Automatic detection of Epileptic seizure based on EEG dataset

Kashif Usmani

Abstract: In this project, my aim is to classify the epileptic seizure using the EEG dataset. Epilepsy is a neurological disorder which occurs simultaneously, repeatedly, and suddenly. The cause of epilepsy seizure has not been identified; however, it has many symptoms like loss of awareness, unusual behaviors, and confusion. To find the seizure before it happens is a challenging task because in most cases it occurs unexpectedly [1]. EEG is significantly used in the diagnosis, classification, and treatment of epileptic seizures. The characteristics of EEG waveform play an important role in the field of epileptic seizure detection and classification [2]. In this work, we have used support vector machine (SVM) and long-short term memory (LSTM) neural network to classify the non-seizure and seizure EEG signals. We used wavelet transform to extract the features of EEG signal. For feature extraction the EEG signal is decomposed using Daubechies wavelet. In addition, we also used multi-class classification of epilepsy EEGs signals using SVM and LSTM models. The performances of classifiers are measured in terms of precision, recall, f1-score and confusion matrix.

Dataset: We use the public available dataset provided by the University of Bonn as acquired by Andrzejak et al [3, 4]. The datasets contain five subsets (A,B,C,D, and E) of single channel EEG segments. Set A and B are the surface EEG recorded during eye closed and open of healthy patients, respectively. Set C and D are the intracranial EEG recorded from unhealthy patients during the seizure free period within the seizure generating area and from outside seizure generating area of epileptic patients. Set E is the intracranial EEG of an epileptic patient during epileptic seizure. Each dataset contains 100-single channel EEG segments with duration of 23.6 seconds and the corresponding time series is sampled into 4097 data points. We divided 4097 datapoint into 23 chunks and each contains 178 datapoints for 1 second, and each data point is the value of the EEG recording at a different point in time. The total dataset consists of 11,500 samples with 178 features. The only dataset (E) has an epileptic seizure and other datasets (A,B C, and D) have samples of non-epileptic seizure datapoints. We categorized non-epileptic as class 0 (Y=0) and the datapoints of epileptic seizure categorized as class 1 (Y=1) and then take binary shape for epileptic seizure and non- epileptic seizure classification.

For feature extraction in case of SVM model, we used discrete wavelet transform. The EEG signal is decomposed using Daubechies wavelet of order 4 (db4) and the calculated wavelet coefficients are used to form the feature vectors and statistical features are extracted from the wavelet coefficients in each sub-band [5]. Total 25 statistical features are extracted from the wavelet coefficients.

Since the dataset is highly imbalanced in which majority of the samples are from non-seizure, we used SMOTE to handle the imbalanced dataset. We compared the performance of our models with balanced dataset and imbalanced dataset using these techniques. The next part of this work is multi-class classification of epilepsy EEGs signals using SVM and LSTM models. The samples are categorized into five different classes (0,1,2,3, and 4) based on the criteria (A, B, C, D and E), respectively. 60% of total dataset is used as the training dataset, 20 % of dataset is used as validation data set and the remaining 20% of dataset is used as test dataset.

References

1. I. Osorio, Hitten P. Zaveri, Mark G. Frei, and Susan Arthurs, Epilepsy: the intersection of neurosciences, biology, mathematics, engineering, and physics., CRC press, 2016.
2. S. Usman, M. Usman, and S. Fong, "Epileptic seizures prediction using machine learning methods", Comput. Math. Meth. Med., vol. 2017, pp. 1-10, 2017.
3. Andrzejak et al., "Indications of nonlinear deterministic and finite - dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state", Physical Review E, vol. 64, no. 6, pp. 061907, 2001.
4. Data link: https://repositori.upf.edu/handle/10230/42894
5. Adeli H, Zhou Z, Dadmehr N. “Analysis of EEG records in an epileptic patient using wavelet transform”, Journal of Neuroscience Methods 2003;123(1):69–87.