**Development of an AI-Based Triage System for Mental Health Assessment in Telehealth Text Communications**

**Introduction**

Mental health is critical for human development and economic stability. World Health Organization(WHO) describes mental health to be a state in which individuals realizes their potential, cope with normal life stressors, works productively, and contributes to their communities(World Health Organization, 2013). However for most of the global population attaining this state is challenging, contributing to over one billion people living with mental or addictive disorders resulting in comorbidities and significant risk of premature death(Rehm & Shield, 2019). The economic impact of mental health disorders results from loss of income and productivity and from medical costs. This drives poverty at the individual level and at the national level it diminishes capital because abled labor diminishes (Canavan et al., 2013) as was demonstrated by a study conducted in 2010 that estimated losses due to mental disorders at 1.3 trillion USD and projected the loss to reach 2.5 trillion USD by 2030 (Bloom et al., 2012) which is considered an under estimate by new studies(GBD 2019 Diseases and Injuries Collaborators, 2020; GBD 2019 Mental Disorders Collaborators, 2022).

The onset of COVID-19 pandemic triggered nationwide lockdowns that triggered known mental health risk factors including social isolation, stress and anticipated economic hardships (Economou et al., 2019; Winkler et al., 2020). The risk factors disproportionately affects individuals who have known mental health problems, high risk groups such as individuals with pre-existing chronic diseases and of certain occupations like health care workers, COVID-19 patient survivors and unemployed individuals (Galea et al., 2020; Hao et al., 2020; Kang et al., 2020; Lu et al., 2020; Wang et al., 2020; Yao et al., 2020; J. Zhang et al., 2020; S. X. Zhang et al., 2020). The risk factors also triggered new ones (Pierce et al., 2020). Adolescents and children also showed high prevalence of mental health by type between 28% and 48% of adolescents (Ma et al., 2021). These were attributed to prolonged school closures and strict social isolation that affected their sleep and wake cycles, physical exercise routines and excessive use of technology (Ma et al., 2021). During the lock down there was also the risk of children witnessing or experiencing violence (Costello et al., 2003)

Several barriers to identification of persons requiring mental health services exists. Mental disorders are subject to negative judgment and stigmatization exposing the patients to cope with the devastating illness, exclusion and prejudice (Rössler, 2016). Stigma therefore reduces individuals’ self-esteem and self-efficacy limiting their prospects for seeking help and prospects of recovery. A second important barrier is cost resulting in increased rate of forgoing medical care (Dedania & Gonzales, 2019) and lastly the difficulty of finding physical facilities for Mental Health care (Abar et al., 2017) which in part is because of the persistence and severity of Mental Health cases that complicates delivery of care.

As the prevalence of mental health gets higher and progress towards identification and access to care is not improving at the same rate, solutions are required to augment the health systems dedicated to mental health management and care. One such solution leverages in technology. Machine learning and artificial intelligences have been used to help solve challenging diagnostic problems in medical field. Its application in multiple psychological treatments have shown excellent potential for predicting occurrence and treatment of Mental health conditions (Iyortsuun et al., 2023). ML and AI have been used in schizophrenia prediction(Srinivasagopalan et al., 2019), depression and anxiety detection (Aleem et al., 2022; Sau & Bhakta, 2017), bipolar disorder detection (Li et al., 2020) etc.

In recent years, the rise of telehealth services has provided unprecedented access to healthcare support, including mental health services. However, the increasing volume of text-based communications between patients and healthcare providers a lot of qualitative data, often brief and therefore poses a significant challenge in promptly identifying and addressing mental health issues such as depression and anxiety. This is because the brevity of such text would include subtle trances of mental health issues that can escape the health providers.

Advancement in Natural language processing (NLP) promises avenues for early detection and intervention in mental health care. By analyzing linguistic cues that may indicate psychological distress, NLP provides a powerful, non-invasive tool for triaging individuals who might not seek help due to stigma or lack of access to care. This technology can process diverse textual datasets, including social media posts, healthcare narratives, and digital communication, identifying patterns that human practitioners might overlook. As such, integrating NLP into mental health services represents a forward leap in preemptive and personalized care, enabling timely support for those at risk (Di Cara et al., 2023; Guntuku et al., 2017; Miner et al., 2017).

The initiative aims to utilize natural language processing to swiftly pinpoint potential mental health issues from text-based interactions on the one2one platform, operated by LVCT Health. This digital health service offers a toll-free line for health-related inquiries, managed by counselors providing initial counseling. The current manual review of texts for urgent mental health follow-ups hinders prompt support. An automated system is proposed to streamline this process, allowing counselors to promptly identify and address clients' mental health needs.

**Objectives**

1. Develop an AI Model: To create a machine learning model that can analyze text messages for indicators of mental health issues.

2. Accuracy and Sensitivity: Ensure the model is both accurate and sensitive in identifying potential mental health concerns.

**Methods**

**Methodology**

**Data Collection and Preparation**

**Dataset**: Utilize an existing dataset of text messages from a telehealth platform, ensuring compliance with data protection regulations like GDPR and HIPAA. This dataset will be anonymized to maintain patient confidentiality.

**Data Preprocessing**: Implement preprocessing steps including tokenization, stemming, and lemmatization to structure the data for NLP analysis. Remove irrelevant features and handle missing values to ensure data quality and consistency.

**Model Development**

**Natural Language Processing Techniques**:

1. Text Representation: Use techniques like TF-IDF (Term Frequency-Inverse Document Frequency) and word embeddings (e.g., Word2Vec, GloVe) to convert text data into a format suitable for machine learning algorithms.
2. Sentiment Analysis: Apply sentiment analysis to gauge the emotional tone of the messages, which can be indicative of mental health states.
3. Topic Modeling: Employ algorithms like Latent Dirichlet Allocation (LDA) to identify prevalent themes or topics within the text data, which could signal specific mental health issues.

**Machine Learning Models:**

1. Experiment with various machine learning models such as Support Vector Machines (SVM), Random Forests, and Neural Networks. Evaluate the performance of each model in classifying text messages with potential mental health concerns.
2. Deep Learning Approaches: Investigate the use of deep learning techniques like Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), which have shown promise in text classification tasks.

**Training and Validation**:

Split the dataset into training, validation, and test sets to evaluate the model's performance. Use cross-validation techniques to ensure the model's robustness and generalizability.

**Evaluation**

1. Performance Metrics: Assess the model using metrics specific to NLP tasks, such as BLEU scores for language generation quality, alongside traditional metrics like accuracy, precision, recall, and F1-score.
2. Iterative Testing and Refinement: Conduct iterative testing cycles, incorporating feedback from healthcare professionals to refine the model and improve its accuracy and usability.