AI4MH Tutorial

Aseem Srivastava, Neeraj, Yash Kumar Atri, Shivani Kumar, Md Shad Akhtar, Tanmoy Chakraborty

https://ai4mh.github.io/

Agenda

- Introduction
- Discussion on Resources
- NLP Applications for Mental Health
- Hands-on + interactive session
- NLP to Enhance Online Peer Therapy
- Current state of mental health x NLP
- Ethical Considerations
- Conclusion

Mental Health

Mental health is a state of mental well-being that enables people to cope with the stresses of life, realize their abilities, learn well and work well, and contribute to their community.

World Health Organization

Key Facts

- Strategies to promote, protect and restore mental health.
- Mark the actions on mental health as urgent.
- Mental health has intrinsic and instrumental value.

Reality of Mental Health Conditions

WIDESPREAD

1 in 8

live with a mental health condition

UNDERTREATED

71%

people with psychosis do not receive mental health services

UNDER-RESOURCED

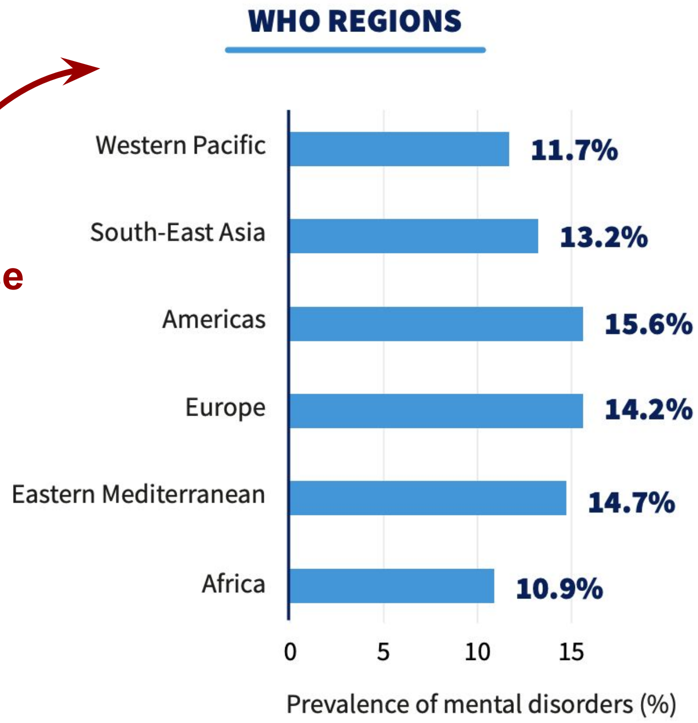
2%

of health budgets, on average, go to mental health

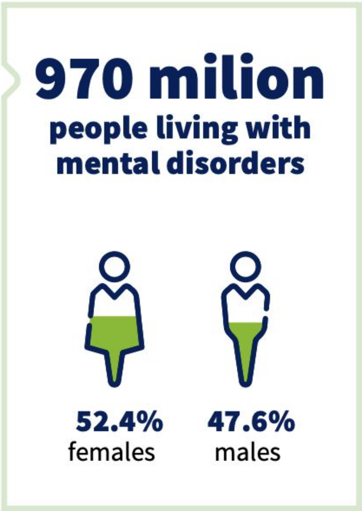
Source: IHME, 2019 (98); WHO, 2021 (5).

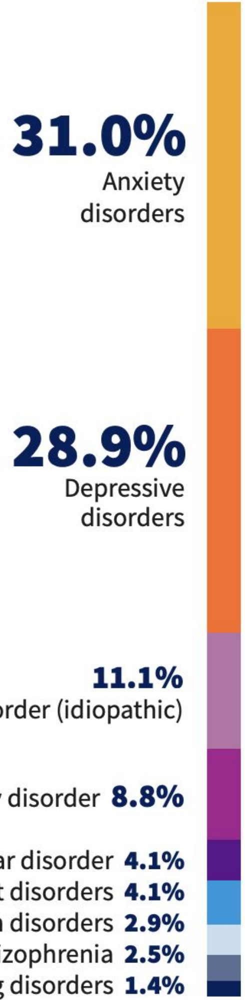
Reality of Mental Health Conditions

Region Wise



Gender vs Symptoms





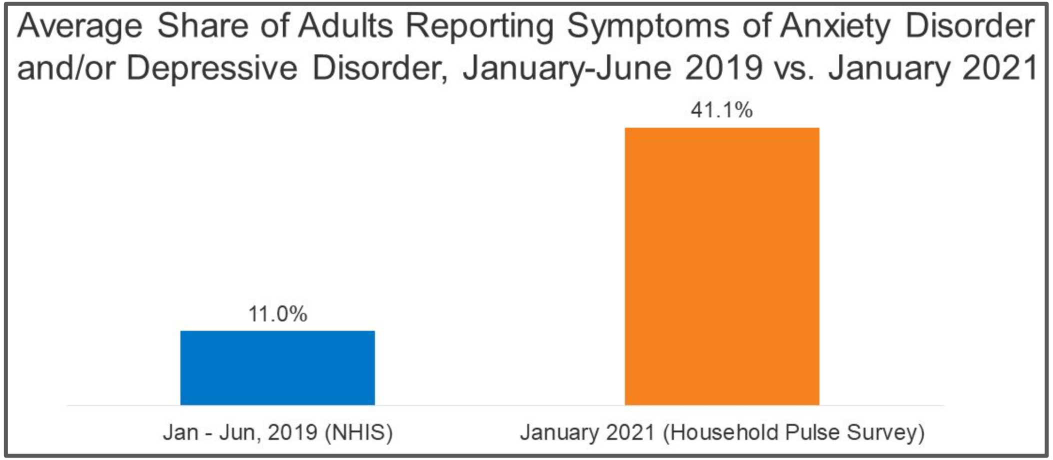
Source: IHME, 2019 (99)

Pandemic and Rising Mental Health Issues

“‘Nobody Has Openings’: Mental Health Providers Struggle to Meet Demand”

– New York Times

Post COVID Surge



https://www.nytimes.com/2021/02/17/well/mind/therapy-appointments-shortages-pandemic.html
https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/

India has 0.75 psychiatrists per 100,000 people. Can telepsychiatry bridge the gap between mental health experts & patients?

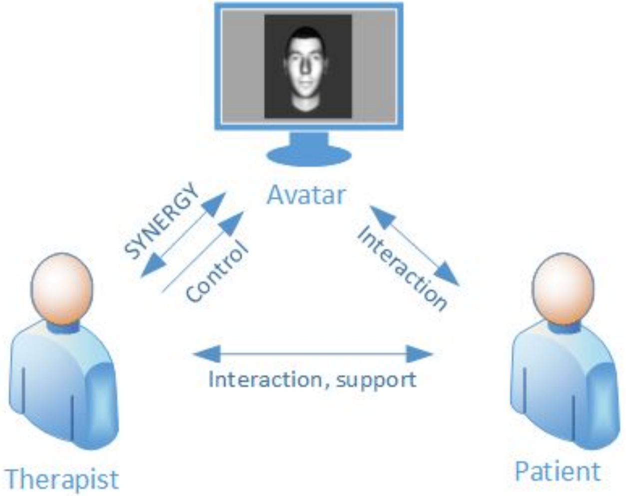
telepsychiatry

https://economictimes.indiatimes.com/magazines/panache/india-has-0-75-psychiatrists-per-100000-people-can-telepsychiatry-bridge-the-gap-between-mental-health-experts-patients/

Avatar Therapy

What Role Can Avatars Play in e-Mental Health Interventions? Exploring New Models of Client–Therapist Interaction

Imogen C. Rehm[1], Emily Foenander[1], Klaire Wallace[1], Jo-Anne M. Abbott[1], Michael Kyrios[2] and Neil Thomas[1,3]*



https://www.frontiersin.org/articles/10.3389/fpsyt.2016.00186/full

Avatar therapy: early trial results 'very encouraging'

A new Wellcome-funded study has shown that avatar therapy may help to reduce auditory hallucinations in people with schizophrenia when used alongside other treatments.

https://wellcome.org/news/avatar-therapy-early-trial-results-very-encouraging

Tech is working out

Manual

- Individual Counseling
- Peer Counseling

Automatic Tech Savvy

- Virtual Mental Health Assistants

Hybrid Tech Savvy

- Human-AI Collaboration

Automatic-cum-hybrid

Human-AI Collaboration | VMHAs

- Many new **virtual mental health assistants** emerged.

- Generally assistants focus on generating helpful responses and understanding patients.

ChatGPT ?

A one stop solution

VICE

VICE World News

ChatGPT Gave Me Advice on How To Join a Cartel and Smuggle Cocaine Into Europe

Here's what happened when VICE's Global Drugs Editor spent 12 hours speaking to OpenAI's chatbot about drugs.

Technology News / News-Analysis

AI BOT AS A THERAPIST: US MENTAL HEALTH PLATFORM USING CHATGPT IN COUNSELLING LEADS TO CONTROVERSY

Koko, a mental health platform used ChatGPT in counselling sessions with over 4,000 users, raising ethical concerns about using AI bots to treat mental health.

Source: News Media vice.com

We **don’t need** an **All-in-One package**

We need **modular** AI-based solution

Modular Approach?

- Specialized Dialog Systems
- Dialog Understanding
- Dialog Summarization
- Datasets

Early age chatbots: ELIZA

New age: Dialog GPT, etc.

Recent Development in Neural Dialogue Systems

Study of works related to

- Dialogue-acts
- Emotion Recognition
- Dialogue Topic

Summarization Systems

- Extractive & Abstractive
- *Standard Abstractive Summarization Approaches and metrics.*

Modular Approach?

- Specialized Dialog Systems
- Dialog Understanding
- Dialog Summarization
- **Datasets**

Dialog Understanding	Dialog Summarization	Mental Health Counseling
Switchboard Corpus	AMI Meeting	Crisis TextLine (Private)
MRDA Corpus	SAMSum	Talklife (Private)
Reddit Corpus	DialogSum	Reddit MH Dataset
	HET Data (Chinese)	
	D[illegible] Summarize Data	
	MIE Dialogue	

CONVERSATIONAL DATASETS

UNAVAILABLE

NEED

for quality counseling conversational

DATASET

-- Jeremy Ang, Yang Liu, and Elizabeth Shriberg. 2005.Automatic dialog act segmentation and classification in multiparty meetings. In 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '05, pages 1061–1064.
-- Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Marie Meteer, and Carol VanEss-Dykema. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech.Computational Linguistics, 26(3):339–371.
-- Jean Carletta, Simone Ashby, Sebastien Bourban, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, Vasilis Karaiskos, Wessel Kraaij, Melissa Kronenthal, Guillaume Lathoud, Mike Lincoln, Agnes Lisowska, Iain McCowan, Wilfried Post, Dennis Reidsma, and Pierre Wellner. 2006. The ami meeting corpus: A pre-announcement. In Machine Learning for Multimodal Interaction, pages 28–39, Berlin,
-- Yulong Chen, Yang Liu, Liang Chen, and Yue Zhang. 2021. DialogSum: A real-life scenario dialogue summarization dataset.
-- Yuanzhe Zhang, Zhongtao Jiang, Tao Zhang, Shiwan060Liu, Jiarun Cao, Kang Liu, Shengping Liu, and Jun Zhao. 2020. MIE: A medical information extractor towards medical dialogues. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 6460–6469, Online. Association for Computational Linguistics.

Current Literature

Data Domain	Purpose	Motivation	MH Target
Audio	Detecting Symptoms	Assist in the early diagnosis and longitudinal monitoring of mental illness symptoms in everyday speech conversation.	Depression
Accelerometer	Detecting symptoms	To accurately detect depression from very easy to obtain motor activity.	Depression
Body (skin conductance)	Detecting symptoms	To aid non-intrusive measures of PTSD symptom severity through skin conductance responses; reducing need for self-report.	PTSD
Audio (counselling session)	Improving treatment	To effectively assess therapist performance to aid their skills development and retention for better patient outcomes.	Substance abuse
Text	Understanding mental health content	To improve understanding of mental illnesses.	Multiple

Current Literature

Data Domain	Purpose	Motivation	MH Target
Audio	Detecting Symptoms	Assist in the early diagnosis and longitudinal monitoring of mental illness symptoms in everyday speech conversation.	Depression
Accelerometer	Detecting symptoms	To accurately detect depression from very easy to obtain motor activity.	Depression
Body (skin conductance)	Detecting symptoms	To aid non-intrusive measures of PTSD symptom severity through skin conductance responses; reducing need for self-report.	PTSD
Audio (counselling session)	Improving treatment	To effectively assess therapist performance to aid their skills development and retention for better patient outcomes.	Substance abuse
Text	Understanding mental health content	To improve understanding of mental illnesses.	Multiple

NLP

Current state-of-the-art **Datasets**

NLP X Mental Health

Dataset	Size	Description
Counseling Conversation [10]	80,885 dialogues, out of which 15,555 were used for analyses	SMS text-based collected from non-profit organization for handling crisis situations such as depression, suicidal thoughts etc. This dataset is not publicly available for research.
Counseling Conversation [17]	259 conversations	Motivational Interviewing collected from video titles of Youtube and Vimeo on topics such as smoking cessation, quit drinking etc.
CBT Dialogue [9]	882 dialogues	Online text-based system for conversations between patients (suffering from depression or anxiety) and their therapists
Online Synchronous Chat [15]	49 transcripts	Young Australian people suffering from mental health issues such as depression, anxiety, suicidal tendency, relationship problems etc.
Hope & Expectations Online counseling [14]	1033 online questionnaire	Supportive counseling, Self-help mechanisms and psychoeducation
DAIC-WOZ Dataset [18]	189 dialogues	Face-to-face counseling conversations between the interviewer and patient suffering from depression, anxiety etc.
Twitter Dataset [6]	7048 users with 21 million tweets	Identifying clinical depression of users from their tweets
Reddit Dataset [19]	1.1 million posts	Identifying mental health information on topics such as schizophrenia, anxiety, autism, self-harm, personality disorder etc.
UAH Dialogue [20]	100 dialogues	Identifying mental health of the speakers from conversations
Twitter Dataset [8]	50 million tweets from 4 lakh users	Gender and Cross-Cultural Differences from mental heath disclosures

Source: T. Saha, S. Chopra, S. Saha, P. Bhattacharyya and P. Kumar, "A Large-Scale Dataset for Motivational Dialogue System" 2021 International Joint Conference on Neural Networks (IJCNN)

Speaker and Time-aware Joint Contextual Learning for Dialogue-act Classification in Counseling Conversations

WSDM 2022

Ganeshan Malhotra[♦], Abdul Waheed[♠], Aseem Srivastava[❖], Md. Shad Akhtar[◆], Tanmoy Chakraborty[◆]

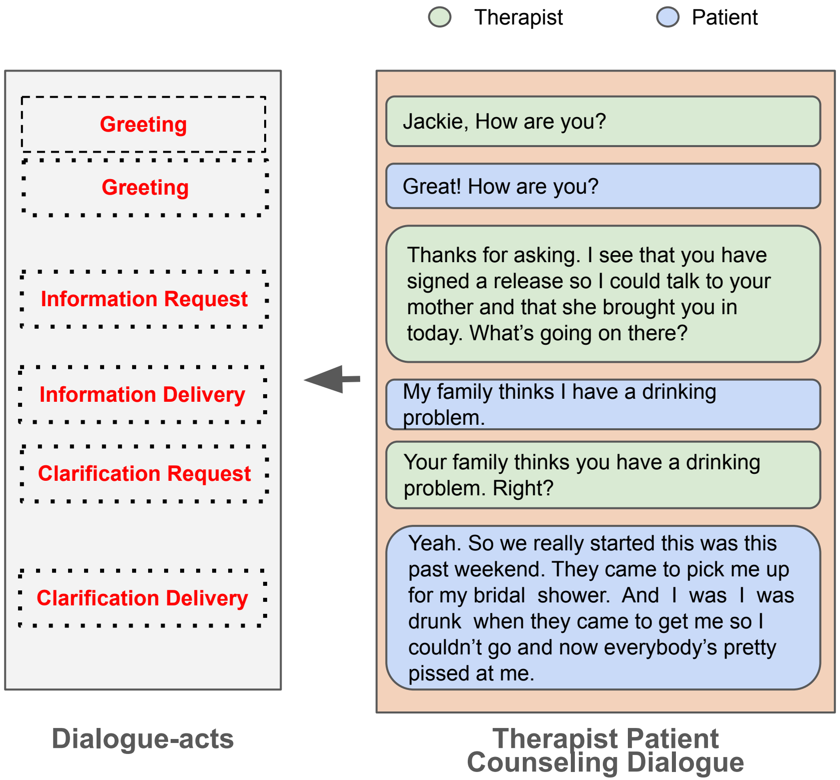
[♦] BITS Pilani [♠] MAIT Delhi [◆] IIIT - Delhi

HOPE Dataset

HOPE: Mental **H**ealth c**O**unselling of **P**ati**E**nts.

12.9k Utterances from 212 dialogue sessions with 12 dialogue act labels divided into 3 hierarchies designed carefully to cater to the specific needs of counselling session.

Sources: Youtube, Counseling Training Portals, Public Counseling Channels.



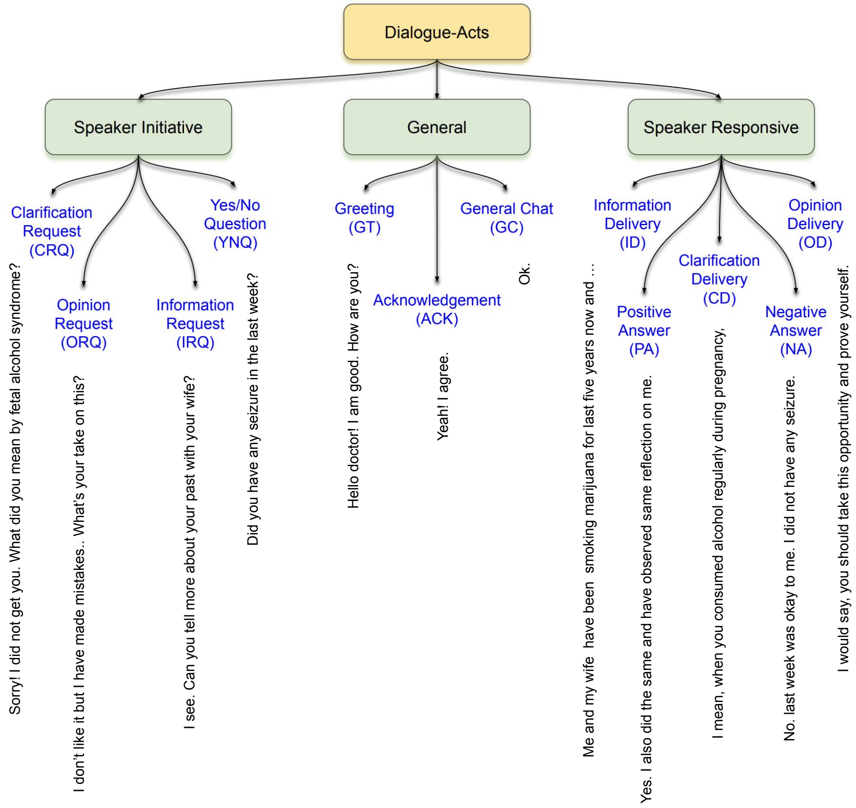
Speaker and Time-aware Joint Contextual Learning for Dialogue-act Classification in Counselling Conversations *In Proceedings of the* WSDM '22.

HOPE Dataset

HOPE: Mental **H**ealth c**O**unselling of **P**ati**E**nts.

12.9k Utterances from 212 dialogue sessions with 12 dialogue act labels divided into
3 hierarchies designed carefully to cater to the specific needs of counselling session.

Sources: Youtube, Counseling Training Portals, Public Counseling Channels.

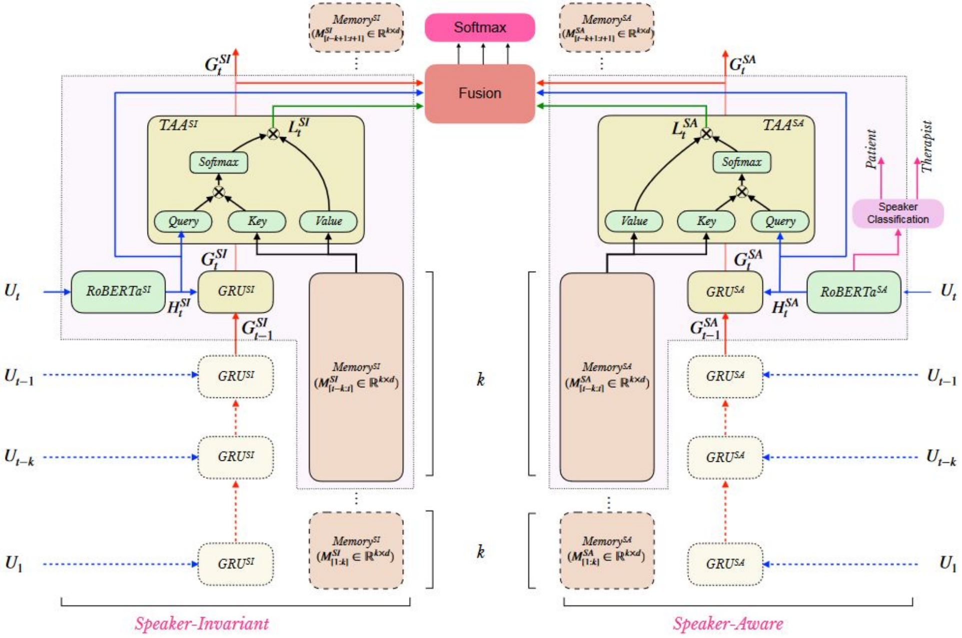


Speaker and Time-aware Joint Contextual Learning for Dialogue-act Classification in Counselling Conversations *In Proceedings of the* WSDM '22.

Task: Dialogue Act Classification

Understanding the importance of

- Speaker knowledge
- Contextual Knowledge



Speaker and Time-aware Joint Contextual Learning for Dialogue-act Classification in Counselling Conversations *In Proceedings of the* WSDM '22.

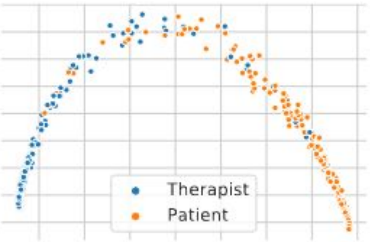
Results

Model	Type of Modelling	Precision		Recall		F1		Accuracy
		Macro	Weighted	Macro	Weighted	Macro	Weighted	
TextVDCNN [8]	U^t	11.01	21.02	19.53	38.53	13.37	36.81	41.77
ProSeqo [15]	U^t	9.77	17.60	11.20	27.90	7.11	14.29	27.35
RoBERTa [23]	U^t	51.01	58.12	47.14	52.97	43.97	49.13	52.97
TextRNN [22]	U^t + GC	30.27	37.92	27.9	41.76	25.55	36.81	41.77
DRNN [42]	U^t + GC	28.39	36.72	31.87	44.32	28.12	37.82	44.32
CASA [34]	U^t + GC	59.78	62.56	51.22	58.46	51.65	55.95	58.46
SPARTA-BS	U^t + GC	58.94	62.31	52.02	57.70	51.83	54.98	57.70
SA-CRF [39]	U^t + LC +SA	33.30	38.97	26.18	45.07	35.97	24.20	45.07
SPARTA-BS	U^t + GC + SA	58.87	63.02	53.28	58.41	52.22	55.57	58.41
SPARTA-MHA (3-fold CV)	U^t + LC + GC + SA	69.60	71.77	59.45	62.67	59.00	62.12	62.67
SPARTA-TAA (3-fold CV)		71.01	72.36	60.49	63.82	60.74	63.38	63.82
SPARTA-MHA	U^t + LC + GC + SA	60.24	66.53	59.64	63.45	58.16	63.26	63.45
SPARTA-TAA		62.15	67.36	61.13	64.75	60.29†	64.53†	64.75†
Significance T-test† (p-value)						0.009	0.014	0.048

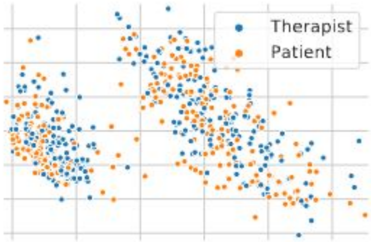
More than **6%** relative improvement

marginal yet effective

Result Analysis



Speaker Aware Utterance Representations



Speaker Invariant Utterance Representations

	ack	cd	crq	gc	gt	id	irq	od	na	pa	orq	ynq
ack	0.52	0.00	0.02	0.02	0.05	0.28	0.05	0.00	0.00	0.05	0.01	0.00
cd	0.16	0.48	0.01	0.01	0.00	0.22	0.00	0.00	0.04	0.03	0.02	0.01
crq	0.00	0.00	0.50	0.03	0.01	0.14	0.16	0.03	0.00	0.01	0.01	0.11
gc	0.05	0.00	0.07	0.54	0.04	0.17	0.07	0.01	0.01	0.02	0.00	0.02
gt	0.12	0.00	0.01	0.00	0.66	0.13	0.01	0.01	0.00	0.01	0.04	0.00
id	0.02	0.01	0.05	0.01	0.01	0.81	0.02	0.00	0.04	0.00	0.01	0.01
irq	0.00	0.01	0.07	0.01	0.01	0.03	0.78	0.02	0.00	0.02	0.02	0.04
od	0.02	0.00	0.04	0.04	0.00	0.43	0.02	0.45	0.00	0.02	0.00	0.00
na	0.00	0.00	0.03	0.02	0.02	0.20	0.02	0.00	0.71	0.00	0.02	0.00
pa	0.22	0.00	0.00	0.01	0.00	0.19	0.01	0.01	0.00	0.54	0.01	0.00
orq	0.00	0.00	0.03	0.00	0.00	0.03	0.15	0.00	0.00	0.03	0.76	0.00
ynq	0.00	0.01	0.11	0.00	0.01	0.03	0.26	0.01	0.00	0.00	0.01	0.58

The better we understand, the more likely we are to create something of significance.

- Simon Sinek

An essential component is **understanding directives** of utterances

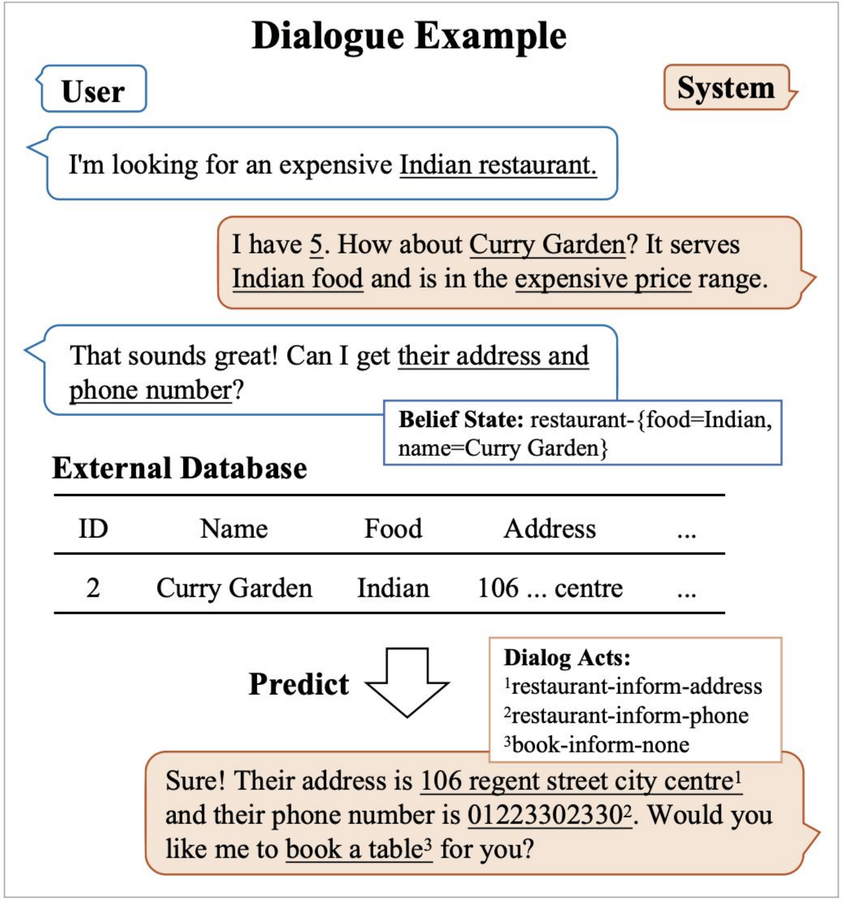
Dialogue Act: It deals with understanding the intended requirements of the utterances.

This essentially act as one of the precursors for the dialogue response generation.

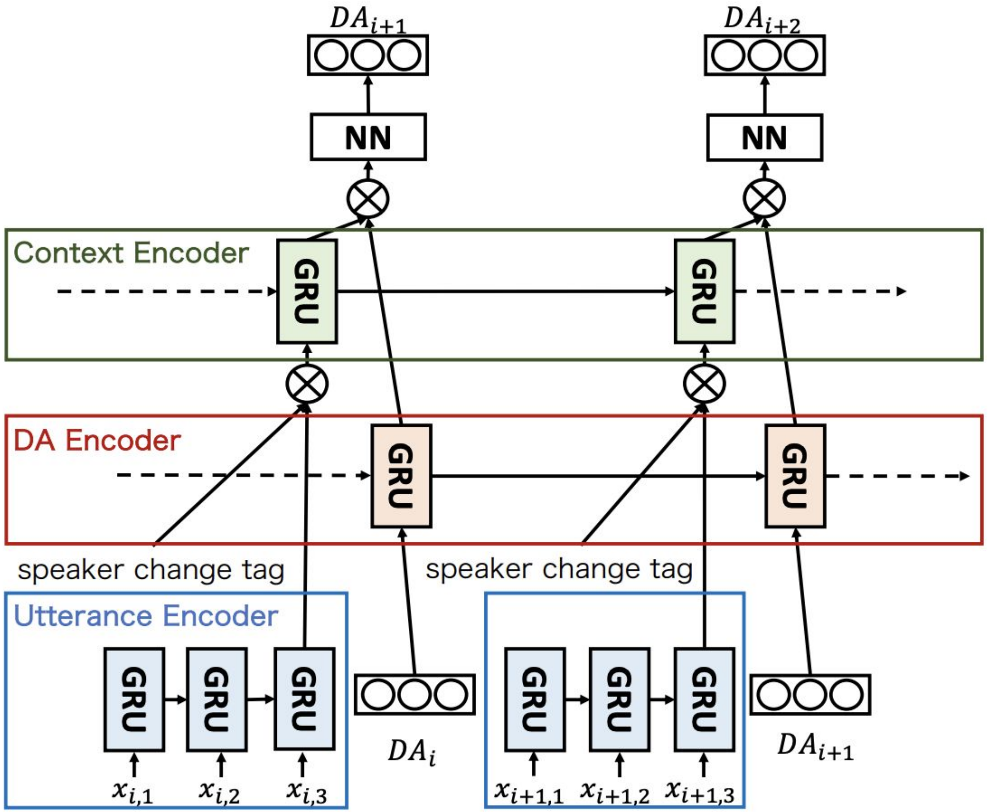
An essential component is **understanding directives** of utterances

Dialogue Act: It deals with understanding the intended requirements of the utterances.

This essentially act as one of the precursors for the Dialogue Response Generation



Multi-Domain Dialogue Acts and Response Co-Generation (Wang et al., ACL 2020)



Dialogue-Act Prediction of Future Responses Based on Conversation History (Tanaka et al., ACL 2019)

Response-act Guided Reinforced Dialogue Generation for Mental Health Counseling

WWW 2023

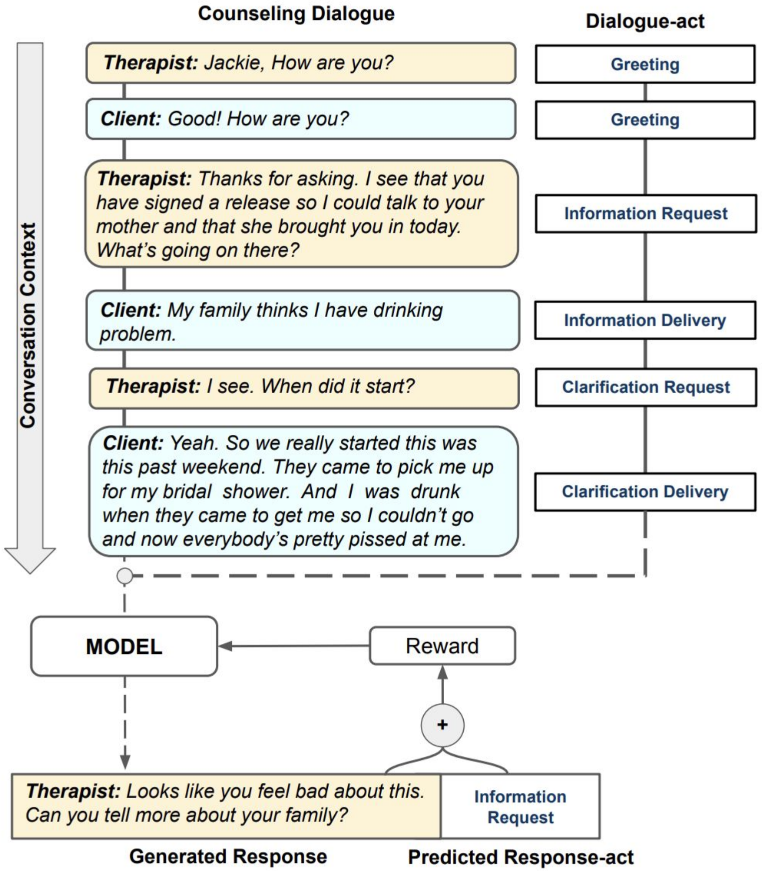
Aseem Srivastava[♦], Ishan Pandey[♦], Md. Shad Akhtar[♦], Tanmoy Chakraborty[♠]

[♦] IIIT - Delhi [♠] IIT Delhi

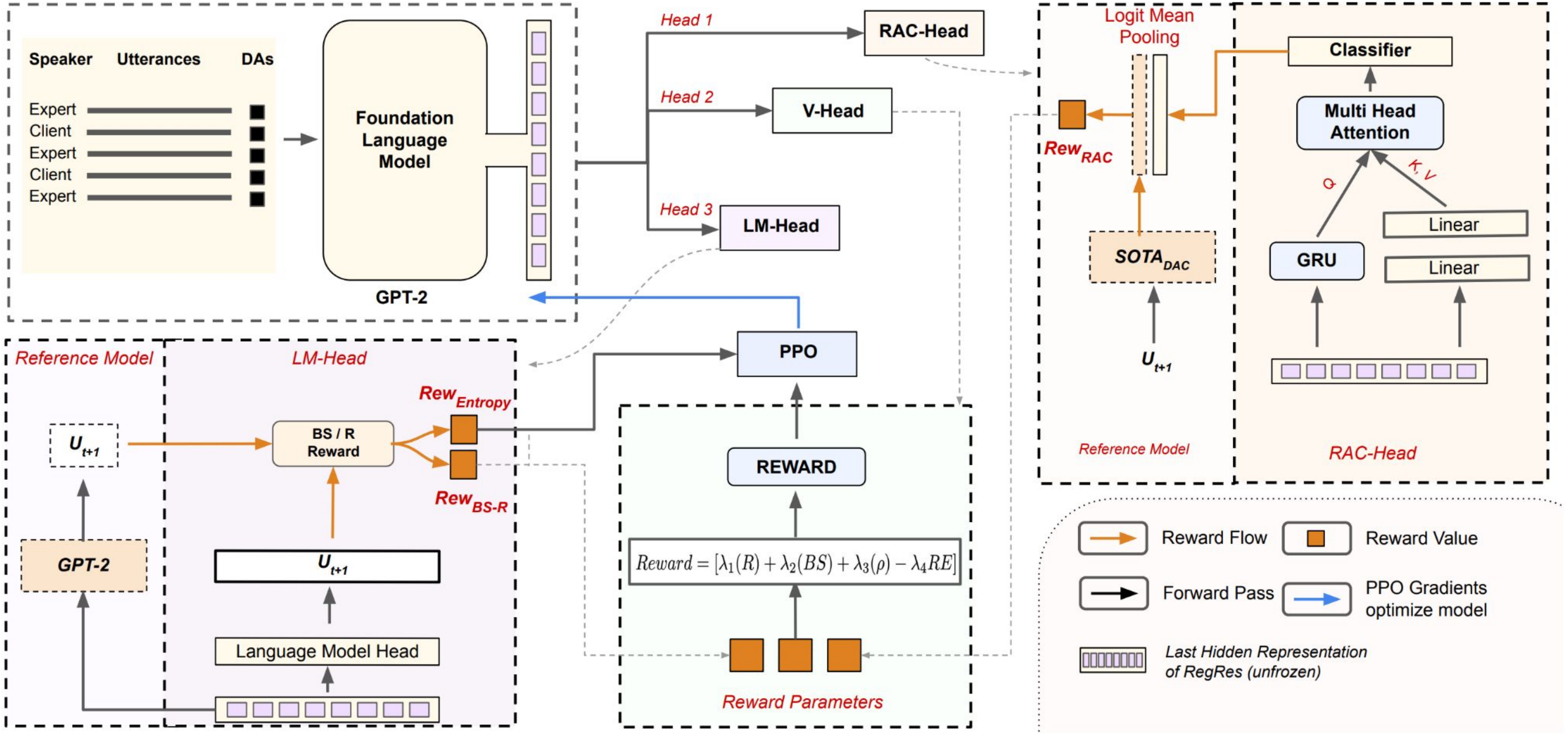
The Task: Response Generation

- A simple response generation task.

- Exploit future dialogue-acts (response-acts) in guiding the RL-model to generate the intended response and maintain the flow of counselling conversation.



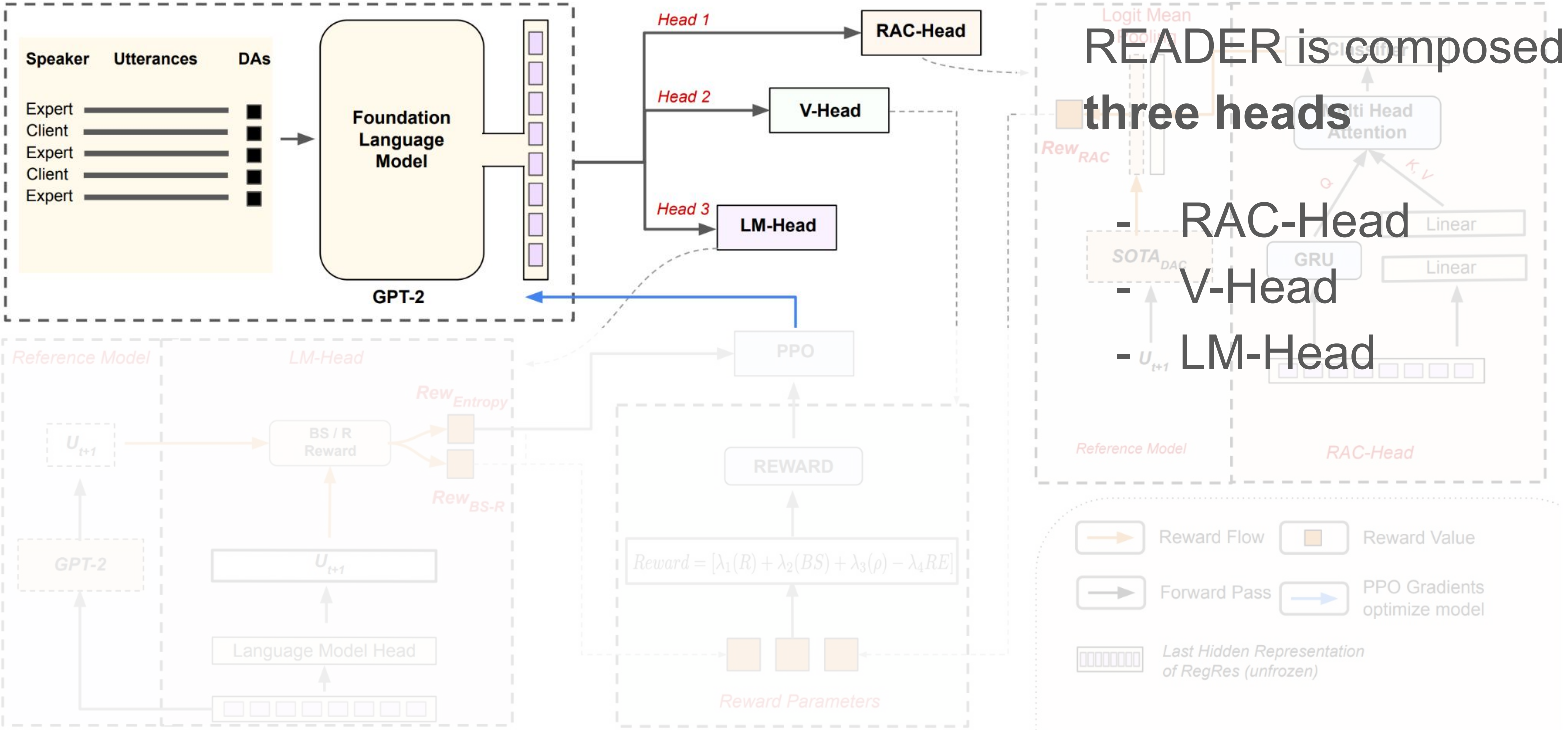
Proposed Model: **READER**



Response-act Guided Reinforced Dialogue Generation for Mental Health Counseling *of* WWW '23

*READER: **RE**sponse-**A**ct guided reinforced **D**ialogue gen**ER**ation model*

Proposed Model: **READER**



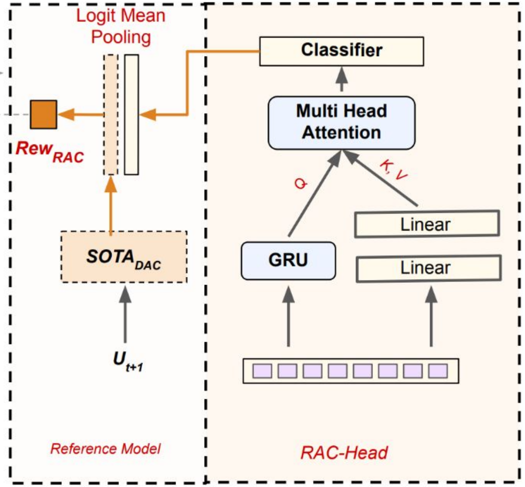
READER is composed of **three heads**

- RAC-Head
- V-Head
- LM-Head

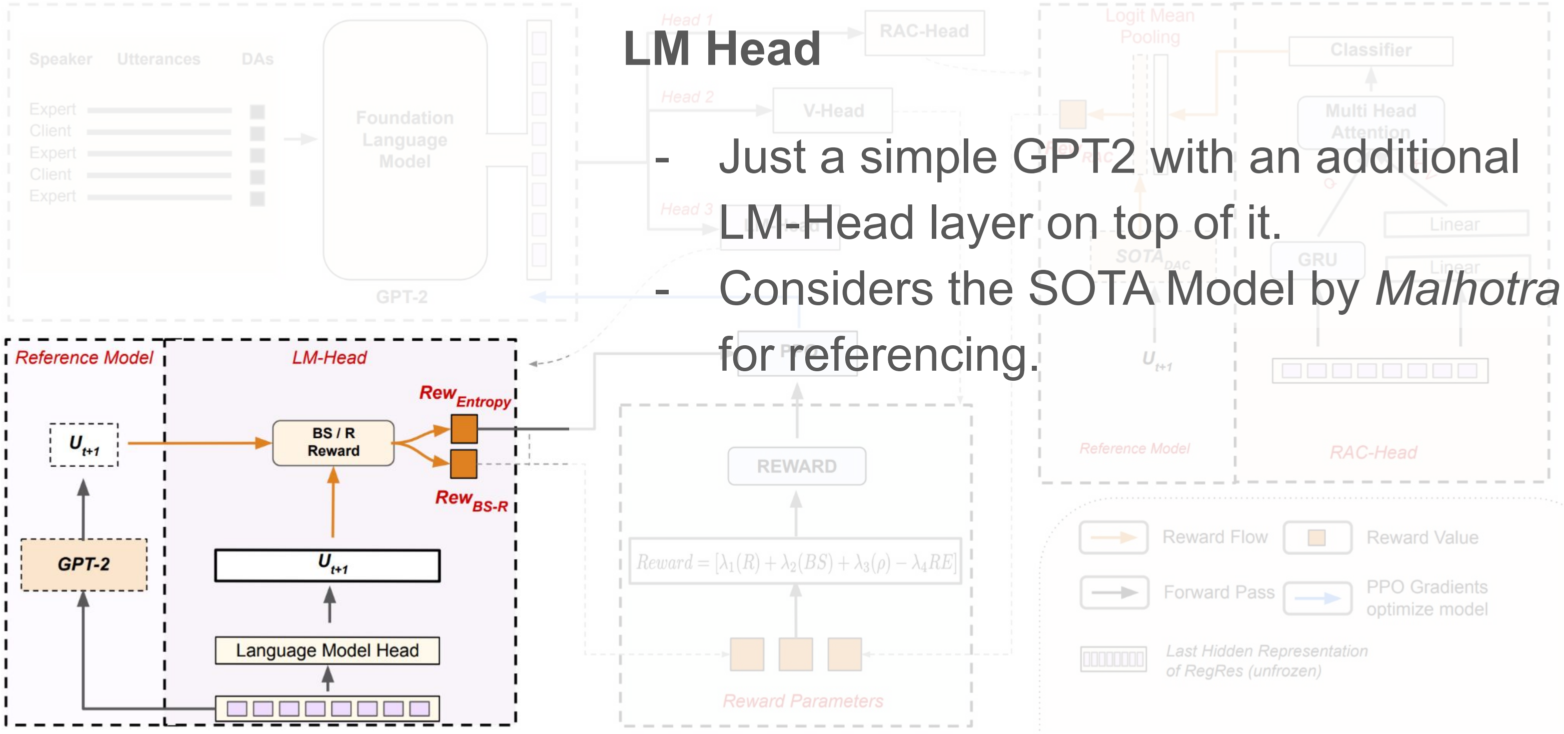
Proposed Model: **READER**

RAC Head

- Responsible for Response-act classification
- Considers the SOTA Model by *Malhotra et al.* for referencing.



Proposed Model: **READER**



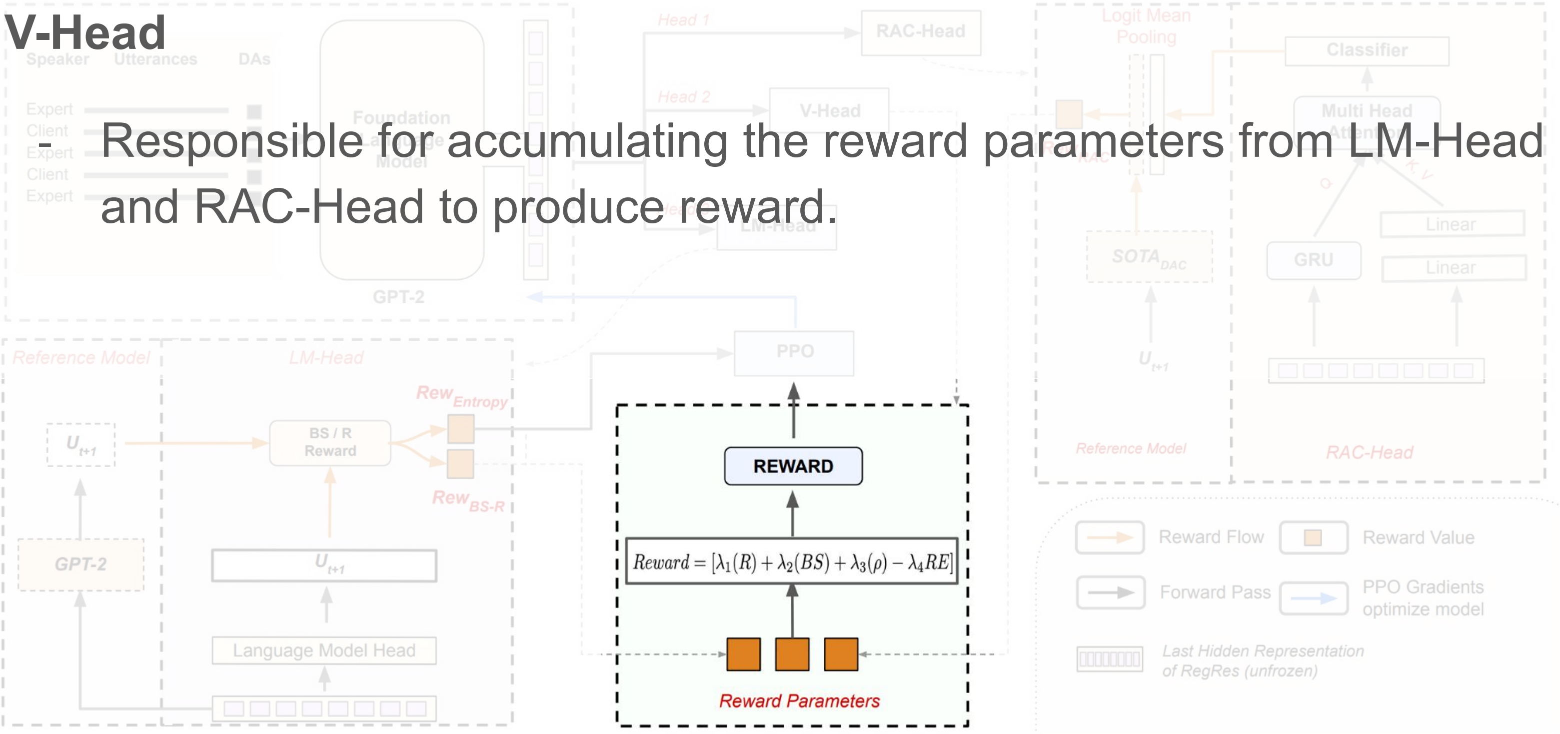
LM Head

- Just a simple GPT2 with an additional LM-Head layer on top of it.
- Considers the SOTA Model by *Malhotra et al.* for referencing.

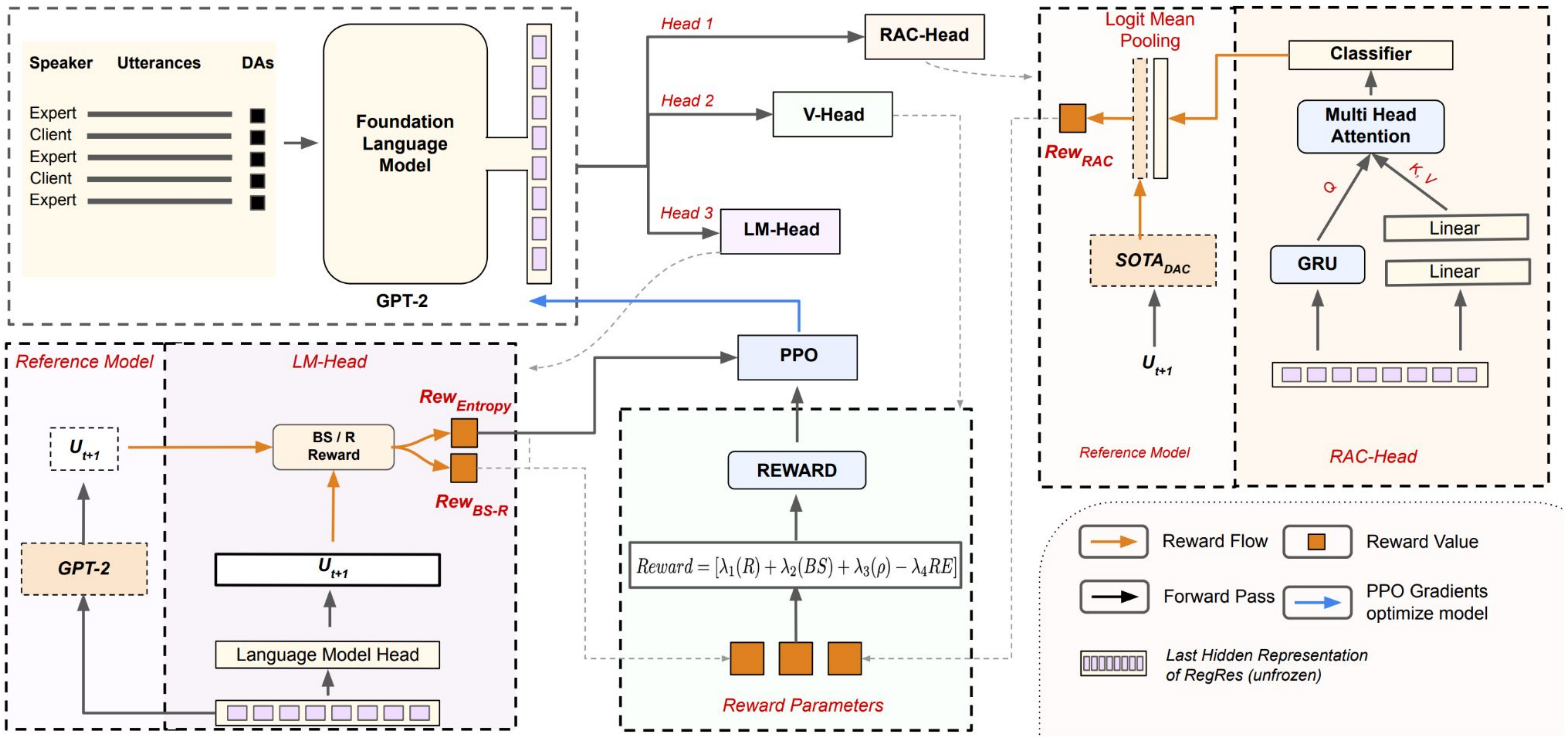
Proposed Model: **READER**

V-Head

- Responsible for accumulating the reward parameters from LM-Head and RAC-Head to produce reward.



Proposed Model: **READER**



Response-act Guided Reinforced Dialogue Generation for Mental Health Counseling *of* WWW '23

*READER: **RE**sponse-**A**ct guided reinforced **D**ialogue gen**ER**ation model*

Results & Ablation

	R1			R2			RL			BS	METEOR
	P	R	F1	P	R	F1	P	R	F1		
DialoGPT	12.34	40.48	15.72	2.92	11.83	4.42	12.23	38.60	15.76	0.7603	0.2021
GPT2	12.70	32.63	14.98	3.08	7.92	3.51	13.74	32.05	15.87	0.7445	0.1754
HRED	11.52	21.51	10.72	1.89	6.42	2.92	12.12	24.36	13.56	0.6259	0.1425
HRED w/ Sp. Utt. Encoder	11.77	28.63	10.08	1.29	4.19	2.06	12.25	21.27	12.72	0.6171	0.1801
RagRes w/ DialoGPT	12.41	43.91	16.12	3.70	13.72	4.98	11.92	41.02	16.30	0.7656	0.2098
READER – RAC-Head	12.64	41.48	15.78	3.60	11.83	4.58	12.3	38.64	15.90	0.7628	0.2039
READER	**12.82**	**43.93**	**16.15**	**3.77**	**13.67**	**4.93**	**12.51**	**40.82**	**16.32**	**0.7666**	**0.2103**
w/ Rew(RAC)	12.48	41.13	15.57	3.52	11.85	4.47	12.22	38.29	15.77	0.7527	0.2092
w/ Rew(RAC + BS)	11.73	38.82	14.65	2.28	8.45	2.96	11.21	35.76	14.53	0.7561	0.1840
w/ Rew(R)	12.01	40.45	15.18	2.72	9.93	3.52	11.46	37.05	14.97	0.7577	0.1908
w/ Rew(R + BS)	12.36	40.71	15.43	3.13	11.12	4.06	11.91	37.63	15.40	0.7609	0.2000
w/ Rew(BS)	11.92	38.06	14.70	2.43	8.26	3.11	11.40	34.98	14.58	0.7530	0.1874
$\Delta_{\text{READER}-BEST}(\%)$	↑ 0.94	↑ 8.5	↑ 2.73	↑ 22.40	↑ 15.50	↑ 11.53	↓ 8.90	↑ 5.69	↑ 2.83	↑ 0.82	↑ 4.05

- READER beats the best-performing baseline across 10 out of 11 metric scores with a **significant 22% increase in R1 Score**

Task: Dialogue-act Classification

https://colab.research.google.com/drive/1BEbhMRKKmytfx5MSthdHsinYVhuovIsh?usp=sharing

We will continue the tutorial from 11.50.

A SHORT BREAK

LCS2
LABORATORY FOR COMPUTATIONAL SOCIAL SYSTEMS

ICON 2023

AI4MH Tutorial

Aseem Srivastava, Neeraj, Yash Kumar Atri, Shivani Kumar,
Md Shad Akhtar, Tanmoy Chakraborty

https://ai4mh.github.io/

Is understanding **module enough?**

Experts maintain
Counseling Notes

Counseling Summarization using Mental Health Knowledge Guided Utterance Filtering

KDD 2022

Aseem Srivastava[♦], Tharun Suresh[♦], Sarah (Grin) Lord[♠❖] Md. Shad Akhtar[♦], Tanmoy Chakraborty[♦]

[♦] IIIT - Delhi [♠] University of Washington [❖] Mpathic.ai

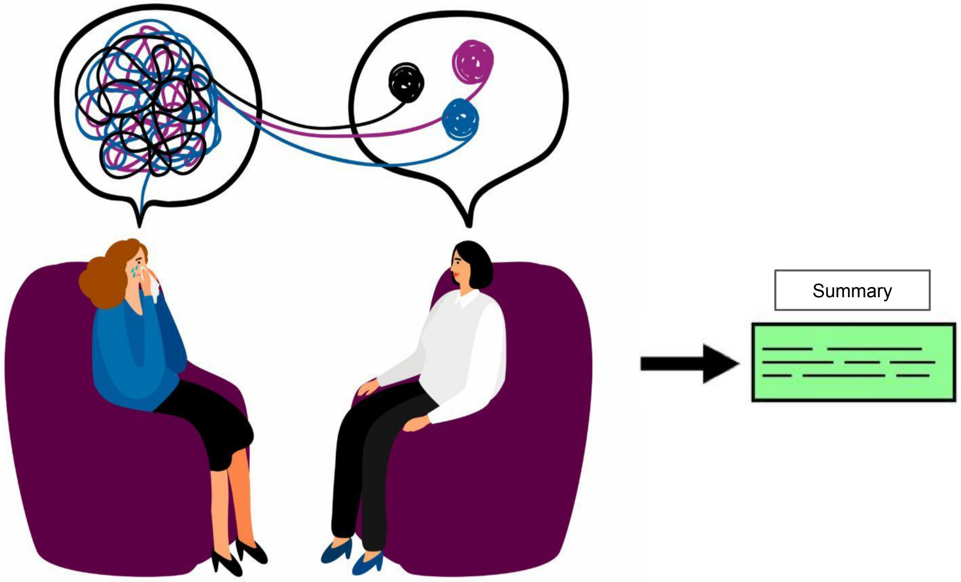
Understanding Counseling Conversations

- During a counseling session, an expert engages in a conversation with the client.

- When providing counseling, a mental health professional also has to summarize the key points in a summary aka ‘counseling note’.

- To meet the shortage in the mental health providers, there is a need to build AI-based summarization modules.

Dialogue Summary Generation

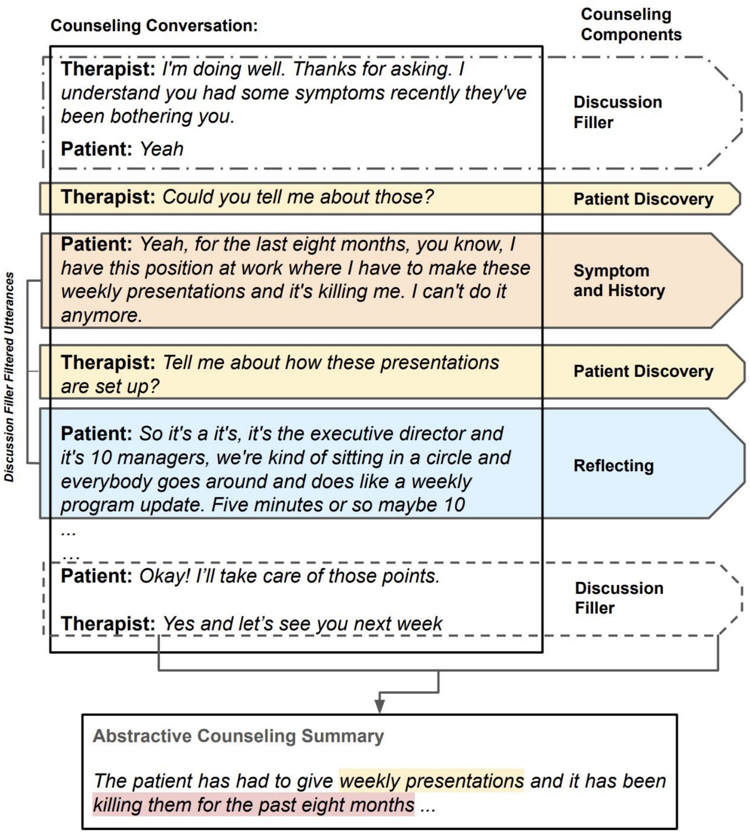


- **Understanding Client**
  - Topic
  - Symptoms / Reasoning
  - Story
  - Routine Conversation

- **Counseling Summary**

  Use of state-of-the-art deep learning models to generate quality summary.

Dialogue Summary Generation



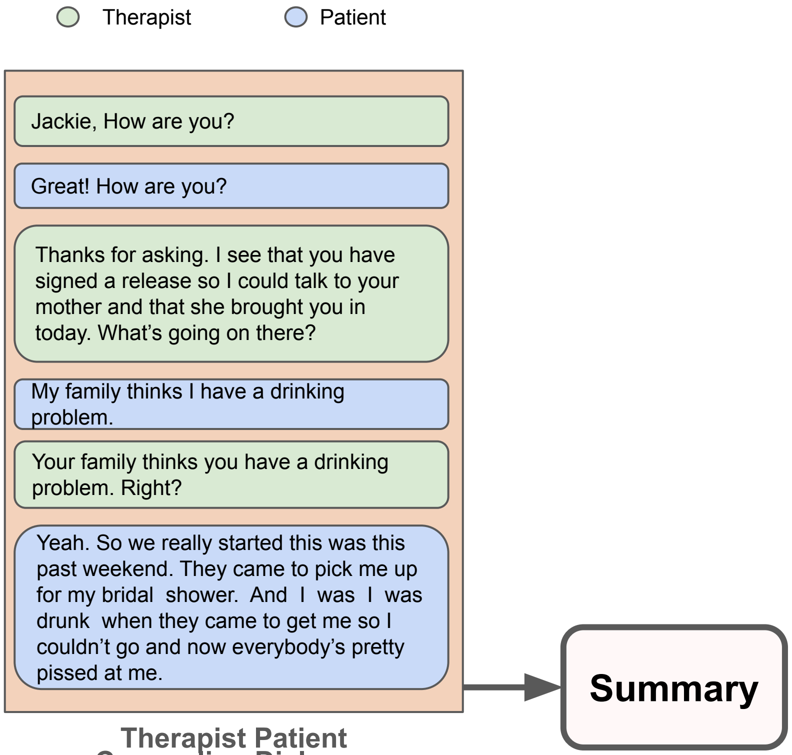
- HOPE - Cognitive behaviour therapy

- A session contains psychotherapy elements *viz.* symptoms, history of mental health issues (reflecting), or the discovery of the patient's behavior along with some discussion filler.

- We refer to these important components as **counseling components**.

Counseling Summarization Using Mental Health Knowledge Guided Utterance Filtering *In Proceedings of* KDD '22

Understanding Counseling Conversations

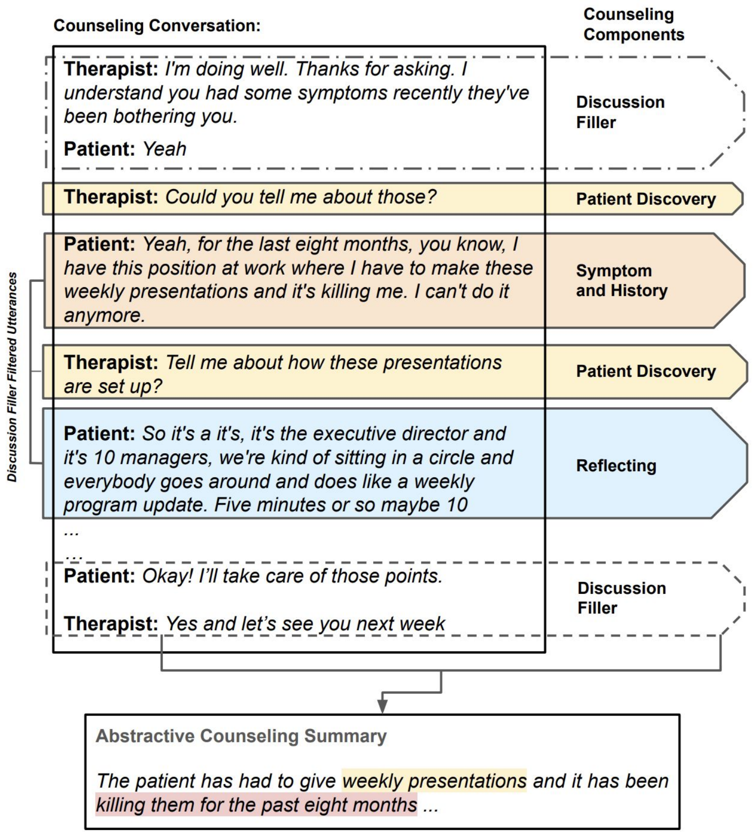


Therapist Patient Counseling Dialogue

- We worked on counseling based conversations are dyadic.

- Session ends with a counseling not containing summary.

Summarization of Counseling Conversations



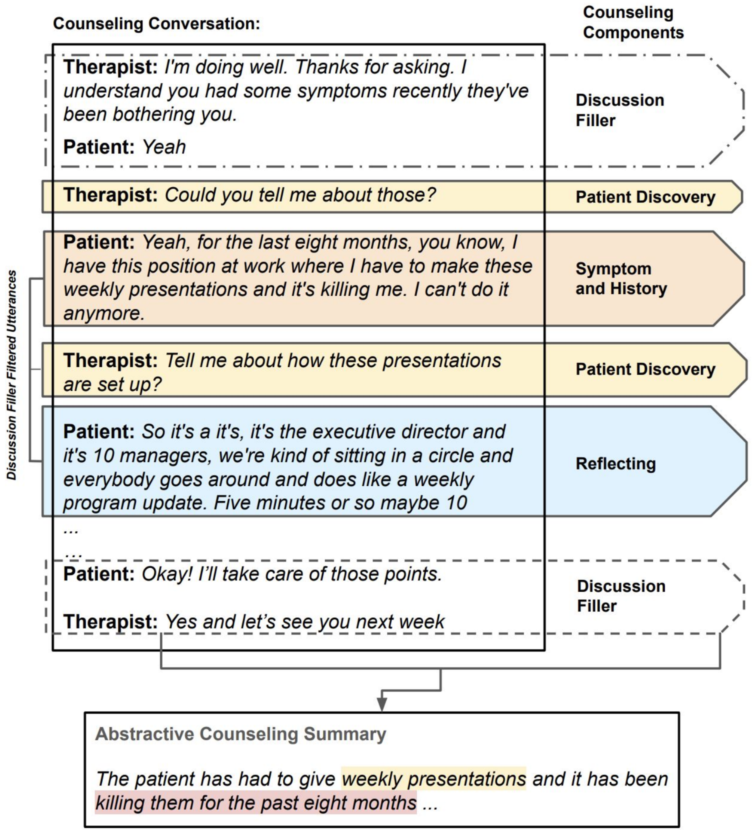
- HOPE - Cognitive behaviour therapy
- A session contains psychotherapy elements *viz.* symptoms, history of mental health issues, or the discovery of the patient's behavior along with some discussion filler.
- We refer to these important components as **counseling components**.

Summary of Our Contributions

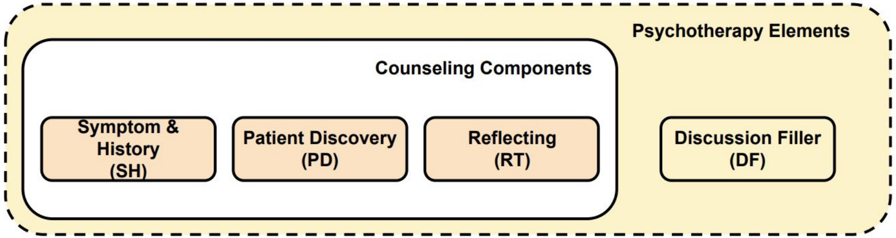
- **MEMO:** A novel counseling summarization dataset - MEMO. We also curate a novel annotation scheme for psychotherapy elements in utterances of counseling dialogue. (MEMO: Mental hEalth suMmarizatOn dataset)

- **ConSum:** A novel summarization model that exploits mental health domain knowledge and counseling components. (ConSum: Counseling Summarization)
- **MHIC:** We propose a new problem specific metric to evaluate summaries, MHIC metric which reasonably evaluates summaries that are most useful from a counseling perspective. (MHIC: Mental Health Information Capture)

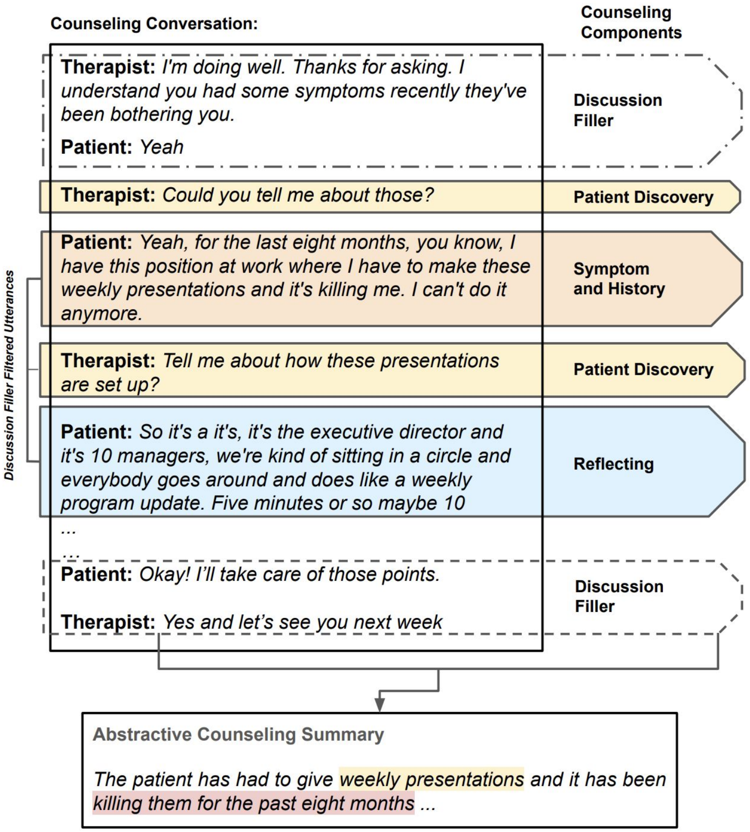
MEMO: Mental Health Summarization Dataset



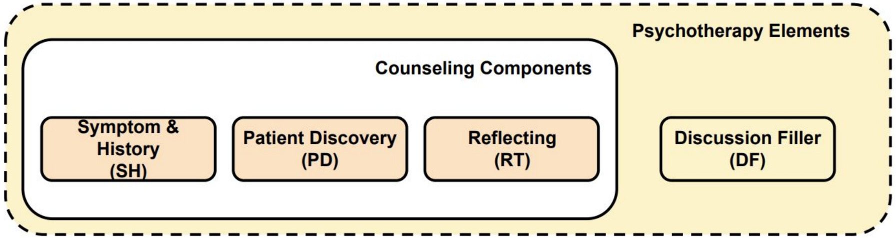
- Counseling components mostly contribute towards successful interventions.
- Discussion barely add relevance to the summary generation.
- We labeled utterances with four fine-grained labels



MEMO: Mental Health Summarization Dataset



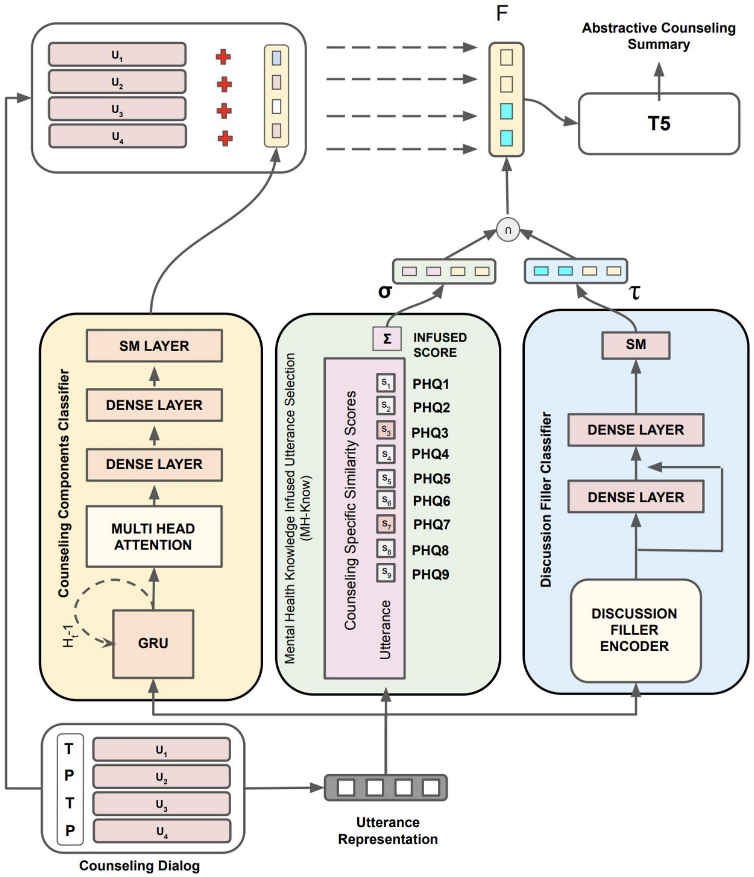
			Discussion Filler		Counseling Components						
					SH		RT		PD		
Split	**#D**	**Sp.**	U/D	#U	U/D	#U	U/D	#U	U/D	#U	**Total**
Train	152	Pt	5.02	764	5.67	862	2.52	383	18.84	2863	4766
		Th	8.10	1232	8.17	1243	4.22	642	10.80	1643	4877
Test	39	Pt	3.43	134	2.95	115	1.00	39	18.38	717	1004
		Th	5.46	213	4.51	176	4.43	173	11.28	440	1006
Val	21	Pt	8.09	170	4.23	89	2.28	48	13.80	290	594
		Th	10.48	220	5.57	117	4.66	98	7.38	155	597
Total	212	Pt	5.51	1068	4.28	1066	1.60	470	17.00	3870	6364
		Th	8.01	1665	6.08	1536	4.44	913	9.82	2238	6480



The Task: Counseling Summarization

- We represent the task as a summary generation task.
- We filter essential utterances to exploit essential knowledge while generating summaries.
- ConSum, a domain knowledge guided encoder-decoder deep learning model generates summary corresponding to each counseling conversation.

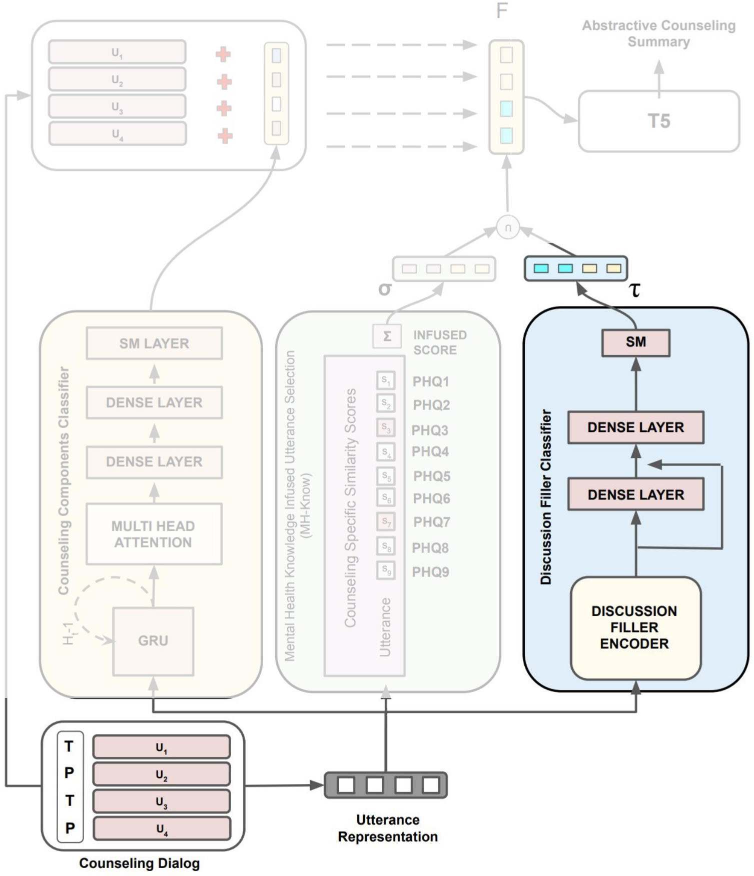
ConSum: Counseling Summarization Model



Complete architecture is divided into three filtering modules:

1. *Discussion Filler Classifier (DFC)*
2. *Mental Health Knowledge Infused Utterance Selection (MHKnow)*
3. *Counseling Components Classifier (CCC)*

ConSum: Counseling Summarization Model

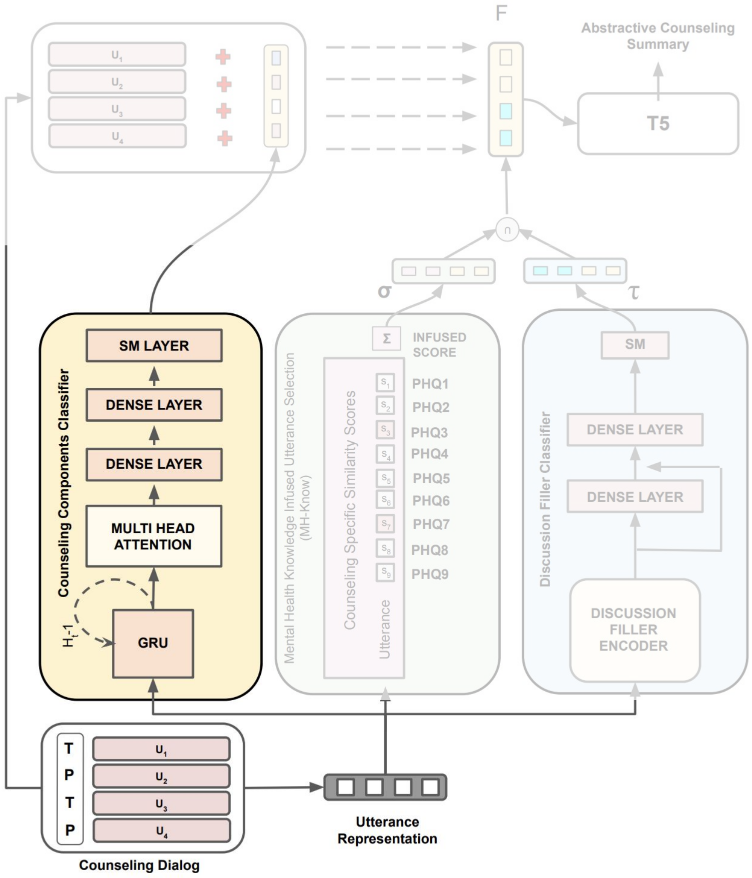


Discussion Filler Classifier (DFC)

DFC performs a binary classification task to mask utterance with relevant or irrelevant.

The output is a mask array [τ], where τ_i is a mask value of utterance U_i

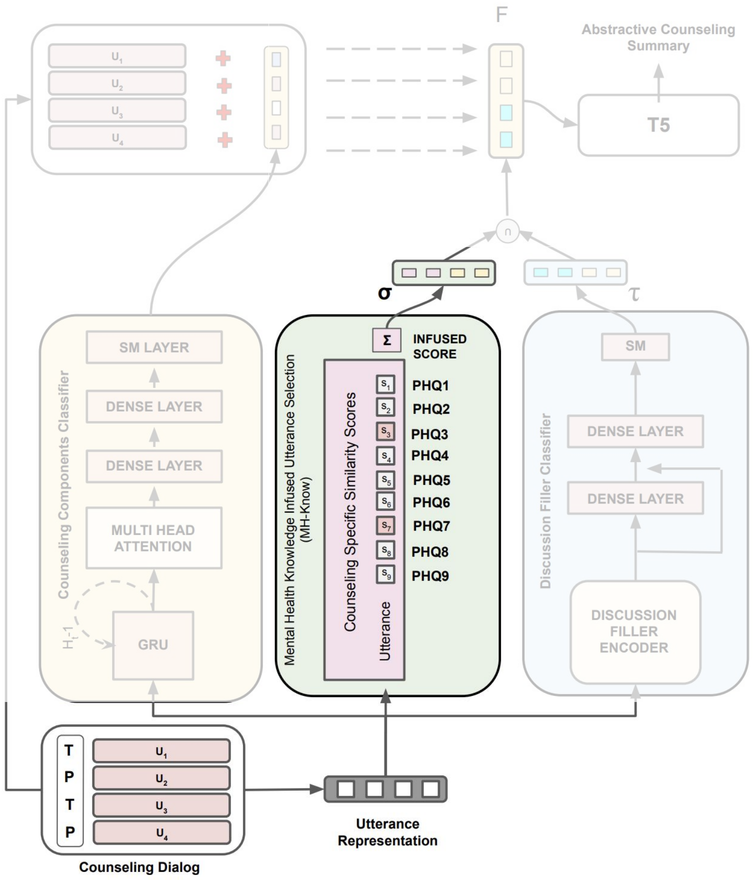
ConSum: Counseling Summarization Model



Counseling Components Classifier (CCC)

CCC uses contextual knowledge and attention to classify four counseling components.

ConSum: Counseling Summarization Model



Mental Health Knowledge Infused Utterance Selection (MHKnow)

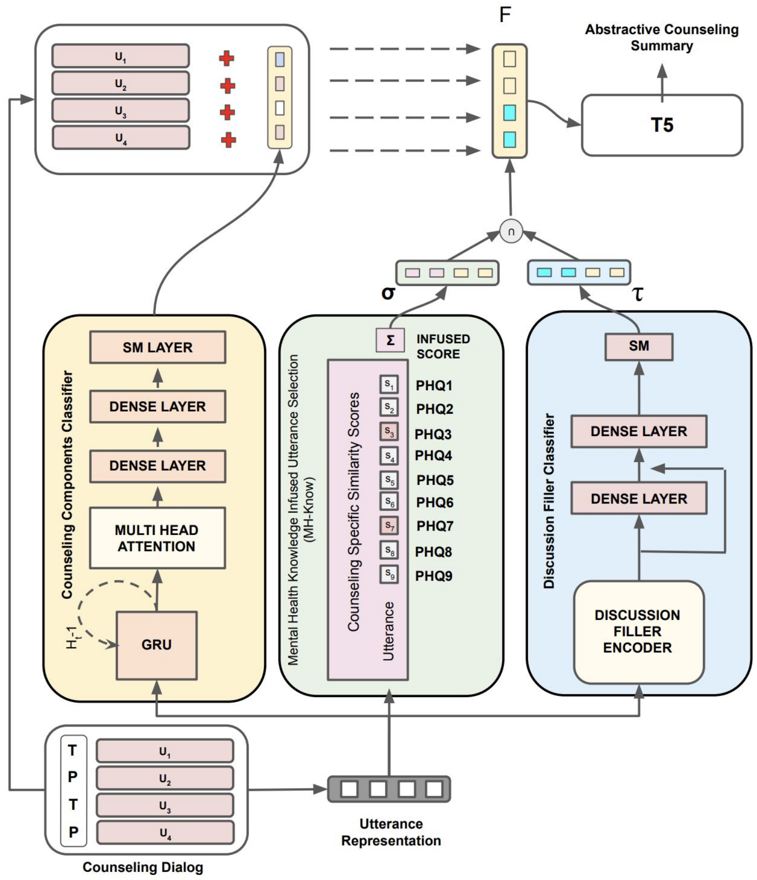
ConSum uses PHQ-9 lexicons to compute BERTScore similarity between lexicons and input utterance.

$$s_i = bertscore(u_i, phq_i) \Rightarrow \psi_i = \sum_{m=1}^{m=9} s_m$$

This creates a mask-array, σ_i containing 1 for cases where intervention similarity score is less than the threshold and 0 otherwise.

$$\sigma_i = \begin{cases} 1, & \text{if } \psi_i \leq \phi. \\ 0, & \text{otherwise} \end{cases}$$

ConSum: Counseling Summarization Model



Summary Decoder – T5

ConSum filters utterances with the help of all mask arrays.

The final subset of utterances with domain knowledge and counseling components are passed through the decoder to generate summary.

Results & Ablation: Intrinsic Eval

Model	R-1	R-2	R-L	QAE	BS
PLM	34.24	11.19	**33.35**	24.34	-0.8678
RankAE	25.57	3.43	24.16	29.98	-1.063
SM	20.46	3.80	18.87	20.22	-0.9454
Pegasus	29.71	7.77	27.57	**36.80**	-0.6130
T5	31.44	5.63	27.38	33.55	-0.5655
ConSum	**45.36**	**15.71**	24.75	25.42	**0.3407**

Counselling Label	R-1	R-2	R-L	QAE	BS
ConSum − SH − PD	20.92	5.00	7.44	20.71	0.0019
ConSum − PD − RT	36.00	9.00	9.14	20.47	0.2032
ConSum − RT − SH	28.63	8.06	9.55	23.02	−0.0209
ConSum − SH	39.77	9.55	8.98	24.11	0.1908
ConSum − PD	36.87	10.02	11.22	**33.38**	0.2420
ConSum − RT	42.01	9.83	16.50	18.03	0.2060
ConSum − MH-Know − CCC	39.67	9.95	12.69	21.19	0.2003
ConSum − MH-Know	40.42	10.09	11.00	23.97	0.2429
ConSum	**45.36**	**15.71**	**24.75**	25.42	**0.3407**

ConSum beats the best baselines by a margin of + 11.12 R1 and + 0.905 BS.

Results & Ablation: MHIC Metric

Mental Health Information Capture (MHIC)

- We propose a new **domain-centric** metric.
- Exploits the counseling components and R1 score to evaluate the generated summaries.

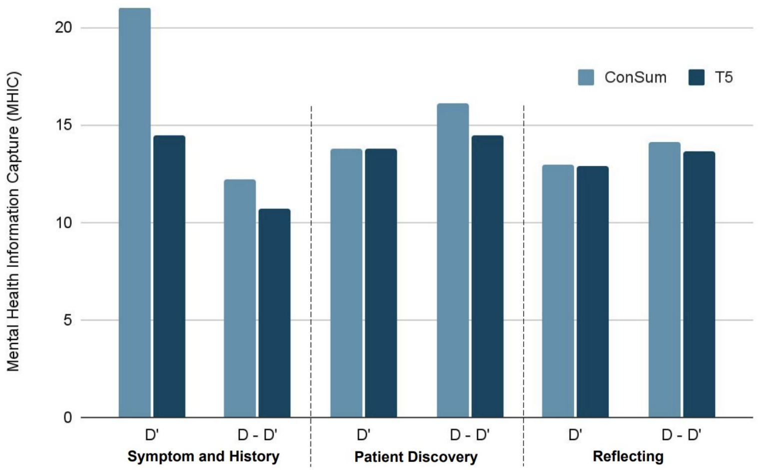
where

$$MHIC(CL_i) = R1\_Score(S_g, H_i)$$

$$CL_i \in \{SH, PD, RT\}$$

Counseling Components

Plot shows that our model, ConSum is able to capture counseling components better than T5



Result Analysis: Human Eval

Model	Summaries
T5	*Patient is asked a series of questions to test their housing. Patients do not want to take drugs, police arrest them and others in their area. They secretly talk about the incident. The patient lives with friends from different places including the city, where they are arrested or taken into custody. The patients feel they have no medical condition such as diabetes or stroke. These occured recently. The patient has had problems with housing due augensis. Patients feel they have been victimized by drug abusers since they were 12 years old. They know people who don't use marijuana to steal*
ConSum	*Whenever there is a basketball or or football, the patient does not really hear what to take. Patient feels overwhelming and feels like having this issue since childhood. Therapist made him realise his problems and patient was curing gradually. Therapist suggested him use. The patient was raised Catholic and they feel guilty about spending more money on drugs and parties. The patient has switched 12 houses in the past year and feel it must have affected their kids. The patient occasionally borrows money and gets less sleep in general*
T5	*The patient is tired when they have to go to work and it's a bit frustrating for them. They feel tired throughout the day without any food, no panic attacks, no medical condition such as diabetes or stroke. Patients are in a position where they can focus on anything. They do not want pills to reduce stress hence their life is limited by diet. The patient has lost three pounds in recent weeks due to this fatigue*
ConSum	*Whenever the patient goes to work. The patient is worried that they might have ADHD. The patient does not suffer from depression, anxiety nor use drugs or call it a metaphysical stuff. The patient wishes to get better and needs something to hold. The patient feel they sway at things, and they have two options. The patient was sent in by a counselor fearing they might hurt themselves. The patient's dad had committed suicide 15 years ago and their sister had attempted once. The patient has been diagnosed with depression and anxiety. The patient lives alone*

Model	Relevance	Consistency	Fluency	Coherence
RankAE	2.80	2.91	3.02	2.98
T5	2.99	3.05	3.04	2.95
ConSum	3.37	3.22	3.11	3.13

Understanding OMHCs and peer interactions

Online Mental Health Communities (OMHCs) are virtual spaces where individuals come together to discuss, share, and support each other in matters related to mental health.

Peer therapy involves individuals with shared experiences providing mutual support and understanding.

Hybrid

Human-AI Collaboration | Peer Counseling

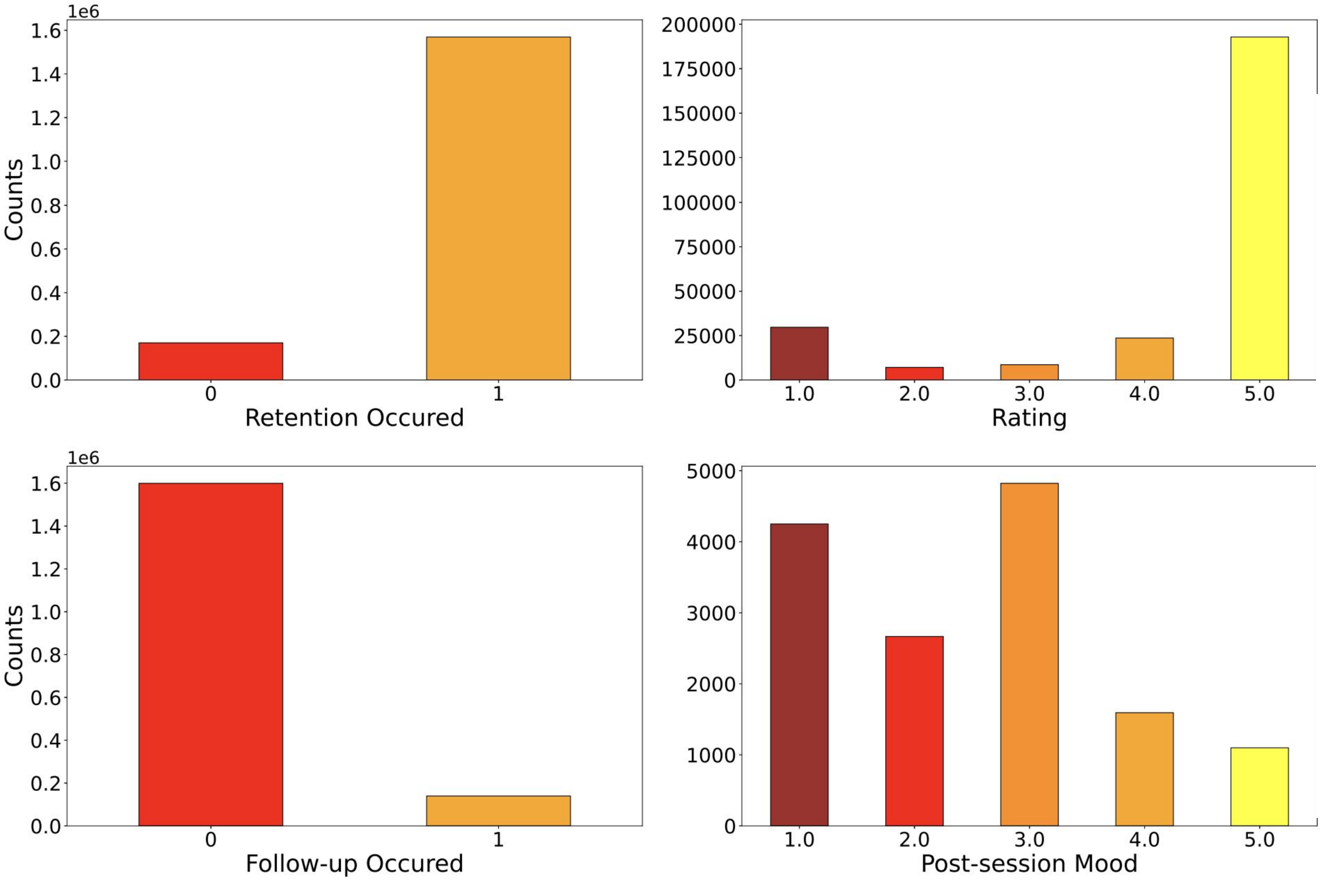
People post their problems online and other peers provide support online.

Super Expensive

Free! But has its own challenges

		Construct Level		
		Individual	**Conversation**	**Community**
Method of Measure	**Attitudinal**	BAI [4] BPRS [4] CDSS [4] GAD-7 [20] PHQ-9 [20, 41] Mood [3, 10, 41]	Rating [62, 83] Satisfaction [10] Session rating scale (SRS) [10] Support provision [80]	Attachment [77] Ease of use [4] Helpfulness [4] Information utility [39] Participation [39] Perceived support [39] Patient empowerment [39] Satisfaction [82] Social interaction [4]
	Behavioral	Affective word use [61] Complexity or repeatability [61] Psycholinguistic keywords [61] Readibility [61] Readibility [61] Symptomatic word use [61]	Conversation length [7, 8] Engagement [63, 73] Frequency [8] Support provision [79]	Engagement [45, 63, 73, 74, 82] Length of participation [77] Number of posts [61] Number of topics [61] Number of responses [61, 77] Support seeking [78]
	Annotation	Moment of change [41, 57]	Satisfaction [70] Self-disclosure [7, 8, 78, 79] Support provision [64]	Support provision [14, 80]

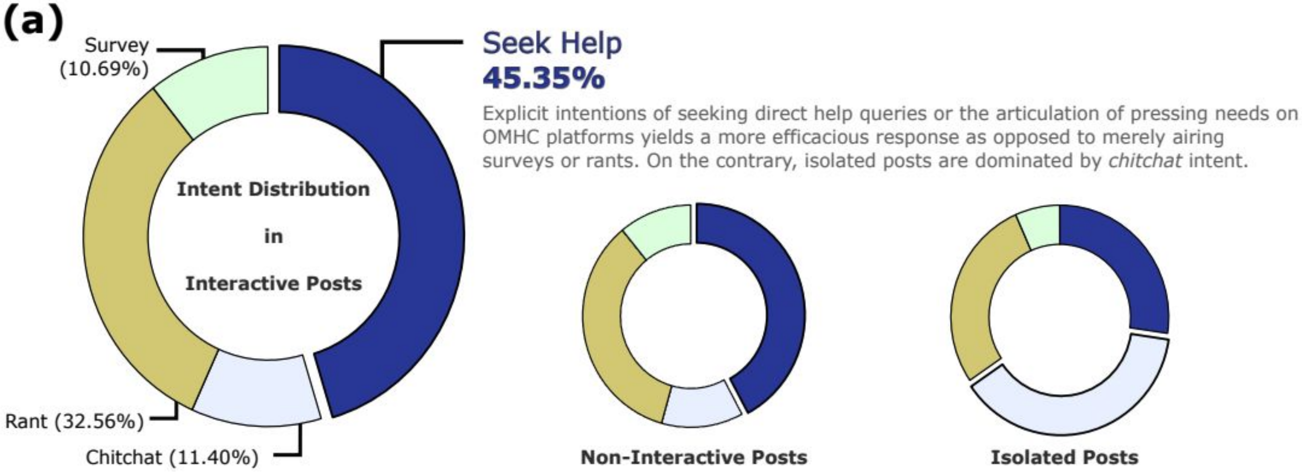
Tony Wang, Haard K Shah, Raj Sanjay Shah, Yi-Chia Wang, Robert E Kraut, and Diyi Yang. 2023. Metrics for Peer Counseling: Triangulating Success Outcomes for Online Therapy Platforms. (CHI '23) USA



Topic	Session Count	Percent (%)
Self Improvement	579192	33.30
Dating	297088	17.08
Parents	200584	11.53
Depression	162014	9.31
Romantic Relationship	135624	7.80
Lonely	120538	6.93
Suicide	67428	3.88
Pandemic	38114	2.19
Anxiety	28012	1.61
Home	26715	1.54
Family	20285	1.17
Sexuality	20235	1.16
LGBTQ	11331	0.65
Dissociative Identity	9856	0.57
Overwhelming	8004	0.46
Stress	5188	0.30
Health	4920	0.28
Intimacy	4270	0.25
Total	1,739,398	100

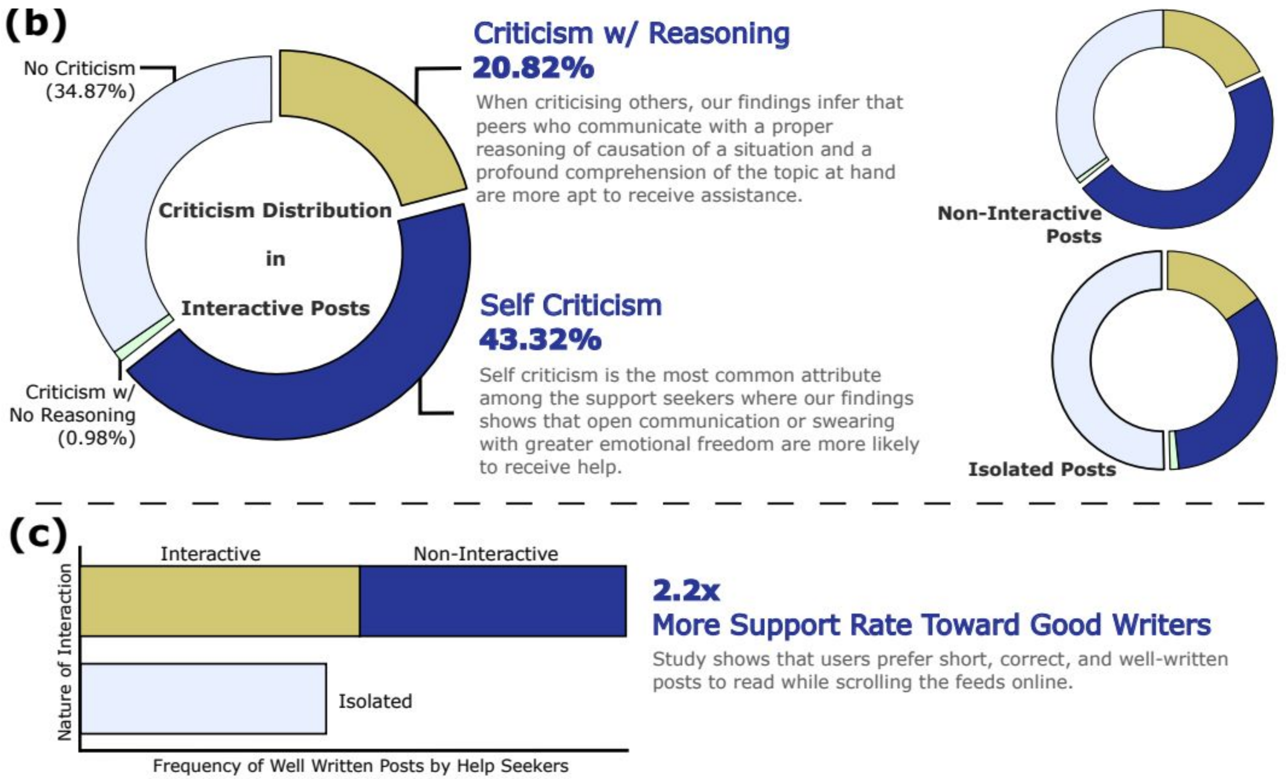
Tony Wang, Haard K Shah, Raj Sanjay Shah, Yi-Chia Wang, Robert E Kraut, and Diyi Yang. 2023. Metrics for Peer Counseling: Triangulating Success Outcomes for Online Therapy Platforms. (CHI '23) USA

How user behavior affects support delivery?



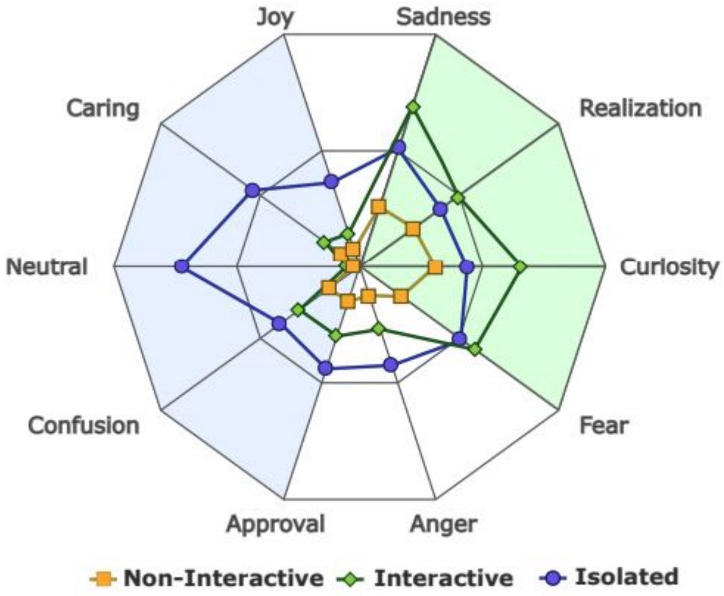
Critical Behavioral Traits Foster Peer Engagement in Online Mental Health Communities

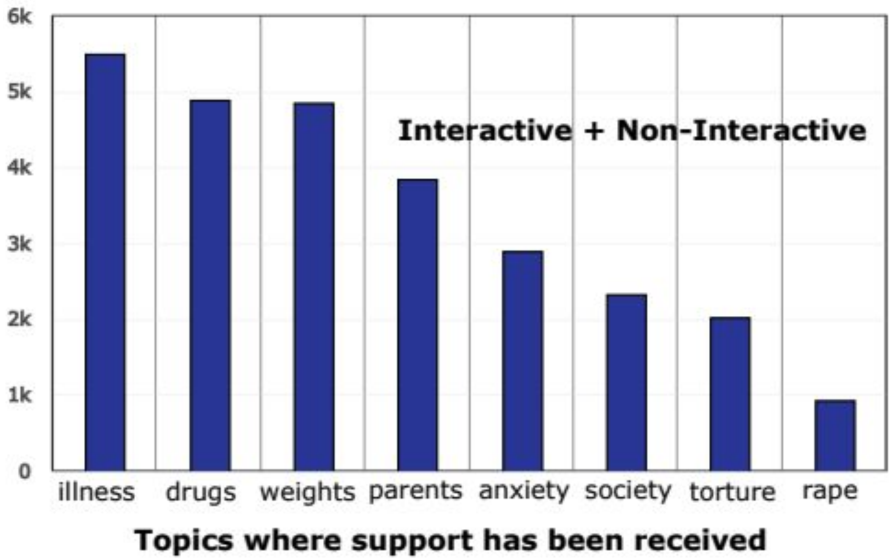
How user behavior affects support delivery?



Critical Behavioral Traits Foster Peer Engagement in Online Mental Health Communities

How user behavior affects support delivery?





Critical Behavioral Traits Foster Peer Engagement in Online Mental Health Communities

Current state of mental health x NLP

Current state-of-the-art LLMs in NLP X Mental Health

MentaLLamA

MentalBART

MentalT5

Fine-tuned Vicuna-33B covering 8 mental health analysis tasks.

Fine-tuned BART-Large covering 8 mental health analysis tasks.

Fine-tuned T5-Large covering 8 mental health analysis tasks.

MentaLLaMA - the current SOTA

Dataset

LLMs

Training Style

Evaluation

IMHI Dataset

Data	Task
DR	depression detection
Dreaddit	stress detection
CLP	depression detection
SWMH	mental disorders detection
T-SID	mental disorders detection
SAD	stress cause detection
CAMS	depression/suicide cause detection
loneliness	loneliness detection
MultiWD	Wellness dimensions detection
IRF	interpersonal risk factors detection

MentaLLaMA - the current SOTA

Dataset

LLMs

Training Style

Evaluation

LLaMa - 7B
LLama - 13B
LLama - 70B

70 billion parameters at 16-bit precision (2 bytes) equals about 140 GB in memory (70,000,000,000 x 2B = 140 GB)

MentaLLaMA - the current SOTA

Dataset

LLMs

Training Style → PeFT technique LoRA

Evaluation

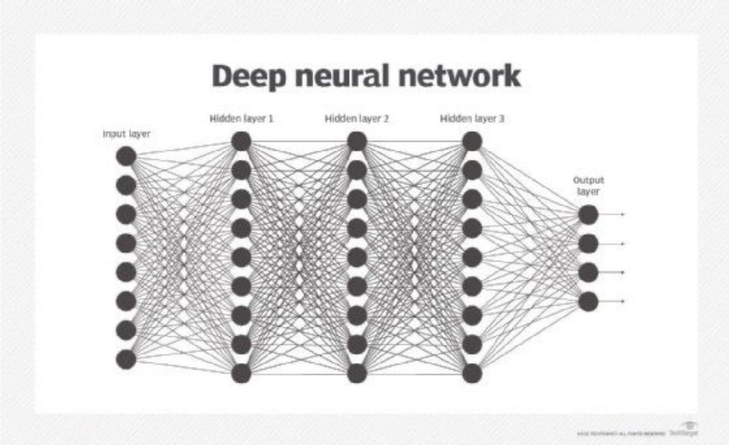
MentaLLaMA - the current SOTA

Dataset

LLMs

Training Style

Evaluation



{ PeFT technique LoRA }

MentaLLaMA - the current SOTA

Dataset

LLMs

Training Style

Evaluation

BERTScore and Human Ratings

Ethical Considerations

Major types of considerations

- Dataset Sensitivity.
- Moral Considerations.
- Legal Considerations.
- Other Considerations.

Major types of considerations

- **Dataset Sensitivity.**
- Moral Considerations.
- Legal Considerations.
- Other Considerations.

Data Collection and Privacy

1. Collection of sensitive mental health data.
2. Important to anonymize and de-identify.
3. Mitigating potential risks of re-identification.

Major types of considerations

- **Dataset Sensitivity.**
- Moral Considerations.
- Legal Considerations.
- Other Considerations.

Representativeness

1. Biases in mental health datasets results in bias in model.
2. Under-representation or over-representation of certain groups.

Major types of considerations

- **Dataset Sensitivity.**
- Moral Considerations.
- Legal Considerations.
- Other Considerations.

Informed Consent

1. Clear and informed consent processes.
2. Obtaining informed consent for mental health data is challenging.

Major types of considerations

- Dataset Sensitivity.
- Moral Considerations.
- **Legal Considerations.**
- Other Considerations.

Data Protection Laws

1. Such as GDPR apply to mental health data in NLP research.

Major types of considerations

- Dataset Sensitivity.
- Moral Considerations.
- **Legal Considerations.**
- Other Considerations.

Patient Rights

1. Rights of individuals in the context of mental health data.
2. Balance between research progress and individual rights.

Major types of considerations

- Dataset Sensitivity.
- Moral Considerations.
- **Legal Considerations.**
- Other Considerations.

Monitoring

1. The role of institutional review boards (IRBs) in overseeing research.

Major types of considerations

- Dataset Sensitivity.
- **Moral Considerations.**
- Legal Considerations.
- Other Considerations.

Impact on Vulnerable Populations

1. Effect of models on vulnerable individuals or communities.
2. Harms and benefits of NLP applications in mental health.

Major types of considerations

- Dataset Sensitivity.
- **Moral Considerations.**
- Legal Considerations.
- Other Considerations.

Stigmatization

1. Biased models may reinforce mental health stigmas.

Major types of considerations

- Dataset Sensitivity.
- **Moral Considerations.**
- Legal Considerations.
- Other Considerations.

Dual-Use Dilemma

1. Technology developed for research can be used for non-beneficial or harmful purposes.

Major types of considerations

- Dataset Sensitivity.
- Moral Considerations.
- Legal Considerations.
- **Other Considerations.**

Explainability and Transparency

1. Model interpretability is important in mental health applications.
2. Black-box models in clinical settings might be harmful.

Major types of considerations

- Dataset Sensitivity.
- Moral Considerations.
- Legal Considerations.
- **Other Considerations.**

Collaboration with Domain Experts

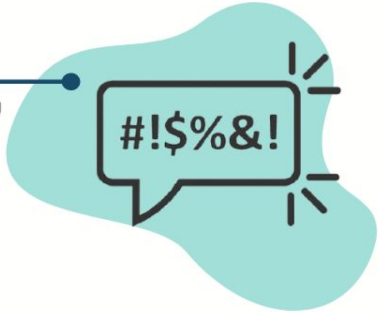
1. Need for collaboration between NLP researchers and mental health professionals to ensure ethically sound practices.

Research @ LCS2

Broad Research Areas @ LCS2

Cyber informatics

We design models for mitigating various cybercrimes on online social media including the spread of fake news, fraud activities, black-market driven collusion, hate speech.



Multimodality

We study the domain of speech and vision modalities with textual processing for a number of tasks such as summarization, emotion analysis, sarcasm detection, offensive post detection.

Conversational Dialogs

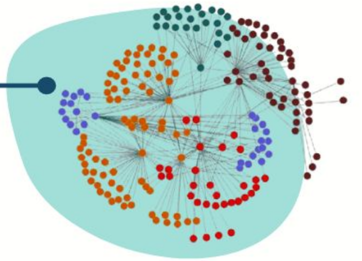
We regard various modules of a dialog system in dialog understanding and generation for both goal-oriented and generic chatbots.

Code-mixed & Low Resource NLP

We explore a wide range of NLP tasks in code-mixed and low-resource languages in Indian context. The prime objective is to

Complex Networks

We study social, information and scientific networks to explore various structural and behavioural properties of nodes and edges.



Team @ LCS2

Faculty:

- Dr. Md Shad Akhtar, Assistant Professor, IIITD
- Dr. Tanmoy Chakraborty, Associate Professor, IITD

- 15 PhDs
- 8 RAs
- Multiple MTechs, BTechs, and Interns
- Multiple Collaborators

Thank You!

Q/A

http://lcs2.iiitd.edu.in

@lcs2-iiitd

@lcs2iiitd