

GROUP 39

# Digital Counsellor

Generative based chatbot using a rank based approach



**Manav Saini**

manav20518@iiitd.ac.in

**Navidha Jain**

navidha20223@iiitd.ac.in

**Tanishqa Shital Singh**

tanishqa20411@iiitd.ac.in

**Anishka**

anishka20282@iiitd.ac.in

**Ashwin Tomer**

ashwin20289@iiitd.ac.in

**Dux Pal Singh**

dux20297@iiitd.ac.in

# What is Digital Counselling?

Digital counselling refers to the use of digital technologies, such as chatbots, online platforms, and virtual communication tools, to provide mental health support and counselling services.

# Importance of Digital Counselling



- **Increased Accessibility:** accessible way for individuals to seek mental health support regardless of their geographical location, physical ability, or time constraints, particularly beneficial for those in remote areas facing transportation challenges.
- **Enhanced Convenience:** convenience of scheduling sessions at flexible times that work best for the individual.
- **Confidentiality and Privacy:** prioritize confidentiality and privacy, allowing individuals to seek support in a discreet and confidential manner using secure and encrypted communication methods to protect client information
- **Expanded Reach:** reach a wider audience, including those who may be hesitant to seek traditional in-person counselling due to stigma, cultural or language barriers, or other reasons.

# Abstract

This paper proposes the development of a digital counselling chatbot that utilizes a generative model with a rank-based approach to training data. The chatbot aims to provide accessible and affordable mental health support to a wide range of people by providing a safe and confidential space for users to discuss their concerns and receive practical advice and information. The chatbot is trained on a dataset scraped from Counsel Chat, a platform that hosts certified counselors whose answers are ranked with upvotes. A medium version of GPT2 has been used and fine-tuned on the dataset to generate intelligent responses. The paper compares the proposed chatbot with a baseline model and evaluates the technical performance using metrics such as BLEU, METEOR, and NIST. Overall, the proposed chatbot with the rank-based approach and generative model shows promising results for providing digital counselling. The paper contributes to the growing body of literature on the use of digital technology for mental health care and highlights the potential of chatbots to provide accessible and affordable mental health support.

# Introduction

The use of digital technology has transformed the field of mental health care, particularly in the development of digital counselling chatbots. These chatbots offer accessible and affordable mental health support to those who may be hesitant to seek help due to stigma, lack of access to services, or concerns about the cost of care. The project aims to develop a digital counselling chatbot that provides a safe and confidential space for users to discuss their concerns and receive practical advice and information.

# Problem Statement

Our aim was to develop a chatbot using a generative model that used a rank-based approach training data. To obtain this training data, we scraped data from Counsel Chat which is a platform that hosts certified counselors whose answers are ranked with upvotes.

# Motivation

The motivation behind creating a digital counselor with a dataset that incorporates human judgement is to provide individuals with a more accessible and efficient means of seeking mental health support. While digital counseling can be effective for some people, it is not a substitute for traditional face-to-face therapy. That is why incorporating a human judgment dataset can enhance the effectiveness of digital counseling by providing personalized and empathetic responses to individuals. This technology has the potential to fill the gaps in traditional mental health services and help individuals overcome barriers to care.

# Related Works

## Chatbot for mental well-being

- The paper proposes a project to develop a generative chatbot for mental well-being.
- The chatbot uses an SVM classifier for mood detection and a Seq2Seq model for generating appropriate responses.
- The proposed chatbot demonstrates the potential of using machine learning techniques for effective communication and stress relief.

## Counsellor Chatbot

- The article presents a study on a chatbot named Xen that provides counseling services and advice based on user input.
- Xen uses Retrieval and Generative techniques, including AIML and K-Means self-learning to deliver its services.
- The chatbot is trained on three datasets, resulting in three types of chatbots: Retrieval Pattern Matching, Retrieval Rule-Based AIML, and Generative.
- The article discusses different chatbot architectures, including retrieval-based, AIML-based, and generative, and how they are used in chatbot development.
- The techniques such as NER and skip-thought vectors are utilized for improving chatbot performance.
- The article provides an overview of various techniques used in chatbot development.

# Related Works (contd.)

## Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation

- The paper discusses the use of AI-based methods for recognizing human emotions and providing conversational counseling services for mental health care using a counseling chatbot.
- The chatbot uses deep learning-based emotion classification models and NLP methods to analyze consultation content and provide appropriate responses.
- It focuses on emotion recognition and monitoring, conversation understanding on the chat assistant, and developing a personalized dialog system that communicates emotionally with the user using information such as facial expression, age, sex, spatial context, location context, and bio-signals collected via wearable devices.
- The goal is to develop a personalized, user-customized correspondence technology that communicates with users through speech, text, audio, and visual representation based on the user's age, gender, and recognized emotions.
- The paper emphasizes the importance of tracking persistent emotional changes for improving the effectiveness of counseling.

# Novelty

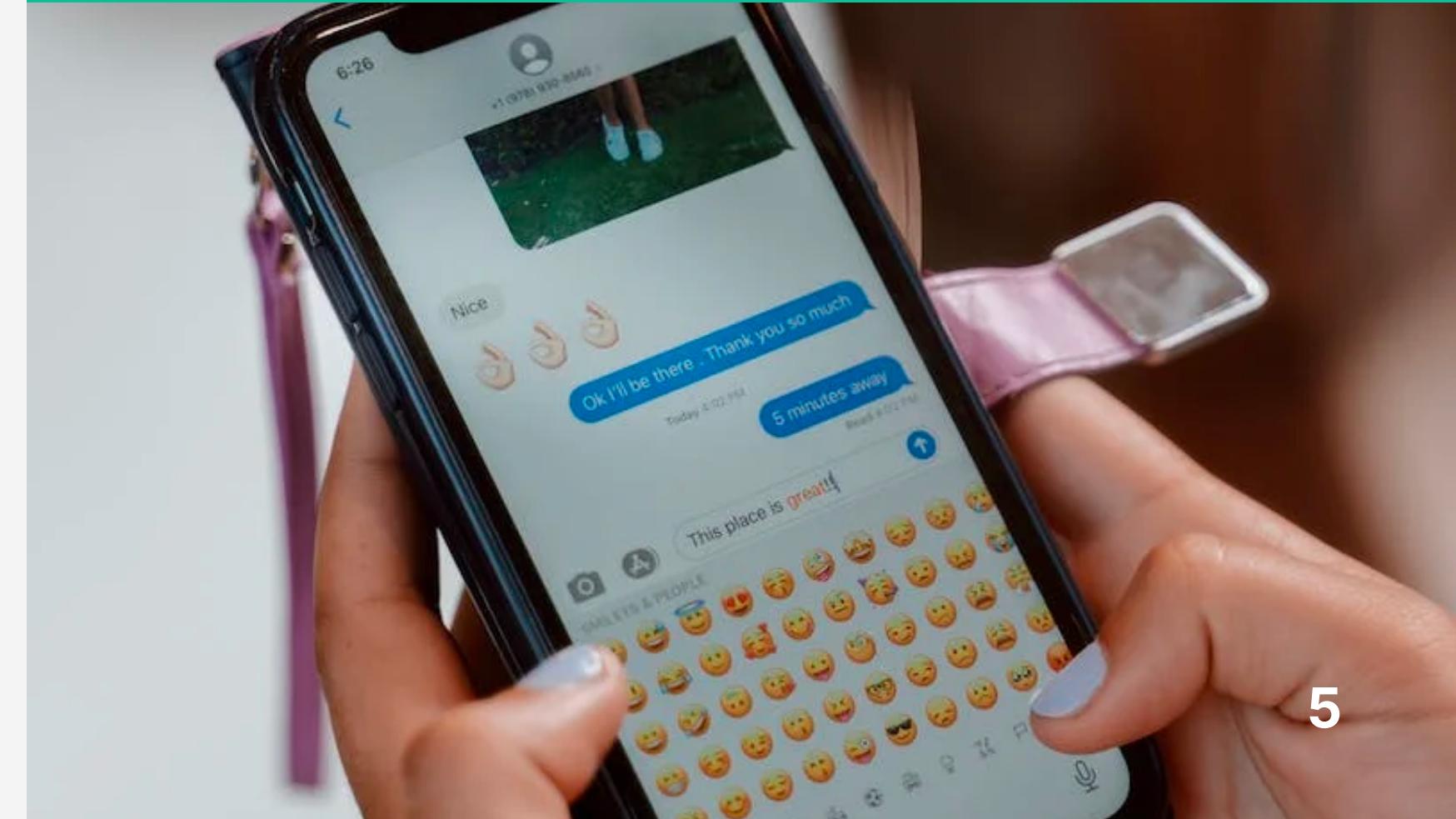
According to the literature review that we did the two main problems identified in the digital counselling sector is a lack of high quality data and a lack of human judgement metric. Through Digital Counsellor we are giving a solution for both the problems. We have scraped and created a database of questions and answers given by certified counsellors with years of experience with the data of up votes on each answer which can be used as a pre result human input metric.

# More Common than You Think

- The idea is to develop a digital counseling chatbot that utilizes NLP algorithms, ML techniques, and personalized counseling sessions.
- A generative model-based design is chosen for conversational flexibility as retrieval-based and rule-based chatbot designs are insufficient for multilinear conversations. This model allows the chatbot to learn and generate responses based on a large amount of training data, and it can be fine-tuned with different domain data for unique purposes.
- Metrics like BLEU, METEOR, and ROUGE-N cannot define whether a chatbot is good or effective, so human judgement criteria in the form of upvotes are used before training the model. Upvotes provide a metric of how appropriate responses are according to human judgement, and this data can be used with the profile data of counsellors to create a high quality dataset needed for generative models.
- The novelty lies in using generative models to provide a solution for digital counselling.

## Problems

- Lack of a bootstrap high quality data due to confidentiality laws of counsellors
- Lack of human judgement in the evaluation of digital counselling chatbots

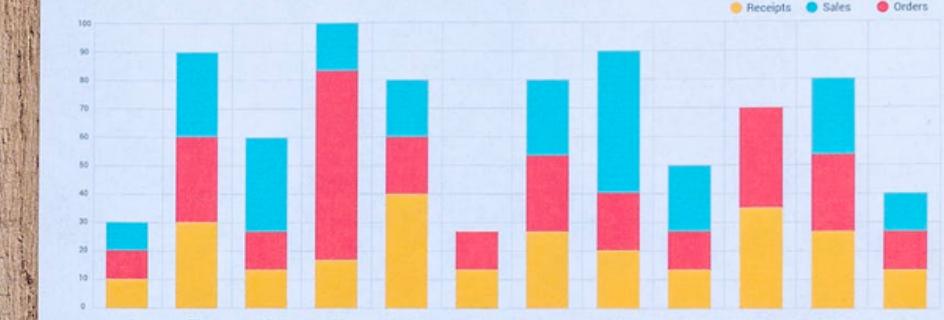


# Proposed Solution

- **Dataset**
  - Identified a website on which a number of verified counsellors give answers to the problems people are facing in day to day life
  - We have developed a web scraper that can scrape all this data with the counsellors information along with the upvotes that a particular answer has got
  - We follow a rank based approach according to the number of upvotes for a particular question and years of experience of a counsellor and form a high quality dataset and in which the upvotes act as a human input
- **Generative Model**
  - We then transform the data into a question answer format text file
  - We use that text file to finetune a generative model like GPT2

# Methodology

1. Scraped data from counselchat.com: We collected data in the form of questions, answers, profile links of counsellors, and upvotes, and built a database of 45,000 entries.
2. Pre-processed data: To remove redundant characters, we pre-processed all the questions and answers and ranked the data based on the upvotes.
3. Splitted and converted data: We split the data into train, test, and validation sets and converted each of them into a text file to fine-tune GPT-2.
4. Fine-tuned GPT-2: We fine-tuned GPT-2 on different hyperparameters and saved the models.
5. Best hyperparameters and model: Based on human judgement and wordcloud analysis, the model with hyperparameters top\_k=10, top\_p=0.7, temperature=0.8, nsamples=1, batch\_size=1, run\_name='run1', truncate=' ', and include\_prefix=False worked the best.



Business items



# Baseline Evaluation and Metrics

Dataset: Rank based dataset on the basis of upvotes

Model: GPT2

METRIC	SCORE
bleu_2	0.3681392715705741
bleu_4	0.2504935456423313
meteor	0.35994374360752923
nist_2	2.399633818486239
nist_4	2.7366025490061476

# Future Work

- Hybrid Retrieval based and generative based chatbot
  - We can finetune a retrieval based model on the dataset formulated by us
  - Generate output on same set of questions from generative and retrieval based model
  - Annotate the data for particular question which output was better out of the two models
  - Create a hybrid based chatbot

Thank you.

<b>Task</b>	<b>Members</b>
Data collection, pre-processing of data, identification of different models to be tested, final evaluation	Dux Pal Singh, Anishka
Implementation of different models for generation task and choosing the best one, Improving baseline scores	Manav Saini, Navidha Jain
Implementation of different models for response generation and choosing the best one, Improving baseline scores	Ashwin Tomer, Tanishqa Shital Singh
Report writing	All members

# Contributions