Kaylyn King
CSCE 5612

Milestone 3

Note: I have both my ideas for my machine learning in my device as well as my app. I wanted to add the app idea here because I think it is a cool idea to incorporate on my own time. For this project, I will probably only have time for the on-device model (Anxiety Detection Model) given it is already almost the end of the semester

Use-Case Scenarios and ML Objectives

Real-World Scenarios

Anxiety is different for each person, therefore there are many scenarios where this system will be useful depending on the person.

The following examples will be when anxiety spikes:

- Testing or presentation anxiety in a school or workplace
- Social anxiety in crowded places
- Travel anxiety at airports or public transport

The following examples will be when anxiety is lowered:

- Therapy sessions
- Deep breathing
- Self-care sessions

Detection, Prediction, Classification

Anxiety Detection Model

This model will classify whether the user is in an anxious or calm state based on inputs from PPG sensors, pressure sensors, accelerometer, and gyroscope.

The following is what will be measured from each sensor:

- **PPG**: heart rate and blood oxygen
- **Pressure**: muscle pressure
- **Accelerometer and Gyroscope**: shakiness

Kaylyn King
CSCE 5612

Mitigation Strategy Recommendation Model

This model will recommend or predict an effective mitigation strategy based on the user's state and input.

The inputs for this model will be:

- Current anxiety state
- User feedback
- Historical effectiveness of strategies

Purpose of ML

My machine learning will have two purposes:

1. To classify the user's current anxiety level with the biometric data given from the sensors
2. To predict the most effective mitigation strategies for the user when in a state of stress through reinforcement learning

Expected Outcomes

1. **Real-time anxiety detection**: The wearable system should continuously monitor the user's actions to alert the user of rising stress in case they are not self-aware of their state in the moment.
2. **Adaptive learning for mitigation strategies**: The app should be able to personalize the user's strategies by learning user preferences, behavior, and response patterns.

Data Collection Plan

Data Collection

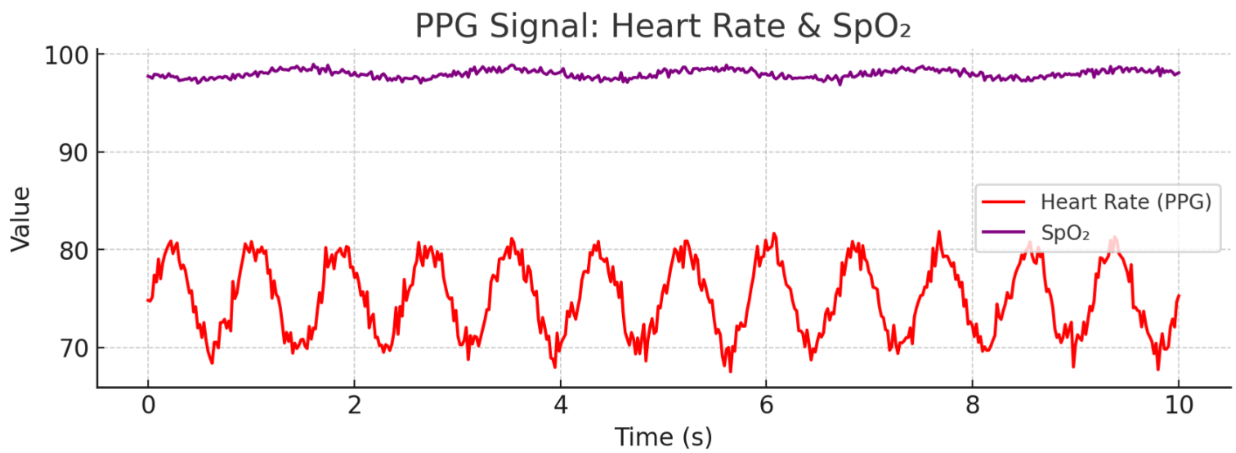
Input	Data Collected	Purpose
PPG Sensor	Heart rate Blood oxygen saturation	Detect elevated heart rate Detect reduced blood oxygen saturation
Pressure Sensor	Force distribution on muscles	Detect physical tension
Accelerometer	Linear acceleration	Measure tremors
Gyroscope	Angular velocity	Measure tremors
User Input	Stress levels Strategy feedback Context annotations	Support personalization and labeling

Kaylyn King
CSCE 5612

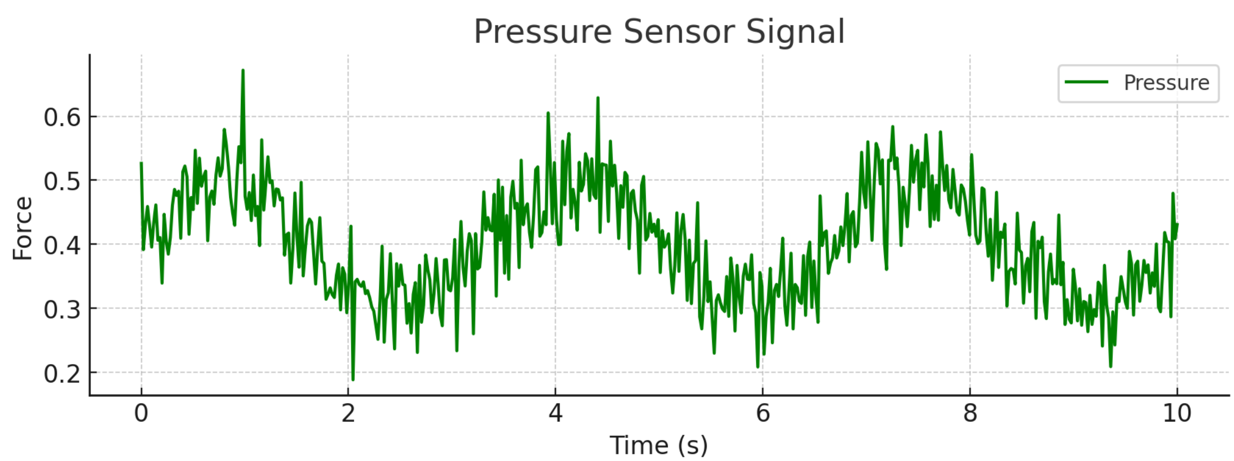
Data Visualization

The following are examples of signal traces that will be gathered from each sensor:

- **PPG**:

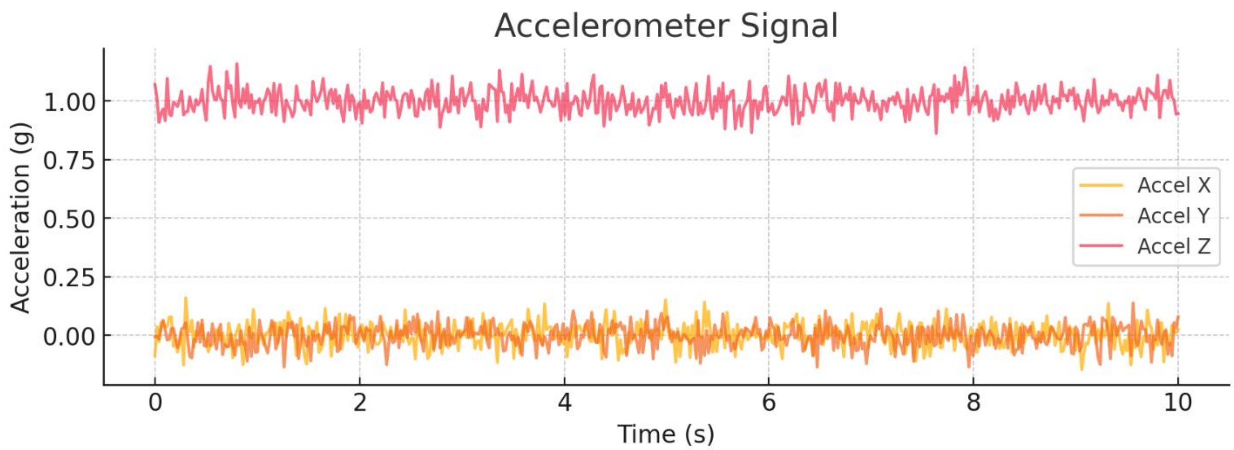


- **Pressure**:

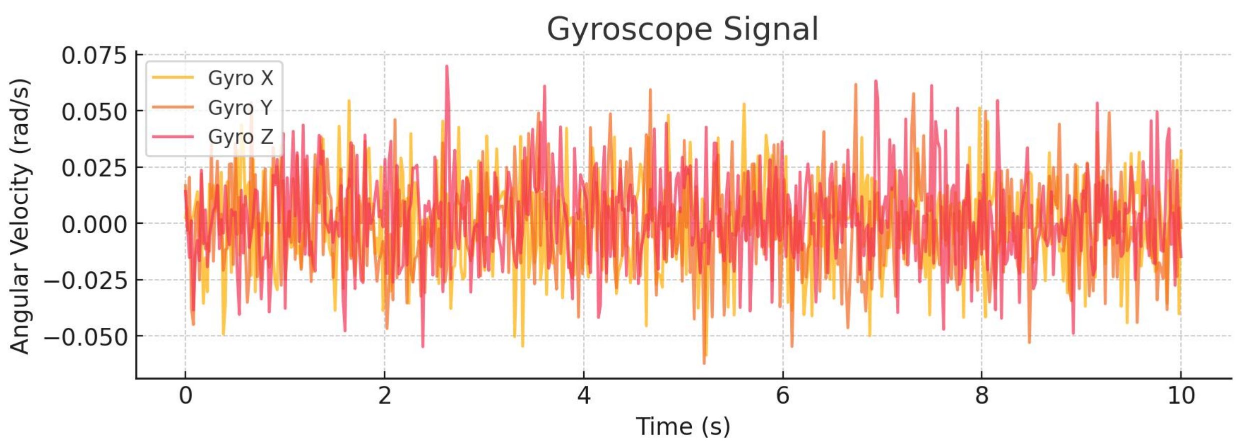


- **Accelerometer**:

Kaylyn King
CSCE 5612



- **Gyroscope**:



Hardware and Data Format

Component	Part	Format
PPG Sensor	MAXREFDES117#	Integer
Pressure Sensor	2Pcs RP-L Film Pressure Sensor Pressure Detector Flexible	Double
IMU	GY-521 MPU-6050 MPU6050	Double Array
Microcontroller	Seeed Studio XIAO ESP32C6	String

The following is an example of how the raw data will be formatted and stored in a JSON file for flexible logging and real-time storage:

{

//these are sensor inputs “X” means a number

“timestamp”: “XXXX-XX-XXTXX:XX:XX.XXXZ”

Kaylyn King
CSCE 5612

"heart_rate": XX,

"blood_oxy": XX,

"accel": [X.XX, X.XX, X.XX],

"gyro": [X.XX, X.XX, X.XX],

"pressure": X.XX,

//these are user inputs

"user_input": "mood"

"location": "place"

"activity": "description"

}

This data will then be preprocessed and transformed into a CSV file like this:

timestamp	mean_hr	hrv	mean_pressure	accel_mag	gyro_sd	mood	place	activity	label
2025-04-10T12:00:00	69.6	29.28	0.61	0.9	0.016	relaxed	office	browsing	calm
2025-04-10T12:01:00	89.9	27.97	0.75	0.22	0.028	neutral	school	presentation	calm
2025-04-10T12:02:00	97.2	13.76	0.27	0.08	0.021	neutral	public	browsing	anxious
2025-04-10T12:03:00	94.4	13.37	0.57	0.92	0.025	anxious	office	presentation	anxious
2025-04-10T12:04:00	98.1	27.39	0.33	0.15	0.013	tense	home	commuting	anxious
2025-04-10T12:05:00	78.1	15.64	0.43	0.54	0.012	relaxed	home	meditating	anxious
2025-04-10T12:06:00	79.8	25.12	0.44	0.1	0.018	tense	home	browsing	calm
2025-04-10T12:07:00	69.7	28.5	0.51	0.36	0.025	tense	public	commuting	anxious
2025-04-10T12:08:00	92.2	23.12	0.45	0.77	0.024	relaxed	public	commuting	calm
2025-04-10T12:09:00	90.7	12.77	0.53	1.31	0.018	anxious	home	presentation	calm
2025-04-10T12:10:00	74.3	19.23	0.4	1.36	0.013	neutral	school	browsing	anxious
2025-04-10T12:11:00	84.2	10.47	0.47	0.19	0.008	tense	home	presentation	anxious
2025-04-10T12:12:00	70.1	19.74	0.82	0.69	0.029	neutral	school	presentation	anxious
2025-04-10T12:13:00	74.6	24.12	0.62	1.49	0.012	anxious	home	presentation	anxious
2025-04-10T12:14:00	88.1	8.85	0.66	0.98	0.012	relaxed	school	browsing	calm

- **"Timestamp"**: ISO-formatted time when the window was captured

- **"Mean_hr"**: Average heart reate during the time window
- **"Hrv"**: Heart rate variability
- **"Mean_pressure"**: Average pressure value indicating muscle tension
- **"Accel_mag"**: Overall acceleration magnitude from the IMU
- **"Gyro_sd"**: Standard deviation of gyroscope values
- **"Mood"**: Self-reported emotional state
- **"Place"**: location of the user
- **"Activity"**: Task being performed
- **"Label"**: Annotated class ("anxious" or "calm")

Collection Protocols

Sampling Frequency:

- 50 Hz for PPG, Accelerometer, and Gyroscope
- 1 Hz for pressure

Sessions:

- 10-15 diverse participants
- Around 30 sessions per participant
- 10-20 minutes of diverse activities for each session
    - Neutral = sitting
    - Mild Stress = mental arithmetic or social interaction simulations
    - Relaxation = breathing or other mindfulness exercises
- If 10 participants and 30 15-minute sessions were chosen, then there would be about 75 hours of total data with about 15-20 hours of high-quality labeled data

Annotation or Labeling Methods

Self-Reported Labels

Participants will rate their perceived anxiety level using the State-Trait Anxiety Inventory-S psychological assessment tool before and after each session. This will give a baseline of the participant's state anxiety and see whether the tool was effective in reducing their anxiety.

Event Tagging

The users will also annotate events (i.e. onset of anxiety, strategy applied, perceived effectiveness) throughout the session to see if the data collected is correlated with the user's perception of events.

Kaylyn King
CSCE 5612

Edge Case Handling

Edge cases in sensor data will be filtered out to improve model reliability during preprocessing which includes:

- Sensor dropouts such as missing or corrupted readings due to poor contact or disconnection
- Periods of inactivity where no meaningful physiological data can be extracted like sleeping or being idle
- Noise from intense physical activity that may trigger anxiety-like signals like a high heart rate from exercise

Data Privacy

Privacy Measures:

- Participants will be assigned anonymous IDs
- Personal data will be encrypted and stored separately
- Informed consent will be obtained, and users can opt-out at any time
- Will comply with GDPR and IRB guidelines

Data Storage:

- Data will be stored in encrypted storage
- Local backups will be stored on encrypted external SSDs

Feature Extraction and Preprocessing

Signal Processing or Preprocessing Steps

Filtering:

- *PPG Sensor*: Bandpass filter of 0.5-4 Hz to isolate heart rate signal and moving average or "smoothing" for pulse waveform clarity
- *Pressure Sensor*: Moving average or "smoothing" to reduce jitter
- *Accelerometer and Gyroscope*: Low-pass filter with a cutoff of around 20 Hz to remove high-frequency noise

Segmentation: Data is segmented into sliding windows of 5-10 seconds with 50% overlap for time-series feature extraction, and each window is treated as a sample for training or inference.

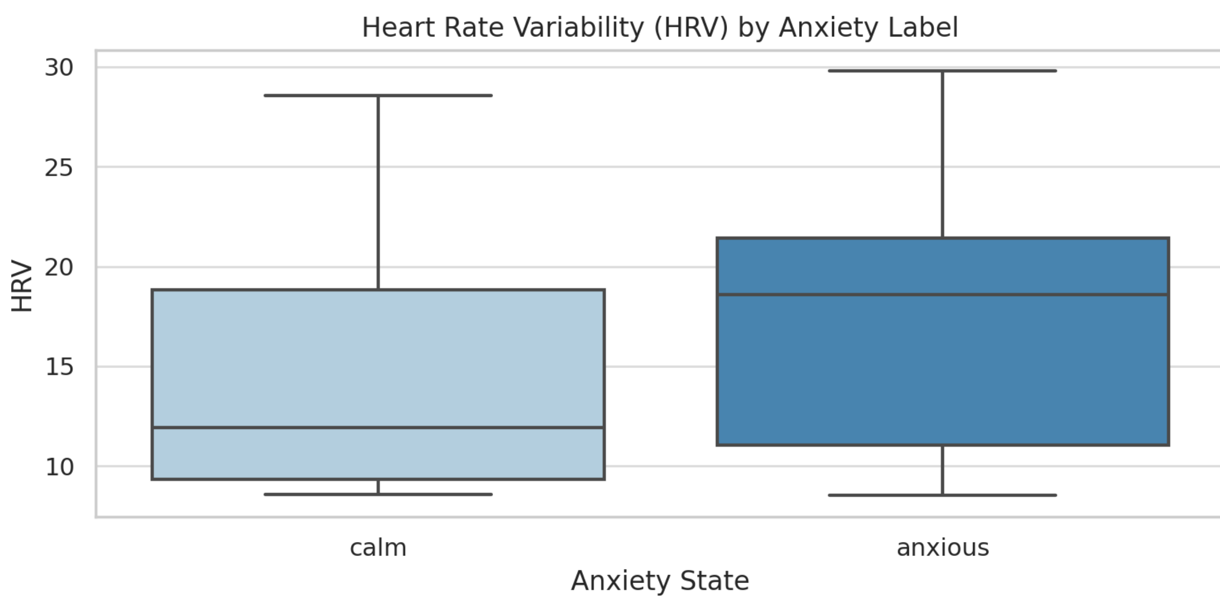
Kaylyn King
CSCE 5612

Normalization: Sensor readings are z-scored normalized with the mean = 0 and standard deviation = 1.

Features to be Extracted

PPG Sensor:

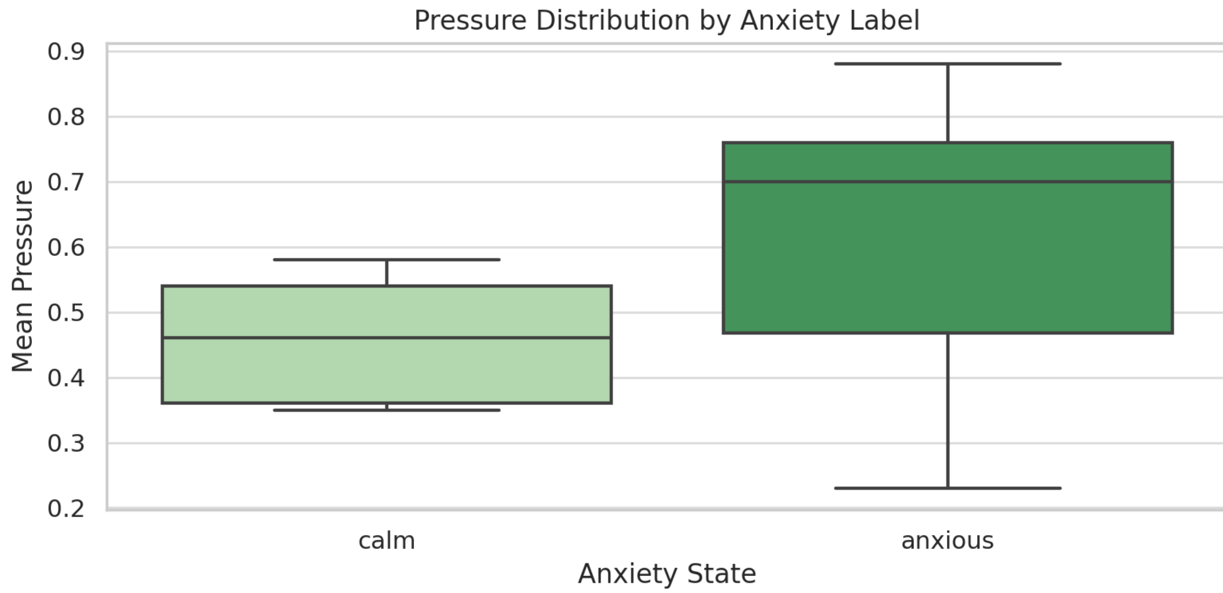
- Mean heart rate
- Heart rate variability
- Root mean square
- Signal entropy
- Peak intervals
- Pulse amplitude
- Example feature box plot:



Pressure Sensor:

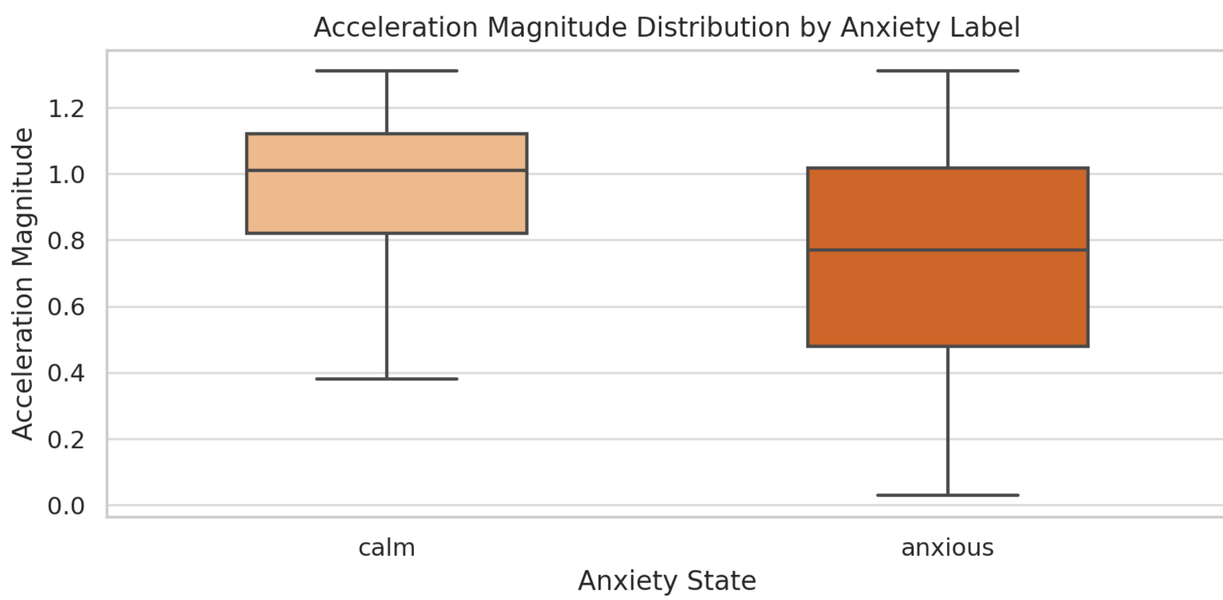
- Average pressure
- Pressure variance
- Peak force
- Example feature box plot:

Kaylyn King
CSCE 5612



Accelerometer:

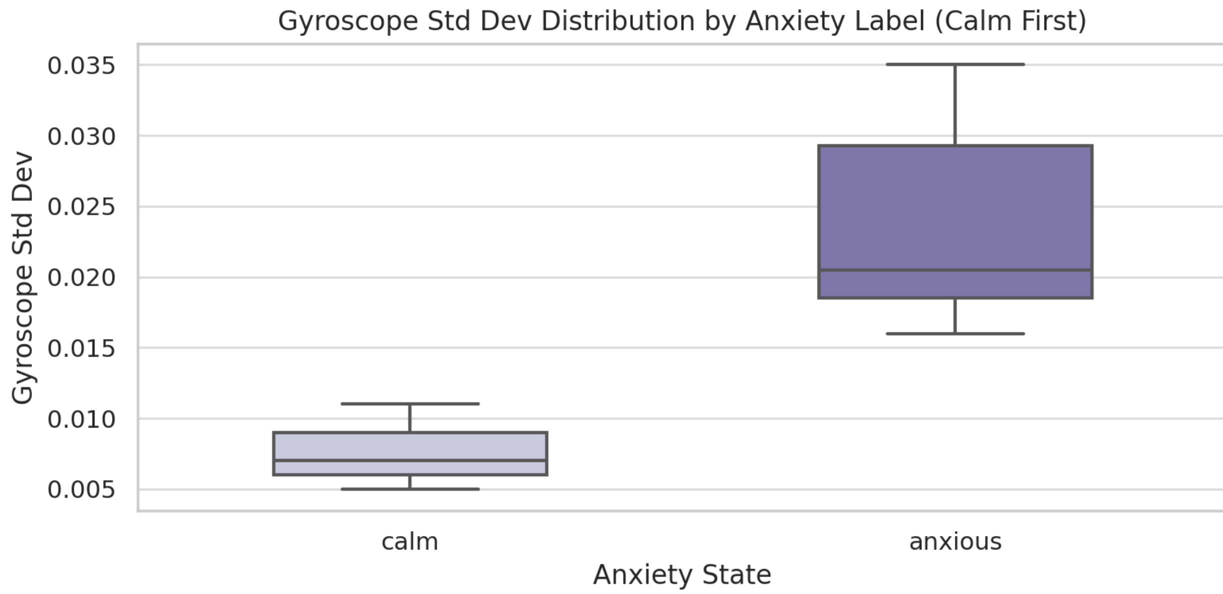
- Mean
- Standard deviation
- Signal magnitude area
- Jerk
- Energy
- Example feature box plot:



Gyroscope:

Kaylyn King
CSCE 5612

- Mean angular velocity
- Rotational energy
- Standard deviation
- Example feature box plot:



User Input: Categorical values encoded as embedded vectors

On-device vs. Offline Computation

Task	**Location**	**Reason**
Signal filtering	On-device	Lightweight filters can run on the microcontroller
Window segmentation	On-device	Enables real-time anxiety detection
Feature extraction	On-device	Efficient features like mean, RMS, and variance
FFT and complex features	Offline	Required more processing power than low-power MCU
Model inference	On-device	For real-time feedback
Strategy recommendation	Offline	Heavier reinforcement learning logic runs on paired app

Kaylyn King
CSCE 5612

Model Selection and Justification

Machine Learning Algorithms

Anxiety Detection Model

Hybrid deep learning model with long short-term memory and convolutional neural network to classify anxiety states based on time-series biometric data

- *Long Short-Term Memory (LSTM)*: A type of recurrent neural network used to learn temporal patterns in data, making it good for sensor readings over time
- *Convolutional Neural Network (CNN)*: Detects local patterns in short time windows and automatically extract powerful features like frequency patterns, signal shapes, or pulse waveforms, making it good for 1D signals on biometric sensors

Strategy Recommendation Model

Deep Q-Network for a reinforcement learning algorithm to recommend a personalized mitigation strategy based on the user's current state and past outcomes

Justification

Anxiety Detection Model

Accuracy: Combines temporal patterns with LSTM with local feature detection with CNN to handle time-series data effectively

Inference Speed: Inference can be done in real-time on edge devices with moderate CPU with windowed input of 5-10 seconds

Embedded Compatibility: Supported via TensorFlow Lite on the XIAO ESP32C6

Model Size: Can be kept under 200 KB – 1 MB using quantization after compression

Memory Use: Optimized LSTM layers and 1D CNNs are efficient for time-series sensor data

Deployment Tools: TensorFlow Lite Micro, CMSIS-NN, or TinyML-compatible converters

Strategy Recommendation Model

Accuracy: Learns optimal strategies over time based on real-world feedback

Inference Speed: Low-latency decision-making once the policy is trained

Embedded Compatibility: Training occurs offline, and only the trained policy lightweight neural network is deployed to the app

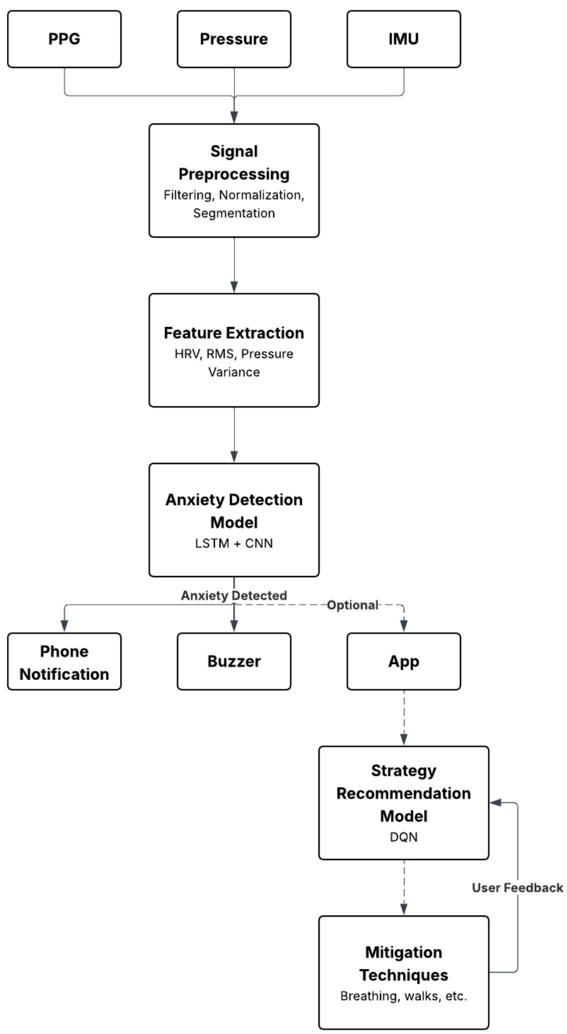
Kaylyn King
CSCE 5612

Model Size: Final policy network can be small with around 2-3 dense layers or under 100 KB

Memory Use: Minimal during inferences, and no on-device training required

Deployment Tools: TensorFlow Lite or ONNX

ML Pipeline Diagram



Model Compression or Quantization

Post-Training Quantization: Reduce weights from float32 to int8 or float16 to reduce model size by roughly 4 times.

Pruning: Remove low-importance neurons and connections to reduce inference load and improve memory usage

Model Distillation: Use a large model to train a smaller model with similar output behavior

Kaylyn King
CSCE 5612

Evaluation Plan

Evaluation of Model's Performance

Metrics

The Anxiety Detection Model with LSTM and CNN will be for binary classification, and the DQN Recommendation Model will be evaluated using a reward-based metric with a reduction in anxiety level being a positive feedback score.

Both models will use these metrics:

- **Accuracy**: Overall correctness with how often the model predicts the right anxiety state
- **Precision**: how many anxious states were correct out of all predicted anxious states
- **Recall**: how many anxious states did the model identify correctly out of all actual anxious states
- **F1-Score**: Harmonic mean of precision and recall to balance false positives and negatives
- **Area under the ROC Curve-Receiver Operating Characteristic (AUC-ROC)**: Measures how well the model distinguishes between classes at different thresholds which is useful for imbalanced data
    - *Area under the ROC Curve*: One-number summary of overall model quality
    - *Receiver Operating Characteristic*: Curve showing trade-off between true positive rate and false positive rate to help visualize model behavior at all thresholds
- **Reward Function Design for DQN**: Reinforcement learning model will be trained with a reward function that reflects the effectiveness of recommended mitigation strategies, allowing the strategy recommendation system to adapt over time.
    - *Primary method*: Reward = (reduction in post-strategy STAI-S score) + (positive user feedback score)
    - *Lightweight alternative for early stages*: Reward = (Drop in heart rate) + (user reports feeling better)

Methodology

To ensure temporal separation, the data set will be split into three sets:

- Training Set: 70%
- Validation Set: 15%
- Test Set: 15%

If there is enough time, a 5-fold cross-validation will be applied for robustness on smaller datasets. These cross-validation folds will ensure that data from the same user does not appear in both training and test sets to prevent overfitting to individual physiology.

Model Testing in Real-World Scenarios

On-Hardware Testing with XIAO ESP32C6:

- Live streaming sensor data
- On-device inference using the quantized TensorFlow Lite model
- Metrics recorded such as inference time, memory usage, battery drain, and latency

User Trials:

- Sessions will simulate real environments such as home, school, and work
- Data gathered will evaluate how accurately the model detects transitions in anxiety and test how effectively the system recommends strategies with a reduction in anxiety.

Iterate or Improve the Model

Model Tuning and Retraining:

- Tune model hyperparameters such as LSTM units, CNN filters, dropout, and learning rate
- Engineer new features such as moving heart rate variability
- User data augmentation to increase sample diversity

User Feedback Loop:

- Collect feedback from users after each session
- Use responses to adjust the DQN reward function and improve policy learning
- Maybe implement personalized thresholds or models for different users based on their baseline anxiety profiles

Deployment Feedback Monitoring:

- Log real-world predictions, confidence levels, and strategy outcomes
- Analyze false positives/negatives for retraining and improving generalization