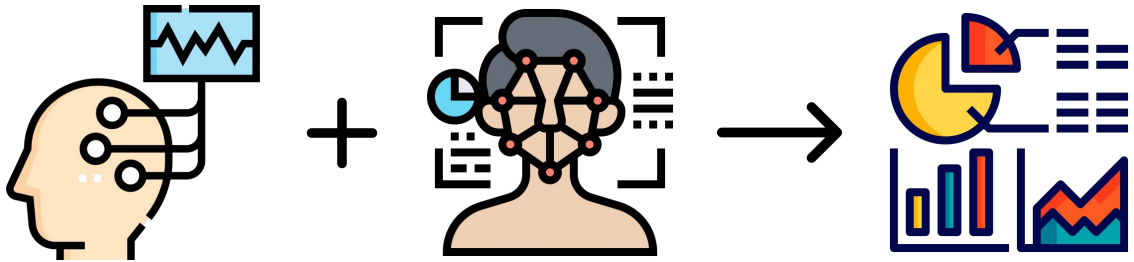
Carnegie Mellon University

Team E0: Focus Tracker App

Arnav Arora, Karen Li, Rohan Sonecha

Solution Approach

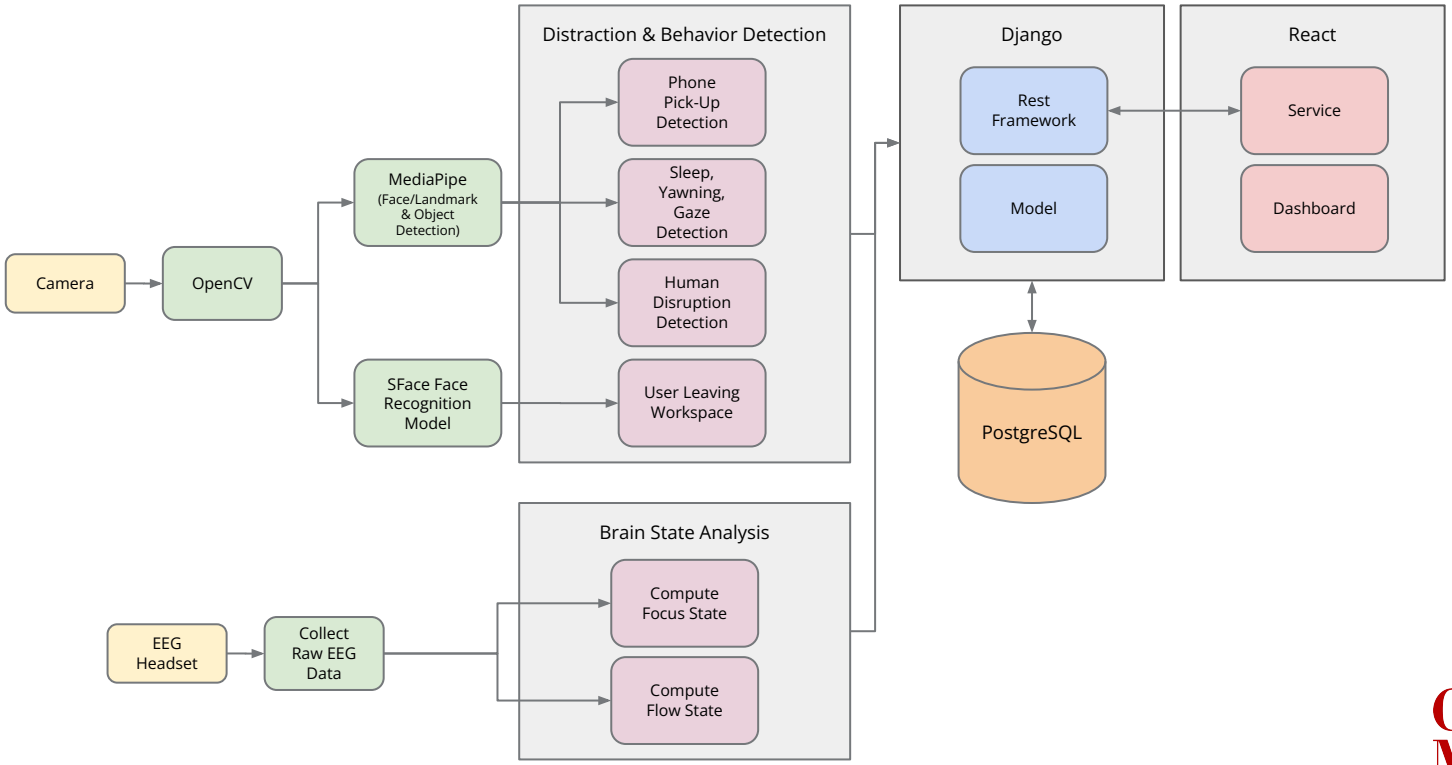
- **Real-time monitoring** of flow and focus states, distracted behaviors, and environmental distractions
- Measure flow and focus states using **EEG headset**, and distractions using **camera**
- Identify **distracted behaviors and environmental distractions**
    - Yawning, sleeping, off-screen gazing, phone pick-ups, disruptions from others, user leaving the workspace
- Dashboard to visualize focus level and distractions over **historical work sessions**
- Summarize **top distractions and behaviors** for a given work session



Carnegie Mellon University

https://www.flaticon.com

Solution Approach



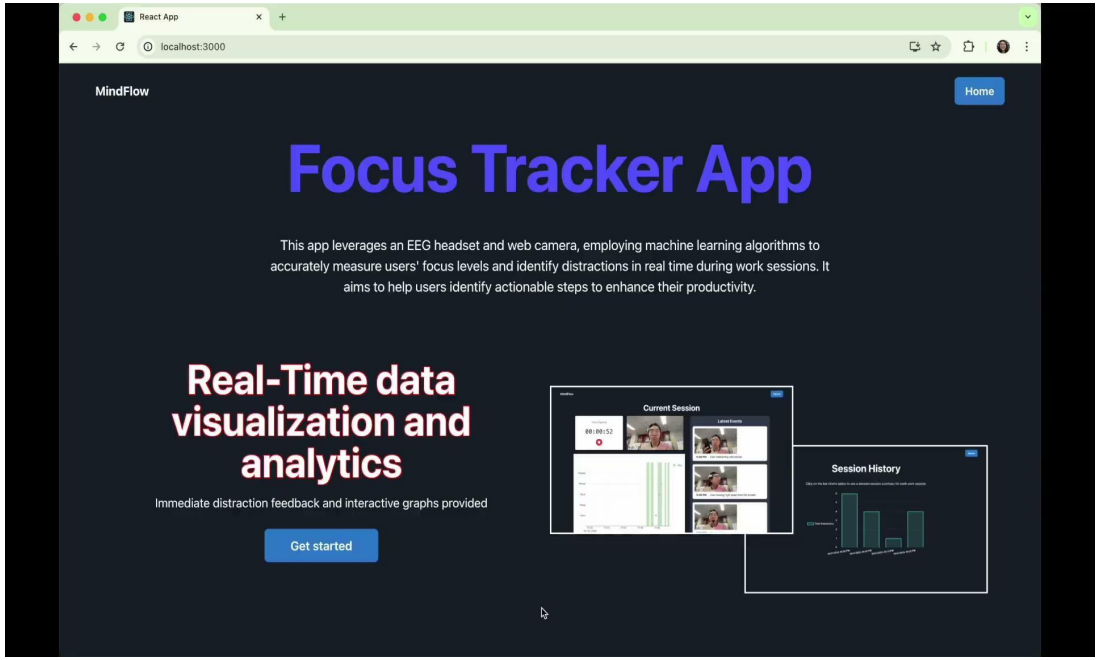
Carnegie Mellon University

Use-Case and Design Requirements

Flow State Accuracy	F-score ≥ 0.7 (across multiple test sets) Recall ≥ 0.9 (across multiple test sets) Prominent features in model match existing EEG flow state research
Focus State Accuracy	F-score ≥ 0.7 (across multiple test sets) Recall ≥ 0.9 (across multiple test sets) Prominent features in model match existing EEG focus state research
Usability and Usefulness	≥90% of users find the user experience to be seamless and easy to use
Distraction and Distracted Behavior Detection	F-score ≥ 0.7 Recall ≥ 0.9
Real-time Monitoring	≤ 3s delay between data capture and analysis (some latency is acceptable)

Carnegie Mellon University

Complete Solution



[AI meets opera: A new blended class at CMU yields insights on music and flow | Pittsburgh Post-Gazette]

Carnegie Mellon University

Quantitative Tests

	Yawning	Sleeping	Off-Screen Gaze	Phone Pick-Up	Other People	User Away
Recall	0.96	0.98	0.92	0.90	0.92	1.00
F-Score	0.94	0.97	0.96	0.95	0.96	0.84
Met	✅	✅	✅	✅	✅	✅

- Test among 5 different users
- Engage in each behavior 10 times over 10-minute session
- Record number of true positives, false positives, and false negatives

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **Target**
  - **Recall ≥ 0.9**
  - **F-score ≥ 0.7**

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Carnegie Mellon University

Quantitative Tests (cont.)

Metric	Test	Target	Actual	Met
YOLOv8 Phone Object Detection	Evaluate model on test set (15 different phones)	Recall ≥ 0.9 F-score ≥ 0.7	Recall = 0.9 F-score = 0.93	✅
Data Capture & Analysis Latency	Measure video processing speed Measure model evaluation time	≤ 3s delay between data capture and analysis	0.1s to process each video frame focus state model: 4 µs flow state model: 7 µs	✅
Usability and Usefulness	Survey users on usability and usefulness	≥90% of users find app usable and useful	90% of users find app usable and useful	✅

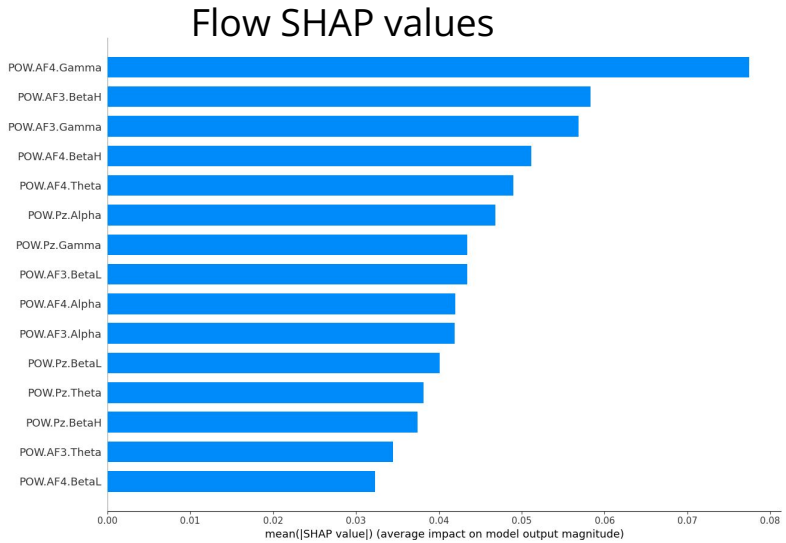
Carnegie Mellon University

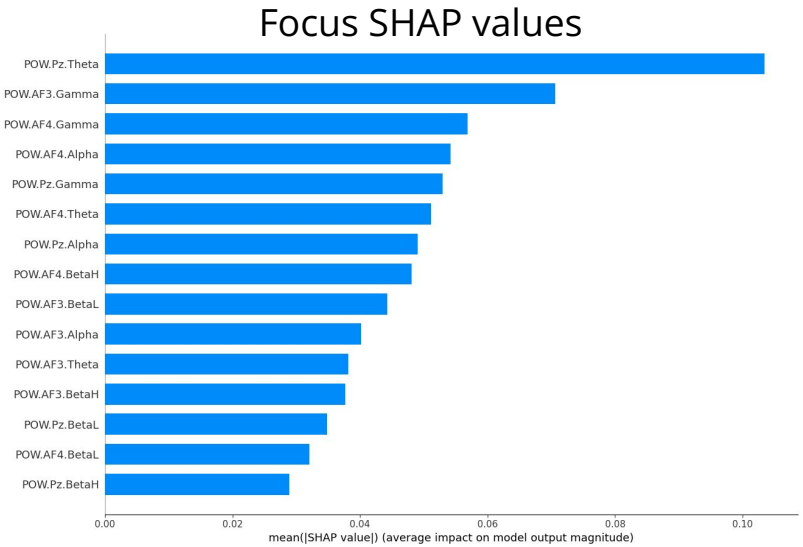
Quantitative Tests (cont.)

Metric	Test	Target	Actual	Met
Flow State	Test data from same recordings as train/validation (music setting) Test data from new recording (music setting, subject not included in training)	Recall ≥ 0.7 F1 Score ≥ 0.9	Recall: 0.9205 F1 Score: 0.9151 Recall: 0.6415 F1 Score: 0.6896	✅ ❌
Focus State	Test data from same recordings as train/validation (work setting) Test data from new recording (work setting, subject included in training)	Recall ≥ 0.7 F1 Score ≥ 0.9	Recall: 0.9141 F1 Score: 0.9199 Recall: 0.5570 F1 Score: 0.6004	✅ ❌

Carnegie Mellon University

Brain State Model Explainability





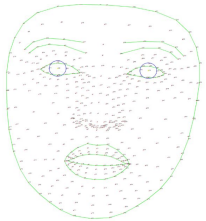
Carnegie Mellon University

Trade-Offs (Flow State Classification)

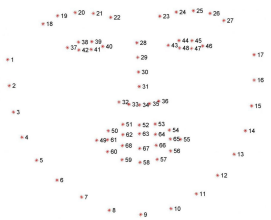
	Thresholding Approach	SVM Classifier	Neural Network Classifier
Pros	Simple to implement and understand Backed by existing research	Linear classifier is easy to implement, visualize, and understand Backed by existing research	Can pick up on non-linear relationships between features and flow vs not in flow brain state Explainable via SHAP values
Cons	The average/stdev power values for flow vs not in flow were not distinct enough	Our input features for flow vs not in flow were not linearly distinguishable	Prone to overfitting

Trade-Offs (Distraction Detection)

Binary Image Classification	DLib Face Detection and Landmarking	MediaPipe Face Detection and Landmarking
• Requires large amounts of data collection and labelling • No need to understand facial features/key facial landmarks	• Requires calibration/thresholding • Requires understanding of key facial landmarks • 60 fps • 68 2D face landmarks	• Requires calibration/thresholding • Requires understanding of key facial landmarks • Hand landmarking for phone pick-ups • 120 fps • 468 3D face landmarks



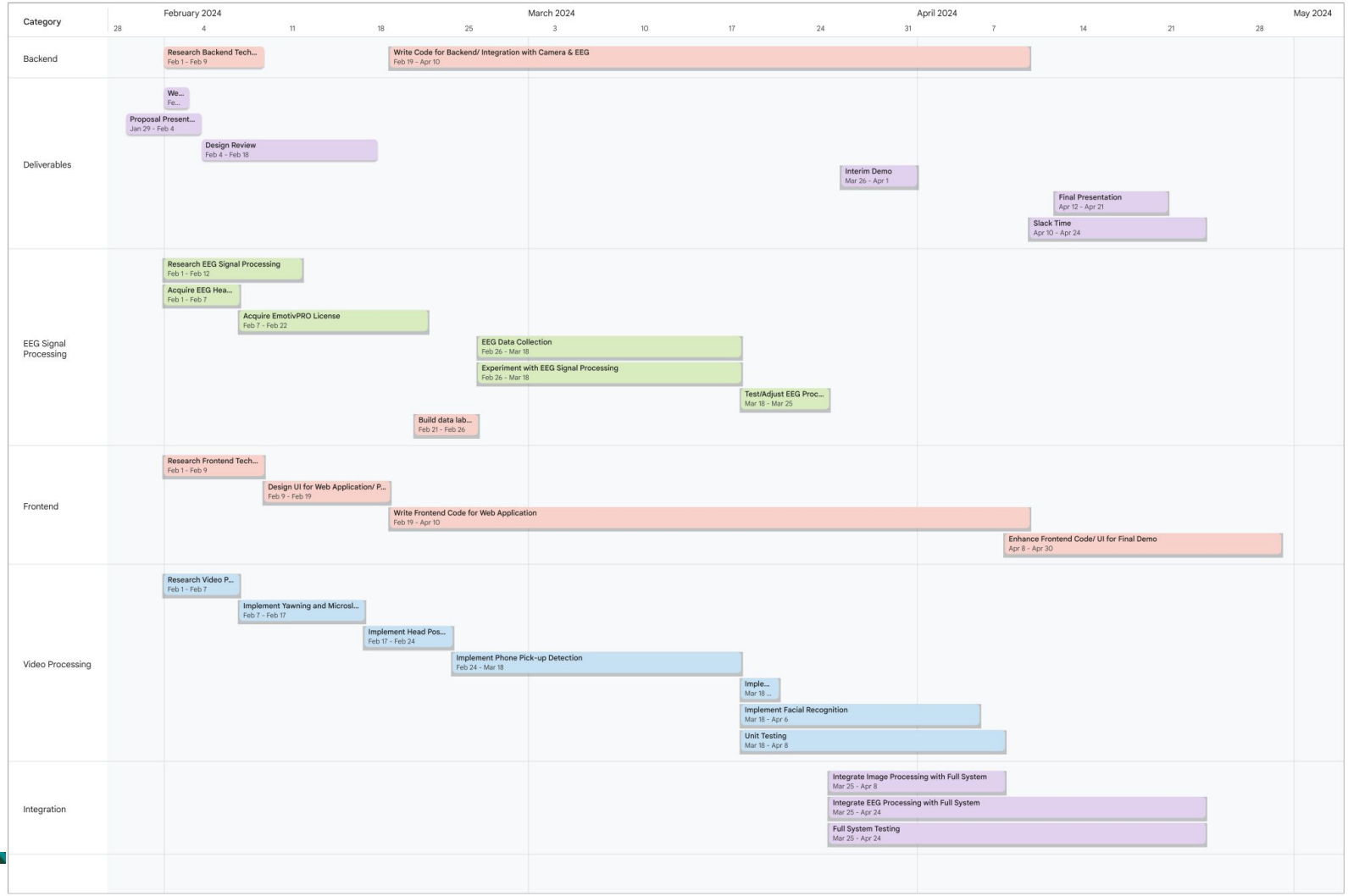
MediaPipe



Dlib

Carnegie Mellon University

Schedule


Category
February 2024
March 2024
April 2024
May 2024
28
4
11
18
25
3
10
17
24
31
7
14
21
28
Backend
Research Backend Tech...
Feb 1 - Feb 9
Write Code for Backend/ Integration with Camera & EEG
Feb 19 - Apr 10
Deliverables
We...
Fe...
Proposal Present...
Jan 29 - Feb 4
Design Review
Feb 4 - Feb 18
Interim Demo
Mar 26 - Apr 1
Final Presentation
Apr 12 - Apr 21
Slack Time
Apr 10 - Apr 24
EEG Signal Processing
Research EEG Signal Processing
Feb 1 - Feb 12
Acquire EEG Hea...
Feb 1 - Feb 7
Acquire EmotivPRO License
Feb 7 - Feb 22
EEG Data Collection
Feb 26 - Mar 18
Experiment with EEG Signal Processing
Feb 26 - Mar 18
Test/Adjust EEG Proc...
Mar 18 - Mar 25
Build data lab...
Feb 21 - Feb 26
Frontend
Research Frontend Tech...
Feb 1 - Feb 9
Design UI for Web Application/ P...
Feb 9 - Feb 19
Write Frontend Code for Web Application
Feb 19 - Apr 10
Enhance Frontend Code/ UI for Final Demo
Apr 8 - Apr 30
Video Processing
Research Video P...
Feb 1 - Feb 7
Implement Yawning and Microsl...
Feb 7 - Feb 17
Implement Head Pos...
Feb 17 - Feb 24
Implement Phone Pick-up Detection
Feb 24 - Mar 18
Imple...
Mar 18 ...
Implement Facial Recognition
Mar 18 - Apr 6
Unit Testing
Mar 18 - Apr 8
Integration
Integrate Image Processing with Full System
Mar 25 - Apr 8
Integrate EEG Processing with Full System
Mar 25 - Apr 24
Full System Testing
Mar 25 - Apr 24