected for the testing purposes while the other participants were used for training the model. This procedure was repeated until all the participants had been used as test dataset.

# 4 RESULTS

The overall test accuracy for the 80/20 training/testing split is 90.83% for predicting two levels of mental stress ("stress", "no stress"). Table 1 shows the classification results of the developed ANN. The precision, recall, and f1-score for the no-stress state are 88%, 98%, 91% respectively. For the stress, the the precision, recall, and f1-score are 95%, 87%, 90% respectively. The unweighted mean precision, recall, and f1-score for the two classes is 91%.

Figure 4 demonstrates the performance of the developed ANN using a confusion matrix for the two levels of stress. The diagonal line shows the percentage of the correct data our model predicted while the off- diagonal boxes represent the percentage of the misclassified labels by the ANN.

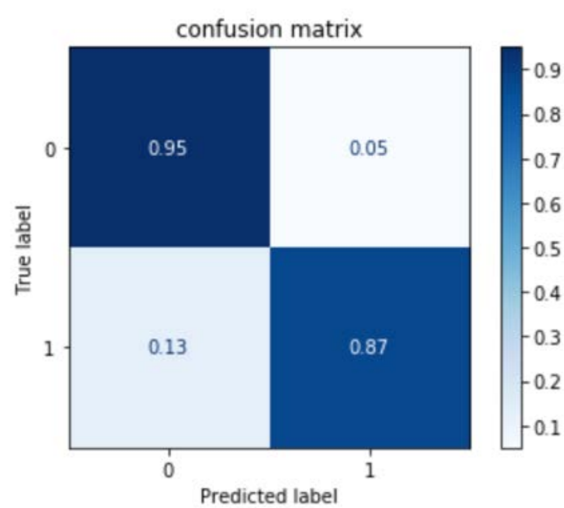


Figure 4: Confusion Matrix for the Two Levels of Stress

The mean accuracy of the leave-one-subject-out cross validation is 84.38%. The developed model performed well for the majority of the participants. The accuracy of thirty-four participants was lying between 90% and 100%. The developed model achieved accuracy from 80% to 89% for nine participants and accuracy from 70% to 79% for six participants. Finally, the accuracy of 5 participants was lying between 60% and 69% and an accuracy from 50% to 59% was achieved for three participants.

## 5 DISCUSSION

In this experiment we developed a laboratory model with known context labels and tests in the same environment. In order to classify the mental stress state only the ECG signal was used and only one machine learning algorithm was tried. Our future work will include the demographic information (age, body mass index) and several machine learning algorithms will be used to classify the mental stress state of an older adult while performing a motor task. Also, from the leave-one-subject-out cross validation results we noticed variations in the accuracy across the participants. The model performed very well for the majority of the participants. For a small percentage of the participants the accuracy was low, we assume that the stress induction was less effective to these participants.

## 6 CONCLUSION

Mental Stress can negatively affect our health and especially the health of older adults. Automated stress detection can be proven very helpful in combating the negative implications of stress, it can lead to better stress management and can have beneficial effects in the quality of their life. In this study we developed an ANN to predict two levels of mental stress in older adults while performing a motor task. The results demonstrate that our optimized model can accurately predict stress levels from ECG signals. We achieved overall test accuracy of 90.83%.

## ACKNOWLEDGMENTS

We thank Texas A&M for providing the data and the director of the Artificial Intelligence and Data science lab at California State University, Los Angeles Dr. Mohammad Pourhomayoun for the comments raised from the discussions with him. Research reported in this publication was partially supported by the National Institute on Aging of the National Institutes of Health under Award Number R15AG047553 and from the National Science Foundation under award number 1900704. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health and National Science Foundation.

## REFERENCES

[1] Paul S Addison. 2005. Wavelet transforms and the ECG: a review. *Physiological measurement* 26, 5 (2005), R155.

[2] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. 2009. Pearson correlation coefficient. In *Noise reduction in speech processing*. Springer, 1–4.

[3] Rossana Castaldo, William Xu, Paolo Melillo, Leandro Pecchia, Lorena Santamaria, and C James. 2016. Detection of mental stress due to oral academic examination via ultra-short-term HRV analysis. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 3805–3808.

[4] Lukas Falat, Zuzana Stanikova, Maria Durisova, Beata Holkova, and Tatiana Potkanova. 2015. Application of neural network models in modelling economic time series with non-constant volatility. *Procedia Economics and Finance* 34 (2015), 600–607.

[5] Jianzhong Fan, Haihui Li, Yu Zhan, and Yajun Yu. 2019. An Electrocardiogram Acquisition and Analysis System for Detection of Human Stress. In *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 1–6.

[6] Yan Gao, Xinyu Zhang, Shouyang Wang, and Guohua Zou. 2016. Model averaging based on leave-subject-out cross-validation. *Journal of Econometrics* 192, 1 (2016), 139–151.

[7] Michael D Giandrea, Kevin E Cahill, and Joseph F Quinn. 2009. Bridge jobs: A comparison across cohorts. *Research on Aging* 31, 5 (2009), 549–576.

[8] DH Hellhammer, C Kirschbaum, and K Pirke. 1993. The 'trier social stress test'-A tool for investigating psychobiological stress response in a laboratory setting. *Neuropsychobiology* 28 (1993), 76–81.

[9] N Keshan, PV Parimi, and Isabelle Bichindaritz. 2015. Machine learning for stress detection from ECG signals in automobile drivers. In *2015 IEEE International Conference on Big Data (Big Data)*. IEEE, 2661–2669.

[10] Desok Kim, Yunhwan Seo, Jaegeol Cho, and Chul-Ho Cho. 2008. Detection of subjects with higher self-reporting stress scores using heart rate variability patterns during the day. (2008), 682–685.

[11] Hye-Geum Kim, Eun-Jin Cheon, Dai-Seg Bai, Young Hwan Lee, and Bon-Hoon Koo. 2018. Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry investigation* 15, 3 (2018), 235.

[12] Kristjan Korjus, Martin N Hebart, and Raul Vicente. 2016. An efficient data partitioning to improve classification performance while keeping parameters interpretable. *PloS one* 11, 8 (2016), e0161788.

[13] Braedyn Kromer and David Howard. 2013. Labor force participation and work status of people 65 years and older. (2013).

[14] Zhibin Li and Harold Snieder. [n.d.]. Shaoyong Su, Xiuhua Ding, Julian F. Thayer, Frank A. Treiber, and Xiaoling Wang. 2009."A Longitudinal Study in Youth of Heart Rate Variability at Rest and in Response to Stress.". *International Journal of Psychophysiology* 73, 3 ([n. d.]), 212–217.

[15] NEAL Lippman, KENNETH M Stein, and BRUCE B Lerman. 1994. Comparison of methods for removal of ectopy in measurement of heart rate variability. *American Journal of Physiology-Heart and Circulatory Physiology* 267, 1 (1994), H411–H418.

[16] Ulf Lundberg, Roland Kadefors, Bo Melin, Gunnar Palmerud, Peter Hassmén, Margareta Engström, and Ingela Elfsberg Dohns. 1994. Psychophysiological stress and EMG activity of the trapezius muscle. *International journal of behavioral medicine* 1, 4 (1994), 354–370.

[17] Marek Malik, J Thomas Bigger, A John Camm, Robert E Kleiger, Alberto Malliani, Arthur J Moss, and Peter J Schwartz. 1996. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European heart journal* 17, 3 (1996), 354–381.

[18] World Health Organization. 2016. World health statistics 2016: monitoring health for the SDGs sustainable development goals. (2016).

[19] Y-S Park and S Lek. 2016. Artificial Neural Networks: Multilayer Perceptron for Ecological Modeling. In *Developments in environmental modelling*. Vol. 28. Elsevier, 123–140.

[20] Mirja Peltola. 2012. Role of editing of RR intervals in the analysis of heart rate variability. *Frontiers in physiology* 3 (2012), 148.

[21] Mohammad Pourhomayoun, Nabil Alshurafa, Bobak Mortazavi, Hassan Ghasemzadeh, Konstantinos Sideris, Bahman Sadeghi, Michael Ong, Lorraine Evangelista, Patrick Romano, Andrew Auerbach, et al. 2014. Multiple model analytics for adverse event prediction in remote health monitoring systems. In *2014 IEEE Healthcare Innovation Conference (HIC)*. IEEE, 106–110.

[22] Lizawati Salahuddin, Jaegeol Cho, Myeong Gi Jeong, and Desok Kim. 2007. Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. In *2007 29th annual international conference of the ieee engineering in medicine and biology society*. IEEE, 4656–4659.

[23] Claude Sammut and Geoffrey I Webb. 2011. *Encyclopedia of machine learning*. Springer Science & Business Media.

[24] Anne-Louise Smith, Harry Owen, and Karen J Reynolds. 2013. Heart rate variability indices for very short-term (30 beat) analysis. Part 1: survey and toolbox. *Journal of clinical monitoring and computing* 27, 5 (2013), 569–576.

[25] Kai Ming Ting. 2010. Precision and Recall. *Encyclopedia of machine learning* 781 (2010).

[26] Joaquin Vanschoren. 2019. Meta-learning. In *Automated Machine Learning*. Springer, Cham, 35–61.

[27] Andreas Voss, Rico Schroeder, Andreas Heitmann, Annette Peters, and Siegfried Perz. 2015. Short-term heart rate variability—influence of gender and age in healthy subjects. *PloS one* 10, 3 (2015), e0118308.

[28] W West, A Hicks, L Clements, and J Dowling. 1995. The relationship between voluntary electromyogram, endurance time and intensity of effort in isometric handgrip exercise. *European journal of applied physiology and occupational physiology* 71, 4 (1995), 301–305.

[29] Ada H Zohar, C Robert Cloninger, Rollin McCraty, et al. 2013. Personality and heart rate variability: exploring pathways from personality to cardiac coherence and health. *Open Journal of Social Sciences* 1, 06 (2013), 32.