## PROPOSING MENTAL-HEALTH INTERVENTION STRATEGIES BASED ON CLIENT INTAKE FORM DATA (CogniXR)

**Business Problem:**

Depression and PTSD (post-traumatic stress disorder) are both mental health conditions that can have a significant impact on a person's daily life. Depression is characterized by persistent feelings of sadness, hopelessness, and a lack of interest or pleasure in activities while PTSD is a condition that can develop after a person experiences or witnesses a traumatic event, such as a natural disaster, combat, sexual or physical assault, or a serious accident. Both depression and PTSD can be treated with therapy, medication, and other forms of support. It's important for anyone experiencing symptoms of either condition to seek professional help.

Mental health intervention strategies can vary based on the specific needs and concerns of the client as identified through their input data. Through the form we are trying to identify Cognitive Behavioral Therapy (CBT) that can help clients identify their problem, seek the help in order to cope with their situation and get better.

We will be dealing with the business problem by following the below steps:

**Step 1: Gathering Data -** Firstly, we will be gathering data by creating questionnaires, using the information collected from the mental health intake form to attain insights of the demographic, behavioral information of clients for appropriate mental therapy.

**Step 2: Organizing Data:** Next, we will be organizing the data by storing, cleaning and retrieving data for analysis. This will allow us easy access, retrieval and use of the information for the purpose of developing user personas [2] like Anxiety, PTSD, Trauma and Substance abuse for targeted mental health interventions.

**Step 3: Developing Machine Learning Models:** Finally, We will use transformer like BERT for extracting features from textual inputs and combine with appropriate ML models like SVM, decision trees and random forest, which will help us in the classifying user personas like Anxiety, PTSD, Trauma and Substance abuse and therapy based on data collected from the form.

**Evaluation Metrics:**

As we will be collecting the data using user intake forms, we will be dealing with text inputs which leads us to Natural Language Processing (NLP) [1].For the evaluation of the performance of the classification problem, there are several metrics. Such as:

* Accuracy: Accuracy is used to measure how close is the to the known value by the model.
* Precision and Recall: Precision and Recall are related metrics that are used to measure the ability of the model to correctly identify positive cases (precision) and all relevant cases (recall).
* F1 Score: F1 Score is the mean of precision and recall and is often used as a single metric to summarize the trade-off between them.
* Confusion Matrix: This is a table that provides a detailed breakdown of the model's performance, showing the number of true positives, true negatives, false positives, and false negatives that will help us calculate the accuracy, precision and recall to evaluate the model.
* ROC Curve: Receiver Operating Characteristic (ROC) curves are often used in binary classification problems and show the trade-off between the true positive rate and false positive rate as the threshold for classifying examples is varied in monitoring the performance of the model.

**Research Questions:**

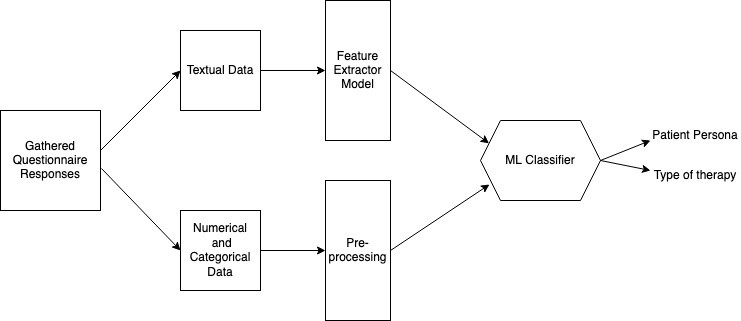
* What are the mental health problems people have?
* What triggers them the most?
* What kind of help do they need?

**Analytical Methods:**

**Development of Psychometric Questionnaire:** As the questionnaire will be used as the method of data collection, the questions will be carefully designed to capture data points required for identifying depression or PTSD and the type of therapy required.

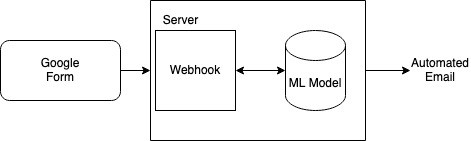
**Data Gathering and Labeling:** The data will be gathered through Google Forms. The data received will be processed and clustered to identify groups with similar responses and thus persona. Each group identified from the cluster analysis will be closely analyzed by mental-health experts to create appropriate labels from personal and type of therapy required. The labeled data will be then fed into the ML Model for classification.

**Development of ML Model:** The textual data will be separated from quantitative and categorical data gathered. A NLP model such as BERT will be used to extract features from textual inputs and convert them into a feature vector, we will combine feature vector from BERT with other inputs and do classification with ML models like SVM(These models can be used to classify clients into different intervention groups based on their intake form data, using a combination of linear and non-linear decision boundaries), decision trees(These models can be used to classify the type of intervention that would be most effective for a particular client based on their intake form data) and random forest(These models are an extension of decision trees and can help to improve their accuracy by combining the results of many decision trees)



*Figure: Workflow diagram*

**Integration of ML Model with Google Form:** The ML Model will be deployed and accessible through Webhook [3]. Once any new submissions are made in google form, the webhook will receive data submitted and classify survey response. An automated email will be sent to mental-health practitioners and concerned authorities notifying about the survey response and predicted mental health condition and appropriate therapy for the user.



*Figure: Application Architecture*

**Pipeline:**

Understand client’s requirements → Prepare questionnaires → Gather necessary data → Preprocess the data → Develop ML models → Create user personas based on Depression and PTSD → Discuss with client for their opinions and inputs → Include changes suggested by client → Finalize the changes

→ Deploy ML Model → Integrate ML Model with Google Form → Send automated email to therapist

# **References:**

1. The Most Common Evaluation Metrics In NLP. (2021). Retrieved 26 January 2023, from <https://towardsdatascience.com/the-most-common-evaluation-metrics-in-nlp-ced6a763ac8b>
2. ten Klooster, I., Wentzel, J., Sieverink, F., Linssen, G., Wesselink, R., & van Gemert-Pijnen, L. (2022). Personas for Better Targeted eHealth Technologies: User-Centered Design Approach. JMIR Human Factors, 9(1), e24172. doi: 10.2196/2417. Retrieved 26 January 2023, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8965674/
3. Sending a Webhook for each Google Forms Submission. (2020). Retrieved 27 January 2023, from <https://medium.com/@eyalgershon/sending-a-webhook-for-each-google-forms-submission-a0e73f72b397>