Optimizing Sleep Stage Classification

February 21, 2025

STAT 4830 - Numerical Optimization

GOALZ Project Group

Electroencephalography

- Electroencephalography(**EEG**) is a major tool for diagnosing sleep disorders and development of wearable sleep trackers.
- EEG measures **electrical activity** of the brain
- The **frequency** of the recorded waves reflects different mental and physiological states, ranging from wakefulness to deep sleep

PolySomnography

- PolySomnography(**PSG**), a comprehensive multi-modal diagnosis tool, is considered the "gold standard" for sleep study and sleep stage classification
- The physiological signals PSG records include:
  - EEG (brain activity)
  - EOG (eye movements)
  - EMG (muscle activity)
  - ECG (heart activity)
  - Respiratory signals (i.e oxygen saturation)

Problem Description

- PSG has a number of drawbacks:
  - Detection between Wakefulness and N1
  - Manual scoring
  - Generaliztion of patient information
  - Complex and Resource Intensive
  - Practicality

Project Objectives

- Develop an optimized classifier for sleep staging from **scalp EEG** that can produce **addititve** or **comprable** classification results as multi-modal PSG
- Improve **wake vs. N1 sleep classification**, particularly during **sleep onset transitions**.
- Address **limitations of existing classifiers**, such as:
  - Over-reliance on **in-bed intervals**.
  - **sensitivity** to sleep and wake epochs.
  - Resolve **Generalization issues** across different subjects and time-of-day variations.

Dataset formats

- **Data Source:** EEG recordings stored in **EDF format**.
- **Sleep Stage Labels:** Provided in **text files**, with **one label per 30-second epoch**.
- **Preprocessing:**
  - **Bandpass filtering** (0.5–40 Hz), **Epoch segmentation** (30s windows), **Normalization per subject**
- **Feature Extraction:**
  - **Power Spectral Density (PSD)** features for **delta, theta, alpha, beta** bands.
  - **Catch22 time-series** capturing autocorrelation, entropy, and complexity.

Current Datasets

- **ANPHY** : Small dataset of 29 healthy adults
  - High density EEG
  - PSG overnight data
- **MESA** : large dataset of patients with cardiovascular disease
  - Overnight PSG
  - Longitudinal data (2000-present)
  - Actigraphy

Methodology

Mathematical Formulation of the Optimization Problem

Given an EEG time series Y with associated covariates Z, we seek to predict the sleep stage S for future time steps using a learned function:

$$\hat{S}_t = f_\theta(Y_{t-l:t}, Z_{t-l:t+h})$$

Methodology

Objective Function

- We train f_θ to maximize the **log-likelihood** of the predicted sleep stage distribution:

$$\max_{\theta} \mathbb{E}_{(Y,Z)\sim p(D)} \mathbb{E}_{(t,l,h)\sim p(T|D)} \log p(S_{t:t+h} \mid \hat{\phi})$$

where $ \hat{\phi} $ are the learned parameters of the predictive distribution.

Implementation of MOIRA

- **Handles multivariate time series** effectively using a **masked encoder architecture**, enabling the model to concatenate the time series models and run them in a single data input.
- **Learns from unobserved data (zero-shot)** to improve generalizability.
- **Uses multi-patch projections** for capturing hierarchical temporal structures.
- **Outputs probabilistic distributions** rather than single-point predictions.

Classification Pipeline

1. **Preprocess EEG Data**
   - Load EDF recordings and segment into **5-second windows**.
   - Apply **bandpass filtering (0.5–40 Hz)**.
   - Extract **PSD features + Catch22 features** for each epoch.
2. **Feature Transformation for MOIRA**
   - Apply **multi-patch projection layers** to generate vector embeddings.
   - Use **learnable masked embeddings** to encode future sleep stage predictions.

MOIRA-Based Classification Pipeline

3. **Training Objective**

   - Train the model using **cross-entropy loss**:

$$\text{loss} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_i^k \log \hat{y}_i^k$$

where:

- K is the number of sleep stages (at least wake and N1).
- $\hat{y}_i^k$ is the predicted probability of sample i belonging to class k.

MOIRA-Based Classification Pipeline

4. **Evaluation Metrics**
    - **Accuracy** of wake vs. N1 classification.
    - **Precision & recall** to balance specificity and sensitivity.
    - **Cross-validation across subjects** to ensure generalizability.

Initial Approaches

- Time-series forecasting models like the Google timesFM approach do not fit our classification problem.
- MOIRA model's masked encoder approach effectively captures **temporal** dependencies
- Understand catch22: Use with MOIRA model.
- Processing full 24-hour EEG recordings is currently infeasible without significant optimization or additional hardware (e.g., GPUs). We expect data aggregation to be necessary in completing this project.
- Dwnsampling is likely required. Decision of sampling rate method is important

Progress

- Acquisition and processing of datasets
- Visualizion of ANPHY time series from individual EEG channels
- Implementation of FFT with PSD plots
- Sleep staging feature extraction

Future Implementations

- **Parallelization Strategies:**
  - Implement `joblib` or `multiprocessing` to speed up **feature extraction**.
- **Deep Learning Extensions:**
  - Evaluate performance of **TCNs, CNNs, and GNNs** and potentially deploy ensemble methods on raw EEG signals.
- **Hyperparameter Tuning:**
  - Optimize **epoch length, learning rate, embedding dimensions**.
- Decision on EEG collection scaling