College of Engineering
University of Missouri

University of Missouri

Computer Science
University of Missouri

AMD: Analysis of Mood Dysregulation

A Machine Learning Approach

Master's Thesis Defense

By:
Nickolas M. Wergeles
Advisor: Dr. Yi Shang

Components

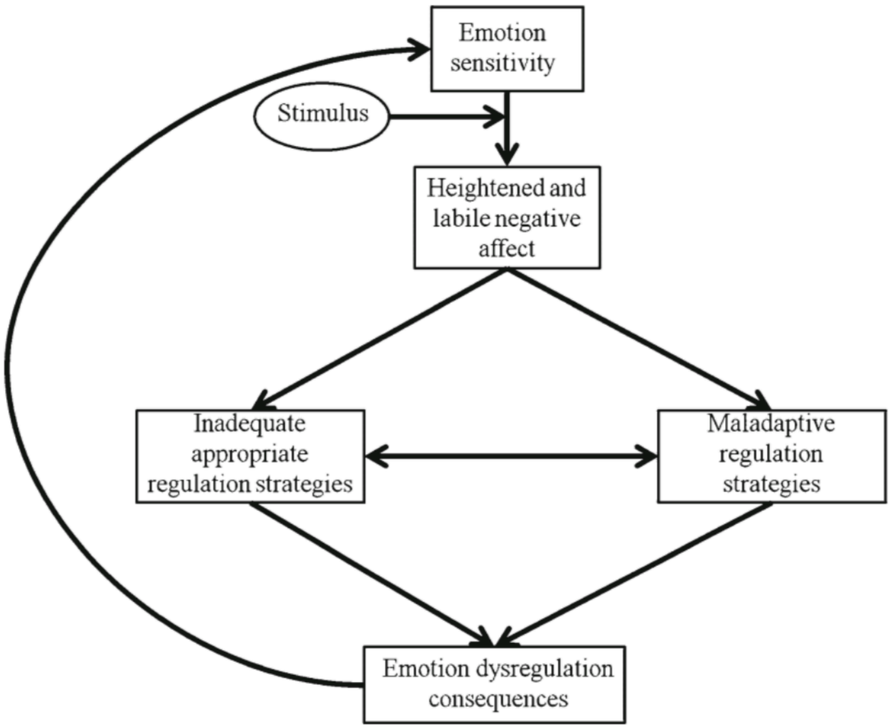
- Introduction
- Related Work
- mAAS System Improvements
- Mood Study Overview
- AMD Pipeline
- Results and Model Comparison
- Knowledge Gained and Future Work

Components

- Introduction
  - Problem
  - Solution
  - Contributions
  - AMD Overview
- Related Work
- mAAS System Improvements
- Mood Study Overview
- AMD Pipeline
- Results and Model Comparison
- Knowledge Gained and Future Work

Introduction

- Mood dysregulation is the inability to flexibly respond to and manage emotions
  - Theorized to be sensitive to emotion from birth
- As a result, mood dysregulation consequences occurs
  - Such as stress or other emotion sensitivity



Carpenter, R.W. & Trull, T.J. Curr Psychiatry Rep (2013) 15: 335. doi:10.1007/s11920-012-0335-2

Introduction

- If you are alive, then you have mood variability
  - Stress kills, can lead to or worsen heart diseases and cancer
- Human studies show mood dysregulation causes psychological and behavioral problems
  - Such as stress, depression, addiction, rage, anxiety and other mood dysregulation issues
- How can we analyze and study mood dysregulation ?

Problem

1. Most methods in clinical psychology research primarily rely on questionnaires and interviews
   – With examiners in lab setting
2. Humans have a hard time understanding mood dysregulation using these methods
   – To much data to analyze by hand
   – Difficult to see trends in the data

Solution

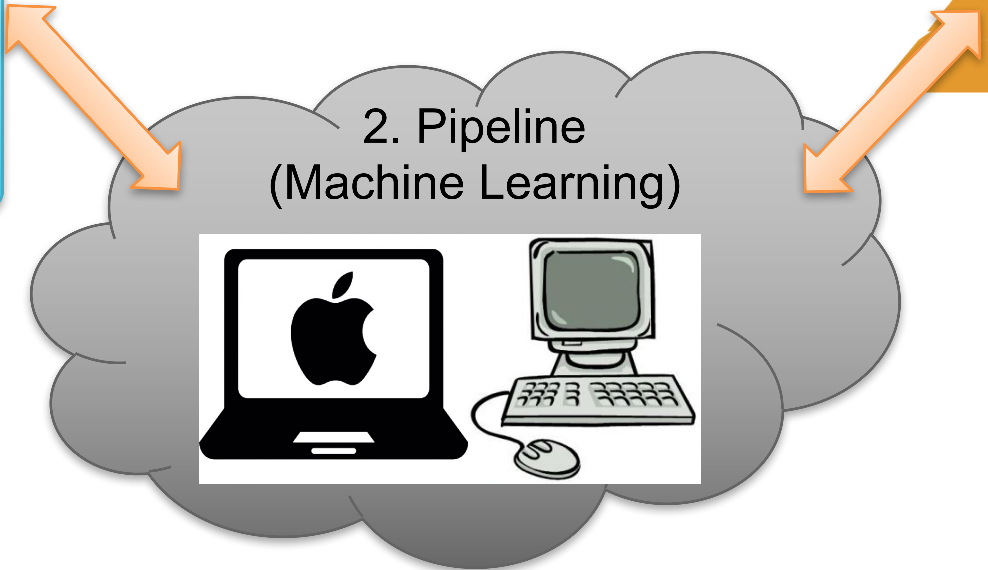
1. Mobile ambulatory assessment in natural real-world environments
   - Using smartphones and physiological wearable body metrics
2. Machine learning models can be developed to identify changes in mood, stress, as well as other psychological problems
   - Possible by combining information about:
     - External environment, participants' physiological and mental states
   - Collected through random system-generated and user-initiated self-reports surveys

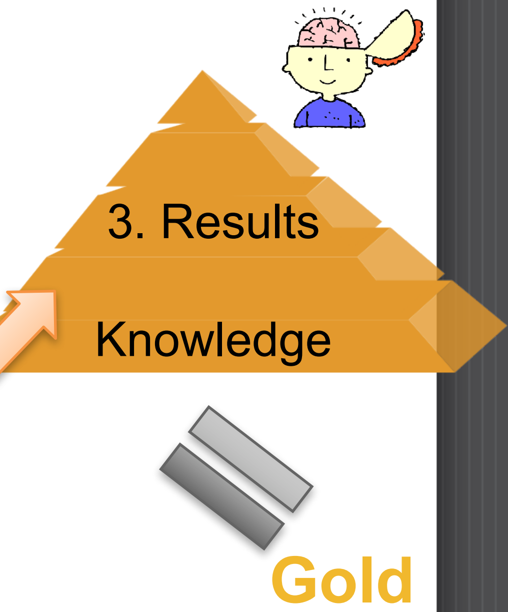
Contributions

- 1. mAAS System Improvements
  - Hardware and user information during study
  - Upload missing data from mobile device
  - Obtaining and displaying data from Hexoskin
- 2. AMD Pipeline
  - Data combination and organization, data selection and removal, data smoothing, feature extraction and creation
  - Classification of Mood Dysregulation in the Natural Environment
    - Based on multiple machine learning models
    - Achieve over 90% accuracy when including time attributes

AMD Workflow Overview

“Logic is not the end of wisdom, it is just the beginning” -- Spock





Components

- Introduction
- Related Work
  - Continuous inference, *2011*.
  - Recognizing mental stress, 2013.
  - Profiling visual and verbal stress responses, 2013.
  - Finding Significant Stress Episodes, 2016.
- mAAS System Improvements
- Mood Study Overview
- AMD Pipeline
- Results and Model Comparison
- Knowledge Gained and Future Work

Related Work

1. Plarre, K., Raij, A., Hossain, S. M., Ali, A. A., Nakajima, M., Al'absi, M., ... & Siewiorek, D. (2011, April). Continuous inference of psychological stress from sensory measurements collected in the natural environment. In *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on* (pp. 97-108). IEEE.
   - Similar project except they are studying stress
   - Helped me with the data cleaning process and feature selection
   - Received approximately 90% accuracy when removing around 70% of the data due to physical activity, poor quality or losses in transmission, etc.

Related Work (cont.)

2. Eggert, C., Lara, O. D., & Labrador, M. A. (2013, April). Recognizing mental stress in chess players using vital sign data. In *Southeastcon, 2013 Proceedings of IEEE* (pp. 1-4). IEEE.
    - Developed system to recognize psychological stress without human subjective bias
    - Use physiological sensors during chess match
    - Determine which moves caused the player to become stressed during the chess match

Related Work (cont.)

3. Bouarfa, L., Bembnowicz, P., Crewther, B., Jarchi, D., & Yang, G. Z. (2013, May). Profiling visual and verbal stress responses using electrodermal heart rate and hormonal measures. In *2013 IEEE International Conference on Body Sensor Networks* (pp. 1-7). IEEE.
   - Discriminate stress responses from watching videos and speaking using Electro-dermal activity and heart rate variability measures
   - Using 17 features calculated from signals
   - 4 Classifiers were investigated with maximum accuracy of 92% achieved for differentiating stress responses

Related Work (cont.)

4. Sarker, H., Tyburski, M., Rahman, M. M., Hovsepian, K., Sharmin, M., Epstein, D. H., ... & al'Absi, M. (2016, May). Finding Significant Stress Episodes in a Discontinuous Time Series of Rapidly Varying Mobile Sensor Data. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 4489-4501). ACM.

   – Delivering sensor-triggered just-in-time interventions (JITIs) on mobile devices for the management of daily stress

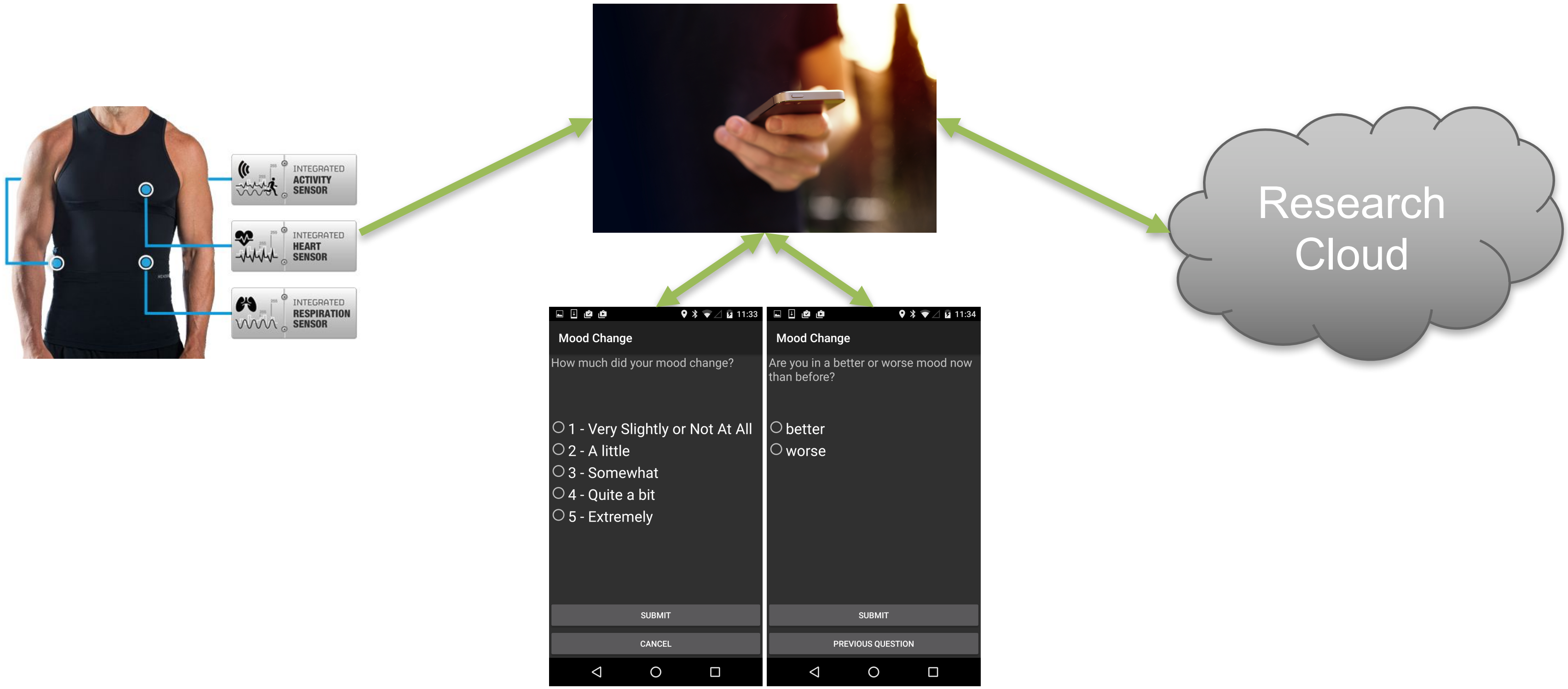
   – 38 users with 4 weeks of physiological, GPS, and activity data

   – Found that the duration of a prior stress episode predicts the duration of the next stress episode

Components

- Introduction
- Related Work
- Mobile Ambulatory Assessment System (mAAS) Improvements
  - mAAS Overview and Introduction
  - mAAS Hardware Information Collection
  - Uploading Missing Survey Data from Mobile Devices
  - Obtaining and Displaying Sensor Data from Hexoskin
- Mood Study Overview
- AMD Pipeline
- Results and Model Comparison
- Knowledge Gained and Future Work

mAAS Overview

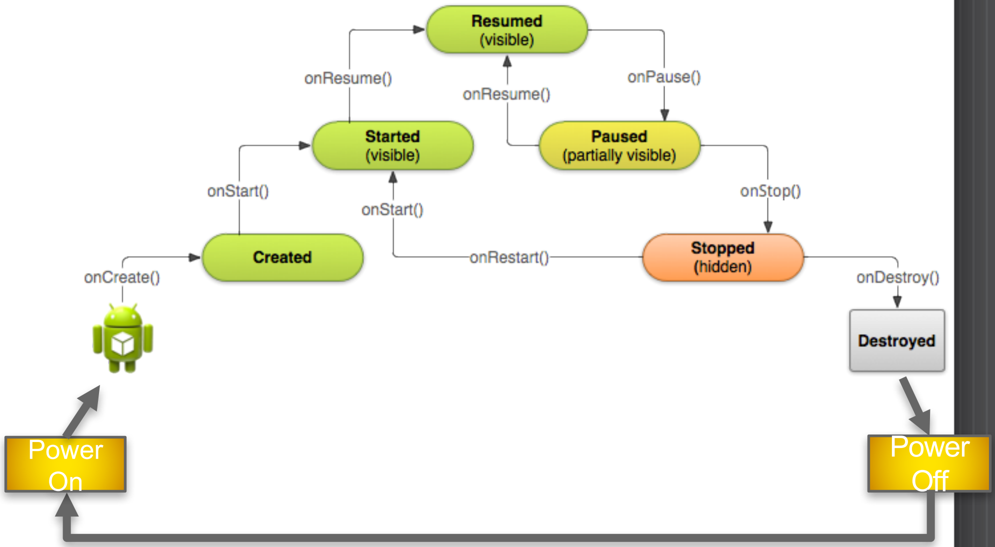


mAAS System Problems

1. Unknown application errors
   – Generated by user or mobile device application
2. Missing survey data
   – Caused when application was not connected to Internet
3. Obtaining physiological data from Hexoskin
   – Could only go through dashboard downloading one file at a time

mAAS Hardware Information Collection

- Monitoring system for hardware on mobile phone
  - Can determine if user turns on/off phone, starts/stops application, Internet connectivity, and many more
  - Used when we did not know if the user was causing error or the application itself
    - Some users claim app not working when in reality, they were not connected to internet, or had the phone off



Example Output

- Every 5 Minutes:
  - Data collected here was sent to the server on 5 minute intervals
  - Effected the battery use minimally

```
Tue Sep 29 04:50:39 CDT 2015

Is Phone Charging? -> true
  Charging By: AC Outlet
Battery Level: 1.0
Network Connection Status: Connected
  MOBILE: CONNECTED
  WIFI: UNKNOWN
  BLUETOOTH for Network: UNKNOWN
  Is Phone Connected To Active Network? -> true
GPS Mode in Phone Settings: LOCATION_MODE_SENSORS_ONLY
  Is There an Active GPS Signal? -> true
  Longitude and Latitude: 38.88794724, -92.34900453
  GPS Provider: gps
  GPS Accuracy: 17.0 meters
  Is the GPS Accuracy Good? -> true
  Will this GPS location be recorded? -> true
Is BLUETOOTH for Device Supported? -> true
  Is BLUETOOTH for Device On? -> false
Airplane Mode Is On: false
---------------------------------------
```

Example Output

- Asynchronous:

```
----------------------------------------
Fri May 29 09:54:26 CDT 2015
Device was TURNED ON by user! And just finished starting up.
----------------------------------------
Fri May 29 09:54:28 CDT 2015
Bluetooth is TURNING ON and was activated by the user !!
----------------------------------------
Fri May 29 09:54:29 CDT 2015
Bluetooth's Current State: ON
----------------------------------------
```

```
Tue May 26 17:55:11 CDT 2015
User has just STARTED the app!
---------------------------------------
```

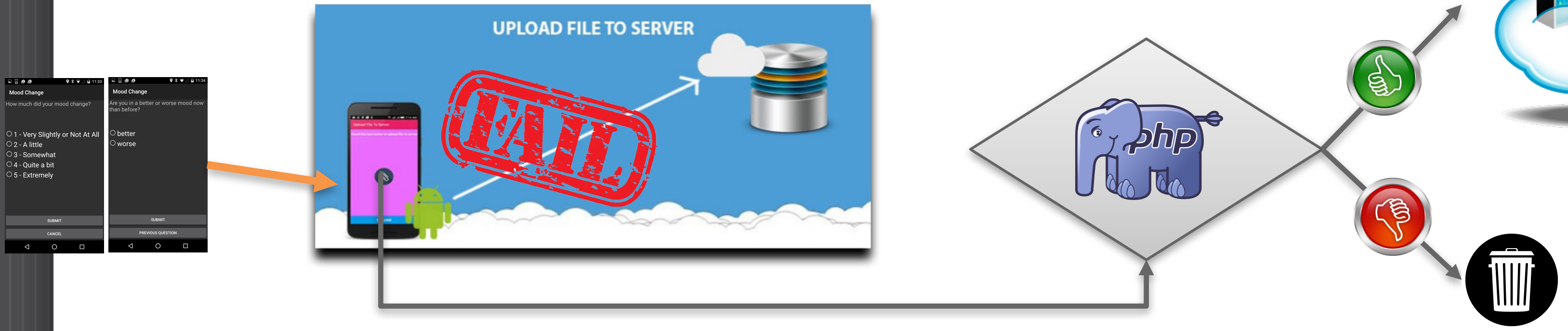
```
Sun Apr 12 16:49:15 CDT 2015
Bluetooth Pairing State is CONNECTING to the Device Named 'Nickolas's MacBook Pro' !!
---------------------------------------
Sun Apr 12 16:49:18 CDT 2015
Active Bluetooth has just been CONNECTED to the Device Named 'Nickolas's MacBook Pro' !!
---------------------------------------
```

```
            User: 0101, Thu Mar 03 17:47:07 CST 2016

The System Clock was JUST CHANGED by user!

The previous time was: Thu Mar 03 19:58:08 CST 2016
The new time is: Thu Mar 03 17:47:07 CST 2016
---------------------------------------
```

```
---------------------------------------
Tue May 26 18:18:22 CDT 2015
User has just CLOSED the app!
---------------------------------------
```

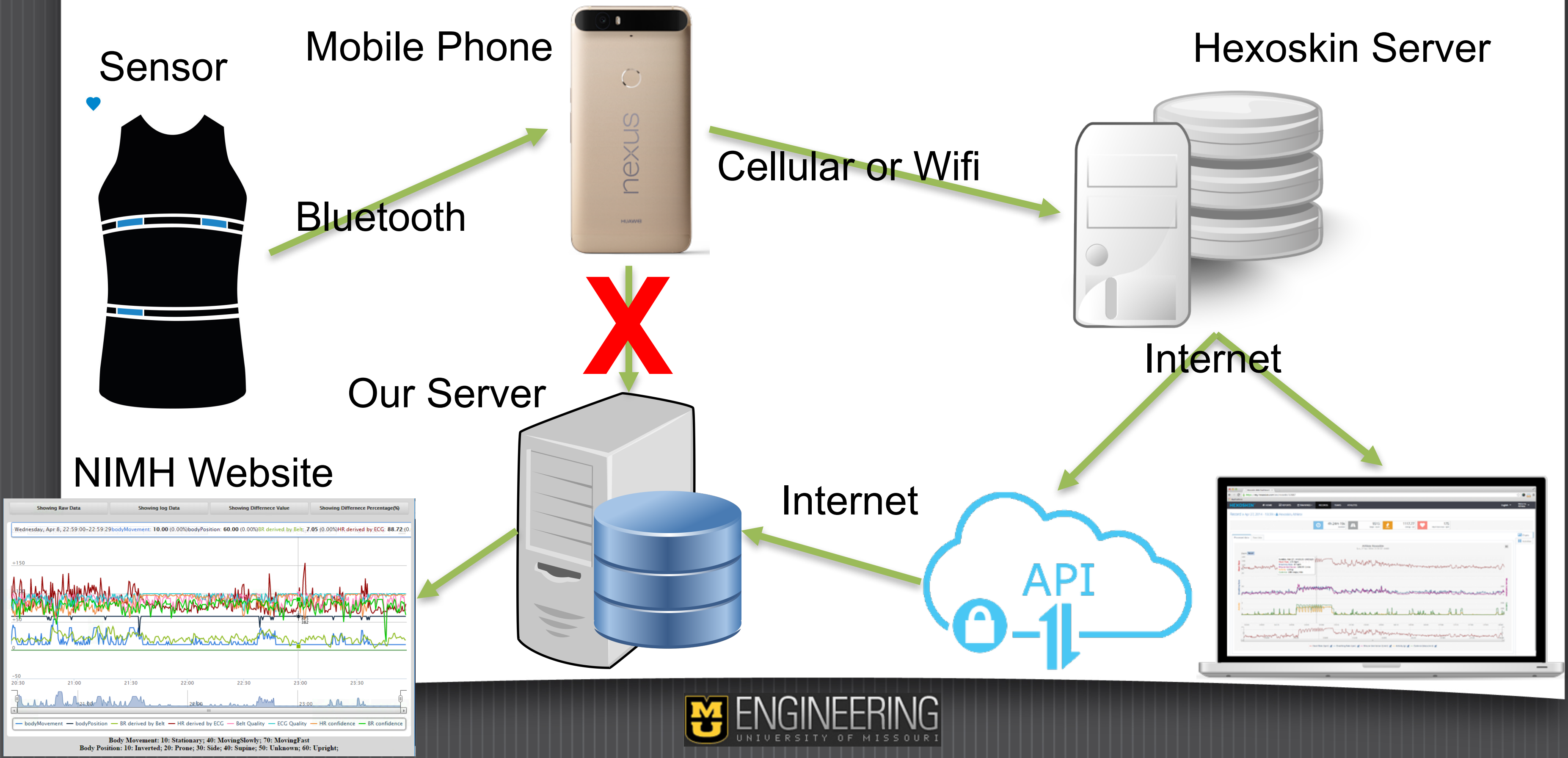
```
----------------------------------------
Fri May 29 14:54:46 CDT 2015
Device is TURNING OFF! And was activated by user!
----------------------------------------
Fri May 29 14:54:47 CDT 2015
Bluetooth is TURNING OFF and was activated by the user !!
----------------------------------------
Fri May 29 14:54:54 CDT 2015
Bluetooth's Current State: OFF
----------------------------------------
```

Upload Missing Survey Data



- <u>Problem:</u> If data does not make it to server, app will not re-upload data and data will remain on phone
  - Psychology department would have to pull data off phone manually
- <u>Solution:</u> An easy to use function on phone, activated by button, to upload all data on phone
  - Server program to handle duplicates and other situations

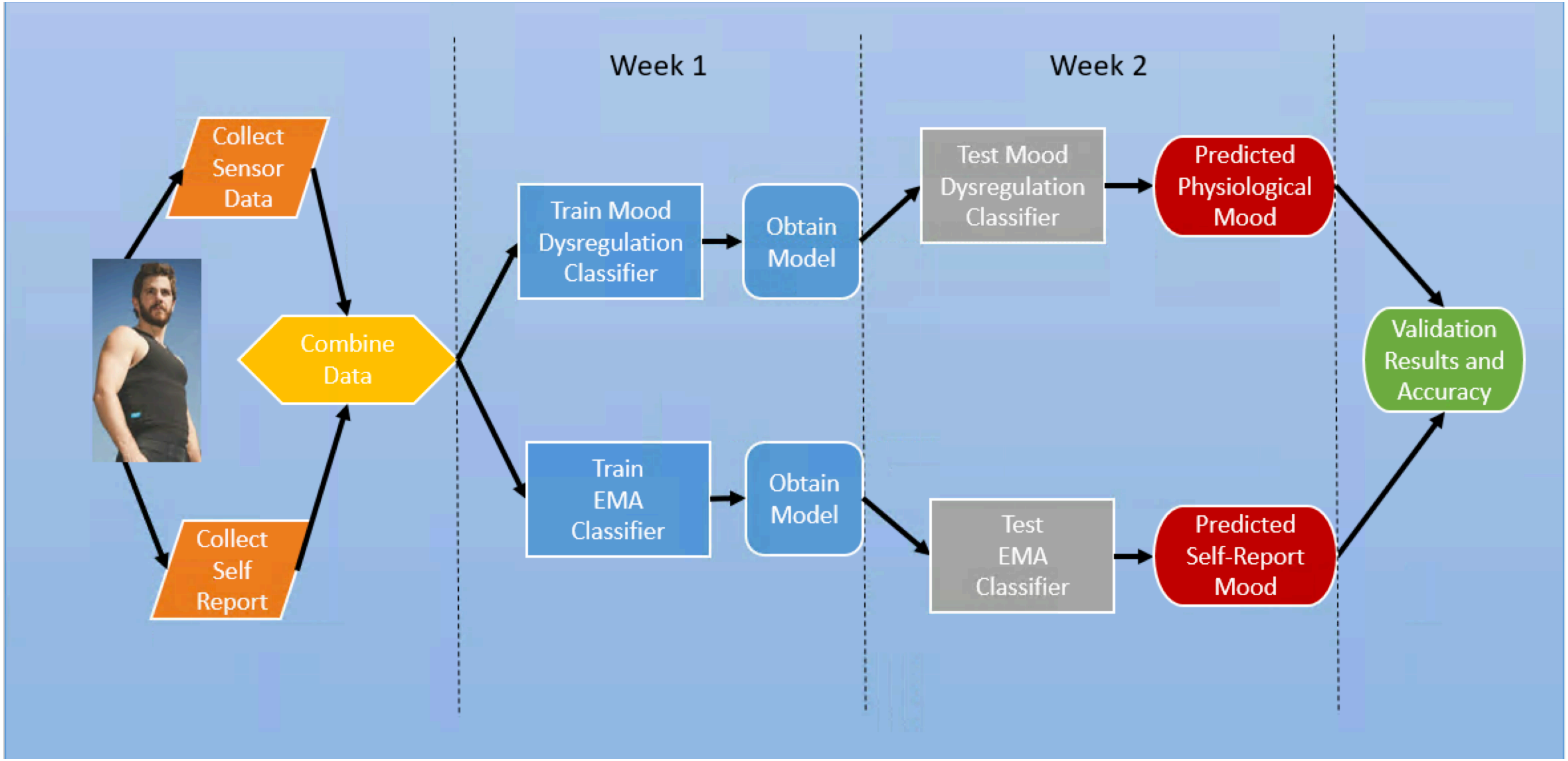
Obtain Hexoskin Data Workflow



Components

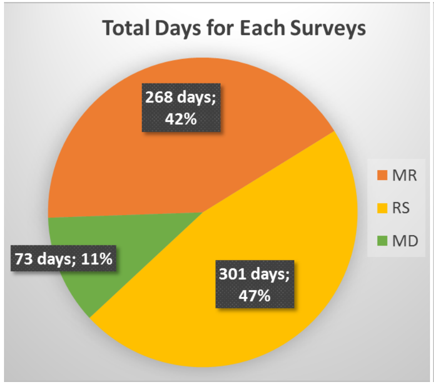
- Introduction
- Related Work
- mAAS System Improvements
- Mood Study Overview
  - Mood Study Procedure
  - Statistics Analysis
  - Visual Statistics
  - Machine Learning (ML) Overview in Study
- AMD Pipeline
- Results and Model Comparison
- Knowledge Gained and Future Work

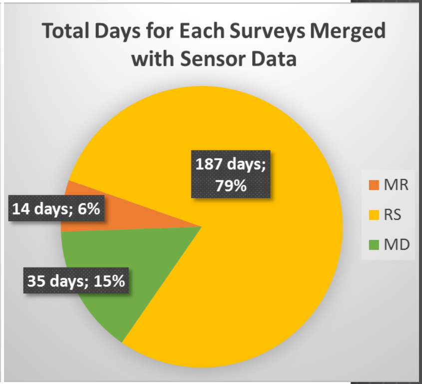
Mood Study Procedure

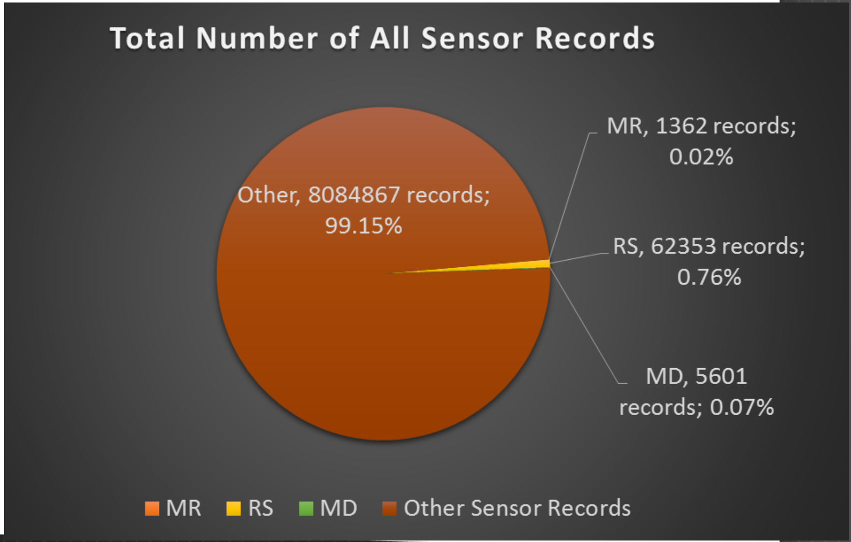


Statistical Analysis

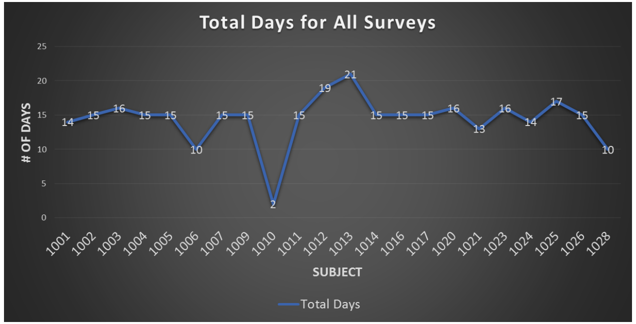
- Hexoskin Sensor Attributes
  - Heart and breathing rate, activity, RR interval, minute ventilation, tidal volume, inspiration and expiration, cadence
- 22 subjects in natural environment
- 318 days of surveys
- 258 days of sensor data
- 8,154,183 sensor records
- After combining with survey data
  - 69,316 associated survey and sensor records
  - -99.15% decrease
  - Less than 1% of sensor data corresponds to survey data

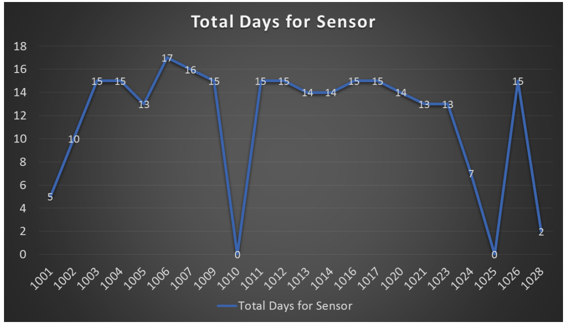


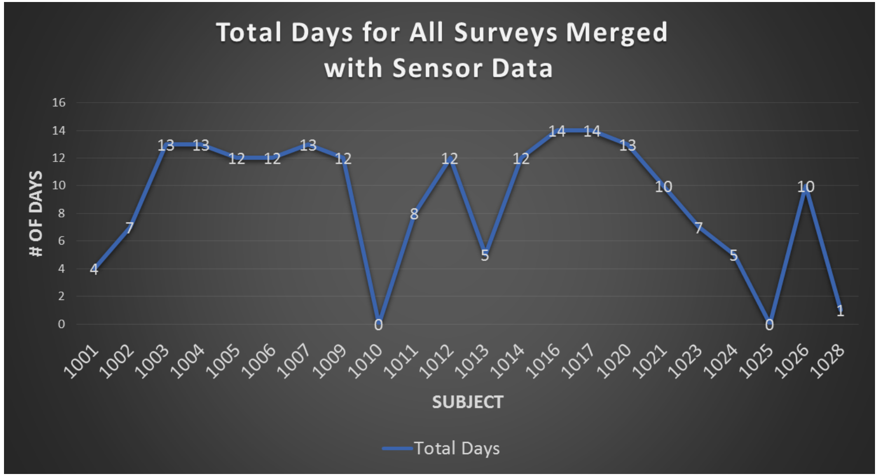


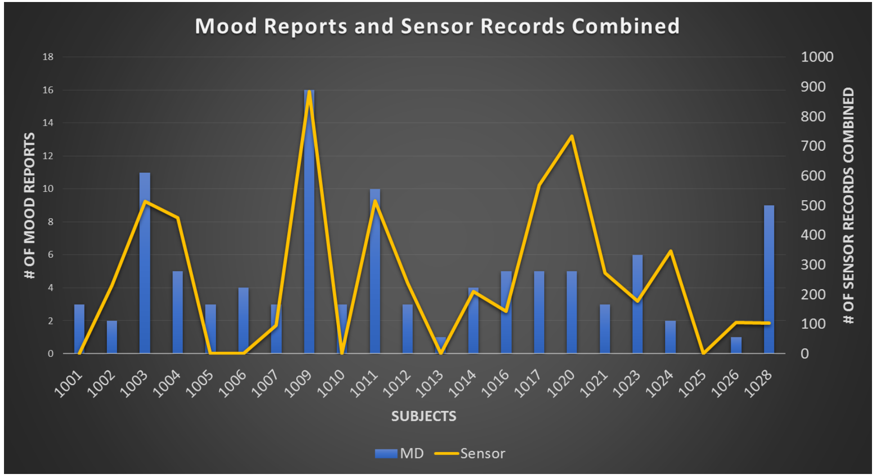


Visual Statistics Physiological

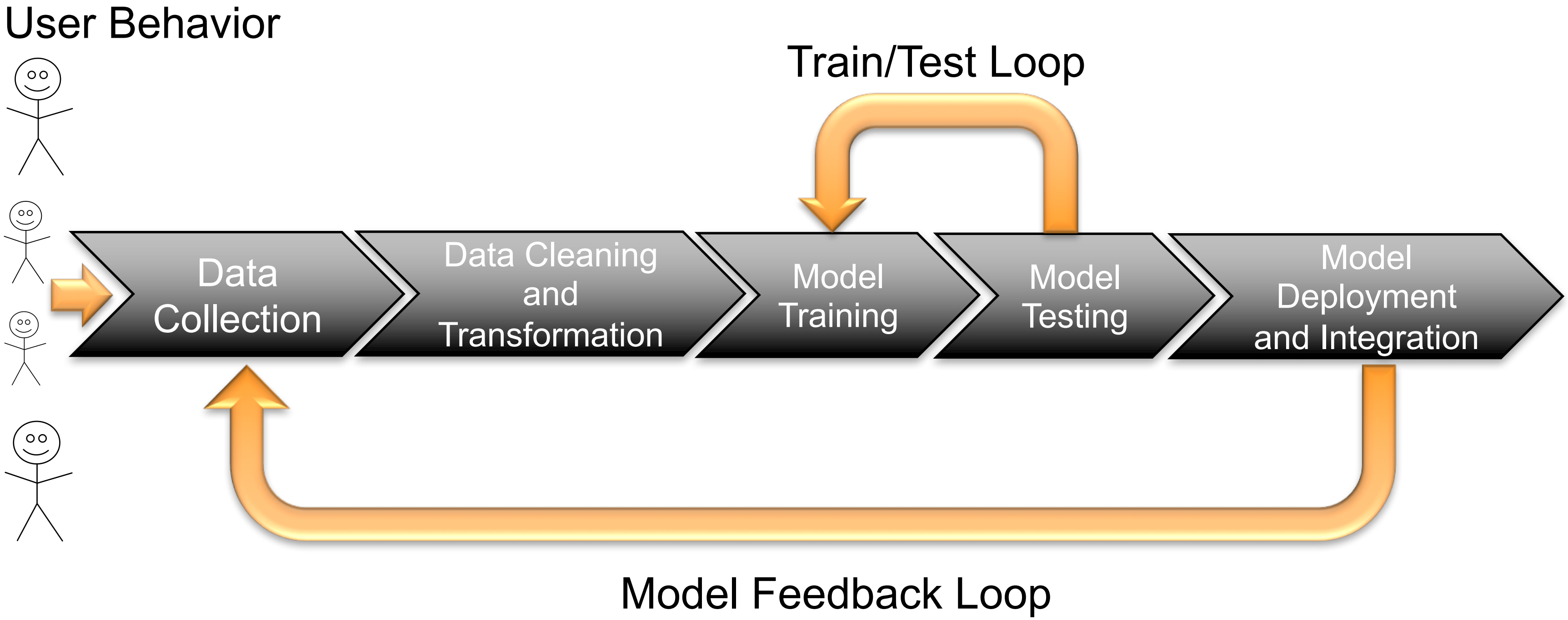








ML Incorporated in Study



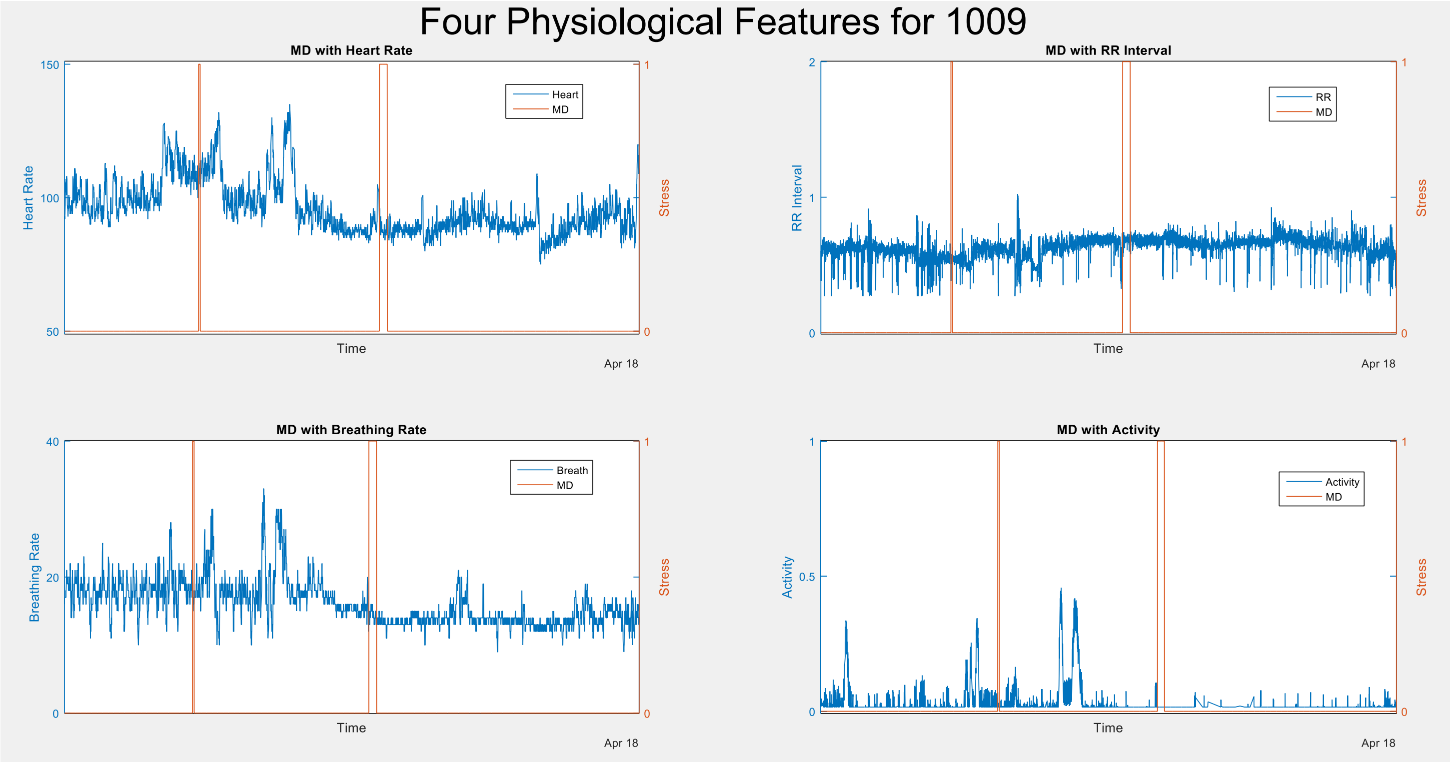
Components

- Introduction
- Related Work
- mAAS System Improvements
- Mood Study Overview
- AMD Pipeline
- Results and Model Comparison
- Knowledge Gained and Future Work

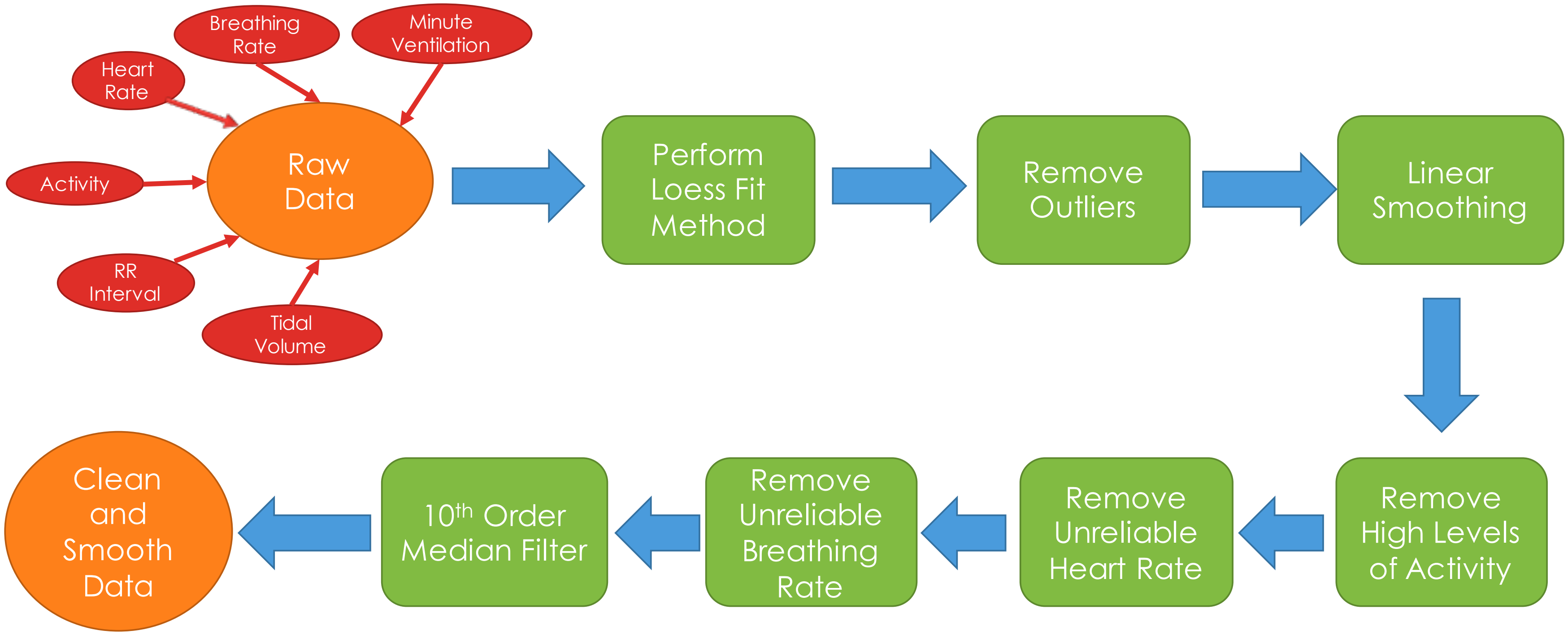
Components

- AMD Pipeline
  - Different Approaches
  - Finalized Pipeline
  - Data Selection and Combination
  - Data Cleaning
  - Feature Creation and Extraction
  - Training and Testing
  - Prediction
  - Experiments with Prediction

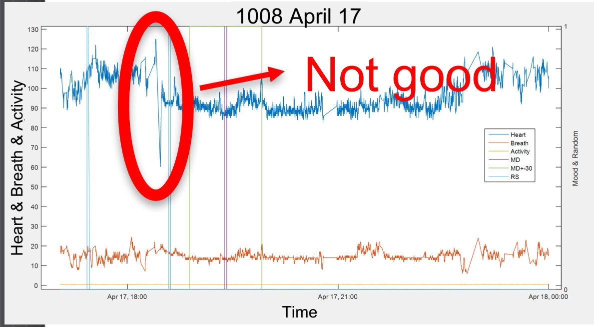
Hexoskin Data

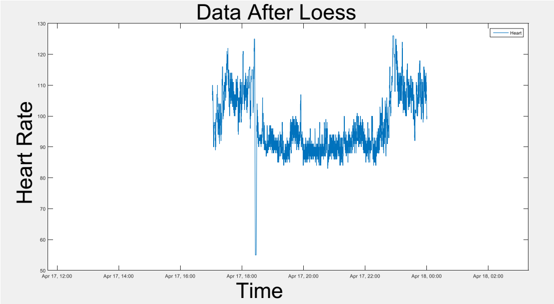


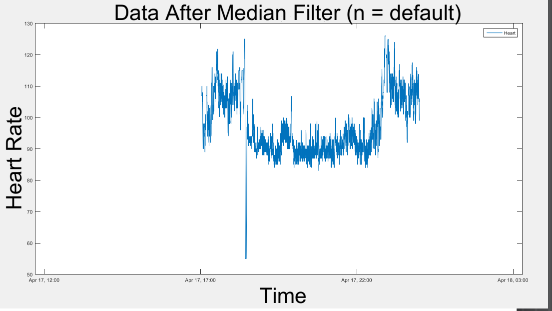
Naïve Cleaning Pipeline

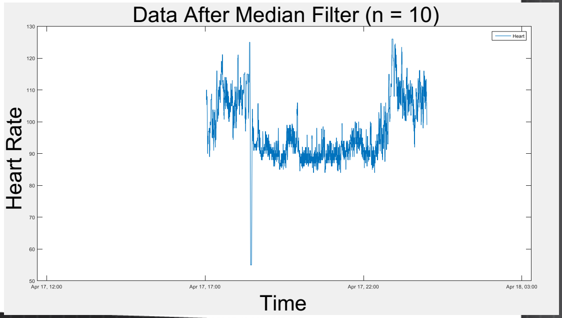


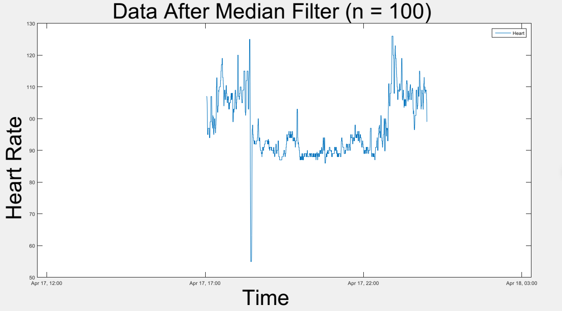
Naïve Cleaning Pipeline

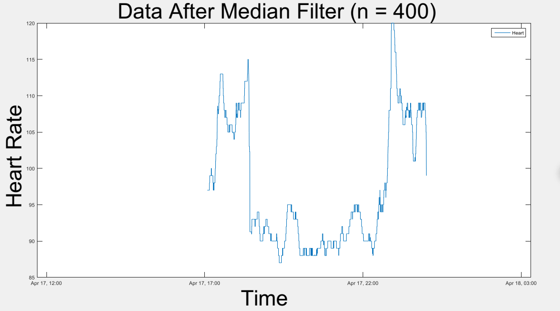




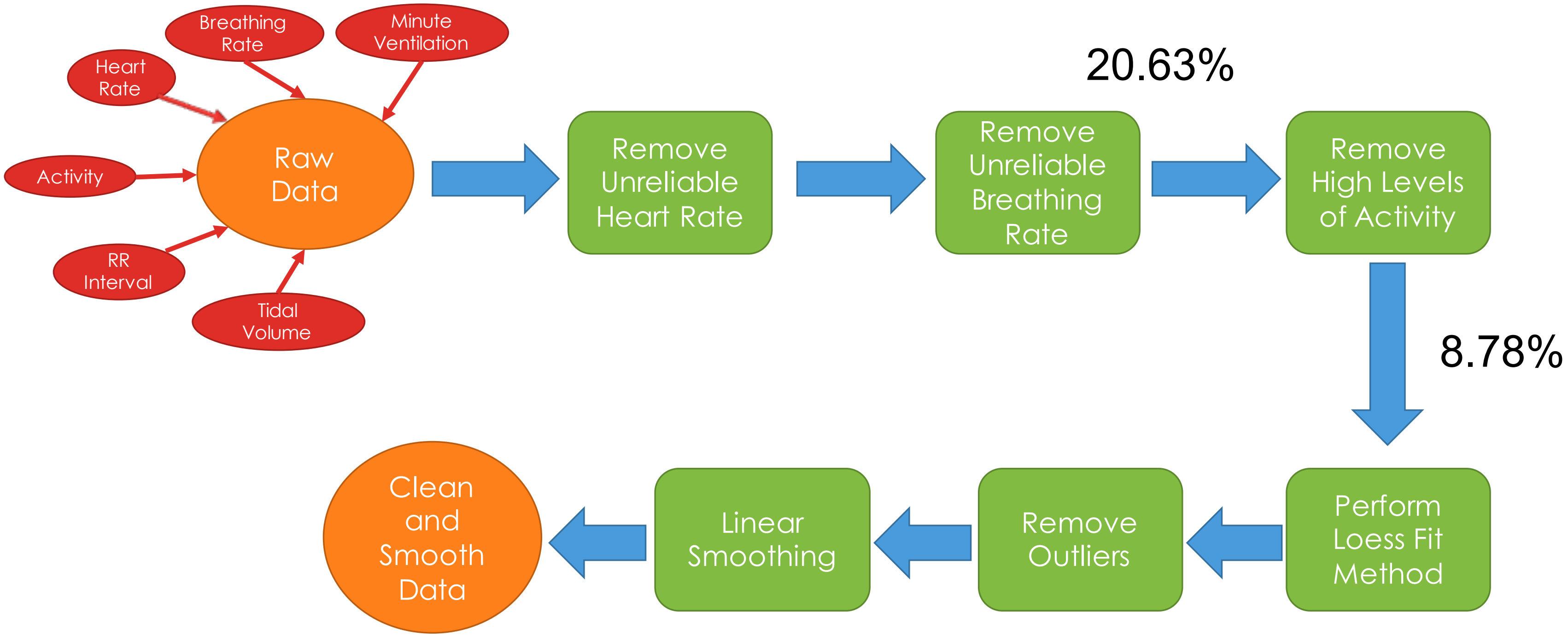








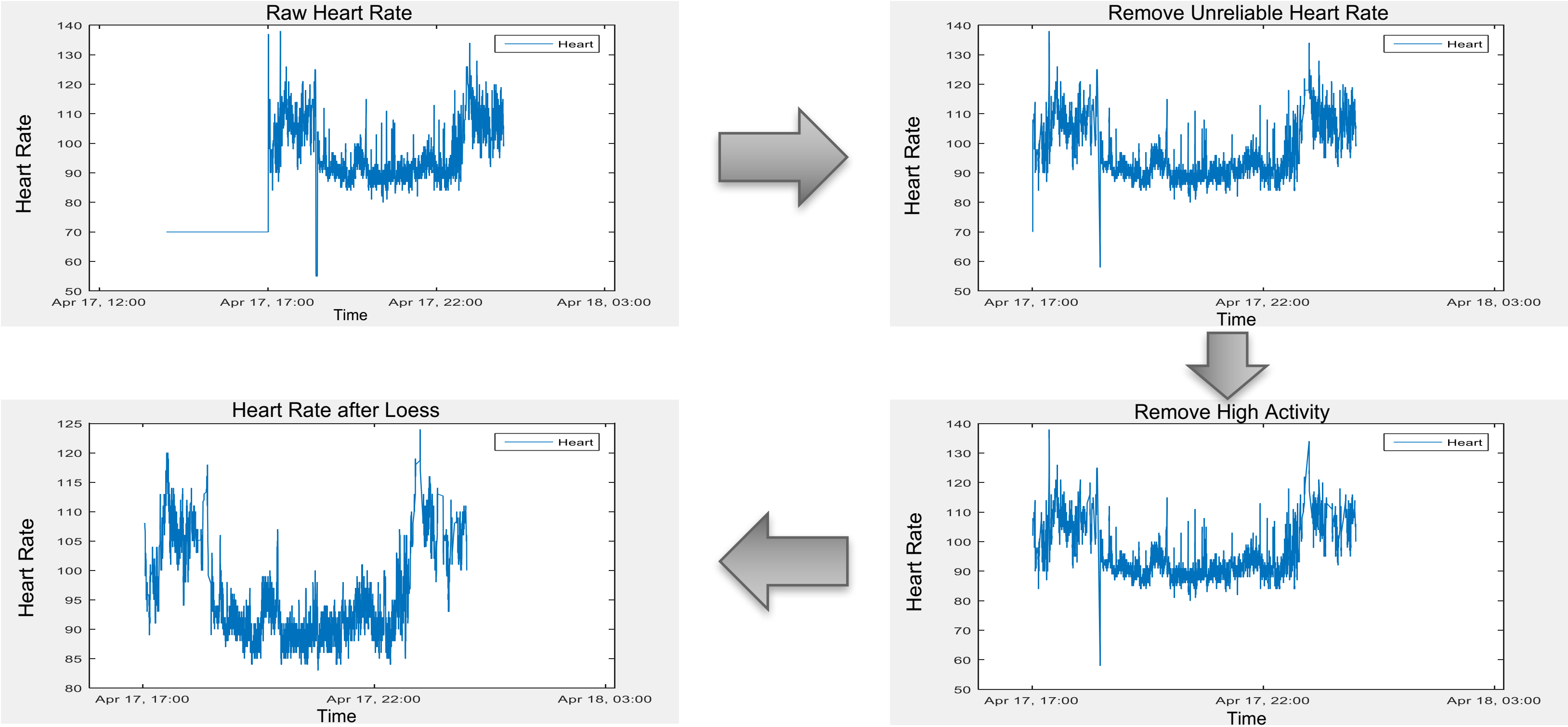
Better Approach



11.27%

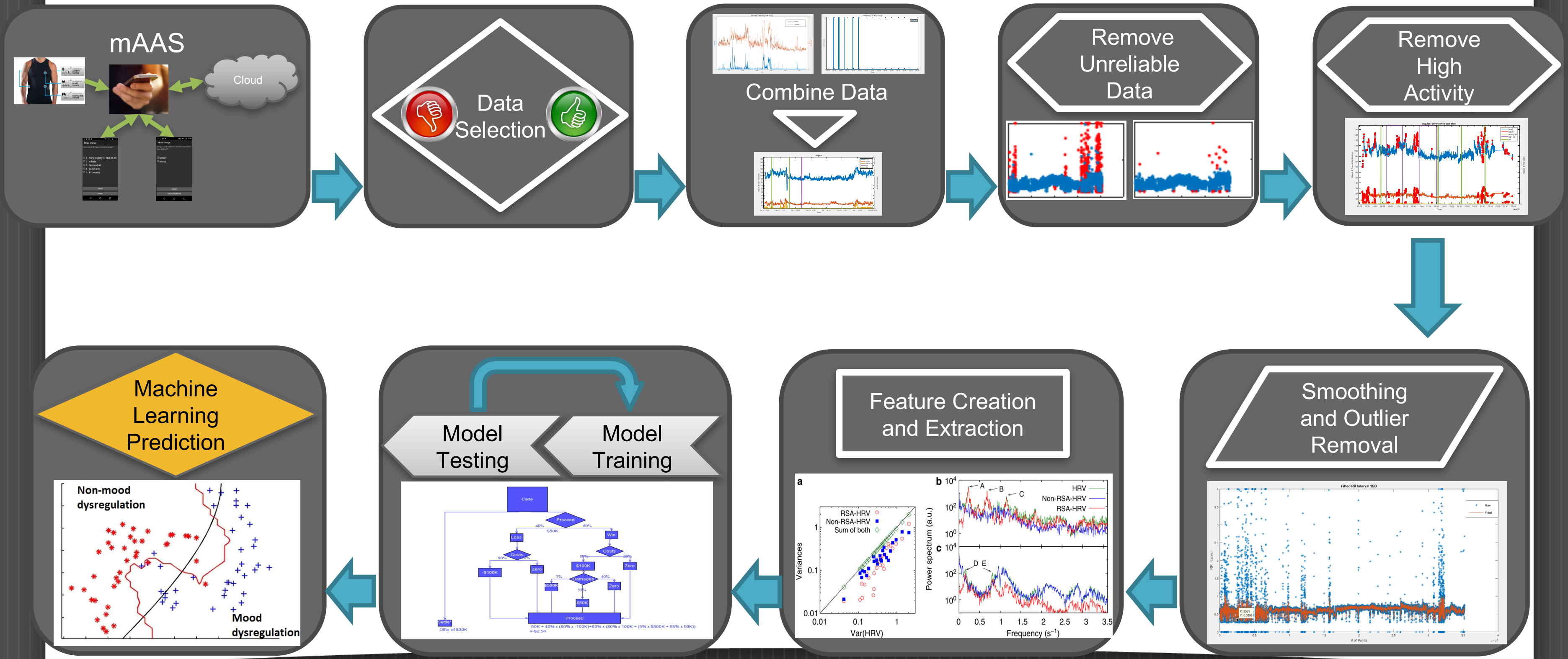
Total % of Data Removed: 40.68%

Better Approach



AMD’s Pipeline

Pipeline was implemented in MATLAB

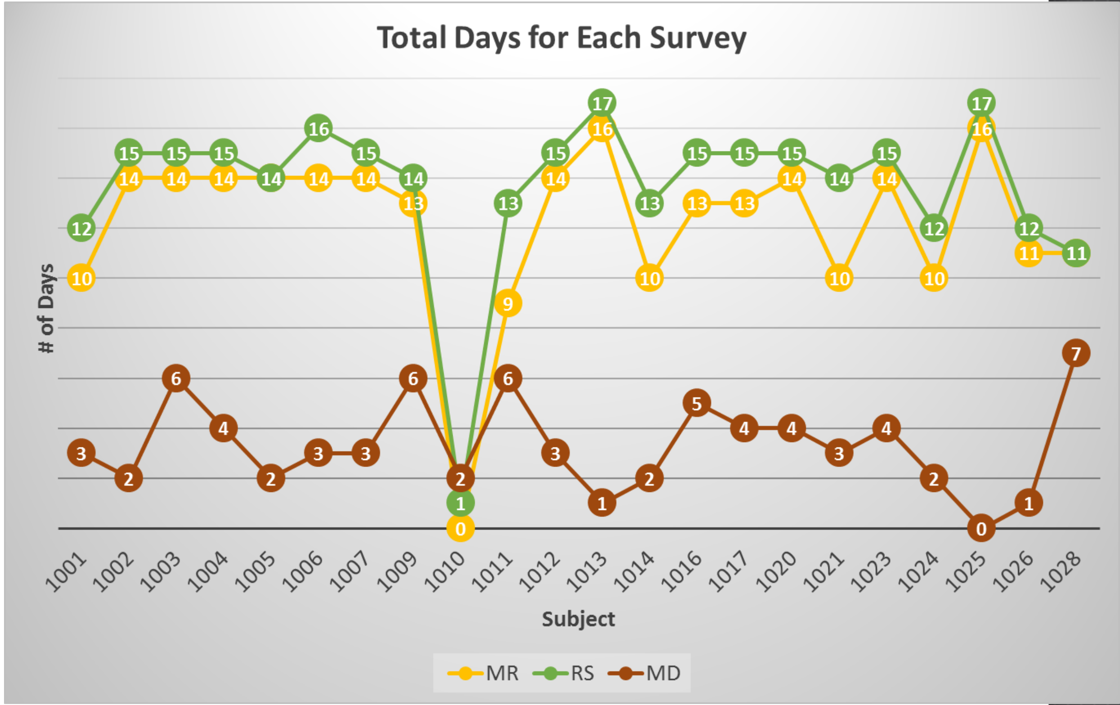


Components

- AMD Pipeline
  - Different Approaches
  - Finalized Pipeline
  - Data Selection and Combination
  - Data Cleaning
  - Feature Creation and Extraction
  - Training and Testing
  - Prediction
  - Experiments with Prediction

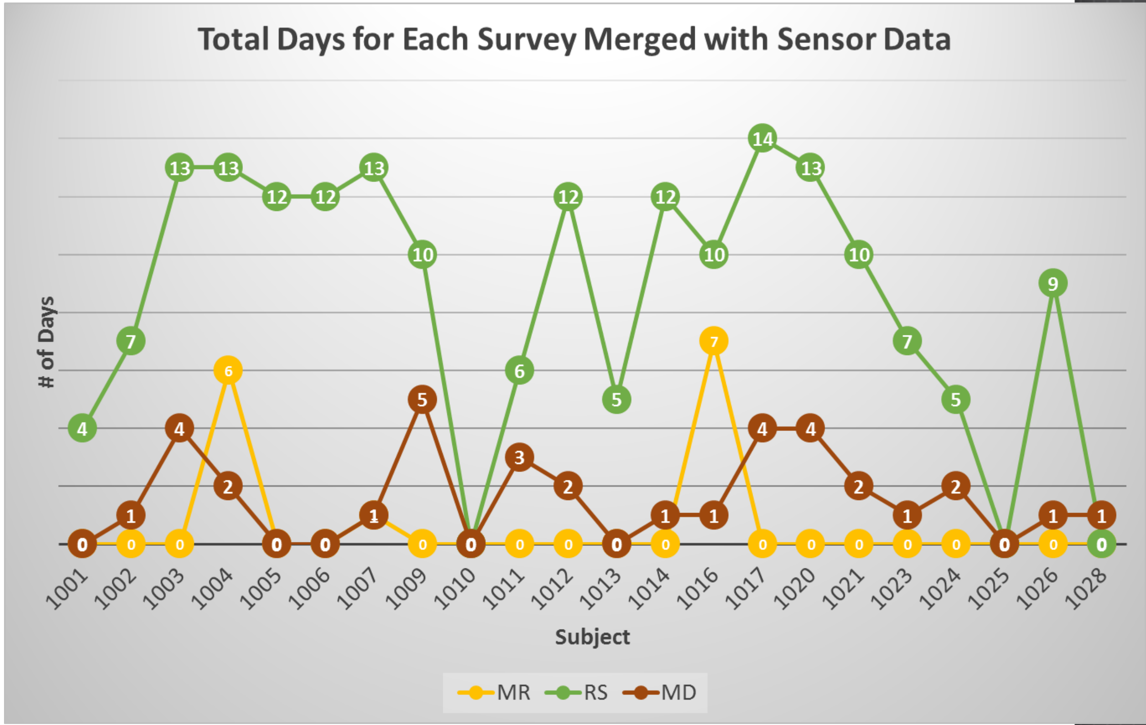
Data Selection

- During my research, a total of 22 participants' data available
  - Consisting of approximately two weeks of sequentially collecting physiological data and self-report mood data
- Only 16 were selected, must have:
  - 9 out of 14 days of self-report data
  - At least one mood dysregulation sample
  - At least one mood sample with corresponding sensor data
  - Valuable sensor data determined by status on sensor



Combine Data

- Hexoskin sensor data and survey data from two different sources
  - Need to be combined into one for machine learning model
- There were approx. 8.5 million sensor data records in DB
  - 1 record corresponds to 1 second
- After combining data, only 5601 sensor records matched in between mood survey start and end times for the 16 users

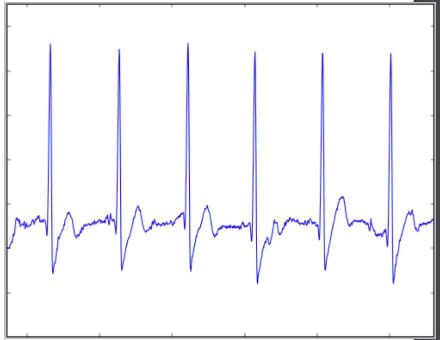


Combine Data (cont.)

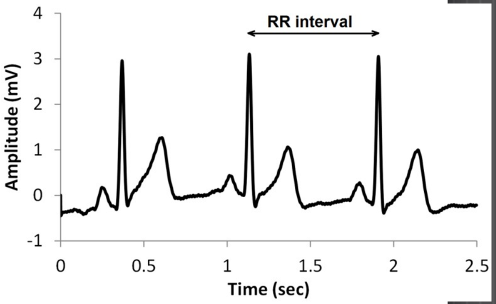
- Reasons?
  - Sensor was not worn by user during time of mood survey
  - Not good connection of sensor during time of survey
- After combining, more participants could not be used
  - Only 10 out of the 16 subjects now able to analyze
  - 6 of them did not have any corresponding survey with sensor data

Remove Unreliable Data

- Hexoskin sensor comes with attributes: Heart Rate Status and Breath Rate Status
  - Heart rate status is confidence of the ECG sensors
  - Breath rate status is confidence of RIP sensors
- If either of these status is unreliable, the data is removed, avg. removed 20.63% of 10 subjects
- This caused more participants to be removed from analysis
  - 7 subjects will continue for machine learning analysis

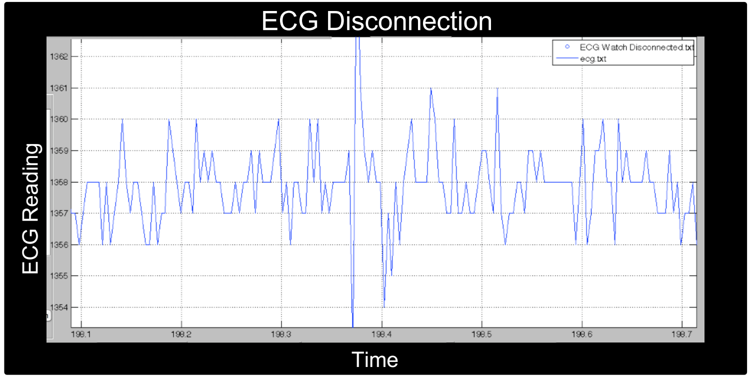


Normal Hexoskin ECG reading

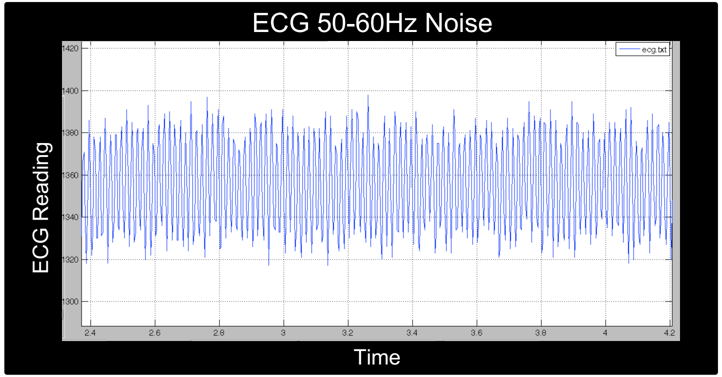


Examples of Unreliable Data

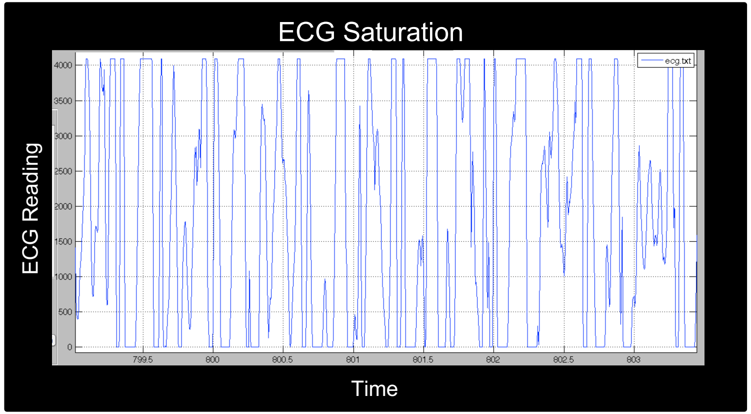
1.



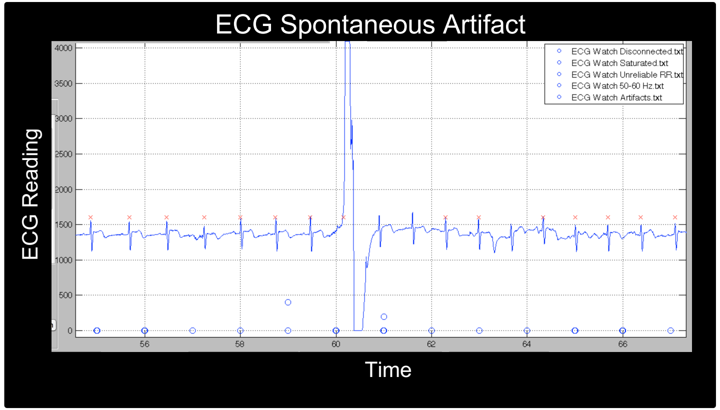
2.



3.



4.

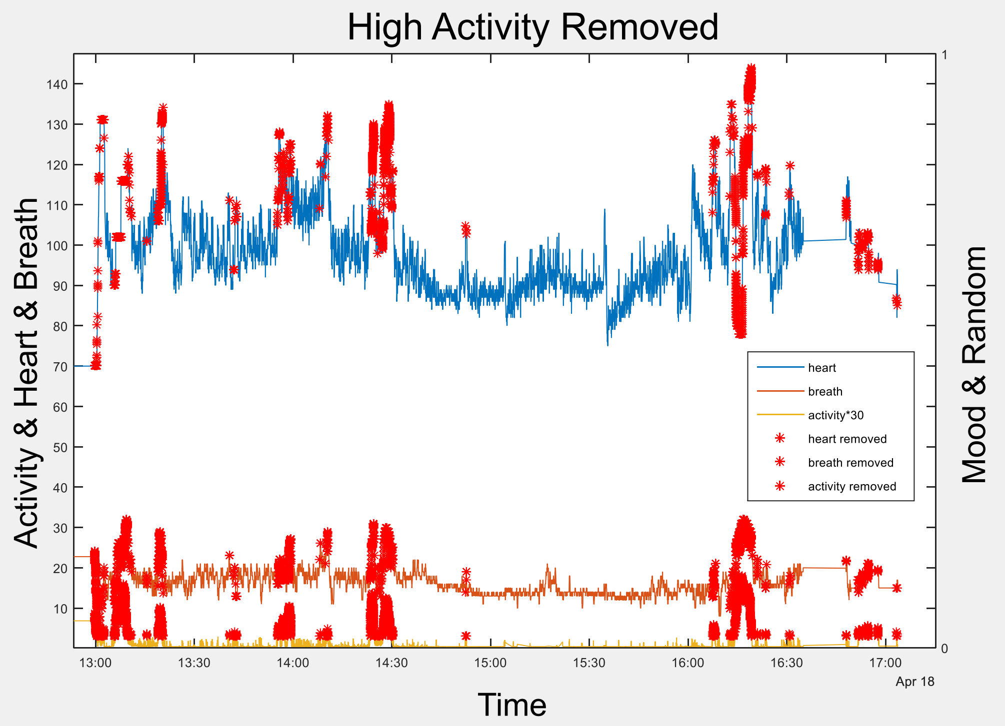


Remove High Activity

- When activity was high, sensor measurements were not reliable and often masks mood dysregulation

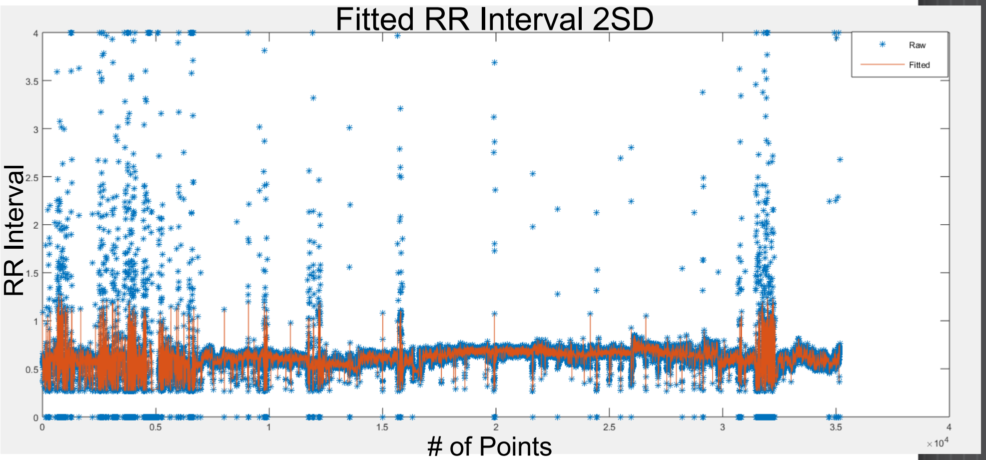
- Two minutes following physical activity were also removed, as described in literature
  - Physiology returns to baseline within two minutes after activity

- Approximately 8.78% removed of 7 subjects



Smoothing and Outlier Removal

- With remaining data, perform Loess Regression
  - Remove outliers 1SD and 2SD away from curve, suggested in literature
    - Better results obtained with 2SD
  - Perform regression imputation to fill in the outliers

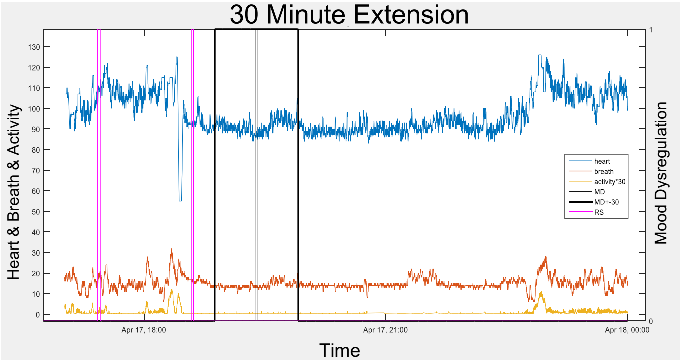


Components

- AMD Pipeline
  - Different Approaches
  - Finalized Pipeline
  - Data Selection and Combination
  - Data Cleaning
  - Feature Creation and Extraction
  - Training and Testing
  - Prediction
  - Experiments with Prediction

Feature Creation and Extraction

- 30 minute extension was created on start and end time of Mood Surveys
  - Now there are 74,042 mood samples instead of 5,601
- 9 Physiological Features were used in Prediction
  - Activity, RR Interval, Heart Rate, Breathing Rate, Minute Ventilation, Expiration, Inspiration, Tidal Volume, and Cadence



- 5 time attributes were created from the TimeStamp
  - Time Category(int), Only Hour, Only Hour and Minute, Only Time, and Time Category(String)
  - Time Categories: 5 < Morning < 11; 11 < Afternoon < 17; 17 < Evening < 23; 23 < Night < 5
- 11 Features were selected, the 9 Physiological Features and 2 Time Features

Training and Testing

- First, splitting the data sequentially was implemented
  - 80% for training, 10% for testing, and 10% for validation
- Next, 10 Crossfold Validation was used for Training and Testing
  - 10 crossfold had better results
  - Approximately 2-3% increase in accuracy
- Four models were implemented to compare results
  - Naïve Bayes, Bayesian Network, J48 Decision Tree, and Random Forest

Components

- AMD Pipeline
  - Different Approaches
  - Finalized Pipeline
  - Data Selection and Combination
  - Data Cleaning
  - Feature Creation and Extraction
  - Training and Testing
  - Prediction
  - Experiments with Prediction

Prediction

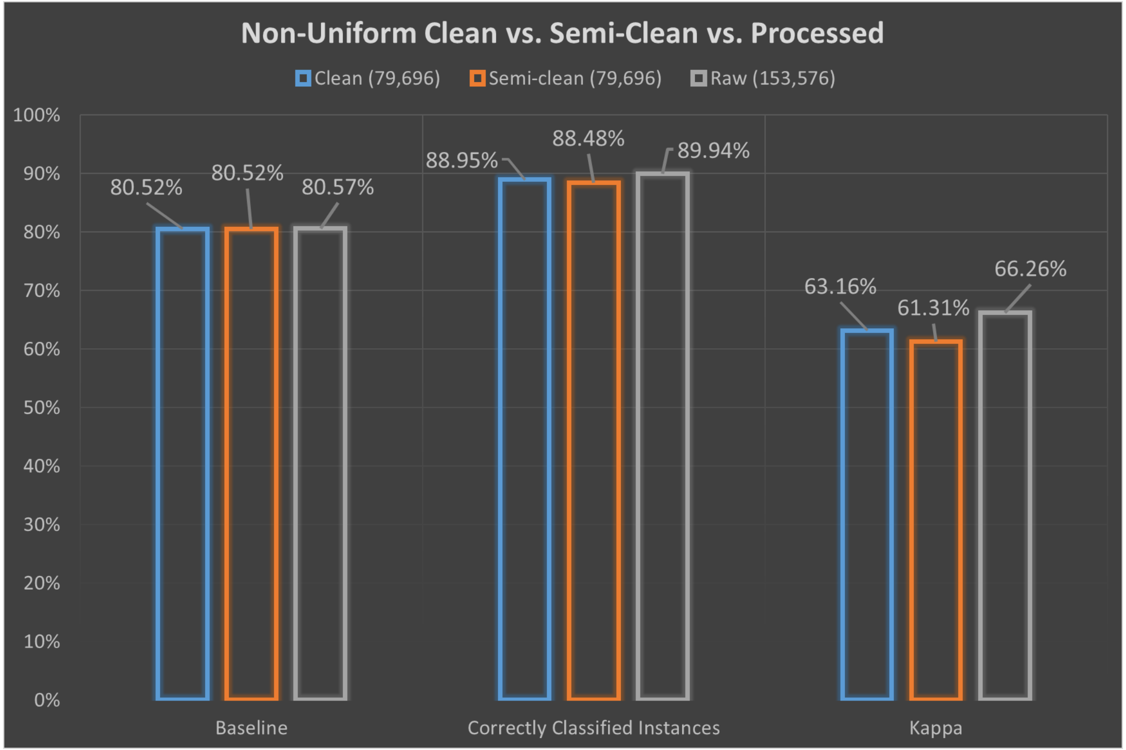
- Purpose: Use machine learning to predict when user is having mood dysregulation based off physiological features input
  - True Positive Labels are Mood Dysregulation Surveys self-initated by user from mAAS
- WEKA was used for machine learning prediction
- Evaluation: Confusion Matrix, Accuracy, Kappa, and ROC Area
  - Kappa > 0.5 acceptable, Overall Accuracy > .80

Components

- Introduction
- Related Work
- mAAS System Improvements
- Mood Study Overview
- AMD Pipeline
- Results and Model Comparison
  - Clean vs. Semi-Clean vs. Raw Data
  - Non-Balanced vs. Balanced Class
  - Time vs. Without Time Features
  - Model Comparison
  - Model Results Across All Subjects
  - Tree Structures and Best Features
- Knowledge Gained and Future Work

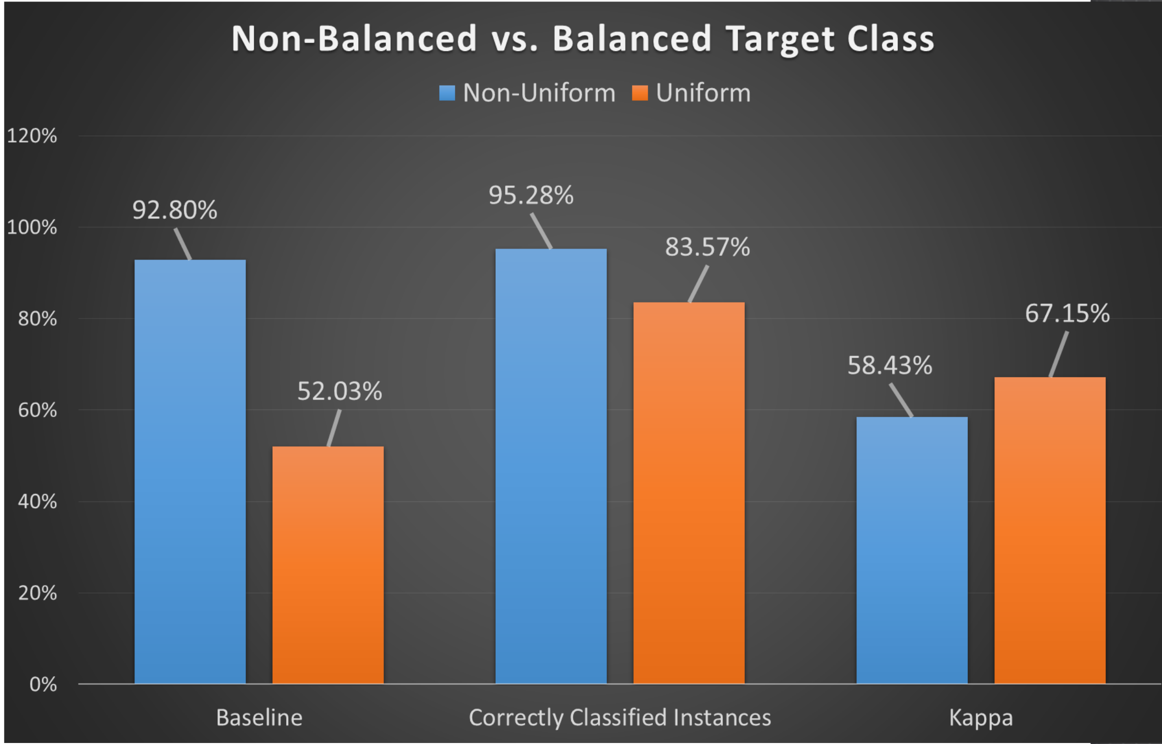
Clean vs. Semi-Clean vs. Raw Data

- Clean Kappa: 63.16%
- Semi Kappa: 61.31%
- Raw Kappa: 66.26%
- Clean > Semi-Clean
- Raw > Clean
- Raw has best result
- Raw is used in further analysis



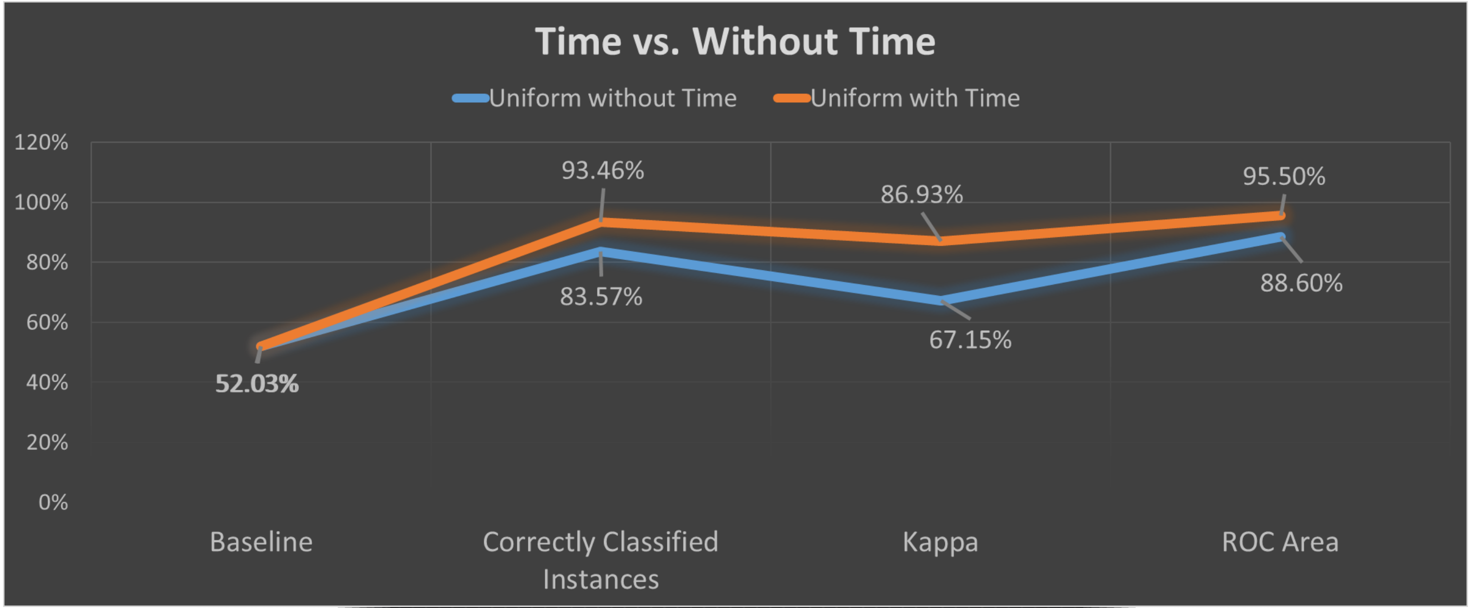
Non-Balanced vs. Balanced Class

- Without uniform
  - Baseline: 92.80%
  - Accuracy: 95.28%
  - Does not make sense
- Uniform has higher Kappa results

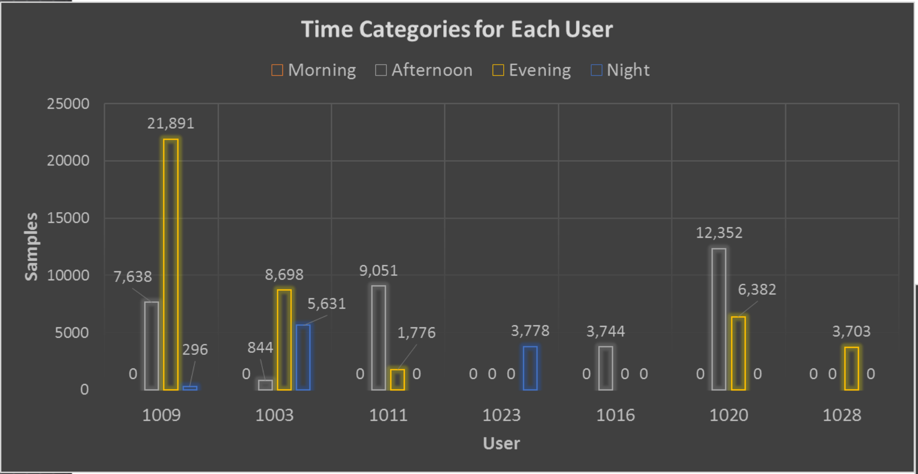


Time vs. Without Time

- Results improved approximately 11%
  - With my custom created features, Hour of Day (Integer) and Time Category (Integer)

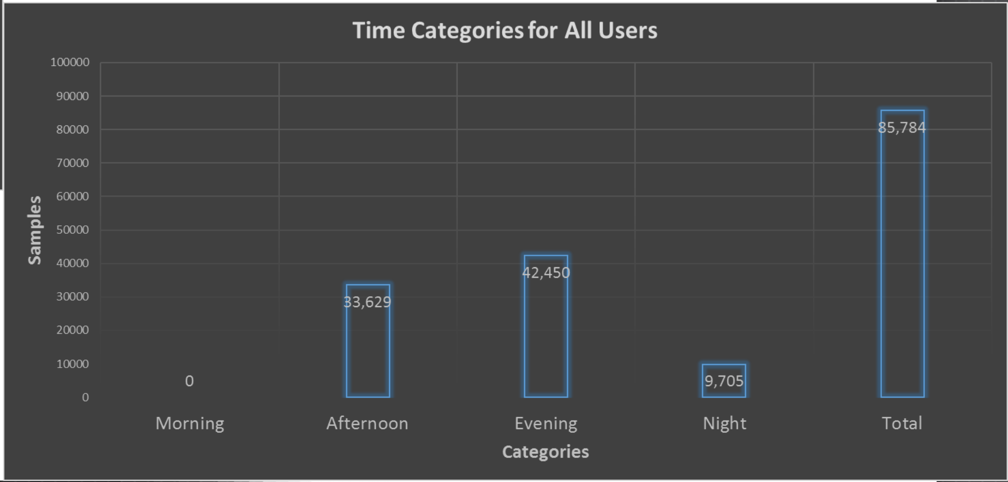


Discussion

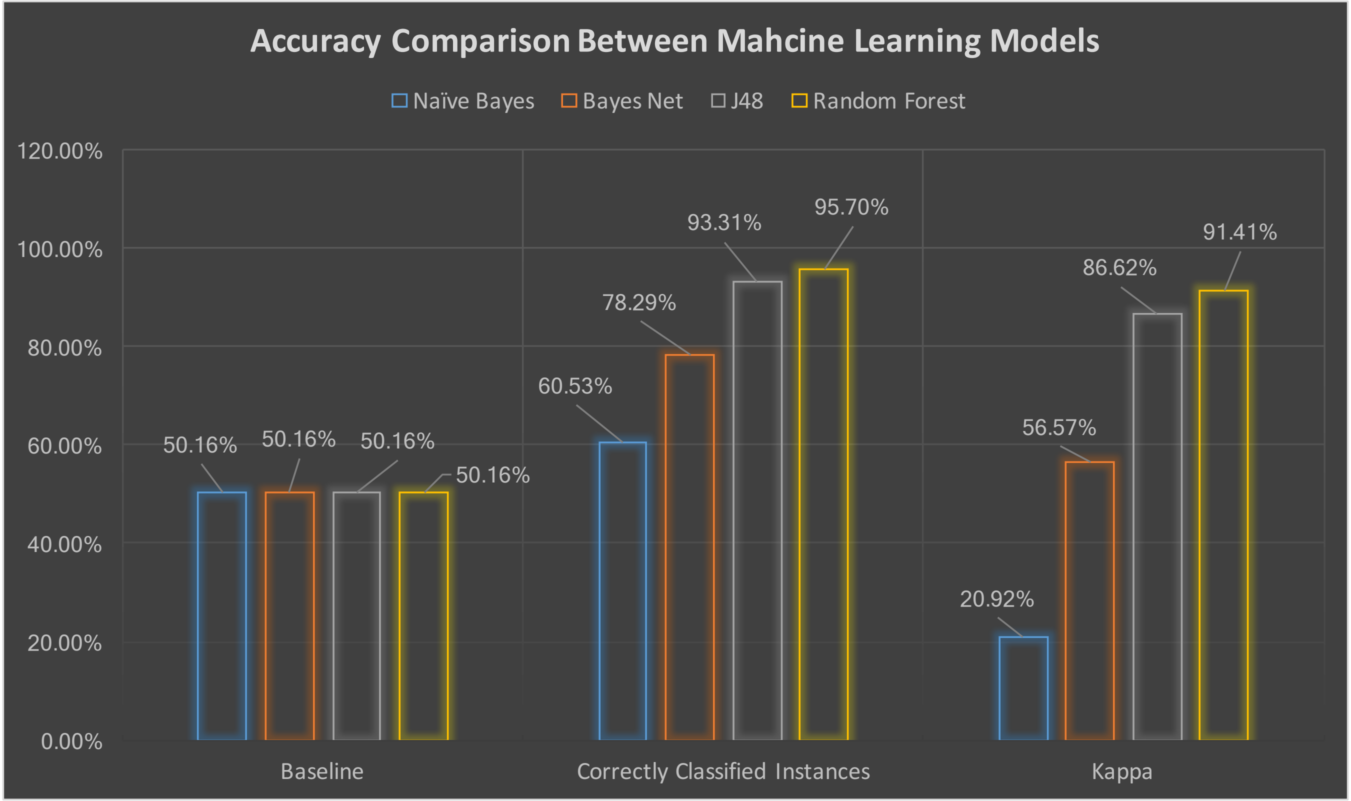


- For each user, most mood dysregulation happens during the Afternoon and Evening Categories

- For all users, you can see the frequency distribution for Afternoon and Evening categories, 39.20% and 49.48% respectively

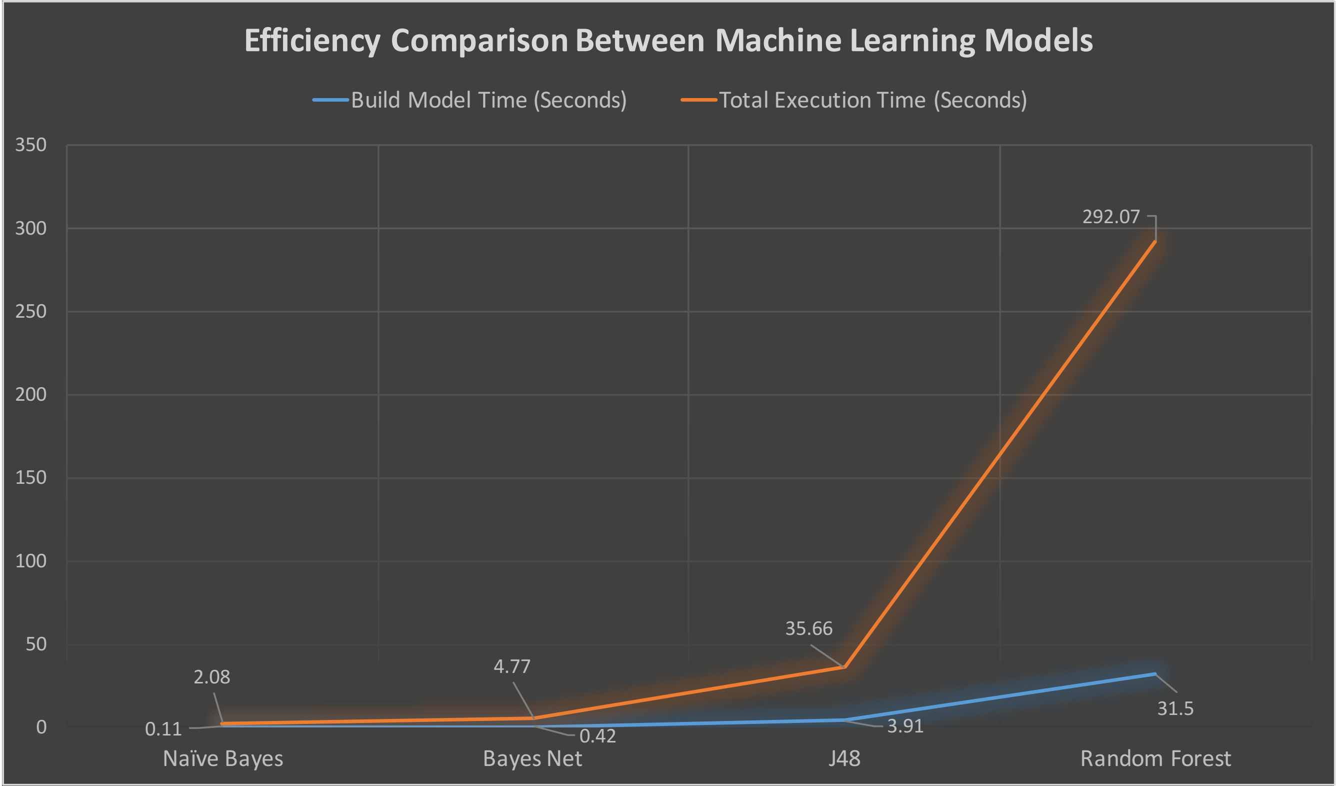


Bayes vs. J48 vs. Random Forest Accuracy



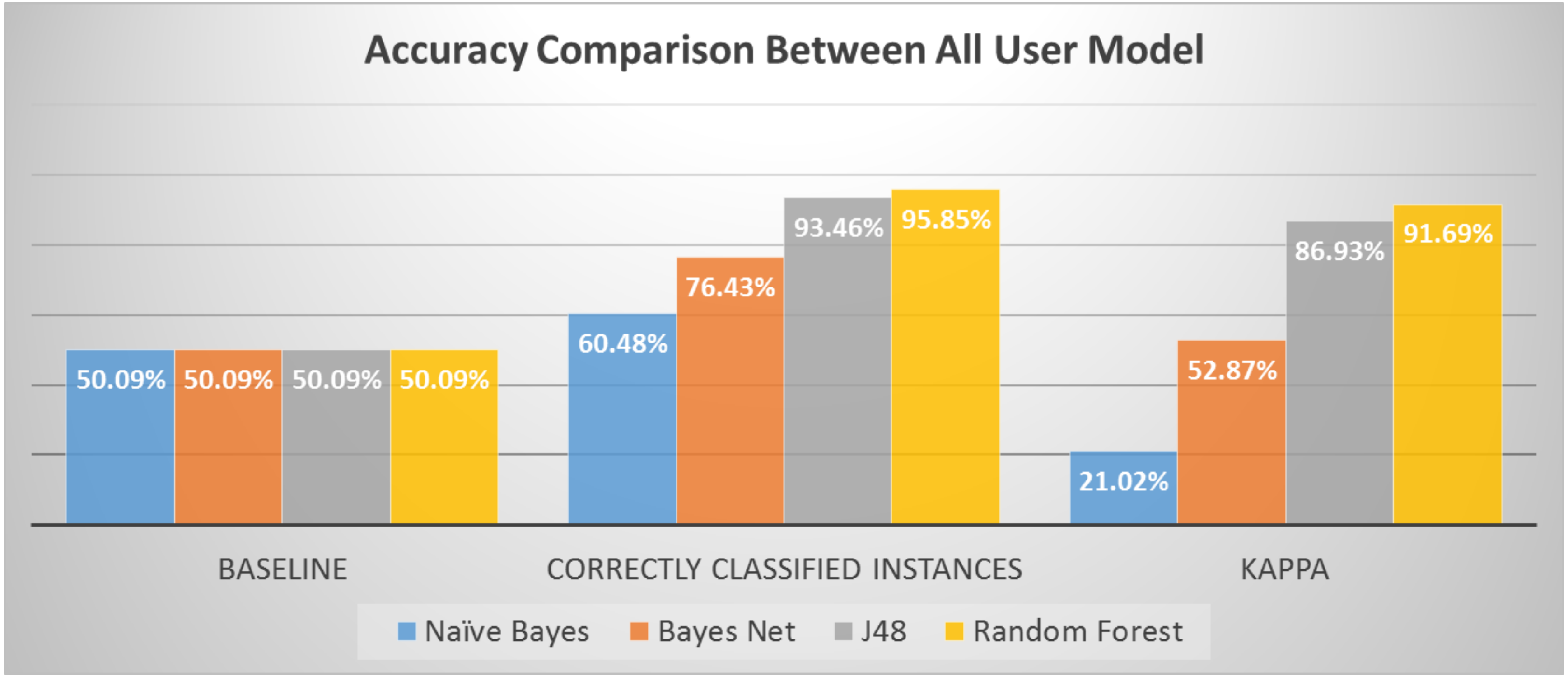
ENGINEERING
UNIVERSITY OF MISSOURI

Bayes vs. J48 vs. Random Forest Efficiency

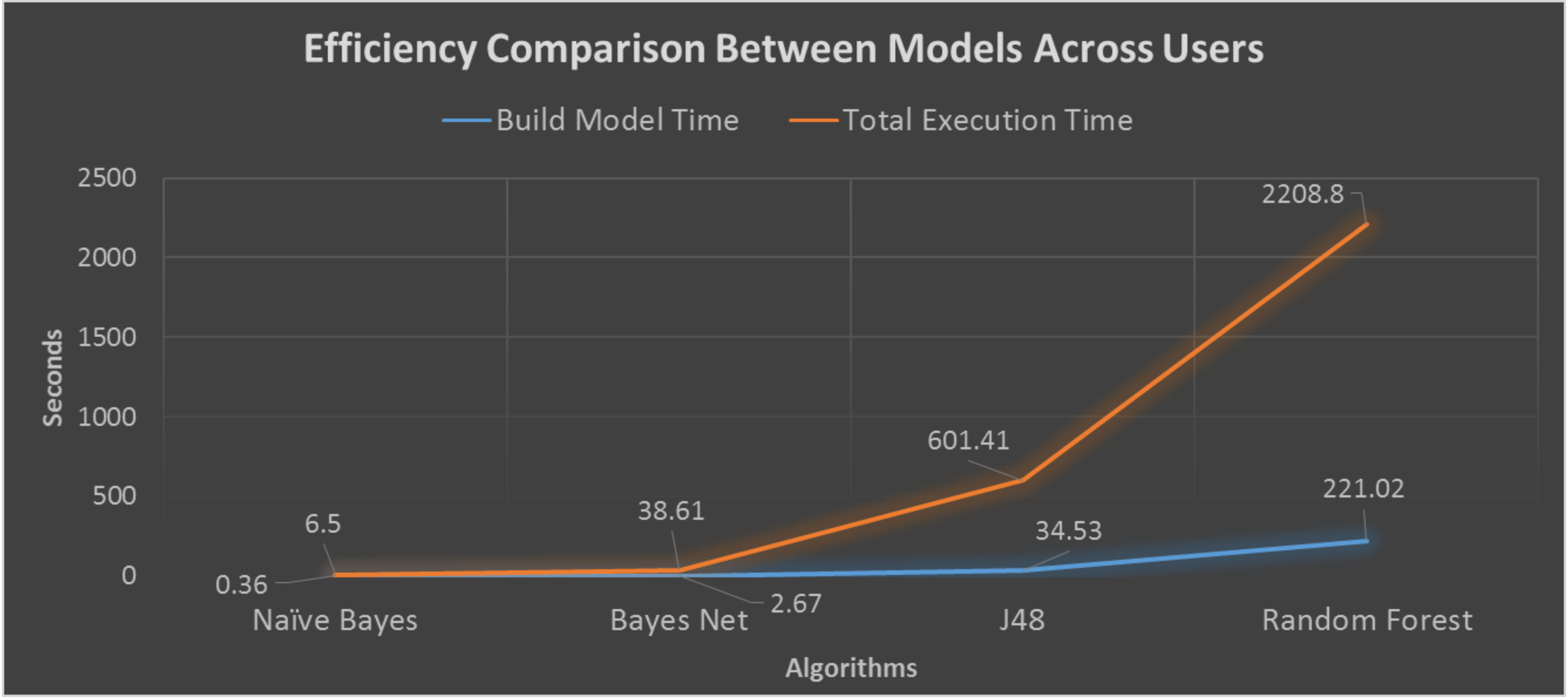


ENGINEERING
UNIVERSITY OF MISSOURI

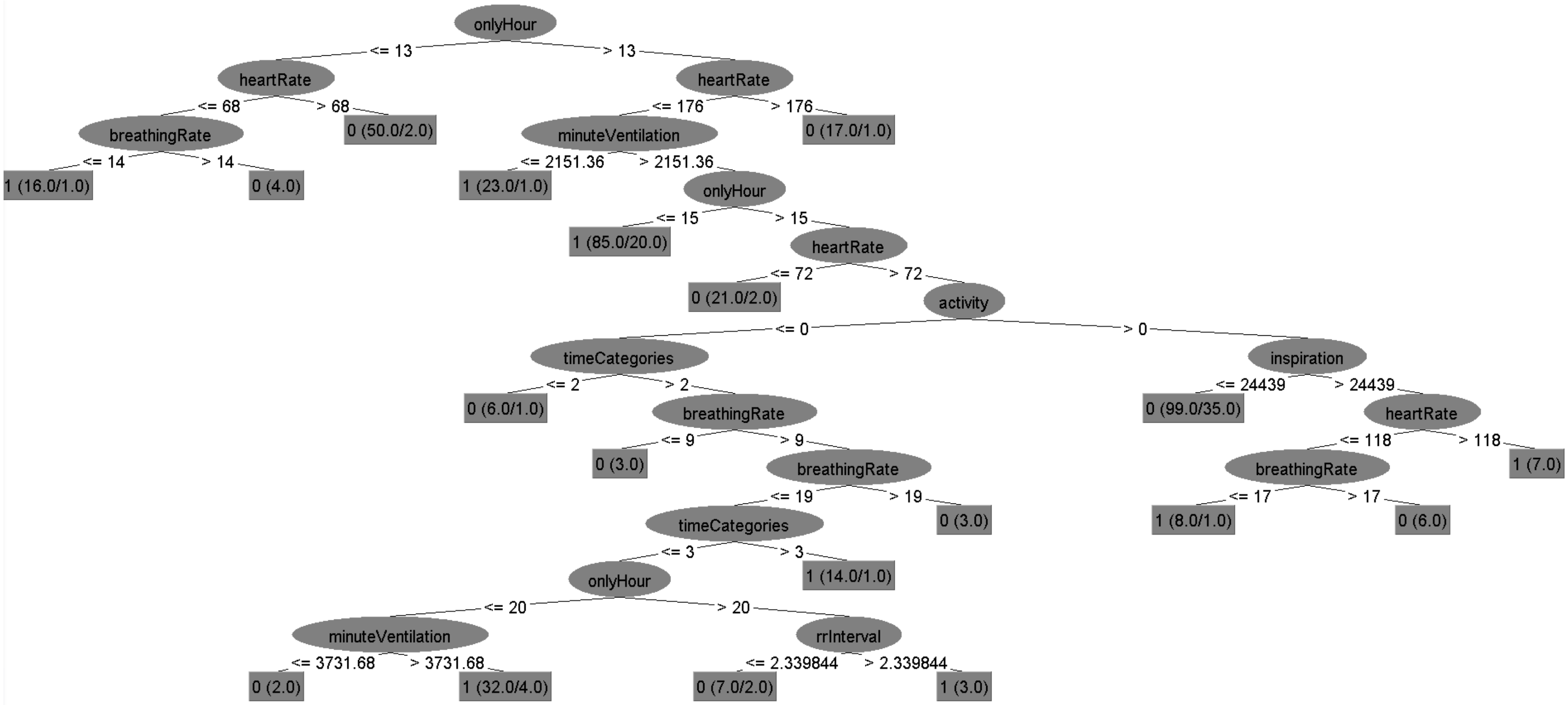
Accuracy Across All Subjects



Efficiency Across All Subjects



Tree Structure for One User



Tree Structure for All Users

- Time Categories is most valuable
- Next is inspiration and expiration
- Proceeding is Heart rate and Hour of Day

```
J48 pruned tree
------------------

timeCategories <= 1: 0 (76001.0)
timeCategories > 1
|   expiration <= 26165
|   |   inspiration <= 26202
|   |   |   heartRate <= 179
|   |   |   |   onlyHour <= 13
|   |   |   |   |   heartRate <= 68
|   |   |   |   |   |   onlyHour <= 1
|   |   |   |   |   |   |   expiration <= 7355
|   |   |   |   |   |   |   |   breathingRate <= 14
|   |   |   |   |   |   |   |   |   minuteVentilation <= 6706.4
|   |   |   |   |   |   |   |   |   |   activity <= 0
|   |   |   |   |   |   |   |   |   |   |   onlyHour <= 0
|   |   |   |   |   |   |   |   |   |   |   |   minuteVentilation <= 3439.52: 0 (1425.0)
|   |   |   |   |   |   |   |   |   |   |   |   minuteVentilation > 3439.52
|   |   |   |   |   |   |   |   |   |   |   |   |   heartRate <= 60
|   |   |   |   |   |   |   |   |   |   |   |   |   |   breathingRate <= 7: 1 (2006.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   breathingRate > 7
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   minuteVentilation <= 6148.64: 0 (69.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   minuteVentilation > 6148.64
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   heartRate <= 56: 1 (354.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   heartRate > 56
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   rrInterval <= 1.148438: 0 (17.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   rrInterval > 1.148438: 1 (17.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   heartRate > 60
|   |   |   |   |   |   |   |   |   |   |   |   |   |   breathingRate <= 6
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   tidalVolume <= 717.12: 0 (63.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   tidalVolume > 717.12
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   rrInterval <= 0.800781
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   breathingRate <= 5: 0 (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   breathingRate > 5: 1 (12.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   rrInterval > 0.800781: 1 (169.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   breathingRate > 6: 0 (650.0)
|   |   |   |   |   |   |   |   |   |   |   onlyHour > 0
|   |   |   |   |   |   |   |   |   |   |   |   tidalVolume <= 252.32
|   |   |   |   |   |   |   |   |   |   |   |   |   minuteVentilation <= 2204.48
|   |   |   |   |   |   |   |   |   |   |   |   |   |   breathingRate <= 6: 0 (35.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   breathingRate > 6
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   heartRate <= 61
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   heartRate <= 54
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   minuteVentilation <= 1938.88: 1 (42.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   minuteVentilation > 1938.88: 0 (40.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   heartRate > 54
```

Components

- Introduction
- Related Work
- mAAS System Improvements
- Mood Study Overview
- AMD Pipeline
- Results and Model Comparison
- Knowledge Gained and Future Work

Knowledge Gained

- Creating time categories for each day improves accuracy of AMD by more than 10%
- Mood dysregulation happens during afternoon and evening categories 60.54% more than the morning and night categories
- Using pre-processed data from Hexoskin achieved higher accuracy than using cleaned data
- Accuracy and Efficiency are consistent across Single User Model and All User Model
  - J48 would be the model of choice since accuracy is high, similar to Random Forest, and 10 times faster than Random Forest

More Knowledge Gained

- For all user model, most important attribute is Time Category
  - The second most important attribute is Inspiration
- For single user model, most important attribute is Hour of Day
  - The second most important attribute is Inspiration
- Out of physiological measurements
  - Inspiration is the most valuable measure for mood
  - Heart Rate is second most valuable measure for mood

Future Work

- Implement AMD on a mobile device
  - Current the pipeline is on a server with unlimited resources
- Implement Deep Learning pipeline and compare it to AMD's results
  - Next, implement the deep learning pipeline on a mobile device
- More users have collected data since my analysis
  - My research can be duplicated on the new users
- More surveys can be analyzed for all users
  - There are 62,353 Random Prompt Surveys out of 69,316 surveys this research analyzed
- My pipeline can be used on 4 other Psychology studies being implemented
  - Mood Toolkit Study, TigerAware Study, SLU HIV Study, and Alcohol Craving Study

Special Thanks Goes To:

Professor Yi Shang

and

Professor Tim Trull

Discussion