Generalized Epileptic Seizure Prediction using Machine Learning Method

Zarqa Altaf¹*, Mukhtiar Ali Unar², Sanam Narejo³, Muhammad Ahmed Zaki⁴, Naseer-u-Din⁵ Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan^{1, 2, 3, 5}

NED University of Engineering and Technology, Karachi, Sindh, Pakistan⁴

Abstract—In recent years, the electroencephalography (EEG) signal identification of epileptic seizures has developed into a routine procedure to determine epilepsy. Since physically identifying epileptic seizures by expert neurologists becomes a labor-intensive, time-consuming procedure that also produces several errors. Thus, efficient, and computerized detection of epileptic seizures is required. The disordered brain function that causes epileptic seizures can have an impact on a patient's condition. Epileptic seizures can be prevented by medicine with great success if they are predicted before they start. Electroencephalogram (EEG) signals are utilized to predict epileptic seizures by using machine learning algorithms and complex computational methodologies. Furthermore, two significant challenges that affect both expectancy time and genuine positive forecast rate are feature extraction from EEG signals and noise removal from EEG signals. As a result, we suggest a model that offers trustworthy preprocessing and feature extraction techniques. To automatically identify epileptic seizures, a variety of ensemble learning-based classifiers were utilized to extract frequency-based features from the EEG signal. Our algorithm offers a higher true positive rate and diagnoses epileptic episodes with enough foresight before they begin. On the scalp EEG CHB-MIT dataset on 24 subjects, this suggested framework detects the beginning of the preictal state, the state that occurs before a few minutes of the onset of the detention, resulting in an elevated true positive rate of (91%) than conventional methods and an optimum estimation time of 33 minutes and an average time of prediction is 23 minutes and 36 seconds. Depending on the experimental findings' The maximum accuracy, sensitivity, and specificity rates in this research were 91 %, 98%, and 84%.

Keywords—Epilepsy; electroencephalogram; artificial intelligence; machine learning; CHB-MIT

I. INTRODUCTION

A set of neurological illnesses known as epilepsy can afflict people of any age and are defined by a persistent propensity to cause repeated seizures. The progressive neurobiological process known as "epileptogenic" causes epilepsy [1]. The aberrant synchronized electrical activity of brain neurons is the primary cause of epilepsy, a persistent, non-communicable condition [2, 3]. The oldest and most prevalent neurological condition in the globe is epilepsy [4, 5]. Epilepsy is the third most prevalent neurological condition in the world, affecting 50 million individuals worldwide, based on a World Health Organization (WHO) study from June 2019 [6]-[10]. An abnormality of the brain characterized by recurrent seizures is called Epilepsy. Typically, a seizure is described as a sudden (abrupt) shift in behavior because of an

*Corresponding Author.

abnormal disturbance in the electrical activity of the human brain [11]. Some minute electrical impulses are continuously produced by the brain resulting in a consistent pattern. Neurotransmitters are the chemical signals which carry electrical signals along with neurons, and neural networks in the brain and throughout the entire body [12].

Fig. 1 illustrates how epilepsy causes the brain's electrical cycles to become unbalanced and cause recurring seizures. Individuals having seizures must face synchronized electrical energy bursts that may alter their cognition, movements, or perceptions and disturb the regular brain electrical sequence for a period. The main symptoms of epilepsy are varied and complex due to variations in the beginning location and method of propagation of aberrant electrical activity in the brain [13]. Recurrent seizures can have a long-lasting severe impact on a patient's psychological and cognitive abilities and pose a serious risk to their lives [14]. Investigation into the treatment and diagnosis of epilepsy, therefore, has huge therapeutic implications.

Epileptic seizures can be prevented by medication if they are predicted early, giving ample time before they happen. Four distinct states occur during epileptic seizures. The first state that emerges before the beginning of the seizure is the prodromal (pre-ictal) state, the second state i.e., the ictal state, starts with the exact beginning of the seizure and completes leaving a threat, after the ending of the ictal state comes the third state i.e., postictal state, and last is an interictal state, that begins after the postictal state of the first seizure and is finished before the begin of the preictal state of the subsequent seizure. The various input conditions for three distinct channels are depicted in Fig. 2. Additionally, the onset of the preictal state can be used to anticipate seizures [15].



Fig. 1. Epilepsy hotpot.



Fig. 2. Different input states of epileptic seizure [15].

The remainder of the essay has been structured as follows: The complaints are covered in Section II, the background of epilepsy seizures is covered in Section III, and the proposed technique is covered in Section IV. The experimental results are reported in Section V. Section VI brings the essay to a close and discusses unfinished business.

II. TYPES AND SYMPTOMS OF EPILEPTIC SEIZURE

Neurologically epilepsy is characterized by abnormal activity of the brain that results in seizures resulting in strange behavior, emotional sensations, and most of the time total loss of conscious [16]. When a person experiences at least 2 seizures that are not related to another established medical problem, such as opiate withdrawal or exceptionally low blood sugar, an epilepsy diagnosis is typically made [17]. That part of the brain from which the seizure frequently originates in early phases causes disturbance in functions of the affected part. The right side of the body is governed by the left half of the brain, while the left side of the body is governed by the right half of the brain. Typically, Doctors determine seizure as either generalized or focal depending on where and how the abnormal activity of the brain starts [18]. Focal seizures are caused by the aberrant activity of the brain in a specific part of the brain, while Generalized seizures appear to be involved in the entire brain [19]. Neuro-experts have divided seizures into two main groups, partial and generalized, depending on the signs, as depicted in Fig. 3 [20, 21]. The symptoms of a partial seizure, which are mostly brought on by damage to the cerebral hemisphere, can be utilized to define it. Additionally, there are two basic categories of partial seizures: simplepartial and complex-partial. In simple-partial, the person appears cognizant and can typically speak, whereas, in complex-partial, patients behave erratically, become disoriented, and frequently mumble and chew. A generalized seizure comprises two main components as well. While definitive seizures are challenging to detect because they lack motor signals, non-conclusive seizures can be identified by their clear motor symptoms. The person is unable to move or say anything other than to gaze [22, 23].



Fig. 3. Types of epileptic seizures.



Fig. 4. Symptoms of epileptic seizure [23].

Fig. 4 illustrates the wide range of seizure signs. Throughout a seizure, some patients just stare aimlessly for a specific period, while others continuously jerk their limbs or legs. One seizure may not necessarily indicate epilepsy. For an epileptic classification, at least 2 unprovoked seizures (seizures caused by unknown reasons) must be occurred within 24 hours away [24, 25]. Any brain-coordinated process can be disturbed by seizures since aberrant brain activity causes Epilepsy. Some specific symptoms determine the epilepsy type. Some of the below-mentioned sensations will be realized from time to time while others become consistent. Most of the time, an individual having epilepsy experiences the same type of seizure every other time. Symbols and Seizure indications may contain [26]:

- Brief uncertainty (confusion).
- Steady Eye spell.
- Rigid Body Movement.
- Uncontrollable spasmodic motion.
- Unawareness and incognizance.
- Spiritual indications such as fear or nervousness.

III. RELATED WORK

Early studies on Epilepsy prediction were conducted in the 1970s utilizing feature extraction methods that were linear [27]. Because of the non-linear character of EEG signals in the 1980s, researchers were able to apply these approaches for feature extraction thanks to the advent of non-linear methods [28, 29]. The utilization of the pre-ictal phase for epilepsy identification was also implemented in this decade with the diagnosis of the EEG patterns associated with epilepsy, including preictal, ictal, and interictal patterns. Salant et al. conducted early ES prediction almost 6 seconds before the seizure began in 1998 [30], and Drogenlen et al. 2003 expanded on this work [31]. They employed a feature called Kolmogorov entropy to forecast epilepsy 2 - 40 minutes before it began. The very first worldwide session on epilepsy forecasting took place in 2002, and several epilepsy facilities contributed a database of multi-day EEG recordings. Eventually, this database was the subject of other investigations [32]. Mormann et al. discovered in 2003 that the periodic synchronization of various EEG channels diminishes before seizure onset [33] using this theory that the hyper-synchronous discharge of the brain's neurons causes ES. Research studies on substantial EEG data have cast doubt on the accuracy of metrics computed in the past century during the first 10 years of the ongoing century. Some researchers discovered that these findings belonging to past studies were based on a limited number of carefully chosen data that could not be replicated on a large amount of previously unreported data. In worldwide workshops held on the subject, it was determined to hold contests on seizure prediction. These contests were created to make it easier to compare the effectiveness of algorithms that had been trained on the same dataset [34, 35]. The International Workshop on Seizure Prediction 3 (IWSP3) and the International Workshop on Seizure Prediction 4 (IWSP4) collaborated on the inaugural seizure prediction competition, which took place in 2007. The participants in both events received continuous iEEG recordings from 3 epilepsy patients. The algorithms' results obtained, however, fell short of expectations.

The 2014 American Epilepsy Society Seizure Forecasting Trial used long-term iEEG recordings of epileptic canines as well as short-term human iEEG containing 942 seizures acquired over more than 500 days. The same training and testing data, lasting 10 minutes, was given to each contestant. An evaluation metric for effectiveness was the Area Under the Curve (AUC). Another competition by Melbourne University with a similar format comprised long-term iEEG recording with 1139 seizures [36]. Any algorithm estimating the fundamental properties of EEG signals for epilepsy predictions or machine learning algorithms based on these basic properties was eligible for the competition. In any scenario, we are still unsure of the ideal characteristics or techniques. People entered algorithms that were excessively complex in the competitions. Therefore, it is challenging to determine which attribute or ML method was better. A novel solution presented by Maturana et al. [37] may be effective for a variety of patients. They determined that the crucial slowing of neural activity served as an ES prediction indication. Fig. 5 shows a timeline for the evolution of EEG data measurements. Readers who are interested in learning more about the background of these advances should consult [38] for additional details. Fig. 6 from Natu et al. [39] discussion on the development of technology for epileptic seizure detection.



Fig. 5. History of epileptic seizure prediction [38].



Fig. 6. Machine learning for generalized epileptic seizure prediction.

IV. MACHINE LEARNING FOR GENERALIZED EPILEPTIC SEIZURE PREDICTION

Since the turn of the century, scientists have been trying to get past the difficulties in diagnosing and predicting epilepsy. The initial emphasis of the ES forecast study was mainly on the evaluation of EEG recordings because EEG data are an important source to observe brain function before, throughout, and after epilepsy. Eye rotations, blinks, heart signals, and muscular noise contaminate EEG signals. To lessen the impact of these numerous sources of interference and distortions, a variety of filtration and noise reduction techniques are employed [40]. Substantial features are required for developing Machine learning models for the classification and identification of interictal and pre-ictal phases once artifacts have been removed. Fig. 7 illustrates the traditional Machine learning approach for epilepsy forecasting and emphasizes the key distinction between the application of Machine learning and Deep Learning methods.

A. Signal Processing

One important step in the analysis of raw biological signals is the identification of noise and artifacts. Filtering of these artifacts is required to lessen their impact on feature extraction. For filtering, a variety of methods have been used, including many filters such as Wavelet, Band-Pass, Finite Impulse response, and adaptive filters. Additionally, such processing is done to make the data standardized so that it may be compared to the records of other patients.

B. Feature Extraction and Collection

Reliable features are a requirement for all prediction models. These features can be divided into unilateral (steps undertaken on each EEG channel independently) and multimodal (measurements taken on two or more EEG channels) categories according to the quantity of EEG channels. Numerous techniques recommended in the literature were used to do the EEG study. As shown in Fig. 8, these methods were widely divided into 4 categories: frequency domain, time domain, nonlinear approaches, and timefrequency domain.

C. Classification

Artificial neural networks (ANN), fuzzy logic, k-means clustering, support vector machines (SVM), and decision trees are used to ensure the identification of epileptic seizures from provided EEG data. Most of the time, feature values with thresholds are used to draw inferences.



Fig. 7. Different input states of epileptic seizure [15].



Fig. 8. Classification of feature extraction approaches [42].

V. GENERALIZED EPILEPSY PREDICTION USING ML METHOD EXPERIMENTAL DETAILS

The primary goal of this research is to use computer vision algorithms to classify EEG signals as epileptic signals (preictal phase) or non-epileptic signals for the diagnosis of epilepsy. The signals in this are centered on ictal release for epileptic signals, whereas non-epileptic signals are consisting of both normal and pathological inter-ictal discharges for nonepileptic signals. The technique utilized to accomplish this is as follows:

1) EEG signal normalization and signal extraction.

2) To generate a feature collection, extract statistical features.

3) Apply wavelet decomposition to the signal to break it down.

4) To reduce the runtime, using k-means clustering for reducing the number of features in the feature set.

5) Training of the Support Vector Machine using the condensed feature set.

6) On a test data set, compare how well the SVM is based on the entire and modified feature set performed in separating epileptic from non-epileptic signals.



Fig. 9. Machine learning proposed model.

Fig. 9 displays the suggested method's block diagram. The gathering of EEG datasets, the pre-processing of these signals, feature extraction, and classification are the four primary stages of this approach. Below is a detailed explanation of these actions.

D. EEG Dataset Collection

The CHB-MIT database, which contains EEG recordings, was used in this investigation. All signals were captured and made accessible to the public by Boston Children's Hospital. Many recordings last an hour, but others go on for two or four hours. 24 portions of an EEG recording are separated and recorded in the EDF database schema. An EEG recording is represented by each EDF file. The CHB-MIT dataset signals include 686 EEG recordings from 23 people ranging in age from 1.5 to 17 years old. Each participant is represented by several EEG signals from various channels, and the dataset's sample frequency is 256 Hz. The Chb01 (1st subject) and Chb21 (second subject) in this database are the same individuals who were enrolled over 1.5 years. Information from the CHB-MIT dataset is shown in Table I.

Fig. 10 depicts a sample EEG signal from a person experiencing an epileptic seizure. In this image, the Seizure time window is indicated by red lines. Instances of (a) non-seizure records and (b) seizure records from the CHB-MIT database are shown in Fig. 11 [41]. Dataset information from EEGLAB executing in MATLAB is shown in Fig. 12.

TABLE I. CHB-MIT DATASET

Case No.	Gender	Age	No. of Seizure
1	F	11	7
2	М	9	3
3	М	61	7
4	F	14	4
5	М	45	5
6	М	3	10
7	F	13	3
8	М	76	5
9	F	36	4
10	М	55	7
11	F	6	3
12	F	14	27
13	М	44	10
14	М	4	8
15	F	6	3
16	М	6	6
17	F	5	3
18	М	9	8
19	F	8	4
20	F	27	3
21	М	23	7
22	F	33	16
23	F	7	7
24	М	21	4



Fig. 10. EEG signal with epilepsy seizure [40].



Fig. 11. CHB-MIT database: (a) Non-Seizure; (b) Seizure.

	EEGLA	3 v2022.	0					-	×
File	Edit	Tools	Plot	Study	Datasets	Help			Y
	—#'	I: ED	F fil	e—					
	Fi	lenam	e: .	se-1	.0.0\ch	b01\chb01_0	3.set		
	Ch	annel	s per	fram	e	23			
	Frames per epoch				921600				
	Epochs				1				
	Events				none				
	Sampling rate (Hz)				256				
	Epoch start (sec)				0.000				
	Epoch end (sec)				3599.99	6			
	Reference				unknown				
	Ch	annel	loca	ations		No (lab	els on	ly)	
	IC	A wei	ghts			No			
	Da	taset	size	e (Mb)		92.2			

Fig. 12. Dataset details from EEGLAB running in MATLAB.

E. Pre-Processing of EEG Signals

Preprocessing is the procedure of transforming raw data into a format that is more suitable for further analysis and interpretable for the user. In the case of EEG data, preprocessing usually refers to removing noise from the data to get closer to the true neural signals.

There are several reasons for preprocessing of EEG data is necessary. First, the signals that are picked up from the scalp are not necessarily an accurate representation of the signals originating from the brain, as the spatial information gets lost. Secondly, EEG data tends to contain a lot of noise which can obscure weaker EEG signals. Artifacts such as blinking, or muscle movement can contaminate the data and distort the picture. Finally, we want to separate the relevant neural signals from random neural activity that occurs during EEG recordings. Fig. 13 represents the preprocessing pipeline that is followed in this research.



Fig. 13. Preprocessing pipeline.

After following the preprocessing pipeline, we reduce the 23 channels into 8 channels which are:

- 1) FP1-F7
- 2) P3-O1
- 3) P4-O2
- 4) FP2-F8
- 5) P8-O2
- 6) FZ-CZ
- 7) CZ-PZ
- 8) P7-T7

The results of these 8 channels are shown in Fig. 14. The dataset with the following setting is shown in Fig. 15. Save the dataset as a '.set' file extension for the next step which is feature extraction.



Fig. 14. Visualization of 8-selected channels.

<u> </u>	EEGLA	3 v2022.0	0							-		\times
ile	Edit	Tools	Plot	Study	Datasets	Help	ERPLAB	ERPsets	Neurolode			1
_		±1 - s	lcha	nnel	EDE fi							
					201 11							
	F	lilena	me:	hes	is\8 Ch	annel	l datas	et.set				
	c	hanne	ls p	er fra	ume		8					
	E	rames	per	epoch	1		9	21600				
	Epochs				1							
	Events				r	none						
	Sampling rate (Hz)				2	56						
	E	poch	star	t (sec	:)			0.000				
	E	poch	end	(sec)			-	599.99	6			
	F	lefere	nce				ā	verage				
	c	hanne	1 10	catior	ıs		3	es				
	3	CA we	ight	S			3	es				
	I	atase	t si	ze (Mb)		-	6.9				

Fig. 15. Data of 8 selected channels.

F. Feature Extraction

To collect the abstract information required for the classification procedure at this point, feature extraction was used to remove the duplicate information from the EEG signals. When analyzing signals using wavelet transform, it's crucial to choose the right wavelets and the right number of layers of decomposition. The signal's prominent frequency components are used to determine the number of decomposition levels. The amount of decay is selected so that the wavelet coefficients preserve the frequencies necessary for the identification of the signal. The MATLAB software program was used to calculate the wavelet coefficients. In this study, we extract several features including the Fast Fourier transform, wavelet transforms, Mean, and Standard Deviation for alpha, beta, theta, delta, and gamma frequencies as shown in Fig. 16. 12 Features value extracted and saved as '.mat' for normal and epilepsy as shown in Fig. 17.



Fig. 16. Alpha, beta, and gamma features.

and the second se	C. Distances	And a second sec						 310 (11) (01) 	0012	10000	1		-	-
								0.01-0	M 10 -=		PRESSOR	SCENES.	~	
Culture - Me Publice -	Better Castle sets													
TextReture Pithen Idata St	the state of the state of the													
Canat . "Abstrice														
NU BELIEF NU														
中世間21+C+Ibes+4bro	1 Zali + Costog +	zona noti, + unobe test code:												
Datables - Sectore College														10 1
Frankare Tpillegrag 10														
SALAR EPOIN E														
1 2 3	4	A Charles and Charles	9	11 12	12	3.8	15	15	11	18	19	.20	23	
8 5400-00 2016-105 53435-45	0.464 0.0	THE CONST. 1.5574-05 12	en or 1777, ds 1217-0	4/09/04/24 7:5483/4-01										
2 121254-06 242254-06 182954-06	0.548 0.0	125 00176 UNITS OF 52	428K D4 8.5255K D4 1.2185K -02	13/0710-04 120401-04										
8 24530c-00 83209c-05 10062c-05	0.1405 0.0	812 DEBEN 5,53274 CS 43	0060 04 3,3929c 04 7,5608z 04	250Mz-04 44156-05										
4 E4062e-00 65480e-04 12955e-04	0.1757 0.0	EST COSTN 14748e OK AS	File OS ETASIE ON ETTATE-S	174820-06 192850-00										
5 A1512e-05 SAG2e-05 13907e-05	0.1668 0.0	THE DOUBLE LINUE OF AL	rile os 1.0218e os 1.1963e-3	2/2018-34 6.0258-31										
8 219/5e-05 1.192/e-05 2.897/e-01	0.1101 0.0	198 00017 2397ae-01 20	stelle of \$280m-of \$550m-of	\$25We-32 2828e-32										
7 389/0e+00 0.7/0e+00 0.7/2/e+00	01409 0.0	NO DOUTS 23/29-01 11	tarin associa anne-b	23948-01 12028-01										
# 200/terit) 8.033/er01 1.010er01	07/99 00	RA COURT 234538-01 90	useon som sstate->	/905/e=00 18032e=00										
\$ 243/erth 395/erth 5.00erth	078-75 000	500 COURT 15/59e-01 11	North Editoria 1211ers)	50350e=07 28800e=03										
10 1380a-lth 6188a-45 1811a-lth	40 5660	N/ 00108 45284-0/ 53	10+04 2028e-05 7.252e-5	2.05/e-01.58/Se-01										
18 2130auth 5136auth 5457auth	0.945 0.0	BYL 00404 2,5825+04 3.1	Sileds 1.853edt 8577eed	2010kau04 (15555eu0)										
17 2007auto 2005auto 1207auto	0.1146 0.0	IN COURT UNDER 22	Electe 12402-on 1485-or	2404e=04 33104e=0)										
15: 4205r-05 4247r-84 5357r-84	0.165 0.0	THE CONDE 22/37%-05 M	then anther tables	5.795.9=31 235359=33										
4 42005+05 1720+05 L4661+05	0.485 0.0	111 02164 1.90%-04 1J	HIN-ON TANDA ON ADMINIO	1.546 (a=54) #377 (a=33										
15 331 (kr-85 1256) + 66 1/267-46	0.108 0.0	171 DOICH 4/103/04 7/	toke on 2.0116r on 2.4717e-b	270252-04 150024-04										
16 UN25c-05 33545c-05 1240k-05	0.039 0.0	488 0.0409 2.7289x 05 L	iste os i salate de salate de	12200u-34 85834u-33										
17 83540c-00 84254c-00 10487c-00	0.560 0.0	THE DOWN LEPTING OF L	WHE DA LEASTE DE ASTRO-X	500We-bit 18726-bi										
TE 2204140 12404140 13104140	0.1898 0.0	RSE 0.0018 338294-05 14	FIRM LITHERS 2400-0	181426-04 821818-00										
19. 034624-05. 122458-05. 1611/4-05	0.1437 0.0	60 0000 200108-00 13	NTH-OK \$1/998-05 110008-02	3 2008 2-24 6-4822 -20										
25 3.43418+40 2.65638+64 1.34538+64	0.1979 0.0	77 DOMS 40270e-01 23	0800 208200 107e0	5-5357e-52 15657e-53										
21 6580e+ID 20031e+ID 20128+01	0.7967 20	10 10-10-100 T	STRID LIFERO JANA -O	1502fe=04 2,2883e=03										
22 BADSeeto 22901eeto 22201eeto	0.753/ 0.0	IDJ COURS TOROARDS TU	ble-of datafe-7 5325e-7	1/52/8-01 259808-03										
 124/benih 185/benih 282/benih 	0.5412 0.0	000ef 2,300e-04 31	08+05 35/Fie-01 62018+-7	1.2,989+-36 332/(+-3)										
28 1825auth 2282auth 1293auth	0.9825 0.6	00 0001 22096-04 10	Re-of 11978-01 83879-0	(1/055e=04 55105e=0)										
PL 22001ex00134525ex85121548ex051	0.1945 0.0	NF 0/383 1/386-04 32	Kite-Cit AAR/Se-Cit BORPER-D	E-MDaroft SMIDer01										
26 64Hierd5 44Kthrot6 54Minute	0.1778 0.0	10 0/107 17HD-04 1/	Read 1044-05 Cold tool	61264-01 13060-01										
1 1 Can AC 1 CAN AC 3 C. C. P.		and sheet of all	The same of the P	ACA. IS COMP. IL										

Fig. 17. Feature extraction.

G. Classification

The goal of the model was to determine the most effective dimensionality reduction method that, when combined with SVM, would provide the maximum degree of sensitivity and validity for gathering statistics as either epileptic or not. In a high-dimensional space, the support vector machine (SVM) creates a hyperplane or series of hyperplanes that can be utilized for classification. SVM has been demonstrated to be a useful supervised model based on a statistical learning tool with high generalization. The principle underlying SVM is the separation of two data sets. This separation can be linear or non-linear. In the case of linear separation, SVM uses a discriminant hyperplane to distinguish classes. However, in the case of nonlinear separation, SVM uses the kernel function to identify decision boundaries. Compared with that of other supervised algorithms, such as ANNs [42, 43] and KNN, the computational complexity of SVM is low [44]-[46].

In this study, the model of all data of 23 people is used for training each time, and data of the remaining 1 subject is used for the test. We explore multiple training, validation, and testing divisions of the dataset to see the effect on the performance achieved on these subsets. With an increase in training data as compared to testing data, an increase in performance for accuracy and sensitivity is observed. In our experimentation, a train-validation-test ratio of (70%-20%-10%) is followed. This ratio resulted in a total of 50 epochs for training. Fig. 18 shows the Epileptic seizure detection training, validation, and testing of 24 patients.



Fig. 18. Epileptic seizure detection training, validation, and testing.

VI. RESULTS

The clinical employment of ES prediction methods requires a sufficient performance and quality check and different evaluation metrics have been discussed in this section. Our end goal is to classify data into two classes nonseizure and seizure. To measure the performance of the proposed method, a confusion matrix, shown in Fig. 19 is obtained. In this table, (TP) represents true positive (epileptic region predicted as epileptic), TN represents true negative (non-epileptic region predicted as non-epileptic), FP represents false positive (non-epileptic region predicted as epileptic), and FN represents false negative (epileptic region predicted as non-epileptic).

One main challenge in classifying seizure data is the imbalance of the dataset. This comes from the fact that seizures (and so preictal data) do not occur frequently and the size of the interictal class is much larger than the preictal class. This may cause naive classification, which means that the classifier labels all the data as interictal and completely ignores the other class, and still reports a good precision. To avoid this, we propose a few contingency plans. First, we do not rely only on accuracy as the main factor to choose the best classifier. More informative factors can be sensitivity and specificity. Here, accuracy is the correct classification rate, sensitivity is the proportion of the epileptic regions that are correctly classified and specificity is the proportion of the non-epileptic regions that are correctly classified. Sensitivity is defined as the ratio of the total number of true positives (TP) to the sum of the total number of true positives and false negatives (FN). True positive is defined as the detection of a seizure in a segment which is also identified as a seizure segment by experts. Whereas false negatives represent a seizure segment not being classified as so by the algorithm, while the segment is identified as a seizure segment by experts. Specificity is defined as the ratio of the total number of true negatives (TN) to the sum of the total number of true negatives and false positives (FP). True negative is defined as the detection of a non-seizure segment which is also identified as a non-seizure segment by experts. Whereas false positives represent a seizure segment being classified by the algorithm, while the segment is identified to be a non-seizure segment by experts Accuracy is defined as the ratio of the sum of TP and TN to the sum of TP, TN, FP, and FN. Hence, the higher the value, the better the performance is achieved. Classification results are shown in Table II.

Measure	Value	Derivations
Sensitivity	0.980	TPR = TP / (TP + FN)
Specificity	0.840	SPC = TN / (FP + TN)
Precision	0.859	PPV = TP / (TP + FP)
Negative Predictive Value	0.976	NPV = TN / (TN + FN)
False Positive Rate	0.160	FPR = FP / (FP + TN)
False Discovery Rate	0.140	FDR = FP / (FP + TP)
False Negative Rate	0.020	FNR = FN / (FN + TP)
Accuracy	0.910	ACC = (TP + TN) / (P + N)
F1 Score	0.915	F1 = 2TP / (2TP + FP + FN)

TABLE II. CLASSIFICATION RESULTS

In Fig. 19, the first two diagonal cells show the number and percentage of correct classifications by the trained network. For example, 98 recordings are correctly classified as benign. This corresponds to 49.0% of all 200 recordings. Similarly, 84 cases are correctly classified as malignant. This corresponds to 42.0% of all recordings. 16 of the malignant recordings are incorrectly classified as benign and this corresponds to 8.0% of all 200 recordings in the data. Similarly, 2 of the benign recordings are incorrectly classified as malignant and this corresponds to 1.0% of all data. Out of 114 benign predictions, 86.0% are correct and 14.0% are wrong. Out of 86 malignant predictions, 97.7% are correct and 2.3% are wrong. Out of 100 benign cases, 98.0% are correctly

predicted as benign and 2.0% are predicted as malignant. Out of 100 malignant cases, 84.0% are correctly classified as malignant and 16.0% are classified as benign. Overall, 91.0% of the predictions are correct and 9.0 % are wrong. Considering the proposed approach system's specificity (SP), sensitivity (SN), and accuracy (AC) allows for an evaluation of its performance. The proportion of the total number of true positives (TP) to the total number of false negatives and true positives is known as the sensitivity (FN). True positive is the identification of a seizure in a part that is also recognized by professionals as a seizure section. False negatives, on the other hand, refer to a seizure segment that is recognized as a seizure segment by specialists but is not classified as such by the algorithm. Therefore, better performance is obtained as the greater the value. The classification time of SVM is shown in Table III.



Fig. 19. Confusion matrix.

Classifier	Training Time (Sec)	Testing Time (Sec)			
Support Vector Machine (SVM)	0.150	0.050			

VII. CONCLUSION

The automatic approaches for detecting epileptic seizures have been suggested in this paper. Data from CHB MIT were utilized to detect seizure events. An SVM classifier was used for classification, and maximum accuracy of 90.7% was attained. Training of the classification algorithm was carried out across patients to assess the effectiveness of the suggested method, and the experimental findings were as a result. The maximum accuracy, sensitivity, and specificity rates in this research were 91.0%, 98%, and 84% correspondingly as shown in Fig. 20.



Fig. 20. Classification result.

REFERENCES

- [1] Sanjay N Rakhade and Frances E Jensen. Epileptogenesis in the immature brain: emerging mechanisms. Nature Reviews Neurology, 5(7):380, 2009.
- [2] A.S. Zandi, M. Javidan, G.A. Dumont, R. Tafreshi, Automated real-time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform, IEEE Trans. Biomed. Eng. 57 (7) 1639–1651, 2010.
- [3] L.S. Vidyaratne, K.M. Iftekharuddin, Real-time epileptic seizure detection using EEG, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (11), 2146–2156, 2017.
- [4] Wang, L.; Xue, W.; Li, Y.; Luo, M.; Huang, J.; Cui, W.; Huang, C. Automatic Epileptic Seizure Detection in EEG Signals Using Multi-Domain Feature Extraction and Nonlinear Analysis. Entropy 2017, 19, 222. https://doi.org/10.3390/e19060222.
- [5] Alotaiby, T.N., Alshebeili, S.A., Alshawi, T. et al. EEG seizure detection and prediction algorithms: a survey. EURASIP J. Adv. Signal Process. 2014, 183 (2014). https://doi.org/10.1186/1687-6180-2014-183.
- [6] H O Lekshmy, Dhanyalaxmi Panickar and Sandhya Harikumar, "Comparative analysis of multiple machine learning algorithms for epileptic seizure prediction", Journal of Physics: Conference Series, 2022.
- [7] World Health Organization, Epilepsy: A Public Health Imperative. Geneva, Switzerland: WHO, 2019.
- [8] E. Alickovic, J. Kevric, and A. Subasi, "Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction," Biomedical signal processing and control, vol. 39, pp. 94– 102, 2018.
- [9] Zack M M and Kobau R, National and state estimates of the numbers of adults and children with active epilepsy—United States, 2015 MMWR Morb. Mortal. Wkly. Rep. 66 821, 2017.
- [10] Mary Jane England, Catharyn T Liverman, Andrea M Schultz, and Larisa M Strawbridge. Epilepsy across the spectrum: Promoting health and understanding.: A summary of the institute of medicine report. Epilepsy & Behavior, 25(2):266–276, 2012.
- [11] L. E. Hebert, P. A. Scherr, J. L. Bienias, D. A. Bennett, and D. A. Evans, "Alzheimer disease in the US population: prevalence estimates using the 2000 census," JAMA Neurology, vol. 60, no. 8, pp. 1119–1122, 2003.
- [12] J. Xiang, C. Li, H. Li, R. Cao, B. Wang, X. Han, J. Chen, the detection of epileptic seizure signals based on fuzzy entropy, J. Neurosci. Methods 243, 18–25, 2015.
- [13] D.C. Bergen, do seizures harm the brain? Epilepsy Curr. 6 (4), 117–118, 2006.

- [14] Usman SM, Usman M, Fong S. Epileptic Seizures Prediction Using Machine Learning Methods. Comput Math Methods Med. 2017.
- [15] Hart, Y. M. Epidemiology, natural history and classification of epilepsy. Medicine, 40(9), 471-476, 2012.
- [16] Yousefi, M. R., Golnejad, S., & Hosseini, M. M. (2022). Comparison of EEG based epilepsy diagnosis using neural networks and wavelet transform. arXiv preprint arXiv:2204.04488.
- [17] Mattar, E. Application of Machine Learning for Electroencephalography Classification of Epilepsy Seizure Prediction, 2022.
- [18] G. Alarc´on and A. Valent´ın, Introduction to Epilepsy, Cambridge University Press, Cambridge, UK, 2012.
- [19] D. J. /urman, E. Beghi, C. E. Begley et al., "Standards for epidemiologic studies and surveillance of epilepsy," Epilepsia, vol. 52, pp. 2–26, 2011.
- [20] R. S. Fisher, "/e new classification of seizures by the International League against Epilepsy 2017," Current Neurology and Neuroscience Reports, vol. 17, pp. 48–56, 2017.
- [21] Ahmad, I., Wang, X., Zhu, M., Wang, C., Pi, Y., Khan, J. A., ... & Li, G. EEG-based epileptic seizure detection via machine/deep learning approaches: A Systematic Review. *Computational Intelligence and Neuroscience*, 2022.
- [22] Namazi, H., Kulish, V. V., Hussaini, J., Hussaini, J., Delaviz, A., Delaviz, F., & Ramezanpoor, S., A signal processing based analysis and prediction of seizure onset in patients with epilepsy. Oncotarget, 7(1), 342, 2016.
- [23] Annuar, R. M. H. R. K., Shahbudin, S., Kassim, M., & Rahman, F. Y. A., Epilepsy Seizure Detection and Classification Analysis using Residual Neural Network. 11th International Conference on System Engineering and Technology (ICSET) (pp. 347-351), 2021.
- [24] Valeta, T. Epileptic Seizures. In The Epilepsy Book: A Companion for Patients (pp. 23-32), 2017.
- [25] SS Viglione and GO Walsh. Proceedings: Epileptic seizure prediction. Electroencephalography and clinical neurophysiology, 39(4):435–436, 1975.
- [26] Z Rogowski, I Gath, and E Bental. On the prediction of epileptic seizures. Biological cybernetics, 42(1):9–15, 1981.
- [27] Mary Ann F Harrison, Ivan Osorio, Mark G Frei, Srividhya Asuri, and Ying-Cheng Lai. Correlation dimension and integral do not predict epileptic seizures. Chaos: An Interdisciplinary Journal of Nonlinear Science, 15(3):033106, 2005.
- [28] Y Salant, I Gath, and O Henriksen. Prediction of epileptic seizures from two-channel eeg. Medical and Biological Engineering and Computing, 36(5):549–556, 1998.
- [29] Wim van Drongelen, Sujatha Nayak, David M Frim, Michael H Kohrman, Vernon L Towle, Hyong C Lee, Arnetta B McGee, Maria S Chico, and Kurt E Hecox. Seizure anticipation in pediatric epilepsy: use of kolmogorov entropy. Pediatric neurology, 29(3):207–213, 2003.
- [30] Klaus Lehnertz and Brian Litt. The first international collaborative workshop on seizure prediction: summary and data description. Clinical neurophysiology, 116(3):493–505, 2005.
- [31] Florian Mormann, Thomas Kreuz, Ralph G Andrzejak, Peter David, Klaus Lehnertz, and Christian E Elger. Epileptic seizures are preceded by a decrease in synchronization. Epilepsy research, 53(3):173–185, 2003.
- [32] Rasheed, K., Qayyum, A., Qadir, J., Sivathamboo, S., Kwan, P., Kuhlmann, L., ... & Razi, A.. Machine learning for predicting epileptic

seizures using EEG signals: A review. IEEE Reviews in Biomedical Engineering, 14, 139-155, 2020.

- [33] Andhale, P., & Patil, D. (2022). Machine Learning Approach for Predicting Epileptic Seizures using EEG Signals: A Review. Machine Learning Approach for Predicting Epileptic Seizures using EEG Signals: A Review (February 25, 2022).
- [34] Levin Kuhlmann, Philippa Karoly, Dean R Freestone, Benjamin H Brinkmann, Andriy Temko, Alexandre Barachant, Feng Li, Gilberto Titericz Jr, Brian W Lang, Daniel Lavery, et al. Epilepsyecosystem. org: crowd-sourcing reproducible seizure prediction with long-term human intracranial eeg. Brain, 141(9):2619–2630, 2018.
- [35] Matias I Maturana, Christian Meisel, Katrina Dell, Philippa J Karoly, Wendyl D'Souza, David B Grayden, Anthony N Burkitt, Premysl Jiruska, Jan Kudlacek, Jaroslav Hlinka, et al. Critical slowing as a biomarker for seizure susceptibility. bioRxiv, page 689893, 2019.
- [36] Editor: Bjorn Schelter, Editor: Jens Timmer, and Schulze-Bonhag. Seizure Prediction in Epilepsy: From Basic Mechanisms to Clinical Applicati. John Wiley & Sons, 2008.
- [37] Natu, M., Bachute, M., Gite, S., Kotecha, K., & Vidyarthi, A., Review on epileptic seizure prediction: machine learning and deep learning approaches. *Computational and Mathematical Methods in Medicine*, 2022.
- [38] M. P. Hosseini, T. X. Tran, D. Pompili, K. Elisevich, and H. S. Zadeh, "Multimodal data analysis of epileptic EEG and rs-fMRI via deep learning and edge computing", Artificial Intelligence in Medicine, Vol. 104, p. 101813, 2020.
- [39] H. Khan, L. Marcuse, M. Fields, K. Swann, and B. Yener, "Focal onset seizure prediction using convolutional networks", IEEE Transactions on Biomedical Engineering, Vol. 65, No. 9, pp. 2109-2118, 2017.
- [40] N. M POUR, I.Y ÖZBEK, "Epileptic Seizure Detection based on EEG Signal using Boosting Classifiers" Erzincan University Journal of Science and Technology, 14(1), 159-167 2021.
- [41] Prasanna, J., Subathra, M. S. P., Mohammed, M. A., Damaševičius, R., Sairamya, N. J., & George, S. T., Automated Epileptic Seizure Detection in Pediatric Subjects of CHB-MIT EEG Database—A Survey. *Journal* of Personalized Medicine, 11(10), 1028, 2021.
- [42] U. Zaki, M. A. Zaki, and S. Narejo, "Chest X-Ray Image-based COVID-19 Recognition using Modified Artificial Neural Network," 3rd International Conference on Computational Sciences and Technologies, MUET Jamshoro, 2022.
- [43] M. A. Zaki, S. Narejo, M. Ahsan, S. Zai, M. R. Anjum, and N. U. Din, "Image-based Onion Disease (Purple Blotch) Detection using Deep Convolutional Neural Network," Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 5, pp. 448–458, 2021.
- [44] Kulsoom, F., Narejo, S., Mehmood, Z. et al. A review of machine learning-based human activity recognition for diverse applications. Neural Comput & Applic 34, 18289–18324 (2022).
- [45] S. Imtiaz, S. -F. Horchidan, Z. Abbas, M. Arsalan, H. N. Chaudhry and V. Vlassov, "Privacy Preserving Time-Series Forecasting of User Health Data Streams," 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 2020, pp. 3428-3437, doi: 10.1109/BigData50022.2020.9378186.
- [46] Narejo, Sanam; Pasero, Eros; Kulsoom, Farzana (2016). EEG Based Eye State Classification using Deep Belief Network and Stacked AutoEncoder. In: International Journal of Electrical and Computer Engineering, vol. 6 n. 6, pp. 3131-3141. - ISSN 2088-8708.