

# A Novel Method to Perform Search Query using Brainwave Signals



Deekonda Sai Manish, Nabeel Shahid, Anand Raj, P. Mohamed Fathimal

**Abstract—** The main aim of the proposed paper is to search the information using brain waves instead of searching using the text query. Electroencephalography (EEG) is an electrophysiological checking strategy to record the electrical activity of the mind. EEG estimates voltage changes coming about because of ionic current inside the neurons of the mind. EEG alludes to the account of the mind's unconstrained electrical movement over a time frame, as recorded from different nodes put on the scalp. For training, When the user has read a question, the brain waves are recorded with the help of sensors like Neurosky device in the form of EEG signals values like Alpha, Beta, and Gamma and are stored in a dataset. The brainwaves of different persons are recorded for different questions and are stored. The features are reduced using PCA and the centroid of the values are calculated using the K-means clustering algorithm. For testing, when the user thinks about a question in the list, the brain waves are recorded and compared with the values available in the dataset. Using KNN Algorithm, the proposed system outputs the respective question which will be submitted to the search engine. K-Means clustering algorithm is used to calculate the cluster centroid. Once the centroid is calculated for each question, we plot each centroid in a 2-D plane. For a random question from the pool of existing questions, we use the KNN algorithm to find the nearest match. When the match is found, the question corresponding to it is submitted to the search engine.

**Index Terms-** Brain Waves, EEG(Electroencephalography), PCA, K-Means, KNN, Neurosky Device.

## I. INTRODUCTION

Brain Wave is among the few innovations the world is looking into nowadays. Brain waves as the name suggest are produced in our brain. Our feelings, chemical activities, thoughts are imparted through the neurons. Neurons, which are the Brain cells and building blocks of the focal sensory system, talk to one another as electrical heartbeats and signs. A huge number of signs speak with one another delivering electrical action in our mind. Brainwaves are liable to change contingent on the progressions of our contemplations and emotions and are basic because their examples enlighten a great deal concerning an individual. Any awkwardness in

**Manuscript published on 30 July 2019.**

\* Correspondence Author (s)

**Deekonda Sai Manish**, Department of Computer Science, SRM Institute of Science & Technology, Chennai, India.

**Nabeel Shahid**, Department of Computer Science, SRM Institute of Science & Technology, Chennai, India.

**Anand Raj** Department of Computer Science, SRM Institute of Science & Technology, Chennai, India.

**Dr.P.Mohamed Fathimal**, Department of Computer Science, SRM Institute of Science & Technology, Chennai, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](#) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

brainwaves can seriously influence our emotional or neuro-physical well-being. Any sort of insecurity in the brainwaves can cause conditions like over the top enthusiastic confusion, alarm assaults, unstable conduct, and headaches and among others. There are two states of unsteadiness in brainwaves. Over-excitement: -This prompts sleep anxiety, imprudent conduct, fits, uneasiness issues, outrage, and animosity. Under-excitement: -This prompts sleep deprivation, gloom, an absence of the ability to focus and unending torment. Brainwaves can be identified through sensors that are put on the scalp through a methodology known as an electroencephalogram (EEG).

## II. UNDERSTANDING BRAIN WAVES

The Brainwaves contain a range of frequencies primarily between 1Hz to 100Hz. An image of the range, which I had taken while learning about the Brain wave, is shared underneath.

Delta Waves movement in babies is substantially more than grown-ups. This can likewise be seen for all intents and purposes as babies invest much energy in resting. The study demonstrates that delta wave is in truth the most ruling mind wave in youngsters till 5 years of their introduction to the world. Delta wave decays with a drop rate of 25% from 11 to 14 years and continues declining entire life expectancy and at 75 years old, they are practically missing.

Theta waves are fundamentally the internal feeling of the body rather than an outer factor. It happens mostly when we are going to rest or when we simply wake up. In theta, we are in the dream and it is a similar spot where we hold our feelings of trepidation, history and bad dreams. It is additionally associated with remedial rest and momentary memory. The recurrence of theta waves lies between 3 to 8 Hz.

Alpha waves were found by Hans Berger, the acclaimed German nervous system specialist who is well known for his development of EEG. Berger likewise found delta waves alongside alpha. Alpha waves emerge from electrical movement of thalamic pacemaker cells. They start from the occipital projection of the mind. Alpha comes in the middle of the beta and theta and crosses over any barrier among cognizant and intuitive personality. It's related with serene, wonderful and practically gliding inclination and is co-related with loosened up attentiveness and innovative musings.

Alpha brainwaves are prevalent in the midst of attentively streaming considerations, and in some intelligent states.





Fig 1. Device Description (NeuroSky, Inc.).

Alpha waves help in maintaining mental coordination, peacefulness, mind/body blend and learning. The recurrence scope of alpha waves is between 8 to 12Hz. Alpha waves OVERSEE THE PSYCHOLOGICAL COORDINATION AND MIND-BODY COMBINATION.

Presently comes the wave related to the most astounding consideration level. At whatever point our mind is effectively occupied with exceptionally participatory mental exercises, it creates beta waves. These waves are of low adequacy and high frequencies. The recurrence of beta waves ranges from 15 to 40 cycles each second. The time span of the beta wave is typical for exceedingly drawn as the main priority.

Beta waves are the image of readiness, thinking, rationale and waking awareness of our own. While Beta Brain waves are essential for viable working for the duration of the day, they additionally can convert into pressure, uneasiness, and eagerness.

The beta waves are additionally subdivided into three groups:

Table I: Raw data collected from the user.

	Delta	Theta	Alpha1	Alpha2	Beta1	Beta2	Gamma1	Gamma2
6	1411200.0	31215.0	30568.0	12991.0	12464.0	5938	4656	2072
24	843190.0	186860.0	35421.0	8970.0	7512.0	5111	3339	4086
48	236580.0	32735.0	7681.0	7282.0	1850.0	11548	4698	3495
10	2680800.0	319430.0	17253.0	8877.0	27184.0	6233	7286	13950
44	948140.0	628360.0	71560.0	55035.0	75705.0	59780	36710	17951

- Low Beta (12-15Hz) - These brainwaves are described as quick, inert or thoughts.
- Beta (15-22Hz) - These brainwaves are characterized as high commitment.
- Hi-Beta (22-38Hz) - These brainwaves are exceedingly unpredictable and describe nervousness or energy.

The Gamma waves were prior disregarded in simple EEG where frequencies under 25Hz were utilized. Researcher and specialists around the idea of gamma waves as a commotion. An almost no have been thought about these waves however it is evident that these waves are utilized in rapid data handling and are related with the preparing of visual, sound-related improvement.

Gamma waves are additionally connected with our capacity to adjust and learn. They are headed without recognition for new things. Studies demonstrate that individuals who are simple-minded will, in general, have low gamma exercises.

### III. ABOUT THE DEVICE

Official Headsets from Neurosky - Neurosky is all up with an accumulation of EEG Headsets. These modules can give us Meditation/Attention signals, EEG signals control band signals and crude wave information too.

It imparts remotely and Neurosky gives free designer instruments moreover as shown in the above Fig.1. It is a decent choice according to display details. Learn more about the device from website (Refer [8]).

### IV. RELATED WORK

Patrick Suppers calculates electrical and magnetic brain waves of seven subjects under different environmental conditions using sixteen different EEG sensors [2]. The results were obtained around 97% and 91% correct encourage research. The performance is at a level comparable to that of speech recognition at its early days. Even though 16 EEG sensors were used the accuracy ranged around 50%. It is clear from the research that the global electric and magnetic brain waves carry significant information about words being processed. Dipti Pawar [6] presented a probabilistic model which increases the efficiency of existing KNN system by considering the prediction probability of each EEG channel. Thus, the efficiency of the system enhances with the increase in data sets. Ahmad Azhari have used k-means clustering for brain waves feature classification. The data is acquired using a Neurosky device [5] and with the electrode placed in the frontal lobe position. The data of 6 people are collected with each person having his data collected twice in the frequency range of 128 Hz per second. There are 9 different cognitive tasks that are asked to be carried out each time the data is being collected. Also presented PCA implementation [7] for brain wave feature reduction.

### V. PROPOSED METHOD AND CONTRIBUTION

At first, the brain waves for each question is collected via the help of the Neurosky device. The data is in the form of EEG signals and is stored in the form of a Nx8 matrix as a CSV file. Once this is done, the feature reduction takes place using the PCA algorithm. The Nx8 matrix is reduced to an Nx2 matrix. Then we apply the K-means clustering algorithm to find the cluster centroid for each question and plot them in a 2-D space. For the question to be searched we repeat the same steps as mentioned above. Once we have the cluster centroid for each Query, we use the KNN algorithm to find the nearest cluster centroid and display it as the question to be searched. The Fig.2 represents the overall System Description.

- Data Acquisition – The data is recorded with the help of a Neurosky device in the form of EEG signals. The values for Alpha, Beta, Gamma are recorded and stored in a CSV file.
- PCA - The original number of features is 8 which results in an Nx8 matrix. To reduce the size of the matrix from Nx8 to Nx2, we use PCA. It helps reduce the number of features and makes it easier to work on the data.



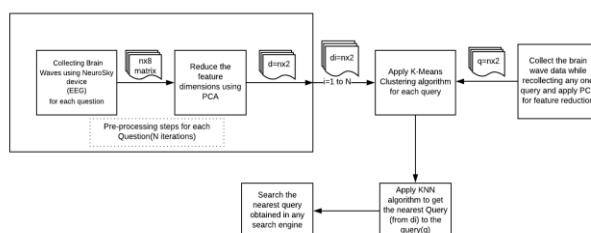


Fig 2. Architecture of the System.

- K-Means – It is used to find the cluster centroid for each question. The cluster centroid is plotted in a 2-D plane.
- KNN – It is used to find the nearest cluster centroid to the cluster centroid representing the question the subject is thinking about.
- Testing – In this phase, the subject is asked to think of a random question from the pool of existing questions. When the subject is doing so, the first three steps are repeated, and the cluster centroid is plotted. Once the centroid has been plotted, the K-Nearest Neighbor algorithm is used to find the closest cluster centroid to that question. The question is then searched in the search engine.

Major contributions of this project are as follows:

- i. It helps users to search using brain waves.
- ii. It can be a boon for disabled people, helping them to communicate.
- iii. It can be used to record the brain waves of various persons, which can be compared to gain some insight about their neurological assessments.

## VI. PROCEDURE

After the successful collection of the EEG Data from the user, it is feed to the PCA module where the dimensions of the data are reduced. The fundamental thought of Principal Component Analysis(PCA) is to decrease the dimensionality of an informational collection comprising of numerous factors related with one another, either vigorously or delicately, while holding the variety present in the dataset, up to the most extreme degree. The equivalent is finished by changing the factors to another arrangement of factors, which are known as the principal components (or basically, the PCs) and are symmetrical, requested with the end goal that the maintenance of variety present in the first factors diminishes as we move down in the request. Along these lines, the first Principal Component holds the most extreme variety that was available in the first segments. The central segments are the eigenvectors of a covariance network, and consequently, they are symmetrical.

The dataset on which PCA is applied must be scaled. The outcomes are sensitive to the relative scaling. PCA obtains a decrease in dimensions by making principal components. Every principal component is comprised of all the original factors present before dimensionality reduction.

PCA is an unsupervised technique, implying that no data gatherings are utilized in the measurement decrease. This implies PCA demonstrates a visual portrayal of the overwhelming examples in an informational index. The Fig.3 represents the data after PCA application. Initially,

there were 8 dimensions for the data which were reduced to 2, these scatter plots represent the two dimensions of the reduced data as x and y co-ordinates. Then we apply the K-Means Clustering algorithm to get the cluster center.

K-Means is a clustering algorithm that attempts to segment a lot of inquiries into K sets (bunches) with the end goal that the questions in each bunch will, in general, be close to one another. Each bunch comprehensively speaks to a question.

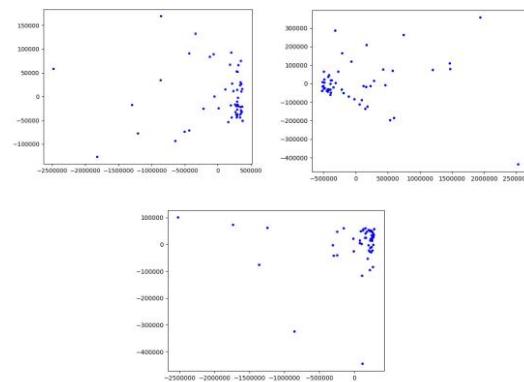


Fig 3. These figures represents the scatter plot of the reduced dimensions obtained from after applying PCA.

It is unsupervised in light of the fact that the focuses have no outside characterization. This process is repeated for the testing data, the center calculated for testing data is known as an assumed center which is used along with trained data center of each query to calculate the nearest neighbor of the assumed center from the pool of testing centers or the query centers.

K-Nearest neighbors are classification (or relapse) algorithm that is used to find the nearest point to a set of given points. It uses a distance measure such as Euclidean distance or Manhattan distance for the same.

Once the cluster centroid for both test query and train query centers are all plotted in the 2-D space, KNN is used to find the nearest query to the assumed center or test query center we want to search, further, this nearest query from the label file is feed to the search engine.

## VII. RESULTS AND ANALYSIS

While collecting the EEG data from the user sometimes it may happen that the device gets disconnected from the forehead, then the values may be unconditional or wrongly saved. When the device gets disconnected the EEG values are very low (<100), this can be used to make the device stop automatically, the system Will raise exception and will not record the data i.e.,



**Table II: Model Analysis.**

	<b>precision</b>	<b>recall</b>	<b>f1-score</b>	<b>support</b>
1	1.00	1.00	1.00	1
10	1.00	1.00	1.00	1
2	0.33	0.50	0.40	2
3	1.00	0.67	0.80	3
4	0.50	0.50	0.50	2
5	0.50	0.50	0.50	2
6	1.00	1.00	1.00	1
7	1.00	0.50	0.67	2
8	1.00	0.50	0.67	2
9	0.00	0.00	0.00	0
<b>micro avg</b>	<b>0.62</b>	<b>0.62</b>	<b>0.62</b>	<b>16</b>
<b>macro avg</b>	<b>0.73</b>	<b>0.62</b>	<b>0.65</b>	<b>16</b>
<b>weighted avg</b>	<b>0.79</b>	<b>0.62</b>	<b>0.68</b>	<b>16</b>

warn the user to adjust the device properly and restart the process. Feedback from the user is used to get the accuracy of the model in the testing phase. The true query number is given by the user and user is asked to give correct/incorrect feed about the query searched. The Results of the different persons are recorded as shown in Table II.

Accuracy is defined as the quality of being correct and precise. Accuracy refers to how to correct the readings of a system are. The accuracy of our model was found out to be 72% (approx.) The above report suggests the model efficiency and the above Table II depicts precision, recall, F1 score and support as recorded for different questions. The average and the weighted average are also calculated. Precision is the number of relevant instances among the total instances present in the system. Recall refers to the number of relevant instances out of the total relevant instances. F1 score is used to calculate the mean of precision and recall. Support refers to the number of true instances that lie in a class. The macro average computes the metric independently for each query, and the average is taken, whereas micro average aggregates outcomes of all the queries and then compute the average metric.

## VIII. CONCLUSION

The Brain waves can be used as an efficient tool to enhance the current searching mechanisms. This searching method is more secure compared to the previous methods. However, the subject must always maintain focus which may take some time getting used to it. Also, the device limits the usage to a few questions. With a better device, we will be able to implement and search for more queries.

## ACKNOWLEDGMENT

We would like to take an opportunity to thank our guide Dr.P.Mohamed Fathimal for his support and motivation due to which this project is a success. We would extend our hearty thanks to our coordinators Dr. J. Arunnehr and Mrs. G. Chamundeeswari for extending their support in writing this paper. We would also like to thank our colleagues for their support and critical appreciation throughout the project.

## REFERENCES

1. Separation Of Alpha, Beta, Gamma & Theta Activities In EEG To Measure The Depth Of Sleep And Mental Status Shah Aqueel Ahmed, Syed Abdul Sattar, D. Elizabeth Rani-2013
2. Brain Wave Recognition of Words Article in Proceedings of the National Academy of Sciences · January 1998 by PATRICK SUPPES, ZHONGLIN LU AND BING HAN
3. BRAIN WAVE SENSOR SYSTEM FOR ACCIDENT PREVENTION IN VEHICLES S. Pradeep Kumar and A. Wisemin Lins, VELS University, Pallava-8, April 2017
4. STUDY AND APPLICATION OF BRAIN WAVES (ALPHA, BETA) FOR USER AMBIENT ENVIRONMENT CONTROL G. Ambica et al, International Journal of Computer Science and Mobile Computing, Vol.4 Issue.10, October- 2015.
5. Ahmad Azhari, Leonel Hernandez,(2016)-Brainwaves feature classification by applying K-Means clustering using single-sensor EEG, International Journal of Advances in Intelligent Informatics, ISSN: 2442-6571.
6. A Survey on Thoughts to Text Using Brainwave Pattern Dipti Pawar, Kapil Kamale, Ishaan Singh, Vishwajeet Kharote, Prajwal Goswami Assistant Professor, Dept. of Computer, Pune, India May-2017.
7. Ahmad Azhari, Murein Miksa Mardhia,(2017)-Principal component analysis implementation for brainwave signal reduction based on cognitive activity, International Journal of Advances in Intelligent Informatics,ISSN: 2442-6571.
8. Mind Wave Neurosky device description: <https://store.neurosky.com/pages/mindwave>

## AUTHORS PROFILE



**Deekonda Sai Manish** completed his B.Tech. in Computer Science and Engineering from SRM institute of science and technology, Chennai, India. His research interests text recognition using machine learning, IOT and NLP.



**Nabeel Shahid** completed his B.Tech. in Computer Science and Engineering from SRM institute of science and technology, Chennai, India. He has a keen interest in AI and is intrigued by the developments in the same field. His research interests include Machine Learning, Deep Learning and IOT.



**Anand Raj** completed his B.Tech. in Computer Science and Engineering from SRM institute of science and technology, Chennai, India. His areas of interest include Android Development, IOT and Advanced Programming.



**Dr. P. Mohamed Fathimal** received her BE, ME and Ph.D. in Computer Science and Engineering from Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu. She has more than 16 years of Teaching Experience. Currently, she is working as Assistant Professor in the Department of Computer Science and Engineering at SRM Institute of Science and Technology. She has published more than 12 papers in various reputed journals. Her research interests include Digital Image Processing, Data Mining, Information Security and Machine Learning.