

# Case 2 Report: Detecting Stress Using Wearables

Bob Ding, Cathy Lee, Malavi Ravindran

```
library(zoo)
```

```
##  
## Attaching package: 'zoo'  
## The following objects are masked from 'package:base':  
##  
##      as.Date, as.Date.numeric
```

```
library(varian)
```

```
## Loading required package: rstan  
## Loading required package: StanHeaders  
## Loading required package: ggplot2  
## rstan (Version 2.21.2, GitRev: 2e1f913d3ca3)  
## For execution on a local, multicore CPU with excess RAM we recommend calling  
## options(mc.cores = parallel::detectCores()).  
## To avoid recompilation of unchanged Stan programs, we recommend calling  
## rstan_options(auto_write = TRUE)
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --  
## v tibble  3.0.3      v dplyr    1.0.0  
## v tidyr   1.1.0      v stringr 1.4.0  
## v readr   1.3.1      v forcats 0.5.0  
## v purrr   0.3.4  
  
## -- Conflicts ----- tidyverse_conflicts() --  
## x tidyr::extract() masks rstan::extract()  
## x dplyr::filter()  masks stats::filter()  
## x dplyr::lag()     masks stats::lag()
```

```
library(corrplot)
```

```
## corrplot 0.84 loaded
```

```
library(glmnet)
```

```
## Loading required package: Matrix  
##  
## Attaching package: 'Matrix'  
## The following objects are masked from 'package:tidyr':
```

```
##
##      expand, pack, unpack
## Loaded glmnet 4.0-2
library(FactoMineR)
library(factoextra)

## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
library(MASS)

##
## Attaching package: 'MASS'
## The following object is masked from 'package:dplyr':
##
##      select
library(caret)

## Loading required package: lattice
##
## Attaching package: 'caret'
## The following object is masked from 'package:purrr':
##
##      lift
library(MLmetrics)

##
## Attaching package: 'MLmetrics'
## The following objects are masked from 'package:caret':
##
##      MAE, RMSE
## The following object is masked from 'package:base':
##
##      Recall
library(pROC)

## Type 'citation("pROC")' for a citation.
##
## Attaching package: 'pROC'
## The following objects are masked from 'package:stats':
##
##      cov, smooth, var
library(gridExtra)

##
## Attaching package: 'gridExtra'
## The following object is masked from 'package:dplyr':
##
```

```
##      combine

library(grid)
library(broom)
library(knitr)
library(sjPlot)

## Install package "strengjacke" from GitHub (`devtools::install_github("strengjacke/strengjacke")`)

library(RColorBrewer)
library(kableExtra)

##
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':
##
##      group_rows

features_all_modalities = function(fn, win_size, shift) {
  df_all = read.csv(fn) %>%
    mutate(ACC_chest_3D = sqrt(ACC_chest_X^2+ACC_chest_Y^2+ACC_chest_Z^2)) %>%
    mutate(ACC_wrist_3D = sqrt(ACC_wrist_x^2+ACC_wrist_y^2+ACC_wrist_z^2))
  drops <- c("X", "Label", "subject")
  df = df_all[ , !(names(df_all) %in% drops)]
  replace_rows = length(rollapply(df[,1], width = win_size*4, by = shift, FUN = mean, align = "left"))
  features_df <- data.frame(matrix(ncol = ncol(df)*4, nrow = replace_rows))
  new_names = sapply(1:length(df), function(c) {
    c(paste0(colnames(df)[c], "_mean"),
      paste0(colnames(df)[c], "_sd"),
      #paste0(colnames(df)[c], "_range"),
      paste0(colnames(df)[c], "_min"),
      paste0(colnames(df)[c], "_max")
      # paste0(colnames(df)[c], "_skew")
    )
  })
  colnames(features_df) = new_names

  for (c in 1:length(df)) {

    # finding mu
    mu_vals = rollapply(df[,c], width = win_size*4, by = shift, FUN = mean, align = "left")
    new_col_name = paste0(colnames(df)[c], "_mean")
    cindx = which(colnames(features_df)==new_col_name)
    features_df[, cindx] = mu_vals

    # finding sd
    sd_vals = rollapply(df[,c], width = win_size*4, by = shift, FUN = var, align = "left")
    new_col_name = paste0(colnames(df)[c], "_sd")
    cindx = which(colnames(features_df)==new_col_name)
    features_df[, cindx] = sqrt(sd_vals)
  }
}
```

```

# finding max
max_vals = rollapply(df[,c], width = win_size*4, by = shift, FUN = max, align = "left")
new_col_name = paste0(colnames(df)[c], "_max")
cindx = which(colnames(features_df)==new_col_name)
features_df[, cindx] = max_vals

# finding min
min_vals = rollapply(df[,c], width = win_size*4, by = shift, FUN = min, align = "left")
new_col_name = paste0(colnames(df)[c], "_min")
cindx = which(colnames(features_df)==new_col_name)
features_df[, cindx] = min_vals

}

# make sure merge label, subject back into feature dataframe
features_df$Label = df_all[1:replace_rows, "Label"]
features_df$Subject = df_all[1:replace_rows, "subject"]
return(features_df)
}

HR_calc = function(df, win_size, shift) {
  df <- df[rep(seq_len(nrow(df)), each = 4), ]
  mu_vals = rollapply(df, width = win_size*4, by = shift, FUN = mean, align = "left")
  sd_vals = rollapply(df, width = win_size*4, by = shift, FUN = sd, align = "left")
  min_vals = rollapply(df, width = win_size*4, by = shift, FUN = min, align = "left")
  max_vals = rollapply(df, width = win_size*4, by = shift, FUN = max, align = "left")

  return(list(mu = mu_vals, sd = sd_vals,
             min = min_vals, max = max_vals))
}

file_list <- list.files()
subject_data = file_list[grepl("df_S", file_list, fixed = TRUE)]

ALL_df = NULL
for (i in 1:length(subject_data)) {
  print(i)
  feature_data = features_all_modalities(subject_data[i], 5, 1) #CHANGE FILE PATH
  subject_no = gsub("\\\\.\\.*", "", sub('.*_-', '', subject_data[i]))
  HR_fn = paste0("~/case-study-2/WESAD/", subject_no, "/", subject_no, "_E4_Data/HR.csv") #CHANGE FILE PATH
  HR_data = read.csv(HR_fn)[-1,]
  hr = HR_calc(as.data.frame(HR_data), 5, 1)

  df_hr = data.frame(matrix(unlist(hr), nrow= length(hr$mu),
                             ncol=4, byrow = F))
  colnames(df_hr) = c("hr_wrist_mu", "hr_wrist_sd", "hr_wrist_min", "hr_wrist_max")#, "hr_wrist_range")
  df_hr$ID =seq.int(nrow(df_hr))

  feature_data = feature_data[-c(1:40),]
  feature_data$ID <- seq.int(nrow(feature_data))

  S_df = merge(feature_data, df_hr, by="ID")
}

```

```

ALL_df = rbind(ALL_df, S_df)
}

## [1] 1
## [1] 2
## [1] 3
## [1] 4
## [1] 5
## [1] 6
## [1] 7
## [1] 8
## [1] 9
## [1] 10
## [1] 11
## [1] 12
## [1] 13
## [1] 14
## [1] 15

ALL_df = ALL_df %>% filter(Label %in% c("2","3"))
write.csv(ALL_df, "master_df.csv")

```

## 1 Introduction

Stress can have disastrous long-term effects on the human body [1]. In fact, in 2015, the British Health and Safety Executive found that 37% of work-related illnesses were attributed to stress alone. One way to recognize and mitigate stress is through automated detection methods [2]. Existing wearable devices, which can be wrist or chest-worn, are able to gather relevant physiological data on wearers. Statistical methods can then be applied to construct a mapping from observed sensor data to the wearer’s affective state (e.g. stress). As certain affective states, namely amusement, have similar physiological markers to stress, it is often a difficult task to discriminate between them.

This case study seeks to determine whether sensor data is useful in discriminating between stress and amusement conditions as well as understand the relationship between various physiological features and stress. It further aims to discover which types of sensor data are most useful in discriminating between amusement and stress—that is, can a model built from only wrist sensor data adequately detect stress, or is a combination of wrist and chest-worn sensor data considerably better? Finally, the study seeks to provide a quantification of heterogeneity across different individuals in the response to stress versus amusement. In order to address each of these objectives, we will use a database provided by Schmidt et al [2]. In section , **we will provide a comprehensive overview of the data as well as describe our feature engineering process.** Section explores the efficacy of wrist-only versus combined sensor data in detecting stress and proposes a logistic regression model with principal components as features. Section \_\_\_\_ describes the heterogeneity in stress response among subjects in the study. Our final section discusses limitations and conclusions.

## 2 Data

### 2.1 Description of Data

A total of 17 individuals participated in the original study. However, due to sensor malfunction for two subjects, only data for 15 subjects was considered in our analysis. Raw data was recorded by two sensor devices: the RespiBAN [3], which is chest worn, and the Empatica E4 [4], which is wrist worn. From the RespiBAN, the following modalities were measured for each individual at 700 Hz: *Electrocardiogram (ECG)*, *Electrodermal*

*Activity (EDA)*, *Electromyogram (EMG)*, *Skin temperature (TEMP)*, and *3-axis accelerometry (ACC)*. From the Empatica E4, the following modalities were measured for each individual: *3-axis accelerometry (ACC, 32 Hz)*, *Blood Volume Pulse (BVP, 64 Hz)*, *Electrodermal Activity (EDA, 4 Hz)*, *Skin temperature (TEMP, 4 Hz)*, *Heart Rate (HR, 1 Hz)*.

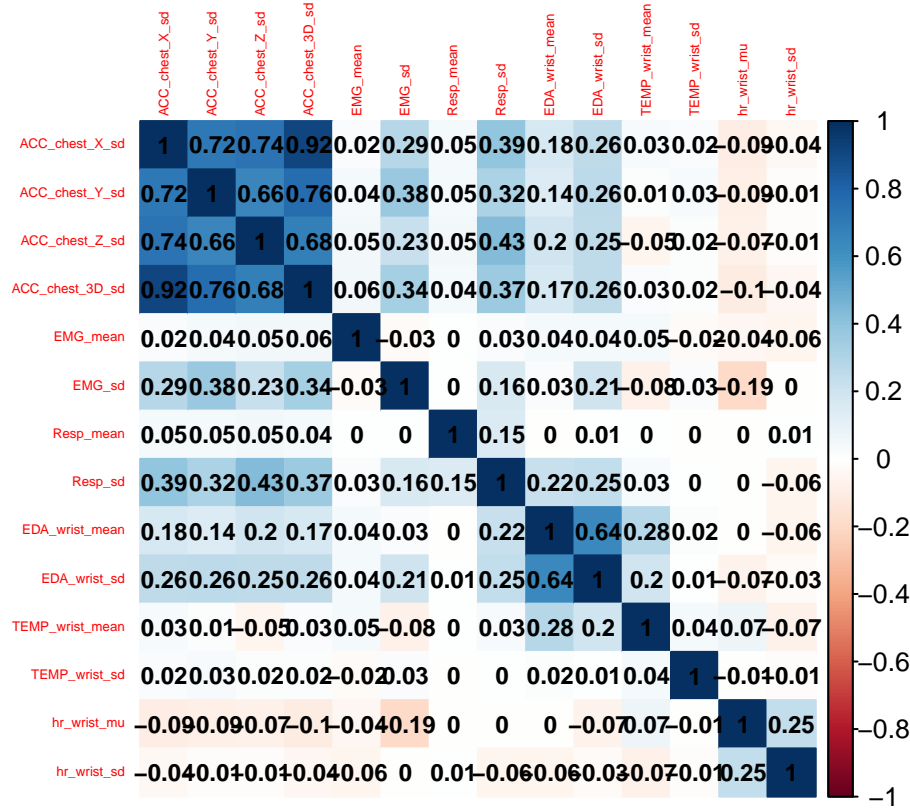
ECG measurements record electrical signals in the heart, and are useful in monitoring heart health [5]. EMG for the chest measures muscle response to brain signals [6]. EDA for both wrist and chest are measures of neurally mediated effects on sweat gland permeability [7]. TEMP measures the skin’s temperature, in which variability can be an indicator of stress [8]. ACC for both wrist and chest is used to record horizontal, vertical, and forward-backward acceleration of object movement. Studies have shown certain additional predictive power for stress detection can be gained by incorporating ACC data into the predictive model [9]. BVP measures the volume of blood that passes through tissues with each beat of the heart [10]. Finally, HR, or the number of heart beats per minute, has been found to vary empirically with affective state [11]. It is important to note that HR was not directly measured by the Empatica E4 device, but was instead derived from the BVP measure using a proprietary algorithm [12].

## 2.2 Feature Engineering

In order to combine each of the modalities from the two sensors, we first downsampled each modality (barring heart rate) to 4 Hz. As heart rate was recorded only once per second, we repeated each value four times in order to provide a proxy for a 4 Hz measurement. Using the accelerometer data for the individual X, Y, and Z axes, we derived an additional measure, ACC 3D, representing the magnitude of total acceleration. We then proceeded to segment these sensor signals into window sizes of 5 seconds with 0.25 second shifts. Within each window the following features were engineered for each modality: *mean*, *standard deviation*, *minimum value*, and *maximum value*. We adapted this feature engineering process from [13]. Finally, as our study seeks to discriminate between stress and amusement, we filtered our data for only those observations in which the two states were observed.

## 2.3 Exploratory Data Analysis

**Fig. 1 Correlation Matrix of Engineered Features**



From the correlation matrix (Fig. 1), we see that the ACC variables (variables related to motion) are highly correlated. We also see some moderate correlation between Resp and EDA variables, as well as between EDA and TEMP variables.

Fig. 2 Distribution of Mean

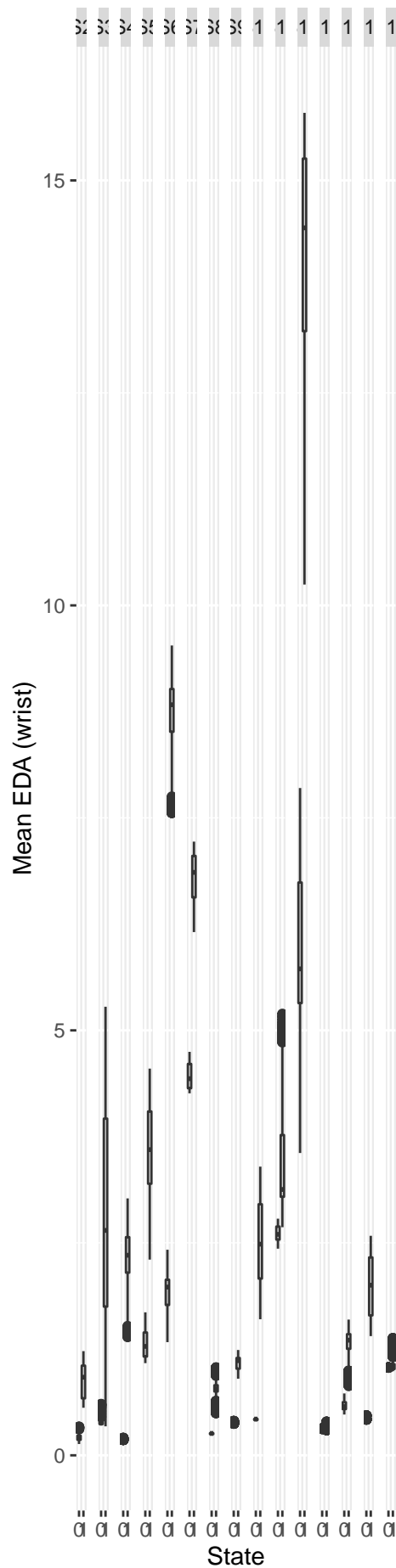
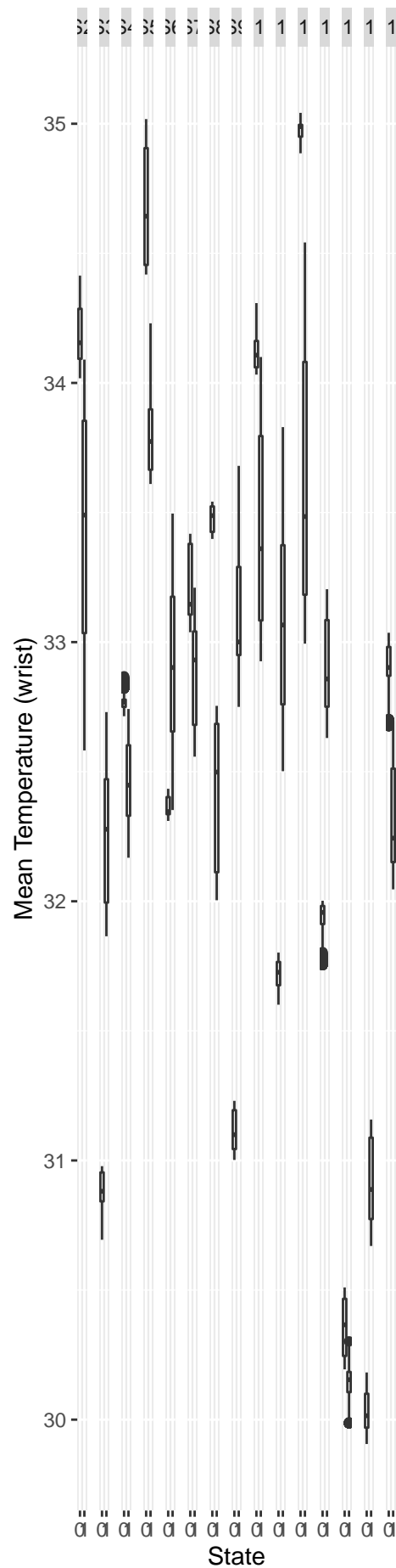


Fig. 3 Distribution of Mean





After examining boxplots of the distribution of the physiological measures by state for each subject (see Appendix A.1 for all plots of this type), the variables that had the greatest difference in distribution between states and across subjects are `EDA_wrist_mean`, `Temp_wrist_mean`, and `hr_wrist_mu`. For all subjects, mean EDA values (Fig. 2) are higher in the stress state, and for subjects 3, 5, 6, 7, 10, 13, the difference in the distribution of EDA between the two states is more drastic than that of the other subjects. For `Temp_wrist_mean` (Fig. 3), we can see that for some subjects, their mean temperature tends to be higher in the stress state but for others, their mean temperature tends to be lower when stressed. Likewise, for `hr_wrist_mu` (Fig. 4), we see that some subjects have higher average heart rate when stressed while others exhibit the opposite behavior. Among all of the other variables, it appears that the motion (`ACC`) variables tend to show a higher distribution for each subject under the stress state, and the `EMG` and `Resp` variables tend not to vary much in distribution between states across subjects. In addition, it is important to note that for certain subjects (such as subject 13), there is no overlap in mean EDA values or mean Temperature values between the two states. This is likely to cause perfect separation in logistic regression (see Conclusion for a discussion on potential changes to the experimental design to deal with this issue).

As our exploratory analysis suggests, high collinearity exists among engineered features. Collinearity among covariates can be dangerous in providing unreasonable coefficient estimates with inflated standard errors [14]. As one of our main objectives is to understand the relationship between certain physiological attributes and affective state, it is imperative that model estimates are trustworthy. For this reason, domain driven insights were useful in selecting certain features to keep in the model, while omitting those contributing to greater multicollinearity. For example, features related to BVP and ECG measurements were ultimately excluded from our analysis due to their strong relationship with heart rate. As noted earlier, the Empatica E4 provides a measure of heart rate that is algorithmically derived from BVP. Heart rate can similarly be derived from ECG measurements [15]. Thus, in avoiding redundancy of information, BVP and ECG features were removed from the analysis, while those regarding heart rate remained. We further opted to remove the mean values of each accelerometry measurement (X, Y, Z, and 3D) over each of the five second windows. Our reasoning for excluding these covariates was due to our belief that variance in motion, which was captured by the standard deviation of accelerometry measurements in each window, is more strongly associated with affective state than average motion [16]. Finally, although we originally computed the minimum and maximum values for each physiological modality, we ultimately omitted these due to issues of underdispersion [17]. Specifically, our exploratory data analysis found that the minimum and maximum measurements across various modalities exhibited small variance. Figure \_\_\_ in the appendix provides a concrete visualization of these low variance features.

## 3 Methods

### 3.1 Model Selection

After reducing the dimensionality of our feature space, we sought to determine the efficacy of the wrist sensor alone in discriminating stress and amusement. For this purpose, we created two separate logistic regression models- one with data coming only from the wrist sensor, and the other using a combination of wrist and chest data. In these two datasets, we continued to explore methods to reduce dimensionality and account for high correlation between covariates. Techniques such as stepwise variable selection, lasso regression, and elastic net regression were considered, but found ineffective in reducing multicollinearity (see \_\_\_\_\_ in appendix for more detail). In both the wrist-only and combined data models, we ultimately settled on principal component analysis (PCA) as a means of reducing dimensionality and multicollinearity. Based on figures \_\_\_\_\_ and \_\_\_\_\_, we chose to utilize 4 components in both models. For both wrist-only and combined data, original features were normalized before the extraction of principal components, as is recommended for PCA (<https://www.theanalysisfactor.com/tips-principal-component-analysis/#:~:text=2.,have%20similar%20scales%20of%20measurement.&text=Variables%20whose%20numbers%20are%20just,th>). Our wrist data considered the following original features, before PCA was performed. Wrist ACC (X, Y, Z, and 3D): standard deviation Wrist EDA: standard deviation and mean Wrist temperature: standard deviation and mean Heart Rate: standard deviation and mean

Our combined wrist and chest data considered the following original features, before PCA was performed: Chest ACC (X, Y, Z, and 3D): standard deviation EMG: standard deviation and mean Wrist EDA: standard deviation and mean Wrist temperature: standard deviation and mean Heart Rate: standard deviation and mean In the combined data, ACC measurements coming from the chest sensor were considered, as opposed to the wrist. This choice was motivated by Table 6 in Schmidt et. al, which shows the importance of ACC features derived from the chest sensor, but not wrist. The same table highlights the importance of temperature and EDA related features coming from the wrist worn device, but not chest, warranting our inclusion of temperature and EDA features coming from the wrist sensor in our combined data model. The other features for the combined data come from heart rate, which was only provided by the wrist-worn device, and EMG, which comes from the chest-worn.

PCA yielded four components, which are simply linear combinations of the original features. Figures \_\_\_\_ and \_\_\_\_ show how much each original covariate loaded onto, or contributed to, each of the four components in the wrist-only and combined data models. Darker shades indicate a stronger contribution from a particular covariate to the principal component. As can be seen, in both the wrist-only and combined data, resulting principal components are fairly interpretable. For example, in the first components, standard deviations of ACC features have a strong loading. Thus, this component can be interpreted as a proxy for movement. Temperature and EDA are the main contributors to the second component, showing that it represents information related to skin temperature and sweat. The third component relates to heart rate, while the fourth \_\_\_\_\_.

We fit two separate logistic regression models using the principal components extracted from the wrist-only and combined data, respectively. We investigated the prediction accuracy and f-1 scores of each model using a 5-fold cross validation approach. In order to ensure that certain subjects were not over or underrepresented in the testing and training sets, we created stratified folds in which each individual’s representation was proportional to that of her representation in the entire dataset. Results are summarized in Table \_\_\_\_.

## 3.2 Cross Validation

## 4 References

1. [https://www.nlm.nih.gov/health/publications/stress/index.shtml?fbclid=IwAR0dCMkPveUAxNd\\_JNih1SQ3-a8wGBJ66Z9EkVZ4lBH\\_Ayw4jzTSeN4M3Y#:~:text=Over%20time%2C%20continued%20strain%20](https://www.nlm.nih.gov/health/publications/stress/index.shtml?fbclid=IwAR0dCMkPveUAxNd_JNih1SQ3-a8wGBJ66Z9EkVZ4lBH_Ayw4jzTSeN4M3Y#:~:text=Over%20time%2C%20continued%20strain%20)
2. SCHMIT PAPER
3. <https://www.mayoclinic.org/tests-procedures/ekg/about/pac-20384983>
4. <https://www.webmd.com/brain/emg-and-nerve-conduction-study#1>
5. [https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-1005-9\\_13#:~:text=Electrodermal%20activ](https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-1005-9_13#:~:text=Electrodermal%20activ)
6. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4664114/#:~:text=Acute%20stress%20triggers%20peripheral%2>
7. <https://ieeexplore.ieee.org/document/7949491>
8. <https://www.biofeedback-tech.com/articles/2016/3/24/the-blood-volume-pulse-biofeedback-basics#:~:text=The%20>
9. <https://www.heart.org/en/healthy-living/healthy-lifestyle/stress-management/stress-and-heart-health>
10. (<https://support.empatica.com/hc/en-us/articles/360029719792-E4-data-BVP-expected-signal>)
11. [https://github.com/DigitalBiomarkerDiscoveryPipeline/Human-Activity-Recognition/blob/master/10\\_code/30\\_end\\_pre\\_processing/32\\_engineer\\_features/33\\_feature\\_engineering.ipynb](https://github.com/DigitalBiomarkerDiscoveryPipeline/Human-Activity-Recognition/blob/master/10_code/30_end_pre_processing/32_engineer_features/33_feature_engineering.ipynb)
12. (<https://statisticsbyjim.com/regression/multicollinearity-in-regression-analysis/#:~:text=Multicollinearity%20re>
13. (<https://medicalxpress.com/news/2018-07-heart-affects.html>).
14. (<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.705.8139&rep=rep1&type=pdf>)

15. (<https://support.minitab.com/en-us/minitab/19/help-and-how-to/quality-and-process-improvement/control-charts/supporting-topics/understanding-attributes-control-charts/overdispersion-and-underdispersion/#:~:text=Underdispersion%20exists%20when%20data%20exhibit,other%2C%20also%20known%20as%20autocorr>)