Epilepsy/Seizure Detection Using Machine Learning

Jigisha Srivastava[1] Akanksha Singh[2] Alfia[3] Akanksha Srivastava[4]
[1,2,3,4]Department of Computer Science and Engineering
[1,2,3,4]Institute of Technology And Management, India

Abstract — Epilepsy is a serious chronic neurological disorder, can be detected by analysing the brain signals produced by brain neurons. Neurons are connected to each other in a complex way to communicate with human organs and generate signals. The monitoring of these brain signals is commonly done using Electroencephalogram (EEG) and Electrocorticography (ECoG) media. These signals are complex, noisy, non-linear, non-stationary and produce a high volume of data. Hence, the detection of seizures and discovery of the brain-related knowledge is a challenging task. Machine learning classifiers are able to classify EEG data and detect seizures along with revealing relevant sensible patterns without compromising performance. As such, various researchers have developed number of approaches to seizure detection using machine learning classifiers and statistical features. The main challenges are selecting appropriate classifiers and features. The aim of this paper is to present an overview of the wide varieties of these techniques over the last few years based on the taxonomy of statistical features and machine learning classifiers— 'black-box' and 'non-black-box'. The presented state-of-the-art methods and ideas will give a detailed understanding about seizure detection and classification, and research directions in the future.

Keywords: Heartbeat Sensor, Accelerometer Sensor, Brain Sensor, GPS, GSM

I. INTRODUCTION

The word epilepsy originates from the Latin and Greek word 'epilepsia' which means 'seizure' or 'to seize upon'. It is a serious neurological disorder with unique characteristics, tending of recurrent seizures. The context of epilepsy, found in the Babylonian text on medicine, was written over 3000 years ago. This disease is not limited to human beings, but extends to cover all species of mammals such as dogs, cats and rats. However, the word epilepsy does not give any types of clues about the cause or severity of the seizures; it is unremarkable and uniformly distributed around the world.

Several theories about the cause are already available. The main cause is electrical activity disturbance inside a brain, which could be originated by several reasons such as malformations, shortage of oxygen during childbirth, and low sugar level in blood. Globally, epilepsy affects approximately 50 million people, with 100 million being affected at least once in their lifetime. Overall, it accounts for 1% of the world's burden of diseases, and the prevalence rate is reported at 0.5–1%. The main symptom of epilepsy is to experience more than one seizure by a patient. It causes a sudden breakdown or unusual activity in the brain that impulses an involuntary alteration in a patient's behaviour, sensation, and loss of momentary consciousness. Typically, seizures last from seconds to a few minute(s), and can happen at any time without any aura. This leads to serious injuries including fractures, burns, and sometimes death.

Based on the symptoms, seizures are categorized by neuro-experts into two main categories—partial and generalized. Partial seizure, also called 'focal seizure', causes only a section of the cerebral hemisphere to be affected. There are two types of Partial seizure: simple-partial and complex-partial. In the simple-partial, a patient does not lose consciousness but cannot communicate properly. In the complex-partial, a person gets confused about the surroundings and starts behaving abnormally like chewing and mumbling; this is known as 'focal impaired awareness seizure'. On the contrary, in the generalized seizures, all regions of the brain suffer and entire brain networks get affected quickly. Generalized seizures are of many types, but they are broadly divided into two categories: convulsive and non-convulsive.

1) We have done the review according to five main dimensions. First, researchers who adopted the EEG, ECoG or both for seizure detection; second, significant features; third, machine learning classifiers; fourth, the performance of the classifier during a seizure, and last, knowledge discovery (e.g., seizure localization).
2) Through study, it has been explored that an ensemble of decision trees (i.e., decision forest–random forest) classifier outperforms other classifiers (ANN, KNN, SVM, single Decision Tree).
3) We also suggest, how decision forest algorithms could be more effective for other knowledge discovery tasks besides seizure detection.
4) This study will help the researchers with their data science backgrounds to identify which statistical and machine learning classifiers are more relevant for further improvement to the existing methods for seizure detection.
5) The study will also help the readers for understanding about the publicly available epilepsy datasets.
6) In the end, we have provided our observations by the current review and suggestions for future research in this area.

Applications of machine learning are significantly seen on health and biological data sets for better outcomes. Researchers/scientists on different areas, specifically, data mining and machine learning, are actively involved in proposing solutions for better seizure detection. Machine learning has been significantly applied to discover sensible and meaningful patterns from different domain datasets. It plays a significant and potential role in solving the problems of various disciplines like healthcare. Applications of machine learning can also be seen on brain datasets for seizure detection, epilepsy lateralization, differentiating seizure sates, and localization. This has been done by various machine learning classifiers such as ANN, SVM, decision tree, decision forest, and random forest Certainly, in the past, numerous reviews have been carried out on seizure detection along with applied features, classifiers, and claimed accuracy without focusing on the challenges faced by the data scientists

whilst doing research on datasets of neurological disorders. Therefore, this article provides a detailed study of machine learning applications on epileptic seizure detection and other related knowledge discovery tasks. In this review, the collected articles are from well-known journals of their relevant field. These references are either indexed by *SCOPUS* or *Web of Science (WOS)*. Besides, we also considered some good ranked conference papers. Extensive literature is available covering the deep analysis of different features and classifiers applied on EEG datasets for seizure detection. Both, feature extraction and applying classification techniques are challenging tasks. Previous literature reveals that for the past few years, interest has been increased in the application of machine learning classifiers for extracting meaningful patterns from EEG signals, which helps for detecting seizures, its location in the brain, and other impressive related knowledge discoveries. Three decades ago, Jean Gottman, analysed and proposed the model for effective usage of EEG signals by applying different computational and statistical techniques for automatic seizure detection. Furthermore, the research has been carried out by different signal processing methods and data science methods to provide better outcomes.

A. Data Collection

The initial requirement is to collect the dataset of brain signals. For this, different monitoring tools are used. Typically, the mostly used devices are EEG and ECoG, because their channels or electrodes are implanted by glue on the surface of the scalp as per 10–20 International system [48] at different lobes. Each of them has a wire connection to the EEG device, providing timely information about the variations in voltage, along with temporal and spatial information [49]. As highlighted in Fig. 2, the EEG channels are placed on the subject's scalp, and the electrical signals are read by the EEG monitoring tool and it displays these raw signals over the screen. Further, these raw signals have been carefully monitored by the analyst and classified into 'seizure' and 'non-seizure' states.

B. Data Transformation

After data collection, the next crucial step is to transform the signal data into a 2-D Table format. The reason for this is to make it easier for analysis and provide necessary knowledge like seizure detection. This datum is raw because it has not been processed yet. Therefore, it will not be suitable to give relevant information. To do the processing, different feature selection modalities have been applied. This step also presents the dataset as supervised, which means that it provides the class attribute with possible class-values.

C. Dataset Preparation

For data transformation, data processing is a decisive step to extract meaningful information from the collected raw dataset. As such, different feature extraction techniques have been used; as shown in Table 1. These methods are generally applied to the extracted EEG signal data set. The raw dataset becomes rich in terms of different statistical measure values.

II. LITERATURE REVIEW

1) A disruption in the brain's electrical activity causes epileptic seizures. Many forms of epileptic seizures occur. Study says there are many forms of epileptic seizures occur. These seizures occur with little or no warning and about one third of people with epilepsy are resistant to treatment that controls these seizures Machine learning techniques have been shown to be successful in the creation of predictive seizure therapies, the development of Internet of Things (IOT), portable low-power devices with network connectivity to help patients. Researchers utilized deep machine learning and based on the data, created a complex analytical tool for reading the patient's electroencephalogram (EEG) data from a wearable cap or other portable device for collecting EEG information. So, it could predict algorithm with sensitivity of up to 81.4 per cent and false prediction rate as low as 0.06 an hour.
2) However, while this still leaves some uncertainty, but expect that as access to seizure data increases, sensitivity rates can be improved. The algorithm can produce optimized features for everyone, using convolutional neural network that is highly attuned to observing changes in brain activity based on EEG readings.
3) Actually nearly 50 million people in world affected from the Epilepsy disease hence it is most chronological disorder. Now a days most of research has been started to avoid and diagnosis the Epilepsy. Here, based on both signal and image processing can predict along with the comparison of performance of different extraction method and classifiers, we can detect the epilepsy. To know the more information about the disorder in the brain we use the muti model approach which improves the specificity of the system as soon as possible to avoid fault detection.
4) [3] EEG signal has been developed by using short time Fourier transform (STFT) and convolution neural network (CNN). The actual research data and parameter setting are the 2 main components which verifies feasibility. As per the comparison the average accuracy of the signal is 86% by using support vector machine (SVN). The epileptic seizure detection will be stable and will perform high using this signal as it increases accuracy to 90%. The connection between EEG and epileptic seizure is based on the EEG signal has been a main approach. They are only a few applications based on deep learning they are EEG signal, re-processing of signals, Time-frequency analysis, these are the overall framework of the algorithm.

III. IMPLEMENTATION

Upload Epilepsy Dataset
Using this module, we will upload dataset to application

A. Dataset Pre-processing:

Dataset often contains missing values and non-numeric data such as patient ID so we need to process dataset to remove patient ID and missing values. Process data will be split into 80% training data and 20% testing data

B. Train Naive Bayes Algorithm:

Process data will be input to Naive Bayes algorithm to train a model

C. Predict Epilepsy from Test Data:

Using this module, we will upload new test data and then apply trained Naïve Bayes model to predict whether test data is normal or contains Epilepsy disease

D. Comparison Graph:

Using this module we will plot Naïve Bayes performance graph in terms of precision, recall, accuracy and FSCORE

IV. RESULT AND FUTURE WORK

In this project we are using machine learning Naïve Bayes algorithm to detect epilepsy disease as all machine learning algorithm require training using past data to build a trained prediction model and this trained model can be applied on new test data to predict whether patient is NORMAL or suffering from epilepsy disease.

To train Naïve Bayes we have used ECG Epilepsy dataset from KAGGLE below website URL https://www.kaggle.com/chaditya95/epileptic-seizures-dataset

Given figure is the home page of the detection system.

For future work, we propose to investigate the usage of state-of-the-art deep learning networks to overcome the limitations of classical learning models. Classical learning models are sensitive to the feature selection and extraction phases. Moreover, we propose to investigate the effect of the given dataset on the classification results. For this, we propose to examine different datasets from different regions of the world.

V. CONCLUSION

Epilepsy is one of the most diseases that affect human lives, so they need to diagnose; it is one of the lives needed. The diagnosis process is not a simple task. In this work, we propose a novel approach to diagnosis the EEG signals using Multi-DWT, and Genetic algorithm coupled with four classifiers such as SVM, ANN, KNN, and Naive Bayes. The experimental results showed that the DWT features coupled with some machine learning algorithms had provided noticeable results, and the ANN classifier outperforms all tested classifiers. The new automated system can detect epilepsy with high accuracy. The detection process of epilepsy seizure passes through different stages. The first step is the pre-processing of the EEG signals that are considered the primary step, which will increase the system performance. This step aims to remove the noises. The second step features extraction. This step is previously implemented with different methods; in this work, we apply multiple DWT. The main aim of this step is to decompose the signals into sub-bands, then compute different features functions on each sub-band. In our work, we use multiple DWT to extract various features, and then these features are reduced by the genetic algorithm to select the best features from a vast number of features. The output of this stage is a features matrix that will be used later in EEG signals classification. In EEG signals classification, the decision is made, and the system performance will be evaluated. The success of the suggested approach is verified by implementing the same procedure for 14 combinations of datasets. The proposed system was tested under different measurement metrics such as Accuracy, Sensitivity, and Specificity.

The results showed that our approach achieved good results in terms of these metrics, and it can be concluded that DWT analysis give satisfactory results compared with the previous studies, and the best performance was gained by artificial neural network classifier. The ANN was compared with the different classifiers, and it performs better in terms of the evaluation metrics in most cases of 14 dataset combinations.

VI. REFERENCES

[1] L. Hussain, W. Aziz, A. S. Khan, A. Q. Abbasi, and S. Z. Hassan, ‘‘Classification of electroencephlography (EEG) alcoholic and control subjects using machine learning ensemble methods,’’ J. Multidiscip. Eng. Sci. Technol., vol. 2, no. 1, pp. 126–131, Jan. 2015.

[2] Hamad, E. H. Houssein, A. E. Hassanien, and A. A. Fahmy, ‘‘Feature extraction of epilepsy EEG using discrete wavelet transform,’’ in Proc. 12th Int. Comput. Eng. Conf. (ICENCO), Dec. 2016, pp. 190–195.

[3] U. R. Acharya, S. Vinitha Sree, G. Swapna, R. J. Martis, and J. S. Suri, ‘‘Automated EEG analysis of epilepsy: A review,’’ Knowl.-Based Syst., vol. 45, pp. 147–165, Jun. 2013.

[4] P. Sarma, P. Tripathi, M. P. Sarma, and K. K. Sarma, ‘‘Pre-processing and feature extraction techniques for EEGBCI applications-a review of recent research,’’ ADBU J. Eng. Technol., vol. 5, no. 1, pp. 1–8, 2016.

[5] C. Umale, A. Vaidya, S. Shirude, and A. Raut, ‘‘Feature extraction techniques and classification algorithms for EEG signals to detect human stress-a review,’’ Int. J. Comput. Appl. Technol. Res., vol. 5, no. 1, pp. 8–14, Jan. 2016.

[6] K. Polat and S. Güneş, ‘‘Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals,’’ Expert Syst. Appl., vol. 34, no. 3, pp. 2039–2048, Apr. 2008.

[7] O. Salem, A. Naseem, and A. Mehaoua, ‘‘Epileptic seizure detection from eeg signal using discrete wavelet transform and ant colony classifier,’’ in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2014, pp. 3529–3534.

[8] S. A. Aljawarneh, V. Radhakrishna, and A. Cheruvu, ‘‘VRKSHA: A novel tree structure for time-profiled temporal association mining,’’ in Neural Computing and Applications. Cham, Switzerland: Springer, 2017, pp. 1–29. [Online]. Available: https://link.springer.com/article/10.1007%2Fs00521-018-3776-7
[9] S. A. Aljawarneh, R. Vangipuram, V. K. Puligadda, and J. Vinjamuri, ‘‘G-SPAMINE: An approach to discover temporal association patterns and trends in Internet of Things,’’ Future Gener. Comput. Syst., vol. 74, pp. 430–443, Sep. 2017.
[10] Radhakrishna, S. A. Aljawarneh, P. Veereswara Kumar, and V. Janaki, ‘‘ASTRA—A novel interest measure for unearthing latent temporal associations and trends through extending basic Gaussian membership function,’’ Multimedia Tools Appl., vol. 78, no. 4, pp. 4217–4265, Feb. 2019.
[11] Radhakrishna, S. A. Aljawarneh, V. Janaki, and P. Kumar, ‘‘Looking into the possibility for designing normal distribution based dissimilarity measure to discover time profiled association patterns,’’ in Proc. Int. Conf. Eng. MIS (ICEMIS), May 2017, pp. 1–5.
[12] S. M. Akareddy and P. K. Kulkarni, ‘‘EEG signal classification for epilepsy seizure detection using improved approximate entropy,’’ Int. J. Public Health Sci., vol. 2, no. 1, pp. 23–32, Feb. 2013.
[13] Baldominos and C. Ramon-Lozano, ‘‘Optimizing EEG energy-based seizure detection using genetic algorithms,’’ in Proc. IEEE Congr. Evol. Comput. (CEC), Jun. 2017, pp. 2338–2345.
[14] N. Williams, S. Zander, and G. Armitage, ‘‘A preliminary performance comparison of five machine learning algorithms for practical IP traffic flow classification,’’ ACM SIGCOMM Comput. Commun. Rev., vol. 36, no. 5, pp. 5–16, 2006.
[15] L. Guo, D. Rivero, and A. Pazos, ‘‘Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks,’’ J. Neurosci. Methods, vol. 193, no. 1, pp. 156–163, 2010.
[16] R. Moshrefi, M. G. Mahjani, and M. Jafarian, ‘‘Application of wavelet entropy in analysis of electrochemical noise for corrosion type identification,’’ Electrochem. Commun., vol. 48, pp. 49–51, Nov. 2014.
[17] D. Chen, S. Wan, J. Xiang, and F. S. Bao, ‘‘A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG,’’ PLoS ONE, vol. 12, no. 3, Mar. 2017, Art. no. e0173138.
[18] Sharmila and P. Geethanjali, ‘‘DWT based detection of epileptic seizure from eeg signals using naive Bayes and k-NN classifiers,’’ IEEE Access, vol. 4, pp. 7716–7727, 2016.
[19] S. Madan, K. Srivastava, A. Sharmila, and P. Mahalakshmi, ‘‘A case study on discrete wavelet transform based hurst exponent for epilepsy detection,’’ J. Med. Eng. Technol., vol. 42, no. 1, pp. 9–17, Jan. 2018.
[20] D. Selvathi and V. K. Meera, ‘‘Realization of epileptic seizure detection in EEG signal using wavelet transform and SVM classifier,’’ in Proc. Int. Conf. Signal Process. Commun. (ICSPC), Coimbatore, India, Jul. 2017, pp. 18–22, doi: 10.1109/cspc.2017.8305848.
[21] B. Harender and R. K. Sharma, ‘‘DWT based epileptic seizure detection from EEG signal using k-NN classifier,’’ in Proc. Int. Conf. Trends Electron. Informat. (ICEI), Tirunelveli, India, May 2017, pp. 762–765, doi: 10.1109/icoei.2017.8300806.
[22] S. Lahmiri and A. Shmuel, ‘‘Accurate classification of seizure and seizurefree intervals of intracranial EEG signals from epileptic patients,’’ IEEE Trans. Instrum. Meas., vol. 68, no. 3, pp. 791–796, Mar. 2019, doi: 10.1109/tim.2018.2855518.
[23] G. Wang, D. Ren, K. Li, D. Wang, M. Wang, and X. Yan, ‘‘EEG-based detection of epileptic seizures through the use of a directed transfer function method,’’ IEEE Access, vol. 6, pp. 47189–47198, 2018, doi: 10.1109/access.2018.2867008.
[24] Z. Zakeri, S. Assecondi, A. P. Bagshaw, and T. N. Arvanitis, ‘‘Influence of signal preprocessing on ICA-based EEG decomposition,’’ in Proc. 13th Medit. Conf. Med. Biol. Eng. Comput. Cham, Switzerland: Springer, 2013, pp. 734–737.
[25] Hamad, E. H. Houssein, A. E. Hassanien, and A. A. Fahmy, ‘‘A hybrid EEG signals classification approach based on grey wolf optimizer enhanced SVMs for epileptic detection,’’ in Proc. Int. Conf. Adv. Intell. Syst. Inform. Cham, Switzerland: Springer, Sep. 2017, pp. 108–117.
[26] Hamad, E. H. Houssein, A. E. Hassanien, and A. A. Fahmy, ‘‘Hybrid grasshopper optimization algorithm and support vector machines for automatic seizure detection in EEG signals,’’ in Proc. Int. Conf. Adv. Mach. Learn. Technol. Appl. Cham, Switzerland: Springer, Feb. 2018, pp. 82–91.
[27] M. Kołodziej, A. Majkowski, and R. J. Rak, ‘‘A new method of EEG classification for BCI with feature extraction based on higher order statistics of wavelet components and selection with genetic algorithms,’’ in Proc. Int. Conf. Adapt. Natural Comput. Algorithms. Berlin, Germany: Springer, Apr. 2011, pp. 280–289.

