

# In Search of Consciousness with Artificial Intelligence

Guided by Dr. Jinan Fiaidhi

Tenzin Nyima (1119946)  
MSc Computer Science  
Lakehead University  
Thunder Bay, ON  
tnyima@lakeheadu.ca

Veera Venkata Nagendra Murthy Chandu (1093805)  
MSc Computer Science  
Lakehead University  
Thunder Bay, ON  
nagendra1693@gmail.ca

**Abstract— Artificial Intelligence (AI) is on the rise. Humans have created the computers. These computers have been making our lives easier, helping us complete tasks quicker and the automations. These computers in return may be able to help us understand who we really are. AI techniques can be applied to read our brains and decipher each and every signal of our brain. Can AI be useful in search of human consciousness? This is the right time to embark on this journey starting with a short paper review on the current status of the application of AI on Brain Signal Analysis.**

**Keywords— Consciousness, Artificial Intelligence, Machine Learning, Deep Learning, Convolution Neural Network, EEG, fMRI, MRI, Epilepsy**

## I. INTRODUCTION

Study of human consciousness is crucial in understanding who we are as human species and to understand how we make decisions collectively, and this could help make better decisions. The Google dictionary gives three similar definitions for the word consciousness as; a) “the state of being awake and aware of one’s surroundings.” b) “the awareness or perception of something by a person.” c) “the fact of awareness by the mind of itself and the world.” The first two are straight forward, but the third definition is quite convoluted. It is saying that the mind is observing the mind itself, and watching how it sees and understands the world. Does that mean, there is a brain of a brain? How does the inner brain look like? How does it work? Is it possible to see the signals from the inner brain in Electroencephalography (EEG) or in Magnetic Resonance Imaging (MRI)? Can the Artificial Intelligent (AI) approach be helpful to decipher the brain signals to detect the workings of the “inner brain”.

On 15th June, 2012 at the Caltech commencement speech, Elon Musk said if we are able to expand the scope and the scale of our understanding of human consciousness, we will be able to ask better questions. This could be the reason he founded the Neuralink company which is a research firm on Brain Machine Interface (BMI) in 2016.

This project is a quick survey of how AI is used in understanding of human thoughts and emotions. Questions like; Does deep meditation give access to the understanding of the workings of the consciousness? Can machines read our minds? Are we able to understand where human

consciousness is stemming from? Can we make our brains always do the right thing with the help of AI? Is it going to make the concept of God obsolete? This is the first project of a series of upcoming multiple projects.

The first step towards understanding the brain and its working could be understanding the ancient practice of deep meditation. Deep meditation is a self-analysis mainly of the brain on how the brain wonders and then to bring it back to a single focal point, and making it stay on course for a prolonged period of time. This is the study of one’s own brain by thought experiment.

EEG is the record or data of the action potential measurements of the neurons in the brain. It is measured through the scalp. The other brain activity measuring techniques are fMRI (functional Magnetic Resonance Imaging), PET (Positron Emission Tomography) and MEG (Magnetoencephalography).

To carry out experiments with Brain Computer Interface (BCI) may not be as difficult compared to earlier times. BCI is a headset device that is mounted on a human head to read the brain signals with varying numbers of electrodes. Each rod serves as one channel. These days, low-cost consumer product BCI can be ordered online. A company like OpenBCI sells BCI with 4, 8, 16 channels boards. BCI related experiments can be done at leisure times. Many of the research papers were done with 62 channels BCI.

This paper also includes a python implementation choosing EEG dataset on Epilepsy. Not only understanding the ultimate source of the thinking, many people have brain damage issues or general brain-health issues. The way data collection and the processing between an Epileptic patient or in pursuit of the “inner brain” signal are similar.

AI is possible mainly by two ways of computing approaches; Traditional method or Statistical Based Machine Learning approach, and Neural Network (NN) based Approach. Statistical based use statistics and probability logics to compute the solution but the NN is adapted from the biological phenomenon of how the neuron cells work hence the name. A neuron has the axon part and the dendrite part. The message is either sent or not sent based on the threshold energy. These algorithms are set up in similar fashion. The values are passed or not based on the chosen activation

functions. The computing of a solution is a matrix based operation instead of statistics and probability.

One type of NN is a Convolution Neural Network (CNN). ANN is a newer paradigm of computation taking advantage of the GPU. Before there were traditional ML like Naive Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), etc. CNN is made up of many layers, which is a sequential stack of linear or non-linear matrix operations. A layer is convoluted if there matrix operations involve filter, kernel, and activation functions like relu, tanh and sigmoid. Other layers can be pooling-layer to reduce dimensions, flattened-layer to turn two or three dimensional matrix into one dimensional, and finally the classification or regression layer for pattern recognition. CNN approach is a great solution for problems having images as input including the videos, and it can also be effective for Text Classification.

#### Epilepsy Background:

Epilepsy is characterised as having frequent seizures which are unprovoked [4]. The main instrument for identifying seizures is the electroencephalogram (EEG). EEG continuously tests the brain's electrical activity through electrodes that are mounted on the scalp or brain surface. For both clinical and experimental settings manual examination of long, continuous EEGs for seizure detection is a time-consuming and laborious process. This will take several hours for the patients admitted to treat epilepsy and closely review days of EEG recordings. Long term EEG recordings (even up to several months) are always to be checked in an experimental environment. In addition, EEG readings obtained by different testers can be incoherent because the standards for irregular EEG results are experiential. Hence it is important to establish an automated system for seizure detection.

Many machine-learning techniques have been implemented for decades to automatically detect seizures [5]. The automatic detection of seizure is hard because of the extreme variability in both inter- and intra- patient EEG. Due to the severe variation of both inter- and intra-patient EEG, the difficulty of automatic seizure detection is. Furthermore, EEG signals are strongly non-stationary and non-linear. Thus, discriminative features between seizure and non-seizure EEGs should be extracted for the development of a generalised seizure detector. Many existing methods were based on hand-made techniques for extracting features in the time domain, frequency domain, time-frequency domain, and using multiple domain combinations from EEG signals[6]. Then, these extracted characteristics were statistically evaluated, ranked and graded. The best classifier for the selected features was calculated by comparing the output of different classifiers[7]. Artificial neural network, k-nearest neighbor, logistic regression, naïve Bayes, random forest, and support vector machine have been included in the classification categories. Thus, autodetection by seizure based on machine learning has historically been composed of two separate procedures. The first part was the process of extraction of the element, and the other part was the process of classification applied to the extracted features.

Automatic feature learning thus has major advantages over conventional machine learning approaches based on the extraction and selection of manual features[8]. It can be accomplished by integrating deep learning, which automatically discovers and learns the discriminative features necessary for classification of inputs. Recently several researchers have investigated deep learning for seizure detection. Such experiments were focused on various deep neural network architectures, such as a fully connected neural network (FCNN), convolution neural network (CNN) and recurrent neural network (RNN).

## II. RELATED WORKS

Dr. Richard Davidson and his colleagues at University of Wisconsin conducted a research, and in their paper [1] submitted in 2004 asked the question, can the long-term meditators like Buddhist monks produce high-amplitude gamma-band without outside stimulation like movie or an event. The research found a resounding positive result that the monks were able to produce sustain prolonged period of time maintaining their high-amplitude gamma-band (25-42 Hz or higher).

Once the EEG signal approach is established instead of fMRI, and other approaches, data collection from a simple headset gives rise to the experiment [2] done by Rohan Dixit. This experiment used a simple EEG system produced by Neurosky, Inc. With this simple headset, a used of most basic ML model SVM became viable to classify the existence higher amplitude of deep meditation with 75.7% confidence.

To use DL approach, the authors of paper [3] did a comparative analysis between a traditional ML with neural network DL. The traditional ML used was FBCSP (Filter Bank Common Spatial Patterns) was compared with a shallow CNN and a Deep CNN. FBCSP produced 91.15%, shallow CNN produced 89.28%, and the deep CNN produced 92.40%.

#### Epilepsy Related Works:

The performance of traditional variance-based approach for detecting epileptic seizures in EEG signals is contrasted with specific approaches focused on nonlinear time series analysis, entropies, logistic regression, discrete wavelet translation and time frequency distributions. We observed that the variance-based approach provided the best outcome (100 %) [5].

It is important to check on out - of-sample data sets to determine the clinical application for any seizure identification scheme. This examination will involve evaluation of a given data on various EEG records that are known not to include an epileptic seizure. Calculating the number of false positives may encourage the similarities of various statistics. To decide how any of the techniques should be used in a therapeutic environment would require either the collection of a very large archive with appropriate time records (many hours) and/or enhanced cooperation between numbers with different study groups [5].

### III. DATASET SOURCE

The source of this Epileptic Seizure Recognition dataset is provided University of California Irvine (UCI). This repository is intended towards Machine Learning projects.

The dataset could be downloaded from the following link;

URL:<https://archive.ics.uci.edu/ml/datasets/Epileptic+Seizure+Recognition>

The dataset comes as .csv format, which stands for “comma separated values”. The file size of this dataset is 7.9 MB. It was donated by the School of Mathematical Science, Rochester Institute of Technology, NY, USA in 2017.

### IV. DATASET STRUCTURE

This dataset comes relatively pre-processed as a tabular form. EEG dataset initially comes as raw time-series sequential data type. The current tabular formatted features can be further reduced into a normalized form.

This dataset is a Multivariate having 179 attributes. The attributes or the feature names go from X1 to X179 with no particular meaning. These attribute values were actually coming from signals the electrodes from the Brain Computer interface headset. All these attributes have real and integer type values.

Along the Y-axis. The instance names are coming from subject and the time stamp based indexed values. There are 115,000 instances. These instances are obviously time series based on the number of reading per seconds. As mentioned earlier, this time-series based data is converted into a tabular format.

This is a categorical type dataset, so each instance belongs to a certain category. There are five categories {1, 2, 3, 4, 5}. These five categories can be further categorized into seizure {1}, and non-seizure as {0}. In the non-seizure {2,3,4,5} clubbed into {0}.

This 7.9 MB dataset generated from 500 participants for 23.5 seconds and have been published preprocessed into this tabular format by the donor for general purpose use.

The CNN is built with Keras library. The model summary is depicted in Figure 2. The model ran on google colab with the GPU capability.

### V. MODEL DEFINITION

Proposed Network Architecture:

A CNN recognizes simple patterns within the data, which can then be used in higher layers to shape more complex patterns. When the position of the feature within the segment is not of great significance, a 1D CNN which is powerful in extracting interesting features of the dataset from its shorter segments, can be used efficiently.

Input Layer (1D CNN): Every instance of the input has 11,392 features by limiting the vectorized maximum feature. It defines a 64 kernel of size 1. One filter would allow the

neural network to learn a single feature in the first layer which is not adequate, hence 64 filters are defined to detect 64 features.

1D CNN layer: The second layer defines a 177 kernel of size 64. To extract more complex feature, more depth is created with subsequent layers.

Batch Normalization: Batch Normalization is a technique for training very deep neural networks that standardizes the inputs to a layer for each mini batch. This has the effect of stabilizing the learning process and dramatically reducing the number of training epochs required to train deep networks.

Pooling layer: A Pooling Layer is usually used after a CNN layer in order to subsample the input to provide a simpler output that helps prevent overfitting of the data. The proposed model uses max-pooling layer after two consecutive CNN layers.

ReLU Activation Layer: The activation function is responsible for the transformation of the combined weighted input from the node into the output node. The rectified linear activation function is a piecewise linear function that will output the input directly if is positive, otherwise, it will output zero. For many neural networks, ReLU has become the defacto activation function, because models that use ReLU are easier to train and often are better at the performance.

Fully Connected (FC) Layers: The FC layers use the flattened output from the CNN layers in the vector of height 178. The proposed model has two FC layers and the last FC layer is responsible for providing a five output to predict five different sentiment. The last FC layer take a softmax as a classifier for it's a multiclass classifier.

A total of 388,546 hyperparameters is learned during training epochs for the proposed model. Different combination of network architecture was tried which necessitated the need to add the Batch Normalization layers and Dropout layers. Without these regularization techniques, the network was over- fitting, and therefore wasn't reliable.

Convolution Neural Network for a multi-class classification has three main parts; 1) convolution layer, 2) flattened layer and 3) classification or softmax layer, as shown in Fig.1.

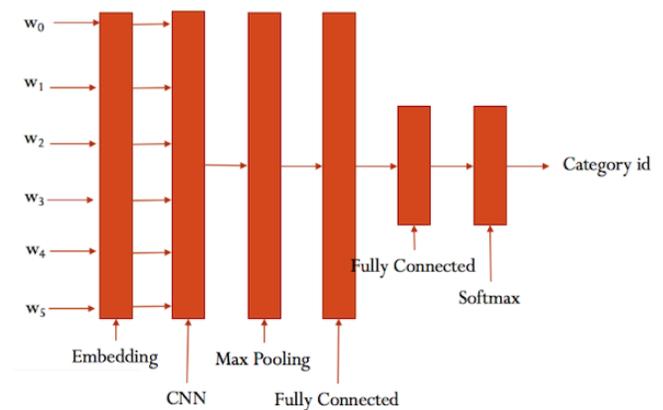


Fig. 1. Model architecture showing the CNN layers

## VI. EXPERIMENTAL SETUP

The CNN model is constructed with the help of Keras library which is a Tensor based library shown in Figure 2. This is a shallow CNN having only two Conv1D layers. It has embedding layer which is the input layer, one Max Pooling layer, a flattened layer and two dense layers. The last dense layer has only two neurons due to binary classification. The model has softmax as classifier with binary-cross-entropy for loss function and rmsprop for the optimizer.

The dataset is split into 8050 instances for training and 3450 for testing. It has two labels {0 and 1}.

Model: "sequential\_1"

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 178, 64)	11392
conv1d_1 (Conv1D)	(None, 177, 64)	8256
conv1d_2 (Conv1D)	(None, 176, 64)	8256
max_pooling1d_1 (MaxPooling1d)	(None, 88, 64)	0
flatten_1 (Flatten)	(None, 5632)	0
dense_1 (Dense)	(None, 64)	360512
dense_2 (Dense)	(None, 2)	130
Total params: 388,546		
Trainable params: 388,546		
Non-trainable params: 0		

Fig. 2: CNN model summary

## VII. RESULT AND DISCUSSIONS

The dataset has 5 labels but only one category is epileptic, so the dataset has been treated as binary classification. For the binary classification, the normal accuracy performance is sufficient.

The dataset was tried on few models to check their performances although the project is mostly dedicated to the CNN and DL models. The loss and accuracy graphs are included only from the CNN due to space constraints. A graph showing the accuracies of all the models is presented in Figure 5.

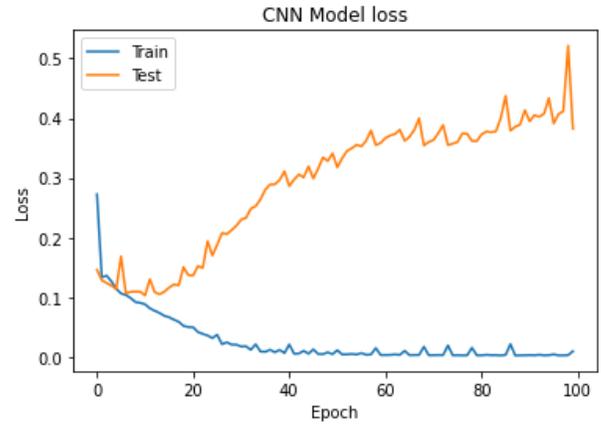


Fig. 3. Loss graph of the CNN model

The training loss graph has smooth down curve, but the testing curve shows a big gap from the training shown in Figure 3 and Figure 4. The gap indicates the model is overfitting and required tuning.

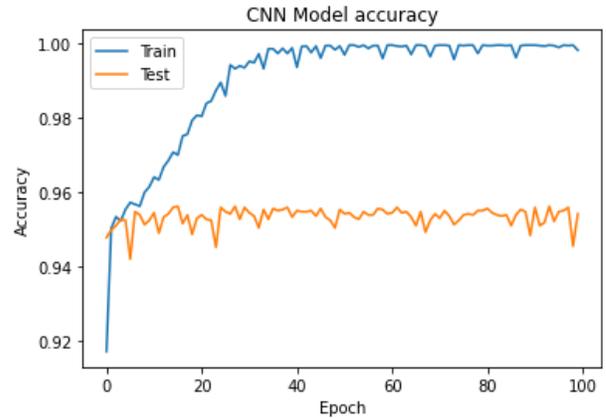


Fig. 4. Accuracy graph of the CNN model

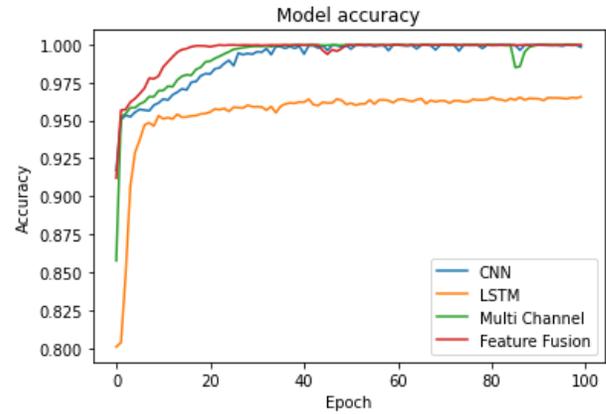


Fig. 5. Accuracy graph of four different ML models

In Figure 5, out of all the models, Feature Fusion model provided the highest accuracy except SVM which is not included here. SVM does not have epochs, so it cannot be depicted as graphs.

TABLE – I

Model	Accuracy %
SVM	97.01
LSTM	95.80
CNN	96.46
Feature Fusion	96.43
Multi-Channel	96.54

SVM is a traditional simple and direct approach unlike CNN and DL, yet it produced the highest result 97.01%.

### VIII. CONCLUSION

The epileptic dataset is downloaded from the UCI dataset. Fortunately, the dataset came preprocessed, otherwise preprocessing EEG dataset can be challenging due to high signal-to-noise ratio. This project able to reproduce the classification with high accuracy. SVM yielded the highest accuracy of 97.01%.

The CNN model showed a big gap in the curves in both loss and accuracy graphs. This indicates model tuning is necessary and this can be done in our future work. This issue can be address though hyper-parameter tunings like, batch-size, number of neurons, number of CNN layers, learning-rate and number of epochs.

From the paper reviews; one, it is established that high-gamma wave can be produced for a prolonged period of time by long-term meditators without the outside influence; these signals are purely self-induced just from the deep concentration. Two, once these brain signals are identified, now ML and DL classification models can be applied. This review found that binary classifier SVM is applied to categorize the long-term meditators' brain signal versus the ones from the novice. Three, the review also witnessed the application of DL on brain signal analysis. Although the DL approach produced only small improvement in the performance over the traditional models, the fact that DL can take raw input gives a huge advantage over other models that cannot. DL approach can also be robust towards the outliers.

If application of AI especially DL on the EEG data types renders useful, then this can bolster the other analytical tools in tandem with the well-established imaging approach like fMRI which is one of the most common approach of data collection. EEG approach of brain wave signal analysis is easy to do experiment and portal and can be carried out in non-traditional settings. Therefore, success on AI on EEG dataset makes the professionals and students more accessible to the cause. All the efforts on this type of researches can get us one step closer to demystify what is consciousness. When a social entrepreneur icon like Elon Musk is invested heavily on such cause makes it very encouraging for the students in the engineering and computer science fields to follow the trend. Finally, the experiment done by Rohan proves that he single-handedly with a simple headset carried out a proper brain signal analysis with a ML approach. And this proves any university student can carry out such experiments.

### REFERENCES

- [1] Antoine Lutz, Lawrence L. Greischar, Nancy B. Rawlings, Matthieu Ricard, and Richard J. Davidson Long-term meditators self-induce high-amplitude gamma synchrony during mental practice PNAS November 16, 2004
- [2] Dixit, Rohan. (2013). Meditation Training and Neurofeedback Using a Personal EEG Device.
- [3] Schirrneister, Robin Tibor et al. "Deep Learning with Convolutional Neural Networks for EEG Decoding and Visualization." Human Brain Mapping 38.11 (2017): 5391–5420.
- [4] Fisher, R. S. et al. ILAE official report: a practical clinical definition epilepsy. *Epilepsia* 55,475482, <https://doi.org/10.1111/epi.12550> (2014).
- [5] Mohseni, H. R., Maghsoudi, A. & Shamsollahi, M. B. Seizure detection in EEG signals: a comparison of different approaches. In: Proceedings of the IEEE Engineering in Medicine and Biology Society Suppl,67246727, <https://doi.org/10.1109/IEMBS.2006.260931> (2006).
- [6] Acharya, U. R., Sree, S. V., Swapna, G., Martis, R. J. & Suri, J. S. Automated EEG analysis of epilepsy: a review. *Knowledge-Based Systems* 45, 147–165 (2013).
- [7] Wang, L. et al. Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis. *Entropy* 19, 1–17 (2017).
- [8] Ullah, I., Hussain, M., Qazi, E. U. H. & Aboalsamh, H. An Automated System for epilepsy detection using eeg brain signals based on deep learning approach. *Expert Systems with Applications* 107,6171, <https://doi.org/10.1016/j.eswa.2018.04.021> (2018).