EEG based Epileptic data classification using Long Short Term Memory (LSTM)

Soyiba Jawed[a], Aamir Saeed Malik[a*]

[a]*Cognitive and Neural Engineering Research Group, Department of Computer systems, Brno University of Technology*

Abstract

This study aims to analyze the Epileptic Electroencephalography (EEG) signals. The goal is to differentiate EEG patterns of epilepsy and non-epilepsy in subjects. The EEG of the subjects was recorded during and before the seizure occurred. Next, the features are automatically extracted from raw EEG data; later, these features are used for classification. Here, the long short-term memory (LSTM) algorithm was implemented as a classifier. The classification accuracy of more than 0.98 is achieved for both classes. These promising results make our model reliable for practical implementations.

Keywords: EEG, LSTM, deep learning, epilepsy.

ARTICLE INFO		* Corresponding author.
RECEIVED	November 12, 2021	Email: malik@fit.vut.cz
REVIEWED	March 15, 2022	
ACCEPTED	April 06, 2022	

1. Introduction

There are 65 million people in the world that are affected by epilepsy. Epilepsy is the most severe neurological disease. It is also recognized as the most common chronic disease (Thurman et al., 2011). Epileptic patients usually suffer from misunderstanding, discrimination, and social stigma (Quintas et al., 2012) (Meyer et al., 2010). It is not easy to live with an unpredictable chronic disease as it can disrupt autonomy for daily life activities. With advances in medical science, epilepsy can be treated successfully in many cases. Still, the treatment is long-term, which affects the treatment process, especially in middle-income and low-income countries (Cameron et al., 2012). Furthermore, all patients do not respond to the medical treatment; sometimes, surgery is required, and other treatments such as neurostimulation to treat the disease. Therefore, looking at brain patterns can help in the treatment of epilepsy.

Electroencephalography (EEG) signals can be used for the identification of brain abnormalities associated with epilepsy seizures. These are represented as sharp spikes in EEG data. EEG epilepsy signals can be classified into four main categories: 1) interictal, 2) postictal, 3) seizure, and 4) preictal. Preictal is the time before a seizure occurs (Rahman et al., 2021). It remains activated in the brain for about fifteen minutes and can go up to one hour. Postictal is the time just after the seizure occurs. Interictal is the time between the two states, that is, postictal and preictal. Seizure depicts the phase where an actual seizure occurs (Rahman et al., 2021). However, many methods exist in the literature for epilepsy classification using EEG. The drawback of these EEG based epileptic seizure detection methods include low sensitivity and specificity. Sensitivity is usually ranging from 26 to 56%. Specificity is a bit better but still not reliable enough as, in some cases, it's as low as 78% (Asadzadeh et al., 2020). One possible reason is the diversity of patients in clinical settings (Benbadis et al., 2003). This paper uses the BONN university dataset to classify epilepsy and control groups.

2. Materials and Methods

2.1. Block Diagram

We propose using a deep learning algorithm called long short-term memory (LSTM) to recognize epileptic brain activity. The block diagram of the system is shown in figure 1. Here the input is the EEG dataset of the Control group and epileptic group. For the pipeline, the EEG dataset is given as input from the given data; features are extracted automatically and then classified into the two classes epilepsy and non-epilepsy.

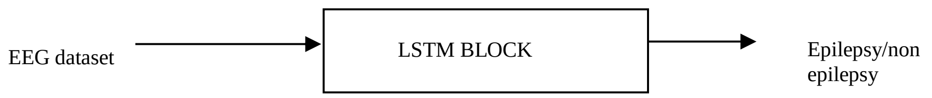

Figure 1. Block diagram of the system.

2.2. Dataset (EEG)

We utilize the publicly available Bonn epileptic database. Bonn dataset is categorized into five subsets (A - E). Data from 5 patients were recorded, and each subgroup contained 100 single-channel EEG signals of 23.6 seconds. The sampling frequency for the EEG recording was 173.61 Hz, and the signal resolution was 12-bits. The raw EEG dataset E contained epileptic seizure activity.

2.3. Recurrent Neural Network (RNN)

The significance of a Recurrent Neural Network (RNN) is that it utilizes sequential information, especially the time series data. If all the inputs and the corresponding outputs are independent of each other, then a traditional neural network can be used. However, this assumption of independence is not valid for the EEG time-series data, and hence for many tasks, that is not a good approach to follow. For example, consider a task that looks at the words in a sentence and then performs a prediction for the next word in that sentence. In order to achieve a good prediction of the upcoming word, it is better to have the knowledge and memory of the words preceding it. Here, for this example, RNN can be employed because they are recurrent in the sense that they perform the same task for every word of the sentence. The output depends on the previous computations, and it is already known that they have a "memory" that is involved in capturing and holding information about the previous computations. Issues of RNN include vanishing problem and exploding gradient problem. In this paper, we propose to use LSTM, which is a type of RNN.

2.4. *Long short-term memory (LSTM)*

1. Using LSTM for classification

LSTM is a type of RNN and RNN is a type of Neural Network. Therefore, just as a neural network can be used for classification, both RNN and LSTM can be used as classifiers. For the proposed LSTM scheme as a classifier, in addition to the input from the current time step, a hidden state is also used to hold the information from the previous time step. Hence, LSTM can perform classification with better results than RNN and Neural Network because of LSTM's good performance for EEG time series data. Since the Bonn dataset contains EEG time series data, LSTM is chosen among all other deep learning techniques.

2. Working Principle of LSTM

Traditional neural networks do not have persistence, and that is their major shortcoming. A recurrent neural network addresses this problem by having loops in them, which makes the information persist. The problem of RNN is long-term dependencies and hence LSTM is proposed to be used to address these dependencies. As mentioned above, the Long short-term memory (LSTM) deep learning networks are a type of RNN and they are involved in learning long-term dependencies. By default, LSTM remembers information because it has memory and can retain it for a long duration. The chain-like structure of LSTM with four layers is similar to RNN, though RNN has one layer only. The LSTM has gate structures, and this empowers LSTM to manipulate information by adding or removing them. Here, in this paper, the LSTM-based deep learning model has been proposed for the classification of epileptic and non-epileptic states using EEG brain waves. There are many techniques that can be used for further encoding, but this paper proposes to use one-hot encoding. Similarly, for the implementation of LSTM, tensor flow is used. The performance of the proposed method is evaluated using a confusion matrix. Figure 2 shows the LSTM architecture used in this paper. In figure 2, X_{t+1} is the input vector, H_t is the hidden state vector, W_f Is the weight vector, $tahh$ is the gate function, C_t is the Cell input activation vector, and finally σ is the forget gate activation function.

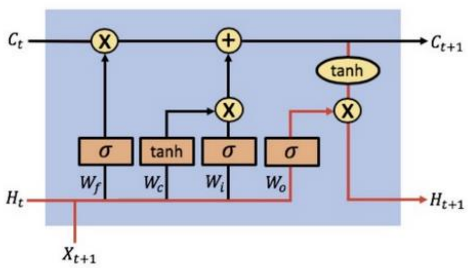

Figure 2. The LSTM architecture

3. Results

The dataset is split into 80% for training and 20% for testing. The labels are converted into one hot encoder as:

0 >> [1,0] and 1 >> [0,1].

The model is trained using the following parameters, and figure 3 shows the learning curve of the model.

- Batch size: which is the number of training examples defined in one iteration,
- Number of epochs: is a hyperparameter that explains how many times the learning algorithm works through the entire training dataset,
- Hidden size: It's the number of hidden units,
- Optimizer: It optimizes the attributes of the deep learning network.

For the implementation of the proposed method, following values are used for the above-mentioned parameters:

- Batch size: 16,
- Number of epochs: 25,
- Hidden size: 128,
- Optimizer: Adam optimizer is used for optimizing.

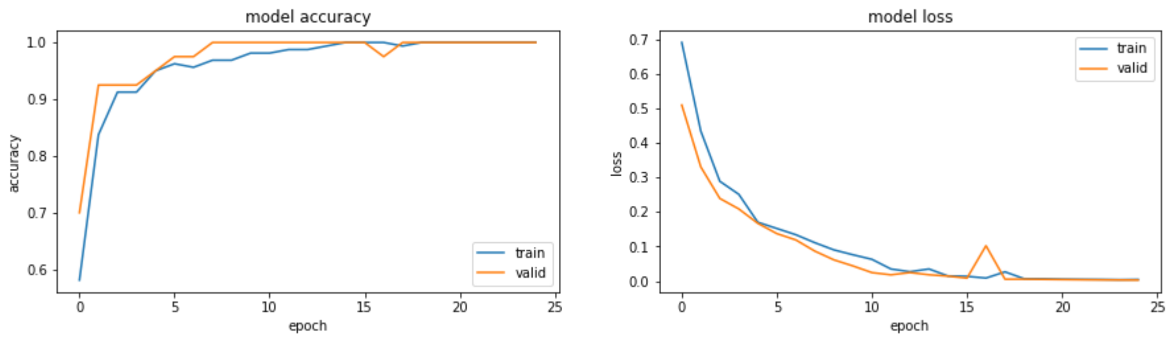

Figure 3. The learning curve of the LSTM model

The model is evaluated using the standard evaluation metrics of precision-recall and F-score using the formula stated is equation 1 to equation 3. Table 1 shows the confusion matrix and Figure 3 shows the ROC curve.

$$Precision = \frac{TP}{TP+FP} \qquad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

$$F1 = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (3)$$

Table 1: Confusion matrix of the model

	precision	recall	f1-score	support
normal	1.00	1.00	1.00	20
epileptic	1.00	1.00	1.00	20
accuracy			1.00	40
macro avg	1.00	1.00	1.00	40
weighted avg	1.00	1.00	1.00	40

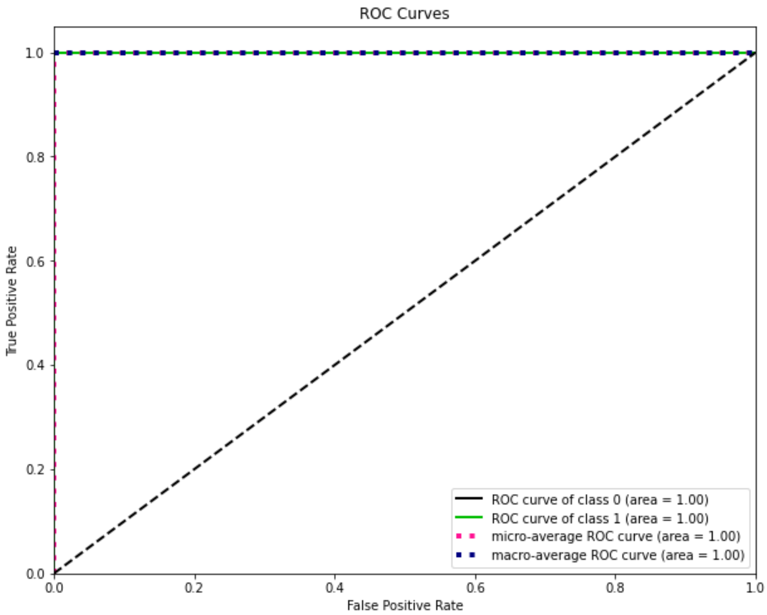

Figure 4. ROC curve of LSTM classifier

This study analysed epileptic and normal classes using EEG signals and Long short-term memory (LSTM) as a deep learning classifier. Participants were classified as epileptic or normal. The EEG signals were recorded from normal subjects and from subjects with epilepsy. Using EEG signals to identify epilepsy is important as it can help subjects with epilepsy in the future to predict seizures before time so that necessary preventive measures can be taken. However, this is just proof of concept with one subject only. This classification accuracy is expected to decrease with more subjects. As the dataset is small, overfitting is possible, which is the reason for obtaining100% accuracy. Including more subjects in the dataset will resolve overfitting and it is expected that the accuracy will be decreased from 100% to 95%-98%.

4. Conclusions

We have proposed a deep-learning model to distinguish epileptic EEG signatures from the Control in this work. EEG is recorded from 21 scalp sites using EEG electrodes. In this paper, we have proposed to use LSTM deep learning algorithm for classification. High classification accuracy was recorded for epilepsy and Control. The above results conclude that LSTM has good potential for classifying the EEG time-series signals.

References

Asadzadeh, S., Yousefi Rezaii, T., Beheshti, S., Delpak, A., and Meshgini, S. (2020), "A systematic review of EEG source localization techniques and their applications on diagnosis of brain abnormalities," *Journal of Neuroscience Methods,* vol. 339, p. 108740, 2020/06/01/, doi: https://doi.org/10.1016/j.jneumeth.2020.108740.

Benbadis, S. R. and Tatum, W. O. (2003), "Overintepretation of EEGs and misdiagnosis of epilepsy," *Journal of clinical neurophysiology,* vol. 20, no. 1, pp. 42-44.

Cameron, A. *et al.* (2012), "Mapping the availability, price, and affordability of antiepileptic drugs in 46 countries," (in eng), *Epilepsia,* vol. 53, no. 6, pp. 962-9, doi: 10.1111/j.1528-1167.2012.03446.x.

Meyer, A. C., Dua, T., Ma, J., Saxena, S., and Birbeck, G. (2010), "Global disparities in the epilepsy treatment gap: a systematic review," (in eng), *Bull World Health Organ,* vol. 88, no. 4, pp. 260-6, doi: 10.2471/blt.09.064147.

Quintas, R. *et al.* (2012), "Psychosocial difficulties in people with epilepsy: a systematic review of literature from 2005 until 2010," (in eng), *Epilepsy Behav,* vol. 25, no. 1, pp. 60-7, doi: 10.1016/j.yebeh.2012.05.016.

Rahman R. *et al.* (2021), "Comprehensive Analysis of EEG Datasets for Epileptic Seizure Prediction," in *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*, 22-28, pp. 1-5, doi: 10.1109/ISCAS51556.2021.9401766.

Thurman, D. J. *et al* (2011)., "Standards for epidemiologic studies and surveillance of epilepsy," (in eng), *Epilepsia,* vol. 52 Suppl 7, pp. 2-26, doi: 10.1111/j.1528-1167.2011.03121.x.