17

Validation and Verification Report

PREDICTION OF SEIZURES DURING SLEEP

GROUP 17: JACK LIN, JOSH OLICK-GIBSON, NIKHIL PATEL

MARCH 2ND, 2018

Table of Contents

Changes from Progress Report

Project Need 1

Project Scope 1

Team Responsibilities 1

Design Schedule 1

Design Specifications 2

Validation and Verification Plan lo

Testable specifications 3

Non-Testable Specifications 4

Proof of Concept

ECG interpretation to obtain heart rate and heart rate variability 5

Training of a decision tree model to perform data classification 5

Simulation of ECG data with variable parameters 6

Collection of ECG data via Arduino 6

Overall Status and Next Steps 7

Appendix A: Design schedule

Appendix B: User satisfaction assessment form

Appendix C: Sample output of ECG interpretation software

Appendix D: Sample decision tree and prediction

Appendix E: Sample ECG simulations using MATLAB

Appendix F: View of Arduino Circuit

Appendix G: Sample ECG data collected using Arduino

References

1.0 Project Update

1.1 Project Need

No change

1.2 Project Scope

This project proposes to deliver a safe, comfortable, and easy to use device that can be worn overnight. This updated project scope incorporates the use of heart rate, heart rate variability, and other measurements from available ECG data, which will be fed into a decision tree algorithm to accurately predict an impending seizure, and will alert the user a short time (on the order of minutes) before the seizure begins. The proposed device will be delivered to Dr. David Lardizabal and the BME 401 instructors, including the necessary software for seizure prediction and a user manual for safe operation by the end of April 2018.

1.3 Team Responsibilities

Most team responsibilities have remained the same. All three members have been working on different pieces of the device that will be integrated together to make a final prototype and all members have contributed to the validation and verification report. Nikhil is responsible for processing the raw ECG data and developing the decision tree machine learning software. Jack is responsible for testing the existing patient data, as well as simulating ECG data for patients with and without seizures. Josh is responsible for signal conditioning and all issues related to hardware and interfacing.

1.4 Design Schedule

An updated design schedule is presented in Appendix A.

1.5 Design Specifications

The specification for range was removed because all of the components in our device are connected by wires and do not receive any signals wirelessly. We modified part of our specification for detection alert in that the device will no longer alert EMS but instead just the user.

2.0 Validation and Verification Plan

The validation and verification plan proposed is broken down according to each design specification. Because some specifications are not testable, they will be verified by inspection at the end of the project. Because we have a very limited set of data from real patients, we will prepare computer simulations of ECGs from patients both at baseline and during a seizure. The characteristics of the ECGs will be determined using information from the literature. Using a range of values for features such as heart rate, heart rate variability, QRS amplitude, QRS width, etc. we will simulate 100 ECGs for 10 minutes each that have randomized features, durations, and seizure start times. Once the simulations have been produced, they will be used to train the decision tree algorithm. Each 5 second window is considered a sample and will be labeled as baseline, pre-ictal, or ictal. After this, another set of simulations will be produced and will be used to test various components of the software. While this will not perfectly simulate a typical human ECG, it will allow us to test if our algorithm can identify the differences between baseline and seizure, and predict the seizure ahead of time. In order to test a human ECG, the Arduino setup will be used. However, because this device cannot be tested on a person having a real seizure, the Arduino will not be a sufficient simulation for our device to prove it’s accuracy. The data collected from our client will be used for additional testing after the model has been trained. A user feedback form that will be used to further validate the device is presented in Appendix B.

2.1 Testable Specifications

2.1.1 Operating time:

The prototype for this device is going to require a computer and an Arduino that will remain operational overnight. Because the computer will be running a power-intensive prediction algorithm, it will likely need to remain connected to a power source overnight. This will be tested by examining the power consumption of the device over the course of 8 hours using a 15 inch MacBook Pro with a 2.2GHz Intel Core i7 processor to see if it is possible for the device to remain disconnected overnight. Because the Arduino will be wired to the computer, it will have a consistent power supply so long as the computer is operational.

2.1.2 Prediction interval:

The prediction interval for a given simulation is defined as the time between the prediction made by the software and the time of seizure onset. After the software has been trained using 100 simulations, the prediction interval will be computed for each successive training simulation. The prediction interval for the device will be defined as the 10-90% range of prediction intervals for the test simulations.

2.1.3 Prediction alert:

Using the simulations, we will identify the ability of the software to predict a seizure when one is about to occur.

2.1.4 Detection alert:

Using the simulations, we will identify the ability of the software to detect a seizure.

2.1.5 Sensitivity:

Using the simulations, we will identify the sensitivity with which the software was able to predict a seizure before one occured. This will be done using the following formula:

$$Sensitivity = \frac{True\ positives}{Positives} = \frac{True\ positives}{True\ positives + False\ negatives}$$

2.1.6 Size:

We will measure the prototype with a ruler to ensure that it is within 100 cm^2.

2.1.7 Weight:

We will mass the prototype with a scale in order to ensure that it is under 250 g.

2.2 Non - Testable Specifications

2.2.1 Safety:

Safety will be tested by inspection to ensure that the device will have minimal sharp edges and cords that can entangle the user. Wires will be properly insulated to avoid any possible electrical shock.

2.2.2 Comfort:

Comfort will be tested by inspection. Similar to safety, the device should have minimal sharp edges, cords that can tangle and insulated wires so that the patient is able to sleep comfortably.

2.2.3 Consumer/Production Cost:

The production cost should not exceed $300. The consumer cost should not exceed $1000.

2.2.4 Transmission:

Device should transmit data to the receiver with minimal wiring and time delay.

2.2.5 Security:

The assessment tool must have 100% security and data safety so that all data is only accessible by the user.

2.2.6 Time of Completion:

A prototype must be completed by May 2018.

3.0 Proof of Concept

3.1 ECG interpretation to obtain heart rate and heart rate variability

Using the data that was given to us by Dr. Lardizabal, the following steps were taken to compute the heart rate and heart rate variability:

1. Data was read using the eeglab software on MATLAB [1]
2. The ECG data was pulled out of the resulting cell
3. The derivative of the data with respect to time was obtained
4. For each (overlapping) 5 second window of the derivative, the following computations were performed
    a. The peaks were found using the findPeaks function with a minimum peak height of 10000 mV/s
    b. The time between each consecutive pair of QRS peaks was computed
    c. The heart rate for this window was computed as the average of these time intervals
    d. The heart rate variability for this interval was computed using the following formula
    e. $HRV = \sqrt{\Sigma(Time\ between\ consecutive\ peaks)^2}$

The heart rate and heart rate variability were plotted as a function of time for the entire sample and the values were within a reasonable range, indicating that the computation was performed properly. A sample output of this computation is presented in Appendix C.

3.2 Training of a decision tree model to perform data classification

A test data set was constructed using the EDF files that were provided to us by our client. Each file was approximately one hour long, meaning that roughly 3600 5-second samples were able to be constructed from the data, with each sample consisting of the following four values:

1. Time point (starting time)
2. Label (-1, 1, or 0 indicating pre-ictal, ictal, and baseline respectively)
3. Average heart rate for the 5-second interval
4. Heart rate variability for the 5-second window

Because there were no known samples that had seizures correctly labeled, half of the data was labeled as baseline (0) one fourth was labeled as pre-ical (-1) and one fourth was labeled as ictal (1). This was done in order to verify that the code is able to read in the data and construct a binary classification tree using the “fitctree” algorithm in MATLAB. The software was run and was able to produce a classification tree. The tree starts with all of the samples together in the root node and purifies the groupings using cutoffs for the heart rate or HRV and splitting the group into two smaller ones at each node until there are only pure groups left (groups labeled with only -1’s, 1’s, or 0’s). The software was not only able to construct a classification tree, but was also able to make predictions using the tree. A sample view of both the classification tree and its predictions are presented in Appendix D.

3.3 Simulation of ECG data with variable parameters

A simulated ECG simulation using Matlab was constructed from a program found on MATLAB’s file exchange [2]. The ECG simulator is able to produce a normal lead II ECG waveform and by changing certain parameters it is able to produce abnormal ECG waveforms. This simulator allows us to save the time and difficulties involved in obtaining real ECG signals. The adjustable features of the simulator include:

1. Adjustable heart rate
2. Adjustable interval between any two waves
3. Adjustable amplitude for any waves
4. Fibrillation can be simulated

Because a single period of a ECG signal can be modeled as a mixture of triangular and sinusoidal wave forms, the main principle behind the creation of the simulated ECG signal is the use of fourier series. See Appendix E for sample simulations.

3.4 Collection of ECG data via Arduino

An Arduino heart monitor was built using a circuit configuration from Theory Circuit [3]. The specific heart monitor component of the circuit is an AD8232 board built by SparkFun

Electronics. The circuit was able to accurately capture an ECG signal from a living test subject. A view of the Arduino setup is presented in Appendix F and a sample view of data collected using the Arduino is presented in Appendix G. The data from the signal was offloaded into saved text files on a computer, enabling access by the machine learning algorithm described above. While it was originally thought that two inverting amplifiers in series would be required to amplify the ECG signal, it was discovered that the internal machinery of the AD8232 board accomplishes this goal on its own.

4.0 Overall Status and Next Steps

Using the methods described above, our primary goal for the next several weeks is to integrate the several pieces of our device together such that ECG data can be recorded from the patient using the Arduino, saved to the computer, and fed into MATLAB to be classified as baseline, pre-ictal, or ictal. These will be used both to train the machine learning model and to be classified using the trained model.

After this is done, several rounds of testing will be done using simulated data, seizure data from our client, and data collected using the Arduino. Several combinations of the various data sets will be used to train and test the model. Because simulated data can be generated in large quantities, this will be the data set used for training and testing, while the client and Arduino data will be used for testing only. If enough data becomes available from our client, we will attempt to use this to train the model to improve its physiological accuracy.

Throughout the testing process, we will explore various types of measurements that can be extracted from the ECG data to make predictions and will determine which ones the model relies most heavily on. In doing so, the model will be refined so as to focus on the most powerful predictive features in order to keep the process efficient and accurate. This process will rely on consistent communication with our client throughout the testing process.

Appendix A: Design schedule

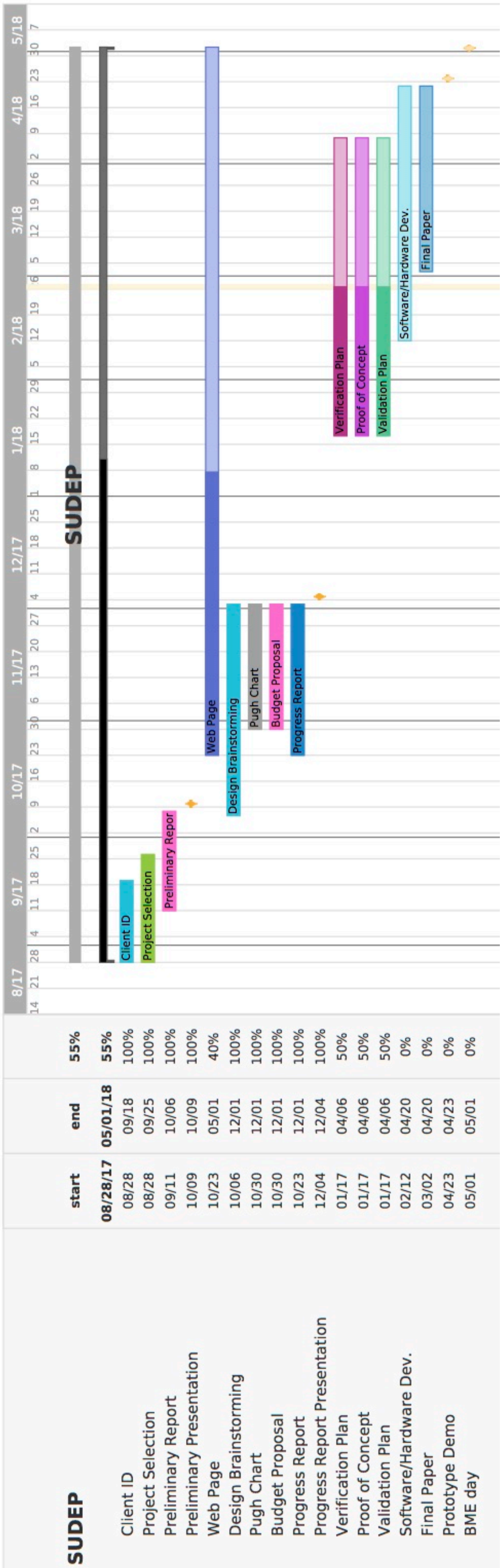
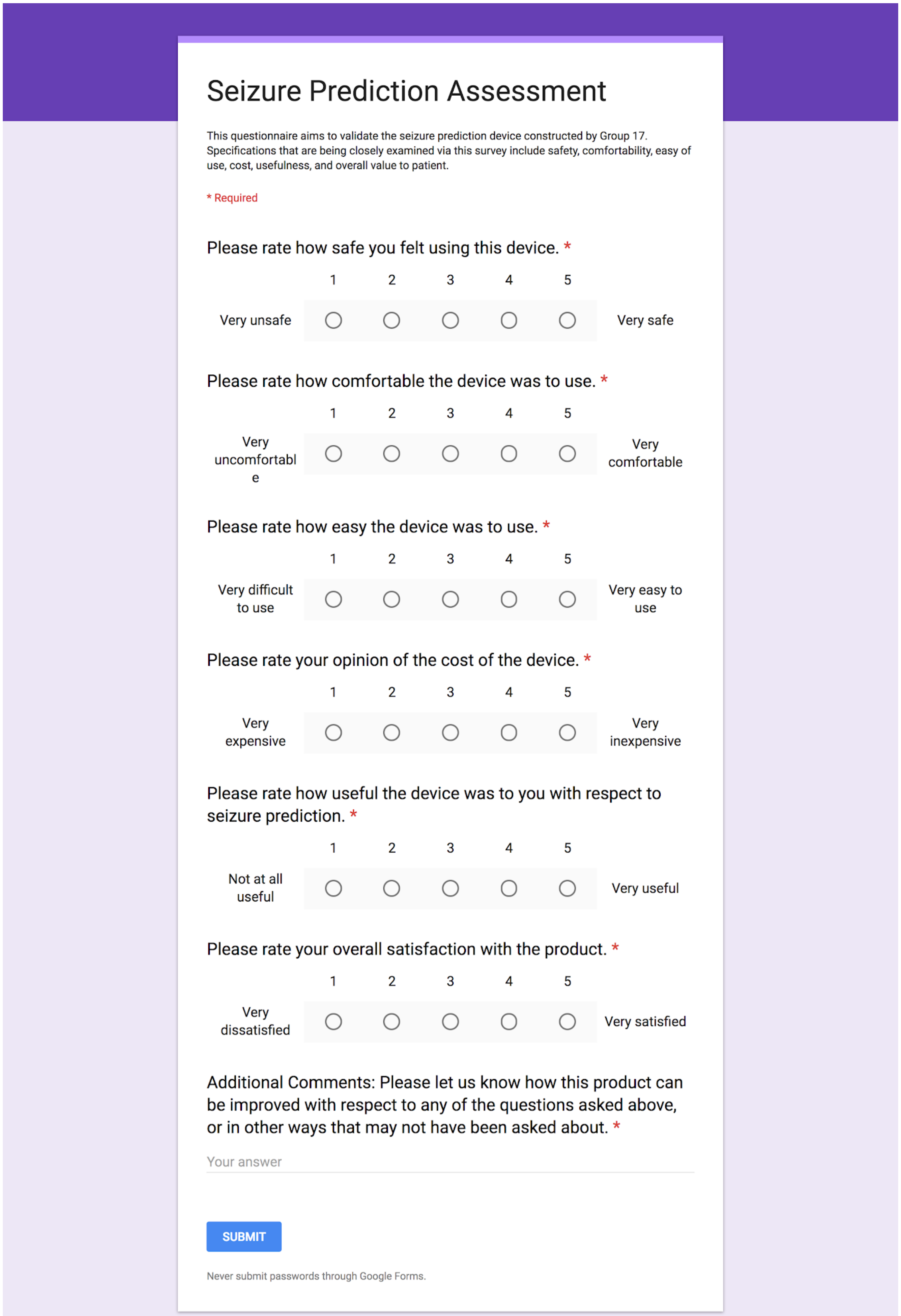


Figure 1: Updated design schedule

Appendix B: User satisfaction assessment form



Seizure Prediction Assessment

This questionnaire aims to validate the seizure prediction device constructed by Group 17. Specifications that are being closely examined via this survey include safety, comfortability, easy of use, cost, usefulness, and overall value to patient.

* Required

Please rate how safe you felt using this device. *

	1	2	3	4	5	
Very unsafe	○	○	○	○	○	Very safe

Please rate how comfortable the device was to use. *

	1	2	3	4	5	
Very uncomfortabl e	○	○	○	○	○	Very comfortable

Please rate how easy the device was to use. *

	1	2	3	4	5	
Very difficult to use	○	○	○	○	○	Very easy to use

Please rate your opinion of the cost of the device. *

	1	2	3	4	5	
Very expensive	○	○	○	○	○	Very inexpensive

Please rate how useful the device was to you with respect to seizure prediction. *

	1	2	3	4	5	
Not at all useful	○	○	○	○	○	Very useful

Please rate your overall satisfaction with the product. *

	1	2	3	4	5	
Very dissatisfied	○	○	○	○	○	Very satisfied

Additional Comments: Please let us know how this product can be improved with respect to any of the questions asked above, or in other ways that may not have been asked about. *

Your answer

SUBMIT

Never submit passwords through Google Forms.

Figure 2: This is the user satisfaction form that will be used to validate that this seizure prediction device meets the needs of those using it.

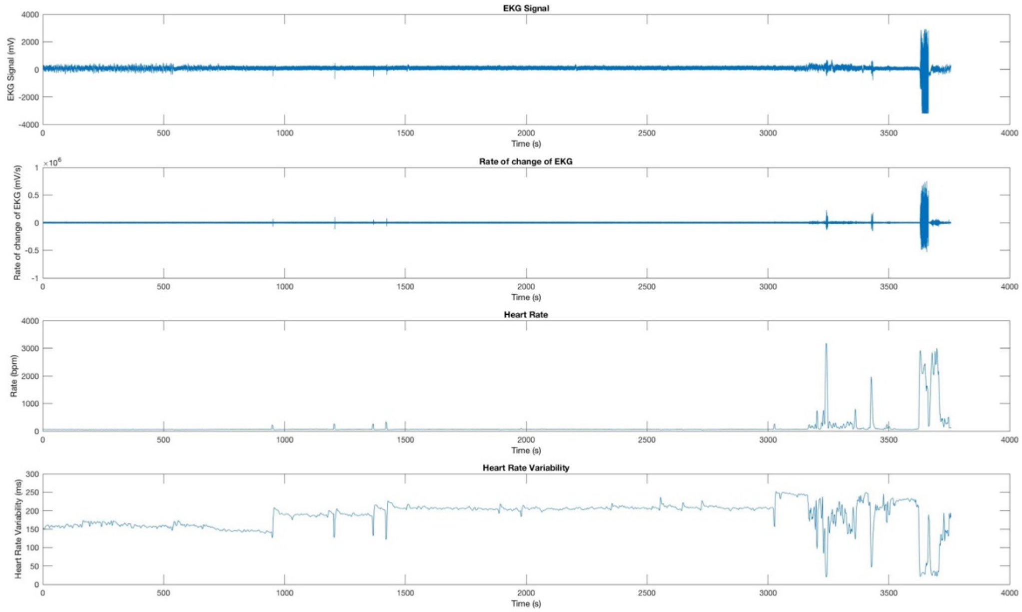


Figure 3a: This is a view of the output of the ECG interpretation software written for this device for an entire (approximately one hour) sample recording provided by our client. From top to bottom, the panels show the raw ECG data, the rate of change of the ECG, the computed heart rate over a 5 second interval, and the computed heart rate variability over a 5 second interval.

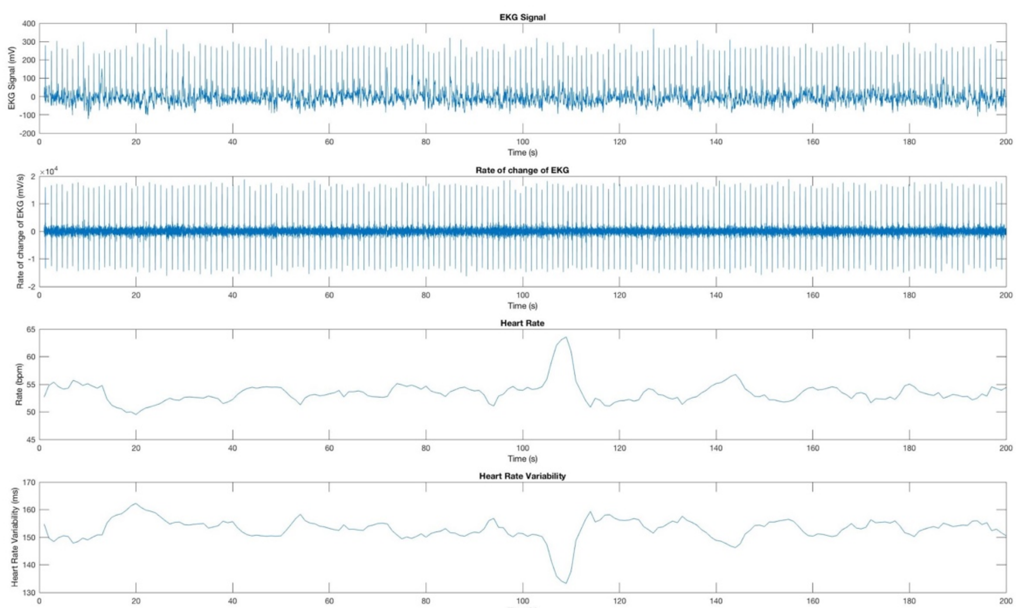


Figure 3b: This is a view of the output of the ECG interpretation software written for this device for 200 seconds of a sample recording provided by our client. From top to bottom, the panels show the raw ECG data, the rate of change of the ECG, the computed heart rate over a 5 second interval, and the computed heart rate variability over a 5 second interval.

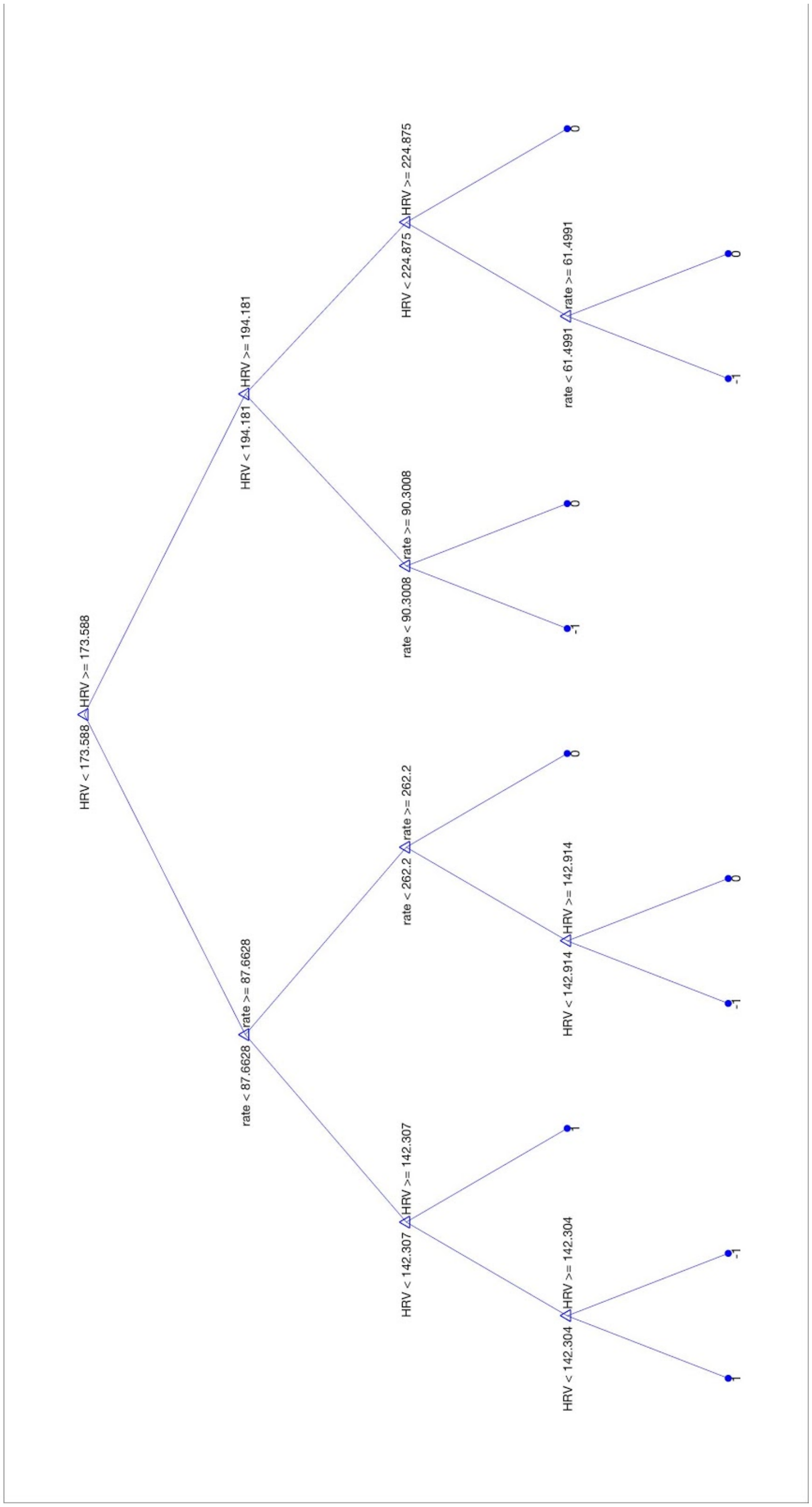


Figure 4: This is a sample view of the classification tree constructed using MATLAB. This tree was constructed using data that was randomly labeled with ½ baseline, ¼ pre-ictal, and ¼ ictal samples. The samples start in one combined group in the top node and are separated at each node based on the criterion specified until they are purified in the leaf nodes. “HRV” indicated heart rate variabilitv and “rate” indicates heart rate.

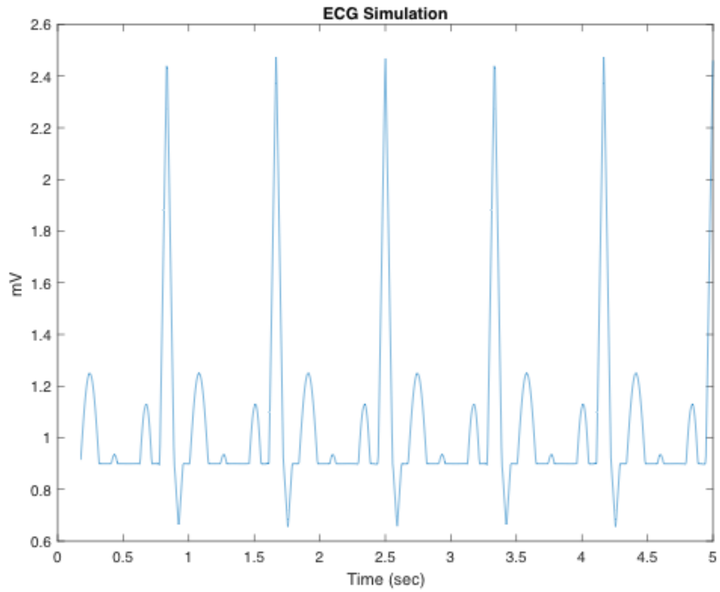


Figure 5a: This is a sample 5-second ECG simulation constructed in MATLAB using the default parameters provided by the software.

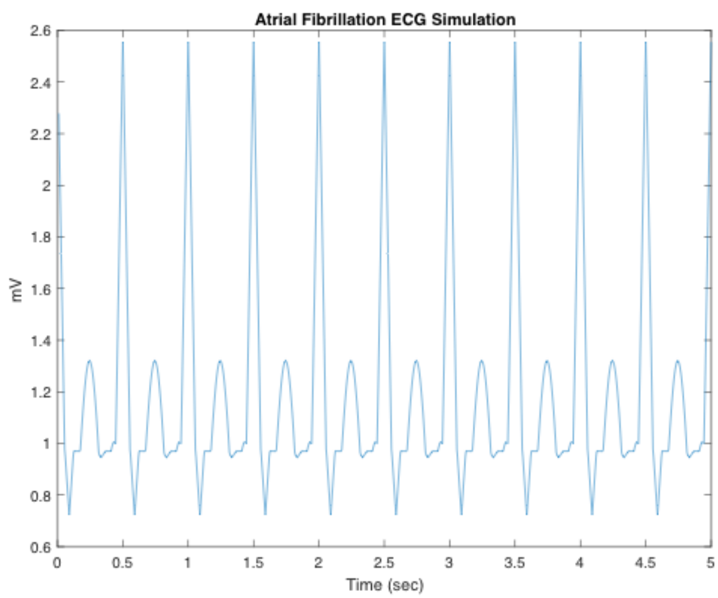


Figure 5b: This is a sample 5-second ECG simulation constructed in MATLAB with varied parameters to simulate atrial fibrillation (Heart rate = 120 bpm and no p-waves).

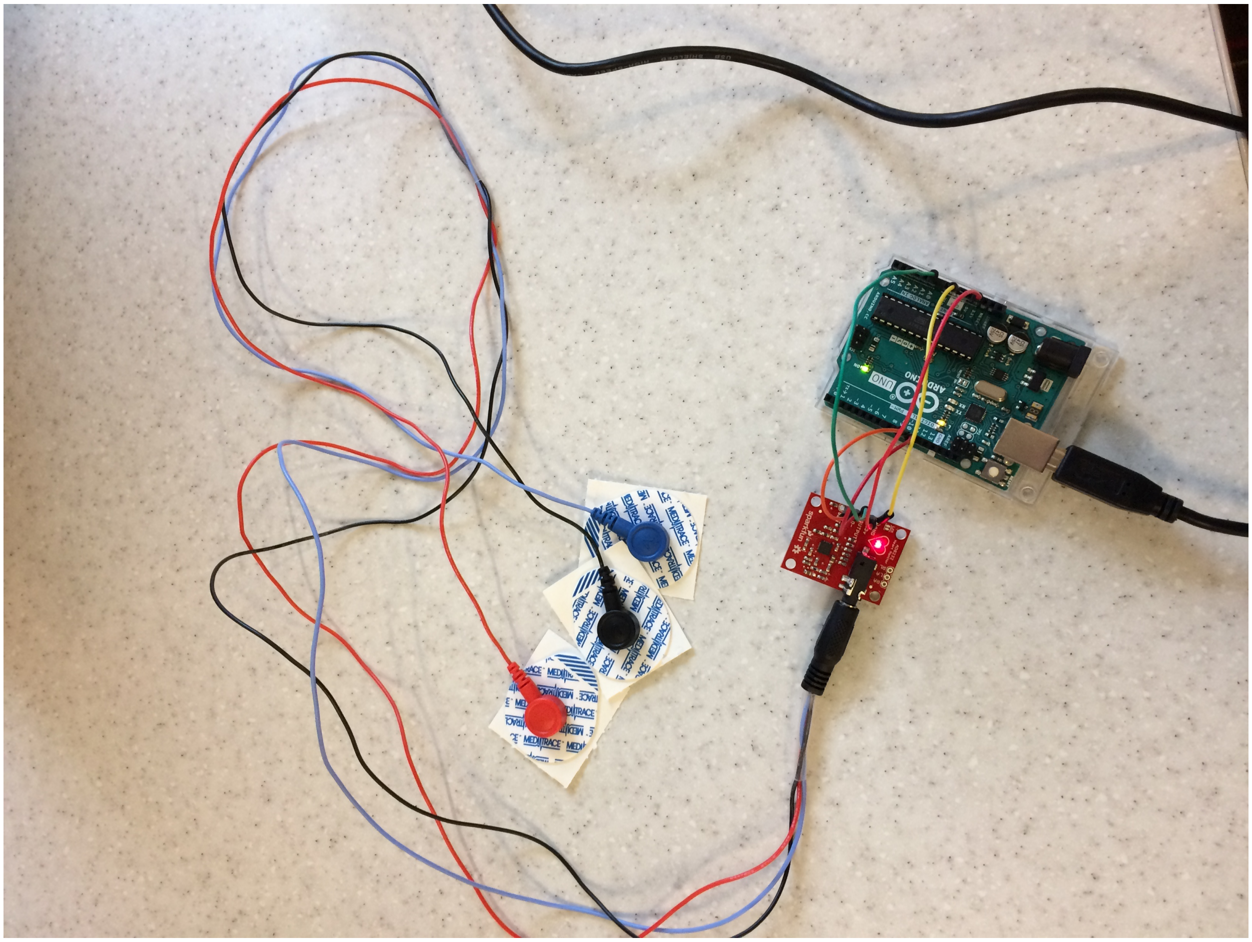


Figure 6: This is an image of the Arduino circuit that was constructed to record human ECG data using the heart rate monitor available from SparkFun.

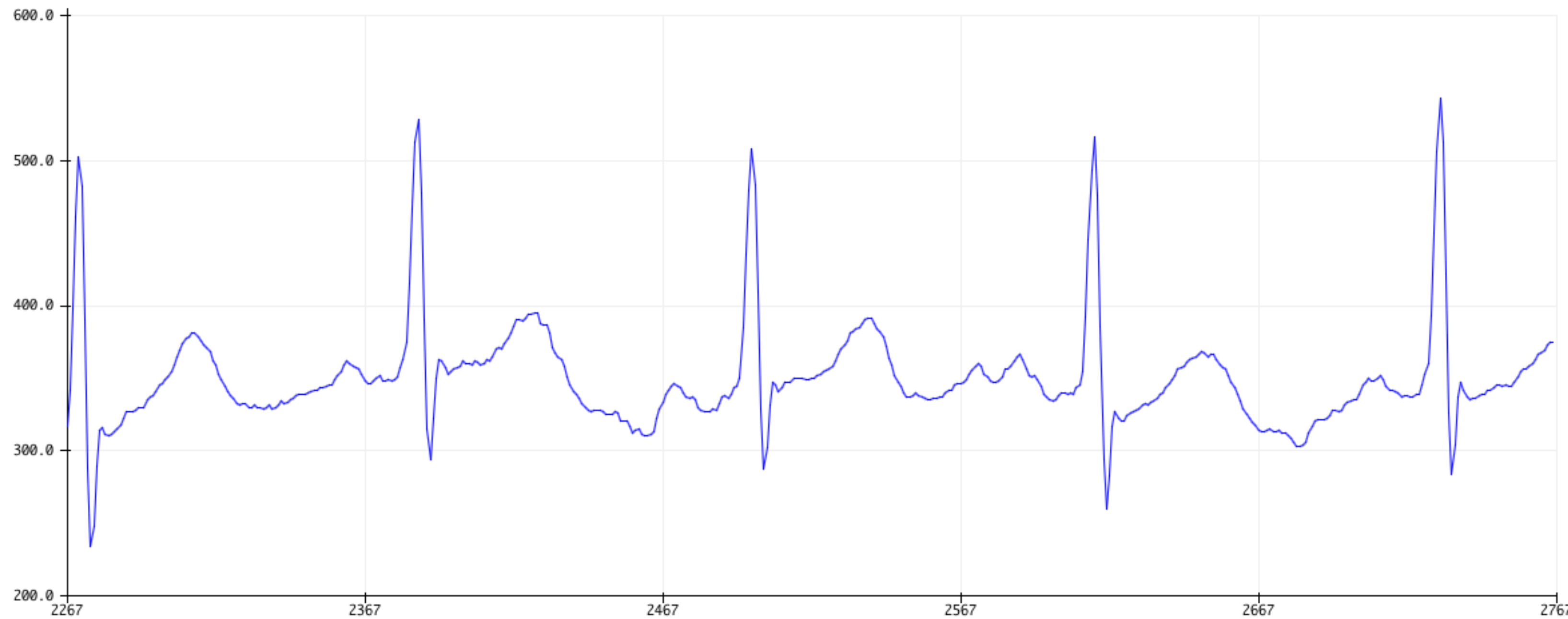


Figure 7: This is a sample view of ECG data collected using the Arduino plotted using the Arduino serial plotter. The y-axis is shown in millivolts and the x-axis is the sample number.

References

1. ECG simulation using MATLAB. (2006). Retrieved from https://www.mathworks.com/matlabcentral/fileexchange/10858-ecg-simulation-using-matlab?s_tid=gn_loc_drop
2. EEGLAB. (2018). San Diego: UC San Diego. Retrieved from https://sccn.ucsd.edu/eeglab/index.php
3. Eisenman, L. Personal interview with third party [D. Lardizabal].
4. Heart Rate Monitor AD8232 Interface Arduino. (n.d.). Retrieved from http://www.theorycircuit.com/heart-rate-monitor-ad8232-interface-arduino/
5. Lardizabal, D. Personal interview.