**Epileptic Seizure Prediction Using Intracranial Brain**

**Records (EEG)**

Pranav Karnani, Awais Mullah, Adarsh Pandey and Prof. Kirti Parthe

*Student, Information Technology,*

*K.C. College of Engineering and Management Studies and Research, Mumbai University, India*

E-mail: [pranavkarnani@kccemsr.edu.in](mailto:pranavkarnani@kccemsr.edu.in)

*Student, Information Technology,*

*K.C. College of Engineering and Management Studies and Research, Mumbai University, India*

E-mail: [awaismullah@kccemsr.edu.in](mailto:awaismullah@kccemsr.edu.in)

*Student, Information Technology,*

*K.C. College of Engineering and Management Studies and Research, Mumbai University, India.*

E-mail: [adarshpandey@kccemsr.edu.in](mailto:adarshpandey@kccemsr.edu.in)

*Assistant Professor, Information Technology,*

*K.C. College of Engineering and Management Studies and Research, Mumbai University, India*

E-mail:

[kirti.parthe@kccemsr.edu.in](mailto:kirti.parthe@kccemsr.edu.in)

***Abstract—*** ***Epilepsy is the second most common brain disorder after migraine. Automatic detection of epileptic seizures can considerably improve the patients' quality of life. Current Electroencephalogram (EEG)-based seizure detection systems encounter many challenges in real life situations. The EEGs are non-stationary signals and seizure patterns vary across patients and recording sessions. Moreover, EEG data are prone to numerous noise types that negatively affect the detection accuracy of epileptic seizures. To address these challenges, we introduce the use of a deep learning-based approach that automatically learns the discriminative EEG features of epileptic seizures. Specifically, to reveal the correlation between successive data samples, the timeseries EEG data are first segmented into a sequence of non-overlapping epochs. Second, Long Short-Term Memory (LSTM) network is used to learn the high-level representations of the normal and the seizure EEG patterns. Third, these representations are fed into Soft max function for training and classification. The results on a well-known benchmark clinical dataset demonstrate the superiority of the proposed approach over the existing state-of-the-art methods. Furthermore, our approach is shown to be robust in noisy and real-life conditions. Compared to current methods that are quite sensitive to noise, the proposed method maintains its high detection performance in the presence of common EEG artifacts (muscle activities and eye-blinking) as well as white noise.***

# *Keywords: LSTM), EEG, Seizure.*

1. **INTRODUCTION**

People suffering from epilepsy or also called seizure disorder, a disorder in which nerve cell activity in the brain is disturbed, causing seizures. These seizures do not have a fixed time they occur at random times causing great disturbance to emotional state and cognitive state of the mind. Their ability to work, social, economic situations come to a sudden halt. Having the option to foresee epileptic seizures will incredibly improve the personal satisfaction of individuals with epilepsy by either giving them an admonition of an approaching seizure so they can move to security or enacting an embedded seizure control gadget that can turn away seizures through medication conveyance or electrical incitement of the cerebrum. How might we arrange the stage epileptic seizures utilizing electrical accounts of mind action?

This world consists of 70 million people having PWE i.e., epilepsy and nearly 12 million PWE patients reside in India; which amounts to about a sixth of the world. The general pervasiveness (3.5-12.1 per 1,600 populace) which has a frequency (0.3-0.9 per 1,400 populaces for every year) information from ongoing examinations in India on all-inclusive community are similar to the paces of high-income nations regardless of checked varieties in populace attributes & information systems. Having a differential conveyance of PWE among different socio-demographic & monetary gatherings with higher rates detailed for the male, country populace, and low financial status. A changing example in the age-explicit event of epilepsy with prevalence towards the more established age bunch is seen due to sociodemographic and epidemiological progress. Neuro-infections, neuro-cysticercosis, and neurotrauma alongside birth wounds have risen as significant hazard factors for optional PWE. In spite of its shifted etiology, PWE are reasonable in outlook.

Programmed identification of epileptic seizures can impressively improve the patients' personal satisfaction. Current Electroencephalogram (EEG)- based seizure discovery frameworks experience numerous difficulties, all things considered, circumstances. The EEGs are non-fixed signs and seizure designs fluctuate across patients and recording meetings. Besides, EEG information are inclined to various clamor types that contrarily influence the recognition precision of epileptic seizures. To address these difficulties, we present the utilization of a profound learning-based methodology that naturally learns the discriminative EEG highlights of epileptic seizures. In particular, to uncover the connection between progressive information tests, the time-arrangement EEG information are first sectioned into a succession of nonoverlapping ages. Second, Long Short-Term Memory (LSTM) arrange is utilized to become familiar with the elevated level portrayals of the ordinary and the seizure EEG designs. Third, these portrayals are taken care of into SoftMax function for training. The outcomes on a notable benchmark clinical dataset show the prevalence of the proposed approach over the current best in class strategies. Moreover, our methodology is demonstrated to be hearty in uproarious and genuine conditions. Contrasted with current techniques that are very touchy to clamor, the proposed strategy keeps up its high recognition execution within the sight of basic EEG ancient rarities (muscle exercises and eye-squinting).

1. **LITERATURE SURVEY**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sr**  **No** | **Title of Paper** | **Methodo- logy** | **Advantages** | **Limitations** |
| 1. | Algorithm & Feature extraction: CNN Used for both | Cconvolutional Neural Network (CNN) | affectability of 90% plus a lower bogus expectation pace of 0.122 FP/h | 90% reliability |
| 2. | Algorithm: Sparse Classifier with Window Based Feature Extraction | VTest | Least errors  High Speed | Need more reliability |
| 3. | Algorithm: Random Forest, WPF used for feature extraction. | Logistic regression & SVM  PCA analysis. | Variance is better & is in differentiation preictal & interictal samples. | None |
| 4 | Journal Paper**,** J. Schroeterand M. M. Sondhi | Techniques for estimating vocal-tract shapes from the speech signal | Higher computational speed | Only voice spelched |
| 5 | M. Pardo, “Vocal tract shape analysis for children. | Vocal tract shape analysis for children | Accurate depiction on children | Only for  children |
| 6 | IEEE standard Book**,** J. F. Curtis, Processes and Disorders of Human Communication | Processes and Disorders of Human Comm- unication | Second Degree is useful | Only vocal use |

# METHODOLOGY

**3.1 WPFs**

The electroencephalogram signal measures the mind wave designs. It has been demonstrated to be viable in following issues related telectrical action of the cerebrum, for example, analyzing with checking the seizure issue along with the rest issue. In the previous experiments, serious execution has been accomplished for epilepsy seizure grouping dependent on EEG. However, little consideration has been given to the preictal state order, also which is extremely useful for the seizure forecast and along these lines is essentially progressively significant in forestalling the coincidental losses brought about by the seizure.

**3.2 WPD**

It is also called the optimal sub-band tree organizing, applies the wavelet change on the high pass filtering results along with the estimation attributes which are the low pass filtering results coming from the past level to accomplish an ideal portrayal of the signal. In the DWT, each level is determined by passing just the past wavelet guess attributes through discrete-time low and high pass quadrature reflect channels. For n levels of disintegration, the WPD produces 2n various arrangements of coefficients (or hubs) instead of (3n + 1) sets for the DWT. Be that as it may, because of the down sampling procedure the general number of coefficients is as yet the equivalent and there is no repetition.

**3.2.1 Random Forest**

Random forest, similar to its name infers, comprises of countless individual choice trees that work as an outfit. Every individual tree in the forest lets out a class expectation and the class with the most agreement turns into our model's prediction. The crucial idea driving arbitrary woods is a basic however amazing one — the shrewdness of groups. In data science talk, the explanation that the arbitrary timberland model works so well is: An enormous number of generally uncorrelated models (trees) working as an advisory group will beat any of the individual constituent models.

**3.3 Convolutional Neural Network (CNN)**

A CNN is a sort of engineered neural system utilized in picture notoriety and preparing that is uniquely intended to way pixel information.

CNNs are ground-breaking photo preparing, engineered insight (AI) that utilizes profound acing to play out each generative and enlightening undertakings, routinely utilizing machine vison that incorporates picture and video notoriety, related to recommender frameworks and natural language processing (NLP).

# PROPOSED SYSTEM

In contrast to increasingly customary strategies for machine learning methods, deep learning classifiers are prepared through component adapting as opposed to task-explicit calculations. This means the machine would understand the patterns in the pictures as it is given instead of requiring the human administrator to characterize the examples which the machines should look for in the picture. The component learning system is utilized each day by the way we show a kid to perceive distinctive objects. Deep learning is a sort of AI that impersonates the neuron of the neural systems present in the human cerebrum. PC Vision Deep learning models are prepared on a lot of pictures a.k.a preparing information, to explain an undertaking. These deep learning models are for the most part utilized in the field of Computer Vision which permits a PC to see and picture as a human would. This kind of system is like the organic sensory system, with every hub going about as a neuron inside a bigger network. Thus, these models are a class of fake neural systems. This model’s calculations adapt dynamically about the picture as it experiences each neural system layer. Prior layers figure out how to recognize low-level highlights like edges, and ensuing layers consolidate highlights from prior layers into an increasingly all-encompassing and complete portrayal.

Chapter 1Introduction

**4.1 Deep Learning vs Machine Learning Algorithms**

Deep learning is a sort of AI that emulates the neuron of the neural systems which human mind has. Deep learning models are prepared considering lot of pictures for preparing information, to comprehend an assignment. These deep learning models are for the most part utilized in the sector of Computer Vision, which allows a PC to compare and envision as a real person would.

These models could be visualized as a group of points wherein each of those points makes a decision based on the nodal input. The human biological neural system as has the same type of network, wherein each node has a neural within a network which is larger than itself. Thus, these models are a class of fake neural systems. Deep learning algorithms adapt continuously about the picture as it experiences each neural system layer. Early layers figure out how to distinguish low-level highlights like edges, and resulting layers consolidate highlights from prior classes into a progressively all-encompassing and thorough portrayal. ANN is a system that simulates the neural networks present in living beings. The work in a similar way as to that of animals or more closely like humans. They aren’t programmed for a specific task. We use examples to make them learn. The best use of ANN may be image recognition. Labelled pictures of an object such as dogs or cats are manually fed to the system. No information of the constituents of the picture are entered. The ANN takes the features into account while creating the model.

Artificial neurons are connected to each other. It mimics the human brain. Each neuron can transmit signal to the other neurons. A neuron receives a signal then processes it and sends this processed signal further. The input signal is a real number. A non-linear function computes the sum of inputs and passes it ahead. The connections are known as edges.

**4.2 Principle**

The basic working is based on the principle that a small change in input brings a large change in output. A given neuron can have multiple inputs or a single input. The same is true for outputs as well.

Propagation function computes the output by processing the inputs from the previous neuron.

Deep Learning may use multiple layers as it is more advanced in computing results. Multiple connections are also possible between the layers. It could be a many-to-one connection or one-to-one.

**4.3 Learning**

The actual process of making a system which computes data correctly is this step. Sample data is used to adjust the weights of the network to improve accuracy. Errors are minimized. The process happens over and over again till the error cannot be reduced any further. The error rate cannot be reduced to zero.

We may redesign the network if the output does not prove to be accurate. Hundred percent accuracy is not possible in such a system but ninety plus is generally considered good.

Learning rate is the steps the model takes to adjust errors in each observation. A high learning rate gets the work done faster but the accuracy suffers.

**4.3 Feedforward ANN**

The information flow is unidirectional. The information does not return to a neuron once it leaves it. It is used in Simple Pattern recognition. The inputs and outputs are fixed.

**4.3.1 Feedback ANN**

The same neuron may get a signal multiple time

**4.4 Supervised Learning**

Here the system comes up with guesses as to what the correct answer may be. The correct answer is given to the system. According to the difference between the answer and guess the network makes changes. The cost function here eliminates incorrect deductions. The most commonly used cost is mean-squared error. Pattern recognition and regression are done using this method. Handwriting and gesture recognition problems can be dealt with using this technique.

**4.4.1 Unsupervised Learning**

Here the output in not known therefore the correct answers to the dataset aren’t available. This could be used when searching for new patterns. The input data along with cost function are provided at the input.

**4.4.2 Reinforcement Learning**

It works on the principle of observing the environment. The best example of this method is playing video game. The environment could give you unpredictable response in every game but you start learning from each game. The goal is to get the lowest cost possible. The goal has to be achieved as early as possible. Initially everything is estimated but as the game progresses things start getting better.

**4.5 Backpropagation**

Backpropagation is used to modify the weights to compensate for the errors. Mathematically the derivative of the cost function is linked with a given state.

**4.6 Applications**

Used is quantum chemistry, playing games, face identification, signal classification, 3D construction sequence recognition, medical diagnosis, controlling car, trajectory prediction, finance, trading, data mining, social network filtering and spam filtering. It is also used in diagnosis of cancer.

It is also used in geoscience, coastal engineering, cybersecurity and geomorphology.

It can also be used to simulate properties of quantum systems although it is still in research stage.

EEG, ECG analysis, designing prosthetics, speech classification, text to speech conversion, image compression, automated information services, braking vehicles, audio signal processing, anomaly detection, guidance system, weapons, electronics, manufacturing products etc.

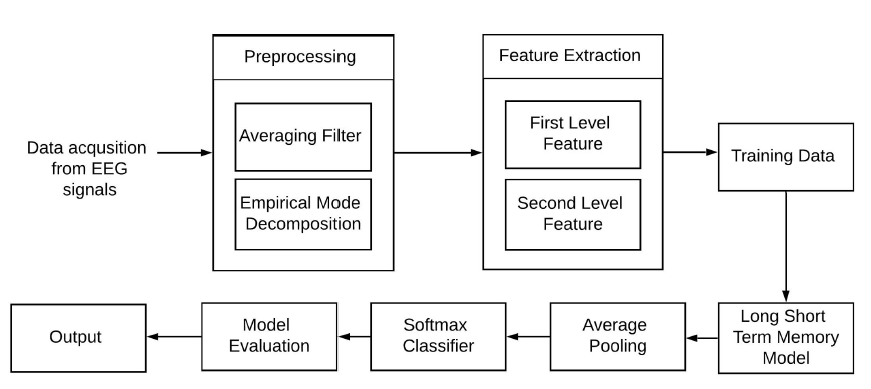
**4.7 Processing Power Required**

Huge and viable neural systems require significant computer resources. While the brain has equipment custom fitted to the undertaking of preparing signals through a chart of neurons, recreating even a rearranged neuron on von Neumann design may expend immense measures of memory and capacity. Moreover, the fashioner frequently needs to transmit flags through a large number of these associations and their related neurons – which require colossal CPU force and time.

The resurgence of neural systems in the twenty-first century is to a great extent owing to propels in equipment: from 1991 to 2015, registering power, particularly as conveyed by GPUs has expanded around a million-overlay, making the standard backpropagation calculation achievable for preparing systems that are a few layers further than before. The utilization of quickening agents, for example, FPGAs and GPUs can diminish preparing times from months to days.

Neuromorphic building tends to the equipment trouble straightforwardly, by developing non-von-Neumann chips to legitimately actualize neural systems in hardware. Another kind of chip upgraded for neural system handling is known as a Tensor Processing Unit, or TPU.

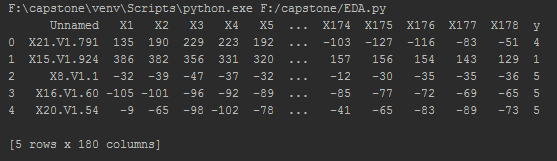
# Project Design

****

1. **IMPLEMENTATION**
   1. **Exploratory Data Analysis**

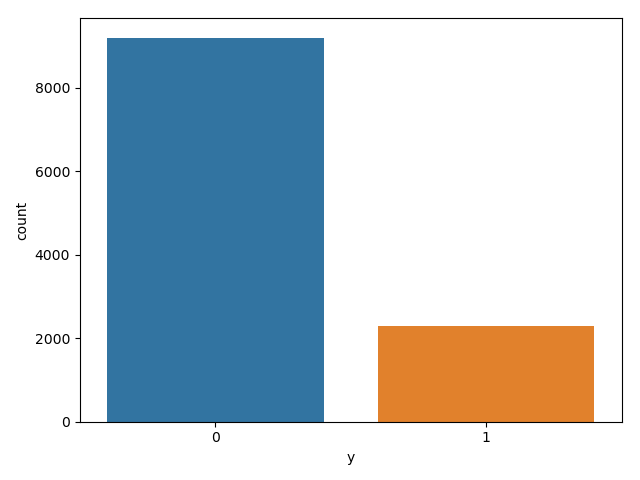
The Dataset here corresponds to what looks as a set of 11500 patients. Each data point in the dataset for one particular patient corresponds to a time value which was recorded at the time. Each patient corresponds to 178 distinct values of point in time at the time of recording or curating of the dataset.

The last column in the dataset is the labelled prediction of one patient. It originally revolves around 5 distinct labels i.e. 1, 2, 3, 4, 5. But since we’re making a binary classification of the dataset we will have to truncate it down to 2 labels namely 1.0.



**5.1. EDA**

As we can see that we have 178 eeg features with 5 possible classes. The main aim of the project here is to classify the correctly whether an eeg data provided has a possibility of seizure or not. Thus, this problem deals with the binary classification.

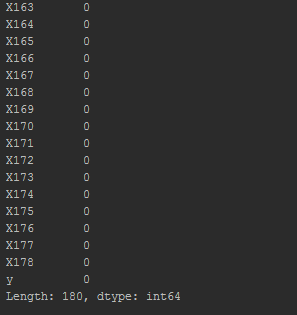


**5.1. data distribution of data among the two classes**

The following graph shows the data distribution of data among the two classes which would be used for binary classification later in the project.

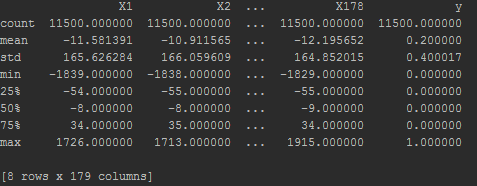
**5.1.** **Preprocessing of data**

1. **Null Value Analysis**

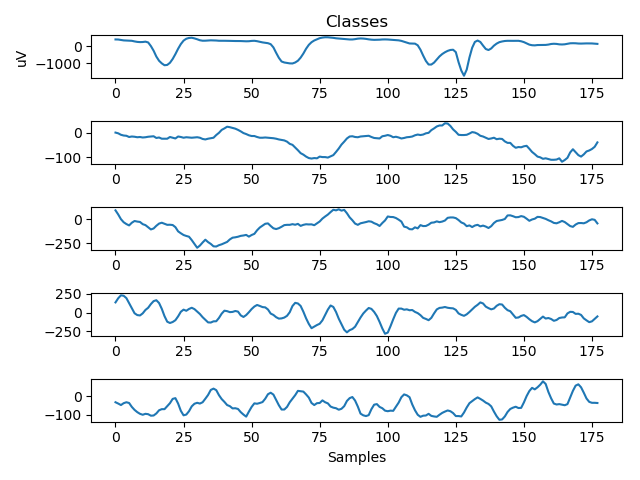


By the figure above it’s completely evident that there are 180 null values in our sample of 11500 tuples with 178 columns.

1. **General Arithmetic Analysis**



1. **Plotting outputs across inputs**



Classes of data here are categorically placed across Y axis as samples of time within a given range.

**5.2. CNN Model decisions**

1. Gaussian Noise was added to improve generalization error which could be problem if not managed well. Adding noise for augmentation can prevent overfitting.
2. Batch Normalization and Dropout were used to decrease the overfitting of the model.
3. Adam optimizer was used instead of the classic stochastic gradient descent as it was especially designed for updating weights of deep neural networks. It is definitely more efficient.
4. The number of hidden convolution layers decrease gradually as the data passed across the network continually complexes. The number of neurons in the first layer is 24, the second has 16 and the third has 8.
5. Dense layers were used to make a robust fully densely connected neural network

**5.3. LSTM Model Decisions**

* 1. Gaussian Noise was added to improve generalization error which could be problem if not managed well. Adding noise for augmentation can prevent overfitting.
  2. Batch Normalization and Dropout were used to decrease the overfitting of the model.
  3. 60 neurons of lstm were used as it was 2/3rd of the number of input features.
  4. Adam optimizer was used instead of the classic stochastic gradient descent as it was especially designed for updating weights of deep neural networks. It is definitely more efficient. Dense layers were used to make a robust fully densely connected neural network.

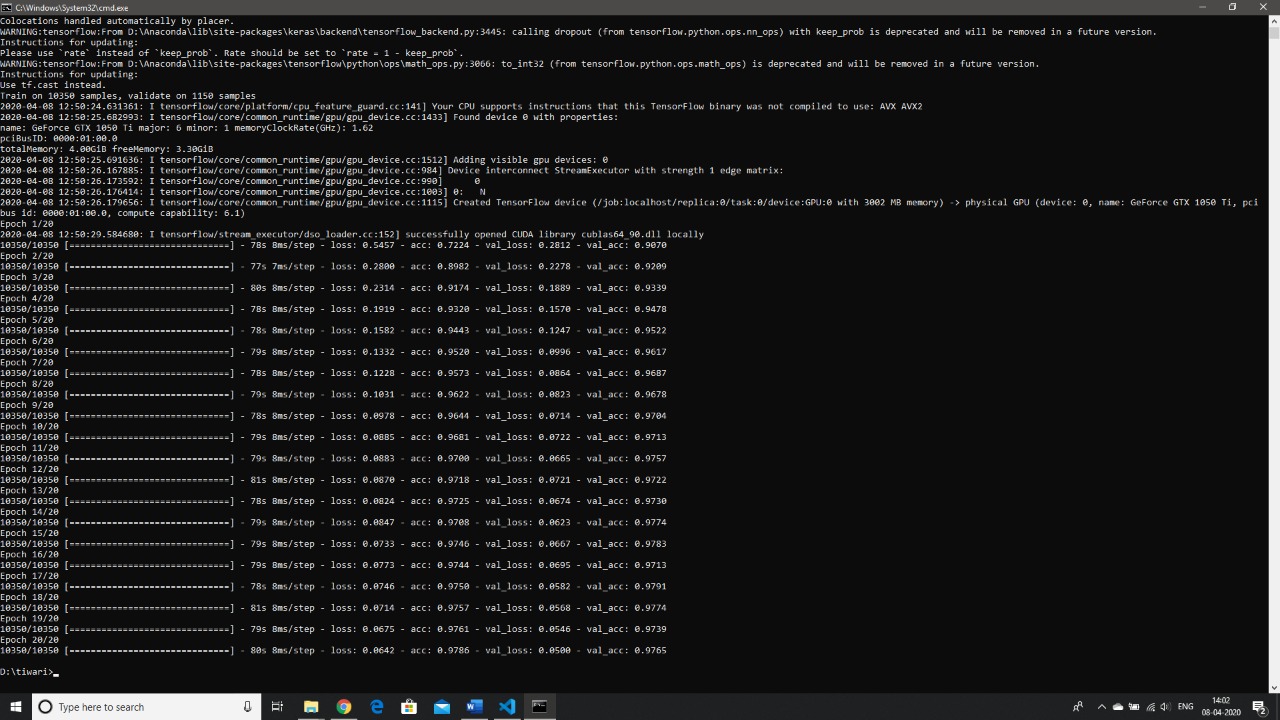
# 6. RESULT

We were aiming for accuracies as good as possible. When we ran the training, we were elated to see 90+ accuracies in both our approaches. It’s worthwhile to state that both the accuracies were achieved in fewer epochs, which in turn means that our models in both the approaches i.e., CNN and LSTM were strong enough to fetch us these accuracies.

LSTM

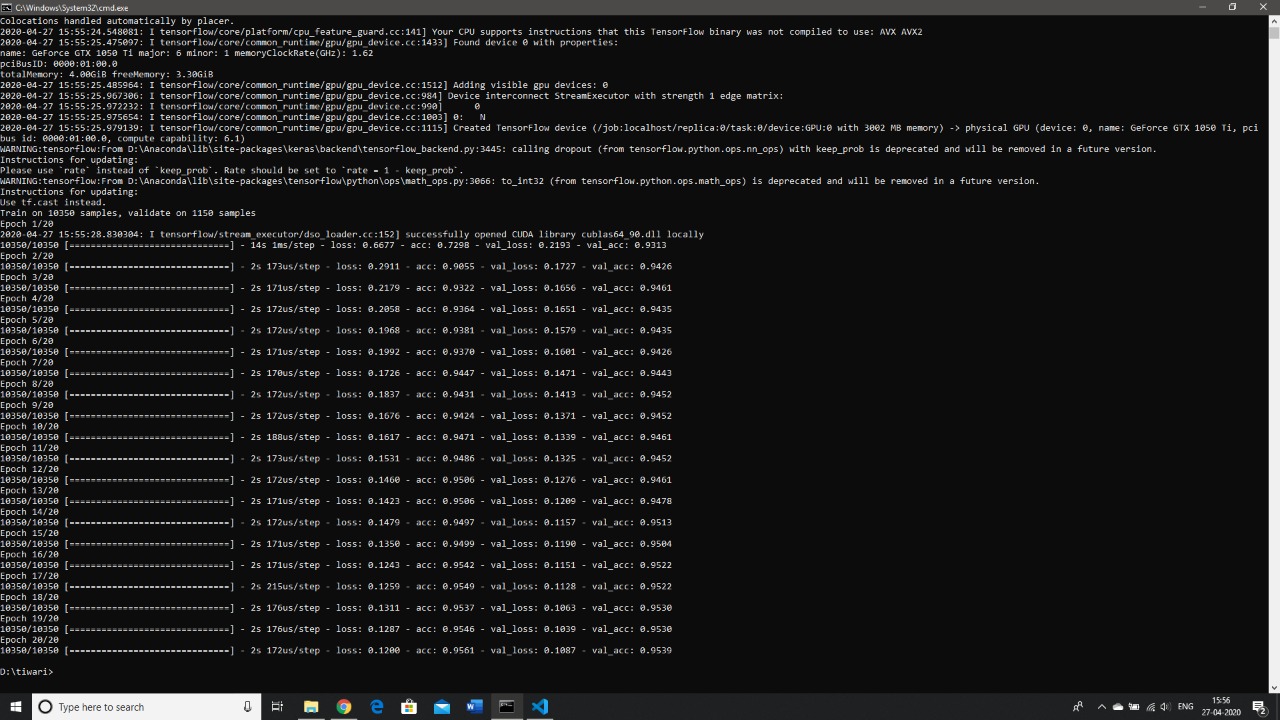
As we can see, the lstm model fetched us an accuracy of 97% in just 20 epochs.

|  |  |
| --- | --- |
| Loss while training | 0.079 |
| Accuracy | 0.97 |
| F1 Score | 0.97 |
| Recall | 0.95 |
| Precision | 0.98 |



**CNN- Accuracy for CNN**

We can see that we achieved 95% accuracy with the given dataset and layer configurations and in 20 epochs.



1. **CONCLUSION**

We introduced a deep learning approach for the automatic detection of epileptic seizures using EEG signals. Compared to the state-of-the-art methods, this approach can learn the high-level representations, and can effectively discriminate between the normal and seizure EEG activities. Another advantage of this approach lies in its robustness against common EEG artifacts (e.g., muscle activities and eye blinking) and white noise. The proposed approach has been examined on the Bonn EEG dataset and compared to several baseline methods. The experimental results demonstrate the effectiveness and superiority of the proposed method in detecting epileptic seizures. It achieves the superior detection accuracies under ideal and imperfect conditions.

**8. FUTURE SCOPE**

In contrast to increasingly customary strategies for machine learning methods, deep learning classifiers are prepared through component adapting as opposed to task-explicit calculations. This means the machine would understand the patterns in the pictures as it is given instead of requiring the human administrator to characterize the examples which the machines should look for in the picture.

The component learning system is utilized each day by the way we show a kid to perceive distinctive objects.

Deep learning is a sort of AI that impersonates the neuron of the neural systems present in the human cerebrum. PC Vision Deep learning models are prepared on a lot of pictures a.k.a preparing information, to explain an undertaking. These deep learning models are for the most part utilized in the field of Computer Vision which permits a PC to see and picture as a human would. This kind of system is like the organic sensory system, with every hub going about as a neuron inside a bigger network. Thus, these models are a class of fake neural systems. This models’ calculations adapt dynamically about the picture as it experiences each neural system layer.

**9. REFERENCES**

1] M. D’Alessandro, R. Esteller, G. Vachtsevanos, A. Hinson, J. Echauz and B. Litt,” Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients, in” IEEE Transactions on Biomedical Engineering, vol. 50, no. 5, pp. 603-615, May 2003.

[2] L. D. Iasemidis et al.,” Adaptive epileptic seizure prediction system, in” IEEE Transactions on Biomedical Engineering, vol. 50, no. 5, pp. 616-627, May 2003.

[3] S. Li, W. Zhou, Q. Yuan and Y. Liu,” Seizure Prediction Using Spike Rate of Intracranial EEG, in” IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol.21, no. 6, pp. 880-886, Nov. 2013.

[4] N. Wang and M. R. Lyu,” Extracting and Selecting Distinctive EEG Features for Efficient Epileptic Seizure Prediction, in” IEEE Journal of Biomedical and Health Informatics. 19, no. 5, pp. 1648-1659, Sept. 2015.

[5] **IEEE standard Book,** J. F. Curtis, (Ed.), *Processes and Disorders of Human Comm- unication.* New York: Harper and Row, 1978.

[6] **Journal Paper,** J. Schroeterand M. M. Sondhi, “Techniques for estimating vocal-tract shapes from the speech signal,”*IEEE Trans. Speech Audio Process.*, vol. 2, no. 1, pp. 133–150, 1994.

[7] **Proceeding Paper, J.** M. Pardo, “Vocal tract shape analysis for children,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 1982, pp. 763–766.