Machine Learning recent Advancements and Innovation for the project

"ECG-Driven emotion detection in unconscious people"

1. Advanced Preprocessing Techniques for ECG Signals

To extract meaningful features from ECG signals in the DREAMER dataset, you need robust preprocessing to handle noise, artifacts, and variability. Here are state-of-the-art techniques as of 2025:

- **Adaptive Noise Cancellation with Wavelet Transform**:
  - **Why?** ECG signals often have noise (e.g., muscle artifacts, baseline wander). Wavelet transforms are great for denoising while preserving signal integrity.
  - **How?** Use a Discrete Wavelet Transform (DWT) with a Daubechies wavelet (e.g., db4) to decompose the signal into frequency bands, remove noise components, and reconstruct the clean signal.
  - **Tool**: Use pywt (PyWavelets) in Python.

- **Pan-Tompkins Algorithm for R-Peak Detection**:
  - **Why?** Accurate R-peak detection is critical for heart rate variability (HRV) features. The Pan-Tompkins algorithm is a standard, but recent enhancements use adaptive thresholding for robustness.
  - **How?** Apply bandpass filtering (5-15 Hz), differentiation, squaring, and moving window integration, then use adaptive thresholds to detect peaks.
  - **Tool**: Use neurokit2 for a modern implementation

- **Baseline Wander Removal with Polynomial Fitting**:
  - **Why?** Baseline wander can skew features like HRV. Polynomial fitting removes low-frequency trends without distorting the signal.
  - **How?** Fit a low-order polynomial (e.g., order 3) to the signal and subtract it.

- **Data Augmentation for ECG**:
  - **Why?** To improve model robustness, especially with limited DREAMER data (23 subjects, ~10 minutes each).
  - **How?** Apply techniques like time warping, noise injection, or synthetic ECG generation using GANs (e.g., ECG-GAN).

2. Advanced Feature Extraction from ECG Signals

To align with recent advancements, focus on extracting both traditional and novel ECG features for emotion detection:

- **Traditional HRV Features**:
  - **Time-Domain**: Mean RR interval, SDNN (standard deviation of NN intervals), RMSSD (root mean square of successive differences), pNN50.
  - **Frequency-Domain**: Power in low-frequency (LF, 0.04-0.15 Hz) and high-frequency (HF, 0.15-0.4 Hz) bands, LF/HF ratio.
  - **Tool**: Use hrv-analysis library
- **Nonlinear Features**:
  - **Why?** Emotions induce complex, nonlinear patterns in ECG. Recent studies emphasize nonlinear features for better emotion classification.
  - **Features**: Sample Entropy, Detrended Fluctuation Analysis (DFA), Poincaré plot metrics (SD1, SD2).
  - **Tool**: Use nolds for nonlinear analysis
- **Morphological Features**:
  - **Why?** ECG waveform shapes (P, QRS, T waves) carry emotion-related information.
  - **How?** Extract amplitudes and durations of P, QRS, T waves using segmentation techniques.
  - **Tool**: neurokit2 can segment ECG
- **Time-Frequency Features with CWT**:
  - **Why?** Continuous Wavelet Transform (CWT) captures time-varying frequency content, ideal for non-stationary ECG signals.
  - **How?** Use CWT to generate scalograms, which can be fed into deep learning models

3. Modern ML Architectures for ECG Emotion Detection

Your professor wants cutting-edge models, so let's explore recent architectures (up to 2025) that excel with time-series data like ECG:

- **1D Convolutional Neural Networks (CNNs) with Attention**:
  - **Why?** 1D CNNs are great for capturing local patterns in ECG signals, and attention mechanisms focus on emotionally salient features.
  - **Architecture**: Stack 1D convolutional layers, batch normalization, ReLU activations, and a self-attention layer.
- **Transformers for Time-Series**:
  - **Why?** Transformers, especially Time-Series Transformers (TST), are state-of-the-art for sequential data in 2025, capturing long-range dependencies in ECG signals.
  - **Architecture**: Use a transformer encoder with positional encoding and multi-head attention.
  - **Library**: Use tsai (PyTorch-based) for a pre-built Time-Series Transformer
- **Graph Neural Networks (GNNs)**:
  - **Why?** If you model ECG features as a graph (e.g., nodes as R-peaks, edges as temporal relationships), GNNs can capture complex interactions.
  - **How?** Convert ECG segments into a graph and use a GNN like GraphConv.
  - **Tool**: Use spektral or PyTorch Geometric
- **Hybrid CNN-LSTM with Residual Connections**:
  - **Why?** Combines CNNs for local feature extraction and LSTMs for temporal dependencies, with residual connections to prevent vanishing gradients.

4. Evaluation:

- Metrics: Accuracy, F1-score, AUC-ROC for imbalanced classes.
- Visualize feature importance with SHAP.