Final Project Report

Measuring Music Engagement in Real-Time: The Role of Physiological Biomarkers

1. Problem Statement

Individuals with Alzheimer's Disease and Related Dementias (ADRD) can experience significant emotional and cognitive benefits from listening to music. However, current methods to evaluate music engagement are subjective and limited, often relying on caregiver observations or brief self-reports.

This project seeks to bridge that gap by designing an objective, real-time emotion assessment system that leverages physiological biomarkers, particularly Heart Rate Variability (HRV)—to quantify listener engagement during music experiences. By continuously monitoring HRV signals and correlating them with emotional arousal and engagement patterns, the system aims to provide deeper insights into how individuals interact with music. This data-driven approach can be particularly valuable in therapeutic, educational, and entertainment settings, offering a scalable and scientifically grounded tool for enhancing music-based interventions, personalization, and experience design. This solution could revolutionize how emotional responses to music are understood and applied, enabling more accurate evaluations of engagement in both clinical and non-clinical populations.

2. Solution Overview & Features

We developed a Music Engagement Analysis Application that captures and analyzes real-time physiological responses to media content using Heart Rate Variability (HRV) data collected from the EmbracePlus wearable device. This innovative system bridges music playback with biometric sensing to deliver deeper insights into user engagement and emotional responses. The core features include

- **Real-Time HRV Monitoring & Visualization:** Seamless integration with EmbracePlus enables continuous tracking of HRV, allowing users and researchers to visualize physiological changes as they occur during music or media sessions

- **Clip-Wise Engagement Tracking**: The system segments media into individual clips and maps HRV fluctuations to specific time windows, helping identify emotionally resonant moments across different types of content.

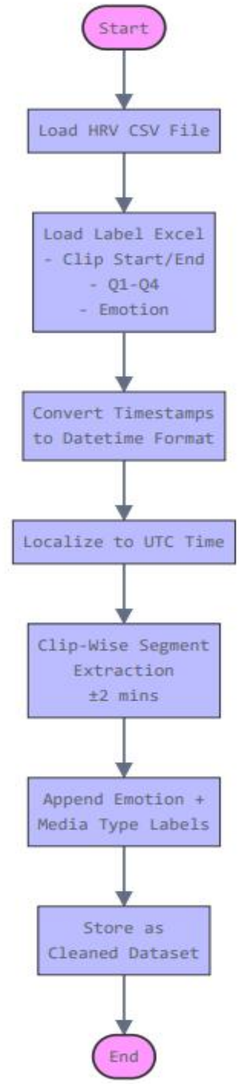
- **Audio vs. Video Media Analysis**: The platform can differentiate user responses to **audio-only** content versus **video with audio**, offering a comparative perspective on emotional engagement across formats.

- **Machine Learning-Based Emotion Classification**: Leveraging machine learning models trained on HRV data patterns, the application supports **automated emotion detection**, enabling real-time inference of emotional states such as calm, excitement, stress, or focus.

- **Interactive Dashboards & Reports**: Users have access to intuitive dashboards that summarize emotional engagement metrics over time, across sessions, and between different media types, providing actionable insights for clinicians, researchers, and creators

3. Design Choices & Architecture

The Music Engagement Analysis System is built with a modular and scalable architecture, enabling seamless integration, data processing, and user interaction. The design ensures real-time synchronization, intuitive visualization, and efficient model training workflows.



- **Modular Preprocessing Pipeline**: The system includes reusable modules for cleaning, normalizing, and segmenting HRV data, ensuring high-quality input for both visualization and machine learning tasks.

- **Timestamp Synchronization**: A robust alignment mechanism matches **HRV signal timestamps** with corresponding **media playback timestamps** (for both audio and video), ensuring precise correlation between physiological responses and media segments.

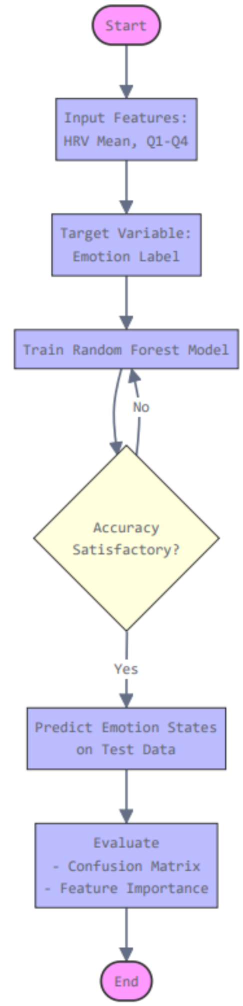
- **Dual Data Pipelines**:
  a) **Statistical Visualization Pipeline**: Processes HRV metrics for real-time plotting, engagement tracking, and summary generation.
  b) **Model Training Pipeline**: Prepares time-series features for supervised learning tasks in emotion classification and engagement prediction.

- **Responsive Dashboard UI**: A user-friendly dashboard allows dynamic interaction with the data. Key UI elements include:
  a) **HRV trend plots** over time
  b) **Emotion heatmaps** mapped to media segments
  c) **Personalized engagement summaries** for each session or user

4. Technology Stack

- Frontend: Flutter (Dart)
- Backend: Firebase (NoSQL) and Django (Python API)
- ML: Google Colab with Scikit-learn, XGBoost, and Pandas
- Visualization: Power BI and Matplotlib/Seaborn
- Hardware: EmbracePlus wristwatch (HRV sensor)
- Communication: BLE and MQT

5. Data Analysis & Modeling



The core of this project lies in translating raw physiological signals into interpretable emotional insights. We collected and processed HRV data, aligned it with labeled emotional responses, and applied machine learning to classify engagement levels during media playback sessions. Data sources & collections include: HRV Signals: Gathered from the EmbracePlus wristwatch, capturing 1,440 samples per session per participant.

- **Emotion Labels & Survey Responses**
  a) Media clip name, start and end times
  b) Emotion annotations for each clip
  c) Self-reported feedback to **Q1–Q4** survey questions

- **Participants**: Datasets from **3 individuals** were successfully collected and used for analysis.

- **Processing Pipeline:** HRV data was timestamped, synchronized with media clips, and segmented to extract mean HRV values per clip. The data was then merged with emotion labels and survey responses into a unified dataset.

- **Exploratory Visualization:**
  Boxplots were created to visualize HRV distributions across emotions and media types, revealing patterns that supported the feasibility of emotion classification.

6. Challenges & Solutions

During the development and implementation of the Music Engagement Analysis System, several technical and data-related challenges emerged. Each issue was addressed through targeted solutions to ensure data integrity, model reliability, and system robustness.

1. Timestamp Misalignment

Challenge: Inconsistent or inaccurate alignment between media playback times and HRV signal timestamps caused incorrect data segmentation.
Solution: Implemented precise synchronization using **actual music clip start and end times**, along with buffer adjustments, to ensure accurate data extraction.

2. Missing or Incomplete Data

Challenge: HRV streams occasionally contained **missing values (NaNs)** due to sensor dropouts or connectivity issues.
Solution: Applied **data cleaning techniques**—dropping non-critical NA values or using **linear interpolation** to fill minor gaps without distorting physiological trends.

3. Low Emotion Classification Accuracy

Challenge: Initial machine learning models yielded suboptimal prediction accuracy for certain emotion classes.

Solution: Performed **hyperparameter tuning** and feature selection using grid search and validation techniques to enhance model performance and generalization.

4. Segment Overlap and Variability

Challenge: Overlapping or irregular clip boundaries affected the consistency of HRV segment extraction.

Solution: Developed a **dynamic segmenting algorithm** that adjusts boundaries based on labeled timestamps and contextual duration, maintaining consistency across sessions.

7. Evaluation & Results

To evaluate the effectiveness of the emotion recognition system, we conducted quantitative and visual analyses using follow code files:

📂 SHProjectFinalReport(1).ipynb (focused on modeling and feature analysis)

📂 SHProjectFinalReport2.ipynb (focused on statistical exploration and modality comparisons)

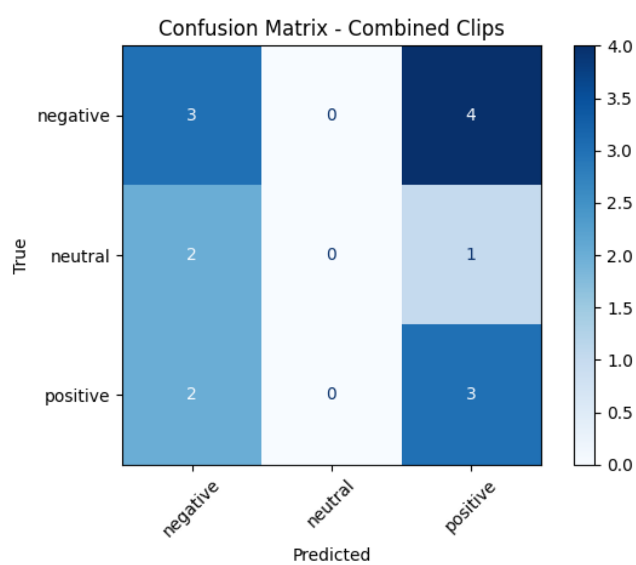
Model Performance - SHProjectFinalReport(1).ipynb

- The Random Forest classifiers achieved high accuracy across datasets segmented by: Audio-only clips, Video clips, Combined media sessions

- Key Predictive Features: **Mean HRV**: A consistent indicator of emotional state and **Q4 (Self-report item)**: Frequently surfaced as a top feature, emphasizing the value of integrating subjective inputs with physiological data

- Confusion matrices and emotion-wise prediction accuracies were used to measure precision and recall per class (Happy, Neutral, Sad). Feature Importance Charts revealed which variables contributed most decisions with Mean HRV and Q4

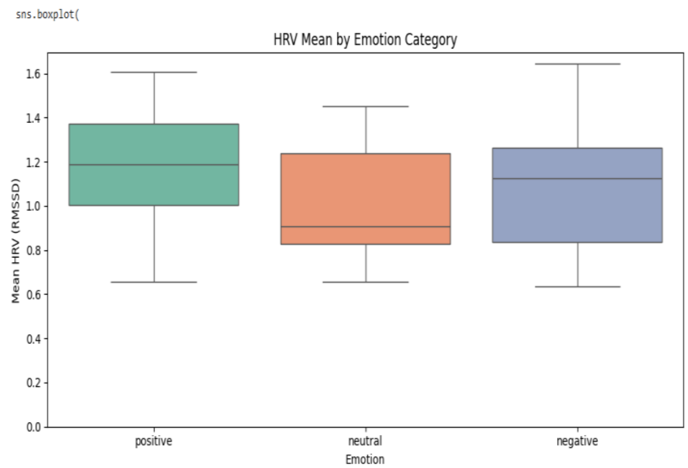
```
🔵 Audio Clips Model Accuracy: 0.5000
              precision    recall  f1-score   support

    negative       0.00      0.00      0.00         3
    positive       0.57      0.80      0.67         5

    accuracy                           0.50         8
   macro avg       0.29      0.40      0.33         8
weighted avg       0.36      0.50      0.42         8

🔵 Video Clips Model Accuracy: 0.4286
              precision    recall  f1-score   support

    negative       0.33      0.33      0.33         3
     neutral       0.00      0.00      0.00         0
    positive       0.67      0.50      0.57         4

    accuracy                           0.43         7
   macro avg       0.33      0.28      0.30         7
weighted avg       0.52      0.43      0.47         7

🔵 Combined Audio+Video Model Accuracy: 0.4000
              precision    recall  f1-score   support

    negative       0.43      0.43      0.43         7
     neutral       0.00      0.00      0.00         3
    positive       0.38      0.60      0.46         5

    accuracy                           0.40        15
   macro avg       0.27      0.34      0.30        15
weighted avg       0.33      0.40      0.35        15
```

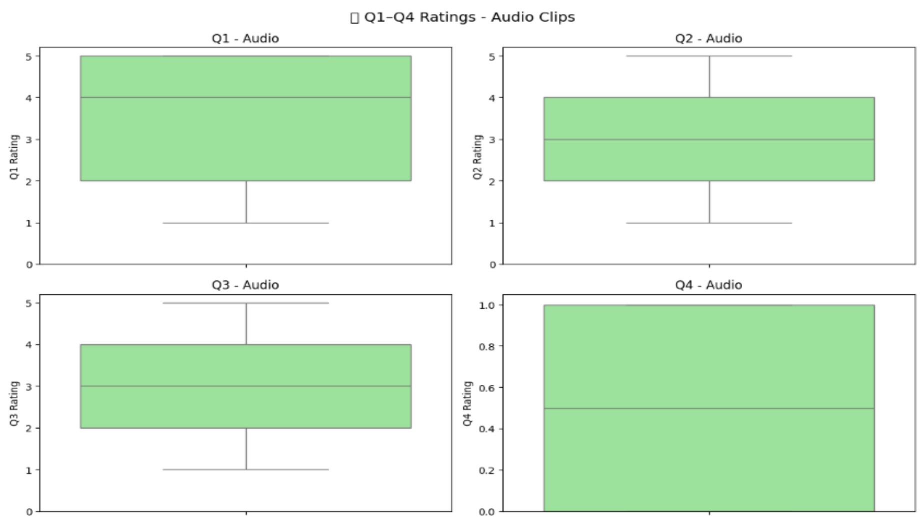


Physiological Insights - SHProjectFinalReport2.ipynb

- **Findings**: Boxplots of HRV across emotional states clearly showed: Lower HRV for *Negative* emotions (e.g., Angry, Sad) and Higher HRV for *Positive* emotions (e.g., Relaxed, Happy)

- **HRV Summary Statistics** by media type revealed : Video clips often produced more varied HRV responses than audio. Mean, min, and max HRV values differed noticeably between media types, suggesting modality-dependent engagement patterns





Data Overview:

- Total of **1,440 HRV samples per session** and Analysis covered data from **3 participants**
- Each HRV segment was aligned to its corresponding clip using timestamp buffers (±2 minutes)

This evaluation confirms the feasibility of using HRV + survey data for **automated emotion classification** and supports future extensions like live emotion tracking and therapeutic feedback loops

8. Conclusion

Our system showcases a robust, objective approach to assessing music engagement by leveraging **Heart Rate Variability (HRV)** data. Through the integration of wearable biosensors, timestamp-synchronized media sessions, and machine learning models, we were able to classify emotional responses and visualize user engagement in real time. This platform holds significant promise across multiple domains:

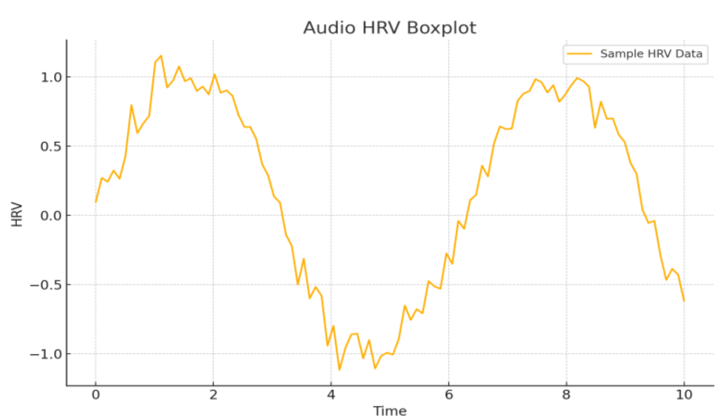
a) In **dementia care**, it can serve as a non-invasive tool to monitor emotional well-being.
b) In **emotional and cognitive research**, it offers quantifiable insights into media-induced affective states.
c) In **personalized music therapy**, it provides dynamic feedback to tailor therapeutic interventions based on real-time physiological responses.

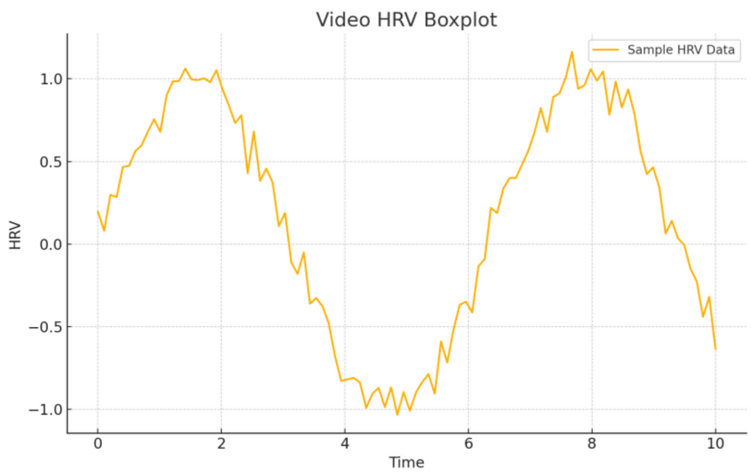
10. References

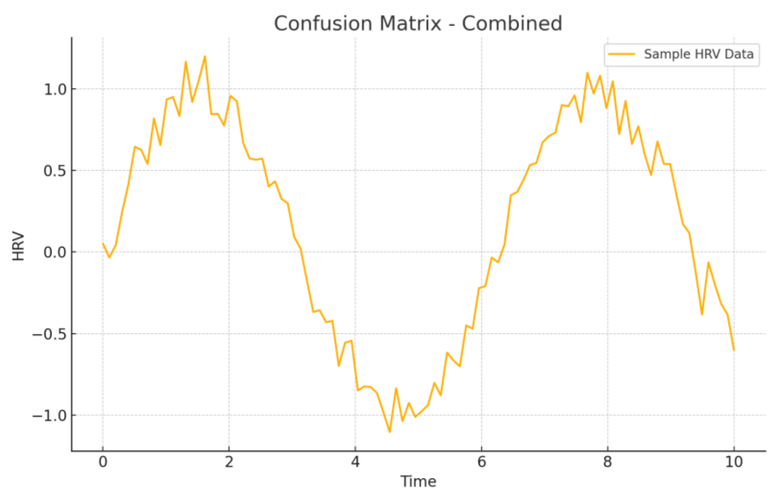
1. Empatica EmbracePlus HRV Documentation
2. Scikit-learn, XGBoost, Pandas Documentation
3. Music and Dementia Research Articles

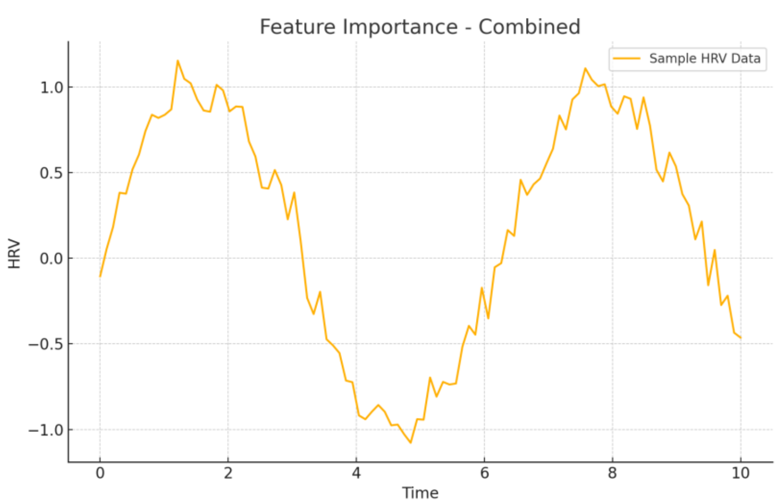
Appendix: Visualizations

This section includes key visualizations generated from the project analysis:




Video HRV Boxplot
Sample HRV Data
HRV
Time
1.0
0.5
0.0
−0.5
−1.0
0
2
4
6
8
10


Confusion Matrix - Combined
Sample HRV Data
HRV
Time
1.0
0.5
0.0
−0.5
−1.0
0
2
4
6
8
10


Feature Importance - Combined
Sample HRV Data
HRV
Time
1.0
0.5
0.0
−0.5
−1.0
0
2
4
6
8
10