Machine learning–XGBoost analysis of language networks to classify patients with epilepsy

Authors: L.Torlay, M. Perrone-Bertolotti, E. Thomas, M. Baciu

[Brain Inform](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5563301/). 2017 Sep; 4(3): 159–169.

Published online 2017 Apr 22. doi: [10.1007/s40708-017-0065-7](https://dx.doi.org/10.1007%2Fs40708-017-0065-7)

I. Statement of the Problem:

a. Research Topic:

The research topic is the combination of the non-linear statistical ML approach, XGBoost algorithm, together with features that extracted form fMRI examination in order to achieve better classification results between patients with epilepsy and healthy people.

b. Research Problem

The research problem was, how applying statistical approach on patients with epilepsy and healthy subjects, will help to classify between them. The statistical approach will allow the identification of atypical language patterns, based on their cerebral activity assessed by fMRI (functional magnetic resonance imaging). The main objective of this study is to evaluate an objective method to distinguish patients (with epilepsy) and healthy people, based on language networks mapping with fMRI, by using machine learning approach.

Another problem this study facing is the identification of regions in the brain, who responsible for cognitive functions, that must be preserved in the pre-surgical phase in order to remove epileptic zone that causing seizures.

c. Research Hypothesis:

The research tested the hypothesis about how well the ML classifier XGBoost (Extreme Gradient Boosting algorithm), that was successfully used in other studies and domains, will identify robust patterns of language representations which are able to distinguish patients and healthy people. The authors claimed that there are no other studies (according to their knowledge) that used XGBoost to objectively classify two groups of people based on their neurophysiological features. The authors also based their hypothesis on previous range of cognitive studies that showed successful use of ML classification.

Another hypothesis examined in this paper was that language representation in the brain for healthy people the left hemisphere showed predominance (typical behavior) versus patients with epilepsy that right hemisphere or bilateral showed predominance (atypical behavior).

\* comments: The purpose of this study was clearly and concisely stated and agreed with the article title. The second hypothesis was hidden at the first reading and it required several readings to understand that the researchers explored this hypothesis.

II. Review of Literature

The authors begin with a brief of literatures that explain about focal epilepsy symptoms induced by lesion or dysfunction of a specific cerebral region, called ‘epileptic zone (EZ). Patients with focal epilepsy show reorganization or plasticity of brain networks involved in cognitive function, inducing atypical brain profiles compared to healthy subjects with typical brain profiles. In addition, they discussed on patients that becomes resistant to epileptic medications and they must undergo surgery to stop seizures. In order to remove EZ and preserved language regions, the authors mention a few cognitive tests (Wada test, fMRI with phonological and semantic tasks) that will help to identify language regions in the pre-surgical phase that must be preserved during surgery to remove epileptic regions.

The authors clearly point out the successful use of fMRI to map language network activity in patients with epilepsy by using large variety of tasks and protocols. The researchers point that in clinical practice, phonological and lexico-semantic task are generally used to maximize the amount of relevant information for language network activity. Another important claim of the authors is that the information extracting from fMRI is very helpful for detecting activated regions in the brain, but not sufficient and needs to be further processed and included more robust statistical analysis.

Next the authors also discussed on literatures about cognitive studies that showed successful use of ML classification. They reviewed previous studies that focus on patients with epilepsy, that apply ML approach based on probabilistic regression that used fMRI data to evaluate hemispheric specialization for language before surgery. The authors mention that those studies showed successful classification accuracy with dissociation between typical and atypical patterns lateralization. The authors also showed other studies that used ML approach with different data (concerning the integrity of white matter fibers) resulting in lower accuracy compared to the ML approach that used fMRI data.

The researches also reviewed literatures that strengthening their choice in XGBoost classifier in order to discriminate the fMRI from epileptic patients and healthy subjects. They mention the significant advantages of XGBoost classifier compared to other ML classifiers: (1) dealing with missing values, (2) requiring data scaling, (3) computationally efficient (use of the gradient boosting algorithm), (4) providing satisfactory results in ML competitions. They also mention that there are no other studies that used XGBoost to classify 2 groups of people based on their data features.

III. Methods

1. Data and Data preparation:

The authors examined 55 participants for this study, 16 patients with focal epilepsy and 39 healthy subjects. All were native French speakers, had normal vision and right and left handed. They used fMRI examination with two language tasks on two separate runs, phonological task and a semantic task. Each task comprised language and control conditions. The authors used a MR facility to perform the experiment. Statistical analysis was performed on the preprocessed data, they used the GLM (Generalized Linear Model) to generate the parameter estimates of activity of each voxel (volume pixel), each condition and each participant. The statistical analysis performed at individual level by calculating the main contrasts between the two pairs of regressors (phonological and control- phonological, semantic and control-semantic). The authors based on previous studies on phonological and semantic task processing to determined ten symmetrical frontal and temporal regions of interest (ROIs), five in left hemisphere and five in the right hemisphere. Those 10 symmetrical features will define the 20 features for the ML classification task, ten for phonological and ten for semantic (five ROIs in the left hemispheres and five ROIs in the right hemispheres).

\* Comments: The methods for creating and processing the data for this study were clearly explained. The instruments and development were deeply explained. The number of participants that used in this study was quite low in order to establish a scientific diagnosis and was limited only for French speakers.

b. Feature Selection Methods and ML Classification:

The authors focused on filter and wrapper methods for feature selection. They used the filter algorithms low variance, CFS, Fisher Score, Laplacian Score and Spectral Score for features ranking and for wrapper algorithms they used sequential forward selection (SFS) and sequential backward selection (SBS). all these methods did not give them stability and significance results which leads them to use exhaustive search among selection of 135 subsets that manually determined by them. For Classification the researchers used the Extreme Gradient Boosting (XGBoost) algorithm, they did not performed fine-tuning to optimize the model hyper parameters.

\* Comments: The feature selection methods were clearly detailed. The selection of the 135 subsets of features were not explained. The ML Classification XGBoost model parameters values were supplied in the article.

c. Validation Strategy and Scoring Metrics:

The authors used a validation strategy in order to prevent overfitting and to have a good assessment of model validity. They used a nested cross validation with an outer iteration of Monte Carlo cross validation which splits the data to 80% training set and 20% test set. On the training set of each iteration of Monte Carlo cross validation, for each subset of features from the 135 subsets an inner iteration of k-fold cross validation was performed on the selected subset. The outer iteration repeated twelve times in order to reduce variance and the inner iteration repeated five times for feature selection. At the end of each inner iteration a classifier was used to evaluate the selected subset of features and scored by AUC score. Finally, the XGBoost model fitted on the training set with the features subset that had the best performance and the prediction evaluated on the validation set that held out from the feature selection step.

\* Comments: The validation strategy that used by the authors was clearly explained and visualized by Tables and Figures in the article. The authors also explained the algorithms that used in the validation step and scoring step.

IV. Results and Conclusions

a. Results and Discussion

The authors begin by mention that this study is a proof of concept that illustrating the ability of the ML approach, XGBoost algorithm, to classify two distinct types, healthy subjects versus patients with epilepsy according to their language representation determined by fMRI. They presented results that showed a certain subset of features that was best to distinguished between the participants. This subset belongs to the Semantic task that performed during the fMRI test at the left hemisphere, regions BA21 and BA47. By those findings the authors showed that the left fronto-temporal activation induced by the Semantic task was the most relevant to classify patients. The authors also showed that the majority of the healthy subjects showed left hemispheric predominance versus patients with focal epilepsy that showed higher variability of language representation within and between hemispheres. These results confirmed the third profile that induced by the chronic development which says that the best distinction between patients and healthy people is based on the changes that occurring in the predominant left hemisphere.

The researchers also argued that the regions BA21 and BA47 should be considered in interaction together rather than separately. According to their findings these two regions are anatomically and functionally connected. The left fronto-temporal regions revealed by activation of BA47 and BA21 are reorganized in patients with epilepsy compared to healthy subjects. An important contribution to this reorganization is added by the interaction between semantic language and memory processes, which explains why these two cognitive functions examined together rather than separately. Furthermore, the authors results showed that compared to phonological task, the semantic task is more reliable for classifying patients even if both tasks activate fronto-temporal regions. In terms of clinical impact, the authors claimed that the differential intra-hemispheric reorganization as reflected by the regions BA21 and BA47, could suggest that the left fronto-temporal regions are sensitive to surgery and should spared during surgery to avoid postsurgical language deficits. They also claimed that XGBoost algorithm is able to compare cognitively plausible patterns, highlight the best one and able to separate categories of participants.

b. Conclusions

In conclusion, the authors state that the XGBoost algorithm is a powerful statistical method for classification tasks that showed a significant potential for classifying patients with focal epilepsy based on the cerebral region and processing their language representation. They note that the importance of the chosen features subset, BA47 and BA21 in the left hemisphere for Semantic task, was plausible given the cognitive and clinical observations that were made with these patients.

V. Overall Critique

This paper was a very in-depth research, in its most parts it was well written, well organized and well explained. There was a definite need for a short review of literature, that this paper relies on, to understand the study performed on the participants. This research critique has appraised each main section of the article to determine strength and weaknesses of the research process. Overall, this paper is a proof of concept, has mention by the authors, it mainly focused in the demonstration of the XGBoost classifier and its advantages rather than determining if XGBoost is a better classifier for classify between healthy people and patients with epilepsy. The data construction and data analysis that performed in this study was deeply detailed by the authors but not fully explained why the parameters used in the modeling part were the best for this data set. The size of participants in this study was small and focused on a specific group of participants that come from the same background. Due to these limitations of the data, it is difficult to determine a scientific fact on patients with focal epilepsy. The paper answered and confirmed its main questions and hypothesis, but it should take a larger scale of study with a much larger group of participants and from a variety of backgrounds to determine a scientific fact on patients with focal epilepsy. Also, on data with larger scale we can know better if XGBoost is a better classifier than other classifiers to classify between group of healthy people and group of patients with epilepsy.