A Project Report On

Automated detection of schizophrenia using nonlinear signal processing methods

Submitted for Summer Internship (CS7492) of 6th semester for partial fulfillment of the requirements for the award of the degree of **Bachelors of Technology**

Computer Science and Engineering

Submitted by:

Rishita Singh(1906022)

Aradhana Kumari(1906033)

Lovely Kumari (1906034)

Under the Supervision of:

Dr.Rajib Ghosh

Assistant Professor Grade-I
Computer Science and Engineering

National Institute of Technology Patna

Department of Computer Science and Engineering

May-July, 2022

राष्ट्रीय प्रौद्योगिकी संस्थान पटना
NATIONAL INSTITUTE OF TECHNOLOGY PATNA

DECLARATION

We hereby declare that this project work for Industrial Training at NIT Patna entitled **"Automated detection of schizophrenia disease using nonlinear signal processing methods"** has been carried out by us in the Department of Computer Science and Engineering under the supervision of **Dr. Rajib Ghosh, Assistant Professor Grade-I, Department of Computer Science and Engineering, NIT Patna**. No part of this project has been submitted for the award degree or diploma to any other Institute.

Name	**Roll**	**Signature**
Rishita Singh	1906022	
Aradhana Kumari	1906033	
Lovely Kumari	1906034	

ACKNOWLEDGEMENT

We hereby take the privilege to express our gratitude to all the people who were directly or indirectly involved in the execution of this work, without whom this project would not have been a success.

We extend our deep gratitude, respect and obligation to our project supervisor, **Dr. Rajib Ghosh**, Assistant Professor Grade-I, Department of Computer Science and Engineering, for her timely suggestions and encouragement.

We further express our gratitude to the Head of the Computer Science and Engineering Department, **Dr. M.P. Singh** for being a constant source of inspiration.

Rishita Singh(1906022)

Aradhana Kumari(1906033)

Lovely Kumari (1906034)

CONTENTS

ABSTRACT

Industrial Training project focused on Automated detection of schizophrenia disease using nonlinear signal processing methods.A disorder that affects a person's ability to think, feel and behave clearly .The exact cause of schizophrenia isn't known, but a combination of genetics, environment and altered brain chemistry and structure may play a role.

The composition of the brain is complex. Therefore, explaining and understanding the brain is a challenge. However, the study of the brain has proved useful for the prevention, detection, treatment and behavioral understanding of neurodegenerative diseases. In order to measure brain activity, many techniques have been used. one of the most widely used method is electroencephalogram (EEG). These technologies play an important role in understanding the risk factors and neurological basis of diseases.

Deep learning methods have been widely used in the field of neuroimaging and analysis of large data. In diagnostic imaging, the classification and prediction models of Schizophrenia disease staging have been extensively studied.

In this paper, a deep learning model has been proposed which was trained and teste on EEG signals of 14 patientsS . This way proposed approach protects our e-health management system from different types of security attacks such as Replay,Man in the Middle Man attacks, e.t.c.

OBJECTIVE

The objective of our project is to investigate and classify the EEG signal patterns into normal and schizophrenia classes. CNN is used for feature extraction from EEG signals which was then implemented to mine 157 features from each EEG pattern,from which 14 of the principal features were identified based on significance. The proposed model involves the extraction of nonlinear features from signals, t-test based feature selection, and performance validation of the different classifiers .The model is employed to evaluate a 19-channel EEG signal collected from normal and schizophrenia class subjects.

INTRODUCTION

Schizophrenia is a severe mental disorder which is characterized by a wide range of unusual behaviors: hearing voices (hallucinations) and distorted or false perception, often bizarre beliefs. They are unable to distinguish between reality and imaginative events. These unusual experiences seem real to the person whereas others assume that the person is lost in their own world.The reason for SZ is not fully understood, though most research has demonstrated that the brain's structural and functional abnormalities play a role in its creation [4]. According to the World Health Organization (WHO) reports, nearly 21 million individuals suffer from such a brain disorder worldwide. The average age starting to get affected by this disorder is in youth age; in males 18 years old, and females 25 years old, and it is more prevalent among males.

EEG is one of the most practical and inexpensive functional neuroimaging modalities, specifically capturing specialist physicians' interests. In this modality, the brain's electrical activities are recorded from the head surface with a high temporal resolution and an appropriate spatial resolution, which is influential in SZ diagnosis [17]. In addition to the mentioned merits, EEG signals regularly have various channels recorded in long-term [17]. In some cases, these reasons make specialist physicians face serious challenges in SZ diagnosis via EEG signals

Automated SZ diagnosis via EEG signals using Deep Learning methods includes preprocessing sections, features extraction and selection, and in the end, classification. Feature extraction is the most important part of SZ diagnosis via EEG signals. The extracted features from EEG signals are mainly categorized into four groups: time , frequency , time- frequency , and non-linear fields.

PROBLEM STATEMENT

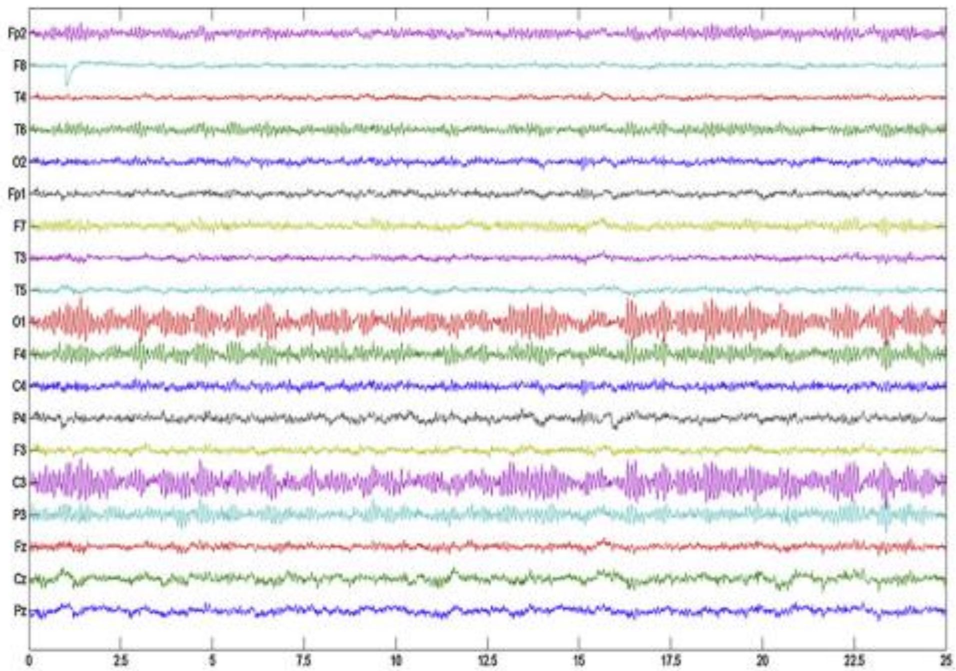
The objective of our project is to investigate and classify the EEG signal patterns into normal and schizophrenia classes. Malfunction of the brain by disease or disorder affects normal activity in humans. Schizophrenia(sz) is a chronic disorder which affects the thinking ability as well as general behavior. Early detection is one of the ways to mitigate this problem. CNN is used for feature extraction from EGG signals.

The dataset includes EEG signals of 28 subjects which consisted of fifteen minutes of EEG signals acquired from 14 patients with paranoid sz, encompassing seven males and seven females, with a mean age of 27.9 ± 3.3 and 28.3 ± 4.1 years, respectively and 14 healthy subjects with same age and gender ratio.

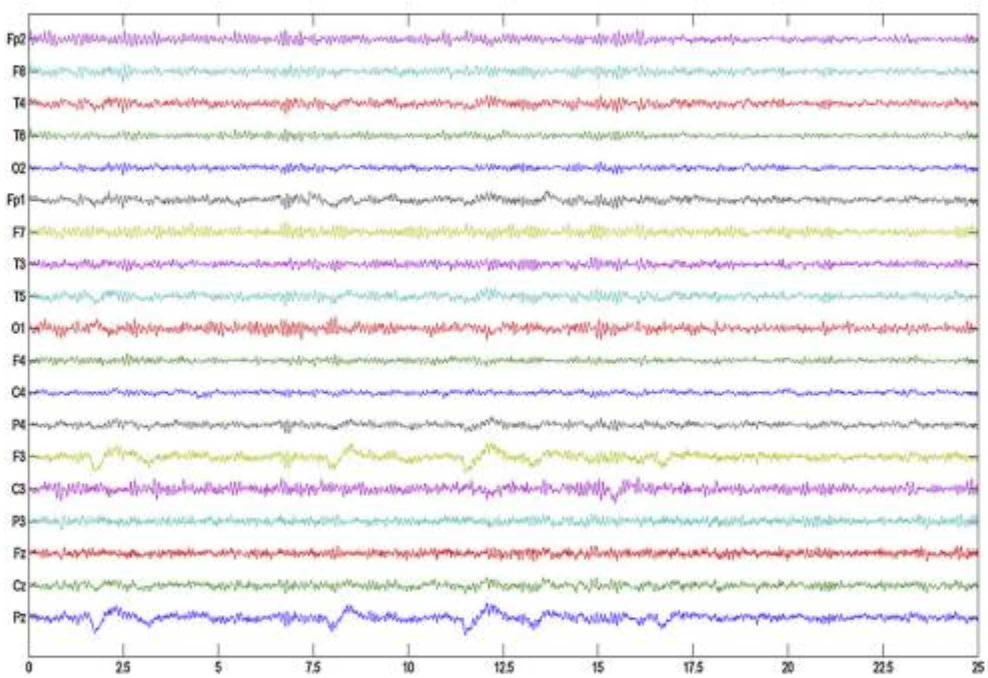
DATASET

This paper uses dataset which contains Fifteen minutes of EEG signals acquired from 14 patients with paranoid schizopherenia ,seven of them were males and rest seven were females.All the patients were of almost 27.9 to 23.3 years old.

This dataset was collected from the Institute of Psychiatry and Neurology in Warsaw, Poland.Fourteen healthy subjects within similar age and gender ratio were taken from the same institute.While taking the brain signal data of a patient 19 electrodes were placed on the scalp of patients which produces a multi-channel (19-channel) EEG signal.The electrodes used wereFp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1 and O2.Fifteen minutes of EEG signals were recorded from the participants at a frequency of 250 Hz, as they remained in a composed state with eyes closed.Table given below explains the EEG segments studied from the two classes. The sample EEG signal of normal and sz cases are depicted in Fig. 1. Fig. 1(a) presents the normal EEG signal, which shows enhanced amplitude values in most of the channels, as compared to the sz signals depicted in Fig. 1(b).



1(a)



1(b)

Fig. 1. Pre-processed EEG signals of (a) normal and (b) schizophrenia.

MODEL

DATA

We have used the "**Automated detection of schizophrenia using nonlinear signal processing methods**" dataset provided by https://repod.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/repod.0107441.

The dataset comprised 14 patients with paranoid schizophrenia and 14 healthy controls. Data were acquired with the sampling frequency of 250 Hz using the standard 10-20 EEG montage with 19 EEG channels: Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2. The reference electrode was placed between electrodes Fz and Cz.

In subject base testing, validation of the system was carried out in three phases: training, testing and validation of data. During training, k-fold validation was used, whereby the entire data was split into fourteen equal parts. Twelve parts were used for training, one was used for validation and another for testing .

PREPROCESSING

Thirty second segments without artefacts were used for analysis. The signals were segmented into nonoverlapping segments of 25 s, such that each segment consisted of 6250 × 19 sample points.

This segmentation gave rise to 1142 EEG patterns, which were then grouped to form a new database of normal and sz class EEGs with a fixed length.

Type	Number of EEG segments
Normal	516
Schizophrenia	626
Total	1142

METHODOLOGY

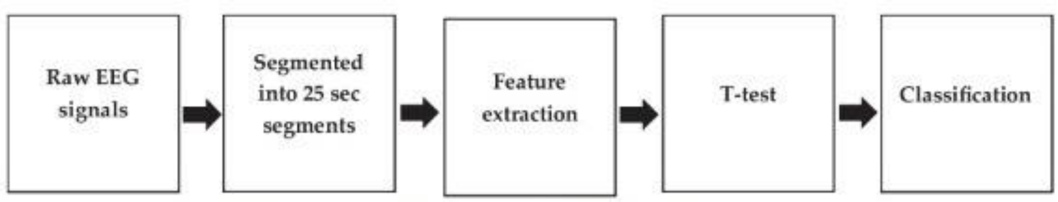


Fig. 2. Organization of the proposed automated tool to detect Schizophrenia.

The raw EEG signals is segmented into 25 second segments.Then CNN is used for feature extraction , out of 157 non linear features extracted from both EEG classes ,the set of 14 features were then selected using Student’s t-test[1] .The subjects are then classifies into two classes healthy and schizophrenia.

14 optimal features:

Activity Entropy(ae):

Entropy is the disorder of a system, described by the distribution probabilities of molecules of gaseous or fluid systems.

Entropy simply quantifies the uncertainty or ignorance in a statistical sense, which is equivalent to the possible configurations of possibilities of the system.

Applying the concept of entropy to time series like electroencephalography (EEG) is a way to quantify, in a statistical sense, the amount of uncertainty or randomness in the pattern, which is also roughly equivalent to the amount of information contained in the signal.

Entropy measures in the time domain generally break up the signal into segments that are then compared for similarity either directly (in the time domain) or after some kind of transformation of the signal (such as the power spectral density).

This typically depends on a few fundamental parameters – the length of segment chosen, the transformation of the signal (if any) and the distance metric or the way the segments are compared.

Two of the popular choices in time domain include approximate entropy and sample entropy. These are used to quantify the amount of repeatability or predictability in the waveform patterns of an EEG signal.

SampEn(k,r,N)=-ln(A(k)/B(k-1) for k=0,1,...,m-1 with B(0)=N, the length of the input series.

The algorithm to find runs starts by finding all points that match the first point within a tolerance r.

Largest Lyapunov exponent(lx):

Lyapunov exponents,can serve as indicators of specific brain function suffered from the disease.The Lyapunov spectrum provides an opportunity to increase the classification accuracy due to a larger number of significant channels. Thus, non-linear EEG analysis, based on Lyapunov exponents, is considered useful for neurodynamic research of patients with schizophrenia.

The Largest Lyapunov exponent and spectrum were computed from the time series of each channel. Wolf , Rosenstein and Kantz algorithms were implemented.

Each EEG signal was represented as a time series. The reconstruction procedure was applied to each EEG data by embedding the signal into phase space. Time delays of 5 ms and embedding dimensions of 4 were used for calculation.

Wolf's method

The idea of the method is the calculation of the maximal Lyapunov exponent by single coordinate sampling.

It is useful when system evolution equations are unknown, and phase coordinates are not measurable.

$$x_i = (x(i), x(i-\tau), \ldots, x(i-(m-1)*\tau)) = (x_1(i), x_2(i), \ldots, x_m(i)),$$

$$\text{where } i = \overline{((m-1)\tau+1), N}.$$

Kantz's method:

Kantz's algorithm [18] finds all the points in the neighborhood of the reference trajectory and estimates an average distance between the neighbors and reference trajectory as a function of time (or non-dimensional time, multiplied by the sampling rate). The algorithm solves the following equation.

$$S(\tau) = \frac{1}{T}\sum_{t=1}^{T} \ln\left(\frac{1}{|U_t|}\sum_{i\in U_t} |x_{t+\tau} - x_{i+\tau}|\right)$$

Rosenstein's method:

Rosenstein's Method is relatively simple and shows good computational speed. However, its result is not a numeric value λ1, but rather is some function of time.

$$y(i, \Delta t) = \frac{1}{\Delta t}\ln d_j(i), d_j(i) = \min_{x_j} x_j - x'_j,$$

Kolmogorov Sinai(k-s) entropy:

The Kolmogorov–Sinai entropy serves as the central complexity measure for dynamical systems and can be considered as a reference for other complexity measures, including in data analysis. Roughly speaking, it measures the mean information obtained by each iteration step.

The Kolmogorov–Sinai (K–S) entropy is used to quantify the average amount of uncertainty of a dynamical system through a sequence of observations.

Sequence probabilities therefore play a central role for the computation of the entropy rate to determine if the dynamical system under study is deterministically non-chaotic, deterministically chaotic, or random.

After applying the detrended fluctuation analysis and Kolmogorov-Sinai entropy (KSE) in the EEG signals, the detrended fluctuation functions follow a power law with Hurst exponents larger than 1/2. The Hurst exponents enhanced at all EEG channels in the state of mathematical activities. The KSE in the relaxed state is larger than those in the state of the mathematical activities. These indicate that the entropy is enhanced in the disorder state of the brain.

Hjorth complexity(hc) and mobility(hm):

Hjorth parameters are indicators of statistical properties used in signal processing in the time domain introduced by Bo Hjorth in 1970.[1] The parameters are Activity, Mobility, and Complexity. They are commonly used in the analysis of electroencephalography signals for feature extraction. The parameters are normalised slope descriptors (NSDs) used in EEG.

Hjorth Activity

The activity parameter represents the signal power, the variance of a time function. This can indicate the surface of power spectrum in the frequency domain. This is represented by the following equation.

$$\mathrm{Activity} = \mathrm{var}(y(t)).$$

Where *y(t)* represents the signal.

Hjorth Mobility

The mobility parameter represents the mean frequency or the proportion of standard deviation of the power spectrum. This is defined as the square root of variance of the first derivative of the signal y(t) divided by variance of the signal y(t).

$$\mathrm{Mobility} = \sqrt{\frac{\mathrm{var}(\frac{dy(t)}{dt})}{\mathrm{var}(y(t))}}.$$

Hjorth Complexity

The Complexity parameter represents the change in frequency. The parameter compares the signal's similarity to a pure sine wave, where the value converges to 1 if the signal is more similar.

$$\mathrm{Complexity} = \frac{\mathrm{Mobility}(\frac{dy(t)}{dt})}{\mathrm{Mobility}(y(t))}.$$

Rényi(re):

The Rényi entropy generalizes the Hartley entropy, the Shannon entropy, the collision entropy and the min-entropy. Entropies quantify the diversity, uncertainty, or randomness of a system. The entropy is named after Alfréd Rényi, who looked for the most general definition of information measures that preserve additivity for independent events.

In the context of fractal dimension estimation, the Rényi entropy forms the basis of the concept of generalized dimensions.

The Rényi entropy of order α, where $\alpha \geq 0$ and $\alpha \neq 1$, is defined as

$$\mathrm{H}_\alpha(X) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{n} p_i^\alpha\right)$$ [1]

Here, X is a discrete random variable with possible outcomes in the set $\mathcal{A} = \{x_1, x_2, \ldots, x_n\}$ and corresponding probabilities $p_i \doteq \Pr(X = x_i)$ for $i = 1, \ldots, n$. The logarithm is conventionally taken to be base 2, especially in the context of information theory where bits are used. If the probabilities are $p_i = 1/n$ for all $i = 1, \ldots, n$, then all the Rényi entropies of the distribution are equal: $\mathrm{H}_\alpha(X) = \log n$. In general, for all discrete random variables X, $\mathrm{H}_\alpha(X)$ is a non-increasing function in α.

Applications often exploit the following relation between the Rényi entropy and the *p*-norm of the vector of probabilities:

$$\mathrm{H}_\alpha(X) = \frac{\alpha}{1-\alpha} \log(\|P\|_\alpha).$$

Here, the discrete probability distribution $P = (p_1, \ldots, p_n)$ is interpreted as a vector in $\mathbb{R}^n$ with $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$.

The Rényi entropy for any $\alpha \geq 0$ is Schur concave.

Shannon(sn):

Instead of assessing the power density distribution on the time-frequency plane, like previously proposed measures of signal nonstationarity, Shannon(sn) measure is based on the shift of the dominant frequency of the EEG signal over time.SEPFS measure can be applied to assess the properties of EEG nonstationarity in subjects before and shortly after they suffered Schizopherenia. Patient having Schizopherenia are tested prior to concussive episodes as a baseline. From this subject pool,those who suffered from Schizopherenia are re-tested . Additional subjects pool (student-athletes without history of concussion, n=30) were recruited and test-retested.

To compute the Shannon entropy of the peak frequency shifting, the spectrum of signal x is divided into M sub-bands from the lowest frequency fmin to the highest frequency fmax, denoted as Sb1, Sb2 … SbM, with central frequencies

fc1< fc2…< fcM. The peak frequency at any time point τ is defined as follows:

$$f_p(\tau) = \frac{f_s f_c}{a_k}, a_k = \max_{i=1,..M} (\mathrm{CWT}_x(a_i, \tau))$$

The SEPFS can be computed as follows:

$$SEPFS = \sum_{i=1,p(i)>0}^{M} p_i \,\boldsymbol{log}(p_i), p_i = \frac{n_i}{N}$$

Tsallis(ts):

Tsallis entropy (TsEn) plays a central role in nonextensive statistical mechanics. It is successful at describing systems with long-range interactions, multifractal space-time constraints, or long-term memory effects . TsEn also allows incorporation of an entropy scaling parameter with which short- and long-range interactions can be probed. EEG spikes, bursts, and continuous or fused rhythms may thus be differentiated with the help of Tsallis statistics.

To quantify the EEG data,Tsallis entropy (TsEn) are examined. More specifically, an entropy-based measure named TsEn area (TsEnA) is proposed to reveal the presence and the extent of development of BS following brain injury. The methodology of TsEnA and the selection of its parameter are explained in detail.

TsEnA reliably quantifies the complex dynamics in BS EEG, and is useful as an experimental or clinical tool for the dignosis of brain disorder like Schizoherenia.

A nonextensive statistics, known now as TsEn, was proposed by Tsallis, which was defined as

$$\boxed{\mathrm{TsEn} = \frac{1 - \sum_{i=1}^{W} p_i^q}{q-1}}$$

Kolmogorov complexity(kc):

It is a measure of the computational resources needed to specify the object, and is also known as algorithmic complexity, Solomonoff–Kolmogorov–Chaitin complexity, program-size complexity, descriptive complexity, or algorithmic entropy.

Complexity measures have been enormously used in schizophrenia patients to estimate brain dynamics. the nonlinear brain dynamics of chronic and medicated schizophrenia patients can be detected using distinct complexity estimators.

The EEG complexity of participants can be investigated using lota of method like approximate entropy (ApEn), Shannon entropy (ShEn), Kolmogorov complexity (KC) and Lempel-Ziv complexity (LZC).

But it is found that KC was more sensitive for detecting EEG complexity of patients than other estimators in all investigated brain regions. Moreover, significant inter-hemispheric complexity differences were found in the frontal and parietal areas of schizophrenics' brain. the utilizing of sensitive complexity estimators to analyze brain dynamics of patients might be a useful discriminative tool for diagnostic purposes.

Therefore nonlinear analysis will gives a deeper understanding of schizophrenics' brain.

Bispectrum (entropy 1, 2 and phase)(bs):

In the feature extraction process, to extract the information contained in the bispectrum matrices, a 3D pyramid filter is used for sampling and quantifying the bispectrum value. Experiment results show that the mean percentage of the bispectrum value from 5 × 5 nonoverlapped 3D pyramid filters produces the highest recognition rate. We found that reducing the number of EEG channels down to only eight in the frontal area of the brain does not significantly affect the recognition rate, and the number of data samples used in the training process is then increased to improve the recognition rate of the system

In the mathematical formula of bispectrum analysis, there is a correlation calculation between the frequency components thus, the phase coupling components of the EEG signals could be revealed. Some characteristics of bispectrum analysis are the ability to extract deviations due to Gaussianity, to suppress the additive colored Gaussian noise of an unknown power spectrum and to detect nonlinearity properties . Because of its superiority, bispectrum analysis is used in this research as the signal processing technique in the feature extraction step. We expect that by using bispectrum analysis, the recognition rate of the EEG-based automatic emotion recognition system will improve.

The autocorrelation of a signal is the correlation between the signal and itself at a different time; for example, at time t and at time $t + m$. The autocorrelation function of $x(n)$ can be expressed as the expectation of stationary process, defined as:

$$R_{xx}(m) := E\{x^*(n)x(n + m)\}. \quad (1)$$

Permutation entropy(pe):

PE is a linear complexity measure that tracks changes in spontaneous brain activity resulting from the administration of anesthetic agents.

The information analysis of the electroencephalogram (EEG) signal is carried out by granulation and reciprocal entropy

(PeEn). The analysis of the EEG signal is obtained by experimental activity. Due to its complexity and multichannel characteristic, together with granular computing (GrC) and PeEn are used to analyze the EEG signal.

The EEG signal consists of 19 channels of data and the experimental data are used to discriminate patterns, with experimental focus on considering real and thinking actions.The time-series EEG signals were granularized according to the changes in the signal and analyzed by PeEn coding and Fuzzy C-Means (FCM) algorithm. Because there are two main actions, i.e., left-handed, and right-handed actions were clearly delineated.

The relative frequency of each permutation can be calculated by counting the number of times the permutation is found in the time series divided by the total number of sequences.

$$Y_i = \left[y_i^{(s)},\ y_{i+\tau}^{(s)},\ y_{i+2\tau}^{(s)},\ \cdots,\ y_{i+(m-1)\tau}^{(s)}\right],\ i = 1,\ 2, \ldots,\ Ns - (m-1)\tau$$

The features were ranked based on the p-values. The Hjorth complexity, with the lowest p-value, is ranked first, portraying to be the most significant feature.

Entropy, with the highest p-value, is ranked fourteenth, portraying to be the worst feature, for the classification of EEG signals.

RESULT

To evaluate the performance of the proposed model, a set of quantitative metrics comprising of accuracy, precision, recall and F1-score have been used. The results are reported in Table 1. They show the highest values of the quantitative metrics obtained until the corresponding epoch number.

Evaluation Metrics :

For evaluating the proposed models we used precision , recall , F1 Score , support and accuracy.We have used classification report and confusion matrix , these metrics are widely used for evaluating supervised machine learning models for classification when the dataset is multi labelled . Suppose a multi-labeled dataset consists of N number of instances, where each instance Ni is given by (xi,yi) , where xi represents the set of attributes and yi represents the set of labels . Let us say that yi and yi’ represent the labels which are true and labels which are predicted respectively for the ith instance then the metrics can be described for the ith instance which are by the following formulae.

Precision :

It is the ratio of accurately predicted authors as hate speech spreaders or not to the total number of predicted authors. It is computed as given in the equation below. The range of precision varies between 0 and 1 , where 1 is the best value and 0 is the worst value.

$$\text{Precision} = \frac{\text{Number of accurately predicted authors}}{\text{Total number of predicted authors}} = \frac{|yi \cap yi'|}{|yi'|}$$

Recall :

It is the ratio of accurately predicted authors as hate speech spreaders or not to the total number of authors. It is computed as is given in the below equation. The range of recall varies between 0 and 1 , where 1 is the best and 0 is the worst value .

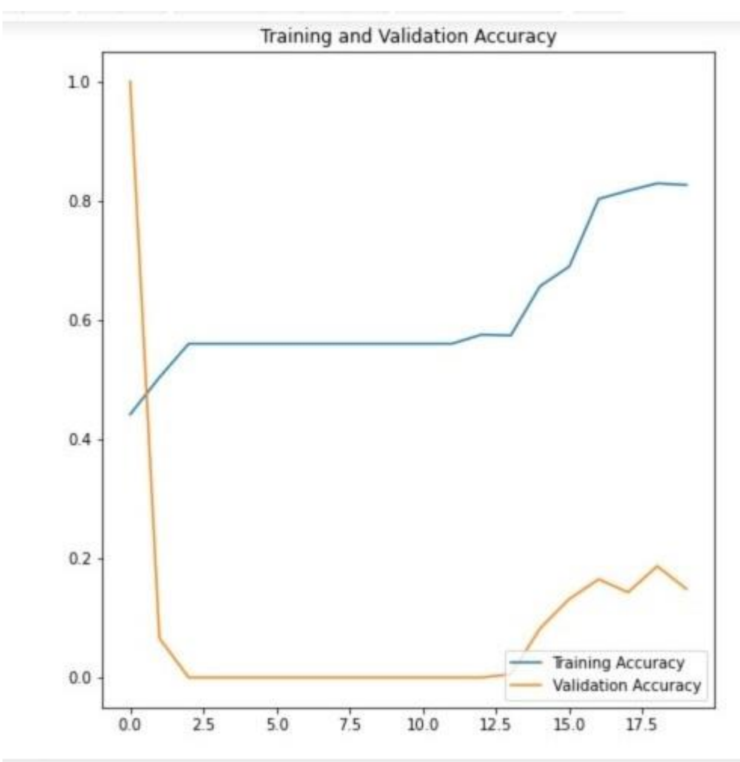
$$\text{Recall} = \frac{\text{Number of accurately predicted authors}}{\text{Total number of authors}} = \frac{|\, yi \cap yi' \,|}{|yi|}$$

F1-Score :

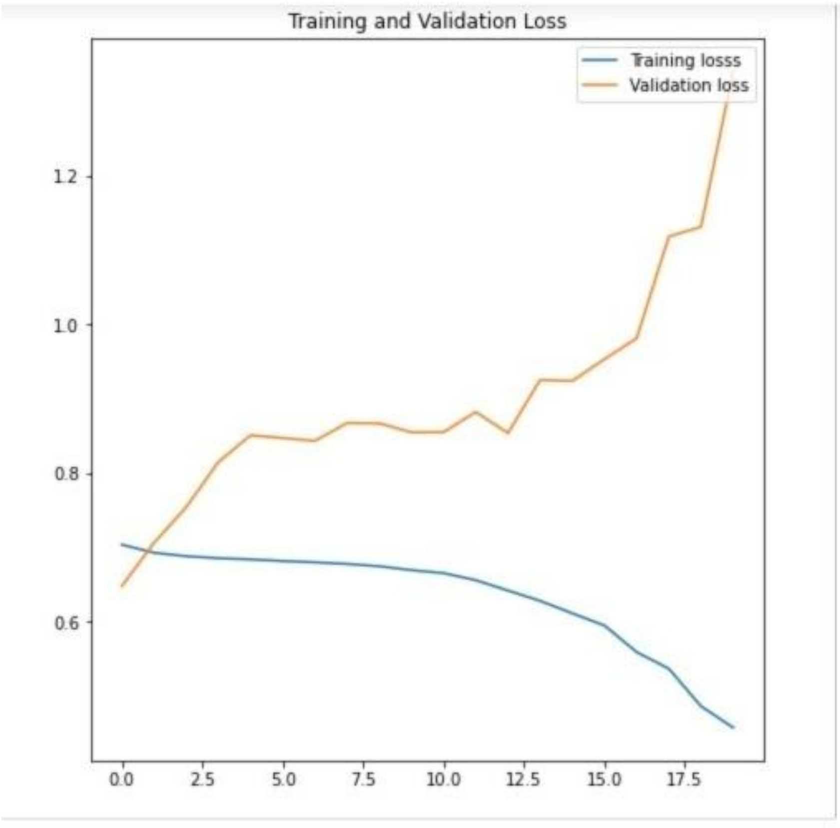
The harmonic mean between Precision and Recall is called F1-Score, which gives the balanced equation between them . It can be represented by the below equation . The range of F1-score varies between 0 and 1 , where 1 is the best and 0 is the worst value.

$$\text{F1-score} = 2 \times \frac{\text{Precision x Recall}}{\text{Precision + Recall}}$$

Comparing training and validation accuracy with number of trials for the model



Comparing training and validation accuracy with number of trials for the model



A highest training accuracy of 90% was obtained over 20 epochs of training and overall testing accuracy was obtained to be 63%.
This is an effective measure of the classification made by the deep learning model .The plot of train and test accuracy and loss against the epochs provide a means of visualization and indication of the speed of model convergence.

CONCLUSION

A CNN based classification has been done on the dataset which consists EEG brain signals(19 channels) of 14 healthy subjects and 14 subjects suffering from schizophrenia ,with an accuracy of 63% .
The accuracy achieved could be further improved using RNN and LSTM as according to studies LSTM provides the best accuracy on sequence inputs.

REFERENCES

[1] Zhou N, Wang L. A modified t-test feature selection method and its application on the hapmap genotype data. Genom Proteom Bioinf 2007;5(3–4):242–9. https://doi.org/10.1016/S1672-0229(08)60011-X.

[2] Jahmunah, V., Oh, S.L., Rajinikanth, V., Ciaccio, E.J., Cheong, K.H., Arunkumar, N. and Acharya, U.R., 2019. Automated detection of schizophrenia using nonlinear signal processing methods. Artificial intelligence in medicine, 100, p.101698.

[3] Shoeibi, A., Sadeghi, D., Moridian, P., Ghassemi, N., Heras, J., Alizadehsani, R., Khadem, A., Kong, Y., Nahavandi, S., Zhang, Y.D. and Gorriz, J.M., 2021. Automatic diagnosis of schizophrenia in EEG signals using CNN-LSTM models. Frontiers in Neuroinformatics, 15.

[4] Dvey-Aharon, Z., Fogelson, N., Peled, A. and Intrator, N., 2015. Schizophrenia detection and classification by advanced analysis of EEG recordings using a single electrode approach. PloS one, 10(4), p.e0123033.

[5] Shoeibi, A., Sadeghi, D., Moridian, P., Ghassemi, N., Heras, J., Alizadehsani, R., Khadem, A., Kong, Y., Nahavandi, S., Zhang, Y.D. and Gorriz, J.M., 2021. Automatic diagnosis of schizophrenia in EEG signals using CNN-LSTM models. Frontiers in Neuroinformatics, 15.

[6] Kolmogorov AN. New metric invariant of transitive dynamical systems and endomorphisms of Lebesgue spaces. Dokl Russ Acad Sci 1958;119:861–4.

7.Sweller J, Van Merrienboer JJG, Paas FGWC. Cognitive architecture and instructional design. Educ Psychol Rev 1998;10:251–96. https://doi.org/10.1023/A:1022193728205.

[8] Sweller J, Van Merrienboer JJG, Paas FGWC. Cognitive architecture and instructional design. Educ Psychol Rev 1998;10:251–96. https://doi.org/10.1023/A:1022193728205A.

[9] Renyi A. On measures of entropy and information. Fourth Berkeley Symposium. 1961.

[10] Shannon CE. A mathematical theory of communication. Bell Syst Tech J1948;27(3):379–423.