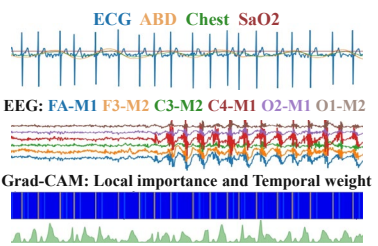
Multimodal PSG-Based Prediction of Cognitive Impairment Using an Interpretable XGBoost Ensemble

Sofia Romagnoli, Diego Cajal, Rodrigo Lozano-Puñet, Josseline Madrid

Universidad de Zaragoza
Zaragoza, Spain

Early identification of cognitive impairment (CI) is essential for mitigating the progression of neurodegenerative diseases. This work presents REMedy's contribution to the George B. Moody PhysioNet Challenge 2026, proposing a machine learning framework to predict CI using polysomnographic (PSG) data. Given the importance of interpretability for understanding physiological mechanisms and enabling clinical translation, we performed a preliminary analysis by adopting an explainable AI (XAI) approach to guide feature selection across PSG modalities. We first trained a neural network on a 662-subject dataset from the Human Sleep Project. An integrated attention mechanism combined with Grad-CAM identified the most relevant signals and features, highlighting cardiac dynamics and EEG slowing as key contributors. Based on these findings, we extracted 327 features capturing spectral dynamics, heart rate variability, and statistical descriptors from ECG, EEG, and respiration signals. Features were computed over 5-minute windows, sampled every 15 minutes across recordings. A feature selection strategy was applied by removing features with more than 10% missing data and those highly correlated ($r > 0.8$) to reduce multicollinearity. Missing values were imputed using K-nearest neighbors. To address noise and heterogeneity in PSG data, we developed an ensemble XGBoost model robust to missing modalities, generating signal-specific predictions that were aggregated into a final probability estimate. Hyperparameters were optimized using 5-fold cross-validation with an inner 3-fold scheme for threshold calibration. On the validation dataset, the proposed pipeline achieved an AUROC of 0.712, an AUPRC of 0.091, an accuracy of 0.707, and an F-measure of 0.136. These results highlight the potential of XAI-driven feature selection to improve predictive performance. However, as optimization was primarily focused on AUROC, future work will focus on improving threshold-dependent metrics, such as accuracy and F1-score, as well as refining feature extraction and enhancing ensemble performance through modality-aware weighted aggregation strategies.



PSG signals and attention outputs from Grad-CAM