Seminar Report

On

Automated Detection of Major Depressive Disorder (MDD) through Question-Answering

Srijanak De (19CS30047)

Guided By: Prof. Partha Pratim Chakrabarti

Jointly Guided by: Prof. Aritra Hazra, Prof. Rajlakshmi Guha

Department of Computer Science and Engineering

Indian Institute of Technology, Kharagpur

Spring Semester, 2021-22

April 6, 2022

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
DAIC	**D**istress **A**nalysis **I**nterview **C**orp
WoZ	**W**izard-**o**f-**Z**
DAIC-WOZ	**D**istress **A**nalysis **I**nterview **C**orp **W**izard-**o**f-**Z**
PHQ-8	**P**atient **H**ealth **Q**uestionnaire - **8**
APA	**A**merican **P**sychiatric **A**ssociation
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

Chapter 1

Major Depressive Disorder (MDD)

1.1 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)

1.2 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 symptoms is 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 represent 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)

1.3 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, nearly every day? (Has anyone said that you look sad, down, or depressed?)** IF NO: **How about feeling sad, empty, or hopeless, most of the day, nearly every day?** 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: **Has it been nearly every day? 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? Has that been nearly every day? 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)

Chapter 2

Literature

2.1 Problem Statements

Auto-structuring of questions that fit the diagnostic criteria (DSM-5)

SCID-5-CV gives a proper structuring of questions to be asked to the subject based on DSM-5. But there is a lot of scope of varying the question order while keeping the structure itself intact. Re-ordering of questions making use of the non-rigid parts of the SCID-5-CV structure although fits the DSM-5 criteria inherently, is yet untested in real world patient-clinician environment. Moreover, if more questions are added to the set of questions in SCID-5-CV, integrating them in the existing structure might turn out to be challenging. The task at hand is to automatically structure questions utilizing the semi-rigid structure of SCID-5-CV as the baseline.

Auto-detection of MDD

SCID-5-CV questionnaire gives a deterministic detection of MDD based on the DSM-5 criteria. As long as the structure and other heuristics pertaining to SCID-5-CV is maintained, there can be a deterministic detection of MDD. Since the pool of questions is quite low in SCID-5-CV, an extension on the baseline is inevitable. The practical applicability, scalability, accuracy, and interpretability of such a model is to be tested.

Auto-generation of questions dynamically as a reply to subject's response based on subject's interests, culture, and demographic variables

Currently, the SCID-5-CV questionnaire puts no focus on the subject's interests, culture, and demographic variables which play an important role in the psychological development of a person. Additionally, there is less to negligible context between consecutive questions when SCID-5-CV jumps from one DSM-5 criteria to another. The problem statement involves incorporating these features into the questionnaire as well as generating the questions dynamically at run time while maintaining the inherent spontaneity of the conversation.

2.2 Literature Survey

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.

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 is 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.

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 pair, 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.

Chapter 3

My Plans

3.1 Proposals

We plan to create a baseline questionnaire graph based on SCID-5-CV. Each node in the graph corresponds to a single set of questions from SCID-5-CV. A single branch pertains to a single DSM-5 criteria. The final leaf node in a branch depicts the satisfiability of the DSM-5 criteria that particular branch denotes. A demo representation of a part of the graph is shown in Figure 2.

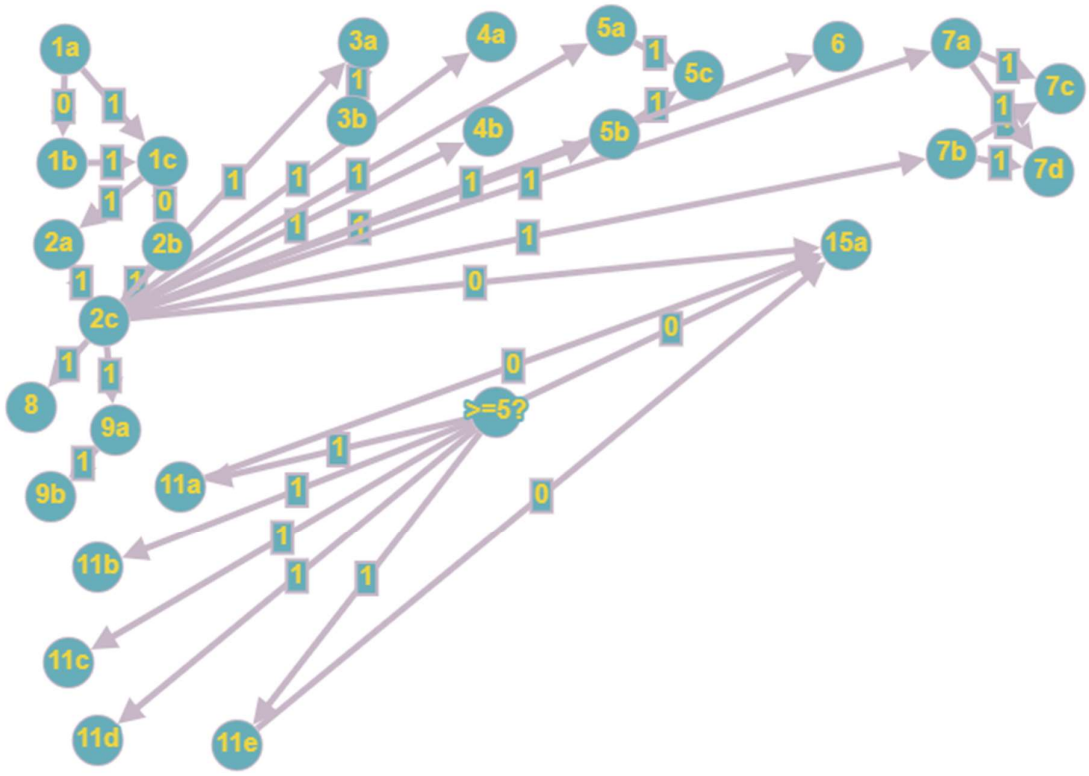

Figure 2: Baseline Questionnaire Graph Based on SCID-5-CV

1a, 1b and 1c, all three nodes carry questions pertaining to DSM-5 criteria 1. Edge weight 0 means that path should be traversed if answer to the set of questions in the previous node is negative and for 1 it is positive. Since there is no particular ordering to be chosen for questions pertaining to DSM-5 criteria 3 to 9, the order of the questions can be changed by traversing the graph through different adjacent nodes.

Subsequently, more clinically relevant questions will be added to the collection introducing personal, cultural, and demographic context to the questionnaire. We also plan to introduce the questionnaire in other languages like Bengali which fits the SCID-5-CV structure. Including this new set of questions inside the existing graph would be challenging. Other alternatives to it can be extending the semi-rigid SCID-5-CV structure so that it can incorporate addition of further questions while still adhering to the DSM-5 criteria.

Once we have sufficient number of questions in our questionnaire pool, the next step would be to auto select the follow-up question from the pool based on the subject's response. We also plan to develop dynamic generation of questions containing the context and continuity of the conversation much like an intelligent chatbot.

The final step would be to analyze the subject's textual response and detect the satisfiability of each DSM-5 criteria by using NLP and not just by determining the positivity or negativity of the response towards fulfilling the DSM-5 criteria. As an alternative to this we could also build an ML model which learns how the DSM-5 criteria detects if a person has MDD or not. This model can then predict the presence of MDD just by taking the whole conversation as an input. We also plan to incorporate the other two modalities (visual and acoustic) into our depression detection process.

References

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.

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.

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.

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.