Using Machine Learning Models to Predict Epileptic Seizures in Patients Using EEG Readings Using Python

Content

1. Abstract

2. Formulation of the Problem

2.1 Problem Statement

3. Methodology

3.1 Data Collection and Preparation

3.2 Logistic Regression

3.3 Support Vector Machine

4. Results

5. Conclusion

6. References

7. Appendix

**1. Abstract**

This report focuses on using machine learning models namely Linear Regression and Support Vector Machine (SVM) to detect seizure in epileptic patients using the readings from electroencephalogram, or EEG, a device used by hospitals to detect whether the patient is suffering from epileptic seizures or not. Since, it requires highly skilled medical professionals to analyze the reading on electroencephalogram, or EEG and get the correct diagnosis, the seizures often go undiagnosed or misdiagnosed. Using machine learning, to detect the seizure can help in the detection as well as an earlier diagnose of epileptic seizures.

2. Formulation of the Problem

2.1 Problem Statement

Epilepsy affects almost 65 million people globally today. Seizures are caused by sudden, uncontrolled electrical disturbances in the brain and they can present in humans as changes in behavior, movements or feelings, and altered consciousness. Epileptic seizures occur in hospitalized patients to newly born infants. Most of epileptic patients go undiagnosed. Nearly 80% of people with epilepsy live in low- and middle-income countries and 75% of epilepsy patients living in low-income countries do not get optimal treatment. The electroencephalogram, or EEG, is a device that can detect abnormal electrical impulses in the brain during a seizure. But it takes highly trained professionals understands and accurately evaluate the reading on the EEG. Since, most of ICUs don’t aren’t equipped with machines like EEG the monitor brain, its hard to detect epileptic seizure. In addition to that doctors or nurses can’t survey the EEG 24/7 and there is also high room for human error. Early detection of epilepsy is known to improve a patient’s condition more efficiently than late detection. Using machine learning that can easily examine if the patient is having an epileptic seizure can make the diagnose readily available as well as easier to get.

3.Methodology

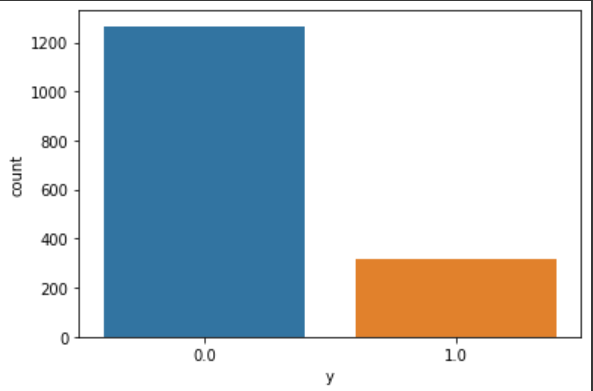
3.1 Data Collection and Preparation

Following Lateef et al. the first step for building a machine learning model is getting the data. The dataset taken consists of 5 different folders, each with 100 files, with each file representing a single subject/person. Each file is a recording of brain activity for 23.6 seconds. The corresponding time-series is sampled into 4097 data points. Each data point is the value of the EEG recording at a different point in time. So we have total 500 individuals with each has 4097 data points for 23.5 seconds. It is further 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(row), each information contains 178 data points for 1 second(column), the last column represents the label y {1,2,3,4,5}.y contains the category of the 178-dimensional input vector. Specifically y in {1, 2, 3, 4, 5}:

* 5 - eyes open, means when they were recording the EEG signal of the brain the patient had their eyes open
* 4 - eyes closed, means when they were recording the EEG signal the patient had their eyes closed
* 3 - they identify where the region of the tumor was in the brain and recording the EEG activity from the healthy brain area
* 2 – recording of the EEG from the area where the tumor was located
* 1 - recording of seizure activity

All subjects falling in classes 2, 3, 4, and 5 are subjects who did not have epileptic seizure. Only subjects in class 1 have epileptic seizure.

We’ll use the 80:20 rule, using 80% of the datapoints as training set while the other 20% as the testing set.



Graph 3.1 Non-Seizure and Seizure datapoints in the dataset

3.2 Linear Regression

Linear regression is a supervised learning algorithm used when target/dependent variable continues real number. It establishes relationship between dependent variable y and one or more independent variable x using best fit line.

3.3 Support Vector Machine

Support Vector Machines (SVM) is another simple, yet foundational, machine learning algorithm. The goal of SVM is to find an optimal hyperplane, across the n-dimensional space created by the independent variables that distinctly classifies the data points.

4. Results

The accuracy of Logistic Regression model was about 76% while that of Support Vector Machine (SVM) model 97%.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | f1-score | Support |
| Non-seizure | 0.82 | 1.00 | 0.90 | 1853 |
| Seizure | 0.98 | 0.11 | 0.19 | 447 |
| Accuracy | - | - | 0.83 | 2300 |
| Macro Avg | 0.90 | 0.55 | 0.55 | 2300 |
| Weighted Avg | 0.85 | 0.83 | 0.76 | 2300 |

Table 4.1 Accuracy and Precision of Logistic Regression Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | f1-score | Support |
| Non-seizure | 0.97 | 0.99 | 0.98 | 1853 |
| Seizure | 0.97 | 0.87 | 0.92 | 447 |
| Accuracy | - | - | 0.97 | 2300 |
| Macro Avg | 0.97 | 0.93 | 0.95 | 2300 |
| Weighted Avg | 0.97 | 0.97 | 0.97 | 2300 |

Table 4.2 Accuracy and Precision of SVM Model

5. Conclusion

Machine learning can be used detect if a patient is having seizure or not based on the EEG reading. The Logistic Regression model is less efficient than Support Vector Machine model. We can more accuracy using more sophisticated Machine Learning models like Long Short Term Memory model.

6. References

* Lateef HA, Ralston G, Bright T, Soundarajan A, Carpenter J. 2022. SeizureSeeker: A Novel Approach to Epileptic Seizure Detection Using Machine Learning. *J Neurol Exp Neurosci* 8(1): 1-8.
* Github repository on Epileptic detection using ML: [*https://github.com/millengustavo/epileptic-seizure-recognition*](https://github.com/millengustavo/epileptic-seizure-recognition)[accessed on 10 December 2022]
* Support Vector Machine Tutorial [*https://www.youtube.com/watch?v=7sz4WpkUIIs&t=2155s&ab\_channel=CloudandMLOnline*](https://www.youtube.com/watch?v=7sz4WpkUIIs&t=2155s&ab_channel=CloudandMLOnline) [accessed on 10 December 2022]
* Support Vector Machine Tutorial 2 *https://www.youtube.com/watch?v=\_BdBJOOqMes&ab\_channel=TalhaAnwar* [accessed on 10 December 2022]

7. Appendix

