A Medical AI Agent as a Tool for Neuropsychiatric Diagnoses

Kotesh Rao
Chief Technology Officer
ADAK Digital Inc.
Houston, TX, USA
kotesh.rao@ADAKdigital.com

Ramarao Inguva
Vice President
East West Inc
Huntsville, AL, USA
rama.inguva@ewehsv.com

Anita Rao
Department of Computer Science
Georgia Institute of Technology
Atlanta, GA, USA
arao338@gatech.edu

***Abstract*— A Medical AI agent perceives its environment with sensors and aids diagnoses, communicating results to the physician and providing feedback to the patient. The agent augments existing clinician interaction by learning the evolving status of the patient over time. With limited contact between patients and physicians due to the onset of Covid-19, there is an increased need for telemedicine and the use of medical AI agents.**

A neuropsychiatric diagnostic system utilizing a medical AI agent is proposed to capture physiological perturbations and patient responses in order to diagnose and track patients with psychiatric disorders like PTSD. Sensors are used to capture vital physiological parameters such as skin conductance, pulse rate, and blood oxygenation. These could be augmented with cameras and microphones that are used to record patient responses such as speech perturbations, oral responses, facial expressions, eye and pupil changes, and body movements to medical questions and stimuli.

Keywords—medical AI agent, neuropsychiatric disorders, wearable sensors, telemedicine, PTSD

I. Introduction

The use of robotics and AI in providing medical care has been proposed to extend the reach of medical professionals in their diagnostic capabilities [1]. One such technology is the use of a Medical AI agent which would have the ability to process sensor data from patients, flag key physiological and self-reported features, and communicate findings to the physician followed by providing medical feedback to the patient. A primary use case for this type of AI agent is augmenting physician interaction and diagnosis that is becoming increasingly virtual through telemedicine with the onset of Covid-19 [2]. Improving accuracy, precise diagnosis, remote treatment, augmenting human abilities, and supporting mental health are just some of the benefits of this technology approach.

Using telehealth, physicians and other healthcare professionals strive to follow evidence-based, trauma-focused therapies. The Veterans Administration (VA) healthcare system in the United States has recognized the need for this and has been leading the efforts to implement telemedicine solutions for managing patients with post-traumatic stress disorder (PTSD) and depression. Suicide rates for Veterans are higher than that of the general population, and suicide prevention especially for those living in rural settings remains a high priority for the VA nationally [3]. The VA's Quality Enhancement Research Initiative is supporting the Telemedicine Outreach for PTSD (TOP) program to deliver therapy and other medical care to rural patients through phone and interactive video contact.

The infrastructure in place for providing psychiatric services using telehealth techniques such as Video to Home (VTH) has been in use for several years now and has been found to be effective [4]. There is an opportunity to augment this infrastructure with the use of additional sensors, artificial intelligence, and machine learning technologies. We present a framework and structure to implement this approach.

II. Neuropsychiatric Disorder Indicators

PTSD is a heterogeneous psychiatric disorder characterized by re-experiencing, avoidance/numbing, and hyper-arousal symptoms [5]. The VA system has a large target patient population consisting of those who have experienced trauma associated with exposure to violence. Increased psychophysiological arousal driven by the activation of the autonomic nervous system is a hallmark observation in PTSD. Outputs of autonomic activation, including heart rate (HR), blood pressure (BP), skin conductance (SC), and respiration rate (RR) are all heightened in individuals with PTSD [6]. Using physiological data, such as HR and SC, offers the potential for quantitative objective assessment of physiological reactivity related to PTSD symptoms. A 2017 study examining Veterans with PTSD employs a method based on script-driven imagery, which uses trauma-related stimuli to evoke physiological responses. The participants described an experienced traumatic event in detail, which was then transcribed and played back to the individual while physiological responses were recorded [7]. Those with a diagnosis of PTSD exhibited stronger HR and SC responses to scripts than non-PTSD trauma survivors.

While this data suggests that PTSD can be characterized by physiological reactivity, the script-driven imagery approach can be burdensome for clinicians. This approach requires elaborate equipment, dedicated space, specialized training, and substantial financial investment. Efforts are currently underway to standardize these methods, such that a less tedious data collection process is required and overall complexity is reduced. A 2017 study using skin-conductance to assess PTSD employs a trauma interview-based method instead of the script-driven imagery approach, which is more conducive to being administered remotely [6]. During a trauma interview session, SC, HR and other activity parameters are recorded continuously using mobile applications on smartphones and tablets.

The potential for medical AI agents to utilize edge-based patient monitoring in neuropsychiatric diagnostics is worthy of further investigation. Additionally, using machine learning techniques to analyze physiological parameters of remote patients is a growing opportunity as well. The following section outlines the methodology for data analysis that would be used by our medical AI agent.

III. Proposed Data Analytic Methodology

For characterization of psychosomatic disorders like PTSD it has been suggested to make the following measurements:

1. Blood Pressure
2. Pulse rate
3. Breathing rate
4. Perspiration and sweat texture on face
5. Skin Conductance
6. Eye movements
7. Facial features
8. Posture changes

To measure these parameters many of the hardware items needed are available as Commercial Off The Shelf (COTS) devices whose data can be consolidated in the mobile device as the data logger. In addition to digital data, it would be useful for completeness to capture verbal responses to orally administered questionnaire in text form using standard content management tools. This has not been addressed in much detail in this paper. Proper security and encryption measures may be needed for protection of integrity of data and privacy rights.

The net outcome of the testing on a human subject by measuring the parameters is a multidimensional vector which serves as an indicator of the stress\mental state of the individual. The data analytics applied to this multidimensional vector will lead to an identification of the PTSD tendency and its severity. This multi-dimensional vector data will be measured during every electronic visit of the patient\human subject. Collected over several visits for subjects with diverse racial, gender and other backgrounds will result in a large database to assist clinicians to identify early diagnostic indicators of suicidal tendencies.

The following are the steps involved in collecting and analyzing the data

A. *Feature Extraction*

This consists of feature definition, feature extraction, feature selection, feature space reduction and finally use the chosen features to perform cluster identification and classification. The classification data would then be used extract the emotional state of the subject.

A set of features can be extracted from the data. For those pieces of data which involve numerical data acquisition over periods of time, one can collect means, standard deviations, root mean square or other similar metrics like mean of first differences.

For those acquisitions which involve time series of data such as blood pressure, pulse rate, skin conductance, facial sweat levels and speech data, features could consist of deconvolution of the standard time series and estimate power spectrum and identification of characteristic frequencies at which large amplitudes occur [8][9]. The estimated frequencies and their amplitudes (absolute values) as features. For those data acquisitions that involve responses to stimuli, they could be classified on a standardized scale of 1-10.

Feature extraction involves processing each independent data set and extracting the features with the entire set of features collectively denoted by the feature vector **X.** If we denote the emotional state of the subject by the vector **Y,** there is a highly nonlinear mapping **f** from the feature vector X.

$$\mathbf{Y = f(X)}$$

B. *Clustering*

There are many means of clustering that can be used. One method called K-means Clustering is used to convert the entire feature set data into clusters. In this one starts with a set of distinct K features. Treating these as centers one can map all the other features that occur around each of these feature centers. As the features get divided into clusters or emerging clusters, we redefine the centers of these clusters as the mapping indicates. At the end of the process we would have divided the feature data into clusters. It is quite possible that all the data will only form into one cluster which will point to the fidelity of the data acquisition and the data acquisition and planning may have to be revisited.

C. *Dimensionality Reduction with PCA*

Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly_uncorrelated variables called **principal components**. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to preceding components. Principal components are guaranteed to be independent only if the data set is jointly normally distributed. The following are steps for the doing PCA analysis

PCA finds orthogonal variables that are linear combinations of original variables and can be used to find clusters in a set of data. It has been shown that the relaxed solution of K-means clustering, specified by the cluster indicators, is given by the PCA principal components, and the PCA subspace spanned by the principal directions is identical to the cluster centroid subspace specified by the between-class scatter matrix [10]. Thus PCA automatically projects to the subspace where the global solution of K-means clustering lies, and thus facilitates K-means clustering to find near-optimal solutions. Thus a combination of K-means clustering and PCA can be efficiently used to complete the delineation of the data into independent clusters. Once clusters are found they can then be used for pattern classification. The patterns can then be mapped to mental states as discussed below.

D. *Pattern Classification Analysis*

The PCA dimensionality reduction step prior to data segmentation improves the performance of the algorithm while reducing the noise because fewer features are considered. Pattern classification is the organization of patterns into groups where each group sharing the same properties, such as a given emotional state. We would then define a set of schemas to be used for classification. These schemas should be independent of each other. Once we chose a set of schemas, these can be used to classify into patterns. The classified patterns can then be used to extract the emotional state **Y** for the subject.

IV. Proof Of Concept With Non-EEG Data

For proof of concept validation of this approach, we used a non-EEG dataset from Physionet for twenty healthy individuals placed under physical stress, cognitive stress, and emotional stress with intermediate periods of relaxation. The data was collected using non-invasive wrist worn biosensors and consists of electrodermal activity (EDA), temperature, acceleration, heart rate (HR), and arterial oxygen level (SpO2) [11].

The procedure above was implemented in Python using the **sklearn** and **numpy** toolboxes, In addition, the WFDB python software package for reading and analyzing physiological signals is used [12].

The data-set had 8 phases per patient which are labelled as 'Relax1', 'PhysicalStress1', 'Relax2', 'MildEmotionalStress1', 'CognitiveStress1', 'Relax3', 'RealEmotionalStress1', 'Relax4' where each Relaxation phase is about five minutes long. The seven features corresponding to the seven signal channels are the three acceleration channels ('ax', 'ay', 'az'), temperature ('temp'), skin conductance ('EDA'), blood oxygenation ('SpO2') and Heart Rate ('HR'). The first five channels are sampled at 8 Hz and the last two channels at 1 sample/second. In an attempt to draw reasonable conclusions between the scatter plot results and the original features in pattern classification analysis stage, we chose to not consider the 'MildEmotionalstress1' Phase which was retrospectively classified a 'Stress' Phase by the team that collected the data.

For the analysis, a 12.5-minute segment of data was taken from each phase. We uniformly chose the first 12.5 minutes after the median sample of each phase, in order to account for outlier behavior at the start and end of each phase. We then calculated the root mean squared value for each segment, creating 8 RMS values for each of the 20 patients. The transformed dataset resulted in 160 rows across the 7 features above for subsequent processing. We also used other metrics such as mean and attained similar results. Picking the right features is critical for getting greater diagnostic fidelity.

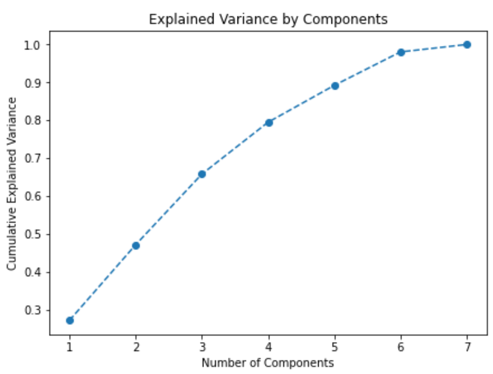

Figure 1: Cumulative Explained Variance Plot

The PCA operations were first performed to determine the optimal number of components to retain for the analysis in order to reduce dimensionality of the dataset, while maintaining 80% of the variance. The cumulative explained variances are plotted in Figure 1 and the minimum number of components that meets the 0.8 threshold sits around the first four components.

Next, we performed the PCA again, this time fitting the model to the 4 components and applying the dimensionality reduction to data. The resulting array contains the PCA scores across all 160 data points for the first 4 principal components.

K-Means clustering was performed using the PCA scores obtained to determine the optimal number of clusters for the dataset. This was determined using the Elbow-Method by observing the cluster WCSS (Within Cluster Sum of Squares) [13]. This process can be automated by looking at the slope of the cluster WCSS plot shown in Figure 2.

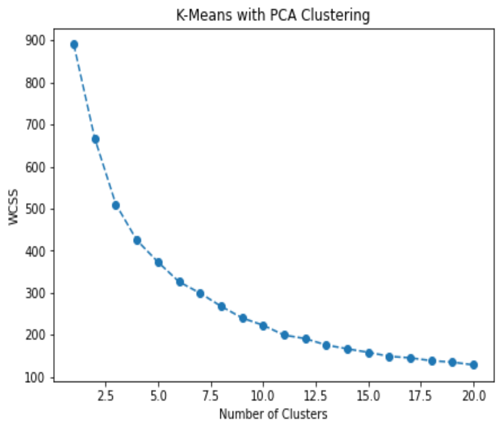

Figure 2: K-Means with PCA Clustering

We selected four clusters due to the "kink" in the graph where the slope changed character such that a drastic difference is not observed as the number of clusters increases. The cluster center coordinates in the PCA coordinate system are as shown in Table 1:

Cluster	Centers in PCA coordinates
1	(0.5399061, -0.779999, 0.9671625, 0.0691647)
2	(-1.5557462, 0.662821, 0.0603008, 0.0879936)
3	(1.6447472, 1.4751591, -0.2117194, -0.242909)
4	(0.0902372, -1.102798, -1.2077576, -0.0148019)

Table 1: Cluster Coordinates on four-dimensional plane

The inertia value, i.e. the measurement of "cluster spread" calculated by taking the sum of squared distances of samples to their closest cluster center, associated with these four clusters is 424.74. The tradeoff between minimizing the number of clusters while bringing the inertia as close to zero as possible yielded the choice of four clusters.

We can observe the separation of the four clusters using two-dimensional plots across two of the four PCA components at a time. The comparison of Component 2 vs Component 1, containing 160 data points separated into four clusters (distinguished by color) and identified by Phase (distinguished by marker) is shown in Figure 3.

Figure 3 shows a distinct separation between the second and third clusters, indicated by red and cyan colors, respectively. With pattern classification analysis, we see a strong correlation between the Physical Stress phase, indicated by the ‘Y-shaped' marker, and Cluster 3 in cyan.

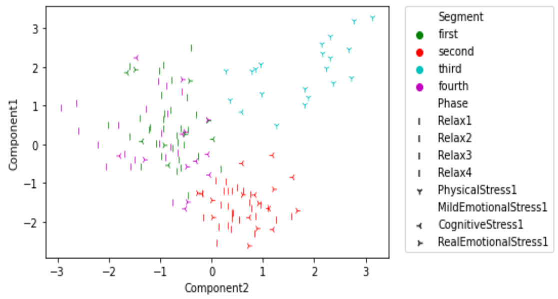

Figure 3: Clustered results Component2 vs Component1

This view does not indicate visible separation between the first and fourth clusters, so another two-dimensional plot was generated across Component3 and Component2 shown in Figure 4. This view shows distinct clusters between the first and fourth clusters, indicated by green and purple markers, respectively. For clusters 1 and 4, we were not able to identify any direct correlation with the original features in pattern classification analysis at this stage, as the data points within the first and fourth clusters appear to span across Relax and Stress states equally.

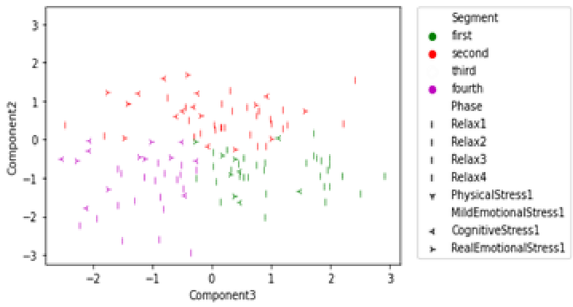

Figure 4: Clustered results - Component 2 vs Component 3

Other categorical data such as gender was used to see if any other indicators could be observed from the cluster scatter plot. Figure 5 shows the results of the same PCA/K-means analysis applied but using gender as the distinguishing factor across the four clusters.

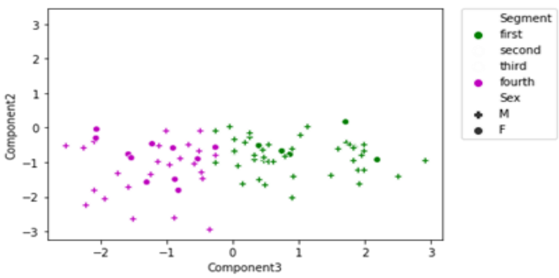

Figure 5: Clustered Results segmented by Gender

Observing the distribution of gender, we see that the first cluster consists mainly of male data points. The presence of female data points in Cluster 1 is marginal compared to the 48 total points representing female patients that appear in fully populated scatter plot

Figure 5 shows the distinct separation between clusters 1 and 4 in the two-dimensional plot of Component 3 vs Component 2, as seen in Figure 4 as well. However, analyzing the cluster separation on a gender-basis suggests that the first cluster, indicated by the color green, has a high correlation with the data points associated with male patients, indicated by the ‘+’ marker.

Lastly, we performed a K-means clustering on all seven features, not using principal component analysis for dimensionality reduction. This was done to justify the use of PCA through an observation of more separation between clusters and reduced inertia within clusters. Figure 6 shows the result of K-means clustering with 4 clusters, without PCA, on temperature and arterial oxygen level metrics.

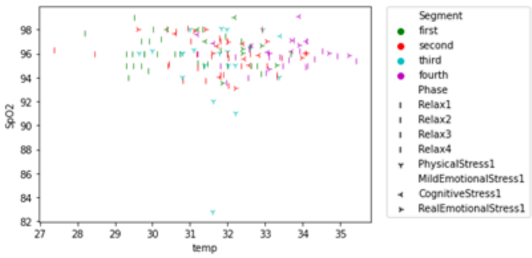

Figure 6: Clustered Temperature vs SpO2 Results without PCA

Figure 7 below applies the same K-means fit of the data, without undergoing PCA, on temperature and heart rate physiological data.

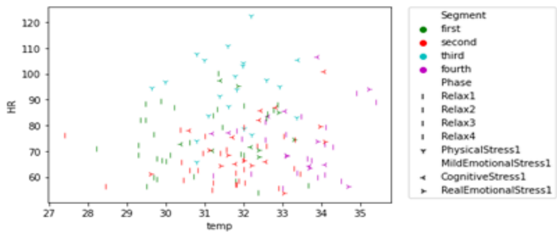

Figure 7: Clustered Temperature vs Heart Rate Results without PCA

Both Figure 6 and 7 show poor separation across the four clusters when passing in the RMS data across all seven metrics, which is essentially a PCA of seven components. Additionally, the inertia value obtained without PCA is 629.42, which is almost 50% larger than the inertia calculated for using four PCA components, and therefore a less optimal representation of the data.

We propose that the analysis steps performed above can be automated within an AI agent, yielding conclusions through pattern classification analysis that simply serve to augment the diagnostic approach of a subject-matter expert or physician.

V. PROPOSED FRAMEWORK

The validation of the approach using the non-EEG data set gives us the basis for a framework using a Medical AI Agent which could reside in a tablet or smartphone at the edge.

The algorithm output of the agent would be a multidimensional vector indicating the stress/mental state of the individual which would be measured during every electronic visit of the patient. Real-time visualizations of significant clinical inferences can yield insights into patients' mental states, enabling higher diagnostic specificity and targeted risk assessments. Data analytics is used to identify patterns and variations in the recorded data through signal processing algorithms, classification techniques, and natural language processing. The data from these devices is consolidated at the edge on the patient side and later stored in a central database with the appropriate identification stamp.

The clusters could be mapped to patterns corresponding to the neuropsychiatric state of the patient. Once an inference has been made about the mental/physiological state of the individual, the severity of the state has to be assessed and the individual is grouped into a low, medium or high-risk category. The diagnostic results can then be delivered to the physician for further evaluation.

Collection over several visits for subjects across diverse backgrounds will result in a progressively more intelligent agent with an increasingly sophisticated learning model. This will assist clinicians in identifying early diagnostic indicators of psychiatric conditions and proposing proper treatment plans in an effective manner. At the end of each session, a report would be generated and sent in a consolidated form to the physician and potentially interfaced with a Electronic Health Record system.

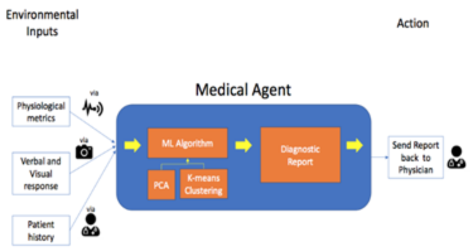

Figure 8: Proposed Medical Agent

Figure 8 represents the overall framework containing the Medical AI Agent. While our analysis *only* sampled physiological metrics, the agent would retrieve a combination of physiological parameters discussed above, verbal and visual inputs from the patient, and an annotated report of the patient history. The agent would process this data using machine learning techniques such as PCA and K-means, automatically invoking the feature extraction, dimensionality reduction, and clustering algorithms utilized in the prior section. Ultimately, the agent would generate a diagnostic report containing recommendations to the physician of potential areas requiring further investigation.

Through a smart and iterative approach, the data platform would integrate prior knowledge of the patient with these behavioral health measurements, enabling the agent to make intra-patient and inter-patient longitudinal comparisons to provide objective records of patient progress to the physician. The platform would provide real-time, quantifiable measures of cognition, behavior, and functional effectiveness by delivering actionable, objective clinical information of higher specificity and accuracy for behavioral healthcare and wellness.

A recent US patent issued to Rau et al. [14] provides comprehensive algorithms for monitoring emotions and stresses with cognitive biometric systems. Some of the diagnostic elements of their invention go beyond using additional sensors and provide methodologies for incorporating questionnaires and physician scored evaluations. As suggested in the patent, the system can be enhanced to have the following features:

- Patient's past behavioral health history, medical history, individual gender, socio-economic backgrounds and available family and social support systems are integrated to refine the standardized tests presented to them
- Provides objective records to provide quantifiable measures of longitudinal patient changes (improvement / deterioration
- to minimize inter-clinician and inter-health setting diagnosis and treatment variations with the support of objective records and inferences
- Time duration differences by patient in answering types of questions; and verbal descriptions --to provide inferences and changes in intensity/severity from each of the significant stressors and triggers unique to each patient.

VI. CONCLUSION AND NEXT STEPS

We presented a method to classify emotional states and implemented it using a publicly available non-EEG dataset of patients with neuro-psychiatric stressors. The dataset was collected on twenty patients with seven sensors from which multiple physiological features were extracted. The entire feature set was separated into clusters using Principal Component Analysis and K-means clustering.Analysis showed that one of the clusters was strongly correlated with Physical Stress as visualized in Figure 3. We also observed that there was a gender basis for another cluster which predominantly contained data points obtained from males patients, shown in Figure 5.

The medical agent and data platform is being implemented using the Galen Data Cloud ecosystem.[15] In a Medical AI Agent based framework, physicians and subject matter experts would initially need to decipher reports from the AI agent relaying strong correlations found in the clustered results. Using these findings, the physician can potentially narrow down the scope of their potential diagnosis or be led to pursue these pattern characteristics (e.g. gender, emotional state) further. If the agent's findings are weak, the physician may need to examine the clusters and the various features/data points to deduce patterns missed by the agent. Additional sensors for facial recognition, gestures, and voice recognition can enhance clustering and pattern recognition enabling meaningful diagnostic and risk group classification.

ACKNOWLEDGMENT

We wish to acknowledge psychiatrists Dr. Vikram Mehra and Dr. Aruna Gottumukkala for their insights and guidance in the diagnoses and treatment of patients with post traumatic stress disorders.

References

[1]. "6 Ways AI and Robotics Are Improving Healthcare." *Robotics Business Review,* https://www.roboticsbusinessreview.com/health-medical/6-ways-ai-and-robotics-are-improving-healthcare/

[2]. "Covid-19 health-care crisis could drive new developments in robotics, editorial says." *The Washington Post,* https://www.washingtonpost.com/health/covid-19-health-care-crisis-could-drive-new-developments-in-robotics-editorial-says/2020/03/27/5208e6ae-6ee7-11ea-a3ec-70d7479d83f0_story.html

[3]. "QUERI – Quality Enhancement Research Initiative." *Implementing Caring Contacts for Suicide Prevention in Non-Mental Health Setting,* www.queri.research.va.gov/visn_initiatives/suicide.cfm.

[4]. Boykin, D.M., Keegan, F., Thompson, K.E., Voelkel, E., and Fletcher, T.L. "Video to Home Delivery of Evidence-Based Psychotherapy to Veterans With Posttraumatic Stress Disorder", Frontiers of Psychiatry, December 2019

[5]. Bryan, Craig; Rudd David (2006). "Advances in the Assessment of Suicide Risk". Journal of Clinical Psychology 62 (2): 185–200

[6]. Hinrichs, R., Michopoulos V,, Winters S., Rothbaum, A.O, Rothbaum BO, Ressler KJ, and Jovanovic, T. "Mobile Assessment of Heightened Skin Conductance in Posttraumatic Stress Disorder", Depress Anxiety. 2017 June ; 34(6): 502–507

[7]. Wout, Mascha & Spofford, Christopher & Unger, William & Sevin, Elizabeth & Shea, Tracie. (2017). Skin Conductance Reactivity to Standardized Virtual Reality Combat Scenes in Veterans with PTSD. Applied psychophysiology and biofeedback. 42. 10.1007/s10484-017-9366-0.

[8]. Wickramasuriya, D.S., Faghih, R., "A Bayesian Filtering Approach for Tracking Arousal from Binary and Continuous Skin Conductance Features" IEEE Trans Biomed Eng 2020 Jun;67(6):1749-1760

[9]. Wickramasuriya, D.S., Amin, M.R. Faghih, R., "Skin Conductance as a Viable Alternative for Closing the Deep Brain Stimulation Loop in Neuropsychiatric Disorder" Frontiers in Neuroscience August 2019, Article 780

[10]. Ding, C., He, X, "K-means Clustering via Principal Component Analysis", Proceedings of the 21st International Conference on Machine Learning, Banff, Canada, 2004.

[11]. Birjandtalab, Javad, Diana Cogan, Maziyar Baran Pouyan, and Mehrdad Nourani, A Non-EEG Biosignals Dataset for Assessment and Visualization of Neurological Status, 2016 IEEE International Workshop on Signal Processing Systems (SiPS), Dallas, TX, 2016, pp. 110-114. doi: 10.1109/SiPS.2016.27

[12]. Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220

[13]. "How to Combine PCA and K-means in Python?" *365 Datascience* , https://365datascience.com/pca-k-means/

[14]. Rau, N., Rau, H., Inguva, R., Baskaran, V., Story, R., "Cognitive Biometric System to Monitor Emotions and Stress",United States Patent and Trademark Office 10, 617,351B2, 14 April 2020

[15]. Cloud Platform Purpose Built for Medical Devices. (2020, September 03). Retrieved September 21, 2020, from https://www.galendata.com/