

## Motivation

### Background:

There are a variety of different wearable fitness/cardiac monitoring devices that are currently used in many people's day-to-day life. The primary cardiac function of these devices is to monitor heart rate; however, we believe that they could be utilized to detect different forms of arrhythmia.

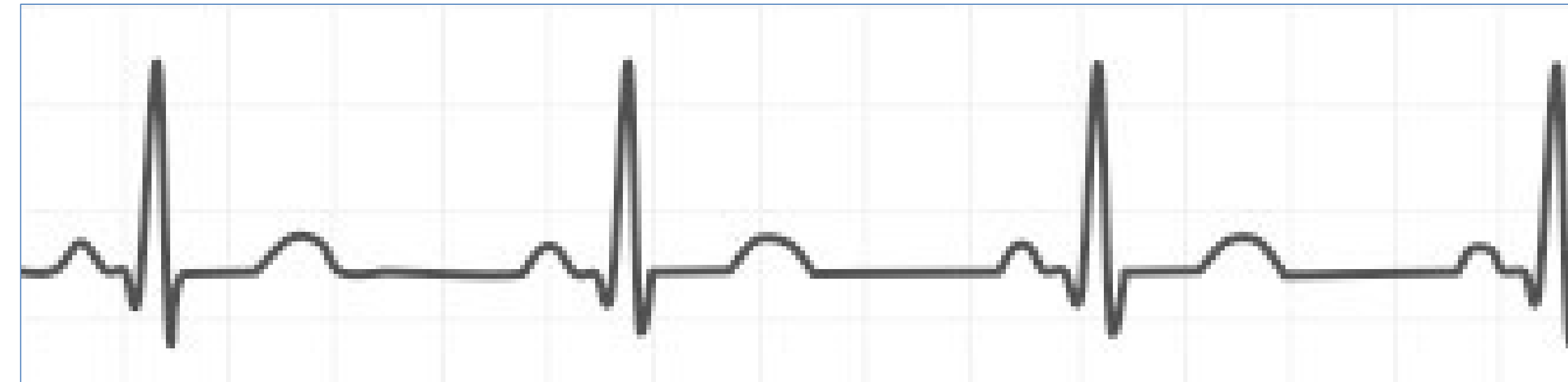
### Challenge:

In order to categorize and identify different forms of arrhythmia, we are utilizing existing electrocardiogram (EKG) databases as a basis for machine learning. The challenge that comes from the existing datasets is that their formats cannot be directly imported into machine learning algorithms, which requires data to be in a vector format. This makes the process of converting the existing datasets into workable vector formats long and tedious.

### Proposed Solution:

Develop a post-processing algorithm that maps to vector form the data from multiple different datasets such that researchers wishing to apply machine learning to EKG data, are able to quickly and accurately access workable datasets.

## Baseline Heartbeat



Normal complex, evenly spaced, 60-100 BPM

## Bradycardia



## Tachycardia



Normal complex, evenly spaced, >100 BPM

## Atrial Fibrillation



## Arrhythmia



## Methods and Results

### Methods:

1. For machine learning to be successful, we first standardize voltage recordings at 100 Hz. This in of itself is challenging as different datasets provide different variations in their timestamps given, ranging form random. intervals to set intervals of 300 Hz to 1 kHz.
2. From there, we take the voltage values that are associated with each of these standardized times and map them into a concise vector form. This form includes those voltages and a set spacing between each value to create an array-like unit.
3. Lastly, due to the timestamp variations among and within online EKG databases, we add different cases to automatically detect these variations and then allow the algorithm to proceed with the automated vector formatting.

### Results:

Our manual method of pre-processing the online EKG datasets is able to create a vector that can be imported into machine learning algorithms. Further, we have created a base algorithm using the Java programming language that automates the vector format mapping for dozens of the databases and has been verified to be a direct match to the manually created method. More importantly, this automated process averages less than a second for this complete database vector mapping.