Automatic Epileptic Seizure Detection With Time And Frequency Analysis On EEG

Islam Gabilayev and Yaşar Uz
Electrical-Electronics Engineering
Gazi University Ankara, Türkiye
igabila00@gmail.com and yasaruz@outlook.com

Assoc.Prof.Mehmet Feyzi AKŞAHİN
Electrical-Electronics Engineering
Gazi University Ankara, Türkiye
maksahin@gazi.edu.tr

***Abstract*—Epilepsy is a chronic symptom characterized by frequent, unexpected and severe seizures. Seizures can manifest as motor, sensory, autonomic, or mental abnormalities, depending on the type of epilepsy. Epileptic seizures are caused by sudden, temporary electrical damage to the brain and can be seen in electroencephalograms (EEGs) recorded by electrodes placed directly on the exposed surface of the brain. EEG is a method of recording the electrical activity of the brain generated by the firing of neurons in the brain. This makes it difficult to recognize the seizure and examine the brain. Machine learning classifiers can classify EEG data, identify seizures, and highlight related patterns without impacting performance. It is said to automatically detect epileptic seizures based on brain wave signals. In this article, EEG signals are analyzed in the time and frequency domain to extract features. we applied the Logistic Regression, Random Forest, Extreme Gradient Boosting Classifier, Multilayer Perceptron Classifier, and Extra Tree Classifier algorithms. According to the results, this research model can accurately predict epilepsy. The model uses the accuracy of this analysis, but it is different. This research model can accurately predict epilepsy. EEG signals are analyzed in the time and frequency domain to extract features. We aim for a correct answer rate of 90% or higher.**

Keywords— EEG; Seizure; Epilepsy; Machine Learning.

I. INTRODUCTION

Epilepsy may be a recurrent seizure caused by injury to brain cells. It will occur at any stage of human life and occurs intermittently or ofttimes throughout life. Some sorts of brain disorder solely seem at one stage of life, whereas others retain their effects for the remainder of their life. Once epilepsy occurs when bound events, corresponding to head trauma or meningitis, it's called symptomatic epilepsy. The sort of epilepsy that happens while not distinguishing the cause is named idiopathic epilepsy [1]. Excessive discharge of brain cells causes seizures. A loss or abnormality in physical function, such as loss of consciousness or excessive muscle activity, is ascertained throughout a seizure [4]. Epilepsy is a chronic neurological disorder marked by frequent, unexpectedly severe epileptic seizures. Seizures can appear as motor, sensory, autonomic, or mental abnormalities, depending on the kind of epilepsy. Attacks can be accompanied by varied degrees of violation of the condition. Some types of epilepsy are characterized solely or mostly by impaired awareness—absences with limited motor symptoms.[2]. Seizures can be global or localized (partial). With widespread seizures with epileptic activity in both hemispheres of the brain practically simultaneously. With focal, there are several focuses in the so-called epileptogenic zone. It is not restricted to this activity, in the circumstances of neighboring sections of the brain, creating the typical symptoms of an attack. Secondary generalized seizures can occur in some cases of focal epilepsy. [2]. The International League Against Epilepsy (ILAE) defines epilepsy in two ways. Epilepsy, according to the 2005 conceptual definition, is a brain illness defined by a persistent proclivity to epileptic seizures, as well as the neurobiological, cognitive, psychosocial, and social repercussions of this condition. [3]. Although it was originally widely held that persons with epilepsy should be treated by psychiatrists, epilepsy now belongs to the science of neurology. [6]. It is preferable that epileptologists be involved in the diagnosis and treatment of epilepsy, but because epileptologists are frequently insufficient, neurologists and psychiatrists are frequently involved in its diagnosis and treatment, unable to navigate all the nuances associated with this disease [7]. Seizures are caused by sudden, temporary electrical damage to the brain and can be seen in electroencephalograms (EEGs) through electrodes placed directly on the exposed surface of the brain. The various types of EEG are routine EEG, sleep EEG or non-sleep EEG, ambulatory EEG, and video telemetry. Throughout the study, the patient is asked to stay seated and to open and shut the patient' eyes as needed. n most circumstances, it is advised to take many deep breaths (known as hyperventilation) over a period of various minutes. At the end of therapy, a flashing light may be put nearby to see if this impacts the patient's brain activity. [5, 25]. A sleep EEG is performed while you are sleeping. It can be used to screen for sleep problems or if a standard EEG does not offer enough information. In certain situations, you may be required to stay up the night before the exam to guarantee that you can sleep throughout it. This is known as a sleep-deprived EEG [2]. An ambulatory EEG is when brain activity is recorded 24 hours a day, seven days a week for one or more days. The electrodes will be coupled to a compact, portable EEG recorder that may be worn on the patient's clothes. Video telemetry, often known as video EEG, is a sort of EEG in which a patient is filmed while an EEG recording is made. This can give further information about a patient's brain function. The procedure is normally carried out over a few days while the patient is housed in a specially designed hospital suite. With graphical record, brain activity is recorded twenty-four hours a day, seven days a week, one day or a lot of. Graphical record signals are classified in line with completely different frequency varies such as delta, theta, alpha, beta, and gamma [11- 14, 16].

Some scientists are using machine learning techniques to study epilepsy and its potential effects on humans. These scientists have investigated various aspects of the disorder, including its causes and cures. Recent studies have shown that brain wave signals can be used to diagnose epilepsy. This is a very important development as it helps to improve the accuracy of the diagnosis [8, 9, 17-21]. Some of these algorithm uses neural network classification and wavelet techniques [22-24].

We use differently from the literature, time domain, frequency domain, and time-frequency domain analyses together to extract features on EEG signals with several machine learning techniques to determine epileptic seizures.

II. Material and Methods

Its goal is to detect epileptic seizures from EEG data automatically. To extract characteristics, EEG data will be evaluated in the time and frequency domains. Machine learning or artificial neural networks will be used for classification.

A. *Dataset Description*

The epileptic seizure dataset used in this paper is from MIT. Recordings were gathered from 5 participants (2 males, ages 3-22; and 3 females, ages 10-19). For these recordings, the International 10-20 system of EEG electrode locations and naming was employed. It was decided to build the entire functioning data set using the FZ-CZ channels. Figure 1. shows proposed model of the epileptic seizure dataset.

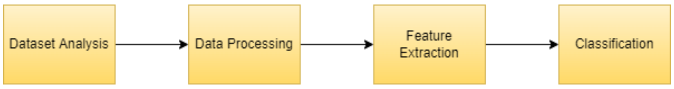

Figure 1. Proposed Model

B. *Data Processing*

1. Time Analysis

Data were divided into 10-second epochs in the form of windows. The necessary data set for other features was prepared. The required parameters for the input groups used here were determined. Analysis methods performed in the time domain feature listed in the feature extraction section.

2. Frequency Analysis

The Welch algorithm could be a non-parametric technique of calculating the PSD that produces a smoother frequency spectrum than the uncooked FFT result. Assume the FFT is computed for the whole sign period. Rather than strolling the FFT at some point of the general time domain, the Welch approach divides the sign into equal-sized windows. By clipping frequencies with intervals longer than the window, the window length affects the readability of the findings. Windowing takes samples from bigger facts set and tapers the sign up the interval's boundaries. This smoothest the sign and removes abrupt transitions that might intervene with the frequency spectrum display [26, 27]. To estimate the power spectrum, divide the time signal into successive blocks, create a periodogram for each block, and average it. Denote the **m** th windowed, zero-padded body from the sign **x** by

$$x_m(n) \triangleq \omega(n)x(n+mR), \mathrm{n} = 0,1\ldots,\mathrm{M}\text{-}1, \mathrm{m} = 0,1,\ldots,\mathrm{K}\text{-}1,$$

where R is defined as the window hop size and K denotes the number of frames available. Then the periodogram for the **m** th block is given by

$$\mathrm{P}_{x_m}, M(\omega_k) = \frac{1}{M}\left|\mathrm{FFT}_N, k(x_m)\right|^2 \triangleq \frac{1}{M}\left|\left(\sum_{n=0}^{N-1} x_m(n)e^{-j2\pi nk/N}\right)\right|^2$$

as before, the Welch estimate of the power spectral density is given by

$$\hat{S}_x^W(\omega_k) \triangleq \frac{1}{K}\sum_{m=0}^{K-1} P_{x_m}, M(\omega_k)$$

[10].

C. *Feature Extraction*

After analyzing the EEG data, Following the formation of feature teams, the input file required for each cluster, as well as the optimum method for retrieving that data, were identified.

The nonlinear function provides a metric that characterizes the chaotic behavior of the vibration signal. These functions are useful for analyzing vibration and acoustic signals from systems such as bearings, gears, and engines [15]. Generating nonlinear features is more computationally intensive than generating other features. In nonlinear feature Lempel-Ziv complexity, Detrended fluctuation, Katz fractal dimension, Petrosian fractal dimension, and Higuchi fractal dimension extracted.

Coming to the entropy features section, the extracted features are Shannon Entropy, Spectral Entropy, Singular Value Decomposition Entropy, Approximate Entropy, Sample Entropy, and Permutation Entropy. Entropy-based feature selection for data clustering utilizing the k-Means and k-Medoids algorithms. The clustering process separates a huge dataset into smaller subsets, each of which is referred to as a cluster. Each cluster has the same features as the others, yet each cluster is unique.

Time-frequency domain features are obtained from the spectrogram data, the obtained features are the mean, quantiles, and total variation. A spectrogram is an approximation of visually representing the strength or "loudness" of a signal at various frequencies found in a given waveform. It is used to show the difference in signal strength and loudness over time. It can also be used to indicate the frequency of a particular signal [18].

One of the aims of the project was to analyze EEG data in time, frequency and time-frequency domain. Time domain features are derived from raw EEG data. These features are RMS, Fano factor, Kurtosis, Skewness, Hjorth complexity, Hjorth mobility, Hjorth activity, Mean, Median, Standard deviation, Variance, Energy, Zero crossing rate, Peak amplitude, Peak to peak amplitude, Min amplitude, Total variation, Line length, and Quantile.

It is used in physics, engineering and applied mathematics for estimating the power of a signal at different frequencies. It's named after Peter D. Welch and is used to estimate the spectral density of signals.[26] Band powers, Standard deviation, Variance, Median, Mean, Quantile, Line, and Wavelet coefficients extracted from PSD data in the frequency domain features.

When this phase, the function utilized for all characteristics is specified in EEG raw data (time series data). For each epoch, all intended outputs are computed separately and adjusted to be employed inside ML model phase. In total, 19-time domain features, 8 frequency domain features, 3 time-frequency domain features, 5 non-linear and 6 entropy features, a total of 41 features were extracted.

D. Machine Learning Model and Algorithms used for Classification

Prior to model training, feature information sets were blended with training, validation, and test data in 0.6, 0.2, and 0.2 proportions, respectively. The most important of these pre-treatments is to reduce the standard deviation to one and the mean to zero by scaling each feature. After that, the cubic centimeter model was trained and assessed utilizing validation and data analysis. The ML models developed in this study, as well as brief information about them and the findings gained during the research phase, are displayed.

- Extra-Trees Classifier:

The Extremely Randomized Trees Classifier (Extra Trees Classifier) is a form of ensemble learning approach that combines the classification results of several de-correlated decision trees [28].

- Random Forest:

Random forests, also known as random decision forests, are an ensemble learning approach for classification, regression, and other problems that works by training by constructing diverse decision trees [32].

- Logistic Regression:

Logistic regression uses previous observations from a data set to predict a binary outcome such as yes or no. Logistic regression models estimate dependent data variables by examining correlations between one or more available independent variables [31].

- Decision Tree:

Decision Trees (DTs) are nonparametric supervised learning algorithms for classification and regression. The objective is to create a model that predicts the value of the target variable using basic decision rules generated from data attributes [19].

- Extreme Gradient Boosting Classifier:

Extreme gradient boosting algorithms a class of ensemble machine learning algorithms that can be used for classification or predictive regression modeling problems [29].

- Multi-layer Perceptron Classifier:

A Multi-Layer Perceptron (MLP) is a kind of artificial neural network that can generate a set of outputs from a set of inputs. MLP consists of several layers of input nodes connected by directed graphs between the input and output layers [30].

III. Results and Discussion

In this section, we analyze the results and classification performance of classifying the epilepsy dataset using multiple classifiers. The feature dataset was mixed with the training, validation, and test data at ratios of 0.6, 0.2, and 0.2, respectively. Feature databases have very high standard deviations because they contain features at very different scales. Therefore, model processing is required. Otherwise, the evaluation test results, and new data can be very misleading. The epoch times used in the tests are 10, 15, 20, 25, and 30 seconds. In addition to this test, after completing the data set, the epoch period that gives the best seizure detection results is determined. Tests performed on different patients with different models showed that 10 seconds was the epoch time where the model performed best 100% of the time. As a result of testing on many different patients. Figure 4. shows a comparison of different classification techniques. Here, we can see ETC is in first place with a score of 100%, followed by RF and MLP with scores of 97% and 95.5% respectively. The Extra-Trees Classifier has proven to be the best model for detecting epileptic seizures. Completing this test took 110 minutes, running on a computer with a 16-core Neural Engine hardware processor capable of 11 trillion operations per second.

Figure 3. Shows Confusion Matrix for ETC.

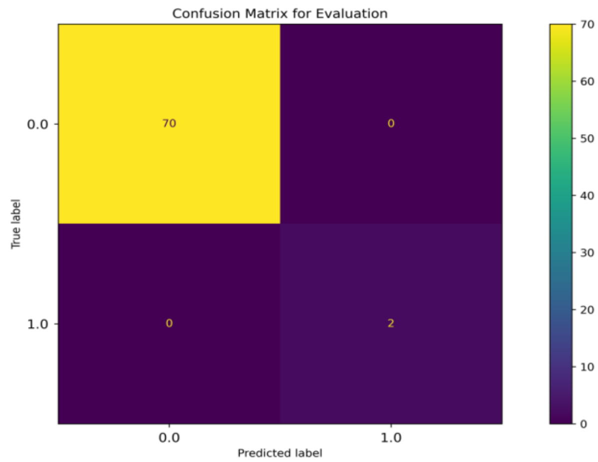

Figure 3. Confusion Matrix of ETC

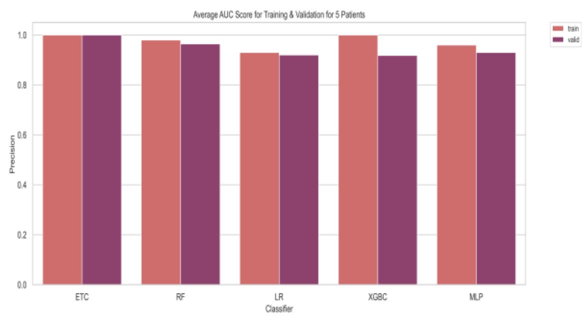

Figure 4. Comparison of the different Classification Techniques

IV. CONCLUSION

With the increase in seizures worldwide and the incredible impact they have on people, early detection and accurate diagnosis are more important than ever. The main goal of this work is to determine the most suitable and efficient ML classifiers for the diagnosis of seizures. This paper used non-tree classifier, random forest, linear regression, decision tree, extreme gradient boosting classifier and multi-level perceptron classification algorithms. The dataset shows that ETC (Extra Tree Classifier) provides better results with less computation and 100% accuracy when averaged over 5 patients. RF (Random Forest) also showed good results.

REFERENCES

[1] World Health Organization & Dekker, P. A. (2002). Epilepsy: a manual for medical and clinical officers in Africa. Epilepsy: a manual for medical and clinical officers in Africa. Retrieved 10 25, 2021, from Epilepsy: a manual for medical and clinical officers in Africa / PA Dekker

[2] S.B. Wilson, et al., Seizure detection: correlation of human experts, Clin. Neurophysiol. 114 (11) (2003) 2156–2164.

[3] Fisher, Robert S.; Boas, Walter van Emde; Blume, Warren; Elger, Christian; Genton, Pierre; Lee, Phillip; Engel, Jerome (2005). "Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)". Epilepsia. Wiley. 46 (4): 470—472. DOI:10.1111/j.0013-9580.2005.66104.x. ISSN 0013-9580

[4] Li, Y.; Wei, H.L.; Billings, S.A.; Liao, X.F. Time-varying linear and nonlinear parametric model for granger causality analysis. Phys. Rev. E 2012, 85, 041906.

[5] Ocak, H. Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy. Expert Syst. Appl. 2009, 36, 2027–2036.

[6] Gordeeva L., Difficult, expensive, but possible: How epilepsy is treated today // Medicinal Review. — 05/08/2014. — № 6.

[7] Semyonova E., A disease with a thousand names: 10 myths about epilepsy // Arguments and facts. - 23/06/2015.

[8] Li, Y.; Liu, Q.;Tan, S.R.; Chan, R.H.M. High-resolution time-frequency analysis of eeg signals using multiscale radial basis functions. Neurocomputing 2016, 195, 96–103.

[9] Li, Y.; Wei, H.L.; Billings, S.A.; Sarrigiannis, P.G. Time-varying model identification for time-frequency feature extraction from eeg data. J. Neurosci. Method. 2011, 196, 151–158.

[10] Julius O. Smith III, "Spectral Audio Signal Processing." W3K Publishing, 2011, https://ccrma.stanford.edu/~jos/sasp/Welch_s_Method.html.

[11] Craig, A., Nguyen, H., Tran, Y., & Wijesuriya, N. (2012, February 10). Regional brain wave activity changes associated with fatigue. Psychophysiology,4(49),574-582.https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-8986.2011.01329.x

[12] Klimesch W. (2012). α-band oscillations, attention, and controlled access to stored information. Trends in cognitive sciences, 16(12), 606–617. https://doi.org/10.1016/j.tics.2012.10.007

[13] Takahashi, K., Saleh, M., Penn, R. D., & Hatsopoulos, N. G. (2011). Propagating waves in the human motor cortex. Frontiers in human neuroscience, 5, 40. https://doi.org/10.3389/fnhum.2011.00040

[14] Electroencephalogram (EEG). (n.d.). NHS. Retrieved November 18, 2021, from https://www.nhs.uk/conditions/electroencephalogram/

[15] Ippolito, P. P. (2019). Feature Extraction Techniques - Towards Data Science. Retrieved, from https://towardsdatascience.com/feature-extraction-techniques-d619b56e31be, (Date of access: December 27, 2020)

[16] Tekin, R. & Kaya, Y. (2018). Epileptik EEG İşaretlerinin Sınıflandırılması İçin Yeni Bir Öznitelik Çıkarım Yöntemi . Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi , Cilt: 22 Sayı: Özel , 529-535. Retrieved from https://dergipark.org.tr/tr/pub/sdufenbed/issue/39695/470097

[17] Almustafa, K. M. (2020). Classification of epileptic seizure dataset using different machine learning algorithms. Informatics in Medicine Unlocked, 21, 100444.

[18] Srinivasan, V., Eswaran, C., Sriraam, N. 2005. Artificial neural network based epileptic detection using time-domain and frequency-domain features. Journal of Medical Systems, 29(6), 647-660.

[19] Polat, K., Güneş, S. 2007. Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Applied Mathematics and Computation, 187(2), 1017–1026.

[20] Subasi, A. 2007. EEG signal classification using wavelet feature extraction and a mixture of expert models. Expert Systems with Applications, 32(4), 1084-1093.

[21] Guo, L., Rivero, D., Dorado, J., Rabunal, J. R., Pazos, A. 2010. Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. Journal of neuroscience methods, 191(1), 101-109.

[22] Guo, L., Rivero, D., Dorado, J., Munteanu, C. R., Pazos, A. 2011. Automatic feature extraction using genetic programming: An application to epileptic EEG classification. Expert Systems with Applications, 38(8), 10425-10436.

[23] Gandhi, T. K., Chakraborty, P., Roy, G. G., Panigrahi, B. K. 2012. Discrete harmony search based expert model for epileptic seizure detection in electroencephalography. Expert Systems with Applications, 39(4), 4055-4062.

[24] Kaya, Y., Uyar, M., Tekin, R., Yıldırım, S. 2014. 1Dlocal binary pattern-based feature extraction for classification of epileptic EEG signals. Applied Mathematics and Computation, 243(2014), 209– 219.

[25] Using Virtual Reality to Examine the Correlation between Balance Function and Anxiety in Stance - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Frontal-channels-in-the-64-channel-EEG-head-cap-Here-green-color-denotes-electrode_fig1_339094908 [accessed 17 Dec, 2021]

[26] Parhi, K.K.; Ayinala, M. Low-Complexity Welch Power Spectral Density Computation. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2014**, *61*, 172–182. [**Google Scholar**] [**CrossRef**]

[27] Zou, P.; Lv, M. A fast algorithm for blind estimation of frequency domain parameters of Direct Sequence Spread Spectrum signal. In Proceedings of the 2009 IEEE International Conference on Communications Technology and Applications, Beijing, China, 16–18 October 2009; pp. 430–435. [**Google Scholar**] [**CrossRef**]

[28] Abhishek, L., 2020, June. Optical character recognition using ensemble of SVM, MLP and extra trees classifier. In 2020 International Conference for Emerging Technology (INCET) (pp. 1-4). IEEE.

[29] Wu, J., Zhou, T. and Li, T., 2020. Detecting epileptic seizures in EEG signals with complementary ensemble empirical mode decomposition and extreme gradient boosting. Entropy, 22(2), p.140.

[30] Prabhakar, S.K. and Rajaguru, H., 2016. Performance analysis of ApEn as a feature extraction technique and time delay neural networks, multi layer perceptron as post classifiers for the classification of epilepsy risk levels from EEG signals. In Computational Intelligence, Cyber Security and Computational Models (pp. 89-97). Springer, Singapore.

[31] Park, H.A., 2013. An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain. Journal of Korean Academy of Nursing, 43(2), pp.154-164.

[32] Verhoeven, T., Coito, A., van Mierlo, P., Seeck, M., Michel, C., Plomp, G., Dambre, J. and Vulliemoz, S., 2016. Using random forest for diagnosis and lateralization of temporal lobe epilepsy from EEG-based directed functional connectivity. In 12th European Congress on Epileptology (Vol. 57, p. 64). Wiley-Blackwell.