Counseling Summarization using Mental Health Knowledge Guided Utterance Filtering

Aseem Srivastava[1], Tharun Suresh[1], Sarah Peregrine (Grin) Lord[2], Md. Shad Akhtar[1], Tanmoy Chakraborty[1]

[1]IIIT Delhi, India [2]University of Washington [2]Mpathic.ai

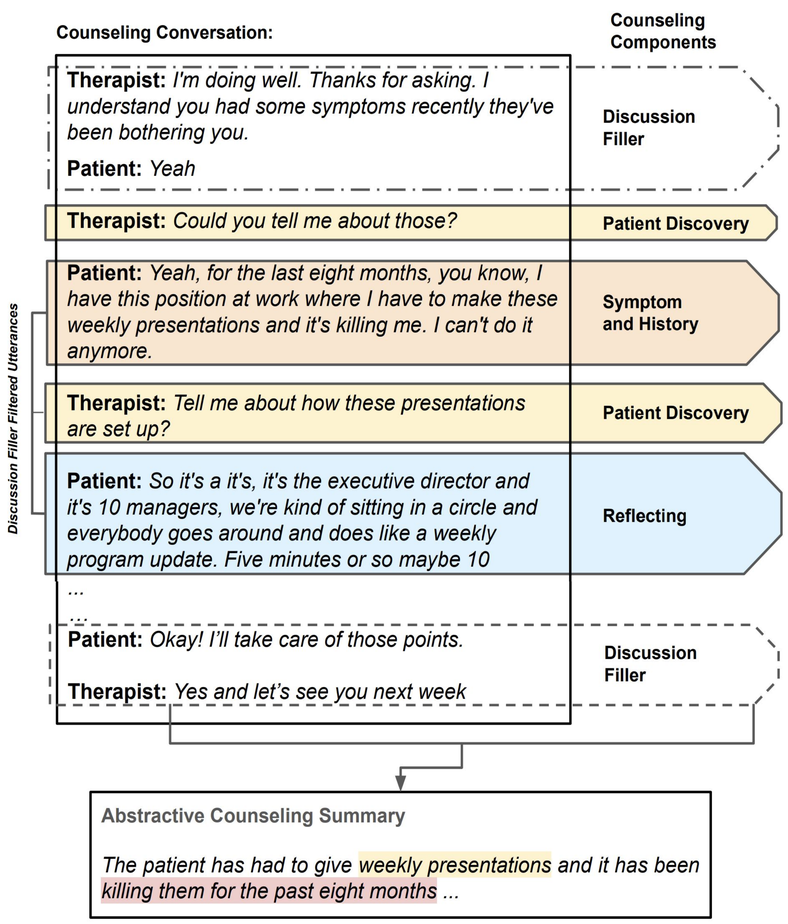
{aseems, tharun20119, shad.akhtar, tanmoy}@iiitd.ac.in; grin@mpathic.ai

Highlights

- This work aims to summarize mental health counseling to build upon domain knowledge and help clinicians quickly glean meaning.
- Counseling sessions contain discussions about symptoms, history of issues, or the discovery of behavior *aka* counseling components.
- We propose a novel counseling summarization dataset - MEMO and a novel annotation scheme for psychotherapy elements viz. symptom and history, patient discovery, reflecting aspects, and discussion filler in utterances of counseling dialogue.
- We propose ConSum, a novel counseling component guided summarization model which exploits knowledge from depressive symptoms, PHQ-9 Questionnaire.
- We propose a new problem specific metric to evaluate summaries i.e., Mental Health Information Capture (MHIC) metric.

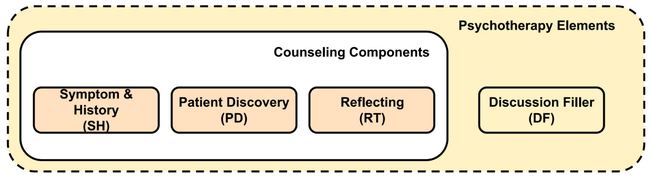
Problem Definition

Counseling Conversation Summarization:
Our work incorporates domain knowledge to generate counseling conversation summaries that inherit essential psychotherapy elements.



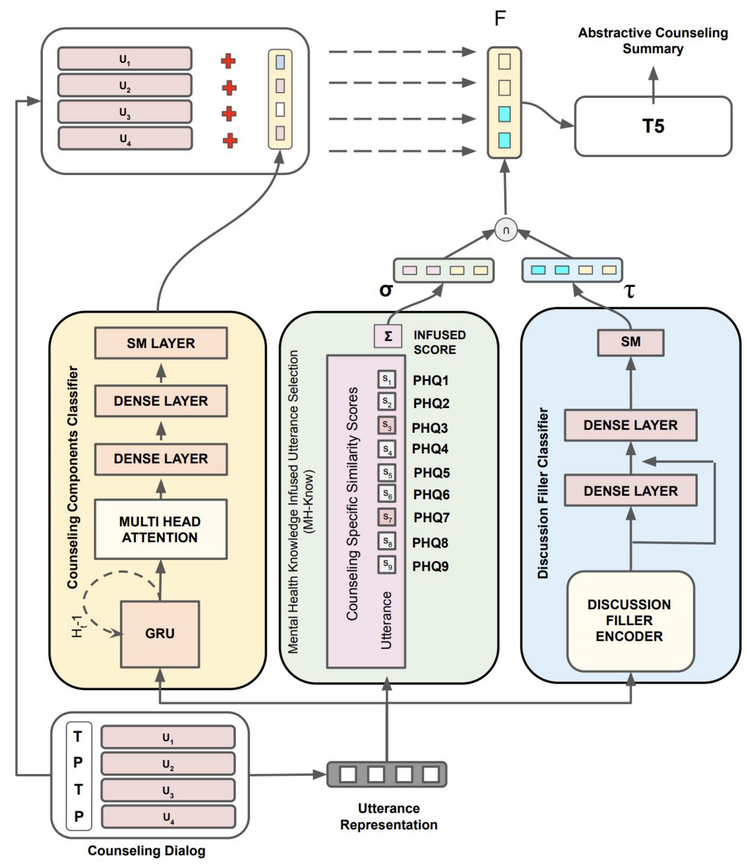
Dataset

Mental Health Summarization Dataset (MEMO) contains 12.9K utterances from 212 counseling conversations annotated with psychotherapy elements and summary.



Split	#D	Sp.	Discussion Filler U/D	Discussion Filler #U	SH U/D	SH #U	RT U/D	RT #U	PD U/D	PD #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

Methodology



We propose ConSum, a three module setup operating independently on the selection and filtering of utterances.

Module 1: Discussion Filler Classifier (DFC)
DFC operates on binary classification task using a feed-forward network to classify each utterance u_i into 'counseling' or 'discussion filler' label. The output is a mask array labeled τ, where τ_i represents mask array from dialogue D_i.

Module 2: Knowledge Infused Utterance Selection (MHKnow)
Here we exploit knowledge from PHQ-9 lexicons to compute BERTScore similarity with input utterance. The summation of similarity (ψ) is compared with the hyperparameter to decide the relevance of utterance.

$$s_i = bertscore(u_i, phq_i) \Rightarrow \psi_i = \sum_{m=1}^{m=9} s_m \quad \sigma_i = \begin{cases} 1, & \text{if } \psi_i \leq \phi. \\ 0, & \text{otherwise} \end{cases} \quad F = (\sigma \cap \tau)'; G = (U_i \oplus CL_i) \otimes F$$

Module 3: Counseling Components Classifier (CCC)
CCC is a contextually rich model which classifies utterances among SH, RT, PD, and DF for further selection of utterances.

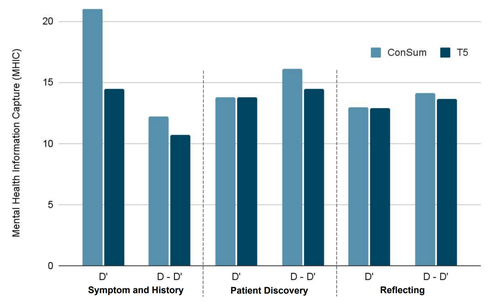
Summary Decoder: The final subset of utterances are used as an input to T5 model for abstractive summary generation

Results & Analysis

Model	R1	R2	RL	QAE	BS
PLM	34.24	11.19	33.35	24.34	- 0.867
RankAE	25.57	3.43	24.16	29.98	- 1.063
SM	20.46	3.80	18.87	20.22	- 0.945
Pegasus	29.71	7.77	27.57	36.80	- 0.613
T5	31.44	5.63	27.38	33.55	- 0.565
ConSum	45.36	15.71	24.75	25.42	+ 0.340

- ConSum beats the best baselines by a margin of + 11.12 R1 and + 0.905 BS.
- We propose Mental Health Information Capture (MHIC) metric which reasonably evaluates summaries that are most useful from a counseling's perspective.

Human Evaluation	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



Paper: https://arxiv.org/abs/2206.03886
Web: www.as3eem.github.io
Contact: aseems@iiitd.ac.in

References

[1] Juan C Quiroz, Liliana Laranjo, Ahmet Baki Kocaballi, Agustina Briatore, Shlomo Berkovsky, Dana Rezazadegan, and Enrico Coiera Identifying relevant information in medical conversations to summarize a clinician-patient encounter. Health Informatics Journal (2020).
[2] Kundan Krishna, Sopan Khosla, Jeffrey Bigham, and Zachary C. Lipton. Generating SOAP Notes from Doctor-Patient Conversations Using Modular Summarization Techniques. In Proceedings of the 59th ACL and the 11th IJCNLP 2021.

LCS2 LABORATORY FOR COMPUTATIONAL SOCIAL SYSTEMS

KDD 2022