An Approach For Finding Emotions Using Seed Dataset With Knn Classifier

K.N.V. Satyanarayana[1], Dr.T. Shankar[2], Dr.P.V.Rama Raju[3]

Research scholar Annamalai University1, Assistant Professor Annamalai University2, Professor S.R.K.R. Engineering College3
knv.satyan@gmail.com

Abstract: Emotions are an indispensable component of our daily lives. Nonetheless, brain-computer interface (BCI) systems have not achieved the requisite level to interpret emotions. Programmed feeling acknowledgment based on BCI frameworks has been a point of extraordinary inquiry within the final few decades. Electroencephalogram (EEG) signals are one of the significant assets for these frameworks EEG may be a physiological flag recorded from brain work out within the frame of brain waves through the scalp. The most advantage of utilizing EEG signals is that it reflects the genuine feeling and can effectively be prepared by computer frameworks. A dummy dataset can be used and filled with EEG data to compute and categorize these signals generated from EEG signals. The dataset used here is Seed, and it can be accustomed by a machine learning technique called the K-nearest neighborhood (KNN) algorithm to systematize the data. Experimental performance achieved through the categorizing values of 94.06% during the classification in the Seed. This proposed method shows that emotion recognition like positive, neutral, and negative is possible through EEG signals.

1. Introduction:

Emotions are organic states related to all of the nerve frameworks brought on by neurophysiological changes differently related to contemplations, sentiments, behavioral reactions, and a degree of joy or disappointment.[1] Emotions can furthermore be grouped as negative, positive, and neutral feelings. Essential positive feelings of cheery and joy are fundamental for constancy, improvement, advancement. Essential negative feelings, counting pity, annoyance, nauseate, and panic, ordinarily work consequently about inside a short period. Brain-computer Interface is a developing inquiry about the field for the past few long years.[2] BCI frameworks include an examination of [3] Electro-Encephalogram signals from the brain. The concept of BCI started as helping people with physical and physiological clutters over a decade. Presently the research in BCI has amplified in different areas including normal individuals. A few of such applications are classification of unusual brain movement, epilepsy discovery, emotion detection, etc. Admitting the feeling from physiological signals in a general sense with EEG has provided consideration from investigators. EEG is the program which is more apt for the signal values collection because of its more transient Resolution, Reliability, and security and free to utilize. EEG has high capability of determination and is highly electric. EEG signals endure originating out affectability created by aperture blinking, eye developments, Pulsation, fibrous works out, and control way obstructions. An EEG could be a particular moderately an organic signal. It can be the degree of the electrical movement of the cerebrum, carried out by arranging many cathodes above the skull. Recently, EEG signals have picked up consideration due to their accessibility. Nowadays, new EEGs are emerging using mobile technology, EEG gadgets within advertisements that are versatile, reasonable, which are simple to utilize. Examining EEG signals is an intriguing way which comprises distinctive investigate zones in information retrieval, data processing, cognitive brain study, wellbeing, and therapeutic study and profile medical building EEG deployed signal acknowledgment is widely utilized for amusement, online course study, and medical nursing usages. EEG is employed for distinctive motives used in illustration, moment informing, virtual sports, reinforced treatment, and brain research. EEG signals which are obtained after feeding it to the SEED dataset in order to determine a feeling concurring to the valence method and various other models. [4]Support Vector Machine (SVM),[5] k-nearest neighbor (k-NN), and artificial neural arrange (ANN) classifiers are used in the grouping the feelings and acknowledge the emotion.SJTU Emotion EEG Dataset in which here it used seed dataset, is a stacking of EEGs datasets driven by the Some intelligence Lab inquire about the facility which is driven by Prof.Bao-Liang Lu.

The title is acquired from the primary adaptation of the dataset, but presently we offer not as it was emotion, but moreover carefulness of the dataset. [5] The drama recordings are deliberately chosen to initiate distinctive sorts of emotions, which are positive, negative, and impartial ones.[6] The SEED-IV is an advancement of the first SEED dataset. The category number of emotions alters to four: cheerful, sad, fearful, and unbiased. In SEED-IV, we offer not as it were EEG signals, but moreover, eye development highlights recorded by SMI eye-tracking glasses, which makes it a well-shaped multi-modal dataset for feeling acknowledgment. The SEED-IV is an advancement of the first SEED dataset. The category number of emotions alters to four: cheerful, sad, fearful, and unbiased. The SEED-VIG dataset is orientated at investigating the carefulness estimation issue. We built a virtual driving framework, in which a tremendous screen is put before a genuine car. Subjects can play a driving amusement within the car, just as in case driving within the real-world environment. In SEED-IV, we offer not as it were EEG signals, but moreover, eye development highlights recorded by SMI eye-tracking glasses, which makes it a well-shaped multi-modal dataset for feeling acknowledgment.[7]The classifier utilized is the K-nearest neighbor Classifier. Better precision of almost 95% is obtained in gamma frequency groups and exactness was expanded when the number of channels was expanded When the number of channels was expanded. Based on the classifiers and data sets a survey paper is helpful to find the emotions in the best way with the help of existing and futuristic parameters [9].

APPLICATIONS OF EEG SIGNALS:

The major applications can be listed as Neuromarketing, Medicalpurposes, Socialintercourse, functional design and designing of user-friendly systems.
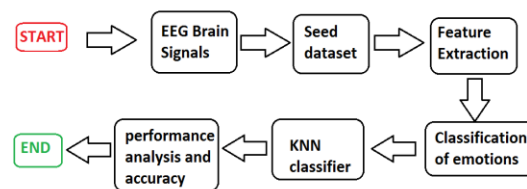
2. Process Flow:

Fig:1 flow diagram

3. Datasets:

For the Classification of emotion Purpose, there are many datasets available like deep dataset, seed dataset, etc.;

• SEED dataset: SEED dataset:[5]SEED is brief for the SJTU Feeling EEG Dataset. The SEED consists of 15 Chinese Subjects' (7males and eight females, cruel matured: 23.27, std: 2.37) EEG signals reported and categorized as they taken after 15 film clips. Downsampled the EEG information to 200 Hz. A bandpass rate filter from to 75 Hz was actualized. Amid criticism, individuals were told to report their enthusiastic responses to each film clip by completing a survey instantly after observing every clip.The chosen recordings may be caught on by not giving clarification and inspire a unique carvedquarry feeling. Hence, in these tests, it can be utilized as the names of trials rather than the surveys' data. The passionate titles contain positive, unbiased, and negative qualities.

4. Classifying of emotions:

They are utilizing the preparing dataset to encourage superior boundary circumstances which may be utilized to choose each quarry course. If the   boundary circumstances are decided, the following assignment is to anticipate the quarry course. The complete handle can be declared as a classification process.
Emotions are being majorly divided into three parts.

- Positive
- Neutral
- Negative

Positive emotion includes Joy, Happiness, Curiosity, etc., and Neutral indicates Normal Calm behavior. Whereas Negative has Anger, sadness, anxiety, depression, etc.; These can be classified using learning algorithms. This can include   mainly triplet groups of learning.

<u>4.1. Supervised learning</u>:

This is a process of study which includes a teacher for categorizing the data. Administered process of study, as the title illustrates, the director's closeness as a teacher. They are coordinated learning instruments or instruct the device utilizing information that's named in a clarity manner, which proposes much information that is presently marked as a right solution.
Besides that the device is provided with recent spreads. The administered study calculation investigates the preparing information(location for preparing cases) along with the   delivered and rectified result obtained by the labeled information.
Directed learning bargains with or knows with "labeled" information, which suggests that a few information is as of now labeled with the right reply.
Types:-
- Retrogression.
- Linear Retrogression
- Categorizing and Grouping
- Naive Bayes Classifiers
- K-NN
- Order nodes

4.2. UnSupervised learning:
This learning can be defined as the preparation in which machines utilizing data that are not one or the other grouped or named and permitting the calculation to perform upon that data which is directionless.In this process the device's assignment is for the purpose of gathering the unarranged data agreeing to likenesses, designs and for preparing the information.Here there is no teacher or guide present to categorize the data.Not at all like administered learning, no instructor is given meaning no preparation will be given to the machine. Subsequently, machines are confined to discover the covered up process in unnamed information on one's own.

Types :-
- Grouping.
- Dividing based on the certain parameters.
- Clustering with incoherence property
- Parallelism.
- Likelihood

4.3. Reinforcement learning:

This learning could be a zone of Machine Learning. It is around taking direct action to maximize in a particular circumstance. It is utilized by diverse computer programs and machines to find the driving conceivable behavior or way in a specific case Fortification learning contrasts from administered learning so that in distributed learning, the preparing information has the reply key with it, so the show is prepared with the proper reply itself. In differentiate, there's no reply in reinforcement learning, but the fortification specialist chooses what to do to perform the given errand. Within the nonattendance of a preparing dataset, it is bound to memorize from its experience.In this for the unsupervised data, Linear regression is being performed.

5. Emotion classifiers:
A calculation which plots and joins the provided data of information within a particular group is known as a classifier. This process of grouping shows that the data can be characterized as a demonstration which strives to pull out a few results obtained from the provided data for the means of preparing. This will foresee the course denomination/group in the purpose of unused information.

5.1.Organized Regression:-
Like the confounding as the title would be, it will be guaranteed. Calculated Relapse may be a categorizing and relapse free calculation.This gages non continuous values (Parallel answer like zero or one, True/False, accept/reject)depending upon the provided information of the dependent free fickle(s). To explain it in a more brief manner it is the process of forecasting the likelihood of the event of a situation or an occasion by making the information fit into a model of work. Subsequently, this can be additionally recognized as model relapse.These answers could be continuously present inside, and one as a reason it forecasts the occurrence of events..

5.2.KNN (k- Nearest Neighbors):-
K closest neighbors may be a straightforward calculation utilized for two of categorizing and relapse issues.It stacks each and every accessible way to categorize the latest circumstances by a lion's share electing its nearest persons and things. The subject allocated to the lesson is the more traditional among the K closest values or things which are considered and calculated by removed work of certain models.Whereas the other triplet past partitioned capacities are utilized for non discrete values, A model of partitioned work is utilized for encoding the numbers and estimating the model.

6. KNN Algorithm:
K-nearest neighbors (KNN) calculation could be a directed ML calculation that can be utilized for both classifications and prescient relapse issues. In any case, it is primarily used for proactive classification issues in the industry. The taking after two properties would characterize KNN well −
6.1.Sluggish study of Computing − KNN may be the sluggish study calculation as it does not include a specialized preparation stage and employs each piece of information for training purpose during categorizing the data.
6.2.Parametric free process for the study of grouping the data − KNN is a parametric free study computation method because it avoids accepting anything around the general data.

KNN works by finding the separations between an inquiry and all the cases within the information, selecting the desired examples (K) closest to the query at that point votes for the original visit name (within the case of classification) or midpoints the terms (within the chance of Relapse). The k-nearest neighbors (KNN) calculation could be a machine learning calculation that can be utilized to unravel both classification and relapse problems.

7. Working of KNN Algorithm:
K-nearest neighbors (KNN) calculation employments 'feature similarity' to foresee the values of current information focuses, which shows such unused information can be allotted esteem based on how approximately it can equivalent with the instructing the data . This can be done by it's working with the assistance of taking after steps −

Step 1 − To carry out any calculation, any one requires a dataset. For that amid the primary way of K nearest neighbour ,it must be arranged for checking the data which has information in it.

Step 2 − The next process is to choose an esteem value of K, i.e., the nearest information focuses. K may be any integer.

Step 3 − The point within the each testing data select the data and perform the below methods −

   3.1 − Compute the separate in middle of the testing data and for every separation of trained data including the assistance for some specific process. The foremost commonly exercised program to calculate separately is the Euclidean process.

   3.2 − Presently, depending on the separate esteem, categorize them in climbing structure..

   3.3 − Following, it can select the best K lines from the categorized and arranged structure.

   3.4 − Presently, knncan relegate the lesson in the testing process based on these rows.

Step 4 − End.
- The exactness of the over classifier increments as we increment the number of information centers interior the arranging set
- In KNN calculation is based on including closeness choosing the proper regard of K may be a handle called parameter tuning and is essential for more superior exactness. Finding the respect of k isn't simple.

8. Focal points of KNN algorithm:
- Quick calculation time.
- Simple calculation – to interpret.
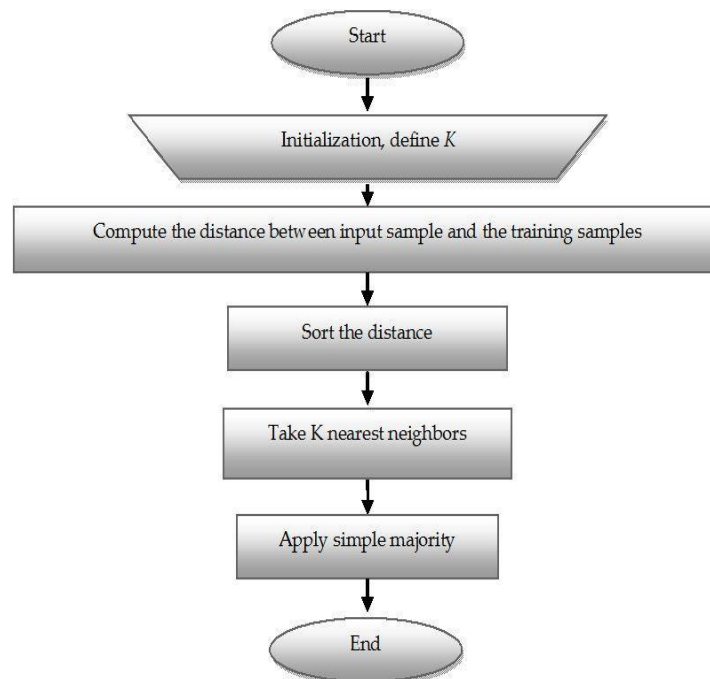- High accuracy.

9. Flow Chart:

Fig: 2 flow chart

10. Classifier results and analysis:

Feelings are classified utilizing a synchronized brainwave dataset; the dataset used here contains a file named seed data.csv.Precision can be defined as the correctness in the middle of facts,figures of particular data and forecasting the details , then recall is the similar to that of True Positive rate and Values, where it can be defined as the degree of victory in compartmenting the feelings, .

The obtained best results are given below:

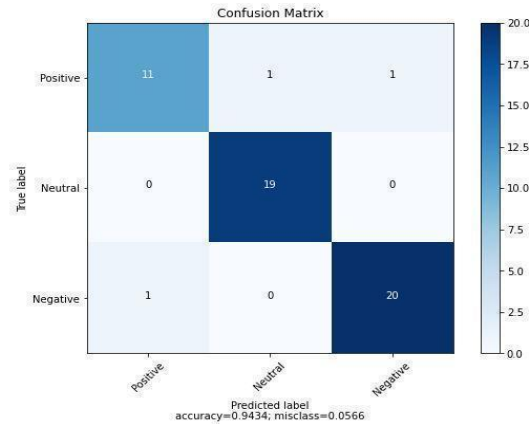● **Training Accuracy Score:0.9433962264150944**

10.1For confusion matrix:

Fig: 3confusion matrix 1

Classification:

	Precision	recall	f1-score	support
0		0.83	0.91	18
1	1.00	1.00	1.00	16
2	0.86	1.00	0.93	19
accuracy			0.94	53
macro avg	0.95	0.94	0.95	53
weighted avg	0.95	0.94	0.94	53

Table 1 classification results

● **Training Accuracy Score: 0.88645321**

10.2.For confusion matrix:
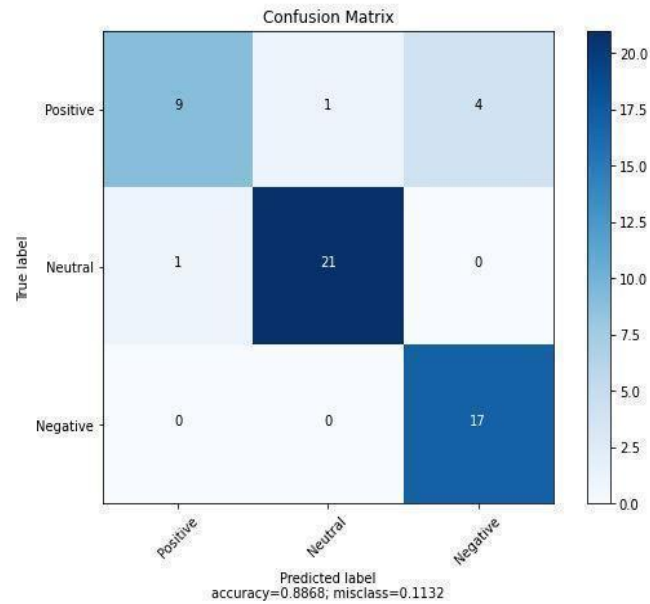
Fig:4 confusion matrix 2

Classification

	Precision	Recall	f1-score	support
0	0.79	0.85	0.81	13
1	0.95	0.91	0.93	23
2	0.94	0.94	0.94	17
accuracy			0.91	53
macro avg	0.89	0.90	0.90	53
weighted avg	0.91	0.91	0.91	53

Table 2 classification results

The disarray network from classification result appeared on Tables and the esteem of f1 score, Recall, Precision were obtained from the most excellent parameter of each test.

Emotion classification results using KNN:-
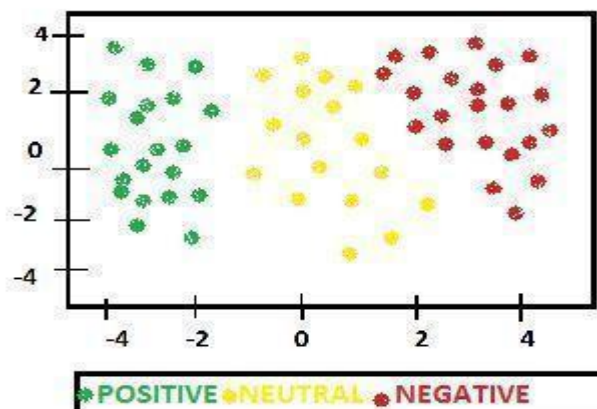
- The above results are obtained by using the Knn classifier.

11. Conclusion and future scope:-

Feelings are exceptionally imperative in human choice dealing with interaction and cognitive handle. Here, we have proposed an EEG based feeling classification strategy that has effectively recognized three emotional states, precisely normal, outrage, and joy. Hence the feelings are classified utilizing Machine learning classification calculations such as Softmax classifier, K-NN classifier, Choice tree classifier, SVM(Support Vector Machine)with the assistance of neural systems. The result's exactness must be considered to urge better performance by having more covered up layers within the neural organization. The accuracy obtained from the KNN classifier 94.06% gives superior results compared to other machine learning classification calculations. Thus, utilizing EEG information might be more exact than behavior information when used to foresee human feelings.The ponder will be done to look at more feeling classes with more extensive datasets to analyze its effect on the brain in the future. It has been challenging to discover highlights that superbly work on feelings as the number of subjects increases. We would be interested in investigating highlights that work on time-frequency space signals. We'll connect and test with other different level classifiers, combination approaches, and Development overhauls done with the sentiments, organized utilizing picture and video dataset using significant learning computation to improve the classification execution.

REFERENCES:

[1].Giannakaki K, Farmaki C. Emotional state recognition using advanced machine learning techniques on EEG data.

[2] Jenke, Robert, Angelika Peer, and Martin Buss. "Feature extraction and selection for emotion recognition from EEG." IEEE Transactions on Affective Computing 5.3 (2014)

[3].Guo G., Wang H., Bell D., Bi Y., Greer K. (2003) KNN Model-Based Approach in Classification. In: Meersman R., Tari Z., Schmidt D.C. (eds) On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE.

[4].Wei-Long Zheng, and Bao-Liang Lu, Investigating Critical Frequency Bands and Channels for EEG-based Emotion Recognition with Deep Neural Networks, IEEE Transactions on Autonomous Mental Development (IEEE TAMD) 7(3): 162-175, 2015.

[5].Short RD, Fukunaga K. The optimal distance measure for nearest neighbor classification.IEEE Transactions on Information Theory 1981;27:622-7. 10.1109/TIT.1981.1056403 [CrossRef] [Google Scholar]

[6].SEED Dataset Homepage. Available online: http://bcmi.sjtu.edu.cn/~seed/seed.html

[7].Practical Approach to Electroencephalography-handbook

[8].F. Siraj, N. Yusoff and L. C. Kee, "Emotion classification using neural network," 2006 International Conference on Computing & Informatics, Kuala Lumpur, Malaysia, 2006, pp. 1-7, doi: 10.1109/ICOCI.2006.5276419.

[9].K.N.V.Satyanarayana,Dr.T.shankar,Dr.P.V.Ramaraju "A Survey on Assessment of Sentiments and Emotions Using EEG Signals"2020 International Journal of Advanced Science and Technology Vol. 29, No. 3s, (2020), pp. 717 - 728