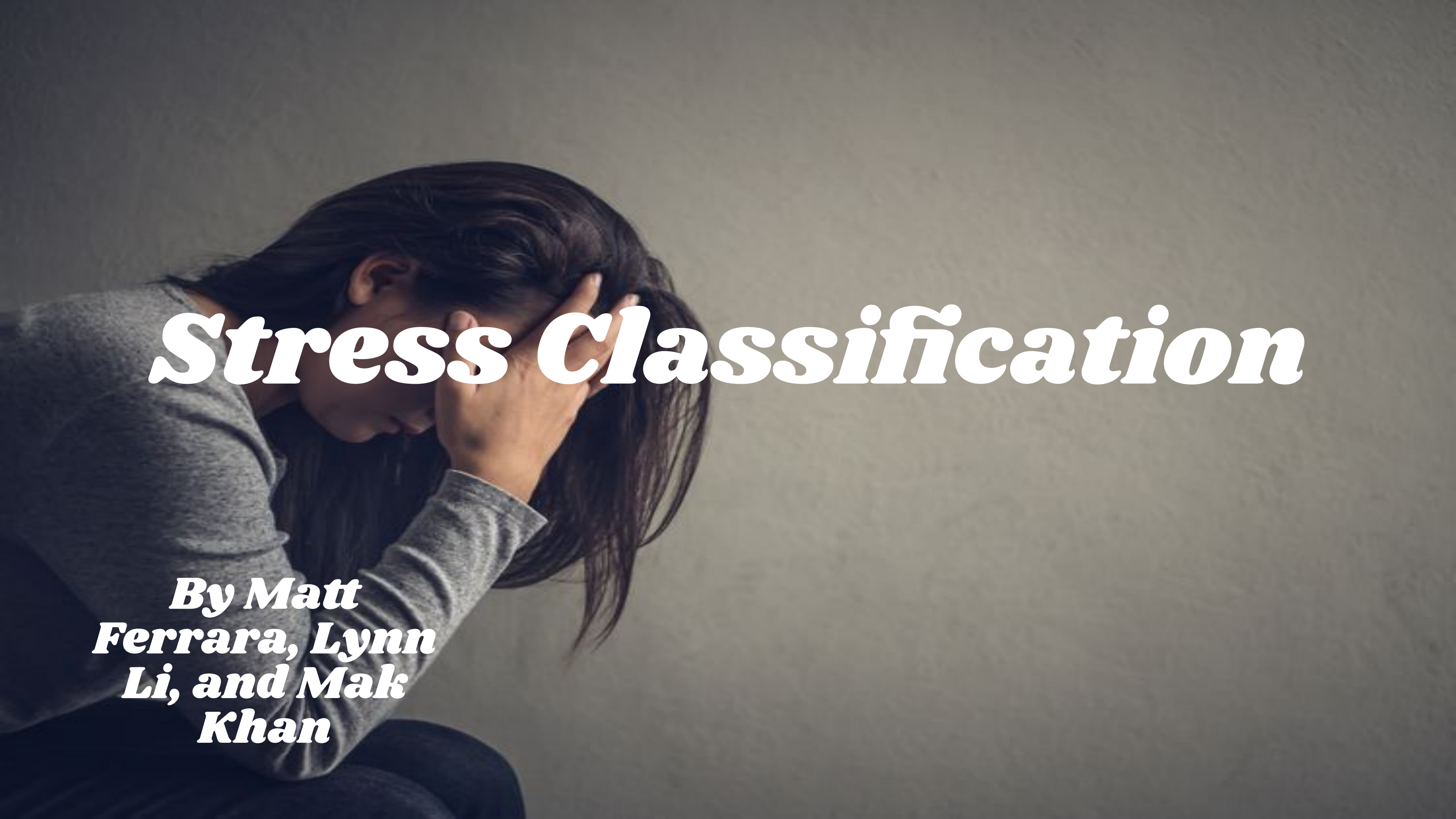
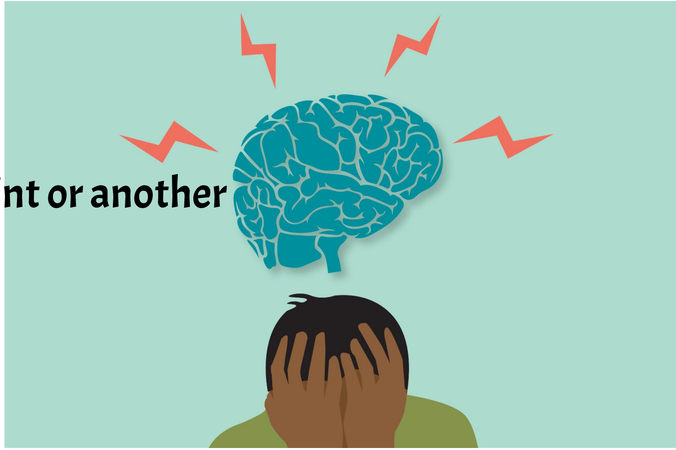


Stress Classification

By Matt Ferrara, Lynn Li, and Mak Khan

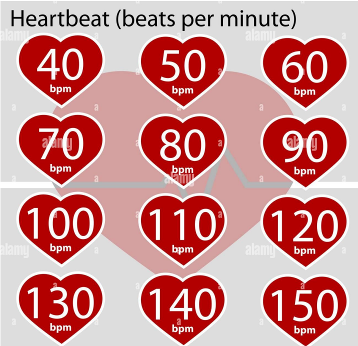
Hypothesis / Motivation for Idea

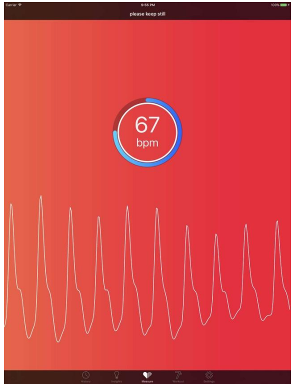


- Stress is something that everyone experiences at one point or another
- A few consequences of stress include:
  - Low energy
  - Headaches
  - Chest pain
  - Insomnia
  - Cardiovascular disease
  - Diabetes
  - Skin and hair problems
- Being able to identify high stress levels and monitor how often they occur is important when it comes to ensuring health throughout life
- If individuals could better identify when they were stressed, they could integrate more intervention methods to decompress and take care of themselves
- Going in, we hypothesized that being stressed would increase BPM on average and thus we could use features extracted from BPM values to train classifiers

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)
- We witnesses the same results across all members where no matter what our resting BPM values were, the stressed values increased by around 20-30 Beats per minute

Data Collected

	Beats per minute	1 = Resting / No Stress ; 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

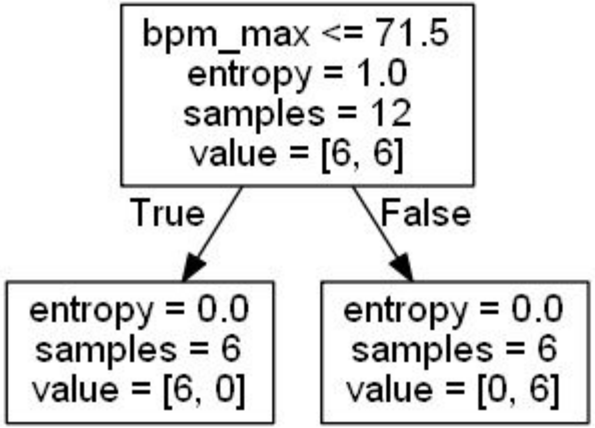
- Features we have implemented:
  - Beats Per Minute
  - Max BPM
  - Min BPM
  - Mean BPM
  - Heart Rate Variability (Standard Deviation of BPM)
- We then trained classifiers based off these features
- As the professor recommended, we tried multiple combinations of features including all features and only heart rate variability

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

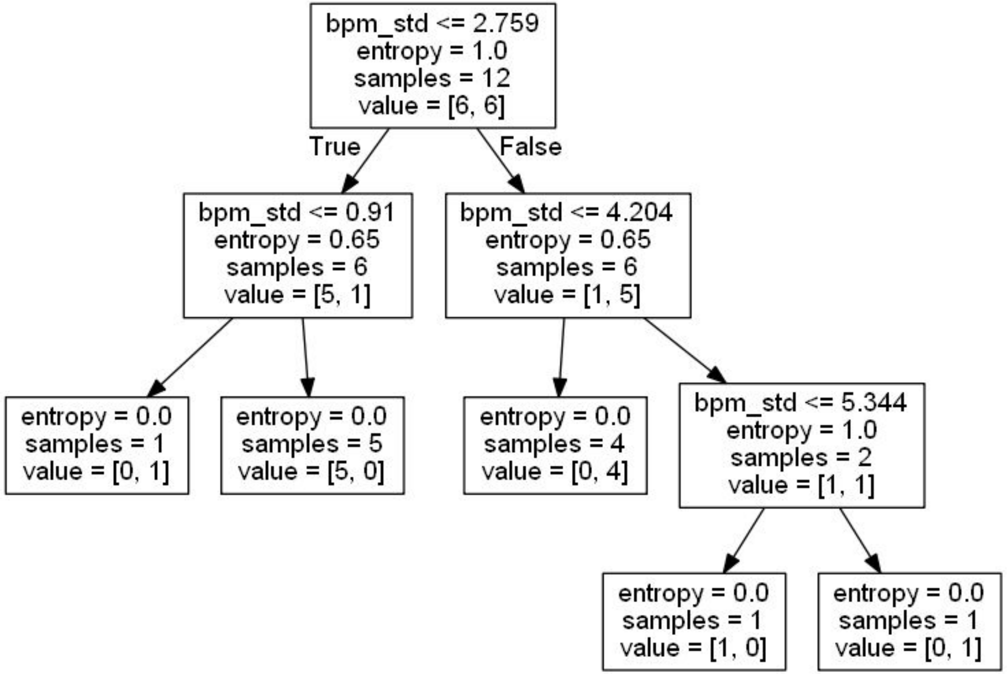
Classification Results-All Features

```
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

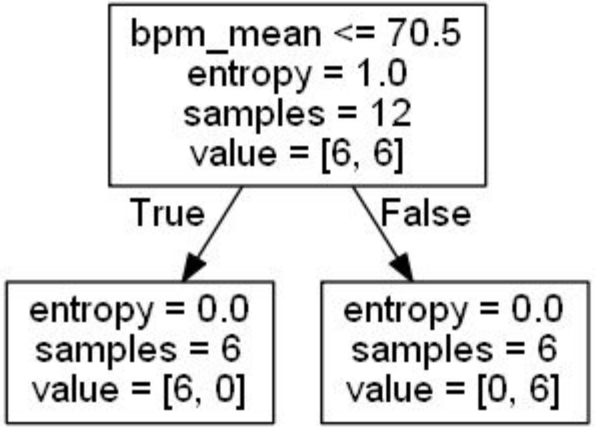
Classification Results-Heart Rate Variability Only




```
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

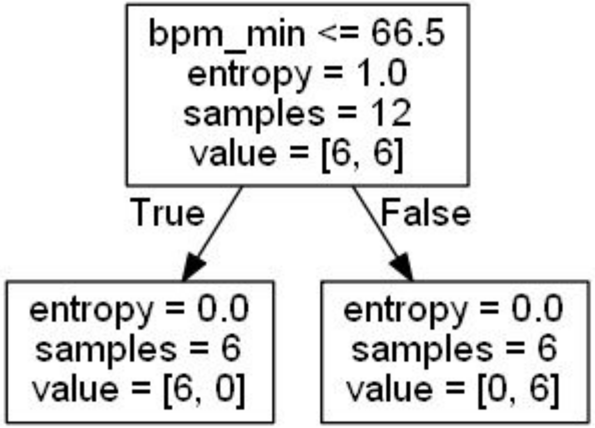



```
Average accuracy is 1.0
Average precision is [1. 1.]
Average recall is [1. 1.]
```

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

Classification Results-No Max No Mean

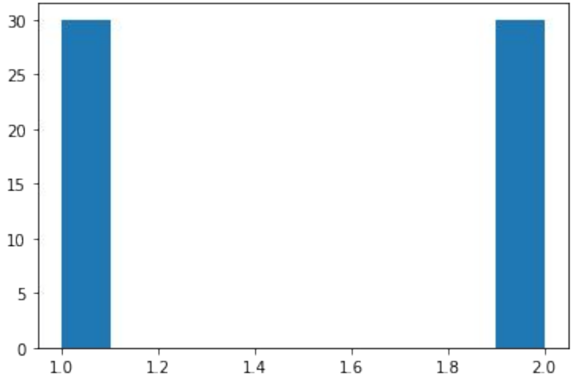
```
Average accuracy is 1.0
Average precision is [1. 1.]
Average recall is [1. 1.]
```



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

Testing Results

- 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 those 60 samples



- We plotted the prediction results on a histogram
- 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

Conclusions

- In the future, it would be interesting to use an actual PPG signal and compare the classifiers
- Because of really only using BPM values and some features of them, the classifiers tended to overfit in most scenarios → there was clearly a BPM increase in stressed vs not stressed so the classifiers were 100 percent accurate
- Heart rate variability, as the professor alluded to previously, is likely the best predictor of stress and best feature for stress classification
- Based on our data and testing results, stress and increased heart rate have a strong correlation with one another and confirm what we hypothesized when we first came up with the project idea