

CSE556: Natural Language Processing 2025
Project Mid-Evaluation
Mental Health Counselling Summarization

Himanshu Raj
2022216

Ishita
2022224

Ritika Thakur
2022408

Abstract

This project focuses on developing an NLP-based system for generating concise and informative summaries of mental health counselling conversations. We aim to capture the key insights of therapeutic dialogues while preserving their contextual nuances.

1 Problem Definition

The primary objective of this project is to design an NLP system that generates concise and informative summaries of counselling dialogues. Key sub-tasks include:

- Utterance Filtering:** Identify and remove parts of the conversation that do not contribute meaningfully (discussion fillers).
- Counselling Component Classification:** Detect therapy-related elements such as symptom description, patient discovery, and reflective insights.
- Summary Generation:** Use the filtered and classified utterances to produce a coherent and context-aware summary.

The generated summaries are evaluated using BLEU and BERT scores to ensure both lexical precision and semantic fidelity to the reference summaries.

2 High-level Plan

We propose improving counseling dialogue summarization by augmenting existing models (like T5) with emotion-aware filtering. While T5 generates summaries from raw text, it overlooks emotional cues critical to mental health contexts. Our goal is to prioritize emotionally salient utterances (e.g., patient distress) during filtering and generation, producing summaries that better reflect patient states and therapeutic progress.

3 Approach

3.1 Baseline Model

- T5 Baseline:** We have used T5-base model from Hugging Face. The original paper fine-tuned the model for 200 epochs; we did it for 20 epochs with similar hyperparameters. The results achieved are as follows:

| Metric | Reported | Reproduced |
|---------------|----------|------------|
| ROUGE-1 | 31.44 | 37.65 |
| ROUGE-2 | 5.63 | 12.91 |
| ROUGE-L | 27.38 | 22.14 |
| BLEURT | -0.5655 | -0.6688 |
| BLEU | NA | 12.20 |
| BERT F1 Score | NA | 79.41 |

3.2 Emotion-Aware Enhancement

| Aspect | T5 Baseline | Proposed Model |
|----------------|--|--|
| Filtering | Removes inactive dialogues using the subtopic field in dataset | Emotion-aware filtering: Retains utterances with high emotional salience (sadness, fear) + PHQ-9 relevance |
| Generation | Vanilla T5 on filtered text | T5 with emotion-contextualized embeddings (e.g., "[sadness] Patient: 'I feel hopeless'") |
| Clinical Focus | Generic symptom reporting | Emotion-aware insights (e.g., links "crying spells" to depression severity) |

References

- [EmoBERT: A BERT-based emotion recognition model for mental health dialogues](#)