EEG-Based Brain Disorders Diagnosis through Deep Neural Networks

Gabriel A. Maggiotti (gmaggiotti@gmail.com)

Jampp.com

Abstract. In most cases, the diagnosis of brain disorders such as epilepsy or a brain tumor is slow and requires endless visits to doctors and electroencephalogram (EEG) technicians. This project aims to automate brain disorder diagnosis by using Artificial Intelligence and deep learning. There are many brain disorders can be detected by reading an Electroencephalography. Using an EEG device and collecting the electrical signals directly from the brain with a noninvasive procedure gives significant information about its health. Classifying and detecting anomalies on these signals is what doctors currently do when reading an Electroencephalography. With the right amount of data and the use of machine learning models, it could be possible to learn and classify these signals into groups like (i.e: anxiety, epilepsy spikes, abnormal tumor activity, etc). Subsequently, a trained Neural Network would interpret those signals and identify evidence of a disorder to automate the detection and classification of those disorders found. Results are promising, with classification accuracy of 99.69% for epilepsy and 85.04% for brain tumor.

Keywords: Brain Disorders, EEG, Deep Neural Networks

1 Introduction

This paper explores the use of a supervised machine learning approach to automate the detection of specific disorders on the brain by reading the EEG signals. Primarily, it focuses on Epilepsy and abnormal tumor activities. Further research could extrapolate this approach to other brain conditions.

Epilepsy is a chronic disorder caused by an imbalance in the electrical activity of neurons in one or several areas of the brain. In most epilepsies, an anomaly in electrical activity can be observed through EEG by registering spikes in the affected areas. These spikes have a unique pattern that can be seen with the naked eye on an electroencephalogram (spikes or peaks are registered with some frequency associated in the amplitudes of the electrical signals recorded). Also brain tumors presents a unique pattern in the affected area that can be observed by an EEG.

These marks are indicators of the presence of the disorder. Patients carry this pattern of spikes almost all the time. Seizures or epileptic seizures are events of short duration, being the spikes the catalysts thereof.

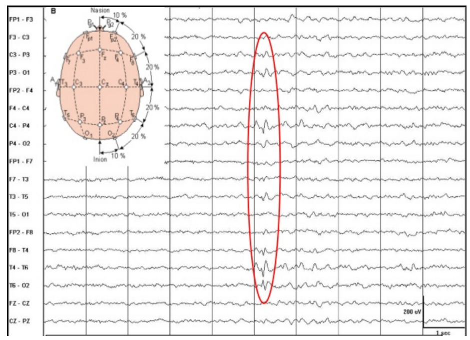

Fig. 1. Representative abnormal EEG waveforms.

This anomalous brain activity generates an observable mark or pattern. That footprint can be learned through a deep neural network. The following section will elaborate the whole process of data extraction and processing, as well as, the proposed five layers of fully connected Neural Network architecture for feature extraction. Furthermore, the process of training the network with a training-set followed by a validation of the results using a testing/validation set.

2 Related work on the subject

In the past, similar studies have been conducted using EEG datasets to analyze and make predictions on epilectic seizures and brain tumor related activity. The most common classifiers used were support vector machine (SVM) and K Near Neighborhoods (KNN) for datasets like CHB-MIT and UCI database.

In the recent years more work has been done using neural networks and deep learning. Most of these works present a much better results regardless dataset been used.

The following are the most relevant works on the subject .

Author		
Webber	[5] - ANN classification system SEN of 76% and FPR of 1 event/h	1996
A.H. Shoeb	[3] - Application of Machine Learning to Epileptic Seizure Onset and Treatment	2009
Yuan, Zhou, Li, Cai	[4] - Epileptic EEG classification based on extreme learning machine and nonlinear features	2011
Liu	[6] - Wavelet decomposition-based feature extraction and by SVM SEN of 94.5% and SPEC of 95.3%	2012
Direito	[7] - Markov modeling classification system. Identified four states - accuracy of 89.3%	2012
Rabbi	[8] - Used fuzzy algorithms for feature extraction for classification SEN of 95.8%	2012
Sharanreddy, Kulkarni	[9] - Automated EEG signal analysis for identification of epilepsy seizures and brain tumour	2015
Krisztian Buza, Jlia Kollerh	[10] - Classification of Electroencephalograph Data: A Hubness-aware Approach	2016
Thodorof, Pineau, Lim	[11] - recurrent convolutional architecture designed to capture spectral, temporal and spatial patterns representing a seizure	2016
Ullah, Hussain, Qazi, Aboalsamh,	[12] - system based on deep learning, which is an ensemble of pyramidal one-dimensional convolutional neural network (	2018
Diyuan Lu, Jochen Triesch	[13] - Image classification and object recognition methods based on convolutional neural networks	2019

Table 1. Related work

3 Methodology: Dataset Processing

Dataset used was taken from The University of California Irvine [1][2]. UCI contains an Epileptic Seizure Data Set supported by 11500 measurements from a total of 500 individuals with each has 4097 data points for 23.5 seconds and sampling rate of the data was 173.61 Hz. Then divided and shuffled every 4097 data points into 23 chunks, each chunk contains 178 data points for 1 second, and each data point is the value of the EEG recording at a different point in time. So now we have 23 x 500 = 11500 pieces of information(rows), each information contains 178 data points for 1 second(columns), the last column represents the labels. The dataset contains five different classes of 2300 samples each. Labels 1,2 and 5 were used respectively: class (1 for seizure activity; class (2 for abnormal tumor activity and class (5 for patients without seizures. Finally, two dataset were constructed. For epileptic seizures samples of classes 1 and 5 were used

(4600 samples). And for brain tumor activity classes 2 and 5 were used with the same number of samples.

To avoid saturation on the activation function and to make the gradient descent converge faster, the features were normalized to a range of values between -1 and 1 so that all features have a similar scale Eq.(1). The method used was standardization, which makes every feature have a zero mean Eq.(2) value and unit variance Eq.(3). It is calculated for each feature as follows:

$$x' = \frac{x - \hat{x}}{\sigma} \tag{1}$$

$$\mu(x_i) = 0 \tag{2}$$

$$\sigma(x_i) = \sigma(x_j) \tag{3}$$

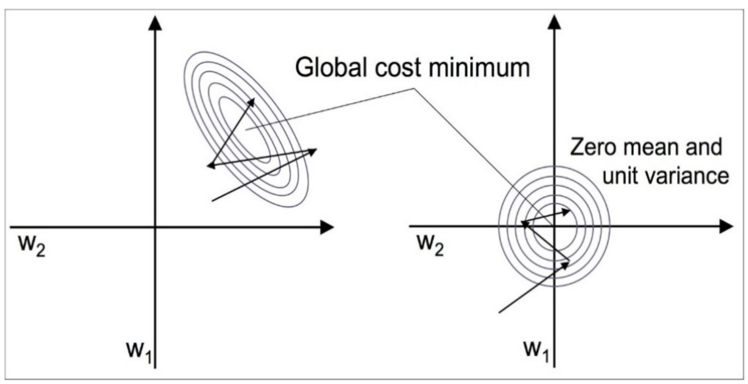

Fig. 2. Feature scaling is a method used to standardize the range of independent variables or features of data. In data processing, it is also known as data normalization and is generally performed during the data preprocessing step.

Lastly, the dataset was further split into training and validation sets. It is very important that dataset is shuffled well to avoid any element of bias before training the ML model.

4 Method / The Solution

Deep learning algorithms are composed of multiple processing layers that learn data representations with multiple levels of abstraction[15]. Using a deep learning network (DNN) implemented in Python (TensorFlow library), we classified the subjects based on each label. Design a fully connected Neural Network to capture the nonlinearity of the signals.

The proposed architecture consists of five layers of fully connected neural networks (Fig. 3) to capture data nonlinearity. An Adam optimizer[51] was used because it is an efficient extension of stochastic gradient descent optimizers. The Adam optimizer achieves good results faster than other approaches and is used for objective function minimization by iteration. It computes individual adaptive learning rates from estimates of the first and second moments of the gradients. A reference to the project and code can be found at [14].

Parameter initialization included assigning random values between 0 and 1 to the weights Eq.(4) and zero values to the biases. Also, Xavier initialization [16] was applied to the weights following Eq.(5) to make the variance to remain the same as we pass through each layer and preserve the back propagated signal as well. This helpes to reach the minimum of the cost function faster and more efficiently:

$$\theta \Rightarrow \theta = \{W_0, W_1, W_2 ..., W_L\} \tag{4}$$

$$Xavier = \sqrt{\frac{2}{features}} \tag{5}$$

The weights were still random, but positive and negative values close to 0 were assigned to produce outputs that followed a similar distribution across all neurons.

The nonlinear sigmoid function Eq.(7) was applied as the activation function of hidden layers. The objective function used measures the error between the neural networks output and the actual target, as shown in Eq.(8):

Iterate for N epochs, for each training example Xi, Yi

$$g(x)^{i+1} = \sum_{j}^{n}(x_j * w_j) \Rightarrow X^i * W^i \tag{6}$$

Hidden activation layers are components that introduce non-linearity to the system. That Allows to capture and perform very sophisticated type of classification functions.

$$L^{i+1} = sigmoid(g(x)^{i+1}) \tag{7}$$

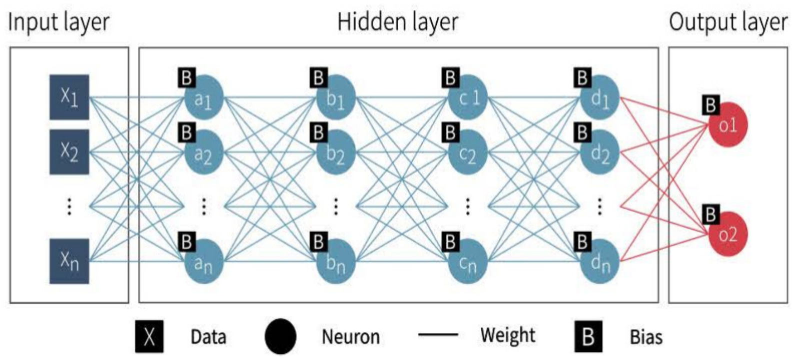

Fig. 3. Architecture for a four five fully connected Neural Network

Calculate the error comparing the output of the NN with the actual target

$$Error = \frac{1}{2}\sum_{i}^{n}(y - \widehat{y})^2 \tag{8}$$

$$\widehat{y} = Sigmoid(x_i \times w_i) \tag{9}$$

Use the chain rule to efficiently compute gradients, top to bottom

$$\frac{\partial E}{\partial w} = \frac{\partial}{\partial w}\frac{1}{2}\sum_{i}^{n}(y - \widehat{y})^2 \tag{10}$$

$$\frac{\partial E}{\partial w} = \sum_{i}^{n}(y - \widehat{y})(-\frac{\partial E}{\partial w}\widehat{t}y) \tag{11}$$

$$\Rightarrow (\frac{\partial E}{\partial w}\widehat{y}) = \widehat{y}(1 - \widehat{y}) \tag{12}$$

Back propagation of errors using the chain rule

$$\nabla = \frac{\partial E}{\partial w} \tag{13}$$

$$\nabla_{n-1} = \nabla n * W_{n-1}^{T} \tag{14}$$

As a regularization procedure for avoiding overfitting, a dropout approach was employed in the fourth hidden layer with a keep probability of 0.5[60]. The optimization procedure was iterated until the minimum error on the training set and the maximum accuracy on the validation set (the number of observations that were correctly classified) were reached (Fig. 4).

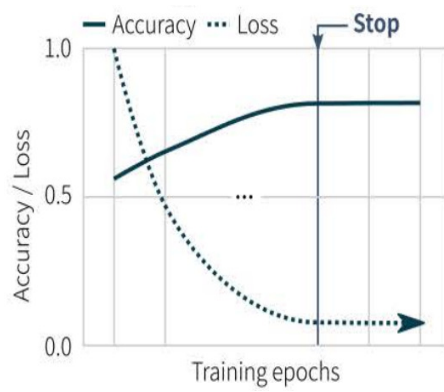

Fig. 4. Training process

5 Results

The main model used in the experiments was a five-layer fully-connected neural networks with a learning rate of 0.0001, Xavier parameter of 0.8, dropout with a keep probability of 50%, L2 regularization with a beta of 0.0001, sigmoid activation and exponential decay.

The model was trained with two datasets, one with patients who had a brain tumor, for whom brain activity was collected in the affected area. The second group was made up of patients with epileptic seizures.

On the UCI dataset,a three-class classification task was performed. The first group, group A, was comprised of 2300 samples of healthy recordings. The second, group B, was a set of 2300 samples of brain tumor activity recording, The same approach was implemented for epilepsy, where a set of recordings with epileptic seizures, group C. Then two datasets were created, A + B for brain tumor classification and A + C for epileptic seizure classification. Both datasets were shuffled and the data was normalized. For each dataset 80% was taken for training and the remaining 20% for validation.

Country List				
Model	Dataset	Accuracy %	Error %	Number of Iterations
3 Layer NN	Epilepsy	97.06 +/- 0.14	1.2 +/- 0.9	300
5 Layer NN	Epilepsy	99.69 +/- 0.05	0.3 +/- 0.6	300
5 Layer NN	Brain Tumor	85.04 +/- 0.08	2.99 +/- 0.16	2760

Table 2. Error and accuracy results for both Epilepsy and Brain Tumor Datasets. +/- is the standard deviation on each set

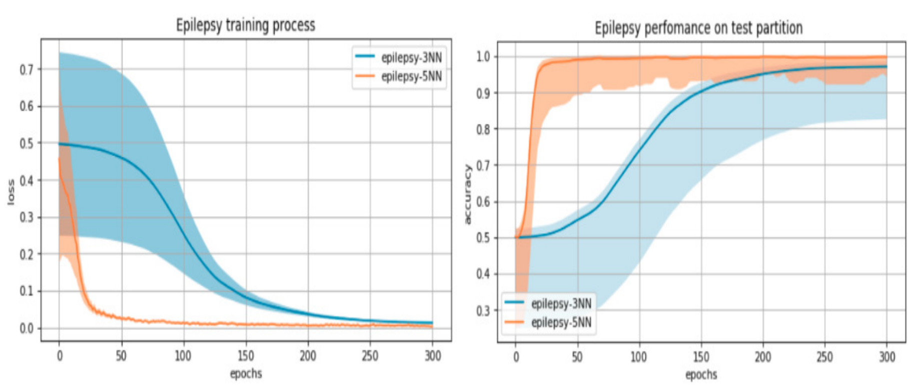

Fig. 5. Error and Accuracy during the training process for the epilepsy dataset.

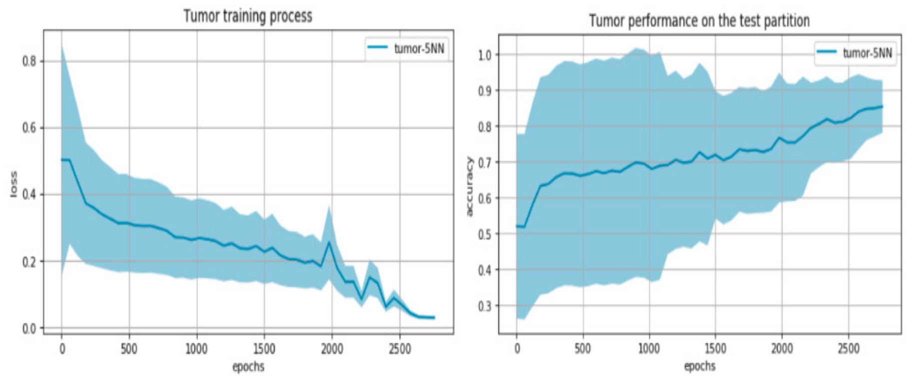

Fig. 6. Error and Accuracy during the training process for the abnormal tumor activity dataset

6 Conclusion and Future Directions

A successful automated detection and prediction of disorders introduces new innovative opportunities for diagnosis and preventive health care. This paper proposes a fast and lightway learning procedure for building a predictive model that satisfies the assignment. The use of deep neural networks in the subject turned out to be an excellent solution that presents high accuracy.

The results are prominent and suggest that the model with existing clinical systems and practices may enable clinicians to make accurate epilepsy diagnosis and start treatments earlier.

Moreover, it opens a door to extend the work on other areas like diagnosis of dementia, brain damage, brain diseases, psychiatric disorders, brain tumors, stroke, seizure forecasting from the study of interictal, preictal and ictal states and other focal brain disorders.

Another area of interest would be Electrocardiogram signals. Further works can also be done on predicting heart attacks from ECG signals (people carrying holter monitors).

References

1. [1] - https://archive.ics.uci.edu/ml/datasets/Epileptic+Seizure+Recognition
[2] - Andrzejak RG, Lehnertz K, Rieke C, Mormann F, David P, Elger CE (2001) Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, Phys. Rev. E, 64, 061907
[3] - A.H. Shoeb, Application of Machine Learning to Epileptic Seizure Onset and Treatment, 2009.
[4] Yuan, Zhou, Li, Cai- Epileptic EEG classification based on extreme learning machine and nonlinear features, 2011
[5] W.R. Webber, R.P. Lesser, R.T. Richardson, K. Wilson - An approach to seizure detection using an artificial neural network (ANN)
[6] Y. Liu, W. Zhou, Q. Yuan, S. Chen - Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG
[7] B. Direito, C. Teixeira, B. Ribeiro, M. Castelo-Branco, F. Sales, A. Dourado - Modeling epileptic brain states using EEG spectral analysis and topographic mapping
[8] A.F. Rabbi, R. Fazel-Rezai - A fuzzy logic system for seizure onset detection in intracranial EEG
[9] Sharanreddy, Kulkarni - Automated EEG signal analysis for identification of epilepsy seizures and brain tumour 2015
[10] Krisztian Buza, Jlia Koller - Classification of Electroencephalograph Data, A Hubness-aware Approach - https://www.uni-obuda.hu/journal/Buza_Koller_66.pdf
[11] Pierre Thodorof, Joelle Pineau, Andrew Lim - Learning Robust Features using Deep Learning for Automatic Seizure Detection - https://arxiv.org/pdf/1608.00220.pdf
[12] Ihsan Ullah, Muhammad Hussain, Emad-ul-Haq Qaziand Hatim Aboalsamh - An Automated System for Epilepsy Detection using EEG Brain Signals based on Deep Learning Approach - https://arxiv.org/pdf/1801.05412.pdf
[13] Diyuan Lu, Jochen Triesch - Residual Deep Convolutional Neural Network for EEG Signal Classification in Epilepsy - https://arxiv.org/pdf/1903.08100.pdf
[14] ”source code” - https://github.com/gmaggiotti/brain-disorders-prediction
[15] LeCun, Y., Bengio, Y. and Hinton, G. Deep learning. Nature 521, 436, doi:doi:10.1038/nature14539 (2015).
[16] Glorot, Xavier and Bengio, Yoshua. (2010) Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS10). Society for Artificial Intelligence and Statistics.