

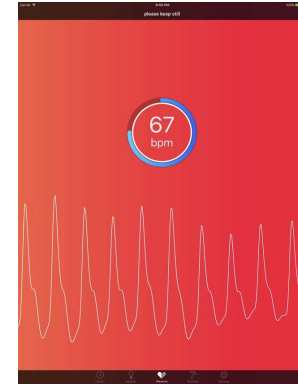
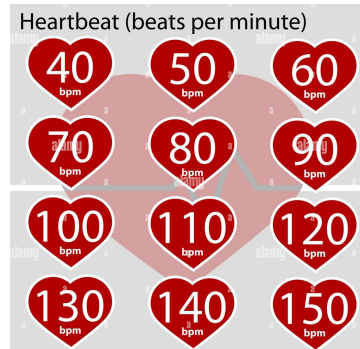


# ***Stress Classification***

***By Matt  
Ferrara, Lynn  
Li, and Mak  
Khan***

# Data Collection

- Due to not having an Android phone to use with the PPG wristband, and the fact that no trusted apps provided the ability to export raw PPG signals, we decided to skip the raw PPG signal entirely, and trust Beats Per Minute outputs → Even if we had a raw PPG, we would likely be finding beats per minute anyway so this is just skipping that step
- We used a combination of Apple Watch readings and readings from the app “Cardiio”
- Cardiio syncs to Apple Health app, as do Apple Watch readings (obviously), and from there we could export the data to our computers as a CSV



## Data Collection (continued)

- Each group member took 10 resting heart rate readings and 10 readings sticking our hands in ice water to simulate stress
- Upon completion, we combined all our data into a single CSV, labelled the samples with integers (1 for no stress, 2 for stress)



# Data Collected

Beats per minute

1 = Resting ; 2 = Stressed

22	56	1
23	61	1
24	60	1
25	58	1
26	59	1
27	56	1
28	60	1
29	58	1
30	59	1
31	78	2
32	78	2
33	77	2
34	78	2
35	76	2
36	84	2
37	77	2
38	86	2
39	73	2
40	71	2
41	87	2

# Feature Extraction

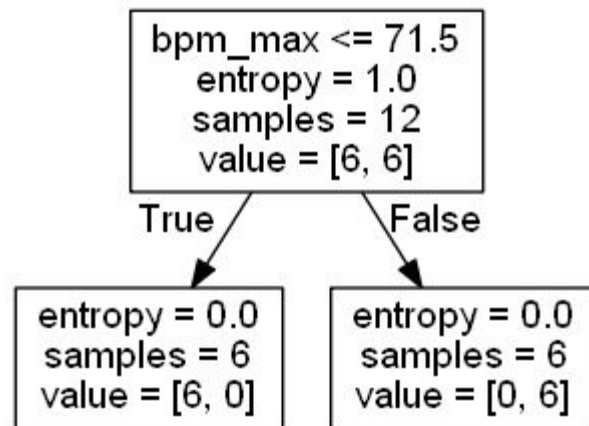
- **Features we currently have implemented:**
  - **Beats Per Minute**
  - **Max BPM**
  - **Min BPM**
  - **Mean BPM**
  - **Heart Rate Variability (Standard Deviation of BPM)**
- **We're trying to train a classifier based off these features**

# Classification

- We used a window size and step size of 5 and decision trees
- Our app did not provide timestamps, so we entered dummy values similar to the ones in A2 P1 for our data
- We extracted features and split into test and train sets in the same manner we did in A2 P1 as well
- Overall, the way our data was structured, having only one value (BPM) made it easy to overfit
  - All features had an accuracy of 1.0, as did not using max or mean, which shows clear overfitting
  - The best feature to use, as suggested by professor was heart rate variability (sd). This did not overfit.
- Decision tree and accuracy/precision/recall outputs for each classifier on next slides

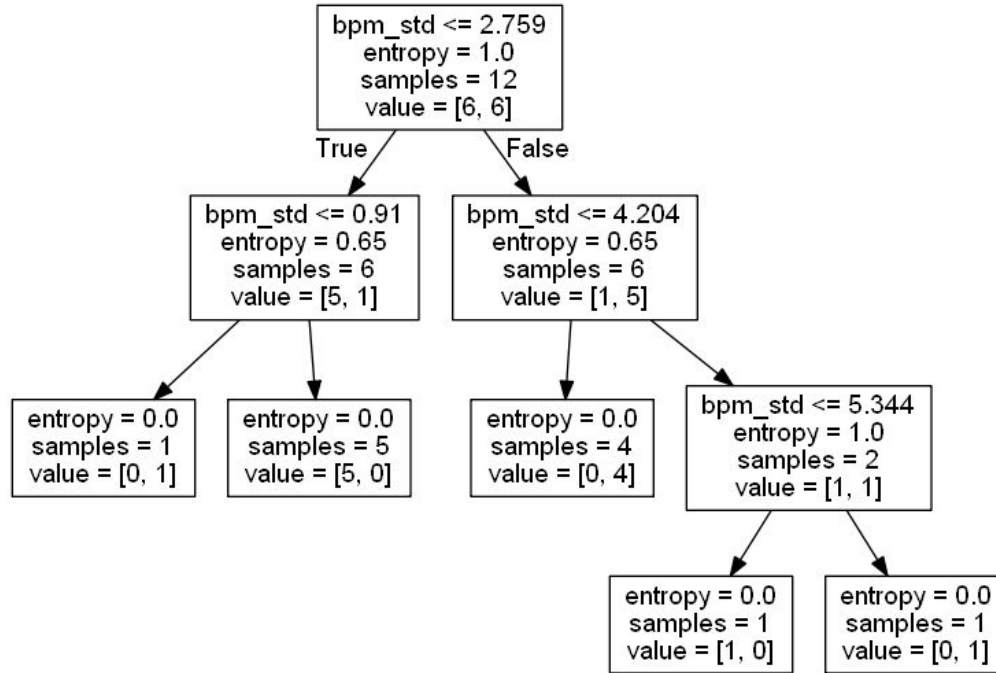
# Classification Results-All Features

```
Average accuracy is 1.0  
Average precision is [1. 1.]  
Average recall is [1. 1.]
```



As you can see, this overfits badly and only needs max as a splitting feature

# Classification Results-Heart Rate Variability Only

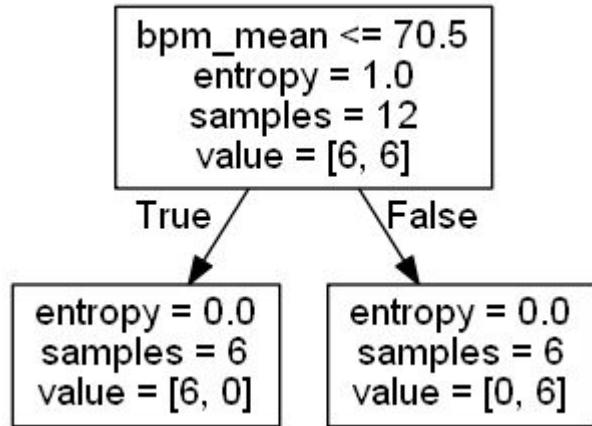


Average accuracy is 0.7  
Average precision is [0.6 0.8]  
Average recall is [0.6 0.7]

No overfitting, fairly accurate and precise, best feature to use in our opinion



# Classification Results-No Max

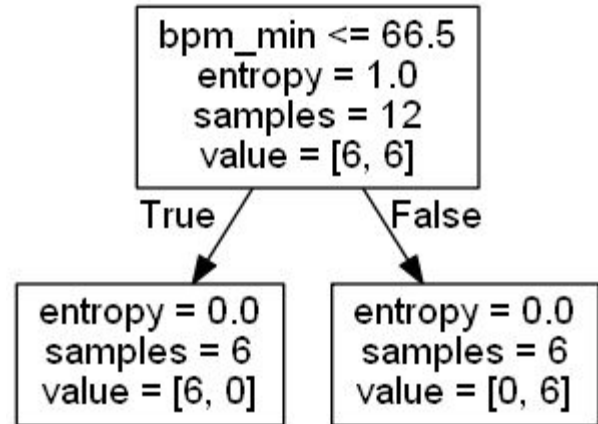


```
Average accuracy is 1.0  
Average precision is [1. 1.]  
Average recall is [1. 1.]
```

Overfits badly again, uses mean instead of max as main splitting feature

# Classification Results-No Max No Mean

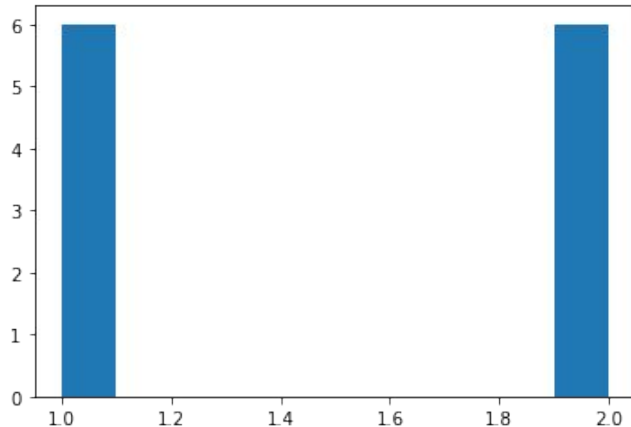
```
Average accuracy is 1.0  
Average precision is [1. 1.]  
Average recall is [1. 1.]
```



Overfits badly again splitting off min, only feature left is sd which was the best feature to use

# Testing Results-Heart Rate Variability Only Classifier

- Because the other classifiers all appeared to overfit, we decided to only use the HRV only classifier on our test data
- Just like our training data, we collected 30 samples of resting BPM values and 30 samples of stressed BPM values and ran had the pickled classifier make predictions over 12 windows with 5 samples in each



- As we would expect, there are about an even amount of 1s for no stress and 2s for stressed predicted by our classifier
- This matches our test data that had 30 of each type of sample