Machine Learning Analysis of EEG Signals to Predict Working Memory Performance in Healthy and Schizophrenic Adults

Niteesh Kalangi

Mentored by Jinbo Bi, Professor of Computer Science at University of Connecticut,

Aaron Palmer, Graduate Computer Science Student and Machine-learning Researcher at University of Connecticut

Introduction

There is structural plasticity in the human brain even for adults based on experiences or events that they perceive (event-related), and this phenomenon is known as neuroplasticity. Neural rehabilitation can be possible because of this neuroplasticity (1). Brain-Computer Interface (BCI) enables direct communication between a machine and the brain by passing and analyzing brain signals (2). Generally, electroencephalography (EEG) can be used to predict real-time changes to neural activity and dynamics when paired with experimentation conducive to meaningful change in that data. EEG data often comes in the form of time-series data, sampled at short intervals to create active and accurate representations of neural activity. The common way EEG data is used to analyze psychological processes is through event-related potentiation (ERP). The ERP involves invoking planned, organized stimuli for the purpose of analyzing the neural responses and activity at the point of the stimulus. However, the information that is contained in the resting EEG data pre-stimulus is also of immense value. In fact, the resting dynamics of the brain can indeed predict oddball evoked potential, and is important to understand the way the brain sets up and prepares before the next stimulus (3). Additionally, prestimulus activity can be controlled to alter the amplitudes of ERPs (4). Moreover, another distinction to be made is that individual neural processes have been related to various oscillations in EEG responses. It has also been proven that different frequency bands in EEG data can reveal different aspects of cognitive processes. These individualistic features of neural, oscillatory activity can be left out in traditional ERP analyses (5, 6, 7). However, it can be challenging to identify the correct features and process large amounts of data without any computational assistance (1). This is where the techniques of machine-learning can come into play. Machine learning is a study area of designing statistical approaches to make computers learn from observations and data. Machine learning techniques can be used to extract patterns from data and quantify the patterns into models that can generalize to new data. They have been widely used in medical science in recent years, and are creating opportunities for more accurate classification and more efficient processing of medical data than ever before (8). In terms of machine learning and signal processing, the EEG data naturally takes the form of time series data. EEG data is generally temporarily high-resolution but spatially low-resolution. To process EEG, the signal data are often segmented into small intervals, for instance, according to the information processing stages of the brain. Afterwards, magnitudes of different frequency bands are extracted from the EEG segments. The feature extraction step is used to summarize the complex EEG signals into more interpretable features that can be potentially useful for prediction or forecast of a task performance. Automatic feature extraction can ease the process and versatility of classification models because the models can be applied to a wide range of data (9). Many different types of analysis can be applied to EEG data. An example of classifiers used for EEG data are convolutional neural networks which can be applied to analyze visual imagery (2). Another set of classification methods build on top of hand-crafted features extracted by human experts such as those frequency bands, and then apply methods such as Support Vector Machines or K-Nearest-Neighbors, to create classifiers (15). The manner in which the collection of EEG data is also very important because it can allow us to make specific conclusions about the meaning outcome of the classification. In schizophrenia, working memory is one of the most prominent and core areas of cognitive impairments (11). One method of experimentation used when collecting EEG data for the purpose of testing memory is called the Sternberg Working Memory Task (SWMT). In this task, subjects are shown a list of letters. After seeing this list, the patients are given a probed letter and are asked if the probed letter was in the original list. The neural data collected during this task can reveal a lot about characteristics of working memory when analyzed properly (9, 10). Additionally, it has been proven that features of EEG data like alpha and theta oscillations are linked to working memory specifically (12). However, when data is collected from the Sternberg Working Memory Test the data preprocessing happens in a way in which the actual letter that is probed is not often included in the data analysis, and doing this is excluding a potentially important feature as the encoding data of the letter that was probed in the list is really the important part of the EEG working memory data. This step in preprocessing can help the model understand the working of the brain in the encoding memory process much better and narrow the focus of the data that a model uses. My goal with this research is to improve on current machine learning models' understanding of working memory and the complex details involved in the process.

Proposed Methods

Preprocessing of the data

The EEG data from many subjects when they were taking a Sternberg Working Memory test has already been collected and the single-trial data has been stored. Additionally, the data regarding the results and details of the test taken in each trial has also been collected and stored. For example, this data includes whether or not they got a question right or wrong and also if the probed letter was actually in the list or not and also what letter was probed during a certain time. The first step is to separate EEG data from the trial based on whether or not the probed letter was actually there or not (positive and negative trials). After this, we can connect the probed letter of positive trials to the eeg data during the time the letter was probed. The EEG data from the letter that was probed and the corresponding letter that was shown can be used as our feature and can thus be converted to time-frequency plots, etc. This specific temporal data is crucial because it shows us how the letter probed is being encoded and the data associated with the encoding process. This allows us to increase the specificity of features of the model. Additionally, creating time frequency plots of the data will allow our model to localize the features as these plots can depict which features have importance in different times.

Deep learning classification model creation

We will be using multi-layer convolutional neural networks as our classification algorithm. In this context, we will be able to use computer vision techniques to classify the preprocessed EEG data in the form of the time-frequency plots. These time-frequency plots will be created using wavelet transforms. This technique of using CNNs on time-frequency plots will prove to be much easier and more efficient than normal because the extraction and analysis of patterns and features in the time-frequency representations will be automated, thus, allowing us to use our model on a diverse array of EEG data and saves us the time and knowledge needed to manually select features (13). The CNNs will be trained and tested on the cleaned and preprocessed time frequency data, as discussed in the last section.

Understanding the extracted features for clinical application

In order to understand and interpret the results of our deep learning model, we will use a multitude of resources to understand the broad patterns recognized by the layers in our network and also the weights that are being activated given an input. One way we will accomplish this is by using activation maximization methods. These methods allow us to search for an input pattern of data for the model that will produce the highest class prediction output response of our choosing. Each class that is part of the classification process of our model has a probability function with a gradient. This means we can perform gradient ascent on these functions to optimize the probability of a class response. Interpreting this point can allow us to figure out what features cause the strongest class responses, allowing us to form clinical conclusions about important features. We will also create saliency plots to understand which features of the EEG data prominently affect the model at specific times (14).

Research Role

I will be actively completing tasks for the research in all stages, such as the preprocessing of the data, the creation of the model, and the interpretation of the model.

References:

1. Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews*, *90*(3), 1195-1268.
2. Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2018). EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of neural engineering*, *15*(5), 056013.
3. Lee, T. W., Younger, W. Y., Wu, H. C., & Chen, T. J. (2011). Do resting brain dynamics predict oddball evoked-potential?. *BMC neuroscience*, *12*(1), 1-10.
4. Keeser, D., Padberg, F., Reisinger, E., Pogarell, O., Kirsch, V., Palm, U., ... & Mulert, C. (2011). Prefrontal direct current stimulation modulates resting EEG and event-related potentials in healthy subjects: a standardized low resolution tomography (sLORETA) study. *Neuroimage*, *55*(2), 644-657.
5. Herrmann, C. S., Senkowski, D., & Röttger, S. (2004). Phase-locking and amplitude modulations of EEG alpha: Two measures reflect different cognitive processes in a working memory task. *Experimental psychology*, *51*(4), 311-318.
6. Roach, B. J., & Mathalon, D. H. (2008). Event-related EEG time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia bulletin*, *34*(5), 907-926.
7. Chen, C. M. A., Stanford, A. D., Mao, X., Abi-Dargham, A., Shungu, D. C., Lisanby, S. H., ... & Kegeles, L. S. (2014). GABA level, gamma oscillation, and working memory performance in schizophrenia. *NeuroImage: Clinical*, *4*, 531-539.
8. Singal, A. G., Mukherjee, A., Elmunzer, B. J., Higgins, P. D., Lok, A. S., Zhu, J., ... & Waljee, A. K. (2013). Machine learning algorithms outperform conventional regression models in predicting development of hepatocellular carcinoma. *The American journal of gastroenterology*, *108*(11), 1723.
9. Johannesen, J. K., Bi, J., Jiang, R., Kenney, J. G., & Chen, C. M. A. (2016). Machine learning identification of EEG features predicting working memory performance in schizophrenia and healthy adults. *Neuropsychiatric electrophysiology*, *2*(1), 1-21.

10. Prakash, B., Baboo, G. K., & Baths, V. (2021). A Novel Approach to Learning Models on EEG Data Using Graph Theory Features—A Comparative Study. *Big Data and Cognitive Computing*, *5*(3), 39.
11. Lee, J., & Park, S. (2005). Working memory impairments in schizophrenia: a meta-analysis. *Journal of abnormal psychology*, *114*(4), 599.
12. Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, *29*(2-3), 169-195.
13. Alaskar, H. (2018). Deep learning-based model architecture for time-frequency images analysis. *International Journal of Advanced Computer Science and Applications*, *9*(12).
14. Montavon, G., Samek, W., & Müller, K. R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, *73*, 1-15.
15. Wulsin, D. F., Gupta, J. R., Mani, R., Blanco, J. A., & Litt, B. (2011). Modeling electroencephalography waveforms with semi-supervised deep belief nets: fast classification and anomaly measurement. *Journal of neural engineering*, *8*(3), 036015.