Index No: CT-01-28

BINARY DECISIONS AND EMOTIONAL EEG BASED VIRTUAL ASSISTANT FOR SPEECH IMPAIRMENTS

Prashadi H. Halpe
Dept.of Information technology
Sri Lanka Institute of Information Technology(SLIIT) Colombo, Sri Lanka
it19017716@my.sliit.lk

Probodha.S. Aththanayake
Dept.of Information technology
Sri Lanka Institute of Information Technology(SLIIT) Colombo, Sri Lanka
it19176666@my.sliit.lk

Prasanna S. Haddela
Dept.of Information technology
Sri Lanka Institute of Information Technology(SLIIT) Colombo, Sri Lanka
prasanna.s@sliit.lk

Lakmini Abeywardhana
Dept.of Information technology
Sri Lanka Institute of Information Technology(SLIIT) Colombo, Sri Lanka
lakmini.d@sliit.lk

Jinendra Bogahawatte
Dept.of Information technology
Sri Lanka Institute of Information Technology(SLIIT) Colombo, Sri Lanka
it19237978@my.sliit.lk

Sanduni Disanayaka
Dept.of Information technology
Sri Lanka Institute of Information Technology(SLIIT) Colombo, Sri Lanka
it19098760@my.sliit.lk

***Abstract*— The potential for brain computer intervention to provide effective and exciting technology for persons with impairments is growing because of the constant advancements in information technology. To help those who have difficulty communicating in a high school environment, our research will develop a wearable assistive device called SPEAK. It is a compact, portable equipment that is lightweight and simple to use. It comes with the Muse 2 headband and a deep learning-based assistance program that operates on a mobile device.**

Sometimes people are unable to communicate vocally for a variety of reasons, yet their minds (brains) nonetheless speak normally. This approach would be really helpful and perfectly acceptable to help such mute individuals who are unable to communicate.

This research puts forth the effort to detect and collect the frequency ranges of two significant analytical components that have been particularly chosen, such as binary decision yes/no analysis and emotional analysis. The examined frequencies are compared to the freshly emerging brain signal excitements in order to recognize and comprehend the phrases in accordance with the relevant brain signal. The virtual assistant produces voice for the muted person whenever a word is recognized from a brain signal and is transformed into speech sound.

Keywords - Yes / No Binary decision, Emotional Analysis, EEG, BCI

I. INTRODUCTION

The objective of this project is to create SPEAK, a wearable speech assistance. Over one hundred thousand Sri Lankans have speech difficulties [1]. This would enable communication. Disabilities in speech vary. Included under "Stuttering" are "Apraxia" and "speech sound anomalies." This gadget exclusively aids speech-impaired individuals who are silent [2]. Using machine learning and deep learning technologies with the Muse 2 headband, a lightweight, user-friendly, portable support application will be developed. First, the frequency ranges of analytical components are read. Included among them are emotion-based analysis, satisfaction rate analysis, directional analysis, and binary analysis. The assessed frequencies are compared to recent brain signal excitations. People who are silent may not be able to express themselves vocally, but their thoughts are clear. The brain understands precisely what a person wants to say aloud since it is aware of all their actions, feelings, and senses. Using Electroencephalogram (EEG) signals, the brain's activity may be read by a computer fitted with Brain-Computer Interface (BCI) technology. Based on brain signals, it detects words. Brain signals are converted into vocal sounds via the mobile device's virtual assistant. This study for binary decision analysis uses "Yes/No" as replies to between-two-choice questions. This makes it easier to convey basic two-choice selections to others. Based on three emotions called "Happy", "Sad", and "Fear," instructors may recognize students' mental states and create an optimal learning environment.

II. BACKGROUND

The effective modality of electroencephalography (EEG) makes it possible to collect brain signals that are related to different states from the surface of the scalp. Based on signal frequencies that span from 0.1 Hz to more than 100 Hz, these signals are often divided into five categories: delta, theta, alpha, beta, and gamma [14]. The electroencephalogram (EEG) is a typical neuroimaging instrument used by clinicians to examine neural activity in the human brain. Data gathered from EEG readings indicate an individual's information processing process [15]. One of the latest research hotspots in the disciplines of biomedicine and signal processing is the brain-computer interface (BCI). A human-computer interaction technique based on brain impulses is known as brain-computer interface technology. It offers a route of communication for non-neuromuscular control. Using a brain-computer interface, the human brain can communicate with the outside world without the need of the muscles or the peripheral nervous system. Memory interfaces are often made using physical vision, or multiple pictures of a motor movement. Using this method, signals generated while imagining the movement of different body parts are collected. For brain-computer interactions, this is accurate (BCI) [16].

III. LITERATURE REVIEW

Majority of the previous research has been used visual based [3] or hand movement [4] [5] [6]based brain – controlled solutions as the techniques. Within the visual based techniques, graphical representations of letters / numbers, motor imaginary, flashing or colored stimuli are used to generate ocular movements to produce brain signal inducements [3]. The hand movements-based research requires remembering that the left-hand imagined movement corresponds to “yes” and the right-hand imagined movement corresponds to “no” [7]. Some other BCI research experiments require many extensive subject biofeedback training sessions for a considerably long time for the subject to gain some degree of voluntary influence over binary decision (yes /no choices) EEG features such as slow cortical potentials [8]. Also, previous research studies have proven that the most direct approach would be simply the subject to imagine “yes” or “no” as the response to a simple question and that would not require any extensive biofeedback training [7]. There are some studies where analyzed spatial patterns of Alpha signals and Theta signals are essential for Yes / No binary decision making [9]. The human brain’s right frontal region and right Centro-parietal region are responsible for Theta and Alpha signals accordingly.

The potential and significance of emotion recognition are vast. The environment is only one of many variables that may have an impact on the complicated psychophysiological process that is emotion. Face, voice, text, heart rate, behavior and physiological signs may all be used to identify that [10]. EEG based approaches are promoted to do emotional analysis in previous research. Video clips, soundtracks, images kind of physical objects are used in previous research to stimulate happy, sad, or fierce related EEG brain signal inducement [11]. Many studies have already used SVM, Gaussian classifier, Bayesian, and Random rainforest as emotion classification algorithms to build the models. As most of the EEG-based research, alpha, gamma, and beta waves seem to be the most discriminative, whereas the frontal and parietal lobes tend to retain the most knowledge about emotional states. This demonstrates that emotion categorization studies, including both lower frequency bands (delta and theta) and higher frequency bands (alpha, beta, and gamma), are of similar importance and are the favored method for acquiring brainwave features among researchers [12].

In the mobile application system, not only the interface reaction but also a clear speech sound output of “yes / no” decisions or “happy/sad/fear” will be produced according to the speech impaired individual’s decision and emotions as a response. Also, simply the user only needs to focus directly on their “yes/no” choice or “happy/sad / fear” emotion, and the device directly detects EEG signals of their response. Therefore, user does not need to remember specific responding methods. Mobile application is not using any graphical representations to stimulate EEG inducement. It detects the direct focus of the user's “yes / no / happy / sad/ fear” thought. Not only that but also, as the mobile application does not require ocular movements or ocular stimuli methods to output “yes/no” decisions or “happy/sad /fear” emotions, that would be very convenient for the users than the BCI spellers. Using BCI spellers like P300, GIBS, Chroma Speller, SSVEP and accessories like thought helmets would be very time-consuming as generating the response would be very slow. But the mobile application is a real-time assistive device. Therefore, the response would be faster than those BCI spellers and accessories.

IV. System Overview

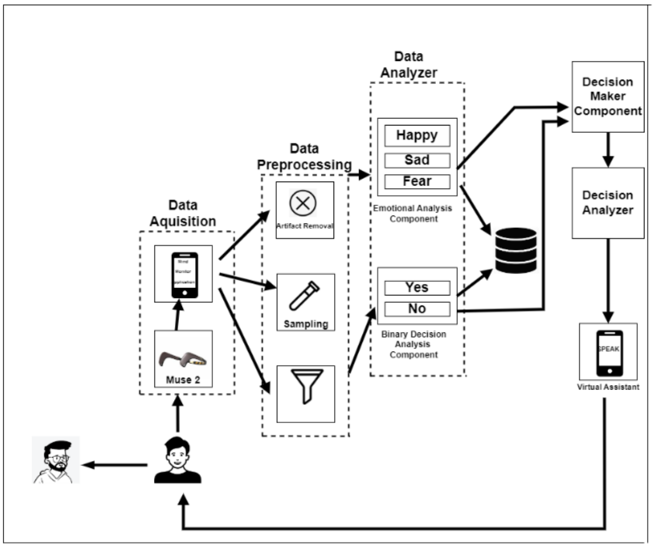


Figure 1: depicts the system overview for the “yes” / “no” binary decision analysis research study.

V. Methodology

A. *Data Acquisition*

Collecting data for the yes / no binary decision analysis is initially decided to use the MUSE2 headband according to a defined experimental practice framework. The framework is structured to use static images of a tick mark to present “yes” and a cross mark to present “no.” This framework is only designed to train the sample for positive and negative responses. The strategy of the static images – tick mark is saying yes for most of the positive responses encapsulating yes, ok, agree, and good. Cross mark is saying no for most of the negative responses encapsulating no, not ok, disagree, and not good. The Muse2 device recorded a considerably large number of null values and noise data during each experiment. As a solution, an existing dataset from a previous research study is being selected [13]. The selected dataset is being recorded as both EEG signals and EOG signals since the existing research is conducted to analyze the performance of gaze independent BCI based on attention to respond “yes” or “no” [13]. For this research component study, the EOG signal data is being eliminated in the feature extraction.

For the emotion analysis study, electrical activity in the brain was measured using a MUSE 2 headband, which has four channels (TP9, AF7, AF8 and TP10) that touch the forehead and gather and monitor electrical activity in the brain. Several times, data on the emotional states of six students were collected at various time intervals for the data acquisition. Three of them are speech-impaired students at age 16-19. Other students are regular. All the participants were in good health and not under the influence of any drugs. Before collecting the EEG brain signal data, a practice framework is used to train the student sample quickly. The framework is used three specific static images a happy face, a sad face, and a frightened face. The positive, happy, or relaxed mind states are focused as the strategy used for happy emotion simulation. Likewise, negative, sad, or tired mindful emotions are focused on the sad face and extreme negative, tensed, or disturbed mindful emotions are focused on the frightened face. Using these static images, the student sample is trained to focus on the emotions they feel when the questioner is being asked at EEG data recording. While asking the participants questions, EEG data are collected on focus and mood of emotion changes to capture the emotions. For everyone, an EEG recording is made as soon as a question is asked, based on how each person's emotions are shown.

B. *Data Preprocessing*

The selected dataset from the existing research for "yes / no" binary decision analysis has so much data that this study needs to avoid as noise data. The collected EOG data cause the noise, captured ocular movements and muscle movements, electronic amplifiers, battery power interference, etc. These irrelevant columns are eliminated at the feature extraction. The dataset did not consist of null values to eliminate. However, there were much of outliers checked by the Boxplot test to eliminate from the dataset. The dataset is split into a training set and a testing set to a ratio of 67:33.

The collected dataset for emotional analysis is prepared before being fed into the classification model by removing outliers identified from the Boxplot test to increase the model's accuracy. Aberrations of EEG readings decreased by eliminating the noise data and artifacts. The null values collected due to the defective operation of Muse2 and irrelevant columns like battery power, timestamp, headband on, etc., are removed from the dataset in the feature extraction. The class of emotional analysis, "happy/sad / fear," are used the one-hot encoding technique to handle the categorical variables to extract the categorical features.

C. *Feature Extraction*

Assumption-made-only EEG-based features are selected to consider in the "yes" / "no" binary decision-making analysis. The features are analyzed with Pearson correlation and Recursive feature selection, which provides higher prediction accuracy and chooses the highly correlated features with labeled data. Hence EOG data and other irrelevant data are illuminated considering extra columns.

D. *Model Building*

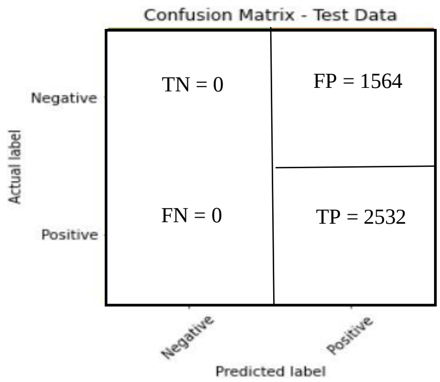
Following a feature extraction process, a considerable number of classification models and regression models are being tested for this "yes" / "no" binary decision-making study. Decision tree index classifier, XGB classifier, gradient boost classifier, random forest classifier, K-nearest neighbor classifier and logistic regression regressor are used to build and test the classification models for "yes / no" binary decision analysis.

Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) and K-Nearest Neighbor Classifiers are used to build and test machine learning models and deep learning models for emotional analysis.

VI. Result Analysis and Discussion

TABLE I: "YES / NO" BINARY DECISION ANALYSIS RESULTS

Component	Models	Accuracy	Loss
"Yes" / "No" binary decision analysis	Decision Tree Index	0.61	0.39
	XGB	0.57	0.43
	Random Forest	0.61	0.39
	KNN	0.61	0.39
	Gradient Boosting	0.60	0.40
	Logistic Regression	0.62	0.38




```
Accuracy (overall correct predictions): 0.62
Auc: 0.5
Recall (all 1s predicted right): 1.0
Precision (confidence when predicting a 1): 0.62
F1 score : 0.76
Detail:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00      1564
           1       0.62      1.00      0.76      2532

    accuracy                           0.62      4096
   macro avg       0.31      0.50      0.38      4096
weighted avg       0.38      0.62      0.47      4096
```

Figure 2: Yes / No binary decision analysis statistics – Logistic Regression

As the dataset behaves as having linearly separable data, which indicates a linear relationship between data points, the data is tested for the simple linear regression model. The simple linear regression model is not suitable for classification; hence SLR deals with continuous values and as well as classification mandates discrete values. The logistic regression model is tested for the collected data since it's a supervised machine learning algorithm and outcomes a binary output, hence can be used for binary classification directly. Several nonlinear ensemble classification models, such as Decision Tree Index, and Random Forest, are also being tested (*Table 1*). Hence those models are nonlinear, and the dataset is linear; the accuracy for each model is lower. The highest accuracy claimed the logistic regression model, which provided 62% accuracy.

TABLE II: "HAPPY / SAD / FEAR" EMOTIONAL ANALYSIS RESULTS

Component	Models	Accuracy	Loss
"Happy / Sad / Fear" emotional analysis	KNN	0.93	0.07
	MLP	0.81	0.19

	precision	recall	f1-score	support
0	0.95	1.00	0.97	491
1	0.93	0.85	0.89	514
2	0.90	0.93	0.91	503
accuracy			0.93	1508
macro avg	0.93	0.93	0.93	1508
weighted avg	0.93	0.93	0.92	1508

Figure 3: Happy/ Sad / Fear Emotional Analysis – KNN

Multi-Layer Perceptron (MLP) and K-Nearest Neighbor (KNN) models are tested in the emotional analysis as a comparison between deep learning and machine learning classifications. To classify the emotion out of the considered "happy/sad/fear," the best classification model is the KNN model which scores the highest accuracy compared to the MLP model. At the same time, the KNN model scores 92.04% highest accuracy, and MLP scores only 81.56% of accuracy.

Real world user testing for the experiment carried out has a success rate of near 66%. Four students out of six student participants confirmed that the application has output the accurate result of their intensions. As for the data acquisition, the Muse2 device is used for the real world user test with confident that there is no harm using the device on human, as the device is a publicly existing wearable equipment in Europe which is used for the meditation purposes. The user test success rate of near 66% is a good evaluation for an experiment of an initial stage, yet expecting to expand the study areas and upgrade the tools and technologies used as working on the improvements.

VII. Conclusion

This research provides a beneficial alternative communication method for speech impaired people who are unfamiliar with any sign language, have a sudden muteness caused by an accident, paralysis, or as a result of an illness. The taken EEG readings are being used to train the models to differentiate the selected responses of yes or no decisions and emotion types. An experiment with a sample of six muted students in the age range of 16 – 19 evaluates the application's main functions: the accuracy of 61% and 92.04% for yes / no binary decision and emotion, respectively. Due to the other existing noise caused by muscle movements, accuracy degradation is happening. The understanding is to increase the sample size to increase the accuracy. Thereby, the overall performance of the application could be improved.

It is found that there is a high demand for such types of helpful equipment especially which can improve and indeed support needed person's life. The main problem a suddenly muted person faces is not being able to communicate with others like before. The mobile application is introduced to fill this gap for such a person personally. The most crucial part of the application is to detect and categorize the EEG signals of the user. For that, the MUSE2 headband is used, and the mobile application-based virtual assistant is introduced along with the MUSE2 headband. The strategy is for the user to wear the MUSE2 head band and focus on the response they need to convey for a question that has been asked. The mobile application will record the EEG and analyze those readings and generate the output by running the specific models. The virtual assistant of the mobile application will output the analyzed response as a speech output, becoming the voice for the muted person.

In the future, this project will be improved by adding features to identify response directions, satisfaction, etc. Also, the virtual assistant will be improved to filter out the users' whole set of thoughts or words to produce a proper full sentence response, like the binary decision and emotional EEG based virtual assistant speaks the user's thoughts out loud.

REFERENCES

[1] D. o. C. a. S. I. S. Lanka, "Disability in Sri Lanka," 2012. [Online]. Available: https://unstats.un.org/unsd/demographic-social/sconcerns/disability/.

[2] W. Fargo, "Speech Disabilities," Deque University, [Online]. Available: https://dequeuniversity.com/class/archive/basic-concepts1/types-of-disabilities/speech.

[3] M. B. P. S. F. G. A. S. a. I. V. A. Rezeika, "Brain–Computer Interface Spellers: A Review," Brain Sci. 2018, 8(4), 57;

[4] W. D. R. S. J. C. E. A. &. S. M. J. Penny, "EEG-based communication: A pattern recognition approach," IEEE Transactions on Rehabilitation Engineering, vol. 8, no. 2, p. 2, 2010.

[5] G. N. C. G. C. H. W. R. H. S. A. O. B. &. P. M. Pfurtscheller, "Current trends in Graz Brain-Computer Interface (BCI) research," IEEE Transactions on Rehabilitation Engineering, vol. 8, no. 2, pp. 216-219, 2000.

[6] F. C. F. L. L. M. J. M. J. V. M. H. J. B. L. &. M. M. G. Babiloni, "Linear classification of low-resolution EEG patterns produced by imagined hand movements," IEEE Transactions on Rehabilitation Engineering, vol. 8, no. 2, pp. 186-188, 2000.

[7] J. N. K. M. J. K. C. W. A. a. M. H. T. D. A. Peterson, "Featuer Selection and Blind Source

Separation in an EEG-Based Brain-Computer Interface," 2005.

[8] A. K. N. G. T. H. J. P. J. K. I. I. B. K. N. N. H. F. N Birbaumer, "The thought translation device (TTD) for completely paralyzed patients," IEEE Trans Rehabil Eng, vol. 8, no. 2, pp. 190-193, 2000.

[9] K. H. K. Jeong Woo Choi, "Covert Intention to Answer "Yes" or "No" Can Be Decoded from Single-Trial Electroencephalograms (EEGs)," Computational Intelligence and Neuroscience, vol. 2019, p. 11, 2019.

[10] X. H. H. W. X. Y. Jinru Yang, "EEG-based emotion classification based on Bidirectional Long Short-Term Memory Network," Procedia Computer Science, vol. 174, pp. 491-504, 2020.

[11] A. B. S. K. Y. V. S. P. H. K. B. S. K. Y. V. S. P. S. T. V. B. Anala Hari Krishna, "Emotion classification using EEG signals based on tunable-Q wavelet transform," IET Science, Measurement and Technology, 2019.

[12] J. M. a. J. T. Nazmi Sofian Suhaimi, "EEG-Based Emotion Recognition: A State-of-the-Art Review of Current Trends and Opportunities," Comput Intell Neurosci, 2020.

[13] I. F. T. C. C. M. S.-R. H.-J. H. H. H, S. D. Christoph Reichert, "Impact of Stimulus Features on the Performance of a Gaze-Independent Brain-Computer Interface Based on Covert Spatial Attention Shifts," 2020.

[14] J. Satheesh Kumar, P. Bhuvaneswari, "Analysis of Electroencephalography (EEG) Signals and Its Categorization–A Study, "Procedia *Engineering*, 2020.

[15] [15] P. Jahankhani, V. Kodogiannis and K. Revett, "EEG Signal Classification Using Wavelet Feature Extraction and Neural Networks," *IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA'06)*, 2006

[16] Rodriguez, Paula & Mejia, Jose & Mederos, Boris & Moreno, Nayeli & Mendoza, Victor. (2018). Acquisition, analysis, and classification of EEG signals for control design.