**Project Title:** *Classifying Engagement States Using Wearable-Compatible Biometrics*

Maintaining attention is critical for cognitive performance and daily functioning, yet individuals frequently experience lapses in focus during tasks like studying, working, or driving. From a bioengineering perspective, being able to detect a user's engagement level using passive biometric signals opens the door to real-time, personalized feedback through wearable technology. A device that can recognize when a user is becoming disengaged could prompt them to take breaks,  
reduce distractions, or switch tasks. In this project, I aim to explore whether engagement states can be classified using a limited set of physiological and environmental features that are measurable with current or near-future consumer wearables. The long-term vision would be to integrate such an algorithm into a wearable tool that passively monitors focus and promotes sustained attention throughout the day.

This work falls under the second category of the project guidelines: implementing an exploratory analysis on a dataset that has not yet been analyzed in this specific way. The dataset I am using—[Kaggle’s Emotional Monitoring Dataset](https://www.kaggle.com/datasets/ziya07/emotional-monitoring-dataset/data)—contains biometric and environmental measurements paired with labels of user engagement level. While some past work has focused on predicting emotional states using this dataset, my project is unique in its focus on engagement state classification using only features that are wearable-compatible, excluding inputs like EEG or facial expression analysis.

My approach will involve first filtering the dataset to retain only wearable-compatible features such as heart rate, skin conductance, temperature, activity level, ambient noise, and lighting. I will clean and normalize the data, then train a logistic regression model to classify whether a user is “Highly Engaged” versus “Not Engaged.” This binary classification task reflects a practical use case for real-time wearable monitoring. I will also explore Principal Component Analysis (PCA) to reduce dimensionality and test its effect on model performance. For comparison, I will evaluate other classification models, including Random Forests, XGBoost, and a shallow neural network. Model performance will be assessed using accuracy, F1-score, and ROC-AUC. Although some of these techniques go beyond what we have covered in class, I am confident that my background in data science will enable me to implement and interpret them effectively.

This project is well-scoped to be completed within a one-month timeframe. The first week will focus on data preprocessing, exploratory analysis, and feature selection. In the second week, I will implement the baseline model using logistic regression. The third week will be dedicated to running PCA and testing additional models for performance comparison. In the final week, I will synthesize the results, draw conclusions, and prepare a final report and visualizations. The overall difficulty of the project is moderate. While the machine learning techniques are relatively standard, key challenges will include ensuring feature compatibility with wearable use cases, handling any class imbalance in the engagement labels, and interpreting physiological signals in a meaningful way.

Expected outcomes include a validated machine learning model capable of classifying a user’s engagement state based on wearable-compatible input features. Additionally, the analysis will offer insight into which physiological and environmental signals are most predictive of disengagement. The final result will serve as a conceptual framework for future wearable applications aimed at sustaining attention and optimizing user productivity. This is an individual project, and I will complete all analysis, implementation, and reporting independently.