# Can machine learning be used as an effective tool to improve the diagnosis of Prolonged Disorders of Consciousness based on resting-state EEG data?

**Sandra Frycz[1,2], Marek Binder[2]**

[1]Doctoral School of Social Sciences, Jagiellonian University, Krakow, Poland [2]Institute of Psychology, Jagiellonian University, Krakow, Poland

## BACKGROUND

- **Disorders of consciousness (DoC)** refer to a group of clinical conditions in which **consciousness is severely impaired** due to an extensive injury to the central nervous system. Most often is caused by traumatic brain injury or anoxia.
- The DoC **misdiagnosis rate hovers around 40%,** thus posing a serious clinical challenge.
- The DoC includes unresponsive wakefulness syndrome (UWS), the minimally conscious state minus/plus (MCS-/+) and the emergence from the minimally conscious state (EMCS). Each diagnosis describes a **different level of consciousness** which is also correlated with it's **neurobehavioral manifestation.**
- Besides neurobehavioral assessments like **Coma Recovery Scale-Revised (CRS-R), electroencephalography (EEG)** is one of the most recommended tools used in DoC diagnosis.

## AIMS

- **The aim of the study was to check, whether there are distinctive EEG spectral properties, allowing to distinguishing between conscious and unconscious patients.**
- **Also, to make the diagnosis easier and possibly more relevant, we wanted to find out, if machine learning will improve the diagnostic process.**

## METHODS

### Participants

- The original sample of 42 DoC patients with varying diagnoses (UWS, MCS-/+, EMCS) and etiology, multiple behavioral assessments with Coma Recovery Scale-Revised (CRS-R)
- Participant were divided into two groups: **"unaware" (n=22, negative class)** including UWS, and **"aware" (n=20, positive class)** consisting of patients diagnosed with MCS-/+ or EMCS

### Procedure

- CRS-R assessments for 5 consecutive days, the final diagnosis was the most frequently occurring diagnosis.
- 10-minute 64-channel EEG in **resting-state** paradigm.

### Data analysis

- The EEG data was **standardly** preprocess. The electrodes chosen for the analysis were from the central and posterior areas of the scalp.

**Parametrizing the spectra with FOOOF algorithm:**

- The FOOOF algorithm enables precise identification of the frequency peaks in the EEG spectral signal while accounting for the underlying background aperiodic signal component.

### Machine learning

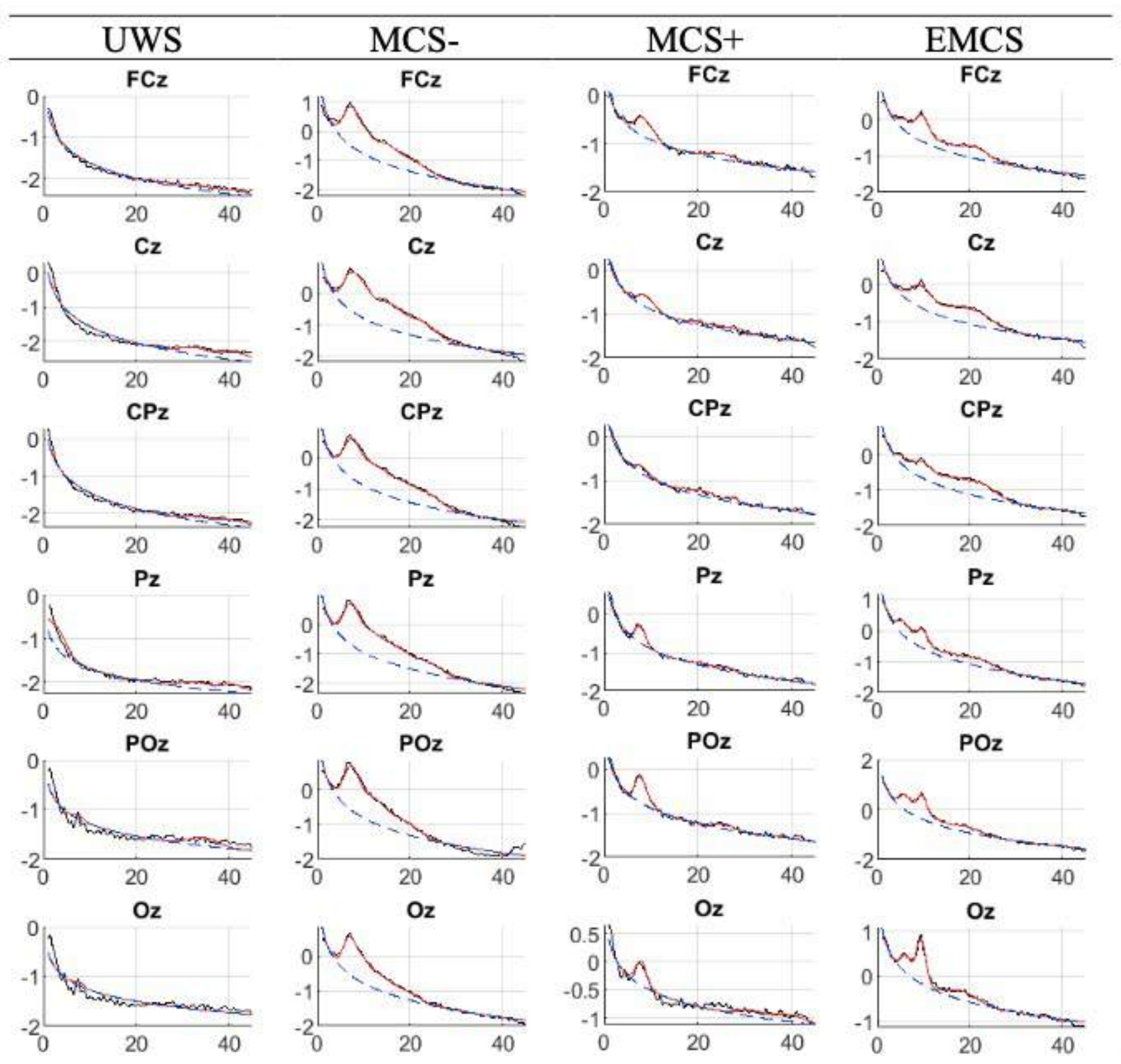
- For the machine learning, the **Bagged Trees Classifier** was used, implemented in Classification Learner app (R2023a) within the MATLAB environment
- The type of learning model = a decision tree; the maximum number of splits = 41, Trees learning models = 30, PCA = off.
- The final performance values of the classifier are the averaged values maintained over **10 consecutive measurement sessions**.

## DISCUSSION

- Our study presents, that there are distinctive EEG signal properties, that can be attributed to different DoC diagnoses, so possibly different levels of consciousness.
- The EEG spectrum from patients diagnosed with MCS-/+ and EMCS tends to present diverse oscillatory component, with higher HL ratio and MaxPeakFrequency, while the spectrum achieved from UWS patients presents the 1/f characteristic, with little or no oscillatory activity.
- Based on the ML performance, the aperiodic component, MaxPeakFrequency and HL ratio have contributed the most to the classifier performance. Moreover, the AP gradient and information from frontocentral channels seems to disturb the classifier accuracy.
- **Overall, the results indicate, that classification of DoC patients into conscious and unconscious using the ML approach, based on resting-state EEG can be helpful in assessing their level of consciousness.**

## RESULTS

### Inspection of EEG spectra and selection of ML features

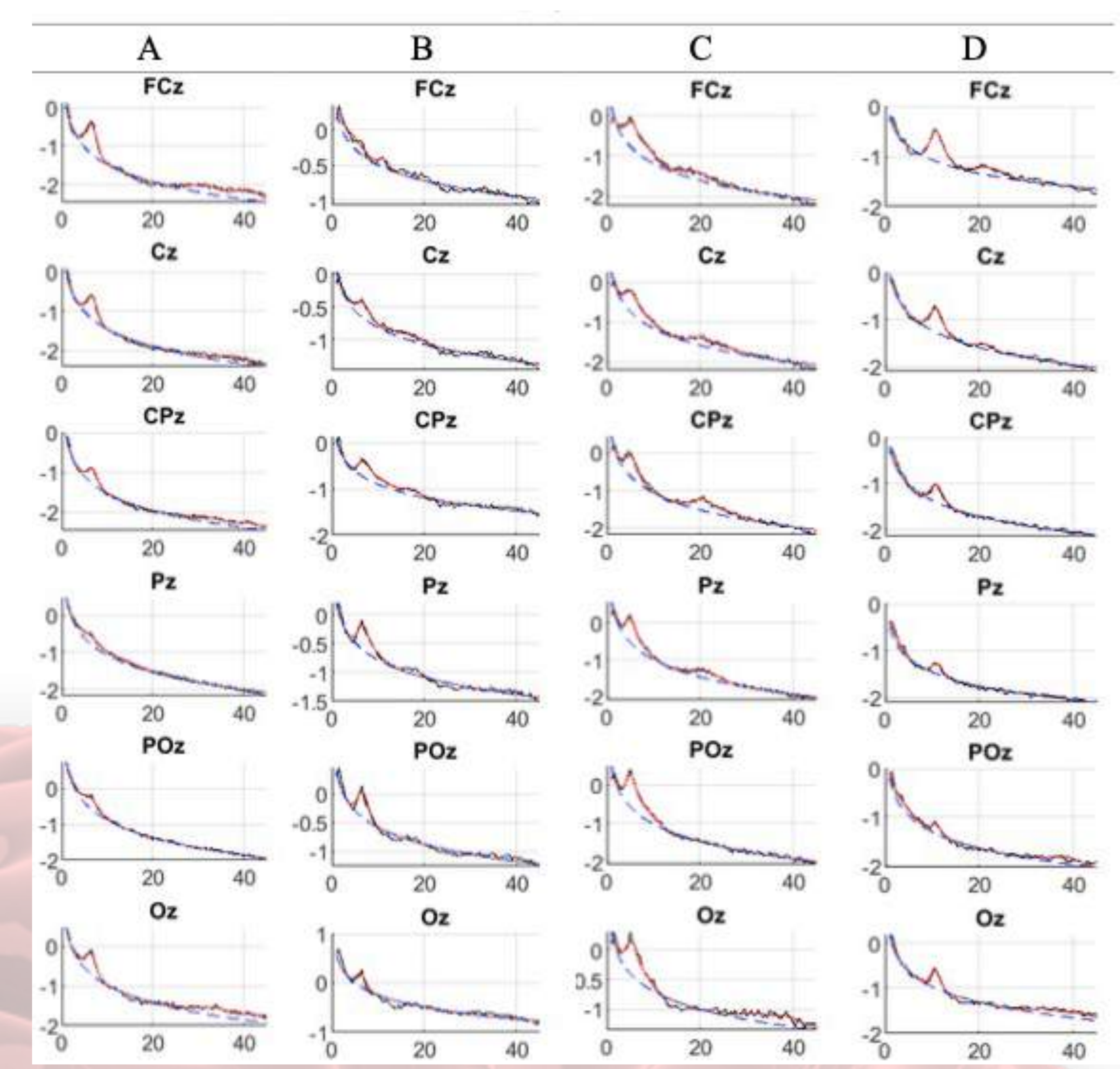



The x-axis represents the oscillations frequencies (Hz), and the y-axis demonstrates the signal power in log power units.

FIG. 1. Representative EEG spectra after FOOOF fitting for UWS, MCS-, MCS+ and EMCS patients. The individuals within "aware" group, tends to show distinctive oscillatory component, while the EEG spectra achieved from UWS patients are more likely to follow 1/f characteristic (aperiodic component), so with no oscillatory component. Furthermore, in the "aware" group the oscillations differ in frequency range and the amplitude. Also, the amplitude of the peak with the highest power decreased from posterior to the anterior areas of the brain. Other than that, patients with less favourable diagnosis tend to present higher power within low frequencies compared to high frequencies. After the EEG spectra analysis, the chosen features (within 1-45 Hz range) for ML session were:

- **Aperiodic component**
- **MaxPeakFrequency**
- **HL ratio**
- **AP gradient**

FIG. 2. Power spectra from patients with UWS diagnosis, that was based on the behavioural assessments. Some of the UWS patients showed a clear, oscillatory component, arising above the aperiodic component. The presence of the oscillatory component might suggest, that this group of patients had higher level of consciousness than it was established based on CRS-R final diagnosis. PET studies of Stender et al (2014) suggest that up to 33% patients with behavioral CRS-R diagnosis of UWS, show metabolic consumption levels comparable to MCS+/- patients.




The x-axis represents the oscillations frequencies (Hz), and the y-axis demonstrates the signal power in log power units.

### Machine learning performance

| Model number | Model properties | Classificator average performance measures | | | | | |
|---|---|---|---|---|---|---|---|
| | | Accuracy | AUC | F-score | Sensitivity | Specificity | Precision |
| 1 | **Aperiodic component MaxPeakFrequency, and HL ratio of the selected channels excluding FC1, FCz, and FC2** | **77.9% (SD=4.1)** | **0.79 (SD=0.01)** | **76.9% (SD=3.7)** | **76.5% (SD=4.1)** | **79.5% (SD=6.9)** | **77.5% (SD=6.1)** |
| 2 | AP gradient, aperiodic component, MaxPeakFrequency, and HL ratio of the selected channels excluding FC1, FCz, and FC2 | 74.3% (SD=3.5) | 0.80 (SD=0.02) | 73.9% (SD=4.5) | 73.5% (SD=4.1) | 75.0% (SD=4.9) | 72.9% (SD=4.3) |
| 3 | Aperiodic component, MaxPeakFrequency, and HL ratio of the selected channels | 74.1% (SD=4.1) | 0.80 (SD=0.02) | 74.1% (SD=3.0) | 77.5% (SD=5.4) | 70.9% (SD=10.3) | 71.5% (SD=5.9) |
| 4 | AP gradient, aperiodic component, MaxPeakFrequency, and HL ratio of the selected channels | 72.4% (SD=5.3) | 0.80 (SD=0.03) | 72.4% (SD=4.7) | 76.0% (SD=6.1) | 69.1% (SD=8.8) | 69.5% (SD=6.0) |

FIG. 2. Detailed Bagged Trees classifier average performance measures obtained for a different configuration of features. The model with the highest overall classificatory performance was Model 1. This model achieved the highest accuracy, F-score, specificity and precision. Please note that in this model, the properties from FC1, FCz and FC2 channels were excluded. Also, the AP gradient were not analyzed.

JAGIELLONIAN UNIVERSITY IN KRAKÓW

C-LAB

SANDRAFRYCZ@GMAIL.COM

ML in PL CONFERENCE 2023
26-29 OCTOBER / WARSAW, POLAND

Research supported by:
NATIONAL SCIENCE CENTRE POLAND
2018/31/B/HS6/03920