Migraine Prediction Using Deep Learning Model

Anjana Jalannavar
School of Computer Science and Engineering Department
KLE Technological University
Hubli, India
anjanajalannavar7@gmail.com

Suvarna G Kanakaraddi
School of Computer Science and Engineering Department
KLE Technological University
Hubli, India
suvarna_gk@kletech.ac.in

Vidya S Handur
School of Computer Engineering Department
KLE Technological University
Hubli, India
vidya_handur@kletech.ac.in

2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT) | 978-1-6654-5635-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICERECT56837.2022.10059843

Abstract— **A migraine is a headache which could reason very terrible palpitate ache, commonly on one part of the head. Symptoms of migraine are nausea, vomiting, and intense sensitivity to brightness and sound. Migraine attacks that may cause for hours to days, and the ache may be so intense that it interferes with each day ventures. Migraine and episodic headache are one of the most common kinds of primary headache. Migraine is considered as a severe health trouble that impacts the pleasant life. Migraines, which have an effect on youngsters and teens in addition to adults, can developed via 4 stages: prodrome, atmosphere(aura), charge and post-drome. Not everybody who has migraines is going via all stages. The paper tries to focus on the migraine prediction by the use of deep learning model. This paper proposes or provides a Machine Learning technique and model to discover the sort of migraine.**

Keywords— Migraine detection, Machine Learning, deep learning, sequential model, Relu, SoftMax, Adam.

I. INTRODUCTION

In India, the well-being or health-care enterprise is one in every biggest industry, and it has an immediate impact at the nice lifestyles of humans in every country. Health care comprises of diagnosis, taking care of the patients with diseases treatment, and prevention of disease, illness, injury, and different bodily and intellectual impairments in humans. Health care is brought via way of means of general doctors in medicine, dentistry, nursing, pharmacy, allied well-being, and other healthcare providers. The fitness care enterprise, or clinical enterprise, is a region that offers items and offerings to deal with sufferers with curative, preventive, rehabilitative or palliative care.

Migraine is a genetically prompted complicated sickness characterised via way of means of episodes of mild-to intense headache, maximum and regularly causing unilateral and typically related to nausea and brightness and sound sensitivity. EEG (electroencephalography) alerts had been used to symbolize migraine sufferers with (MwA) that is migraine with aura.[2]. The term migraine is derived from the Greek term hemicrania which later transformed into Latin as hemigranea. The French translation of this type of term is migraine. The non-stop tracking of the mind actions is completed via way of means of the usage of Somatosensory Evoke Potentials (SEP) [1]. Migraine is a common reason of incapacity and lack of exertion. Migraine ambushes or attacks the complicated mind activities that spread over hours to days in a recurrent matter. The introduction of private IoT smart gadgets makes it feasible to autonomously gather the environmental data, and locate correlation among various factors and migraine strikes in actual time in order that an correct prediction of an imminent migraine IoT machine learning system may be made, and steps may be taken to save from such migraine triggers.[3]. The maximum kind of migraine is migraine with aura which is (75 of cases). This task evaluates the etiology and pathophysiology of migraine and spotlights the function of the collaborative group in and treating sufferers with this situation as there are numerous sufferers affected by migraine with aura.

Migraine may be labelled into subtypes in keeping with the headache category according to committee which is called as International Headache Society:
Migraine without aura is a reoccurring headache of four to seventy-two hours; commonly unilateral in location, fluctuating in great, mild to intense in intensity, irritated through bodily pastime, and related to nausea and brightness and sound sensitivity (phonophobia and photophobia). Migraine with aura has reoccurring completely reversible triggers, lasting minutes, commonly one or extra of those unilateral signs and symptoms: visualy impaired, sensory, speech, language, motor, brainstem, and retinal which is related to eyes, commonly observed via way of means of headache and migraine signs and symptoms. Chronic migraine is a headache which takes place in fifteen or extra days in a month or greater than three months and has migraine capabilities on at the least eight days in a month. basilar-kind migraine is an unusual form of migraine with aura characterised via way of means of brainstem signs and symptoms including dysarthria, vertigo, or ataxia, without proof of motor weak spot this sort of migraine is commonly related to blood arteries. Sporadic Hemiplegic migraine is a unprecedented subtype of migraine with aura, characterised via way of means of the presence of motor weak point as aura ,it can every so often cause intense circumstance in affected person into coma situation. The motor weak point is frequently followed via way of means of different sorts of aura, like impairment in vision, speech, or sensation. Familial hemiplegic migraine is a kind of migraine headache that pass-via generations to generations. This Familial Hemiplegic migraine takes place due to a change (variation) in a particular gene. This is referred to as familial hemiplegic migraine.Bi-Annual Online Journal (ISSN:2581-611X),Volume 5-Issue 2,Paper 39.

II. RELATED WORK

The non-stop tracking of the mind actions can permit the affected person to reply on time earlier than the incidence of

978-1-6654-5635-7/22/$31.00 ©2022 IEEE

the imminent migraine episode to decrease the severity using Somatosensory Evoked Potentials (SEP) which are applied to display the migraine signs in an ambulatory surroundings aiming to have a processor incorporated on-sensor for power-green and well timed intervention .The proposed early migraine detection classifier is carried out in a CMOS 180nm[1] .A deep learning framework – Migraine-Cloud -which makes use of a front-end migraine detection app for private IoT cell gadgets, and a back-end deep learning neural network to study and eventually expect the onset of migraine for a selected affected person.[3]. Classifying The category of headaches using the Random Forest algorithm carried out in numerous stages [4]. These semi- automatic strategies can probably be utilized in free-residing situations offering directional attention, including computer unit on the office - a improvement which might extensively amplify the operational wrapping of migraine studies. with this technique compared to the JPEG set of rules 8 topics finished the study [5]. The dataset attributable to every approach includes all of the functions selected, and could introduce into the 3 classifiers: SVM, Boosting and enhance the performance of the classifiers with the aid of using themselves and with the committee.[6].Headache disorder is one of the most often illness. At least 50% of the world's population has experienced a headache. Primary headaches have several types; migraine, tension, cluster, and medication overuse. Computer aid for diagnosis could help people locate the headache type without the need to meet the doctor. The Random Forest algorithm was used in this study to produce a reliable model for classifying the headache types and generate feature importance.[9]. To inspect migraine signs and symptoms, there is a advancement in a visible migraine. simulation in digital surroundings (VE) the usage of a VR head-installed display screen. For the migraine detection, advancement of migraine aura blended with an relevance impact [12].Disease Prediction Using Data Mining and Machine Learning Techniques. Advanced Prognostic Predictive Modelling in Healthcare Data Analytics[13]. Sentiment Analysis of Covid-19 Tweets Using Machine Learning and Natural Language Processing[14].COVID-19 Outbreak Prediction using Additive Time Series Forecasting Model[15].Health care is important to live a happy and healthy life.So Chatbots can be helpful in monitoring current health status before visiting the doctors physically. Doctor Chatbot is a means to communicate with man and machines[16].

III. PROPOSED METHOD

The proposed system architecture in Fig 1 depicts the high-level design of migraine prediction starting from giving the input of migraine dataset, sampling the dataset, data Splicing. Classification And Feature extraction.

Fig.1. High-Level Design of Migraine Prediction Model

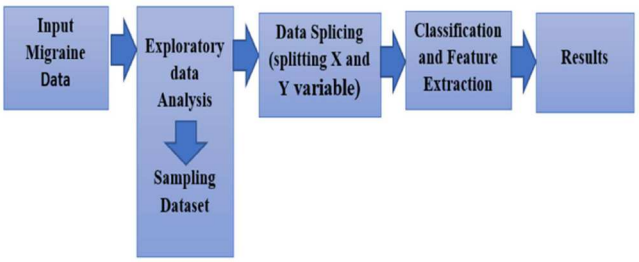


Fig. 2. low-Level design of Migraine Prediction Model

Fig 2 depicts the block diagram of low-level design of migraine prediction. The process includes input of data source that is set of features and selection of correct features based on the object data will occur. After that pre-processing method takes place which will prepare the raw data which will be suitable for machine learning algorithm. In next step classification and identification of the set of features takes place.in last step interpretation of the data takes place then the result is produced and there are some new insights. Here migraine data is given as input, and exploratory data analysis takes place, next sampling of the dataset according to its columns and rows takes place. In next step there is data splicing which is called as splitting of X and Y variables which are called as independent and dependent variables will takes place. classification and identification of types and severity of migraine along with feature extraction takes place. In last and final step Result analysis will take place. Migraine database is collected from the Kaggle repository. It contains 24 set of features for migraine prediction. Using this the training set and testing set is created to train the model. Set of migraine features or symptoms are given as input to predict which type of Migraine it is. Among the type of migraines such as Migraine with aura, migraine without aura, Familial Hemiplegic Migraine, Basilar Typed Aura etc. Diagnosis of migraine can usually be made by medical history of a person. Treatment which includes daily lifestyles changes like diet, exercise, sleeping habits medications including analgesics, serotonin receptor and agonists, calcium channel blockers, and antiemetic. There are set of features of migraine. Age, duration, frequency, location, character, Intensity, Nausea, Vomit, Phonophobia, Photophobia, Dysartia, Vertigo, Tinnitus, Hypocesis, Diplopia, Visual defect, Conscience, Paraesthesia, DPF.

First step in Migraine prediction is data loading, here the number of training examples is 400 and number of features are 24. In next step Exploratory Data Analysis takes place and the graphs are plotted for intensity, frequency and type of migraine. Lossless Sampling of the dataset takes place in next step according to columns and rows. Data splicing of X and Y variables which are independent and dependent variables takes place in the next step. after all this steps Identification and classification of Type and severity and feature extraction of migraine takes place.in Final step there will result analysis.

A. DNN Model Details

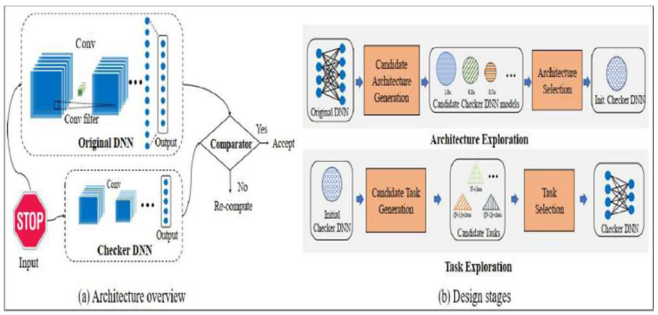


Fig .3. Architecture of DNN (Deep Neural Networks)

The Fig 3 shows the overview of deep neural networks neural community wishes to research all of the time to clear up responsibilities in a extra certified way or maybe to apply numerous strategies to offer a higher result. When it receives new statistics withinside the device, it learns a way to act therefore to a brand-new situation. Learning will become deeper whilst responsibilities can clear up get harder. Deep neural community represents the sort of system gaining knowledge of whilst the device makes use of many layers of nodes to derive high-stage features from enter statistics. It approaches reworking the statistics right into an extra innovative and summary component.

Deep Neural Networks model gives deep understanding of implementation of deep neural networks model and its activation functions. A deep neural network model (DNN) is an Artificial Neural Networks with multiple or many hidden layers which will be between the input and output layers.

B. Hidden layers of DNN

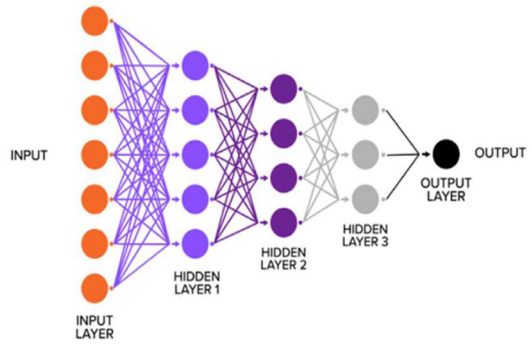


Fig. 4. Hidden layers of DNN Model

The Fig 4 shows hidden layers of (DNN) deep neural networks model. Nodes are little elements of the machine, and they're like neurons of the human brain. When a stimulus hits them, a system takes location in those nodes. Some of them are linked and marked, and a few are not, however in general, nodes are grouped into layers. The machine has to system layers of facts among the enter and output to remedy a task. The greater layers it has to system to get the result, the deeper the community is considered. There is an idea of Credit Assignment Path (CAP) this means that the quantity of such layers wished for the machine to finish the task. The neural community is deep if the CAP index is greater than two. A deep neural community is useful while you want to

update human exertions with self-sustaining paintings without compromising its efficiency. The deep neural community utilization can locate diverse applications.

C.Activation Functions

1). ReLU

The relu function or upright activation function is a piecewise upright function that will give the result input directly, if it is positive or optimal value, and it will output zero if its negative.

Mathematically, it can be written as:

$$f(x) = \max(0, x) \qquad (1)$$

Graphically it can be given as,

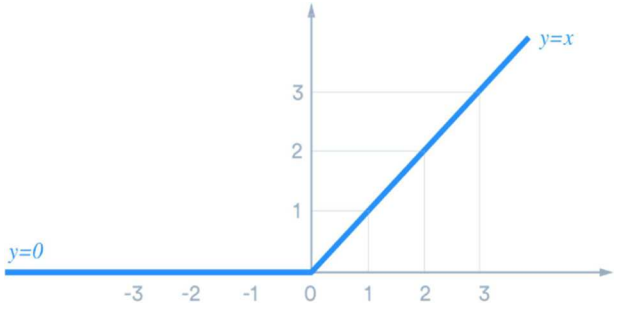


Fig.5. ReLU activation function graph

Advantages of ReLU:

1)Simpler Calculation: Derivative will remain same i.e., 1 for a optimal input which will decrease the time taken for the model to learn and to minimize the errors.

2)Representation Spareness: It is efficient in giving true value as zero as the output.

3)Length: upright activation functions are easier to optimize and allow for a smooth flow. Therefore, it is best suited for supervised tasks on the bigger sets of labelled data.

2)SoftMax function:

The SoftMax feature takes a vector z of K actual numbers, and normalizes it right into a possibility distribution which includes k number of possibilities so as to be proportional to the exponentials of the entered numbers. That is, previous to making use of SoftMax activation feature, a few vector additives may be negative, or more than one; and won't sums to one. however, after using SoftMax activation feature each, element may be with inside the interval, and the additives will be added as much as 1, so they may be interpreted as possibilities. Furthermore, the bigger enter additives will correspond to large possibilities.

The standard (unit) SoftMax function is defined when by the formula

$$\operatorname{softmax}(z_j) = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \text{ for j} = 1,\ldots,K \tag{2}$$

σ = softmax

$\vec{z}$ = input vector

e^{z_i} = standard exponential function for input vector

K = number of classes in the multi-class classifier

e^{z_j} = standard exponential function for output vector

e^{z_j} = standard exponential function for output vector

2)Stochastic Gradient:

Fig 6 indicates the graph of stochastic gradient descent. Stochastic gradient descent (SGD) is a form of gradient descent that runs one education instance consistent with iteration. Or in different words, it approaches a education epoch for every instance inside a dataset and updates every education instance's parameters one at a time. As it calls for most effective one education instance at a time, as a result it's miles less complicated to keep in allotted memory. However, it indicates a few computational performance losses in evaluation to batch gradient structures because it indicates common updates that require greater element and speed. Further, because of common updates, it's also handled as a loud gradient. However, from time to time it may be useful in locating the worldwide minimal and additionally escaping the neighborhood minimal.

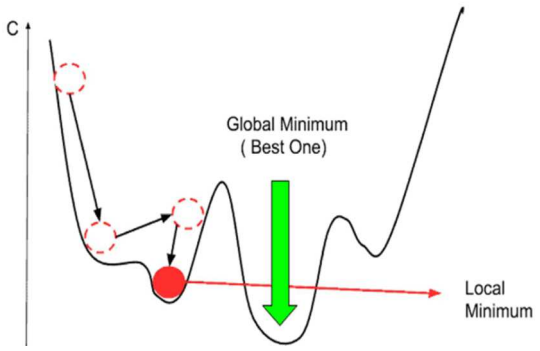
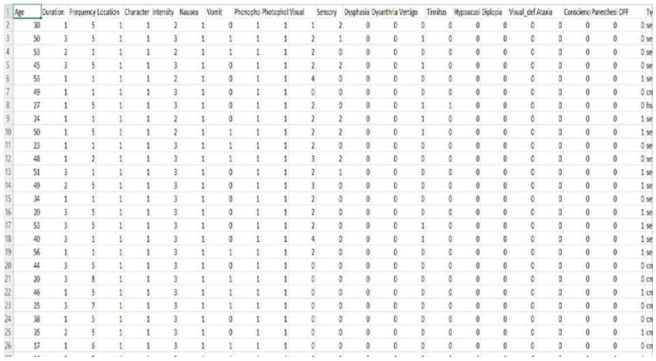


Fig.6. Stochastic gradient descent

IV. RESULT ANALYSIS

Result gives an analysis of Migraine features and its type, and severity. In Exploratory Data Analysis, correct identification of data takes place and graphs are plotted for intensity, frequency, graph of migraine type and severity, bar graph of age and count and duration of migraine.

Table .1. Migraine dataset



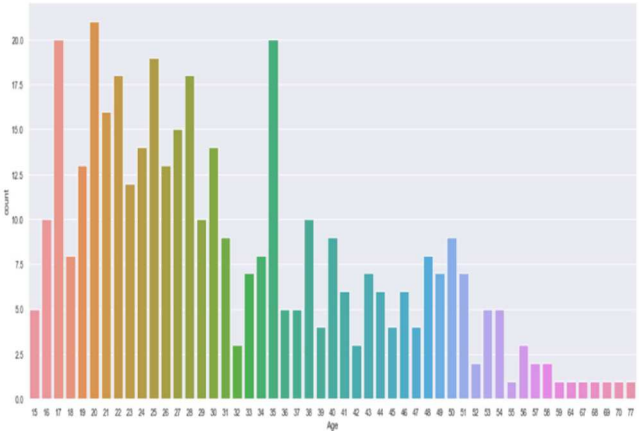


Fig .7. Age and Count of Patients Migraine dataset

Fig 7. shows the Bar graph of the age and count based on patient's migraine dataset after the extraction of the features from the provided dataset.

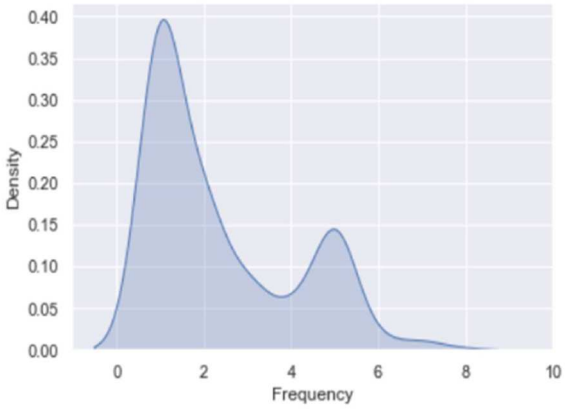


Fig .8. Frequency and density of Migraine

Fig 8 shows the graph of the frequency and its density based on patient's migraine dataset after the extraction of the features from the provided dataset.

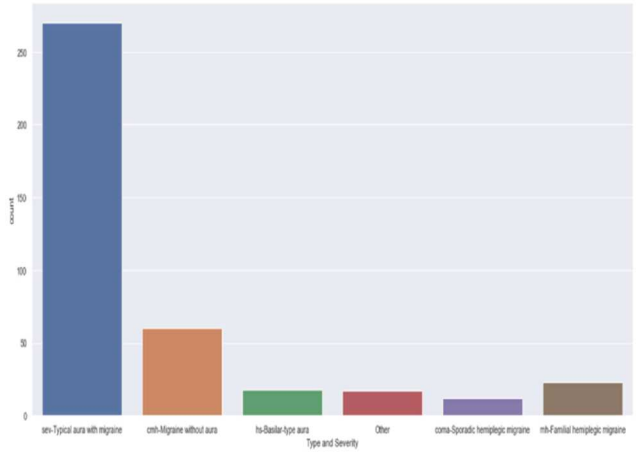


Fig. 9. migraine Type and Severity

Fig 9 shows the graph of the migraine type and its severity based on patients' migraine dataset after the extraction of the features from the provided dataset.

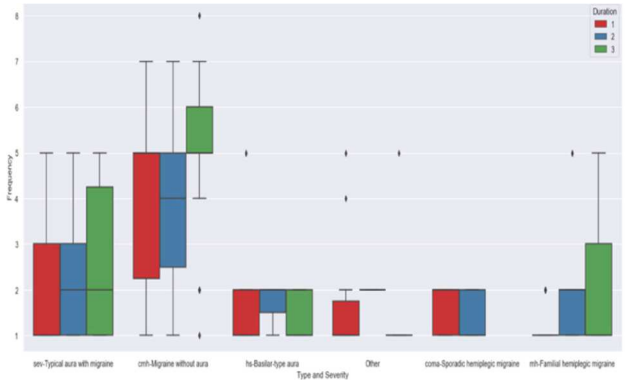


Fig .10. Duration of Migraine

Fig 10 shows the graph of the frequency and its duration based on patient's migraine dataset after the extraction of the features from the provided dataset.

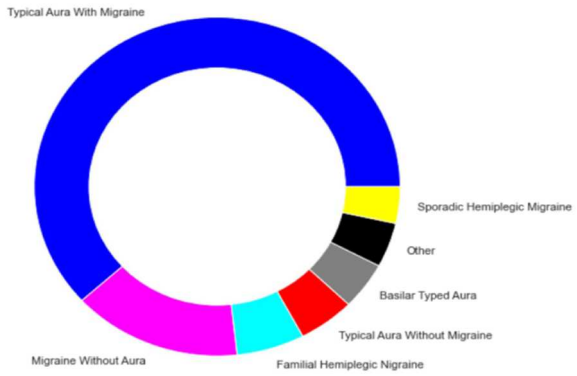


Fig .11. Migraine Types

Fig 11 shows the Pie graph of the Migraine Type based on patient's migraine dataset after extraction of the features from the provided dataset.

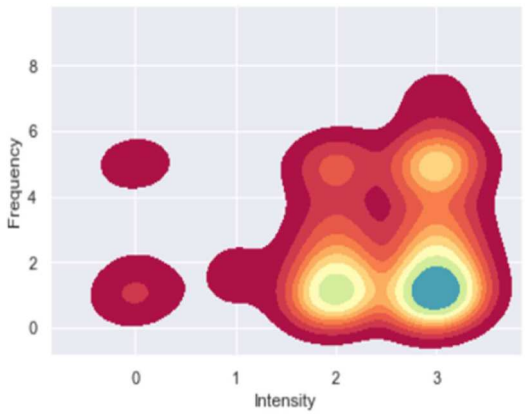


Fig .12. Intensity of Migraine

Fig 12 shows the Intensity graph of the Migraine Type based on patient's migraine dataset after the extraction of the features from the provided dataset.

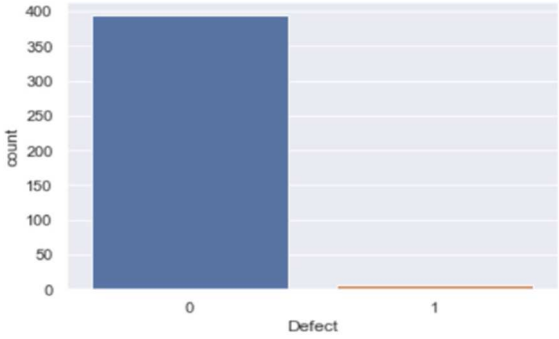


Fig.13. Defect and its count in Migraine

Fig 13 shows the defect and count graph for migraine based on patient's migraine dataset after the extraction of the features from the provided dataset.

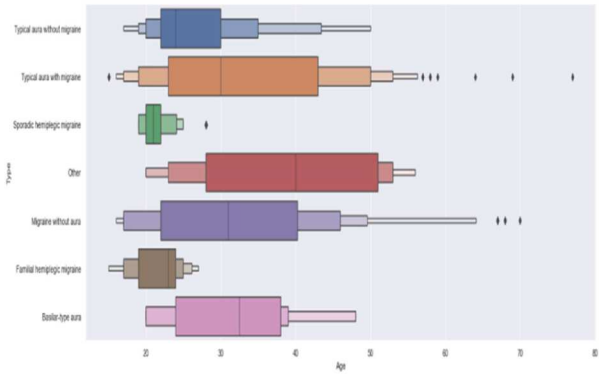


Fig .14. Age and Type of Migraine:

Fig 14 shows the Age and its type graph for migraine based on patient's migraine dataset after extraction of the features from the provided dataset.

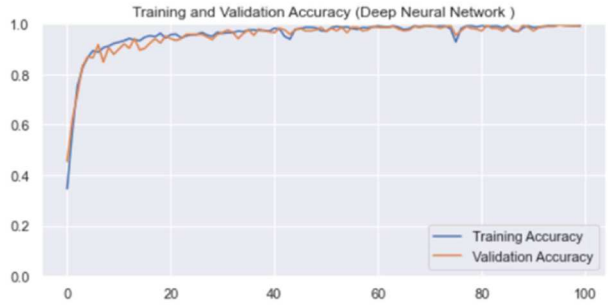


Fig.15. Training and Validation Accurate results of DNN model

Fig 15 shows the Graph of validation and accuracy of DNN Model and blue line in the graph will show the Training Accuracy and red line in the graph will show the Validation accuracy.

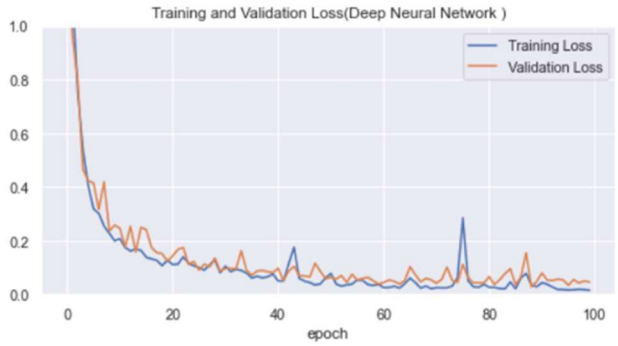


Fig .16. Training and Validation loss of DNN model

Fig 16 shows the Graph of validation and accuracy loss of Deep Neural Network Model and blue line in the graph will show the Training Accuracy loss and red line in the graph will show the Validation accuracy loss.

Classification Report of Deep Neural Network

	precision	recall	f1-score	support
0	1.00	1.00	1.00	47
1	0.91	1.00	0.95	31
2	1.00	1.00	1.00	40
3	1.00	1.00	1.00	41
4	1.00	1.00	1.00	36
5	1.00	0.94	0.97	52
6	1.00	1.00	1.00	43
accuracy			0.99	290
macro avg	0.99	0.99	0.99	290
weighted avg	0.99	0.99	0.99	290

Fig .17. Accuracy, precision, recall, f1-score, support obtained by DNN model for all classification

Fig.17 shows the classification report of the Deep Neural Network which contains precision recall, f1-score and support with their accurate results, then it gives the macro accuracy and weighted average which are considered for all classification.

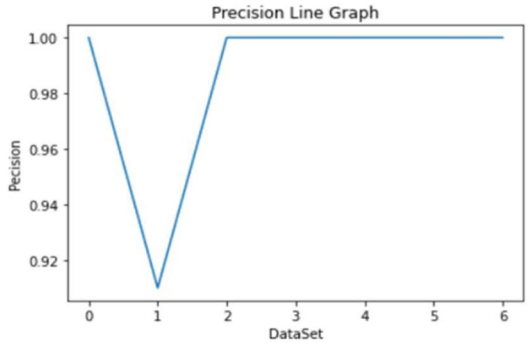


Fig .18. Precision Line Graph

Fig 18 Fig 18 indicates the precision line graph. Precision is one degree of a device mastering model's execution which gives excellent prediction made via way of means of the model.
Precisions can be referred as (the number of true positives plus number of false positives).

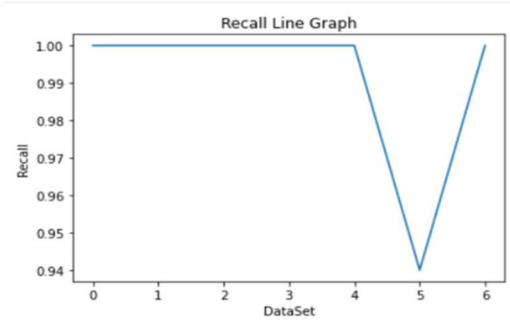


Fig .19. Recall Line Graph

Fig 19 depicts the Recall line graph. Recall is a standard that specify the no of correct positive or optimal predictions which are made of all positive or optimal predictions that could have been made.

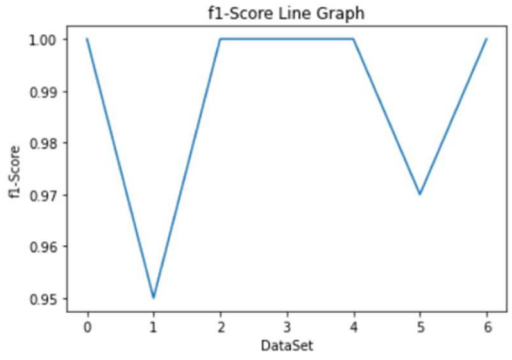


Fig .20. f1-score Line Graph

Fig 20 depicts the f1-score Line graph. The F1-score will combine the precision along with recall of a classifier into one standard by taking their harmonic mean.

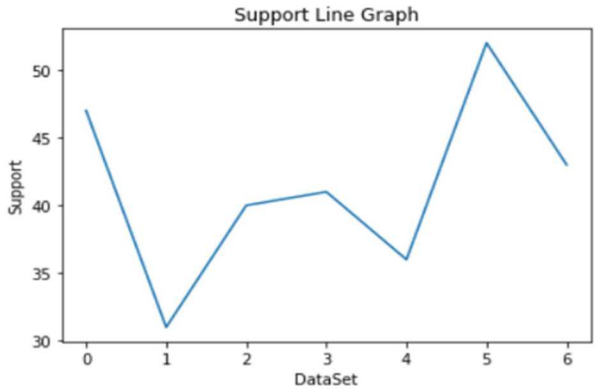


Fig .21. Support Line Graph

Fig 21 depicts the Support Line graph. Support is the number of real occurrences of the class in the defined dataset. Imbalanced support in the training data may indicate weaknesses in the scores of the classifier and could indicate the need for sampling or rebalancing.

```
...
Epoch 96/100
37/37 [==============================] - 0s 6ms/step - loss: 0.0158 - accuracy: 0.9965 - val_loss: 0.0335 - val_accuracy: 0.9
966
Epoch 97/100
37/37 [==============================] - 0s 5ms/step - loss: 0.0166 - accuracy: 0.9948 - val_loss: 0.0545 - val_accuracy: 0.9
931
Epoch 98/100
37/37 [==============================] - 0s 6ms/step - loss: 0.0185 - accuracy: 0.9931 - val_loss: 0.0418 - val_accuracy: 0.9
931
Epoch 99/100
37/37 [==============================] - 0s 7ms/step - loss: 0.0172 - accuracy: 0.9922 - val_loss: 0.0495 - val_accuracy: 0.9
931
Epoch 100/100
37/37 [==============================] - 0s 6ms/step - loss: 0.0147 - accuracy: 0.9948 - val_loss: 0.0457 - val_accuracy: 0.9
897
```

Fig.22. Loss and Accuracy after 100 Epochs of DNN Model.

The following Fig 22 shows the Loss and Accuracy after the 100 Epochs of Deep Neural Network Model. After 100 Epochs the value accuracy will be 0.9 and the value loss will be 0.0457. These epochs will increase the accuracy of the deep neural network model.

```
Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_8 (Dense)             (None, 512)               12288

 dense_9 (Dense)             (None, 256)               131328

 dense_10 (Dense)            (None, 128)               32896

 dense_11 (Dense)            (None, 7)                 903

=================================================================
Total params: 177,415
Trainable params: 177,415
Non-trainable params: 0
```

Fig .23. Model Summary with Trainable and Non-Trainable Params of DNN Model.

This Following Fig 23 shows Model summary with Trainable and Non-Trainable Parameters of deep Neural Networks. It consists of layer(type) there are dense layers and output shape and number of parameters. The Total parameters are 177,415 and Trainable parameters are 177,415 and 0 non-Trainable parameters.

V. CONCLUSIONS

Through this research developed a deep learning model to identify the type of Migraine and its severity. Here sequential model is used to train the dataset. and the features are extracted to identify the type of migraine. In Exploratory Data Analysis, correct identification of data takes place and graphs are plotted for intensity, frequency and type of migraine. Lossless Sampling of the dataset takes place in next step according to columns and rows. Data slicing of X and Y variables which are independent and dependent variables takes place in the next step. after all this steps Identification and classification of Type and severity and feature extraction of migraine takes place.in Final step there will result analysis. The deep learning model provides the best approach and has effective results in the field of machine learning which makes a great difference classification and identification of features of migraine. The dataset here considered were less so in the coming days more dataset around 1000 or more than that will be produced. Future scope is to collect more migraine dataset from the hospitals and train and test them using Deep learning sequential model.

REFERENCES

[1] Garingo, Mario, Sahba, Farhang, Ma, Weijing, Published in: International Conference on Computational Science and Computational Intelligence (CSCI) 06 December 2018.

[2] “Migraine-Cloud: A Machine-Learning IoT Framework for Capturing Triggers and Predicting Migraines”.

[3] Shrey Mohan, Arindam Mukherjee, Published in Conference: Southeast Con 04 October 2018. Dharamsalam, Adi Wibowo, Fajar Agung Nugroho, Eko Adi Sarwoko.Published in: 2020 4th International Conference on Informatics and Computational Sciences (ICICoS) 24 December 2020.

[4] Ioannis Pavlidis, Ivan Garza, Panagiotis Tsiamyrtzis, Malcolm Dcosta, Jerry W Swanson. Published in: IEEE journal on biomedical and health informatics in-May 23, 2019.

[5] Yolanda Garcia-Chimen, Begonya Garcia Zapirain, Marian Gomez Published in: BMC Medical Informatics and Decision-Making 13 April 2017.

[6] Akhila jagarlapudi; amey patil; d. p. rathod Published in 2021 IEE international conference on distributed computing, vlsi, electrical circuits and robotics (discover).

[7] Li-Wei Ko; Kuan-Lin Lai; Pei-Hua Huang; Chin- Teng Lin; Shuu-Jiun Wang Published in: 2013 6th International IEEE/EMBC Conference on Neural Engineering (NER).

[8] Çigdem Gülüzar Altıntop; Fatma Latifoğlu; Aykut Yalabık; Kübra Keser; Ersin Kasım.Ulusoy. Published in: 2017 10th International Conference on Electrical and Electronics Engineering (ELECO).

[9] Dhiyaussalam, Adi Wibowo,Fajar Agung Nugroho Published in: 2020 4th International Conference on Informatics and Computational Sciences (ICICoS) 24 December.

[10] Larysa Malanchuk, Mariia Riabokon,Artem Malanchuk Published in: 2020 10th International Conference on Advanced Computer Information Technologies (ACIT) 30 September 2020 .

[11] Leiqiang Yao; Rong Chen; Hui Ji Published in: IEEE Transactions on Neural Systems and Rehabilitation Engineering in 18 August 2022.

[12] Sebastian Misztal; Guillermo Carbonell; Lysann Zander; Jonas Schild Published in: 2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH).

[13] Kanakaraddi, S.G., Gull, K.C., Bali, J., Chikaraddi, A.K., Giraddi, S. (2021). Disease Prediction Using Data Mining and Machine Learning Techniques. Advanced Prognostic Predictive Modelling in Healthcare Data Analytics. Lecture Notes on Data

Engineering and Communications Technologies, vol64. Springer,Singapore.
[14] Kanakaraddi, S.G., Chikaraddi, A.K., Aivalli, N., Maniyar, J., Singh, N. (2022). Sentiment Analysis of Covid-19 Tweets Using Machine Learning and Natural Language Processing.Proceedings of Third International Conference on Intelligent Computing, Information and Control Systems. Advances in Intelligent Systems and Computing, vol 1415. Springer,Singapore.
[15] Gull, K.Kanakaraddi, S.G, & Chikaraddi, A. (2022). COVID-19 Outbreak Prediction using Additive Time Series Forecasting Model. Trends in Sciences, 19(22), 1919.
[16] Nehabanu H. Harlapur & Vidya Handur ,Smart Innovation, Systems and Technologies book series (SIST,volume 311),First Online: 01 October 2022.
[17] Nehabanu H. Harlapur & Vidya Handur,Bi-Annual Online Journal(ISSN:2581-611X),Volume 5-Issue 2,Paper 39 August 2022.