Automated Detection of Major Depressive Disorder through Question-Answering

Project Report-2 submitted to
Indian Institute of Technology Kharagpur
in partial fulfilment for the award of the degree of
Bachelor of Technology
in
Computer Science and Engineering

by

Srijanak De

(19CS30047)

Under the supervision of Professor Partha Pratim Chakrabarti

Jointly Guided by: Professor Aritra Hazra, Professor Rajlakshmi Guha

Department of Computer Science and Engineering

Indian Institute of Technology Kharagpur

Spring Semester, 2022-23

April 10, 2023

DECLARATION

I certify that

(a) The work contained in this report has been done by me under the guidance of my supervisor.

(b) The work has not been submitted to any other Institute for any degree or diploma.

(c) I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

(d) Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the textof the thesis and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.

Date: April 10, 2023 (Srijanak De)

Place: Kharagpur (19CS30047)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

KHARAGPUR - 721302, INDIA

CERTIFICATE

This is to certify that the project report entitled "**Automated Detection of Major Depressive Disorder through Question-Answering**" submitted by **Srijanak De** (Roll No. 19CS30047) to Indian Institute of Technology Kharagpur towards partial fulfilment of requirements for the award of degree of Bachelor of Technology in Computer Science and Engineering is a record of bona fide work carried out by him under my supervision and guidance during Spring Semester, 2022-23.

Professor Partha Pratim Chakrabarti
Department of Computer Science and Engineering
Indian Institute of Technology Kharagpur
Kharagpur – 721302, India

Date: April 10, 2023

Place: Kharagpur

Abstract

Name of the student: **Srijanak De** Roll No: **19CS30047**

Degree for which submitted: **Bachelor of Technology**

Department: **Department of Computer Science and Engineering**

Project title: **Automated Detection of Major Depressive Disorder through Question-Answering**

Project supervisor: **Professor Partha Pratim Chakrabarti**

Month and year of report submission: **April 10, 2023**

Automated detection of depression entails asking patients' relevant questions in accordance with an established diagnostic criterion and thereby analyzing the corresponding responses. In this work, we propose a novel automated depression detection pathway which is adaptive to newer contexts, personified to patient's responses, and flexible to rules as per user's choices. Users can flexibly choose from or add to a pool of questions and can modify their default order as and how they wish to within the diagnostic criteria guidelines. Additionally, they can choose between various diagnostic criteria and what outcomes of diagnosis they want to see. They can also define their own rules of diagnosis, that is, which order of responses lead to which outcome category. The goal of the project was to make the depression detection tool as flexible as possible for the user while keeping its usage simplicity to the maximum.

Acknowledgements

I would first like to thank my project supervisor Professor Partha Pratim Chakrabarti for his constant guidance and support. He consistently allowed this project to be a result of my own thoughts and efforts, and at the same time kept on steering me in the right direction. Furthermore, this project would not have been possible without the apt guidance of Professor Aritra Hazra and Professor Rajlakshmi Guha.

Finally, I must express my profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this report. This accomplishment would not have been possible without them. Thank you.

Srijanak De

Contents

Abbreviations

MDD	**M**ajor **D**epressive **D**isorder
DSM-5	**D**iagnostic and **S**tatistical **M**anual of Mental Disorders, **Fifth** Edition
SCID-5-CV	**S**tructured **C**linical **I**nterview For DSM-5 **D**isorders **C**linical **V**ersion
WHO	**W**orld **H**ealth **O**rganization
DARPA	**D**efense **A**dvanced **R**esearch **P**rojects **A**gency
DAIC	**D**istress **A**nalysis **I**nterview **C**orp
WoZ	**Wizard-of-Oz**
DAIC-WOZ	**D**istress **A**nalysis **I**nterview **C**orp **Wizard-of-Oz**
PHQ-8	**P**atient **H**ealth **Q**uestionnaire - **8**
BDI	**B**eck's **D**epression **I**nventory
DASS	**D**epression **A**nxiety and **S**tress **S**cale
DRS	**D**epression **R**ating **S**cale
APA	**A**merican **P**sychiatric **A**ssociation
AMC	**A**nother **M**edical **C**ondition
RMSE	**R**oot **M**ean **S**quared **E**rror
MAE	**M**ean **A**bsolute **E**rror
BGRU	**B**idirectional **G**ated **R**ecurrent **U**nit
ML	**M**achine **L**earning
NLP	**N**atural **L**anguage **P**rocessing
NLU	**N**atural **L**anguage **U**nderstanding

Chapter 1

Introduction

Depression is one of the most common mental illnesses, impacting billions of people worldwide. Often people suffering from depression don't realize that they are undergoing a mental health condition demanding clinical attention. Sometimes people don't feel comfortable going up to or sharing information with an unknown person to talk about their mental health. There are instances when mental health treatment is not even accessible to people who need it. The main aim of this work is to help clinicians in their diagnosis by automatically detecting depression in patients. Using this automated depression detection tool, we attempt to provide a remedy to all the above scenarios and more. We hope that once launched, this tool will be able to make the lives of both the clinicians as well as patients suffering from depression easier.

1.1 Objective and Problem Definition

Automated Detection of Major Depressive Disorder through Question-Answering

The objective of this project is to create a fully automated system that engages in interaction with a patient mainly through asking a series of questions and analyzes their corresponding responses to detect if the patient is diagnosed with depression or not, what might be its reason or the severity of it. The questions, their order, the diagnostic outcomes, the diagnostic criteria to be followed for detection can all be preselected by the clinician conducting the diagnosis or they can choose to use the default option for each choice. The automated depression detection system should be adaptive to newer contexts and demographics, personified to patient's responses, and flexible to making conversations on the go. This task is segregated into the following three stages of development:

1. Creating a baseline questionnaire structure
2. Making the baseline adaptive to newer contexts and rules
3. Handling full blown textual interactions and responding on the go
4. Analysis of responses at the backend

Creating a baseline questionnaire structure

SCID-5-CV gives a proper structuring of questions to be asked to the subject based on the DSM-5 diagnostic criteria. It follows a rigid graph structure which leads to a deterministic detection of MDD based on the diagnostic criteria. The SCID-5-CV questionnaire was used as the baseline model while accepting only 'yes' or 'no' answers to automatically detect depression based on the subject's responses.

Making the baseline adaptive to newer contexts and rules

At first, the pool of questions was increased by creating relevant questions pertaining to personified contexts and Indian demographics vetted by professional clinicians. The pool of questions was further increased by adding questions pertaining to different diagnostic criteria and questionnaires namely PHQ-8, BDI, DASS-21 and Hamilton DRS. DSM-5 is used as the default diagnostic criteria and SCID-5-CV is used as the default question pool and order. But clinicians who want to use this tool to diagnose patients can add more questions, change their order, define their own diagnostic criteria they want to use for detection of MDD, and also change the outcome classes they want to see their patient's diagnosed with. Additionally, they can also define their own rules using which responses to each question lead to the output category of their choice. The default of SCID-5-CV has been set at all instances to make the use of the tool easier for the users.

Handling full blown textual interactions and responding on the go

Since, as of now, the baseline model only accepts 'yes' or 'no' as an answer to each of the questions, the immediate next goal should be to accept textual responses. The main challenge which arises from accepting textual answers is that a predefined question order no longer makes sense to be followed. Textual responses should be analyzed and replied on the go while at the same time the answers should ultimately lead to a proper diagnosis of depression based on the used diagnostic criteria. Moreover, since now full textual responses will be accepted the model should be adaptive to newer contexts which may arise on the go and try to zero-shot learn their analysis and come up with the next response and diagnosis.

Analysis of responses at the backend

Analyses of types of responses, duration at each kind of question, at which question subject's mostly leave the platform without answering and anything else that might help clinicians will be done using the data collected at the backend. Further machine learning based predictive analysis and data analysis can be performed on the data to find anomalies and patterns which might reveal additional relevant information about patient mental health. The first version of the generalized model is now ready to be deployed, possibly with the scope of making the handling of the tool easier with some modifications. Once the data is collected from users using this tool, the respective analyses can be started.

1.2 Motivation

Depression is a common illness worldwide, with an estimated 3.8% of the population affected, including 5.0% among adults (WHO). At its worst, depression can lead to suicide. Over 700,000 people die due to suicide every year (WHO). The lack of adequate resources is an immediate cause of awareness. With the advent of AI, automated depression detection tools can now ease the work of clinicians by flagging people with the risk of depression. But as AI is a black box algorithm, the automated diagnosis can be less practical or rigid in nature. Also, an AI is as strong as the data it is fed with and since the data is not exhaustive to different contexts and demographics, the same tool is not the best fit to be used across the world. To give the clinicians flexibility to choose questions while understanding the underlying mechanism of depression detection as well as to bring adaptability to the tool with respect to newer contexts, demographics and rules, a novel pathway is proposed in this project.

In the remainder of this report, we provide a detailed literature survey of related works done so far and outline the literature gaps in them. We then give a brief background required to understand the problem definition better. In the following sections we explain the work done so far and the future works that we plan to do.

Chapter 2

Literature Survey

2.1 Related Work

In this section, several related works are presented which are probable alternatives to a similar goal as this project. We also outline the literature gaps in them. DARPA, by using previous clinical interview videos, created a virtual automated interviewer to diagnose depression in patients. They also maintain a database of clinical interviews including their audio and video recordings as well as their transcripts and extensive questionnaire responses. The state-of-the-art in auto-detection of depression uses all three modalities, audio, visual and text to detect depression using attention-based deep neural networks. Other notable methods using only the transcripts of interviews include building a multi-scale bidirectional GRU with pretrained word embeddings and using a hierarchical attention network.

2.2 Literature Gaps

Among the existing literature very less focus is given to zero-shot detection of MDD. Zero-shot detection is of utmost importance in this field due to the high variability in the patient-clinician interaction. Even when the ML model encounters a new set of question-answer pairs, using a pretrained model it should be able to do zero-shot learning and detect the presence of MDD dynamically at runtime for more scalability and practicality of the model.

Interpretability has been an age-old problem with ML models, more now with the increasing complexity of the models. This problem is further enlightened in the clinical field as the models have questionable practical applicability without minimal clinical justifiability. The current ML models for depression detection mostly lack interpretability and clinical justifiability.

Furthermore, the current literature puts no focus on the importance of personal, cultural, and demographic variables in depression detection. There exist several models for text generation with demographic context, but they have no relation to the detection of depression at present. Moreover, no notable work has been done in the field of auto-generation and auto-structuring of questions for detection of depression. This is important for personalizing the questionnaire for each subject as well as for maintaining the context and continuity of the interaction.

2.3 Dataset

The Distress Analysis Interview Corpus (DAIC) [Gratch et al., 2014]

This is the most widely used dataset across all existing works in the field of depression detection. The database contains clinical interviews designed to facilitate the diagnosis of psychological distress conditions such as anxiety, depression, and post-traumatic stress disorder. The dataset includes audio and video recordings, their transcripts and extensive questionnaire responses. The DAIC-WOZ part of the corpus includes data from the Wizard-of-Oz (WoZ) interviews, conducted by an animated virtual interviewer called Ellie, controlled by a human interviewer in another room.

2.4 Detection of Depression

1. **The Verbal and Non-Verbal Signals of Depression -- Combining Acoustics, Text and Visuals for Estimating Depression Level [Qureshi et al., 2019]**

This paper proposes a novel attention based deep neural network to regress depression level. It facilitates the fusion of all three modalities, acoustic, text and visual. The model has been experimented with on the DAIC-WOZ dataset. From the results, it is empirically justified that the fusion of all three modalities helps in giving the most accurate estimation of depression level. The proposed approach outperforms the state-of-the-art by 7.17% on RMSE and 8.08% on MAE.

2. **Text-based depression detection on sparse data [Dinkel et al., 2019]**

This paper proposes a text-based multi-task BGRU network with pretrained word embeddings to model patients' responses during clinical interviews. The focus of the paper is on handling the sparse data scenario of clinical interviews. The main approach uses a novel multi-task loss function, aiming at modeling both depression severity and binary health state. Word and sentence-level word-embeddings as well as the use of large-data pretraining for depression detection are independently investigated. To strengthen the findings, mean-averaged results for a multitude of independent runs on sparse data are reported. It is experimentally verified that pretraining is helpful for word-level text-based depression detection. Additionally, the results demonstrate that sentence-level word-embeddings should be mostly preferred over word-level ones. While the choice of pooling function is less crucial, mean and attention pooling should be preferred over last-timestep pooling. The method outputs depression presence results as well as predicted severity score, culminating a macro F1 score of 0.84 and MAE of 3.48 on the DAIC-WOZ development set. It is important to note that the F1 score of 0.84 is for single-fold runs, whereas for five-fold runs the best F1-score is 0.69.

3. Affective Conditioning on Hierarchical Attention Networks applied to Depression Detection from Transcribed Clinical Interviews [Xezonaki et al., 2020]

This paper proposes an ML model for depression detection from transcribed clinical interviews. According to the paper depression is a mental disorder that impacts not only the subject's mood but also the use of language. To this end, the paper uses a Hierarchical Attention Network to classify interviews of depressed subjects. The attention layer of the model is augmented with a conditioning mechanism on linguistic features, extracted from affective lexica. A detailed analysis was performed, and the results show that individuals diagnosed with depression use affective language to a greater extent than not depressed. The experiments show that external affective information improves the performance of the proposed architecture in the General Psychotherapy Corpus and the DAIC-WOZ 2017 depression datasets, achieving state-of-the-art 71.6 and 68.6 F1 scores (for five-fold runs) respectively.

Chapter 3

Background and Preliminaries

3.1 Major Depressive Disorder (MDD)

Definition - *"In DSM–5, a mood disorder characterized by persistent sadness and other symptoms of a major depressive episode but without accompanying episodes of mania or hypomania or mixed episodes of depressive and manic or hypomanic symptoms is called Major Depressive Disorder."* (Source: APA)

3.2 Diagnostic Criteria for MDD

Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)

Table 1 contains the DSM-5 diagnostic criteria for MDD. The column "Sustained" is marked if the symptom has been sustained for at least two weeks, every day, most of the day. If the symptom is "clearly present" then that column is marked. For a diagnosis of MDD to be present, 5 out of 9 criteria from Section A must be marked as BOTH "clearly present" and "sustained" as well as, criteria B and criteria C must be met. Simultaneously, items C, D and E must be clearly present.

Clearly Present	Sustained	
		A) Five (or more) of the following symptoms have been present during the same 2-week period and represent a change from previous functioning; at least one of the symptomsis either (1) depressed mood or (2) loss of interest or pleasure. *(Note: Do not include symptoms that are clearly attributable to another medical condition)*
		1) Depressed mood most of the day, nearly every day as indicated by either subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful). *(Note: In children and adolescents, can be irritable mood).*
		2) Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation).

		3) Significant weight loss when not dieting or weight gain (e.g., a change of more than 5%of body weight in a month), or decrease or increase in appetite nearly every day. *(Note: In children, consider failure to make expected weight gain.)*
		4) Insomnia or hypersomnia nearly every day.
		5) Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down).
		6) Fatigue or loss of energy nearly every day.
		7) Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).
		8) Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others).
		9) Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation withouta specific plan, or a suicide attempt or a specific plan for committing suicide.
		B) The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
	C) The episode is not attributable to the physiological effects of a substance or to another medical condition.	
	Note: Criteria A-C represents a major depressive episode *Note: Responses to a significant loss (e.g., bereavement, financial ruin, losses from a natural disaster, a serious medical illness or disability) may include the feelings of intense sadness, rumination about the loss, insomnia, poor appetite, and weight loss noted in Criterion A, which may resemble a depressive episode. Although such symptoms may be understandable or considered appropriate to the loss, the presence of a major depressive episode in addition to the normal response to a significant loss should also be carefully considered. This decision inevitably requires the exercise of clinical judgment based on the individual's history and the cultural norms for the expression of distress in the context of loss.*	
	D) The occurrence of the major depressive episode is not better explained by schizoaffective disorder,schizophrenia, schizophreniform disorder, delusional disorder, or other specified and unspecified schizophrenia spectrum and other psychotic disorders.	
	E) There has never been a manic episode or a hypomanic-like episode. *Note: This exclusion does not apply if all of the manic-like or hypomanic-like episodes are substance-induced or are attributes to the physiological effects of another medical condition.*	

Table 1: DSM-5 Diagnostic Criteria (Source: APA) [5]

3.3 Clinical Interview Questions for Diagnosis of MDD

Structured Clinical Interview for DSM-5 Disorders Clinician Version (SCID-5-CV)

The SCID-5-CV structure of questioning directly follows from the DSM-5 criteria. Figure 1 shows some questions from the SCID-5-CV questionnaire along with the DSM-5 diagnostic criteria they pertain to. Some salient features of the structuring of the questions are as follows:

i. For each DSM-5 criteria some questions focus on asserting if the criteria is “clearly present” nearly every day, most of the day. If it is “clearly present”, then the next set of questions try to figure out if it “sustained” for at least 2 weeks. If and only if both the conditions are met, the “+” mark is selected for that DSM-5 criteria or in other words, that criteria is satisfied.

   Example: “In the past month, since (ONE MONTH AGO), has there been a period of time when you were feeling depressed or down most of the day, nearly every day? (Has anyone said that you look sad, down, or depressed?)” – This question tries to ascertain if the criteria is “clearly present”.

   “IF YES TO EITHER OF ABOVE: What has it been like? How long has it lasted? (As long as 2 weeks?)” – If the criteria is “clearly present” then this question tries to figure out if the criteria sustained for at least 2 weeks.

ii. There are some instances where the order of the DSM-5 criteria is important but in most cases that order is flexible. It is also important to note that due to the presence of Sections B, C, D and E of DSM-5 there are some additional heuristics to the structure of the questioning which makes part of the structures more rigid.

   Example: “IF PREVIOUS ITEM RATED ‘+’: During that time, did you have less interest or pleasure in things you usually enjoyed? (What has that been like?)” – The first question pertaining to DSM-5 criteria 2 requires the clear and sustained presence of DSM-5 criteria 1 and follows directly after it. No such ordering in most cases.

iii. Multiple questions along with different variations of them are asked to assert the presence or sustenance of each DSM-5 criteria.

   Example: “How has your appetite been? (What about compared to your usual appetite? Have you had to force yourself to eat? Eat [less/more] than usual? Has that been nearly every day? Have you lost or gained any weight?)” – Multiple questions and their variations are asked to ascertain marked change in appetite of the subject.

A. MOOD EPISODES

	CURRENT MAJOR DEPRESSIVE EPISODE	MAJOR DEPRESSIVE EPISODE CRITERIA		
	Now I am going to ask you some more questions about your mood.	A. Five (or more) of the following symptoms have been present during the same 2-week period and represent a change from previous functioning; at least one of the symptoms is either (1) depressed mood or (2) loss of interest or pleasure.		
A1	**In the past month, since** (ONE MONTH AGO), **has there been a period of time when you were feeling depressed or down most of the day, <u>nearly every day</u>? (Has anyone said that you look sad, down, or depressed?)** IF NO: **<u>How about feeling sad, empty, or hopeless, most of the day, nearly every day?</u>** IF YES TO EITHER OF ABOVE: **What has it been like? How long has it lasted? (As long as 2 weeks?)**	1. Depressed mood most of the day, nearly every day, as indicated by either subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful).	— +	A1
A2	→ IF PREVIOUS ITEM RATED "+": **During that time, did you have less interest or pleasure in things you usually enjoyed? (What has that been like?)** → IF PREVIOUS ITEM RATED "—": **What about a time since** (ONE MONTH AGO) **when you lost interest or pleasure in things you usually enjoyed? (What has that been like?)** IF YES TO EITHER OF ABOVE: **<u>Has it been nearly every day?</u> How long has it lasted? (As long as 2 weeks?)**	2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation).	— +	A2
	IF BOTH **A1** AND **A2** ARE RATED AS "—" FOR THE CURRENT MONTH, Continue with **A15** (Past Major Depressive Episode), **page 13.**			
	FOR THE FOLLOWING QUESTIONS, FOCUS ON THE WORST 2-WEEK PERIOD OF THE PAST MONTH: **During** (2-WEEK PERIOD)**...**			
A3	**...how has your appetite been? (What about compared to your usual appetite? Have you had to force yourself to eat? Eat [less/more] than usual? <u>Has that been nearly every day?</u> Have you lost or gained any weight?)** **IF YES: How much? (Had you been trying to [lose/gain] weight?)**	3. Significant weight loss when not dieting or weight gain (e.g., a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day.	— +	A3

Figure 1: SCID-5-CV Questionnaire (Source: SCID-5-CV) [6]

Chapter 4

Summary of Work Done

In this work, we try to create a fully automated depression detection tool. This entails asking patients' relevant questions in accordance with an established diagnostic criterion and thereby analyzing the corresponding responses. We propose a novel automated depression detection pathway which is adaptive to newer contexts, personified to patient's responses, and flexible to rules as per user's choices. Users can flexibly choose from or add to a pool of questions and can modify their default order as and how they wish to within the diagnostic criteria guidelines. Additionally, they can choose between various diagnostic criteria and what outcomes of diagnosis they want to see. They can also define their own rules through which a set of responses lead to a desired outcome. In this work we use DSM-5 as the default diagnostic criteria and SCID-5-CV as the default question pool and order. In this section we outline the work done until now per segregation of the task as depicted in the problem definition.

4.1 Creating a baseline questionnaire structure

The SCID-5-CV questionnaire was used as the baseline model while accepting only 'yes' or 'no' answers to automatically detect depression based on the subject's responses. Initially, a questionnaire was created using SCID-5-CV and the diagnosis was exactly according to its guidelines. This baseline accepted only 'yes' (y) or 'no' (n) answers. The image below shows one such question-and-answer pathway based on sample user response.

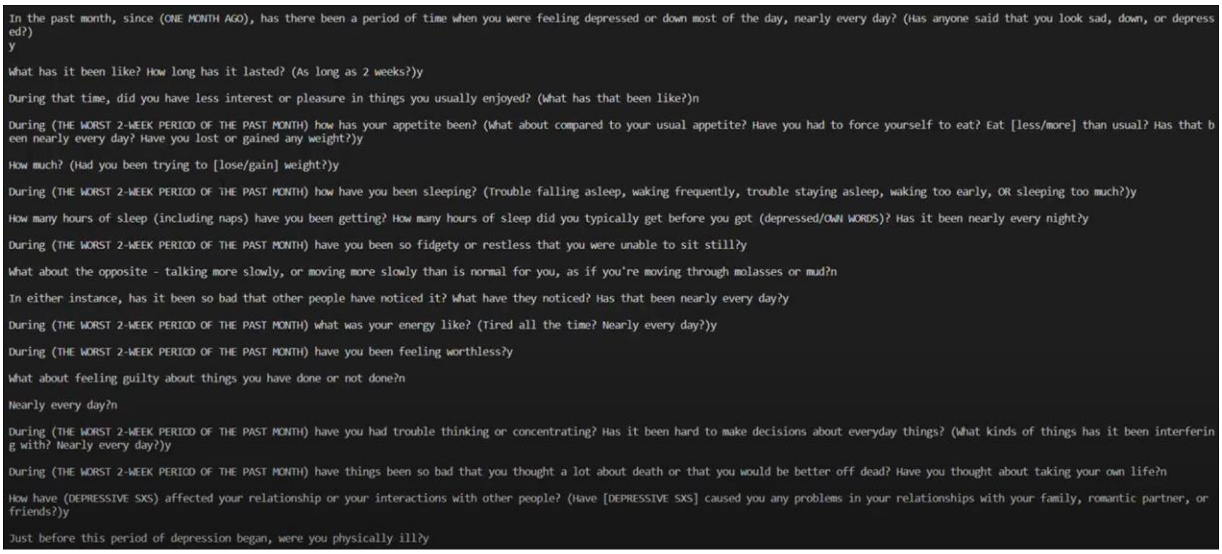

Figure 2: Example SCID-5-CV Baseline Question-Answer Pathway

Here, there are 12 nodes namely A1 to A12 denoting the 12 nodes of questions leading to diagnosis of MDD according to SCID-5-CV. Each node contains a set of ordered questions, answers to which either lead to '+' or '–' correspondingly signifying either presence or absence of that particular symptom pertaining to that node for a 2-week period in the past month for most of the day, nearly every day. Thereafter, according to SCID-5-CV guidelines, a person is either diagnosed with MDD or substance-induced depression or depression due to AMC or not depressed.

This tool was then used as a webapp to collect responses by users. The link to the webapp can be found here.

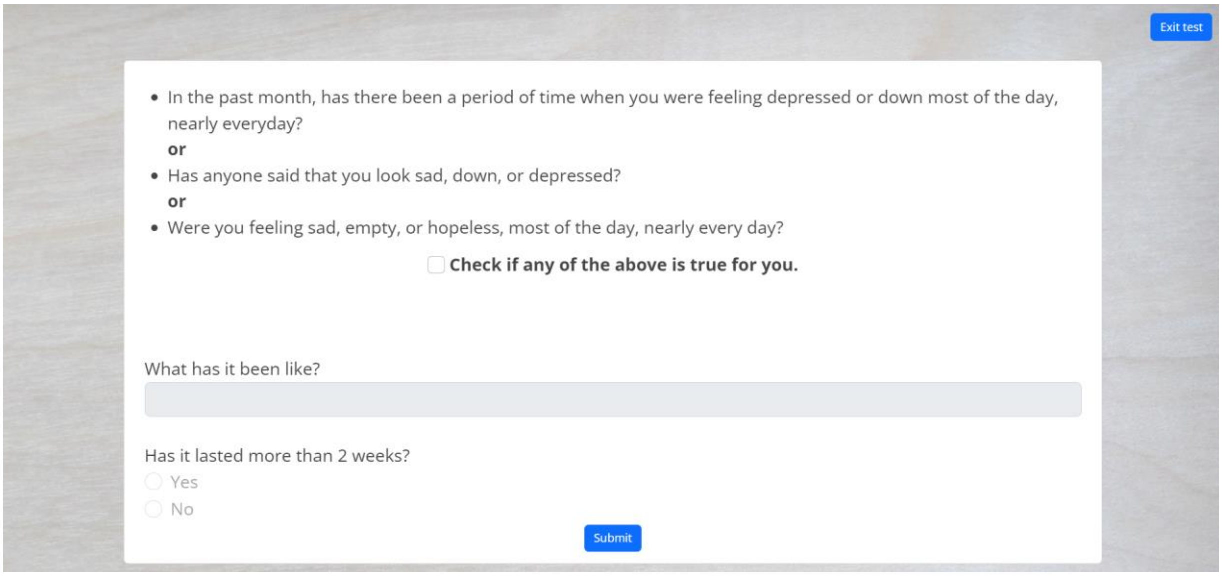

Figure 3: Sample Question Asked in Webapp

In the image above, it is clear how the questions were framed in order pertaining to each node so that each 'yes' or 'no' answer suffices to determine their contribution in diagnosis of depression. Additionally, the question – "What has it been like?" – has a text box field for its response. This is to incorporate future textual interactions as the scope of the model. The user can also exit the test at any point of time. The duration for each question and the answers in order are stored on the backend along with relevant details of each user. All this information is stored with consent from the user at the start of the test.

4.2 Making the baseline adaptive to newer contexts and rules

At first, the pool of questions was increased by creating relevant questions pertaining to personified contexts and Indian demographics vetted by professional clinicians. The pool of questions was further increased by adding questions pertaining to different diagnostic criteria and questionnaires namely PHQ-8, BDI, DASS-21 and Hamilton DRS.

The goal is to allow the clinicians, using this tool to diagnose patients, flexibility to add or remove questions as they deem fit. They can also predetermine the order of these questions, the diagnostic criteria they want the tool to use to detect depression and the classes of outcome they want as diagnosis.

DSM-5 is used as the default diagnostic criteria and SCID-5-CV is used as the default question pool and order. While SCID-5-CV has – not diagnosed with MDD, diagnosed with MDD, substance-induced depression, and depression due to AMC as its diagnosis classes, BDI has – not depressed, mild-moderate depression, moderate-severe depression, and severe depression as its classes. The clinician can predetermine which questionnaire to follow, and correspondingly which diagnostic classes are to be followed by the tool.

The output categories are stored in an array and can be modified based on clinician's choice. Nodes A1 to A12 in SCID-5-CV can be generalized to categories like lack of sleep, loss of appetite, lack of pleasure/interest and so on. This entire questionnaire is represented using a tree structure where each node class contains root to a question tree in it and a pointer to the left and right child where the left child corresponds to the '+' instances of the node and the right child corresponds to the '–' instances of the node. Each node in the question tree per questionnaire tree node again contains the specific question and a pointer to the left and right child where the left child corresponds to the 'yes' instances of the question and the right child corresponds to the 'no' instances of the question.

Using the tree structure enables easy insertion and deletion of nodes/questions. If the clinician chooses to create the entire structure from scratch, he can do so too. To dynamically create the questionnaire structure, the webapp will later open the pool of questions for the clinician to choose from or to type in a newly formed question which will also be added to the pool. After that, which node/question to go to from that point can again be clicked pictorially on the webapp. For now, this entire structure and its diagnosis follows SCID-5-CV by default.

Finally, the conservation of the tree structure directly follows from the assumption that all issues related to MDD can be grouped under one of the several major topics as used in existing diagnostic criteria. If the tree structure doesn't change even if newer diagnostic criteria come at a later point, this tool can adapt itself to newer rules as required by that criterion.

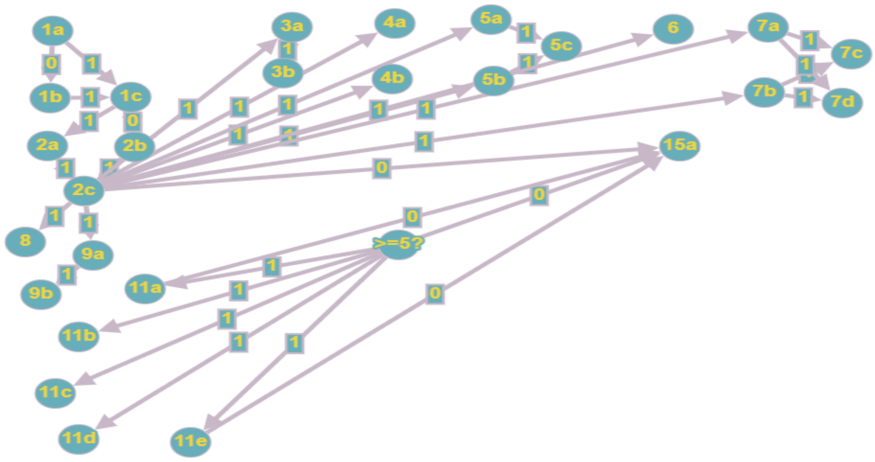

Figure 4: A Part of the Questionnaire Graph Based on SCID-5-CV

At first a graph is created using number of nodes and edges and nodes who have edges between them as inputs. Then for each node a set of questions is taken input. Thereafter, the rules by which count is increased for each node or an output is generated is taken input. Following this, the patient's can use this tool to diagnose themselves for depression. By traversing the graph using DFS, the questions are printed one by one and the answers input by the users are stored. Once the questions in all nodes are exhausted, using a string matching between the responses with the rules defined by the clinician, the output is generated.

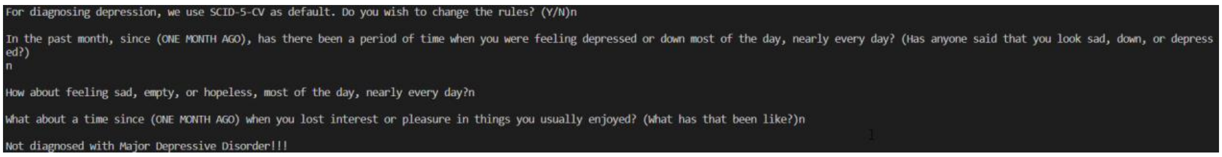

Figure 5: Generalization Tool with SCID-5-CV as Default

The output in Figure 5 above shows that the generalization tool uses SCID-5-CV as the default questionnaire and DSM-5 as the default diagnostic criteria. Further the user can change the output categories from SCID-5-CV, although they can use the same as default. The user can define rules for questions in each node giving a direct pathway through which count can be increased for that node or the pathway leads to a output category. By default, it is assumed that if count exceeds 5, then the output points to the last output category. Figure 6 below shows such a demonstration.

```
For diagnosing depression, we use SCID-5-CV as default. Do you wish to change the rules? (Y/N)y
Do you wish to change the output categories? (Y/N)n
Enter the number of nodes and edges:3 2
Enter the node numbers (starting from 1) which are connected in groups of two separated by a new line. E.g.- 1 2
2 3
1 2 2 3
Enter number of questions for 1th node:
1
Please enter the questions
a
Enter number of questions for 2th node:
2
Please enter the questions
b
c
Enter number of questions for 3th node:
2
Please enter the questions
d
e

Now for each node (newline separated) input the question number along with the corresponding response that increments count for that node or add an extra index to point to the output category, as neede
d.
E.g.- 1121312
105161
111
11212
10203
Please answer the following questions with either 0 or 1.
a
0
b
1
c
1
d
0
e
0
Diagnose: Substance-Induced Depressive Disorder!!!
```

Figure 6: Demonstration of the Working Generalization Tool

The code below shows the DFS implementation where the responses for each question in a node is stored. For different nodes, responses are stored in different locations for ease of string matching with the rules later.

```
void dfs(int a,int p) {
    vector<int> temp;
   for(string i:q[a]){
    cout<<i<<"\n";
    int t;
    cin>>t;
    temp.push_back(t);
   }
   response.push_back(temp);
   temp.clear();
   for(int i:g[a])if(i!=p)dfs(i,a);
}
```

String matching was used with the rules input by the clinician and the responses given to the questions for each node and correspondingly either an output was generated, or count was increased for the node.

```
if(response[k][(rules[k][t]-'0')-1]!=(rules[k][t+1])-'0')
            {
                flag = 1;
                break;
            }
...
if(flag==0)
            {
                cout<<endl<<out[(rules[k][j-1]-'0')];
                return 0;
            }
            else
                cnt++;
```

Note: A problem with the above implementation is that it doesn't handle the case where combination of responses to questions in more than one node leads to a specific output category. Future versions of the above implementation should handle the above case.

4.3 Handling full blown textual interactions and responding on the go

For now, both the baseline and generalization tool only accepts binary answer to each of the questions. The next goal is to adapt this tool to full blown textual interactions. The webapp has been built such that it can accept textual responses in its input field but for now the results of the diagnosis don't depend on it. To analyze textual responses and diagnose depression based on how the conversation flows requires the tool to respond on the go and zero-shot learn the diagnosis of depression. This work is yet to be implemented.

4.4 Analysis of responses at the backend

Analyses of types of responses, duration at each kind of question, at which question subject's mostly leave the platform without answering and anything else that might help clinicians will be done using the data collected at the backend. Further experimentation is yet to be done to collect enough data at the backend for analysis purposes.

Chapter 5

Future Work

5.1 Creating a baseline questionnaire structure

Most of the work is done in this section. After a few final touches the webapp will be ready to launch and we can have our baseline tool ready for automated depression detection albeit with the restriction of answers being only ‘yes’ or ‘no’.

Further Scope – For the baseline model, the questions were directly replicated from SCID-5-CV. For later versions, the questions can be modified to be user friendly while keeping their diagnostic purpose the same.

5.2 Making the baseline adaptive to newer contexts and rules

Most of the work is done in this section. The webapp version of the tool is yet to be implemented. Some more flexibility and ease of use features might be implemented later.

Further Scope – For now the implementation design is such that either the clinician can choose the default or create the entire questionnaire structure from scratch. At a later stage, it might be beneficial to facilitate remodeling the structure from any point and providing suggestions or narrowed options for the clinicians to choose from. Moreover, classifying a set of questions into symptom nodes like lack of sleep is done manually as of now. For questionnaires like DASS-21, they don’t essentially fall into a similar structure and also owing to the increase in the pool size, automating this process would also make more sense.

5.3 Handling full blown textual interactions and responding on the go

For now, the baseline model only accepts ‘yes’ or ‘no’ as an answer to each of the questions. The methodology to adapt this tool to full blown textual interactions is proposed as follows. To detect ‘+’ or ‘–‘ instance for a symptom, similar rule based methodologies can be followed based on the diagnostic criteria. Each response can be categorized as positive or negative which would correspondingly mean a ‘yes’ or ‘no’ instance for that particular question. This is sufficient to deterministically diagnose depression based on selected diagnostic criteria.

To make the tool adaptive to responding on the go NLU can be used. Since questions are grouped under categories like lack of sleep and loss of appetite, responses can also be classified into intents like lack of sleep and loss of appetite. This requires use of NLP to train the model under specific data with proper contexts and demographics. On intent classification the model can next chose which question to ask or what to respond based on NLU. But in this case, since the number of questions or set of responses is no longer fixed, the model needs to be further trained to correctly diagnose depression as the model determinism is lost. Experimenting this work by merging NLP with our novel rule-based approach remains to be done in the future. It should be noted that there will be further tradeoff between making the model more robust and generalizable to newer contexts and zero-shot detection of depression.

5.4 Analysis of responses at the backend

This work is yet to be started as the webapp hasn't been launched yet. Once data is collected at the backend, we aim to do all sorts of analysis with it. The goal is to use machine learning based predictive analysis and data analysis to find anomalies and patterns which might reveal additional relevant information about patient mental health. Also, simple statistical analysis like duration of responding each question, at which question user's mostly exit the platform, most common answer orders pertaining to certain age groups, most common words used by people with severe depression and similar can be performed. The analysis mainly aims to aid clinicians in their diagnosis of depression.

Conclusion

In this work, we present a novel solution for clinicians to automatically detect depression in patients with varying degrees of flexibility. We aim to combine NLP and rule-based approaches based on diagnostic criteria to interact with users and diagnose depression and/or levels of severity of it. We further hope that our statistical and machine learning driven analysis of data collected at the backend will aid clinicians understand and diagnose depression better.

The tool is yet to be launched in practice, so we lack experimental results. But this work was done closely with mental health professionals and is well vetted. We hope that, once launched, this tool will likely aid clinicians in detecting depression as well as help reduce the number of depressed people across the world by diagnosing depression at a larger rate than ever before.

Srijanak De

References

[1] Gratch et al. (May, 2014). "*The Distress Analysis Interview Corpus of human and computer interviews*". In: Proceedings of Language Resources and Evaluation Conference (LREC), 2014, pp. 3123–3128. url: http://www.lrec-conf.org/proceedings/lrec2014/pdf/508_Paper.pdf.

[2] Syed Arbaaz Qureshi, Mohammed Hasanuzzaman, Sriparna Saha, Gaël Dias (April 2019). "*The Verbal and Non Verbal Signals of Depression -- Combining Acoustics, Text and Visuals for Estimating Depression Level*". doi: 10.48550/arXiv.1904.07656. url: https://arxiv.org/pdf/1904.07656.pdf.

[3] Heinrich Dinkel, Mengyue Wu, Kai Yu (April, 2019). "*Text-based depression detection on sparse data*". doi: 10.48550/arXiv.1904.0515476. url: https://arxiv.org/pdf/1904.05154.pdf.

[4] D. Xezonaki, G. Paraskevopoulos, A. Potamianos, S. Narayanan (June 2019). "*Affective Conditioning on Hierarchical Networks applied to Depression Detection from Transcribed Clinical Interviews*". doi: 10.48850/arXiv.2006.08336. url: https://arxiv.org/pdf/2006.08336.pdf.

[5] American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

[6] First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. (2016). *User's guide for the SCID-5-CV Structured Clinical Interview for DSM-5® disorders: Clinical version* (pp 10). American Psychiatric Publishing, Inc.