Exploring EEG Signal Features for Predicting Post Cardiac Arrest Prognosis

Antonio G C Santos[1], Joao A L Marques[2], Luís O Rigo Jr[3], João P V Madeiro[1]

[1] Federal University of Ceará, Fortaleza, Brazil
[2] Laboratory of Applied Neurosciences, University of Saint Joseph, Macao SAR, China
[3] University for International Integration of the Afro-Brazilian Lusophony, Redenção, Brazil

Abstract

In the George B. Moody PhysioNet Challenge 2023 on 'Predicting Neurological Recovery from Coma After Cardiac Arrest', our team, UFC_MDCC, employed machine learning techniques to predict patient prognosis based on electroencefalogram (EEG) signals. Our strategy was to extract features from the EEG signals, capturing both linear and non-linear characteristics from time and frequency domains. The chosen model was Random Forest, trained with various feature extraction strategies. Our team's performance on the test set at different time intervals are as follows: At 12 hours - Rank 10, Challenge Score 0.312; at 24 hours - Rank 24, Challenge Score 0.312; at 48 hours - Rank 28, Challenge Score 0.272; and at 72 hours - Rank 32, Challenge Score 0.272.

1. Introduction

Cardiac arrest remains a major medical emergency, representing a significant health challenge and a crucial decision-making point for physicians. After such a life-threatening event, the need for a reliable prognosis becomes paramount. Given the significant implications of this task, which often informs decisions about treatment goals and often withdrawal of life support, it is essential to use reliable prognostic tools. EEG, in particular, has been used since the 1960s and has proven to be a valuable tool, especially when used in conjunction with other prognostic tools [1]. Although EEG is the most commonly used prognostic tool after cardiac arrest [2], its interpretation can be challenging and somewhat subjective.

In light of the importance of accurate prognostication, the George B. Moody PhysioNet[3] Challenge 2023 was established to foster the development of algorithms predicting neurological recovery from coma after cardiac arrest[4]. This work is our contribution to the challenge, aiming to explore a comprehensive list of characteristics present in EEG signals for more accurate post-cardiac arrest prognoses. By leveraging a rich dataset from the International Cardiac Arrest Research (I-CARE) consortium [5], we aim to unveil a tool with potential clinical utility in guiding post-cardiac arrest care decisions.

2. Methodology

2.1. Data Description

The dataset used in this work is sourced from seven academic institutions in the U.S. and Europe, all participants of the International Cardiac Arrest Research consortium (I-CARE). The data encompasses 1,020 adult patients who experienced either out-of-hospital or in-hospital cardiac arrest, subsequently gained return of spontaneous circulation (ROSC), but remained in a coma.

The data was partitioned into training (60%), validation (10%), and test (30%) sets. Only the training set, consisting of 607 samples, was used in this work. From this training set, which was the only set publicly shared by the Challenge, we generated two new subsets. The first subset, which we refer to as the "local training set," containing 70% of the official training data. The second subset, termed the "local test set," comprises 30% of the official training data. Each patient may have been subjected to up to 72 hours of EEG recordings. Finally, the signal, while continuous, may exhibit gaps or deteriorations in quality due to ICU conditions or other non-physiological factors.

Throughout this work, two versions of the dataset were used. In the first version [6], only the cleanest 5 minutes of EEG data per hour are provided for each patient. These signals were extracted from electrodes placed according to the popular international standardized 10-20 system via 18 bipolar channel pairs [7], in which each electrode's voltage is linked and compared to an adjacent one to form a chain of electrodes. Finally, all EEG data were downsampled to 100 Hz.

However, in the second version of the released data [8], each full hour of EEG signal is provided, and then presented via the monopolar electrode placement [7], which involves the use of a recording electrode positioned away from the area of interest, ensuring capture of minimal activity or no relevant activity (e.g., an EEG electrode on the

 ISSN: 2325-887X DOI: 10.22489/CinC.2023.312

scalp with the reference electrode placed off the scalp). Additionally, other biosignals and annotations are introduced, like Electromyogram (EMG), Electrocardiogram (ECG) and Blood oxygen saturation (SpO2).

Clinical data elements include patient demographics (age, sex), hospital identifier, location of the arrest event, type of cardiac rhythm at the time of resuscitation, and the interval from cardiac arrest to ROSC. Patient outcomes were assessed using the Cerebral Performance Category (CPC) scale, an ordinal scale ranging from 1 (optimal neurological function) to 5 (death). For this study, outcomes were bifurcated as "Good outcome" and "Poor outcome".

2.2. Preprocessing

Data pre-processing is a pivotal step in any signal analysis work, especially when dealing with medical data. Such datasets often contain artifacts, noise, and other undesired elements that can compromise the quality of subsequent analyses. One of the consistent challenges we faced across all runs was the presence of blank or missing values. Handling these appropriately is paramount to ensure data integrity and to prevent distortions in the analyses.

To address this issue, notably in clinical data, we employed the SimpleImputer function from the scikit-learn [9] library available on Python software language. By default, this function replaces missing values with the mean of the column, ensuring a smooth imputation that doesn't introduce outliers or skew the dataset distribution.

Additionally, categorical attributes, such as gender, were converted into numerical representations. Specifically, we utilized one-hot encoding techniques to transform these categorical values into binary vectors. This transformation ensures that the algorithms treat these attributes correctly without making inappropriate ordinal assumptions.

2.3. Feature Extraction from EEG

To facilitate the process of extracting features from the EEG data, we designed a Python class called `EEGFeatureExtractor`, tailored for this purpose. This class primarily operates on a real-valued matrix, $\mathbf{X}$, defined as:

$$\mathbf{X} \in \mathbb{R}^{\mathbb{N} \times \mathbb{M}}$$

Where:

- N represents the number of channels (rows of the matrix).
- M represents the number of time-points or events (columns of the matrix).

The class is equipped with a plethora of methods, each of them dedicated to extracting a specific feature from the data matrix. The features were categorized into two main domains. Time domain, represented by the following features: Mean, Standard Deviation, Skewness, Kurtosis, Signal Energy, Zero Crossing Rate, Peak-to-Peak Amplitude, Root Mean Square, Shannon Entropy, Hjorth Parameters. Frequency domain, represented by the features: CVIF (Coefficient of Variation of Instantaneous Frequency), Total Amplitude Rhythm, Total Rhythm Index, Channels Coherence, Phase Lag Index, Envelope Correlation, Detrended Fluctuation Analysis. Upon processing, the class outputs a vector consolidating all the extracted attributes. To help and ensure the accuracy of the feature extraction process, we utilized several renowned Python libraries. Specifically, we used the statistical and signal processing functions from `SciPy` [10] and the `nolds` library for non-linear analysis [11].

2.4. Machine Learning Model

Two machine learning models were employed for distinct tasks throughout this work. The first model, a classifier, was utilized to classify the EEG signal based on the extracted features and provide a prognosis as either "Good" or "Poor". The second model, a regressor, aimed to predict the Cerebral Performance Category (CPC) ranging from 1 to 5.

Both models were implemented using the Random Forest algorithm from the `scikit-learn` library with the same set of hyperparameters as shown in Table 1. The decision to maintain the same model throughout the work was not driven by the intent to find the best model per se. Instead, the primary objective was to evaluate the quality and robustness of the extracted features. By keeping the model consistent, any variations in performance could be attributed primarily to the features, thereby providing a clearer insight into their effectiveness and relevance.

Parameter	Value
n_estimators	123
max_leaf_nodes	456
random_state	789

Table 1: Hyperparameters used on the Random Forest models.

2.5. Execution Strategies

2.5.1. Extracting features from whole EEG recording

The initial strategy employed involved extracting features from the entire EEG recording. Specifically, for each hour in which a signal was recorded, the respective signal was concatenated with the signals from the previous hours, thus generating a single consolidated signal for each patient.

Even with the multichannel nature of EEG recording, we decided to extract features from only the first channel (Fp1-F7), which specifically targets the fronto-parietal regions of the brain, which are associated with consciousness and cognitive activity, making them relevant for prognosis. of cardiac arrest. Furthermore, we extract additional features by applying the same extraction methods to the frequency bands derived from the main signal: Delta (δ): 0.5 - 4.0 Hz, Theta (θ): 4.0 - 8.0 Hz, Alpha (α): 8.0 - 12.0 Hz, Beta (β): 12.0 - 30.0 Hz. Consequently, this resulted in a set of 200 distinct attributes derived from the EEG signal. This strategy was applied in the first version of the dataset.

2.5.2. Feature extraction with additional signal preprocessing techniques

Based on the second version of the dataset, we adopted a more robust approach to signal preprocessing, aiming to get rid of unwanted noise.

The initial preprocessing phase involved applying a bandpass filter to the signals, restricting the frequency components to a range between 0.1 Hz and 30 Hz. Notably, if the mains frequency resided in this band, a filter notch would be used to eliminate possible interference. Then, the signals underwent a resampling procedure. Depending on the original sampling frequency, the data was resampled to 128 Hz or 125 Hz. After resampling, normalization was applied to constrain the data within the range [-1, 1].

To extract features from the EEG, the data were transposed into a bipolar montage using the channels F3-P3 and F4-P4. These channels also specifically target the fronto-parietal regions of the brain, and was summarized by average amplitudel. From the resulting signal the Feature Extraction class (EEGFeatureExtractor) was applied.

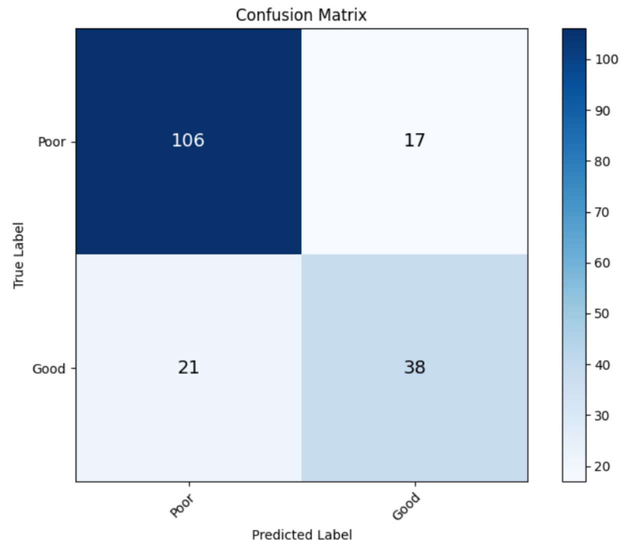

Figure 1: Confusion Matrix from the first strategy applied to the local test set.

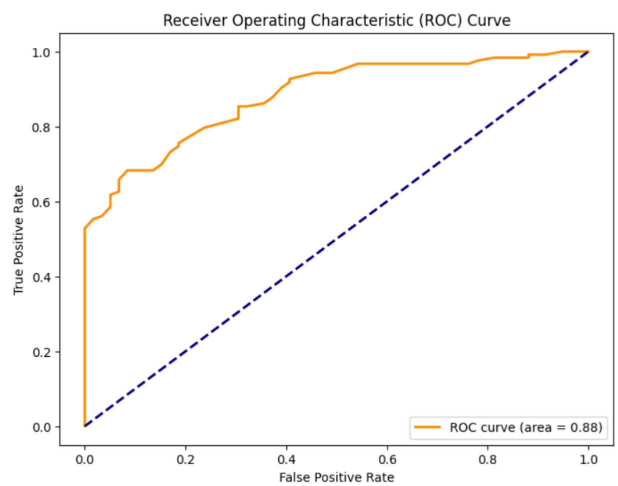

Figure 2: ROC Curve from the first strategy applied to the local test set.

3. Results

Using the first strategy mentioned in this work, we achieved the following results by using the local training and test set: Challenge Score: 0.561, AUROC: 0.880, Accuracy: 0.791, F1 Score: 0.6667, Precision: 0.6909, Recall: 0.6441, MSE: 2.025, MAE: 1.240. The Confusion Matrix and ROC Curve obtained from the first strategy when applied to the local test set are illustrated in Figure 1 and 2. No entries using this strategy were submitted during the official phase of the Challenge and therefore were not officially scored or ranked.

With the second execution strategy mentioned in this work, our `UFC_MDCC` team was ranked at the 32nd position in the official phase of the PhysioNet Challenge 2023. Metrics obtained from official validation set: Challenge Score: 0.357, AUROC: 0.682, Accuracy: 0.681, F1 Score: 0.676, Precision: 0.5758, Recall: 0.7037, MSE: 3.289, MAE: 1.557.

On the official test set, we achieved the following metrics: Challenge Score: 0.272, AUROC: 0.702, Accuracy: 0.716, F1 Score: 0.614, MSE: 2.671, MAE: 1.403.

Detailed results achieved by official data set be shown on Tables 2, 3, 4 and 5.

In short, for official phase, our team achieved a score of 0.373 on the official validation set and 0.272 on official test set.

Official Dataset	Score
Training	0.437
Validation	0.373
Test	0.312

Table 2: Scores (12 hours after ROSC).

Official Dataset	Score
Training	0.647
Validation	0.433
Test	0.312

Table 3: Scores (24 hours after ROSC).

Official Dataset	Score
Training	0.783
Validation	0.254
Test	0.342

Table 4: Scores (48 hours after ROSC).

Official Dataset	Score
Training	0.893
Validation	0.373
Test	0.272

Table 5: Scores (72 hours after ROSC. This is the time point for the Challenge).

4. Conclusion

Reflecting on our approach and results, we discern several areas that offer paths for improvement. The decision to only utilize the frontal EEG channels used from the second version of the dataset potentially limited the national richness representation of our extracted features. Extracting features from all EEG channels could have given us a more holistic view of the patient's brain activity, improving prognostic accuracy.

In retrospect, combining the feature extraction approach from first strategy with processing steps for every EEG channel of the second strategy would have likely resulted in a more robust set of features. By amalgamating the strengths of both strategies, we could have created a comprehensive feature set that better encapsulates the intricacies of the EEG signals.

In conclusion, while our results in the Challenge were promising, the insights gained from this experience provide a clear roadmap for future refinements in our approach.

References

[1] Dougherty M. The Utility of EEG in Prognosis Post-Cardiac Arrest. Practical Neurology 2017; URL https://practicalneurology.com/articles/2017-apr/the-utility-of-eeg-in-prognosis-post-cardiac-arrest. Accessed: Sep. 1, 2023.

[2] Friberg H, Cronberg T, Dünser MW, Duranteau J, Horn J, Oddo M. Survey on current practices for neurological prognostication after cardiac arrest. Resuscitation 2015;90:158–162.

[3] Goldberger A, Amaral L, Glass L, Hausdorff J, Ivanov PC, Mark R, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000;101(23):e215–e220. Online.

[4] Reyna MA, Amorim E, Sameni S, Weigle J, Elola A, Bahrami Rad A, Seyedi S, Kwon H, Zheng, WL and Ghassemi M, Putten MJAM, Hofmeijer J, Gaspard N, Sivaraju A, Herman S, Lee JW, Westover MB, Clifford GD. Predicting Neurological Recovery from Coma After Cardiac Arrest: The George B. Moody PhysioNet Challenge 2023. Computing in Cardiology 2023;50:1–4.

[5] Amorim E, Zheng WL, Ghassemi MM, Aghaeeaval M, Kandhare P, Karukonda V, Lee JW, Herman ST, Adithya S, Gaspard N, Hofmeijer J, Putten MJAM, Sameni R, Reyna MA, Clifford GD, Westover MB. The International Cardiac Arrest Research (I-CARE) Consortium Electroencephalography Database. Critical Care Medicine 2023;In press.

[6] I-CARE Database - First version (5 minutes epochs). https://physionet.org/files/i-care/1.0/, 2023. Accessed: 2023-05-01.

[7] Fehmi LG, Collura T. Effects of Electrode Placement Upon EEG Biofeedback Training: The Monopolar-Bipolar Controversy. J Neurother 2007;11(2):45–63. URL https://doi.org/10.1300/J184v11n02_04.

[8] I-CARE Database - Second version (Full hour of EEG with ECG, EMG, Spo2). https://physionet.org/files/i-care/2.0/, 2023. Accessed: 2023-08-01.

[9] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, al. Scikit-learn: Machine Learning in Python 2011;12:2825–2830.

[10] Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods 2020; 17:261–272.

[11] Christopher Schölzel. Nonlinear measures for dynamical systems, 6 2019. URL http://doi.org/10.5281/zenodo.3814723.

Address for correspondence:

Dr. Joao Paulo do Vale Madeiro
Department of Computing Science, Federal University of Ceara, Campus do Pici, 60440-900, Fortaleza, Ceara, Brazil.
E-mail:jpaulo.vale@dc.ufc.br/