Time-Embedded EEG Sequence Learning for Comatose Patients' Prognosis

Simanto Saha, Raquib-ul Alam, Andrea Samore, Andrew Goodwin, Michael Loong-Siong Wong,
Alistair McEwan, Collin Anderson

The University of Sydney, NSW, Australia

Abstract

In an intensive care unit (ICU), an accurate prognosis of comatose patients' recovery is critical for ongoing medical interventions. Patient prognosis guides decisions around continuation of care. Patients may recover from a coma despite poor initial prognosis; thus, more reliable predictors for recovery are needed. Electroencephalography (EEG)-based neurological markers may complement the current prognosis. The PhysioNet Challenge 2023 includes a dataset of EEG signals and clinical attributes from a total of 1020 adult patients in ICUs that remained in a coma after cardiac arrest, 607 of whom were dedicated to algorithmic training. We conceptualized a novel time-embedded feature space for continuous EEG followed by a bidirectional long short-term memory for learning any temporal patterns associated with comatose patients' recovery. We extracted EEG-related attributes: dynamic range, skewness, kurtosis, and subband (δ, θ, α and β) power after selecting 1 minute/hour EEG using a preprocessing algorithm. With a false positive rate < 0.05, the true positive rate (TPR) was the scoring metric at the 72nd hour post cardiac arrest. Our team, USYD_BrainBuzz ranked 27th and achieved scores of 0.26, 0.51 and 0.40 on the training, validation and testing sets, respectively. Results implicated that our approach has shown promise for continuous monitoring of comatose patients.

1. Introduction

Electroencephalography (EEG) is a clinical marker for investigating various neurological conditions, including predicting comatose patients' recovery after cardiac arrest [1, 2]. Expert visual inspection of multi-day EEG signals is tedious. Results are often subjective, and inaccurate clinical assessment may lead to a catastrophic outcome. For example, the decision to continue life support depends on the accurate prognosis, but erroneously withdrawing life support can cause death [3]. Advances in signal processing and artificial intelligence algorithms offer an alternative to manual clinical assessment of EEG signals [4]. Recent studies developed predictive models based on 1-dimensional convolutional neural network (1D-CNN), long short-term memory (LSTM) and transformer architectures for time-series or sequential data analysis [5]. Other methods include transforming data sequences into images to apply 2D-CNN for feature learning associated with coma and healthy states [4]. In this study, we have demonstrated a novel algorithm featuring signal-processing-based EEG feature extraction followed by LSTM-based time-embedded sequence analysis for predicting comatose patients' recovery.

2. The PhysioNet Challenge 2023

2.1. The Dataset

A group of researchers from the USA and Europe collected the data from seven different hospitals as part of the International Cardiac Arrest REsearch consortium (I-CARE) [1, 6, 7]. The dataset consists of 1020 patients who were in coma post cardiac arrest. The data were divided into three subsets: training, validation and testing. All participants had access to only the training set ($n = 607$) to develop their algorithms. The organizers then evaluated the submitted algorithms on the validation set, followed by the final evaluation on a testing set for selecting the competition winners.

Cerebral Performance Category (CPC) was used to classify the recovery on a scale of 1-5. While scores 1 and 2 refer to good outcomes (good neurological functions to moderate disabilities), scores 3, 4 and 5 demonstrate poor recovery (severe neurological disabilities, coma or vegetative state to death). For this competition, the aim was to develop algorithms for predicting good versus poor outcomes. The number of patients with good and poor outcomes in the training set was 225 and 382, respectively.

2.2. The Proposed Algorithm

Figure 1 illustrates the clinical environment and the block diagram of the proposed classification algorithm. Continuous EEG signals were collected up to several days, but for the competition we used data up to 72 hours. We
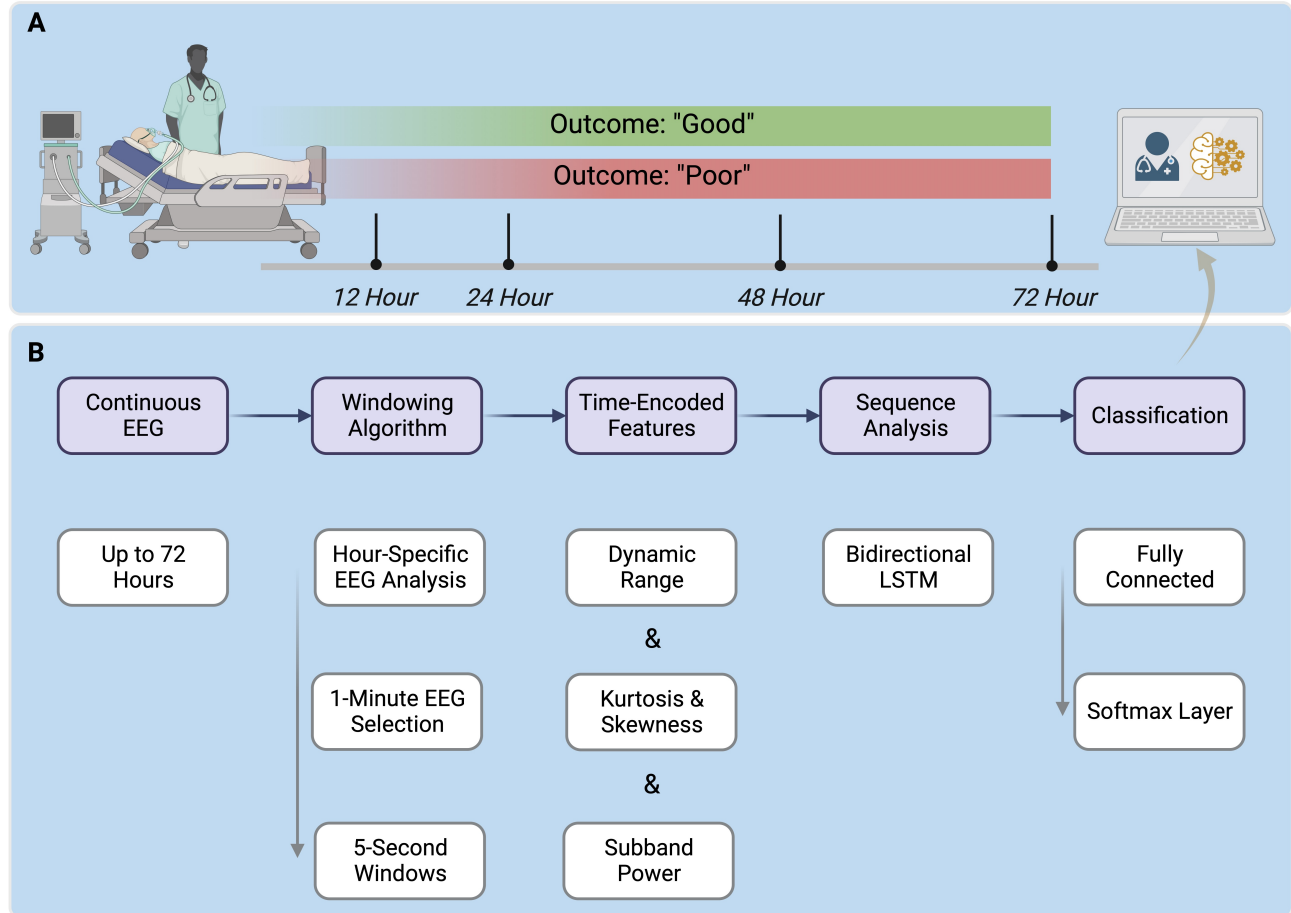
Figure 1. A schematic illustration of (A) continuous monitoring in an intensive care unit and (B) the block diagram of the proposed EEG Sequence Learning Algorithm. (Figure created with BioRender.com)

utilized only 9 unipolar EEG to construct 12 bipolar EEG signals (*i.e.*, F3-C3, C3-P3, Fz-Cz, Cz-Pz, F4-C4, C4-P4, Fz-F3, Fz-F4, Cz-C3, Cz-C4, Pz-P3 and Pz-P4). The signals were resampled to a sample frequency $Fs = 100$ Hz followed by band pass filtering with corner frequencies 0.5 and 32 Hz. We used a zero-phase Butterworth filter with an order 4.

2.2.1. Windowing Algorithm and Features

We divided each hour data into 1-minute segments followed by calculating 99% power (PW) and corresponding bandwidth (BW) for each segment. For each hour, the mean values were calculated for BW and PW. The Euclidian distance between each 1-minute segment's BW and PW, and mean BW and PW was then estimated. The distance metric was averaged over 12 bipolar channels. The segment with minimum distance was then selected. Finally, each 1-minute signal was further divided into 5-second windows for feature extraction.

Dynamic Range: Dynamic range was defined as the difference between 5th and 95th percentiles.
Skewness and Kurtosis: The skewness and kurtosis measure higher order statistical properties of a random signal and are defined as follows, respectively:

$$s = \frac{E(x - \mu)^3}{\sigma^3}, \tag{1}$$

$$k = \frac{E(x - \mu)^4}{\sigma^4}. \tag{2}$$

Here μ and σ are mean and standard deviation of x, and E refers to the expected value.
Subband Power: EEG signals were divided into four sub-bands, *i.e.*, δ ($0.5 - 4$ Hz), θ ($4 - 8$ Hz), α ($8 - 14$ Hz) and β ($14 - 30$ Hz). The power of each subband signal was then calculated as sum of the absolute squares of their time domain samples divided by the signal length.

We also utilized the clinical information such as age (in years), sex, time to return of spontaneous circulation

Table 1. Specifications of the Bidirectional LSTM model.

Layer Type	Information
Input	Sequence data size 85×864
Bi-LSTM	No. of hidden units 128
Dropout	Probability 0.2
Bi-LSTM	No. of hidden units 64
Dropout	Probability 0.2
Bi-LSTM	No. of hidden units 32
Dropout	Probability 0.2
Fully-connected	No. of neurons 2
Softmax	Output

from cardiac arrest, in-hospital or out of hospital cardiac arrest, targeted temperature (in Celsius). The clinical features were normalized to 0 to 1.

2.2.2. Time-Embedded Feature Space

All features were represented with time-embedding. For any feature, the first 12 values extracted from 5-second windows for any hour were embedded in chronological order, resulting a total of $12 \times 72 = 864$ attributes for each channel up to 72 hours. There were cases when no data were available for different hours contributing not a number (NaN) values. Data imputation, *i.e.*, NaN values were replaced with 0 to keep the feature space equal size for all patients while preserving temporal information.

2.2.3. Sequence Analysis for Classification

We then applied bidirectional LSTM for analyzing whether the time-embedded feature space depict any trends for differentiating between good and poor outcomes. LSTM is a recurrent neural network-based feature learning algorithm suitable for sequence data analysis. The output of our proposed LSTM model was fed into a fully-connected layer following classification with a softmax layer. Table 1 lists specifications of the bidirectional LSTM architecture. The LSTM training options are optimizer - Adam, gradient threshold - 1, no of epochs - 250, learning rate - 0.001, $L2$ regularization - 0.1, shuffle - every epoch and environment - GPU. Dividing the training set ($n = 607$) into two subsets, features from only 500 patients (held-out subset of the training set) were used to train the model, and the remaining data from 107 patients (training set indices: $201 - 307$) were used for validation. The output network was evaluated on the training and unseen validation and testing sets.

Table 2. Scores on training, validation and testing sets.

	Training	Validation	Testing
Score	0.26	0.51	0.40

2.3. Computational Resources

We implemented the proposed algorithm in MATLAB 2023a. The codes were then run on a g4dn.4xlarge instance on AWS featuring 16 vCPUs, 64 GB RAM (60 GB usable), 300 GB of local storage excluding the dataset, and an optional NVIDIA T4 GPU. For computing with GPU, we had a limit of 48 hours to train our proposed model followed by 24 hours for validation.

2.4. Performance Metrics

Assuming that poor outcome and good outcome represent positive and negative classes, respectively, the true positive rate (TPR) was calculated at a decision threshold using the following equation:

$$TPR = \frac{TP}{FP + FN}. \tag{3}$$

The false positive rate (FPR) was calculated such that

$$FPR = \frac{FP}{FP + TN} < 0.05. \tag{4}$$

Here, TP, FP, TN and FN are the total number of true positives, false positives, true negatives and false negatives, respectively.

3. Results and Discussion

We achieved scores of 0.26, 0.51 and 0.40 on the training, validation and testing sets when constrained to FPR < 0.05 (Table 2). The performance needs to be improved before an automated algorithm can be used in a clinical settings. Figure 2 illustrates time-embedded representation of dynamics range, skewness , kurtosis and subband (δ, θ, α and β) powers. Further investigation is essential if the selected 1-minute data reflect associated physiological changes for differentiating comatose condition from healthy state.

There were missing data posing challenges for feature representation as input to the proposed sequence learning model. Imputing missing data streams, *i.e.*, replacing NaN values with zero may have biased the training of the proposed algorithm. In some cases, there were only a few hours of EEG signals available, which seems inadequate for learning useful temporal patterns as we hypothesized in our study.

Due to the time constraints of the PhysioNet Challenge 2023, there have been unexplored avenues of the dataset.
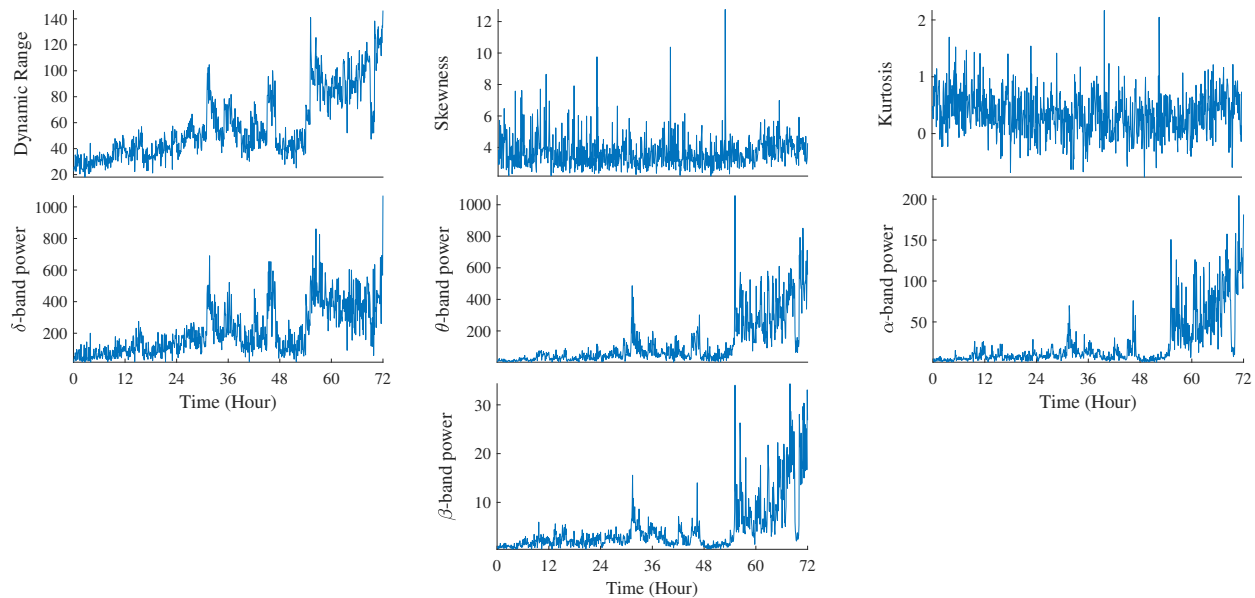
Figure 2. The time-embedded feature representation extracted from bipolar C3-P3 channel for a subject (ID #0553).

For example, supplementary signals like electrocardiography (ECG) may offer complementary information to the proposed time-embedded feature representation. Future studies may include heart rate variability-related features extracted from ECG. The proposed time-embeddings of the EEG-extracted attributes demand further investigation into whether the sequence learning for predicting comatose patients' recovery is suitable in the clinical context. The training time limit of 48 hours restricted our opportunities to explore more complex artificial intelligence-based algorithms. A more extensive dataset than only 1 minute/hour EEG signals may accommodate training a complex model.

4. Conclusion

We ranked 27th and achieved scores of 0.26, 0.51 and 0.40 on the training, validation and testing sets, suggesting that time-embedded feature representation enables recovery prediction based on EEG alterations over time. Further studies are needed to evaluate the clinical utility of our proposed algorithm and whether it is adaptable for real-time continuous monitoring of comatose patients' recovery.

References

[1] Reyna MA, Amorim E, Sameni R, Weigle J, Elola A, Bahrami Rad A, Seyedi S, Kwon H, Zheng WL, Ghassemi M, van Putten, J.A.M. M, 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.

[2] Saha S, Mamun KA, Ahmed K, Mostafa R, Naik GR, Darvishi S, Khandoker AH, Baumert M. Progress in brain computer interface: Challenges and opportunities. Frontiers in Systems Neuroscience 2021;15:578875.

[3] Kondziella D, Amiri M, Othman MH, Jakobsen EW, Jansen T, Møller K. Understanding, detecting, and stimulating consciousness recovery in the icu. Acta Neurochirurgica 2023; 165(4):809–828.

[4] Zubler F, Tzovara A. Deep learning for eeg-based prognostication after cardiac arrest: from current research to future clinical applications. Frontiers in Neurology 2023;14.

[5] Hua Y, Zhao Z, Li R, Chen X, Liu Z, Zhang H. Deep learning with long short-term memory for time series prediction. IEEE Communications Magazine 2019;57(6):114–119.

[6] Amorim E, Zheng WL, Ghassemi MM, Aghaeeaval M, Kandhare P, Karukonda V, Lee JW, Herman ST, Sivaraju A, Gaspard N, Hofmeijer J, van Putten MJ, Sameni R, Reyna MA, Clifford GD, Westover MB. The international cardiac arrest research (i-care) consortium electroencephalography database. Critical Care Medicine 2023 (in press); doi:10.1097/CCM.0000000000006074.

[7] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, 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.

Address for correspondence:

Simanto Saha
The University of Sydney, NSW 2008, Australia
simanto.saha@ieee.org